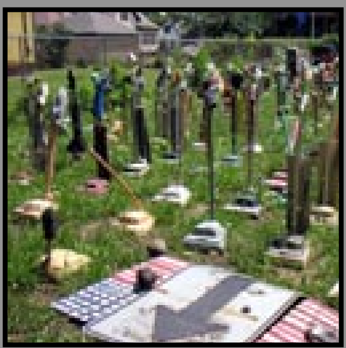


# Roses, Icebergs, Hoovers and all that Language

## an investigation of the cognitive foundations of our comprehension of object mediated communication

PhD thesis by
**Kristian Tylén**

The Institute of Language and Communication,
University of Southern Denmark

nov. 2008

# Preface

This PhD thesis is the culmination of three years of hard work and great fun. I must admit that I have had a great time exploring this interesting strand of human cognition and communicative practice and acquainting myself with ever new research methodologies, communities and traditions. Hopefully it will show.

My initial choice of subject for the project – our comprehension of material signals – owes itself to a lecture held once by professor Per Aage Brandt at the Center for Semiotics, University of Aarhus. In a parenthesis, Professor Brandt noted his curiosity that we can sometime experience aspects of our otherwise inanimate, mute and static material environment as addressing us. Some material objects seem to "speak to us" in ways that we would normally only expect people to. When, shortly after, I was invited to make a contribution to the journal *Cognitive Semiotics* on the topic of language and agency I decided to pursue the idea about "communicating objects" in more conceptual detail. Later, these initial ideas became the subject of a PhD application sent to The Institute of Language and Communication, SDU, and were granted a stipendium.

The project was initially intended as a purely conceptual, analytical investigation of the cognitive foundations of material signals. This type of approach is also reflected in the early part of the investigation, manifested in *Article I* of the thesis. But influenced by the generally more empirically oriented environment at the Institute of Language and Communication I became increasingly interested in ways of validating the proposed conceptual framework. In the autumn 2005, I coincidentally bumped into the semiotician, anthropologist and neuroscientist Andreas Roepstorff at a workshop on 'Language, Mind and Brain' at the University of Aarhus. He showed great interest in my project which he found was tightly related to some of the work in experimental cognitive neuroscience going on at The Danish National Research Foundation's Center of Functionally Integrative Neuroscience (CFIN), Aarhus University Hospital. I was invited to take part in the establishment of a new project group called 'Interacting Minds – a biological basis' and was given the opportunity to work with some of the new methodologies and technologies within the field of cognitive neuroscience. The product was a functional brain imaging experiment on material signals conducted in cooperation with Andreas Roepstorff and my old fellow student from semiotics Mikkel Wallentin, who meanwhile had become an experienced neurolinguist. Most of what I know today about designing and programming experiments and analyzing brain imaging data is due to Mikkel's unquenchable patience and Andreas' immense ambitions.

The fruitful research environment at the CFIN was even more enhanced when The Danish National Research Foundation granted the UCL professors Chris D. Frith and Uta Frith a 'Niels Bohr professorship' to participate in the Interacting Minds project. Suddenly I found myself discussing material signals over morning coffee with some of the world leading people within the fields of social cognition and brain sciences. Together with the rest of the 'Interacting Minds' group they have been invaluable sources of inspiration and insights for this project.

Besides the brain imaging study I have advanced more humanistic style experiments. One of these was conducted together with master student in linguistics and cognitive semiotics Johanne

Stege Philipsen. The study (reported in *Article III*) owes a lot to Johanne's insightful observations, comments and critiques. In the course of analysis we included another member of the 'Interacting Minds' group, Ethan Weed, as a third coauthor. Ethan has also been a tremendous help both as a sparring partner in, among other things, choices of statistical and experimental procedures and as a proofreader of the articles.

While much time has been spent on empirical studies in Aarhus, I also participated in the establishment of the Language and Cognition Research Group at the University of Southern Denmark. With its more linguistic orientation this group has been an invaluable source of insight into aspects of usage-based and cognitive language studies and a great place to discuss a lot of cognitive issues related to my project.

I must thank my supervisor professor Carl Bache who has been most supportive and given me a lot of confidence in my perhaps somewhat unusual project. He has also been a great help in navigating through the PhD program, with all its forms, norms, and regulations.

Last but not least, I will thank my fantastic family, my wife Vibeke and our kids Jonas, Bastian and Maja for all their support and patience.

# Table of Contents

# 1 Introduction

This PhD project addresses the cognitive foundations of a special kind of human communicative practice, that is, when we employ aspects of our material surroundings in mediation of communicative interaction. As such the study can be considered an extension of current trends in cognitive language studies occupied with various genres of non-verbal communication including gesture and 'body language'. This project thus extends the scope to material objects when these are employed as signifiers in human signaling activities. In a recent article, Herbert Clark (2005) proposes to treat these various genres of non-verbal communication under the unifying concept of *material signals.* In this thesis I will adopt his concept but not in the original broad sense. Instead - unless otherwise specified – 'material signals' will be taken to designate material objects (or other aspects of our inanimate physical surroundings) that are used or perceived as signifiers in intentional semiotic interactions.

A central assumption guiding the project is thus that material signals are used and comprehended in functional continuity with other modes of human communicative mediation. That is, they are 'a kind of language' functioning complementary to (and sometimes even interchangeable with) gesture, bodily expressions, and verbal language (whether spoken, written or signed). Given the substantial structural differences, this somewhat fragile analogy entails a certain approach to the concept of language. I will thus adopt a *language-as-an-activity* perspective akin to H. Clark (1996). The idea is to approach communicative mediation in actual usage situations, that is, when various types of signals are used among people to coordinate activities and engage in joint meaning constructing activities. In these situations of concrete language use the structural (*language-as-an-object*) details would seem secondary to their mediating functionality. These matters will be more thoroughly discussed in a later section (cf. section 2.5.1).

In the following, I will attempt to sketch out what will be and will not be understood by 'material signals' in this thesis. A number of existing studies within developmental, comparative or discourse studies have addressed issues of material signals in various kinds of online ostensive social contexts (cf. e.g. Ganea *et al.*., 2008; Myers & Libens, 2008; Preissler & Bloom, 2008; Younger & Johnson, 2004; Tolar *et al.*, 2007; Simcock & DeLoache, 2006; H. Clark, 2005; Tomasello, 1997; Troseth *et al.*, 2007; Suddendorf, 2003; Johnson *et al.*, 2005; Striano *et al.*, 2003; Rakoczy *et al.*, 2005; DeLoache, 2004. See also the section 2.1.2 and 2.1.3 for a review). In this project I focus on a slightly different kind of material communication – the *non*-ostensive, *monological* situation where we encounter an inanimate, static, mute material scene that is experienced as manipulated or arranged in a striking and deliberate manner that calls for a special kind of social meaning-exploring interpretation. Various traces of intentional manipulations seem in efficient ways to frame material objects as communicative signals, either in combination or isolation. To briefly exemplify, consider for instance chairs put out in the street to reserve a parking space, an iceberg entirely covered in red paint to induce an aesthetic experience, or a bunch of roses left on the sidewalk in memory of someone who died there. In all cases some person(s) intentionally left behind conspicuous traces of her (their) actions in the material environment to be recognized by others as meaningful social signals.

Generally, I have aimed a broad definition of object mediated communication and have in my experiments drawn on a large repertoire of quite different types of material signals, from artistic aesthetic installations to more instrumental instructional object manipulations and conventional material symbols. Thereby this project differs from most existing studies that typically focus exclusively on either *iconic* (Ganea *et al.*, 2008; Myers & Libens, 2008; Preissler & Bloom, 2008; Younger & Johnson, 2004; Tolar *et al.*, 2007; Simcock & DeLoache, 2006), *indexical* (H. Clark, 2005; Tomasello, 1997; Troseth *et al.*, 2007; Suddendorf, 2003) or *symbolic* (Johnson *et al.*, 2005; Striano *et al.*, 2003; Roepstorff, 2008; A. Clark, 2006; Rakoczy *et al.*, 2005; DeLoache, 2004) styles of communicative signals.

When I have advanced this more integrative approach to material signals it is in pursuit of general phenomenological aspects of the recognition of the intended representational nature of material objects. The basic categorical question of the recognition of someone's intention that an object is perceptually approached as a source of social meaning seems neutral to the type of reference (iconic, indexical or symbolic) at hand. Another motivation is that one and the same material signal often offers itself for several possible and quite different meaning exploring interpretations, which may be lost in studies that are limited to a single perspective (cf. e.g. Myers & Libens, 2008, and *Article III* of this thesis for a discussion). Still, some limitations have been necessary to focus the study. These are explicated in the following.

I have in this project not considered aspects of artifact *design*. That is, the permanent shape given to an everyday object to complement its instrumental functionality with extra connotational value. I have thus only considered situations where communication can be considered the primary function of the contextualized configuration of material object. Nor have I incorporated examples that are fully dependent on extra contextual knowledge and inaccessible to a casual perceiver. A lot of our everyday use of objects for communicative purposes takes place under highly contextualized circumstances between informed initiates, as when ones partner leaves letters on the entrance hall table to remind us to mail them or when the unexpected *absence* of piles of dirty dishes in the kitchen sink is experienced as a communicative signal (Tylén, 2007). On the one hand, these kinds of situations are highly relevant instances of object mediated communication, but since they critically depend on prior explicit negotiations or attunement to implicit local (intra-family or otherwise sub-cultural) norms they are not suited to the kind of investigative manipulations (e.g. experiments) that I have advanced in this project. Importantly, this does not mean that the kind of cognitive issues that I have pursued in the various investigations making up this thesis do not generalize to these kinds of communicative situations as well.

A third type of non-verbal object mediated communication I will not consider is that involving maximally entrenched conventional symbolic artifacts, such as traffic lights and road signs, national symbols (e.g. flags) and religious symbols (e.g. crucifixes) in their standard use (i.e. when they are not integrated in new composite signals). Again these are related matters but still seem to have a slightly different cognitive profile. For instance it is not clear if such signals are comprehended by orientation to the intentions of a personalized addresser.

To summarize, 'material signals' is operationalized in this thesis as follows: *Material objects or properties of our environment that are intentionally manipulated by a person to function as* *a signal in order to make another person explore the object/properties as a source of representational social/semantic meaning.* And since the concrete investigations tend to be preoccupied with a receptive perspective on material signaling, the definition can even be further focused as: *Material objects or properties of our environment that are recognized as intended for the exploration of representational social/semantic meaning.*

Two central questions have structured the various investigations of material signals constituting this project. The first is a categorical question: (1) *Which cognitive processes subserve our recognition of the intended representational status of certain types of configurations of material objects*? The second is a semantic question: (2) *Which cognitive processes subserve our attribution of social/semantic meaning to these material signals*?

These questions are pursued in a range of multi-methodological and cross-disciplinary investigations from introspective conceptual analytical studies to hypothesis-driven experimental paradigms involving functional brain imaging. Overall, the investigative strategy can be framed as a version of Shaun Gallagher's 'front-loaded phenomenology' (2003). To front-load phenomenology in Gallagher's sense means to introspectively identify phenomenologically relevant distinctions and then operationalize them in the design of experimental paradigms that can test concrete hypotheses.

Roughly, this has been the structure of this investigation. In the early phases of the project I have identified distinctions relevant for our contextualized perception and experience of material objects in various configurations. These predictions have subsequently been formulated as concrete hypotheses that are tested in various experimental designs on samples of subjects. The general experimental strategy has been to record subjects' individual introspective categorizations of a stimulus set and subsequently analyze these with quantitative and qualitative models. The findings of each of these studies have again led to further discussions, elaborations and revisions of the conceptual starting point (cf. e.g. the discussion sections of *Article II* & *III*).

More concretely, the project comprises four very different studies reported in three articles: (1) When Agents Become Expressive: toward a theory of semiotic agency, (2) Say it with flowers! A fMRI study of object mediated communication, and (3) Taking the Language Stance in a Material World: a comprehension study, all of which are represented in the final part of the thesis. Throughout the introductory chapters of the thesis I will refer to these as *Article I*, *II* & *III* respectively.

## 1.1 The articles

*Article I* is a conceptual investigation that uses analysis of others' theories and introspection as the primary methodological approach. I introduce the notion of *semiotic agency* as an enactive framework for talking about language and communication. I pose a critique of Alan Leslie's developmental theory of agency that ignores an important distinction between two types of social encounters: third person *observational* encounters and first/second person *participatory* encounters. I use this distinction to discuss two ways of perceiving aspects of our material environment. We can either be 'passive' observers to material objects or we can recognize traces of someone's intention such that we take a 'participatory' interpretative stance and explore the object as a source of social meaning.

*Article II* takes up the idea of our recognition of material signals and reformulates it as specific hypotheses that can be experimentally tested. The first experiment reported in the article is a classification study where a population of 63 subjects evaluated a stimulus set of 152 photographs depicting naturally occurring scenes containing everyday material objects, some of which are manipulated in purposeful functional ways and some of which are not. In a questionnaire design subjects were asked to identify the material object configurations that they found to be social communicative signals. This allowed me to choose the best exemplars of material signals and the best controls (images consistently rated *not* to be signals) and use these as stimuli in a functional brain imaging experiment. Here a new population of 22 subjects was presented with the two conditions of images while their functional brain activity was recorded using magnetic resonance imaging technology (MRI). The idea was thus to investigate if the conceptual distinction (e.g. signal vs. non-signal) found to make a phenomenological difference to us also would make a difference to our biological brain (Gallagher, 2003). Moreover, the techniques enabled us to investigate which parts of the brain are involved in the recognition and comprehension of material signals. Though this localization strategy is not uncontroversial it might give us a hint to the cognitive foundations of our recognition of material signals by comparing our brain scanning results to other existing imaging investigations in human cognition, language, and communication.

In *Article III* I used another type of experimental design to address the issue of our recognition and comprehension of material signals. 20 subjects were presented to a subset of the images used in the previous study. These subjects were not however told anything about the purpose of the study but were asked to make short oral introspective descriptions of the scenes. These were subsequently coded by three experimenters in relation to a series of predefined parameters, such as the subjects' explicit reference to communication or social meaning. This allowed me to assess aspects of subjects' casual and naïve understanding of the scenes. Furthermore, while the brain imaging study was primarily preoccupied with the categorical contrast - signal vs. non-signal – the rich verbal data from the latter study allowed me to pose questions concerning the kind of signal-*meaning* subjects would attribute to various kinds of material scenes.

## 1.2 Structure of the thesis

In the following chapters of this thesis I will briefly introduce and discuss issues related to theories and methods used in my investigations. In the following chapter (*Chapter 2*) I outline some of the existing research on language, cognition and material signals that has had an impact on my project. This includes diverse fields of inquiry from behavioral studies of young children to aesthetic philosophy and brain imaging. Since my project does not commit itself to any particular discipline I will try to make apparent how inspirations from these various fields have informed my investigations both methodologically and in the interpretative framing of my findings. The third chapter (*Chapter 3*) discusses the different methodologies I have used in my investigations. My project is cross-disciplinary and several methodologies have been applied in pursuit of fundamentally the same research questions. This also allows me to discuss problems of cross-method validation which will be the one of the topics of *Chapter 5* that will also address issues of converging evidence across studies. *Chapter 4* presents additional comments to each of the three articles reporting from the investigations of the project. In some cases, my conceptions of the overall subject of material signals have gradually changed since the investigations, and in other cases, the style and genre of the report do not lend space for thorough elaborations of problems and findings. In *Chapter 6* I sketch out some future perspectives for the investigations in our comprehension of material signals and *Chapter 7* features the three articles. In addition, I have included an appendix with an English and a Danish summery, statements concerning multiple authorship, and the hundred stimulus images that are used in all experiments reported in *Articles II* & *III*.

# 2 Theoretical Foundations

The main interest of this project is the way material objects can come to be perceived as communicative signals in various kinds of human interactions. Though one could argue that the problem resides in the heart of human cultural cognition, no particular, well-established branch of scientific investigation is currently dedicated to this research question. Rather, the existing research addressing more or less central aspects of the problem is distributed between various disciplines, research traditions, and paradigms, often occupying a somewhat peripheral position. This has had some consequences for the organization of the present investigation. Since no obvious, unifying research tradition or methodology has offered itself as a rock-steady point of departure, the current investigation can not be classified under any particular branch of science. Rather, as symptomatic for many humanistic studies of human cognition, it can be said to inhabit a somewhat ambiguous position between cognitive language studies with an often introspective and qualitative style of inquiry, and psychology often characterized more by experimental and quantitative approaches (Núñez, 2007).

The project is informed by a long series of conceptual and empirical investigations from such diverse disciplines as aesthetic philosophy, developmental psychology, cognitive neuroscience, and psycholinguistics. In the following I attempt to give a brief overview of the various fields of research that have had an impact on my project and comment on some of the problems and possibilities that each of them pose for this particular field of interest.

## 2.1 Evidence from developmental psychology

Interestingly, the majority of empirical studies of non-linguistic communication appears to be conducted within a developmental paradigm. This might be due to the idea that these 'alternative' communication forms are most prominent in pre-linguistic stages of development and is later taken over by a more linguistic style of signification. Another reason could be the general belief that to capture the true cognitive nature of a phenomenon you need to study its ontogenetic emergence or cases of its deficits (cf. e.g. Bates & Tomasello, 2001). Though my investigations of material signals do not in any way address issues of cognitive and behavioral development, this branch of science has had a great impact on my project, among other reasons because this is where I have found the largest experimental literature.

Generally, the field of developmental psychology is characterized by carefully controlled behavioral studies which often include large populations of subjects and which use a quantitative style of data analysis. In the following I will briefly introduce two strands of developmental investigation that have had great impact on my project. The first concerns general aspects of communicative intentionality and ostensive cues in ontogeny, the other treats children's emerging understanding and use of material signals.

### 2.1.1 Joint attention, ostensive cues, and communicative intentions

An important source of inspiration, especially in the earlier stages of this project, comes from the field of developmental and comparative psychology and in particular the group surrounding Michael Tomasello at the Max Planck Institute of Evolutionary Anthropology. One of the main aims of this group has been to identify and investigate the special properties of human cognition that have facilitated our species unique and advanced social, cultural, and technological developments, including verbal language. In an extensive series of empirical developmental and comparative studies they have found that human children in even early stages of development show species unique motivations to engage in social activities and special abilities in recognizing the mental states of other people, an ability often referred to as 'theory of mind' (ToM) (See Tomasello *et al.*, 2005, for a discussion). Special attention has been given to children's emerging ability to recognize the intentions of others. Following Tomasello (1999a), human intentionality can be divided into two kinds: 1) simple intentions, and 2) complex communicative intentions. Simple intentions are primarily directed at objects in the environment and underlie for instance such instrumental goal directed behaviors as striving for food, protection and reproduction. It has a simple structural inventory of an agent and an object/patient with an inherently stable and unidirectional relation. To recognize the simple intention of another subject means to perceive his or her behavior as meaningful in respect to some project or goal. We recognize a certain direction in the behavior of others that makes us reason about the pursued (but yet unrealized) goal state and makes us capable of roughly predicting the course of the action (Tomasello *et al.*, 2005).

Complex communicative intentions on the other hand have a far more complex structure. They have in their internal structure an agent, a patient, and an object of shared attention. The agent/patient-roles are reversible and interchangeable which requires that both participating individuals are able to structurally represent (and shift between) both roles. The intentional action of the agent is directed at the patient with the goal that the patient should alter his intentional states with respect to a third entity, i.e. the object of shared interest. To understand complex intentions means to appreciate how an agent's goals and intentions are directed at another agent's psychological states, manipulating his/her attention with respect to an object of shared attention and/or joint activity. Ostention and joint attention is thus thought to be among the important socio-cognitive features enabling true symbolic interaction eventually leading to verbal language (Tomasello, 1999a; 1999b; 2003; Tomasello *et al.*, 1997; 2005).

An interesting twist to the joint attention story has very recently come from a team of Hungarian developmental psychologists with the leading figures György Gergely and Gergely Csibra. They find Tomasello and colleagues' mind-reading account too mentalistic and have in a series of studies pursued a slightly different idea that they term *natural pedagogy* (Csibra & Gergely, 2006; Gergely *et al.*, 2007; Senju *et al.*, 2008; Csibra & Volein, 2008). Their experimental findings can be roughly summarized as follows: Human infants are from birth adapted for the acquisition of cultural knowledge. This adaptation facilitates a special innate sensitivity to ostensive cues that signal adult caregivers' intentions to share relevant knowledge about the object world. According to this pedagogical approach to early triadic interactions infants learn *from* rather than *about* other minds. While Tomasello and colleagues find that infants learn about objects by attention to the adult caregiver's subjective mental states (intentions, emotions and dispositions), Csibra, Gergely and colleagues find the early triadic interactions to have a more epistemic function. They thus take an *object-centered* rather than a *person-centered* approach stating that children in these triadic interactions pick up on generalizable properties of referent objects and thereby acquire relevant cultural knowledge about the word rather than knowledge about the mental states of the individual adult caregiver (Csibra & Gergely, 2006; Gergely *et al.*, 2007). The key feature in this regard is ostensive cues (Sperber & Wilson, 1986). Thus, when an infant in a non-ostensive context observes an adult manipulating an object it might learn something about the person, for instance "my father likes broccoli" (i.e. *person-centered* representation). In contrast, when infants are addressed by ostensive cues such as eye-contact, motherese, or calling their name the same manipulation of an object is picked up by the child as relevant generalizable information about the object, for instance "broccoli is good!" (Gergely *et al.*, 2007). Csibra and Gergely hold that this kind of object-centered cultural learning may function efficiently without reference to complex mental inferential abilities and a full-blown theory-of-mind.

In all, Csibra and Gergely's alternative approach to object related communicative interactions introduces an important element that could be interpreted as the foundation of humans' general tendency to orient to social norms (though they do not themselves state this explicitly), differentiating local 'subjective' pragmatics and universal 'objective' semantics (Gergely *et al.*, 2007). This issue is somewhat downplayed in Tomasello's work, which instead focuses on long term conventionalization and entrenchment (1999a; 2003).

Though I generally appreciate the problems of applying concepts from developmental studies on adult competence, Tomasello and colleagues' distinction between simple subject-object oriented intentions and complex social or communicative intentions has played an important role both in the conceptual (*Article I*) and empirical parts of this project (*Article II* & *III*). For instance these types of distinctions were used to distinguish test and control conditions in the stimulus image corpus (see *Chapter 3* on methodology for details). Applied to the current object of interest the idea has been that a general sensitivity to intentionality might guide the inferential interpretation of various types of static material scenes. Compositional properties could thus either motivate an unintentional (natural or accidental), a simple intentional (a "private", instrumental) or a complex communicative understanding of the scene. In the latter case properties of the static material scene are comprehended in terms of an agent's intention that the perceiver adopt a special interpretative attitude when approaching the scene (see *Article III*). The latter point brings in a social normative perspective that could be argued to follow from Csibra and Gergely's critique. When an object is recognized to be intended for communicative exploration the question ceases to be about the specific mental states of the specific addresser-agent (i.e. person-centered representations). Instead, the object is explored as a source of 'public' semantic meaning that potentially may be detached from the specific ostensive context and brought into new ones (object-centered representations). This idea was introduced in the discussion section of *Article II* to suggest why the stimulus contrast 'communicative vs. non-communicative objects' did not elicit activation in areas of the brain normally associated with theory-of-mind and communicative intentions.

### 2.1.2 Children's understanding and use of material signals

A series of behavioral studies has addressed different aspects of young children's emergent comprehension of the communicative representational status of material objects. For instance, several studies have investigated children's emerging representational understanding of toy replicas when they are used to designate their real object or action counterpart (Tomasello *et al.*, 1999; Younger & Johnson, 2004; Johnson *et al.*, 2005). Similarly, a series of experiments have investigated children's understanding of the representational status of scale models, for instance when these are used to locate a target object in a real-size environment (DeLoache, 2000; 1996; Deloache *et al.*, 1997; Troseth *et al.*, 2007). Furthermore, a study has assessed children's comprehension of objects used as 'pointers' to find a target object (Tomasello *et al.*, 1997). Yet another related strand of experiments has investigated young children's emerging comprehension of the representational status of iconic depictions and photographs (cf. e.g. Ganea *et al.*, 2008; Simcock & DeLoache, 2006; Suddendorf, 2003).

Together these studies map the ontogenesis of children's early symbolic skills and understanding. This practice is thought to be related to, for instance, their ability to understand the social intentions of others (Tomasello *et al.*, 1997), and to hold *dual representations*, that is to simultaneously appreciate the physical object as such and its representational status as pointing to a reference object or action (see DeLoache, 2004, for a review).

Generally, these studies are of great interest to this project as they target the comprehension of the intended representational status of some material objects. They show how even young children come to engage in the cultural and cognitive practice of using material objects as sources of social meaning and tools for coordination of shared activities. There are however a number of important aspects of the experimental designs that make comparisons to the current project troublesome. I will address three issues that potentially disqualify such comparison, some of which can simultaneously be read as a critique of the studies.

The first issue is related to the developmental perspective of the investigations. The fact that the subjects are young children makes the findings difficult to generalize to an adult population.

Dual representations are for instance not known to be a cognitive challenge to competent adult symbol users (DeLoache, 2004).

The second issue has to do with the notion of symbolic reference operationalized in these studies. Without much explicit elaboration, the kind of symbolic tasks chosen for the experimental design in all these studies (except one of the conditions in Tomasello *et al.*, 1997) is modeled on iconic resemblance (Peirce, 1998; Deacon, 1997), which may in fact not be the best test case for the comprehension of symbolic reference (cf. Siple *et al.*, 1982 and Tolar *et al.*, 2007, for at discussion). Furthermore, in a large potion of the studies (Tomasello *et al.*, 1997; 1999; DeLoache, 2000; 1996; Deloache *et al.*, 1997; Troseth *et al.*, 2007; Simcock & DeLoache, 2006) iconic signs are used in the context of instruction (i.e. embedded in an *indexical* practice), either to request objects or activities or to indicate the location of hidden objects. Again, these are interesting issues, but far from uncontroversial. When a child fails the test in these kinds of experiments it is often obscure whether it is due to the misinterpretation of the iconic or indexical part of the message. In fact, a recent study Myers & Liben, 2008, point to the fact that iconic understanding and proper reaction to an instruction may be cognitively separate issues and findings in some studies mixing these matters may thus not reliably reflect children's (or adults' for that matter) more general symbolic abilities. Related to this is the fact that the only measured indicator of children's representational insights in these experiments is their observable behavior (i.e. performance in the tests), which may not reveal these subtleties.

A third point of relevance for present purposes is the fact that the experimental designs employed in the studies referred to largely rely on explicit instructions and ostensive gestural cues. The children were thus either explicitly instructed to use objects as symbols, or the object use had an ostensive gestural component. It is thus often obscure how much of the work was actually carried out by the objects and how much by explicit instruction and accompanying gestural and situational cues (cf. e.g. DeLoache *et al.*, 1999; Striano *et al.*, 2003), a point also relevant for some studies of adult material communication (cf. e.g. H. Clark, 2005). The experiments thus did not assess the central issue of my project: subjects' unaided teleological intuitions in regard to communicative objects, that is, how ostensive cues and communicative intentions are perceived as externalized in compositional features of inanimate, static material scenes.

### 2.1.3 Artifact categorization and children's historical intuitions

Another group of developmental experiments has been an equally important source of inspiration for the present project. These are concerned with the impact of children's historical intuitions on their categorization of artifacts (Bloom, 1996; Gutheil *et al.*, 2004; Gelman and Bloom, 2000; Gelman & Ebeling, 1998; Preissler & Bloom, 2008; Bloom & Markson, 1998; Kelemen, 1999). The hypothesis tested in these experiments was originally adapted from a series of essays on aesthetic philosophy by Jerrold Levinson (1979; 1989; 1993; see separate section below). While there has generally been some dispute in psychology about whether human subjects rely on form or function in their categorization of artifacts (cf. e.g. Landau *et al.*, 1998 and Miller & Johnson-Laird, 1976, respectively), this group of studies explored the alternative hypothesis that subjects orient to the history of the object and the original intentions of the creator in their categorization of artifacts. The argument is conducted in the following terms: When we categorize an object as a chair we might orient to its physical features. Some theories imagine an object category as consisting of its perceptual characteristics (e.g. shape and size) or even feature lists (cf. e.g. Fodor, 1975). But chairs come in a wide variety of forms from beanbag chairs to chairs that are suspended from the ceiling on chains, and this diversity of appearance makes form a doubtful critical criterion for artifact categorization (Bloom, 1996). Another cluster of theories has emphasized function as playing the decisive role. Artifacts are categorized in relation to their functional use (cf. Malt & Johnson, 1992 for a review). But a table does not become a chair because I sit on it and a miniature paper chair is still a chair though I certainly cannot sit on it (Bloom, 1996). Neither form nor function are thus alone sufficient criteria for the categorization of artifacts. Inspired by Levinson (1979, 1989, 1993) Bloom thus introduces the alternative hypothesis that artifacts are categorized in terms of the original intention of their creator. We categorize the miniature paper chair as a chair because we appreciate the creator's intention that it should be recognized as belonging to the category 'chair'. This hypothesized role of historicity and intentionality in relation to artifact categorization seems to solve some of the definitional problems in relation to form and function. It allows both for deviant forms such as the beanbag chair and 'non-functional' exemplars such as the paper chair, as both are still recognized to be successfully made with the intention that they belong to the category 'chair' (Bloom, 1996). The conceptual prediction has motivated a series of developmental experiments in which children's categorical intuition have been manipulated as a function of their orientation to object history and creator's intentionality (Diesendruck *et al.*, 2003; Gutheil *et al.*, 2004; Gelman and Bloom, 2000; Gelman & Ebeling, 1998; Preissler & Bloom, 2008; Bloom & Markson, 1998). The findings largely support the predictions. In all the referred experiments children consistently orient to available knowledge of object history and creator's intentions in guidance of their categorical resolutions.

Again the studies of children's historical intuitions have been of great interest and inspiration for my project. Although my interest is more in the direction of social and symbolic meaning construction than categorization and naming, the general idea that we orient to historical and intentional intuitions in our interpretation of artifacts is potentially of great explanatory value in relation to my experiments. Furthermore, the studies at least peripherally address some of the important issues about externalization and materialization of communicative intentions that were not treated by the other cluster of developmental studies referred to above. While in the instructional types of communicative situations addressed in the studies above the infant subjects had to recognize the ostensive intentions of the *present* object-manipulating communicator, in the studies by Bloom, Gelman and others the children had to use other material sources to make inferences about the intentions of *absent* agents/addressers. In this respect, these studies are very much in line with the kind of problems treated in my project. But there are also some issues that make Bloom *et al.*.'s studies less applicable to my work. In the following, I will address some of the most conspicuous.

A general problem in the studies of categorization and historical intuition is that they are ambiguous with respect to the important semiotic distinction between symbolic and non-symbolic (utilitarian) artifacts. Without explicit theoretical elaboration the same experimental paradigm is applied to the categorization of ordinary, non-symbolic artifacts (Diesendruck *et al.*, 2003; Gutheil *et al.*, 2004;

Gelman and Bloom, 2000; Gelman & Ebeling, 1998), pictorial representations (Bloom & Markson, 1998), and art (Preissler and Bloom, 2008). The consequence is that the type-token relationship of the category 'chair' is treated in the same manner as the representational relationship between for instance a picture and its referent. Thereby important differences in the implied intentionalities are blurred. It could seem that some of these ambiguities are due to the ignorance of the crucial distinction made by Levinson (most explicitly in the 1992 paper) between *categorical* and *semantic* intentions. I will return to this issue in a special section below (see section 2.3.2).

## 2.2 Philosophical Considerations

The primary ambition of this project has been to advance empirical cognitive investigations of our engagement with material signals. An alternative direction could potentially have been to address the issue from purely conceptual, philosophical and theoretical psychological perspectives. A project like that would probably have included discussions of material signals in relation to the theories of C. S. Pierce, J. J. Gibson, L. Vygotsky and N. Goodman among many others. No doubt, these are relevant and influential thinkers and their ideas deserve attention. Ideally, any empirical investigation should be based on such elaborated philosophical work when settling on proper definitions of conditions and experimental contrasts. But in practice, it might not always be possible to cope with it all. In the exacting process of designing, conducting and analyzing an experiment one may often find oneself settling on more coarse-grained and easily operationalizable concepts and definitions that primarily meet the practical challenges of data collection and analysis. The constitutive conceptual apparatus is thus often adapted from other experimental paradigms rather than from the philosophical source itself. This project is no exception. Still, a series of more philosophically oriented works have had a great inspirational impact on my project the most important of which will be discussed below.

### 2.2.1 The extended mind

Two recent programmatic works, Hutchins' *Cognition in the Wild* (1995), and Clark & Chalmers short manifesto "The Extended Mind" (1998), have been very influential lately in the discussion of the relation between human cognition and the material environment. The idea has been to study ways that our material surroundings integrate with, augment, and influence cognitive processes for solving everyday tasks. Such perspectives challenge standard internalist ideas of the mind as 'contained' within the skin and skull of an individual person. The alternative view proposed is termed *active externalism* (Clark & Chalmers, 1998). In the following brief introduction special emphasis will be put on the work of Andy Clark (1997; 2006) and Clark & Chalmers (1998) as they have been important sources of inspiration for my project, especially in relation to *Article I* and *III*.

A. Clark and Chalmers challenge the idea the cognition is a purely internal mental matter. They point to the many types of situation where humans employ aspects of their external material environment to facilitate mental processes and solve epistemic problems. Like certain instrumental artifacts like scissors of spades augment the capacities of our body enabling new manipulation of the material environment, other types of artifacts may augment the capacity of our mind enabling new types of mental processes (A. Clark, 1997). For instance, a simple multiplication task can be aided by the use of fingers, paper and pencil, or a bead frame, we can tie a piece of string around our finger to help ourselves recall an important appointment or we can employ external material symbol systems like diagrams, maps and flowcharts to facilitate complex and abstract thought processes (A. Clark, 2006; Roepstorff, 2008; Stjernfelt, 2007). In such cases the human organism seems to couple with the external material environment in complex, integrated, two-way interaction, in which some operations are performed by internal cognitive processes while others are carried out by the manipulation of an external material environment, and it may not be obvious where cognition stops and the material environment starts (Clark & Chalmers, 1998).

### 2.2.2 Language - the ultimate artifact

An especially interesting case of extended cognition is related to material symbols systems such as natural languages. While the primary function of language may be to facilitate interpersonal communication A. Clark argues that its implications for human culture and cognitions go far beyond that. Besides being a means for communication verbal language is an external cultural artifact that in ontogeny is internalized and maximally integrated in our cognitive apparatus (cf. A. Clark's notion of *the ultimate artifact*, 1997). Once there, language enables a totally new style of thinking. It does so by (re)structuring thought processes and guiding the attention and action of the language user (cf. e.g. Vygotsky's famous studies of children's self-directed speech, 1986).

The special "cognitive power" of language has to do with its material nature (A. Clark, 2006). Though we might be inclined to think of language in terms of its biological and mental ontology A. Clark stresses the cognitive impact of language's physical *expressions* whether these come in the form of spoken, signed or written words (or other types of material symbols for that matter). Disputing the widespread assumption that language's material expressions work by mechanisms of full translation into some kind of content-capturing inner code (cf. e.g. Fodor, 1998), A. Clark proposes a *complementary* view of language processing. This entails that language is not only encoded by translation to some kind of mentalese, but that the material structures of the symbol tokens themselves (or image-like representations thereof) continues to impact thought in an irreducible and complementary fashion facilitating a special kind of "hybrid thought". Rather than thinking in an abstract internal language of representation people think through the manipulation of concrete material symbols (Roepstorff, 2008; A. Clark, 2006). Illustrative examples can be found in the history of mathematics, where for instance the replacement of Roman by Hindu-Arabic numerals (Middle Ages) or the later invention of the comma or period as a decimal separator have played a tremendous role for the development of mathematical thinking as these new material representational units allow for new and more efficient styles of number manipulations (Østergaard, 2006).

The idea at offer could seem to be very productive for the kind of cognitive issues I am pursuing in my project. When people approach everyday objects as sources of social coordination and meaning construction, the aspect of materiality comes to play an important cognitive role. In some cases, meaning is assigned to material objects in a totally arbitrary fashion by way of cultural convention (c.f. conventional symbolic artifacts such as traffic lights), which might put more weight on what A. Clark terms a translational style of representation. But in the far majority of cases of material signals investigated in this project, other representational strategies seem crucial. When a casual observer is confronted with novel and unconventional styles of material signals s/he is to a larger degree dependent on the material structure itself as the source of meaning construction. This entails an exploration of more iconic/diagrammatic and indexical motivations of meaning (Peirce 1998; Stjernfelt, 2007), where concrete structural properties of objects come to function as the main scaffolding. Aspects of such different meaning exploring stances are discussed in *Article III*.

But there are some fundamental issues that make the theories of distributed cognition referred to less applicable to my work. The proclaimed objective in A. Clark's approach to language is to ignore its communicative functions to investigate how its material nature impacts thought (1997; 2006). Thereby language becomes a tool for the individual mind – a relationship between an individual person and an object. Generally, A. Clark does appreciate communication as a primary function of and an evolutionary motivation for language (A. Clark, 1997). But he argues that once language was already in place it enabled other kinds of manipulations leading to new styles of symbolic thinking as a kind of byproduct. Language in A. Clark's terms thus has a double nature as 1) "*a medium for communication*" and 2) "*a tool that alters the nature of the computational task involved in various kinds of problem solving*" (A. Clark, 1997:193). It could seem that these two functions of language are unrelated and work by different mechanisms, and by separating them A. Clark fails to discuss the distributed and complementary character of language as a means for communication (see a related critique in Spurrett & Cowley, 2004).

I am generally very sympathetic to A. Clark's idea about the cognitive powers of the complementary nature of language, but I find it hard to see the motivation for A. Clark's inclination to separate '*language for communication'* and '*language for thought*' as two distinctive phenomena. On the contrary, I would claim that all the cognitive aspects of language related to its materiality discussed by A. Clark play an equally central role when language is used as a medium for communication. But as evident from the kind of applications I have advanced (briefly summarized above) there are complications to the story. In the kinds of communicative situations I address in my studies the material object is no longer just a tool for the individual mind. Rather, it comes to mediate between two or more minds as a tool for *social coordination* and *collective thinking*. As a consequence, the "extended mind" (Clark & Chalmers, 1998) engaged in these cognitive activities is not satisfactorily described in terms of coupled systems distributed between an agent and an object, but involves several agents that must attune and align their meaning exploring attitudes in complex ways. For an individual mind to successfully hook up to this system it will need to adjust both to the special structural properties of the object affording cognitive exploration and to orient to intended or socially defined normative strategies for doing so. Consider for example DeLoache's studies of children's comprehension of scale models (Deloache, 2000; Deloache *et al.*, 1997), or Myers and Liben's of study of children's comprehension of cartographic symbols (2008), both mentioned in an earlier section. In these studies it is evident that the same material object (e.g. a miniature model or a schematic depiction of a scene) can potentially afford several types of explorations. For the child to use the object in the intended representational way, for instance as a map to find a hidden toy, it needs to understand this instructional context and use it to regulate the exploration of the object by attunement to the special structural properties that may accommodate the rewarded purpose. Once the child gets a grasp at the manipulative cognitive procedure it may be able to internalize it in a Vygotskian fashion (1986) and later on use objects as maps in private problem solving activities as described by A. Clark (2006). The examples are, though, thought to illustrate one of the ways that our seemingly individual approach to material symbols are often explicitly or implicitly scaffolded and regulated by the intentions of other people and orientation to social norms.

The findings reported in *Article III* of this project support this emphasis on the impact of social context on distributed cognition and the complementary nature of language. The pronounced alignment of subjects' representational approach to specific object configurations recognized as signals is interpreted as a strong indication that these explorations are not a matter of subjects' private problem solving attitudes, but rather point to the way subjects consistently pick up on and attune to any cue that could guide their meaning exploring efforts in relation to interpretative norms and addressers' intentions.

## 2.3 Historical Intuitions

Under this headline I will briefly discuss two conceptually grounded theories that both, though in radically different terms, concern the way that material structures cognitively point to their causal past. Both theories have been important sources of inspiration for my project as they may help to explain how we can have meaningful experiences of novel and unconventional object manipulations. The first theory finds expression in Michael Leyton's extensive work *Symmetry, Causality, Mind* (1992).

### 2.3.1 Shape, causality and cognition

Leyton's point of departure is a cognitive rethinking of geometry with the introduction of time and causality as crucial parameters. But once the argument is settled Leyton generalizes the fundamental cognitive relationship between shape and time to a broad scale of everyday human phenomena, from the comprehension of pictorial arts to the syntactical structures of language. In the following, I will leave out mathematical details and assumptions and focus on the more conceptual cognitive aspects of Leyton's idea that has been applicable to my project, most explicitly in relation to *Article I*.

The fundamental idea in Leyton's theory is that we perceive object shapes as holding information about their causal past. In Leyton's words 'shape is time'. A given shape (say a rectangle) is thus perceived not as a fixed representation comprising a series of static relations, but as a dynamic form condensing the memory of the series of causal events and forces that broad it about. And to recognize or understand the shape is to recover the special process-history that it has undergone. Our representations of objects are thus an example of how, in Leyton's view, our cognition more generally is inextricably linked to time and causality. In his words, *"a cognitive representation is a causal history hypothesized to explain a stimulus"* (Leyton 1992:156). Leyton's concept of process-history and memory is strongly related to the abstract dichotomy *symmetry* vs. *asymmetry*. While the perfect symmetry is memory-less, asymmetry *"is the memory that processes leave on objects*" (Leyton, 1992:7). A perceived asymmetry is in other words understood as a past symmetry that has undergone a series of manipulations. Leyton mentions scars on the surface of the moon, chips on vases and graffiti on subway trains as concrete examples of asymmetries that tell us a story of past causal events. Symmetries on the other hand tell no such stories. If the graffiti is successfully cleaned off the subway train, a casual perceiver would subsequently not be able to recover the process-history of the train from its mere appearance. Leyton goes on to describe an extensive repertoire of processes that can leave traces on objects and that can be recovered in our process-recovering cognitive approach to objects. I will leave that part out here to briefly comment on possible implications of Leyton's ideas for my work and a couple of problems that show up in that respect.

Leyton's theory elegantly shows how we cognitively come to infer conceptual dynamic meaning (such as forces and causes) from non-conceptual static shape (e.g. curvatures). The idea that part of the understanding of a present object shape is to recover the history of causes and forces that have influenced it seems to be very productive for the kinds of subject-object encounters that I have addressed in my project. When subjects in the various experiments reported in *Article II* and *III* recognize novel and puzzling static object configurations as intentional communicative signals it is probably because they are able to recover the story of a series of purposeful, deliberate acts that caused this scene. Accordingly, Leyton's concepts of symmetry vs. asymmetry bear relations to the control vs. test conditions of my classification experiments. The control stimulus scenes primarily depict everyday objects in canonical, natural or utilitarian settings. These scenes are symmetrical in the sense that they do not bear traces of any manipulative acts or conspicuous causal events. They are thus *memory-less* in Leyton's sense. Test stimulus scenes on the other hand primarily comprise objects that have been intentionally manipulated in often uncanonical ways. They are either introduced in new strange contexts, central properties (e.g. the color) have been changed or they have been arranged in new puzzling configurations. They are thus material *memory stores* in the sense that aspects of their form point to the acts and events that have preceded and caused the scenes.

When, despite these seemingly nice correspondences, I have not adapted Leyton's theory of shape, causality and cognition as the main frame of reference for my later studies (*Article II* & *III*) it is due to some unresolved issues concerning 1) the relationship between the symmetry/asymmetry distinction in mathematics and geometry vs. the real world, and 2) the theoretical framework's apparent insensitivity to various kinds of causes and motivations (e.g. accidental vs. intentional).

As stated above, Leyton grounds his theory in mathematics and geometry. This allows him to hypothesize how, by systematically recovering the sometimes long series of protrusions, indentations, squashings, internal resistances or bifurcations that constitute their process-history, any shape can be traced back to a couple of *absolute* symmetrical forms such as the circle or the square (see Leyton, 1992:chapter 1). When applied to geometrical shapes the idea has a lot of intuitive power.

It also seems applicable to systematic investigations and descriptions of shapes, especially in abstract art (see also Leyton, 2006), architecture, and some fields of medicine (e.g. in the case of describing the growth or reduction of tumors), among others. In contrast, some of the alternative applications proposed by Leyton, especially in relation to everyday visual perception, are more troublesome. These are the cases when Leyton seemingly without any explicit elaboration confuses two very different definitions of symmetry, an absolute geometrical and a more pragmatic epistemic one. Compare for instance how Leyton in one paragraph describes the shape of a human hand as a series of curvatures, protrusions and indentations apparently bearing relationship with an abstract geometrical circle (1992:30-32), and in another asks us to appreciate how an abandoned crumpled newspaper on a subway station represents a history of past crumpling in relation to our knowledge of the default shape of newspapers (1992:1). In the first example the shape described only bears a very abstract relation to time and causality. In this absolute geometrical reading, a hand missing a couple of fingers is more symmetrical than an ordinary hand. In contrast, in the latter case a crumpled newspaper is an asymmetrical version of a normal newspaper, though its ball-like shape may actually be more symmetrical in the geometrical sense after the manipulation. Here (like in the graffiti example above) we are apparently supposed to think in terms of real world causes and events. A symmetrical object is now the object as we expect to find it, while the asymmetrical version is one that bears traces of causal (unexpected) manipulation. If we apply this more pragmatic understanding of symmetry (as defined by our knowledge and expectations to the world) to the hand-example, a hand missing two fingers is certainly more *a*symmetrical than a normal hand (as we would expect to find it), pointing to the possible past events that could have caused their partition. This substantial ambiguity is not satisfactorily treated by Leyton. And in relation to the kind of issues that I am concerned with in my studies this constitutes a serious problem. I will now point to a couple of examples.

In the stimulus corpus used in the experiment of this project (see the section 9.5 of the appendix) we encounter numerous examples of both geometrical and pragmatic symmetries and asymmetries. But curiously, geometrical symmetry seems to be one of the strategies frequently used by communicating agents to make objects become signals. By arranging everyday objects in a symmetrical shape (in Leyton's terms a memory-less figure) they create unexpected pragmatic asymmetries, that is, new shapes that point to a series of past manipulative acts. Examples include a bunch of clamps put together in a symmetrical composition to form a sculpture, beer caps arranged in a perfect spiral shape to create an aesthetic effect, and tomato slices systematically distributed in an ordered pattern on the front pane of a car perhaps to make a practical joke. Ironically, in these cases (among several others) geometrical symmetry is thus exactly recognized as 'memory stores', pointing to the careful and subtle manipulation of an addressing agent, while the matching asymmetrical control scenes are experienced as random, unintentional, and thus memory-less.

Another point of critique relevant at this point is that Leyton's otherwise clever and fine-grained analyses do not distinguish between the various kinds of motivations that can bring about causal events. I can thus not use his typology of shape-manipulating forces to establish if a specific manipulation was due to for instance accidental or intentional events. This distinction is crucial to my project. In the first experimental studies reported in *Article II* I thus asked subjects to classify a corpus of stimulus scenes in relation to four types of causal events: natural events, accidental events, instrumental intentional events and intentional communicative event. In other words, the task was not only to establish the transitional process-history of a given object or scene but in addition to recognize which kind of underlying motivation had been causally involved.

While Leyton's hypothesized intimate relationship between shape, causality and cognition has been of general explanatory value and an important source of inspiration for the project his proposed framework for more detailed analysis turned out to not to be well suited to my present purposes. My sense is that his theories will work more effectively in fields where pure geometrical issues are at stake.

### 2.3.2 Categorization, semantics and creators intentions

Another theory related to the impact of historical intuitions on cognitive aspects of perception comes from the field of aesthetic philosophy and art criticism, and finds expression especially in the works of Jerrold Levinson (1979, 1989, 1993), Mark Rollins (2004), and Paisley Livingston (2005). The following short introduction will primarily be concerned with three essays on aesthetic philosophy by Jerrold Levinson entitled "Defining Art Historically" (1979), "Redefining Art Historically" (1989) and "Extending Art Historically" (1993).

The kind of ideas put forth by Levinson could seem only of peripheral interest to this project as they exclusively discuss the categorization and comprehension of art. But as mentioned in a previous section, Levinson's ideas have in fact been of major impact outside this particular field. In his 1996 article "Intention, history and artifact concepts" Paul Bloom generalized aspects of Levinson's ideas to the broader field of artifact categorization and in the following decade the hypotheses were pursued in an extensive series of experimental and developmental studies (Bloom, 1996; Gutheil *et al.*, 2004; Gelman and Bloom, 2000; Gelman & Ebeling, 1998; Preissler & Bloom, 2008; Bloom & Markson, 1998; Kelemen, 1999) most of which are discussed in more detail above.

Like Leyton, Levinson is concerned with the way that inferences about the causal past of an object (in this case a work of art) influence our present categorization and conception of it. Contrary to Leyton, though, Levinson does not discuss this on the level of structural analysis, but advocates for an *internal historicism* focusing on the intentional motivational forces that can cause things to happen and how we orient to them in our exploration of the material world. More concretely, Levinson asks the question: what constitutes the definition (and the cognitive category) of art? Which kind of constitutive properties can be said to be shared between the classic antique Greek sculpture *Aphrodite of Melos* (c. 150 BC), Dostoevsky's novel *Crime and Punishment* (1866) and Bjørn Nørgaard's piece of performance art, *The Horse Sacrifice* (1970), where Danish artists did a ritual dissection of a horse (*my examples*)? It is clear that when these quite different types of phenomena all happen to be categorized as *art* it is not due to the sharing of any intrinsic features. Alternatively, Levinson suggests a solution that takes as its point of departure the artist's original intention, that is, what the object is *intended for*, or with which *kind of regard* the spectator is expected to approach the object (Levinson, 1979, 1993). This grounding of the concept of art inevitably bounds up with history. The artist thus wants the surroundings to appreciate his/her creations in a similar regard to that "normally" applied to works of art. It is thus the conscious and intentional purpose of the artist to create an object to be appreciated in some kind of continuity with preceding pieces of art that is the primary criterion for its subsequent categorization. The unifying property constituting various tokens belonging to the concept of art is thus neither similarity of form, nor affordance for a particular style of experience (different works of art afford for quite different types of experiences), but recognition of the artist's intention that this object is made for *regard-as-a-work-of-art* (Levinson, 1979).

While this part of the theory concerns the artist's intentions that an object is categorized and accordingly approached in a particular manner, hence *categorical intentions*, Levinson introduces another concept, *semantic intentions*, designating the intended meaning of the art work (Levinson, 1992, 1993; Rollins, 2004). An artist can for instance create an object intending it to be conceived of as a sculpture which would constitute the categorical intention, and could then intend the sculpture to symbolize the oppression of the working class, which would be its semantic intention (Levinson, 1993). When we recognize an object as intended as a work of art, we thus pitch our attentional system toward a special kind of aesthetic exploration, searching properties of its structural compositionality as cues to the semantic intentions of the artist. In other words the intentions of the artist manifested in the art work come to have a regulative influence on our interpretation (Rollins, 2004). Levinson suggests that the two types of intentions (categorical and semantic) might not be of the same nature. While categorical intentions are always the intentions of the actual artist responsible for the artistic creation, semantic intentions can be hypothetical or virtual. That is, the meaning of an object categorized as a work of art can be defined not necessarily in relation to *the* artist who actually created the object, but *an* (generic) artist employing the artistic conceptions and devises of a particular time and place (Levinson, 1996). The latter point might be interpreted in the direction, that even the "languages of art" are structured by an orientation to some kind of social normativity, manifested both in the production and reception of artistic meaning.

Levinson's ideas have had a tremendous impact on my project. Though he is exclusively occupied with the interpretation of art, nothing in his theory suggests that the same kind of cognitive processes should not apply to other kinds of object mediated communications (including indexical and symbolic), a fact also reflected in Bloom's alternative applications (see separate section above). Of special importance is Levinson's distinction between categorical and semantic intentions. In *Article I* it is reflected in the comparable though maybe not fully compatible idea that a signal simultaneously indexically points to the original communicative intention of its addresser (categorical intentions) and symbolically to its meaning (semantic intentions). In *Article II*, the distinction is used again to explain patterns of brain activation elicited by material signals. While bilateral activations of structures in the fusiform gyrus were associated with categorical resolutions (signal vs. non-signal), corresponding to the recognition of categorical intentions, other bilateral activations in prefrontal areas (including Broca's area), were thought to express more fine-grained semantic explorations corresponding to Levinson's recognition of semantic intentions. In *Article III*, Levinson's distinction is further elaborated to accommodate the broader scope of material signals (and not only art works) in general. As material signals can be intended for either aesthetic iconic, instructional/indexical, and symbolic exploration I thus split Levinson's concept of categorical intention in two: (1) a general signal vs. non-signal, and (2) a distinction between three representational meaning constructing strategies, structured by the communicative use of Pierce's sign types *icon*, *index* and *symbol*.

Furthermore, Levinson and Rollins' ideas about the regulatory role of artists' intentions on our interpretation of art have been a great source of inspiration for my project, most explicitly reflected in the conclusions of *Article III*. Applied to the broader field of material signals, the force of their proposition is that it grants asylum for a theory of "monological communication" that is not based on old fashioned Sausurean transmission metaphors (Pickering & Garrod, 2004). Since many contemporary functional cognitive accounts of communication tend to be preoccupied with a dialogical (conversational) style of interaction, the question of a social cognitive grounding of monologue have not yet received the same scientific attention (cf. section 2.5.3 for a discussion). In the case of art, Levinson and colleagues suggest that despite the physical absence of a responsible addresser-creator, his or her communicative intentionality is still at stake, externalized and manifested in the structural compositionality of the object. By this external intentional intervention the artist is not "conveying" ready-made content by some kind of arbitrary code, but is implicitly instructing or regulating the perceptual exploration and interpretative processes of the addressee in relation to the object. In other words the categorical and semantic intentions of the artist only scaffold the active meaning constructing work of the perceiver.

While the suggestion appears to be very promising and potentially has a lot of explanatory value to the kind of material scenes I am investigating in my project, there are also shortcomings. Levinson's theory is primarily a philosophically elaborated idea and is not to my knowledge accompanied by concrete analyses or at least illustrative examples. It is thus not entirely clear which structural properties of an art work that come to guide the interpretations of the spectator and to what extent. While Leyton's idea about cognitive process-history recoveries from structural shape has the drawback that it does not systematically differentiate between intentional and unintentional causes, Levinson's theory focuses on this latter point but unfortunately tends to miss the detailed structural analyses.

## 2.4 Material signals in the brain

Since brain imaging takes a central position in the experimental part of my project (see *Article II*) I now turn to the various theories about the brain that have structured these investigations.

Many of the ideas about the socio-cognitive foundation of various kinds of communicative mediations introduced in the previous sections are also reflected in the rapidly growing field of cognitive neuroscience. A still increasing number of published functional brain imaging studies address fundamental questions about the biology of social interaction, communication and language. With the gradually more widespread propagation of imaging techniques such as functional magnetic resonance imaging (fMRI), brain imaging is becoming an ever more integrated part of the field of cognitive science, and findings are becoming increasingly referred to and cited in other branches of science, including the humanities. It can be argued whether functional brain imaging is really mature for this kind of general impact, and there are still many reasons to be skeptical (cf. e.g. Uttal, 2001). With the enhanced sophistication of measurements and new standards for sample sizes and subject populations, the paradigm is itself struggling to stamp out some of the initial technical and conceptual difficulties (see also section 3.3.5 for a discussion). Personally, I believe that brain imaging can be a valuable tool for testing motivated hypotheses about human cognitive processes when used in a sensible and careful way. Most importantly, no imaging study should stand alone. Ideally, brain data should always be systematically related to and supported by other types of experimental measurements such as behavioral data. Only to the extent that we can find mutually supportive results across several types of measurements or investigative methodologies targeting the same cognitive phenomena can brain imaging data be granted validity.

No other study has yet used brain imaging technologies to address issues of our use and comprehension of material signals, and since the subject tends to occupy a perfectly ambiguous position somewhere between language studies, cognitive semiotics, and social psychology, it is not completely obvious how it should be framed in relation to existing brain imaging studies. There exist however a number of studies that have investigated related aspects of "theory of mind", social cognition, and non-verbal communication. In the following, I will briefly introduce some of the functional studies and anatomical theories that have played a role in the design and analysis of my fMRI experiment. Moreover, I will make an attempt to discuss how my findings can be used to suggest an alternative way of thinking about social cognition and the brain.

### 2.4.1 Mentalizing and the brain

A by now well-established paradigm in cognitive neuroscience is sometimes treated under the somewhat presumptuous headline "the social brain". It more or less implicitly defines 'social' in term of our species unique abilities to 'mentalize', that is, to make inferences about the mental states of others. In an extensive series of studies of 'theory of mind' and 'mentalizing' (cf. e.g. Castelli at al, 2000; Walter *et al.*. 2004; Gallagher *et al.*, 2000, 2002; Gallagher & Frith, 2003, 2004; Amodio & Frith, 2006; Fletcher *et al.*, 1995; Kampe *et al.*, 2003; Saxe, 2006) it has been suggested that a distinct pattern of brain areas including (among others) the temporal poles bilaterally, the right temporo-parietal junction (TPJ), the right superior temporal sulcus (STS) and the medial prefrontal cortex (MPFC) are involved in explaining and predicting the behavior of other people regardless of the sensory modality or task involved.

I will now focus on some of the studies that seem especially relevant for present purposes as they are preoccupied with our recognition of communicative intentions. For instance, in a fMRI experiment Walter *et al.*. (2004) used cartoon strips depicting various scenes containing two people either engaged in each their private affair (e.g. hanging up laundry and digging the garden) or in a social activity (e.g. one asking the other to point out directions on a map). They found that the anterior paracingulate cortex (part of MPFC) was only sensitive to situations containing socially directed intentions, but was not activated by private intentions. A second study revealed furthermore that even a single person's intentional activities directed at anticipated social interaction (e.g. setting the table for a romantic candle light dinner) could activate this neural structure. The authors thus conclude that the anterior paracingulate cortex is the key to our recognition of socially directed intentions.

Another relevant study used fMRI to investigate the neural underpinnings of our recognition of *ostensive signals*, that is, signals that express an agent's intent to make contact with another agent with the purpose to communicate or interact (Kampe *et al.*, 2003). In everyday interactions such ostensive signals are often made by calling someone's name or making eye contact. Again, in Kampe *et al.*.'s study the experimenters found activations in areas of the mentalizing network, more specifically the anterior paracingulate cortex and the temporal poles bilaterally, when subjects heard someone calling their own name in contrast to another name or when they saw someone directing their gaze directly at them vs. not making eye contact.

Ostensive signals could be argued to play a central role in my project. In *Article I* the recognition of communicative intentions is staged as the central problem in relation to material signals: how do we come to recognize a static, mute, material scene as purposefully addressing us? And in *Article II* & *III* the categorical 'signal vs. non-signal' contrast could easily be argued to be a matter of recognizing 'externalized' ostensive cues. Accordingly, in the early preparatory phases of my fMRI experiment, I had strong expectations that this particular network of brain areas would be activated. Since these brain areas are shown to be sensitive to future anticipated interactions (Walter *et al.*, 2004), they could be imagined to be equally sensitive to the 'externalized communicative' intentions that are at stake in my experiments. I return to this issue in the discussion below.

### 2.4.2 Non-verbal communication in the brain

The studies of mentalizing (the social brain) do not directly treat issues of communication, but rather some of its hypothesized prerequisites in terms of our more general abilities to e.g. recognize the mental states of other people. Other studies have however investigated how the brain responses to different kinds of communicative mediations. The great majority of these are dedicated to aspects of verbal language, from phonology to syntax (cf. e.g. Price, 2000; Rodd *et al.*, 2005; Hickok *et al.*, 2007; Papathanassiou *et al.*, 2000; Vigneau *et al.*, 2006; Wagner *et al.*, 2001; Thompson-Schill *et al.*, 1997; Bedny *et al.*, 2007; Goldberg *et al.*, 2007; Wallentin *et al.*, 2005, 2006, among many others), but recently there has been an enhanced interest in the neural foundations for non-verbal communication as well (cf. e.g. Lotze *et al.*, 2006; Skipper *et al.*, 2007; Dietrich *et al.*, 2007; Lawrence *et al.*, 2006). These studies, which for instance target communicative gestures, body language and non-verbal semiotic vocalizations, could be of great interest for my study as another instance of non-verbal communication. In the following I will briefly introduce a series of studies that seem especially relevant as they address different modalities of non-verbal communicative mediation in contrast to non-communication. These studies are as such comparable to my own. Interestingly, the brain areas involved in these studies of non-verbal communication do not consistently overlap with the areas found in studies of mentalizing as might have been expected. Rather, they gather in areas traditionally associated with language and semantics.

In an fMRI experiment, Lotze *et al.* (2006) contrasted the neural responses to three different kinds of right hand movements: isolated hand movements (e.g. twisting a lid), body-referred movements (brushing teeth) and expressive gestures (e.g. holding up the index finger to threat someone). While the first two are not intended as communicative signals the third mediates social communicative meaning. In the body-referred and the expressive movement conditions the experimenters found activity in the STS of the social brain network. This area has previously been associated with the perception of another person's action (Gallagher & Frith, 2004). But, when the expressive gesture condition was contrasted with the body-referred movement condition only the left ventro-lateral prefrontal cortex (VLPRC)/Brodmann Area (BA) 47, was active. Interestingly, this area has previously been associated with semantic processing in verbal language (cf. e.g. Dapretto & Bookheimer, 1999; Fiez, 1997).

In another imaging study of brain regions associated with the perception of non-verbal vocalizations Dietrich *et al.* (2007) found a similar pattern. They were interested in the fact that in addition to words in a natural language people make use of other types of vocalizations to express social meaning. For instance, affective bursts like laughter or cries have an intrinsic (indexical) social communicative function in signaling other individuals' affective states, but in addition vegetative sounds such and belchs and yawns are sometimes produced deliberately to signal arbitrary (symbolic) communicative meanings. This fact was manipulated in Dietrich *et al.*.'s experiment. While being scanned subjects were instructed to classify different types of vocalizations including affective bursts (e.g. laughter) and vegetative sounds (e.g. snoring or belching) in terms of their presumed function: communicative signals vs. non-signals.

As in the previous study, we could have expected the task in this study to involve a lot of mentalizing. It seems to be about the recognition of communicative intentions. But when the experimenters contrast vocal signals with non-signals none of the mentalizing areas are elicited. Rather the hemodynamic responses reside in areas of the left hemisphere temporo-parietal junction, proposed to be an verbal 'auditory-to-meaning interface' in previous studies (cf. e.g. Hickok & Poeppel, 2000), together with some structures of the inferior frontal gyrus (left BA 47 and right BA 44 and 45). These are all areas consistently found in studies of verbal language and semantics. Besides, note that one of the prefrontal components (left BA 47) overlaps with the findings of Lotze *et al.*. (2006).

A third study of interest in this connection addressed aspects of 'body language'; ways we direct motor acts at other people as signals recruiting "shared representations" (Lawrence *et al.*., 2006). The experimenters used an adapted version of the Profile of Non-Verbal Sensitivity or PONS test (Rosenthal *et al.*., 1979) in a fMRI experiment assessing subjects' understanding of non-verbal communication. In the scanner subjects saw a series of short video clips showing an actor conveying various types of facial and bodily expressions and gestures, and were asked to rate them on either a non-social (face or body?) or a social (angry or happy?) decision task. The experimenters interpreted subjects' behavioral responses (button presses) to the social condition in terms of their ability to empathize. Empathy is also normally considered part of our 'theory of mind' (Lawrence *et al.*, 2006), yet in the main contrast between *social perception* and *non-social perception,* none of the regions supposed to be part of the 'social brain' is activated. Rather, similar to the previous examples the primary main effects are found in the dorso-lateral prefrontal cortex (DLPFC/BA 46) and inferior frontal gyrus (BA 44) including Broca's area bilaterally (among a couple of other sites). The authors interpreted these patterns of activation in terms of an orientation to shared representations. The same areas are however part of a common network involved in various aspects of language processing (cf. e.g. Hagoort, 2005; Price, 2000).

### 2.4.3 The tale of two social brains

All studies introduced above (Lotze *et al.*, 2006; Dietrich *et al.*, 2007 and Lawrence *et al.*, 2006) are concerned with a special type of social encounter; situations where bodily movements, facial expressions, vocal sounds and hand gestures are recognized as signals mediating social meaning. While from a conceptual point of view, 'theory of mind' and mentalizing should be central components in interactions related to for instance the recognition of communicative intent (Tomasello, 1999a; Tomasello *et al.*, 2005), areas of the social brain network seem to play a somewhat marginal role in these studies. Instead, we find activation in a series of brain areas traditionally associated with language and semantics. This curious fact is reproduced in my own fMRI study (see *Article II*). While based on the studies of e.g. Walter *et al.* (2004) and Kampe *et al.* (2003), I would have expected 'mentalizing'-activations due to the recognition of communicative intentions, these were conspicuous in their absence. But there might be good (post hoc) explanations for that. In the following I will attempt a framework that can potentially explain the kind of puzzling patterns of brain activity related to various kinds of social phenomena, eventually linking up to some of my initial ideas and criticisms expressed in *Article I.* The result will be a tale of not one but two social brains – or networks – that in complementary ways subserve our engagement with the surrounding social environment.

The notion of 'theory of mind' that in a number of studies has been adapted for brain imaging experiments is often defined in terms of a series of distinct inferential and predictive abilities in relation to other minds (see e.g. Leslie, 1987, 1993; Frith & Frith, 2003). The most commonly listed abilities belonging to the 'theory of mind' paradigm are *empathy* (the ability to identify with another person's affective states), the ability to understand and dissociate other agents' *beliefs* from ones own, the ability to read the *intentions* of other agents, and the ability to understand *pretence* and *deceit*. In practice, an agent's theory-of-mind-abilities are often reduced to her performance in specific psychological tests, such as the false belief test or pretence play, etc. (cf. Leslie, 1993). It is however, or so I will argue, not uncontroversial if these various types of mind reading should be considered in terms of conceptual and developmental continua. One important distinction seems to be blurred in the traditional psychological literature and is likewise imported in cognitive neuroscience. In *Article I,* I thus propose to distinguish between two types of social encounters: (1) the third person observational encounter, and (2) the first/second person participatory encounter (cf. also Di Paolo *et al.*, 2008; De Jaegher & Di Paolo, 2007; Gergely *et al.*, 2007). Though, both can be argued to have an element of mentalizing there are great differences in the kind of mental attributions the perceiver makes and their role in these two types of encounters. Accordingly, we might expect somehow dissociated neural pathways to be involved.

The first type is at stake in the vast majority of psychological and neuropsychological studies of theory of mind and social cognition in general. A lot of studies have addressed issues of 'social perception' such as our recognition of human faces (see Kanwisher & Yovel, 2006, for a review), and biological motion (see Vaina *et al.*, 2001, for a review) in contrast to non-social stimuli. Others have been concerned with the way that we explain and predict other peoples' behavior by making inferences about their state of mind, such as in false belief tests (Frith & Frith, 2005; Fletcher *et al.*, 1995) or for instance Walter *et al.*.'s study on the recognition of socially directed intentions (2005). Two traits are characteristic of these kinds of experimental paradigms: (1) they feature one agent – often the subject - passively observing the behavior of another agent that is unaware of and unresponsive to the observer, and (2) the kind of 'meaning' that the observing agent attributes to the stimuli is of an *indexical* sort, in that it establishes likely mental causes and explanatory and predictive patterns for the observed behavior.

In contrast, a few recent studies have taken a more enactive or interactive approach to social cognition. In these studies the subject it not just a passive observer, but is part of an ongoing interaction as an addressee for intentional sense-making. Thereby behavioral cues are not only read as indices of the other person's state of mind, but rather as mediations of joint meaning construction (notice the correspondence to Csibra and Gegerly's concepts of person-centered vs. object-centered representations, Csibra & Gergely, 2006; Gergely *et al.*, 2007; Senju *et al.*, 2008; Csibra & Volein, 2008). It is not so much about the contents of another mind, but about the shared representational meaning that is jointly negotiated somewhere between minds (Di Paolo *et al.*, 2008; De Jaegher & Di Paolo, 2007). This description could seem to fit at least one of the tasks traditionally considered part of the "theory of mind battery", that is, the recognition of pretence. Contrary to false belief, pretence requires the agent is aware of and responsive to her observer. We do not just pretend – we pretend for someone. Pretence is thus an interactive endeavor, a semiotic activity that calls for an active participatory style of perception and meaning construction (see *Article I,* for further analyses).

So far, no social cognitive neuroscience studies have systematically manipulated the difference between third person observational and a first/second person participatory social encounters. But a cautious comparative meta-analyses of neurocognitive theory of mind studies and the studies of various types of non-verbal communicative mediation (Lotze *et al.*, 2006; Dietrich *et al.*, 2007; Lawrence *et al.*, 2006) suggests that dissociative neural networks may subserve the two strands of social cognition. Though I realize that I might twist Csibra and Gegerly's notions of *person-* vs. *object-centered representations* (introduced in an earlier section) beyond their preference, these seem useful in this context to capture the substantial differences between the two networks. The first is thus a *person-centered system* responsible for our inferential explanation and prediction of the behavior of other agents (Frith & Frith, 2006). It comprises the network of areas such as the rTPJ, the STS and MPFC sketched out in an earlier section. This network allows us to trace the dispositions of other minds and eventually to recognize possible agents with whom we can interact. It may as well subserve our 'indexical' recognition of ostensive cues (cf. *Article I* and Kampe *et al.*, 2003). The second is an *object-centered system*, primarily concerned with social meaning construction. The system comprises a network of areas responsible for our attribution of representational social and semantic meanings to perceivable aspects of the external environment, such as for instance verbal language structure, gestures and material signals. While this is still somewhat presumptuous, I hypothesize that the core of this network is the inferior frontal gyrus (IFG), especially left hemisphere BA 44, 45, 47 and 6 comprising the “unification area” proposed by Paul Hagoort (cf. e.g. Hagoort, 2005; see also Müller & Basho, 2004, for an interesting discussion). Other areas possibly include Werniche’s area, the left TPJ, and part of the fusiform gyrus.

The proposed ‘dual model’ of social cognition is supported by my own findings reported in *Article II.* One could have imagined the material signals depicted in the stimulus images of this study to have elicited neural activations in traditional theory of mind areas due to the recognition of communicative intentions manifested in the external media. Instead, material signals, like e.g. the studies on gestures and vocalizations, were found to tap into areas of the other network associated with language and social meaning.

Of course the idea of two distinct social networks is still quite speculative and more research is needed for instance to establish the precise role of the language network in relation to social interaction and the complementary relationship between the two “social brains” – the person-centered and the object-(or meaning)-centered. Ideally both networks should be activated both separately and in orchestra in a single experimental design to investigate the conceptual borders. This study is currently on the sketchpad.

## 2.5 Language, communication and semiotics

As should be evident by now, this project is not firmly grounded in a single and well-established theoretical framework, but has rather explores a broad range of different research fields to establish what they each can contribute to the discussion of the cognitive nature of material signals. Still, coming from the language sciences and semiotics, I myself see this project first and foremost as a study of human language and communication, and this stance tends to pervade the conceptual parts of all three articles. On the other hand I of course realize that this ‘language perspective’ applied on strange material object configurations may not be obvious to the casual reader and may even provoke some parts of the linguistic society. In the following, I will thus briefly introduce some of the ideas on human language and semiotic interaction that have inspired my work and furthermore make an attempt to explore the borderlines of the proposed analogy between material signals and verbal language.

### 2.5.1 Language as an activity

In the introductory parts of his comprehensive work *Using Language* (1996), Herbert H. Clark discusses two quite different perspectives on language: *language-as-an-object* vs. *language-as-an-activity*. Most traditional structural, functional and generative paradigms within linguistics (more or less informed by the Saussurean tradition) have been preoccupied with a structural description of the ‘object of language’. That is, they pursue a description of the various ‘units’ of speech and writing and their interrelationships in a ‘language system’. These kinds of studies have given us valuable insight into the for instance the structure of phonology, morphology, semantics and syntax. However, H. Clark invites us to appreciate the alternative perspective, that is, what we *do* with language. He thus studies language in the context of its use as a form of joint action and a means for social coordination and negotiation of meaning. This approach also offers new perspectives on the relationship between verbal language and other styles of (non-verbal) mediation. While structural *language-as-an-object* investigations will inevitably lead to a focus on the differences between verbal language and non-linguistic means of communication, the *language-as-an-activity*-approach highlights ways that verbal and non-verbal signals often function fully interchangeably or integrate and complement each other in various types of composite signals (H. Clark, 1996). H. Clark illustrates this point by listing a series of everyday situations where a hand gesture, a nod or the manipulation of an object (e.g. waving a flag) take the function of a communicative signals that could just as well have been carried out in a verbal style of mediation. Thereby he challenges the presumption that nonlinguistic signals are "*crude, unsystematic, ad hoc, and marginal, and deserve to be relegated to the periphery of language use*". On the contrary, H. Clark argues, "*the non-linguistic methods* [of signaling] *are subtle, highly systematic, and not at all ad hoc. (...) Ignoring nonlinguistic methods has distorted people's picture of language use, and it is important to put that picture right.*" (H. Clark, 1996:156).

To launch a theory of language use that can deal with various modes of linguistic and nonlinguistic communicative actions from spoken and written words to gestures, material objects and even pauses, H. Clark introduces the general notion of *signals*. In short, a signal is "*an act by which one person means something to another*" (H. Clark, 1996:155). Signals are thus neutral to the linguistic/nonlinguistic distinction that H. Clark reduces to a matter of '*methods of signaling*'. H. Clark grounds his concept of signals in semiotics and more specifically on Charles S. Peirce's theory of signs. While I will not go into great length about Peirce's theory of signs, it should be noted that it can be somewhat ambiguous in relation to intentional language and communication (H. Clark, 1996). H. Clark thus proposes we see signals as the deliberate employment of signs for purposes of communicative interaction. Peirce found that signs can represent their meaning by three types of relationship: *iconic*, *indexical* or *symbolic* reference (Peirce, 1998; see also Deacon, 1997, for a neurocognitive approach to sign types). Roughly, icons represent their reference object by perceptual resemblance. Indices on the other hand, represent their reference objects by a dynamic causal or spatial relationship. Symbols represent their objects by conventional rules. Symbols are thus intentional by nature while icons and indices can also describe non-intentional representational situations, that is, instances of 'natural meaning' (Grice, 1957). When these sign types are employed in intentional communicative interactions they come to characterize three different representational relationships between a communicative signal and its meaning. H. Clark proposes we phrase these types of intentional relationships as follows: demonstrating a thing (icon), indicating a thing (index), and describing as a type of thing (symbol). Verbal language is normally treated as a symbolic system – that is as system of signs that represents their meanings by conventional rules. But in a *language-as-an-activity*-perspective, H. Clark states, an utterance is not just a string of symbols. It largely depends on iconic demonstration and indexical indication as well and often several sign types work in orchestra, as in composite signals (H. Clark, 1996).

The kind of semiotic multi-modal approach to language use launched by H. Clark seems able to deal with some of the central issues in human linguistic interaction that is often ungraspable to more structural approaches to language. Other influential strands of linguistic research have pursued the idea that verbal language constitutes an isolated self-enclosed system, and even occupies a separate module in the brain (cf. e.g. the chomskian paradigm, Chomsky, 1972; Pinker, 1994). From theories like these necessarily follows the prediction that linguistic behavior is somehow detached from non-linguistic communicative behavior, leading to an often troublesome distinction between language and communication. Rather than dissociating various methods of signaling (e.g. verbal language from gesture, etc.) H. Clark's account opens up for an exploration of the functional similarities and interactions between quite different expressive means from words to gestures, bodily actions and material signals. A broad range of recent findings in language acquisition research (e.g. Tomasello, 2003; Spurrett & Cowley, 2004), developmental science (e.g., Deloache, 2004) and brain imaging ((Lotze *et al.*, 2006; Dietrich *et al.*, 2007; Lawrence *et al.*, 2006) suggest that this kind of multimodal interactional approach to language has more psychological validity.

### 2.5.2 Material signals

In his article "Coordinating with each other in a material world" (2005) H. Clark develops this broad scope on language use in relation to material signals. In qualitative microanalyses of everyday joint actions and some experimental situations he identifies two main types of signal activities related to material objects: *directing-to* and *placing-for*. 'Directing-to' is a person's acts and activities oriented at other persons to direct their attention to objects, places, or events. Examples of 'directing-to' acts include *pointing to*, *touching, exhibiting*, or *poising* an object. Common to all these types of nonlinguistic 'directing-to' communicative acts is that they require addresser and addressee to coordinate their attention and intentions in relation to a shared perceptual space to establish the referent (H. Clark, 2005). Another kind of 'directing-to' activity is when persons direct an addressee's attention to themselves and their actions as a source of social meaning or to coordinate activities. Among the signals we can employ to direct other peoples' attention to ourselves is H. Clark's notion of *manifested action*. Manifesting an action is to perform it in a "*time, place,* or *manner intended to be recognized as marked or special.*" (H. Clark, 2005:514). By for instance *exaggerating* an everyday canonical act we direct other peoples' attention to it and thereby make it into a communicative signal for others to interpret.

The second kind of material signaling activity proposed by H. Clark is 'placing-for'. In 'placing-for' a person places an object or action at a special site for an addressee to interpret or give appropriate response to. When objects are placed for addressees at specific sites it is not only to make other people attend to the objects. The site is chosen to motivate a specific kind of interpretations of the object or act. Placing an object at a supermarket checkout counter will frame the object as an 'item to be bought', while placing it at a museum stand will, at least potentially, frame it as an 'item to be aesthetically explored' (*the latter is my example*). 'Placing-for' thus has much in common with 'directing-to' in that various joint intentional manipulations of objects will induce them with additional signal qualities and affordances (H. Clark, 2005).

H. Clark's theory of language and material signal use has had a great impact on my work and is used as the primary theoretical framing of my studies reported in *Article II & III.* The multi-modal and process-oriented view of human signaling practices seems to fit hand in glove with the kind of cognitive semiotic approach I have advanced in my project. Though not directly applicable, his ideas about '*manifested action*' and '*placing-for*' potentially have great explanatory value in relation to the problem of how static, mute object configuration become perceived as dynamic communicative signals.

It should, though, be added at this point that H. Clark is mostly concerned with various types of online conversation and dialogue dynamics. Even in his study of material signals he investigates situations where two or more people engage in joint activities involving material objects that they either point to or place for each other in signaling ways. His concept of material signals thus does not fully coincide with mine as he does not consider situations where the static material scene is actually doing most of the job. By 'material' H. Clark thus means not only objects but likewise gestures and acts. This could be formulated as a critique of H. Clark paper. It is actually not always obvious which role the material object plays in his material signals. When someone points to an object for instance to direct a recipient's attention to it, what is then the signal 'meaning' of that object? Is it not really the gesture that is meaningful while the object stays the same? At least it is not totally clear in many of H. Clark's examples if the objects themselves actually gain new referential status as signals or if the signal meaning is associated with the gestural activities alone. In contrast, I have only investigated monological material signals where the addresser is not present but has arranged the material scene in an intentional way that calls for a meaning exploring receptive attitude. In these kinds of situations it is thus not the ostensive and gestural acts of a present addresser, but the compositionality of the material objects alone that guide the meaning constructing activities of the addressee.

Like many of the other theories presented in earlier sections, H. Clark's theory of material signals has proven a useful and valuable source of inspiration, but as evident from the critique above it turns out to be not fully compatible with the kind of communicative situations I am addressing in this project, and thus it cannot alone function as the theoretical grounding of the project.

### 2.5.3 Enactment, monologues and dialogues

As apparent from all articles (*I-III*), in my investigation of our comprehension of material signals I have tended to take an interactive, communicative approach. Though the kinds of situations I have studied only feature a single agent (the perceiver-addressee) and a static material scene, my claim has been that the addresser is still somehow cognitively present in the receptive situation due to certain compositional features of the material scene. Obviously, though, these signaling situations are not online conversations or dialogues with continuous feedback, turn taking and joint action sequencing (Pickering & Garrod, 2004). They are monologues with quite fixed addresser-addressee roles and thus could rather be said to resemble written style texts. The challenge has been to describe in adequate ways how these kinds of signaling activities still can be approached as a kind of social encounter with some element of interaction. One attempt has already been made in the section *Material signal in the brain* above reflecting ideas from *Article I*, where I contrast third-person (person-centered) observational encounters and first-second person (object-centered) participatory encounters. But here I will continue the discussion relating my ideas to recent theories of language, meaning construction and interaction.

Lately, a lot of theories have been launched that aim to ground language and communication in human intersubjectivity and social cognition. One example is H. Clark's ideas expressed in *Using Language* (1996) introduced above. Other general theories of language and communication that have inspired this project include works such as Pickering & Garrod's "Toward a Mechanistic Psychology of Dialogue" (2004) and De Jaegher and Di Paolo's "Participatory Sense-Making: an enactive approach to social cognition" (2007) and "Horizons of the Enactive Mind: values, social interaction, and play" (2008). Common to these various approaches is that they aim to leave behind old conduit and transmission metaphors of communication characteristic of traditional structural linguistics and semiotics (cf. e.g. Jakobson 1960 and Saussure, 1974), and reframe communication in the context of joint action where two or more participants cooperatively negotiate the social meaning of utterances in orientation to their shared perceptual and embodied world. In this framework meaning is thus not an object "transmitted" from one mind to the other through some kind of arbitrary "code" that is subsequently "translated" into internal representations (Kravchenko, 2007; De Jaegher & Di Paolo, 2007; see also A. Clark, 2006). Rather, meaning is a process in which two agents coordinate and align their representations through continuous interaction (Pickering & Garrod, 2004). From more radical versions of this approach it follows that language meaning is not primarily denotational, but connotational; no meaning is context-free and every contextualization is unique. Rather than abstract symbolic conventions, language is tightly grounded in concrete instances of use linking it in indexical ways to human coordinative actions and activities (Cowley, 2004; Kravchenko, 2007). While I am generally sympathetic to these approaches, which also tend to pervade my own investigations and theories of material signals, I also sometimes worry that they will end in the same kind of pitfalls that have characterized more traditional accounts of communication. Generally, the field of more traditional linguistic paradigms has been strongly affected by an implicit 'written language bias' (Linell, 2005). The intense occupation with written style language has resulted in general theories of language that are inherently monological and individualistic in scope. According to Pickering and Garrod (2004) such theories cannot cope with the more natural and basic forms of language, that are inherently social and dialogical. Like H. Clark (1996) and others they thus launch alternative theories of language based on online social interaction and dialogical dynamics.

A problem related to my work is that the alternative social and enactive theories of language are often exclusively based on studies of conversations, dialogs and other online face-to-face communicative situations. As such they cannot function as a general theory of language applying to both online dialogue and offline monological styles of language, but only solve half of the problem. Being interested in offline monological material signals I thus easily get stuck using old fashioned conduit metaphors of language and communication. This is most explicitly the case in *Article I*, where I tend to end up in a messy Cartesian based transmission model of language and communication that for instance imply a quite passive addressee-receiver role. In *Article III,* I make an attempt to solve the issue by showing how addressee-subjects' individual meaning exploration is regulated by orientation to shared normative frames and addresser's manifested intentions. The meaning of the monological material signals is thus not "transmitted" to the perceiver. Rather aspects of the composition of the scenes are picked up as cues proposing the proper strategy for perceptual exploration and meaning attribution. To account for this phenomenon I propose to adopt a version of Roepstorff & Frith (2004) notion of 'top-top interaction'. In an article on different types of experimental designs in neurocognitive studies they discuss how experimenters' pre-experimental instruction to subjects can be seen as facilitating an alignment of representational strategies. Often the experimenter not only instructs the subject to attend to a certain stimulus but in addition to do it in a certain fashion, taking on a certain attitude to the situation or task at hand. In this way, experimenter and subject come to align and share common top-down scripts or frames for the proper understanding of the situation, a practice putatively termed 'top-top interaction' (Roepstorff & Frith, 2004). There is apparently nothing to suggest that this orientation to shared top-down scripts and frames should not apply more generally to human social cognition (Frith *et al.*, 2008) including the comprehension of material signals.

# 3 Methodological Considerations

The main strategy characterizing the investigations of this project has been to approach a single overall research question (our comprehension of material signals) from a series of complementary methodological perspectives. The project is thus inter-disciplinary in nature and investigations are based on multiple qualitative and quantitative methodologies. This reflects an important ground for the investigation: the aim has been to attain new knowledge of a particular human practice and any methodological question has been secondary to that aim. No particular method was privileged in the outset of the project, and the initial research questions were not formulated to fit a particular research procedure. Thus, the inclusion and application of any new method of investigation has been motivated by the desire to gain new perspectives on the initial research question.

This multi-methodological approach has had several advantages. Each new method applied has contributed with new interesting facets of the overall research object. Ambiguous results obtained in an experiment using one method were explained in another experiment based on other methods. And potentially important findings that turned out to be opaque to one set of measures were picked up by other measures. I will discuss concrete examples of such cross-method variability of results in greater detail in a separate chapter (see *Chapter 5*).

There are, though, also some disadvantages to a multi-methodological approach. It is generally quite time-consuming to familiarize oneself with the procedures of multiple methodologies. Each of them is embedded within extensive research traditions and large literatures. They each have their advantages and pitfalls that have to be taken into account. A lot of time has been invested in the acquisition of basic tools and procedures such as how to program stimulus presentation in MatLab, how to statistically measure inter-coder reliability in content analysis and how to analyze brain-imaging data. Some of the time spent on acquisition of new research methods, tools and procedures might have been used in more investigations of my research object had I settled on a single methodology. But as will become evident in the discussion of cross-method variability and converging and complementary evidence, the time and effort invested in the multi-methodological approach seems justified.

The methodological challenge in relation to all the various investigations constituting this project has been to address the issue of our (inter)subjective experience of material signals in a systematic, controlled and generalizable fashion. Since we cannot access peoples' subjective experiences in any direct way, we are left with different indirect measures. In very general terms, there are three different kinds of indirect measurable indices for subjective experiences. First, we can observe peoples' behaviors and infer their 'contents of mind' from defined behavioral correlates. Second, we can do physiological and neurological measurements and hope that they somehow index subjects' states of mind. Finally, a third possibility is simply to ask people and trust that their meta-reports reflect their actual subjective experiences (Jack & Roepstorff, 2003). I have chosen to model most of my investigations on various kinds of introspectively derived categories, whether my own (*Article I*) or my subjects' (*Article II* & *III*). As such, the overall methodological strategy of the project can be argued to follow the lines of a "phenomenologically enlightened experimental science" proposed by e.g. Gallagher (2003) and Jack & Roepstorff (2002, 2003). The point of departure is thus a conceptually and phenomenologically motivated idea elaborated in *Article I*. This is then further elaborated and turned into concrete hypotheses that are tested in a series of experimental paradigms reported in *Article II* and *III*.

In the following, I will briefly introduce the main methods used and reflect over some of the choices, implications and problems. An overview of the methodological architecture of the investigations across the three articles is represented below:

Table 1: schematic representation of the methodological structure of the investigation across studies

| **Investigation** | **Method** | **Report** |
|---|---|---|
| Conceptual investigations | Experimenter's theoretical conceptual analyses based on introspection | *Article I* |
| Stimuli selection | Experimenter's introspective analyses and classifications | *Article II* |
| Stimuli validation | Controlled elicitation of subjects' introspective classifications | *Article II* |
| fMRI brain imaging experiment | Subjects naïve phenomenological experiences combined with their a posteriori controlled classifications and quantitative statistical analyses | *Article II* |
| Categorical content analysis | Subjects naïve phenomenological reports combined with experimenters' a posteriori controlled classifications and quantitative statistical analyses | *Article III* |
| Semantic content analysis | Experimenter's introspectively motivated semantic analyses | *Article III* |

As evident from Table 1 above, phenomenology and introspection are used in two quite different ways in this project: (1) as a method for the development and elaboration of conceptual and theoretical frameworks and formulation of hypotheses, and (2) as a way of empirically testing predefined concepts and hypotheses. These qualitatively different uses will be treated in separate sections below.

## 3.1 Theoretical and conceptual analyses

One of the central methods of this investigation belongs to battery of more traditional humanistic research methods, e.g. qualitative and conceptual investigations based primarily on introspection and rational analyses of others' theoretical work. The conjunction of theoretical analysis and introspection can be regarded as the golden methodological standard within most language sciences (Talmy, 2007), though its subjective foundation is not uncontroversial (cf. e.g. Gibbs, 2007; Schwitzgebel, 2008). While the status of pure rational, conceptual 'evidence' itself can sometimes be disputed, it is important to remember that no empirical investigation can do without conceptual grounds. Any experimental design or analysis of data rests on some kind of theoretical and conceptual assumption that guides the systematic data treatment, whether this is implicit or explicitly stated.

In the present project, these more theoretical methods have been especially important in the early hypothesis-generating stages of the project, represented in *Article I*. One of the basic procedures of this type of theoretical work is to contrast theoretical ideas from various domains and see how they complement or reconstruct each other. In the present study, the point of departure was the theoretical approach to language as an activity rather than a formal system of signs (cf. e.g. H. Clark, 1996). In itself this is not controversial. However, in *Article I*, I used the *language-as-an-activity* idea to challenge other theories of our understanding of intentionality, causation and language. The outcome was the notion of 'semiotic agency' that potentially could confound some of the existing developmental studies of agency. Furthermore, it offered specific testable predictions in relation to the cognitive processes implied in our comprehension of material symbols.

## 3.2 Empirical methods

Generally, empirical investigations of human cognitive function can be based on three types of data: introspective, naturalistic/corpus or experimental data, and each has its advantages and limitations. All types of data were used to various degrees and in combination in the present investigation. In the following, I will briefly comment on some of the important aspects of each method in turn, before returning to a more detailed account of the methodological empirical procedures used in the present investigation.

### 3.2.1 Introspection

Introspection has been widely used in studies of the structure of language (Talmy, 2007), and may be an especially valuable method for hypothesis-generation (see above). Though I generally acknowledge the potential of this research method, its foundation on the purely subjective experiences of a single experimenter makes it a potentially unreliable method for the verification of many types of research questions, and should probably not stand alone (cf. e.g. Schwitzgebel, 2008). Still, when implemented in a more controlled experimental setting, it can be a valuable (and even necessary) tool for assessing subjects' otherwise inaccessible phenomenological experiences (Jack & Roepstorff, 2002, 2003). Introspection has thus been a central component in both the fMRI brain imaging study (*Article II*) and the content analysis-study reported in *Article III*. Here multiple subjects' phenomenological experiences of a stimulus set constituted the "raw data" for more controlled hypothesis-driven analyses (see below for details). In the empirical part of the project introspection did not constitute the experimenters' method of investigation, but was used as a mechanism for the generation of experimental data.

### 3.2.2 Naturalistic/Corpus approaches:

Lately within the language sciences, there has been a growing interest for studies of naturalistic data (cf. corpus studies or usage-based approaches, Tummers *et al.*, 2005; Tomasello, 2000, not to mention Conversation Analyses and Discourse Studies). The term 'naturalistic data' is here used to designate data that is collected in a non-interventional way from real world ecological environments. Naturalistic data are often materialized in text or image corpora or video/audio recordings of naturally occurring events. Data are preferably recorded in a non-elicited way without the subjects being aware of it, which ensures that their behavior/language is not affected by the experimental intervention. The great advantage of this kind of data is that it allows researchers to study their object of interest in close connection to the actual environmental context and functional use. This is thought to lend more ecological validity to the data (Tummers *et al.*, 2005). There are however also limitations to the study of naturalistic data corpora. In the following I will concentrate on a couple of problems special to the kind of questions that I want to address in this project.

One way to assess the comprehension of material communicative objects could potentially be to collect data from peoples' actual, spontaneous and non-elicited encounters with such scenes, e.g. in the form of video recordings. A practical issue here is to get a sufficient amount of comparable samples to enable meaningful quantitative analyses. This is preferred in order to allow general theorizing about our comprehension of material signals. The non-interventional and uncontrolled style of enquiry inevitably leads to great variability in the data that challenge stringent comparative approaches. A related problem is that data would have to be recorded at multiple sites. Thus, the subjects that would be involved in an encounter with one object configuration would probably differ from those involved in another causing too many independent variables (degrees of freedom) in the analysis. Yet another problem concerns the assessment of the data. It is very problematic from purely behavioral observations to infer any cognitive or mental profile of subjects' comprehension (Tummers *et al.*, 2005; Jack & Roepstorff, 2002). Though I generally favor any aim at ecological validity including collections of naturalistic data, these methods seem inappropriate to answer the questions I have posed in this project concerning our comprehension of material signals.

Still, a kind of corpus approach has been involved in the experimental part of the project. The hundred stimulus images used in all experiments are photographs of actually occurring scenes, collected from the experimenters' local environment and from web-based image databases. No images were artificially constructed for the experiments. This stimulus set thus constitutes a corpus of naturalistic data in its own right (see the separate section on stimulus collection below). I of course acknowledge the fact that the images were carefully selected in relation to a series of predefined premises and thus do not reliably reflect the distributive nature of actual object use which probably disqualifies it as a 'real corpus' (Tummers *et al.*, 2005). It is worth noting, however, that many researchers do not even orient to naturally occurring data in their selection of stimuli for experiments, but construct this material themselves, thereby seriously jeopardizing the ecological validity of their experiments.

### 3.2.3 Experimental approaches

In the empirical parts of this project I have primarily relied on various experimental methods of data collection. The main advantage of the experiment is that it allows you to isolate or elicit specific online data in a controlled environment (Talmy, 2007; Tummers *et al.*, 2005). In this way you can collect a large amount of relevant measurements from a large population of subjects in a relatively short time. In addition, experiments allow you to assess such online phenomena as subjects' perception and comprehension of a stimulus in a systematic and carefully controlled way. Potential disadvantages include the risk that such explicitly elicited and interventional data collection creates strong bias effects and thus does not reliably reflect subjects' actual patterns of cognitive behavior as they would occur in real-world contextualized situations. Furthermore, the fact that discrete components of e.g. perception and comprehension are often treated in isolation may blur important interaction effects (Talmy, 2007). Although these kinds of problems are to some degree inevitable, in the present studies I have endeavored to take at least some measures to avoid such serious confounds. One of these precautionary measures has been to essentially repeat the same experimental question in a series of different experiments employing quite different methods of data collection, and then evaluate the cross-method reliability. This issue is discussed in the last part of this chapter.

## 3.3 The organization of empirical investigations in the project:

The theoretical, conceptual analyses characterizing the initial phase of this project (cf. *Article I*) were reformulated in the second part into more concrete predictions/hypotheses that could be tested in controlled experimental ways.

In all the experiments of this study, I have exposed a sample of subjects to a series of stimulus images depicting scenes with various configurations of objects to assess the subjects' comprehension of these scenes using different kinds of measurement. On a very general plane, the experiments can be characterized and compared in relation to two parameters: explicitly vs. implicitly elicited classification designs and quantitative vs. qualitative analyses.

### 3.3.1 Explicitly vs. implicitly elicited classification designs

In the following, I will use the terms explicit vs. implicit elicitation to categorize different kinds of experimental designs in terms of the degree to which the overall purpose and focus of an experiment is disclosed to the enrolled subjects prior to the experiment. An explicit elicitation can be thought of as an experimental design that places regulating constraints on the cognitive performance of the subject in the experiment. If an experiment is designed in a way that limits the possible behavioral repertoire of a subject (as in e.g. 'multiple choice' or 'forced choice' tasks) it will create strongly predictable patterns of behavioral output (e.g. the answer 'yes' or 'no'). Such explicitly elicited designs are advantageous if you want to address very subtle distinctions in a stimulus. By cautious instructions, masking of stimulus and carefully designed response tasks you can focus the attention of the subject on precisely the issue of research interest, leaving out most of the interfering noise constituted by e.g. subjects' free associations and distractions. A serious drawback to this approach, however, is that you cannot be certain whether the kind of distinction you ask subjects to make reliably reflects their casual, spontaneous perception of the same stimulus in an ecological environment. Subjects' knowledge of the purpose of the experiment and wish to please the experimenters can potentially influence his/her performance and thus be a considerable confounding factor.

Another possibility is to do a more implicit elicitation of classification. Here you keep the overall purpose and focus of the experiment undisclosed to the subject to prevent influences on his/her performance and preferences. The task is also made more open and unregulated. In this way, you can enhance the 'naturalness' of the experimental situation and potentially get a more realistic image of the cognitive implications of the subject-stimulus encounter (Gallagher, 2003). Often the subjects in these designs do not know they are classifying the stimuli, as the classifications are derived from the subjects' behavioral or verbal responses in post-experimental analyses (cf. e.g. the section on *content analysis* for an example). A great problem in these kinds of experimental designs is of course the enhanced amount of noise and variability in the data that challenge stringent and systematic analyses. When subjects' attention is not explicitly drawn to a particular focus of experimental interest, you will often find a great 'spread' in responses across subjects.

In the empirical part of this project I used a combination of explicit and implicit classification experiments to accommodate their individual pitfalls. The two classification studies reported in *Article II* (and partly described in *Article III*), were thus of the explicit kind. In these studies subjects received careful instruction concerning the purpose of the study and subsequently performed an explicit task with a very limited repertoire of behavioral responses (of the multiple choice kind). In contrast, the main experimental task of the brain imaging study (*Article II*) and the content analysis study (*Article III*) can be characterized as implicit in the present use of the term. In these studies subjects' knowledge of the theoretical focus of the experiment was kept at a minimum and their responses were not guided or regulated by explicit measures and questions. The complementary approaches thus allowed me both to address very specific questions concerning subjects' comprehension and to study aspects of their more casual and spontaneous reaction to the stimuli.

### 3.3.2 Qualitative vs. quantitative approaches

While the distinctive pair ‘explicit’ vs. ‘implicit’ classification is relevant in the description of the different experimental designs, another distinction becomes relevant for the characterization of the analytic approach to the experimental data, i.e. qualitative vs. quantitative analyses. Both types of analytic approaches have been used to various degrees and in combination in this project and I will in the following shortly comment on some relevant advantages and limitations of both approaches in this context.

Generally, the quantitative approach is hypothesis-driven and aims at precise classifications and statistical verification. The experimental design and investigative methods are often carefully planned and investigations are thus primarily deductive. A quantitative approach requires that collected data can be reduced to numerical indices. Among the great advantages is the generalizability of results and the formalized and stringent approaches to the question of validity. Limitations include a general insensitivity to contextual factors and an incapability to deal with at lot of more subjectively grounded questions.

In contrast, the qualitative approach is explorative in nature and not necessarily guided by a priori hypotheses. The interpretative design often emerges ad hoc and is generally more inductive. The qualitative analysis aims at a complete, detailed and contextualized description of data. Its advantages include the openness to new unanticipated directions in the analysis and capability to pick up some of the rich and unquantifiable nature of data. Qualitative results are, though, often less generalizable and the subjective style of inquiry has often led to discussions of validity (cf. e.g. Winter, 2000).

In the empirical part of this project I have faced the methodological challenge of wanting to address peoples’ subjective, phenomenological experience in a systematic and generalizable way. The aspired objective has been to capture very general cognitive trends in subjects’ comprehension of material communicative signals. While the phenomenological aspects of the research focus thus call for more qualitative styles of inquiry, such as e.g. subjects’ introspective reports, the ambition to state something very general about subjects’ dealings with material symbols calls for a quantitative, statistically based approach.

The strategy has thus been to combine these two approaches in various ways. In the brain imaging experiment (*Article II*), quantitative methods were used in the collection and analyses of brain data. But the statistics were modeled on subjects’ introspective responses to the stimuli (cf. Jack & Roepstorff, 2002). Similarly, in the content analysis experiment (*Article III*) subjects’ introspective reports constituted the raw data, which was subsequently quantified for statistical analyses (cf. Neuendorf, 2002). Moreover, part of the analysis in this study was based on a more explorative, qualitative approach to capture in depth some of the subtleties that were opaque to the quantitative approach.

### 3.3.3 The collection and classification of stimuli

A central component of all experiments in this project is the corpus of stimulus images (see the appendix). All experiments reported in the *Article II & III* are concerned with various aspects of subjects' comprehension of material scenes depicted in this collection of images. It should be stressed that currently, to my knowledge, no established corpus of images tagged for properties such as communicative intentionality exists. In the preparatory phases of the empirical part of the project I thus had to establish one. In the following I will reflect on the methods and criteria used to collect and classify images for the stimulus corpus.

The aim of the experimental investigations was to assess subjects' comprehension of scenes featuring material objects used as signals in communicative interactions. While a number of existing studies of material signals use have relied on a rather narrow definition of communication as being about 'instructing someone to do something' (cf. e.g. Striano *et al.*, 2003; Tomasello *et al.*, 1997, 1999; DeLoache, 2000; Suddendorf, 2003; H. Clark, 2005), I wanted to broaden the scope of investigation to capture more general cognitive semiotic trends. The definition of communication operationalized in this project thus includes both instrumental/instructional, aesthetic and symbolic genres (see the Introduction and section 2.5 for further elaboration). In addition, I wanted to address the various degrees of conventionality of material signals, again making demands on the definition and classification of stimuli.

The objective to deal with a very general notion of communication (including several subgenres) has had some consequences for the choice of experimental stimuli and design. Since I wanted to test subjects on a wide variety of samples it would involve great practical challenges to use real world naturalistic subject-object encounters (see the section on naturalistic data above for a discussion). Instead, I settled on the use static photographic representations of such naturally occurring scenes. One of the great advantages of this kind of stimuli is that it is very 'manipulable'. Images can be normalized and balanced on a long list of variables and parameters that are not themselves within the scope of interest, but ensure comparability of measurements. Such variables include e.g. light, contrast and color saturation, not to mention the important issue of controlling stimulus presentation time and place. Without control of these kinds of presentation parameters quantitative analyses become fragile.

There are also problems involved in using photographs to simulate real world experiences. A major potentially confounding factor that should not go unmentioned in this context is the fact that a photograph is always taken by someone from a certain perspective and with a specific purpose. Thus, it can rightly be considered to constitute an intentional communicative object in itself. Behind even the most neutral-looking photograph is always the intention of a photographer to capture a specific situation or section of the world and display it to other people to share the aesthetic or factual (political, historical, etc.) meaning of the scene (cf. e.g. Quan, 1979 for a discussion). The use of photography thus brings to the design an extra layer of iconicity and intentionality that could potentially be a source of undesirable noise. It should however be stressed that experiments addressing similar theoretical issues have been successful in their use of photographic (and video) stimuli (cf. e.g. Lotze *et al.*, 2006; Lawrence *et al.*, 2006; Abell *et al.*, 2000; Willems *et al.*, 2007; Kampe *et al.*, 2003).

One of the measures that I have taken to counter this problem in the collection of stimulus images is to avoid photos with properties that in an appreciable way give relevance to the presence and intentions of the photographer. These include properties of the form, such as skewed perspectives or exaggerated close-ups, and in some cases the depictions; for instance a photo without an obvious figure or subject-matter would seem unnatural and might make the perceiver wonder about the relevance and thus indirectly the intentions of the photographer. Ambiguities caused by these kinds of image properties were also used as exclusion criteria in the later quantitative classification and rating procedures described below.

The objective of the experimental studies was to make contrastive analyses of subjects' comprehension of communicative objects vs. objects without any communicative relevance. I order to accommodate such contrastive studies I needed a corpus of images depicting two types of object configurations: configurations that would consistently be associated with a communicative intent and images that did not motivate such readings. Initially I collected 152 images from the local environment, using a Canon Eos D300 digital reflex camera, and from web-based photo databases and photo sharing sites such as http://www.polfoto.dk and http://www.flickr.com. In this stage of the process images were selected on purely introspective criteria. A tight contrastive match was premeditated by the arrangement of stimuli in pairs. Two kinds of pairs were formed: 1) objects manipulated for purposes of interpersonal communication vs. the same objects in canonical contexts without any conspicuous intentional communicative relevance, and 2) communicative objects in context vs. the same contextual scene without the communicative object. That is, in the first type of contrast the object was stable while the surrounding context would change, and in the second type, the context would be stable while the object would change or simply disappear (see the appendix for examples). While this could be seen as introducing unnecessary complexities (and extra degrees of freedom in statistical terms), both types were needed to capture a broad repertoire of communicative object uses at various degrees of conventionality. While the first type of contrast was suited to pick up everyday objects that would gain communicative value, for instance by virtue of their placement in a non-canonical context (cf. e.g. the fire extinguisher on a museum stand vs. the fire extinguisher in the messy cupboard), the second type would capture uses of more conventional symbolic objects that would be experienced to have such communicative significance as an inherent property and thus would rarely be featured in non-communicative uses (e.g. a door sealed off with police cordon vs. a door without any cordon).

The initial collection of 76 contrastive image pairs (152 images) was subsequently tested in a pilot classification study (see *Article II* for details). The study was designed as an electronic questionnaire and used an explicit multiple-choice task. The purpose of the classification study was twofold: 1) to test whether my introspectively based selection of images and distinction between material signals and non-communicative object configurations would find quantitative support in a larger population, and 2) to use informants' ratings to pick the best exemplars for a test and a control condition of stimulus images. I anticipated that a generally high agreement between my introspective classification and informants responses in the test would lend validity to the distinctive categories applied, while if informants failed to identify these categories it would be an indication that these distinctions were artificial and did not have any relevance for the casual perception of such scenes. On the other hand, if some images showed a conspicuous category effect (Haushofer & Kanwisher, 2007) and others did not, I interpreted that as a general verification of the predefined categories but at the same time an indication these particular images would be unsuited for the study.

As evident from the analysis of classification data presented in *Article II,* I found that informants generally performed well and their responses largely supported the a priori division of stimulus images. The study thus allowed me to select the 50 stimulus image pairs that scored best for a stimulus corpus. Subsequently, this corpus was used as a foundation for all the experimental investigation of the project.

It should of course be noted that the explicit style of inquiry used in these kinds of classification studies has the risk of inducing subjects with attentional biases that do not reliably reflect their naïve, casual perception of such scenes (see section 3.3.1 on explicit vs. implicit classification designs above). Examples of evidence for these kinds of problems in relation to the stimulus image corpus are discussed in further detail in section 5.2 on cross-method validation.

### 3.3.4 Experimental Subjects

In the various experimental paradigms reported in *Article II* & *III,* I have enrolled various populations of subjects. Unlike the majority of existing studies on material signals I have only used healthy adult subjects. This choice was mainly a matter of practicality. It is generally easier to include adults in experimental settings as they are more receptive to instructions. Furthermore, functional brain imaging with children is exceedingly difficult (unless they are anaesthetized), due to the unnaturalness of the situation that involves fixation and immense levels of noise. Another important motivation for using healthy adult subjects is the central role of phenomenological reports and introspection in the studies. These methodologies require subjects with at least some training in accessing and reporting the content of their own thought processes. This is not a trivial thing and cannot even be presupposed for all healthy adult subjects (Gallagher, 2003). Ideally, a subject population in an experiment should be representative of a common population with all its diversities of culture, age, gender, education etc. In practice, however, this is rarely a feasible strategy for recruitment of experimental subjects in cognitive experiments. There are several reasons for this. In brain imaging experiments recorded images of subjects' brains are normalized to a single template to make functional activations between subjects comparable. To get the best results from this pre-analytic processing you need brains that are as structurally (anatomically) similar as possible. One of the parameters that are known to have great influence on brain structure is age. Making young and elderly brains comparable in terms of activity is very difficult, due to the gradual degeneration of the cortex. In most functional imaging studies subjects are thus recruited among a somewhat homogeneous group of young people, and most often undergraduate psychology students that have professional motivations in participating. My fMRI study is no exception. Another argument for choosing a more homogeneous group is that fMRI experiments currently employ small sample-sizes (in most cases $n<30$) and thus we cannot afford the potential extra degrees of statistical freedom that follow from the introduction of ever-new socio-cultural parameters in the choice of subjects (unless these parameters themselves are the main focus of the study). Likewise, in the content analysis study reported in *Article III,* we ended up recruiting most of the twenty subjects from populations of undergraduates attending education at the faculty of the humanities, as we found that these were the most comfortable and trained in perceptually exploring images and verbally reporting their experiences.

I of course recognize that these recruitment strategies are not unproblematic and potentially challenge the attempted generalization of results. These criticisms notwithstanding, however, such practices seem to be the (implicit) rule rather than the exception within most of the cognitive and neurocognitive research that is currently published.

### 3.3.5 Functional Magnetic Resonance Imaging (fMRI)

The main methodology applied in the brain imaging experiment reported in *Article II* is functional magnetic resonance imaging, hence fMRI. During the last couples of decades fMRI has become one of the most widely used techniques for addressing question of functional brain activations and localization of brain functions. In the following I will give a very brief introduction to some of the main issues concerning the data acquisition and analyses with focus on some of the advantages and limitations of this method in relation to the current object of interest.

Functional brain imaging data is collected by scanning subjects in a magnetic resonance (MR) system. Roughly, fMRI measures changes in the oxygenation of blood in the brain, the so-called Blood Oxygen Level Dependent (BOLD) contrast (Ogawa *et al.*. 1990). The relative blood oxygen concentration in local areas of the brain is thought to reflect increases and decreases of functional neuronal activity in the same areas, although the exact mechanisms involved are complex, and as yet not fully understood (Logothetis *et al.*, 2001). The dynamic signal changes are recorded by the MR-system by rapid acquisition of multi-sliced images of the volume of interest. The technique thus allows you to – at least indirectly - detect patterns of localized neuronal activity that are statistically correlated with cognitive, motor or perceptual tasks performed by the subject in the scanner (Amaro & Barker, 2006).

Functional magnetic resonance imaging has some obvious advantages in comparison with other brain imaging technologies. In contrast to positron emission tomography (PET), fMRI is conducted without the injection of radionuclide tracers, which makes the procedure much easier (and cheaper). FMRI also has no known side effects, which has made it one of the most prevalent methods for studying brain function in normal healthy volunteer subjects (Toga & Mazziotta, 2002). Another advantage compared with, for instance, electroencephalography (EEG) is the high spatial resolution, which makes this technique (together with PET) especially suited for addressing issues of localization of brain function (Amaro & Barker, 2006). Depending on scanning parameters, the technique can be sensitive to signal changes in voxel sizes of few millimeters. With an acquisition rate of 1 – 5 seconds the temporal resolution is in the middle range, far better than PET (30 – 45 min.), but slower than EEG and MEG (milliseconds). Still, these resolution parameters should allow researchers to address most issues concerning higher cognitive brain functions (Cohen & Bookheimer, 1994).

A few limitations and disadvantages of fMRI should be mentioned as well. Beside the indirect and somewhat unresolved index of measurement there are a number of limitations to the kind of tasks and research questions that can be addressed using the fMRI technique due to practical constraining factors in the scanning situation. During the fMRI scanning procedure the subject solves a cognitive task while located inside a c. 60 cm. wide, 120 cm long magnetic tube and exposed to 120 db of acoustic background noise with his/her head fixated to reduce movement artifacts (Amaro & Barker, 2006). These very unnatural conditions for studying neural correlates of "general human cognitive functions" are in themselves not uncontroversial (cf. e.g. Bloom, 2006) and as a consequence it is currently not possible to scan people while they play the grand piano or tennis, and even scanning small children and animals poses great challenges, unless they are anaesthetized.

When designing an fMRI study you need to settle on specific experimental designs and strategies, each with their advantages and limitations. As most of these 'choices' are rather implicit in *Article II*, I will briefly introduce some important issues.

An fMRI study design can be based on various comparison strategies (cf. e.g. subtraction, factorial, parametric and conjunction). The comparison strategy adapted for the study reported in *Article II* is often called *cognitive subtraction* (Posner *et al.*, 1988; Aguirre, 2003). The basic principle is to have two conditions, a test and a control, varying only on the parameter of experimental interest while all other features are kept stable. Thereby the special cognitive properties of interest can be isolated by subtracting the control condition from the test condition. An illustrative example could be the subtractive analysis of brain activations related to the perception of two stimulus images; one depicting a normal strawberry, the other depicting a blue strawberry (cf. Zeki & Marini, 1998). Here the subtractive procedure would enable us to isolate the difference in neural processing caused only by the unnatural color while keeping other elements stable such as e.g. the contour and image composition. Similarly, in the fMRI experiment presented in *Article* II, the comparative manipulation was, roughly, to subtract canonical instrumental uses of everyday objects from communicative uses. Thereby, I could isolate the "extra" neuronal activity associated with the communicative functionality of the object depicted in the test condition.

While cognitive subtraction probably still is the most prevalent comparison strategy used in fMRI studies (Aguirre, 2003; Amaro & Barker, 2006) it is not uncontroversial. It relies on the principle of "pure insertion", that is, the idea that two cognitive operations can be added or withdrawn without affecting each other. Generally, this assumption is probably not valid (Friston *et al.*, 1996). Still, the procedure has the advantage of being simple in comparison with alternative strategies and has produced quite robust and reproducible results (Friston *et al.*, 1999).

In addition to the question of comparison strategy there are various types of stimulus presentation strategies, the most common of which are *blocked*, *event-related* and *mixed* designs (Amaro & Baker, 2006). Again each has its advantages and limitations, often dependent on the kind of cognitive task involved in the experiment. In the following I will mainly focus on aspects of the event-related design which was the stimulus presentation strategy adapted in the study reported in *Article II.*

Blocked designs have been the mainstay of fMRI for a long time mainly due to the heritage from PET that does not allow rapid image acquisition (Amaro & Baker, 2006). In blocked designs, multiple stimuli from the same condition are presented in a block and the BOLD signal is composed of the multiple individual hemodynamic response function (HRF) from these activations. In contrast, in event-related designs, stimuli from all conditions are presented intermingled in a randomized order and the BOLD signal from each individual HRF is detected and modeled in its own right. This technique has some noticeable advantages over blocked designs. For instance, the fact that stimuli can be carefully randomized can be used to counter practice effects and reduce subjects' ability to predict the next stimulus event (D'Esposito *et al.*, 1999). It is also known to reduce sensitivity to head movement artifacts (Amaro & Baker, 2006). Furthermore, the modulation of the BOLD signal for each HRF enables the detection of transient temporal variation in hemodynamic responses allowing a more detailed (e.g. weighted) analysis (Buxton *et al.*, 2004). Another great advantage that is especially relevant in this context is that conditions can be defined on post hoc behavioral criteria. The event-related design thus allow subjects via button presses or post-scan debriefing to individually and introspectively classify the stimuli in relation to the various conditions (Jack & Roepstorff, 2002). This was the strategy in the fMRI study reported in *Article II*. While subjects were exposed to a randomized series of stimulus images in the scanner, the statistical subtractive analyses of brain data were modeled on their individual post-scan introspective classification of the stimuli (see the article for details).

#### 3.3.6 Content Analysis

The experiment reported in *Article III* is modeled on a methodological approach often referred to as *content analysis* (Neuendorf, 2002). Content analysis is a standardized methodology in the social sciences for studying the content of human interactions. It has though also found wide application outside this field, e.g. in more humanistic textual analyses (Holsti, 1969). The method allows you to address the often open-ended and qualitative questions of communicative content in a deductive, objective and quantitative way. It does so by identifying tokens of a limited set of predefined keywords or meanings (types) that are thought to somehow organize the message in a dataset. The collected samples of keywords are analyzed statistically to test premeditated hypotheses, such as the relative frequencies of variables, differences in distributions of variables, bivariate relationships (the mutual attraction or repulsion of two variables), etc. (Neuendorf, 2002). In that respect, content analysis *summarizes* critical points rather than scrutinize subtle details of the data set (Neuendorf, 2002). The search procedure can either be automated using computer software, or can make use of a human multiple-coder approach. While the automated approach has the advantaged of being probably the most objective, it may suffer from the inability to capture more nuanced and contextually modulated meanings (cf. e.g. Mohler & Zuell, 2000). This may be solved by using multiple human expert coders. The use of human coders enables the identification of more fuzzy and contextualized tokens of the variables/keywords. In this approach, several informed coders independently assign the predefined variables to specific tokens in the data set. Subsequently, the individual codings are summarized and statistically analyzed as in the automated version.

The use of human coders also brings an element of subjectivity to the analysis. A critical issue is thus the measurement of inter-coder reliability, that is, the degree to which coders agree in their coding of the material. Inter-coder reliability can be seen as a common index of three factors: how well the variables of interest are defined at the outset of analysis, how well these variables actually relate to the data, and how well the individual coders perform in the coding procedure. If the inter-coder reliability is high it is thought to strengthen the validity of the data prior to analysis and vice versa (Neuendorf, 2002).

A range of statistical tools for measuring inter-coder reliability are available. Many studies only measure inter-coder reliability in terms of percent agreement (Lombard *et al.*, 2002). This is problematic, as such simple calculations do not take into account the agreement that could be due to pure chance. Among the more conservative tests, which do control for the impact of chance, are Cohen's *kappa* (Cohen, 1960), Scott's *pi* (Scott, 1955) and Krippendorff's *alpha* (Krippendorff, 1995).

A potential major confounding factor in the application of content analysis to at least some areas of human communicative practice is related to its declared positivistic point of departure. One of the grand ambitions in content analysis has been to "meet the standards of the scientific method" (Neuendorf, 2002; Gunter, 2000). In concordance, the grounding assumption that seems to pervade this methodology is that basically any aspect of human social life has an objective existence or at least an objective counterpart that can be measured (for a discussion related to the positivist vs. interpretivist dichotomy, cf. e.g. Ramberg, 2004; Dennett 1990). Though I am generally sympathetic to such approaches, they could easily end up operationalized in the prediction that there is something like a one-to-one relationship between human social structures and ideas and their material manifestations. It is unlikely that this is true in the case of human language. Language (whether spoken, signed or written) is an extremely flexible and pragmatic means of interpersonal meaning construction. A single idea can be expressed in almost an infinite number of ways and a single linguistic expression can be interpreted in an almost infinite number of ways. This points to the way language critically depends on pragmatic settings (common ground) and functional contexts (H. Clark, 1996). When we study language material we should thus be very cautious of procedures in which specific lexical items (keywords/variables) are blindly counted. Or put in other words, there exists a potential conflict between the ambition to use very stringent and objective measures to capture fuzzy and ungraspable, context-dependent meanings. This conflict raises serious challenges for the researcher.

In the experiment reported in *Article III,* I applied an adapted version of the content analysis procedure. Since the general structure of the design is accounted for in detail in the article itself, I will here only discuss some issues related to the above-mentioned challenges. In the first part of the experiment I used three expert coders to assess subjects' image descriptions. I settled on a (human) multi-coder approach rather than the fully automated computer approach in order to allow room for an interpretive aspect in the otherwise quantitative analysis. In addition, the variables of interest were defined in terms of semantically defined categories of words rather than specific lexical items. This allowed experimenters to approach the coding procedure in a somewhat more pragmatic and interpretative manner. In the second part I used a single-expert-coder approach to identify subtle aspects of meaning attribution in a subset of the data material identified in the previous analysis. In this second analysis I was thus not interested in a statistical comparison between conditions but rather the degree of agreement among the subjects.

### 3.3.7 Cancelled Studies

In the course of the empirical investigation for this project some experimental setups were cancelled halfway through as they appeared unproductive or not properly designed to answer the intended research questions. They are briefly mentioned here to provide an honest picture of the whole process of investigation.

In the post-scan, debriefing study used to model the individual responses in the fMRI design, an attempt was made to make subjects classify the stimulus images in relation to three communicative genres: aesthetic, instructional and symbolic. The classification data could potentially have been used to model contrastive analyses of brain activations related to these genres. Unfortunately, the questionnaire design allowed very ambiguous responses and the collected data thus proved very difficult to interpret. Therefore, this part of the analysis was abandoned.

Furthermore, an attempt was made to use reaction time as index for the comprehension of communicative objects of various degrees of conventionality. In a separate experiment subjects were exposed to the series of hundred stimulus images (corresponding to the corpus used in other experiments). For each image they would perform a forced choice classification task (communicative vs. non-communicative object configuration) while their reaction time was carefully measured. Subsequently, they were asked to rate the images they had just classified as 'signals' on an analogue 1 - 4 conventionality scale. The measurements allowed me to analyze and test specific hypotheses regarding the interaction between the reaction time and the degree of conventionality for each image classified as communicative. After the incorporation of ten subjects (half of the premeditated population), preliminary analyses revealed complex tendencies that did not directly support the hypotheses. This experiment was thus also subsequently abandoned.

In retrospect, I would suggest that the kind of stimuli and classification criteria applied in this experiment were too complex to be reliably indexed by reaction time measurements. Rather, the reaction time might have reflected the uncontrolled and very varied degree of complexity of the visual display (cf. Long & Wurst, 1984, for a discussion).

# 4 Commentaries to the articles

In the following I will briefly comment on each of the three articles constituting this dissertation. The aim is primarily to establish how they relate to each other and to the overall scope of the project. Since most of the positive results are presented in the articles, the purpose of these comments is also to point to some of the critical issues and problems that are not treated in detail in the articles. Furthermore, I will try to establish how my conception of material signals has developed through out the project.

## 4.1 Article I: When Agents Become Expressive: towards a theory of semiotic agency

This article was published in the spring 2007 in the journal *Cognitive Semiotics* (Peter Lang AG) in a theme issue on 'agency'. The article treats conceptual issues connected to the cognitive structure of our comprehension of material signals. The work on the article was initiated before I started this PhD project and central parts of the text were reflected in the project description of the PhD grant application. However, the editorial work and funding issues dragged on for several years while my concept of the subject gradually changed. In the early phases of my stipendium I was thus given the opportunity to do a major revision. Substantial parts of the article were revised given new insights in the research field and literatures. As such, the article presents some of the fundamental conceptual issues and problems that have guided the later empirical investigations in this project. It should though be emphasized that part of these early reflections on the cognitive foundations for material signals tend to be somewhat premature and have been continuously developed in later stages of the project. In the following I will point out some of the strengths and weaknesses of the article and try to relate it to the later investigations, presented in *Article II* & *III*.

On at least two central points my conceptual apparatus has gone through substantial changes from this early conceptual investigation. Firstly, some passages of the article seem to reflect a very Cartesian and un-embodied style of thinking about cognition. Much emphasis is put on the dissociation between internal and external processes in the account of mental and semiotic causality. Furthermore, the concept of an internal, readymade and perhaps even innate 'governing cognitive schema' that, when 'triggered' by external means, structures interactional encounters is inherently Cartesian in scope. This perspective is though somewhat softened in the last sections of the article where I introduce a more distributed perspective. Likewise, in the conceptual foundations of the following two articles I take a much more embodied phenomenological approach inspired by Gibsonian and Vygotskian approaches to human cognition (Vygotsky, 1986; Gibson, 1979) also manifested in the theories on distributed cognition and extended mind (Hutchins, 1995; Clark & Chalmers, 1998). In this alternative view, the internal-external boundary ceases to be the most relevant in the description of cognitive processes (see section 2.2.1, 2.2.2 or *Article III,* for details).

Secondly, the theory of 'communication as causality' that I launch in the first part of the article has not yet finally come to terms with old conduit metaphors of language and communication (cf. e.g. Jakobson 1960 and Saussure, 1974). The emphasis on the 'directionality' of the hypothesized causal structure of the communicative act leaves the addressee as a more or less passive participant who's attentional and intentional states are being manipulated (except maybe for the very last paragraphs where I introduce the notion of a 'symbolic stance'). Furthermore, the material signals I consider in this context tend to be framed as empty vehicles for some kind of dissociated mental contents. In the later investigations reported in *Article II* & *III,* I favor a much more enactive approach to communication entailing participants' joint participation in the meaning construction. In this process, aspects of the materiality of signals continue to impact the meaning exploration in a complementary fashion (see section 2.2.2 and 2.5.3 for a discussion).

Besides these more theoretical problems I will briefly comment on the issue of so-called 'mirror neurons' in relation to human social cognition that is superficially introduced in the article. Some parts of the cognitive sciences have been very optimistic since the first single cell recordings of mirror neurons in macaque monkeys were published (Gallese *et al.*, 1996; Rizzolatti *et al.*., 1996). Single cell recordings are currently not possible in humans, but a comparable 'mirror system' has been hypothesized in humans based on the much more indirect measurement of functional brain imaging (cf. e.g. Jones, 2002). The general idea is that this 'human mirror system' should facilitate simulation and that simulation is the foundation of social understanding (cf. e.g. Gallese & Goldman, 1998). Mirror neurons are not mentioned elsewhere in this project as I have become very skeptical about their proposed involvement in a number of social practices. If something like a mirror system is involved in social cognition, I would expect it to be as part of a third-person observational system (as seemingly the case with the 'theory of mind' network) neutral to interaction and signaling practices (for a similar perspective, see e.g. De Jaegher & Di Paolo, 2007).

But the article also has many insightful points that turn out to predict some of the later empirical findings of *Article II* & *III.* The hypothesis that the casual perceiver does "historical readings" of certain 'asymmetrical' aspects of their static material environment, and in doing so distinguish between various types of causal motivations, seems to be supported by all the experimental findings. In addition, aspects of the critique of Leslie's theory of agency become relevant in the analyses of the relationship between social cognition and communication. The often unappreciated distinction between third person observational encounters and first- and second-person participatory (or interactional) encounters in approaches to theory of mind and agency may thus have great explanatory value in relation to the interpretation of many fMRI brain imaging experiments, including my own (cf. section 2.4.3). Furthermore, the same distinction is reflected in very recent findings in developmental behavioral studies of Csibra and Gergely and colleagues (cf. section 2.1.1). Their ideas about ostensive signals and their effect on the cognitive framing of social encounters seem fully compatible with my theory of the dissociated, though complementary, indexical (ostensive) and symbolic (semantic) aspects of signal meanings. Finally, the idea that ostensive cues can be externalized in the material environment as fixed traces of purposeful semiotic acts appears to coincide with some of Jerrold Levinson's ideas about the recognition of *categorical* and *semantic intentions* in works of art (cf. section 2.3.2).

Interestingly, in the concluding section of the article, I briefly introduce the notion of a "symbolic stance" as a certain attitude by which we approach encountered purposeful variations in our material environment. This idea is taken up again and further elaborated in the interpretation of experimental findings in *Article III,* though now rephrased as a "language stance".

## 4.2 Article II: Say It With Flowers! An fMRI study of object mediated communication

This article was accepted for publication in July 2008 by the journal *Brain and Language* (Elsevier).

The article reports from a fMRI brain imaging experiment targeting the neurocognitive foundations of our comprehension of material signals. Unlike *Article I,* the conceptual theoretical elaborations are here kept brief, while focus is on the empirical methods, results and interpretations. To meet the academic traditions of this particular genre of research, the version of the article accepted for publication thus frames both the hypotheses and the interpretation of result in relation to other fMRI brain imaging experiments on related issues. This can make the article seem a bit detached from the rest of the project. Still, the purpose of the study was to test some of the predictions sketched out in *Article I,* in relation to existing knowledge of social cognition, language and communication in the brain. In the following, I will point to some of the deficiencies and strengths of the study and try to relate the findings to the other two articles and to the project as such.

As evident from the analyses, my initial hypotheses were only partly supported by the findings in this study. Roughly, the overall prediction was that the recognition and comprehension of material signals would be cognitively processed in a manner comparable to other kinds of communicative mediations. This hypothesis is largely supported by the findings. But one of the more precise predictions put forth in *Article I* was related to theories of agency and 'theory of mind'. The recognition of another person's *communicative intentions* from the compositionality of static material scenes plays a central role in the theoretical framing of material signals in *Article I* and accordingly motivated concrete expectations in relation to certain anatomical sites in the early phases of designing the experiment. Interestingly, though, we did not find activations in areas of the brain normally associated with 'theory of mind', as we initially expected. Other studies on the recognition of social intentions and ostensive cues report activations in structures such as the temporo-parietal junction and medial prefrontal cortex (see section 2.4.1 for details). These are conspicuous by their absence in our study. Still, there may be very good conceptual reasons for this that are actually already latently implied in some of the elaborations of *Article I*. A closer look at the paradigms that are employed in the majority of studies on 'the social brain' and 'theory of mind' are concerned with a third-person observational type of inferential processes. Subjects in these studies are to *explain* and *predict* other people's behavior by reference to inferences about their mental states. But joint meaning construction might not be so much about the mental states of other people. Rather, it is more likely to be about how we attend to meanings in a shared material world. This idea is supported by other brain imaging experiments that study various kinds of communicative mediations, that is, paradigms that imply a first-second person interactional type of attitude in the subject (see section 2.4.2 and 2.4.3). The critique of Leslie's typology of agency launched in *Article I* is thus echoed in the interpretations of findings in *Article II*: interaction may actually not be so much about reading the mental states of others as about orienting to a shared world of meaning. Accordingly, we might find two dissociate social networks: one person-centered 'theory of mind' network and another object- or meaning-centered interaction network comprising some of the areas traditionally associated with language and communication including those reported in *Article II*. These are of course empirical questions that need to be addressed in future experiments, preferably designed in a way that invokes both networks independently and in combination in the same controlled paradigm.

Another relevant issue in relation to this study is of a more methodological kind. In the experimental design we used a passive viewing task in the scanner combined with a post-scan explicitly elicited classification task. An implicit assumption guiding the statistical modeling of the data is thus that subjects' responses in the explicit debriefing classification task reflect their prior naïve and casual experience of the stimuli during functional brain scanning. This might be the case. But it is also possible that the explicit questions in the debriefing task change subjects' conception of the stimuli as they impose a new frame of reference and interpretation (Jack and Roepstorff, 2002). As a consequence, we simply do not know how the subjects actually perceived and comprehended the stimulus images while in the scanner. This fact was part of the motivation for the experimental design launched in the comprehension study reported in *Article III*.

## 4.3 Article III: Taking the language stance in a material world: a comprehension study

This article is currently submitted to the journal *Pragmatics and Cognition* (John Benjamin's Publishing Company), to become part of a theme issue entitled 'The Dynamics of Symbolic Matter'. It reports from an experiment targeting subjects' casual understanding of material signals. Contrary to the procedure in *Article II*, this experiment does not just address the categorical question of the recognition of material signals (signal vs. non-signal contrast), but also investigates the kind of *social meaning* that subjects attribute to various kinds of recognized material signals. Another important difference is that in this experiment we do not statistically model subjects' responses on an explicit classification task that might potentially create undesirable interpretative biases, but instead use a coding procedure to classify subjects' unbiased, casual and introspective reports post hoc (see section 3.3.6 for details). This provides me with a very rich dataset that allows for both hypothesis-driven quantitative analyses and more explorative post hoc approaches. In addition, the more humanistic orientation of this research journal allows me to advance more complex conceptual elaborations and discussions. There are, however, a couple of issues related to the experimental design and the findings that could deserve a critical comment.

To limit the raw audio dataset to a size (six hours of recordings) that could be handled within a reasonable timeframe for transcription and coding etc. we chose six stimulus images for each of the three experimental conditions giving a stimulus set of in all eighteen images. It could be argued that six stimulus events are a bit on the low side to make the kind of generalizations that I propose in the article. Furthermore, the two test conditions (conventional and unconventional material signals) end up separated not only "horizontally" by reference to conventionality, but also "vertically" (across the conditions) by reference to various interpretative affordances (iconic, indexical and symbolic). Thereby the number of stimulus images in each discrete category (e.g. conventional/symbolic, unconventional/symbolic, etc.) becomes so small that quantitative findings have to be interpreted with caution.

Another related issue is that the hypothesized impact of conventionality that constitutes a central parameter in the study dividing the two test conditions could actually turn out to be secondary to the derived parameter of interpretive affordances (sign types). Though both are found to have an effect, the latter parameter may seem to be a stronger predictor of and more conspicuously manifested in the image descriptions. This intuition is supported by the fact that symbolic material signals are primarily attracted to the conventional condition while most iconic material signals are found in the unconventional condition indicating that the conventionality parameter is not neutral to the sign type distinction. Of course more experiments are needed to ascertain these matters and one of my coauthors is currently planning a follow up study involving more stimulus images.

A final comment is dedicated to the curious finding that not all stimulus images found to be good exemplars of material signals in the explicitly elicited classification studies reported in *Article II* got the same high ratings in the content analysis study using ‘implicit’ elicitation. Although we generally found good correspondences between ratings in the explicit classification and the description tasks, a couple images consistently rated to depict material signals in the classification studies did not motivate descriptions that could subsequently be coded as ‘material signals’ in the description design. This could be an indication that the explicit experimental elicitation (questionnaires and force-choice paradigms) sometimes create unintended response biases. The problem is further discussed in the *Chapter 5* below.

# 5 Converging evidence - cross method validations and variations

In a multi-methodological investigation like the present project, a crucial question concerns the way the different experiments and applied methodologies mutually inform, support and complement each other. Since such important meta-perspectives are beyond the scope of any of the individual research articles making up this dissertation I have devoted a special section to the issue here.

## 5.1 Positive evidence

The general hypothesis of this project derived from the early conceptual work in *Article I* seems to find support in all empirical investigations manifested in *Article II* and *III*: When confronted with certain types of purposeful object configurations people consistently change their interpretative attitude (they “take a language stance”) and explore the material scene as a source of social meaning. In *Article II*, this finds expression in the pronounced category effects in the explicit classification study and in the significant brain activation patterns in the fMRI experiment. In *Article III*, the signal/non-signal classification was derived from subjects’ naïve and casual image descriptions by a coding procedure and a significant effect was found when conditions were statistically compared.

The fact that significant signal/non-signal category-effects show up in all experiments across various methodologies and measurements can be considered both an indication that the contrast does in fact capture substantial cognitive (and neurocognitive) aspects of our dealings with the material and social environment, and as a validation of the empirical methodologies applied in this project.

There are however also parts of the initial hypothesis that were not supported by the empirical investigations. In *Article I*, I hypothesize that an indexical, person-centered, 'theory of mind'-like component plays a conspicuous role in relation to our recognition of material signals. Considering recent findings in the neurocognitive paradigm, my primary expectation was that this component would be related to brain areas consistently associated with communicative intentions and ostension, for instance the temporo-parietal junction and the anterior paracingulate cortex of medial prefrontal cortex. No such areas were activated in our fMRI experiment. Rather, we found neural activations in a network of areas traditionally associated with language and gesture. This could be argued to correspond to the second 'object-centered' component of the prediction in *Article I* – that is, the claim that signals also symbolically refer to entities in a shared world of meaning.

While the double nature of material signals hypothesized in *Article I* was not fully supported by the fMRI experiment, another related distinction was derived from the patterns of brain activations: *categorical* vs. *semantic* resolutions. Two bilateral sets of areas were interpreted as playing a central role in the recognition and comprehension of material signals: the fusiform gyrus and the inferior frontal cortex. Traditionally, fusiform gyrus is considered part of the visual 'what'-stream involved in object recognition and categorization. But the structure also includes certain subareas with, supposedly, quite specialized functions. The most well-validated of these is the so-called *fusiform face area* that is consistently shown to be involved in the recognition of faces (Kanwisher & Yovel, 2006), and the *parahippocampal place area* which is seemingly sensitive to the perception and recognition of buildings and sites (Epstein & Kanwisher, 1998). But a third subarea has been suggested – the *visual word from area* (VWFA) - that is supposedly involved in the recognition of letter strings (Cohen *et al.*, 2000; Cohen *et al.*, 2002; Dehaene *et al.*, 2002) and perhaps symbols more generally (Reinke *et al.*, 2008; Kronbichler *et al.*, 2004). Interestingly, the fusiform gyrus activation peaks of our study reside in exactly this area. This suggests that the proposed function of the VWFA could be extended from conventional letters and symbols to include material structures perceived as signals. Accordingly, the interpretation we propose in *Article II* is that the fusiform activations in our study are responsible for the categorical resolution 'signal vs. non-signal'. Another bilateral activation was found in the inferior frontal cortex (IFC), in the superior part of Broca's area. Since these areas have been consistently associated with verbal semantics, we interpret these activations as involved in the more subtle contextualized meanings of the material signals. Together, these patterns of activation thus reflect Jerrold Levinson's concepts of dissociate *categorical* and *semantic* intentions (Levinson 1979, 1989, 1993, cf. section 2.3.2).

Interestingly, this finding is replicated in *Article III*. Here we identify various 'stages of comprehension' comparable to the proposed distinction between the categorical and semantic activation patterns seen in the fMRI experiment, though we hypothesize a intermediate stage of orientation to certain interpretative strategies related to Peirce's sign types *icon, index* and *symbol.* Furthermore, compatible trends are found in relation to the effect of conventionality. In *Article III* we only found the conventionality parameter to affect the more subtle semantic stages of comprehension, while the categorical stages seemed neutral. Accordingly, when applying a 'conventionality regression analysis' to the fMRI brain imaging data, we do not find the structures of fusiform gyrus hypothesized to be involved in the categorical resolution to be sensitive to this parameter; rather, we find significant activations in the inferior frontal structures thought to facilitate contextualized social meanings/ semantics. Across these quite different types of experimental measurements, conventionality is found not to affect subjects' recognition (categorization) of signals, but has an effect on their exploration of signal meanings.

Another example of cross-study compatibility between *Article II* and *III* deserves a brief comment. In the study reported in *Article III*, subjects' descriptions of control scenes were generally characterized by a high degree of variability. This was interpreted in terms of subjects' inclination toward a "private" and unregulated style of perception and comprehension of these non-communicative scenes. This could seem to correspond to the finding that the 'non-signal>signal' contrast in the fMRI experiment did not show any significant activations. As subjects in their approach to control scenes do not orient to shared top-down frames for the attentional and intentional exploration of the scenes (cf. *Article III*), their neural activations do not appear to align in any patterned fashion.

## 5.2 Grit in the machinery

Although the results across various empirical investigations and methodologies in general tend to complement each other nicely, there are a couple of problems that should not pass unnoticed. In Chapter 3 on methodology I briefly discussed the issue of biases in explicit vs. implicit classification elicitation (section 3.3.1). A detailed analysis of the descriptions reported in *Article III* points to some of the potential problems in this regard (also discussed in the commentaries to *Article III*). In one of the test conditions (unconventional material signals) used in this paradigm, we included an image depicting an axe placed in a HP printer. In the explicitly elicited classification tasks this image was consistently classified by subjects to be a material signal and in these studies ended up as one of the best exemplars of the category (among the 15 highest scoring). Curiously, in the implicit elicitation paradigm employed in *Article III* however, the same image scored very low as a material signal. Only a few of the twenty subjects identified the axe scene as a signal, while the rest did non-signal interpretations of the scene. A similar pattern applies to an image depicting a chair put up on a wall to create an aesthetic effect.

It could be argued that the implicitly elicited classifications give us a more realistic view of subjects' naïve and casual understanding of the material scenes. And since only 12 of the original 50 test images were used in the last experiment, we do not know how prevalent the problem is in relation to the corpus as a whole. One reason to worry about inconsistencies between the two types of classifications is that the implicit elicitation paradigm resembles the passive viewing task used in the scanner more closely than the explicit elicitation paradigm. It might thus be the case that subjects' casual viewing of images in the scanner and responses in the explicitly elicited post-scan classification task used to model the analyses of brain data do not match especially well. This question cannot be answered at present.

Another potential problem I will briefly comment on is the role given the concept of conventionality in this project. No doubt conventionality plays a role, but as discussed in the commentaries to *Article III*, it might be that conventionality applies quite differently to the various types of signals (*iconic, indexical* and *symbolic*). Symbolic reference, by definition, rests solely on convention while iconic and indexical signs are more 'naturally' (non-arbitrarily) bound to their meanings. While the successful comprehension of symbols entails some prior knowledge of and experience with the symbols in question, iconic and indexical styles of reference might be more transparent and accessible to the casual perceiver. This could be one of the reasons why most of the signals rated as conventional tend to employ a symbolic (and partly indexical) style of reference, while the majority of the unconventional signals tend to be of the iconic/aesthetic kind. We can easily imagine that, when naïve to a conventional symbolic use of a signal, we would tend to go for an iconic or indexical reading. These parameters might thus not be fully independent. Again, this is not counterbalanced in the fMRI experiment and might thus potentially confound the conventionality regression analysis of the study, or at least the interpretation of it.

# 6 Future Perspectives

While I consider a few questions regarding the cognitive foundations of our comprehension of material signals to be answered in this thesis, there are of course a lot of issues that still need to be addressed. Some follow-up studies are already initiated, while others depend on unresolved issues of future funding. Here I list a few of the perspectives I hope to pursue in future studies.

One of the findings of this study was that the recognition of material signals elicits activity in brain areas commonly associated with verbal language and gesture. Interestingly, we did not find any activation in brain areas found to be associated with social and communicative intentions in previous studies. I have hypothesized that this is due to Csibra and Gergely's notion of dissociate object-oriented and person-oriented representations (Csibra and Gergely, 2006; Gergely *et al.*, 2007). I thus imagine that the 'mentalizing network' thought to facilitate our understanding of communicative intentions and ostension is at work when we observe other people in order to predict or explain their behaviors by trying to figure out their minds. It might be that the moment we not only observe but engage ourselves as participants in a joint sense-making effort the question of person-centered mental states is not as important as the construction of object-centered intersubjective meaning. Together with Shaun Gallagher's student Micah Allen, I am currently working on a new experiment to test this hypothesis. The current preliminary plan is to use the fMRI technique to monitor subjects' brain activations while they perceive simple object manipulations either accompanied or not accompanied by ostensive cues. The experiment is designed to address both our recognition of ostensive eye cues, and the impact of ostention on our comprehension of simple object manipulations resembling H. Clark's material signals 'placing for' and 'directing to' (H. Clark, 2005). The idea is that when these factors (ostensive eye gaze and motor object manipulation) interact, a new object of attention emerges – a joint sense-making that does not depend on mentalizing per se (De Jaegher & Di Paolo, 2007). The design will thus, at least potentially, allow us to study the interaction between functional brain networks related to mentalizing, simulation and social signaling in one experiment.

Another aspect of the current investigation that needs to be further explored is the issue of the multiple regulating interpretative stances toward material signals found in the studies in *Article III*. While I generally found it productive to merge quite different styles of communicative object use (from indexical imperatives to symbolic artifacts and works of art) in the studies constituting the thesis, a natural next step would be to systematically contrast these. One of my coauthors, Johanne Stege Philipsen, is already advancing a follow-up study on the content analysis experiment presented in *Article III* as part of her master program in Cognitive Semiotics. Her idea is to add more stimulus images and subjects to the design, and then to again code for the 'interpretative attitude' of the subjects to various types of stimuli. Finally, I have applied for a post doctoral position to address questions regarding the cognitive and neurocognitive foundations of our perceptual approaches to aesthetic objects in various contrastive contexts. Part of this application concerns the way that our motivated adoption of an aesthetic perceptual attitude toward the material environment affects our phenomenological experiences.

# 7 The Articles

Kristian Tylén

# When Agents Become Expressive: A theory of semiotic agency

In this paper I will outline a theory of agency proposed by Alan M. Leslie. It focuses on the cognitive mechanisms underlying our immediate recognition of different types of causal agency. I will argue that a fourth kind of agency should be added to Leslie's list, i.e. semiotic agency. Semiotic agency designates our ability to recognize and interpret intended expressive behaviour and objects in our surroundings.

Whenever we use language we perform an intentional act of a causal nature. By producing signs we induce a change in the attentional orientation of the addressee. This mental change is an event, and the addresser is responsible for the event, i.e. is recognized as the causing agent. Thus, communication is agency.

In the following I will introduce the notion of *semiotic agency* focusing on our ability to recognize communicative intentions in the behaviour of other agents and symbolic artifacts – an approach to the act of enunciation that will likely be seen as a challenge to the prevailing use of the notion causation and agency.

## Semiotic intention and attention

Sometimes a rock is just a rock. We may or may not notice it there on the ground. We may be annoyed by its presence just ahead of the lawnmower or admire it for its aesthetic 'rocky' properties and pick it up for our rock collection. Sometimes a rock is a tool. We may consider the rock useful in some instrumental context (breaking windows

CORRESPONDENCE: Kristian Tylén, Institute of Language and Communication, University of Southern Denmark. Email: kristian.tylen@language.sdu.dk.

or building fences, etc.). When we recognize a rock as a useful instrument for some activity or purpose, it is because it fits an internal intentional program. Our intentional attitude makes the cognitive attentional system 'transform' the surrounding objects into potential helpers (or harmers) of the project. If the project is a window that needs to be broken or a fence that needs to be built suitable rocks will suddenly constitute the foreground of our visual attention.

But sometimes the rock is a message. We recognize the rock as referring to something other than itself. Take a gravestone. It 'tells' us that someone died and is buried there. Perhaps it 'tells' us to honour this person's memory. This is puzzling. How can a rock possibly tell us anything at all? What is it that suddenly makes the rock *communicatively significant* to us?

When we experience the rock as a message, it cannot be a product of our own intentional attitude toward rocks alone. It is not a 'top-down attention' in the sense that we are explicitly searching for 'message rocks' as could be said in the case with the 'instrument rock'[1]. In some fantastic way, the rock addresses us (bottom-up attention). Or more precisely – we are able to recognize the addressing intention of someone else in the rock. We understand that some cognitive being has acted upon this rock in order for it to represent something totally different from itself. And this immediate understanding detaches our attention from the rock as such and directs it to the content of the message.

As I would prefer to avoid any hocus-pocus supernatural explanation of this information exchange, I suppose that the communicative intentional action of the addresser left a recognizable trace on the rock. It could be some kind of manipulative modification of the surface of the stone like an inscription, a modification of the immediate surroundings of the rock with cut flowers, etc., or a simple displacement.

Another example: We often go around lifting our eyebrows. It doesn't mean anything. But sometimes a lifted eyebrow means a lot; it is intended as a message and is somehow immediately understood as such. Generally it seems that what would appear to be the smallest, most insignificant displacement, transformation or manipulation of an object can make it a medium for communication and is immediately interpreted as so by the addressee. Verbal language is made out of such micro-differentiations in the phonology of speech and the corresponding graphic representations in written

[1] About top-down and bottom up attention see e.g. Oakley (2003).

language, even though this ontology does not solve the mystery. Our everyday life leaves dozens of traces in the surroundings whereas only a minor part is intended and recognized as communication. Still, we manage to sort out the immense amount of background noise and focus attention on the semiotically significant details of the surroundings.

Thus, to sum up the first problem, it would seem that we are somehow cognitively equipped to recognize human semiotic intention and that our attentional system is extremely sensitive to semiotic acts and expressions in preference to other kinds of world phenomena. However, this complicated problem already breeds a lot of related questions. For instance, what is a semiotic act, and how does it differ from other kinds of human acts?

## A Theory of Agency

In his article "A Theory of Agency" (1993) Alan M. Leslie proposes a tri-partite theory of agency, with a cognitive and developmental psychological approach. His theory may prove valuable in this context, as it concerns our ability to distinguish a set of properties that differentiates agents from other physical objects and allows us to track and interpret different aspects of agency in our surroundings.

To Leslie, agency is understood as '*the core constraints that organize our early learning about the behaviour of agents*' (Leslie 1993: 1). The notion of 'core' constraint indicates a modular, biological approach to the problem of agency. Thus, Leslie believes that, as a result of adaptive evolution, humans have developed a highly specialized information processing system that provides the basis for learning and understanding the behaviour of agents. The result is a '*sophisticated capacity to explain, predict and interpret their [the agents'] behaviour*' that has adaptive advantages in a socio-physical world (1993: 1).

Though we could already at this point contest Leslie's theoretical grounding, his work makes a suitable platform for further investigations into the notion of agency. In the following I shall give a brief overview of Leslie's theory, skipping most details and focusing especially on the last part, which unfortunately is the least elaborated and documented.

Leslie's experimental observations of children from the age of 6 – 18 months motivates a typology of causality, hierarchically ordered in *mechanical causality, intentional causality* and *psychological causality,* in which each of these components corresponds to a set of properties that characterize agents and distinguish them from other kinds of physical objects. These are as follows:

> Agents have *mechanical* properties. They have an internal and renewable source of energy or force, i.e. they do not need to rely on external sources. They are 'self-propelled' and capable of bringing about changes in physical space.
>
> Agents have *actional properties*. Agents act intentionally, in pursuit of goals and re-act to the environment as a result of perceiving. Furthermore, the acting and re-acting agents can get together and *inter-act*.
>
> Agents have *cognitive properties*. The behaviour of agents is determined by internal cognitive properties, e.g. holding a certain attitude to the truth of a proposition.
>
> (Slightly adapted version of Leslie 1993: 2, italics added)

While Leslie is most concerned about the first category and the physical force representation (and this is where his theory is most thoroughly documented) I will concentrate on the latter two.

## Mechanical agency

During the first eighteen months of life, children gradually acquire the ability to recognize these different aspects of agency. From about six months, children show sensitivity to mechanical causality and the ability to distinguish internal-force bearing, self-propelled agents form mere physical objects. Central to the understanding of mechanical causation is the transmission of force, e.g. from an agent to an object, through a spatial (and temporal) contiguity. Without physical contact, no mechanical causal influence is possible.

## Actional agency

Within the second half of the first year, children begin to recognize intention in the acts of other individuals, i.e. an agent acts in pursuit of an internal goal or reacts to an external aspect of the environment through perception. Furthermore, two or more goal-pursuing agents can join in different kinds of cooperative or competitive interaction, such as helping or harming, depending on the mesh of goals. In this period of development, children also begin to follow the eye gaze of other people. Thereby the child will seek the reason for the agent's behaviour in the focus of her attention and action. Thus, the agent is now approached not only as a transmitter of force (in the physical/mechanical sense) but also as a possessor, a transmitter and a recipient of *information* (Leslie 1994: 143).

An important aspect of actional agency in opposition to physical/mechanical agency is the spatial and temporal detachment of, or 'distance' between, cause and effect. Thus, we don't find the same transmission of force through direct contiguity. When we witness an intentional goal-directed action of an agent, the effect or goal-state has not yet come about. It is still a little ahead in time, and may in fact never be realized at all. Likewise when we react to some perceived occurrence in the surrounding environment, this 'cause' may be spatially distanced from the perceiving and reacting subject. As this non-contiguous relation between cause and effect is only enabled by the interference of a cognizing mind, perhaps it would be more precise to follow Østergaard (2000) and term it *mental causation*[2]. As Østergaard stresses, the difference in ontology between mental cause and physical effect makes it impossible to talk about a 'transmission of force' in the traditional sense (Østergaard 2000: 10). Cautiously, however, one could propose the notion of *pregnance* in René Thom's sense of the word (Thom 1990) as a kind of force underlying actional agency: Whenever a perceived situation leads to a reaction of an agent, it is because some element has a potential to influence the agent. Likewise, when an agent out of internal motivation directs her intentional actions towards a certain goal, e.g. the acquisition of an object, it is because the mental visualization of the goal-state represents a pregnance to the agent (in this case an attraction towards conjunction with the object) (Østergaard 1998: chap. 1)[3]. This 'goal-status,' with its potential influence, is not (necessarily) a property of the objective physical world but is due to the biological and cognitive *values* of the agent. In a well-elaborated attempt to naturalize human intentionality, Gallese and Metzinger

---

[2] In his article "Mental Causation" (2000), Østergaard gives a detailed description of the mental processes intervening between a perceived physical event and a corresponding physical reaction from the observer-agent. These processes include perception, memory, belief, pro-attitude and motor control.

[3] This is a very superficial presentation of an otherwise very elaborate theory; a more detailed presentation is beyond the scope of the present article. I should add, though, that René Thom distinguishes between internal and external pregnancies, that more or less correspond to my notions of biological and cognitive values (see e.g. Østergaard 1998: 15-33)

(2003) describe the tight neural connections between motor action schemes (e.g. the so-called F5 area of premotor cortex), the goal-representational system and the value system. Their conclusion is that we seem to possess certain 'goal-related' neurons. When confronted with a pregnant object, this system immediately triggers action simulations, i.e. motor schemes that correspond to the most rewarding or satisfactory behaviour towards the specific object. Interestingly, a set of these neurons seemingly do not mind the special perspective of action goal representation but are equally active in first-person goal-pursuing action and the third-person observation of other subjects directing their intentional behaviour towards an external object. These are the so-called 'mirror neurons', which would seem to play an important role in our understanding of others' behaviour, thereby making it possible to adjust our behaviour to our social context. These neurons probably constitute the base of social cognition (Gallese and Metzinger 2003: 29)[4].

## Cognitive agency

From about the age of eighteen months, children gradually gain the ability to 'detect' volition, belief and pretence in agents. That is, they begin to understand the cognitive properties of other subjects. According to Leslie, this requires a sensibility to the relationship between agents and information (Leslie 1993: 11). Attending solely to the 'novel meaning' of the behaviour of an agent, the child becomes aware of the propositional attitude of the agent. She can understand some actions of the agent as caused by representational, fictional circumstances. Thus, when the agent asks the child to say goodnight to a teddy bear (my example) it is not because she misinterprets the actual situation, believing the toy bear to be cognizant of verbal language. The agent is not delusional. The child recognizes the intentional behaviour of the agent as related to (i.e. 'caused' by) an imaginative reality in which the bear is animate and cognitively equipped. To do that, the child must not only direct its attention to the actual referent of the behavioural expression, but likewise to the agent's own attitude to the behaviour. In other words, the child must pay special attention to the way it is performed, i.e. intonations, special gestures or the facial expression of the agent that signifies fictional attitude. Furthermore, the child must itself be capable of simultaneously holding and

[4] Other important brain areas in this respect are the so-called 'Theory of Mind' areas described by Amodio and Frith (2006).

keeping track of two representational inputs; one constituting the actual scenario in which the teddy bear is just an inanimate, 'dead' object, and another featuring the content of the agent's behaviour, i.e. an animate and cognitively equipped bear[5] (DeLoache 2004).

To Leslie, the notion of 'cognitive agency' only designates our ability to understand the behaviour of an agent as caused by fictive (non-actual) circumstances. Principally he tries to keep the third-person perspective on the agent, having the child on the sideline as an observer, interpreter and recognizer of cognitive agency. This project succeeds in some of his false-belief tasks where a child is asked to predict the behaviour of another subject in a classic 'hidden object task' (Leslie 2000). Nevertheless, most of his examples contain a lot more. Take his favourite example, repeated in all his writings on the subject (e.g. Leslie 1987: 417, 1993: 12, 1994: 141, 2000: 1236): A mother hands the child a banana saying, "The telephone is ringing. It's for you". As in my own teddy bear example, the child is not only expected to understand the behaviour of the adult agent as caused by a cognitive attitude, but is invited to take part in the pretence, that is, to attend to the specific *content* of the pretence and *act* accordingly. Most eighteen-month-old children will readily adapt the propositional attitude of the adult agent and follow it into the representational world of banana telephones. In the same way that the child in the earlier stage of development followed the eye gaze of the adult in actual space, it now understands the 'invitation' of the adult to follow her attention into a mental, representational space. But how does this come about? The recognition of pretence does not in itself explain the causal structure of the representational and participatory effect in the child. Leslie wilily escapes this obvious problem of causality and agency. Though he mentions 'sensitivity to information' and even 'communication' as elements of cognitive agency, he goes to great lengths to leave out remarks on language. This is probably due to Leslie's conviction that language forms a discrete and isolated cognitive module in itself and thus cannot be directly connected to the cognitive mechanisms of causality. Only in a couple of lines in his 1994 article (p. 142) does he mention language explicitly, confirming his belief in separate processes. However, this strict separation of language and communication must strike one as counterintuitive when we focus on the causal structures of semiotic exchange. How does an agent cause a subject to create mental

[5] The act of pretending thus yields a capacity for holding dual representations, not losing track of the specific labels of the inputs: what is real and what is fictive (cf. DeLoache 2004: 4).

representations of non-present realities and afterward to attend to certain aspects of these? Trying to account for this intriguing problem, I will suggest a supplementary notion to Leslie's typology of agency, i.e. *Semiotic Agency*.

## Semiotic agency

One could wonder why I chose to spend several pages of this paper lining up an elaborate theory of agency that does not once mention language and thus proves insufficient in solving our initial problem. Nevertheless, I approach the communicative act as a type of causation, which at the same time makes the question of agency relevant. Furthermore it might turn out that these 'information processing systems' are not at all as autonomous as Leslie assumes. The introduction of semiotic agency at least yields tight cognitive connections and 'overlaps' with the cognitive mechanisms underlying our recognition of other kinds of causality and agency, e.g. Leslie's notions of *actional* and *cognitive* agency.

As briefly noted in the prior sections, our ability to recognize actional and cognitive agency constitutes the foundation for the development of *social cognitive skills*. Some of these we share with our fellow primates, but at some crucial point in evolution something happened that made the species of homo sapiens develop in a unique way. Not hesitating to avoid the still problematic cognitive and neural details, this species-unique aspect of social cognition can be phrased more or less like an understanding that *"I am an intentional and cognitive 'self' among other intentional and cognitive 'selves' like me"* (Tomasello 1999: 4). This understanding enables us to identify with other subjects, their intentional and mental states – something that dogs, squirrels and hyenas probably do not do. Thereby it becomes possible for us to exchange and share experience and knowledge, i.e. it enables an especially effective *cultural transmission* and *cumulative cultural evolution* (I return to these notions below). Only humans seem to possess this ability that is the primary reason for our high degree of evolutionary adaptation (ibid.).

One of the most interesting and important aspects of social cognition is *joint* and *shared attention*, which probably constitutes the core foundation of semiotic agency and exchange. As noted in an earlier section, from early stages of ontogenesis, children tend to follow the gaze direction of others. But only at the age of nine to twelve months will they begin to recognize an intentional communicative use of attentional orientation; sounds and gestures that previously were only comprehended as dyadic emotional exchanges are now understood as a volitional attempt to manipulate the attention of the child with respect to external objects (Tomasello 2000: 64).

The child understands the special behaviour of the adult as an invitation to participate in an intentional act. Not as an agent directing its own intentional acts towards a pregnant object, as when she is reaching out for a toy. Nor as a direct object of some other agent's intentional acting, as in the act of changing diapers, for instance.

The recognition of semiotic agency thus triggers a special governing cognitive schema (Brandt 2000: 4)[6]: A syntactic structure constituting a constrained set of actantial roles; first-person subject (addresser), a second person dative-object (addressee) and a third-person direct object (content referent). A basic property of this schema is the 'directedness' of the act, going from the first-person agent via the second-person patient and further on to the third-person object of attention. It can be phrased simply: "I want you to attend to X [= the object of joint attention]".

The structural properties of the schema can be illustrated with a simple example. Say I want to direct the attention of a child to a specific point in space, e.g. to share the experience of a passing dog in the street. Then one of the ways to go about it is to establish eye contact with the child and point to the dog using my arm and index finger. If my intention is recognized, the child will redirect her attention away from my eyes, along the imaginative line extending from my finger and aiming at the dog. Declarative pointing is in fact very illustrative in explaining the proposed structure of semiosis, as it features a decomposition of the communicative intention and the content specification of the message in two discrete acts: The eye gaze of the addresser can be interpreted as signifying: "This is a message! I want you to attend to this [the object pointed to]", while the act of pointing does the work of filling in the brackets by naming the referent content – the dog.

Comprehension of the directionality of communicational acts is crucial to the acquisition of intentional semiotic behaviour. In early pre-linguistic language acquisition, the child must be able not only to recognize the intentional state of an adult-agent with respect to a third-person goal-object, but also to recognize it as directed toward her own attentional states. This requires that the child is able to monitor herself as another intentional agent intended to engage in a specific 'role' of a

[6] Brandt (2000) suggests that this schema is innately determined. I am a little more hesitant in that respect, but it is a fact that even very young infants are a lot better at grasping the basic "architecture of interaction" than any of our fellow primates (see for instance Tomasello *et al.* (1997).

joint activity. When the child comprehends the intentional state of adult as an invitation to share attention, it will follow the adult agent's attentional direction with respect to the third object.

The next step is to understand this fixed set of roles as interchangeable. Thus, to fully acquire semiotic behaviour, the child must engage in role reversal imitation. Again, this is not a simple task. The child cannot merely substitute herself for the adult with respect to a third-person goal-object as in other kinds of intentional doings. In the act of communication, the primary intentional goal is actually the attentional state of the second-person addressee. The child must thus also substitute the adult for herself in the role of second-person dative object, i.e. she must understand that the act of signification is directed towards the second-person and not directly at the reference object (Tomasello 1999: 103).

The triadic structure of semiotic exchange is what distinguishes it from the other object-object and subject-object oriented types of agency and is maybe the best argument for the establishment of semiotic agency as a discrete type.

Yet another constraining property of semiotic agency is its spatial structure. In opposition to the spatial structure of actional (mental) and cognitive causality that would have either a mental cause and a physical effect (intentional acting) or a physical cause and a mental effect (re-acting), both the cause and effect of semiotic causality seem to have a mental ontology: The addresser's inclination to share an experience constitutes the cause, while the reorientation of the addressee's attention is the effect. But the causal connection can only be established through an intermediate chain of causal events in physical space. The very act of signification must be physical. All kinds of signifying gestures need to have a minimal physical ontology, i.e. they have to be perceptually traceable in physical space in order to reach the addressee (Sinha 1999: 11). On the other hand, any object or physical property can gain referential symbolic meaning if an agent intends it to do so and the recipient recognizes this intention. Intentionality is thus both a necessary and sufficient criteria for symbolicity (DeLoache 2004: 67). We can easily establish ad hoc symbols, agreeing that the spoon is a car or (to take up Leslie's example) that a banana is a telephone, in a specific game or situation. But like the case of declarative pointing and other communicative gestures treated above, the intention that 'this object or gesture is now to be understood symbolically' is mediated through expressive behaviour such as gaze and facial expression[7]. It is the intentional *acting with the object* and not the object itself that makes it a symbol. And a moment later, the spoon will be a spoon again and the banana will be eaten.

While we still have all three participants of the communicative situation present in the same setting, with the addresser and the addressee facing each other, we can explain the signification and recognition of communicative intention by extra-linguistic conditions such as eye gaze, gesture etc. But much of our everyday communication does not proceed in these face-to-face intersubjective contexts. Throughout cultural history, oral communication and ad hoc symbolization has been supplemented with other modes of communicative exchange, using conventional cultural artifacts such as graphically represented symbolic sign systems, figurative and auditory media, aesthetic objects, etc. But when the sign does not any longer have the intentional, behavioural support, the accompanying gaze and gestures, what then constitutes the 'addressing' mechanism of symbolic artifacts? Or put differently: How is the communicative intention signified in such a way that the potential addressee knows that the object at hand is to be approached as a piece of symbolic communication? That she is supposed to take on the role of a dative-object in relation to the object? Let us first have a look at written language.

There is no doubt that the 'invention' of linguistic sign-vehicles has been of tremendous importance for human cultural evolution, as it made an especially effective accumulation and exchange of cultural knowledge possible. By means of 'materializing' the oral, symbolic gestures, communicational exchange can be detached from the moment of utterance, i.e. a message can be produced in the absence of the addressee and received in the absence of the addresser. The message is less transient: potentially numerous subjects can continuously consult it over time. But another important property of such conventional symbol systems is that they are artifacts, (almost) exclusively employed as a medium for interaction. Unlike the situation with the spoon-car and the banana-telephone, we are not dependent on the 'acting out' to know that they are intended symbolically. The symbol-objects are not of any particular interest in their own respect as they do not lend themselves to other instrumental activities (for instance they can't be eaten). When fully acquired, linguistic signs are so cognitively

---

[7] The recognition and interpretation of facial expression probably constitutes the most basic level of primate expressive behavior and we even seem to be equipped with special neural facilities to recognize and interpret facial expressions.

integrated it almost takes an academic degree to be able to disregard the referential aspects and appreciate their phonological or graphic object-like properties[8].

In the process of conventionalization or entrenchment of a symbolic sign system, the sign-objects themselves seem to become 'bearers' of the communicative intentionality. Consequently the communicative intention of the addresser is *not* explicit and immediately distinguishable from the referential function, the 'pointing'. They do not, it would seem, constitute discrete properties. Let us take the example of linguistic texts: Confronted with a piece of written verbal text, we will probably not doubt that it is a symbolically coded message, intentionally produced by someone for interactional purposes. We recognize this even before reading it. In fact, we are usually capable of recognizing intended linguistic communication even when we cannot 'access' the semantics of the specific sign system, e.g. a foreign language.

But far from all of our communicative uses of objects and artifacts are conventionalized in the way that verbal language is. Cultural convention as such cannot solve our problem. For instance, we readily recognize the addressing intention in works of conceptual art that do not in any way reproduce a cultural convention. If by coincidence we should come across Marco Evaristti's red iceberg[9] (not knowing that it was an artistic installation) we would probably spontaneously categorize such an unexpected variation of our surroundings as socially motivated and communicatively significant. This is due to a dynamic temporal reading of the seemingly static scenario. In other words, we understand it in terms of its causal history (Leyton 1992: 157). From the organization of the scene we infer the type of events, acts and agents that could have caused it. Naturally, these kinds of inferences are not one-to-one reconstructions of the past but rather work as a kind of natural probability calculation: What is the most likely causal process leading to the present state?

Non-expected variations of our surroundings can be caused by natural forces or by accidental or intentional instrumental human acts and each of these kinds of cause leave their more or less specific traces on the scenario. A set of such traces will motivate a communicative reading. That is, the scenario actualizes a previous intentional act that is directed at the manipulation of our attention. If we return to the

---

[8] This is probably the main motivation when Andy Clark (1997) states that verbal language is "the ultimate artefact". He hereby stresses the well-integratedness of an external symbol system in our cognitive processes and compares it to other kinds of (more instrumental) tool-manipulations.

[9] See pictures of Marco Evaristti's installation "The Ice Cube Project" at the URL: http://www.evaristti.com/Work/performance/ICE/index.htm

iceberg scenario, it tells the story of an agent giving at least some effort and concentration on (re)organizing the scene in a way that strikes us as purposeful. Icebergs are simply *not* red. Of course we could imagine some one spilling some paint or a dining polar bear leaving some traces of seals' blood. But a crucial fact about this particular scene is that the iceberg is entirely covered with red colour. The probability that this could have happened by chance is infinitesimally small. We thus have two principles guiding our intuition that this is intended by some agent to be a piece of aesthetic communication: 1) The bringing together of two elements, an iceberg and red colour, that are not normally part the same scenario, and 2) a certain 'well-orderedness' or symmetry in the way the elements are arranged or combined. Analyses of a series of other communicative scenarios reveal a similar pattern: The combination of the two principles seems to be of a general character[10].

Communicative signs constituted by scenarios such as the Ice Cube Project do not gain symbolic meaning by bare recognition of cultural conventions. They do so by reference to the intentional causal history we are able to reconstruct from them. This is due to the interpretative nature of sign types. The symbolicity of a sign-object is not an intrinsic property of the object, but is dependent on the attitude of the interpreter. We can thus approach the same object with regard to its similarities with other objects (iconic reference), or its correlations with other things (indexical reference). Furthermore, we can approach the object with regard to its "*involvement in systems of conventional relationships*", i.e. as a symbol (Deacon, 1997: 71). Though a sign-object is intentionally used in only one of the referential functions, say symbolically, and interpreted as such, there might be intermediate interpretative processes in the reception of the sign. This is at least indicated by the structure of symbolic reference proposed by Terrence Deacon.

## The hierarchical nature of symbolic reference

In his interesting work *The Symbolic Species* (1997), Terrence Deacon proposes a cognitive model of symbolic reference hierarchically based upon sign systems of icons and indices. Founded in evolutionary neurobiology, his point is that the enigmatic phenomenon of symbolic reference did not emerge out of nothing but can be

[10] Here I refer to a work in progress related to my PhD thesis on communicative intentions and symbolic artifacts.

approached as a higher order extension of existing sign systems shared with our fellow primates. The building blocks for symbolicity were already present in earlier stages of our neural biological development. To roughly sketch Deacon's model, indices rely on relations among icons: Our immediate recognition of a perceptual phenomenon, say the smell of smoke, is due to its iconic relation to previous experiences of smoke. But it also tells us that something is burning. Former experiences of spatial contingency of fire and smoke create a stable correlation between the icons. Whenever confronted with one of them we will expect the co-presence of the other. Smoke means fire. Central to the conditional nature of indexical reference is thus the continuous spatial or temporal contiguity of sign and reference. Should we suddenly begin to have experiences of smoke that are not accompanied by fire, the reference will soon break down (Deacon 1997: 82). Likewise, when we teach our dog to respond to linguistic signs like "food", "sit" or "time for a walk" it will simply understand the correlation of that phonological string and some activity or object. The reference is indexical (though we may expect it to be symbolic) and can only be maintained through a continuous correlation. That is the funny thing about symbols: They do not depend on contiguity of sign and reference. We can easily point to objects that we do not have any prior experiences of, or which do not exist, such as dragons, UFO's and teenage mutant ninja turtles. Many children have words like "tiger" and "witch" in their early vocabulary, despite the fact that they have not had any (non-fictional) experiences of these phenomena.

Deacon explains this problem by the special relational character of a single symbolic reference in relation to a system of symbols. Symbolic reference is thus formed by systematic relationships between index tokens (that again are formed by icons). When a logical combinatorial system of relationships among indexical tokens is stabilized, we begin to rely on the relation of tokens and the contiguity of sign and object is no longer necessary for the maintenance of reference (Deacon 1997: 87)[11].

[11] In the early stages of language acquisition the child learns words for things already familiar. In some sense, the process goes from perceived object to symbol. But gradually, as the relational system of symbols is adapted, the symbolization process can be inverted. It now goes from word to object. The child will learn new words for which the reference still needs to be stabilized. The reference object may be less familiar or even unknown to the child. Often she will temporarily mix up the novel word's reference object (e.g. mistake 'lion' for 'tiger' etc.).

## The index of semiotic agency

The hierarchical dependency of sign types in symbolic representation may help to explain our initial problem concerning the signification of semiotic agency, though we may end up somewhere else than Deacon intended. When we are confronted with an intended symbolic sign-object, our first and immediate cognitive routine is probably one of recognition. We recognize it as a symbolic sign because of its similarities with previously experienced symbolic signs. This basic iconic reading precedes the reference aspect. But the recognition of the special 'sign features' of the object changes our attentional attitude toward the object. We no longer perceive the object as such, but relate it to its causal history; the act that created or manipulated it. We thus experience the object or scenario as a kind of metonymical extension of the intentional agent responsible for it. In this sense the sign-object is primarily indexical; the object points to its cause: The communicative intention of an agent. The triadic actantial schema is triggered and we become dative-objects. Only then does the question of symbolic reference become relevant – it is secondary to the indexical comprehension of semiotic agency.

Now the remaining question is of course the nature of 'sign features', that is, the perceivable features of an object that tell us that some agent has acted upon it in order to make it stand for something else, to signify something other than itself.

Though some of the structural properties of semiotic agency may in fact be neurobiologically hardwired, Tomasello stresses the importance of *cultural transmission* and *learning* in the acquisition and understanding of symbolic sign systems. The evolution of the sign-symbol is thus probably not to be traced in the neural architecture of the individual, but in the shared structures of cultural knowledge and cognition (Tomasello 1999: 11). Symbolic sign-objects are cultural artifacts. They cannot be compared with the ways our fellow primates communicate using smells and the like. These behaviours are innate and have evolved in evolutionary time. Thus, symbolic sign-objects are products of an accumulative cultural process: By a gradual and steady stabilization of conventional object-meaning pairings, handed over and further developed from generation to generation, symbolic language finds its present form.

In a socio-cultural world, survival is to a large degree dependent on one's ability to get on in symbolic exchanges. From the earliest stages of ontogenesis children get an intuitive understanding of symbolic signs as valuable objects. Through a kind of 'conditioning process,' symbolic signs are highly invested with pregnance. In acquiring language, the child is continuously rewarded by an increasing adaptation to its social surroundings. It is satisfactory for the child to understand and be understood, and, most importantly, to be able to participate in ever new social practices.

The general human attentional sensitivity to semiotic gesture is probably due to these kinds of socio-cultural adaptive processes. To reach a high level of adaptation, we have to attend to meaning by differentiating micro-details in the phonological and graphic behaviours that constitute language. We gradually acquire an attentional sensitivity towards a broad range of conventionalized sign-systems, from the 'language' of traffic lights and aesthetic artifacts to culturally determined behavioural practices, dress codes and complex grammatical constructions.

It might be that our experience and increasing familiarity with various conventional symbol systems motivates a more abstract understanding of ways that objects can be manipulated for symbolic purposes. We begin to recognize certain kinds of acts and object manipulations as communicatively intended, not because they reproduce a convention but because they reproduce a more abstract behavioural pattern. The bringing together of objects and new contexts, combined with the principle of well-orderedness, seems to be a powerful way of expressing oneself. In the examples treated above, it is the unconventional combination of objects, properties and context that motivates the indexical reading of communicative intentionality. It is very unlikely that these different elements would be brought together by natural forces or accidental acts. They are most probably the product of an intentional expressive act.

The same goes for our initial problem of the rock. Even though we might not be familiar with the specific conventional uses of rocks (on graves, mountain peaks, etc.), we are probably still sensitized to the perception of rocks, standing in an upright position against all probability calculations of its 'natural tendencies' (forces of nature tend to have oblong rocks lying down). The static rock scenario is thus interpreted in a dynamic fashion as part of the causal history. And the well-ordered, non-natural and non-coincidental characteristics of the scenario point to the intervention of human intentionality. We recognize the otherwise superfluous human act of erecting rocks as expressive and significant, as an index of semiotic agency[12].

---

[12] It would seem that instrumentally superfluous acts are especially suited for communicative use. Banging hammers against walls, for instance, would not qualify as a good way of creating a language, as the communicational use would be hard to separate from the instrumental. We would have to cancel all kinds of

It would seem that the borders of intersubjectivity extend far into the physical domain. What at first glance seem to be static physical states and properties of our surroundings are experienced as highly dynamic processes of interpersonal communication. Almost any variation in physical space (or even lack of expected variation) can be interpreted as socially grounded depending on the context. The 'variation' thus bears the trace of a socially motivated intentional act, which changes our attitude to the scenario – we take the symbolic stance.

Knowing the specific cultural conventional use of rocks on graves we can thus access the intended reference of the symbol. But the recognition of a communicative intentional act is always prior to the decoding of referential meaning.

## References:

Amodio, M.A. & Frith, C.D. (2006). 'Meeting of Minds: the medial frontal cortex and social cognition'. *Nature Reviews, Neuroscience,* vol. 7: 268-77.

Brandt, P.Aa. (2004). 'From gesture to theatricality'; in P.Aa. Brandt. *Spaces, Domains, and Meaning.* Bern: Peter Lang Verlag.

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

Deacon, T.W. (1997). *The Symbolic Species – The Co-evolution of Language and the Brain.* New York: W.W. Norton.

DeLoache, J.S. (2004). 'Becoming symbol-minded'. *Trends in Cognitive Sciences,* 8 (2): 66-70.

Gallese, V., & Metzinger, T. (2003). 'Motor ontology: the representational reality of goals, actions and selves'. *Philosophical Psychology* 16 (3): 365-388.

Leslie, A.M. (1987). 'Pretense and representation: The origin of "theory of mind"'. *Psychological Review*, 94 (4): 412-426.

Leslie, A.M. (1993). 'A theory of agency'. *Technical Reports of the Rutgers University.* London: Center for Cognitive Science.

Leslie, A.M. (1994). 'ToMM, ToBy, and Agency: Core architecture and domain specificity'; in L.A. Hirschfeld & S.G. Gelman (eds.). *Mapping the Mind: Domain specificity in cognition and culture* (pp.129-148). Cambridge: Cambridge University Press.

Leslie, A.M. (2000). '"Theory of Mind" as a Mechanism of Selective Attention'; in M.S. Gazzaniga (ed.). *The New Cognitive Neurosciences, 2nd edition* (pp. 1235-1247). Cambridge, MA: MIT Press.

Leyton, M. (1992). *Symmetry, Causality, Mind.* Cambridge, MA: MIT Press

Oakley, T. (2003). *A Grammar of Attention. A Treatise on the Problem of Meaning.* http://www.cwru.edu/artsci/engl/oakley/papers.htm

Sinha, C. (1999). 'Grounding, Mapping and Acts of Meaning'; in T. Janssen & G. Redeker (eds.). *Cognitive Linguistics: Foundations, Scope and Methodology* (pp. 223-255). Berlin & New York: Mouton de Gruyter.

Talmy, L. (2000a). *Toward a Cognitive Semantics.* Vol. I: *Concept Structuring Systems.* Cambridge, MA: MIT Press.

---

instrumental banging. Small finger rings, big bronze statues, silly dance steps and strange oral behaviors etc., on the other hand, are excellent for expressive use as we could not imagine them employed in any other sensible connections.

Talmy, L. (2000b). *Toward a Cognitive Semantics*. Vol. II: *Typology and Process in Concept Structuring*. Cambridge; MA: MIT Press.

Thom, R. (1990). *Semio Physics: A sketch*. US: Addison-Wesley.

Tomasello, M., Call, J. & Gluckman, A. (1997). 'Comprehension of Novel Communicative Signs by Apes and Human Children'. *Child Development*, vol. 68 (6): 1067-80

Tomasello, M. (1999). *The Cultural Origin of Human Cognition*. Cambridge, MA: Harvard University Press.

Tomasello, M. (2000). 'First steps towards a usage-based theory of language acquisition'. *Cognitive Linguistics*, 11 (1-2): 61-82.

Østergaard, S. (2000). 'Mental Causation'. Århus: Center for Semiotics, University of Aarhus.

Østergaard, S. (1998). *Kognition og katastrofer – studier i dynamisk semiotik*. Copenhagen: Basilisk.

ARTICLE IN PRESS

Brain and Language xxx (2008) xxx–xxx

Contents lists available at ScienceDirect

Brain and Language

journal homepage: www.elsevier.com/locate/b&l

# Say it with flowers! An fMRI study of object mediated communication

Kristian Tylén [a,b,], Mikkel Wallentin [b,c], Andreas Roepstorff [b,d]

[a] *Institute of Language and Communication, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark*
[b] *Center for Functionally Integrative Neuroscience, Aarhus University Hospital, Nørrebrogade, 8000 Aarhus C, Denmark*
[c] *Center for Semiotics, University of Aarhus, Jens Chr. Skous Vej 7, 8000 Aarhus, Denmark*
[d] *Institute for Anthropology, Archaeology and Linguistics, University of Aarhus, Moesgaard, 8270 Højbjerg, Denmark*

A R T I C L E I N F O

*Article history:*
Accepted 24 July 2008
Available online xxxx

*Keywords:*
Material objects
Mediation
Non-verbal communication
Pars triangularis
Conventionality
Material signals

A B S T R A C T

Human communicational interaction can be mediated by a host of expressive means from words in a natural language to gestures and material symbols. Given the proper contextual setting even an everyday object can gain a mediating function in a communicational situation. In this study we used event-related fMRI to study the brain activity caused by everyday material objects when they are perceived as signals. We found that comprehension of material signals activates bilaterally areas of the ventral stream and pars triangularis of the inferior frontal cortex, that is, areas traditionally associated with verbal language and semantics. In addition, we found that right-hemisphere inferior frontal cortex is recruited as a function of the increasing unconventionality of communicative objects. Together these findings support an interpretation of the traditional language areas as playing a more general role across modalities in relation to communicational mediation of social semantic meaning.



## 1. Introduction

Imagine two scenes. The first is a cluster of wild flowers on the roadside. The second is a similar bunch of flowers but cut, tied, and placed on the doorstep of a private house. Though both scenes are relatively common, there are great differences in the way we make sense of them. While the first is a natural and thus 'unintentional' scene, the latter seems to call for another kind of interpretation. Flowers do not normally end up on doorsteps by coincidence or forces of nature. But we can easily imagine that someone left them intending another person to find them and associate them with social semantic meaning—perhaps a declaration of love. In this case, the flowers are not just flowers; they are also mediating meaning in an interpersonal communication. It is as if this static constellation of objects does a job usually done by words in conveying meaning between interacting agents (H. Clark, 1996, 2005).

There is a host of ways that two human agents can engage in joint meaning constructing activities exploiting different resources of their shared material space. An extensive number of studies have addressed various aspects of the way we use verbal language to mediate interpersonal semantic meaning and the neural structures underlying such processes (cf. e.g. Bedny, Hulbert, & Thompson-Schill, 2007; Goldberg, Perfetti, Fiez, & Schneider, 2007; Hickok & Poeppel, 2007; Moss et al., 2005; Papathanassiou et al., 2000; Pietrini et al., 2004; Price, 2000; Rodd, Davis, & Johnsrude, 2005; Rodd et al., 2005; Stowe, Paans, Wijers, & Zwarts, 2004; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997; Vigneau et al., 2006; Wagner, Paré-Blagoev, Clark, & Poldrack, 2001; Wallentin, Roepstorff, Glover, & Burgess, 2006). But language may not be the only means by which we engage in the joint meaning constructing activities. Recent brain imaging studies of communicative gestures (Lotze et al., 2006; Nishitani, Schürmann, Amunts, & Hari, 2005; Willems, Özyürek, & Hagoort, 2007), nonverbal vocalizations (Dietrich, Hertrich, Alter, Ischebeck, & Ackermann, 2007) and facial and bodily expressions (Lawrence et al., 2006) suggest that nonverbal mediators of interpersonal meaning might not only function in ways comparable to linguistic words, but may also be associated with activation of brain regions traditionally associated with language functions (e.g. left inferior frontal gyrus). The current investigation expands upon this line of thought by addressing another kind of non-verbal human communication—the way we exploit everyday material objects as signals in joint meaning constructing activities (H. Clark, 2005).

The notion of *material signals* is adapted from Herbert H. Clark's article, "Coordinating with each other in a material world", in which he discusses various ways that material objects can be integrated in communicative interactions to accomplish coordination and signal meaning between people. Clark's general finding is that many non-verbal signals, including intentional manipulations of material objects, are used in ways comparable to and interchangeable with linguistic signals (2005, 1996). In this context we will concentrate on a specific subtype of communication mediated by material signals, i.e. a person's encounter with a non-ostensive,

* Corresponding author. Fax: +45 65932483.
*E-mail address:* kristian.tylen@language.sdu.dk (K. Tylén).

0093-934X/$ - see front matter 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.bandl.2008.07.002

Please cite this article in press as: Tylén, K., et al. Say it with flowers! An fMRI study of object mediated communication. *Brain and Language* (2008), doi:10.1016/j.bandl.2008.07.002

static material scene arranged in a purposeful manner to signal social meaning. In the absence of an actively communicating addresser, the comprehension of such material signals relies on the categorical and teleological inferences of the addressee (Bloom, 1996; Goodman, 1976; Kelemen, 1999; Pickering & Garrod, 2004). In this context, successful material signaling thus depends on the addressee's inclination to pick up on intentional materialized cues left behind by the addresser and to take the appropriate interpretative stance towards the scenes, that is, explore the scenes as a means for social meaning construction (A. Clark, 2006; Dennett, 1987; H. Clark, 1996; Levinson, 1993).

Like with other forms of communicative mediation (including verbal language), the use of material signals can be aided by a mutual orientation to conventions (H. Clark, 1996). As noted by Pickering and Garrod (2004), conventionality might play an especially important role in cases of offline one-way communications where interlocutors are not simultaneously present and thus cannot rely on the continuous negotiation and alignment of representation. The recurrent and stable use of a specific material item as a communicative signal is in these situations likely to enhance the efficiency of the communication, solve ambiguities and prevent misunderstandings (H. Clark, 1996). The comprehension of unconventional material signals, on the other hand, will tend to rely heavily on contextual integration and inferential modeling of the addresser's intentions and may generally be more costly to process (Pickering & Garrod, 2004).

In this study we used event-related fMRI to assess the neural correlates of subjects' categorical intuitions in relation to various kinds of material signals, contrasted with non-communicative material scenes. Motivated by the conceptual considerations outlined above (e.g. H. Clark, 2005) we hypothesized that material signals are processed in ways comparable to other kinds of communicative signals, including language, gesture and bodily expressions. Furthermore, we predicted that the degree of conventionality of the material signals would influence these activation patterns.

The study was carried out in two parts. In a categorization study we addressed the issue of which particular configurations of objects would be consistently perceived as communicatively relevant. The study provided us with a corpus of 50 paired images depicting actual occurring scenarios where one was experienced as communicative and the other as non-communicative. These image pairs were then used as stimuli in an event-related fMRI experiment to investigate the neural activations underlying the comprehension of static material objects and artifacts used as communicational signals.

## 2. Methods and materials

### *2.1. The categorization study*

Prior to the brain imaging experiment we carried out a web-based electronic questionnaire study to test the comprehension of various configurations of everyday objects. A dataset of 152 color photos of actually occurring scenarios containing material objects and artifacts was used as test material. All images were collected by the experimenters from the local environment, using a Canon Eos D300 digital reflex camera, and from web-based photo databases and photo sharing sites such as http://www.polfoto.dk and http://www.flickr.com. Images were picked and arranged in pairs where one depicted a scenario containing everyday objects manipulated in a striking and deliberate way for purposes of interpersonal communication while the other depicted the same objects in natural, utilitarian or accidental situations without any conspicuous intentional communicative relevance.

Images of communicative objects were chosen to represent a broad scale of different expressive means and degrees of conventionality from material symbols (e.g. arrows) to unconventional objects configurations (e.g. aesthetic installations) (for examples of image pairs, see Fig. 1). All images were preprocessed in Adobe Photoshop 7.0 image editing software for normalization of color saturation, light and contrast. For the sake of ecological validity, no mask or other kind of manipulation was applied to the images.

Sixty three informants (21 females and 42 males, mean age 34.6 ± 10.3 (STD)) participated in the categorization study. Before entering the questionnaire section participants received careful instructions regarding the focus of the study and the specific task at hand. A link allowed participants to return to these instructions at any time during the questionnaire section. In the questionnaire section an image (in a pseudo-randomized order) would appear at the top of the screen while four written questions (the same throughout the experiment) were displayed at the bottom. Responses were given by checking radio buttons. Only one response could be given for each image. The Danish questions would translate into something like: (1) "Do you experience the objects depicted as part of a natural (non-intentional) scene?", (2) "Do you experience the objects as part of an accidental (non-intentional) scene?", (3) "Do you experience the objects as part of a 'private' intentional project (not directed at an addressee)?", or (4) "Do you experience that the objects depicted are arranged in order to communicate something to someone (directed at an addressee)?". Alternatively the participants could choose to answer "I don't know". While question 1–3 was given primarily to help participants make the relevant categorial distinctions, the responses to question four (the communicative vs. non-communicative distinction) was analyzed and used as a measure for the selection of stimuli for the fMRI-study.

### *2.2. Results of the categorization study*

Generally there was a high degree of consistency in the responses across informants. That is, the informants largely agreed which objects are comprehended as communicatively relevant. In a binominal test carried out in MATLAB (Mathworks Inc., Sherborn, MA, USA) with a significance threshold at $p < .05$, 70/152 (or 46 pct.) of the images were rated as significantly communicative while 62/152 (or 40 pct.) were rated as non-communicative showing a conspicuous category effect (Haushofer & Kanwisher, 2007). This is interpreted by the experimenters to support the relevance of the categorical distinction adapted from the studies of gesture (e.g. Lotze et al., 2006) and vocalization (Dietrich et al., 2007) in the perception of object configurations.

### *2.3. The brain imaging experiment*

#### *2.3.1. Subjects*

Twenty-two healthy, right handed volunteers (10 females and 12 males) mean age 24.5 ± 2.3 (STD) participated in the experiment. All participants gave informed written consent in accordance with requirements of the local medical Ethics Committee. Subjects received no fee for participation.

#### *2.3.2. Stimuli*

Based on the results from the categorization study we included the 100 images (50 image pairs) with the highest consistency in ratings across informants, as stimuli in the brain imaging experiment. The stimuli were divided into a test and a control condition according to their ratings in the communication vs. non-communication trial ('yes' for test and 'no' for controls). In the control condition, images depicted objects found to be without conspicuous communicational significance, while the test condition depicted

Please cite this article in press as: Tylén, K., et al. Say it with flowers! An fMRI study of object mediated communication. *Brain and Language* (2008), doi:10.1016/j.bandl.2008.07.002

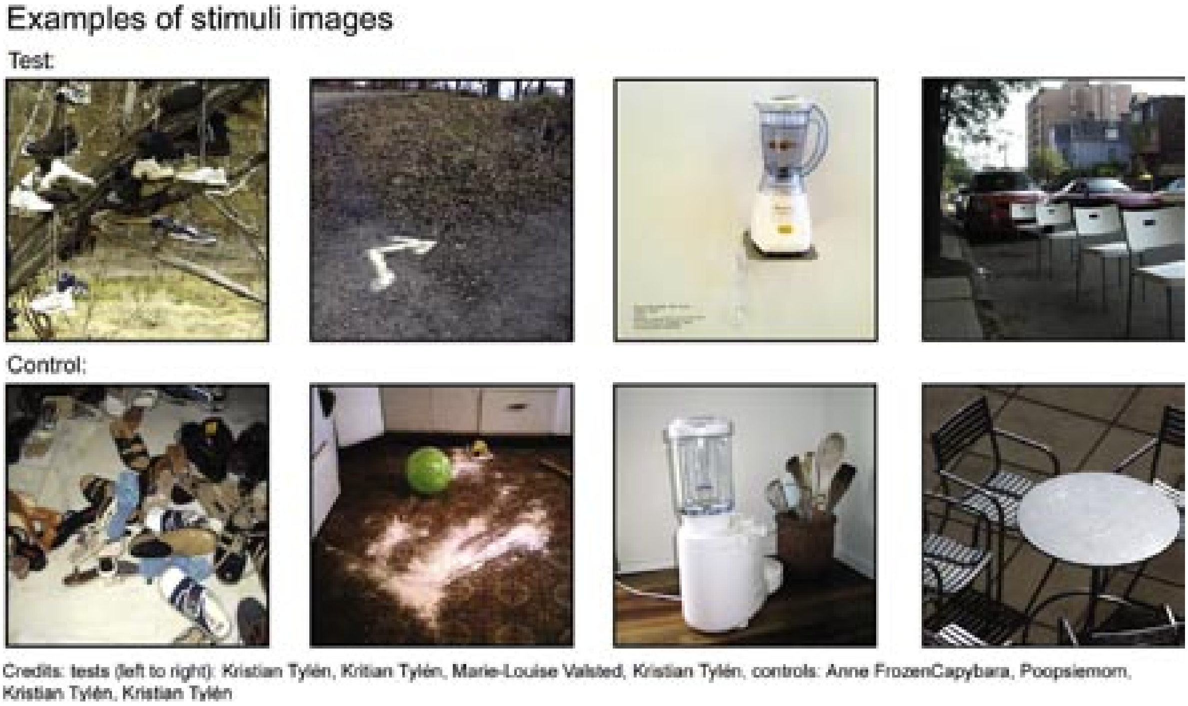



**Fig. 1.** Examples of stimuli image-pairs. Test images include static objects and artifacts perceived as means for intentional communication. Control images include objects without any conspicuous communicational significance.

the same objects found to be intentionally manipulated for the sake of interpersonal communication.

Examples of control stimuli included photos of shoes scattered on the floor, flour accidentally spilled on the floor, a blender on a kitchen table, and chairs around a table.

Examples of test stimuli included photos of shoes hung by their laces from a tree, flour sprinkled in the shape of an arrow in the forest floor to indicate a direction, a controversial piece of art, consisting of a blender placed on a museum stand with a goldfish in it, and chairs put out in the street to reserve a parking lot (see Fig. 1).

*2.4. Experimental design*

Stimuli were presented using the software extension Cogent 2000 (Wellcome Department of Imaging Neuroscience, University College London, London, UK; http://www.fil.ion.ucl.ac.uk) for MATLAB (Mathworks Inc., Sherborn, MA, USA). Control and test stimuli were semi-randomized for each subject, with the constraint that control images (e.g. the blender without the goldfish) would always precede the corresponding test images (the blender with the goldfish in it) by a mean spread of 12 images. This was done to counter the 'contagious effect' of the communicatively relevant test images on the neutral controls that we found would potentially be a stronger confounding factor than the bare 'repetition effect'.

We used a passive viewing task combined with subjects' introspective post-scan reports (Jack & Roepstorff, 2002). Subjects received no information regarding the theoretical focus of the investigation prior to scanning, but were instructed to keep their eyes open and concentrate on the images. Each image was shown for 5000 ms, followed by a grey screen with a black fixation cross for a jittered delay of 2000–5000 ms (mean 3500 ms). Stimuli were projected onto a screen and then reflected into the visual field of the subject by a mirror attached to the head coil of the scanner.

After scanning sessions, subjects were presented with the same stimuli outside the scanner, and were tested in an electronic questionnaire targeting their comprehension of each image with respect to parameters such as the communication vs. non-communication distinction, degree of conventionality and emotional intensity. Thus, as in the categorization study subjects were asked to asses whether or not the depicted scenes contained objects or properties they believed were manipulated with the purpose of intentional communication. Affirmative responses were followed up with a question asking subjects to rate on a four-point scale the degree of conventionality of the communication (with 1 being most unconvential and 4 being most conventional). Conventionality was defined in terms of relative frequency, that is, how often subjects would expect to encounter a certain type of material signal. Again, this was followed by a question assessing the emotional intensity of stimuli, which was included to rule out the influence of this parameter on the analysis. These responses were later used in the statistical modeling of scanning data for each subject.

*2.5. Scanning parameters*

Functional images were acquired using a General Electronics 3-T MR system with a 16 channel NOVA head coil. For each subject, 320 contiguous multislice T2*-weighted images were obtained using an echo-planar imaging sequence with the following parameters: repetition time (TR): 3000 ms, echo time (TE): 30 ms, and flip angle: 90 . Thirty-nine sequential, interleaved 3.5 mm axial slices were obtained per volume with a 128 128 pixel resolution matrix and a field of view of 240 240 mm.

*2.6. Data analysis*

FMRI data were spatially realigned (Friston et al., 1995a), unwarped (Andersson, Hutton, Ashburner, Turner, & Friston, 2001),

ARTICLE IN PRESS

4 K. Tylén et al. / Brain and Language xxx (2008) xxx–xxx

slice-time corrected, normalized (Ashburner & Friston, 1999) to the MNI template, and smoothed (10 mm FWHM) using SPM5 (Statistical Parametric Mapping, Wellcome Department of Imaging Neuroscience, University College London, London, UK; http://www.fil.ion.ucl.ac.uk) executed in MATLAB (Mathworks Inc., Sherborn, MA, USA). Task-related BOLD-responses for each subject were estimated using a general linear model (Friston, Holmes, Poline, Frith, & Frackowiak, 1995b) with a 128-s highpass filter. All events were modeled using the standard hemodynamic response function of SPM5. Individual *t*-contrasts were created from the estimated -weights and sent to a second level random effect analysis. Significance levels were set to $p < .05$, FDR-corrected for multiple comparisons (Nichols, 2007).

Two first-level analyses were conducted. The first of these (hence 'FLA1') modeled onsets for the two conditions (communication/non-communication), in accordance with the predetermined organization of stimuli in two equal-size groups of each 50 events, based on results form the categorization study. The second analysis, hence 'FLA2', was modeled in concordance with the subjects' individual responses to the post-scan questionnaire. The model included a regressor for all non-communicative images and a regressor for all communicative images. A separate parametric regressor modeled each subject's rating of the degree of conventionality/entrenchment of the communicative use for each of the images rated to be communicational, i.e. this analysis would pick up on regions of the brain where the BOLD response linearly increased/decreased as a function of increased or decreased conventionality ratings (on a 1–4 point scale). The FLA1 contrast was submitted to a second-level RFX analysis using the one-sample *t*-test option in SPM5 with the significance threshold set to $p < .05$ (FDR corrected for multiple comparisons). The result was subsequently used as to mask the FLA2 analysis, so that only statistical differences found in both analyses were counted as activations. Results were again thresholded at $p < .05$ (FDR corrected for multiple comparisons).

In addition, a contrast was made from the conventionality regressors and masked by the main effect to test which areas were sensitive to the rated degree of conventionality of the communications. Threshold was set to $p < .05$ (FDR corrected for multiple comparisons).

Putative anatomical regions were located using WFU (Wake Forest University School of Medicine) Pickatlas (Maldjian, Laurienti, & Burdette, 2004; Maldjian, Laurienti, Kraft, & Burdette, 2003) referencing the aal (automatic anatomical labeling) atlas (Tzourio-Mazoyer et al., 2002).

### 2.7. Debriefing results

There was a significant correspondence between responses to the post-scan questionnaire and the pre-experimental categorization study (two-sample unequal variance *t*-test, $p < .001$, cf. Ruxton, 2006, mean correspondence = 80.2% ± 12.6 (STD)). However, in some cases the introspective procedure for classifying test and control events resulted in an unbalanced fit (e.g. more events rated more communicative than non-communicative). Together with the pronounced variation in responses this motivated the use of masking in the statistical analysis of scanning data as described above. Results from the conventionality ratings from each subject were included as a regressor in the first-level SPM analyses as described above. As expected, the questionnaire ratings of the emotional intensity of stimuli did not show any significant interactions with the other measures and were discharged from further analysis.

### 2.8. fMRI results

Communicative > non-communicative: viewing images of objects and artifacts experienced as communicatively relevant, relative to objects and artifacts with no communicative relevance, activated a network in the ventral visual stream bilaterally, and

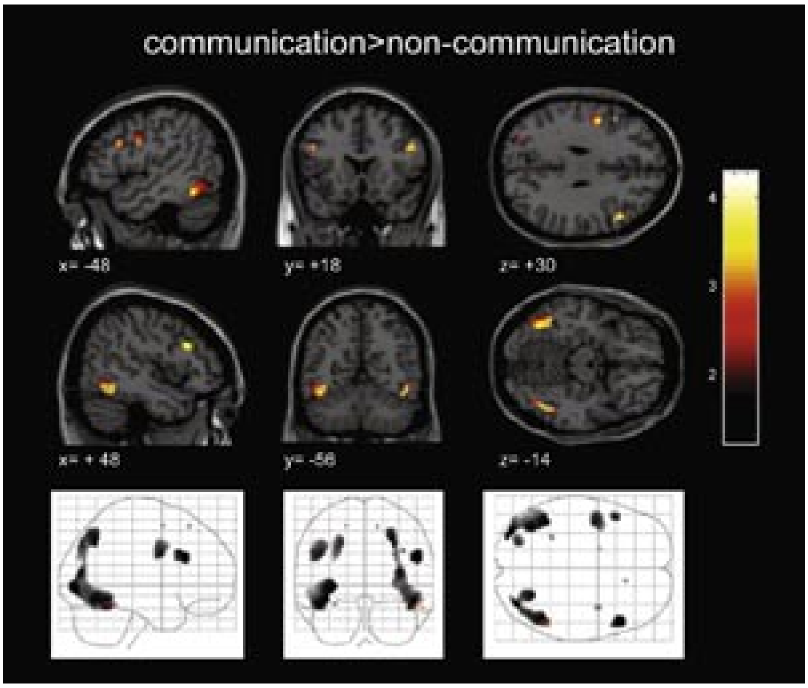



**Fig. 2.** Communication > non-communication: subjects' comprehension of static objects and artifacts used for purposes of interpersonal communication. Significance threshold: $p < .05$, FDR-corrected for multiple comparisons. (Left hemisphere is projected onto the left side of the depicted brain and vice versa.)

Please cite this article in press as: Tylén, K., et al. Say it with flowers! An fMRI study of object mediated communication. *Brain and Language* (2008), doi:10.1016/j.bandl.2008.07.002

**Table 1**
Communication > non-communication

| Putative anatomical regions | BA | Hemisphere | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Left | | | | Right | | | |
| | | *Z*-score | Coordinates | | | *Z*-score | Coordinates | | |
| | | | *X* | *Y* | *Z* | | *X* | *Y* | *Z* |
| Inferior frontal, PTr | 45 | 2.96 | 48 | 18 | 30 | 3.40 | 52 | 22 | 26 |
| Precentral sulcus | 6 | 3.06 | 46 | 2 | 30 | 2.28 | 38 | 2 | 34 |
| Inferior temporal lobe | 37 | 3.44 | 44 | 50 | 18 | 3.22 | 48 | 56 | 14 |
| Fusiform gyrus | 19/37 | 3.24 | 38 | 64 | 12 | – | | | |
| Superior occipital lobe | 19/39 | 2.86 | 24 | 78 | 38 | 3.32 | 28 | 64 | 38 |
| Inferior occipital lobe | 18/19 | 3.57 | 34 | 86 | 4 | 3.16 | 34 | 80 | 12 |

Subjects' comprehension of static objects and artifacts perceived as means for intentional communication. Significance threshold: $p < .05$, FDR-corrected for multiple comparisons. PTr is short for pars triangularis.

inferior prefrontal areas bilaterally. Peak regions included the posterior part of inferior temporal lobe bilaterally (BA 37, MNI [ 44, 50, 18] and [48, 56, 14]), the left fusiform gyrus (BA 19, MNI [ 38, 64, 12]), the superior occipital cortex bilaterally (BA 18/19, MNI [ 24, 78, 38] and [28, 64, 38]) and the pars triangularis of the inferior prefrontal lobe bilaterally (BA 45, MNI [ 48, 18, 30] and [52, 22, 26]) (see Fig. 2 and Table 1). The reverse contrast revealed no significant activity.

Non-conventional > conventional regression analysis: Images rated by subjects as generally more unconventional activated regions in the inferior occipital lobes bilaterally (BA 18/19, MNI [ 32, 88, 4] and [32, 80, 12]), in the right fusiform gyrus (BA 37, MNI [32, 46, 16]) and in the pars triangularis of right inferior prefrontal lobe (BA 45, MNI [50, 20, 30]) (see Fig. 3 and Table 2).

## 3. Discussion

We found that the perception of material objects used for means of interpersonal communication elicits activity in brain areas commonly associated with human language and communicative behavior. Both the pars triangularis in the inferior prefrontal lobes and the posterior inferior temporal lobes (including the so-called Visual Word Form Area of fusiform gyrus) are widely held to be engaged in the processing of human language and gesture (Bedny et al., 2007; Cohen et al., 2000, 2002; Dehaene, Le Clec, Poline, Le Bihan, & Cohen, 2002; Goldberg et al., 2007; Hickok & Poeppel, 2007; Lotze et al., 2006; Moss et al., 2005; Papathanassiou et al., 2000; Pietrini et al., 2004; Price, 2000; Rodd et al., 2005; Skipper, Goldin-Meadow, Nusbaum, & Small, 2007; Thompson-Schill et al., 1997; Vigneau et al., 2006; Wagner et al., 2001; Wallentin et al., 2006; Stowe et al., 2004). The inferior temporal lobes are part of the ventral visual pathway that is consistently associated with visual object recognition and semantic integration (Desimone & Ungerleider, 1989; Ungerleider & Mishkin, 1982). Recent brain imaging studies suggest that the ventral pathway is not modality specific, but largely overlaps with other sensory modalities and thus may be related to more abstract semantic processing, including the retrieval of lexical semantics (Hickok & Poeppel, 2007; Pietrini et al., 2004; Price, 2000; Rodd et al., 2005). More specifically, the ventral stream activation peaks in the current study seem to reside in an area in medial fusiform gyrus often referred to as the Visual Word Form Area (VWFA) (Cohen et al., 2000, 2002; Dehaene et al., 2002). Though still the subject of lively debate (cf. e.g. Price & Devlin, 2003), this area is thought to play a decisive role in pre-lexical representations of words, letter strings, and possibly symbols in general (Kronbichler et al., 2004; Reinke, Fernandes, Schwindt, O'Craven, & Grady, 2008).

In the present study, we reason that the activations in the inferior temporal lobe and fusiform gyrus were associated with the categorical resolution: communication vs. non-communication (Levinson, 1993; Tylén, 2007). Objects and artifacts become

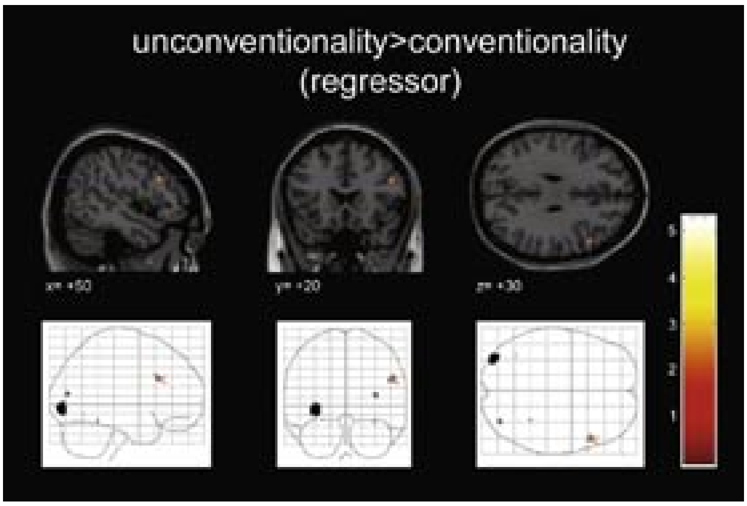



**Fig. 3.** Unconventional > conventional communication, regression analysis: brain areas activated as a function of increasing unconventionality of the communicational objects. Significance threshold: $p < 0.05$, FDR-corrected for multiple comparisons. (Left hemisphere is projected onto the left side of the depicted brain and vice versa.)

Please cite this article in press as: Tylén, K., et al. Say it with flowers! An fMRI study of object mediated communication. *Brain and Language* (2008), doi:10.1016/j.bandl.2008.07.002

**Table 2**
Unconventional > conventional communication, regression analysis

| Putative anatomical regions | BA | Hemisphere | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Left | | | | Right | | | |
| | | Z-score | Coordinates | | | Z-score | Coordinates | | |
| | | | X | Y | Z | | X | Y | Z |
| Inferior frontal, PTr | 45 | – | | | | 2.59 | 50 | 20 | 30 |
| Fusiform gyrus | 37 | | | | | 2.54 | 32 | 46 | 16 |
| Inferior occipital lobe | 18/19 | 4.17 | 32 | 88 | 4 | 3.17 | 32 | 80 | 12 |

Brain areas activated as a function of increasing unconventionality of the communicational objects, based on the subjects' post-scan ratings. Significance threshold: $p < .05$, FDR-corrected for multiple comparisons. PTr is short for pars triangularis.

interestingly complex when employed in new ways to mediate social semantic interactions. In addition to their inherent properties, they gain new referential functions due to their intentional use. They are both objects in their own right and devices for the mediation of semantic meaning (A. Clark, 2006; DeLoache, 2004; Tomasello, 1999). Interestingly, the recognition of such emergent referential functions seems to be facilitated by brain areas associated with the decoding of other kinds of visual symbol systems, such as conventional writing systems. And this could challenge the original idea that the VWFA is an area specific to verbal language (Cohen et al., 2000, 2002; Dehaene et al., 2002).

The involvement of the pars triangularis (PTr) in the dorsal part of inferior frontal cortex (IFC) in language processing has been recognized since Broca's famous lesion studies in the 1860's (Broca, 1861). The area has traditionally been associated with language production (see e.g. Geschwind, 1965) but in recent years this specialization of PTr has been widely disputed. Evidence from a range of neuroimaging studies suggests that PTr is most likely to play a significant role in language comprehension as well (Papathanassiou et al., 2000; for a review see Vigneau et al., 2006). In particular, it has been suggested that these areas are involved in aspects of semantic processing, such as semantic retrieval or selection among multiple competing alternatives (Bedny et al., 2007; Goldberg et al., 2007; Moss et al., 2005; Rodd et al., 2005; Skipper et al., 2007; Stowe et al., 2004; Thompson-Schill et al., 1997; Wagner et al., 2001). With a few (bilateral) exceptions, these studies report left lateralized specialization for PTr semantic processing, in accordance with the traditional view that language processing is left lateralized. However, there is growing evidence that right hemisphere PTr is involved in the processing of semantic ambiguity (Grindrod & Baum, 2003, 2005; Mason & Just, 2007; Rodd et al., 2005; Strange, Henson, Friston, & Dolan, 2000; Wallentin, Østergaard, Lund, Østergaard, & Roepstorff, 2005; Zempleni, Renken, Hoeks, Hoogduin, & Stowe, 2007). More specifically, Zempleni et al. (2007) and Grindrod and Baum (2003), Grindrod and Baum (2005) suggest that the left IFC processes semantic selection, while the right IFC is responsible for the integration of contextual knowledge when needed to resolve semantic ambiguity. A group of recent studies comparing the neural processing of conventional and novel unconventional metaphors (Faust & Mashal, 2007; Mashal, Faust, & Hendler, 2005; Mashal, Faust, Hendler, & Jung-Beeman, 2007; Schmidt, DeBuse, & Seger, 2007) support this interpretation, demonstrating that the right hemisphere, especially right IFC, is involved in comprehension of novel metaphors (Faust & Mashal, 2007, p. 860).

In line with this interpretation we speculate that the activation of the PTr in the present study is generally related to semantic retrieval and selection. While the activations of the inferior temporal lobes (incl. VWFA) are involved in more categorial questions concerning the representational status of the objects (communicative vs. non-communicative), PTr may be involved in the more fine-grained semantics, i.e. the representational meaning of material symbols. For instance, Stowe et al. (2004) report PTr activity in coordinates largely overlapping with the present study in a task concerning the reading of ambiguous and unambiguous written sentences. Though these areas are commonly associated with verbal language and lexical semantics, it would seem they have a more abstract, multimodal function in relation to the mediation of communicative meaning.

We reason that the significant recruitment of right hemisphere PTr in this study may be due to the crucial role of pragmatic and contextual integration involved in the comprehension of more ambiguous and unconventional signs. Conventional material symbols, like a buoy or a police cordon, are probably experienced as embodying their symbolic referential qualities more or less independent of context. The symbolic reading is ritualized, yielding less categorical and semantic ambiguity. Conversely, in the case of novel and unconventional communicational uses of objects, the meaning is not experienced as intrinsic to the object as such but heavily dependent on its situation in the surrounding context. The comprehension of such scenes will probably impose challenges on the processing mechanisms, requiring additional contextual information to resolve problems of multiple categorical and semantic alternatives.

If the function of right IFC is connected to the integration of context to resolve semantic ambiguity as suggested above, we expect the effect in this region to be most prominent for the novel, unconventional, and non-entrenched communicative uses of objects. To verify the hypothesis that the effect in right IFC is modulated by the degree of conventionality of the communication, we conducted a second analysis including a regressor that modeled the subjects' ratings of conventionality for each of the material symbols. The right IFC showed significant activation as a function of increasing unconventionality, supporting our predictions. This supports the interpretation of this area as playing a role in relation to novel, unconventional, and therefore ambiguous form-meaning pairings in accordance with the above-referred literature, in particular Rodd et al., 2005.

We did not in this study find activations in any of the areas related to the so-called 'social brain' such as the medial prefrontal cortex and temporo-parietal junction, that in a series of recent studies have been associated with the recognition of communicative intentions (cf. e.g. Amodio & Frith, 2006; Kampe, Frith, & Frith, 2003; Walter et al., 2004). We speculate that this might be due to the special kind of 'offline' communicative mediation that is addressed in this study. Since no persons or otherwise animated items are depicted in the stimuli, subjects in this study may to a larger extent rely on symbolic, *object-centered* strategies rather than mentalizing, theory of mind and other *person-centered* strategies (cf. e.g. A. Clark, 2006; Gergely et al., 2007; Pickering & Garrod, 2004).

## 4. Conclusion

This study targeted the neural substrates for our comprehension of material objects used for purposes of interpersonal communication. It takes as its point of departure the idea that we humans can exploit various expressive means and devices when we engage in communicative interactions. While words in a natural language are a forceful means of such semantic exchange it is far from the only one. We gesture, vocalize and employ various material symbols and artifacts in functional ways that make them referentially relevant to other people. In concordance previous functional brain imaging studies of the comprehension of e.g. communicative gestures (Lotze et al., 2006), vocalizations (Dietrich et al., 2007) and facial and bodily expression (Lawrence et al., 2006) have found activity in brain regions largely overlapping with those normally associated with verbal language. In the current study we report activity in similar language related brain areas (especially the pars triangularis of the inferior frontal gyrus and the visual word form area of inferior temporal lobe) when subjects recognize everyday material objects to play a mediating role in a communicative exchange.

The results thus suggest that in certain situations material objects are doing the job of words: they are perceived as meaning-conveying communicative symbols comparable with other modes of symbolic expression, including verbal language.

In addition, we found areas in the pars triangularis of the right inferior frontal cortex that are modulated by the degree of conventionality and entrenchment of a communicative object. The findings seem to coincide with a number of studies reporting activations in these areas of the brain in tasks concerning the resolution of semantic ambiguity. When new and unconventional communicative expressions are introduced, this poses certain challenges for the underlying processing mechanisms. Often integration of extra contextual information is needed to solve categorical and semantic ambiguities. This may explains the relatively enhanced recruitment of right-hemisphere frontal areas compared with studies of conventional verbal language. Our results thus expand upon previous studies that consistently report a role for right pars triangularis in tasks involving contextual integration, the resolution of semantic ambiguity, and inhibition of competing semantic alternatives.

## Acknowledgments

The authors thank Chris Frith, Ethan Weed, Svend Østergaard, Henrik Lund, The Danish National Research Foundation's Center for Functionally Integrative Neuroscience and Politiken Fonden.

## References

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience, 7*, 268–277.

Andersson, J. L. R., Hutton, C., Ashburner, J., Turner, R., & Friston, K. (2001). Modeling geometric deformations in EPI time series. *NeuroImage, 13*(5), 903–919.

Ashburner, J., & Friston, K. J. (1999). Nonlinear spatial normalization using basis functions. *Human Brain Mapping, 7*(4), 254–266.

Bedny, M., Hulbert, J. C., & Thompson-Schill, S. L. (2007). Understanding words in context: The role of Broca's area in word comprehension. *Brain Research, 1146*, 101–114.

Bloom, P. (1996). Intention, history and artifact concepts. *Cognition, 60*(1), 1–29.

Broca, P. (1861). Remarques sur le siége de la faculté du langage articulé; suives d'une observation d'aphemie. *Bulletin de la Société Anatomique de Paris, 6*, 330–357.

Clark, H. H. (1996). *Using language*. Great Britain: Cambridge University Press.

Clark, H. H. (2005). Coordinating with each other in a material world. *Discourse Studies, 7*, 507–525.

Clark, A. (2006). Material symbols. *Philosophical Psychology, 19*(3), 291–307.

Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M. A., et al. (2000). The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain, 123*, 291–307.

Cohen, L., Lehericy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain, 125*, 1054–1069.

Dehaene, S., Le Clec, H. G., Poline, J. B., Le Bihan, D., & Cohen, L. (2002). The Visual Word Form Area: A prelexical representation of visual words in the fusiform gyrus. *NeuroReport, 13*, 321–325.

DeLoache, J. S. (2004). Becoming symbol-minded. *Trends in Cognitive Sciences, 8*, 66–70.

Dennett, D. C. (1987). *The intentional stance*. Cambridge, Mass: MIT Press.

Desimone, R., & Ungerleider, L. (1989). Neural mechanisms of visual processing in monkeys. In F. Boller & J. Grafman (Eds.). *Handbook of neuropsychology* (2, pp. 267–299). Amsterdam: Elsevier.

Dietrich, S., Hertrich, I., Alter, K., Ischebeck, A., & Ackermann, H. (2007). Semiotic aspects of nonverbal vocalizations: a functional imaging study. *NeuroReport, 18*(18), 1891–1894.

Faust, M., & Mashal, N. (2007). The role of the right cerebral hemisphere in processing novel metaphoric expressions taken from poetry: A divided visual field study. *Neuropsychologia, 45*, 860–870.

Friston, K. J., Ashburner, J., Poline, J. B., Frith, C. D., Heather, J. D., & Frackowiak, R. S. (1995a). Spatial registration and normalization of images. *Human Brain Mapping, 2*, 165–189.

Friston, K. J., Holmes, A. P. J. W. K., Poline, J. B., Frith, C. D., & Frackowiak, R. S. (1995b). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping, 2*, 189–210.

Geschwind, N. (1965). The organization of language and the brain. *Science, 170*, 940–944.

Goldberg, R. F., Perfetti, C. A., Fiez, J. A., & Schneider, W. (2007). Selective retrieval of abstract semantic knowledge in left prefrontal cortex. *Journal of Neuroscience, 27*(14), 3790–3798. 4.

Goodman, N. (1976). *Languages of art: An approach to a theory of symbols* (2nd ed.). Indianapolis: Hackett Publishing Company.

Grindrod, C. M., & Baum, S. R. (2003). Sensitivity to local sentence context information in lexical ambiguity resolution: Evidence from left- and right-hemisphere-damaged individuals. *Brain and Language, 85*, 503–523.

Grindrod, C. M., & Baum, S. R. (2005). Hemispheric contributions to lexical ambiguity resolution in a discourse context: Evidence from individuals with unilateral left and right hemisphere lesions. *Brain and Cognition, 57*, 70–83.

Haushofer, J., & Kanwisher, N. (2007). In the eye of the beholder: Visual experience and categorization in the human brain. *Neuron, 53*, 773–775.

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

Jack, A. I., & Roepstorff, A. (2002). Introspection and cognitive brain mapping: From stimulus–response to script-report. *Trends in Cognitive Science, 6*(8), 333–339.

Kampe, K., Frith, C. D., & Frith, U. (2003). "Hey John": Signal conveying communicative intention toward the self activate brain regions associated with "Mentalizing," regardless of modality. *The Journal of Neuroscience, 23*(12), 5258–5263.

Kelemen, D. (1999). Function, goal and intention: Children's teleological reasoning about objects. *Trends in Cognitive Sciences, 3*(12), 461–468.

Kronbichler, M., Hutzler, F., Wimmer, H., Mair, A., Staffen, W., & Ladurner, G. (2004). The visual word form area and the frequency with which words are encountered: evidence form a parametric fMRI study. *NeuroImage, 21*, 946–953.

Lawrence, E. J., Shaw, P., Giampietro, V. P., Surguladze, S., Brammer, M. J., & Davis, A. S. (2006). The role of 'shared representations' in social perception and empathy: An fMRI study. *NeuroImage, 29*(4), 1173–1184.

Levinson, J. (1993). Extending art historically. *The Journal of Aesthetics and Art Criticism, 51*(3), 411–423.

Lotze, M., Heymans, U., Birbaumer, N., Veit, R., Erb, M., Flor, H., et al. (2006). Differential cerebral activation during observation of expressive gestures and motor acts. *Neuropsychologia, 44*, 1787–1795.

Maldjian, J. A., Laurienti, P. J., & Burdette, J. H. (2004). Precentral gyrus discrepancy in electronic versions of the Talairach atlas. *NeuroImage, 21*(1), 450–455.

Maldjian, J. A., Laurienti, P. J., Kraft, R. A., & Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage, 19*(3), 1233–1239.

Mashal, N., Faust, M., & Hendler, T. (2005). The role of the right hemisphere in processing nonsalient metaphorical meanings: Application of principal components analysis to fMRI data. *Neuropsychologia, 47*, 2084–2100.

Mashal, N., Faust, M., Hendler, T., & Jung-Beeman, M. (2007). An fMRI investigation of the neural correlates underlying the processing of novel metaphoric expressions. *Brain and Language, 100*, 115–126.

Mason, R. A., & Just, M. A. (2007). Lexical ambiguity in sentence comprehension. *Brain Research, 1146*, 115–127.

Moss, H. E., Abdallah, S., Fletcher, P., Bright, P., Pilgrim, L., Acres, K., et al. (2005). Selecting among competing alternatives: Selection and retrieval in the left inferior frontal gyrus. *Cerebral Cortex, 15*(11), 1723–1735.

Nichols, T. E. (2007). False discovery rate procedures. In K. J. Friston, J. T. Ashburner, S. J. Kiebel, T. E. Nichols, & W. D. Penny (Eds.), *Statistical parametric mapping: The analysis of functional brain images*. UK: Elsevier.

Nishitani, N., Schürmann, M., Amunts, K., & Hari, R. (2005). Broca's region: From action to language. *Physiology, 20*, 60–69.

Papathanassiou, D., Etard, O., Mellet, E., Zago, L., Mazoyer, B., & Tzourio-Mazoyer, N. (2000). A common language network for comprehension and production: A contribution to the definition of language epicenters with PET. *NeuroImage, 11*, 347–357.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27*, 169–226.

Pietrini, P., Furey, M. L., Ricciardi, E., Gobbini, M. I., Wu, W.-H. C., Cohen, L., et al. (2004). Beyond sensory images: Object-based representation in the human ventral pathway. *PNAS, 101*(15), 5658–5663.

Price, C. J. (2000). The Anatomy of Language: Contributions from functional neuroimaging. *Journal of Anatomy, 197*, 335–359.

Price, C. J., & Devlin, J. T. (2003). The myth of the visual word form area. *NeuroImage, 19*, 473–481.

Reinke, K., Fernandes, M., Schwindt, G., O'Craven, K., & Grady, C. L. (2008). Functional specificity of the visual word form area: General activation for words and symbols but specific network activation for words. *Brain and Language, 104*(2), 180–189.

Rodd, J. M., Davis, M. H., & Johnsrude, I. S. (2005). The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cerebral Cortex, 15*, 1261–1269.

Ruxton, G. D. (2006). The unequal variance *t*-test is an underused alternative to Student's *t*-test and the Mann–Whitney *U* test. *Behavioral Ecology, 17*(4), 688–690.

Schmidt, G. L., DeBuse, C. J., & Seger, C. A. (2007). Right hemisphere metaphor processing? Characterizing the lateralization of semantic processes. *Brain and Language, 100*, 127–141.

Skipper, J. I., Goldin-Meadow, S., Nusbaum, H. C., & Small, S. L. (2007). Speech-associated gestures, Broca's area, and the human mirror system. *Brain and Language, 101*, 260–277.

Stowe, L. A., Paans, A. M. J., Wijers, A. A., & Zwarts, F. (2004). Activations of "motor" and other non-language structures during sentence comprehension. *Brain and Language, 89*, 290–299.

Strange, B. A., Henson, R. N. A., Friston, K. J., & Dolan, R. J. (2000). Brain mechanisms for detecting perceptual, semantic and emotional deviance. *NeuroImage, 12*, 425–433.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *PNAS, 94*, 14792–14797.

Tomasello, M. (1999). The cultural ecology of young children's interaction with objects and artifacts. In E. Winograd, R. Fivush, & W. Hirsh (Eds.), *Ecological approaches to cognition: Essays in honor of Ulric Neisser*. MahWah, NJ: Erlbaum.

Tylén, K., (2007). When agents get expressive: A theory of semiotic agency, *Cognitive Semiotics, 0*, Aarhus: special promotional issue.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage, 15*(1), 273–289.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. Ingle, M. Goodale, & R. Mansfield (Eds.), *Analysis of visual behavior*. Cambridge, Mass: MIT Press.

Vigneau, M., Beaucousin, V., Hervé, P. Y., Duffau, H., Crivello, F., Houdé, O., et al. (2006). Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *NeuroImage, 30*, 1414–1432.

Wagner, A. D., Paré-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning: Left frontal cortex guides controlled semantic retrieval. *Neuron, 31*, 329–338.

Wallentin, M., Østergaard, S., Lund, T. E., Østergaard, L., & Roepstorff, A. (2005). Concrete spatial language: See what I mean? *Brain Language, 92*(3), 221–233.

Wallentin, M., Roepstorff, A., Glover, R., & Burgess, N. (2006). Parallel memory systems for talking about location and age in precuneus, caudate and Broca's region. *NeuroImage, 32*(4), 1850–1864.

Walter, H., Adenzato, M., Ciaramidaro, A., Enrici, I., Pia, L., & Bara, G. B. (2004). Understanding intentions in social interaction: The role of the anterior paracingulate cortex. *Journal of Cognitive Neuroscience, 16*(10), 1854–1863.

Willems, R. M., Özyürek, A., & Hagoort, P. (2007). When language meets action: The neural integration of gesture and speech. *Cerebral Cortex, 17*(10), 2322–2333.

Zempleni, M.-Z., Renken, R., Hoeks, J. C. J., Hoogduin, J. M., & Stowe, L. A. (2007). Semantic ambiguity in sentence context: Evidence from event-related fMRI. *NeuroImage, 34*, 1270–1279.

# Taking the language stance in a material world: a comprehension study

Kristian Tylén[a,b], Johanne Stege Phillipsen[c], Ethan Weed[b]

a. Institute of Language and Communication, University of Southern Denmark; b. Center for Functionally Integrative Neuroscience, Aarhus University Hospital; c. Center for Semiotics, University of Aarhus

Two basic traits profoundly characterize human nature and distinguish us from even our closest relatives in the animal kingdom: our advanced social skills and our extensive engagement with material artifacts and technologies. This article explores the ways in which these two strands of human life interweave when we systematically exploit our material environment for means of social coordination and meaning construction. We report from an experiment assessing subjects' comprehension of everyday static objects like vacuum cleaners, sliced tomatoes, and sneakers when these are intentionally manipulated or placed at special sites to accomplish communicative functions. Our findings suggest that when subjects are confronted with such intentional object manipulations – hence material signals - they tend to take a 'language stance', that is, explore the scenes as means for intersubjective meaning construction. In addition, a pronounced alignment effect in subjects' meaning-exploring attitudes indicate that the interpretation of material signals is not a "private" endeavor, but a process heavily regulated by an orientation to "shared" norms, conventions and interpretative strategies. We speculate that such shared top-down perceptual strategies are indirectly regulated by the priming efforts of an agent-addresser, hence a type of 'top-top interaction' (Roepstorff & Frith, 2004).

## 1. Introduction

Homo sapiens is the materialistic species. Like no other primate our life has come to depend on all kinds of external matter. From stone axes to wheels, skyscrapers, newspapers and art, not to mention more abstract social objects such as governments, money and marriage (Searle, 1995). No other species employs material objects and technologies in the complex and systematic way that humans do for purposes of for instance external memory storage (Donald, 2002), cognitive augmentation (A. Clark, 1997; A. Clark & Chalmers, 1998; Hutchins, 1995), and interpersonal communication (H. Clark, 2005). In the current study, we focus on the latter aspect, investigating how everyday material objects come to play a mediating role as signals in interpersonal communicative interactions. The claim will be that given a particular contextual setting even vacuum cleaners, icebergs or trainers can become '*a-kind-of-language*'. This happens when we actively approach such material structures as sources of intersubjective meaning construction (Tylén, 2007; A. Clark, 2006; Roepstorff, 2008).

Based on experimental data we will argue that this is not a private endeavor going on between a single person and an object but is better explained as a social process heavily regulated by orientation to interpretative norms and conventions primed by the semiotic agency of addressing agents (Tylén, 2007; Rollins, 2004). Material signals may thus prove a case of how the human mind does not only extend into the material world (A. Clark & Chalmers, 1998; Hutchins, 1995) but does so in a regulated fashion by orientation to shared, normative top-down perceptual and cognitive strategies, putatively termed top-top interactions (Roepstorff and Frith, 2004).

## Meaning and Materiality

In a recent, influential article (A. Clark, 2006), Andy Clark proposes a so-called *complementary view of language*. The idea is to appreciate the material nature of symbolic language not as an empty vehicle that in comprehension is fully translated into some kind of qualitatively different 'content-capturing inner code'. Rather, when we encounter linguistic signals the materiality of the symbolic language itself (or image-like internal representations of its material structures) continues to impact the mental processes in a dynamic, complementary and coordinating way as an alternative fulcrum of attention, perception and action (A. Clark, 2006). While A. Clark's focus is on this 'hybrid nature' of verbal language (also called "the ultimate artifact", A. Clark, 1997), there is little to suggest that this does not generalize to other types of material matter, including objects used as signals (H. Clark, 2005; Roepstorff, 2008). The claimed coordinative relationship between aspects of materiality and what we would term the meaning-constructing potential of a material symbol invites us to appreciate how specific kinds of object manipulations and compositions together with their contextual setting come to afford special styles of perceptual and meaning-exploring work. An important aspect seemingly missing in A. Clark's account however, is the impact of social coordination and conformity on such distributed cognitive processes. While A. Clark gives emphasis to the role of the materiality of symbols in a subject's "private" problem-solving activities our investigations will take a slightly different direction, addressing the perhaps more fundamental role of material language, that is, its function as a means of social coordination and meaning construction (Cowley, 2007; Spurrett & Cowley, 2004).

## The language stance extended

In this study we address situations where properties of our static, mute, material environment come to be perceived as dynamic social signals. Some salient, purposeful arrangements of material objects are thus experienced as affording a special type of interpretative attitude, not incompatible with Spurrett and Cowley's notion of a *'language stance'*. In a number of conceptual studies, Cowley has proposed the notion of 'a language stance' designating young children's emerging ability to pick up on normative patterns in spoken and written utterances (Cowley, 2007; Spurrett and Cowley, 2004). While Cowley and colleagues have focused primarily on issues of verbal phonetics and graphics we propose to extend the 'language stance'-concept to the recognition of (non-verbal) material signals. In this context, it will designate a perceiver's motivated inclination to systematically explore a material structure as a source of intersubjective semantic meaning. Taking the language stance toward a conspicuous configuration of everyday material items means to attend not only to their form and inherent functionality but rather to their situated contextualized use as mediating an intersubjective intention. This shift in the top-down attitude and attentional profile of the perceiver thus makes new details of the material scene semiotically relevant. We of course recognize that the proposed analogy between verbal language and these somewhat non-canonical object uses may be somewhat fragile. Still, focusing broadly on the dynamic *activity of signaling* in social interactions allows us to appreciate common aspects of their cognitive functionality across modalities and structural differences (H. Clark, 2005; 1996). Besides recent neurocognitive studies seem to suggest that quite different forms of communicative mediation including material signals may work in ways comparable to words in verbal languages (cf. Tylén et al, in press; Roepstorff, 2008).

### Top-top interaction

A social interactional perspective on material signals introduces additional questions concerning conformity and conventionality in meaning constructing approaches to material structure. Often material structures will potentially lend themselves to several alternative meaning exploring interpretations (Myers & Liben, 2008). The employment of material signals in mediation of joint participatory sense-making entails that the interacting agents attune to shared attitudes and strategies for the perceptual and meaning constructing approach to material structures (Pickering and Garrod, 2004; De Jaegher & Di Paolo, 2007). The negotiations and alignments of such top-down attitudes heavily regulate the kind of meaning that the interacting agents may attribute to a perceptual scene. Roepstorff and Frith (2004) have proposed to term these processes of mutual alignment of top-down attitudes '*top-top interactions*'. Since meaning is not an object 'transferred' between interacting persons in 'code-vehicles' (Kravchenko, 2007) but depends on their active meaning-pursuing approach to aspects of material structure (A. Clark, 2006), intersubjective meanings entail shared perceptual and meaning constructing strategies for such explorations.

## 2. The recognition and comprehension of material signals

In the current study, we focus on a special type of object-mediated signal activity, that is, those instances when we happen to perceive everyday objects like chairs, sprinkles of flour, or vacuum cleaners, as means of coordination and construction of intersubjective meaning. This is accomplished when we recognize such objects or properties of our material environment as manipulated or arranged in a striking and deliberate manner that calls for a special kind of social interpretation (thus, bearing some relationship to H. Clark's notion of *manifested action*, 2005 or Tylén's *semiotic agency*, 2007). For instance, we can encounter chairs put out in the street to reserve a parking lot, flour sprinkled out in the shape of an arrow to indicate a direction, or rows of vacuum cleaners lined up on a lawn to create a conceptual piece of art. For the sake of simplicity, we will in the following consider four types of manipulations that efficiently frame material objects as social signals, either in combination or isolation:

1) The translocation of an object into a new and unfamiliar context.
2) The manipulation of a central property of an object, such as color or shape.
3) The arrangement of several common objects in a composition, for instance a symmetrical configuration.
4) The complementary employment of conventional signs, such as hearts, arrows, flags and text.

Though the typology is neither exhaustive nor systematically, it can serve as an intuitive guide for the kind of situations that are investigated below.

In the following, we report from an experiment investigating naïve subjects' casual comprehension of scenes featuring different types of material object configurations, some of which are social signals of the kind described above. The experiment is based on content analysis (Neuendorf, 2002) and methods are comparable to those used in for instance Abell et al, 2000.

In the experiment, we showed subjects a series of images depicting static configurations of material objects and asked them to orally report their comprehension of the scenes. The stimulus images were arranged to accommodate two types of contrastive analyses of subjects' comprehension of objects: social signal vs. non-signal, and conventional vs. unconventional material signals. Were premeditated a series of different investigative approaches to the collected data including quantitative statistical analyses, and more explorative detailed semantic analyses of subsets of subjects' verbal reports.

In relation to the signal vs. non-signal contrast, we expected that when confronted with specific types of intentional object manipulations (the two types of test stimuli) subjects would be more likely to 'take a language stance', that is, pick up on intentional cues and explore the scene as a means for socially constructed, semantic meaning (Spurrett & Cowley, 2004). We predicted this would be reflected in the distribution of specific lexical items between conditions in subjects' verbal reports such as communication-related terms and discursive markers (e.g. references to addresser and addressee relations). In contrast, we did not expect significant differences between these conditions in the distribution of communication-neutral lexical items such as agency-related words.

In relation to the contrastive analysis between the two test conditions (conventional vs. unconventional communication) we predicted that subjects would recognize both as social signals, but would tend to be more comfortable making explicit semantic interpretations of conventional material signals, while they would be reluctant to do so with the novel unconventional object configurations. We anticipated this effect would be due to reduced *common ground* in the unconventional condition, that is, the interlocutors' mutual attunement to shared knowledge (H. Clark & Brennan, 1991; Pickering & Garrod, 2004).

Furthermore, we expected the more explorative analyses to reveal systematic differences in the interpretative strategies applied by subjects to various types of scenes, which again would have consequences for the kind of meaning-constructing work that would guide the comprehension of various types of material signals.

## 3. Subjects

Twenty healthy informants (10 females/10 males), mean age 26 ± 6 (STD) participated in the study. Only native, Danish speaking subjects, naïve to the purpose of the study, were included. Prior to inclusion all subjects gave informed written consent. Subjects received no fee for participation.

## 4. Stimuli

Stimuli consisted of a set of eighteen images carefully selected from a dataset of 100 color photos used in a previous study on material signals (see Tylén et al, in press, for details). Images depicted actually occurring scenarios containing static configurations of material objects and artifacts. No persons were depicted in the images.

The selection of images for the current study was based on their scores in the previous classification study. Twenty-two subjects rated the images on two parameters: 1) the presence/non-presence of communicative signals, and 2) the degree of conventionality of the signals. Conventionality was defined in terms of relative frequency, that is, how often subjects would expect to encounter a certain type of material signal. Based on statistical analysis of the classification data, the experimenters selected six stimulus images among the best exemplars for each of three predefined conditions.

- Test I: images that were found to depict conventional ways of using material objects as communicative signals (thresholds: >80% in communication rating/>60% in conventionality rating). Examples: a buoy, marking the location of fishing nets, a large crowd of roses left on the pavement in grief of a person murdered there, and flour sprinkled out as an arrow to show directions.
- Test II: images that were found to depict unconventional ways of using material objects as communicative signals (thresholds: >80% in communication rating/<20% in conventionality rating). Examples: a crowd of vacuum cleaners arranged on a lawn to make a conceptual piece of art, thin slices of tomato in a patterned distribution on the windscreen of a car, and an iceberg covered with red paint.
- Control: images that were found by the subjects not to evoke any conspicuous communicative interpretations (threshold: <10% in communication rating). Examples: colorful laundry hung out to dry, a blender on a kitchen table, and a fire extinguisher in a messy cupboard (see fig. 1).

## 5. Experimental design

Stimulus images were presented in a randomized order on a 13.3” computer monitor using Microsoft PowerPoint. Each image was shown for sixty seconds and images were separated by a two second

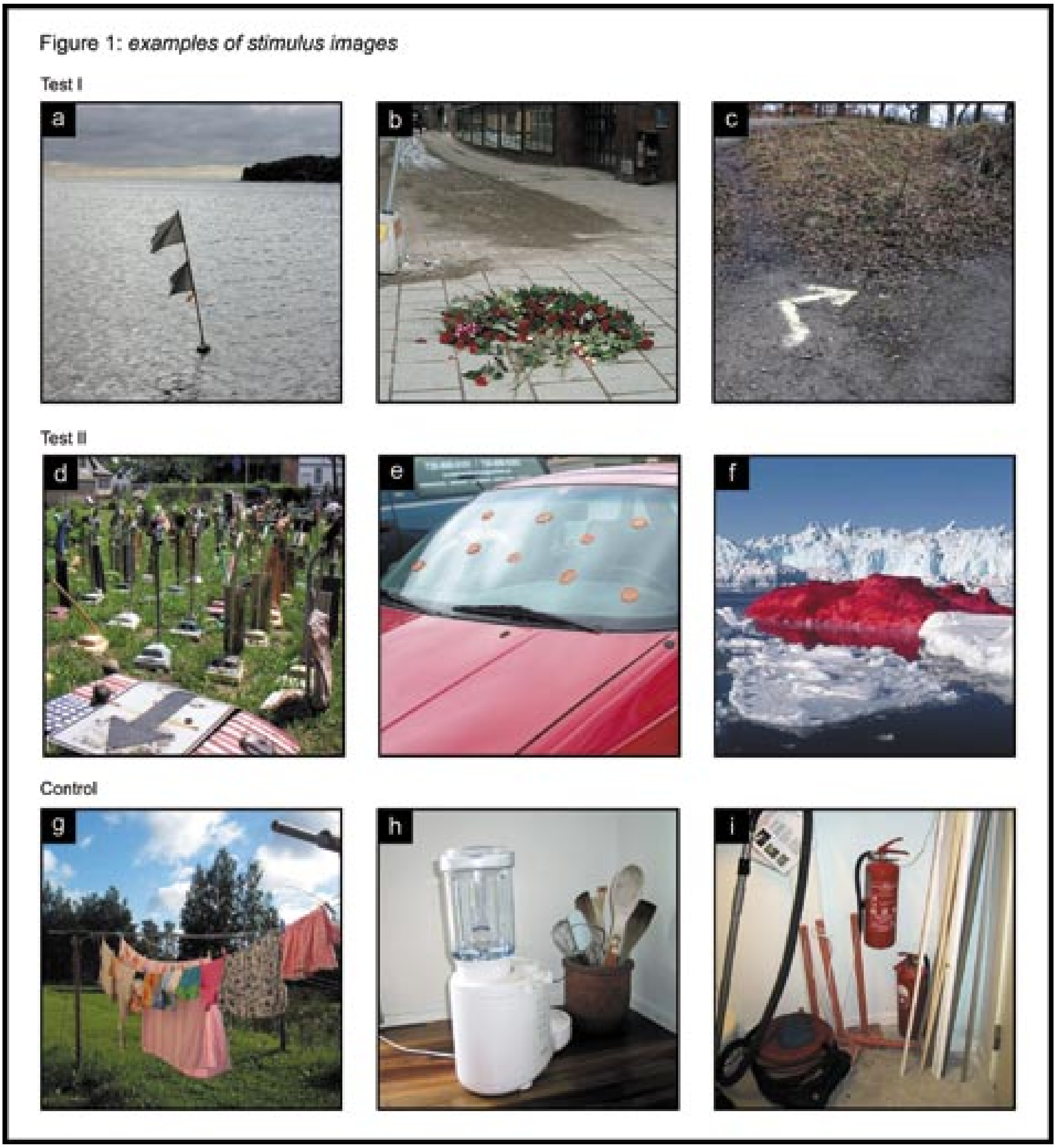



***Figure 1***: Examples of stimulus images from each of the three conditions. First row depicts Test I stimuli: conventional ways of employing objects as signals. Second row depicts Test II stimuli: unconventional ways of employing objects as signals, and third row depicts Control stimuli: scenes without any conspicuous communicative relevance.

black pause screen. Subjects were instructed to speak while watching the images and their oral reports were recorded using a Sony stereo IC Recorder (model ICD-SX56). Prior to the test phase, subjects received careful instructions in Danish concerning the task. The formulation of the main task translates as "*describe what you see in the image and your understanding of it*". This phrasing was chosen to motivate free descriptions that would (preferably) go beyond a mere listing of items depicted, but without establishing any strong interpretative biases (Gallagher, 2003).

Subsequent to the data collection, subjects' oral reports (a corpus of approx. six hours of audio recordings) were transcribed using the software CLAN (MacWhinney, 2000) and all analyses were carried out on the transcribed material.

## 6. Data analysis

Three types of analyses were employed to capture different aspects of the data: 1) a quantitative, statistical analysis based on a set of predefined hypotheses. The purpose of this analysis was to assess subjects' categorical intuitions (signal vs. non-signal) in relation to each stimulus image across conditions. This was carried out by first coding subjects' image descriptions in relation to a set of predetermined variables and then testing for significant distributions of these across the three conditions. In addition, we planned a series of post hoc analyses to test specific hypotheses regarding the distribution of these classes between the two test conditions (conventional vs. unconventional signals). 2) A second, more explorative coding procedure targeted subjects' orientation to specific interpretative strategies in their construction of intersubjective, semantic meaning of the material signals, and 3) a series of detailed qualitative, semantic analyses organized in three different illustrative cases of subjects' exploration of material objects as signals. The three types of analyses are presented in separate sections below.

### Coding procedure I: categorical resolutions

Three experimenters coded the subjects' reports independently in relation to a predefined coding scheme consisting of five variables: A) *communication,* B) *semantics/social meaning*, C) *addresser*, D) *addressee*, E) *agency* (see table 1). In order to accomplish a high level of reliability, the scheme was designed primarily to designate easily recognizable, explicit lexical items. It should be stressed at this point that the authors recognize that verbal language is an extremely flexible means of interpersonal meaning construction that will often express such relations in a pragmatic way without explicit lexicalization, but to accomplish a desirably stringent and objective investigation method we settled on this somewhat conservative search strategy (see table 1).

The individual codings of the three experimenters were summarized for each stimulus image and each subject in a weighted fashion. A subject's description of an image was rated in relation to each variable in percent expressing if no (0%), one (33%), two (66%) or all three experimenters (100%) would assign the given variable (e.g. *communication*) to the description. Ratings were subsequently averaged for each image across subjects. Statistically significant differences across the three conditions (test I, test II and control) were measured for each variable using a one-way ANOVA. Furthermore, a set of planned post hoc analyses were executed using Tukey's HSD multiple-comparison test. Both types of analyses were performed in MatLab (Mathworks inc., Sherborn, Massachusetts, USA).

Inter-coder reliability between the three coders was measured pairwise for each variable and each subject in percent agreement (PA) and Cohen's *kappa* (κ) (Cohen, 1960) performed in PRAM v 0.4.5 (Skymeg Software, Inc.) and consequently averaged across subjects, coders and variables (cf. Neuendorf, 2002, for a discussion on reliability measurements).

Table 1: *coding scheme*

| Code | Description | Examples |
|---|---|---|
| A | explicit references to communication terms | nouns: 'sign', 'art', 'symbol', 'message', etc., or verbs: 'express', 'say', 'instruct', 'show', etc. |
| B | explicit references to semantic meaning | social functions, modals, motivated figurative or metaphoric meanings, motivated references to objects and situations not present in the image |
| C | explicit references to the discursive marker addresser | grammatical subjects for communication related verbs like 'someone', 'an artist', 'you', etc. |
| D | explicit references to the discursive marker addressee | grammatical indirect objects of communication related verbs like 'someone', 'you', 'people', etc. |
| E | explicit references to non-semiotic agency in general | non-semiotic activities and acts like in the verbs: 'build', 'hang up', 'make', 'put on', etc. |

***Table 1***: Coding scheme for the multi-coder analysis of image descriptions. Three coders independently scored subjects' image descriptions in relation to the five variables, defined in the scheme. In order to accomplish a high level of objectivity (and inter-coder reliability) the scheme was designed primarily to designate easily recognizable, explicit lexical items.

## Results of coding procedure I

The one-way ANOVA revealed no significant difference in the number of words used to describe images across conditions ($F(2,15) = 2.86$, $p = .089$). Thus, we can assume that other significant differences are not due to systematic variations in the number of words used in subjects' reports.

There were significant differences between conditions in all variables except 'agency': Communication: $F(2/15) = 8.45$, $p = .0035$, Semantics: $F(2/15) = 13.64$, $p = .0004$, Addresser: $F(2/15) = 6.57$, $p = .0089$, Addressee: $F(2/15) = 3.88$, $p = .0439$, Agency: $F(2/15) = 0.68$, $p = .52$ (see table 2).

A set of pairwise post hoc Tukey's HSD multiple comparison tests (threshold = $p < .05$) were conducted to explore differences between the two test conditions: conventional and unconventional material signals. No significant effects were found. Although the difference between the two conditions in relation to the variable *semantics* seems to show a strong tendency (*semantics*, test I: $M = 59$, vs. test II: $M = 39$), corresponding to our hypothesis, this effect did not reach significance due to the pronounced variability of the data. Thus, while the predefined variables of the coding scheme worked successfully to capture differences between subjects' comprehension of objects as communicative signals vs. non-signals, they seemed less suited to pick up differences between the two test conditions (conventional vs. unconventional material signals).

The overall inter-coder reliability was measured to percent agreement (PA) = 86. The more conservative Cohen's *kappa* (κ) (Cohen, 1960) was measured to 0.56, corresponding to a 'fair to good' agreement beyond chance, according to Banerjee et al. (1999). The reliability was noticeably higher for the communication-related variables (*communication*: PA = 86, κ = 0.62, *semantics*: PA = 85, κ = 0.65, *addresser*: PA = 90, κ = 0.56, *addressee*: PA = 94, κ = 0.62), which indicate that the coding scheme worked efficiently for these variables, while the inter-coder reliability for the non-communicative variable *agency* was rather poor (PA = 74, κ = 0.37). We interpret this as an indication that this variable was either hard to operationalize or not satisfactorily defined in the coding scheme. The results for this variable may therefore have to be interpreted with some caution. Overall, the inter-coder reliability was found to be satisfactory which lends validity to the statistical analyses (Neuendorf, 2002).

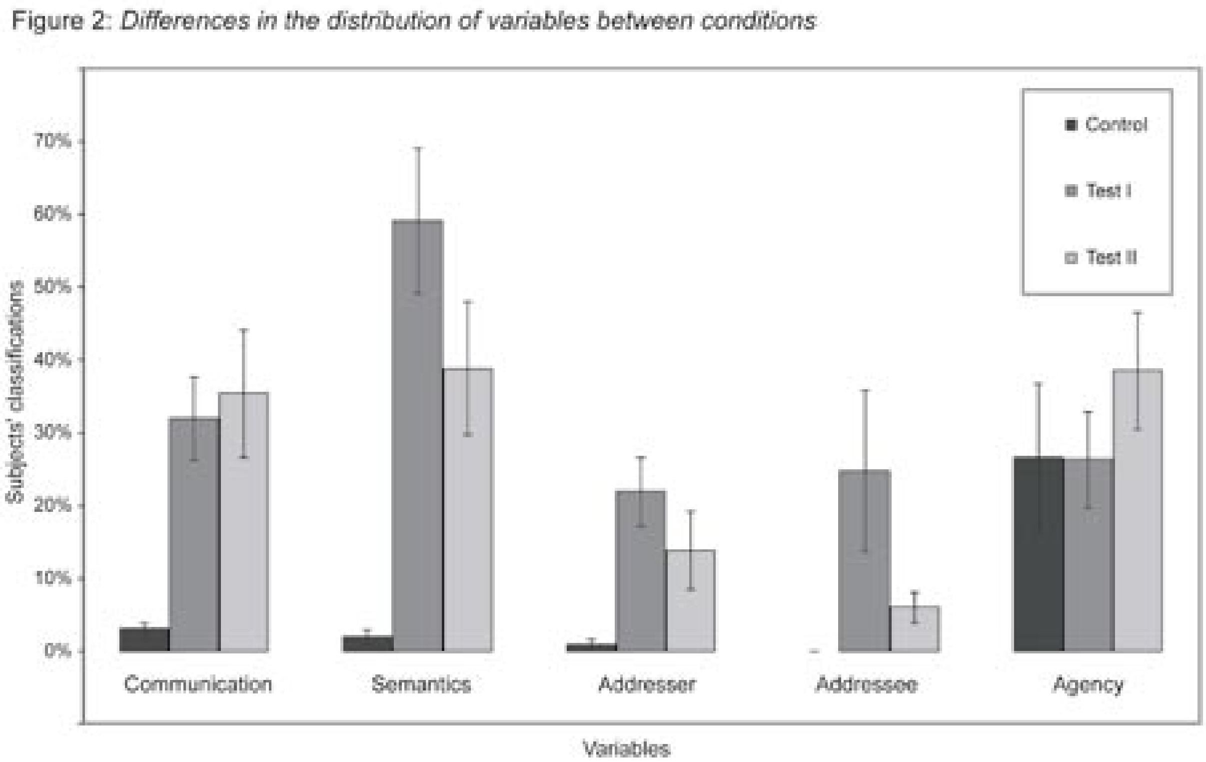



***Figure 2***: Diagram representing differences in the mean distribution of the three conditions in relation to each of the five variables. The y-axis represents a summery of the three coders' weighted ratings of subjects' image descriptions in percent. Error bars express standard error of the mean (SEM). All communication-related variables (*communication, semantics, addresser, and addressee*) showed significant effect in the one-way ANOVA ($p < .05$), while the communication-neutral variable (*agency*) did not.

## Coding procedure II: different interpretive stances

While the question addressed above concerns the categorical contrast between 'signal vs. non-signal', we will in the following section address aspects of the different kinds of social and semantic meaning that seem to be consistently associated with some of the specific object configurations. Thus, we will have a closer look at the parts of subjects' reports that the experimenters in the previous analysis coded as *social or semantic meanings.* Since such questions will relate to various subsamples of the data (i.e. small sample sizes) and are explorative in nature, we will approach these analyses in a more qualitative fashion. Hypotheses are thus formulated post hoc and analyses are guided by more subjectively motivated searches in the data.

Interestingly, preliminary searches revealed that subjects employ quite different meaning constructing strategies in their interpretative approach to the various stimulus images. The kind of representational work involved in assigning social meaning to a red iceberg (cf. figure 1.f) seems fundamentally different from that of assigning meaning to a flour-arrow on a forest path (cf. fig. 1.c) or a crowd of roses left on the street (cf. figure 1.b). This phenomenon was pursued in a second coding procedure targeting the different types of semiotic reference regulating the meaning construction in subjects' image descriptions. We identified three such regulating strategies apparently bearing some relationship to Pierce's classical sign types, *icon, index* and *symbol* (Peirce 1998). These are:

1) An *iconic/aesthetic strategy* where special attention is put on objects' resemblances and diagrammatic affordances (Stjernfelt, 2007), example: the red iceberg is associated with blood.

2) An *instructional strategy*, where material objects are considered as social imperatives, pointing to something or instructing the perceiver to adapt his/her behavior, example: the flour arrow is showing which direction to go.

3) A *symbolic strategy*, where material objects are appreciated for their conventional reference to arbitrary cultural meanings, example: the roses are expressing grief in relation to a sudden death.

In the second coding procedure, each of the subjects' image descriptions previously tagged as expressing *social, semantic meaning* were rated by an expert coder in relation to the three premeditated types of interpretational strategies above. The ratings were subsequently summarized and analyzed for subjects' relative alignment of interpretative strategy. The aim of this analysis was to investigate whether and to what extent the subjects tended to choose the same kind of strategy in their interpretative approach to a particular material signal.

This alignment effect was measured in percent agreement of interpretative attitude among the positive classifications (reports that were assigned the variable '*semantic social meaning*' in the former coding procedure) for each of the images of the test conditions. Separate measurements were applied for two types of agreement/disagreement encountered in the data: 1) *inter-subject agreement* (e.g. two subjects choose the same strategy in their exploration of an image), and 2) *intra-subject agreement* (a subject quickly settles on a single strategy rather than drifting between several alternative interpretative strategies in the exploration of an image).

In the following, both corresponding types of *dis*agreement, that is, disagreement between two or more subjects or disagreement within a single subject (the subject's explicit uncertainty in decision between several possible interpretations) are seen as reflecting the same causal factor: the failure of the addresser of the material signal in establishing (priming) an unambiguous frame for the proper interpretation (see the discussion). Furthermore, since the original communicative intentions of addressers are inaccessible in this study and true representational correspondences between addressers and addressees cannot be assessed, an underlying assumption guiding this investigation is that a strong intra- and inter-subject alignment of representation is an indirect indication of strong addresser–addressee alignment (cf. Tylén et al, *in press*).

### Results of coding procedure II

Analyses of the data from the second coding procedure revealed strong alignment effects in subjects' interpretative approach to images of the test conditions (overall agreement = 91,5%). Interestingly, however, there were substantial differences in the degree of alignment between the two test conditions. While the alignment of interpretative strategy was near perfect in the approach to conventional material signals of the Test I condition (*inter-subject* PA = 99%, *intra-subject* PA = 100%), this effect was somewhat weaker for the unconventional material signals depicted in the Test II condition (*inter-subject* PA = 81%, *intra-subject* PA = 86%) (See figure 3). Generally, subjects' interpretative approach to unconventional material signals show more uncertainty, manifested in both a higher between-subject disagreement and subjects' enhanced inclination to explore several possible interpretative strategies for the same material signal.

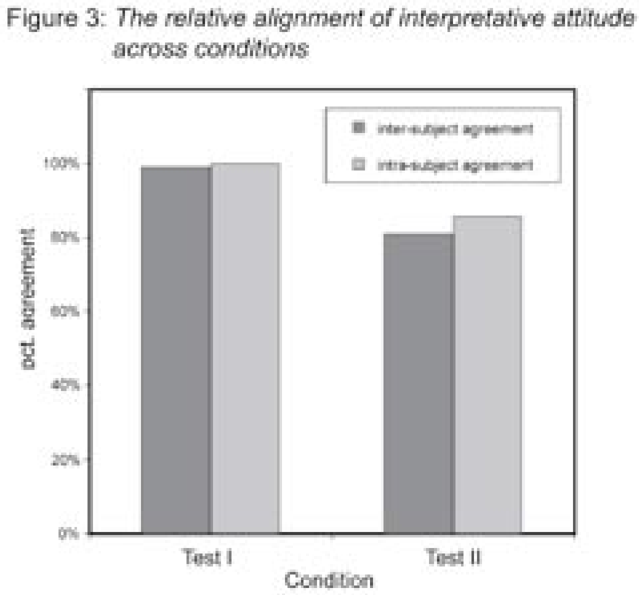



***Figure 3***: Diagram expressing the relative agreement (in percent) of subjects in their choice of interpretative attitude across test conditions. Two types of agreement are represented: 1) inter-subject agreement (the extent to which two or more subjects agree in their choice of strategy), and 2) intra-subject agreement (the extent to which each subject commits herself to a single interpretative strategy rather than indecisively drifts between several alternative strategies).

## 7. Qualitative semantic analyses

In the following, we will use a more qualitative style of analysis to address issues of meaning construction in material signals that elude the previously-employed quantitative methods. As space is limited, it will not be possible to go into lengthy reports of all the interesting findings. Instead, we will present a couple of cases that seem to address fundamental questions of intersubjective meaning construction employing material signals. In the first case study, we will investigate some of the pronounced differences in subjects' meaning exploring approaches to test vs. control scenes. In the second, we will study how subjects' choice of interpretative strategy affects their understanding of two different depictions of arrows, and in the third we will assess systematic variations in subjects' comprehension of a very ambiguous material scene.

## Case 1: alignment of representation in control vs. test scenes

In the following, we will have a closer look at some of the pronounced differences between subjects' approach to control versus test scenes. Consider this short passage from the data:

> (…) *I think this glass or plastic thing on top of the kitchen machine over here, which looks like a blender or something, kind of looks like a face that has long eyes and a nose. It looks like it has some kind of hat on.*

The passage is picked from a subject's description of one of the images (cf. figure 1.h) belonging to the group of control stimuli. The image depicts a blender on a kitchen table - a neutral scene with an everyday object in a canonical context. The description is interesting because it deviates considerably and in an interesting way from the other subjects' descriptions of the same image. In the example, the subject appreciates the iconic resemblance between the glass jug of the kitchen blender and a human face. This is in fact a possible reading of the image. And it illustrates one of the ways that we can potentially approach almost any material object we meet on our way. The kind of meaning that we derive from a material scene is, thus, to a large extent dependent on the perceptual attitude of the perceiver (Peirce, 1998; Stjernfelt, 2007; Deacon, 1997). And it is probably not uncommon from time to time to suddenly notice some aesthetic property of an everyday utilitarian artifact or to appreciate its potential iconic resemblances.

In the case referred to above, the subject's reading of the image was an exceptional case and did not represent any trend of the population. The majority of subjects simply list the items present in the picture, while a few comment on the things that one could imagine putting into the blender or refer to personal memories involving blenders.

However, a couple of other scenes belonging to the test conditions seem consistently to motivate exactly such iconic interpretative attitudes or strategies that are apparent in the in the description cited above. One of these images is figure 1.f. The image depicts an iceberg at sea, covered in red paint and surrounded by white icebergs. Here are fragments of two representative descriptions, both exploring iconic associations related to the unnatural color of the iceberg:

> *"...Of course the red is eye-catching in the blue ocean and with the white icebergs. You might almost think it looks like strawberry sorbet."*
>
> *" ...Could easily be from Greenland, but... It looks strange the way it's kinda floating, it looks kinda like a... a bloody iceberg."*

Interestingly, the majority of subjects (14/20) make these kinds of figurative analogies in their description of this particular image, most of them with association to blood. As indicated by the second coding procedure above, this curious fact seems characteristic for the difference between material objects identified as signals in contrast to objects without any communicative function. While we observe no particular unifying strategy in subjects' approach to control scenes (as the kitchen blender) subjects seem quite consistently to align their interpretational strategies to test scenes like the red iceberg. The iconic reading of the blender cited in the example above would thus seem to imply the same style of representational work as the iceberg-examples but in a private Andy Clarkian fashion, not motivated or regulated by reference to specific intentional manifestations in the composition of the scene. In contrast, the iceberg bears "eye-catching" traces of intentional manipulation affording a particular style of exploration and thus comes to implicitly regulate subjects' interpretative processes (Tylén, 2007; Rollins, 2004), resulting in pronounced inter-subject alignment of representation (Pickering & Garrod, 2004).

Two other images seem to consistently elicit similar iconic, interpretive attitudes: figure 1.d depicting a crowd of vacuum cleaners lined up on a lawn (iconic readings: 14/20), and another image from the test II condition depicting a bunch of sneakers hanging from their laces from a tree (iconic readings: 9/20).

## Case 2: The meaning of an arrow

While a subset of test images thus affords an iconic, figurative style of interpretation, other images of the test conditions seem to call for different kinds of interpretative strategies due to e.g. their compositional character. In the following, we will compare the intersubjective, semantic meaning that subjects associated with two images, figure 1.c and 1.d, that were both consistently perceived as signals, but which were found to motivate different meaning-constructing attitudes. A comparative analysis is motivated by the fact that both images depict an arrow in a central position of the scene. In figure 1.c flour is sprinkled out to form an arrow on a path in a wood, and in image 1.d a road sign depicting an arrow is lying on a lawn surrounded by old vacuum cleaners arranged in rows. In the second coding procedure, we identified differences in subjects' primary interpretative strategies to these two images. While figure 1.c tends to motivate an instructional/indexical style of interpretation, figure 1.d calls for aesthetic/iconic meaning. Though an arrow could be thought of as a very conventional type of symbol with an almost inherent diagrammatical meaning (Stjernfelt, 2007), the differences in interpretative approach to the two scenes have some consequences for the kind of meaning that subjects end up assigning to the two exemplars of arrows. These differences are manifested in distributions of specific recognizable lexical items between subjects' image description.

In their descriptions of both images, subjects consistently use the noun 'arrow' (fig. 1.c: 20/20, fig. 1.d: 18/20) and in both cases about half of the subjects use the arrow-related verb 'to point' (fig. 1.c: 9/20, fig. 1.d: 11/20). This is taken to indicate that the arrows play an equally central role in the description of both images. But there are also noticeable differences in the distribution of lexical elements used to describe the two scenarios that could point to rather distinct interpretations of the arrows. Interestingly, in the descriptions of fig. 1.c we find more than twice as many modal verbs expressing deontic meanings (ex. should, ought to, have to, etc.) than in fig. 1.d (fig. 1.c: 26, fig. 1.d: 9). Thus, while the arrow in fig. 1.c is consistently associated with an instructional functional meaning (something that you should or should not do), this type of interpretation is almost totally lacking in fig. 1.d, as illustrated in the representative phrases below:

Example from a description of fig. 1.c:

*"You are in a forest and there is an arrow drawn in white chalk, I think, that points to the right, so you're probably not supposed to go up the small path that is straight ahead (...)"*

Example from a description of fig. 1.d:

*"(...) In front of the vacuum cleaner formation you see a road sign with an arrow pointing ahead to the foreground and this sign lies upon something that looks like cardboard on which is drawn the American flag".*

The differences in subjects' meaning-exploring attitudes towards the two scenes thus result in different interpretations of otherwise quite similar items – the arrows. While subjects largely agree that the arrow in fig. 1.c has a social deontic function, that it *instructs* you to do something, the arrow in fig. 1.d is seemingly "just" lying on the ground pointing, that is, without this modal meaning.

Again these examples point to the way subjects not only use material structures as a source of meaning but do so in a systematic manner seemingly regulated by orientation to specific styles of intentional manipulation and compositionality.

## Case 3: the ambiguous scene

Another stimuli image found to be a material signal (fig. 1e) depicts a red car with a bunch of thin tomato slices distributed in an ordered pattern on the windscreen. The car is apparently parked among other cars in a parking lot. The stimuli image is a good candidate for a perfectly ambiguous scene that lacks sufficient contextual guidance and thus opens for multiple conflicting interpretations.

Generally, subjects initiate their description in a somewhat reluctant manner expressing a lot of uncertainty in their understanding of the scene. The majority (11/20) explicitly comments on the "oddness" of the scene and some (7/20) note that the special arrangement of items could not have happened by coincidence or "the laws of nature". Still 12/20 subjects ends up making intentional communicative interpretations of the scene. Interestingly, though, subjects do not show the same pattern of consistent preference for one particular interpretative approach for this scene as in the previous cases. Instead, they seem to be equally attracted by two competing interpretative frames. Compare the fragments from two descriptions of the scene:

1) *"I don't think they got there naturally, so I think it's meant to be some kind of artistic statement, the way the slices of tomato have been put on that window. I don't know exactly what it should mean. It could be some kind of parallel with the color red, or, I don't really know. It could be health related, like the car has been infected with tomatoes, or something weird like that".*

2) *"It's some kind of hostile act, or somebody pulling somebody else's leg by putting slices of tomato on the windscreen of their car. Either way it's a pretty strange thing to do. And its not quite clear what's going on. It could of course be somebody playing a practical joke with (or on) some friends*".

The first example is representative for a group of subjects (9/20) who approach the scene applying an iconic, figurative attitude similar to the case with the red iceberg and the vacuum cleaners. Several subjects even suggest that the scene is intended as a "piece of art". Example 2 represents another approach taken by somewhat smaller group (6/20, there is a slight overlap, as three subjects make both interpretations). Here the tomato configuration is not appreciated for its iconic resemblances, nor as a social instructional imperative (like the arrow in fig. 1c). Rather, the scene is understood as a *symbolic message* between interlocutors applying a more arbitrary style of reference. Similar symbolic attitudes are found in descriptions of, for instance, fig. 1b depicting a crowd of roses left on a pavement in memory of a person murdered there.

A tentative explanation for the conflicting interpretative strategies observed above probably resides in the lack of or insufficiency of a regulating 'common ground' (H. Clark and Brennan, 1991) which gives rise to variability in the distribution and weight of attention that subjects apply to different constituent parts of the scene (Rollins, 2004). Though subjects tend to agree that the object configuration is intended as a communicative signal, the meaning seems to be irresolvable without additional priming of a contextual common ground. One could easily imagine the scene as something taken out of a larger context with the preceding events lending a lot of support for the proper interpretation. The object configuration may thus be addressing particular "initiates" that would have special prerequisites for knowing how to approach it. Naïve to this supporting context, the subjects of this study base their interpretation on teleological intuitions about the depicted part of the scene alone. The result is two quite distinct types of interpretations (iconic/aesthetic and symbolic) that seem to be driven by two different perceptual-attentional strategies: If we choose to focus locally on the symmetrical compositionality of the tomato slices and the sameness of color between car and tomatoes this would surely motivate an aesthetic reading. Someone seems to have put at least some effort into the well-ordered arrangement of the items. But if we instead take a more global perspective, disregard compositional details, and attend to the everyday surrounding scenery of the parking lot an artistic interpretation seems less obvious and the symbolic, "practical joke" interpretation more probable.

## 8. Discussion

Alignment of representation and top-top interaction

The results from the quantitative analyses of the first coding procedure seem to support our primary, initial hypothesis: all communication-related variables showed a significant effect across conditions, while the communication-neutral variable (*agency*) did not. This is taken by the authors to suggest that subjects consistently 'take a language stance' towards specific configurations of material objects and not others corresponding to the predefined conditions (Cowley, 2007; Spurrett & Cowley, 2004). In these cases the material objects are not only appreciated for their natural and instrumental affordances; their structural properties are also explored as sources of intersubjective social meaning (A. Clark, 2006; Roepstorff, 2008). Interestingly, subjects not only find these scenes to call for special meaning-exploring attitudes, they also consistently make explicit references to discursive roles such as 'addresser' and 'addressee', reflecting an element of interpersonal interaction in the interpretation of the static material configurations.

Contrary to A. Clark's account of material symbols as scaffolding 'private' thought processes, the approach that subjects take to the meaning-exploration of material objects in this study seems to be an inherently social affair. This fact is reflected in the strong alignment effects in subjects' interpretative strategies. Thus, in the second coding procedure we found that subjects not only agreed to approach certain object configurations as sources of meaning construction, they also quite consistently, though independently, oriented to the same kinds of interpretative strategies (iconic, instructional and symbolic) in their exploration of various test scenes. Interestingly, this alignment effect only applies to the test conditions and is not found in subjects' descriptions of control stimuli. Subjects' description of control stimuli are generally characterized by a high degree of variability. Besides merely listing the items present in the image, subjects focus and comment on very different aspects of the depicted scenes. Comments range from the possible seasons depicted to associations to distant childhood memories, and their interpretations show no inclination to align.

From such great discrepancies in the attitude to control scenes, subjects approach test scenes in far more similar manners, orienting for instance toward shared iconic-aesthetic, indexical-instructional, or symbolic meaning exploring strategies. Thus, it seems that in these cases some intentionally manifested compositional or contextual feature has a regulating and coordinating influence on subjects' responses, creating a considerable inter-subject representational alignment effect. Bearing in mind the hypothesized relationship between addressee-addressee and addressee-addresser alignment, when a large group of subjects agrees in their communicative interpretation of a scene, it is considered likely to be due to the successful regulating 'priming' of the addresser (Rollins, 2004).

We speculate that such priming mechanisms are based on socio-normative cognitive processes, related to Roepstorff and Frith's notion of "top-top" interactions (2004), and Pickering & Garrod's notion of alignment mechanisms (2004).

Both quantitative and more qualitative styles of analyses point to the fact that despite the absence of an addresser-agent or any explicit ostensive markers, the seemingly deliberate arrangement of objects is experienced as a kind of socially intended signal. Some property of the scenes strikes the spectator as an intentional communicative cue (cf. *semiotic agency,* Tylen, 2007) that has been left in the scene by an intentional communicating agent to implicitly "instruct" the addressee in terms of which interpretative strategy should apply for the scene (Levinson, 1979; 1993). Subjects' alignment of perceptual and interpretational strategy is thus motivated by the priming of a socio-normative script or frame for the appropriate interpretative approach to the scene (Pickering & Garrod, 2004). This initial recognition, then, seems to have the effect that subjects switch from a more "private", unregulated, free associative interpretation strategy to an orientation toward a "shared" social or cultural norm. This could be phrased something like the implicit understanding that "I am *supposed to* approach this item *in a particular manner*", or in a more pronounced version "someone *wants* me to *appreciate* this item applying a *symbolic, instrumental,* or *aesthetic attitude*". Rather than conveying a specific semantic content, these cues thus suggest to the addressee a relevant frame for what to attend to or how to investigate the material scene. When subjects in a consistent manner choose an aesthetic-iconic interpretative attitude toward one scene and an instructional-deontic attitude toward another, it is not because these kinds of meanings are themselves more objectively 'present' in the scenes than others. They are made socially relevant by the priming efforts of some addressing agent. In this way, the top-down attitude that comes to guide the interpretation of the scene is being regulated by the intentional intervention of another agent, resulting in a kind of "top-top interaction" - the attunement to and sharing of a conceptual script or frame for the proper investigation of the material scene (Roepstorff & Frith, 2004).

Once the subject-addressee, implicitly informed by the addresser, settles on an interpretative attitude, s/he can now approach the structural properties of a scene in a regulated meaning-constructing manner. If a compositional feature of a scene for instance primes subjects toward an aesthetic reading (cf. the iceberg and the vacuum cleaner scenes), then these consistently pick up on image-inherent iconic properties and associate figurative and diagrammatic meanings to these. On the other hand, when another object configuration primes subjects toward a social instructional reading, subjects tend to orient to image-external social deontic meanings, etc.

The findings tell a story somewhat different from Andy Clark's. Rather describing subjects' meaning-exploring approaches to material objects as a facilitation of new forms of 'private thinking' (A. Clark, 2006), we suggest that these processes becomes socially coordinated and regulated in the ways described above, resulting in a kind of 'collective thinking'.

## Multiple stages of comprehension

In the test conditions, inter-subject alignment effects are found on three levels of analysis, suggesting a layered hierarchy of different stages of comprehension. The primary distinction investigated in the first coding procedure was object configurations experienced as communicative signals vs. non-signals. The recognition of the intended representational status of an object configuration seems necessary for the application of an overall meaning-constructing attitude. But as apparent from the second coding procedure, a subsequent distinction between various 'interpretative strategies' seems derivable from subjects' various efforts at meaning-exploration. In the present study, we identified three such attitudes or strategies: aesthetic, instructional and symbolic (related to Pierce's classical sign types: *icon, index* and *symbol*, Peirce, 1998). The choice of interpretative attitude has a regulating influence on the kind of intersubjective semantic meanings that is attributed to the material signal. But even the motivated choice of a certain interpretative attitude seemingly cannot fully predict the final semantic interpretation of an object configuration. For instance, though subjects tend to agree in their choice of an iconic strategy in their approach to the iceberg scene (case 1), they still end up assigning quite different semantic meanings to the scene (strawberry-sorbet vs. blood, respectively). The comprehension of a material signal can thus be understood as a nested, multi-stage process, from the very general categorical resolutions signal vs. non-signal to the choice of interpretative attitude narrowing down to the more fine-grained semantic resolutions (see figure 4).

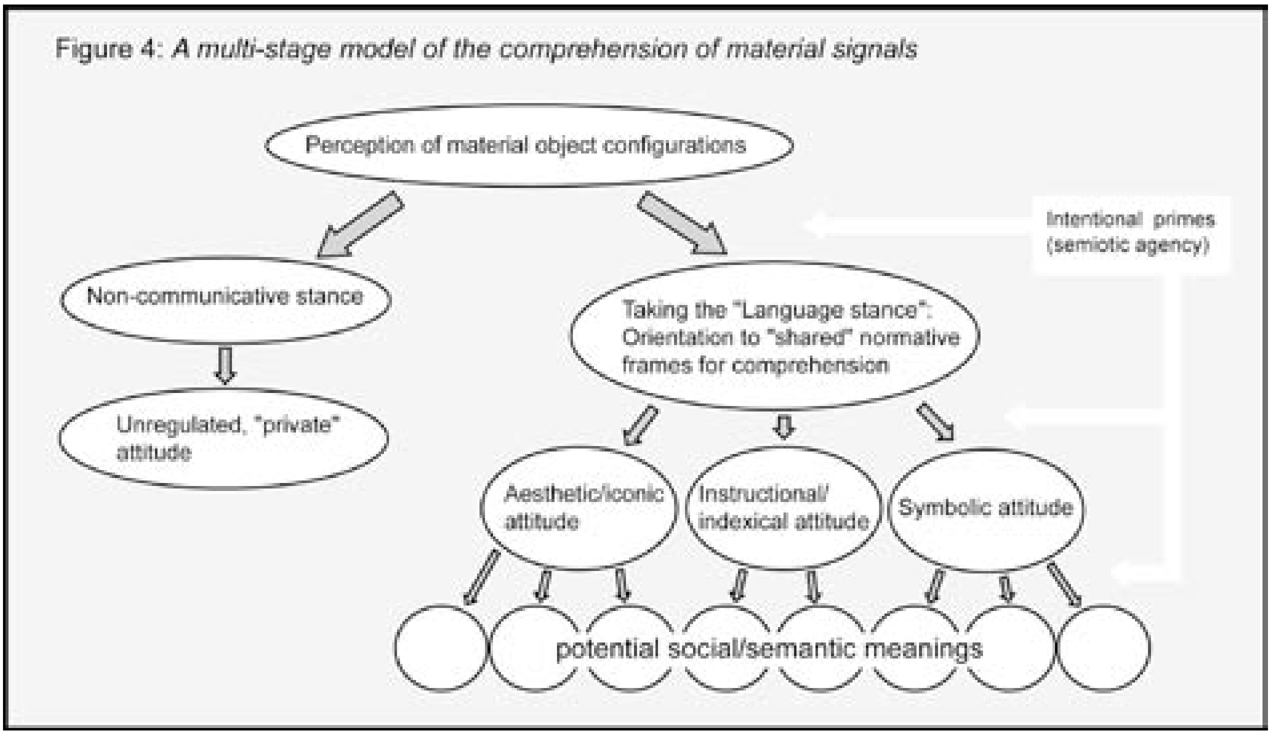



***Figure 4***: A multi-stage model of the comprehension of material signals. The model depicts a nested hierarchy of levels in the interpretative approach to different kinds of material scenes. If the perceiver does not associate the scene with any intentional communicative relevance, the interpretation becomes casual and unregulated (left side). But when perceivers take a "language stance" and explore a scene as means of social meaning construction, they tend to orient to normatively regulating frames for interpretation (right side).

## Conventionality revisited

In this study we observed strong inter-subject alignment especially at the initial categorical level (signal vs. non-signal) of the hypothesized model of comprehension, followed by a slight tendency to decrease in alignment as we proceed towards the more subtle stages of comprehension. Interestingly, this tendency toward decreasing alignment appears to be particularly strong for the unconventional material signals (the test II condition), as revealed by the second coding procedure. Compare for instance the differences in multilevel alignment between fig. 1.b (a crowd of roses left on the pavement in a city environment) rated to be a quite conventional social signal, and fig. 1.d (depicting rows of vacuum cleaners lined up at a lawn) rated to be an unconventional or infrequent kind of signal. In the descriptions of the rose-scene, 18/20 subjects categorize the roses as a signal. Of these, 17 choose a symbolic interpretative strategy (while only one subject makes an aesthetic reading) and again all 17 subjects arrive at a semantic reading relating the roses to grief in relation to the recent, unnatural death of someone at the particular spot. In comparison, 17/20 subjects categorize the vacuum cleaners in fig. 1.d scene as a signal. Out of these, 12 subjects agree in their choice of an iconic-aesthetic interpretative strategy while others are more hesitant in their interpretative strategies. The iconic-aesthetic interpretations can roughly be divided into two groups of more specific semantic readings: 7 subjects associate the vacuum cleaner formation with a cemetery, the other 4 with an army of warriors. Thus, while we observe a general high inter-subject alignment at all stages of comprehension (with only a slight decrease from stage one to two) in the case of the conventional material signal in fig. 1.b, this alignment effect rapidly decreases at the more subtle levels of comprehension in the unconventional material signal depicted in fig. 1.d. The relatively poor alignment of semantic meanings in this case is probably due to representational ambiguities, resulting in a somewhat higher propensity of subjects to "spread out" between several alternative interpretations.

The differences in distributive patterns proposed above seem to apply rather consistently to the two groups of images constituting the test conditions. While the conventionality vs. unconventionality distinction did not show great effect at the primary categorical level of comprehension assessed in the first coding procedure, it turns out to play an important role at lower levels as a predictor of the relative spread in semantic attributions and resulting decreases in inter- and intra-subject alignment. Based on the authors' cautious assumption regarding the complementary relationship between 'addressee-addressee' and 'addresser-addressee' alignment effects, it appears that intersubjective meaning, when mediated by unconventional object configurations, is more effortlessly recognized. However, unconventional object configurations are not as efficient as conventional signals in regulating the interpretative strategies of addressees at more subtle stages of comprehension and thus may not result in particularly well-coordinated meaning constructing processes.

The results reported above find support in a recent brain imaging study targeting the neural substrates of our comprehension of conventional and unconventional material signals (Tylén et al, in press). The finding was that material signals elicit activity in a series of brain areas normally associated with verbal language. Interestingly, some of these areas related to the more categorical resolutions (signal vs. non-signal) seemed insensitive to the relative degree conventionality, while others involved in the more fine-grained semantic processing showed a conspicuous conventionality-effect, corresponding to our findings in the present study.

## Common ground and recipient design

Relative frequency or conventionality might not be the whole story. A related issue is the amount of contextual information surrounding the material signal. In the stimuli of this study, only a minimum of contextual information, constituted by the immediate physical surroundings of the object(s), is directly available. This kind of receptive situation is thus thought to simulate the coincidental, unqualified and unbiased subject-object encounter to assess subjects' naïve and casual understanding. The authors acknowledge that this may not be the best and most realistic way of modeling our everyday usage of material signals. Often such signals are designed to address specific individuals or groups of recipients enrolled in larger contexts that will lend important support to the interpretative process. Taken out of this larger story, the material signal may be almost incomprehensible. This is likely the case with the tomato slice-scene (case 3 above). The unconventional style of the signal and the very limited contextual access challenge the interpretative process. Still, we can easily imagine the arrangement to be unambiguous and meaningful to a specific group of initiates somehow involved in the interaction. Thus, the unconventional style of material expression may not necessarily be less accurate and regulative, but will probably to a larger extent depend on the contextual setting, relevance criteria and recipient design (Sperber & Wilson, 1995; Sacks et al, 1974). In contrast, the force of more conventional communication is that it gradually becomes decontextualized, or rather, the relatively greater experience with these kinds of signaling practices enhance addressees' ability to reconstruct a missing or incomplete contextual frame for the intended message (Tomasello, 2003).

## 9. Summary and Conclusion

In this study, we set out to investigate how configurations of static material objects gain a mediating function in human communicative interaction. In a content analysis experiment, we assessed subjects' unbiased comprehension of a series of images depicting three predefined categories of material scenes: conventional material signals, unconventional material signals and object configurations without any conspicuous communicative relevance. Quantitative and qualitative analyses of subjects' introspective reports suggest that they consistently approach some object configurations by taking a 'language stance' (Cowley, 2007; Spurrett & Cowley, 2004). That is, they systematically explore certain material properties and compositions of objects (depicted in the two test conditions) as a means for intersubjective semantic meaning construction. The impact of the materiality itself on these cognitive processes is considered to reflect Andy Clark's complementary view of material symbols (A. Clark, 2006).

Furthermore, this approach of active meaning-exploration to test scenes seems to be socially regulated. From the great variability in the interpretative strategies that subjects apply in their descriptions of control stimuli, these strategies seem to align in interesting ways when subjects approach images of the test conditions. Subjects thus largely agree to make iconic-aesthetic readings of some scenes and instructional-deontic or symbolic readings of others. We speculate that these alignment effects are due to a motivated orientation to a 'shared interpretative norm'. Some intentional compositional properties of the scene seem to function as primes implicitly "instructing" subjects to take on a certain attitude in relation to a scene. The choice of top-down interpretative strategy is thus indirectly regulated by the intentional intervention of another agent (the addresser). Communicative interaction may thus not be so much about 'conveying' specific semantic meanings, but rather about negotiating and sharing scripts and frames for the proper exploration of the social and material world, a practice termed "top-top interaction" by Roepstorff & Frith (2004).

The alignment of subjects' interpretative strategies was observed at various levels of interpretation, suggesting a nested multi-stage model of comprehension in relation to material signals (illustrated in figure 4). When subjects in a motivated manner take the language stance and approach a material scene as a means for intersubjective meaning construction, they face a series of other interpretative dispositions. Alignment at subtle levels of semantic interpretation thus depends on alignment at higher levels including the choice of an aesthetic, instructional or symbolic interpretative attitude.

While inter-subject alignment effects are comparably strong at the higher and more abstract levels of comprehension for images of both test conditions, there seems to be a 'conventionality effect' at the more fine-grained levels of comprehension. Relative to the conventional scenes, the subjects' descriptions of unconventional material signals tend to show enhanced variability (decrease in alignment) at the lower levels of the comprehension process, indicating a less regulated interpretative process. We speculate this is due to two related factors: 1) the relative dependency on extra contextual support in cases of more unentrenched forms of communicative mediation (Tomasello, 2003), and 2) the aspect of recipient design, that is, the degree to which a material signal is tailored for a specific individual or group of addressees with special interpretative prerequisites (Sacks et al, 1974).

Overall, our findings suggest that materials signals can be understood not only as a case of how cognition extends into the material environment in a continuous fashion (A. Clark & Chalmers, 1998; A. Clark, 2006; Hutchins, 1995), but in addition how *social* cognition (norms, primes and frames) pervades our shared experience of the material environment.

## Acknowledgements

The authors would like to thank Chris Frith, Uta Frith, Mikkel Wallentin, Andreas Roepstorff, Kim Mouridsen, Politiken Fonden and the Danish National Research Foundation's Center for Functionally Integrative Neuroscience

## References:

Abell, F., Happé, F., Frith, U. (2000). Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development, *Cognitive Development,* 15:1-16

Banerjee, M., Capozzoli, M., McSweeney, L., Sinha, D. (1999). Beyond kappa: a review of interrater agreement measures, *Canadian Journal of Statistics,* 27(1):3-23

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

Clark, A. (2006). Material Symbols, *Philosophical Psychology,* 19(3):291-307

Clark, A., Chalmers, D.J. (1998). The Extended Mind, *Analysis,* 58:10-23

Clark, H.H. (1996). *Using Language*, Cambridge University Press, Great Britain

Clark, H.H. (2005). Coordinating with each other in a material world, *Discourse Studies*, 7(4-5):507-525

Clark, H.H., Brennan, S.A. (1991). Grounding in Communication, in L.B. Resnick, J.M. Levine, and S.D. Teasley (eds.), *Perspectives on Socially Shared Cognition*, APA Books, Washington D.C., 127-149

Cohen, J. (1960). A coefficient of agreement for nominal scales, *Educational and Psychological Measurement,* 20(1):37-46

Cowley, S. (2007). How Human Infants Deal with Symbol Grounding, *Interaction Studies,* 8(1): 81-104

Deacon, T.W. (1997). *The Symbolic Species – The Co-evolution of Language and the Brain*. New York: W.W. Norton.

De Jaegher, H., Di Paolo, E. (2007). Participatory sense-making: An enactive approach to social cognition, *Phenomenology and the Cognitive Sciences*, 6:484-507

Donald, M. (2002). *A Mind So Rare: the evolution of human consciousness,* Norton, New York

Gallagher, S. (2003). Phenomenology and Experimental Design: Toward a Phenomenologically Enlightened Experimental Science, in Anthony Jack & Andreas Roepstorff (eds.) *Trusting the Subject? The use of introspective evidence in cognitive science,* vol. 1, Imprint Academic, UK

Hutchins, E. (1995). *Cognition in the Wild*, Cambridge, Mass. MIT Press

Jack, A.I., Roepstorff, A. (2002). Introspection and Cognitive Brain Mapping: From Stimulus-Response to Script-Report, *TRENDS in Cognitive Science,* 6(8):333-339

Kravchenko, A.V. (2007). Essential properties of language, or, why language is not a code, *Language Sciences,* 29:650-671

Landis J.R., Koch, G.G. (1977). The measurement of observer agreement for categorical data, *Biometrics*, 33:159--174

Levinson, J. (1979). Defining art historically, *British Journal of Aesthetics,* 19(3):232-250

Levinson, J. (1993). Extending art historically, *The Journal of Aesthetics and Art Criticism,* 51(3):411-423

MacWhinney, B. (2000). The CHILDES Project: Tools for Analyzing Talk. 3rd Edition. Mahwah, NJ: Lawrence Erlbaum Associates

Myers, L.J., Liben, L.S. (2008). The Role of Intentionality and Iconicity in Children's Developing Comprehension and Production of Cartographic Symbols, *Child Development,* 79(3):668-684

Neuendorf, K.A. (2002). *The content analysis guidebook*, Sage Publications, Inc., USA

Peirce, C.S. (1998) *Collected Papers*, I-VIII, (ed. Hartshorne and Weiss; Burks) London: Thoemmes Press (1931-58)

Pickering, M.J., Garrod, S. (2004). Toward a mechanistic psychology of dialogue, *Behavioral and Brain Sciences,* 27:169-226

Roepstorff, A. (2008). Things to think with: words and objects as material symbols, *Philosophical Transactions of the Royal Society B,* 363:2049-2054

Roepstorff, A., Frith, C.D. (2004). What's at the top in the top-down control of action? Script-sharing and 'top-top' control of action in cognitive experiments, *Psychological Research,* 68: 189198

Rollins, M. (2004). What Monet Meant: Intention and attention in understanding art, *The Journal of Aesthetics and Art Criticism,* 62(2):175-188

Sacks, H., Schegloff, E.A., Jefferson, G. (1974). A Simplest Systematics for the Organisation of Turn-Taking for Conversation, *Language*, 50:696-735

Searle, J. R. (1995). *The Construction of Social Reality*, New York Free Press

Sperber, D., Wilson, D. (1995). *Relevance: communication and cognition,* 2nd edition, Oxford, UK:Blackwell

Spurrett, D., Cowley, S. (2004). How to do things without words: Infants, utterance-activity and distributed cognition, *Language Sciences,* 26(5):443-466

Stjernfelt, F. (2007). *Diagrammatology: an investigation on the borderlines of phenomenology, ontology, and semiotics,* Springer, Nederlands

Tomasello, M. (2003). *Constructing a Language: a usage-based theory of language acquisition,* Harvard University Press, London

Tylén, K. (2007). When agents become expressive: a theory of semiotic agency, *Cognitive Semiotics,* promotional issue 0:84-101

Tylén, K, Wallentin, M., Roepstorff, A. (in press). Say it with flowers! An fMRI study on object mediated communication, *Brain and Language,* doi:10.1016/j.bandl.2008.07.002

# 8 References

Abell, F., Happé, F., Frith, U. (2000). Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development, *Cognitive Development*, 15:1-16

Aguirre, G.K. (2003). Functional Imaging in Behavioral Neurology and Cognitive Neuropsychology, in T. E. Feinberg & M. J. Farah (Eds.), *Behavioral Neurology and Cognitive Neuropsychology*. New York: McGraw Hill.

Amaro, E., Barker, G.J. (2006). Study design in fMRI: Basic Principles, *Brain and Cognition*, 60:220-232

Amodio, D.M., Frith, C.D. (2006). Meeting of Minds: the medial frontal cortex and social cognition, *Nature Reviews, Neuroscience,* 7:268-277

Ashburner, J., Friston, K.J., (1999). Nonlinear spatial normalization using basis functions. *Human Brain Mapping*, 7(4):254– 266.

Bates, E., Tomasello, M., eds. (2001). *Language Development: the essential readings*, Blackwell Publishers, Oxford

Bedny, M., Hulbert, J.C., Thompson-Schill, S.L., (2007). Understanding words in context: The role of Broca's area in word comprehension, *Brain Reseach,* 1146:101-114

Bloom, P. (1996). Intention, history and artifact concepts, *Cognition,* 60(1):1-29

Bloom, P. (2006) Seduced by the flickering lights of the brain, *SeedMagazine.com: Brain and Behavior*, posted june 27

Bloom, P. Markson, L. (1998). Intention and analogy in children's naming of pictorial representations, *Psychological Science,* 9(3):200-205

Buxton, R.B., Uludag, K., Dubowitz, D.J., Liu, T.T. (2004). Modeling the hemodynamic response to brain activation, *NeuroImage,* 23, supplement 1:220-233

Castelli, F., Happé, F., Frith, U., Frith, C.D. (2000). Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *NeuroImage,* 12(3):314–325.

Chomsky, N. (1972). *Language and Mind,* Harcourt Brace Jovanovich, New York

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

Clark, A. (2006a). Material Symbols, *Philosophical Psychology,* 19(3):291-307

Clark, A. (2006b). Language, embodiment and the cognitive niche, *Trends in Cognitive Sciences,* 10:370-374

Clark, A., Chalmers, D.J. (1998). The Extended Mind, *Analysis,* 58:10-23

Clark, H.H. (1996). *Using Language,* Cambridge University Press, UK

Clark, H.H. (2005). Coordinating with each other in a material world, *Discourse Studies,* 7:507-525

Cohan, M.S., Bookheimer, S.Y. (1994). Localization of brain function using magnetic resonance imaging, *Trends in Neurosciences,* 17(7):268-277

Cohen, J. (1960). A coefficient of agreement for nominal scales, *Educational and Psychological Measurement,* 20(1):37-46

Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M.A., Michel,

F. (2000). The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients, *Brain*, 123:291-307

Cohen, L., Lehericy, S., Chochon, F., Lemer, C., Rivaud, S., Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area, *Brain*, 125:1054-1069

Cowley, S. (2004). Contextualizing bodies: human infants and distributed cognition, *Language Sciences,* 26:565–591

Csibra, G., Gergely, G. (2006). Social Learning and Social Cognition: the case for pedagogy, in Y. Munakata & M. H. Johnson (Eds.), *Processes of Change in Brain and Cognitive Development. Attention and Performance*, XXI, pp. 249-274, Oxford: Oxford University Press

Csibra, G., Volein, A. (2008). Infants can infer the presence of hidden objects from referential gaze information, *British Journal of Developmental Psychology*, 26:1–11

D'Esposito, M., Zarahn, E. Aguirre, G.K. (1999). Event-related functional MRI: Implications for cognitive psychology, *Psychological Buletin,* 125:155-164

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

De Jaegher, H., Di Paolo, E. (2007). Participatory Sense Making: an enactive approach to social cognition, *Phenomenology and the Cognitive Sciences,* 6(4):485-507

Deacon, T.W. (1997). *The Symbolic Species: the co-evolution of language and the brain*, W.W. Norton, New York

Dehaene, S., Le Clec, H.G., Poline, J.B., Le Bihan, D., Cohen, L. (2002). The Visual Word Form Area: A prelexical representation of visual words in the fusiform gyrus, *Neuroreport*, 13:321-325

DeLoache, J. (2000). Dual Representation and Young Children's Use of Scale Models, *Child Development,* 71(2):329-338

DeLoache, J.S. (1996). Shrinking trolls and expanding minds: Early symbolic development, *Psychological Science Agenda,* 8-9

DeLoache, J.S. (2000). Dual Representation and Young Children's Use of Scale Models, *Child Development,* 71(2):329-338

DeLoache, J.S. (2004). Becoming symbol-minded, *Trends in Cognitive Sciences, 8,* 66-70

DeLoache, J.S., Miller, K.F., Rosengren, K.S. (1997). The Incredible Shrinking Room: Very young children's performance with symbolic and non-symbolic relations, *Psychological Science,* 8(4): 308-313

DeLoache, J.S., Peralta de Mendoza, O.A., Anderson, K.N. (1999). Multiple Factors in Early Symbol Use: Instructions, Similarity, and Age in Understanding a Symbol-Referent Relation, *Cognitive Development,* 14:299-312

Dennett, D.C. (1990). The interpretation of texts, people and other artifacts, *Philosophy and Phenomenological Research,* 1:177-194

Di Paolo, E., Rohde, M., De Jaegher, H. (2008). Horizons fro the enactive mind: values, social interaction and play, in *Enaction: Towards a New Paradigm for Cognitive Science*, J.

Stewart, O. Gapenne, and E. A. Di Paolo (eds.), Cambridge, MA: MIT Press

Diesendruck, G., Markson, L., & Bloom, P. (2003). Children's reliance on the creator's intent in extending names for artifacts. Psychological Science, 14, 164–168

Dietrich, S., Hertrich, I., Alter, K., Ischebeck, A., Ackermann, H., (2007). Semiotic aspects of nonverbal vocalizations: a functional imaging study, *NeuroReport,* 18(18):1891-1894

Epstein, R., Kanwisher, N. (1998). A cortical representation of the local visual environment, *Nature,* 392:598-601

Fiez, J. A. (1997). Phonology, semantics, and the role of the left inferior prefrontal cortex. *Human Brain Mapping*, 5:79–83

Fletcher, P.C., Happé, F., Frith, U., Baker, S.C., Dolan, R.J., Frackowiak, R.S., Frith C.D. (1995). Other Minds in The Brain: a functional imaging study of "theory of mind" in story comprehension, *Cognition,* 57(2):109-128

Fodor, J.A. (1975). *The language of thought,* Harvard University Press

Fodor, J.A. (1998). Do we think in mentalese: Remarks on some arguments of Peter Carruthers, in *Critical Conditions: Polemical essays on cognitive science and the philosophy of mind,* Cambridge, MA: MIT Press

Friston, K. J., Holmes, A. P., J, W. K., Poline, J. B., Frith, C. D., and Frackowiak, R. S. (1995b). Statistical Parametric Maps in Functional Imaging: A General Linear Approach. *Human Brain Mapping*, 2:189-210.

Friston, K.J., Ashburner, J., Poline, J.B., Frith, C.D., Heather, J.D., Frackowiak, R.S., (1995a). Spatial registration and normalization of images, *Human Brain Mapping,* 2:165– 189.

Friston, K.J., Price, K.J., Fletcher, P., Moore, C., Frackowiak, R.S.J., Dolan, R.J. (1996). The Trouble with Cognitive Subtraction, *NeuroImage*, 4:97-104

Friston, K.J., Zarahn, E., Josephs, O., Henson, R.N., Dale, A.M. (1999). Stochastic designs in event-related fMRI, *NeuroImage,* 10:607–619

Frith, C.D., Frith, U. (2006). How we predict what other people are going to do, *Brain Research,* 1079(1):36-46

Frith, C.D., Frith, U., Roepstorff, A. (2008). Interacting Minds: progress and future, *presentation at the Cognition, Communication and Culture (CCC) Workshop*, University of Aarhus, April 4th.

Frith, U., Frith, C.D., (2003). Development and neurophysiology of mentalizing, *Philosophical Transactions of the Royal Society: Biological Sciences*, 358:459– 473.

Frith, U., Frith, C.D., (2005). Theory of Mind, *Current Biology,* 15(17):644-645

Gallagher, H.L., Frith, C.D. (2003). Functional Imaging of 'Theory of Mind', *TRENDS in Cognitive Science,* 7(2):77-83

Gallagher, H.L., Frith, C.D. (2004). Dissociable neural pathways for the perception and recognition of expressive and instrumental gestures, *Neuropsychologia,* 42(13):1725-36

Gallagher, H.L., Happa, F., Brunswick, N., Fletcher, P.C., Frith, U. Frith, C.D. (2000). Reading the Mind in Cartoons and Stories: an fMRI study of 'theory of mind' in verbal and nonverbal tasks, *Neuropsychologia,* 38(1):11-21

Gallagher, H.L., Jack, A.I., Roepstorff, A., Frith, C.D. (2002). Imaging the Intentional Stance in a

Competitive Game, *NeuroImage,* 16:814-21

Gallagher, S. (2003). Phenomenology and Experimental Design: Toward a phenomenologically enlightened experimental science, in Anthony Jack & Andreas Roepstorff (eds), *Trusting the Subject?: the use of introspective evidence in cognitive science,* vol. 1, Imprint Academic, UK

Gallese, V., Fadiga, L., Fogassi, L., *et al.*. (1996). Action recognition in the premotor cortex, *Brain*, 119:593–609.

Gallese, V., Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading, *TRENDS in Cognitive Sciences,* 2(12):493-501

Ganea, P.A., Pickard, M.B., DeLoache, J.S. (2008). Transfer between picture books and the real world by very young children, *Journal of Cognition and Development,* 9:46-66

Gelman, S.A., Bloom, P. (2000). Young children are sensitive to how an object was created when deciding what to name it, *Cognition,* 76:91-103

Gelman, S.A., Ebeling, K.S. (1998). Shape and representational status in children's early naming, *Cognition*, 66:35-47

Gergely, G., Egyed, K., Király, I. (2007). On Pedagogy, *Developmental Science,* 10(1):139-146

Gibbs, R.W. (2007). Why Cognitive Linguistics Should Care More About Empirical Methods, in Gonzalez-Marquez, M., Mittelberg, I., Coulson, S., & Spivey, M. (eds.), *Methods in Cognitive Linguistics*, Amsterdam: John Benjamins

Gibson, J.J. (1979). *The Ecological Approach to Visual Perception,* Boston: Houghton Mifflin

Goldberg R.F., Perfetti C.A., Fiez J.A., Schneider W. (2007). Selective retrieval of abstract semantic knowledge in left prefrontal cortex, *Journal of Neuroscience*, 4;27(14):3790-3798

Grice, H.P. (1957). Meaning, *Philosophical Review,* 66:377-388

Gunter, B. (2000). *Media research methods: Measuring audiences, reactions and impact,* London: Sage

Gutheil, G., Bloom, P., Valderama, N., Freedman, R. (2004). The role of historical intuitions in children's and adults' naming of artifacts, *Cognition,* 91, pp. 23-42

Hagoort, P. (2005). On Broca, brain, and binding: a new framework, *Trends in Cognitive Sciences*, 9(9):416-423

Haushofer, J., Kanwisher, N. (2007). In the eye of the beholder: Visual experience and categorization in the human brain, *Neuron,* 53:773-775

Hickok, G., Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception, *Trends in Cognitive Sciences*, 4:131–138

Hickok, G., Poeppel, D. (2007). The cortical organization of speech processing, *Nature Neuroscience,* 8:393-402

Holsti, O.R. (1969). *Content Analysis for the Social Sciences and Humanities.* Addison Wesley, Reading, Mass.

Hutchins, E. (1995). *Cognition in the Wild*, Cambridge, Mass. MIT Press

Jack, A.I., Roepstorff, A. (2002). Introspection and Cognitive Brain Mapping: from stimulus-response to script-report, *TRENDS in Cognitive Sciences*, 6(8):333-339

Jack, A.I., Roepstorff, A. (eds.) (2003). *Trusting the Subject? The use of introspective evidence in* *cognitive sciences*, *vol. 1*, Imprint Academic, UK

Jakobson, R. (1960). Closing Statements: Linguistics and Poetics, in Thomas A. Sebeok (ed.), *Style In Language*, Cambridge Massachusetts, MIT Press

Johnson, K.E., Younger, B.A., Furrer, S.D. (2005). Infants' symbolic comprehension of actions modeled with toy replicas, *Developmental Science,* 8(4):299-314

Jones, R. (2002). Functional Imaging: Monkeys see, fMRI do, *Nature Neuroscience,* 3(7)

Kampe, K.K., Frith, C.D., Frith, U. (2003). "Hey John": signals conveying communicative intention toward the self activate brain regions associated with "mentalizing," regardless of modality, *Journal of Neuroscience*, 23(12):5258-5263

Kanwisher, N., Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces, *Philosophical Transactions from the Royal Society B,* 361:2109-2128

Kelemen, D. (1999). Function, goal and intention: children's teleological reasoning about objects, *TRENDS in Cognitive Sciences,* 3(12):461-468

Kravchenko, A.V. (2007). Essential Properties of Language, or, why language is not a code, *Language Sciences,* 29:650-671

Krippendorff, K. (1995). On the reliability of unitizing continuous data, *Sociological Methodology*, 25:47-76

Landau, B., Smith, L.B., Jones, S.S. (1998). Object shape, object function, and object name, *Journal of Memory and Language*, 38:1–27

Lawrence, E.J., Shaw, P., Giampietro, V.P., Surguladze, S., Brammer, M.J., Davis, A.S. (2006). The role of 'shared representations' in social perception and empathy: An fMRI study, *NeuroImage,* 29(4):1173-1184

Lawrence, E.J., Shaw, P., Giampietro, V.P., Surguladze, S., Brammer, M.J., Davis, A.S. (2006). The role of 'shared representations' in social perception and empathy: An fMRI study, *NeuroImage,* 29(4):1173-1184

Leslie, A.M. (1987). Pretence and representation: the origin of "theory of mind", *Psychological Review,* 94(4):412-426

Leslie, A.M. (1993). A Theory of Agency, *Technical Reports from Rutgers University*, London: Center for Cognitive Science

Levinson, J. (1979). Defining art historically, *British Journal of Aesthetics*, 19(3):232-250

Levinson, J. (1989). Redefining art historically, *The Journal of Aesthetics and Art Criticism,* 47:21-33

Levinson, J. (1992). Intention and Interpretation: A Last Look. In G. Iseminger (ed.), *Interpretation, Intention, and Truth*, Temple University Press

Levinson, J. (1993). Extending art historically, *The Journal of Aesthetics and Art Criticism,* 51(3):411-23

Levinson, J. (1996). *The Pleasure of Aesthetics: Philosophical essays,* Cornell University Press, Ithaca, NY

Leyton, M. (1992). *Symmetry, Causality, Mind,* The MIT Press, Cambridge, Massachusetts

Leyton, M. (2006) *The Structure of Paintings*, Springer, New York

Linell, P. (2005). *The written language bias in linguistics: its nature, origins and transformations*, Routledge, London and New York

Livingston, P. (2005). *Art and Intention: a philosophical study,* Oxford: Oxford University Press

Logothetis, N.K., Pauls, J., Augath, M., Trinath, T., Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal, *Nature,* 412:150-157

Lombard, M., Snyder-duch, J., Bracken, C.C. (2002). Content analysis in mass communication: Assessment and reporting of intercoder reliability, *Human Communication Research,* 28(4): 587-604

Long, G.M. & Wurst, S.A. (1984). Complexity Effects on Reaction-Time Measures of Visual Persistence: Evidence for Peripheral and Central Contributions, *The American Journal of Psychology*, 97(4): 537-561

Lotze, M., Heymans, U., Birbaumer, N., Veit, R., Erb, M. Flor, H., Halsband, U. (2006). Differential cerebral activation during observation of expressive gestures and motor acts, *Neuropsychologia,* 44: 1787-1795

Malt, B.C., Johnson, E.C. (1992). Do artifact concepts have cores? *Journal of Memory and Language,* 31:195-217.

Miller, G.A., Johnson-Laird, P. (1976). *Language and perception*. Cambridge, MA: Harvard University Press

Mohler, P.Ph., Zuell, C. (2000). Observe! A Popperian Critique of Automatic Content Analysis, paper form *JADT 2000: 5es Journées Internationales d'Analyse Statistique des Données Textuelles*, available from: http://www.cavi.univparis3.fr/lexicometrica/jadt/jadt2000/pdf/10/10.pdf

Müller, R-A., Basho, S. (2004). Are non-linguistic functions in "Broca's area" prerequisites for language acquisition? FMRI findings from an ontogenetic viewpoint, *Brain and Language,* 89: 329-336

Myers, L.J., Libens, L.S. (2008). The role of intentionality and iconicity in children's developing comprehension and production of cartographic symbols, *Child Development,* 79(3):668-684

Neuendorf, K. (2002). *The Content Analysis Guidebook*, Sage Publications Inc, USA

Núñez, R. (2007). Inferential statistics in the context of empirical cognitive linguistics, in Gonzalez-Marquez, M., Mittelberg, I., Coulson, S., & Spivey, M.J. (eds.), *Methods in Cognitive Linguistics,* John Benjamins, Amsterdam

Ogawa, S., Lee, T., Kay, A., Tank, D. 1990. Brain Magnetic Resonance Imaging with Contrast Dependent on Blood Oxygenation. *PNAS* **87**(24): 9868-9872

Papathanassiou, D., Etard, O., Mellet, E., Zago, L., Mazoyer, B. , Tzourio-Mazoyer, N. (2000). A Common Language Network for Comprehension and Production: A Contribution to the Definition of Language Epicenters with PET, *NeuroImage,* 11:347-357

Peirce, C.S. (1998) *Collected Papers*, I-VIII, (ed. Hartshorne and Weiss; Burks) London: Thoemmes Press (1931-58)

Pickering, M.J., Garrod, S. (2004). Toward a mechanistic psychology of dialogue, *Behavioral and Brain Sciences,* 27:169-226

Pinker, S. (1994). *The Language Instinct: how the mind creates language,* William Morrow, New York

Posner, M.I., Petersen, S.E., Fox, P.T., Raichle, M.E. (1988). Localization of cognitive operations in the human brain, *Science,* 24:1627-31

Preissler, M.A., Bloom, P. (2008). Two-year-olds use artist intention to understand drawings, *Cognition,* 106(1):512-18

Price, C.J. (2000). The Anatomy of Language: contributions from functional neuroimaging, *Journal of Anatomy,* 197:335-359

Quan, R.H. (1979). Photography and the creation of meaning, *Art Education,* 32(2):4-9

Rakoczy, H., Tomasello, M., & Striano, T. (2005). How children turn objects into symbols: A cultural learning account. In L. Namy (Ed.), *Symbol use and symbol representation,* New York: Erlbaum

Ramberg, B. (2004). Naturalizing idealizations: Pragmatism and the interpretivist strategy, *Contemporary Pragmatism,* 1(2):1-63

Rizzolatti, G., Fadiga, L., Gallese, V., Fogassi, L. (1996). Premotor cortex and the recognition of motor actions, *Cognitive Brain Research*, 3:131-141

Rodd, J.M., Davis, M.H., Johnsrude, I.S., (2005). The Neural Mechanisms of Speech Comprehension: fMRI studies of Semantic Ambiguity, *Cerebral Cortex,* 15:1261-1269

Roepstorff, A. (2008). Things to think with: words and objects as material symbols, *Philosophical Transactions of the Royal Society B,* 363:2049-2054

Roepstorff, A., Frith, C.D. (2004). What's at the top in the top-down control of action? Script-sharing and 'top-top' control of action in cognitive experiments, *Psychological Research,* 68: 189198

Rollins, M. (2004). What Monet Meant: intention and attention in understanding art, *The Journal of Aesthetics and Art Criticism*, 62(2):175-88

Rosenthal, H., Hall, J.A., DiMatteo, M.R., Rogers, P.L., Archer, D., (1979). Sensitivity to Nonverbal Communication: The PONS Test. The Johns Hopkins University Press, Baltimore

Saussure, F. (1974). *Course in General Linguistics*, Fontana/Collins, London

Saxe, R. (2006). Uniquely Human Social Cognition, *Current Opinion in Neurobiology,* 16:235-239

Schwitzgebel, E. (2008). The Unreliability of Naive Introspection, *Philosophical Review,* 117:245-273

Scott, W.A. (1955). Reliability of content analysis: the case of nominal scale coding, *Public Opinion Quarterly*, 19:321-325

Senju, A., Csibra, G., Johnson, M.H. (2008). Understanding the referential nature of looking: Infant preference for object-directed gaze, *Cognition,* 108:303-319

Simcock, G., DeLoache, J.S. (2006). Get the Picture?: The effect of iconicity on toddler's reenactment from picture books, *Developmental Psychology,* 42(6):1352-1357

Siple, P., Caccamise, F., & Brewer, L. (1982). Signs as pictures and signs as words: Effect of language knowledge on memory for new vocabulary, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 82:619–625

Skipper, J.I., Goldin-Meadow, S., Nusbaum, H.C., Small, S.L. (2007). Speech-associated gestures, Broca's area, and the human mirror system, *Brain and Language,* 101:260-277

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

Stjernfelt, F. (2007). *Diagrammatology: an investigation on the borderlines of phenomenology, ontology, and semiotics,* Springer, Nederlands

Striano, T., Rochat, P., Legerstee, M. (2003). The role of modeling and request type on symbolic comprehension of objects and gestures in young children, *Journal of Child Language,* 30:27-45

Suddendorf, T. (2003). Early Representational Insight: Twenty-four-month-olds can use a photo to find an object in the world, *Child Development,* vol. 74, No. 3, pp. 896-904

Talmy, L. (2007). Foreword, in Gonzalez-Marquez, M., Mittelberg, I., Coulson, S., & Spivey, M. (eds), *Methods in Cognitive Linguistics*, Amsterdam: John Benjamins

Thompson-Schill, S.L., D'Esposito, M., Aguirre, G.K., Farah, M.J., (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation, *PNAS,* 94:14792-14797

Toga, A. W., Mazziotta, J. C., (eds.) (2002). *Brain Mapping - The Methods - Second Edition,* Amsterdam, the Netherlands: Academic Press

Tolar, T.D., Lederberg, A.R., Gokhale, S., Tomasello, M. (2007). The development of the ability to recognize iconic signs, *The Journal of Deaf Studies and Deaf Education,* published online September 7th

Tomasello, M. (1999a). *The Cultural Origin of Human Cognition*, Harvard University Press, Cambridge Massachusetts, London

Tomasello, M. (1999b). The Cultural Ecology of Young Children's Interaction with Objects and Artifacts, In E. Winograd, R. Fivush, and W. Hirsh (eds.), *Ecological approaches to cognition: Essays in honor of Ulric Neisser,* MahWah, NJ:Erlbaum

Tomasello, M. (2000). First steps toward a usage-based theory of language acquisition, *Cognitive Linguistics,* 11(1/2):61-82

Tomasello, M. (2003). *Constructing a Language: A usage-based theory of language acquisition*, Harvard University Press, Cambridge Massachusetts, London

Tomasello, M. Carpenter, M., Call, J., Behne, T., Moll, H. (2005). Understanding and Sharing Intentions: The origin of cultural cognition, *Behavioral and Brain Sciences*, 28:675-735

Tomasello, M., Call, J., Gluckman, A. (1997). The comprehension of novel communicative signs by apes and human children, *Child Development,* 68, pp. 1067-1081

Tomasello, M., Striano, T., Rochat, P. (1999). Do children use objects as symbols?, *British Journal of Developmental Psychology,* 17, pp. 563-584

Troseth, G.L., Pickard, M.B., DeLoache, J.S. (2007). Young children's use of scale models: testing an alternative to representational insight, *Developmental Science,* 10(6):763-769

Tummers, J., Heylen, K., Geeraerts, D. (2005). Usage-based approaches in Cognitive Linguistics: a technical state of the art, *Corpus Linguistics and Linguistic Theory,* 1(2):225_261

Tylén, K. (2007). Agentivitet, in Bent Sørensen & Torkild L. Thellefsen's (eds.) *Livstegn: encyklopædi semiotic.dk*, Haase & Søn's Forlag, Kbh.

Uttal, W.R. (2001). *The New Phrenology: the limits of localizing cognitive processes in the brain,* The MIT Press, Cambridge Mass. London

Vaina, L.M., Solomon, J., Chowdhury, S., Sinha, P., Belliveau, J.W. (2001). Functional neuroanatomy of biological motion perception in humans, *PNAS*, 98(20):11656-11661

Vigneau, M., Beaucousin, V., Hervé, P.Y., Duffau, H., Crivello, F., Houdé, O., Mazoyer, B., Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing, *NeuroImage,* 30:1414-1432

Vygotsky, L.S. (1986). *Thought and Language*, MIT Press

Wagner, A.D., Paré-Blagoev, E.J., Clark, J., Poldrack, R.A., (2001). Recovering Meaning: Left frontal cortex guides controlled semantic retrieval, *Neuron,* 31:329-338

Wallentin, M., Østergaard, S., Lund, T.E., Østergaard L., Roepstorff, A. (2005). Concrete spatial language: See what I mean?, *Brain Language,* 92(3):221-233

Wallentin, M., Roepstorff, A., Glover, R., Burgess, N. (2006). Parallel memory systems for talking about location and age in precuneus, caudate and Broca's region, *NeuroImage*, 32(4):1850-1864

Walter, H., Adenzato, M. Ciaramidaro, A., Enrici, I., Pia, L., Bara, G.B., (2004). Understanding Intentions in Social Interaction: the role of the anterior paracingulate cortex, *Journal of Cognitive Neuroscience,* 16(10):1854-1863

Willems, R.M., Özyürek, A., Hagoort, P. (2007). When Language Meets Action: The Neural Integration of Gesture and Speech, *Cerebral Cortex*, 17(10):2322-2333

Winter, G. (2000). A comparative discussion of the notion of 'validity' in qualitative and quantitative research, *The Qualitative Report,* 4(3-4), available at http://www.nova.edu/ ssss/ QR/QR4-3/winter.html

Worsley, K.J., Friston, K.J. (1995). Analysis of fMRI time-series revisited – again, *NeuroImage,* 2(1):173-181

Younger, B.A., Johnson, K.E. (2004). Infants' comprehension of toy replicas as symbols for real things, *Cognitive Psychology,* 48(2):207-242

Zeki, S., Marini, L. (1998). Three cortical stages of colour processing in the human brain, *Brain,* 121:1669-1685

Østergaard, S. (2006). Symbolization in linguistic cognition, *commentary paper at the Cognition, Communication and Culture Workshop*, Dec. 1st , University of Aarhus

# 9 Appendix

## 9.1 English Summery

This PhD thesis addresses the way we sometimes come to perceive everyday objects like chairs, flowers and vacuum cleaners not only in their respective natural or instrumental function, but as meaningful communicative signals. This happens when such objects are manipulated in striking new ways that call for a special kind of meaning-exploring interpretation. For instance, chairs can be put out in the street to reserve a parking lot, flowers can be arranged in a bouquet and left on a doorstep to express a declaration of love, and old vacuum cleaners can be lined up in ranks on a lawn to make a conceptual piece of art. In these situations the everyday objects do a job usually done by words: they mediate meaning between communicating individuals. The current investigation addresses the cognitive and neurocognitive foundations for this understanding of objects as communicative signals using theoretical analysis and experimental behavioral and brain imaging techniques. The main findings are that object mediated communication can be considered 'a-kind-of-language' in the sense that it functions in similar and complementary ways to other strands of non-verbal and verbal communication. This is further supported by the finding that our perception of objects used as communicative signals elicit activations in brain areas commonly associated with verbal language and gesture such as Broca's area and the Visual Word Form Area. Another experimental finding is that people tend to orient to shared social norms for their perceptual and interpretational approaches to communicative object configurations. Even without any explicit instruction, people individually choose very similar interpretative strategies in their explorations of material signals. When people are asked how they understand an image depicting a neutral everyday scene with a lawn mower in a garden they tend to wander of in all kinds of diverse directions commenting on the season, distant childhood memories or the possible locality. In contrast, when they are shown an image depicting an iceberg at sea all covered in red paint (Marco Evaristti's art installation "The Ice Cube Project", 2004), their stories seem to align in significant and interesting ways. In this case, all informants are suddenly concerned with resemblance; what "it looks like" - a question that had not come up in the description of the lawn mower scene. Accordingly, when different informants in the experiment responded to other material signals they tended not only to systematically approach these as sources of intersubjective representational meaning, but, furthermore, to agree on very similar strategies for their interpretative meaning explorations. The relative agreement in informants' understanding of material signals, however, decreases when the signals employ a very unconventional and context dependent style of expression - a fact that is supported in the brain imaging study as well.

In all, the findings of this project suggest that human communication goes far beyond words, grammars and gestures. Even everyday material objects can in certain contexts be experienced as meaningful communicative symbols that function in ways comparable to and often even interchangeable with verbal language. In that respect it can be considered "a kind of language". Besides, on a more abstract plane the project points to the way that human cognition is not only extended into the material

environment, but how matter links people in setting the stage for an ‘extended social interaction’. This works by our general capacity to conform to shared norms and strategies in our perceptual exploration of the material world.

## 9.2 Dansk Resume

Denne ph.d.-afhandling beskæftiger sig med hvordan vi indimellem oplever hverdagsting som stole, blomster og støvsugere ikke bare i deres naturlige eller instrumentelle funktion, men som meningsbærende kommunikative symboler. Dette sker når disse objekter er manipuleret på nye og iøjnefaldende måder som synes at påkalde sig en særlig slags betydningsfortolkende opmærksomhed. For eksempel kan man opleve at stole er sat ud på gaden for at reservere en parkeringsplads, blomster kan være bundet i en buket og efterladt på et dørtrin for at udtrykke en kærlighedserklæring og gamle støvsugere kan være stillet op i geledder på en græsplæne som del at et konceptuelt kunstværk. I disse situationer er hverdagsobjekterne brugt i et ærinde hvor vi ellers ofte vil anvende ord: de formidler betydning mellem kommunikerende individer. Nærværende projekt bruger teoretisk analyse, adfærdseksperimenter og hjerneskanninger til at undersøge det kognitive og neurokognitive grundlag for vores forståelse af objekter som kommunikative signaler. Den væsentligste opdagelse er at objekt-medieret kommunikation kan betragtes som en ’slags sprog’, forstået derhen at den synes at fungere på måder der ligner og er komplementære til andre former for non-verbal og verbal sprogbrug. Denne tilgang bestyrkes af den opdagelse at vores perception af objekter der er anvendt som kommunikative signaler aktiverer hjerneområder der normalt forbindes med verbalt sprog og gestik så som Broca’s område og the Visual Word Form-område. En anden eksperimentel opdagelse er at folk generelt synes at orientere sig mod fælles sociale normer i deres perceptuelle og fortolkende tilgange til kommunikative objektkonfigurationer. Selv uden eksplicit instruktion synes folk hver for sig at vælge meget ens strategier i deres fortolkning af materielle signaler. Når de fortæller hvordan de forstår et billede af et neutral hverdagsscenarie med en plæneklipper i en have synes deres historier at fortabe sig i alle mulige retninger fra kommentarer til årstiden til fjerne barndomsminder eller havens mulige lokation. Når de derimod vises et billede af et isbjerg dækket af rød maling (Marco Evaristti’s kunst-installation “The Ice Cube Project”, 2004), synes deres historier på interessante måder at nærme sig hinanden. Pludselig er alle informanterne optaget af hvad isbjerget ’ligner’ – et spørgsmål der ikke var kendetegnende for deres beskrivelser af plæneklipperbilledet. Dette mønster går igen når eksperimentets informanter beskriver andre materielle signaler. De synes ikke bare systematisk at anskue disse som udtryk for intersubjektiv betydning, men yderligere at være meget enige om strategierne i deres betydningstolkning. Imidlertid synes denne relative enighed at aftage når signalerne betjener sig af meget ukonventionelle og kontekstafhængige udtryksformer – hvilket også underbygges af hjerneskanningsstudiet.

Samlet peger dette projekt således på at menneskelig kommunikation er mere end ord, grammatik og gestik. I bestemte kontekstuelle sammenhænge vil selv materielle hverdagsobjekter opleves som betydningsbærende kommunikative symboler, der i deres funktion kan sammenlignes med og ofte endog erstatte det verbale sprog. I dette lys kan de betragtes som ”en slags sprog”. På

et mere abstrakt plan peger projektet desuden på hvordan den menneskelige kognition ikke bare er bundet op med vores eksterne materielle omgivelser, men hvordan aspekter af vores materille verden synes at kæde folk sammen ved at muliggøre en 'udbygget social interaktion'. Denne er funderet i vores generelle orientering mod fælles normer og strategier i vores perceptuelle udforskning af den materielle verden.

## 9.3 Statement concerning the roles of coauthors in relation to *Article II*: Say It With Flowers! A fMRI study of object mediated communication

Kristian Tylén has the full initiative for the study. He has applied the ethics committee, collected and prepared the stimuli, has programmed the web-based questionnaire study and the stimulus presentation for the fMRI experiment. He has recruited the subjects, has conducted the brain scanning experiment, analyzed and interpreted the data and has written the article.

Mikkel Wallentin has assisted especially in the programming of stimuli presentation, technical operation of the MR scanner, and the data analysis.

Andreas Roepstorff has made all the expensive tools and technologies accessible to the investigations and has assisted in the more conceptual parts of the design and interpretations of the results.

Certification:

Kristian Tylén Mikkel Wallentin Andreas Roepstorff

## 9.4 Statement concerning the roles of coauthors in relation to *Article III*: Taking the Language Stance in a Material World: a comprehension study

Kristian Tylén has the full initiative for the study. He has collected and prepared the stimuli, has programmed the stimulus presentation and designed the instructions, has transcribed most of the audio material, has functioned as a coder and has made statistical analyzes and interpreted the data. Besides he has written the article.

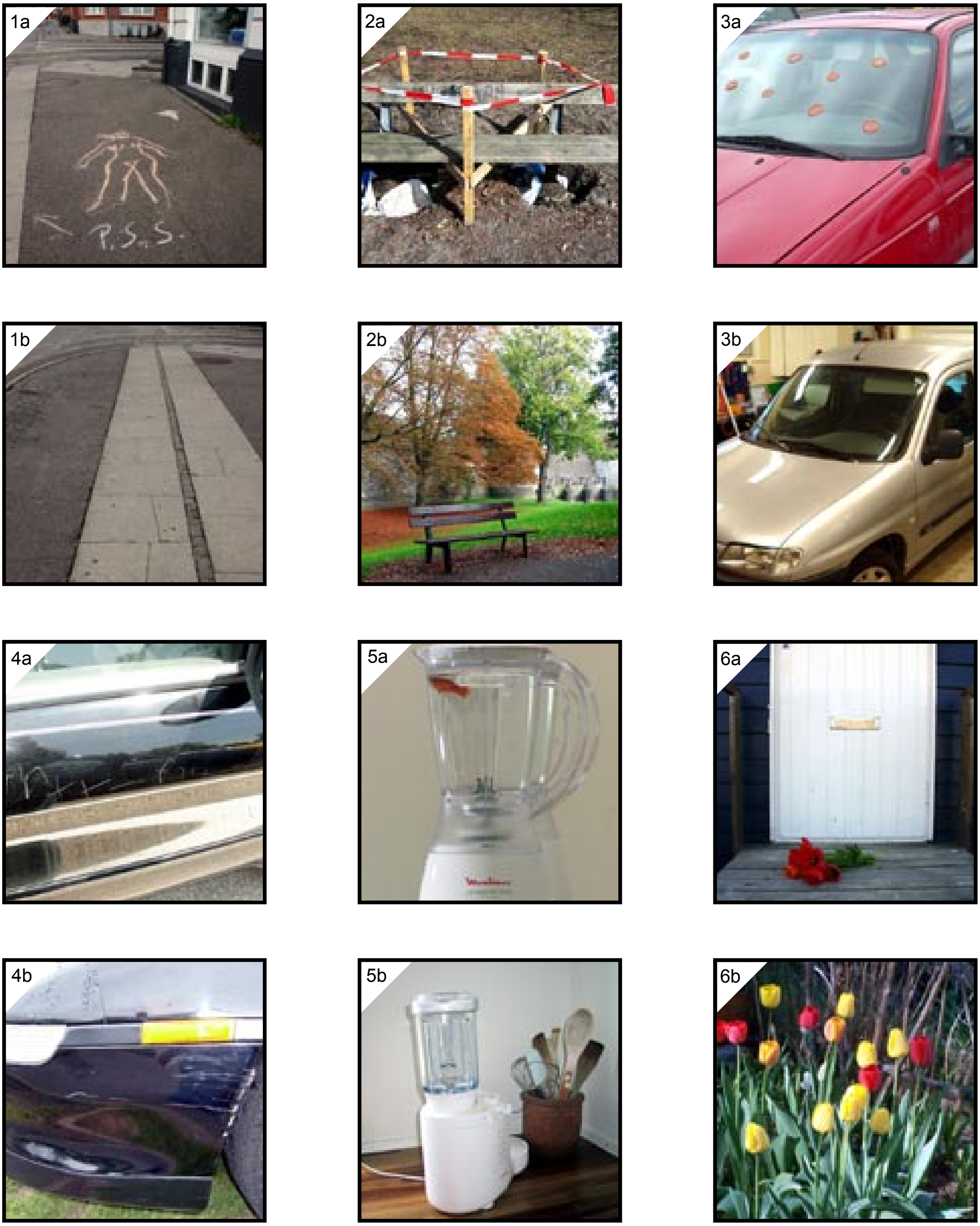
Johanne Stege Philipsen has been employed as a student helper on the experiment. She has assisted in the design of the experiment, has recruited subjects and has conducted the data collections. She has functioned as a coder and has transcribed a part of the audio material.

Ethan Weed has functioned as a coder and as a proofreader of the article.

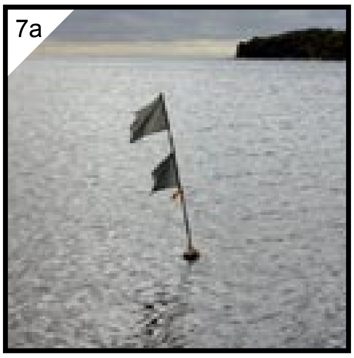
Certification:

Kristian Tylén Johanne Stege Philipsen Ethan Weed

## 9.5 Stimulus image corpus




Stimulus image pairs: all images marked with the letter ‘a’ constituted the test condition and all images marked with compatible numbers followed by ‘b’ constituted the corresponding control condition. For purposes of the current presentation layout the image formats have been altered in some cases resulting in a very tight crop. In the original display used in the experiments greater parts of the scenes were visible.


7a


8a


9a


7b


8b


9b


10a


11a


12a


10b


11b


12b


13a


14a


15a

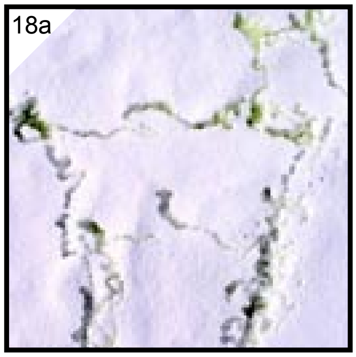

13b


14b


15b

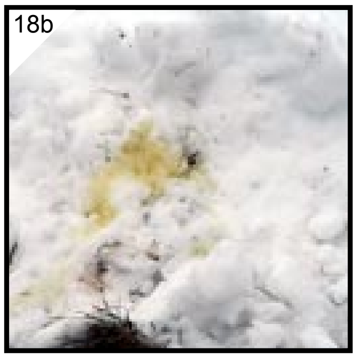

16a


17a


18a


16b

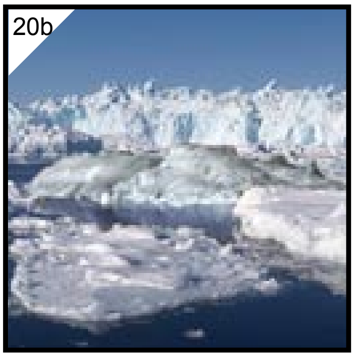

17b


18b


19a

20a

21a

19b

20b

21b

22a

23a

24a

22b

23b

24b


25a


26a


27a


25b


26b


27b

28a


29a


30a


28b


29b


30b


31a


32a


33a


31b


32b


33b


34a


35a


36a


34b


35b


36b


37a

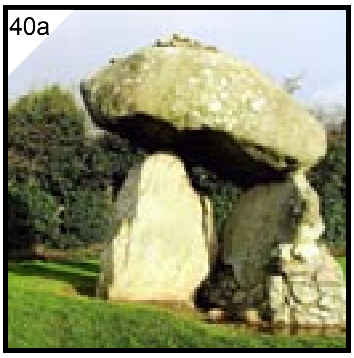

38a


39a


37b

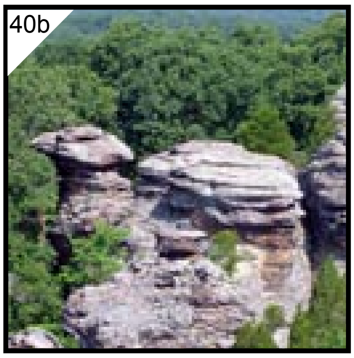

38b


39b


40a


41a


42a


40b


41b

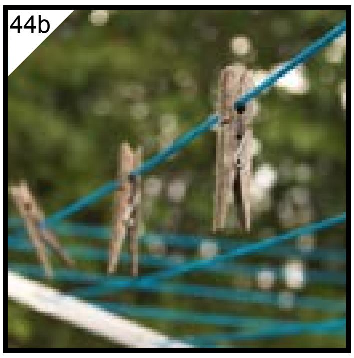

42b


43a


44a


45a


43b


44b


45b


46a


47a

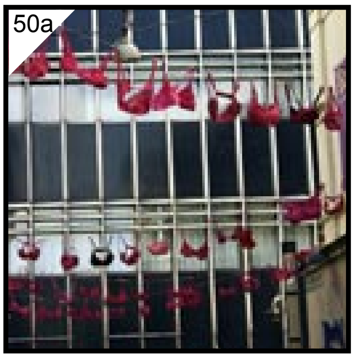

48a


46b


47b


48b

49a



50a

49b

50b