

REFERRING TO OBJECTS

Description:

Various notions of material objects play a key role in the study of different sectors of cognition. Infants appear to parse the world in object-like chunks; language makes predominant reference to objects; visual perception is attuned to objects; primates appear to be able to manipulate and modify objects. Yet the notions of objecthood used in the corresponding disciplines are not always aligned and are at times in need of conceptual clarification. The 2004 Summer School on "Reference to objects" tried to address these and related issues, under the direction of Paul Bloom, Daniel Povinelli, François Recanati, and Zenon Pylyshyn. The present workshop moves from papers developed at the School and furthers the discussion on the theme of the role of objects in cognition.

The workshop as part of the Ecole Thématique 2004 on "Reference to Objects" is sponsored by the CNRS, by Institut Nicod, by RESCIF, by the European Science Foundation as part of the ESF-OMLL project "Mindreading and the Emergence of Human Communication", by the Indo-French programme of the Maison des Sciences de l'Homme in Paris, and by the University of Eastern Piedmont at Vercelli

[French version]

Diverses notions d'objets matériels jouent un rôle clé dans l'étude de différents secteurs de la cognition. Les jeunes enfants semblent segmenter le monde en morceaux assimilables à des objets ; le langage donne une primauté à la référence aux objets ; la perception visuelle privilégie la sélection des objets ; les primates semblent être capables de manipuler et de modifier les objets. Toutefois, les notions d'objectivité utilisées dans les différentes disciplines correspondantes ne sont pas toujours alignées et nécessitent parfois des travaux de clarification conceptuelle. L'Ecole d'Été 2004 sur « La Référence aux Objets » a tenté de prendre en compte ces thèmes ainsi que quelques autres auxquels ils sont associés, sous la direction de Paul Bloom, Daniel Povinelli, François Recanati et Zenon Pylyshyn. Le présent workshop part d'articles développés dans le cadre de l'Ecole et approfondit la discussion sur le thème du rôle des objets dans la cognition.

Comme partie intégrante de l'Ecole Thématique 2004 sur la « La Référence aux Objets », le workshop est patronné par le CNRS, par l'Institut Nicod, par le RESCIF, par la Fondation Européenne de la Science au titre du projet "Mindreading and the Emergence of Human Communication", Origines de l'Homme, des Langages et du Langage (OMLL), par le programme Indo-Français de la Maison des Sciences de l'Homme à Paris, et par l'Université du Piedmont Est à Vercelli.

Moderators:

Roberto Casati (CNRS, Institut Jean-Nicod), **Gloria Origgi** (CNRS, Institut Jean-Nicod)

Guest Panel:

Andre Abath (University of Sheffield), **Erkki Ahlström** (Institut de sciences cognitives, Lyon 2), **Jeannes Aptekmans** (Université Paris VI, Laboratoire Lattice), **Sarah Bendaoud** (Institut Nicod), **Arnaud Blanchard** (Université d'Hertfordshire), **Paul Bloom** (Yale University), **Mario Borillo** (IRIT, CNRS), **Amita Chatterjee** (Jadavpur University, Calcutta), **Camelia-Mihaela Dascalus** (Université Paris 1), **Frédérique de Vignemont** (University College London, Institute of Cognitive Neuroscience), **Jérôme Dokic** (EHESS, Institut Nicod), **Julien Dutant** (Institut Nicod), **Karine Duvignau** (Laboratoire Lordat / Institut des Sciences du Cerveau, Toulouse), **Manuel Gatto** (University of Eastern Piedmont Amedeo Avogadro), **Valeria Giardino** (University of Rome, La Sapienza), **Valentina Gliozi**, **Emmanuelle Glon** (Université Paris 4), **Donald Glowinski** (Conservatoire de Musique, Paris), **Bastien Guerry** (Ecole Normale Supérieure), **Jose Luis Guijarro** (University of Cadiz), **Damian Justo** (EHESS, Institut Jean Nicod), **Basileios Kroustallis** (University of Crete), **Gilles Lafargue** (Université de Lille-3), **Saadi Lahlou** (Laboratoire de Design Cognitif EDF R&D, Laboratoire de Psychologie Sociale, EHESS), **Frédéric Landragin** (THALES,

Research and Technology, Palaiseau.), **Florent Levillain** (Université Paris 8), **Pascal Ludwig** (Association ArtCognition), **Erika Marchetto** (International School For Advanced Studies, Trieste), **Alda Mari** (CNRS, ENST), **Olivier Massin** (Laboratoire Philosophie, langage, et cognition, Grenoble), **Milena Nuti** (University College London), **Gloria Origi** (CNRS, Institut Jean-Nicod), **Elisabeth Pacherie** (CNRS, Institut Jean-Nicod), **Nausicaa Pouscoulous** (Institut Nicod), **Daniel Povinelli** (University of Louisiana at Lafayette), **Zenon Pylyshyn** (Rutgers University), **Susrut Ray** (University of Delhi), **François Recanati** (CNRS, Institut Nicod), **Giacomo Romano** (Technical University of Eindhoven), **Pierre Steiner** (CEPERC, Université de Provence), **John Stewart** (CNRS), **Valérie Tartas** (Université de Toulouse 2), **Giuliano Torrenço** (University of Eastern Piedmont, Vercelli), **Richard Zuber** (CNRS)

In partnership with :



Liste des Articles

- *Deferential Utterances*

- *Les Usages Déférentiels*

Philippe de Brabanter (Institut Nicod)

David Nicolas (CNRS, Institut Nicod)

Isidora Stojanovic (Granada University, Spain and Institut Nicod)

Neftalí Villanueva Fernández (Granada University, Spain and Institut Nicod)

- *Tracking Agency*

- *Traquer l'Agentivité*

Ángeles Eraña (U.N.A.M. (Mexico), Institut Jean Nicod)

Benjamin Sylvand (Université Paris 4 Sorbonne, Institut Jean Nicod)

Dario Taraborelli (CRESS - University of Surrey)

- *Tracking objects, Tracking agents*

- *Poursuivre les objets, Poursuivre les agents*

Nicolas Bullot (Institut Jean Nicod, University of British Columbia)

Patrick Rysiew (University of British Columbia)

Deferential Utterances

Philippe de Brabanter (Institut Nicod), David Nicolas (CNRS, Institut Nicod), Isidora Stojanovic (Granada University, Spain and Institut Nicod), Neftalí Villanueva Fernández (Granada University, Spain and Institut Nicod)

(Date of publication: 3 March 2005)

Abstract: The aim of this paper is to spell out the distinction between two linguistic phenomena, both of which may be called “deference”. We distinguish default deference (the ubiquitous fact that the semantic content of our utterances is determined with respect to some language parameter supplied by the context), from deliberate deference (the intentional, communicative act of using a given expression the way it is used in some contextually specified language). We further distinguish linguistic deference from related epistemic and conceptual phenomena.

Introduction

Our aim in this paper is to clarify the distinctions and the relationships among several phenomena, each of which has certain characteristics of what is generally called “deference”. We distinguish linguistic deference, which concerns the use of language and the meaning of the words we use, from epistemic deference, which concerns our reasons and evidence for making the claims we make. In our in-depth study of linguistic deference, we distinguish two subcategories: default deference (roughly, the ubiquitous fact, noted by externalists like Burge or Putnam, that the truth conditions of our utterances are determined with respect to the language parameter supplied by the context), and deliberate deference (roughly, the intentional, communicative act of using a given expression the way it is used in some contextually specified idiolect or dialect). We also discuss the phenomenon of imperfect mastery, often associated with deference, and which we show to be independent both of linguistic deference and of epistemic deference. If our analysis is correct, then some recent debates on deference (e.g. between Recanati and Woodfield) can be shown to result from a failure to appreciate all the distinctions that we draw here.

1.1 A quick look at the literature

The enthusiastic reception of Hilary Putnam’s dictum –“meanings are not in the head”– triggered a wave of interest in certain kinds of language use that intuitively support Putnam’s thesis. Externalists argued that the contents of our thoughts and utterances are not individuated by processes internal to the agent, but by the way the world happens to be. Tyler Burge (1979) offered examples to show that two distinct individuals could be, qualitatively, in the same internal state, and still have different beliefs. Particularly enlightening are examples involving *imperfect mastery*, that is, meaningful uses of concepts by speakers or thinkers who do not really master them. Such partial understanding, Burge argued, is omnipresent in the use of language. Brian Loar took the lessons from those cases of “falsity-due-to-misunderstanding” one step further, arguing that the satisfaction conditions of many concepts depend not on our personal epistemic abilities, but on language, “that social fact to which we defer” (1990, 118).

Two intuitions are at conflict here. On the one hand, we do not master all the concepts that we use. Still, we would not want to say that it is impossible for us to have or express thoughts about, say, rockets, aluminum or contracts just because our understanding of those concepts is incomplete. On the other hand, it is not clear how we manage to say anything true or false when we don’t master the concepts associated with the words we are using. Talking nonsense is not the same as saying something false.

The idea of *semantic deference* was introduced precisely in an attempt to reconcile these conflicting intuitions. Its rationale is that we can meaningfully use concepts that we do not master because we defer to the public language, whose rules are fixed by our linguistic community.

François Recanati has postulated the presence of a “deferential operator” in all cases of deference. The deferential operator is an unarticulated constituent that affects the truth conditions of an utterance.

It applies to any symbol (whether or not the symbol belongs to the language of the utterance) and delivers a well-formed expression, provided that there is tacit or explicit reference to some user of this symbol, the *deferee* (cf. Recanati 2000, 282, 272; 2002).

However, given that we do not master most of the concepts we use, does that mean that a deferential operator affects the semantics of most terms occurring in our utterances? Recanati thinks so: deference is present in every case of imperfect mastery, from the layman who uses technical terms in court to children learning their mother tongue.

This stance has raised criticism. Andrew Woodfield has argued that deference is not just something that happens to any language user. Deferring, he says, is an “intentional act, done by a person for a reason” (Woodfield 2000, 449-450). Typical cases of deferring involve an ordinary speaker and an expert. The *deferrer* and the expert identify their roles in communication, and they both recognize the existence of an external rule that governs their use of words. Under this description, it is difficult to see how one can accept the idea of ‘unconscious deference’ endorsed by Recanati (2000, 282) – the idea that speakers who do not master a concept unconsciously defer to the linguistic community for its meaning. Furthermore, Woodfield is reluctant to accept that language learning involves semantic deference. [2] Learning, he holds, is a matter of degree, while the deferential operator, which, according to Recanati, must accompany every use of an imperfectly mastered concept, cannot disappear gradually. [3]

We find, then, a cluster of problems around the notion of semantic deference. To address those problems, one had better start by identifying the different cases in which we would want to talk of deference. The tendency has been to treat all those cases in terms of *semantic* deference. This, we argue, is a mistake. Our aim in this paper is to clear up potential confusion by articulating the differences among several phenomena all of which seem to deserve the label ‘deference’.

1.2 The plan

In our taxonomic study, we draw a distinction between *linguistic* and *epistemic* deference. Within linguistic deference itself, we distinguish between *default* deference and *deliberate* deference. A first approximation to those distinctions may be provided using a single example, the arthritis example, a version of which was given by Burge (1979).

Consider a woman who, coming back from the doctor’s, tells her partner “I have arthritis.” Although it is not the first time she has heard of arthritis, she only has a vague idea of what arthritis is, insufficient for distinguishing arthritis from many other diseases. Thus, she may be unable to differentiate between arthritis, which is a condition of the joints, and myositis, which is a condition of the muscles, and she may even say such things as “I have arthritis in the thigh.” Even though the woman’s concept of arthritis is poor and, so to speak, indeterminate, i.e. insufficient for fixing the truth conditions of her utterance, this utterance has a determinate truth value, as Burge, Putnam, and externalists in general have successfully argued. This truth value is determined by appealing to the experts, and to the linguistic community more generally, regarding the question of what counts as “arthritis”.

This general and truly ubiquitous phenomenon corresponds more or less to what we call *default deference*. A speaker who defers by default most often does not have the intention to defer. [4] As a consequence, default deference usually goes unnoticed by speaker and hearer. This contrasts with what we call *deliberate deference*. A speaker who defers deliberately must intend to do so, and her intention must be recognized by her interlocutors.

In general, deliberate deference involves a language-shift. The speaker intends to use an expression in the way in which it is used in some dialect, sociolect or idiolect. She exploits various contextual features to enable her interlocutors to recognize her deferring intention and identify the intended deferee. Consider for instance two doctors who have a patient in common, and suppose that this patient believes that arthritis is an inflammatory condition of the muscles, and keeps saying to her doctors things like “I’ve been suffering so much from the arthritis in my left thigh.” Now suppose that one of the doctors has pain in his calves, and, making it clear to the other doctor that he is alluding to their common patient, he says “My calves really hurt. It must be arthritis.” Here, the doctor intends to use the word “arthritis” in the way the patient does, that is, for muscle inflammation. He makes use of

the context to elicit a language-shift to the patient's idiolect when it comes to interpreting the word "arthritis". Deliberate deferring, in sum, is an intentional act in which the dialect deferred to must be made salient by the speaker and identified by the interpreter.

An important notion that is related to linguistic deference but should not be confused with it is that of *epistemic* deference. Let us go back to the lady who, coming home from the doctor's, tells her partner "I have arthritis." To establish the truth value of her assertion, we need to determine with respect to which language her words must be interpreted. In English, the meaning of "arthritis" is established in connection with the common body of medical knowledge. So, even if the lady has picked up the word "arthritis" from her doctor, it is not quite right to say that she is deferring to *him* for her use of the word. For imagine that the doctor himself is mistaken on the question of what arthritis is, and believes that it is a condition of the muscles. Then, if the lady says "I have arthritis in the thigh," with no overt intention to defer precisely to her doctor, her utterance is false, given that arthritis *is* a disease of the joints. Now, even though the lady defers by default to the English linguistic community, not to the doctor, there is a sense in which she *does* defer to the doctor. But, rather than to his linguistic competence, she defers to his judgment (his diagnosis) that she has arthritis. This is what we call *epistemic deference*.

Below, we present more thoroughly the distinction between default and deliberate deference, and we argue that both phenomena are distinct from epistemic deference. We also discuss in more detail the phenomenon of *imperfect mastery*, already mentioned above. As we will show, partial understanding of a concept implies neither that a speaker using the word associated with the concept will intend to defer to others for the meaning of that word, nor that a speaker using the concept in making a claim will base this claim on someone else's judgment. Thus, even though they are often not differentiated, default deference, deliberate deference, epistemic deference, and imperfect mastery will be shown to be distinct phenomena.

2. Linguistic deference: default deference vs. deliberate deference

Default deference is involved in every communicative act. When interpreting and evaluating an utterance, we must take into account a language parameter (which is typically the language of a larger linguistic community, though it can also be a dialect, sociolect or idiolect), and this language parameter is contextually given a default value. To designate this default value, we will use the term "source language". Default deference takes place whether or not we seek to defer. *Deliberate deference*, on the other hand, is something done intentionally by the speaker. The speaker targets a particular value for the language parameter and exploits the context to help the interpreter identify this value. In this section we will illustrate this distinction with a number of examples and provide a more complex theoretical panorama. We will show that, even if a speaker typically defers by default to the linguistic community, she can also defer by default to a sociolect or idiolect. Similarly, in deliberate deference, even though a speaker typically intends to defer to a certain sociolect or idiolect, she can also deliberately defer to the linguistic community.

There are a few distinctions with which the distinction between default and deliberate deference might be confused, so let us forestall those possible confusions before proceeding. Deliberate deference is intentional and therefore conscious: a speaker who is deferring deliberately must be aware of what she is doing. But this does not make the default/deliberate distinction collapse into the self-conscious/unconscious distinction, [5] for in the case of default deference, too, the speaker may be perfectly aware of the fact that she is deferring *by default*. A second possible confusion consists in seeing default deference as *semantic* and deliberate deference as merely *pragmatic*: in the default case, the truth value of the utterance containing the deferential expression would depend on the source language provided by the context, whereas in the deliberate case, the speaker would merely convey her intention to use an expression in the way in which it is used by the deferee, without this impinging on the actual truth value of the utterance. On our view, however, both default and deliberate deference affect the truth values of utterances.

2.1 Default Deference

When the lady comes home from the doctor's and tells her partner "I have arthritis in the thigh", our intuitions are clear that the truth conditions of the utterance involve arthritis, not any other medical condition. As the lady cannot have arthritis in her thigh, she is saying something false. People often say false things when they use words they do not completely understand. This is the widespread phenomenon that Brian Loar called "falsity-due-to-misunderstanding" (Loar 1990).

Cases like these were used by externalists to show that, if the sense of a term were identified with the set of descriptions available to a competent language user, then this sense could not determine the term's semantic value. Falsity-due-to-misunderstanding is possible only because the terms we use in our utterances make a semantic contribution that is fixed by linguistic conventions that reflect the community's knowledge of the way the world is. In externalist frameworks, this idea is often grounded in a theory of the social division of linguistic labor. In every linguistic community, there are special groups of language users, the *experts*, who are entrusted with an important task: determining the semantic value of the terms of the language. Average members of the linguistic community defer to these experts whenever they have to determine the truth conditions of utterances like "I have arthritis in the thigh". The words used by the lady in the context described above acquire their semantic value through these experts. As arthritis is a disease of the joints and cannot affect the muscles, the proposition expressed by her utterance is false. Our intuitions about the truth conditions of this proposition are justified by her deferential use of the term "arthritis".

The process described above is what we call *default deference*. As suggested earlier, whenever an utterance is produced for the purpose of communication, the participants in the communicative exchange have to settle on a language with respect to which interpretation can be carried out, i.e. the source language. In cases of imperfect mastery, it is the *experts* who determine to which thing or event a given expression applies correctly.

If we resort to Recanati's deferential operator [6] to analyze what the arthritic lady says in the context above, and use it in the manner suggested by Recanati, we get the following representation:

(1) I have R_{doctor} (arthritis) in the thigh,

where " R_{doctor} (arthritis)" is the complex expression that results from the application of the deferential operator to the term "arthritis". The semantic value of this complex expression is arthritis (the actual disease). (1) is false because arthritis is a disease of the joints and there are no joints in the thigh. The subscript specifies who is being deferred to, in this case the doctor whom the lady visited. As we shall see, this kind of analysis does not entirely do justice to our intuitions about the truth conditions of deferential utterances.

2.1.1 Deference by default is not always deference to the "experts at hand"

Let us imagine that the doctor the lady consulted is not a real doctor, but some madman in a white coat who had just escaped from a psychiatric ward. This bogus doctor thinks that arthritis is nothing but a bad hangover. Moreover, he has got it into his mind to pay no attention to his lady patient's symptoms and to tell her that she has arthritis. The lady comes home and reports to her partner that she has arthritis. What she says could conceivably be represented with the help of the deferential operator:

(2) I have $R_{\text{bogus doctor}}$ (arthritis).

But are the truth conditions of the lady's utterance correctly captured by (2)? We do not think so. What the lady says is true if and only if she has in fact arthritis. The semantic contribution of the term "arthritis" to the proposition expressed by her utterance is not a bad hangover, as (2) states, but arthritis. Perhaps the lady had one too many glasses of vodka the night before and was suffering from a bad hangover on the day that she uttered (2), but that would not make the proposition expressed by her utterance any truer. Only arthritis can make that proposition true.

What is happening here? The lady is *deferring epistemically* to the bogus doctor, since she trusts his diagnosis without further questioning. But she is not deferring to the bogus doctor for the meaning of the term “arthritis”. Instead, she is deferring by default to the norms of the linguistic community. She is not deferring by default to the first expert at hand, but to whoever really knows the meaning of “arthritis”. Only this ideal expert can satisfy both the externalist claim that meanings are in the world and our intuitions about the truth conditions of her utterance in this context.

2.1.2 Deference by default is not always deference to the linguistic community

Pedro and María are watching the race walking competition in Beijing 2008. Pedro has not seen a race walking event in his entire life, but María, who knows a thing or two about the rules, has just spelt out to him the difference between walking and running in this Olympic sport. Some time after the start, the following dialogue takes place:

(3) (a) Pedro: “Hey, the second guy is walking so fast he’s gonna catch up with the one in the lead!”

(b) María: “Actually, he’s running... I’d say he’s gonna be disqualified”.

(c) Pedro: “Oh, yes, you’re right, he had both feet off the ground for a fraction of a second.”

If Pedro and María were not in a race walking context, their judgments would probably be different from those expressed in the above conversation. For instance, it is not unreasonable to assume that neither Pedro nor María would distinguish between the first and second contestant, so similar is the way they are moving. It is quite possible that they would judge both to be running rather than walking. Yet, these are realistic assumptions only if we take Pedro and María to be using the verbs “walk” and “run” in their ordinary sense. And our claim is precisely that they are not.

We think that, in the context at hand, the source language is not the common language but the particular sociolect of the race walking community. This community has its own experts, namely IAAF judges. These experts define walking in their rule 230 as “a progression of steps so taken that the walker makes contact with the ground, so that no visible (to the human eye) loss of contact occurs”. Rule 230 is the convention that determines the correct application of the term “walk” in this context. The judgments expressed in (3), and the distinctions underlying them, only make sense with respect to such conventions.

Various elements contribute to making the race walkers’ sociolect the source language here. Pedro and María are watching an Olympic race walking event, a sport whose rules they are now familiar with. They have been talking for some time about the technical interpretation of terms like “walking” and “running”. The meaningfulness of their conversation, including their initial disagreement, is further evidence that they are not speaking everyday English. We can therefore conclude that deference by default can select a source language whose conventions differ from those of the language community as a whole.

2.1.3 Deference by default to a particular ‘local dialect’

Imagine twin sisters, Natalya and Olga, who have been brought up in a very isolated area. Their parents use Standard English, except in one respect: they have a peculiar sense of humor, and thought it would be fun to always use “apple” for “pear” (and conversely) in their daughters’ presence. This is a reclusive family and, by the age of six, the sisters have hardly had any contact with anyone outside the family. On their first day at school, the two six-year-olds share the meal their parents have prepared for them, including some fruit. Looking enviously at her sister’s bigger pear, Natalya says to Olga:

(4) Hey, *that’s* a huge apple!

Any speaker of Standard English would say that the fruit is a pear and would therefore judge Natalya to have uttered a false proposition. But the thing is, in this case, that there is no speaker of Standard English involved in the situation. Both the speaker and her addressee are using the local dialect that

their parents have taught them. The whole of their linguistic community actually amounts to themselves and their parents (in their playful mood). The sisters are not even aware that there is a wider linguistic community whose norms may differ from what they have learnt from their parents. Thus, when Natalya or Olga use “apple” and “pear”, they defer by default to their parents’ invented dialect, not to the norms of a language community of which, strictly speaking, they are not part. Using the deferential operator, one could represent the proposition expressed by Natalya and understood by Olga as:

(4’) The object Natalya is pointing at is a huge R_{parents} (apple), [7]

which is the same as:

(4’’) The object Natalya is pointing at is a huge pear.

We are aware that this analysis is not self-evidently the right one. All the same, we believe it to be plausible: Natalya and Olga have always deferred by default to their parents, who were the purveyors of the linguistic norm in their environment. [8] As long as their linguistic community does not extend beyond their parents and each other, they could not defer to anyone else than their parents. This situation will change if their conversation is overheard by someone who knows for a fact that (according to the conventions of Standard English) the fruit they were talking about is a pear, not an apple, and who feels it is her duty to set the record straight for the kids. If the children accept that they are dealing with someone who is more trustworthy than their parents, they will probably change their minds about apples and pears. In our framework this change of mind would be explained as follows: the sisters will have realised that they belong to a wider language community, and that in the community there are experts who are more reliable (more knowledgeable) than their parents as to what this or that object or event ought to be called. If that is what takes place, then the source language of their utterances will shift from their parents’ local dialect to Standard English. Now, next time Natalya says to Olga that the fruit she is pointing at is an apple, her utterance will no longer be true if the fruit is indeed a pear. Thus, we see that, other things being equal, the truth conditions of an utterance of (4) are affected by a change in the source language.

2.1.4 Woodfield vs. Recanati on deference by default

Some of the issues we have raised here shed light on certain aspects of the debate between Andrew Woodfield and François Recanati. Woodfield’s opposition to an unrestricted application of the deferential operator is partly rooted in his conviction that in cases of partial understanding the expert is not the final source of normativity the deferrer is seeking for:

“Both parties [deferrer and expert] take for granted that there are norms which determine the proper meaning of the word, norms to which they both owe allegiance. D [the deferrer] defers to E [the expert] on a particular issue because D takes E to be a good guide, given the meaning that the word already has. D does not take E to be the *giver* of meaning. No fact about E *constitutes* the word’s meaning what it does. D knows that experts are fallible. D regards E’s judgement as good evidence that the word means such and such, but D does not suppose that E makes it the case that the word means such and such”. (Woodfield 2000, 450)

The example of the bogus doctor in 2.1.1 gives support to this general intuition. Deference does not always convert some speaker into a ‘giver’ of meaning. In our example, even if the lady heard the word “arthritis” for the first time from the bogus doctor, that would not mean that the bogus doctor could impose his peculiar use of “arthritis” on the meaning of the lady’s utterance. The intuitions about the truth conditions of the proposition expressed in that case are that the semantic contribution of the word “arthritis” is the disease arthritis, and not a bad hangover.

To make his point, Woodfield (Woodfield 2000, 448) resorts to the following example. Alf is a boy who has been told by his teacher that Cicero’s prose is full of synecdoches. The boy picked up the word “synecdoche” from his schoolteacher, unaware that the latter systematically called “synecdoches” what are actually metonymies. Alf meets L, an expert who knows what a synecdoche is, and the following conversation takes place:

(5) (a) Alf: "Cicero's prose is full of synecdoches."

(b) L: "No it is not. It's true that his prose is full of figures of speech. But very few of them are synecdoches."

(c) Alf: "I accept what you say. Cicero's prose is not full of synecdoches."

According to Recanati, Alf's utterances should be analyzed in the following way:

(5) (a') Cicero's prose is full of R_{teacher} (synecdoches)

(c') Cicero's prose is full of R_L (synecdoches)

On this view, what " R_{teacher} (synecdoches)" contributes to the proposition expressed by Alf in (5a') is the content the teacher attributes to the word "synecdoches", that is, metonymies. In (5c'), however, the semantic contribution of " R_L (synecdoches)" is synecdoches. It would seem then that Alf and L are 'talking at cross-purposes'. What Alf says in (5c) does not deny what he said in (5a). This, for Woodfield, is an unacceptable situation.

Within our framework, this problem does not arise. When Alf says in (5a) that Cicero's prose is full of "synecdoches", he is not using his schoolteacher as a *sense-giver*; he is deferring by default to the linguistic community through his teacher, whom he takes to be a reliable expert. As in the case of the bogus doctor, deference by default is not always deference to the first expert at hand. Alf is deferring to the linguistic community, and thus the semantic contribution of the word "synecdoches" as used in (5a) is synecdoches, not metonymies. When Alf is corrected by L, he learns something about Cicero's prose. His utterance in (5c) is the negation of (5a), because the semantic contribution of the term "synecdoches" is, in both cases, synecdoches. In (5a) he was deferring *epistemically* to his teacher, since he was taking for granted what the teacher had told him. Now that he has found a more reliable source of knowledge, he decides to defer epistemically to this new source, namely L, and consequently reconsiders his first statement that Cicero's prose is full of synecdoches. [9]

Our discussion of (5) shows that there is a difficulty with the application of the deferential operator to instances of default deference. The question is whether the deferential operator can be used in a proper representation of the truth conditions of utterances like these while, at the same time, preserving the explanatory power that the device has for cases of deliberate deference.

This is a problem that Woodfield has successfully detected. However, his own characterization of deference faces real difficulties. Though it is adequate for what we have called default deference to the linguistic community, it would have a hard time accounting for cases in which the speaker defers to a certain sociolect or local dialect, as illustrated by the examples of the race walkers and the misled sisters. Given that Woodfield recognizes only deference to the linguistic community, it is not clear how he could deal with these examples, where it is plausible to assume that the IAAF judges and Natalya and Olga's parents play a central role in fixing the meaning of certain terms. Furthermore, Woodfield would have some difficulty accounting for deliberate deference too, since typically, as we shall see, one defers deliberately to a certain idiolect or sociolect, rather than to the whole linguistic community.

2.2 Deliberate deference

Imagine that Tineke and Jan know about the bogus doctor who takes arthritis to be a bad hangover. And they like the story. Last night they partied especially hard and had a lot to drink. In the morning, they wake up and Tineke says to Jan:

(6) Jan, I have this bad case of arthritis. Would you close the curtains and hand me some aspirin?

Tineke's head is aching badly and she is feeling sick. But she is making playful use of the bogus doctor's misapplication of the term "arthritis" to say that she has a hangover. She knows she can rely on certain contextual features to make manifest the language with respect to which the term "arthritis" is to be interpreted. Tineke goes even further than that: she engineers a language-shift to a *target*

language different from the *source language* set by default, namely a shift from Standard English to the bogus doctor's idiolect. This, she can afford to do because she can rely on certain features of the context, notably the fact that Jan and herself had a laugh about the story of the bogus doctor and had a lot to drink the previous night. But Tineke could also have made her meaning clear by uttering "I have a bad arthritis", thus using, anomalously, the mass noun "arthritis" as if it were a countable noun. Contextual features of the above kind pretty much ensure that (6) is going to be understood by Jan as expressing the following proposition:

(6') Tineke has a bad case of $R_{\text{bogus doctor}}$ (arthritis).

Unlike what we observed in cases of default deference, the application of the deferential operator does not raise any issues here. Thanks to it, we can show how Tineke managed to express the proposition that she had a bad hangover, even though she uttered the word "arthritis". (6) is a paradigmatic case of deliberate deference. It presents all the characteristics of instances of 'self-conscious' linguistic deference mentioned in the literature: (i) the speaker *chooses* to defer for the interpretation of some of her words; (ii) she defers to *someone's idiolect*; (iii) deference takes the form of a *language-shift* that results from the *exploitation of certain contextual features*. It is to the needs of these examples that Recanati's deferential operator is tailored.

As we shall see, the above characteristics are not exhibited by all cases of deliberate deference. The following sections are devoted to a scrutiny of non-paradigmatic instances of deliberate deference. In sections 2.1.2 and 2.1.3, we established that the status of the source language (common language vs. sociolect and local dialect) was not constitutive of default deference: a speaker can defer by default not just to the whole linguistic community, but also to a sociolect (the race walking example) and even to a very local dialect (the misled sisters example). Similarly, we will see in the next section that characteristic (ii) does not apply systematically: deliberate deference does not have to involve a shift to someone's peculiar idiolect; in some cases, the target language is a sociolect or even a common language like English. This means that the distinction between default and deliberate deference cannot be a matter of the sort of language to which speaker and hearer defer. In section 2.2.2, we will show that characteristic (iii) is not a necessary condition for deliberate deference either. We therefore propose an account of deliberate deference that does not appeal to language-shifts in the strict sense.

2.2.1 Deliberate deference to the linguistic community

Suppose that an interdisciplinary wild bunch are working frantically on a taxonomy of linguistic deference. For several hours now they have been discussing similarities and differences between certain examples of default deference and borderline cases of deliberate deference. All the participants, A, B, C and D, agree on a common characterization for these terms and are now trying to tie up the remaining loose ends. The debate seems never-ending. At a critical moment, realising that lunchtime is almost over, the most obstinate, A, tells the others:

(7) All right, let's say that, *in deference to you*, I'll accept your argument.

We assume that the source language of their discussions is a local dialect that conforms to the definitions on which they had previously agreed. But, if A's utterance is understood by B, C and D, they will not think that A is deferring linguistically or epistemically to any of them, but rather that she is accepting their argument *out of respect for them*. However, *respect* is not what the word "deference" would mean in the source language of this context: it is a meaning it has in a different language, namely Standard English.

In (7), the speaker again exploits contextual features in order to make it manifest that she means to shift out of the source language (the deferentialists' dialect) and into a target language that is Standard English. The co-text plays a central role: in its source-language technical sense, the noun "deference" does not collocate with "in ____ to you". This alone should be enough to induce recognition of the shift. All in all, this example shows that deliberate deference does not necessarily rest on language-shifts to an idiolect or a sociolect. One can defer deliberately to the linguistic community.

2.2.1.1 Deference, polysemy and Humpty-Dumpty

This, at any rate, is the conclusion if our analysis is the right one. Yet, we are aware of another possible account for (7): it could be said that, in uttering (7), A simply exploits the polysemy of the word “deference”. In other words, where our analysis posits a language-shift from a technical dialect into Standard English, others might see no shift at all. Their argument, then, would be that the technical dialect of the interdisciplinary team is nothing but an *extension* of the standard language. In this extension, the ordinary senses of “deference” (*respect* and *compliance with another’s judgment*) are inhibited, while a technical sense is highly activated. On this view, all A does in uttering (7) is reactivate an ordinary sense of “deference”.

We have some sympathy for this analysis. However, we think that its implications are not so straightforward as they look. First, notice that if polysemy is involved in (7), then it is polysemy of a special kind, for the sense that “deference” has in the source language (the technical local dialect) is not (yet) one that is recorded in the lexicon of the target language (Standard English). The problem here is that the deferentialists’ work results in “deference” acquiring a new meaning. Neologisms and meaning-creations always originate in the margins of the common language. Sometimes they catch on, sometimes they don’t. But, if they do, it is always because some aspects of the language spoken by a small group become incorporated into the common core. Until that happens, those aspects cannot be said to *belong* to the common language. Actually, as some lexicographers have shown (e.g. Rey-Debove 1978, 283-286), new words and new lexical meanings, when they occur in utterances of the common language, are often set off by quote marks or special prosody, indicating that they still feel like words in another language. Our analysis in terms of deliberate deference provides an explanation for the diachronic process by which lexical creations may become part of the common core. In the case of new meanings, this will lead to increased polysemy, but *only after the process of extension of the common language has been completed*.

Our analysis is less susceptible to accusations of Humpty-Dumptyism than an account strictly in terms of polysemy. On our view of deliberate deference, a speaker does not decree that this or that expression is to be ascribed a new meaning. Rather, she uses expressions which have already acquired a meaning in a given language (be that a common language, sociolect or idiolect). The only decision the speaker makes is to exploit contextual features in order to induce the appropriate language-shift. That is not Humpty-Dumptyism. By contrast, those who reject the deferential account and argue that examples like (7) exhibit plain polysemy *can* be suspected of Humpty-Dumptyism. On their view, a single language underlies the interpretation of (7), namely an extension of the common language. This extension includes a new meaning of an already existing term. It seems then that, merely as a result of their theoretical debates, the deferentialists have succeeded in creating a new meaning for “deference”. This means that they have acted pretty much like Humpty-Dumpty in the Alice story.

2.2.1.2 Two more examples

We have shown how our account could accommodate the intuition that polysemy is somehow involved in (7). However, we believe that there are examples similar to (7) for which a polysemy-based account is not even a likely contender. We present two such cases below. The first illustrates deliberate deference to the linguistic community, while the second shows that a speaker can deliberately defer to *another* common language.

Imagine a guru who, though using the spelling, the grammar and large chunks of the English lexicon, nevertheless chooses to redefine a whole class of key terms (say, “life”, “love”, “devotion”, etc.) in such a way that the ordinary senses of these terms no longer have currency in the language of the guru’s community. One can hardly say here that the guru’s language is a mere extension of Standard English. Now imagine that the guru is preaching to his flock and that his sermon is broadcast on his own satellite TV channel. For a while, he talks directly to his live audience. At one point, however, he looks straight at the camera and, addressing ‘the rest of the world’, says things like:

(8) You may experience ‘love’ and ‘devotion’ in your hearts, but these are just debased forms of true love and true devotion.

Our suggestion is that, at least for the interpretation of “‘love’” and “‘devotion’”, the guru shifts into Standard English. This, we indicate by means of scare quotes, to reflect the fact that the language-shift engineered by the guru is a deliberate one.

Examples (7) and (8) belong with a class of utterances which display an intrasentential shift into another common language, as in:

(9) Barthes described the book as “un choc historique” and “un repère nouveau et un départ pour l’écriture”. (*Times Literary Supplement*, 03/05/02 : 9)

The shift here may be for the sake of accuracy in quoting, or for local color, or meant as a display of one’s linguistic skills. However that may be, this is a deliberate language-shift into a common language: French. Although we acknowledge that (9) is different from the previous two examples, it provides further evidence suggesting that deliberate deference is not systematically to idiolects and sociolects. [10]

2.2.2 Deliberate deference without language-shift?

We now consider a class of utterances that seem to fall under the same category as the previous ones. Yet, they turn out to lack one important property exhibited by the various examples of deliberate deference studied so far.

Let us assume that Kate, who has no training in law, is attending a trial. Both the judge and the defendant’s counsel use terms of art with which she is not familiar. For instance, it is not obvious to her whether the defendant committed a *felony*, an *offence* or a *misdemeanor*.

During a break, while talking about the proceedings with other members of the audience, Kate is trying to determine the sort of crime that the defendant is guilty of. In so doing she says things like:

(10) I don’t think what he did was a felony. I’d say it was a misdemeanor.

Since she realises that her understanding of these terms is at best sketchy, she often supplements her utterances with a metalinguistic comment, or articulates them with a special intonation pattern, of the sort that can be rendered by means of scare quotes:

(10’) I don’t think what he did was a felony, as the judge put it. I’d say it was a misdemeanor, if I understand the lawyer’s distinction.

(10’’) I don’t think what he did was a ‘felony’. I’d say it was a ‘misdemeanor’.

These comments and extra markers indicate that we are not dealing with instances of default deference. So, are we dealing with deliberate deference, and, if so, to whom? In an externalist framework such as ours, it is generally accepted that the meaning of legal terms is fixed by members of the legal profession for the whole of the linguistic community. There should therefore be no difference between the meaning that the judge ascribes to “felony” and “misdemeanor” and the meaning that these terms have in the lexicon of English. And if there *were* a difference, Kate, as a non-expert member of the audience, would probably choose to trust the norms of the linguistic community (as fixed by the body of experts alluded to above). This suggests that, when Kate utters (10), (10’) or (10’’), she is not (just) deferring to this judge or lawyer, or even to the legal profession, but to the norms of the linguistic community.

It is tempting to conclude that examples (10)-(10’’) are a further illustration of deliberate deference to the linguistic community. But, as hinted above, these examples lack one significant feature exhibited by the other cases: they involve no language-shift. In (10)-(10’’), the language with respect to which terms such as “felony” and “misdemeanor” are interpreted is none other than the source language set by default. This entails that, unlike what can be observed in (6) and (7), Kate’s deferring turns out to have no impact on the truth conditions of her utterances. Still, there is a major difference between (10)-(10’’) and genuine cases of default deference. Kate resorts to metalinguistic comments or special

intonation patterns in order to make the language parameter of the context salient. This does not happen in cases of default deference, where the speaker typically has no communicative intention to bring the language of interpretation into the foreground.

Faced with these facts, we believe that the right theoretical choice consists in maintaining that (10)-(10'') involve deliberate deference. Accordingly, we must relax criterion (iii) of paradigmatic instances like (6), so as not to require the presence of a language-shift in the strict sense. [11] We therefore propose the following definition:

S performs an act of deliberate linguistic deference if and only if:

(a) S produces an utterance u ;

(b) S exploits certain contextual features in order to make salient the linguistic parameter L for the interpretation of u or some segment of u ;

(c) S wants her exploitation of contextual resources to be recognized as part of her communicative intentions by the audience.

Although our definition does not include any requirement for a language-shift, cases of deliberate deference can still be represented by means of the deferential operator. For instance, what happens in (10) can be captured by the following formula:

(10₁) I don't think what he did was a $R_{\text{StandardEnglish}}$ (felony). I'd say it was a $R_{\text{StandardEnglish}}$ (misdemeanor).

The deferential operator indicates that the expressions "felony" and "misdemeanor" are to be interpreted with respect to Standard English. In cases of language-shifts, the only difference is that the value of "x" in " $R_x(s)$ " is distinct from the source language. Deliberate deference with a language-shift is nothing more than an important sub-category of deliberate deference.

3. Non-linguistic deference and other related phenomena

In this section, we study the notions of *epistemic deference* and *imperfect mastery*. Epistemic deference should be carefully distinguished from linguistic deference, and our first comments will be focused on justifying this distinction. We then show that the notions of epistemic *deference* and epistemic *evidence*, though related, must be kept well apart. Finally, we discuss the phenomenon of imperfect mastery and its relationship to linguistic and epistemic deference. [12]

3.1 Epistemic deference

Deference is an issue of interest not only to linguists and philosophers of language, but also to epistemologists and philosophers of science. It is generally admitted that a lot of the knowledge that we possess is acquired deferentially, by testimony. But deference does not only affect the things we *know*: it also affects our *beliefs*, beliefs we are none the less ready to act upon. We receive information from many different sources, and we make choices as to which information to accept and which to reject. Imagine a lady with a rare disease who wants to gather different opinions about her illness before undertaking a medical treatment. Every doctor she meets gives her an opinion, based on evidence and other considerations. If the diagnoses differ, she will have to decide which doctor to trust above all others. But, underlying our beliefs and actions are not just other people's judgments on issues for which there is a fact of the matter. We also defer to others on issues that are largely a matter of personal opinion. Suppose that Takeshi has been told by one friend that Sakura is *the best* sushi-bar in town, and by another that Mikado is the best. If he wants to take his fiancé(e) for a date, Takeshi will have to decide which friend to trust, whose judgments of taste are more reliable.

In what follows, we will focus on those instances of epistemic deference that underlie assertions, because it is in these cases that epistemic deference may be most easily confused with linguistic deference. We will say that a person who makes an assertion is deferring epistemically when she

bases her claim, partly or completely, on someone else's opinion. Typically, a person who asserts that she has arthritis is epistemically deferring to the doctor on whose diagnosis she relies. We argue below that this phenomenon is distinct from linguistic deference. Furthermore, it cannot be reduced to the notion of epistemic evidence, even if a certain correlation exists.

3.1.1 Epistemic and linguistic deference

Whereas linguistic deference is involved in fixing the meaning of a term, epistemic deference occurs when a person defers to someone else concerning a particular judgment. Whether a speaker is deferring epistemically or not is independent of whether she is deferring deliberately or by default for the use of the words occurring in her utterance.

To begin with, it is easy to realize that default linguistic deference must be independent of epistemic deference. As we have argued at length, default linguistic deference is a ubiquitous phenomenon. Epistemic deference, on the other hand, occurs when we rest a claim upon other people's opinions. It is not surprising, then, that default linguistic deference can, but need not, co-occur with epistemic deference. For example, suppose that Tim goes to see a doctor who, having examined him, tells him: "You have myositis." Tim does not know what myositis is. He only understands that it is related to the pain he is feeling in his calves. Back at home, he tells his mother:

(11) I have myositis. It is nothing serious. I should just rest for a while.

In saying (11), Tim is deferring by default to the linguistic community concerning the meaning of the term "myositis", and, at the same time, he is deferring epistemically to the doctor, the truth of whose diagnosis he takes for granted. But when Tim tells his mother: "My calves hurt badly", he is certainly not deferring epistemically to the doctor, for he is best placed to judge whether a part of his own body hurts or not. However, Tim will still be deferring by default to the linguistic community concerning the meanings of the words that he is using to report the pain in his calves, such as "calves", "hurt", etc.

Somewhat more interesting are the connections between epistemic deference and deliberate linguistic deference. Let us approach these through various examples. We have already seen that epistemic deference occurs independently of default linguistic deference. When Tim sincerely asserts that he has myositis, he is deferring epistemically to the doctor, but from a semantic point of view, he is deferring by default to the entire linguistic community. This is even more obvious in his assertion that "it is nothing serious", which is again epistemically based on the doctor's judgment, but involves only terms that Tim, a native English speaker, fully masters.

Conversely, deliberate linguistic deference occurs independently of epistemic deference. This is clear from our analysis of example (5) above:

(5) Jan, I have this bad case of arthritis. Would you close the curtains and hand me some aspirin?

Though Tineke is borrowing the bogus doctor's deviant definition of "arthritis", she is not deferring to any medical diagnosis made by that doctor, or even to any opinion that he might have regarding her condition.

The mutual independence of linguistic and epistemic deference can be given a more complex, and more subtle, illustration. Think again of the doctor who intentionally uses the word "arthritis" with the deviant meaning that his patient attributes to it. Suppose that this doctor asked for a specialist's opinion regarding the symptoms in his calves. Diagnosed with inflammation, which is precisely the condition for which the misguided patient uses the word "arthritis", he tells the colleague with whom he has that patient in common: "My calves hurt. It is arthritis." Although he does not defer epistemically for the claim that his calves hurt, the doctor defers epistemically to the specialist for the claim that his condition is "arthritis," i.e. inflammation of his calf muscles. At the same time, he is deferring deliberately to their patient's idiolect, for the semantic question of what counts as "arthritis." In this case, deliberate linguistic deference occurs together with epistemic deference, but with distinct deferees.

In sum, linguistic deference and epistemic deference are distinct and mutually independent phenomena, though they can combine in various ways, as has been amply illustrated in previous sections.

3.1.2 Epistemic deference and epistemic evidence

The examples given above might suggest that epistemic deference occurs as a direct result of there being insufficient epistemic evidence for making a claim. And it is true that epistemic deference is quite often a matter of the amount of epistemic evidence that one has for making a certain statement. Thus, if you have no independent evidence to assert p , but have been told by someone you trust that p , you are likely to assert p , simply because you rely on that person's judgment. We say in such a case that you are epistemically deferring to that person. On the other hand, when you have the best possible epistemic grounds of your own for asserting p , then in asserting p , you will probably not want to rely on someone else's judgment.

However, lack or poverty of epistemic grounds are neither a sufficient nor a necessary condition for epistemic deference. Someone who has all the evidence that can be had may still choose to defer epistemically to someone else. Thus consider a doctor who happens to be the greatest expert on arthritis, but lacks self-confidence. It is plausible to say that, when he tells a lady patient "You have arthritis," he is epistemically deferring to his colleagues on the issue of whether that woman's condition is indeed arthritis, even though he has enough of his own evidence for this claim. To indicate that he is doing so, he might say "We believe that what you have is arthritis." Or, imagine that Naïma is a first-rate scientist whose research shows how to achieve cold fusion, but is very shy and insecure. She is doing tests in her lab when a senior researcher, whom she deeply respects and admires, tells her: "You are wasting your time. Believe me, cold fusion is something impossible!" Out of sheer insecurity, she decides to defer to his opinion, even though it directly contradicts a claim that she has excellent evidence for, namely, that cold fusion is possible.

Conversely, there are situations in which people may form and express a firm judgment even on an issue for which they have no good epistemic grounds. Consider a woman whose partner tells her "You have arthritis" just out of some inner conviction. He is not, then, deferring epistemically to anyone. People do make claims for which they have no good evidence, and which do not reflect other people's opinions. Such claims –people's best guesses, as we might put it– exemplify the case where one lacks epistemic evidence, and yet abstains from deferring epistemically.

3.2 Imperfect mastery

Many philosophers hold that there are concepts, and that concepts are very much like mental *files* in which information gets stored. Consider the concept that Carmelia has of a certain particular, say François Recanati. Her concept contains three main types of information: *perceptual* information, e.g. that the particular concerned by this concept is *that guy*, whom she sees talking right there in front of her, *descriptive* information, e.g. that he is the author of *Literal Meaning*, and *metalinguistic* information, e.g. that he is called "François Recanati". Our concepts of universals, too, mostly combine those three types of information. But in many cases, the concepts that we associate with words that we use, like "arthritis", "elm" or "hydrogen", are fairly poor, and the information they contain does not enable us to decide on any given occasion whether the word correctly applies to something we are presented with, or to draw certain inferences that someone more knowledgeable could draw. If the concept that a person associates with a term is poor or, at any rate, not as rich as the concept that experts associate with it, we talk of *imperfect mastery*. Note that mastery is very much a matter of degree, and that it is not obvious that anyone ever achieves *perfect* mastery. But to bring the issue home, one might want to know how the phenomenon of imperfect mastery relates to linguistic and to epistemic deference, and ask questions like the following. What information must be present in a concept for one to be able to defer, deliberately or by default, using the associated term? Conversely, could the presence of some information make deferring impossible? Does epistemic deference arise whenever we make assertions using concepts that we do not perfectly master? And will the wealth of information in our concepts prevent us from deferring epistemically?

By way of giving a single answer to these questions, we hold that imperfect mastery is a phenomenon that must be kept separate from linguistic deference and from epistemic deference. In other words,

whatever a person's mastery of the concept associated with some term, whatever the amount and quality of the information contained in the mental file, the following options all remain viable: the person will defer by default when using the term, or she will defer deliberately to some contextually salient dialect, be it or not the dialect from which she picked up the term. Likewise, the person may or may not defer epistemically for assertions that she makes using the term.

3.2.1 Imperfect mastery and linguistic deference

Consider a medical expert whose concept of "arthritis" is as rich and determinate as can be. Does such a person defer linguistically when she uses the term "arthritis"? The intuition is that the meaning of such terms is determined precisely by such experts. So if we say that the expert defers in turn, who could she possibly defer to?

Though there may be a grain of truth in this intuition, our account of default deference does not require the speaker to have the intention to defer, or to know which source language is contextually selected. This means that even our medical expert defers by default to the linguistic community when using "arthritis". It just happens that she is among the experts who ultimately determine the meaning of the term.

We have just shown that "perfect" mastery is compatible with deference by default. Is it also compatible with deliberate deference? Again, the answer is "Yes". To see this, just recall our example of the doctor who suffered from an inflamed calf muscle. This doctor could be assumed to know as much about arthritis as possible. Yet, this did not prevent him to wittily exploit the ignorance of a patient and tell his colleague "It must be arthritis."

In a similar way, imperfect mastery allows both for default and deliberate linguistic deference. A woman who knows virtually nothing about arthritis, except that there is something called "arthritis", can use this word to say true or false things, whether or not she has any intention to defer linguistically at all. Or she may defer deliberately, indicating the source from which she got the word and making it clear that she intends to apply the word to whatever it is that her source applies it to, even though she might have no idea what that is.

3.2.2 Imperfect mastery and epistemic deference

Our level of mastery of a given concept can neither force us to defer epistemically nor prevent us from doing so. Someone who has perfect mastery can still choose to defer epistemically, like the shy scientist who endorses her senior colleague's opinion that cold fusion is impossible. Most often, though, people with excellent mastery of a concept make assertions without deferring to other agents, provided that they have strong enough epistemic grounds for their assertions. It is true, too, that, if we know hardly anything about myositis, we are unlikely to go around making unwarranted claims about it. Thus, if we report that Tim has myositis, we will typically do so because someone whose judgment we trust told us that Tim had myositis, or because we read it in Tim's medical file. In those cases, we defer epistemically. But others with the same level of mastery may make the very same claims without deferring epistemically, e.g. out of some inner conviction, however odd this may seem. In any case, even cautious speakers aware of their poor mastery of a given concept will be ready to make certain assertions about myositis without deferring epistemically. For example, they will confidently assert that myositis is a condition called "myositis", or that they would not like to be diagnosed with myositis, even though they have no idea what that is.

In sum, even though there is probably a correlation between an agent's imperfect mastery of a certain concept and her being *inclined* to defer epistemically, epistemic deference and imperfect mastery are distinct phenomena, irreducible to one another.

4. Conclusion

In this paper, we have argued for the mutual independence of three related phenomena, namely *linguistic deference*, *epistemic deference* and *imperfect mastery*. One of our initial questions has been what kind of framework could accommodate instances of falsity-due-to-misunderstanding and cases in which a speaker overtly chooses to use an expression the way someone else uses it. Our answer has been that, in both types of cases, an expression or segment of discourse is used deferentially. This is what grounds the category of *linguistic deference*. Within this category, we have distinguished two varieties, *default deference* and *deliberate deference*, which, together, cover a significant proportion of the examples addressed in the literature.

Default deference has been shown not to be restricted to those cases where the speaker defers to the linguistic community as a whole. We have supplied examples of default deference to a sociolect and even to a local dialect. As for deliberate deference, which is usually understood as deference to an idiolect or local dialect, we have given evidence that *it ain't necessarily so*. Our examples suggest that speakers sometimes defer deliberately to the norms of the common language. The account we provide differs in one further respect from the picture that emerges from what little literature has been written on the subject. Deliberate deference does not always involve a genuine language-shift.

Concerning the related notions of epistemic deference and imperfect mastery, we have contended that they are distinct from each other and orthogonal to linguistic deference. Someone defers epistemically when they base a claim on someone else's judgment, but this does not entail that they are deferring for the meaning of the words they are using. With respect to imperfect mastery, we have shown that the partial understanding of a concept does not constrain a speaker to defer either linguistically or epistemically.

Providing a taxonomy of the various cases of deference discussed in the literature is like drawing a map of the tip of an iceberg. Any stable theory of deference requires awareness of the philosophical and linguistic issues of titanic proportions that underlie it. The submerged part of the iceberg comprises issues such as quotation, simulation, echoic uses, irony, polysemy, knowledge acquisition, justification, cognitive architecture and concepts.

[1] The order of the names is not intended to reflect differences in the work done by each author.

[2] Willing to take up these challenges, though, Recanati tried to provide a model of concept learning that would incorporate both the idea of gradual improvement and an explanation of deference in terms of the deferential operator (2000, 282-285).

[3] Note that Burge explicitly excluded learning from his explananda (1979, 90). A child who memorizes " $e=mc^2$ ", he held, does not have beliefs about the theory of relativity.

[4] Of course, whenever we engage in communication, we *implicitly* intend to conform to the rules of language use.

[5] We are thinking of Recanati (2000, 281ff) here, even though Recanati himself opposes *self-conscious* deference to *imperfect mastery*, and does not speak of *unconscious* deference.

[6] Here is how Recanati initially characterizes the deferential operator: "The deferential operator is translinguistic. It belongs to a certain language, say L , but applies to any symbol s whether or not s also belongs to L . The result of applying the deferential operator to a symbol s , whether or not that symbol exists in L , is a well-formed expression of L , which I write as ' $R_x[s]$ ', where x is a user of the symbol s [...]" (Recanati 2000, 272). Let us note that we will be using the variable x either for users of a language or for the language itself.

[7] We are not, at this stage, claiming that the deferential operator is suitable for analyzing instances of default deference. Here, it is simply a convenient means for representing the manner in which "apple" is to be interpreted. See section 2.1.4 for a discussion of some problems raised by the application of the deferential operator in default cases.

[8] Had the conventions of their parents' language been the same as those of the common language, Natalya and Olga would eo ipso have deferred by default to the linguistic community as a whole. But the point here is precisely that (a few of) the conventions set by the parents clash with those of the whole linguistic community.

[9] We are ignoring the possibility that Alf might be deferring deliberately to his teacher, in which case the proposition expressed by his first utterance would contain metonymies, not synecdoches. This is not the way the example was originally framed by Woodfield.

[10] There are plenty of examples like (9), and they are usually taken to be related to quotation. (See issue 17 of *The Belgian Journal of Linguistics* for various discussions.)

[11] Another option would be to leave the criterion for deliberate deference unaltered. As a result, examples like (10)-(10'') would come under a third category of linguistic deference, intermediate between default and deliberate. In our view, however, these examples have much more in common with deliberate deference; hence we shall not pursue this line of reasoning further.

[12] Let it be clear from the outset that our goal in this section is not so much to make a new contribution to the existing literature on epistemic and cognitive issues related to deference, as to clearly distinguish those issues from the ones that arise in relation to *linguistic* deference.

French version

- Les Usages Différentiels

(Traduction de l'original en anglais de David Nicolas)

Résumé : Nous tentons de clarifier la distinction entre deux phénomènes linguistiques qui peuvent l'un et l'autre porter le nom de "déférence". Nous distinguons la déférence par défaut de la déférence délibérée. Par ailleurs, nous distinguons également la déférence linguistique de phénomènes épistémiques et conceptuels qui s'en rapprochent.

Tracking Agency

Ángeles Eraña (U.N.A.M. (Mexico), Institut Jean Nicod),

Benjamin Sylvand (Université Paris 4 Sorbonne, Institut Jean Nicod),

Dario Taraborelli (CRESS - University of Surrey)

(Date of publication: 3 March 2005)

Abstract: The aim of this paper is to propose an extension of the object file notion to the study of nonconceptual individuation of agents. Robust evidence in psychology supports the hypothesis that dedicated perceptual mechanisms mediate sensitivity to objecthood as opposed to agency. Object individuation and perceptual animacy have been largely studied in recent literature, whereas little is known about mechanisms mediating individuation and tracking of perceptual entities endowed with agency. By introducing a notion of an agent file, we aim to provide a theoretical framework for more constrained empirical investigations into the ability to perceptually track agency.

1. Individuation. Creating and Maintaining Reference

A large number of cognitive skills rely on the perceptual ability to single out individuals. In order to interact with, ascribe properties to, or reason about particular entities, we need to be able to pick out individuals, establish a referential link with them and maintain it over time. Many authors have pointed out that such ability must be grounded in the deployment of some nonconceptual skills: for picking out and maintaining reference to an individual, simply detecting some of its perceptual properties is not sufficient. We need to postulate a referential mechanism that provides a direct (nonconceptual and unmediated) link to this individual in order to be able to ascribe perceptual properties to it or to make perceptual judgments about it.

1.1 Reference to Objects vs. Reference to Agents in Infants

There is large evidence in the developmental literature indicating that, from early on, infants are able to deal with two distinct classes of individuals.

On the one hand, they are able to interact with and reason about objects, i.e. individuals that behave according to physical constraints. Understanding the observable behavior of objects requires — at least *prima facie* — a grasp of some of the principles underlying physical phenomena.

On the other hand, infants can interact with and reason about agents, i.e. entities endowed with intentionality, whose observable behavior cannot be reduced to those constraints governing objecthood. Being able to detect agency requires being sensitive to some specific observable cues of intentionality, such as purposefulness.

A prevailing working hypothesis in developmental studies claims that, when dealing with entities belonging to each of these domains (objects vs. agents), different individuation mechanisms are at work. Paul Bloom (Bloom, 2004) asserts that infants are commonsense dualists: 'who have two ways of looking at the world: in terms of bodies and in terms of souls' (p.191). Kuhlmeier et al. (2004) have designed and carried out a number of experiments to support this idea: infants display precocious abilities to distinguish inanimate objects (entities that behave according only to physical constraints) from animate entities (e.g. humans). These abilities suggest that infants' early understanding of animate entities does not rely on the typical individuation principles and constraints on which the individuation of objects depends (p.7).

Most studies have focused so far on the ability to distinguish animacy cues from objecthood cues, assuming that this distinction - resulting in significantly different observable patterns - is one of the main sources of infants' early knowledge about animate beings as opposed to inanimate entities. Yet little is known about how the animate vs. inanimate distinction relates to another more fundamental one, that of agents vs. objects. As the authors of this study (Kuhlmeier et al. 2004) themselves

explicitly acknowledge, it is debatable how empirical data should be interpreted with respect to this distinction:

It [...] remains an open question whether the results of the present study are due to a distinction between animate versus inanimate entities, intentional agents versus non-intentional objects, or humans versus other entities. (ibid., p.7)

If on the one hand, infants' perceptual representation of objects has been largely studied, little is known, on the other hand, about infants' perceptual representation of agents. Spelke (Spelke, 1994) has suggested a number of principles circumscribing the notion of perceptual objecthood in infants. These principles can be formulated, following Bloom (2004, , p.12), as:

1. Cohesion. Objects are connected masses of stuff that move as a whole. If you want to know where the boundaries of an object are, an easy test is to grab some portion of stuff and pull—what comes with what you are pulling belongs to the same object; what remains does not.
2. Solidity. Objects are not easily permeable by other objects; if you tap at an object with your finger, your finger does not penetrate.
3. Continuity. Objects move in continuous paths; they travel through space without gaps. An object would violate this rule if it disappeared from one location and reappeared in another.
4. Contact. Objects move through contact. A ball on a pool table is not going to move unless something contacts it; it will not run from the cue or come when it is called. The exceptions to this rule are animate creatures, like people and dogs, and also certain complex artifacts, such as robots and cars.

Taken together these principles define what counts as an 'object' for children. More precisely, they characterize a cluster of properties that children systematically privilege in picking out a certain kind of entities in the environment, prior to any conceptual identification. For this reason such clusters (often referred to as 'Spelke Objects') have been described as proto-objects.

An analogous characterization of what counts as a perceptual 'agent' for infants has not yet been clearly proposed. We argue that a principled inquiry into infants' ability to distinguish objects from agents cannot be based on mere sensitivity to animacy vs. inanimacy cues, since the ability to perceptually individuate and track agents relies on a much more fundamental distinction: the ability to individuate and track bearers of intentionality as opposed to objects[1].

Are there, beyond mere animacy sensitivity, perceptual mechanisms for tracking agency? To which extent can agents (as opposed to objects) be perceptually individuated, parsed and tracked by children?

We submit that understanding the infant's perceptual representation of agents might benefit from introducing the notion of a proto-agent, i.e. a cluster of properties that determine the way in which infants single out intentional entities prior to conceptual identification.

In order to achieve this aim, we will propose an extension of the object individuation paradigm that relies on Pylyshyn's proposal for visual object tracking. To construe a notion of proto-agent, we will assume:

(a) following Pylyshyn (2000) and Pylyshyn (2001), that human beings individuate and track single entities via a nonconceptual mechanism which can be operationalized by appealing to the notion of an object file;

(b) following Carey and Xu (2001), that the same nonconceptual mechanisms underly infants' early understanding of objects and adults' ability to keep track of perceptual entities;

(c) according to the evidence mentioned above, that there are plausibly different processes that allow children to individuate agents as opposed to objects.

Our main contention is that while tracking perceptual entities endowed with agency (or *proto-agents*) and their persistence over time, the human cognitive system opens a special sort of file, which we will call an agent file. Furthermore, we will suggest that the same mechanisms underlie infants' individuation of agents and adults' ability to perceptually keep track of agents before explicit identification.

2. Perceptual Tracking of Individuals

Nonconceptual mechanisms for individuating and tracking objects have been largely studied in the perceptual literature. In particular, humans' abilities to visually track individual entities have been extensively studied in recent years by Pylyshyn and collaborators Pylyshyn (2000); Blaser et al. (2000); Pylyshyn (2001); Scholl et al. (1999); Scholl et al. (2001).

There is robust evidence that a referential link to a particular entity, an individual, is established in virtue of low-level indexing mechanisms (which Pylyshyn dubs 'finsts') that:

- precede the deployment of focal attention;
- operate before and independently from the perceptual identification of the tracked item;
- are pretty insensitive to major featural modifications in the visual appearance of the tracked items;

These three aspects suggest the nonconceptual nature of such referential mechanisms: they allow a subject to keep track of an item without necessarily encoding specific features or attributes of that item. The nonconceptual nature of this referential link has to be considered, according to defendants of the finsts theory, as a precondition to any form of identification or perceptual judgment about objects. In this respect, the visual index theory can be considered as an extension and partial amendment of the hypothesis according to which humans access and store information about perceptual objects in terms of 'files'.

2.1 Object Files and Their Dynamics

The notion of an object file is due to the work of Kahneman and Treisman (1992). They proposed that object perception is mediated by the opening of episodic 'files' within which object tokens are constructed. Information about particular perceptual items is thus selected from the sensory array, integrated over time, and stored in such files.

Pylyshyn has suggested that object files can account for the way in which perceptual information is stored, but not for the creation and maintenance of the referential link to the object. Pylyshyn's model enriches and extends the previous theory at two different levels: first, it introduces a mechanism of nonconceptual reference as a requirement for any theory of perceptual reference to individuals; second, it embeds the traditional object file notion in a larger framework that accounts for the whole dynamics of perceptual items. Following Pylyshyn, hence, we should distinguish three independent classes of properties that are relevant for understanding the dynamics of object tracking, namely:

(1) Index-Grabbing Features

properties that cause the assignment of an index (and enable an object file to be opened for the indexed item);

(2) Index-Preservation Features

properties that allow the indexed item to be tracked (and the file to persist over time);

(3) Encoded Features

properties that can be ascribed to the indexed item (information that can be stored in the file).

Taken together, (1)-(3) characterize how 'files' work: the content of a file, i.e. the information attached to the indexed individual, is defined by properties of type (3); the dynamics of the tracking is defined by properties of type (1) and (2). It should be noted that properties belonging to (1) and (2) need not be encoded in the file: they need not be used for identifying entities as objects, i.e. they do not provide per se any basis for identifying the object or making perceptual judgments about it. Properties belonging to (3), on the contrary, are properties stored in the file, and they can be used for further qualification (categorization/identification) of the object at a conceptual level.

2.2 From Object Files to Infants' Representation of Objects

Up to now we have been referring to perceptual objects, i.e. individuals picked out by our perceptual systems without benefit from concepts or prior knowledge. Carey and Xu (2001) have argued that adults' nonconceptual representation of perceptual items and infants' object representation share a number of characteristics:

- they privilege spatiotemporal information in decisions of individuation;
- they are subject to the same set of size limitations for parallel individuation;
- they survive occlusion and are sensitive to the distinction between cessation of existence and temporary loss of visual contact (ibid., p.186).

In particular, the individuation mechanisms described in developmental psychology literature and those described by the object tracking literature both seem to privilege entities that are bounded, cohesive and that persist over time.

According to these authors, these strong similarities suggest that the account provided by studies regarding how human adults track perceptual objects and the widely discussed results in developmental psychology regarding how infants represent objects are two descriptions of the same system. If their proposal is correct, then it is plausible to assume that the same mechanisms are put to work when adults track perceptual objects and when infants track objectual entities: nonconceptual representations of individual entities involved in both domains can hence be described as the same kind, which we refer to using the notion of a proto-object (p.23 Pylyshyn, 2004).

2.2.1 Proto-objects vs. commonsense objects. It should be noted that the notion of nonconceptual representation of objectual entities (proto-objects) need not be relevant for our commonsense understanding of physical objects. While Pylyshyn (2004) acknowledges that trackable individual items are typically the proximal counterpart of commonsense physical objects, it is debatable whether this is always the case. We should hence be prudent not to conflate the theoretical notion of proto-object which is relevant for perceptual and developmental psychology with our shared intuitions about objects, which can be shaped by conceptual, linguistic and cultural factors (Casati, 2004).

The legitimacy of the distinction between a theoretical notion of (proto)objecthood and the commonsense notion of an object motivates our claim that the study of agency and agent individuation should respect a similar prescription: what we intend to suggest is that - no matter what our shared intuitions are about agents in everyday life, how we individuate them or identify them - there might be a theoretical notion, comparable to the notion of a proto-object, (we might call it the notion of a 'proto-agent') that is relevant for the understanding of nonconceptual individuation and tracking of entities endowed with agency and intentionality. Such a notion should, then, not be conflated with its commonsense counterpart: from now on we will refer to 'agents' tout court for this theoretical notion of a 'proto-agent'.

3. Perceptual tracking of agency

Picking out and maintaining a referential link to agents seems *prima facie* to raise the very same problems involved in referring to objects. We argue that merely being sensitive to agency cues is not sufficient to account for the way agents are individuated, tracked and referred to. What is needed, much as in the case of object tracking, is to understand how a subject is able not only to detect agency, but to maintain reference to an individual which:

- is unique in spite of multiple agency cues;
- can persist over time;
- can survive to changes in some of its features;
- can cease to exist, split or merge with other entities;
- can be tracked in parallel and independently of other entities of the same kind.

The need for such a mechanism that allows agent tracking can be illustrated through a case like the following:

Playing 'footsie'. John is invited for dinner at a friend's place. He is seated across from a number of guests. At a certain point in the dinner he notices that something is going on under the table: he detects an agency cue, like someone willing to play 'footsie' with him. After a while, a second agency cue is detected. John has a problem understanding whether there actually is an agent behind the detected cues, and if it there is, whether one and the same agent is responsible for both cues or more than one agent is involved.

Figure 1. Playing Footsie

Thus, keeping track of agents seems to require some sort of mechanism for the selection of individuals, the creation of a referential link and its maintenance over time. It is unclear, though, whether and how keeping track of agents could be done using the mechanism used for ordinary object tracking. The mechanism involved in object tracking is not triggered by properties such as purposefulness which could plausibly be a necessary condition for parsing an agent as persistent. If this is so, then such mechanism would not detect the agent's willingness of playing footsie with John and, thus, would not allow him to keep track of the agent behind the agency cue (not to say that it would be helpless for John's understanding of whether one and the same agent is responsible for both, the first and the second, cues). Furthermore, there may be cases in which objectual cues conflict with agency cues and it is hard to establish whether individual entities are tracked in virtue of the former or the latter. The literature on infants' perception of intentionality in self-propelling shapes (Heider and Simmel, 1944; Premack, 1990; Schöll and Tremoulet, 2000) represents a good case study for investigating such issues. Heider and Simmel (Heider and Simmel, 1944), for instance, showed that people tend to associate particular intentions or intentionality (doing some good to someone, being good, etc.) to certain systematic ways in which some geometrical figures (circles, triangles, etc.) move while they "tell a story" in a movie. One might ask under which conditions criteria of individuation and perceptual tracking are based on objectual cues rather than agency cues. It is an interesting empirical issue to study whether:

- an item preserving its agency cues and changing its objectual features can still be tracked as the same agent (e.g., an 'evil' triangle turning into an 'evil' square without violation of other constraints or an 'evil' triangle disappearing and reappearing at another place with a violation of spatio-temporal continuity);
- an item preserving its objectual features and changing its agency features can be still tracked as the same object (e.g. an 'evil' triangle suddenly turning into a 'good' triangle).

Evidence from such experiments supports the claim that there are specific individuation processes that depend on agency which can be dissociated from individuation of the same entities as objects (Bloom, 1996). In what follows we will argue that the question of whether children are able to individuate and refer to agents by appealing to mechanisms similar to those of object tracking might benefit from an extension of the notion of 'file' to the case of intentional agents.

3.1 Extending the File Notion to Agents

An interesting way to operationalize perceptual individuation and the tracking of agents involves an appeal to the notion of file described above, extending it from the domain of (proto-)objects to the domain of (proto-)agents by postulating the notion of an *agent file*.

If the ability to track agents is not reducible to mere sensitivity to agency cues, in virtue of the agent's continuity and persistency over time, it seems reasonable to assume that in order to track such entities and ascribe properties to them a subject might use files. Their dynamics can be then described by three classes of properties analogous to those we introduced for object tracking.

1. Agency-grabbing Properties

properties that enable individuation of an agent (i.e., properties that are necessary to open an agent file);

2. Agency-preserving Properties

properties that allow an individual agent to be tracked (and the file to persist over time);

3. Encoded Properties

properties that can be stored and retrieved from the agent file.

In the example introduced earlier, such properties could be: John's detection of an animacy cue on his leg (an agency-grabbing property); the spatio-temporal coherence of this stimulation (for instance its spatial orientation or its frequency: agency-preservation features); any other information that can be ascribed to the agent (including non-perceptual information such as beliefs about the agent's explicit intention of playing 'footsie': encoded properties).

The main idea is that files are domain-specific so that the kind of information and properties that an agent file can store is different from that which is stored in an object file: the encoded information in the former type of file can include intentions, which can be retrieved in reasoning or while making judgments about the agent.

3.1.1 Agent Files vs. Agency Cues. It should be noted that by positing the existence of a class of properties that allow an individual agent to be picked out we are not endorsing the idea that as soon as such properties are detected an agent file is automatically opened.

In the case of objects, it has been shown that detection of objectual cues is not per se sufficient for tracking an entity: well-formed perceptual objects might still lack conditions for being tracked over time (Scholl et al., 2001). The mere presence of objectual cues does not entail the presence of reference fixation cues.

Similarly, there can be cases in which merely detecting an animacy or agency cue, although sufficient to respond to its presence, does not imply that an agent is individuated and susceptible of perceptual tracking. John can for example withdraw his leg as a reaction to the stimulation, without necessarily individuating an agent. Since agent tracking is not required in such cases (either because there is no need to individuate the source of the cue or because there is no need to represent this entity as persistent over time), we assume that opening an agent file is not needed. Actually, we claim that agency-grabbing properties are necessary conditions for agent tracking but they are not per se

sufficient for the creation of an agent file. Distinguishing between simple detection of properties and full-blooded perceptual tracking - as in the case of detection of objectual cues vs. object tracking - is a natural consequence of our approach.

4. Object vs. Agent Tracking: Empirical Directions

Introducing the notion of an agent file raises the problem of understanding whether and how such files are related to files that apply to tracking of objects. If we accept the hypothesis that there are two distinct mechanisms for agent vs. object tracking, it is reasonable to ask how such mechanisms can interact or be mutually related in specific experimental conditions. Broadly speaking, there can be two general options:

(A) Independence view

At each level of description, object files and agent files share no common features. Properties relevant for describing the two kinds of files and their dynamics (file fixation, preservation and content) are necessary and sufficient to account for the two distinct kinds of tracking.

(B) Dependence view

At some level of description, object files and agent files might share common features. Properties relevant for describing the two kinds of files and their dynamics are sufficient but not necessary to account for the two distinct kinds of tracking. Some properties of one kind of file can be exploited at a given level by the other kind of file.

We will outline in what follows some cases in which the relation between the two kinds of files can be empirically studied at each level.

4.1 Object vs. Agent File Fixation Cues

Once we acknowledge that object vs. agent tracking can be articulated at three distinct levels, we can raise the question of whether, in the case of tracking,^[2] different classes of properties allow fixing of reference to an item in the case of agents and in the case of objects. Let us consider an example drawn from a classic arcade game of the '80.

The Asteroids Game. A player must drive a space vessel in order to avoid and destroy both asteroids and enemy vessels. Asteroids are characterized by passive physical movements, while enemy vessels are characterized by motor patterns revealing purposeful behavior (like avoiding asteroids, actively changing speed and direction, and shooting other vessels). The player must be able to track both asteroids and enemy vessels and react appropriately to their movements in order to destroy them, survive and win the game.

Figure 2: The Asteroids Game: objects vs. agents

This example illustrates a case in which a subject needs to pick out and maintain reference to two different types of individual (objectual entities, like asteroids, vs. intentional entities, like enemy vessels). We assume that in order to do this, the subject must detect two classes of cues prior to any further identification or categorization. It seems that, to establish reference, the subject does not need to access properties that might be used to identify objects and agents (e.g., asteroids and enemy targets might have the same shape). In short, we are claiming that a subject becomes able to individuate and maintain reference to entities belonging to two distinct classes (proto-objects vs. proto-agents) in virtue of his ability to detect objectual vs. agency cues, and of the specific task requirements that force him to maintain perceptual reference to individuals over time. If it can be empirically

demonstrated that in similar conditions subjects display robust capabilities to differentially detect reference-grabbing properties belonging to two mutually exclusive classes as a condition for tracking items, then we might plausibly conclude that two distinct and independent individuation processes are at work. This distinction would support the claim that object tracking and agent tracking are independent at the level of reference fixation properties.

The existence of two distinct classes of reference fixing properties — although sufficient to support the independence view at this level (agency cues can be segregated from objectual cues) — is not sufficient to conclude that the two mechanisms of agent and object tracking are necessarily independent. We must also consider the relation between agents and objects at two other levels, viz. reference preservation and feature encoding.

4.2 Object vs. Agent File Preserving Properties

We have insisted on the fact that fixing perceptual reference is still not enough for tracking individual entities over time. Tracking implies preserving a referential link to a perceptual item already picked out. This raises the question of determining in virtue of which properties reference to a single individual can be maintained over time.

In the example above, in order to be able to avoid an asteroid, a subject must be able to track it as persisting over time. Following Pylyshyn, we assume that the properties used for fixing reference when the individual is picked out need not be the same as those that preserve reference. Once an item is individuated, reference fixing properties can be discarded without breaking up the referential link itself. Provided there are some file preserving properties, an item can undergo major changes without losing its singularity: we argue that if preservation conditions are met, an individual will not cease to be treated as a single perceptual item even if the properties initially used for its individuation have disappeared. Assuming that an asteroid was picked out as an individual object through its shape, it is not necessary that its shape be maintained over time in order for the object to persist in a perceptual tracking task.

Recent literature on Multiple Object Tracking has demonstrated that tracked items can survive several kinds of disruption of their features. It seems, though, that certain properties are required for an item to preserve its individuality. It is reasonable to assume, on the basis of this literature, that in our asteroids game example, although asteroids might 'survive' temporary occlusions which do not alter their trajectory, they would fail to maintain a perceptual link in cases of spatio-temporal incoherence, like sudden shrinking or disappearing and reappearing at a different location (Scholl and Pylyshyn, 1999). Coherence of trajectory as well as cohesiveness (Van Marle and Scholl, 2003) are hence examples of properties that seem to be required in order to maintain reference to perceptual items in tracking tasks.

We might then ask whether the properties used to keep reference alive are different in the case of agent or object tracking.

4.2.1 An Objectual Bias in Agent Files Preservation? Many observable properties might in principle be recruited to preserve agency. For instance, the reiteration of animacy cues (e.g. an enemy vessel periodically shooting at the player) or the lack of cues of agency disruption (e.g. the lack of observable indicators of an enemy vessel being destroyed) are good candidates for the properties that contribute to the preservation of agent files. This might suggest that at the level of file preservation as well, agents and objects do not share any property (and hence that object files are independent from agent files). Nevertheless, given the fact that intentional entities are usually coinstantiated with objectual entities ('bodies') in our environment, it is plausible to assume that agents can be perceptually tracked via objectual preservation properties. We call this phenomenon an objectual bias in agent files preservation. This amounts to saying that:

(a) there is a class of properties that are sufficient to maintain reference to an object, once an object file is opened.

(b) there is another class of properties that are sufficient to maintain reference to an agent, once an agent file is opened;

(c) among each of the above classes there are single properties that are more or less strongly correlated with preservation of the file: in the case of objects, there might be properties strongly or weakly correlated with object persistence;

(d) one and the same property can be correlated with different degrees of reliability to agency or objecthood preservation;

(e) object preservation properties that are weakly correlated with preservation of agency might nonetheless be recruited for preserving agency when other agency preservation cues are absent or not available.

The fact that one and the same property P might be relevant to different degrees for maintaining reference either to an object or to an agent does not threaten per se the very possibility that there are two distinct kinds of file. We are just suggesting that it is empirically possible that, in virtue of the robustness of the agent-body correlation in our environment, nonconceptual tracking of agents might exploit objectual properties.

Which classes of properties are actually being used for tracking a proto-agent (whether they are properly described as agency-related rather than object-related) is, thus, an empirical issue that calls for experimental work. We maintain that — until a number of explicit conditions for distinguishing between these two classes are formulated — empirical results supporting the hypothesis that infants track individual entities as persistent objects, as persistent agents or as agents persisting in virtue of their objectual features might be seriously undermined. As a possible suggestion, one might test whether the fact of using objectual properties for keeping track of agents' persistence has any consequences on the number of items an infant is able to individuate and track at the same time.

4.3 Object vs. Agent File Content

Once the conditions for fixing and preserving an item are met, we have at our disposal a device (a file) to store information attached to this individual. The way in which a file attaches properties to an entity is radically different from the way in which the nonconceptual mechanism at work while tracking an individual establishes and preserves reference to it.

On the one hand, nonconceptual tracking provides a direct link to an individual, i.e. allows a particular entity to be grabbed and segregated from other individuals or other properties of the scene: tracking is thus a necessary condition for parsing an individual as persisting.

On the other hand, having a file attached to an individual is required in order to ascribe some properties to it.

In the object file literature, a file content typically consists of properties that can be predicated of an object and used in perceptual judgment, categorization and identification. Following Scholl et al. (1999, p.2):

[This kind] of property determines the object's appearance — what a particular object looks like — including its color, shape, lightness, and texture. We call these featural properties.

If our hypothesis on the extensibility of the object file paradigm to the case of agent tracking is valid, then we can think of the content of an agent file as a temporary structure attached to an individual entity which can store information about this intentional entity.

It is an empirical issue to understand whether and what are the particular constraints on properties that can be stored in an agent file as opposed to an object file. Yet, we submit that the extension of the file notion to the case of agency can explain how perceptual agency properties can be ascribed to a particular agent and used in perceptual judgment.

When an infant is asked to identify among a number of perceptually available agents which is the agent displaying, say, aggressive behavior, we claim that she is making use of information stored in a file to perform this task. It should be noted that since properties stored in a file are those properties that are conceptually accessible, they need not have any relevance for understanding how the nonconceptual parsing and tracking of individual entities is done.

Are there really Agent Tracking Mechanisms?

The previous paragraphs were meant to outline a number of empirical issues related to the possible distinction of object vs. agent tracking at three different levels: (a) the level of reference-fixing, (b) the level of keeping reference alive and, (c) the level of ascribing properties. It might be objected that our proposal of an extension from the domain of objects to that of perceptual agents is in fact a mere redescription of the functional role of object files and thus the idea that we are able to track agents in virtue of dedicated mechanisms would be brought into question. The proposal underlying such objection can be called the deflationary view on agent tracking.

Our reply to this objection can be articulated at different levels.

- The rationale for the existence of agent tracking mechanisms is that if we want to account not only for detection of and sensitivity to agency or animacy cues as opposed to objectual cues, but also for the ability to maintain reference to an intentional entity persisting over time, then we need to explain how this representation of persistence is achieved.
- From the fact that there might be significant similarities in the dynamics and nature of object vs. agent tracking (what we called the dependence view), it does not follow that the latter should be reduced to the former. In the previous paragraph we made some suggestions about possible empirical ways to assess the similarity/difference between the two mechanisms.
- Our proposal is consistent with a large literature in developmental psychology that has demonstrated the existence in children of two distinct domains of perceptual properties: the domain of animacy and the domain of objecthood. Our contribution can be considered as a framework for extending these investigations to the question of how individuals endowed with animacy and agency can be grabbed as entities persisting over time.
- The hypothesis of the existence of agent files and agent-related tracking abilities opens up some interesting research directions leading to the study of possible conflicts and dissociations between object and agent tracking. Empirical research might shed light on the fact that:

§ intentional and objectual entities can compete for the same attentional or computational resources: it is possible that the limit on the number of items a subject can track at one time is dependent on (or independent of) the class of tracked items (objects only, agents only or objects plus agents);

§ there may be interesting cases of dissociations, i.e. selective impairments of either of these abilities without functional consequences for the other: we might imagine cases of subjects being able to track objectual entities but not agents or viceversa.

§ inattention blindness studies might investigate whether the existence of two distinct classes of entities has any effect on their neglect: this might suggest that the traditional list of 'styles of attention' (object-based vs. space-based attention) should be extended to include a third kind of style (agent-based attention);

§ developmental investigations might benefit from the notion of an agent tracking mechanism to establish the conditions under which infants represent the persistence of a perceptual item over time.

The above considerations suggest that dedicated mechanisms for agent tracking are likely to have psychological reality and account for a number of capacities involved in perceptual reference to persistent entities endowed with agency. The relation of such mechanisms to those involved in

perceptual reference to objects remains, however, an open empirical issue deserving further investigation. We have given arguments in favor of a moderate dependence view suggesting that in some cases objectual properties might be recruited to establish and maintain reference to agents.

6. Conclusions

A crucial step for understanding our nonconceptual abilities to refer to individual entities consists in explaining how such individuals are picked out and tracked over time, prior to any form of categorization or conceptualization. Such mechanisms of direct reference to individuals lay probably at the basis of both infants' abilities to parse objects and adults' capacity to keep track of multiple perceptual items. We have proposed an extension of the studies on object individuation and tracking to the domain of perceptual individuation and tracking of entities endowed with agency. This extension is intended to fill a gap between the study of perceptual sensitivity to agency cues and a full-fledged understanding of how perceptual items tracked as agents (what we called proto-agents) can persist over time. We have argued that unless some explanation of the mechanisms underlying agent persistence is provided, many empirical results concerning infants abilities to track animate entities could be dramatically undermined. Our proposal of three distinct levels that might be involved in agency fixing, preservation and ascription is meant to provide the basic requirements for any explanation of perceptual capabilities to track agents. We argue that by analyzing the dynamics of agents at these three different levels, more principled answers might be given to the question of possible interferences and biases between mechanisms dedicated to agent vs. object individuation.

References

- Blaser, E., Pylyshyn, Z. W., and Holcombe, A. O. (2000). Tracking an object through feature space. *Nature*, 408:196–199.
- Bloom, P. (1996). Possible individuals in language and cognition. *Current Directions in Psychological Science*, 50:90–94.
- Bloom, P. (2004). *Descartes' Baby. How The Science Of Child Development Explains What Makes Us Human*. Basic Books.
- Carey, S. and Xu, F. (2001). Infant's knowledge of objects: beyond object files and object tracking? *Cognition*, 80:179–213.
- Casati, R. (2004). Common sense, philosophical and theoretical notions of an object: some methodological problems. In *Barcelona ESPP-SPP Joint Conference*.
- Heider, F. and Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology*, 57:243–249.
- Kahneman, D. and Treisman, A. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24:175–219.
- Kuhlmeier, V., Bloom, P., and Wynn, K. (2004). Do 5-month-old infants see humans as material objects? *Cognition*, 94(1):95–103.
- Premack, D. (1990). The infant's theory of self-propelled objects. *Cognition*, 36(1):1–16.
- Pylyshyn, Z. W. (2000). Situating vision in the world. *Trends in Cognitive Sciences*, 4:197–207.
- Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, 80:127–158.
- Pylyshyn, Z. W. (2004). Notes for the oleron summer workshop on objects.
- Scholl, B. J. and Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology*, pages 259–290.
- Scholl, B. J., Pylyshyn, Z. W., and Feldman, J. (2001). What is a visual object: Evidence from target-merging in multiple-object tracking. *Cognition*, 80(1-2):159–177.
- Scholl, B. J., Pylyshyn, Z. W., and Franconeri, S. L. (1999). The relationship between property-encoding and object based attention: Evidence from multiple object tracking.
- Scholl, B. J. and Tremoulet, P. (2000). Perceptual causality and animacy. *Trends Cogn Sci*, 4(8):299–309.
- Spelke, E. (1994). Initial knowledge: Six suggestions. *Cognition*, 50:431–445.
- Van Marle, K. and Scholl, B. J. (2003). Attentive tracking of objects versus substances. *Psychological Science*, 14(5):498–504.

French Version

- Traquer l'agentivité

Résumé: Dans ce papier nous apportons quelques résultats méthodologiques pour l'étude de la traque nonconceptuelle des «agents perceptuels», c'est-à-dire des entités perceptuelles pourvue d'agentivité. Le cadre théorique proposé permet de préciser la notion de 'proto-agent' et d'expliquer les relations possibles entre les capacités de traque des objets et des agents.

**Acknowledgments:* We wish to thank participants in the *2004 Oleron Workshop on Objects and Reference* as well as Dick Carter, Roberto Casati and Erika Marchetto for valuable suggestions on earlier versions of this paper. Ángeles Erañas work was supported by the research project CONACYT number 41196, "Filosofía de las prácticas científicas". Dario Taraborelli's work was partly supported by the European Commission IST-2002-002114 "Enactive" Network of Excellence. The *playing footsie* picture (p.9) is artwork by Stefano Mandracchia.

[1]We use hereafter the term 'tracking' in a technical sense to refer to the perceptual ability to parse and maintain reference to individual entities (see Pylyshyn, 2000). Issues related to *reidentification* of individuals (mediated by conceptual representations) are beyond the scope of the present analysis and should be kept distinct from the narrow notion of tracking that we adopt in this article.

[2]We do not intend to discuss here cases of mere *sensitivity* to objectual cues vs. agency cues. See the paragraph in section 3.1.1 on the distinction between detecting and tracking agency.

Tracking objects, Tracking agents

Nicolas Bulot (Institut Jean Nicod, University of British Columbia),

Patrick Rysiew (University of British Columbia)

(Date of publication: 3 March 2005)

Abstract: Assuming a realist ontology for the world of physical objects and an object-dependent epistemology of singular knowledge, we compare the tracking of objects without mental states with the tracking of intentional agents. Our contention for studying the uniqueness problem in situated agent tracking, referred to as the 'object-dependence view', is that in the most basic case, humans track intentional agents by tracking their bodies. However, we uphold that tracking intentional agents as intentional agents requires specific capacities for detecting and understanding biological motion and intentional states.

Introduction

Animals and humans have to keep track of *individuals* in their environment, both in perception (sensorimotor tracking) and in cognition (e.g., spatio-temporal localization and linguistic reference via memory, communication and reasoning). Items that are typical targets for tracking are things such as stationary physical objects (e.g., rocks, plants, trees, buildings, or attached artifacts), moving physical objects (e.g., animals, certain artifacts) and human beings. All such items are located in a spatial environment, persist over time, and are – or at least closely related to, in the case of conspecifics' bodies – *physical objects* that respect non trivial *objecthood criteria* such as having cohesive parts, following continuous spatio-temporal paths, or possessing causal powers and dispositions. Perceptually tracking these objects through space and over time requires possessing sensorimotor systems (e.g., the oculomotor system, the visuo-haptic system or the auditory system) that can anchor into and smoothly pursue objects' properties. Nonetheless, one may suspect that tracking *intentional agents* (i.e. creatures to whom it is natural to attribute intentional states such as beliefs, plans, desires, and who may even participate in shared intentionality), as opposed to *physical objects without mental states* (i.e. objects which are not intentional agents) exploits or requires further abilities and strategies. In particular, at least for humans, tracking conspecifics amounts to tracking intentional agents. This raises the question of how the perceptual tracking of non-intentional objects relates to the keeping track of intentional agents. Here, we propose an extension and augmentation of recent work on object-tracking to the tracking of intentional agents. Based on the examination of the elementary procedures available for pursuing agency, our principal suggestions are as follows.

First, identifying intentional agents is significantly dependent upon the perceptual abilities of physical-object tracking, and might therefore be explained by the 'object-file' hypothesis (this hypothesis is explained in section 3), which we suggest to use in the study of the tracking of intentional agents (section 4). In the most elementary case, humans track intentional agents *as physical objects*: they track such agents by tracking their bodies (section 5). Even though this kind of tracking is insufficient for keeping track of human individuals 'as' intentional agents (and explaining their behavior with an intentional stance), it may suffice to explain a number of interacting and situated behaviors in social contexts *with* intentional agents. Second, however, tracking intentional agents 'as' *intentional agents* requires additional capacities for detecting and understanding intentional states and certain further properties which creatures with such states can exhibit (section 6) – e.g., (ir)rationality, and the capacity to participate in shared intentionality. We note, however, that tracking a human individual as an intentional agent may require an appeal to specific perceptual cues and may even recruit basic sensorimotor skills – such as the detection of biological motions – whose tracking might be independent of the understanding of conspecifics' mental states. For reasons of parsimony and computational economy, unless we have reason to think that there is a separate system devoted to tracking intentional agents, we should suppose that the same mechanism used to track physical objects is recruited for tracking intentional agents.

Keeping track of token individuals and the uniqueness problem for object tracking

We shall assume a *realist ontology* of the material world. According to this ontology, the world is composed by individual physical objects/bodies – that is, tokens of external mind-independent individuals with identity conditions (e.g., Strawson, 1959). If this ontology is correct, the identity conditions of physical objects (including the bodies of human agents or animal organisms) can be explained on the basis of mind-independent causal properties and composition principles – such as chemical composition and genetic properties for living organisms, continuity criteria for objects in general, and so on. Let us call *objecthood properties* those properties – such as cohesion, continuity and impenetrability – which are criteria of physical objecthood, and which are directly relevant to the identification and re-identification of particular (physical) objects. Objecthood properties and criteria have been studied and discussed in both philosophy and psychology.

We will accept also an *object-dependant epistemology*. According to this epistemology, human knowledge is based on skills and singular representations/contents that enable us to single out and track mind-independent objects.[i] In order to interact with objects, to access their properties, to ascribe to them abstract properties or to reason about them, humans need to be able to pick out individual entities, fix a referential link with them and maintain that link over time. However, because the available properties or descriptions of an individual *x* (at a given moment) are not always sufficient to single *x* out and refer to *x* as the same individual, cognizers are confronted with *the uniqueness problem of object tracking* in many situations: For any given unique physical object *x*, how can the cognizer keep track over space and time of *x*'s 'sameness and uniqueness'? The object-dependant epistemology has thus to postulate that cognizers possess capacities which enable them to track perceptually *x* and to keep track cognitively over time of *x* as the same continuing individual – i.e. as a function of *x*'s sameness and uniqueness. For these capacities to be the basis of keeping track of *x*'s uniqueness, they must enable the cognizer to *distinguish* over time *x* from all the other individuals of the same kind. When used for conceptual thinking, this ability is what philosophers have referred to as singular representations and judgments fulfilling the 'Russell's principle' (Evans, 1982) or the 'know which' requirement (e.g., McDowell, 1990; McDowell, 1998; Peacocke, 1991; Strawson, 1997). According to Pylyshyn (2003), there is also reason to think that such an ability must be ascribed to the visual system itself.[ii]

The notion of 'object-file' as an explanatory notion in understanding the cognitive relation to the object uniqueness

Taken together, this realist ontology and this object-dependant epistemology have a remarkable implication: token physical objects can be *identified* and *re-identified* by cognizers, because objects fulfill sameness criteria which are traceable over time by cognitive means. A fundamental challenge is thus to determine *how* cognizers solve the uniqueness problem for object tracking, that is to represent and track token objects in a way which preserves the link with their identity (or uniqueness). Several analyses[iii] have introduced an intuitive answer to this question, either as a useful metaphor (Perry, 2001a: 51-2) or as explanatory concept. The answer consists in the supposition that agents can store information about individuals in specialized, unique 'repertories' or 'files'. The two seminal ideas are that, in standard cases, (1) the uniqueness of the mental repertory ought to be related to the uniqueness of the target individuals in the world and that (2) one has to distinguish the creation and maintenance of a file from its purely descriptive content. These repertories are frequently called 'object files', and have been discussed in the theory of perception and language. Several notions of 'object file' remain to be distinguished, specifically in terms of the relation they bear to memory.

With respect to object perception, a number of authors have hypothesized the existence of a perceptual ability to manipulate (short term) *perceptual* object files (Kahneman et al., 1992; Pylyshyn, 2003). An *object file* in this sense is a mechanism that allows keeping track of an individual in a perceptual field. Kahneman and his colleagues (Kahneman & Treisman, 1984; Kahneman et al., 1992) have suggested that we store information about visually presented objects in *visual* object files. They introduced the notion in these terms:

"Some time ago we proposed an account of object perception as the process of setting up and utilizing temporary 'episodic' representations of real world objects, which we call object files (Kahneman &

Treisman, 1984). These are separate from the representations stored in a long-term recognition network, which are used in identifying and classifying objects.” (Kahneman et al., 1992: 176)

These authors introduced this notion of object file along with the distinction between *temporary representations* of objects and *long-term recognition* of objects. This distinction is justified by Kahneman et al. (1992: 176) mostly by empirical evidence, including the primacy of objects in determining the allocation of attention (e.g., Baylis & Driver, 1993; Duncan, 1984; Egly, Driver, & Rafal, 1994; Pylyshyn, 2003: 181-200; Scholl, 2001). Attention to any one property of an object causes irrelevant properties of that object to be attended to, as in the Stroop effect (Stroop, 1935). Moreover, the division of attention between relevant attributes is facilitated if the attributes belong to the same object (Treisman, Kahneman, & Burkell, 1983). Kahneman et al. (1992) have also found an ‘object-specific matching effect’: the focusing of attention of a target object not only enhances the salience of all its current properties – it also selectively reactivates the recent history of that object.

In the same tradition, Pylyshyn (2001; 2003) has further suggested that these object files are ‘linked to’ individual distal items via an indexing mechanism called a ‘visual index’ (or ‘FINST’). This analysis allows not only to account for the encoding of object-property information in a file, but also for its production and persistence. According to this kind of analysis, there are at least three classes of properties relevant to the understanding of the dynamics of tracking individual objects. First, *index-capture features/cues* are properties that cause the assignment of an index (and an object file to be opened). Second, *index-preservation features/cues* are properties that allow the indexed item to be tracked over time (and the file to persist over time). Third, *encoded features/cues* are properties that can be ascribed to the indexed item (information that can be stored in the file). The encoded features or cues are typically considered as *descriptions* that are suitable for the individual whose appearance has triggered the opening of the file; they are frequently said to be responsible for the visual awareness of the object. Taken together, these three kinds of properties characterize how ‘object files’ function. The *narrow* content of a file, i.e. information attached to the indexed individual, is defined by encoded features/cues. The dynamic anchoring of the tracking is defined by index-grabbing and index-preservation cues. The *broad* content of a file is defined by its existential dependency with respect to the causal properties of the external target object.

Extending the ‘object-file’ notion to agent tracking

Human beings seem to be able to deal with two distinct classes of individuals. On the one hand, they are able to interact with and reason about *non-intentional* physical objects (e.g., Baillargeon, 2001; Spelke, 1990; Spelke et al., 1995). Tracking the observable behavior of material objects requires a detection or understanding of some of the regularities underlying physical phenomena. On the other hand, humans can also interact with and reason about *intentional agents*, i.e. individuals to whom we can ascribe *intentional states* such as beliefs, desires and intentions (e.g., Bloom, 2002, 2004; Dennett, 1971; Malle, Moses, & Baldwin, 2001). In order to account for the observable behavior of agents it seems that we need to add further principles, principles not reducible to those used in tracking physical objects, and which depend on intentional agency. This distinction between two domains in which different individuation mechanisms are at work, is what Bloom calls “commonsense dualism”. According to his analysis of common sense (Bloom, 2004: 191), dualists have two ways of looking at the world: in terms of bodies and in terms of souls; a direct consequence of this dualism is the idea that bodies and souls are separate.

There is an interesting way in which singling out and tracking intentional agents might be explained by appealing to the notion of ‘agent file’ akin to the notion of ‘object file’ in vision. This approach has already been mentioned at least in philosophy. Perry (2001a: 51-2; 2001b: 123-46) suggests that information about agent identity can be stored in ‘files’ and that acquiring further perceptual information about an agent can lead to the opening, merging, splitting, or enriching a file. *Prima facie*, it seems reasonable to assume that in tracking agents humans might exploit the same taxonomy of properties that we have introduced for object tracking. We need to *pick out* a single intentional entity, to *keep track* of it across time and *ascribe* information to it. According to this analysis, one should enumerate *agent-indexing features/cues* that are properties that enable the anchoring in an intentional agent – and to open an ‘agent file’; *agent-preservation features/cues* that are properties that allow the indexed agent to be tracked – and the file to persist over time; *encoded features/cues* that are

properties that are described by symbols within the file and that are ascribed by the system to the tracked agent (information that is stored in the agent file).

The conjecture that there exists such a dedicated mechanism for tracking an agent – an ‘agent file’ – raises the question of understanding whether and why such files are related to ‘object files’ that apply to tracking non-intentional objects. Here we are presented with at least three distinct options. According to an *independence view of agent and object tracking*, object files and agent files do not share any common feature and are dedicated to distinct domains. According to a *deflationary view of agent tracking*, there is only one type of mechanism that keeps track of and refers to either an object or an agent; agent files are thus reduced to object files. According to an *object-dependence view of agent tracking*, although there can be important differences between tracking physical objects and intentional agents, we track intentional agents by tracking objects. In the following sections, we will argue in favor of the third option, the object-dependence view.

The object-dependent view of agent tracking: tracking agents by tracking physical objects

A number of arguments support the idea that tracking an agent requires tracking it on the basis of its status as a physical object because ordinary tracking of intentional agents seems to rely on a robust correlation between agents and bodies. One attractive argument runs as follows: given the usual connection between intentional agents and their bodies, on the one hand, and agents’ bodies and objecthood properties, on the other, being able to perceptually track objecthood properties (i.e. physical objects/bodies) might provide a means of accessing and keeping track of intentional agents.[iv] Somewhat more formally, one argument for the object-dependence view is thus as follows:

(1) *Thesis of the co-instantiation*: A human agent *a* is co-instantiated with his/her human body *b*.

(2) *Thesis of body tracking*: The tracking of a human body *b* requires tracking *b*’s object properties.

Conclusion: Hence, the tracking of a human agent *a* requires the tracking of the object properties of his/her body *b*. (Note that the conclusion should hold not only for humans but also *mutatis mutandis* for primates and other animal species.)

The thesis of the co-instantiation is the one most open to conflicting interpretations and ontological debates (see the next section). We shall first comment on the thesis of body tracking, which is far less subject to controversy. It seems simply to fall out of an examination of the common sense assumptions about the spatio-temporal characteristics of human bodies, in a way that does not conflict with common assumptions in cognitive science. For, as a biological organism, the body of a human being fulfills a number of objecthood criteria.[v] Such a body has *cohesive* parts (limbs, organs) that are functionally and hierarchically organized. It possesses a defined *shape* which is delimited by surfaces (skin) that are boundaries between internal parts and external phenomena or objects. It is mostly impenetrable in a weak sense: a human body cannot strictly coincide in space with another human body. Several adult human bodies can only be in contact or close contact but cannot fuse or overlap in space. Each body can be counted among groups of several individuals because it is a unit. In addition, phenomena which threaten the cohesion and unity of the body are threatening the very existence of that body as a living unit. This unity is linked to its uniqueness. One of the clues to this uniqueness is that, as with ordinary physical objects, each body follows a continuous and unique spatio-temporal path from birth to death. As a result, this body can be located within different frames of reference (egocentric, allocentric) and many of the inferences that one can make about a human body are based on an implicit knowledge of the previous characteristics. For instance, if one knows that Paul is in the aircraft number 743, one can infer the spatio-temporal location of Paul’s body on the basis of the knowledge of that aircraft’s allocentric position (this is a cognitive way of keeping track of a physical target object, in absence of present perception). In addition, the eventuality of the aircraft’s crash has its deadly implications *because of* our persistent assumptions of all the previous characteristics (that is, the physics of the human body is known to be such that it *will be* destroyed by any event that would cause destruction of the aircraft).

Given that human bodies fulfill these spatio-temporal objecthood criteria, the premise of body tracking seems unproblematic. For object properties/cues are the main available target for object perceptual tracking and the main evidence for reasoning about the spatio-temporal paths of objects which are not

currently perceived. In addition, this last assumption seems consistent with recent evidence on perceptual tracking systems. Consider first the case of visual tracking and visual attention. Many of the properties which are assumed to be used in order to open and maintain a visual object file (or to assign a FINST) are object properties in the sense that they fulfill objecthood criteria (Carey & Xu, 2001; Kahneman et al., 1992; Pylyshyn, 2003). In particular, they are related to *cohesion* and *spatio-temporal continuity*. For instance, in the standard MOT experiment (Pylyshyn & Storm, 1988), the only available feature to distinguish targets from non targets is the spatiotemporal path followed by each element. Although the system for visual tracking can deal with short periods of occlusion, tracking abilities are clearly impaired as soon as the target element infringes on the cohesion principle via transformations similar to “pouring” (vanMarle & Scholl, 2003) or exploding. It is also difficult to track particular parts of distinct objects (Scholl, Pylyshyn, & Feldman, 2001). These results are evidence that physical objects may be typical targets for visual tracking. Hence, given that human bodies fulfill the major objecthood criteria, human bodies may be typical targets for visual tracking. If the object-dependence view is true, this observation about visual/perceptual tracking may generalize beyond mere visual tracking to multimodal tracking, in a way which should be congruent with a variety of cognitive strategies for keeping track of individuals in the world.

The change blindness paradigm provides further evidence for the view that the perceptual tracking of intentional agents can be on occasion performed on the basis of the tracking of basic (spatio-temporal) objecthood cues. In one experiment by Simons & Levin (1998) an experimenter initiated a conversation with a pedestrian and during the interaction he was surreptitiously replaced by a different experimenter. Only half of the pedestrians detected the change. What this seems to suggest is the fact that the pedestrian might use a primitive mechanism such as a (visual) object file (or a FINST) to perceptually track the experimenter agent – a mechanism that uses spatio-temporal features for tracking instead of elaborate unique visual descriptions. In the case of detection failure, the basic mechanism might have been blind to the person change during the ‘surreptitious replacement’ because this event would be (incorrectly) interpreted as a temporary occlusion of the *same* visible object. This type of analysis is consistent with the deictic or interactive approaches to vision (Ballard, Hayhoe, Pook, & Rao, 1997; Churchland, Ramachandran, & Sejnowski, 1994; O’Regan & Noë, 2001; Pylyshyn, 2000), according to which situated vision is not a rich description of the visible scene but essentially involves direct relations to objects in the vicinity, which are probed or explored only when required.

Moreover, while we think there is good reason to resist the stronger, deflationary view of tracking agents (section 6), it can seem that *at least* the object-dependence view has got to be correct. Recall Bloom’s claim that humans are natural dualists who have two ways of looking at the world: in terms of bodies and in terms of souls and that a direct consequence of this dualism is the idea that bodies and souls are separate (Bloom, 2004: 191). Not only are they separate (ontologically distinct), but ‘souls’ are held to be immaterial, and hence not perceivable by the usual means. However, if this be the case then souls cannot be tracked directly; whereas, agents’ bodies – which peoples’ souls somehow ‘inhabit’ – *are* among ‘the furniture of the world’. Hence, agents’ bodies are the only public, tractable objects relating to agency even from the dualist viewpoint. (Cf. Wittgenstein: “The best picture of the human soul is the human body”.) Given the assumption of common sense dualism of ‘one body, one soul’ (which seems to imply accepting the premise of the co-instantiation for the period of ‘terrestrial life’), thus, ‘commonsense dualism’ lends itself to the object-dependence view.

It is essential to note, however, that essentially the same reasoning holds even once dualism is rejected. For while materialism about the mind entails that *intentional states* (agents’ states such as beliefs or emotions), ultimately, are (or supervene upon) *physical states*, there is no suggestion among current materialist theories (simple behaviorists aside) that the intentional states are just the familiar, observable states of a person’s body. Rather, it is usually said that what constitutes a given intentional state are certain states of a person’s brain (e.g., Dennett, 1978a; Shoemaker, 1997), perhaps together with facts about those states’ history, their relation to other such states, and so on. Materialists about the mind, of course, disagree over *just which* such physical states and facts constitute intentional states. But the important point is that none of the most promising and popular candidates are any more readily accessible by ordinary, unaided perceptual means than the immaterial ‘souls’ posited by the dualist. Once again, though, given that those states are states of an intentional agent whose body *is* directly observable, humans can track the former by tracking the latter.

Beyond the object-dependence view: Differences between tracking *intentional agents* and *non intentional objects*

The evidence reviewed so far is compatible with what we called the object-dependence view of agent tracking. In the terminology of the object-file theory, the conclusion would be that object files may, on occasion, be put to use in tracking intentional agents; agent files are thus sometimes reduced to object files. However, there are reasons for thinking that, while we do track intentional agents by tracking their bodies, we should resist the stronger, deflationary view. For there appear to be important differences between tracking physical objects and tracking intentional agents.

The variety of ontological theories of personal identity

First, the distinction between object and agent tracking is required by the ontological theories of personal/self identity which are not consistent with the *thesis of the co-instantiation* expressed in section 5, which states that a human agent is co-instantiated with his/her material human body. The conflict would heighten if one replaces 'agent' by 'identical person' in the thesis of the co-instantiation. The latter version of the thesis is rejected by a number of metaphysicians for which the criteria of personal identity are not reducible to objecthood/bodiliness criteria, but pertain to other criteria such as psychological continuity (e.g., Shoemaker, 1997) or the understanding specific to shared intentionality and social folkpsychology (e.g., Ames et al., 2001; Tomasello, Carpenter, Call, Behne, & Moll, in press). For instance, Shoemaker (1997) defends a functionalist analysis of personal identity which rests on conditions of psychological continuity that persist in spite of dramatic bodily transformations such as brain transplantations or teleportation. On the credit of an object-dependence view of personal identity, it is worth noting however that the compatibility of Shoemaker's analysis with materialism remains subject to caution (van Inwagen, 1997). In addition, a number of metaphysicians defend ontological analyses of personal identity which seem consistent with the thesis of the co-instantiation for personal identity. For instance van Inwagen (1990) and Merricks (2001) uphold that each person is a material object and persists over time because this person is identical to the biological organism he/she is (*strictly*). In the present article, our general use of 'agent' instead of 'person' is an attempt to remain as neutral as possible with respect to the puzzles of personal identity. Our claim focuses on agency and can remain true with a variety of different ontologies of personal identity.

Argument related to means of direct perceptual anchoring into agency (agent indexing)

Apart from the previous ontological considerations, arguments can also be found with respect to an examination of the perceptual indexing of agents. Even though this leaves intact the thesis of body tracking (section 5), there are important differences in the properties that can trigger and maintain singular agent perceptions – i.e. the opening/indexing and maintenance of 'agent files' –, and not just in the descriptive information carried out by the singular perceptual representation. One argument supporting this claim is that specific mechanisms seem to anchor the mind on properties that are usually co-instantiated with agency, and that appear as *direct* means for perceptually keeping track of agents (i.e. physical objects that instantiate agency properties).

For instance, humans can efficiently detect and track biological motions which are specific to agents. A tradition dating from the studies by G. Johansson (1973; 1975) showed how moving organisms can be detected purely from motion information. Johansson hypothesized that people would be able to perceive the movement of the human body from just the motion of the body's *joints*. To test this hypothesis, he filmed an actor in the dark with small lights attached to his joints (ankles, knees, hips, shoulders, elbows, and wrists) so that nothing was visible except the lights. When the actor was seated motionless in a chair, observers perceived a meaningless configuration of points, rather like a constellation of stars; nobody perceived the lights to be connected to a human agent. But within fractions of a second after the actor began to move, first standing up and then walking, he was immediately and unmistakably perceived as a human agent in motion. In other further studies, other researchers have found that observers are able to discriminate between male and female walkers who have lights placed just on their ankles, knees, and hips (Cutting, Proffitt, & Kozlowski, 1978) – see also Troje (2002) and the simulation of this phenomenon at

<http://www.journalofvision.org/2/5/2/genderclass.html>.

Also, there is evidence that certain *types of motion* lead to the attribution of specific types of intentional states (e.g. Gergely, Nadasdy, Czibra, & Biro, 1995; Heider & Simmel, 1944; Jacob & Jeannerod, 2003: pp. 222-6). Heider and Simmel (1944) showed human adults an animation on a screen involving three geometric objects: a small circle, a small triangle and a large triangle moving in the vicinity of a very large non-moving square. Unlike the perception of static display of the three geometric objects, the perception of the kinetic structure of the patterns of motion of the objects conveyed psychological and even social information about the objects. Subjects used highly intentional verbs to describe the behavior of the triangles such as they 'chased', 'attacked', 'caressed' or 'comforted' the circle. In the same tradition, in reply to Atran (1998), Todd & López (1998: p. 592) describe experiments similar to the famous previous ones: 'We have explored a simple visual cue-based algorithm for judging intention from motion in just such instances (Blythe, Miller, & Todd, 1996). We had participants generate motions of two moving "bugs" on a computer screen, corresponding to simple intentional categories including pursuit, evasion, fighting, courtship, and play. Other participants were later able to categorize the intentions of the "bugs" with high accuracy from their trajectories alone. This study supports the notion that animate intention can be determined using only a few simple spatiotemporal cues (which include, from trajectory analysis, relative heading, relative distance, relative velocity, and vorticity or "loopiness"). Knowing the intention (as opposed to the general intentionality that Atran mentions) of another organism can trigger the appropriate domain-specific mechanism for response, including species-level categorization and recall of relevant traits.'

Moreover, in the visual domain, there is evidence that the visual system does not rely on the same resources for recognizing non-face objects as opposed to faces (e.g. Grill-Spector, Knouf, & Kanwisher, 2004). This might suggest that the presence of a face – an agent specific cue – can contribute to a specific file indexing, an agent-based indexing. This could be the case not only in vision, but also in auditory perception and cross-modal integration. It seems plausible, for instance, that noises that bear the specific signature of human or animal behavior trigger the opening of an agent file – i.e. we can detect and track a person just by hearing a sound that bears the specific signature of a human body (voice, footsteps). This is the case when one is detecting a sound related to the phonological part of the body (e.g. Handel, 1995).

Argument from parsimony for the direct access to agency

The empirical data drawn from the study of early motor and sensory skills seem to support the idea that there may even be two distinct kinds of files. What would be the use of such a distinction from a computational perspective – as introduced by Ballard et al. (1997) and Pylyshyn (2001)? It is often necessary to direct the focus of attention specifically toward agents, as opposed to mere non intentional objects, for instance in order to achieve collaborative actions (e.g. shared intentionality in word learning or collective sports), or to evaluate judgments involving intentional relations. If the deflation view is true, the only difference between object and agent files would be in terms of the descriptive content they encode. This would imply that there would have to be a (descriptive- or content-driven) search for intentional properties in object files, which seems computationally costly. As opposed to this approach, one can speculate that the addressing mechanism of each file was typed as *object* or as *agent*; however, such a search would be made easier from a computational perspective. For instance, one might speculate that files could be sorted by a non conceptual 'tag' signaling directly that it is a '(mere) object file' or an 'agent file'; such a tag would avoid the need to access the descriptive content of a file so as to target one of the two kinds of entity for any attentional or motor routine (e.g. routines such as 'escaping from a predator', 'seeking for help' or 'searching for a team member').

Rationality and agency

Evidence for/against rationality is also relevant for tracking intentional agents, but not objects (which, since they are not intentional agents at all, are neither rational nor irrational). Thus, e.g., if someone consistently violates obvious norms of reasoning (e.g., *modus ponens*), it becomes difficult to even make sense of them as intentional agents [e.g., Dennett (1969; 1971), Davidson (1980; 1984)]; though of course, it may be not difficult at all to continue to track their bodies. Similarly, when someone consistently exhibits markedly different patterns of behavior/reasoning in different domains

(cf. discussions of ‘compartmentalized thinking’, and/or ‘false consciousness’), even if these differences do not correlate to marked differences in features of their body, we sometimes find it necessary to suppose, in effect, that there are two agents there (less dramatically, that markedly different psychological states and traits characterize the agent in the two domains) even though there may be seamless continuity in his/her single body. This assumption is frequently made about people supposed to be ‘possessed’ as depicted in ghost movies or stories; it is manifested in a pathological way in syndromes such as the Capgras syndrome where the patient believes that his/her near ones have been replaced by substitutes occupying the same body. In short, considerations of a conspecific’s (ir-)rationality is a criterion relevant to the opening, maintenance, and updating of agent files, but it is not relevant at all to (non intentional) object files; and perceived discontinuities or dramatic differences in an agent’s intentional states themselves can lead us to ‘split’ or seriously amend the relevant agent file, even though the corresponding object file (used in tracking his/her body) remains intact, and does not undergo comparable changes in its contents.

Conclusions

Assuming a realist ontology of the material world and an object-dependent epistemology, we have examined the possibility of tracking not only physical objects but intentional agents. Our contention, referred to as the object-dependence view, is that tracking intentional agents requires the same kind of processes used in tracking physical objects. Furthermore, we postulated, in a manner consistent with the findings of experiments conducted on visual perception, that it is possible to draw a parallel between the classes of properties that characterize and produce agent and object files (indexing, preservation and encoded properties or cues). Given the similar – though *not*, we have argued, identical – anchoring, content and architecture of (non intentional) object files and agentfiles, the exact relation between them remains an open question which we have only begun to explore. However, at least the following seems plausible: (1) we perceptually and cognitively keep track of intentional agents, (2) object files and agent files are distinguishable in important ways, but nevertheless (3) agent tracking exploits the resources of object tracking by anchoring the agent file on perceptible features of agents’ bodies (thesis of body tracking).

Possible directions for future research

Our object-dependence view suggests directions for further research. A number of classical experiments have been carried out on object perception and multiple-object tracking in vision, in various conditions including conditions of target occlusions. It would be interesting to see what the results would be in analogous experiments in the case of visual multiple-agent tracking. What sorts of changes (in motion? behavior? intention?) are/are not relevant to subjects for maintaining the same agent file rather than opening a new one? Can one produce in an experimental display cases where persons-as-bodies and persons-as-agents ‘come apart’ (as when ‘the soul rises to heaven’, or ‘body-switching’ cases). Presently, the need first lies not in answering these questions but in raising them.

However, there is plenty of anecdotal evidence we could cite. For example, we have no difficulty in conceiving of ‘body-switching’ – an agent simply occupying a different body. For a ‘fictional-indexical’ descriptions of this, see Dennett (1978b); for a metaphysical conception, see the examples given by Parfit (1971) and Shoemaker (1997). This is not incompatible with our view. In fact, it might support it because the object-dependence can be exquisitely discovered in the ways one conceives of or depicts the ‘traveling souls’. For, depictions of souls leaving a particular body often take the form of *ghostly* bodies leaving the body of flesh-and-blood. This is not surprising, for how else could souls be depicted – or, we ask, how else could they be *tracked*? Is it not by depicting human bodies that painters have been representing the souls of ‘The Blessed’ in Paradise and of ‘The Damned’ in Hell?

See for instance ‘The Last Judgement’ (1431) by Fra Angelico (<http://www.abcgallery.com/A/angelico/angelico39.html>, and the details of the blessed (<http://www.abcgallery.com/A/angelico/angelico41.html>) and the damned (<http://www.abcgallery.com/A/angelico/angelico42.html>) in the picture.

We *do* have some trouble making sense of cases of fission/fusion (where one agent becomes two, or two become one – a good illustration for this could be found in the view of Lewis (1976). Here too, the cases might support our position: fission/fusion are puzzling, not merely because they involve splitting/joining minds, but because our usual means of tracking agents (via their bodies) is *frustrated* – the one-body/one-soul correlation is disrupted.

French version

-Poursuivre les objets, Poursuivre les agents

Résumé :

En admettant une ontologie réaliste du monde des objets physiques et une épistémologie de la connaissance singulière dépendante des objets, nous comparons la poursuite ('tracking') des objets dépourvus d'états mentaux avec la poursuite des agents intentionnels. Pour étudier la poursuite située des agents, nous défendons 'l'approche de la dépendance à l'égard de l'objet' qui affirme que, dans le cas le plus rudimentaire, les humains poursuivent les agents intentionnels en poursuivant leur corps. Toutefois, nous soutenons que la poursuite des agents intentionnels comme agents intentionnels requière des capacités spécifiques pour détecter et comprendre le mouvement biologique et les états intentionnels.

Acknowledgments: We would like to thank for their comments on earlier versions of this paper Erika Marchetto, Nivedita Gangopadhyay and the members of the 'agent tracking group' of CNRS Thematic School 'Reference to objects' (<http://www.objectreference.org/>). Part of this research has been funded by the Enactive Network.

References

- Ames, D. R., Knowles, E. D., Morris, M. W., Kalish, C. W., Rosati, A. D., & Gopnik, A. (2001). The social folk theorist: Insights from social and cultural psychology on the contents and contexts of folk theorizing. In B. F. Malle & L. J. Moses & D. A. Baldwin (Eds.), *Intentions and Intentionality* (pp. 307-329). Cambridge, MA: MIT Press.
- Atran, S. (1998). Folk biology and the anthropology of science: Cognitive universals and cultural particulars. *Behavioral and Brain Sciences*, 21, 547-609.
- Bach, K. (1987). *Thought and Reference*. Oxford: Clarendon Press.
- Baillargeon, R. (2001). Infants' Physical Knowledge: Of Acquired Expectations and Core Principles. In E. Dupoux (Ed.), *Language, Brain, and Cognitive Development* (pp. 341-361). Cambridge: MIT Press.
- Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. N. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, 20(4), 723-767.
- Baylis, G. C., & Driver, J. (1993). Visual attention and objects: Evidence for hierarchical coding of location. *Journal of Experimental Psychology : Human Perception and Performance*, 19, 451-470.
- Bloom, P. (2002). Mindreading, communication and the learning of names for things. *Mind & Language*, 17(1&2), 37-54.
- Bloom, P. (2004). *Descartes' Baby*. New York: Basic Books.
- Blythe, P. W., Miller, G. F., & Todd, P. M. (1996). Human simulation of adaptive behavior: Interactive studies of pursuit, evasion, courtship, fighting, and play. In P. Maes & M. J. Mataric & J.-A. Meyer & J. Pollack & S. W. Wilson (Eds.), *From Animals to Animats 4: Proceedings of the Fourth International Conference on Simulation and Adaptive Behavior*. Cambridge, MA: MIT Press.
- Campbell, J. (1994). *Past, Space and Self*. Cambridge, MA: MIT Press.
- Campbell, J. (2002). *Reference and Consciousness*. Oxford: Clarendon Press.
- Carey, S., & Xu, F. (2001). Infant's knowledge of objects: beyond object files and object tracking. *Cognition*, 80, 179-213.
- Churchland, P. S., Ramachandran, V. S., & Sejnowski, T. J. (1994). A critique of pure vision. In C. Koch & J. L. Davis (Eds.), *Large Scale Neuronal Theories of the Brain* (pp. 23-60). Cambridge, MA: MIT Press.
- Clark, A. (2000). *A Theory of Sentience*. Oxford: Clarendon Press.
- Cutting, J. E., Proffitt, D. R., & Kozlowski, L. T. (1978). A biomechanical invariant for gait perception. *Journal of Experimental Psychology: Human Perception & Performance*, 4(3), 357-372.
- Davidson, D. (1980). *Essays on Actions and Events*. Oxford: Clarendon Press, Oxford University Press.

- Davidson, D. (1984). *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press.
- Dennett, D. C. (1969). *Content and Consciousness*. London: Routledge and Kegan Paul.
- Dennett, D. C. (1971). Intentional systems. *The Journal of Philosophy*, 68(4), 87-106.
- Dennett, D. C. (1978a). *Brainstorms*. Montgomery: Bradford Books.
- Dennett, D. C. (1978b). Where am I?, *Brainstorms* (pp. 311-323). Montgomery: Bradford Books.
- Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology : General*, 119, 501-517.
- Egly, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology : General*, 123, 161-177.
- Evans, G. (1981). Understanding demonstratives. In H. Parret & J. Bouveresse (Eds.), *Meaning and Understanding* (pp. 280-303). Berlin: De Gruyter.
- Evans, G. (1982). *The Varieties of Reference*. Oxford: Oxford University Press.
- Gergely, G., Nadasdy, Z., Czibra, G., & Biro, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, 56, 361-370.
- Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience*, 7(5), 555-562.
- Handel, S. (1995). Timbre perception and auditory object identification. In B. C. J. Moore (Ed.), *Hearing* (pp. 425-461). San Diego, CA: Academic Press.
- Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology*, 57, 243-259.
- Jacob, P., & Jeannerod, M. (2003). *Ways of Seeing, The Scope and Limits of Visual Cognition*. Oxford: Oxford University Press.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, 14(210-211).
- Johansson, G. (1975). Visual motion perception. *Scientific American*, 232(6), 76-88.
- Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of Attention* (pp. 29-62). Orlando: Academic Press.
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24(2), 175-219.
- Kripke, S. (1980). *Naming and Necessity*. Cambridge: Harvard University Press.
- Lewis, D. (1976). Survival and identity. In A. Rorty (Ed.), *The Identity of Persons*. Berkeley: University of California Press.
- Malle, B. F., Moses, L. J., & Baldwin, D. A. (2001). Introduction: The significance of intentionality. In B. F. Malle & L. J. Moses & D. A. Baldwin (Eds.), *Intentions and Intentionality* (pp. 1-24). Cambridge, MA: MIT Press.
- McDowell, J. (1984). De Re Senses. *Philosophical Quarterly*, 34(136), 283-294.
- McDowell, J. (1990). Peacocke and Evans on demonstrative content. *Mind*, 99(394), 255-266.
- McDowell, J. (1998). *Meaning, Knowledge, and Reality*. Cambridge, MA: Harvard University Press.
- Merricks, T. (2001). *Objects and Persons*. Oxford: Oxford University Press.
- Millikan, R. G. (1997). Images of identity: In search of modes of presentation. *Mind*, 106(423).
- O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24(5), 939-1031.
- Palmer, S. E. (1999). *Vision Science, Photon to Phenomenology*. Cambridge, MA: MIT Press.
- Parfit, D. (1971). Personal identity. *The Philosophical Review*, 80(1), 3-27.
- Peacocke, C. (1991). Demonstrative content: a reply to John McDowell. *Mind*, 100, 123-133.
- Perry, J. (1980). A problem about continued belief. *Pacific Philosophical Quarterly*, 61, 317-332.
- Perry, J. (2001a). *Knowledge, Possibility, and Consciousness*. Cambridge, MA: MIT Press.
- Perry, J. (2001b). *Reference and Reflexivity*. Stanford: CSLI Publications.
- Pylyshyn, Z. W. (2000). Situating vision in the world. *Trends in Cognitive Sciences*, 4(5), 197-207.
- Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, 80, 127-158.
- Pylyshyn, Z. W. (2003). *Seeing and Visualizing: It's Not What You Think*. Cambridge, MA: MIT Press.
- Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, 3(3), 179-197.
- Quinton, A. (1962). The soul. *The Journal of Philosophy*, 59(15), 393-409.
- Recanati, F. (1993). *Direct Reference: From Language to Thought*. Oxford: Blackwell Publishers.
- Saiki, J. (2003). Feature binding in object-file representations of multiple moving items. *Journal of Vision*, 3, 6-21.
- Scholl, B. J. (2001). Objects and attention: the state of the art. *Cognition*, 80, 1-46.

Scholl, B. J., Pylyshyn, Z. W., & Feldman, J. (2001). What is a visual object? Evidence from target merging in multiple object tracking. *Cognition*, 80, 159-177.

Shoemaker, S. (1997). Self and substance. *Noûs*, 31(Supplement: Philosophical Perspectives, 11, Mind, Causation, and World), 283-304.

Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review*, 5(4), 644-649.

Spelke, E. S. (1990). Principles of object perception. *Cognitive Science*, 14, 29-56.

Spelke, E. S., Guthrie, G., & Van de Walle, G. (1995). The development of object perception. In S. M. Kosslyn & D. N. Osherson (Eds.), *Visual Cognition: An Invitation to Cognitive Science, Second Edition* (pp. 297-330). Cambridge, MA: MIT Press.

Strawson, P. F. (1959). *Individuals: An Essay in Descriptive Metaphysics*. London: Methuen.

Strawson, P. F. (1997). *Entity and Identity and Other Essays*. Oxford: Clarendon Press.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643-662.

Todd, P. M., & López, A. (1998). Pulling the trigger on the living kind module. *Behavioral and Brain Sciences*, 21(4), 592.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (in press). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*.

Treisman, A., Kahneman, D., & Burkell, J. (1983). Perceptual objects and the costs of filtering. *Perception & Psychophysics*, 33, 527-532.

Troje, N. F. (2002). Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *Journal of Vision*, 2, 371-387.

van Inwagen, P. (1990). *Material Beings*. Ithaca, London: Cornell University Press.

van Inwagen, P. (1997). Materialism and the psychological-continuity account of personal identity. *Noûs*, 31(Supplement: Philosophical Perspectives, 11, Mind, Causation, and World), 305-319.

vanMarle, K., & Scholl, B. J. (2003). Attentive tracking of objects versus substances. *Psychological Science*, 14(5), 498-504.

[i]A pioneering philosophical analysis of human knowledge based on the cognition of individual objects can be found in Strawson (1959). Philosophers such as Evans (1981; 1982), Perry (2001b), Campbell (1994; 2002), Clark (2000) have developed this view of cognition in the philosophy of mind and perception. Views expressed by philosophers of language such as Kripke (1980), McDowell (1984; 1990), or Recanati (1993) may also be consistent with it. Lastly, recent work in the psychology of perception by Kahneman et al. (1984; 1992), Spelke et al. (1990; 1995) and Pylyshyn (2003) may provide empirical ground for an understanding of the cognition of physical objects.

[ii]Compare the following assertions. According to Evans on Russell's Principle, "The principle is that a subject cannot make a judgment about something unless he knows which object his judgment is about. (...) [Evans] shall suppose that the knowledge which it requires is what might be called *discriminative knowledge*: the subject must have a capacity to distinguish the object of his judgment from all other things. (...) we have the idea of certain sufficient conditions for being able to discriminate an object from all other things: for example, when one is perceiving it at the present time; when one is recognize it if presented with it; and when one knows distinguishing facts about it" (Evans, 1982: 89). According to Strawson (1997: 22-23), the philosophical slogan 'no entity without identity' can be interpreted as an "admirable maxim" which states, "There is nothing you can sensibly talk about without knowing, at least in principle, how it might be identified." According to McDowell (1990: 256-7) "The essential background for Evans's account of demonstrative mode of presentation is the principle that to entertain a thought, one must know what it would be for the thought to be true. It is an application of this principle that if there is some particular object which must be a certain way for the thought to be true, the subject must know which object it is. This requirement – the 'know which' requirement – is what Evans calls 'Russell's Principle'." Finally, according to Pylyshyn, "(...) if the visual system is to do something concerning some visual object, it must in some sense know *which* object it is doing it *to*" (Pylyshyn, 2003: 201).

[iii]For object files in thought and language, see e.g., Perry (1980; 2001a; 2001b), Bach (1987), Recanati (1993); for object files in vision, see e.g. Kahneman et al. (1984; 1992), Pylyshyn (2003), Palmer (1999), Saiki (2003). Interesting skeptical remarks can be found in Millikan (1997).

[iv]Here, we do not discuss directly the specific cases of self-identification and personal identity (e.g., Parfit, 1971; Quinton, 1962; van Inwagen, 1997) – see section 6.

[v]Cf. e.g. Strawson (1959: 87-116), van Inwagen (1990), Merricks (2001).

[vi]For instance : ‘(...) there is exactly *one* thing where we truly believe there to be a human person and a human organism (and a human body). Obviously, this implies that the person is identical with the organism (is identical with the body).’ Merricks (2001: 86).

Discussion

▼Agent Tracking- a too simplistic account?

Nivedita Gangopadhyay

Mar 24, 2005 9:34 UT

While agreeing with the general idea expounded in this paper, I think agent tracking as it has been described here could be a slightly too simplistic account. When one is tracking an agent one of the vital question concerns the task at hand- “What” is it that is being tracked. It is not sufficient to say that the entity being tracked is an intentional agent. Intentional agents, whether human or non-human, could be tracked merely as an agent performing a particular task or they could be tracked as “persons” i.e. as being endowed with a unique personality. For example suppose a person is viewing a game of soccer in a stadium. From where the viewer is it is not possible to distinguish the players as “persons” i.e. as some unique individual and suppose the viewer is not so familiar with the names, features etc. of the players. Under such circumstance the viewer can certainly follow the game, without tracking the players as “persons”, by merely tracking them as intentional agents who are performing certain actions with a ball. Such tracking of intentional agents is nearly similar to the tracking of mere physical objects. Hence in this case the “agent file” functions in more or less the same way as an “object file”. This supports the object-dependent view. But intentional agents can also be tracked as “persons” and here the task would be cognitively more demanding and not the same as the tracking of agents as merely performing some actions. For example- two very close child-hood friends accidentally meet after a gap of many years during which they had no contact what-so-ever with each other. Both of them have now grown up and it is highly possible that their bodies, including their faces, have undergone significant changes. Yet is it absolutely impossible to imagine that they can still recognize each other? In such cases the criteria for uniqueness and identity seem to include something more than just the physical object-dependent identity. There seems to be something stored in the “agent file” that indicates the uniqueness of the “person” despite physical appearances. Also in cases of Multiple Personality Disorder it is not possible to track the different “persons” or personalities by tracking the physical body. Hence it seems that “agent file” could consist of two distinct components- a) “body file”- agent as body performing some action and b) “person file”- agent as person performing some action. “Body file” is closely similar and in certain cases even completely reducible to “object file” whereas “person file” constitutes a distinct mechanism that helps us track intentional agents as persons. Experiments such as the one performed by Simons and Levin indicate that “agent file” could indeed be comprised of two such distinct components as in such cases the subject tracked the speaker simply as a human shape (“body file”) performing some action. Had she tracked the human shape as a “person” (“person file”) she would have immediately noticed the change.

▼Reply to Nivedita Gangopadhyay

Maria Rossi

Apr 10, 2005 21:40 UT

Thank you very much, Nivedita, for this comment and for introducing these interesting examples. I will reply on the basis of my own understanding of the logic of the object-dependence view. Your position seems to be the following: (A) you agree that the object-dependence view accounts for basic perceptual-motor interactions (e.g. watching individuals/agents playing sport, navigating in a crowd etc.), however (B) the object-dependence view cannot account for more ‘complex’ inter-PERSONal interactions in which “the criteria for uniqueness and identity include something more than just the physical object-dependent identity”.

I do not consider that (B) is an objection for our analysis because you have introduced additional puzzles about the identity conditions of 'persons' and our object-dependent view relates primarily to agents instead of person (cf. the end of the paper and the distinction of the object-dependence view with the deflationist view about agent tracking). We are in agreement to state that identity conditions of 'persons' might differ from identity conditions of '(bodily and embodied) agents' – and that the metaphysics of 'personhood' is a debatable topic. In addition, although it is outside the direct scope of the paper, I find it plausible to consider that 'personhood' may be dependent on cultural and social phenomena in a way that would make it dependent not only on bodily or psychological internal continuity but also on emotional and social relations.

Let me consider your first example, the old-friend recognition case. Even though alternative metaphysical options remain possible, the case can be treated as being a standard case of re-identification. Although the two bodies have changed, each is the SAME body according to the framework of the object-dependence view for (1a) it follows the standard organizations rules of biological organisms and their historical transformations (1b) you assume that the two persons are unique namely because although their organism has changed as a function of time, each continuously changing organism has followed a unique and traceable spatio-temporal path during its own history. Moreover most of the social or technological evidence you can use to judge that this agent is the same agent as the person you knew by acquaintance long before may rely on the anchoring provided by the object-dependence view (cf. for instance my note below on biometrics). Please consider your ID card and the ID cards with fingerprints and the causal/historical theories of proper names (Kripke, Donnellan, Perry etc): proper names may originate in something like the 'baptism' of a physical object/agent which is the starting point of a social mechanism of automatic tracking of a single and continuing physical object via the proper name and the ID card with the proper name written on it.

Finally, I like the idea you mentioned with respect to the result of Simon and Levin. One should account for these cases in terms of a tracking dynamics based on alternating between using different object, agent and person files. This would account for the incremental characteristics of perceptual identification. It remains highly speculative.

▼Biometrics and the object-dependence view

Maria Rossi

Apr 10, 2005 21:48 UT

I would like to point toward the importance of the research on biometrics to understand the implication of the discussion about the object-dependence view. Here is a quote, according to the consortium of biometrics '<http://www.biometrics.org/>': "Biometrics are automated methods of recognizing a person based on a physiological or behavioral characteristic. Among the features measured are: face, fingerprints, hand geometry, handwriting, iris, retinal, vein, and voice. Biometric technologies are becoming the foundation of an extensive array of highly secure identification and personal verification solutions. As the level of security breaches and transaction fraud increases, the need for highly secure identification and personal verification technologies is becoming apparent. Biometric-based solutions are able to provide for confidential financial transactions and personal data privacy. The need for biometrics can be found in federal, state and local governments, in the military, and in commercial applications. Enterprise-wide network security infrastructures, government IDs, secure electronic banking, investing and other financial transactions, retail sales, law enforcement, and health and social services are already benefiting from these technologies." (quote from <http://www.biometrics.org/html/introduction.html>, available on April 09 2005).

Here are some basic comments:

- Biometrics is a group of knowledge and applications for (helping in) keeping track of the identity of unique human intentional agents (with prosthetic devices). 'Face, fingerprints, hand geometry, handwriting, iris, retinal, vein, and voice' are bodily properties (dependent on causal properties of a single organism). The uniqueness problem for agency tracking is basic for those

technologies – hence uniqueness is not only the concern of philosophers, it is a fundamental topic for contemporary sciences and societies.

- 'Transcendental' argument: Ethical and political considerations aside, the possibility of biometrics is (implicitly) based on an object-dependent ontology and epistemology. The efficiency of biometrics (e.g. fingerprint reading) shows how robust the object-dependent ontology and epistemology are. If you are concerned about the ethics of biometrics, you should first acknowledge the object-dependence view.

- The author of the quote above writes 'biometrics are automated methods of recognizing a person based on a physiological or behavioral characteristic'. This sort of deflationist view of personhood is against the object-dependence view upheld by Patrick and me. According to the careful terminology of the object-dependence view, one should replace 'person' by something like 'human agent' so as to approach neutrality.

▼A few remarks

David NICOLAS

Mar 29, 2005 17:16 UT

A simple, deflationary thesis would go like this. Agents are a specific case of material objects: material objects with specific properties. Hence « agent files » are simply « object files » with additional information: information about properties specific to agents. Bullot & Rysiew argue against theses of this kind by presenting three types of experimental evidence and an argument of computational efficiency:

1) « Humans can efficiently detect and track biological motions which are specific to agents. »
2) « Certain types of motion lead to the attribution of specific types of intentional states » (e.g. triangles perceived as being chased, attacked, caressed, etc.). 3) « The visual system does not rely on the same resources for recognizing non-face objects as opposed to faces. » 4) If agent files are tagged as agent files (as opposed to object files), then there is « no need to access the descriptive content of a file so as to target one of the two kinds of entity for any attentional or motor routine. » It is therefore less costly to have files tagged as agent files than not. They conclude that agent files are not merely object files with additional information. They are files marked as dedicated to the tracking of agents. But of course, these files include information about the agents' bodies, and the tracking of agents is often made by tracking their bodies. I have a lot of sympathy for the authors' position. Here are a few questions and remarks. Would they agree with the following, more general picture? Human beings have files about individual things. Any type of information can enter into such files. Indeed, « individual files » can be about many different types of individuals, like material objects, sounds, human beings, gods, electrons, etc. However, among these files, at least one type of file is distinguished because of its importance: files about agents are tagged as such for easier access and use. This kind of picture raises several questions: which types of files are distinguished for special use, why exactly, and what does this « distinguishing » amount to? For instance, it is conceivable that, instead of « agent files », there be files for conspecifics on the one hand, and files for members of other species on the other. It would therefore be interesting if the authors could bring to bear some evolutionary arguments in favor of their thesis. It does seem plausible that evolution selected specific mechanisms for dealing with agents, be they conspecifics or member of other species. Conspecifics will often be mates or competitors for sexual reproduction. While members of other species will often be preys or predators. So accurate monitoring of both conspecifics and members of other species is particularly important for survival. But can more detail be given about how exactly an evolutionary story would go? For instance, would it types of files that would get selected, and if so, which? By the way, which organisms can be said to have « object files » but no « agent files »? Povinelli argues that much of primates' behavior can be explained with simple, non-intentional mechanisms. Would these non-intentional mechanisms involve only « object files »? Nota bene: Some of the experimental evidence adduced above does not necessarily imply the existence of intentional information in the « agent files ». Thus, recognition of certain patterns as corresponding to human motion and recognition of human faces does not, by itself, require the existence of intentional information in the files.

▼**Reply (part 1) to question (1): The hypothesis of the Universal File for a Token (UFT)**

Maria Rossi

Apr 18, 2005 10:37 UT

Thank you very much, David, for this comment and these interesting questions. Let me focus on your first question about the following hypothesis: “Human beings have files about individual things. Any type of information can enter into such files. Indeed, ‘individual files’ can be about many different types of individuals, like material objects, sounds, human beings, gods, electrons, etc.” Let me name this hypothesis the ‘hypothesis of the Universal File for a Token (UFT)’, and formulate it in the following way:

UFT Hypothesis: Each human mind obtains information and singular knowledge about an individual (object, agent, event, sound, fictional entity etc.) because it uses as standard a Universal File for a Token (UFT) (as opposed to an object file, agent file, event file etc.).

Here are some sketchy and skeptical remarks about this hypothesis. The UFT hypothesis is a speculative but interesting version of the file concept/hypothesis, which is found in some conceptual analyses. It postulates a file which would be about any kind of unique entity – a sort of all-purpose and ecumenical uniqueness mechanism. (UFT might be akin to the theoretical construct of “ ‘chunks’ of short-term memory” in early cognitive psychology (Miller, 1956; Simon, 1975) or some uses of the notion of ‘pointer’ in the same contexts.) Intuitively, it seems plausible that one can entertain singular thoughts about many kinds of entity, so why not suppose that we use only such a kind of abstract file? Authors in semantics might even argue that this kind of ability is required by the linguistic processing of indexical and anaphoric terms.

However, as stated, this hypothesis is not so clear. If you admit and use UFT in your theory, you have to deal with a very (too?) abstract notion of ‘file’, which is difficult to assess and to map onto a particular skill or cognitive ability. In fact, the notion of UFT file seems to lack a clear-cut explanatory goal. If its explanatory goal is to give a unified and systematic account of how one keeps track of things in perception, memory, reasoning and conception (Perry, 2001: pp. 50-51) – it is in need of complement so as to make explicit the relationships between perception, memory, reasoning and conception. In this case, Perry (2001) seems right consider this notion as a sheer metaphor. If it just points toward a fundamental problem at the basis of your ability to obtain singular knowledge, it may not have explanatory power.

▼**Reply (part 2) to question (1): The hypothesis of the Universal File for a Token (UFT)**

Maria Rossi

Apr 18, 2005 10:41 UT

As opposed to a UFT, the notion of a ‘visual object file’ in vision/perception seems to have a much more determinate explanatory goal: that is, basically, to contribute to explaining how one can track ‘identical’ tokens within the visual/perceptual field. According to the version of the object-dependence view which appears to me as plausible, the human mind uses an attentional system which is functionally equivalent to a ‘perceptual/visual object file’ because its routine in perception is to re-identify distal tokens of material/physical objects, and that is what the theoretical notion of perceptual/visual object files attempts to clarify. If a perceptual object-file has the function to facilitate the re-identification of object, a UFT file cannot be restricted to this function. For instance, you mentioned the possibility to use a UFT file for keeping track of non standard objects such as a sound and a god. However, can we keep track of them in the same sense that we keep track of physical objects? For instance, it may be difficult or impossible to re-identify token sounds (they may be unique events) – although you can re-identify and track the spatio-temporal path of the object which produces the sound (this phone, this man etc.), which are standard physical objects for perceptual tracking. Finally, if a UFT file is just a chunk in memory, you can call it ‘file’ but it can also be called ‘chunk’ or ‘pointer’ or ‘unit of working memory’, or even ‘deictic variable’ – and this could actually be a terminological question.

References

- Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.
- Perry, J. (2001). *Knowledge, Possibility, and Consciousness*. Cambridge, MA: MIT Press.
- Simon, H. A. (1975). The functional equivalence of problem solving skills. *Cognitive Psychology*, 7(2), 268-288.