Chapter 10

Ruth Millikan

HAL Id: ijn_00000384

https://jeannicod.ccsd.cnrs.fr/ijn_00000384

Submitted on 12 Sep 2003

CHAPTER TEN
TRACKING THE DOMAINS OF CONVENTIONAL SIGNS

I want now to argue that just as no intentional representations of retinal images intervene between physical objects and the seeing of those objects, no representations of speaker intentions in speaking need intervene between world affairs spoken of by speakers and hearers' understandings of those words.[1]

When conventional signs are true or satisfied and when this has come about in the normal way, conventional signs are locally recurrent natural signs. True, tokens of the same conventional sign may have diverse etiologies, through different people's perceptual systems and cognitive systems. They differ from more ordinary recurrent natural signs in that there will usually be numerous different kinds of causal paths to their production, depending on the ways that different speakers have managed to translate diverse prior natural signs into a uniform medium of thought and expression. But there are reasons why the same linguistic form continues to coincide with the same kind of represented affair over a certain domain --it is no accident-- and we have decided to take that as the primary criterion for a locally recurrent sign (Chapter Six). Assuming that this step in the production of a conventional sign has been accomplished through normal mechanisms --the speaker is not confused, does not lie, and so forth-- then reading a conventional sign is mainly a matter of tracking its natural domain, that is, determining what reproducing family it has been copied from. Compare tracking the bird species from which a particular e-track was derived (Chapter Three).

Defining descriptions (Chapter Four) need to be mentioned here separately. "The dog," for example, is not a conventional sign, say, of Fido. But it can be used as a cooperative intentional sign of Fido. To be an intentional sign it must be produced by a speaker whose ways of speaking are cooperatively tuned to fit the capacities of hearers whose ways of interpreting are, in turn, tuned ro fit the sign-making dispositions of speakers (Chapter Eight). I am claiming that hearers' capacities to understand speakers are continuous with their capacities to understand natural signs. Normally the hearer will be able to understand the speaker's token of "the dog" as referring to Fido only if this token, given the domain it is in, is a natural sign of Fido. But "the dog" need not be a locally recurring sign of Fido. Perhaps it has never been used to refer to Fido before and never will be again. Recall the discussion of tracking Scamper the squirrel in Chapter Four. You track Scamper by recognizing squirrel signs which in turn are signs of Chipper, given that they occur in a certain domain. A general capacity to track the presence of squirrel becomes a specific capacity to track Scamper when restricted to a certain domain. Indeed, although squirrels squeak only under special circumstances and it could be that Scamper does so only once in his life, still a certain kind of squeak may be a natural sign of squirrel and hence, in a more restricted domain, of Scamper. Similarly, "the dog" is a recurrent local sign, in a fairly wide domain, of dogs, and hence, in a much more restricted domain, may turn up as a natural sign of Fido. Fido is manifesting himself in the sort of way that dogs commonly do through the medium of an English speaker.

In cases where the speaker is wrong or confused or lies, and so forth, normally a linguistic

sign will still be a natural sign, but not a natural sign of the same thing of which it is an intentional sign. Perhaps it is a sign of the speaker's mental state, of what she believes, or of what it is her intent to induce the hearer to believe, and so forth. It is not, however, a conventional or intentional sign of any of these. On the other hand, a conventional linguistic sign that is true for normal reasons is a natural sign also of what the speaker believes, of what it is the speaker's purpose to communicate about, and so forth, but it is not an intentional sign of any of these. Interpreting language signs as natural signs of speaker beliefs and intentions is a rather sophisticated activity. It requires the interpreter to go well beyond tracking and recognizing the memetic family of the sign and what the sign would normally represent. Especially, the hearer must possess concepts of mental states. Most children don't acquire these before age four or five, by which time they have been using and understanding language for several years. The translation of conventional language forms directly into belief or intentional action does not require concepts of mental states. It is a fairly simple, straightforward affair.

Optometrists ask their patients to look through a stereoscope and tell, for example, whether the dog has jumped all the way through the hoop, half way through the hoop or, say, whether just its front legs have gone through. Children believe they really are looking at a picture of a dog jumping through a hoop. They assume they are looking at whatever it is they apparently see. For them, "seeing is believing." Adults, on the other hand, may be aware that they are not looking at a single picture at all, but looking at a picture of a dog with one eye and a picture of a hoop with the other. They may have no tendency at all to believe they are actually seeing a picture of a dog jumping through a hoop. In this situation, for them, seeing is not believing. Indeed, if they have had a lot of experience of the right kind, they may be able to tell, straight off, by looking, that their left eye is stronger than their right eye, or vice versa. It does not follow that when looking at a picture of a dog jumping through a hoop in the normal case, they first judge that it looks as if a dog were jumping through a hoop, then judge that they are not looking through a stereoscope, and that there are no other unusual influences disturbing the light that arrives at their eyes, finally concluding that they are actually seeing a picture of a dog jumping through a hoop. Similarly, the fact that there are situations in which understanding what someone says does not result in believing it, indeed, in which <u>by</u> understanding what is said one tells straight off that the speaker is confused, wrong, lying, or joking, does not imply that in the normal case one first judges that the speaker is not confused, lying, or joking, and then moves to believing what the speaker says. Normally, hearing that p is believing that p --possibly not in the purely statistical sense of "normally" (although I would argue that that is so too), but believing that p is the default. It is what happens when nothing intrudes.[2]

It does not follow either, however, that understanding conventional language forms is merely "decoding" (compare Sperber and Wilson 1986, Origgi and Sperber 2002). Only signs that have been copied from one another to mean the same coexist within the same conventional sign domain. Conventional signs have domains that must be tracked, just as the domains of local natural signs do. Or if they don't have to be tracked, this is purely contingent. It is a lucky fact if the

interpreter never happens to stumble into another domain where the same surface form means something different, or never stumbles on a speaker using the form who is misinformed, lying or joking.

There is no such thing, for example, as a pattern of sound that is intrinsically dedicated to one function. Just as nothing can prevent some new disease from developing that causes spots that look just like measles, nothing can prevent conventions arising from diverse sources that accidentally cross, producing homonyms or equivocal words, phrases or sentences. Putting this another way, there can be no such thing as a convention that determines when a sound pattern has been produced in accordance with a convention, or in accordance with this convention rather than that where two conventions cross. That a token results from convention-following, from reproduction, rather than from some other source cannot itself be a matter of convention, nor can it be a matter of convention that it came from this source rather than that (Millikan 1998, 2003).

The division between semantics and pragmatics has sometimes been interpreted as resting on whether the context of the sign has to be considered in determining the meaning. But if what is meant by "the sign" is merely the sound or the written pattern, clearly its context is always relevant. It must be recognized as part of a certain language, for example, or as one from among several homonyms, or as representing one from among various familiar senses of a certain word, and so forth. The domain from which it comes may be recognized by the country in which it is spoken (hearing "surgery" or "bonnet" --am I in England or America?) or the language of the words that surround it, or the accent in which it is spoken, or by knowing the background of the person speaking, or by knowing the meanings of surrounding words such that it fits with them to make a complete representation, or makes a representation of something that might come to a speaker's mind in this context, and so forth. "Hit me" means one thing when playing blackjack and another when being instructed in boxing. "Break a leg!" conventionally means good luck but, again, only when copied within a certain kind of context from prior such uses in similar contexts. Similarly, where tokens of sentences identical on the surface are conventionally used to perform different speech acts, as when the indicative mood is used for giving orders in the army or the interrogative mood used for making requests, though these usages are thoroughly conventional, they can only be tracked through context.

A particularly clear example to show the continuity of language-sign tracking with natural-sign tracking is the tracking of proper names. How do you know which John is meant when somebody says "John"? That is, how do you know which of the dozens of memetic families of referring "John" tokens this particular token has been copied from? Just as you may have to take into account what part of the country or the wood you are in to recognize quail tracks, you may have to take into account with whom the speaker is acquainted, or where the speaker has just been, or what general domain he or she has in mental focus, say, family or work or the last hunting trip, to know which "John" domain this token comes from.

Does taking into account the domain the speaker has in mental focus require thinking about the speaker's mind? The tradition of pragmatics following Paul Grice's work assumes that

understanding a communication always involves that the hearer should first "understand what the speaker intends to communicate," and that this implies that the hearer knows, say, that the speaker intends her to believe such and such, or intends her to do such and such.[3] One upshot of this tradition has been a great deal of interest in the question which if any of the non-human animals possess a "theory of mind," for it is assumed by many that this would be necessary for human-style communication. But if you read the phrase "understand what the speaker intends to communicate" transparently, it does not imply that the hearer thinks about the mind of the speaker at all. It describes the content of the hearer's understanding, but not necessarily by using a description of that content that the hearer herself would employ or understand (Chapter Seven). It means, merely, that the hearer thinks <u>the</u> <u>same</u> <u>content</u> that the speaker purposefully communicates. Let me try to make plausible that this is the correct reading of "understand what the speaker intends to communicate" in this context. That is, no thoughts about other people's minds are necessary in order to grasp their meanings during ordinary communication using conventional forms in conventional ways.[4]

Recall the discussion in Chapter Nine of what is involved when one sees things in a mirror. How do you reidentify what you see through a mirror, for example, how do you identify the car you saw through the rear view mirror a moment ago with the car you now see passing on your left and moving in front of you. You may identify it in part by its style and color, of course. But different individual cars may have the same style and color. This style and color will represent reappearance of the same car for you only in so far as you are also tracking this car. Your ability to reidentify depends on the fact that the car projects itself in a continuous path through space and time and that you do the same, a way that makes the relation between the two of you over a short period of time predictable within limits. Not that you couldn't make a mistake. One blue Ford seen in the rearview might, in principle, suddenly be passed by an indistinguishable blue Ford which now passes you without your noticing the switch. How one Normally manages to reidentify is the point; infallibility is always irrelevant. The point is that your ability to track, over space and time, the domain in which presence of a blue Ford will signify <u>that</u> <u>same</u> <u>car</u> <u>again</u> does not require you to have any understanding of how mirrors work, of light energy, of principles of reflectance, and so forth. You don't have to know anything <u>about</u> mirrors. You only have to recognize what place, relative to yourself, you are seeing when you look in the rearview, and be able to continue tracking places roughly continuous with that same place as you move your glance through the side window then through the windshield.

Similarly, how do you tell what you are seeing through a pair of binoculars? Which of the birds lined up on the telephone wire just seen way over there with the naked eye is the one you are now seeing through the binoculars? Perhaps you tell a little bit by what it looks like, but you wouldn't be using the binoculars if you already knew what it looked like in full detail. You can tell there is probably a bird there without the binoculars and with the binoculars you see something that is definitely a bird. But which bird? One that is over there in the same direction the binoculars are pointing. That is the beginning of your tracking. But there seem to be a number of birds over there

lined up on the telephone wire and you can't tell at exactly which one the binoculars are pointing. So you look at the context that surrounds the bird. It is just in front of the Y on the lower branch of the small maple tree just behind the wire. If you can also see that Y without the binoculars, you may be able correctly to track that bird from seeing it with the naked eye to seeing it with the binoculars. Or perhaps you can see that it is the third bird over from the left and you can see which is third also without the binoculars. The bird's general properties, the bird's general direction in relation to you, the direction the binoculars are pointing, and certain features of the bird's context all combine to enable tracking hence (re)identification. And if you now move the binoculars slightly to the right, you may identify the next bird on the right a bit more easily than you did the first bird, by the relation it bears to the first. Again, you don't have to know anything at all about the principles behind the operation of lenses, not even that there ARE lenses in the binoculars, to be able to do all this.

Now consider the phenomenon --clearly a very important one--of joint looking. Very early, infants follow the gaze of another person trying to communicate with them. Soon they also use and comprehend pointing and other showing gestures quite spontaneously. In doing so they are following where the other person's attention is directed. This is like observing in which direction the binoculars are pointed.[5] When talking to an infant, where an adult's attention is directed is likely to be where the subject of the adult's conversation is to be found. This helps the infant correctly to identify what she sees with what she is hearing about. What kitties look like when projected through the medium of ambient light is one thing; what they sound like when projected through the medium of another person's speech is different. But the infant learns to recognize these signs as of the same thing. Kitties can be identified and learned about in various ways, through various media (Millikan 2000, especially Chapter 6). The infant learns what kitties look like in various postures, what they feel like, the sounds they make and what they sound like through language. There seems no reason why this last would require that the infant employ a theory of mind or concepts of mental states. Why would the infant need to understand the innards of minds any more than it will need to understand the innards of binoculars?

Now consider how you might identify what you see in a photograph or a home video if you can't recognize it straight off by its appearance. If you know where the roll of film was taken, you will know something about the domain from which the natural signs it contains emanated, and this may be how you recognize that, yes, that is the Ely Cathedral. Or since the photo falls between the picture of Uncle Robert and the picture of Aunt Sally, it must be an angle on their house in Little Falls. Or, not knowing in advance where the roll was taken, you may recognize baby Willie by recognizing older sister Jane who holds him up, and older brother Tom squinting there in the back. Similarly, watching a home video you identify many of the things you see by their context and by space-time continuities over short periods, as when reidentifying the blue Ford seen in your rear view mirror a moment ago. You identify by recognizing the general domain on which the camera was focusing and/or by first recognizing some of the things in the domain directly. Again, this requires no knowledge and no thoughts about what makes cameras work or what is inside them.

Similarly, I may correctly read the volt meter as telling the voltage of my battery by knowing that it is my battery to which it is connected but without having any interest in what is inside volt meters or how they work. These acts of identification require, merely, knowing how to track various kinds of sign domains.

In exactly similar manner, you recognize which John it is that is manifested through a token of the word "John," by knowing with whom the speaker is acquainted, or where the speaker has just been, or what general domain the speaker is in the middle of talking about, hence on what domain his inner binoculars are focused, what domain he is drawing verbal pictures of for you. And it is by tracking rather than by thinking of speaker beliefs and intentions that you recognize the domains on which speakers are focusing when they use brief definite descriptions such as "the dog," "the boy," "the table," "the lake" and so forth. Knowing on what domains they are focusing, you may know which individual boys, dogs and so forth, are the ones for which occurrence of the properties mentioned in the descriptions are natural signs. That is also how you recognize the domains of quantifiers, as in "Then everyone went out to see the sunrise" (all the people? --<u>which</u> people?) and in "Some people were complaining about the food" (some of the people? --some of <u>which</u> people?). Assuming that yours and the speakers's perceptual/cognitive binoculars are similarly constructed, tending to distinguish the same figures and ground when focusing on the same domain, tracking is also the way you recognize which pairing relation is being expressed by the possessive, for example, whether "John's book" is the one he wrote, or the one he owns, or the one he is looking at or carrying, or the one he has nominated for the Pulitzer prize. And that is how you know from hearing just the sentence "John is too small" for what John is too small. It is in this

sort of manner that you interpret what the speaker's words refer to --not "what the speaker has in mind," for there is no need for you to think of that which is meant under a mental description. No "theory of mind," no representation of the speaker's beliefs and intentions, is required for this.

Understanding language is seeing the world through the cognitive systems of another person who has learned, been trained, been calibrated, to make manifest in a uniform way, things in the world on which she focuses. To know what is manifested through the conventional speech of another, one may have to know on what this human instrument is focused, what it is currently wired up to. But one needn't know anything about its insides.

---

1. My target, of course, is the contemporary neo-Gricean school of pragmatics.

2. For a careful defense of this position and a fuller discussion of the theory that communicative speaking involves having "Gricean intentions," see (Millikan 1984) Chapter 3.

3. ...or intends her to believe he intends her to believe such and such or to do such and such...and so forth....

4. I believe this is largely true for nonconventional uses, such as metaphors and other figures of speech, Gricean implicatures and so forth as well. The argument for that appears in (Millikan forthcoming) as an appendix to the Chapter "Purposes and Cross Purposes."

5. According to Gomez (1991), arguing that there is no evidence that Gorillas have a theory of mind, gorillas reared with humans understand that looking at another's eyes is a means of controlling useful causal contact. They check to see that a human is attending to the same thing that they are, but for them this is merely part of a causal link in an interactional process.