

THEORETICAL ESSAY

COMMUNICATING WITHOUT CONVENTIONS: THE CO- OPERATION MODEL

Michael TOMASELLO  

Duke University | Max Planck Institute of Evolutionary Anthropology

ABSTRACT

For obvious and very good reasons the study of human communication is dominated by the study of language. But from a psychological point of view, the basic structure of human communication – how it works pragmatically in terms of the intentions and inferences involved – is totally independent of language. The most important data here are acts of human communication that do not employ conventions. In situations in which language is for some reason not an option, people often produce spontaneous, non-conventionalized gestures, including most prominently pointing (deictic gestures) and pantomiming (iconic gestures). These gestures are universal among humans and unique to the species, and in human evolution they almost certainly preceded conventional communication, either signed or vocal. For prelinguistic infants to communicate effectively via pointing and pantomiming, they must already possess species-unique and very powerful skills and motivations for shared intentionality as pragmatic infrastructure. Conventional communication is then built on top of this infrastructure – or so I will argue.

RESUMO

Por razões óbvias e muito boas, o estudo da comunicação humana é predominantemente orientado pelo estudo da linguagem. Mas do ponto de vista psicológico, a estrutura básica da comunicação humana – como ela funciona pragmaticamente em termos das intenções e inferências envolvidas – é totalmente independente da linguagem. Os dados mais importantes aqui são atos de comunicação humana que não envolvem



OPEN ACCESS

EDITORS

– Miguel Oliveira, Jr. (UFAL)
– René Almeida (UFS)

REVIEWERS

– Marianne Cavalcante (UFPB)
– Martha Shiro (UCV)

DATES

– Recebido: 09/06/2020
– Aceito: 07/17/2020
– Publicado: 01/13/2021

HOW TO CITE

TOMASELLO, Michael (2021).
Communicating Without Conventions:
The Co-Operation Model. *Cadernos de
Linguística*, v. 2, n. 1, p. 01-27.

convenções. Em situações em que a linguagem, por alguma razão, não é uma opção, as pessoas comumente produzem gestos espontâneos e não convencionais, entre os quais os mais proeminentes são o gesto de apontar (gestos deícticos) e as pantomimas (gestos icônicos). Estes gestos são universais entre os humanos e exclusivos da espécie, e na evolução humana quase certamente precederam a comunicação convencional, seja ela sinalizada ou oral. Para que as crianças em fase anterior à aquisição da linguagem se comuniquem efetivamente através do apontar e de pantomimas, elas já devem possuir habilidades e motivações únicas e muito poderosas para a intencionalidade compartilhada como infraestrutura pragmática. A comunicação convencional é então construída em cima desta infraestrutura - é o que vou argumentar aqui.

KEYWORDS

Gestures; Communication; Language.

PALAVRAS-CHAVE

Gestos; Comunicação; Linguagem.

INTRODUCTION

For obvious and very good reasons the study of human communication is dominated by the study of language. But from a psychological point of view, the basic structure of human communication – how it works in terms of the intentions and inferences involved – is totally independent of language. Pragmatics should be the discipline that studies these basic psychological processes directly, but, unfortunately, traditional pragmatics studies them only as they supplement and support linguistic communication. Traditional pragmatics takes a top-down approach, an approach in which pragmatics gets the leftovers.

But there is also a bottom-up approach, a pragmatics first approach. The most important data here are acts of human communication that do not employ conventions, that is, that do not have a conventional meaning but only a meaning in the moment. Such acts are ubiquitous in situations when language is for some reason not an option, for example, in a foreign land when one does not speak the language, or even in one's own land when in a noisy environment or across a crowded room. In such situations people often produce spontaneous, non-conventionalized gestures, including most prominently pointing (deictic gestures) and pantomiming (iconic gestures). Such gestures may, of course, be used as supplements to spoken language, or they may even themselves become conventionalized or even systematized into a sign language. But for current purposes the key phenomena are non-conventional gestures, used as whole communicative acts on their own without language – what I call natural gestures. With natural gestures, there is no issue of linguistic or conventional meaning, but only the communicator's intended meaning in the moment. With natural gestures, the veil is lifted and we see pragmatics in the raw (Tomasello, 2008).

The natural gestures of pointing and pantomiming are universal among humans and unique to the species (with one small but interesting exception to be discussed below), and in human evolution they almost certainly preceded conventional communication, either signed or vocal. We may infer this both from logical considerations and from observations of prelinguistic infants (and great apes) communicating gesturally. For prelinguistic humans to communicate effectively via pointing and pantomiming – in either phylogeny or ontogeny – they must already possess species-unique and very powerful skills and motivations for shared intentionality as pragmatic infrastructure. Conventional communication is then built on top of this infrastructure – or so I will argue.

1. NATURAL GESTURES AS COMPLETE COMMUNICATIVE ACTS

Wittgenstein famously asks, "Point to a piece of paper. And now point to its shape—now to its color—now to its number. . . . How did you do it?" (1953, p.13) How, indeed? The finger and its spatial orientation are the same in all cases. How is the recipient supposed to distinguish? And assuming it is somehow established that I am pointing to the piece of paper as a whole, why am I doing that? To suggest that you fetch and dispose of that rubbish on the floor? Or to point to the valuable document you just dropped? Or to point out to you that our child has left her homework on the floor again, as she was warned not to, and so (by prior agreement) it is your turn to find and scold her.

The possibilities are endless, of course, and, of course, they are due to "context". But how does that work? How does the act of protruding a finger in a given direction communicate so richly in some contexts? The Gricean maxims are mostly of no help here, the broad categories of speech act theory help only a little, and the formal semantics/pragmatics of indexicals has basically nothing useful to say at all. As far as I can tell, in asking where meaning comes from in such cases, the only relevant theory is relevance theory (Sperber & Wilson, 1986). But relevance theorists never develop the point that if such acts can communicate so richly, then most theories of semantics and pragmatics in modern linguistics and philosophy (e.g., those based on truth conditions) are totally inadequate, if not totally wrong-headed.

My claim here, following Tomasello (2008), is that the semantics and pragmatics of linguistic communication are secondary phenomena, parasitic on already being able to communicate richly without any conventions at all. Stated historically, creating communicative conventions with semantic content presupposes the ability to communicate richly in non-conventional ways first. To quote Wittgenstein (2005, p. 25e) again:

I wouldn't know what I should point to in the picture as a correlate of the word *kiss*. . . or . . . the word *taller*. . . [But] there is an act of "directing attention to the size of people" or to their actions. . . . This shows how it was possible for the general concept of meaning to come about.

The main point is that to conventionalize a way of doing something - in this case a way of communicating something - some original, non-conventional way of doing it must be available first; there must be something to conventionalize.

1.1. SOME EXAMPLES OF NATURAL GESTURES

Here are some actual observations of people communicating with one another by pointing, with no accompanying language. Each communicative act may be glossed in terms of the

communicator's referential intention (what the communicator wants the other to attend to) plus the social intention (what the communicator wants the other to know, do, or feel as a result of this attending).

Example 1

A man in a bar wants another drink; he waits until the bartender looks at him and then points to his empty shotglass. *Gloss*: Attend to the emptiness of the glass; please fill it up with liquor.

Example 2

We are climbing up a steep riverbank, me already on top, and the person following, in order to have her hands free for climbing, hands up to me a book and points to the protruding end of a pen. *Gloss*: Attend to the precariousness of the pen; please be careful and don't let it drop out.

Example 3

People standing in line. The line has moved forward and a man hasn't noticed this because he is turned around talking to the person behind him. Someone from still further back points for him to the newly opened gap. *Gloss*: Attend to the empty space; please move up into it.

Example 4

I am standing in the rear of the airplane, just to stretch for a bit, and this is near the bathroom. A woman approaches, and when she sees me points to the bathroom door with a quizzical look on her face. *Gloss*: Attend to the bathroom; are you waiting for it?

Example 5

A man in the lobby of a hotel unknowingly drops a piece of paper on the floor as he searches his coat pockets. I point to it for him. *Gloss*: Attend to that paper; know that you dropped it (and then do what you wish).

The first thing to notice about these otherwise mundane examples is that the referent of a pointing gesture may be just about anything, from the emptiness of a glass (or perhaps the absent liquor) to the precariousness of a pen. The second thing to notice is that the social intentions involved may be just about anything as well, from "move up in line" to "know about the paper you dropped". Apparently in all of these cases the communicator is confident that the recipient will be able to infer his social intention by attending in the designated direction (and his intention that she so attend) referentially, given their shared context.

One might suppose that only someone who is linguistic already could use a pointing gesture to communicate in such complex ways—that somehow the ability to communicate so richly with a simple pointing gesture is parasitic on linguistic skills. But that is not the case. Human infants before they have much or any language can already use pointing to direct others' attention to all kinds of referents in order to communicate all kinds of complex social intentions. Here are some actual observations of prelinguistic or just-linguistic infants made by mothers keeping systematic diaries (from Carpenter et al., in progress).

Example 6

At age 11.5 months, after Mom had poured water into J's glass at the dinner table, a few minutes later J points to his glass to request that she pour him some more. *Gloss:* Attend to my glass; fill it up.

Example 7

At age 13.5 months, after finishing eating, L points to the bathroom in anticipation of going to wash her hands. *Gloss:* Attend to the bathroom; it's time to go there.

Example 8

At age 13.5 months, while Mom is looking for a missing refrigerator magnet, L points to a basket of fruit where it is (hidden under the fruit). *Gloss:* Attend to the basket of fruit; it's there.

Example 9

At age 14 months, two different children, J and L, have accidents when a parent is not looking; when the parent comes to investigate, the infant points to the offending object (i.e., the thing he bumped his head on, or the thing that fell down). *Gloss:* Attend to that object; it hurt me/fell down.

Example 10

At age 14.5 months, as Mom is bringing the highchair to the table, L points where it goes. *Gloss:* Attend to that place; put it there.

Notice that even these prelinguistic, or just-linguistic, infants are pointing to things like the contents of their glass (Example 6, similar to the adult Example 1), the location of the not-perceivable refrigerator magnet, the location where the chair should go, and things that caused accidents some moments before. Certainly the range of things infants this young can point to referentially is smaller than for adults; but still the range is impressive. Of course these infants have been hearing language for all their lives, but they only show their first signs of comprehending and producing linguistic conventions during this same developmental period (prototypically, 14 months), not before. And deaf children of hearing parents, exposed to basically no conventional language during their first year of life or more (either spoken or signed), still begin to use the pointing gesture normally and at around the same age (Ledeberg and Everhart 1998; Spencer 1993). There are also logical reasons, elaborated below, to believe that infants' comprehension of language depends on their comprehension of pointing - or other nonlinguistic ways of achieving joint attention - rather than the other way around.

Our main focus here will be on pointing, as the primordial act of uniquely human communication, but we may also note very briefly that iconic gestures may also be used without language to communicate in quite complex ways. Iconic gestures - even if they are not conventionalized - have semantic content in a way that pointing gestures do not, and they typically are used to indicate referents (either objects or actions or relations) that are not currently perceptually present: they invite the recipient not to attend to the referent but

imagine it. But the pragmatic infrastructure is basically the same as for pointing, though it is deployed in a slightly different way. Some observed examples:

Example 11

I am in a cheese shop in Italy, and I ask for “parmegiano.” The proprietor asks me something I do not understand, but guessing—and not having the appropriate word—I twiddle my fingers as if sprinkling grated cheese onto my pasta. *Gloss:* Imagine what I am doing this to; and give me some of it.

Example 12

The airport security guard motions his hand in a circular motion to tell me to turn around so he can scan my back. *Gloss:* Imagine your body doing this motion; do it.

Example 13

At a vegetable stand, the proprietor—from a few meters away with back partially turned—is following a customer’s request to fill a bag with potatoes. She pauses with a questioning look to query nonverbally “Should I stop here?” The customer motions his hand in a shoveling motion like the one she was just doing. *Gloss:* Imagine doing this action (which you were just doing); do it (i.e., “Keep going with additional potatoes in the bag”).

Example 14

At a loud construction site, one worker pantomimes to another ten meters away as if he were using a chainsaw. *Gloss:* Imagine me doing this; bring me the thing I need to do it.

Here the process is that the communicator invites the recipient to imagine a referent object or action, and then to know, do, or feel something about it. This acting out of situations means that iconic gestures can be used in cases where pointing cannot, or at least not easily. But iconic gestures still require the identifying of a referent and an inference about why the communicator wants the recipient to attend to that referent. (NB: because iconic gestures are essentially pretense, the recipient must “quarantine” the interpretation, and not see the act as a normal intentional action but rather as a communicative act only; Leslie, 1987.)

Infants do some iconic gesturing, but because they are acquiring language - and iconic gestures, with their semantic content, compete with language in a way that pointing does not - they do not use iconic gestures nearly as frequently or creatively as they do the pointing gesture. But deaf children who have not been exposed to any conventional vocal or signed language invent iconic gestures to communicate in extremely rich and complex ways early in development (Goldin-Meadow 2003). Like pointing, then, communicating with iconic gestures also does not depend on language.

Pointing and pantomiming are “natural” in a way that “arbitrary” linguistic conventions are not because they build on natural human abilities and proclivities. Specifically, pointing is based on humans’ natural tendency to follow the gaze direction of others to external targets (an ability of all primates; Tomasello et al., 1998), and pantomiming is based on humans’ natural tendency to interpret the actions of others intentionally (an ability of at least some

nonhuman primates as well; Call & Tomasello, 2008). In both cases, of course, this natural tendency is not sufficient; to get to nonnatural meaning (Grice, 1957), the communicator must intend certain things and both participants in the communicative act must relate to one another socially in particular ways. Specifying these intentions and social relations is our goal here.

1.2. THE UNIQUENESS OF NATURAL GESTURES

Before proceeding to our account of how human beings communicate so richly with such impoverished communicative vehicles as protruding fingers, it is instructive to look briefly at the gestures of our nearest primate relatives.¹

Great apes communicate regularly using gestures, and they use at least some of them in quite flexible ways, attending to the attentional state of the intended recipient in the process (e.g., only producing visual gestures when the other is looking; see Call & Tomasello, 2007). But these gestures are still a long way from human gestures because they are mostly dyadic, regulating the interaction between two individuals directly (e.g., initiating play by feigning hitting), not about external referents at all. The one exception to this is that great apes produce some gestures designed to get others to attend to themselves (e.g., slapping the ground, poking others whose backs are turned, throwing things at others, etc.), or sometimes to one of their body parts (e.g., "offering" their back for grooming). Simple though these attention-getting gestures may be, they are the only examples I know in the animal communication literature of individuals having the primary goal of directing the other's attention (though vocalizations grab the attention of others, that is not the vocalizer's goal - and indeed vocalizations naturally only attract attention to their source, not to any external referent). And all of these gestures are imperative or directive in the sense that the gesturing individual is attempting to get the other to do what he wants her to.

Great apes do not point or gesture iconically for one another. But great apes who live in one way or another among humans do sometimes "point" to things they want the humans to retrieve for them (typically with whole hand, not with index finger), as a special kind of attention-getting gesture (Leavens & Hopkins, 1998). But again they only do this in imperative mode, to get humans to do things; they do not use their "pointing" to inform

1 There is no time for the full story here, but primate vocalizations cannot be the birthplace of human language, for sure not its pragmatic dimensions (see Tomasello, 2008, for a full treatment). Primate vocalizations are basically hardwired, with vocalizers having virtually no voluntary control, and they are broadcast indiscriminately to all in the area, not directed at other individuals or attuned to their attentional state. We humans naturally interpret their food calls and alarm calls as informing others of things, but although recipients are informed by these calls, that is not what communicators are attempting to do. They are attempting to do things like attract company for their meal (as defense against predators), or attract others to help mob a predator, or frighten a predator away. Seyfarth and Cheney (2003, p. 168) say: "Listeners acquire information from signalers who do not, in the human sense, intend to provide it".

humans of things or just to share attention to things. The richest plausible interpretation of this gesture, therefore, is that the ape wants the other to see something (e.g., the grape) so that it will do something (e.g., fetch it for him). This is the only behavior in animal communication of which I am aware in which there is a clear split between something like the referential intention (that someone attend to something) and the social attention (that someone do something as a result).

But still the inferential process involved here is very different from that of humans; most likely, the ape just knows from past experience that when the human sees food she normally hands it over. As evidence for this difference of process we may look at apes' comprehension of the pointing gesture (i.e., lack thereof). If a human points for an ape to food, the ape will very likely follow the gaze and pointing direction to the food and get it. Fine. But a small change of procedure changes everything. If the ape is searching for hidden food and a human points to a bucket, they do not understand that he is informing them that the food is inside (nor do they understand iconic signs in this context; Grosee et al., 2015; Bohn et al., 2016). Prelinguistic human infants understand this informative gesture quite readily from as young as 12 to 14 months of age (Behne et al., 2005; 2012). The reason that apes do not comprehend pointing in this seemingly simple situation is not because they cannot follow pointing direction - they often do in this experiment (but then choose randomly anyway). And it is also not because they cannot make inferences, because in other experiments they make inferences of the appropriate kind with regularity. For example, if one actually reaches for one of the buckets, attempting to procure it directly for the self in a competitive situation with the ape, they immediately infer that there must be food inside (Hare & Tomasello, 2004). They understand my direct intention to get into the bucket (and so infer there must be something good inside), but they do not understand my cooperative communicative intention to inform them that something good is inside.

But this is still not the full story. Apes do not understand informative pointing also because they cannot create with others the kind of common conceptual ground necessary to interpret simple gestures in complex ways (e.g., the common ground necessary to know that pointing to an empty glass in certain contexts means that one wants it filled, and with a particular liquid). Indeed, the full story is that apes do not comprehend informative pointing because their communication is not structured cooperatively - in terms of shared intentions, common ground, cooperative motives - in the human manner.

2. THE COOPERATION MODEL OF HUMAN COMMUNICATION

Ever since Grice (1975) introduced the cooperative principle, there has been debate about how cooperative human communication really is. But it all depends on one's frame of reference. In an evolutionary perspective, it is clear that human communication is cooperative to an unprecedented degree, though obviously it may be used for non-cooperative purposes as well. The proposal here is that human cooperative (Gricean) communication arose evolutionarily in the context of collaborative activities - to initiate and coordinate them more efficiently - and so in the beginning it was cooperation all the way down, with no use of this form of communication outside of collaborative activities and no possibility of lying at all. Cooperative communication outside of collaborative activities, and non-cooperative uses of cooperative communication, came only later.

Prima facie evidence for this proposal is that the basic structure of human cooperative communication is the same as that of human collaboration more generally. Specifically, in common with collaborative activities in general, human cooperative communication is underlain by skills and motivations of shared intentionality (e.g., Bratman, 1992; Tuomela, 2007; Searle, 1995): (1) human communicators form the joint goal (each with his own role) of getting the communicator's message across, evidenced most clearly by the communicator's active adjustments for the recipient, and the recipient's requests for clarification, confirmations, etc. (i.e., human communication just is a collaborative activity; Clark, 1996); (2) both human collaboration and communication are structured by joint attention and other forms of common conceptual ground (with coordinated perspectives); and (3) both human collaboration and communication are powered by the fundamentally cooperative motives of helping and sharing. The common infrastructure of shared intentionality underlying both the collaborative and communicative activities of contemporary human beings thus provides us with a tangible stamp of their common evolutionary origin.

2.1. COOPERATIVE MOTIVES

Humans communicate for three basic types of motives, and a few more specific ones. The basicness of these motives is established by their early emergence in human ontogeny - with each associated with a distinctive vocal/intonational signature (see below).

The first fundamental human communicative motive, seemingly unique to the species, results from the fact that individuals often want to offer help to others—specifically, by informing others of things helpfully, even when they themselves have no personal interest in the information. Thus, when I point out to you the paper you just dropped or tell you that the main street home is blocked today, it is mutually assumed that I am attempting to be helpful

to you. Obviously, helping others by informing them of things they will find useful or interesting involves cooperative/altruistic motives of a type that requires special evolutionary explanation - especially since other species do not seem to do it whereas even prelinguistic human infants point to inform others of things helpfully, for example, the location of an object they are searching for.

The second fundamental human communicative motive is requesting or directing. All primates communicate imperatively in this way, of course, but the difference is that instead of ordering the other what to do, humans often do something more gentle like requesting help (from someone who likes helping). That is, unlike ape imperatives, human imperatives can range from orders to polite requests to suggestions to hints, depending most fundamentally on the degree to which a cooperative attitude may be assumed of the recipient. Thus, if you are on my land I can order you to leave, or I can simply inform you that I would like you to leave (or even that it is *my* land) if I think you will readily comply. We might call the first type individual imperatives or requests—since I tell you what to do directly—and the second type cooperative imperatives or requests—since I simply inform you of my desire and assume that you will decide to help me fulfill it (i.e., if I simply inform you of my desire that you leave, you must care about my desire if the request is to work). Questions may be thought of also as requests for help, that is, requests for desired information.

Informatives and requestives thus both involve helping: either offering help by providing useful information informatively, or responding positively to a request for help by either acting or providing information. Socializing a well-known formula of Searle's (1999; originally from Anscombe, 1958), we may thus say that requestives reflect a You-to-Me direction of fit, as I want you to conform to my desire, whereas informatives reflect a Me-to-You direction of fit, as I want to conform to your desires and interests.

In addition to these two most basic motives, we must posit a third basic communicative motive as well. People often simply want to share feelings and attitudes about things with others—what I will call an expressive or sharing motive. For example, on a beautiful day it is quite common to say to your officemate upon arrival, “What a beautiful day!”—which derives not from any requestive or informative motive involving help, but rather from a purely social one. This kind of communicative act is simply a sharing of attitudes and feelings with others as a way of bonding socially by expanding our common ground with them. This sharing motive underlies much of the everyday talk of people as they gossip about all kinds of things, expressing opinions and attitudes which they hope the other will to some degree share. It

turns out that this motive emerges ontogenetically quite early in infants' prelinguistic pointing, as they, for example, point for a parent to a colorful clown and squeal with glee.²

Human cooperative communication is thus structured most basically by motives of helping (both requesting and offering help) and sharing attitudes. The informative motive, in particular, is interesting and important because it creates the possibility of both speaking the truth and lying. Truth is not a relevant concept when all we are doing is issuing imperatives.³ And lying is possible only because recipients trust that when someone informs them of something they are usually attempting to be helpful, which in the case of informatives means truthful. Truth is thus not a basic and general epistemological concept, but rather one that emanates from a specific, evolved communicative motive for being cooperative (Tomasello, 2018).

2.2. COMMON GROUND AND COOPERATIVE REASONING

These three most basic of human communicative motives - requesting, informing, and sharing - underlie a virtual infinity of particular social intentions (e.g., that you pick up the paper, that you know your friend has arrived, etc.) in particular social contexts. A communicator's particular social intention on a particular occasion is determined by the recipient through a complex inferential process involving several layers of intentionality, all taking place within the context of some kind of common conceptual ground between interactants.

The explanation for humans' uniquely complex ways of communicating gesturally is that for humans the communicative context is not simply everything in the immediate environment, from the temperature of the room to the sounds of birds in the background, but rather the communicative context is what is "relevant" to the social interaction, that is, what each participant sees as relevant and knows that the other sees as relevant as well—and knows that the other knows this as well, and so on, potentially ad infinitum. This kind of shared, intersubjective context is what we may call, following Clark (1996), common ground or, sometimes (when we wish to emphasize the shared perceptual context), the joint attentional frame. Common ground includes everything we both know (and know that we both know, etc.), from facts about the world, to the way that rational people act in certain situations, to what people typically find salient and interesting (Levinson 1995).

2 There is some correspondence here with the basic speech act functions posited by theorists such as Searle (1999), though the mapping is not totally straightforward. We should also note a number of specialized motives for specialized situations that are ontogenetically early and very likely culturally universal: greeting/leave-taking ("Hello" and "Goodbye"), expressing gratitude ("Thanks"), and expressing regret ("Sorry"). These are special because they are not referential in the normal way and so work a bit differently.

3 So-called assertions, in the current view, are simply informatives in advanced discourse contexts such as science and politics, where my individual goal is to convince you of something. Informing is more basic than asserting, which is a special case.

Common ground is necessary for the recipient to determine both what the communicator is directing attention to (his referential intention) and why he is doing it (his social intention). Thus, in the relatively simple first example of pointing given above (a customer points for the bartender to his empty shotglass to request another drink), without some kind of common ground the bartender cannot know if the customer is pointing to the glass as a whole, or its color, or a small crack in it. Indeed, in the actual example, the customer is pointing not to the glass itself but to its emptiness (imagine the difference if the pointed-to glass were already full—the customer’s meaning would have to be something very different). And even keeping the exact same referent, the social intention may be different depending on common ground. Thus, in the normal situation the customer is pointing to his empty shotglass to request it being filled with liquor—which the bartender understands because they both know together that customers are at the bar because they want to drink, an empty glass does not afford drinking, the bartender has drink if the customer can pay, shot glasses are typically filled with liquor not wine or beer, etc. But, if the customer and bartender are actually buddies who regularly attend Alcoholics Anonymous together, the customer could be pointing to the emptiness of his shot glass in this case to indicate to his buddy that he has still managed, after an hour at the bar, to resist having a drink.

The critical point about common ground is that it takes people beyond their own egocentric perspective on things. For example, modifying an example from Sperber and Wilson (1986), suppose that in a park I point to direct your attention to a location some meters away. There are three people there: an ice-cream vendor, a jogger you have never seen before, and William, who is your lover. If you are being egocentric, you assume in the first instance that I am drawing your attention to William, as he is very relevant for you, whereas the other two are not relevant for you. In the normal case, though, your search for relevance is not egocentric but takes place within the context of our shared common ground from the beginning, for example, taking into account from the beginning whether we both know together that we both know about William. Thus, suppose that I do not know about William and you know this for certain (he is your *secret* lover), and suppose further that you and I both know that we both share a passion for ice cream (we have explicitly discussed this). If I now direct your attention in the general direction of these same three people, no matter how relevant William is for you egocentrically, and even if you were lying to me about the ice cream (so that it is not in reality relevant for you at all), you will still assume that I am indicating for you the ice-cream vendor, since we both “know” from our previous discussion

that we both love ice cream and you think I do not know about you and William. In direct competition, shared common ground trumps individual personal relevance every time.⁴

Of course, you may hypothesize that I really do know about William somehow and proceed on that assumption. But then you are, essentially, guessing about the kind of common ground that would make the process the canonical one. In the normal case, you want to know from the outset why I think that looking in that direction will be relevant for you, with a prerequisite being that we know together about the potential referent and its relevance for you. And so what comes to your mind most readily as an interpretation of my pointing gesture, at the top of the stack as it were (even though you may have your own personal interests as well), will be those things that are in our common ground. Another variation is cases in which we do not have direct personal common ground, but we both, as members of a particular culture or social group, have assumptions about what the other should know (and know I know, etc.). Thus, I might point to a sight for you out the airplane window even though we have never before met, as I assume that you can identify the intended referent based on (presumably) shared assumptions about what people typically find salient, beautiful, and so forth. But note that in both of these cases—guessing and general cultural common ground—the recipient attempts to comprehend the communicative act by, in effect, imagining or assuming some form of common ground that she must share with the communicator if the whole thing is to make sense. The normal case—the one with which young children begin and the one that adults process without hesitation—is thus the case in which we both recognize our common ground within which the communicative act is immediately comprehensible.

Importantly, for all types of human communication including language, the relationship between the overt communicative act and common ground—of whatever type—is complementary. That is, as more can be assumed to be shared between communicator and recipient, less needs to be overtly expressed. Indeed, if enough is shared in common ground, the overt expression of either motive or referent may be totally eliminated without diminishing the message at all. For example, in the dentist's office the dentist may sometimes point to the instrument she wants without overtly expressing her desire per se to the assistant, since her desire to have the instrument is mutually assumed in this mutually known context (cf. Wittgenstein's builders). Conversely, the dentist may simply hold out her hand, indicating that she wants an instrument, and the assistant, based on shared knowledge of the

4 If we want to get really perverse, we can, following Clark and Marshall (1981), imagine a situation in which I have found out, unbeknownst to you, that William is your secret lover. Then we both know that he is extremely relevant for you, but since you do not know that I know this, you will not assume I am referring to him. This recursion may continue ad infinitum, with successful reference not possible with any finite number of iterations. It is required that we both know together, it is in our common ground, that William is especially relevant for you.

procedure, puts the correct one (of many on the table) in her hand without the intended referent ever having been indicated specifically.

In all, then, it is only because humans are able to construct with others various forms of conceptual common ground and joint attention that very simple pointing and iconic gestures can be used to communicate in complex ways. Indeed, in many cases, when the common ground is particularly well defined, simple gestures may communicate as powerfully as language. Most basically, as can be clearly seen in the examples in which the referent of pointing changes with the common ground—for example, pointing to the shot-glass to indicate either the object itself, its color, its emptiness, or its state of repair—a certain kind of perspective shifting is involved. It is thus possible that this kind of reference shifting in gesturing—accomplished by making contact in different ways with communicator-recipient common ground—paves the way for perspectival linguistic conventions both phylogenetically and ontogenetically. Moreover, although reference to entities displaced in space and time has traditionally been seen as the exclusive province of language—and there is no doubt that language does this by far most productively—within an appropriate shared context, people may point or gesture iconically to direct attention to the nonpresence of expected entities (e.g., the absence of drink in a glass) or even to indicate absent entities directly (e.g., the desired chainsaw in example 14), which may also pave the way for displaced reference in language. What this means is that many of the especially powerful properties that people often attribute to language—including referring others to perspectives on things and to absent referents—are actually present more fundamentally in human cooperative communication with very simple gestures. This is possible because of—and only because of—various types of common conceptual ground and joint attention between communicators.

The facts that communicators operate with cooperative motives and that recipients are inclined to respond appropriately (all other things being equal) are part of the common ground between human communicators. Indeed, this is what motivates them to cooperate in getting the message across in the first place—they both assume mutually that it will be to their individual and mutual benefit to do so. Because the communicator knows this, he makes sure that the recipient knows that he is attempting to communicate, as if to say: “You’re going to want to know this” (i.e., that I have a request of you, that I have something I want to inform you about, that I have an attitude I want to share). This additional layer of intentionality—“I want you to know that I want something from you”—is absolutely critical to the process and is most commonly referred to as the (Gricean) communicative intention.

Grice (1957) observed that human communicative acts involve an intention about the communication specifically. That is, when I point to a tree for you, I not only want you to attend to the tree, I also want you to attend to my desire that you attend to the tree (often signaled by eye contact, etc., and also implicit quite often in the expression of motive, as a

sign that this act is done “for you”). This additional intentional layer is necessary to motivate you to make the kinds of relevance inferences required to identify both my referent and my social intention (Sperber & Wilson, 1986). Thus, when you see me pointing to a tree, and clearly wanting you to know that I am pointing it out for you, then naturally you want to know why I am doing that: what I want you to do, think, or feel with respect to the tree. You assume that when I point to the tree for you, I believe it will be interesting or relevant for you in some way: perhaps because it is your favorite kind of tree and I want to inform you of its presence here, or perhaps because I have a request about it that I think you would like to fulfill, or perhaps because I want you to share my enthusiasm for it.

The main point is that this process occurs because both participants know together and trust together the cooperative motivations involved. That is to say, in general, if a human communicator requests help (all other things being equal), the recipient will want to help—and they both know this and trust in this. Similarly, if the communicator offers information, they can mutually assume that he thinks the information will be useful or interesting for the recipient (and that normally means “true” as well)—and so she will accept it. And finally, if the communicator wants to share attitudes, they assume together the prosocial motive of sharing, and the communicator may expect the recipient to share unless there are good reasons against it. The communicator therefore overtly signals his intention to communicate, and they therefore both work together to ensure that the communicative act succeeds.

Importantly, overt expression of the Gricean communicative intention places the communicative act itself—the gesture or the utterance—into the participants’ common ground, specifically, into the ongoing joint attentional frame within which they are communicating. Thus, it is most precise to say not just that I want you to know that I want you to attend to something, but that *I want us to know this together*—I want my communicative act to be a part of our perceptually co-present joint attention (I want it to be mutually manifest, in the terms of Sperber and Wilson 1986, or “wholly overt”). Because human communicators make their communicative intention mutually manifest, this makes this intention, in an important sense, public—which triggers a whole other set of processes (Habermas 1987). Specifically, the fact that I have communicated to you overtly, publicly, actually creates not just expectations of cooperation but actual social norms, whose violation is unacceptable. For example, if I ask you at the dinner table to pass the salt, you cannot really just say no, or if I find out that your child has been injured at school, I cannot intentionally neglect to inform you.

The cooperative motives involved here, and the mutual knowledge of these cooperative motives and even norms, mean that the participants in human communication must reason not just practically, but cooperatively. Thus, when apes observe another ape signaling to them, they try to discern what he wants via individual practical reasoning about his goals and perceptions. But they are not trying to understand the message *because* he wants them

to, since the two of them do not share an assumption that he is trying to be helpful. The communicator thus does not signal or “advertise” his intention specifically, as humans do in signaling their communicative intention. And in choosing a response, ape recipients do not respond in a certain way *because* the other wants or expects them to; rather, they simply try to do what is best for themselves in the situation given what the communicator apparently wants. In contrast, when humans see that someone is attempting to communicate with them, they want to know what he is attempting to communicate at least partly *because* he wants them to (and they trust his cooperative motives), and they choose a response—e.g., complying with a request or accepting offered information or sharing enthusiasm about something—at least partly *because* that is what the other wants them to do. Since human recipients comprehend and respond to communicative bids in certain ways at least partly because that is the way the communicator wants them to (with the communicator relying on this)—and indeed because this way of operating is, if everything is public, normatively prescribed—we call the kind of practical reasoning characteristic of human communication cooperative reasoning.

2.3. EVOLUTION OF SHARED INTENTIONALITY

And so what makes human communication unique in comparison with that of other animals, including our nearest primate relatives, is all of this cooperation. So where did it come from? Our proposal is that human cooperative communication evolved in the context of uniquely human forms of cooperation in general, from dyadic collaboration to cultural organization. Cooperative activities were the “forms of life” within which cooperative communication arose.

First, perhaps some 400,000 years ago, early humans created new forms of collaborative activity, specifically those structured by joint intentionality. In acts of joint intentionality, early humans began to form with one another joint goals toward mutually beneficial ends, structured by joint attention. In pursuing their joint goals structured by joint attention these early humans also recognized simultaneously different individual roles in the collaborative activity and different individual perspectives on their joint focus of attention. This dual-level structure of joint agency created for the individuals involved a shared world comprising distinct individual perspectives.

Within these new kinds of collaborative activities, early humans begin to communicate cooperatively with one another in unique ways, that is, to inform one another of things helpfully, through the natural gestures of pointing and pantomiming. This created the basic ostensive-inferential structure of uniquely human communication. Motivationally, when we are engaged in a mutualistic activity, if there is any information that will help you in your role I should inform you of it, since this helps me too. Following Tomasello (2008, 2014), the

ecological context within which these skills and motivations developed, was likely some kind of cooperative foraging. Humans were put under some kind of selective pressure to collaborate in their gathering of food - they in fact became obligate collaborators - in a way that their closest primate relatives were not.

The proposal is thus that human cooperative communication evolved first inside of collaborative activities – in order to facilitate the creation of a joint goal and to coordinate the different roles involved. The activities themselves provided the needed common ground for establishing jointly focused referents, and they generated the cooperative motives that are essential if the inferential machinery is to work in the human-like manner. Only some time after humans had developed means of cooperative communication inside of collaborative activities did they begin to communicate cooperatively outside of such activities. Skills and motivations for cooperative communication thus co-evolved with collaborative activities, because it both depended on them, but then it also contributed to them by facilitating further cooperative coordination.

Other evolutionary stories are possible, of course, for example, that human cooperative communication involve first for pedagogy (Gergely & Csibra, 2006). But what makes the current story especially plausible is that all of the elements of the underlying infrastructure of cooperative communication and collaborative activity - joint goals, coordinated roles, joint attention, differentiated perspectives, cooperative motives - are identical. Chimpanzees lack both true collaboration and cooperative communication, and human children, as we shall see shortly, develop the two in concert.

2.4. THE EVOLUTIONARY TRANSITION TO LANGUAGE

For the evolution of language, iconic gestures are especially important as they involve symbolic representation, typically of displaced referents. Nevertheless, iconic gestures, like pointing, have communicative limitations as well. If I pantomime for you the act of digging to suggest to you, a novice, what you should now do (assuming you understand it as a communicative act), comprehension relies to some degree on your familiarity with digging in general and your assessment of what is needed now in the current situation.

Human groups at some point went beyond iconic gestures that needed to be invented anew on every occasion, and moved to communicative conventions. Conventions are ways of doing things that are somewhat arbitrary—there are other ways they could be done—but it is to everyone’s advantage if everyone does it in the same way, and so everyone just does what everyone else is doing because that is what everyone is doing (Lewis 1969). This arbitrariness means that one cannot invent conventions on one’s own. One can invent communicatively effective iconic gestures, but arbitrary communicative conventions require that

they be “shared,” so that everyone can rely on everyone else in the group knowing how the convention is used communicatively.

So how did communicative conventions get started? Invoking a process of explicit agreement—as in social contract theories—is not really a viable option, as agreement presupposes an already existing means of communication, more powerful than the one to be invented, in which to formulate the agreement. But among organisms who already possess the cooperative communicative infrastructure we have laid out here, and who are also capable of collaboration and imitation, conventions can arise “naturally” as a result of a combination of shared and unshared experiences. Here is the kind of scenario that must have occurred at the dawn of arbitrary communicative conventions. First came some kind of cooperative iconic gesture. For example, perhaps a female of the genus *Homo* wishes to go digging for tubers. To get others to come with her, she pantomimes digging for them in exaggerated fashion in the direction in which tubers are normally found. The cavemates understand this gesture naturally, that is, they understand that this digging gesture is intended to depict a real instrumental action of digging. It is possible that some of them might then learn this gesture from her, by imitation, thus creating a shared communicative device that is both shared and at least partially arbitrary in the sense that other gestures for this same function could certainly have been used.

But now let us assume the following extension of the scenario. Some individuals not familiar with digging, perhaps children, observe this “Let’s go digging” gesture, and for them the connection between the ritualized digging gesture and the act of digging for tubers is opaque (though they do see that it is intended to be communicative); they think it is just intended to initiate leaving generally. They might then imitatively learn the gesture to initiate leaving (for something other than digging) on some future occasion—so that the original iconic grounding of the gesture is now completely erased. (This is not unlike the way that some motivated linguistic forms, such as metaphors, become opaque [“dead metaphors”] across historical time as new learners are not exposed to the original motivation.) One can possibly imagine in addition some kind of general insight at some later point that most of the communicative signs we use have only arbitrary connections to their intended referents and social intentions, and so, *voilà*, we can if we want make up new arbitrary ones as needed.

The move to communicative conventions is thus, paradoxically, a natural one. No one intends, certainly not initially, to invent any conventions. Communicative conventions happen naturally as organisms who are capable of imitation and who already know how to communicate in fairly sophisticated ways—cooperatively, with gestures—imitatively learn one another’s iconic gestures. Then individuals who are not privy to the iconic relation observe the communicative efficacy of the gesture and use it on that basis only, without any iconic motivation—at which point it has become, for these new users, arbitrary. This is what has

been called a “process of the third kind,” a sociological result of human intentional actions, but not something that any one person actually intended (Keller 1994). Ultimately, for the group-wide establishment of communicative (linguistic) conventions requires skills and motivations of collective intentionality in which everyone in the group shares certain knowledge, including of communicative conventions, in their cultural common ground.

Linguistic conventions thus basically codify the ways that previous individuals in a community have converged upon to manipulate the attention and imagination of others in specific ways. The arbitrary sound or gesture itself carries no message “naturally,” but observing its use reveals—for those with the appropriate cognitive skills and motivations—how those who share the convention use it to direct the attention and imagination of others. The appropriate cognitive skills and motivations are of course none other than (i) the same shared intentionality infrastructure that underlies human pointing and pantomiming, and (ii) a shared learning history in cultural common ground, with the convention that we all know together (implicitly) that we share it—a fact that may be signaled by various kinds of cultural markers (including even use of the convention itself in an appropriate manner). Humans’ creation and use of shared communicative conventions thus means that now even the communicative forms themselves depend on processes of shared intentionality.

3. EVIDENCE FROM ONTOGENY

Often it is easiest to see the components of complex skills and how they work together when we study their emergence in children’s early development. An important source of evidence for the cooperation model of human communication, therefore, is how things work ontogenetically. Moreover, it turns out that gestures used as full communicative acts (without language) have been investigated much more intensively, especially in experiments, in infants and young children than in adults.

3.1. COOPERATIVE MOTIVES

In classic accounts, infants point communicatively for one of two motives: they point to request things (imperatives) and to share experiences and emotions with others (declaratives) (Bates et al., 1975). But, in line with our analysis above, we think that there are actually two subtypes of declaratives: expressives and informatives. Requestives and informatives have to do with helping - either requesting it or offering it - and expressives have to do with sharing one’s emotions or attitudes about the world, and helping and sharing are the basic motives of shared intentionality in general. Infants show all of these motives in their

communication before language, and they all emerge together at around the first birthday (see Tomasello et al., 2007, for a review).

The basic question here is why one year - why not earlier? The specific behavioral form of pointing - distinctive hand shape with extended index finger - actually emerges reliably in infants as young as three months of age (Hannan & Fogel, 1987). However, as far as anyone can tell, infants at this age are not using this hand shape for any communicative function. This is despite the fact that they also seem to have some of the needs that precipitate truly communicative pointing later in development, for example, the need for adults to do things for them, including fetching out-of-reach objects (underlying requests), and the need for adults to share emotions with them in protoconversations (underlying expressives). So why do infants not learn to use the extended index finger for these social functions at 3 to 6 months of age, but only at 12 months of age?

Our basic answer is that 3- to 6-month-old infants do not point for others communicatively because they do not yet have the requisite understanding of joint intentions, joint attention, and cooperative motives; that is, they have not yet developed the skills and motivations for shared intentionality. When did these arrive on the scene? They arrive on the scene at exactly the same time. In what some have called the nine-month revolution, infants as they approach their first birthday begin to interact with others collaboratively in so-called joint attentional engagement (Tomasello, 1999). At about nine months of age infants begin for the first time to do things with adults like roll a ball back and forth, or stack blocks together, that involve a very simple joint goal and joint attention. This is the control parameter (to use dynamic systems language) that creates the initial possibility of human cooperative communication in ontogeny. We can see it at work most clearly in infants' early comprehension of pointing.

3.2. COMMON GROUND AND COOPERATIVE REASONING

We might presume that common ground plays a critical role in infant pointing from the beginning, based on such things as our diary observations involving many different contexts determining many different meanings for the infant's pointing gesture. However, demonstrating a role for common ground requires demonstrating that the context is indeed "shared" or mutually known, and that is (or at least has been) most readily demonstrated in comprehension. This is true for both the social intentions and referential intentions involved.

In terms of social intentions, Liebal et al. (2009) had 14- and 18-month-old infants and an adult clean up together by picking up toys and putting them in a basket. At one point the adult stopped and pointed to a target toy, which infants then picked up and placed in the basket. However, when the infant and adult were cleaning up in exactly this same way, and a second adult who had not shared this context entered the room and pointed toward the

target toy in exactly the same way, infants did not put the toy away into the basket—presumably because the second adult had not shared the cleaning-up game as common ground with them. Rather, because they had not just been interacting with this adult, they seemed most often to interpret the new adult’s pointing gesture as a simple invitation to notice and share attention to the toy (i.e., as an expressive declarative). Infants in both cases were thus directed to the same referent toy—they understood the referential intention in the same way in both cases—but their interpretation of the underlying social intention was different in the two cases. Most importantly, this interpretation did not depend on their own current egocentric interests, but rather on their recently shared experience (joint attention, common ground) with each of the pointing adults.

In the case of the referential intention, Moll et al. (2008) had an adult direct an ambiguous request to 14-month-old infants by gesturing in the general direction of three objects (the target and two distractors) and asking the child to hand “it” over. In different experimental conditions, the infant had had different experiences with the adult previously, and so had different common ground on which to draw in identifying the referent of the request. Specifically, in the experimental condition, prior to the request the adult and infant had shared the target object excitedly as it unexpectedly appeared and reappeared in several places in a hallway (whereas they had handled the two other objects [distractors] in a more normal fashion). In this condition infants responded to the adult’s request by handing him the target object, the one they had shared—based on their common ground with him previously. Importantly, they did not do this in either of two control conditions. In one of these a new adult made the request, and so there was no common ground; infants then chose randomly. In the other the adult who made the request had previously experienced the objects individually (in the same excited fashion) while the infant simply looked on unnoticed; so again there was no common ground and the infants chose randomly. Thus, when faced with a request for an unspecified referent object, infants did not assume that the requestor was asking for the object that she, the child, had been excited about (or else they would have retrieved the target also when the different adult requested it), nor did they assume that the requestor was asking for the object that he himself had been excited about (or else they would have retrieved the target also in the condition in which they simply watched the requestor become excited about the target object on their own). Instead, the infants assumed that the adult was requesting the object about which the two of them together, in their recent common ground, had showed excitement.

Infants thus use their shared common ground with a pointing adult—not their own egocentric interests—to interpret both the adult’s referential intention and his underlying motive and social intention. But do they understand the other aspects of the process in an adult-like way? In particular, do they understand the Gricean communicative intention? One very general piece of evidence that they do is their performance in a simple object

choice task with informative pointing. Thus, Behne et al. (2005) hid an object from 14-month-old infants, and then pointed to its hiding location (the same task that apes fail - see above). Unlike apes, infants understood immediately that the adult was informing them of the location of the hidden object. Important for current purposes, and a control condition the experimenter held her hand in a pointing-like configuration but without any signs that this was a communication for the child (she only inspected the watch on her wrist and looked up to the infant occasionally). This would seem to suggest that they have at least some understanding of when communication is "for me" and when not.

Further, in the study of Shwe and Markman (1997) children at 30 months of age requested a desired object from an adult. They always got the object they wanted, but in one case the adult signaled that she had understood everything correctly whereas in the other case the adult signaled that she thought the child wanted a different (wrong) object sitting nearby—which she said he could not have, so she actually gave him the one he really wanted (by accident, as it were). That is, in this especially interesting condition, the child got what he wanted in terms of the object, but his message was actually not understood correctly. In this case, children corrected these adult misunderstandings nevertheless. This suggests that these children had both the goal of getting the object (as social intention) and the intention of communicating successfully with the adult as a means to that end—which they wanted to accomplish in its own right.

3.3. ICONIC GESTURES, PRETENSE, AND COMMUNICATIVE CONVENTIONS

At the same time they begin pointing for others, human infants also begin to use other kinds of gestures. Some of these are conventionalized (e.g., for goodbye) and presumably are learned by imitation, just like linguistic conventions. But others are used creatively. One example is of a child who wants his mother to dump an object off of her head; so he looks at her and tilts his head. But children do not use such iconic gestures to nearly the same extent that they do pointing gestures. Indeed, over the second year of life nondeictic gestures (conventionalized and iconic) actually go down in frequency in comparison with pointing (Iverson et al., 1994). The explanation most often given is that children are learning language during this time, and conventionalized and iconic gestures compete with linguistic conventions in a way that pointing does not—presumably because iconic and conventionalized gestures and language both involve some kind of symbolic representation and even categorization of a referent.

In terms of linguistic conventions themselves, children begin acquiring their first pieces of natural language precisely on the heels of their first use of gestural communication. They have very powerful skills of imitation, and they can learn to direct the attention and imagination of others to all kinds of referents by doing so in a way that others do it. But they could

not do this if they did not have already in place an underlying infrastructure of shared intentionality. Without some kind of joint attention or common ground between communicator and recipient, how is the young child to comprehend an adult when she utters “Gavagai” if not with reference to their shared common ground? The critical role of the shared intentionality infrastructure in language learning and use, including joint attention and common ground, is the central premise of the social-pragmatic theory of language acquisition, as espoused by Bruner (1983), Nelson (1985, 1996), and Tomasello (1992, 2003). Indeed, one of the best established facts in all of the language acquisition literature (supported by both correlational studies in early acquisition and experimental studies somewhat later) is their children’s learning of linguistic conventions is scaffolded by their participation in joint attentional activities with others (see, Tomasello, 2003, for a review).

Infants’ first acquisition and use of linguistic conventions thus provides further support the cooperation model. The problem of referential indeterminacy arises precisely when an act of reference is removed from the kinds of shared intentionality contexts within which language acquisition normally occurs. When children experience an adult using a linguistic convention outside of such contexts, it is true: they acquire nothing. But when children experience an adult using a linguistic convention within such inherently meaningful contexts, they are quite often able to understand what is being communicated independent of language and so to acquire productive use of that convention.

Many animals can associate sounds with experiences, and human infants can do this from a few months of age. If association or “mapping” were all that is involved in acquiring a linguistic convention, then language would be everywhere in the animal kingdom, and it would start at three months of age in humans. But the fact is that animals and young human infants do not acquire or use linguistic conventions. The reason is that “arbitrary” linguistic conventions can be acquired only in the context of some kind of conceptual common ground with mature speakers, often in collaborative activities with joint goals and joint attention, and this only becomes possible in human ontogeny at around one year based on species-unique skills and motives for shared intentionality

4. CONCLUSION

In a well-known pronouncement, Searle (1969, p. 38) claims that:

Some very simple sorts of illocutionary acts can indeed be performed apart from any use of conventional devices at all. . . . One can in certain special circumstances “request” someone to leave the room without employing any conventions, but unless one has a language one cannot request of someone that he, e.g., undertake a research project on the problem of diagnosing and treating mononucleosis in undergraduates in American universities.

But indeed we can make such a request without language. That is to say, if we have linguistic individuals who have been discussing, in language, the fact that “we need someone to undertake a research project on the problem of diagnosing and treating mononucleosis in undergraduates in American universities,” then I could, at the right moment in the conversation, point to you, and the meaning of that pointing act would be that “*you* should undertake a research project on the problem of diagnosing and treating mononucleosis in undergraduates in American universities.” Of course this cannot happen without linguistic organisms setting up the context linguistically—this is clear. But the key point for current purposes is simply that when the context—the shared conceptual ground—is set up in enough detail, however that is done, a pointing gesture can refer to situations as complex as one wants.

I believe that this fact is not fully appreciated in modern discussions of reference and meaning. The reason is that almost all of the analyses are done in terms of linguistic communication, with all of its semantic content, propositional attitudes, indexical terms, and on and on. What I have tried to do here is to show that the most basic features of uniquely human communication do not derive from our ability to create, learn, and use communicative conventions, but rather they come from our unique ways of engaging with one another socially - that is, cooperatively - more generally. Our evolved skills of cooperation and shared intentionality enable us to create with one another joint attention and other forms of common conceptual ground and to communicate with one another for cooperative motives - which lead to all of the complex inferential machinery of human intentional, cooperative communication. This is not pragmatics as leftovers, but pragmatics as the foundation that creates the possibility of conventional communication in the first place. Pragmatics from the bottom up.

REFERENCES

- Anscombe, G. E. M. (1958). On Brute Facts. *Analysis*, 18:3, p.24
- Bates, E., Camaioni, L., & Volterra, V. (1975). The acquisition of performatives prior to speech. *Merrill-Palmer Quarterly*, 21, 205-224.
- Behne, T., Carpenter, M., & Tomasello, M. (2005). One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Developmental Science*, 8, 492-499.
- Behne, T., Liszkowski, U., Carpenter, M., & Tomasello, M. (2012). 12-month olds' comprehension and production of pointing. *British J. of Developmental Psychology*, 30, 359-75.
- Bohn, M., Call, J., & Tomasello, M. (2016). Comprehension of iconic gestures by chimpanzees and human children. *J. of Experimental Child Psychology*, 142, 1-17.
- Bratman, M. (1992). Shared co-operative activity. *Philosophical Review*, 101(2), 327-341.

- Bruner, J. (1983). *Child's talk*. New York: Norton.
- Call, J. & Tomasello, M. (2008). Do chimpanzees have a theory of mind: 30 years later. *Trends in Cognitive Science*, 12, 187-192
- Clark, H. (1996). *Uses of language*. Cambridge: Cambridge University Press.
- Clark, H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In: A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 10-63). Cambridge University Press, Cambridge.
- Gergely, G., & Csibra, G. (2006). Sylvia's recipe: The role of imitation and pedagogy in the transmission of cultural knowledge. In N. J. Enfield, & S. C. Levinson (Eds.), *Roots of human sociality: Culture, cognition and interaction* (pp. 229-255). Oxford: Berg Press.
- Goldin-Meadow, S. (2003). *The resilience of language: What gesture creation in deaf children can tell us about how all children learn language*. New York: Psychology Press.
- Graefenhain, M., Behne, T., Carpenter, M., & Tomasello, M. (2009). Young children's understanding of joint commitments. *Developmental Psychology*, 45, 1430-43.
- Grice, H. P. (1957). Meaning. *The Philosophical Review*, 64, 377-388.
- Grice, P. (1975). Logic and conversation. In P. Cole, & J. Morgan (Eds.), *Syntax and Semantics, Volume 3: Speech Acts* (pp. 43-58). New York: Academic Press.
- Grosse, K., Carpenter, M., Call, J., & Tomasello, M. (2015). Differences in the ability of apes and children to instruct others using gestures. *Language Learning & Development*, 11, 310-330.
- Grosse, G., Scott-Phillips, T., & Tomasello, M. (2013). Three-year-olds hide their communicative intentions in appropriate contexts. *Developmental Psychology*, 49, 2095-2101.
- Habermas, J. (1987). *The theory of communicative action*. New York: Beacon Press.
- Hannan, T., & Fogel, A. (1987). A case study assessment of "pointing" in the first three months of life. *Perceptual and Motor Skills*, 65, 187-194.
- Hare, B., & Tomasello, M. (2004). Chimpanzees are more skillful in competitive than in co-operative cognitive tasks. *Animal Behaviour*, 68, 571-581.
- Iverson, J. M., Capirci, O., & Caselli, M. C. (1994). From communication to language in two modalities. *Cognitive Development*, 9, 23-43.
- Keller, R. (1994). *On language change: the invisible hand in language*. Routledge.
- Leavens, D. A., and Hopkins, W. D. (1998). Intentional communication by chimpanzees: A cross-sectional study of the use of referential gestures. *Developmental Psychology*, 34, 813-822.
- Lederberg, A., & Everhart, V. (1998). Communication between deaf children and their hearing mothers: The role of language, gesture, and vocalization. *Journal of Speech, Language, and Hearing Research*, 41, 887-899.
- Leslie, A. (1987) Pretense and representation: The origins of "theory of mind". *Psychological Review* 94, 412-426.
- Levinson, S. C. (1995). Interactional biases in human thinking. In: E. Goody (Ed.), *Social intelligence and interaction* (pp. 221-260). Cambridge: Cambridge University Press.
- Lewis, D. (1969). *Convention*. Harvard University Press
- Liebal, K., Behne, T., Carpenter, M., & Tomasello, M. (2009). Infants use shared experience to interpret pointing gestures. *Developmental Science*, 12, 264-71
- Moll, H., Richter, N., Carpenter, M., & Tomasello, M. (2008). 14-month-olds know what "we" have shared in a special way. *Infancy*, 13, 90-101.

- Nelson, K. (1985). *Making sense: The acquisition of shared meaning*. New York: Academic Press.
- Nelson, K. (1996). *Language in cognitive development*. New York: Cambridge University Press.
- Searle, J. (1969). *Speech acts, an essay in the philosophy of language*. Cambridge: Cambridge University Press.
- Searle, J. (1995). *The Construction of Social Reality*. New York, NY: Free Press.
- Searle, J. R. (1999). *Mind, language and society. Philosophy in the real world*. New York: Basic Books
- Seyfarth, R. M., & Cheney, D. L. (2003). Signalers and receivers in animal communication. *Annual Review of Psychology* 54, 145-173.
- Shwe, H. I., & Markman, E. M. (1997). Young children's appreciation of the mental impact of their communicative signals. *Developmental Psychology*, 33(4), 630-636.
- Spencer, P. (1993). Communication behaviors of infants with hearing loss and their hearing mothers. *Journal of Speech and Hearing Research*, 36, 311-321.
- Sperber, D. (1994). Understanding verbal understanding. In J. Khalfa (Ed.), *What is Intelligence?*(pp. 179-198). Cambridge University Press.
- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Harvard University Press.
- Tomasello, M. (1992). The social bases of language acquisition. *Social Development*, 1(1), 67-87.
- Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Harvard U. Press.
- Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press
- Tomasello, M. (2008). *Origins of Human Communication*. Cambridge, MA: MIT Press.
- Tomasello, M. (2014). *A natural history of human thinking*. Harvard University Press.
- Tomasello, M. (2018). Where's the cooperation? Commentary published in Philip Pettit's volume *The Birth of Ethics*. Oxford University Press.
- Tomasello, M., Call, J., & Hare, B. (1998). Five primate species follow the visual gaze of conspecifics. *Animal Behaviour*, 55, 1063-69.
- Tomasello, M., Carpenter, M., & Lizskowski, U., (2007). A new look at infant pointing. *Child Development*, 78, 705-22.
- Tuomela, R. (2007): *The Philosophy of Sociality: From A Shared Point of View*. Oxford: Oxford University Press
- Warneken, F., Chen, F., & Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child Development*, 77, 640-663.
- Wittgenstein, L. (1953). *Philosophical investigations*. Oxford: Basil Blackwell.
- Wittgenstein, L. (2005). *The big typescript: TS 213*. Oxford: Blackwell Publishing.