XV Jornadas Peruanas de Fenomenología y Hermenéutica, 20-22 de noviembre del 2019, Lima

"En la infancia, el niño está atraído por el lenguaje": Merleau-Ponty y los orígenes socio-afectivos del habla

Kee, Hayden
Center for Human Nature, Artificial Intelligence, and Neuroscience, University of Hokkaido
mail@haydenkee.com
ORCID ID: 0000-0003-0893-2075

Resumen: Siguiendo el espíritu de Merleau-Ponty, el presente artículo tiene como objetivo explorar las fuentes afectivas y motivacionales de la unicidad humana en el desarrollo de la cognición social y el lenguaje. Para esto, el autor se propone considerar algunas aproximaciones dominantes a la unicidad humana centradas en el lenguaje y cognición, como las de Chomsky y Tomasello. Luego, postula su alternativa merleau-pontiana, apoyada en una discusión de investigaciones recientes acerca del aspecto afectivo y motivacional en la infancia. Posteriormente, bosqueja una propuesta explicativa de desarrollo que considera estos aspectos como los más básicos y finales de la cognición social, el lenguaje y la cultura. Para cerrar, el autor revisa un caso de estudio que muestra el punto de unión entre Merleau-Ponty y Tomasello.
Palabras clave: lenguaje; Merleau-Ponty; Chomsky; Tomasello

"In infancy, the child is attracted to language": Merleau-Ponty and the social-affective origins of speech

Abstract: Following Merleau-Ponty's spirit, this article aims at exploring the affective and motivational sources of human uniqueness in the development of social cognition and language. In order to do so, the author sets himself to consider some of the dominant language and cognition-centered approaches to human uniqueness in the work of Noam Chomsky and Michael Tomasello. Then, he develops his Merleau-Pontian alternative, supported on a discussion of recent research on uniquely human affective and motivational characteristics in early infancy. After that, he outlines a general explanatory-developmental account that takes into consideration these features as the ultimate and basic origin of species-unique social cognition, language, and culture. He closes discussing a test case which shows a intersection between Merleau-Ponty and Tomasello.

Keywords: language; Merleau-Ponty; Chomsky; Tomasello

§ 1. Introduction

This paper attempts to unearth what Merleau-Ponty called an "existential unity" of human language.[1] It is not, however, an exegetical presentation of his views on the topic. It

[1] See especially (Merleau-Ponty 1969; 1959; 1964).

follows rather the spirit and style of Merleau-Ponty's work, and his unique way of integrating phenomenological and scientific inquiry. After all, as Merleau-Ponty asks,

> How could there be a division between the science of expression, provided it conceives its subject as broadly as it should, and the lived experience of expression, where it is lucid enough? Science is not devoted to another world but to our own; in the end it refers to the same things that we experience in living. (Merleau-Ponty 1969: pp)

This spirit is timely as ever, and it informs much current research at the intersection of phenomenology and the cognitive sciences. However, much has changed in the sciences of mind and language since Merleau-Ponty's day. This paper is meant to be both a phenomenological critique as well as a constructive interpretation and synthesis of recent directions in the empirical inquiry into language and human nature.

My objective here is to explore affective and motivational sources of human uniqueness in the development of social cognition and language. I will ask whether these factors might properly be the most basic and ultimate origins of human uniqueness, as opposed to more specifically cognitive or linguistic factors. The latter on my hypothesis may turn out to be downstream consequences of basic, species-unique affective and motivational traits.

I begin by considering some of the dominant language and cognition-centered approaches to human uniqueness in the work of Noam Chomsky and Michael Tomasello. I then develop my Merleau-Pontian alternative through a discussion of recent research on uniquely human affective and motivational characteristics in early infancy. I then provide an outline of how a general explanatory-developmental account could locate these factors as the ultimate and basic origin of species-unique social cognition, language, and culture that develop from them. I close by briefly discussing a test case: the species-unique form of human perspectival perception and cognition identified by both Merleau-Ponty and Tomasello.

§ 2. Cognitivist Views of Human Uniqueness: The Zoon Logon Echon

Attempts to determine what is unique about the human being within the animal kingdom have long focused on *language* and rationality as the definitive characteristic of our species. This determination was codified in Aristotle's well-known definition of human being as the *zoon logon echon*—the animal that has *logon*: the power of reason, or speech. With this one densely polysemous word, language and reason came into a close association that would be definitive of the human in the Western tradition down to the present day. But what precisely is unique about human language and common to all human languages, such that language could serve to differentiate the human within the animal kingdom? Other animals, we know, communicate in ways that show some similarities to human communication. And, even more fundamentally, what, if anything, is the basic underlying difference in our human nature that allows us to acquire, produce, and comprehend language?

In recent decades, the search for the universal structural characteristics of all human languages and the underlying cognitive architecture of the human mind that allows for language has led to fervid debates centering around the work of Noam Chomsky and Generative Linguistics. Chomsky identifies the "basic property of language," the explanandum for the sciences of language, as follows: each language, from finite resources, can generate an infinite array of hierarchically ordered propositions. (Chomsky 2013: 647, 654) The underlying cognitive ability that allows for this basic property of language would, then, be what is unique to the species and common to all normally-developing members of it. Chomsky identifies this as *Merge*, the computational operation through which two syntactic objects are combined into one.

There are many weighty technical objections to Chomsky's characterization of Merge and the human language faculty. (These objections concern the possibility of its evolution; [Studdert-Kennedy and Terrace 2017] whether language-specific learning mechanisms are even required to explain language acquisition; [Tomasello 2003] and whether in fact all languages exhibit the character of recursion that Chomsky invokes Merge, in part, to explain [Tomasello 2009].) But prior to this, a phenomenologist, or perhaps simply an unbiased onlooker, might ask the following question: How has Chomsky decided that the infinite productivity of language should count as its "basic property," and hence as the central fact to be explained by a science of language? True, once we objectify and formalize language, construing it according to our theoretical demands, we can arrive at this remarkable logico-syntactic property of language: from finite means, natural languages can output logically infinite possible propositions. But why, and in what sense, should this be viewed as the (or even *a*) *basic* property of language? Beginning from a more original experience of language, we might find the *actual* existence of a single expression a much more remarkable and basic fact to be explained than the logical (and quite trivial) *possibility* of infinite propositions.

With Chomsky's claim for a modular cognitive adaptation for language thrown into question on various fronts, it might do well to look elsewhere for the uniquely human characteristic(s) that enable(s) language. Michael Tomasello, a prominent primatologist and developmental psychologist, has done just this. He argues that our linguistic ability should be viewed as secondary in the discussion of human uniqueness, emerging out of more basic social-cognitive abilities that are foundational for it. (Tomasello 2008) Human beings are, as Tomasello puts it, "the world's experts at mind-reading." It is this basic ability to understand the intentions of our fellow humans that allows us to undertake basic collaborative endeavors together in the here and now. It further allows us to acquire a theory of mind, ultimately leading to the macroscale achievements characteristic of our culture, such as intergenerational transmission of knowledge, sophisticated tool use, language, and an objective perspective on reality. The difference that makes the difference between us and other animals, according to Tomasello, is social-cognitive.

And yet, this position, too, has proven increasingly difficult to define and defend in recent years. Other animals, our fellow primates in particular, also possess social-cognitive abilities. What, then, distinguishes human social cognition from that of other species? It was

long believed that the ability to understand *false belief* was the—or at least *a*—decisive differentiator: I understand that other intelligent beings act on the world *as they believe it to be*, rather than as it actually is. If another human has a false belief about the world, she will act in the world as she believes the world to be, not as it in fact is. Non-human primates and other animals, it was thought, are incapable of understanding that another animal may have a *false* belief about the way the world is. But this accepted view has been challenged by recent research (from Tomasello's own research group [Krupenye et al. 2016]) suggesting that several species of non-human primate *are* in fact capable of anticipating behavior based on false beliefs, and hence possess at least an implicit understanding of false belief.

We face the following challenge: Viewing our human cultural-linguistic form of life as a whole, it seems that there is something importantly different about it with respect to the social-communicative behaviors exhibited in non-human animals. Further, even chimps raised in captivity, though they sometimes achieve a degree of symbolic competence, never acquire anything like normal human language. Nonetheless, when we attempt to identify basic *linguistic* or *cognitive* differences to account for these large-scale differences in forms of life and behaviors of individuals and species, we come up emptyhanded. Tomasello himself sums up the explanatory challenge like this:

> [H]ow do human individuals come to the species-unique cognitive and social abilities necessary for participating in cultural coordination and transmission? To answer this question the obvious first step is to establish exactly how human psychology differs from that of other primates—precisely how humans as individuals are unique. The difficulty is that over the past few decades empirical research has established that humans' nearest living relatives, the great apes, possess cognitive and social skills highly similar to those of humans, including many that are seemingly relevant to cultural processes. For example, there is recent research demonstrating that at least some great apes (1) make and use tools, (2) communicate intentionally (or even "linguistically"), (3) have a kind of "theory of mind," (4) acquire some behaviors via social learning (leading to "culture"), (5) hunt together in groups, (6) have "friends" with whom they preferentially groom and form alliances, (7) actively help others, and (8) evaluate and reciprocate one another's social actions. (Tomasello 2019: 4)

Tomasello's solution to this problem is, in brief, more fine-grained description of these abilities and behaviors, with a focus on the subtly different *ways* in which humans and other primates exhibit the "same" behavior or possess the "same" ability. Specifically, close attention to ontogeny will show how a uniquely human variation on great ape cognition emerges in the early years of life. On the balance of evidence, Tomasello infers that "early in ontogeny humans already employ *unique adaptations for social cognition*, which enable them to master, in a way that other apes cannot, the cultural skills of communication, cooperation, and social learning." (Tomasello 2019: 28, emphasis mine)

The ensuing account of human uniqueness remains largely focused on the cognitive aspects of human sociality, however (though Tomasello does mention the importance of human pro-sociality—Tomasello's points 6 and 7 in the quoted passage). It places less emphasis on the role of distinctively human *motivations* and *affectivity* in the phylogenetic

and ontogenetic development of a uniquely human way of life. Tomasello seeks to "invoke basic processes of shared intentionality as *the ultimate source* of human uniqueness," (Tomasello 2019: 86, emphasis mine) where shared intentionality is understood as a social-cognitive ability. In the following I will ask whether species-unique affective and motivational traits might constitute a more basic and ultimate source of human uniqueness. Whether this contribution merely constitutes a complementary account, or whether it is a genuinely contrasting and alternative account to Tomasello's more cognition-centered account is something I will address below.

§ 3. An Alternative View: Affective Origins of Language in Ontogeny

What if—and this is a proposal I glean from Merleau-Ponty—the "basic" difference that makes the difference is not to be found on the level of language (narrowly construed, as per Chomsky) or social cognition (as per Tomasello), but rather on the level of *affectivity and motivation*? What if we *become* the speaking animal not because of a basic adaptation for language or social cognition, but because of the way language and the social world *affect* us and elicit in us deep and persistent *desire* to speak?

Merleau-Ponty's critique of attempts such as Chomsky's or Tomasello's to identify what is unique and universal about human language or cognition is that they seek to situate our uniqueness on the level of *concepts* and *reason*. But the human drama does not unfold initially and for the most part, basically and ultimately, on the level of concepts and reason, but rather on the level of *existence*. Hence Merleau-Ponty invites us to look instead for "a concrete universality that is only realized little by little and is found in the *expressive will* that animates languages more than in the transitory forms which it reaches." (Merleau-Ponty 1964) Or, again, the universal features of language "are simply to be conceived as a dimension which is neither that of the concept nor of essence, but of *existence*." (Merleau-Ponty 1969) At the root of the specific historical forms language takes, and driving the acquisition of social cognition and language in the life of an individual, there is a "*drive* [*poussée*] of speaking subjects who *want* to understand one another." (Merleau-Ponty 1969: pp.) "The driving force of language is the *will* to communicate." (Merleau-Ponty 1964: pp.) And, at the root of this drive, this desire, this motivation, is a basic fact of affectivity: in infancy, there is a "general *attraction* to language." (Merleau-Ponty 1964: pp, emphasis mine in citations throughout this paragraph).

This emphasis on the deep human desire to express and commune suggests a new perspective and hypothesis concerning the uniqueness of the human language faculty: could it be that the human mind is not originally radically different from those of other higher animals where cognition, reason, and conceptuality are concerned—that is, that we do not possess any basic and unique cognitive modules or learning mechanisms that can't be found in other higher animals—but that we *acquire* the kinds of minds we have because of our initial affective receptivity to the surrounding social and linguistic world, and because of our intense desire to participate in this world of expression and communion?

The proposal is the following: I suggest the difference is that human beings are deeply—indeed, innately—*attuned* and *attracted* to the social and linguistic world. As such, we are *motivated* to participate in the extensive training program required for human social, linguistic, and cultural life. In undertaking this training program, we may not require any basic learning mechanisms different in kind from what other higher animals possess. We just have to be (1) surrounded by a cultural environment designed to facilitate the relevant kind of learning, (2) attuned to that cultural world in the right way, and (3) motivated to practice intensively participating in it for a long time.[2]

What positive evidence speaks in favor of my affective-motivational approach? I will here touch on five considerations: (1) the infant's (and, indeed, the fetus's) precocious attunement and attraction to phonetic and musical qualities of language; (2) babbling; (3) proto-conversation; (4) declarative pointing; (5) prosociality.

We begin with the infant's most original affective attunement to language, which is present even before birth. In the womb, the fetus is receptive to rhythmic, melodic, and vocalic features of the mother's language, as is evidenced by variation between the rhythmic, melodic, and vocalic character in the cries of newborn infants with different maternal languages. (Wermke et al. 2016; Mampe et al. 2009; Prochnow et al. 2019; Moon, Lagercrantz, and Kuhl 2013; Trevarthen 2017) At birth, neonates can even learn to discriminate very quickly between some languages other than their maternal language, though by the end of the first year they can only discriminate the phonemes of languages spoken in their environment. They show a significant preference for the language spoken by their mother during pregnancy, and a general preference for their own language over others regardless of who speaks it.

The productive counterpart of these discriminative abilities and preferences is an innate propensity for babbling, from which infants derive immense pleasure. Babbling is, among other things, a practice for the fine motor skills required to produce the phonemes and syllables of the native language. Over the course of the first year, the largely inarticulate vocalizations of the first months begin to converge towards the phonemes and syllables spoken in the ambient language environment. If the infant were not intrinsically motivated to attend to and produce speech sounds in babbling, it would never progress in the extensive training regime of listening and speaking required for acquiring a language. The right kind of affective, receptive, and motivational relationship to the surrounding world of language, then, is at least a necessary condition for the normal course of language acquisition in the first year of life.[3]

[2] A *bigger* brain may furnish the processing power required for the high cognitive demand of human socio-cultural life. However, my hypothesis is that the brain does not need to be different in kind with respect to basic modules or learning mechanisms than the brains of other higher animals. Of course, the human being has also evolved the physiological, perceptual-articulatory apparatuses required for production and perception of speech. Though these factors should not be treated as irrelevant to an embodied and enactive understanding of human uniqueness, they are not the focus of the present inquiry.

[3] I should emphasize that I am here using the term "normal" in a statistical sense, not in an evaluative or normative sense. For example, infants whose first language is a signed language will "babble" in a manual-

Throughout the first year, and still prior to any verbal linguistic behavior in the narrow sense, the infant enjoys practicing its coos and goos, *bas* and *das*, along with its growing repertoire of bodily and facial expressive and imitative abilities, in face-to-face "proto-conversations" with caretakers. In proto-conversation, movements of hands, eyes, and mouths, the infant's expressions are rhythmically timed to alternate with those of the interacting caretaker. (Trevarthen 1979; Stawarska 2009: 103 ff.) In these exchanges, we see a "give-and-take, address-and-reply pattern" that is structurally isomorphic to the adult dialogical conversation format that it anticipates. Verbal imitation begins here, as the infant attempts to adjust its innate disposition to babble to match the experienced voice of the interlocutor. Some vocalizations come to be ritualized and repeated consistently, forging connections between certain verbal-expressive gestures and the affective, intersubjective significance that accompanies them. The infant also gains an appreciation for the *interchangeability* of verbal and gestural communicative exchange: the same gestures are available to both you and me, and can be assumed by either of us. This is a crucial pragmatic insight into the functioning of language, and into human social life more generally, where individuals may take on interchangeable roles within a group activity. Proto-conversation is a form of social interaction unique to humans, with no direct equivalent in higher animals, even though infant chimpanzees generally spend considerably more time in close bodily contact with their mothers than human infants in many cultures do.[4]

When infants advance from the dyadic, I-you structure of the proto-conversation, to a triadic, I-you-object structure of joint attention towards the end of the first year of life, they again display species-unique patterns of motivation in their interactions. In joint attention, you and I are attentionally directed towards the same object, and we are both aware that the other is directed to the same object. We are directed towards the object *together*. A crucial behavior for directing joint attention is the pointing gesture, which human infants acquire

signed modality just as infants whose first language is a spoken language babble in a vocal-linguistic modality. Babbling, then, must be seen as a modality-independent manifestation of a more basic expressive-communicative desire, rather than being narrowly tied to the verbal modality. See Petitto and Marentette 1991. Interestingly, however, even the congenitally blind still produce spontaneous bodily gestures to accompany their speech, which may provide some evidence for a primacy of the bodily-gestural-visual modality in human expression and community, a hypothesis that complements the "gesture-first" view of language evolution.

[4] The developmental psychologist Peter Hobson writes, "What chimpanzees lack is the *capacity to identify with others through emotional engagement*—an ability that, as we have seen, transforms the intellectual life of a baby. My bet is that, *if* chimpanzees acquired this ability, their intellectual life would be transformed, too. I once visited the Yerkes compound for chimpanzees in Atlanta[.] […] I sat gazing at a chimpanzee who sat on the other side of a fence, gazing at me. As a psychoanalyst, I have been taught to analyze the countertransference, which means that I try to formulate how this individual is making me feel. So I sat there and tried my very hardest to do just that. I felt . . . something missing. I could not connect. […] It was as if this chimpanzee was not at home, mentally speaking. Or at least I was not entering a home, in his mind. I wondered if *this* was because we belonged to different species. Would it be different if I were a chimpanzee? I seriously doubt it. Primatologists tell me that chimpanzees indulge in a lot of rough and tumble together[.] […] What they do not do is to spend time gazing into each other's eyes, or engaging in the kinds of intense face-to-face interpersonal communication that we see in human infants. They do not dwell in each other's expressions. […] [T]hey do not enter the subjective lives of their fellow chimpanzees. They never (or almost never) show things to each other, nor do they appear to share experiences of the world with others." (Hobson 2004: 269 ff.)

cross-culturally towards the end of the first year. Merleau-Ponty believed that only humans acquire the pointing gesture, (Merleau-Ponty 1962) but this is not true: many apes reared in captivity learn it with a range of applications, there are some quite limited uses evidenced by primates in the wild, and corvids have been demonstrated to employ a form of deictic signaling.[5] But while primates raised in captivity will point with an *imperative* motive (they want someone to do something for them) or in trained experimental contexts (they know they will get treats if they point to objects on a screen), young children engage in a form of joint attention that appears to interest no other animal on our planet. Call it *demonstrative*, or *declarative*, pointing, as opposed to *imperative* pointing.[6] Young children will often point to an object, or present it to a caretaker or companion, not because they want the other person to *do* something with it, but simply for the sake of establishing and maintaining joint attention itself. From a cognitive development perspective, it is sometimes said that children do this to learn about the attitudes others have towards objects, and to see what spontaneous use others will make of objects. And they undoubtedly *do* learn these things along the way. However, we may question whether it is the correct interpretation of the lived motivations of behavior to say that children do this *in order to* learn these things. From an existential, affective, motivational perspective, infants may be engaging in demonstrative pointing and sharing for the sheer joy of sharing the world with others.[7]

Finally, a word should be said about the prosociality of human children once they begin to understand more of the practical adult behavior going on in the surrounding world. (Silk 2013) Here the findings come from experiments with humans and chimps in prosocial games and instrumental helping tasks. In the prosocial game, one participant gets to choose whether it wants to take a reward just for itself (1-0 option) or whether it would like to have a reward for itself and for another member of its social group from whom it is separated by a transparent partition (1-1 option). In an example of an instrumental helping task, the participant's help is solicited by a familiar caretaker who requires a tool to obtain an objective. Control conditions are used to ensure subjects understand the task. In these experiments, chimpanzees demonstrate a certain level of prosociality. But given how low-cost it is for them to help (in the prosocial game, for example, there is no loss to the test subject whatsoever for choosing the prosocial option), what is curious is that they do not help *more often*. In a meta-analysis of existing studies, Silk calculates that chimpanzees choose prosociality about half the time, and summarizes that "chimpanzees effectively flip a coin

[5] On primate pointing, see Leavens, Hopkins, and Bard 2005. For corvids, see Pika and Bugnyar 2011.
[6] See Tomasello 2008.
[7] It is often said that in autism, this basic desire to share the world may be missing or attenuated. But perhaps, at least in some cases, it simply takes an unusual form. Oliver Sachs tells the story of two autistic twins who would sit in a corner for long periods of time exchanging long strings of numbers. One twin would utter a figure, then after a moment's contemplation, the other would smile in approval and both would relish the figure for a moment before the other continued the game with another figure. There seemed to be some rule and rationale to their exchanges, but for a long time no one could decipher what it was. Eventually, Sachs cracked their code: the twins were exchanging and contemplating prime numbers. The desire to share the world took a most abstract form for the twins. But it was present nonetheless. (Thank to Anthony Steinbock for bringing this example to my attention.)

each time they are presented with the opportunity to help." (Silk 2013: 369) Human children, by contrast, volunteer their help and choose prosocially significantly more often.[8] Again, as the control conditions suggest, there is no difference in cognitive ability. The difference here is strictly in degree of motivation.

If these considerations are on the right track, they draw us back to a *more basic* feature of language than the infinite productivity mentioned by Chomsky and to a more fundamental aspect of the human language faculty and experience than the marvelous mind-reading abilities celebrated by Tomasello. We begin to develop a deeper appreciation of what it means for the human being to be the *zoon logon echon*, the animal that speaks. We can unpack that classic definition now, and say the following:

1. From infancy, the human being is *spoken to by speech* on the level of receptivity and *affectivity*: speech calls to us, appeals to us, attracts us (*die Sprache spricht uns an, le langage me dit quelque chose, el lenguaje nos llama*[9]);
2. Addressed by speech in this way, the human being has a strong *desire* and *motivation* to participate in linguistic life (in existential terms, human being is *dying to speak*);
3. Human being is thus willing to undertake the extensive regime of socio-culturo-linguistic training required for the acquisition of cultural competence and linguistic facility (from babbling and proto-conversation all the way to studying foreign languages, learning professional jargons, and crafting poetry);
4. Human being thus eventually becomes the animal that speaks (and—here a nod to Chomsky—is in principle capable of unlimited linguistic productivity based on the finite means of a historical language).

On this understanding, the higher cognitive, rational, and linguistic (in the narrow sense) features are no longer the basic features of the human language experience. They are downstream consequences. More basic than the rational, conceptual ability to speak language, and the necessary condition thereof, is the affective, existential experience of being spoken to, appealed to, by language, and wanting to respond.

§ 4. A Test Case: The Perspectival Character of Perception in Human and Chimpanzee

Thus far, we have only a hypothesis. It stands in need still of considerable elaboration *qua* hypothesis as well as further empirical corroboration. In this section, by way of conclusion, I will attempt to contribute something to the elaboration of such hypothesis.

I am suggesting that motivational and affective differences between humans and our closest non-human kin may be the *basic* differences between us and them, and the *ultimate* causes of higher-order differences in cognition and language. To elaborate the hypothesis, we must take an example of some such higher-order cognitive or linguistic difference and

[8] Note, however, that these considerations may only indicate a difference in *degree* of motivation and do not necessarily suggest motivations different in *kind* in chimps and humans. It is admittedly unclear what the motivating factors are for the older children studied in these experiments. Hence, I find more compelling the evidence for uniquely human prosociality, affectivity, and motivations from the studies of early infancy.

[9] Thanks to Celeste Vecino for advising about the Spanish.

explain how it can be acquired based on other basic cognitive abilities that we share with other primates *if* the additional, species-unique motivational and affective structures, as well as a suitable environment to scaffold the relevant learning experiences, are in place.

For our test case, let us consider the human ability—unique, so far as we know, in the animal kingdom—for grasping the perspectival nature of perception. Human beings understand that it is possible for one and the same object to appear differently to different individuals—and, further, that beyond the individual perspectives we may have, in many cases there may be something like an independent, objective standard of correctness, however this is to be understood. Our close primate kin somehow seem to come up short of this insight. Astonishingly, more than 75 years ago, Merleau-Ponty was already describing the perception of higher animals along these lines and in terms very much akin to those employed by Tomasello and others today.[10] Merleau-Ponty writes,

> [T]he box-as-seat and the box-as-instrument are two distinct and alternative objects in the behavior of the chimpanzee and not two *aspects* of an identical *thing.* In other words, the animal cannot at each moment adopt a point of view with regard to objects which is chosen at its discretion. (Merleau-Ponty 2002: pp)
> Animal activity […] loses itself in the real transformations which it accomplishes and cannot reiterate them. For man, on the contrary, the tree branch which has become a stick will remain precisely a tree-branch-which-has-become-a-stick, the same *thing* in two different functions and visible *for him* under a plurality of aspects. (Merleau-Ponty 2002: pp; Merleau-Ponty 2010, 430, and Tomasello 2019: 15 ff.)

Assuming this description is correct, it is rather interesting that chimpanzees do not arrive at this aspectual understanding of perception. For, as we have seen, in some cases, they are indeed at least implicitly capable of recognizing that others have false beliefs and of acting on the way others take the world to be, not the way it is. And yet, to arrive at the understanding of the perspectival nature of experience in the full sense, it seems it would be necessary not only to predict the behavior of the other based on the other's false belief, but then further to actively juxtapose the other's false belief with my own correct one, becoming aware of the discrepancy as such. This the chimpanzee never appears to do. It is as though when anticipating the behavior of the other chimp, the observing chimp is so fully immersed in the behavior of the other that it never even feels the conflict between its true belief and the other's false belief. In Merleau-Ponty's terms, it is as though in the moment of anticipating the other's behavior, the observing chimpanzee "loses itself" in a transformation of the perceptual world that it has achieved but cannot reiterate or juxtapose this with its own perspective after the fact. It is never perplexed, that is, by the *paradox of perspective*. Hence Tomasello, concurring with Merleau-Ponty, writes

[10] I say the coincidence is astonishing because in Merleau-Ponty's time there was precious little serious ethological work on non-human primates available. Merleau-Ponty seems to have based his views almost exclusively on Wolfgang Köhler's *The Mentality of Apes* (Köhler 1927), while today's researchers can build on vast experimental and ethological records that have been amassed in the ensuing decades. For more on Merleau-Ponty's appropriation Of Köhler's research with apes, see Kee 2018.

> [P]erspective-taking [in the strict sense] requires that a subject imagine more than one way of perceiving or understanding a given entity or situation; there can be no such thing as a single perspective on something without at least the possibility of other perspectives on it. What we have called "imagining what another sees and knows" [i.e., in chimpanzees' implicit understanding of what others know] does not involve different perspectives in this same way; the subject is imagining the other's experience, but her own experience is not part of her mental processing. (2019: 64, bracketed glosses mine)

In this sense, Tomasello maintains, the chimpanzee's cognition is "nonperspectival." Apes can track what a conspecific directly perceives,

> with their own perception of the situation not an object of attention at all (the participant is seeing "through" his perceptual experience, not examining it or comparing it to something else). When the participant is imagining the competitor seeing something (or not), he is simply tracking her perceptual experience—full stop—irrespective of what he himself is or is not experiencing. Without an awareness of multiple potential ways of seeing the situation, an individual cannot be said to be taking perspectives at all. (2019: 64)

For my purposes, then, the challenge becomes that of describing how properly perspectival human ways of interacting with the world and others can emerge on the basis of cognitive and perceptual abilities that we share with other primates *if* the right motivational factors and environmental scaffolding for learning are in place. A microgenetic account of the acquisition of species-unique perspectivality on the basis of these factors, then, would constitute the elaboration of my hypothesis. In brief, the account would run as follows:

> (1) Humans acquire a perspectival perception and cognition because *motivationally* they are driven to share and show themselves and the world with others in ways that chimpanzees never are. This was illustrated above in the species-unique practices of proto-conversation and declarative pointing. The latter is of especial interest here, for it is through triangulating between myself, an object, and another that I come to learn about perspective. There may be a steady developmental trajectory through which different elements of perspectivality are discovered. An infant may first start to recognize that different individuals may have conflicting *preferences* for one and the same object when they bring something, say a toy that they like, to a caretaker, who then (perhaps playfully) makes a face of disapproval. Later, they learn that different individuals have a different *perceptual perspective*, and hence may have different information, about one and the same object. Only then might they begin explicitly to understand that different individuals may have contrasting or conflicting *beliefs* about one and the same thing.[11]
>
> (2) Because it is intrinsically aspectual, *language* acquisition also seems to play a role in children's gradual acquisition of human perspectivality. In early stages of language

[11] On the sequential development of children's understanding of various theory of mind tasks, see Wellman and Liu 2004; Wellman, Fang, and Peterson 2011; Shahaeian et al. 2011.

> acquisition, younger children struggle tremendously even to accept that one and the same object may have different names. But with increased practice, they learn to navigate the aspectuality of language, and this likely scaffolds their developing understanding of the perspectivality of cognition and perception. This claim is supported by studies using perspective-shifting discourse with children who have not yet mastered false belief tests but who are at about the age when children usually do acquire an understanding of false belief. Children were given training sessions with perspective-shifting discourse, in which the child is asked what she thinks an object is, then is asked again in light of new information suggesting it is something else, then is asked what she previously thought it was. Scaffolding the experience with discussion helps children improve on false belief tests, whereas if the same scenario unfolds without a supporting linguistic exchange, no significant improvement on false belief tests results. Training with propositional attitude verbs ("he knows," "she thinks," "they believe") also led to improvement on false belief tests. (Lohmann and Tomasello 2003)
>
> (3) Finally, it would have to be shown that with the linguistic environment and suitable motivations in place, the child acquires human perspectival cognition and perception without relying on any different-in-kind learning mechanisms or cognitive modules. Making this case might be difficult. However, in brief, the claim here would be that there are perceptual learning mechanisms common to higher animals responsible for the development in question. The child's global patterns of perception and interaction may undergo a restructuration through the course of social and linguistic learning, such that a perspectival way of seeing things, which might be very difficult to achieve at first, comes to be habitualized in perception and thought. Merleau-Ponty discussed similar reorganizations of affective and perceptual structures in his work on developmental psychology, (Merleau-Ponty 2010) and these might provide a model for thinking about the kind of acquisitional development at play here.

My proposal, then, amounts to something more than a mere shift in emphasis compared to Tomasello's view. Where Tomasello claims that it is our social-cognitive abilities that are the ultimate and basic source of the difference between us and other animals, I propose that our social-cognitive abilities may actually emerge as a natural consequence of a more ultimate and basic difference in our social-affective-and-motivational orientation. This proposal comes in strong and weak flavors. The weaker version holds that our social-affective orientation is *a* basic and ultimate cause of human uniqueness. The stronger version holds that it is *the* basic and ultimate cause of human uniqueness. The weak version allows that there may be some basic and ultimate differences in kind between human cognition and the cognition of non-human animals. The strong version insists that any basic cognitive differences are only differences in kind, and that cognitive differences in linguistic and social-cognitive abilities that emerge later in ontogeny simply follow entirely as a

consequence of our unique social-affective orientation. If increased cognitive power plays a role, this is, for the strong version of the proposal, a difference in degree rather than in kind.

References

Chomsky, Noam. "The Dewey Lectures 2013: What Kind of Creatures Are We? Lecture I: What Is Language?". *Journal of Philosophy* 110, 12 (2013): 645–662.

Hobson, Peter. *The Cradle of Thought: Exploring the Origins of Thinking*. Oxford: Oxford University Press, 2004.

Kee, Hayden. "Phenomenology and Naturalism in Autopoietic and Radical Enactivism: Exploring Sense-Making and Continuity from the Top Down". *Synthese*. June, 1-21, 2018.

Köhler, Wolfgang. *The Mentality of Apes*, trans. Ella Winter. London: Routledge and Kegan Paul, 1927.

Krupenye, Christopher, Fumihiro Kano, Satoshi Hirata, Josep Call, and Michael Tomasello. "Great Apes Anticipate That Other Individuals Will Act According to False Beliefs". *Science* 354, 6308 (2016): 110–14.

Leavens, David A., William D. Hopkins, and Kim A. Bard. "Understanding the Point of Chimpanzee Pointing". *Current Directions in Psychological Science* 14, 4 (2005): 185–89.

Lohmann, Heidemarie, and Michael Tomasello. "The Role of Language in the Development of False Belief Understanding: A Training Study" *Child Development* 74, 4 (2003): 1130–44.

Mampe, Birgit, Angela D. Friederici, Anne Christophe, and Kathleen Wermke. "Newborns' Cry Melody Is Shaped by Their Native Language". *Current Biology* 19, 23 (2009): 1994–1997.

Merleau-Ponty, Maurice. *Signes*. Paris: Gallimard, 1959.

Merleau-Ponty, Maurice. "Un Inédit de Maurice Merleau-Ponty". *Revue de Métaphysique et de Morale* 67, 4 (1962): 401–409.

Merleau-Ponty, Maurice. "La Conscience et l'acquisition Du Langage". *Bulletin de Psychologie* 18, 3–6 (1964): 226–59.

Merleau-Ponty, Maurice. *La Prose Du Monde*. Paris: Gallimard, 1969.

Merleau-Ponty, Maurice. *La Structure Du Comportement*. Paris: Gallimard, 2002.

Moon, Christine, Hugo Lagercrantz, and Patricia K Kuhl. "Language Experienced in Utero Affects Vowel Perception after Birth: A Two-Country Study". *Acta Paediatrica* 102, 2 (2013): 156–60.

Petitto, L. A., and P. F. Marentette.. "Babbling in the Manual Mode: Evidence for the Ontogeny of Language". *Science* 251, 5000 (1991): 1493–96.

Pika, Simone, and Thomas Bugnyar. "The Use of Referential Gestures in Ravens (Corvus Corax) in the Wild". *Nature Communications* 2 (2011): 560.

Prochnow, Annette, Soly Erlandsson, Volker Hesse, and Kathleen Wermke. "Does a 'Musical' Mother Tongue Influence Cry Melodies? A Comparative Study of Swedish and German Newborns". *Musicae Scientiae* 23, 2 (2019): 143–56.

Shahaeian, Ameneh, Candida C. Peterson, Virginia Slaughter, and Henry M. Wellman. "Culture and the Sequence of Steps in Theory of Mind Development". *Developmental Psychology* 47, 5 (2011): 1239–47.

Silk, Joan B. 2013. "The Origins of the Prosocial Ape: Insights From Comparative Studies of Social Preferences". In *Navigating the Social World: What Infants, Children, and Other Species Can Teach Us*, edited by Mahzarin R. Banaji and Susan A. Gelman, 367–70. Oxford: Oxford University Press.

Stawarska, Beata. *Between You and I: Dialogical Phenomenology*. Athens, Ohio: Ohio University Press, 2009.

Studdert-Kennedy, Michael, and Herbert Terrace. "In the Beginning: A Review of Robert C. Berwick and Noam Chomsky's Why Only Us". *Journal of Language Evolution* 2, 2 (2017): 114–25.

Tomasello, Michael. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, Mass: Harvard University Press, 2003.

Tomasello, Michael. *Origins of Human Communication*. Cambridge, Mass: MIT Press, 2008.

Tomasello, Michael. "Universal Grammar Is Dead". *Behavioral and Brain Sciences* 32, 5 (2009): 470–471.

Tomasello, Michael. *Becoming Human: A Theory of Ontogeny*. Cambridge, Massachusetts: Harvard University Press, 2019.

Tomasello, Michael, Malinda Carpenter, Josep Call, Tanya Behne, and Henrike Moll. "Understanding and Sharing Intentions: The Origins of Cultural Cognition". *The Behavioral and Brain Sciences* 28, 5 (2005): 675–91; discussion 691–735.

Trevarthen, Colwyn. "Communication and Cooperation in Early Infancy: A Description of Primary Intersubjectivity". In *Before Speech*, edited by M. Bullowa, 321–47. Cambridge: Cambridge University Press, 1979.

Trevarthen, Colwyn. "Maternal Voice and Communicative Musicality: Sharing the Meaning of Life from Before Birth". In *Early Vocal Contact and Preterm Infant Brain Development: Bridging the Gaps Between Research and Practice*, edited by Manuela Filippa, Pierre Kuhn, and Björn Westrup, 3–23. Cham: Springer International Publishing, 2017.

Wellman, Henry M., Fuxi Fang, and Candida C. Peterson. "Sequential Progressions in a Theory-of-Mind Scale: Longitudinal Perspectives". *Child Development* 82, 3 (2011): 780–92.

Wellman, Henry M., and David Liu. "Scaling of Theory-of-Mind Tasks". *Child Development* 75, 2 (2004): 523–41.

Wermke, Kathleen, Johanna Teiser, Eunice Yovsi, Paul Joscha Kohlenberg, Peter Wermke, Michael Robb, Heidi Keller, and Bettina Lamm. "Fundamental Frequency Variation within Neonatal Crying: Does Ambient Language Matter?". *Speech, Language and Hearing* 19, 4 (2016): 211–17.