

Tomasello and the Scene of Language: Intersubjectivity and communicating sensation¹

The following, which would have been a presentation, summarizes several sections pertaining to affliction and alterity from a longer investigation into the implications for literary criticism of Michael Tomasello's work on language acquisition. My broader investigation concerns the relationship between what I call "scene-dependent" uses of language, language whose meaning or function is always inseparable from the spatio-temporal scene or event in which it occurs, and "scene-independent" uses of language, such as the language of private deliberation, silent reading, or the private reflection upon a proposition. I cannot here properly elaborate the subtleties of this distinction, but let it suffice to say that it is not absolute (and Austin's performative–constative distinction, which is similar but not equivalent to my own, is noteworthy). For the sake of investigation, I assume the veracity of my account of Tomasello's work, which, as I hope to make clear, presents an opportunity for disciplinary confluence between language acquisition psychology and the ethics of otherness. A further defence of my method, in which a hypothesis from empirical science is taken as an investigative starting principle, and an examination of its limitations are matters that I must also leave aside for now.

On the basis of his research into child and nonhuman primate language acquisition, experimental psychologist Michael Tomasello argues that the ability to engage in "joint attention scenes" is the unique determinant of the human linguistic capacity. He defines the scene of joint attention as a social interaction "in which [two or more individuals] are jointly attending to some third thing [the object], and to one another's attention to that third thing, for some reasonably extended length of time" (*Cultural Origins* 97). The subject of this scene attends to the object

¹ Please direct any questions, comments, or criticisms to lucassimpson@uvic.ca.

and the other's attention towards that object ("other" being an often excessively broad term in theoretical discourse, I use it in Tomasello's stricter sense of the non-subjective participant(s) in the scene of joint attention, though the relation to the Levinassian and related ethical senses should not here be forgotten entirely). Although Tomasello's earlier (1999) account attributes the human joint attention capacity to a species-unique ability to recognize the intentionality of conspecifics (*Cultural Origins* 19), a more recent (2008) formulation attributes the key difference to the ability to form joint goals and joint attentional scenes (*Origins of Human Communication* 176-7). Humans, Tomasello argues, have a unique advantage in their capacity to intentionally collaborate. Although nonhuman primates can understand the intentionality of conspecifics, "it is almost unimaginable that two chimpanzees might spontaneously do something as simple as carry something heavy together or make a tool together" (176). The general collaborative capacity, when applied to the special case of communication, translates to engagement in joint attention scenes.²

² Instead, nonhuman primate social interaction is based on prediction of behaviours based on past experience but not on an ability to intentionally collaborate with the conspecific. Although they may be able to perceive intention, Tomasello demonstrates that nonhuman primates cannot learn by imitation (*Origins of Human Communication* 25-26). In one of Tomasello's experiments (*Cultural Origins* 29-30), chimps and humans witness a demonstrator obtain (with varying degrees of efficiency) an out-of-reach object by means of a rake-like instrument. Whereas the humans in the rake experiment imitated the demonstrator (even when the method was inefficient), the chimps innovated their own means to achieve the same result (sometimes more efficiently than the demonstrated method). "For humans the goal or intention of the demonstrator is a central part of what they perceive, and indeed the goal is understood as something separate from the various behavioural means that may be used to accomplish the goal" (30). The chimps can recognize results (that the object can be reached with the rake) but not the means of obtaining those results. This path-independent style of learning Tomasello calls "emulation," as opposed to "imitation," its uniquely human counterpart. Similarly in the wild, chimps, after watching their conspecifics turn over a log to eat the insects underneath, will then do the same, but only because their watching allows them to understand that the insects are there; the means of accessing them (i.e., how to turn over a log) can only be learned independently (29).

Because all of its participants are attending to an object, the joint attention scene narrows the field of perceptual information to only the object(s) of significance. The key to this shared perspective is the capacity to perform what Tomasello calls “role reversals.” A child attends to an object while being aware of another’s attendance to that object. The child then reverses roles, that is, perceives the attendance to the object through the other’s own awareness of it. “Human infants understand joint activity from a ‘bird’s-eye view,’ with the joint goal and complementary roles all in a single representational format” (*Origins of Human Communication* 179). With this shared perspective children can use the linguistic symbol, which Tomasello defines as “a communicative device understood intersubjectively from both sides of the interaction” (*Cultural Origins* 106). The most elementary example of the linguistic sign is the pointing gesture: an adult points to an object in order to draw a child’s attention to it. That the roles can be reversed indicates that the sign is not limited to an individual participant’s perspective. Each of the gesture performance’s two participants “knows that they can both comprehend and produce the symbol” (106). This role fluidity distinguishes the linguistic sign from other non-linguistic signals. Infant chimpanzees initiate nursing by reaching for the nurser’s nipple, an action that the nurser begins to associate with nursing (note, the nurser only needs to *associate* the gesture with nursing, not grasp the infant’s intention). Through the recurrence of this action (a process Tomasello calls ontogenetic ritualization), the nurser becomes receptive at only a simple arm-touch gesture (31). The arm-touch is communicative but not symbolic, because there is a fixed relationship between the roles (nursing and nursed) and the gesture, unlike in the pointing

example. By contrast, the role reversals on which the linguistic sign depends makes the sign neutral or shared. Human language begins with “we” rather than “I.”³

The role fluidity in the scene of the linguistic symbol allows each participant to attend not only to the object but also *to the other’s attendance to the object*—hence the “joint attention scene.” The joint attention scene establishes a shared referent among its participants. This referent need not be governed by a pointing gesture. As with gestures, children can perceive when sounds are made with communicative intention. Often early language acquisition occurs with the two working together (e.g., “look: a bird [or any other object],” while pointing), but most language acquisition occurs “in the ongoing flow of social interaction in which both [the child] and the adult are trying to do things” (113). Children track the adults’ doings while following their intentionality and learning when certain utterances are appropriate in certain circumstances. The “object” of attendance should, therefore, be understood more broadly as the aspect(s) of shared experience to which the utterance governs the drawing of shared attention.⁴ A linguistic sign is perspectival. “What makes linguistic symbols truly unique from a cognitive point of view is the fact that each symbol embodies a particular perspective on some entity or event: this object is simultaneously a rose, a flower, and a gift” (107). This perspective emerges from the way in which the adult (whose intentionality the child language-learner perceives) engages with the object.

When a child learns a linguistic sign, the sign’s significance depends on the scene of utterance because that scene contains the other attending to (some aspect of) the object.

³ See Charles Taylor’s discussion of Tomasello, which emphasizes the scene of attention’s communion-establishing function (52-64).

⁴ Although the most intuitive example of a scene of joint attention for non-blind and non-deaf people is the pointing and name-calling performance, Tomasello’s theory of language acquisition is not limited to specific senses.

Therefore, every person's initial experience with language is ineluctably intersubjective and therefore scene-dependent. Whatever language's full shape, Tomasello's evidence suggests that language is fundamentally scene-dependent. Whatever precedes shared language (e.g., non-communicative conceptual thinking) in the process of mental development, then, cannot be called language⁵ use. A child initially uses language only when attending to both the object and the other's attention towards that object while the linguistic sign (e.g. word, gesture) mediates this attentional status. The object of attendance, as I have said, is an aspect of shared experience. How, then, can we communicate aspects of experience that *cannot* be shared, such as emotions, sensations, pains? The shift from communicating aspects of shared experience to aspects of unshared experience would coincide with the shift from scene-dependent to scene-independent language.

If this is a starting assumption, how are we to understand the communication of pain in (or by) language? How does one learn to conceptualize and communicate pain? I can't point my pain out to others so that they will know what I mean. I can't make it an object of joint attention. There is no way for me to know the pain of someone else except on the basis of my own (and that I can have a determinate idea even of my own pain is doubtful; see Wittgenstein 258). I could elaborate my pain in further detail (e.g., "pain, in my right big toe, throbbing") but these, too, would ultimately rely on something that cannot be an object of shared experience. The conclusion, then, is that we cannot convey the thought or sensation that is pain.

⁵ We could, of course, define language differently so as to include non-human communication systems. Following Tomasello's usage, I mean by "language" the communication system that humans seem to uniquely use. Some might prefer to call this "human language."

The word “convey” can mislead here if it is assumed to be equivalent to “communicate.” Such an assumption would reduce acceptable forms of language communication to propositions. Moving beyond language as the conveyance of thoughts resolves this apparent problem (Wittgenstein 305). When we speak of knowing one’s pain, we do not “know” it in the same way that we can “know” the information that “I am in the kitchen” or “the cat is on the mat,” etc. A broader sense of the shape and scope of language therefore requires a broader sense of what knowledge is: what does it mean to “know” someone is in pain? Stanley Cavell makes this point in response to Wittgenstein’s investigation into the knowability and communicability of pain. Of the many senses of the word “know,” Cavell emphasizes the sense of *acknowledgement*: as in, “I know I am late” or “I know I am being childish” (255). Knowing in this sense is not the knowing of information. It is a knowing that cannot be divorced from its shared circumstances and cannot be held privately (i.e., it is scene-dependent knowing). The headache avowal probably intends not to communicate information but rather to request sympathy or to beg someone to stop whistling or to excuse the speaker from a social gathering, etc. And we, as auditors, acknowledge our friend’s suffering by doing any number of these. “Your suffering makes a *claim* upon me. It is not enough that I *know* (am certain) that you suffer,” because such certainty is impossible. Rather, “I must do or reveal something (whatever can be done)” (263).

To return to Tomasello’s formulation, when you acknowledge another’s pain, you make the pain the attended object of the scene of joint attention. Rather than transmitting knowledge of the pain (in the limited, propositional understanding of language), the avowal *establishes* it for joint attention. By following the intention of another and learning to associate a sign (“I am in pain”) with some aspect of the reality (the experience of the speaker/other) towards which the

intention and attention is directed, the language-learner can come to grasp a new perspective on reality. The child can then rehearse this perspective privately, but whatever concept is associated with the sign will always be constituted for intersubjective communication (*Cultural Origins* 149). The “pain” of “I am in pain” is never my pain or your pain but our “pain,” in the sense that it (the concept corresponding to the word in quotations) is constituted in the shared scene of joint attention. The role reversal on which the sign depends guarantees the neutrality of the sign-users with respect to the concept. The linguistic sign grants no special status (in its functioning alone) to the signing participants. “Pain” designates a shared concept by which we navigate our common experience, and this concept defines a set of rules of appropriate response or “acknowledgement.” The normativeness of ordinary language subordinates sincerity to these force-providing rules. As speaker, I can elaborate on my private experience in terms of the shared concept of pain. Because of my articulation, my auditor can better share the perspective on this aspect of experience and respond accordingly. The expression “I am in pain” differs from crying only in that the crying is not delimited as a (and here consideration of the general meaning of “definition” illuminates its specific meaning with respect to words) *definite* concept, established in a scene of joint attention, and therefore cannot be understood by the auditor from the common perspective that words offer. The words give definition to the aspect of reality (pain) that the crying only vaguely communicates.

To learn a language, then, is to learn the language-culture’s given perspective on experience. Language provides a set of concepts and perspectives (in addition to already existing private ones) that are “constructed” (Tomasello’s word) for culturally specific behavioural practices (*Cultural Origins* 166). That a given language generates or structures a culturally

relative perspective on the world is nothing new. What Tomasello's work does offer is a grounding of this structure in its experienced acquisition. Language is not fundamentally about making our private thoughts, appetites, or needs known to others but instead is, despite its abuses, about establishing both communion with our fellows and cooperative engagement with our shared experience.

Building on Cavell's notion of "acknowledgement," Toril Moi suggests that reading itself should be treated as a "practice of acknowledgement" (Moi 216). When you hear your friend's headache avowal, you acknowledge it by navigating your relationship relative to the avowal. The question is not about whether you should believe the avowal but how you should respond. Similarly, the work of literature cannot impart anything to us. Instead, we read it by situating ourselves and our responses with respect to the work.⁶ "Acknowledgement isn't just a matter of accounting for the work's concepts. It also requires us to understand our own position in relation to the work's concerns" (217). As auditors, we often experience the "claim" that the headache avowal has on us as compassion or sympathy. And we may feel similarly when we read, say, elegies, despite their fictionality and despite the fact that there is nothing we can do to improve the poetic speaker's suffering. This kind of engagement with the text, this acknowledgement of our interrelation with it, is, for Moi, a process of self-education, in which you learn to give authority (in your assessment or reaction to the work) to your own experience (219). In explaining the aesthetic experience by means of an ordinary language process (*viz.*,

⁶ Of course, this should not be taken to suggest that the way in which we situate ourselves with respect to the experience is the same in literature as it is for pain avowals or any other ordinary experience. We respond differently when we witness the expression of pain on stage than we would in our everyday lives. "I don't mean to say that [aesthetic experience] doesn't require special insights, knowledge, skills, and judgment. But so do other experiences as well: truly to appreciate great cooking, a particularly excellent fishing spot, or great growing soils also requires specific knowledge, judgement, and skills." (Moi 219).

acknowledgement), Moi grants no special status to the aesthetic experience: “Aesthetic experience is ordinary: to find out what our aesthetic experience means entails the same difficulties and joys as the investigation of other experiences” (218). In both the private experience of reading and the shared experience of acknowledging someone’s pain, we must configure our private experience with respect to the shared information we receive.

Although I do not disagree with Moi’s account, I want to conclude by suggesting a further exploration: to explore what might be a special capacity of literature—to expand our range of understanding beyond the ordinary function of language. In ordinary language, sincerity is relegated to convention (e.g., types of acknowledgement). The imagination plays little to no role in the behaviours that determine language’s concepts. Although the language-user internalizes (I could even say imagines) the sign’s concept, the concept is still delimited only in terms of external behavioural convention. For Moi and Cavell, what matters is not that we know the other’s pain, have experienced it ourselves, or can imagine it on the basis of our own, but that we respond (by words and behaviours) to the “claim” the text makes on us. But do we not often turn to literature, in part, because it articulates and communicates private experience, private pains, that would otherwise go unsaid, *unknown*? I suggest that there is a sense in which the aesthetic experience of language constitutes a sphere of meaning where the boundaries of convention and scene-dependence dissolve to give way to autonomous meaning—meaning that transcends the limits of ordinary language and can then include what we had to exclude (e.g., the private sensations). By means of metaphor, for example, abstract concepts are constituted, thus expanding the scope of the articulable, on the basis of concepts that can be objects of joint attention. If (as my hypothesis requires) we accept a fundamentally scene-dependent and

intersubjective starting point, this expansion will always derive from such objects of joint attention, and therefore always be fundamentally ethical in its constitution by and with the other.

Works Cited

Austin, J.L. *How to Do Things With Words*. Oxford UP, 1962.

Cavell, Stanley. "Knowing and Acknowledging." *Must We Mean What We Say?* Cambridge
1969. Reprint, 2002. 238-266.

Moi, Toril. *The Revolution of the Ordinary: Literary Studies after Wittgenstein, Austin, and
Cavell*. Chicago UP, 2017.

Taylor, Charles. *The Language Animal*. Harvard UP, 2016.

Tomasello, Michael. *The Cultural Origins of Human Cognition*. Harvard UP, 1999.

—. *The Origins of Human Communication*. MIT, 2008.

Wittgenstein, Ludwig. *Philosophical Investigations*. Trans. G.E.M. Anscombe, P.M.S. Hacker,
and Joachim Schulte. Revised fourth edition. Wiley-Blackwell, 2009.