

# Attention (and Joint Attention)

JOHN B. HAVILAND

University of California-San Diego, USA

A prototypical example of “joint attention” involves visual attention in what are referred to as “social communicative routines” (Bruner 1985) – repeated and sometimes prolonged interactions between child and caregiver that are said to scaffold a child’s early language, specifically learning referents for words, along, perhaps, with conversational skills (Tomasello 1988). Joint attention on this account resembles an augmented version of Augustine’s picture – famously criticized by Wittgenstein (see Wittgenstein, Ludwig) – of a hypothetical language-learning child who notices adults’ attention to an object and “grasp[s] that the thing was called by the sound they uttered when they meant to point it out” (Wittgenstein 1958). The notion of “joint attention” is meant to add to this picture an interactive coordination – a child’s attention is *directed* to an object by an adult’s gaze or pointing gesture, and the child attends to the object by recognizing the adult’s intention so to direct her attention. A more reciprocal version is also possible: the child alternates gaze between adult and object, seemingly manipulating – both calling for and checking – the other’s attention.

Scholars apparently take “joint attention” among mature humans as a given but are puzzled about how young infants acquire the requisite cognitive, conceptual, and psychosocial abilities. Special attention has focused on the child’s developing theory of mind (see Language and Mind) as a distinctive and defining part of human consciousness, by contrast, for example, with that of apes or even, say, ravens: other species which display sensitivity to the eye gaze of other beings without reaching what researchers qualify as true “joint attention.” Autistic individuals are also thought to exhibit differences from others with respect to their capacity for joint attention, and developmental studies on the topic abound. There is also a large companion philosophical literature on how to decompose and interpret the essential conceptual ingredients for joint attention, in ways relatively “rich” or “lean” with respect to the implied theories of mind.

Anyone who has spent time with young children and their caregivers will be familiar with canonical displays of joint attention, although not all such routines focus on learning new lexical referents. Alice, at age nine months, speaks no recognizable words, but she has a few emblematic gestures, including an admonitive negative finger wave. She crawls toward a garbage bin, and as she reaches for it, her mother says, “Don’t touch that thing, Alice!” Alice halts and turns to look at her mother, simultaneously performing the negative finger wave, to which her mother replies with a “tsk” and a “no.” Alice again waves her finger and gives a further negative head shake. The infant then stares and waves her finger at the forbidden garbage can, as her mother intones “no, no, no, mustn’t touch that thing.” Here, if there is an “object” being jointly attended to at all, it is the garbage bin; and what is learned is not a referent but a prohibition – spoken and gestured.

*The International Encyclopedia of Linguistic Anthropology*. Edited by James Stanlaw.

© 2021 John Wiley & Sons, Inc. Published 2021 by John Wiley & Sons, Inc.

DOI: 10.1002/9781118786093.iela0018

Whether or not such a familiar example fits the definition of “joint attention” in the literature – an adult and a child conjointly attending to a third “object” with some sort of mutual consciousness of that conjointness – is unclear. Nonetheless, ethnographic interest in joint attention surely need not be bound to such a definition, which is not without its own problems. The etymology of “attention” – involving a metaphor of “stretching toward” (linked conceptually to thin cords, sinews, and looms) – suggests a process, involving duration and effort, rather than a “thing.” The idea of “jointness,” far from implying an infinite regress (the child’s attention depends on her recognizing the adult’s intention to get her to attend, the adult’s recognition of her recognition, and so on) suggests that the yoking together of multiple attendings accomplishes something that would be otherwise impossible. Conjointly attending to something as part of an activity qualitatively different from any single person’s so attending is classic collaborative, or as Charles Goodwin (2018) would have it, “co-operative” action. Restricting the use of “joint attention” – undeniably central to human interaction, whether linguistically achieved or not – to the standard definition above has several disadvantages, to be considered in turn.

First, the emphasis on visual modality assumed in most psychological and experimental work is unnecessarily restrictive (Tomasello and Farrar 1986). Gaze may be reasonably easy for an external party to observe in humans, but there are important alternatives.

Consider a kiss, and not just a simple peck on the cheek (although that may also require considerable joint attention: you and your partner must decide how many pecks to give, and also which cheek to offer first, depending on what country you’re in), but an especially *good* or memorable kiss. You may, indeed, have your eyes closed, centering your attention on the feeling of your lips against the other’s lips, the taste of her tongue, the smell and warmth of her neck and face. All human senses may be involved in co-constructing action. Joint attention can rely on all of them (or, indeed, none, when the object of attention is imagined).

Young Mal, at 16 months, points in the direction of a blast from a distant truck horn, and says “go” to her mother, a performance interpreted as declaring that the truck is about to leave as signaled by its horn. Victor, a hearing but signing child at 23 months, in conversation with his deaf mother, visually draws his mother’s attention to an audible “entity” she cannot hear. Because the canonical joint attention described above links learning spoken words to visual attention to objects, interaction between a deaf caregiver and a hearing signing child is an interesting reversal. Victor had been listening to a brass band playing on the street outside his yard, but the band has now moved off. He wants his mother (who cannot hear the band but is aware that others can) to accompany him to watch it more. He signs to his mother: HEY! (requesting her visual attention); LET’S GO (with a hand flip to show which direction he means); HEAR (or SOUND), covering his left ear to denote the sound of the band; THAT WAY (pointing the way the band, by its sound, is headed). His utterance incorporates a (visible) linguistic reference to a sound, which he invites his mother jointly to attend to.

Familiar, too, are episodes of joint attention involving other modalities: commenting to an interlocutor on the smell of baking bread, or sharing the taste of a spoonful of broth. There is also joint haptic or tactile attention: piano movers balancing a heavy

load; inviting a guest to confirm that the pasta is *al dente*. These are aspects of joint attention that Tomasello anticipated (1995) in characterizing it as “two persons attending to the same aspect of their common environment”; but the sensory scope of empirical investigation has been remarkably truncated.

Second, the notion of an “external object” as target of joint attention is unnecessarily restrictive (and, indeed, ontologically unclear). There is no “external” kiss without the kissing. It can only be experienced (in the relevant way) by the kissers themselves. Nor will their experiences of the kiss be symmetrical: a beardless person may attend to a hirsute partner’s kiss differently than the other way around, even though it remains “the same” kiss. Victor’s experience of the band’s sound is different from his mother’s (since she cannot hear it); she nonetheless is meant to enter a state of joint attention to that sound (however she conceives of it) by virtue of her signed interaction with her son.

Related to the indeterminacy of the “object” of joint attention are several final concerns: the temporal dimensions of joint attention, and the distributed co-construction of the “object” of attention, temporally, interpersonally, and ontologically.

Two final illustrative vignettes: in a string quartet master class, a professional viola player suggests bowing and phrasing to a student quartet. He concentrates on a single measure of the written score of a Mozart quartet, via a variety of instantiations, in different modalities. (Each of the members of the quartet, of course, has a different version of that measure to play.) He mimes the bowing with his arm while singing the passage, and he also plays it multiple times on his viola, experimenting with different phrasings and dynamics. The students are expected to engage in experimental haptic and aural instantiations by playing the passage several times themselves.

In preparation for a Zinacantec wedding, a small army of women forms an assembly line to produce a massive quantity of tortillas. Women leave and join the line over the course of several hours. The process has been organized into component steps assigned to individuals and small groups: kneading the dough into a large mass; forming the *masa* into small individual balls; pressing each ball into shape and placing it on a large griddle; monitoring each tortilla as it cooks, turning it, and encouraging it to puff up; removing the cooked tortillas from the griddle; and finally moving the carefully stacked tortillas for serving. There is neither boss nor conductor, and women move fluidly to different spots on the line (which is distributed around the space, some women sitting at tables, others standing by the griddle, and so on). Together all jointly attend to the overall process and to their own specific parts of it. They coordinate their different actions with one another, noting when a space appears on the griddle where a new tortilla can be placed, when it is time to turn each tortilla, when one is ready to be taken off the griddle, and so on.

In these examples, “joint attention” seems to have a moving target. Indeed, part of the point of the string quartet master class is to refine the conjoint understanding of what the object of attention is: the notes on the page are variously performed and interpreted, over the sustained interaction. The potential fluidity of the abstract object of attention – a musical phrase, part of the string quartet as a whole – is as important as any of its given realizations. And one’s appreciation of what it is one is attending to can

change as interaction proceeds: one comes to sense or hear something new, to discover unexpected perspectives and previously unencountered aspects of the entity.

Goodwin considers language as just one element within what he calls “multiparty embodied participation frameworks” which as he points out resemble Tomasello’s “joint attentional frames.” He notes, however, that the framework goes beyond attention as a mental process. The “objects of attention” in both the master class and the tortilla-making production line just described are organized across time and space, and the possibility of focused, joint attention depends to a large extent on the physical arrangements and interrelationships of objects and people, and on coordinated attention between individuals, who contribute differentially to the activity while they jointly attend both to the circumstances and to each other. Analysis must move beyond the limits of an asymmetric communicative dyad – an imagined semi-competent child and an adult – attending jointly (but unequally) to a supposedly fixed “object,” to encompass instead multiparty interactions of distributed attention, knowledge, skill, and affect, focused on a fluid, jointly construed, mutually and collaboratively constituted dynamic “object” of attention.

SEE ALSO: Child Language; Common Ground; Communication, Animal; Gaze; Husserl, Edmund; Interaction, Face-to-face; Intersubjectivity; Kinesics and Gesture; Language Acquisition; Language Emergence; Language and Mind; Language Socialization; Modality, Multimodality; Quine, W.V.O.; Sign Languages; Wittgenstein, Ludwig

## REFERENCES AND FURTHER READING

---

- Bruner, Jerome. 1985. *Child's Talk: Learning to Use Language*. New York: Norton.
- Carpenter, Malinda, Katherine Nagell, Michael Tomasello, George Butterworth, and Chris Moore. 1998. “Social Cognition, Joint Attention, and Communicative Competence from 9 to 15 Months of Age.” *Monographs of the Society for Research in Child Development* 63 (4): 1–174. <https://doi.org/10.2307/1166214>.
- Duranti, Alessandro. 2010. “Husserl, Intersubjectivity and Anthropology.” *Anthropological Theory* 10 (1–2): 16–35. <https://doi.org/10.1177/1463499610370517>.
- Eilan, Naomi, Christoph Hoerl, Teresa McCormack, and Johannes Roessler, eds. 2005. *Joint Attention: Communication with Other Minds: Issues in Philosophy and Psychology*. New York: Oxford University Press.
- Goodwin, Charles. 2018. *Co-operative Action*. New York: Cambridge University Press. <https://doi.org/10.1017/9781139016735>.
- Tomasello, Michael. 1988. “The Role of Joint Attentional Processes in Early Language Development.” *Language Sciences* 10 (1): 69–88. [https://doi.org/10.1016/0388-0001\(88\)90006-X](https://doi.org/10.1016/0388-0001(88)90006-X).
- Tomasello, Michael. 1995. “Joint Attention as Social Cognition.” In *Joint Attention: Its Origins and Role in Development*, edited by Chris Moore and Philip J. Dunham, 103–130. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Tomasello, Michael, and Michael Jeffrey Farrar. 1986. “Joint Attention and Early Language.” *Child Development* 57 (6): 1454–1463. <https://doi.org/10.2307/1130423>.
- Wittgenstein, Ludwig. 1958. *Philosophical Investigations*. Oxford: Blackwell.