



Social Interaction. Video-Based Studies of Human Sociality.
2021 Vol. 4, Issue 3
ISBN: 2446-3620
DOI: 10.7146/si.v4i3.128152

Social Interaction

Video-Based Studies of Human Sociality

“More than meets the eye”: Accessing senses in social interaction

Emily Hofstetter & Leelo Keevallik

Linköping University

1. Introduction

The papers in this special issue witness the leap in recent years regarding the interactional analysis of the senses. While foundational work in the organization of sociality was carried out within the realm of auditory communication (Sacks et al., 1974), we have now firmly arrived at a stage where the visual details of behavior—the traces of embodiment in recording—are incorporated into analysis, across fields ranging from workplace studies to linguistics (Neville, 2015). Multisensoriality studies, as initially outlined by Mondada (2019), focus on the moments where sensory events are made auditorially and visually accountable and thus capturable with video (Goodwin & Cekaite, 2018; Mondada et al., 2021/this issue). We have started to literally see how sensorial behavior is a regular aspect of sense-making between co-present participants. This inevitably leads us to further research tasks, such as considering what senses to study, how to access them, and what the limitations are.

2. What senses do we analyze?

From the earliest video analyses (C. Goodwin, 1979, 1980; C. Heath, 1982) to the present day, we have been sharpening our understanding of the role of visible and audible behavior in interaction. Most accounts in the current volume provide ample illustrations of how participants rely on both hearing and seeing senses in carrying out action, and argue that close attention to these is therefore essential for a well-grounded analysis. Smith takes on particular challenges of vision at extreme distances, demonstrating how participants systematically handle each other's line-of-sight, while, at a close distance, Merlino discusses how visual highlighting of the therapist's mouth can provide access to speech articulators (e.g., the tongue). Several studies show the affordances of technology in enhancing human seeing, such as imaging techniques in surgery that provide access to the interior of the body (Kuroshima & Ivarsson, 2021/*this issue*) or smartphones that enable visual and audible access to geographically distant families (Gan, 2021/*this issue*).

The senses beyond hearing and seeing, however, remained unanalyzed for decades because they were considered inaccessible for co-participants as well as the analyst, and because naturally video afforded easy access to sight and sound (Streeck, 2003). Within ethnomethodology and conversation analysis (EMCA), this boundary has only recently been surpassed, as studies begin to investigate the socially organized nature of taste, smell and touch (e.g., Fele & Liberman, 2021; Goodwin & Cekaite, 2018; Mondada, 2018; Nishizaka, 2017). As it happens, these innovative accounts have so far mostly centered on specific sensorially focused activities, such as learning to become a connoisseur or professional. The paper by Mondada et al. (2021/*this issue*) addresses such food-and-drink-focused activities and illustrates how the haptic senses are used for disclosing aspects of sausage, cheese, and drying banana peels for co-present others. In this special issue, access to experience is mostly limited to a single participant, such as when only one person in the medical team can sense the reach of a guidewire inside the patient's blood vessels (Kuroshima and Ivarsson, 2021/*this issue*) or how the weight, smell, and taste of noodles is only available to one side of a video chat (Gan, 2021/*this issue*). Likewise, lips touching a screen in remote 'kissing' is an asymmetrical event. At the same time, there is definitely a difference between touching a solid flat surface and parts of the soft human body, but unless made public, these sensations are difficult to pinpoint with EMCA methods. As the authors point out, many sensory experiences, though constantly unfolding, are not made available to others. Some may be systematically avoided as subjects due to prevailing social norms. For instance, it is taboo to mention disliking the food offered by a host or someone next to you smelling bad, even though these very matters can certainly affect ongoing interaction, its duration and outcomes.

Some studies in this issue explore senses beyond hearing, vision, touch, taste, and smell. LaBonte et al. (2021/*this issue*) discuss aspects of strain as revealed

through hearable and visible signs of failure (the rattling of the equipment and the trainee's slumping body) but they also consider how alternative kinds of access to that experience might reveal interactionally relevant aspects of bodies in action. Katila and Turja (2021/*this issue*) address kinaesthesia, specifically discomfort and heaviness as experienced in the torso, which can also be made visible through specific ways of moving. Departing from theoretical standpoint of intercorporeality, they furthermore argue that some of our bodily experiences are "readable" without the participants themselves explicitly indexing a particular sense. Aspects such as being uncomfortably cold, scared, or sleepy can be "given off", or even simply presumable, regardless of attempts to disclose the evidence. The papers in this issue recurringly mention tensions in analyzing solely video-accessible events and analyzing primarily asymmetric sensory availability, which brings us to two substantial issues: a) it is not only the five "traditional" senses that we want to work on, and b) bodily information is not always made accountable in the same way as a spoken word or visible movement.

While the five sensations (vision, hearing, touch, taste, and smell) have by now emerged as topics of interaction analysis, these are far from the only ones to occur in interactional environments. The dominant focus on those senses has long been critiqued in anthropology, phenomenology, philosophy, and biology as not only inaccurate and incomplete, but moreover centered in specific (Western European) cultural ideologies (Ingold, 2000; Pink, 2009). Senses such as balance, proprioception, kinaesthesia, pain, heat, energy, etc. have mostly yet to be addressed from an interactional perspective. Interestingly, we already have evidence that pain, for instance, is socially organized at medical consultations (C. Heath, 1989; La & Weatherall, 2020), and it has now also been shown to interfere with progressivity at both activity and turn level (Weatherall et al., 2021). Likewise, strain can be systematically made public through, e.g., the practice of temporarily suspended syntax, which allows for a strain display to emerge as an all-encompassing preoccupation of the body (Hofstetter et al., 2021). Strain is thus yet another sensorial aspect of the body that is intersubjectively managed.

As a further example, dance teachers may vocalize strain at precise moments when effort is due in students' ongoing performance. Excerpt 1 comes from a class where a Charleston combination is taught to student couples. The lead dancers in the couple need to bring their partners from side to front and back again, which requires a well-timed creation of physical tension to reverse the follow dancers' momentum. During the excerpt, one of the teachers (behind her partner in figure A) provides the rhythm with vocalizations, half-singing (Keevallik, in press).

Excerpt 1.

beats 1 2 3 4 5 6 7 8
01 Tea: tara ta rah ta ta ra taaa

beats 1# 2 3 4 5# 6 7 8
02 **KRRhahhhh** chaga **ΣQUOOΣ** qaaa
 #fig A #fig B



Fig A. Reversing backward motion to a move forward



Fig B. Students performing the strenuous move, teacher "joining in"

In line 01, the students are supposed to dance a basic step, side-by-side, in a rhythmical but relaxed manner, which is also reflected in the teacher's subdued movements and simple singing. In line 02, however, the first syllable is qualitatively different: It features a loud elongated trill across the two first beats of the step pattern, and a heavy outbreath at the end. This is where the leads are supposed to redirect the followers' energy (as shown in Figure A). During the next two beats, the joint forward movement is again relaxed, which is also reflected in the open, less stressed syllables "chaga". Then the most strenuous move is due, highlighted by the teacher's markedly louder voice and an extended high diphthong, *uo* (contrasting with back vowels accompanying the rest of the pattern). The vocalization is furthermore uttered with the air barely seeping through the glottal closure (marked with Σ), and with a clear glottal stop onset (marked with "Q"). In addition, or perhaps in order to produce the specific strain sound, the teacher herself also embodies strain in parallel to the lead dancers (figure B). We can thus both see and hear how strain is collaboratively achieved and therefore becomes accessible through video. By including these further sensorial experiences, such as the ostensibly invisible ones of proprioception, balance, pain, and strain, we can expand the range of phenomena we study, at least in perspicuous settings.

3. How do we analytically access senses?

Each of the papers grapples with how to best access the ongoing sensory work with cameras. Some focus on viewing the participants' joint perspectives (Smith), shared referents like screens (Gan), objects (Kuroshima & Ivarsson, Mondada et al., LaBonte et al.), and mirroring of bodies (Katila & Turja, Merlino). The name of the game is to capture the visual-spatial positions where sensing takes place—and, as LaBonte et al. point out, to anticipate where it will be relevant (and safe to film). While the high stakes settings of, e.g., an operation room, place clear restrictions on camera arrangements (Kuroshima & Ivarsson), in the case of establishing what the patient can access in the speech therapist's mouth (Merlino), a head-mounted camera may not have been too intrusive and might have provided a better view. Body-mounted cameras often made available useful supplements to third-person views (however, see discussion in Smith), though not all footage from the myriad of cameras across the studies resulted in innovative insights. As cameras have become smaller and cheaper, capturing events from various vantage points has largely become default, but useful placement of these multiple angles still relies on forethought and ethnographic insight.

The question of how to access participants' sensory experiences is trickier, and two approaches emerge among the papers. One can access bodily-sensory events with the aid of ethnography, namely through researcher participation (as featured in LaBonte et al.), or one can seek out moments where bodily-sensory

events become accountable (mostly clearly in Merlino, Katila & Turja, Mondada et al.). The former is a process well-documented by ethnomethodology's unique adequacy requirement (see also Garfinkel, 2002), namely, the expectation that a researcher would be able to act as a reasonably competent participant in the activity studied. This provides access not only to information on *how* to sense in a locally appropriate way, but also to lived experience from one's own body (see Streeck, 2013). The specialized activities in this issue make ethnographic understanding particularly relevant, as they involve trained, or training, sensation. It is not just "touching", but particular "ways of the hand" (Sudnow, 1993), not just "seeing" but "professional vision" (C. Goodwin, 1994). Interestingly, the study most poised to take advantage of researcher participation does not report using such techniques. Katila & Turja's study of nurses trying out an exoskeleton does not report the researchers being nurses, trying out the lifting necessary on the ward, or trying on the exoskeletons. Besides sequential evidence, the analysis draws on general intercorporeal information available from simply having a body. Neither do LaBonte et al. discuss how intercorporeal access might be incorporated into their analysis. These two studies, and their approaches to accessing bodily-sensory events, remind us that ethnomethodology and intercorporeality have significant cross-dialogue ahead of them, and that a combination of the two may be fruitful (see Jenkins, 2017; and others in Meyer & v. Wedelstaedt, 2017).

In order to focus on what is verbally-visually available, i.e., available to a camera and thus to analysts (especially analysts without extensive background in an activity), studies of multisensoriality have had to restrict their analysis to moments where sensing is made explicitly relevant and/or accountable. Although a few studies note that sensing is constant, none of them analyzes taken-for-granted moments. By only looking at "available" cases, we necessarily overlook the omnipresence of sense and dualistically imply that there are unavailable sensations. Regardless of one's own ontological take on this matter, this makes for interesting research questions: Do participants assume or expect the occurrence of sensing all the time? How (if at all) does the omnipresence of sensation figure in participant sense-making and interactional order? For instance, significant sweating is notable if a rock climber was expected to achieve something easily, but is not notable (highly unlikely to be made accountable) if the route was expected to be difficult. Furthermore, the hotter or more humid the day, the less notable sweating, slippery hands, and general exhaustion are, even though these elements are consequential for the progressivity (and safety) of the activity.

The readers of our studies also need to access the analyzed phenomena, which rely on how the sensory materials are represented. How and where to incorporate still images from video is critical. For example, zoomed-in images next to their contextual frames helpfully focus the reader to the phenomena. All the papers chose images that capture the "peak" (or stroke) of a motion, and/or the joint attentional focus of a constellation of bodies—for instance, the moment where a

sausage being selected is touched (Mondada et al.), the moment lips kiss a mobile phone camera (Gan), or the moment of pointing (Smith). However, these images do not work for demonstrating movement, especially subtle movements such as squeezing a cheese or sausage. Two papers attempt alternative visualizations of the unfolding of motions over time. LaBonte et al., for example, show an instructor “clapping” upon a tense line to demonstrate that it holds all the weight, and aligns a series of stills-with-background to the transcript, with timestamps. Katila and Turja also use a series of stills to show a nurse’s gait; however, the features are more obvious in some stills than others. For instance, the curved arrows suggested a wide swinging shape to the legs, but these arrows were used both in diagrams that described a normal and a swinging gait, making it harder to see the events as contrastive. Meanwhile, the timing differences noted between gaits provided ideal evidence for the hindrance in walking under discussion.

Such difficulties in representation are to be expected when communicating living events on static surfaces; depth and spatiality are very difficult to ascertain in two dimensions. Highly specialized movements, such as in martial arts or dancing, are virtually impossible to visualize for a wider audience. Diagrams that abstractly present spatial layout or complex features could help rectify this challenge, but few appear in the issue (see Gan for the layout of the video chat screens). Inviting readers (and live audiences) to reenact certain motions can also be used to induce embodied understanding of the phenomena members are managing. The authors’ experiments with this kind of embodied co-analysis include having audience members test the sensation of counterbalancing bodies, and having them pull on rope equipment to feel how tension is a resource for climbing. These exercises provided access to sensation and permitted independent analysis of the data, as is ideal in EMCA.

4. Limitations and opportunities

The captivating collection inevitably raises questions of methodological limitations—what other forms of evidence can be used for EMCA that still fulfil its requirements for relevance for participants (see Schegloff, 1987, 2009)? One option is to rely on an ethnographer’s members’ knowledge of how, for instance, the tactile sensation of weight on a rope, with visually available rope tension, *should* feel or *does* feel. The other is to involve sensor technology. Multisensoriality as a field of investigation invites us to be imaginative in looking for places where we can *discover* traces of member orientations. Technologies beyond video cameras may reveal aspects that are not made verbally or visually accountable. For instance, thermal cameras might be an option when looking at physical exertion and bodily stress. The key, as with other technologies, is to maintain a focus on participant orientation. Eye-tracking glasses should not be viewed as a means to access every detail of vision, but rather a way to capture

with greater ease how gaze direction is accountable (Stukenbrock & Dao, 2019). Instrumental measurements represented with spectrograms and pitch traces, as another example, support the analysis of speech and make acoustic evidence easily representable in texts. While spectrograms and eye-tracking have become relatively commonplace, laryngoscopes, motion capture (Stevanovic et al., 2017), breathing sensors (Aare et al., 2019; Torreira et al., 2015), and other techniques are rare or, to date, unused. We should continue experimenting with how we observe and present interactional evidence if we are to enact what Katila and Turja so neatly call for (echoing Streeck, 2013): analyzing the body as a sentient being, not only as a being that produces sensorial resources.

5. Conclusion

The advancements in multisensoriality are exciting and the studies here exemplify the benefits of observable relevance in video analysis. The realm of the sensorial seems an especially perspicuous area in which to be imaginative regarding both the expansion of the range of phenomena analyzed as well as incorporating evidence beyond video into analysis.

References

- Aare, K., Włodarczak, M., & Heldner, M. (2019). Breath holds in spontaneous speech. *Journal of Estonian and Finno-Ugric Linguistics*, 10(1), 13–34. <https://doi.org/10.12697/jeful.2019.10.1.01>
- Fele, G., & Liberman, K. (2021). Some Discovered Practices of Lay Coffee Drinkers. *Symbolic Interaction*, 44(1), 40–62. <https://doi.org/10.1002/symb.486>
- Gan, Y. (2021). Capturing love at a distance: Multisensoriality in intimate video calls between migrant parents and their left-behind children. *Social Interaction. Video-Based Studies of Human Sociality*, 4(3).
- Garfinkel, H. (2002). *Ethnomethodology's program: Working out Durkeim's aphorism* (A. W. Rawls, Ed.). Rowman & Littlefield Publishers.
- Goodwin, C. (1979). The interactive construction of a sentence in natural conversation. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology* (pp. 97–121). Irvington Publishers.
- Goodwin, C. (1980). Restarts, Pauses, and the Achievement of a State of Mutual Gaze at Turn-Beginning. *Sociological Inquiry*, 50(3–4), 272–302. <https://doi.org/10.1111/j.1475-682X.1980.tb00023.x>

- Goodwin, C. (1994). Professional Vision. *American Anthropologist*, 96(3), 606–633.
- Goodwin, M. H., & Cekaite, A. (2018). *Embodied Family Choreography: Practices of Control, Care, and Mundane Creativity*. Routledge.
- Heath, C. (1989). Pain Talk: The Expression of Suffering in the Medical Consultation. *Social Psychology Quarterly*, 52(2), 113.
<https://doi.org/10.2307/2786911>
- Heath, C. C. (1982). *The display of reciprocity: An instance of a sequential relationship in speech and body movement*. 42(2–4), 147–168.
<https://doi.org/10.1515/semi.1982.42.2-4.147>
- Hofstetter, E., Keevallik, L., & Löfgren, A. (2021). Suspending syntax: Bodily strain and progressivity in talk. *Frontiers in Communication*.
- Ingold, T. (2000). *The Perception of the Environment: Essays on Livelihood, Dwelling and Skill*. Psychology Press.
- Jenkins, K. N. (2017). Rock climbers' communicative and sensory practices: Routine intercorporeality between climbers, rock, and auxiliary technologies. In C. Meyer & U. v. Wedelstaedt (Eds.), *Moving Bodies in Interaction – Interacting Bodies in Motion: Intercorporeality, interkinesthesia, and enaction in sports* (Vol. 8, pp. 149–172). John Benjamins. <https://doi.org/10.1075/ais.8.06jen>
- Katila, J., & Turja, T. (2021). Capturing the nurse's kinesthetic experience of wearing an exoskeleton: The benefits of using intercorporeal perspective to video-analysis. *Social Interaction. Video-Based Studies of Human Sociality*, 4(3).
- Keevallik, L. (in press). Vocalizations in dance classes teach body knowledge. *Linguistic Vanguard*.
- Kuroshima, S., & Ivarsson, J. (2021). Toward a praxeological account of performing surgery: Overcoming methodological and technical constraints. *Social Interaction. Video-Based Studies of Human Sociality*, 4(3).
- La, J., & Weatherall, A. (2020). Pain displays as embodied activity in medical interactions. In S. Wiggins & K. O. Cromdal (Eds.), *Discursive Psychology and Embodiment: Beyond Subject-Object Binaries* (pp. 197–220). Springer Nature.
- LaBonte, A., Hindmarsh, J., & vom Lehn, D. (2021). Data collection at height: Embodied competence, multisensoriality and video-based research in an

extreme context of work. *Social Interaction. Video-Based Studies of Human Sociality*, 4(3).

Merlino, S. (2021). Making sounds visible in speech-language therapy for aphasia. *Social Interaction. Video-Based Studies of Human Sociality*, 4(3).

Meyer, C., & v. Wedelstaedt, U. (Eds.). (2017). *Moving Bodies in Interaction – Interacting Bodies in Motion: Intercorporeality, interkinesthesia, and enaction in sports*. John Benjamins.

Mondada, L. (2018). The multimodal interactional organization of tasting: Practices of tasting cheese in gourmet shops. *Discourse Studies*, 20(6), 743–769. <https://doi.org/10.1177/1461445618793439>

Mondada, L. (2019). Contemporary issues in conversation analysis: Embodiment and materiality, multimodality and multisensoriality in social interaction. *Journal of Pragmatics*, 145, 47–62. <https://doi.org/10.1016/j.pragma.2019.01.016>

Mondada, L., Bouaouina, S. A., Camus, L., Gauthier, G., Svensson, H., & Tekin, B. S. (2021). The local and filmed accountability of sensorial practices: The intersubjectivity of touch as an interactional achievement. *Social Interaction. Video-Based Studies of Human Sociality*, 4(3).

Nevile, M. (2015). The Embodied Turn in Research on Language and Social Interaction. *Research on Language and Social Interaction*, 48(2), 121–151. <https://doi.org/10.1080/08351813.2015.1025499>

Nishizaka, A. (2017). The Perceived Body and Embodied Vision in Interaction. *Mind, Culture, and Activity*, 24(2), 110–128. <https://doi.org/10.1080/10749039.2017.1296465>

Pink, S. (2009). *Doing sensory ethnography*. Sage.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, 50(4), 696. <https://doi.org/10.2307/412243>

Schegloff, E. A. (1987). Analyzing Single Episodes of Interaction: An Exercise in Conversation Analysis. *Social Psychology Quarterly*, 50(2), 101–114. <https://doi.org/10.2307/2786745>

Schegloff, E. A. (2009). One perspective on conversation analysis: Comparative perspectives. In J. Sidnell (Ed.), *Conversation Analysis: Comparative Perspectives* (pp. 357–406). Cambridge University Press.

- Smith, M. S. (2021). Achieving mutual accessibility through the coordination of multiple perspectives in open, unstructured landscapes. *Social Interaction. Video-Based Studies of Human Sociality* 4(3).
- Stevanovic, M., Himberg, T., Niinisalo, M., Kahri, M., Peräkylä, A., Sams, M., & Hari, R. (2017). Sequentiality, Mutual Visibility, and Behavioral Matching: Body Sway and Pitch Register During Joint Decision Making. *Research on Language and Social Interaction*, 50(1), 33–53.
- Streeck, J. (2003). The body taken for granted: Lingering dualism in research on social interaction. In P. J. Glenn, C. D. LeBaron, J. S. Mandelbaum, & R. Hopper (Eds.), *Studies in language and social interaction: In honor of Robert Hopper* (pp. 427–440). Erlbaum.
- Streeck, J. (2013). Interaction and the living body. *Journal of Pragmatics*, 46(1), 69–90. <https://doi.org/10.1016/j.pragma.2012.10.010>
- Stukenbrock, A., & Dao, A. N. (2019). Joint Attention in Passing: What Dual Mobile Eye Tracking Reveals About Gaze in Coordinating Embodied Activities at a Market. In E. Reber & C. Gerhardt (Eds.), *Embodied Activities in Face-to-face and Mediated Settings: Social Encounters in Time and Space* (pp. 177–213). Springer International Publishing. https://doi.org/10.1007/978-3-319-97325-8_6
- Sudnow, D. (1993). *Ways of the Hand: The Organization of Improvised Conduct*. MIT Press.
- Torreira, F., Bögels, S., & Levinson, S. C. (2015). Breathing for answering: The time course of response planning in conversation. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.00284>
- Weatherall, Ann, Leelo Keevallik, Jessica La, Maria Stubbe, Tony Dowell (2021) The multimodality and temporality of pain displays. *Language and Communication*, 80, 56–70.