Sounding for others: Vocal resources for embodied togetherness

Leelo Keevallik, Emily Hofstetter*
Linköping University, Sweden

Article info
Article history:
Available online 11 March 2023

Abstract
Standard models of language and communication depart from the assumption that speakers encode and receive messages individually, while interaction research has shown that utterances are composed jointly (C. Goodwin, 2018), dialogically designed with and for others (Linell, 2009). Furthermore, utterances only achieve their full semantic potential in concrete interactional contexts. This SI investigates various practices of human sounding that achieve their meaning through self and others’ ongoing bodily actions. One person may vocalize to enact someone else’s ongoing bodily experience, to coordinate with another body, or to convey embodied knowledge about something that is ostensibly only accessible to another’s individual body. This illustrates the centrality of distributed action and collaborative agency in communication.

© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

The prototypical use of language and other human vocal tract sounds is to express one’s own ideas, experiences, and actions in a vocal form. Yet every day we also use those resources to empathize with others, co-construct meaning, (co)-enact others’ actions, stories, and various mental and bodily experiences. We respond to stories with the reactions the teller may have performed in the described event (Cantarutti, 2022). We urge others to put in more effort with our own voices full of strain and our own bodies full of tension. We watch someone get injured and suck in our breath at their pain (Heyes and Catmur, 2022). A mother sounds for the tastes their infant is experiencing (Wiggins and Keevallik, 2021), a teacher sounds to remind students of a sequence of dance steps (Keevallik, 2021), a doctor for a patient’s pain (Weatherall et al., 2021). In this special issue, we explore different ways in which we “sound for others”, a phenomenon whereby we instruct and express the experiences of our interlocutors as if they came from our own bodies, re-considering communication at the boundary between language and non-language as joint action with separate physical bodies.

The phrase “sounding for others” is meant as an analytic umbrella term; “sounding” includes a wide variety of vocalizations and various vocal behaviors beyond speaking; “for” includes both for the benefit of others and on their behalf; and “others” includes any co-participants or interlocutors. The studies in this issue differ accordingly; sounding for the benefit of others usually scaffolds the actions of other participants, such as guiding instructees on what they ought to be doing (Mondada, Okada), while sounding on others’ behalf involves sounds that the other participant did or might have done themselves (Albert & vom Lehn; Ben-Moshe; Cantarutti; Nomikou; Weatherall) or such as (co)-sounding for others’ physical

* Corresponding author.
E-mail address: emily.hofstetter@liu.se (E. Hofstetter).

https://doi.org/10.1016/j.langcom.2023.02.002
0271-5309/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
exercise and for collective excitement upon someone’s success (Keevallik & Hofstetter; Tekin), thus there is even sounding with-and-for others.

The ability of humans to vocally enact the experiences of others demonstrates (further) evidence that vocal behavior is not a system of linear communication of one person conveying something to the other, but an embodied, multisensory and collaborative resource of meaning making in interaction. This special issue collects evidence cross-linguistically from several speech communities (English, Estonian, German, Hebrew, Italian, Japanese, Turkish) and settings (teaching, chatting, playing, exercising, grooming, learning to dance), and together the studies recommend a dialogic (Du Bois, 2014; Linell, 2009), action-oriented (Garfinkel, 1967) theory of language. The papers herein undertake this project by asking how the sounding for others phenomenon lets participants accomplish being with others and what contingencies are involved in so doing. The special issue (SI) takes as its springboard the latest methodological and theoretical developments in ethnomethodological conversation analysis, which has demonstrated the empirical benefit of paying close attention to the multimodal details of coordinated behavior (C. Goodwin, 2018; Mondada, 2019b). In this editorial, we overview sounding for others as a family of related phenomena, how it is interconnected with previous conceptualizations of language, and introduce the papers in this issue.

2. Representing other's voices

The traditions of both structural and functional linguistics (beginning in de Saussure, 1916; and living on in, e.g. V. Evans, 2019; Halliday, 2014) examine (mostly hypothetical) utterances as produced by individuals, often with minimal consideration of actual contexts, let alone other participants. However, language is a multiparty activity; not only are utterances done to be heard by others (e.g., Deppermann, 2013), but the use of language involves recycling the voices of others and speaking for others (Bakhtin, 1981). This kind of a complex understanding of a speaker as either a mere animator, the author who selects the wording, or the responsible principal, was first delineated by Goffman (1981, pp. 144–145) and then prominently diversified across speakers and recipients by Goodwin and Goodwin (1986, 1987; though see also Bakhtin, 1981; Volosinov, 1973). Giving a voice to others is an empirical fact to be studied in its own complexities (Holt and Clift, 2007; Wilkinson et al., 2020), but also important for our conceptualization of what language is and how it functions.

Dialogism (Bakhtin, 1981; Du Bois, 2014; Ducrot, 1984; Linell, 2009) provides such a conceptualization, describing how utterances are imbued with the voices of others. Bakhtin (1981) grounded his theory of dialogism in the way that language is inherently polyphonic, comprised of a myriad of contributing voices from history. On a diachronic scale, we acquire a system for communication when we are born into it and begin to participate in the available culture (Raczaszek-Leonardi, 2009). The sum of historical language use provides expected meanings for words and orders for turn design, which we must adopt in order to be seen as using that language. Additionally, on an enchronic scale (Enfield, 2013), language users constantly adapt their utterances to manage recipient attention, intersubjectivity, and co-participation. This occurs live, even in the course of an utterance; Goodwin (C. Goodwin, 1979) classically demonstrated how a sentence unfolds in connection with the involvement of a series of recipients. Since then, the study of emergent syntax (Hopper, 1987) has repeatedly illustrated the ways co-participants are involved in the production of utterances (Maschler et al., 2020). The notion of a speaker (and their utterances) being fully separated from a recipient seems to be a product of Western (especially Cartesian) philosophy, as linguistic anthropological studies report that many other cultures’ conceptualizations of language regularly attend to the shared nature of speakership and recipiency (Liberman, 1985; Meyer, 2010; Robbins and Rumsey, 2008; Rosaldo, 1973; see Johnstone, 2000). Interaction studies in Western cultures have also demonstrated how a properly behaving recipient is required or implied in every utterance, which includes several kinds of actions that are essentially co-constructed (C. Goodwin, 2004; M. H. Goodwin, 1990; Helasvuori, 2004; Lerner, 1991), chorally produced (Lerner, 2002), or where recipients’ actions are otherwise attended to in the ongoing turn (De Stefani, 2021; Iwasaki, 2009). Co-participants’ mutual involvement in each turn and action is thus well-documented, though the ‘speaker’ is typically seen as contributing the majority of the multiple resources used to construct a turn.

Sounding for others distributes those resources in a more radical way. In a literal sense, there are many linguistic practices that sound for others, that is they repeat or enact (or purport to) the utterances and actions of others. Reported speech (Holt and Clif, 2007), for instance, can quote others’ former utterances (Goffman, 1981), repeating prior talk for current situated purposes, though it regularly redesigns others’ words, such as summarizing, exaggerating, calibrating prosody, or inserting laughter particles, which are taken to be the current speaker’s stance on the reported content (C. Goodwin, 2007; Günthner, 1999). The utterances, presented as originating from third parties, thus feature multilayered voices. Reformulation is another way to give voice to another’s words and its interactional use has frequently been studied in therapeutic settings (Ekberg, 2021; Peräkylä, 2019; Weiste and Peräkylä, 2013), as it is done as a means of demonstrating attentiveness, shared understanding, and empathy with co-present participants (Ford et al., 2019; Hepburn and Potter, 2007; Ruusuvuori, 2005; Versteeg and te Molder, 2016). Finally, animations (Cantarutti, 2020), also termed reenactments (Sidnell, 2006) and belonging to the broader category of depictions (Clark, 2016; Löfgren and Hofstetter, 2021), are a practice whereby speakers can enact, perform, or embody others. While the above practices were traditionally studied with respect to vocal, verbal language use, the entire field has gradually moved to explore not only prosody but the use of the body in doing reported speech (Niemelä, 2010; Thompson and Suzuki, 2014) and animations (Cantarutti, 2020), showing how gesture, gaze, body position, and movement can multimodally depict co-participants and organize utterances as being on others’ behalf.

Across the papers in this SI, we showcase a variety of distributions of vocal and bodily resources across different participants, problematizing a simplistic notion of there being one speaker and agent at any given time in interaction. The role of the
body is particularly critical for the studies in this issue, as each deals with moments when participants sound for others’ bodies.

3. Communicating bodily matters

3.1. Voicing participants

Past studies have foregrounded spatially and/or temporally displaced (re)voicings. The handling of bodily events on someone else’s behalf is usually discussed within the context of its limitations (Heritage, 2011; Throop, 2008) or their infringement on agency (Robillard, 1996). One exception has been Keevallik’s (2010) research on bodily quotation, wherein instructors redo the bodily motions of students, often in an exaggerated manner, in order to reflect back how a bodily action should and should not look. However, bodily quotation specifically refers to using one’s own body to redo or ‘quote’ another’s bodily movements, typically for the purpose of demonstration. In this SI, the papers aim to investigate the vocalizing of other’s bodies, which for many papers involves focusing on events that have a strong sensorial and/or emotional component, where the ‘internal’ experience can also be taken up by participants that ostensibly did not personally experience the phenomenon.

For example, Cantarutti (this issue) shows that participants can literally produce sounds on behalf of others in responsive animations, combining non-lexical vocalisations and gestures in ensembles. This entails momentarily shifting from the usual exchange of independent conversational contributions by individuals to a joint engagement with a single participant’s experience. The animations are produced in response to a participant’s self-deprecating disclosure and amplify denouncing components while creating brief moments of heightened other-attentiveness. They constitute a form of distributed agency (Enfield, 2017) between participants, with one offering a described attribute or experience that the other turns into an embodied demonstration.

Nomikou (this issue), likewise, dissect vocalizations that could have been produced by the co-present other, in this case an infant. Grunt-like sounds that accompany caregivers’ haptic handling of infants occur with minute temporal precision, and represent the infant’s putatively ongoing sensation (as has also been discussed across various activity contexts by Keevallik et al., 2023). The paper suggests that this multimodal coordination of bodily-motivated sounds with the infant’s movements, i.e., a way of sounding “on behalf” of another, could form the bases of developing more advanced forms of sociality in later life.

In a different instructional setting, Albert and vom Lehn (this issue) explore how novice dancers use non-lexical vocalizations when learning to navigate unfamiliar dance movements together. The vocalizations are produced along with apologies, accounts, and embodied actions that mark out moments of coordination trouble and provide reference points to account for, evaluate, and re-animate the experiences of otherwise inchoate sequences of joint embodied action. Sounding is here used as a communicative device for representing underspecified bodily experiences and unfamiliar movements, for the benefit of co-present others. The authors argue that vocalizations (both lexical and non-lexical) are used systematically to render the bodily movements available and mutually observable, thus problematizing the boundary between body and language. Communicating body knowledge constitutes an essential challenge for participants (see Brône and Ehmer, 2020) and in this SI we show how vocal resources become crucial in accomplishing that.

3.2. Voicing sensations

The phenomenon of multi-participant involvement in sensing is well-documented in ethnographic studies of sensation (sensory ethnography, Pink, 2009; anthropology of the senses, Classen, 1997; phenomenological studies of everyday sensing, Allen-Collinson and Hockey, 2009). This line of research documents how senses are (re)produced through culture and institutions. The recurrent refrain in ethnographic studies of sensing is that to sense is to participate, and to participate is to sense. Chau (2008) demonstrates this by documenting how participants at a temple festival not only sense the atmosphere of excitement and energy but contribute to it by being there. Ethnographic studies of sport in particular highlight collective participation in sensing, for instance ‘communal sounding’ (Schmitz and Effenberg, 2017), where the sounding can be for the self or other co-participants, and it can also be used in the martial arts practice when jointly undergoing intense exertion (Bar-On Cohen, 2009).

At times, multiple participants chorally engage in sounding for others, a phenomenon explored by Tekin (this issue). Expanding on previous work on choral or collective speaking, such as in collaborative completions (Lerner, 2004), audience responses (Atkinson, 1984; Broth, 2011), and cheering at sports events (Kerrison, 2018), Tekin demonstrates that cheering together as a choir during video gaming is an interactive accomplishment through which participants manifest their co-engagement in the events. By way of establishing, sustaining, modifying, and terminating their choral vocalizations and embodied displays in a concerted manner, participants display collective agency, essentially dissolving the distinction between self and other.

Ethnographic studies of sensation and interactional research share an interest in understanding how human sensation is socially organized, and both focus on describing local, situated sensing. As Pink (2009) discusses, sensory ethnography is more a collection of observations, despite calls for theories that take a non-dualistic, embodied ontology of social interaction (Bargiela-Chiappini, 2013; Crossley, 1995). Recent research in interaction has focused especially on moments where sensation is topicalized for local purposes, such as tasting sessions (Mondada, 2018, 2019a, 2021) or assessing perceptual capabilities (on vision checks, Gibson and vom Lehn, 2020). The sharing of those events is a multimodal accomplishment, where
embodied enactment of sensing such as sniffing or visually inspecting underscores the immediacy and authenticity of the event, while the verbalization provides an opportunity to highlight what categories are relevant for the task at hand (C. Goodwin, 1994), from purchasing (Mondada, 2018) to determining data points (C. Goodwin, 1995) to identifying fetal body positions (Nishizaka, 2011).

A commonality in dealing with more explicit topicalization of sensation is that the verbalizations and labels of sensations are used to assign official or enculturated categories, such as assessing the quality of goods (Mondada, this issue), marking the relevant and correct code when doing fieldwork (C. Goodwin, 1994), or diagnosing medical and procedural difficulties (Ekström and Lindwall, 2014; B. Evans and Reynolds, 2016; Nishizaka, 2014). As a result, instructional environments, where the sensing is made relevant to explain, are both highly perspicuous for interactional research, and (thus) the most common activity in focus in multisensory interactional work (as seen in much of the above work, see also Cromdal et al., 2020, p. 202; Ehmer, 2021; Lindwall and Lymer, 2014; McIlvenny, 2019; Reed, 2021; Simone and Galatolo, 2020; Zemel and Koschmann, 2014). When teaching or showing others how to engage the body in particular ways or to particular ends, the participants display explicit attention to each other’s bodily experiences and draw on intercorporeal understandings of the body to better assist others in tuning into or producing the right sensations (Bäckström, 2014). Such instructions produce sense-oriented interactions that make the mutual attention to each other’s bodies available to the researcher for study (Downey et al., 2015). Building on this stream of work, the current issue targets vocal practices of being “with” other bodies and instructing them. The papers expand on the topicalization of sensing in investigating non-lexical vocalizations, which are limited in conventionalization of meaning (Keevallik and Ogden, 2020), and target temporality, the achievement of sequentiality and simultaneity in the collaborative production of sensing.

Okada (this issue) studies how time-critical moments in boxing moves are co-constructed by the boxer and the coach who produces lexical repetitions pronounced with latching, vowel lengthening, higher pitch or creaky voice. Either occurring before the targeted action or alongside them, those vocal performances can also amount to joint experiencing of the opponent’s punches, when accompanied by fitting body positioning and tactile opportunities by both participants. The study shows how vocal resources by one participant are used to enter the experiential realm of another, thus sound together “with” that person.

Hofstetter and Keevallik (this issue) discuss how participants can coordinate different modalities in order to bring about action. Studying pilates practice, they show how the instructor not only structures others’ bodily practice through locally establishing repetitive prosodic patterns, but also constitutes these very actions by simultaneously representing their temporal and sensorial aspects through voice quality and loudness. At these moments, vocalization by one participant and embodied performance by others become joint action.

Studying experiential body-based sounds at tasting sessions, Mondada (this issue) describes the systematic use of audible sniffs, that emanate from somatic aspects of the body but index smelling. They make the sensing of smell publicly audible, demonstrating the embodied nature of interactional resources and producing sensorial intersubjectivity in engagements with materiality. The paper shows how non-lexical sound resources of various sniffs contribute to the mutual understanding of embodiment and sensoriality, and illustrates how they are methodically mobilized together with the body in social interaction. Since the sniffs are potentially oriented to by co-participants, they can be said to be produced “for the benefit” of the recipients, who can treat them as instructions to smell themselves in this activity context.

Weatherall (this issue) demonstrates that experiential sensory-based vocalizations can furthermore be produced on behalf of non-present others, in the context of self-defence training. Focusing on performances of pain that would be experienced by imaginary assailers, the paper argues that the vocalizations display an understanding of the defence techniques as pain-relevant. The pain cries thus collaboratively complete the performances of the pedagogically constructed scenes, while displaying alignment with the instructor’s stance that impactful self-defence resists attack by inflicting harm to the assailer.

3.3. Voicing emotions

Emotions are a further experience that are traditionally seen as individual and inaccessible without signalling (e.g., Cowen et al., 2019; Ekman, 1992; Keltner et al., 2019). Meanwhile, interactional research has demonstrated that emotion is organized together with co-participants, with verbal expressions (Peräkylä and Sorjonen, 2012), and increasingly with the body (Robles and Weatherall, 2021). Interactional research is not the first to stipulate a social function to emotion displays (see e.g., Fridlund, 1991), however it emphasizes and empirically documents the way that the events we label as emotions arise out of social action rather than individual response. Emotions are achieved together, rather than signalled, and are even intercorporeal (Katila and Philipson, 2019). Targeting the intersubjective understanding of emotion, Ben-Moshe (this issue) shows how ‘gasps’ blur the lines between body and language, self and other: by vocalizing a bodily experience appropriate for another’s point of view, the speaker simultaneously interprets this experience as her own and thereby expresses empathy. Gasps invoke the impression of an overflowing of emotion, gaining additional expressive force from the iconic link to the physiological startle response, thus also blurring the boundary between experience itself and its expression.

We are thus offering an array of analytical considerations at the interface of language, vocal and bodily expression, and the somatic concerns across activity contexts.
4. Future directions

The SI on ‘sounding for others’ showcases a range of phenomena where sensations, emotions, and proprioception are made relevant through vocalization for other participants. Together, these studies demonstrate the inappropriateness of theories that divide individuals into single sources of messages. Though we analytically separate speakers’ bodies, we do this in order to demonstrate their temporal co-ordination and by extension address theoretical shortfalls. The participants themselves in our data take the availability of each other’s bodies for granted, as they can even simultaneously co-enact for and with each other, merely by distributing modalities of expression. Especially in sounding for each other’s bodily events, the participants maximally highlight co-presence, engage in an intense form of sociality, and provide evidence of the dialogic nature of both vocal communication and social action. Under this perspective, individuality is a mode of being, rather than a prerequisite for communication. Participants can of course index rights to speak for themselves (e.g. with mental formulations, Peräkylä and Silverman, 1991; Weiste et al., 2016; Zinken and Kaiser, 2020), and thereby foreground their separateness and individual agency (see also in co-tellings, Dressel and Satti, 2021), however the other end of the spectrum permits them to dissolve such territories (Hayano, 2016) and be maximally together (Goodwin, 2018). Both individuality and collectivity are achieved through language (Enfield and Kockelman, 2017).

Sounding for others is one way to demonstrate attunement to co-participants (in both a lay and phenomenological sense, Merleau-Ponty, 1968) and ratify displayed togetherness, or individuality. Merleau-Ponty (1968) called this sharing and co-sensing intercorporeality, which can be defined as the interconnectedness of our bodies with each other and with the environment, and the way experience is mediated by interactions with others, objects, and places (Weiss, 1999). Intercorporeality is underdeveloped in connection with interactional research (the extant contributions being Katila and Philipsen, 2019; Katila and Raudaskoski, 2020; Meyer et al., 2017; Meyer & v. Wedelstaedt, 2017). The questions of how intercorporeality can be established and maintained, how it is ratified, and whether it is a matter for intersubjectivity, have only just begun to be explored. While phenomenological theory takes intercorporeality as an assumed ongoing phenomenon, interactional studies would want to investigate how and when its presence or absence becomes an issue for their ongoing action – when it becomes relevant for participants. The majority of instances of sounding for others in this special issue could be understood as relying on intercorporeality, an empathic connection to and understanding of another’s body. Further work is needed to explore this connection on the premises of interactional methods.

The extensive observations of social sensing in ethnographies and interactional research demonstrate that co-participation in others’ sensory and perceptual phenomena is a regular occurrence. Perception should thus be conceptualized as cultural, while recognizing the local work of co-participants to share and ratify sensory and affectual events. While empirical evidence is amassed through interaction studies, including the work in this issue, connections to adjacent theoretical fields have, to date, been minimally explored. Not only phenomenology and ethnography of the senses (see above) but also in distributed (Cowley, 2011; Hutchins, 1995) and enactive cognition (Di Paolo et al., 2018) paradigms do we see evidence and theorizing of language and sensing as socially organized, emergent phenomena, though relying on very different theoretical backgrounds and assumptions (De Jaegher et al., 2016). We hope that the papers in this issue can provide empirical bases for analyzing language as dialogic and emergent, to be explored from various ontological premises, but crucially highlighting the everyday work that participants do to achieve being and sensing together.

Acknowledgements

This work has been funded by the Vetenskapsrådet [2016-00827] and Riksbankens Jubileumsfond [P21-0447].

References

