

Walking through the city soundscape: an audio-visual analysis of sensory experience for people with psychosis

SARA MERLINO 

University of Roma Tre, Italy

LORENZA MONDADA

University of Basel, Switzerland

OLA SÖDERSTRÖM

University of Neuchâtel, Switzerland

ABSTRACT

This article discusses how an aspect of urban environments – sound and noise – is experienced by people walking in the city; it particularly focuses on atypical populations such as people diagnosed with psychosis, who are reported to be particularly sensitive to noisy environments. Through an analysis of video-recordings of naturalistic activities in an urban context and of video-elicitations based on these recordings, the study details the way participants orient to sound and noise in naturalistic settings, and how sound and noise are reported and reexperienced during interviews. By bringing together urban context, psychosis and social interaction, this study shows that, thanks to video recordings and conversation analysis, it is possible to analyse in detail the multimodal organization of action (talk, gesture, gaze, walking bodies) and of the sensory experience(s) of aural factors, as well as the way this organization is affected by the ecology of the situation.

KEYWORDS

conversation analysis • multimodal analysis • multisensoriality
• psychosis • sensory aural factors • social interaction • urban
environment • video methods

Visual Communication 2022

Vol. 22(1) 71–95

© The Author(s) 2022



Article reuse guidelines: sagepub.com/journals-permissions

DOI 10.1177/14703572211052638



1. INTRODUCTION

Despite an increasing interest in how local ecology impacts social interactions, the way sensorial features of the environment are oriented to and integrated in social exchanges remains understudied. In particular, surrounding sounds and noise have been neglected in interactional studies. This article offers a video-based approach to aural–visual phenomena characterizing social activities in urban contexts, with a special focus on how people walking in the city orient to the sonic ecology of spatial and material surroundings – their *soundscape*. It focuses on patients with psychosis, who are known to be particularly sensitive to sensory environments, specifically noise. The article contributes to several lines of inquiry. It adds to *sound studies* in an original way, by including not only sounds but a more holistic view of the *soundscape*, considering the bodies of those who hear them and their responses to them. It contributes to a recent interest of *interactional studies* in *sensoriality*, by suggesting that the perception of sounds can be integrated in a multimodal multisensorial analysis. It also contributes to the *study of atypical populations* within the city, namely persons living with a diagnosis of psychosis, who are notoriously affected by noise, but who have not yet been observed as they actually experience noisy situations. These contributions are developed within the perspective of multimodal conversation analysis, where video enables an understanding of how participants in social interaction mobilize talk and their bodies within particular ecologies.

2. STATE OF THE ART

The idea of *soundscape* has been proposed by Schafer (1977) to refer to our sonic environment and our ability of distinguishing nuances of sounds not only in nature but also in cultural and urban spaces. Sound qualities and aural sensoriality have been addressed within *sound studies* (Pinch and Bijsterveld, 2012; Sterne, 2012), including *acoustemology*, an ethnographic exploration of sonic sensibilities (Feld, 2015). The sound quality of the environment is a fundamental dimension of the world we inhabit and of our sensory experiences. Although cities are often considered as noisy places, the array of sounds that populate them is much larger and can be categorized in many ways (Thibaud, 2003), which may be variously considered by citizens: ‘the spaces of the city form an ordered as well as a temporally defined ecology of noise, sound and occasional silence and one which is regularly contested at both the individual and broader political scales’ (Atkinson, 2007: 1908).

How to capture, then, the ways in which the soundscape actually impacts the details of social interaction? Despite developments in conversation analysis concerning multimodality (Goodwin, 2017; Mondada, 2016; Streeck et al., 2011), and multisensoriality (Mondada, 2019), as well as an interest in the spatial surroundings of social interaction (Bergmann, 1990;

Haddington et al., 2013), non-linguistic sounds and their sensory appraisal have been largely neglected in this approach. Recent contributions (Keevallik and Ogden, 2020) focus on the sounds of the body (like grunts, sniffs, etc.), rather than sounds of the environment or noisy contexts. An exception are Heinemann and Rauniomaa (2016) and Rauniomaa and Heinemann (2014) examining how participants adjust to ambient noise, by adapting their talk and by turning on or muting auditory objects – in the latter case, displaying prioritizing talk. By focusing on how noise is oriented to and integrated in actual situated practices, we propose an approach of aural sensoriality in conversation analysis.

Noisy environments have been identified as particularly harmful for the global population in general, and for people with psychosis in particular (Vassos et al., 2012). Noise is a well-documented source of stress in psychiatric research on psychosis. However, research in psychiatry on ‘auditory anomalies’ is mostly lab-based and therefore abstracted from ordinary everyday life contexts. The impact of noise on people with psychosis has been largely identified on the basis of self-reported descriptions (Landon et al., 2016), sound evaluation by patients (Micoulaud-Franchi et al., 2012) or tasks observed in lab settings (Smucny et al., 2013; Wright et al., 2016), showing how noise is linked to delusions and cognitive difficulties. But not much is known about how people experience the soundscape of the city in real time, and in situated activities. Our video-based approach enables us to explore this aspect, not only enriching previous observations with naturalistic data, but also revealing detailed circumstances that may constitute aggravating environments for persons with psychosis, and for how they handle them.

3. METHODS

The stroll is an exemplary ordinary activity in which to observe how people walk across different urban environments; how they orient to the surrounding spatial and sensory details; and how they relate to their companion during the walk, while talking with them or walking silently. Through video-recordings of patients strolling through the city with their partners, we provide a documentation of actual embodied conducts as they happen, and we show how patients orient to noise and report their aural noticings in situ and in real time. This methodological approach contrasts with experimental approaches of noise and psychosis, but also with ethnographic approaches to soundscapes: it allows for observation of how people actually behave in ordinary natural situations and orient to and make relevant, through their embodied conduct, features of the surrounding environment. Moreover, we combine video-recordings of naturalistic activities in an urban context with video-elicitations, based on these recordings, that produce retrospective and general comments related to personal experiences. The first approach is essentially

based on what people *do*, often in an embodied silent way. The second focuses on what people *say*. The first captures the contingencies of action in a way that is more subtle and delicate than within explicit formulations; the second provides for explicit declarations, which escape the context of the experience and rather elaborate on it and generalize it. Documenting sensoriality through video recordings or elicited discourses always encounters some limitations: video enables the capture of intersubjectivity of sensorial experiences, rather than individual private sensations; elicited discourses enable a demonstration of how these experiences are interpreted in a situated way and are socially shared. This article not only presents the results of these methodologies but also shows how to analytically reflect on the approaches themselves.

Conversation analysis uses video recordings as a way of documenting situated activities and social interactions as they happen within their ordinary settings, in the least intrusive way possible and without orchestrating the activities of the participants (Heath et al., 2010; Mondada, 2012), although without hiding any cameras, for ethical purposes. Video has also been used in methodologies using interviews, for supporting recollections, reflections and interpretations of past events: video recordings of activities are shown either to the same participants or to other relevant persons, who watch them, decide on which significant moment to pause, and comment on them (Henry and Fetters, 2012; Pomerantz, 2005).

4. DATA

This study is part of a larger project on psychosis in the city, conducted by an interdisciplinary team including psychiatrists, geographers and linguists (Söderström, 2019). We video recorded a series ($N = 10$) of walks in which a participant affected by psychosis and an accompanying person (a friend or a family member) were going for a stroll across a city in Switzerland. The participating patients ($N = 10$) are part of a larger cohort involved in the Treatment and Early Intervention in Psychosis Program (TIPP) launched in 2004 by the Department of Psychiatry at the University Hospital in Lausanne (Switzerland). The patients have already been diagnosed as suffering from psychosis and the medication they have been prescribed may have an impact on their sensory perceptions. However, a broader survey of a larger group of patients shows that sensitivity to noise is important both before and after onset (Conus et al., 2019). The present study aims to make more complex the description of the sensory orientations of the participants and their partners in naturalistic settings.

Each recorded walk took between 20 and 90 minutes. We asked for permission to record ordinary promenades to places the participants would usually go, without any particular instruction. Three cameras were used: one held by a researcher following the pair; another by a second researcher filming laterally; and a third placed on the chest of the accompanying person, thus capturing both the walk itself and its environment.

The position of the cameras was chosen in order to best capture the participants without being in their frontal field of vision and without influencing their trajectories. The sound was recorded using a wireless microphone positioned on one of the two participants.

The videos were produced and studied by the linguists in the team (see Merlino and Mondada, 2019) using multimodal conversation analysis, providing materials for the first part of this article. They were also re-used in video-elicitations, in which the patients watched the video of their walk, pausing it whenever they wanted to comment on particular moments of that experience (the mean length of video-elicitations was 60 minutes). The video-elicitations were also video-recorded; their verbal transcripts provided material for content analysis by the geographers of the team (see Söderström et al., 2016, 2017). Their detailed multimodal transcription provides materials for the second part of this article.

Multimodal transcripts of video data (Mondada, 2018) annotate the resources (such as language, gesture, gaze, facial expression, head orientation, body posture and movement) made relevant by the participants organizing their actions in a publicly accountable way (Goodwin, 2017; Heath, 1986; Mondada, 2016; Streeck et al., 2011). They allow the analyst and the reader to reconstruct the detailed temporal unfolding and fine-tuned adjustment between a multiplicity of linguistic and embodied resources organizing the coordination of the participants' joint actions.

5. ANALYSIS

Our analysis is based on a sub-set of episodes in which some urban *sound* happens, at least one participant orients to it and often categorizes it as 'noise'. We focus first on what happens in the street as people walk together (§5.1); then on how participants comment on these episodes when viewing the video-recordings (§5.2). The comparison of these two moments shows the intricate relationships between live experiences and re-enacted ones. It also reveals their differences and complementarities: video-recordings in situ show embodied experiences and lively reactions to noise that are rarely commented on in video-elicitations, while the latter generate general comments about recurrent problems, rather than references to specific events and circumstances.

5.1. Hearing sounds while walking and talking in the street

This section focuses on sound events during the video-recorded strolls. It shows that participants' perceptions of/orientations to noise varies: they can either orient to noise or not (§5.1.1), and we demonstrate how this is sensitive to the type of engagement of the participants in the conversation. Furthermore, we highlight that orientation to noise can be either individual or public, either silently embodied or verbally commented upon (§5.1.2).

Finally, even in the absence of comments or embodied orientation, we show how noise can overtly affect interactional conduct, disturbing the exchange (§5.1.3). Video evidence thus enables a fine-grained differentiation between different responses to noise, which will be further discussed on the basis of the retrospective comments in the video-elicitation (§5.2).

5.1.1. Different orientations to noise while walking and speaking/listening

Car horns are pervasive in the urban soundscape. Their occurrence is generally unexpected, their sound quality is difficult to ignore and their normative meaning immediately available (interpretable as signalling a possible dangerous situation or sanctioned urban conduct). While some pedestrians might be indifferent to these sounds, others might clearly orient to them, by turning to them and/or by uttering comments about them. In the following excerpts, we show that this orientation to car horns is sensitive to the type of engagement of the participants in the conversation during the walk.

We join the first excerpt as Benoit¹ is talking about his flat-mates to his friend, Nadia (see lines 1-2 and Figure 1 of Excerpt 1, reproduced here below). The transcript follows the conventions of conversation analysis, respecting the specificities and the timing of the flow of talk and embodied conduct of the participants (body movements are precisely timed in relation to portions of talk or silences, thanks to symbols indicating their length and position, like for example ‘+gaze+’). The sound of a car horn occurs during Benoit’s story (3). Nadia visibly turns her head in the direction of the noise (Figure 2), while Benoit does not display any particular orientation to it:

```

1  BEN      #après je lui ai dit >vas-y prépare ta chambre, y a
           then I told him go on get your room ready there are
2  des gens qui vont ven[ir< (.) tu +veux$]pas la$#montrer&
           some people who are coming (.) you don't want to show it
3  CAR →    [ ((car horn)) ]
           ben      >>looks in front-----+looks right---->
           nad →    >>looks in front-----$,,,,,,$turns back->
           fig      #fig.1                                     #fig.2

```

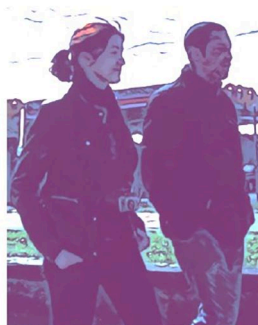


fig.1

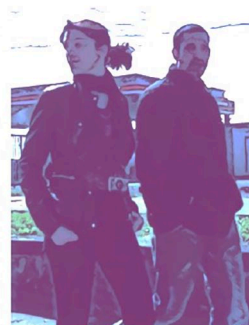


fig.2

4 BEN &comme ça: \$tu #+vois?
 like this you see
 nad -->\$turns\$looks at BEN-->
 ben -->+looks at NAD-->
 fig #fig.3

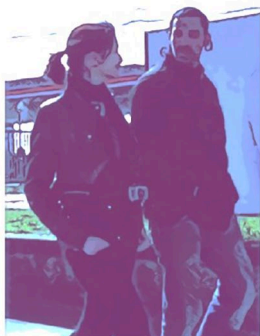


fig.3

5 (0.3)\$ (0.5)\$+(0.4)
 nad \$nods\$
 ben -->+looks in front-->
 6 BEN et puis euh:: lui il m'a dit euh::\$: AH non y a de gens
 and then eh:: he told me eh::: OH no are there some people
 nad -->\$looks in front-->
 7 qui qui viennent
 who who are coming

Excerpt 1 (B_W_1604)

When the car horn sounds, Benoît continues his turn and completes it, without manifesting any concern about the event. Rather, the direction of his gaze and the continuation of his talk show that he is attentively orienting to his interlocutor. Indeed, when Nadia looks back at him, he also looks at her and produces a turn final tag-question (*tu vois?* 4, Figure 3), to which Nadia responds by nodding. Once he has secured his partner's attention and mutual gaze again, he looks ahead and continues his story. So, in this excerpt, the recipient of the story orients to the noise, while the speaker engaged in his story, does not.

A similar case of asymmetric orientation to a car horn is observable in the next conversation (Excerpt 2), in which Christian has announced he has possibly found a flat in a village nearby and Sandra comments about the distance from the city centre:

1 SAN .h ouais puis# vous êtes tout près de ((city name)),
 .h yeah and then you are very close to ((city name))
 fig #fig.4a/b



fig. 4a

= fig.4b

2 (0.9)
 3 SAN +après# avec le train vous êtes à+ un quart d'heure,
 then by train you're fifteen minutes away
 car +looks back-----+
 fig #fig.5a/b



fig.5a

= fig.5b

4 j'crois, pas plus.
 I think, no more
 5 %(0.8) [(0.2)] (2.7)%
 %bus overtakes them--%
 6 eve → [((horn))]
 7 SAN et puis c'est jol[i eu]+h: même ((village name))
 and then it's nice eh: also ((village name))
 8 CAR → [((horn))]
 +looks L/at the road->
 9 c'est+: (2.3)+ c'est vraiment des jolis endroits,
 it's (2.3) it's really a nice place (plural)
 chr ->+, , , , , , , +

Excerpt 2 (C_W_0729)

When Sandra elaborates on the proximity of the proposed area to the city centre (1) they are both walking, looking ahead (Figures 4a/4b). After a lapse of 0.9 seconds (2), Sandra further expands on her positive appreciation (3–4), while Christian looks back (Figures 5a/5b). He seems to orient to the fact that a big bus is approaching from behind (hearable but not yet visible on the video). Their conversation pauses (5) when the bus noisily overtakes them and, during that lapse, a car horn sounds (6). The sound of the horn is not visibly or audibly oriented to by either of them. When Sandra resumes her talk (7), a new horn is clearly audible (8): she does not respond to it, whereas Christian turns to the road for quite a long time.

As in the previous excerpt, sound events happening in the environment are oriented to by one participant but not by the other: unlike Sandra, who is engaged in talking and does not orient to them, Christian, who is listening, turns twice, the first time anticipating the arrival of a noisy bus, and the second time turning back after a car horn sounds. The excerpt confirms that the speaking

participant notices these sounds less than the hearing participant. Moreover, when the bus approaches them, it produces some loud noise – and actually the conversation is suspended at this point (5) – but this event, as well as the horn co-occurring with it, are not bodily oriented to at that moment. This shows how temporality (an anticipated noise is less noticed than a sudden unexpected sound) and contingencies (a single overlapping noise is not noticed as strongly as repeated noise) matter in the definition of sounds as perturbations.

These two first excerpts show that when the participants are engaged in a conversation (vs walking in silence), they might orient differently to noise in the environment (through body displays and directions of gaze). The participants’ asymmetry relies on different forms of conversational engagement at that specific moment: the speaker is less affected than the listener by this noise. In this case, differences are related to the relevance of categories like ‘speaker’ vs ‘listener’, and their actions (Sacks, 1992), rather than to the categories of ‘person with psychosis’ (Benoît, Christian) vs ‘accompanying person’ (Nadia, Sandra).

5.1.2. Commenting on noise

Whereas the previous section focused on embodied orientations to noise, this section focuses on verbal comments. It shows that, when noise is verbally formulated, *explicit comments and explanations* are *initiated* exclusively by the persons with psychosis. In contrast, their partners merely *respond* to the previous comments, and tend to minimize the impact and problematic ascription of noise.

We join the following fragment (Excerpt 3) as Christian and Sandra are walking in silence along a road with traffic. A car horn sounds (2). As soon as they turn into a pedestrian street, Christian comments on it (3).

1	both	(11) [(0.2)#(0.5)] (3.0)ε	(2.2)	ε
		>>walk along a road-----fturn to a pedestr strfcont on it->>		
2	CAR →	[((horn))]		
3	CHR	xx tous les gens i klaxonnennt		
		xx all the people are sounding their horns		
4	SAN	ouai:s, c'est euh (0.6) c'est [l'trafic j'pen]se		
		yeah it's eh (0.6) it's [the traffic I think		
5	CHR	[xxx xx]		
6		(0.5)		
7	SAN	HHeh		
8		(1.3)		
9	SAN	c'est les bus eh		
		it's the bus eh		
10		(6.5)		

Excerpt 3 (C_W_2700)

Christian’s comment is uttered more than 5 seconds after the noise occurred, in a significant location, as they engage in a pedestrian street in which there are no cars. Thus, for him, the horn has a relevance that lasts for a relatively long time; moreover, the quieter atmosphere of the new street might, in contrast, invite a comment about it. The comment is a generalization about the use of the horn (referring to ‘all the people’, 3), which has also a normative character of complaint and blame. Nadia’s response formally aligns with an initial *ouai:s/yeah*’ (4), but then modifies Christian’s moral attribution of responsibility into a more contingent explanation (‘it’s the traffic’, 4), further expanded, after some laughter (7), into the mention of the

buses (9). Whereas the comment by the person with psychosis is a generalization about how people behave, the accompanying person uses an impersonal formulation and minimizes the problem in a fragmented and hesitant turn.

Something similar happens in the next extract (Excerpt 4):

```

1  RAM      (c'est sympathique)
           (it's pleasant)
           iri      ±smiles->
2  IRI      oua- ou#ais °c'est [(°sy±+mp°)#±±
           yeah yeah it's [(pleas-)
3  CAR →    [(horn)
           iri      +gazes to road+
           ->±,,,,,±irritated face->>
           fig      #fig.6          #fig.7

```

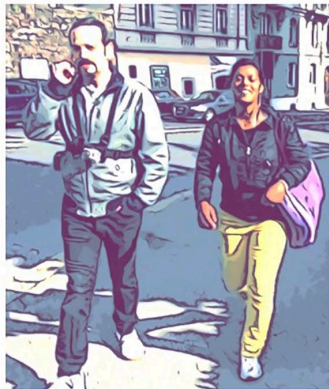


fig.6



fig.7

```

4          (0.1) [(0.25)] (0.3)
5  CAR →  [(horn))]
6  IRI      HE
7          (0.2) + (0.2) *(0.4) # (0.3) ** (1.0)
           iri      ->+upper body turns to road+
           ram      *greet-----*
           fig      #fig.8

```

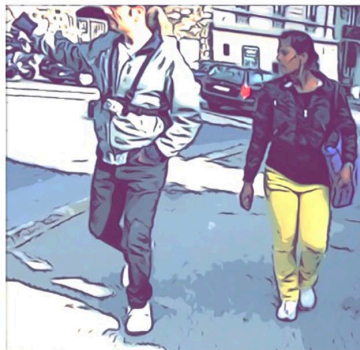


fig.8

```

8  RAM      (peut-être) des amis? he he h
           (maybe) some friends he he h
           (0.2)±(0.5)
           iri      ->±gazes at RAM->>
10 IRI      j'connais pas.
           I don't recognize

```

Excerpt 4 (I-W_2758)

Ramon and Irina are walking along a road with traffic while jokingly talking (1–2). Irina smiles and closes her eyes (Figure 6), fully engaging in and enjoying the conversation. When a car horn sounds twice (3–5), she immediately gazes at the road and her facial expression changes drastically, displaying irritation and anger (Figure 7). She also makes a negative assessment (*HE*, 6) and then bodily turns to the road. This negative reaction to the event contrasts with that of Ramon: he also turns to the road but greets the car by raising his right hand (Figure 8). He offers an account for that to Irina (8), still joking. But she responds curtly (10), thereby not only closing the sequence but also manifesting a disagreeing stance.

In Excerpts 3–4, the persons with psychosis (Christian, Irina) initiate blaming comments and display negative stances towards car horns. In contrast, their partners respond to them by minimizing the event: in Extract 3 by offering a situated and contingent counter-explanation; in Extract 4 by adopting an opposite stance, supposing a friendly reason for the car's horn. The patient and the accompanying person manifest diverging interpretations of the event and the perturbing character of the noise.

Although patients do not always orient to noise (for instance, when they are actively engaged in the conversation – see the previous section), when they do so with explicit comments, they assume a very distinctive posture.

5.1.3. Being affected by noise

Participants might orient to sounds as perturbing noise not only by bodily reacting to them (§5.1.1) or explicitly commenting on a noisy event (§5.1.2), but also by showing a change in their verbal and embodied conduct before and after the noise.

In the next fragment (Excerpt 5), Benoît and Nadia are walking along a sidewalk. Nadia recognizes, on the right side of the sidewalk, the building and signs of a company she works with and she starts to talk about this company. The conversation unfolds in an ordinary way, initiated by Nadia's noticing, and collaboratively developed through Benoît's questions and displays of interest (this part of the transcript is omitted here). Meanwhile, the participants not only continue to walk together, but they also alternatively look together at details of the environment and at each other as they talk.

The verbal and embodied conduct of Benoît is observably altered a few seconds after, when a series of audible noises occurs. At this precise moment, the participants are walking past a big truck (Figure 9), which not only spatially restricts the way, creating a narrow corridor, but also possibly amplifies the effects of noise. The first two noises occur while Nadia is mentioning some presents the company offered to its customers for Christmas:

1 NAD #[mai:s tu sais le coffret que j'avais [+#reçu] là?
 [but you know the small box I [received there?
 2 EVE → [((dull noise)) [((metallic noise))]
 ben >>looks in front-----+looks R-->
 fig #fig.9 #fig.10



fig.9

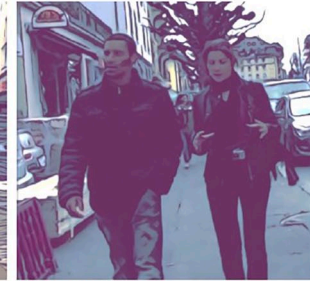


fig.10

3 (0.8)
 4 NAD le coffret en +boi:s avec la \$bouteille *et pui:s* tous
 the small wooden box with the bottle and then all
 ben -->+looks in front-->
 ben *nods---*
 nad -->\$looks at BEN-->
 5 les [xx xx] pour euh\$:: +#(0.6)&
 the [xx xx] for eh:: (0.6)
 6 EVE → [((loud +#metallic noise))]
 ben -->+looks down-----+looks L-->
 nad -->\$looks down-->
 fig #fig.11 #fig.12



fig.11



fig.12

7 NAD &les bouteilles.
 the bottles
 (0.2)
 8 NAD [\$c'est eux qui m'ont offert \$ça
 [it's them who gave this to me
 nad >\$looks at BEN-----\$looks in front-->
 10 BEN [.tsk +MAIS OUI:::
 [.tsk SURE
 ben -->+,,,,,,+looks in front-->

Excerpt 5a (B_W_013756)

The first metallic noise is quite short (2) and seems to capture Benoît's attention (who looks to his right, Figure 10). Nadia continues her turn by representing with her hands, in an iconic gesture, the 'coffret' (small box) she is talking about: while she increments her turn, by adding more details (4), she gazes at

Benoît, who is now looking in front and responds with a minimal nod. Another very loud metallic noise occurs (6) that overlaps with Nadia's turn: while she keeps looking at Benoît, he first looks down (Figure 11), and then to his left (Figure 12). These gaze reorientations are sequentially located during and after the noise occurs. In contrast, the noise does not affect Nadia's turn who, despite a short hesitation/perturbation, continues to talk and completes her turn with a final falling intonation (7). Following a short pause (8), she increments it again with a new development (9). This shows how Nadia is managing the absence of immediate uptake by Benoît. He only responds later (10), in a quite delayed and minimal way; moreover, his turn is characterized by an irritated prosody (high volume and stretched syllables). He does not gaze at Nadia any more but focuses instead on the street in front of him.

His posture becomes even more disengaged in what follows:

```

11          (1)
12 NAD      c'était joli ça
            that was nice
13          (2.3)
14 BEN      ouais
            yeah
15          (1)
16 NAD      °°xxx xxx°
17          (1.6) [ (0.4)+(0.4)](0.5)+(3)
18 CAR →   [((horn)) ]
            ben      -->+looks down+looks in front->
19          [(0.6)+#*(0.2)#(1.6)+(0.4)* + (1.8)] #
20 EVE →   [((strident noise))]
            ben      -->+looks R-----+looks in fr+looks down->
            ben      *grits mouth-----*
            fig      #fig13 #fig.14 #fig15

```

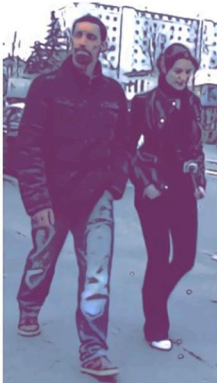


fig.13

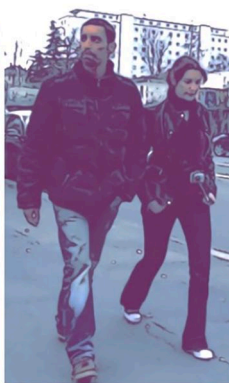


fig.14

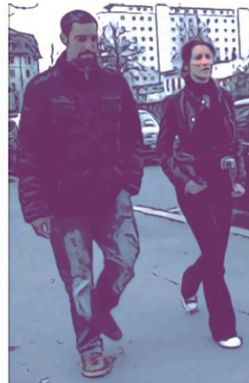


fig.15

```

21          (20.2)
22 NAD      tu fais quoi cet après-midi?
            what do you do this afternoon?

```

Excerpt 5b

Keeping the focus on the same conversational topic, after a long silence (11), Nadia produces an assessment (12). Assessments project the relevance of an answer/acknowledgement by the interlocutor: Benoît aligns to this projection, but only after a pause and with a minimal *ouais* (14). The temporality of his

turns and his minimal contributions show a quite different engagement in the conversation, as compared to what he was doing just before this noise occurred.

The conversation fades out and a very long silence follows (29 seconds: the longest silence in the entire walk-along, 17–21). During this conversational gap, other noise occurs: first, a car horn (18), following which Benoît looks down (this shows again a recurrent sequential pattern, in which gazing down follows a noise); then a very strident and long noise (like a screech of tyres, 20). While Nadia does not seem to change her conduct, Benoît looks to his left and squeezes his mouth (Figures 13–14), progressively walking away from his partner (Figure 15). During the following pause (21), Benoît looks to the front again and then from right to left in a circumspect way. This long silence is interrupted by Nadia who finally introduces a new topic (22), formatted as a question (which prompts her interlocutor to respond).

This excerpt shows how participants might be affected by surrounding noise, even when they do not explicitly turn to or comment on it. The comparison between Benoît's (patient and hearer) conversational conduct before and after a noisy event – especially a loud and strident one – reveals dramatic changes in his capacity to participate in the conversation. These changes are observable in the form of minimal replies, delayed contributions and absence of topic initiation or collaboration. Moreover, they are manifested by a particular solitary gait (distinct from a coordinated walk with the partner), gaze aversion and gaze redirection, browsing the environment and orienting towards the source of the noise. Even if he does not make any comment about the noise, the fine-grained sequential analysis of his conduct reveals important perturbations, which should be taken into consideration to identify possible behavioural effects of urban audible sources of stress.

To sum up, video data make accessible not only the sounds characterizing the city environment, but first and foremost the participants' embodied and verbal conduct in reaction to them, making them relevant and meaningful (or not). This allows the identification of complex, context-dependent features that shape the treatment of the city soundscape among urban dwellers. Moreover, this confirms that persons living with a diagnosis of psychosis are hypersensitive to urban noise, and enables possible differentiation between persons with psychosis and other city users.

5.2. Talking about noise in video elicitations

Persons with psychosis often talk about noise as being stressful. However, as shown above, urban noise is treated as a source of perturbation only in certain circumstances and not in others. So, how can video-elicitations contribute to the study of noise in psychosis? In this section, we turn to video-elicited excerpts, in which patients explicitly talk about noise as harmful, while they are shown the video of their previous stroll, and they are invited to comment on them. As we demonstrate, watching the video in this way is itself a

situated activity that not only generates opinions and comments but can also be submitted to detailed video-based analysis of the interactional contingencies in which comments are produced.

The patient and three researchers are watching the video on a laptop (see Figures 16a/b of Excerpt 6, here below). The video they see comes from the camera following the walkers, making available their prospective trajectory and the ambient sound around them (the conversation is not audible). In this excerpt, the patient comments on what happened in Excerpt 1. We showed above that he was talking during the sound of a horn; and that he did *not* visibly, audibly, verbally or interactionally orient to it. In contrast, in the video elicitation, he briefly acknowledges that noise:

1
2 CAR (23.0)
3 ((horn))
4 BEN (0.5)
 ah,#
 oh
 head up
 fig #fig.16a/b



fig.16a



fig.16b

5 *(1.6)*
 ben *several slow nods->
6 BEN ouais.*
 yeah
->*
7 (5.0)

Excerpt 6 (B_Int_0534/1'05)

After 23 seconds during which Benoît has been motionless while silently viewing the video, the sound of the horn occurs (2). Upon hearing the horn, he utters a simple *oh* and begins to move his head, nodding, before saying *yeah*. In this way, the horn is addressed, treated as remarkable, in contrast with the previous moments on the video. This orientation demonstrates recognition of the noise, treated not as something new or curious, but something familiar (and possibly shared with the researchers, since it is not explicitly commented upon). It does not refer to any recollection or feeling concerning the past event in which the noise is audible. This reflects the specific status of the audible detail picked up by the patient: its salience is enhanced by the fact that the sound of the video he watches is *not* the one documenting the pair's activity during the walk (including their conversation), but the one from the camera following them (favouring ambient sound and offering a view from

behind, showing the local environment and the walking pair seen from the back at some distance). This enhances the chances of video elicitation generating comments about surrounding noise, rather than about the activity of the participants.

In contrast with the previous excerpt, in the following one (Excerpt 7) the patient does not visibly orient to the noise he reacted to during his walk (the event watched in Line 1 corresponds to Excerpt 2):

```

1      (37.0) %(0.8) [ (0.2) ] (2.7)% (0.5) [(0.3)] (20.0)
      bus      %overtakes-----%
2      CAR      [((horn))] [((horn))]
3      [ (0.2 )] (11.0) * (3.5) *
4      CAR      [((horn))]
      chr      *rubs eyes*
5      RES      on s'arrête p't'êt juste là? y a on a y a xxx?
      maybe we stop just there? there're we have there're xxx?
6      REQ      ouais
      yeah
7      (1.0)
8      RES      des choses qui vous passent par la +tête juste+ sur
      any things that come to your mind right about
      chr      +shakes head+
9      cette [séquence là?
      this [sequence there?
10     CHR      [no:n rien du tout.
      [no nothing at all
11     CHR      non i >'s'pas(se rien<)
      no there's >'nothing happening<
12     (5.0)
13     CHR      no:n.
      no:
14     (1.2)
15     CHR      à part euh:: qu'y a des gens qui klaxonnent un peu
      except ehm:: that there're some people sounding their horns a bit
16     parce qu'ils ont vu la caméra (0.3) °c'est tout°.
      cause they've seen the camera (0.3) °that's all°
17     (1.3)
18     CHR      j'crois hein (0.3) après peut-être qu'c'était pas
      I think right (0.3) then maybe it was not
19     pour les caméras
      for the cameras
20     (0.3)
20     RES      ouais.
      yeah

```

Excerpt 7 (C_int_3214)

While watching the video, Christian does not react in any way: his body and head are immobile and his face impassive. After the segment corresponding to Excerpt 2 (1), there is one more car horn sound (2). But this is not addressed by Christian either, who remains silent for another 11 seconds. The only change in his posture is after this pause, when he rubs his eyes – an action that is not directly responding to what happened just before and that could merely display tiredness.

This is seized on by the researcher as an opportunity to pause the video and ask a question about what could have been noticed in the previous fragment (8–9). The first immediate responses of Christian, in overlap, are negative (8, 10, 11, 13). In the absence of an uptake by the researcher, he finally mentions an

exception (15), referring to the horns and offering a possible explanation for them (15–16). The mention of the horn is thus triggered by the researcher’s question and mentioned in response to it. Moreover, its explanation is not related to the participant but to the camera filming the stroll. In this way, the patient’s claims – expressed in a tentative and hesitant way – are responsive to the researcher’s actions. This indicates how careful the interpretation of the elicitation has to be: the fact that the participant mentions (or not) something while watching the video is not per se a retrospective comment about the event the video is documenting, e.g. about what had been experienced or felt in the past. The video works as a trigger – for both the participant *and* the interviewer – offering audible and visible hints for comments that are general statements, rather than recollections referring to the particulars of the past event.

Although the observations above issue a caveat about the use of video-elicitations, which have to be re-cast in their context in order to be interpretable, the discrepancies observed in the previous extracts between what happened in the past and what is pointed to in the video elicitation can be meaningful, as shown in the following and last case, referring to Excerpt 5. We are at the very beginning of the elicitation session: before Benoît starts watching the video, the researcher asks if there is anything particular he would like to focus on (Excerpt 8). In his answer, Benoît refers to an *unpleasant sensory experience* he recalls about his walk-along, when a noise was produced by a ‘pallet’ falling to the ground:

- 1 RES peut-être la la première chose c’est est-ce que vous y a
maybe the the first thing is do you are there
- 2 des séquences que qui vous ont >vous avez< remarqué des
sequences that that have >you have< noticed some
- 3 choses vous aimeriez prendre des séquences, si vous r-
things that you would like to take some sequences, if you r-
- 4 (0.5)
- 5 BEN [euh::
[ehm::
- 6 RES [rappelez du du parcours?
[remember the the walk
- 7 BEN y a: oui: y a à un moment donné ouais où euh: (0.4)
there is: yes: there is at some point yeah where eh: (0.4)
- 8 .tsk (0.3) où: ‘je vais° vous expliquer un petit truc
.tsk (0.3) where °I’m gonna° explain you a little thing
- 9 .h:: (0.3) euh du parcours oui, .h: en revenant, (0.3)
.h:: (0.3) ehm about the walk yes, .h when coming back (0.3)
- 10 → euh du parcours euh: ‘l y a une: palette qui est tombée
ehm from the walk eh: there’s a pallet that fell down
- 11 (0.6)
- 12 RES [oké]
[okay]
- 13 BEN → [et ça] ça a fait un énorme écho
[and this] this made a huge echo

Excerpt 8 (B_int_170415_entretien_cam1)

In answering the general question, Benoît singles out a precise moment he remembers. He does not describe the noise itself but refers to the event causing it (‘a pallet that fell down’, 10) and its audibility, expressed in a hyperbolic way (‘this made a huge echo’, 13), hinting at its unpleasant character.

Despite the researcher proposing to immediately watch that portion of the video, Benoît prefers to start from the beginning of the entire promenade, so that this topic is momentarily abandoned. After approximately 50 minutes, however, the researcher goes back to this topic and Benoît names precisely the place where this noise happened. This occasions the watching of the related video segment, corresponding to Excerpt 5. We join the action (Excerpt 9a) as the group looks silently at the video for 38 seconds (1). At this point, Benoît produces a summons (*attention*, 2), announcing a forthcoming event. This alert is given when the pair on the video is approaching a truck (Figures 17a/b), in the back of which a pallet is visible.

1 (38.0)
 2 PAT #attention
 watch out
 fig #fig.17a/b

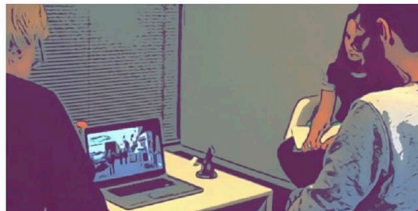


fig.17a



fig.17b

3 (4.4) [(0.5) (1.1) [(0.5) (3.1) [(0.8)
 4 EVE → [((dull sd)) [((metallic sd)) [((strident sd))
 5 RES +ah ouais+
 oh yeah
 pat +moves head back+
 6 (0.2)
 7 PAT là
 there
 8 (0.4)
 9 RES ça
 this
 10 *(0.6)*[(0.3) (0.7) [(0.6) [(0.5)
 11 EVE → [((dull sd)) [((metal sd)) [((people's voices))
 res *nods*
 12 PAT ça aussi
 this too
 13 [(0.9)
 14 EVE → [((metallic sd))
 15 PAT et encore
 and again
 16 [(3.7) [(0.7)
 17 EVE → [((low people's voices)) [((horn))
 18 PAT .tsk +ah+
 .tsk oh
 +raises eyebrows+
 19 +(1)+ *(0.6)
 pat +raises left shoulder+
 res *raises hand-->
 fig #fig.18
 20 RES xxx des bruits, des bruits *ai[ghus hein?*
 xxx some noises, some noises hi[gh-pitched right?
 -----*/////////*
 21 EVE → [((strident sd))
 22 +(0.2)
 pat +...->
 23 PAT #+ouais+
 yeah
 pat ->+double hand gesture+
 fig #fig.19
 24 (+0.4+ 3.6)
 pat +,,,+



fig.18



fig.19

Excerpt 9a (B_int_5058)

After Benoît's summons, several sounds are audible on the video (4): first a dull sound, then metallic, and finally loud and strident. He does not react in any particular way to these sounds. The researcher, instead, displays recognition ('oh yeah', 5) of the last strident sound as the noise Benoît was previously referring to. Benoît briefly aligns with the spatial adverb (7), which is confirmed by the researcher with a demonstrative pronoun (9), both using indexical expressions. At this point, a shared understanding seems to be reached, expressed by minimal indexical resources. Several other types of noise follow (11, 14), acknowledged by the researcher's nods and some comments by Benoît ('this too', 12, 'and again', 15). This shows the incremental nature of the interpretation of the sounds and their progressive identification as belonging to the same category of noise.

Next, a car sounds its horn (17): Benoît reacts with *ah* (as in Excerpt 6) and by raising his eyebrows and left shoulder. While the previous reactions to sounds were highly indexical and non-representational (20), here the researcher provides a summarizing description of what has been audible until then: he offers a category ('noise'), which treats the sounds as a nuisance, and a descriptor ('high-pitched'), which indicates the specific quality of these sounds. The descriptor co-occurs with a gesture (Figure 18). Benoît acknowledges with a simple 'yeah' (23) while making a very similar gesture, using both hands (Figure 19). In this way, the categorization is aligned both verbally (although Benoît does not actually repeat or use a description) and gesturally (he does not merely imitate but enhances and modifies the initial gesture). Again, a shared understanding seems to be reached.

However, after these multiple alignments, Benoît claims not to have heard the sound of the pallet he was referring to at the beginning and that he was expecting to hear in that environment (Excerpt 9b):

25 PAT j'ai pas entendu une palette qui tombe
I have not heard a pallet that falls down
26 (1.9)
27 RES mais peut-être que on l'a moins entendue parce que::
but maybe we have heard it not as much because
28 euh:: on on a le son de derrière pas le son devant
eh:: we we have the sound from behind not the sound from ahead
29 (1.2)
30 PAT peut-être que vous l'avez carrément pas et
maybe you haven't (heard it) at all and
31 y a que moi qui l'a entendu
it's only me who has heard it

Excerpt 9b (B_int_5058)

This generates a puzzle: this noise was well remembered – independently of the video-elicitation and prior to it – and was foreshadowed and expected (by *attention*, 2, projecting it). A series of sounds have been recognized on the video by both participants. But while the researcher was referring to (at least one of) these as including the ‘pallet’ noise, Benoît now says he has *not* heard it. We can notice that the sounds audible on the recording (and transcribed above) are heterogeneous – the researcher’s categorization identifies *strident*, louder noise as relevant, but ignores *dull* sounds, which are lower. We can hypothesize that the noise of the pallet is in fact the dull sound featuring just after Benoît’s *attention* (2), and possibly not heard by the participants watching the video (they produce their comments only after the occurrence of other sounds).

At this point, the participants retrospectively discover their discrepant perceptions. The researcher addresses them by referring to the audio quality of the recording (27–28 in Excerpt 9b): in ‘we have heard it not as much’, he uses the (French) pronoun *on* that encompasses both of them and has a generic value, and treats the sound audibility as variable, thereby accounting for their discrepant perceptions. In contrast, Benoît suggests another explanation, pointing out a radical asymmetry between them, positioning himself as a specific and maybe unique perceptive subject, able to detect sounds that others cannot hear. He retrospectively makes clear that his aligned responses were targeting the noise heard in common, but not the identification of the pallet’s noise. He also contests the researcher’s claim to have heard it. Absence of hearing is here treated as evidence of the ability to hear the sound during the walk-along and of a radical difference between the participant and the researcher. In this way, the participant constructs a specific identity, in contrast to that of the researcher, claiming to have specific aural abilities, possibly related to his illness, and assuming it in a positive, rather than stigmatized, manner.

This last fragment highlights both methodological and substantive issues. It shows that video elicitation requires a careful interactional sequential analysis and cannot be reduced to registering contents and opinions. It also confirms the *situated variability of hearing perception*,

both in the actual walk and in the elicitation session: the indexicality of hearing and listening relies both on the ecology of perception (the material and spatial environment of the walks, the sound quality of the recordings and the audio players used in the elicitation session) and on the context of interaction (during talk vs during silence). Variations refine our understanding of the specific and systematic conditions in which an urban sound is heard or not, categorized or not as a noise, heard as displaying a particular sensory sensibility or as a shared perception of the urban environment. Video analysis enables precise pinpointing of these relevant details and their indexicality.

6. CONCLUSION

This article has offered a multimodal conversation analytic approach to aural sensoriality in noisy urban environments. It provides a better understanding of how sounds are oriented to, integrated in verbal and embodied conducts and possibly commented on in social interaction. Moreover, by focusing on how atypical populations, such as people with psychosis, orient to urban noise, the article specifies some claims in the literature about their sensory hypersensitivity.

Video-analyses locate sound events within specific ecologies, in specific moments of an activity, and within specific participation frameworks. First, sounds are analysed in relation to sensory practices of hearing, listening, looking and seeing in the environment, situated within interactional moments: this allows us to define the orientation to sounds as not purely individual, but as socially shaped and socially shaping. Second, orientations to sounds can take a variety of forms, such as embodied conducts (gaze aversion, absence of responses, facial expressions, signs of irritation, etc.), response cries and evaluations, elaborated comments and formulations, which constitute indicators of how situated sounds are oriented to and categorized (e.g. as noise). Third, the findings highlight how sensitivity to sounds considered as noise depends on systematic differential features: (a) independently of their health condition, participants are more sensitive to noise when they are listening than when they are speaking to their partner; (b) when noise is explicitly addressed and commented on, persons with psychosis elaborate more often on them and with distinct and more detailed arguments than their partners; (c) the tendency for patients in video-elicitations to provide generic summarizing statements highlights their development of specific interpretative capabilities and self-diagnostic discourses, showing themselves to be experts of their own condition. This complements experimental research concerning sensory gating (Collip et al., 2008; Micoulaud-Franchi et al., 2012), by revealing systematic patterns and variations observable in naturalistic (vs laboratory) settings.

The analysis of video-recorded activities and the analysis of filmed video-elicitations generate different findings. The analysis of video-recorded

walk-alongs reveals embodied and affective features of participants' urban practices, in which they orient and react to sounds, although not always speaking about them. Video-elicitations focus on participants' interpretations of themselves in urban situations, leading to more generic statements about how persons living with psychosis experience noise, abstracted from the local specificities documented by videos of situated urban activities. The combined use of these video data shows their tension, that is, a tension between elusive situated embodied conducts and explicit generalizing formulations. The submission of these data to careful analytic scrutiny also shows how problematic it is to extract possible statements about personal experiences from the interactional and sequential context in which they were produced. In other words, actions and statements make sense in relation to previous events during the walk, to previous turns in the conversation, or to previous interviewers' questions during the elicitation: documenting through video-recordings the temporal and sequential development of these different activities makes it possible to analyse (a)typical conducts as they emerge within conversational and embodied practices.

More broadly, taking several interactional environments into consideration enables a systematic characterization of the context-dependent variability of the participants' sensoriality, that is, an understanding of how sound sensitivity is modulated according to the local ecology and the sequential environment within social interaction, and how it is expressed in embodied conducts and/or in noticings and self-reportings. This contributes to a video-based investigation both of the sound dimension of human urban sociality and of psychosis in situated experiences.

FUNDING

This research was funded by the Swiss National Science Foundation (grant number: CR1311_153320).

RESEARCH ETHICS AND PATIENT CONSENT

All participants in the study signed an Informed Consent Form. The entire procedure was approved by a Medical Ethics Committee.

TRANSCRIPTION CONVENTIONS

Transcription of talk follows the conventions established by Jefferson (2004). Multimodal transcripts have been established following conventions established by Mondada (2018).

ORCID ID

Sara Merlino  <https://orcid.org/0000-0002-3335-0328>

NOTE

1. The names used in this article are all pseudonyms.

REFERENCES

- Atkinson R (2007) Ecology of sound: The sonic order of urban space. *Urban Studies* 44(10): 1905–1917.
- Bergmann J (1990) On the local sensitivity of conversation. In: Markova I and Foppa K (eds) *The Dynamics of Dialogue*. Hertfordshire: Harvester, 201–226.
- Collip D, Myin-Germeys I and Van Os J (2008) Does the concept of ‘sensitization’ provide a plausible mechanism for the putative link between the environment and schizophrenia? *Schizophrenia Bulletin* 34(2): 220–225.
- Conus P et al. (2019) City avoidance in the early phase of psychosis: A neglected domain of assessment and a potential target for recovery strategies. *Frontiers in Psychiatry* 10: 342.
- Feld S (2015) Acoustemology. In: Novak D and Sakakeeny M (eds) *Keywords in Sound*. Durham, NC: Duke University Press, 12–21.
- Goodwin C (2017) *Co-operative Action*. Cambridge: Cambridge University Press.
- Haddington P, Mondada L and Nevile M (eds) (2013) *Mobility Interaction*. Berlin: De Gruyter.
- Heath C (1986) *Body Movement and Speech in Medical Interaction*. Cambridge: Cambridge University Press.
- Heath C, Hindmarsh J and Luff P (2010) *Video in Qualitative Research*. London: Sage.
- Heinemann T and Rauniomaa M (2016) Turning down sound to turn to talk: Muting and muffling auditory objects as a resource for displaying involvement. *Gesprächsforschung* 17: 1–28.
- Henry SG and Fetters MD (2012) Video elicitation interviews: A qualitative research method for investigating physician–patient interactions. *Annals of Family Medicine* 10(2): 118–125.
- Jefferson G (2004) Glossary of transcript symbols with an introduction. In: Lerner GH (ed.) *Conversation Analysis: Studies from the First Generation*. Amsterdam: Benjamins, 13–31.
- Keevallik L and Ogden R (2020) Sounds on the margins of language at the heart of interaction. *Research on Language and Social Interaction* 53(1): 1–18.
- Landon J et al. (2016) When it’s quiet, it’s nice: Noise sensitivity in schizophrenia. *American Journal of Psychiatric Rehabilitation* 19(2): 122–135.
- Merlino S and Mondada L (2019) Crossing the street: How pedestrians interact with cars. *Language & Communication* 65: 131–147.

- Micoulaud-Franchi J-A et al. (2012) Toward an exploration of feeling of strangeness in schizophrenia: Perspectives on acousmatic and everyday listening. *Journal of Abnormal Psychology* 121(3): 628–640.
- Mondada L (2012) The conversation analytic approach to data collection. In: Sidnell J and Stivers T (eds) *The Handbook of Conversation Analysis*. Malden: Wiley-Blackwell, 32–56.
- Mondada L (2016) Challenges of multimodality: Language and the body in social interaction. *Journal of Sociolinguistics* 20(3): 336–366.
- Mondada L (2018) Multiple temporalities of language and body in interaction: Challenges for transcribing multimodality. *Research on Language and Social Interaction* 51(1): 85–106.
- Mondada L (2019) Contemporary issues in conversation analysis: Embodiment and materiality, multimodality and multisensoriality in social interaction. *Journal of Pragmatics* 145: 47– 62.
- Pinch T and Bijsterveld K (eds) (2012) *The Oxford Handbook of Sound Studies*. Oxford: Oxford University Press.
- Pomerantz A (2005) Using participants' video stimulated comments to complement analyses of interactional practices. In: te Molder H and Potter J (eds) *Talk and Cognition: Discourse, Mind and Social Interaction*. Cambridge: Cambridge University Press, 93–113.
- Rauniomaa M and Heinemann T (2014) Organising the soundscape: Participants' orientation to impending sound when turning on auditory objects in interaction. In: Nevile M et al. (eds) *Interacting with Objects: Language, Materiality, and Social Activity*. Amsterdam: Benjamins, 145–168.
- Sacks H (1992) *Lectures on Conversation*. Oxford: Blackwell.
- Schafer RM (1977) *The Soundscape: Our Sonic Environment and the Tuning of the World*. Rochester: Destiny Books.
- Smucny J et al. (2013) Early sensory processing deficits predict sensitivity to distraction in schizophrenia. *Schizophrenia Research* 147(1): 196–200.
- Söderström O (2019) Precarious encounters with urban life: The city/psychosis nexus beyond epidemiology and social constructivism. *Geoforum* 101: 80–89.
- Söderström O et al. (2016) Unpacking 'the city': An experience-based approach to the role of urban living in psychosis. *Health & Place* 42: 104–110.
- Söderström O et al. (2017) Emplacing recovery: How persons diagnosed with psychosis handle stress in cities. *Psychosis* 9: 322–329.
- Sterne J (ed.) (2012) *The Sound Studies Reader*. London: Routledge.
- Streeck J, Goodwin C and LeBaron C (eds) (2011) *Embodied Interaction, Language and Body in the Material World*. Cambridge: Cambridge University Press.
- Thibaud J-P (2003) The sonic composition of the city. In: Bull M and Back L (eds) *The Auditory Culture Reader*. London: Berg, 329–342.

- Vassos E et al. (2012) Meta-analysis of the association of urbanicity with schizophrenia. *Schizophrenia Bulletin* 38(6): 1118–1123.
- Wright B et al. (2016) Effects of environmental noise on cognitive (dys) functions in schizophrenia: A pilot within-subjects experimental study. *Schizophrenia Research* 173(1): 101–108.

BIOGRAPHICAL NOTES

SARA MERLINO is Assistant Professor of Linguistics at the University of Roma Tre. She investigates social interaction using conversation analysis and a video-based multimodal approach. Her current research focuses on interactions with people diagnosed with language disorders, such as aphasia. She is particularly interested in the communicative dynamics of speech–language therapy encounters and in the embodied and multimodal dimension of the therapeutic process.

Address: Dipartimento di Studi Umanistici, Università degli Studi Roma Tre, Via Ostiense 234/236, Rome 00146, Italy. [email: sara.merlino@uniroma3.it]

LORENZA MONDADA is a Professor of Linguistics at the University of Basel. Her research deals with social interaction in ordinary, professional and institutional settings, within an ethnomethodological and conversation analytic perspective (EMCA). Her focus is on video analysis and multimodality, integrating language and embodiment in the study of human action. Currently, she works on how interactants engage not only in coordinating their joint actions in publicly accountable manners but also in sensing the material world together – within an EMCA perspective on sensoriality in interaction.

Address: Department of Linguistics, University of Basel, Maiengasse 51, Basel 4056, Switzerland.
[email: lorenza.mondada@unibas.ch]

OLA SÖDERSTRÖM is a Professor of Social and Cultural Geography at the University of Neuchâtel. He has published widely on social and cultural dimensions of urban change and more specifically on visuality in urban planning, gentrification and urban globalization. He has worked on the social construction of heritage, the role of the visual in urban planning, the geography of architecture, processes of urban globalization, smart urbanism and urban geographies of mental health.

Address: Institute of Geography, University of Neuchâtel, Espace Tilo-Frey 1, Neuchâtel 2000, Switzerland. [email: ola.soderstrom@unine.ch]