University of Bath

Alternative formats

If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

Collaborative Music-Making with Digital Audio Workstations: The 'nth member' as a Heuristic Device for Understanding the Role of Technologies in Audio Composition

Phillip Brooker & Wes Sharrock

Abstract

This article examines amateur music-making using a digital audio workstation (DAW), showing how audio and software are used as resources for creating compositions. The article has two aims. Firstly, to depict how digital music-making is formed from routine interactional techniques. Secondly, to probe how researchers might account for such multi-modal activity through a heuristic device: the 'nth member'. Whereas sociology has typically been concerned with the cultural facets of how music is made and consumed, we explore the material practices of collaborative song creation utilising conversation analytic techniques – 'turn-taking' and 'next-selection' – to capture two key interactional moments.

Keywords

Music; software; video; ethnomethodology; conversation analysis

This study reflects a growing interest in the application of ethnomethodology to social activities occurring outside of the workplace (cf. Rouncefield and Tolmie 2013; Schmidt 2011). We pay special regard to one such activity – amateur digital music-making. Although leisure activities are a relatively recent interest to sociology, existing studies agree that as much artful work goes into leisure as into professional occupations.[1] Using video recordings of amateur music-making with a digital audio workstation (DAW), this study focuses on the practices of two software-users composing songs. Our treatment of the video recordings draws on ethnomethodology and conversation analysis (Garfinkel 1967; Sacks 1995), and attends to the visible and audible interactions occurring between people – talk, body orientation, pointing, gazes, etc – and how those interactions organise the activity. The nature of music-making with DAWs ensures that members' interactions become thoroughly intertwined with the sonic properties of the audio being produced, as well as with what is visible on-screen and between participants. This case is, therefore, one where members' interactions deal with (a) ordinary verbal and physical interactional cues, and (b) technical audio and computational features of the workstation. One focus of this paper is on the ways ordinary and technical resources become enmeshed in digital music-making.

Given our difficulties with characterising the role of the material 'objects' – the DAW software and the music it produces – around which digital music-making is organised, a second focus is on how sociology might account for activities which are not achieved through talk-in-interaction alone. To this end, we reflect members' utilisations of multi-modal resources, asking: how can researchers render the materiality of members' usages of the software and audio? We attempt this with a transcription and analytic device: the 'n[th] member'.

This 'nth member' operates as an intentional mis-writing and/or mis-reading of the transcribed data, wherein we – the researchers – playfully consider the music and/or software *as if it were* a member whose outputs can legitimately be placed within members' exchanges. This is strictly a heuristic device in the spirit of Sormani's (2014) mis-readings of videoed moments of scientific discovery work as instructional tutorials for him to copy. However, rather than mis-read our transcripts for instructional purposes as Sormani does, we mis-*write* them in the first place. This leads to a subsequent mis-reading akin to a Garfinkelian (1967) 'breaching experiment', designed to provoke thinking about an interactional order (i.e. our transcripts and analyses mis-represent the episodes they report *as if* the attendant audio and software were agency-bearing members). Since members patently do not share this conception, this encourages viewing the interactional order in such a way as to display the relevance of audio and software to members. We openly acknowledge that 'nth members' *are not really* members equable with our DAW users. In fact, we would distance our approach from those that do go as far as suggesting that inanimate 'objects' themselves demonstrate agency in interactions (e.g. Actor-Network Theory (ANT) (Latour 2005) and its use in the 'New Sociology of Art' (de la Fuente 2007)). Rather than cast music and software in the same terms as humans the 'nth member' as a heuristic methodological device allows us to more clearly ascertain how the role of music and software *does* differ from members' contributions. Alongside the substantive empirical findings, we hope to address this issue throughout.

Setting Out Our Tools

The video analysed here depicts the principal author (hereafter referred to in the third person with the initials PB, for ease of reference) and a friend (LC) engaged in collaborative amateur

music-making activities. Their collaboration involves the production of music with a DAW called Reason. Reason features synthesisers, samplers for importing external sound files, sound effects modules, and digital versions of mixing and mastering equipment (figure 1).

[FIGURE 1 HERE]

Figure 1 - The Reason 'rack', where users can see, connect and control instruments, sound effects and mixing/mastering modules.

Synthesisers and samplers can be programmed to emit sounds, which can be shaped and refined with digital sound effects modules and mixing/mastering equipment. DAWs also allow users to sequence (figure 2) sounds into structured arrangements that formulate full songs.

[FIGURE 2 HERE]

Figure 2 - The Reason 'sequencer', where users graphically edit and arrange musical ideas and automate instrument controls.

Musical phrases (and instrument controls such as volume levels and filter frequencies) can be recorded into the sequencer via MIDI piano-keyboard, or programmed in with a mouse and QWERTY keyboard. Having inputted phrases and instrument controls, users can edit them – cut them up, move them around, change the notes played, change the instrument sound, layer them over other instruments, etc. Playback operates through the tracking bar, which moves from left to right playing recorded and programmed phrases as they occur, coordinating the produced sounds with the visual aspects of the sequencer.

4

The present paper explores two usages of Reason – *talking about musical ideas* and *enacting musical ideas* – which feature PB and LC in a dyadic partnership where the capacity to hear and understand the music and software plays a central role in their decision-making.

In regard to *talking about musical ideas*, PB and LC collaborate verbally to decide which ideas to take forward into practical enactment via the software. Here, the 'nth member' is the *audio* playing alongside the users' conversation. Often, the activities captured on camera include periods of listening wherein PB and LC play back musical ideas, using the audio as a conversational resource for discussing what to do next. The issue here is how PB and LC manage their conversational turn-taking (Sacks, Schegloff and Jefferson 1974) against the audio background which at times involves placement of its sounds within the users' turn.

Enacting musical ideas pertains to the work of inputting and refining ideas to achieved desired effects. Here, the 'nth member' is the *software* itself – whilst PB and LC's conversation is mutually collaborative, the members' DAW usage features a computer-setup with only one mouse and one QWERTY keyboard. This necessitates a division of labour whereby only one person at a time can enact ideas. Hence, the issue is how PB and LC go about next-selecting (Sacks, Schegloff and Jefferson 1974) themselves as candidates for enacting ideas in respect to features pertaining to the software; chiefly, who has relevant expertise.

Aside from investigating the activities substantively, by focussing on the 'nth member' as a device to help characterise collaborative music-making activities we demonstrate how human members draw on non-human resources in the production of their work activities. We are under no illusions; our mis-writings and mis-readings of music and software as an 'nth member' does not reflect how they are treated in practice. It is simply a technical solution to

the problem of how to annotate the position of contextually-relevant (yet non-speaking) features of a setting. Hence, we would distance ourselves from an Actor-Network Theory-influenced reading of the scene (see below). Our treatment of 'n[th] members' as agency-bearing members is superficial and heuristic in nature – 'n[th] members' are *not* treated by PB and LC as having an equable agency to human members, though conducting an analysis *as if they were* can help elicit a more realistic understanding of their material relevance in such interactions. This mis-writing and mis-reading of transcripts results in an account of the cases outlined above which is thoroughly attuned to the material aspects of PB and LC's activities, and which explores the role of the *content* of music and software in their ongoing collaborations. Here, our analyses aim to capture how the *sound of the music* facilitates PB and LC's conversational work, and how *features of the software* inform how PB and LC structure their collaborations.

Background

Historically, sociological studies of music have conceptualised music as an aural representation of social life (Adorno 1978; Blackstone 2009; Kotarba *et al*. 2009; Lerdahl and Jackendoff 1996; Merrill 2010), concentrated on the systemic qualities of 'the music industry' (Laing 1985; Peterson and Berger 1971), explored the goings-on of subcultures with a musical affiliation (Becker 1971; Holyfield *et al*. 2013), or combined elements from all of the above (Frith 1978) – see Martin (1995) for an exemplary history of sociological approaches to music. In recent years, these sociological interests have extended towards digital music; for instance, Kenton and O'Hara's (2006) collection examines the role of digital technologies in facilitating new forms of social interaction around music consumption. Yet Becker, Faulkner and Kirshenblatt-Gimblett succinctly point out what is missing from

such accounts: "There has always been a blind spot in the sociology of art: any discussion of specific artworks" (2006: 1). Less succinctly, they note:

> "An analysis that simply invokes class, race, organization, or any of the commonly summoned 'social variables' does not get to the heart of what social science can contribute to understanding art… Art is social not because social variables affect it but because it is the product of collective work" (Becker, Faulkner and Kirshenblatt-Gimblett 2006: 3).

This drive towards capturing the material content of artworks is taken up by the 'New Sociology of Art' (see de la Fuente (2007) for an overview). Here, researchers deploy concepts from Actor-Network Theory (Latour 2005) to suggest that a sociology of music which does not capture the artwork "misses music's active properties and thus diminishes the potential of music sociology by ignoring the question of music's discursive and material powers" (DeNora 2003: 39). Practically, the New Sociology of Art entails the following:

> "Taking explicit cultural objects – artworks, musical pieces, poems, museums – as a starting point, the field analyzes how these objects are mobilized in social networks, institutions and interactions to shed light on the very building blocks of implicit culture" (Acord and DeNora 2008: 224).

This has been applied in empirical studies such as Strandvad's investigation of a Danish filmmaking project, wherein the aim is to "illustrate how the product [a film] becomes an active participant in the process of its development…[and how] the product and the social relations in the production process are being co-produced" (2012: 164). In these ways, the New Sociology of Art is indebted to Actor-Network Theory, incorporating material elements

into analyses by ascribing agency to the artworks around which artists' collaborative efforts are organised.

Though we support the incorporation of the content of artworks into accounts of their production (and do so ourselves here), we are less convinced by the need to do so by subscribing to Actor-Network Theory with all the conceptual baggage it brings. As such, it is worth outlining just where our approach diverges from the wider ANT project. ANT situates itself as a "sociology of translation" (Latour, 2005: 106) of discrete relationships into wider social networks; an approach which necessitates researchers framing peoples' activities in sociological lexicons to make ontological claims. Callon notes:

> "[For ANT] The vocabulary chosen for these descriptions and explanations
> can be left to the discretion of the observer. He cannot simply repeat the
> analysis suggested by the actors he is studying… It is up to the sociologist to
> choose the one [vocabulary] that seems the best adapted to his task and then
> to convince his colleagues that he made the right choice" (Callon 1986: 200).

However, the decision to interpret peoples' activities from a staunchly sociological mindset (and in terms not necessarily shared by the people doing those activities) serves to misconceive how those activities are to be understood. As Lynch notes in regard to ANT's infamous forays into scientific work:

> "This is suggested when everyday terms like *invention*, *inscription*,
> *manufacture*, *machination*, *manipulation*, and *intervention* are theoretically
> preferred over equally familiar idioms like *discovery*, *description*, *observation*,
> *testing*, *proving*, and the like" (Lynch 1993: 266).

More widely, Rawls notes that for mainstream sociology (of which ANT and the New Sociology of Art are a part):

> "The emphasis is placed on conceptual clarification and operationalizing terms, not on understanding what is going on in the situation from the inside out, or on a detailed documentation of the research site" (Rawls 2002: 27).

Hence, because mainstream approaches are conceptual in nature and theorized in advance, the picture of reality they produce belongs "to the researcher and not to the population being studied" (Rawls 2002: 28), and this raises questions around what their findings can be taken to be about. As such, we explore a non-ANT-originating approach to socio-materiality which is less committed to simply finding usages for sociological terms and more attuned to peoples' own understandings of their activities.[2] To do so, we trace a lineage through several ethnomethodologically-affiliated works.

Schutz moves toward the study of the practices of doing music by highlighting the inadequacy of musical notation as a way of keeping musicians in sync. He proposed instead to look at music-making as a "mutual tuning-in relationship" (1976: 161) whereby members of musical groups recognise and situate other members' actions during their playing – e.g. I see the drummer bob his head more vigorously and increase the volume of my own playing, I watch the fiddle player's elbow move more sharply and know to make my own phrasing staccato, etc.

Weeks takes up Schutz' proposal, asking "how it is that musicians, whether in small chamber groups or in jazz combos, get it together and keep it that way" (1996: 200) – how is synchronous playing possible as a practical matter? Weeks concentrates on a selection of

activities – vocalisations, verbalisations, demonstrations – as constitutive of an achieved synchrony:

> "we can conceive of the conductor's instructions, both verbal and sung, as oriented to 'translating' the text [sheet music/score] with these distinct 'indexical' (or context-dependent) expressions into a *joint practical course of action*" (Weeks 1990: 327).

Hence, the guiding of musicians into musical time is not achieved objectively and formally – there *is no possibility* for musicians to share an exact understanding of how many beats-per-minute the word *andante* signifies. Yet there are interactional techniques for curating musical synchrony; e.g., through the deference of instrumentalists to the conductor's gestures. It is within such practices where Weeks finds order in the activity.

Practices like these are explored empirically by Faulkner and Becker (2009) and Sudnow (1978). Faulkner and Becker examine how jazz musicians competently establish and perform a repertoire of songs together, despite rarely having opportunity to meet or plan a set-list before gigs. Examining several collaborative performances, Faulkner and Becker observed that this was always achieved primarily through musicians' mid-performance interactions rather than through formal planning:

> "Like every other kind of activity people undertake together, what jazz musicians do is *neither* random and disjointed nor totally fixed and predictable...performances always mix the two, the terms of the mixture not a simple application of known ways of reaching agreement, but rather an on-the-spot creation" (Faulkner and Becker 2009: 184).

Similarly, Sudnow outlines the process of teaching himself to play jazz piano improvisations – not through fixed formal sequences, but through the practice of 'feeling around' the keyboard until he finds improvisational skills residing in his fingers:

> "I recall playing one day and finding as I set out into a next course of notes... that I was expressly aiming for the sounds of those particular notes...that I had gone to *do* them" (Sudnow 1978: 37).

To Sudnow, this distinction – *doing* the soundedness of a set of notes – is important:

> "it is one thing to recognize familiar sounds you are making and another to be able to aim for particular sounds to happen. A different sort of directionality of purpose and potential for action is involved in each case" (Sudnow 1978: 38).

Here, Sudnow's newly-acquired "melodic intentionality" (1978: 41) hints towards the usage of audio properties of activities and settings as a way of understanding what is going on (e.g. a moment of learning).

Building on these pioneering ethnomethodological works, we aim to work out the requirements for an account of music-making grounded in members' understandings of what it is they are doing, and which captures the material aspects of their activities without overburdening our descriptions with sociological concepts. To this end, we transpose Coulter and Parsons' (1990) and Greiffenhagen's (2014) accounts of the role of visual practices in forming social activities to practices employed in working with audio. Coulter and Parsons note that "we must acknowledge that 'seeing' is akin to an *achievement* and is not any sort of activity, process or 'undertaking'" (1990: 255). In this sense, an analysis of the activities that make up the final 'seeing' – the 'looking', the 'scanning', the 'examining', the 'searching',

the 'witnessing' and so on – becomes pertinent to the characterisation of the activity. The salient questions are: what *kind* of 'seeing' is achieved, and how so?

Greiffenhagen (2014) demonstrates how this praxiological focus might be applied empirically in an investigation into the visual/material aspects of mathematics. Looking at the use of blackboards as pedagogical tools, Greiffenhagen highlights a tutor's strategies for imparting mathematical knowledge through the visual organisation and handling of elements of a proof on a blackboard – the processual nature of board-writing in demonstrating an unfolding proof, the deleting of obsolete components, the designation of different regions for different purposes, and so on. This work shows how thinking in mathematics, typically considered an immaterial process, is "inextricably interwoven with writing mathematics" (2014: 25) as an accountable material practice. We extend Coulter and Parsons' (1990) and Greiffenhagen's (2014) focus on the visual into another sensory arena – audio – by showing how digital music production is achievable for members via a selection of similarly accountable material practices. Where the aforementioned ethnomethodological works (i.e. Faulkner and Becker, 2009; Sudnow, 1978; Weeks, 1990, 1996) deal primarily with the practices of musicians in the course of their playing, we draw on Coulter and Parsons (1990) and Greiffenhagen (2014) to engage more directly with the 'material' properties of the sounds being produced.

The Methods and Methodologies of the Research

The analyses presented here are based on video recordings of PB and LC taken over several non-continuous days. The excerpts we present are from approximately eight hours of videoed interaction, recording multiple sessions working on different compositions. The camera captured PB and LC's interactions as they took place at the computer, also capturing the

screen itself for analysis. The microphone picks up the talk occurring between them as well as other ambient audio, including that which plays as a result of PB and LC's music-making.

The excerpts have been analysed from an ethnomethodological and conversation analytic perspective, which seeks to study the organisation of everyday activities as those activities are understood by those involved in them (i.e. members) – see Garfinkel (1967) for further elucidation. As is customary to ethnomethodology and conversation analysis, the selected excerpts capture short-but-strategically-relevant segments of interaction – the two transcripts cover interactions taking place over 54 and 27 seconds. Similarly, any individual interactional 'events' (such as handing over a turn-at-talking) are fleeting[3], measured in fractions of seconds. This is appropriate to a project which aims not to make generalisations from accumulated empirical cases, nor to demonstrate the relevance of empirical cases to sociological theories or concepts. Rather, these analyses are descriptive projects designed to unpick the salient features of seemingly innocuous social interactions and uncover what members take for granted as the mundane context of their social activities.

The key ethnomethodological concept we put to work is the "unique adequacy requirement of methods" (Garfinkel and Wieder 1992: 182). This requires that analysts understand the activities of members to the same degree *as* members, and are able to characterise the setting in terms indigenous to the practice. This is facilitated by virtue of one author being also one of the two music-makers. In this regard, the principle author's prior music-making with LC counts equally as fieldwork, which is invaluable for both practical and substantive reasons even when video recordings are available as a 'data resource'. As Heath, Hindmarsh and Luff note:

"Fieldwork… can help you develop a familiarity with the characteristics of the setting that may be critical in deciding when and how to record, where to position equipment and how to deal with problems that might arise in securing a clear visual image and good quality sound" (Heath, Hindmarsh and Luff 2010: 49-50).

However, Bezemer *et al.*, in their study of interaction in surgical theatres, argue that some elements of interaction are not so easily reduced to a set of field notes:

"Much of what nurses and surgeons do is instantiated in the subtle and fine grained detail of body movements such as the positioning of a retractor, or a shift in gaze from operative field to scrub nurse. Thus video analysis produces a much richer and nuanced account of communication than what can be captured on-the-spot and in field notes by researchers" (Bezemer *et al.* 2011a: 315).

Hence, a key affordance of the video recordings is the way in which they "preserve limited but crucial aspects of the spatial and environmental features of a setting, the temporal unfolding organization of talk, the visible display of participants' bodies and changes in relevant phenomena in the setting as relevant courses of action unfold" (Goodwin 2011: 179). Despite PB and LC's music-making being a leisure activity there is nonetheless a large degree of complex technical work which may have been difficult to account for from a single witnessing. Our concentrating on two short video excerpts allows us to pinpoint any "seen-but-unnoticed-features" (Garfinkel 1964: 229) of the ongoing interactions, which zip by unremarkably for those involved, but which for outsiders (i.e. analysts) may require more concentration[4].

The analyses themselves have been conducted with the aim of uncovering the multi-modality with which PB and LC carry out their activities – inasmuch as their activities rely on and feature non-verbal audio elements, our analyses account for those elements alongside the physical and verbal happenings between members and on-screen. In this sense, our video work is akin to others in the field (i.e. Bezemer *et al*. 2011a, 2011b; Heath and vom Lehn 2004; Lindwall 2008; Mondada 2007; Murphy 2012), who draw on visual aspects of members' settings as fundamental to their organisation. For instance, Bezemer *et al* (2011a) investigate the 'implicit' gestural cues involved in surgical work, or Heath and vom Lehn attend to how visitors' bodies "feature in the perception and experience of exhibits in museums and galleries" (2004: 46) as they point at and re-enact scenes in paintings and so on. What unifies these accounts is their use of the visible/'witnessable' qualities of interactions as an analytic resource – for instance, Murphy (2012) captures the tactile communication between designers by referring to the mutually visible hand gestures through which they communicate that the item they are designing should be 'soft'. Here, Murphy relies on a visual proxy for capturing the tactile sense; his video data allows him to *see* this work taking place (though, crucially, not to *feel* it).

Mondada (2007) notes that there are already systems for helping researchers supplement their transcriptions with visual-based interactions, e.g. Goodwin's (1981) gaze and Schegloff's (1984) gesture notation systems. Yet no such system exists for non-spoken audio, creating a technical problem for empirical work where sound is a key organising principle for members. Hence, we probe the possibility of analytic usages of non-spoken sound. We achieve this through our deployment of the 'nth member' as a heuristic device arising as a result of our transcripts being purposefully mis-written and mis-read. In this sense, the transcripts are unsatisfactory in certain respects; for instance, in lieu of a method for integrating audio

15

elements into a written transcript we present the sequentiality of a musical playthrough by referring to the number of the bar currently playing at a given time, which does not adequately capture the 'material' properties of the actual sounds available for hearing. However, this inadequacy is part of what makes our heuristic use of these transcripts possible. These 'missing whats' help provoke questioning around precisely what is *not* captured – the real-time goings-on of listening to and reviewing musical decisions and operating music production software. Exploring what these transcripts (and the analyses they help elicit) can offer will remain a focus throughout the paper.

The Cases

We now turn to the transcribed video excerpts, dually working through features of the organisation of the substantive areas at hand and exploring what the idea of an 'nth member' offers analytically. The cases are transcribed with conversation analysis conventions adapted from ten Have (2007: 215-216) and Hutchby and Wooffitt (2002: vi-vii).

Talking About Musical Ideas

In this first excerpt, PB and LC are discussing new musical ideas whilst simultaneously listening to the realisation of previous ones. The analysis explores how conversational work is managed in relation to a contextually-pervading 'nth member': a continually-playing source of audio which sometimes forms an ambient background to conversation but which also sometimes comes to the fore as something requiring a more attentive listening. We elucidate how the two members work with the material aspects (i.e. the *content*) of the music as an organisational resource which helps structure their collaboration. This is done with reference to a conversational technique – turn-taking (Sacks, Schegloff and Jefferson 1974).

Transcript 1 describes an activity where PB and LC are listening to the latest arrangement of a song, whereupon LC introduces, explains and reviews some recent changes to PB. The song has been generated by LC prior to collaboration with PB, and is being presented to PB as a 'show-and-tell' exercise. Prior to the transcribed episode, PB and LC have played the song from its beginning, listening to elements and commenting on them without stopping the playthrough. The audio of the song is transcribed below, as an 'nth member' (NM), paying regard to its sequential organisation in the wider 'conversation'. Both members are able to *see* the sequencer trackbar traverse bars as the audio plays, though it is not necessarily obvious as to which coloured bars refer to what instrument or what sounds each instrument produces (see figure 2). PB and LC are also able to *hear* indications as to where the playthrough may be up to – this is the subject of the transcript below, wherein LC arranges for PB to hear some recent changes to the sound of an instrument.

The focus here is on a change in topic instigated by LC, from PB's discussion of how to resolve a separate musical issue. LC is sitting at the computer in position to control the mouse and keyboard, and PB is spectating from behind. The focal point of the analysis is this change in topic and how LC manually manages the 'ambient' audio background (by 'interrupting' the playthrough and verbally introducing a 'thing to be listened to') to permit him to discuss a new topic.

Transcript 1: Talking About Musical Ideas

```
 1.    NM:    playing from bar 37
 2.    PB:    like on ESP you don't notice
 3.           (2.5)
 4.    LC:    Yeah
 5.    NM:    bar 39=
 6.    PB:              =and, plus, it's only to work on. cos if you did
 7.           want it to start as soon as you press play (1.0) the last
 8.           thing we'd do on the song.=
 9.    NM:                          =bar 41
10.           (1.2)
11.    PB:    before ripping it (.) to a WAV (.) we'd just move it back.
```

```
12.              (2.3)
13.   PB:    and then we've worked on it so it's done.=
14.   LC:                                        =yeah=
15.   NM:                                              =bar 43
16.              (2.1)
17.   PB:    it's well sweet this though, I've talked all over it but
18.          it's brilliant.=
19.   LC:                   =you di-, you didn't hear- ma- the- the-
20.          my porta=mento.
21.   NM:           =bar 33
22.              (0.5)
23.   LC:    what I had to do for this: (.) I've just made an exact
24.          copy of this ins- er, the organ instrument.=
25.   PB:                                                 =°m::°=
26.   LC:                                                       =but I
27.          wanted portamento on it.=
28.   NM:                            =bar 35
29.   LC:    so: I just (.) that's all it is(.)
30.   PB:                                     (.[)yeah=
31.   LC:                                            =just exactly the
32.          same <but with portamento on it> (0.8) °but I didn't want
33.          it on these bits.°
34.              (0.8)
35.   LC:    ↑n here it is.=
36.   NM:                 =bar 37
37.              (3.9)
38.   LC:    >yuh c'n ba:rely hear it.<=
39.   PB:                              =m::
40.   LC:    °b't° I know it's there.=
41.   PB:                           =yeah.
42.              (1.1)
43.   NM:    bar 39
```

The excerpt begins with the song playing as PB discusses some ancillary issues unrelated to the playthrough (lines 1-18). At lines 19-20, LC finds an opportunity to foreground a recent change he has made – the addition of a portamento effect[5] to an organ instrument, which has already passed by unnoticed by PB – and formulates this as the topic at hand ("you di-, you didn't hear- ma- the- the- my portamento") (cf. Goodwin (2003) on pointing as a gestural method for foregrounding). At this point, LC instigates a return to bar 33; 4 bars before the point where LC knows the portamento can be heard. This allows him time to explain the changes he's made to PB and gives PB an opportunity to compare the pre- and post-portamento instrument. LC holds the floor for a verbal explanation of the portamento effect, then begins to close his turn at line 29, summarising his explanation as the playthrough approaches bar 37 (where he knows the portamento effect is first introduced). At this point,

line 35, LC 'nominates' the playthrough itself for a turn (denoted as NM in the transcript), foregrounding it as something to be listened to closely – "↑n here it is.", immediately prior to bar 37. At this point, the two members are silent for a period (3.9 seconds) while the song plays. LC draws this period to a close by reviewing the change post-hearing – ">yuh c'n ba:rely hear it.< °b't° I know it's there.". PB agrees, and the song continues.

Whilst PB and LC converse over the 'ambient' background audio, there are brief periods when the audio is made something to be more pointedly attended to. At these periods, PB and LC undertake conversational work to orient their attention to the audio. Here, LC wishes to topicalise a portamento effect on an instrument, and being currently in control of the playback he arranges for PB to hear the portamento through his talk and through handling the software to a point before which the effect can be heard. As the song comes up to the point where the portamento can first be heard, LC ensures that the audio – the 'nth member' – is foregrounded as something not to be interrupted by PB. After allowing sufficient time to hear the portamento effect uninterrupted, LC draws the listening to a close by taking a turn himself to review what has just been hearable.

Analytically we take this episode to be one of turn-taking, though it is a special case given the peculiar way in which the playthrough is 'offered' a turn of its own. Sacks, Schegloff and Jefferson note a rule of turn-taking: that "Overwhelmingly, one party talks at a time" (1974: 700). Furthermore:

> "First the system allocates single turns to single speakers: any speaker gets,
> with the turn, exclusive rights to talk to the first possible completion of an
> initial instance of unit type [a question, a sentence, etc]...Second, all turn-
> transfer is coordinated around transition-relevance places, which are

themselves determined by possible completion points for instances of the

unit-types" (Sacks, Schegloff and Jefferson 1974: 706).

Using this rule as a basis for (mis-)understanding the episode, we find that LC grants access of a turn to the song that is playing, thereby also giving it 'exclusive rights' to complete that turn (i.e. reach a point after the portamento effect has been demonstrated). Here, PB and LC work with the projectability of the music's turn, available to both members as an audible property (denoted by a 'different-sounding' instrument) that is also prefaced verbally by LC (cf. Rauniomaa and Heinemann (2014) on prefacing audio 'interruptions', and Broth (2008) on the duration of camera shots in TV production work). The interaction is coordinated around this transition-relevant point – at a point where LC adjudicates that the portamento effect has been sufficiently hearable, he takes a turn himself to close the demonstration.

Enacting Musical Ideas

In our second excerpt, PB and LC are utilising the DAW to enact musical ideas (i.e. programming notes and automating controls). The desktop PC being used, which features one mouse and one QWERTY keyboard, necessitates that carrying out any computing task involves interactional negotiation. In these periods, features of the software (and their visual representation on-screen) come to the fore as the 'nth member' around which PB and LC construct their activity. PB and LC initiate the task of enacting musical ideas with the software by 'next-selecting' (Sacks, Schegloff and Jefferson 1974) one of them to take the controls; this process of selection forms the focus of the analysis. The 'nth member' in this episode is not transcribed – it informs the basis of the whole organisation of the activity and accordingly features more in the analysis of the episode (where we, the authors, tease out how the software becomes relevant to the ongoing interaction).

Transcript 2 depicts PB and LC attempting to utilise a sampling instrument called an 'NN19'

to achieve a 'timestretch' effect[6]. Though both PB and LC understood the concept of

'timestretching', they were yet to hear how it would sound when applied in the specific

context of the current instrument. Furthermore, only PB had previously attempted to enact

'timestretching' with an NN19. This excerpt focuses on how PB and LC decide to hand over

the computer controls (from LC to PB) in regard to the contextualisations afforded by the

software itself.

Transcript 2: Enacting Musical Ideas

```
1.            ((LC is at computer, on the mouse and keyboard))
2.    LC:    so I've loaded that into, C:3::. Do I j- do I n- do I need
3.            a new patch for eera- ev'ry new note then? Cos I don't use
4.            NN19s.
5.            ((PB brings a stool across to the left of the computer and
6.            places it down in a spectators position))
7.    PB:    you need a new: (.) yeah.=
8.    LC:                          =(.[)do I need a [new NN19 fur:]
9.    PB:                                          [an (.) extra]
10.           (0.4)
11.   LC:    °OK°=
12.   PB:       = it is worth it for that (.) well=
13.   LC:                                  =ye:ah.=
14.           [[LC opens the Reason menu for creating new instruments,
15.           then closes it]]
16.   PB:                                     = let's see
17.           if it is worth it.
18.   LC:    oh, by doing the ti[mestretch.]
19.   PB:                       [timestretch yeah cos that's the only
20.           reason to 'ave 'em as far as I [can tell.]
21.   LC:                                   [ye:ah]
22.           ((PB sits on the spectator's stool))
23.   LC:    where (0.4) °where is it.°
24.           (0.9)
25.   LC:    <cos [ah've not ye- ah've never used it.>
26.           [((LC shifts his body away from the computer desk.
27.           Takes his hand off the mouse and puts it by his side. PB
28.           reaches for mouse with right hand and rests left arm on
29.           computer desk))]
30.   PB:    oh yeah, r- erm::]
31.           ((PB removes hand from mouse, orients gaze to camera, and
32.           gets up to reposition body[7]. LC picks up his coffee and
33.           drinks from it))
34.           (1.5)
35.           ((PB kneels in front of computer desk and takes control of
36.           mouse))
37.   PB:    (h)e:::y::
```

This excerpt begins with LC at the computer. At lines 2-4, LC asks PB a question, couching that in terms which indicate his lack of experience with the NN19 sampler. Whilst PB is positioning himself as a spectator to LC's enactment LC requests PB's help, leading to a discussion of the practicalities of doing 'timestretching' with an NN19 (lines 5-22). During this period, LC remains at the controls and attempts to enact the ideas under discussion (lines 14-15), whilst PB and LC jointly imagine and assess the proposed new sound (cf. Buscher (2001) and Laurier and Brown (2014) on similar interactions in architecture and film editing respectively) as well as clarify the aim of the activity (lines 16-20 – "let's see if it is worth it", "oh, by doing the ti[mestretch.]"). At line 22, PB instigates next-selection for the activity by consolidating his position as a spectator – he sits on the stool he has placed in a position only to watch the enactment, a move which projects LC taking the lead. Whilst already knowing where in the menu the NN19 can be located and initialised (lines 14-15), LC problematises the emerging arrangement of bodies and tasks by again indicating his unfamiliarity with NN19s (lines 23-25). LC further extricates himself from the position of enactment physically, by shifting his body away from the computer controls leaving them open for PB to take (lines 27-29). Acknowledging LC's lack of familiarity with NN19s (line 30, "oh yeah, r- erm::]]"), PB finds a comfortable position of enactment (lines 31-36), and begins enacting the 'timestretching'.

In this excerpt, the activity – 'timestretching' a sample – is given shape by virtue of the software requiring expertises that one member, LC, does not yet have. Though both members agree to try using one instance of an NN19 to test the general appeal of 'timestretching', LC states that he does not know how to put the idea into practice. In doing this, LC appeals to skills he knows PB has – PB has done 'timestretching' before, whilst LC has not. Our mis-reading of the transcript characterises this as an 'n[th] member' – the software and specific

elements of it (i.e. NN19s) – which shapes how the activity is conducted. Here, we identify

elements of Sacks, Schegloff and Jefferson's concept of 'next-selection'. As they note:

> "A current speaker may select a next speaker (as when he addresses a
>
> question to another party); or parties may self-select in starting to talk"
>
> (Sacks, Schegloff and Jefferson 1974: 701).

In the process of discussing the practical exigencies of 'timestretching', LC's withdrawing

from the enactment of the task consequentially leads to a next-selection of PB to do it – PB is

better placed to work with the tool at hand, which positions the tool (the NN19) as an

organisational resource for continuing the activity. This unfolds iteratively, with the (human)

members' interactions developing in light of LC's deepening demonstrations that he is

unfamiliar with the NN19 device. Hence, as LC works to make his unfamiliarity with NN19

an interactive resource, it becomes agreed (naturally, though not explicitly or verbally) that

PB should take the controls and enact the idea, which he does.

Concluding Remarks: Exploring Multi-Modal Interactions with 'n^{th} Members'

The present paper has had two aims. Firstly, to explore how verbal and interactional

resources become enmeshed with technical/computational features of digital music-making.

In this regard, we have rendered PB and LC's material practices in such a way as to highlight

how their social activity consists of ordinarily-achieved (yet technically-contextualised)

modes of interaction (i.e. 'turn-taking' and 'next-selection'). Our second aim was to show

how researchers might account for multi-modal activities incorporating the audio field. We

have outlined our conception of 'n^{th} members' and explained what they have added to our

analyses as a heuristic device for drawing out the intelligibility and accountability of the

practices on display.

23

What we argue here is that to understand multi-modal activities, it is necessary and valuable to feature elements of the settings and contexts of interaction in transcriptions and analyses. Our strategy has been to create a new member in our transcriptions and analyses, an 'nth member', which is portrayed, on an *as if* basis, as being oriented to by the human (i.e. *real*) members PB and LC. We do not mean to say that we require settings and contexts to be treated as members with equable agency to humans. PB and LC do not converse *with* 'nth members', only *about* them; they do not interact *with* them, only *through* them as a medium. In this sense, it is misconceived for analysts to treat these contexts differently than PB and LC do. Instead, what 'nth members' demonstrate is the value of an analytic orientation towards the contextual elements within which an interaction unfolds. For multi-modal interactions, settings and contexts are fluid and dynamic – they change, members react to and engage with them, they may be differently experienced by different members: above all, they are drawn on as resources for achieving practical accomplishments. Our usage of 'nth members' as a heuristic device, therefore, provokes an accounting for how these contextual elements feature as relevant to members interactions as they unfold.

In suspending our disbelief and (mis-)treating sonic/contextual elements of settings as members, we open up a way of capturing how they sometimes feature in interactional work. At this point, it becomes possible and appropriate to apply conversation analysis techniques to them as we have here, even where these so-called members do not 'speak' in the conventional sense. Here, however, the 'nth member' is to be dropped – this dropping is the key move, reflecting its status as a mis-writing and mis-reading of events rather than how members naturally characterise their interactions. Unlike those who subscribe to ANT, once we have a sense of their role in structuring and shaping members' interactions we find no need to retain a sense of inanimate objects as having the capacity to exert agency and

influence on interactions. Rather, and more in line with an ethnomethodological approach to socio-materiality, the 'nth member' is deployed to work out how complex technical knowledges and skills from other modes (i.e. audio) feature as part of the ordinary and everyday ways of achieving practical accomplishments.

Endnotes

1. Arguably this focus on amateur leisure pursuits originates in the 'serious leisure perspective' (see Stebbins 1992), though more recent ethnomethodological examples include: Bennerstedt and Ivarsson (2010) on collaboration in massively multiplayer online games; Durrant *et al*. (2012) on designing technology for theme park visitors, and; Tolmie, Benford and Rouncefield (2013) on playing in Irish music sessions.

2. Paradoxically, we achieve this empirically by subverting the idea of non-human members as agency-bearing from being an ontological principle (as ANT would have it) into a heuristic device which is provocative *precisely because* it is so far removed from members' accounts.

3. This should come as no surprise to readers regardless of their familiarity or otherwise with ethnomethodology and conversation analysis – think back to your last phone call and recall how many times and how quickly a change in turns-at-talking occurred!

4. Curiously, this remains the case even when the analyst is one of the members captured on video!

5. Portamento is a controllable effect for some devices within Reason, which alters the degree to which notes glide from one pitch into another. An example of this may be heard at: https://soundcloud.com/phillip-brooker/example-of-portamento-effect

6. 'Timestretching' is a term used to refer to an effect achievable with an NN19 sampling device where small slices of a sample are copied, arranged and replayed as a phrase of 1/64th notes. The NN19's 'sample start' control is gradually increased over the phrase so that each 1/64th note plays the same sample started at a slightly later point than the previous note. This both stretches the sample across time, as well as giving it a 'stuttery' feel. An example of this may be heard at: https://soundcloud.com/phillip-brooker/example-of-timestretching

7. Though this action highlights PB's role as both member and researcher, this duality does not change the outcome in any relevant way – PB's awareness of what the camera can capture does not halt the accomplishment of the task at hand (namely, setting up an NN19 device).

References

Acord, Sophia K. and Tia DeNora. 2008. "Culture and the Arts: From Art Worlds to Arts-in-Action." *Annals of the American Academy of Political and Social Science* 619:223-237.

Adorno, Theodor W. 1978. "On the Social Situation of Music." *Telos* 35:128-164.

Becker, Howard S. 1966. *Outsiders: Studies in the Sociology of Deviance*. London: Collier-Macmillan Limited.

Becker, Howard S., Robert R. Faulkner and Barbara Kirshenblatt-Gimblett. 2006. "Editors' Introduction." Pp. 1-20 in *Art From Start to Finish: Jazz, Painting and Other Improvisations*, edited by H. S. Becker, R. R. Faulkner and B. Kirshenblatt-Gimblett. London: University of Chicago Press.

Bennerstedt, Ulrika and Jonas Ivarsson. 2010. "Knowing the Way. Managing Epistemic Topologies in Virtual Game Worlds." *Computer Supported Cooperative Work* 19:201-230.

Bezemer, Jeff, Alexandra Cope, Gunther Kress and Roger Kneebone. 2011a. ""Do You Have Another Johan?" Negotiating Meaning in the Operating Theatre." *Applied Linguistics Review* 2:313-334.

Bezemer, Jeff, Ged Murtagh, Alexandra Cope, Gunther Kress and Roger Kneebone. 2011b. ""Scissors, Please": The Practical Accomplishment of Surgical Work in the Operating Theater." *Symbolic Interaction* 34(3):398-414.

Blackstone, Lee R. 2009. ""The Spider is Alive": Reassessing Becker's Theory of Artistic Conventions Through Southern Italian Music." *Symbolic Interaction* 32(3):184-206.

Broth, Mathias. 2008. "The "Listening Shot" as a Collaborative Practice for Categorizing Studio Participants in a Live TV-Production." *Ethnographic Studies* 10:69-88.

Büscher, Monika. 2001. *Ideas in the Making: Talk, Vision, Objects and Embodied Action in Multi-media Art and Landscape Architecture*. PhD dissertation, Lancaster University.

Callon, Michel. 1986. "Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fishermen of St. Brieuc Bay.", Pp. 196-223 in *Power, Action and Belief: A New Sociology of Knowledge*, edited by J. Law. London: Routledge & Kegan Paul.

Coulter, Jeff and E. D. Parsons. 1990. "The Praxiology of Perception: Visual Orientations and Practical Action." *Inquiry* 33(3):251-272.

DeNora, Tia. 2003. *After Adorno: Rethinking Music Sociology*. Cambridge: Cambridge University Press.

Durrant, Abigail, David S. Kirk, Steve Benford and Tom Rodden. 2012. "Pursuing Leisure: Reflections on Theme Park Visiting." *Computer Supported Cooperative Work* 21:43-79.

Faulkner, Robert R. and Howard S. Becker. 2009. *"Do You Know...?": The Jazz Repertoire in Action*. London: The University of Chicago Press.

Frith, Simon. 1978. *The Sociology of Rock*. London: Constable.

de la Fuente, Eduardo. 2007. "The 'New Sociology of Art': Putting Art Back into Social Science Approaches to the Arts." *Cultural Sociology* 1(3):409-425.

Garfinkel, Harold. 1964. "Studies of the Routine Grounds of Everyday Activities." *Social Problems* 11(3):225-250.

Garfinkel, Harold. 1967. *Studies in Ethnomethodology*. New Jersey: Prentice Hall, Inc.

Garfinkel, Harold and D. Lawrence Wieder. 1992. "Two Incommensurable, Asymmetrically Alternate Technologies of Social Analysis." Pp. 175-206 in *Text in Context:*

Contributions to Ethnomethodology, edited by G. Watson and R. M. Seiler, London: Sage.

Greiffenhagen, Christian. 2014. "The Materiality of Mathematics: Presenting Mathematics at the Blackboard." *The British Journal of Sociology* 65(3):1-27.

Goodwin, Charles. 1981. *Conversational Organization: Interaction Between Speakers and Hearers*. New York: Academic Press.

Goodwin, Charles. 2001. "Practices of Seeing Visual Analysis: An Ethnomethodological Approach." Pp. 157-182 in *Handbook of Visual Analysis*, edited by T. van Leeuwen and C. Jewitt. London: Sage.

Goodwin, Charles. 2003. "Pointing as Situated Practice." Pp. 217-242 in *Pointing: Where Language, Culture, and Cognition Meet*, edited by S. Kita, London: Lawrence Erlbaum Associates.

Heath, Christian, Jon Hindmarsh and Paul Luff. 2010. *Video in Qualitative Research: Analysing Social Interaction in Everyday Life*. London: Sage.

Heath, Christian and Dirk vom Lehn. 2004. "Configuring Reception (Dis-)Regarding the "Spectator" in Museums and Galleries." *Theory, Culture & Society* 21(6):43-65.

Holyfield, Lori, Maggie Cobb, Kimberly Murray & Ashleigh McKinzie. 2013. "Musical Ties That Bind: Nostalgia, Affect and Heritage in Festival Narratives." *Symbolic Interaction* 36(4):457-477.

Hutchby, Ian and Robin Wooffitt. 2002. *Conversation Analysis: Principles, Practices and Applications*. Cambridge: Polity Press.

Kotarba, Joseph A., Jennifer L. Fackler & Katherine M. Nowotny. 2009. "An Ethnography of Emerging Latino Music Scenes." *Symbolic Interaction* 32(4):310-333.

Laing, Dave. 1985. *One Chord Wonders: Power and Meaning in Punk Rock*. Milton Keynes: Open University Press.

Latour, Bruno. 2005. *Reassembling the Social: An Introduction to Actor-Network Theory*. Oxford: Oxford University Press.

Laurier, Eric and Barry Brown. 2014. "The Mediated Work of Imagination in Film Editing: Proposals, Suggestions, Reiterations, Directions, and Other Ways of Producing Possible Sequences." Pp. 237-261 in *Studies of Video Practices: Video at Work*, edited by M. Broth, E. Laurier and L. Mondada. New York: Routledge/Taylor & Francis Group.

Lerdahl, Fred and Ray Jackendoff. 1996. *A Generative Theory of Tonal Music*. London: The MIT Press.

Lindwall, Oskar. 2008. *Lab Work in Science Education: Instruction, Inscription, and the Practical Achievement of Understanding*. Linköping University: Faculty of Arts and Sciences.

Lynch, Michael. 1993. *Scientific Practice and Ordinary Action: Ethnomethodology and Social Studies of Science*. Cambridge: Cambridge University Press.

Martin, Peter J. 1995. *Sounds and Society: Themes in the Sociology of Music*. Manchester: Manchester University Press.

Merrill, Bryce. 2010. "Music to Remember Me By: Technologies of Memory in Home Recording." *Symbolic Interaction* 33(3):456-474.

Mondada, Lorenza. 2007. "Multimodal Resources for Turn-Taking: Pointing and the Emergence of Possible Next Speakers." *Discourse Studies* 9(2):194-225.

Murphy, Keith M. 2012. "Transmodality and Temporality in Design Interactions." *Journal of Pragmatics* 44:1966-1981.

O'Hara, Kenton and Barry Brown (eds.). 2006. *Consuming Music Together: Social and Collaborative Aspects of Music Consumption Technologies*. Dordrecht, The Netherlands: Springer.

Peterson, Richard A. and David G. Berger. 1971. "Entrepreneurship in Organisations: Evidence From the Popular Music Industry." *Administrative Science Quarterly* 16(1):97-106.

Rauniomaa, Mirka and Trine Heinemann. 2014. "Organising the Soundscape: Participants' Orientation to Impending Sound When Turning On Auditory Objects in Interaction." Pp 145-168 in *Interacting with Objects: Language, Materiality, and Social Activity*, edited by M. Nevile, P. Haddington, T. Heinemann and M. Rauniomaa. Amsterdam: John Benjamins Publishing Company.

Rawls, Anne W. 2002. "Editor's Introduction." Pp 1-64 in Harold Garfinkel, *Ethnomethodology's Program: Working Out Durkheim's Aphorism*, edited by A. W. Rawls. Lanham, Maryland: Rowman and Littlefield.

Rouncefield, Mark and Peter Tolmie. 2013. "Overview: Garfinkel's Bastards at Play." Pp. 1-18 in *Ethnomethodology at Play*, edited by P. Tolmie and M. Rouncefield. Surrey: Ashgate.

Sacks, Harvey. 1995. *Lectures on Conversation, Volumes I & II*. Oxford: Blackwell.

Sacks, Harvey, Emanuel A. Schegloff and Gail Jefferson. 1974. "A Simplest Systematic for the Organisation of Turn-Taking for Conversation." *Language* 50(4):696-735.

Schegloff, Emanuel A. 1984. "On Some Gestures' Relation to Talk." Pp. 266-298 in *Structures of Social Action*, edited by J. M. Atkinson and J. Heritage. Cambridge: Cambridge University Press.

Schmidt, Kjeld. 2011. "The Concept of 'Work' in CSCW." *Computer Supported Cooperative Work* 20:341-401.

Schutz, Alfred. 1976. "Making Music Together: A Study in Social Relationship." Pp. 159-178 in *Collected Papers II: Studies in Social Theory*, edited by A. Broderson. The Hague, Netherlands: Martinus Nijhoff.

Sormani, Philippe. 2014. *Respecifying Lab Ethnography: An Ethnomethodological Study of Experimental Physics*. Surrey, UK: Ashgate Publishing Limited.

Stebbins, Robert A. 1992. *Amateurs, Professionals and Serious Leisure.* Montreal, QC and Kingston, ON: McGill-Queen's University Press.

Strandvad, Sara M. 2012. "Attached by the Product: A Socio-Material Direction in the Sociology of Art." *Cultural Sociology* 6(2):163-176.

Sudnow, David. 1978. *Ways of the Hand: The Organization of Improvised Conduct*. London: Routledge & Kegan Paul.

ten Have, Paul. 2007. *Doing Conversation Analysis: A Practical Guide (2ⁿᵈ edn.)*. London: Sage.

Tolmie, Peter, Steve Benford and Mark Rouncefield. 2013. "Playing in Irish Music Sessions." Pp. 227-256 in *Ethnomethodology at Play*, edited by P. Tolmie and M. Rouncefield. Surrey: Ashgate.

Weeks, Peter. 1990. "Musical Time as a Practical Accomplishment: A Change in Tempo." *Human Studies* 13:323-359.

Weeks, Peter. 1996. "Synchrony Lost, Synchrony Regained: The Achievement of Musical Coordination." *Human Studies* 19:199-228.