Article

Discourse Studies
2017, Vol. 19(5) 515–535

DOI: 10.1177/1461445617730235
journals.sagepub.com/home/dis
SAGE

On the concept of action in the study of interaction

NJ Enfield
The University of Sydney, Australia

Jack Sidnell
University of Toronto, Canada

Abstract

What is the relation between words and action? How does a person decide, based on what someone is saying, what would be an appropriate response? We argue that (1) every move combines independent semiotic features, to be interpreted under an assumption that social behavior is goal directed; (2) responding to actions is not equivalent to describing them; and (3) describing actions invokes rights and duties for which people are explicitly accountable. We conclude that interaction does not involve a 'binning' procedure in which the stream of conduct is sorted into discrete action types. Our argument is grounded in data from recordings of talk-in-interaction.

Keywords

Accountability, action, conversation analysis, philosophy of language, practice, speech acts, talk-in-interaction

> A *picture* held us captive. And we couldn't get outside it, for it lay in our language, and language seemed only to repeat it to us inexorably.
>
> Wittgenstein, *Philosophical Investigations*, §115

When we talk, we are not just saying things – making reference, conveying information – we are also, unavoidably, doing things (Austin, 1962).[1] This insight underlies a vast literature on speech acts in pragmatics and the philosophy of language (inter alia Austin, 1962; Searle, 1969), and a growing literature on actions in interaction in research on talk-in interaction (inter alia Levinson, 2013; Sacks, 1995; Schegloff, 1996, 1997, 2009). The

Corresponding author:
NJ Enfield, Department of Linguistics, The University of Sydney, Building A20, Sydney NSW 2006, Australia.
Email: nick.enfield@sydney.edu.au

idea is simple: when a person says something – for example, the statement 'She's just starting' – she is also, unavoidably, doing something. For example, she is 'noticing something', 'making an excuse', 'criticizing someone' or 'defending someone'. How do we get from the words someone says to the thing(s) that they are doing? What exactly is the relation between words and action?

Research on language and social interaction is making progress in seeking to answer this question (Couper-Kuhlen, 2014; Heritage, 2012; Levinson, 2013; Rossi, 2015; Rossi and Zinken, 2016), and scholarly interest in the domain is growing. But there is a need for conceptual clarity. In what follows, we preview our book *The Concept of Action* (Enfield and Sidnell, 2017), with the goal of clarifying the concept of action as relevant to research on language in social interaction.

Two axioms, three proposals

We start with two axioms about the relation between language and action, which any reasonable account must handle:

> *Axiom 1. What we say* (*vocally and otherwise*) *and what we do by saying it are two different, yet related, things.* The same words (e.g. 'She's just starting') can have different import in different contexts (e.g. giving an account, defending somebody, noticing something). An analysis of action must determine the link between the specific practices used (i.e. words and grammatical constructions, prosody, contextual positioning, bodily conduct) and the action that these practices achieve in any instance of social interaction. This leads to the question: How does saying X count as doing Y?[2]
>
> *Axiom 2. An action, in interaction, is fully consummated only by its relation to the result it brings about, a central aspect of which is the response it engenders*. Since action in interaction is action upon another person, what is effectively done by a move is only fully determined with reference to how that move is responded to (as long as that response is allowed through).[3] An analysis of action must incorporate both formulation and interpretation, and the relation of co-determination between the two. This leads to the question: How is saying X related to the appropriate response Z?[4]

These two axioms lead us to the central puzzle for the study of action in interaction. This is a problem of interpretation that participants constantly face in interaction. How does a person decide, on the basis of what someone is saying, what would be an appropriate response?[5] Our approach to solving this puzzle makes reference to three proposals:

> *Proposal 1. Every move is made up of a combination of independent features from all corners of the semiotic repertoire; all features matter for the interpretation (and thus the construction) of action, and none necessarily signal specific actions*. The task of figuring out how to respond in interaction requires us to view a person's conduct as means to some ends (i.e. as purposive, goal-directed behavior). This

presupposes that people subprehend others' goals from their conduct.[6] The machinery of interaction is organized around recipients' flexible capacity to interpret conduct by viewing it as goal-directed and purposive: that is, people's propensity to interpret controlled conduct as having been produced for a reason.[7] This is what people are doing when they ask – consciously or not – 'Why that now?' in the flow of interaction (Schegloff and Sacks, 1973). We do not mean that people are always mentalizing 'What is her reason?'. We mean that, as Garfinkel (1952: 367) argued, we are guided by a sense of how people in general would likely interpret others' conduct in terms of its reasons.[8]

Proposal 2. An action cannot be identified and talked about without a description being chosen for that action, and any description is one among many possible; conversely responding to an action is not equivalent to describing it, nor does it entail that a description of it has been made at any level. There are two ways in which a move can be categorized in action terms. One is what it is *called* (or effectively called; that is, how it can be described),[9] while the other is in the way an action is (appropriately) *treated* in response (drawing on the inferences of means-to-ends mapping). How we respond to an action and what we call it can be two quite different things.[10] This has important consequences for an understanding of action precisely because categorization by treating it in a certain way does not implicate a pre-existing inventory of discrete actions, while categorization by naming it in a certain way (what we refer to as 'binning') does suggest such an inventory. So when, in response to one participant saying something, a recipient passes the butter, their passing the butter does not describe the prior action. Rather, it treats the prior move as requiring the passing of the butter. Crucially, it does not definitively characterize the initiating conduct as having been an instance of a particular action (e.g. as a request, complaint, command, instruction or suggestion).[11] A definitive categorization of conduct into action types is not required for the orderly flow of interaction.[12]

Proposal 3. Alternate descriptions of action foreground different accountabilities (Schegloff, 1988; Sidnell, 2017). *The verbal description of an action invokes a particular set of rights and duties for which those involved are treated as accountable, and therefore such description is specifically relevant to thematizing the accountability of participants; analysts must be careful as to whether they mean to be thematizing certain accountabilities and not others when they name actions in interaction*. Conduct in interaction is designed by reference to accountability.[13] Accountability harnesses the linguistic capacity for reflexive formulation, that is, the fact that language can be used to say things about itself (Jakobson, 1957). In order to hold someone accountable, you have to be able to say what they have done, and to do this, you have to choose one way of describing or formulating their conduct as a type of (intentional) action (or by word-for-word report, e.g. 'She said, "She's just starting"'). Saying what a person has done is equivalent to saying what they are accountable for having done.

The essence and upshot of these proposals was anticipated in a lecture by Harvey Sacks (1995, vol. 1):

> [T]here's no room in the world to definitively propose formulations of activities, identifications, and settings. They are things that have got to be used with caution. And there may well be ways of allowing things to go on without their necessary invocation. (p. 517)

The argument that interaction can proceed without the use of explicit formulation holds, according to Sacks, for 'activities, identifications and settings', 'role relationships' and 'meaningful actions'.

Here is our key point. Participants in interaction need not definitively formulate or categorize conduct in order to respond to it. Interaction does not involve a kind of binning procedure in which the stream of conduct is sorted into discrete action types. This is because responses do not themselves imply definitive categorization of prior actions into discrete types. It does happen that prior actions are described ('Are you asking or telling me?', 'Don't tell me what to do!', 'You already asked me that', etc.), but even in such cases, there is no guarantee that the 'action' proposed by such a description accurately (or adequately or exhaustively) describes what was done. Indeed, there are good reasons for supposing that it could not possibly do this, that no bit of conduct could be definitively described by a single description.

Can actions be readily categorized?

Some examples might suggest that linguistic practices map onto social actions in straightforward ways. Take the highlighted exchanges in the following cases from a family dinner:

(1)
DAD: **An' may I have thuh -b<u>u</u>tter please.**
MOM: Yes.=hh ((extending hand to pass butter))

(2)
MOM: °Can I have the butter please,°
DAD: Sure. ((extending hand to pass butter))

In these two cases, from a family dinner involving a mother, father, and school-age daughter, an identifiable linguistic structure is used in a way that is readily recognizable to us as analysts, and to the participants, as a request (see e.g. Couper-Kuhlen, 2014; Curl and Drew, 2008; Heinemann, 2006). We can say this on the basis of two kinds of evidence. One is in its internal composition, the measurable features of the highlighted moves themselves. In these cases, we can identify a 'May/Can I have X please' yes/no question construction that is consistent across the cases. Second, there is the sequential status of the move. We can point to the uptake that each of these moves receives: the addressed parties immediately produce a verbal confirmation ('sure', 'yes') in response to the ostensive yes–no question, while at the same time producing a bodily action of passing the named/requested object.

On the basis of these features of position and composition, it is hard to imagine anybody hesitating to label the highlighted moves as requests. Nor is there any reason to think that the interactants had any doubt as to how they should respond, that is, by fulfilling the request.

One view might be that this clarity of action and response – and especially of the analyst's ability to label the action being done – provides evidence that interpreting a move as an action is a matter of categorizing[14] the swatch of behavior 'as action X', and *through that act of categorization* thus knowing what an appropriate response would be. We refer to this as a 'binning' account of action. By this account, the interpreter assigns the move to an action category or bin, and this determines the appropriate category of response (independent from a range of specifics, determined, for example, by the broader context of activity, or the specific details of composition).

But there are problems with this view. One is that the vast majority of moves in interaction are not recognizable like these textbook examples.[15] Anyone who has attempted to work through a conversational transcript and exhaustively 'identify the action' in every line knows that it is a hopeless task. But if 'identifying the action' were a procedure that participants followed, then the task should not be hopeless for analysts. We contend that focusing primarily on such minimally ambiguous cases as our examples of requests does more harm than good in the study of interaction.[16] It not only fails to reveal the significant complications involved in the production of interaction, it also encourages the problematic way of thinking that we term the binning approach. Therefore, seemingly straightforward words–action mappings such as mealtime requests are a poor model for action. The truth is that most moves in interaction lack this clarity of design-to-action mapping.

To be clear, our critique is not a methodological one. That is, we are not proposing that conversation analysts, or other scholars of interaction, necessarily 'bin' actions in the course of their investigations (although some may do, at least to some extent or for certain analytic purposes). Rather, we are suggesting that the binning idea constitutes an underlying, tacitly held conceptual model of the way interaction works. To paraphrase Wittgenstein (see the quotation earlier), this is a picture that 'holds us captive'. Because the binning model is so deeply held and so rooted in our language (through e.g. metalinguistic vocabulary as well as the grammar of reported speech – see Rumsey, 1990; Silverstein, 1979), it is difficult to see past it, to conceptualize the phenomena of action in interaction in any other way. What we are proposing then is that the binning model is a 'concept of action', and one that informs much scholarship into the organization of social interaction.[17]

The many bits of conduct that make up actions

What alternative to a binning approach to action will account for the orderly organization of conduct in interaction? The following fragment from a conversation between a husband and wife having dinner allows us to illustrate our approach to this problem. Our observations will focus on the highlighted lines (03, 13):

[Ravioli dinner]
01 Kim: I don't like it.
02 (35.0)
03 Mark: **Dennis came in today:=uh:m** (0.2) tlk=He wrote this big
04 letter tuh Ford,

05 (1.4)
06 Mark: Cuz his car's been in the shop ya know,
07 (1.5)
08 Mark: tlk for so long,
09 (.)
10 Mark: Took 'em like two weeks tuh fix his transmission.
11 (0.3)
12 Mark: (H[e had tuh ri-)/(he had uh ru-)
13 Kim: [**So he wrote a complaint?**
14 (2.0)
15 Mark: .hh Mm hm?,
16 (4.0)
17 Mark: ≠We:ll, (0.5) he has like: eight hundred dollars in:
18 (.) rental cars fer (.) two weeks.
19 (0.4)

There has been a lapse (for 35 seconds) when Mark produces the talk at line 03: 'Dennis came in today:=uh:m (0.2)'. Just before he does this, Mark has scooped food onto his fork and is raising it toward his mouth. He pauses, looking down at his fork which is midway between his mouth and his plate, and begins speaking. Mark has then visibly suspended his ongoing course of action (eating) in order to produce this first part of the talk. The fork is held in the air during the silence of two-tenths of a second, but when Mark begins to speak again he lowers the fork, thereby disengaging from the activity of eating and prioritizing the activity of telling Kim about Dennis (see Raymond and Lerner, 2014). Within the silence at line 03, Kim momentarily looks up from her own plate, and the ravioli she is dividing with a fork, and glances at Mark. Kim begins to look up at the very last part of Mark's 'uhm' and just glances at him. When Mark starts speaking again (line 03), Kim has returned her gaze to her plate:

Dennis came in today:=uh:m (0.2).

We might ask, '*What action* is Mark doing at line 03'? That question could be answered in a range of ways. We could say that he is 'asserting that X'. We could equally well say that he is 'telling Kim X' or that he is 'proffering a topic'. Alternatively, we could say that he is 'making a bid to tell a story', and so on. The problem with the question is that it is unconstrained and could be answered accurately in any number of ways.[18]

We propose that a better question to ask is this: 'How can Kim appropriately respond to what has Mark done?' This question goes directly to what matters for these participants. It is the question that Kim has to answer. We suggest that in order for Kim to do this, she needs to look for clues in the design of Mark's move. This point is of course well known to analysts of talk-in-interaction, but we emphasize that her inspection of his turn need not be – and will seldom if ever be – aimed at categorizing *what action* has been produced. Her concern is how to appropriately respond.

Taking this as a starting point, we can notice that Mark's turn refers to someone by name, thereby indicating that Kim should be able to recognize who he is talking about (Sacks and Schegloff, 1979). Mark has, then, begun to talk about someone that both

speaker and recipient know. Moreover, by saying that Dennis 'came in today', Mark has placed this person, who both speaker and recipient know, in a setting consisting of a place indicated solely by 'came in' (likely Mark's workplace, drawing on Mark and Kim's common ground) and a time, 'today'. Such recognitional references to persons coupled with descriptions of settings are strongly associated with the beginnings of stories (Dingemanse et al., 2017; Jefferson, 1978; Rossano, 2012; Sacks, 1995; Sidnell, 2010). This does not mean that a participant has to categorize the move as a 'story launching'. The key thing is that they need to respond to such reference–description couplings by showing themselves to be willing recipients (Stivers, 2008).

So, Mark has talked in such a way as to suggest that he has something to tell Kim about a mutual acquaintance – something that he learnt today. The reference–description coupling might well have been enough for Kim to be able to produce an appropriate aligning response which, had the turn reached completion at this point, might have been something like 'uh huh' (Schegloff, 1982). However, Mark does not come to possible completion of this unit at 'today'. Instead he produces 'uhm' latched to 'today' with no perceptible gap. In this way, Mark shows that the unit of talk so far produced is to be continued. Particles such as English 'uhm' in a position like this often have as their basic function 'turn-holding'.[19] Here, Kim can gather that Mark has more to say and thus that this part of the talk is merely a beginning to something more extensive. At the same time, disfluencies such as this one regularly elicit the gaze of a non-gazing recipient (Goodwin, 1979; Rossano, 2012).

How then does Kim respond? Her glance at Mark is timed to coincide with just that moment when he is continuing without actually speaking. With a quick glance, Kim can show Mark that she recognizes a boundary in the talk has been reached but that the talk will nevertheless continue. Moreover, by glancing while not talking (and thus not initiating repair), Kim conveys that she has recognized the person referred to by the name 'Dennis'.

We might say that Kim's glance responds to Mark's talk not as a speech act of asserting, but as the beginning of a larger project, the telling of a story. Rather than say that Kim has at some level decided 'This is a story launching', a simpler account is that by responding with a glance, and with no talk of her own, Kim has found a way of responding that aligns with the activity Mark has initiated, an activity in which she will serve as recipient and he as teller. There is no evidence that Kim has categorized Mark's turn (by ascription or recognition or in some other way) as anything specific at all, whether it be 'story launching' or 'announcement' or 'topic bid' or something else. All she has had to do is produce an appropriate response. At the very least, this is a more parsimonious account than a binning-style action categorization account. Under both accounts, the appropriate response must be formulated. In our account, there is no intermediate level of action labeling.

Any analysis of action must account for how and why Kim produces the response she produces – in other words, why she treats Mark's move in a certain way – and it is not clear that we need to do more than that. Adding a claim that she (consciously or not) ascribed or recognized 'what action' it was is neither necessary nor sufficient, and without such a claim our account is more economical. As we outline in the following, naming the action is not only more than is needed, it also adds something that demonstrably is not there; it adds that Mark is specifically accountable for 'launching a story'

or whatever other action he may be claimed to be doing (and correspondingly it implies that he is not accountable for the alternative – but not offered – descriptions of what he was doing).

Let us further develop this argument as we turn to the next line of interest. After his first mention of Dennis coming in today, Mark continues, saying that Dennis wrote to Ford because of how long his car was in the shop. Then Kim (line 13) says,

So he wrote a complaint?

With 'So he wrote a complaint?', Kim formulates the upshot of what Mark has been saying by glossing in speech act terms the action that Dennis was engaged in by writing 'this big letter', as Mark put it. Naming the action Dennis was engaged in provides a strong demonstration of understanding what Mark has been saying because it requires Kim to piece together the bits of information so far provided. This can be seen as furthering the work of alignment that Kim began with the glance discussed earlier. A demonstration of understanding shows not just that she has acted as a recipient, but also that she has understood.

Different kinds of responses are made relevant at different points within the project Mark is engaged in. So in line 03, the talk 'Dennis came in today:=uh:m (0.2)' only made a display of alignment relevant from Kim. At that point, little else had been revealed to which she might respond. Here, however, with more of the story Kim can respond more substantially, showing both a willingness to align as a story recipient and some of the fruits of that alignment. With 'So he wrote a complaint?', Kim displays an understanding of what has so far been said.

An action is constructed from components, and some actions are leaner than others. For instance, Kim's glance consists solely of a head/gaze reorientation and momentary suspension of her other bodily actions.[20] 'So he wrote a complaint?', conversely, is produced while she looks at her plate (where it landed after Mark's 'two weeks' – Kim gazing at Mark before that) and so does not actively employ gaze for her action. Rather, the action here makes use of the linguistic resources of lexical semantics and grammar. 'So' shows the relation to the prior talk, marking what is said as an inference from the evidence Mark has provided (Bolden, 2006; Raymond, 2004). The rest of the turn draws on English semantics to name the action that Dennis was engaged in. In Peircean terms, this is a representational interpretant of Mark's statement in lines 03–04 (Enfield, 2013; Kockelman, 2005). Notice that although the grammatical construction is declarative, Kim can be interpreted as doing questioning here. In line 15, Mark confirms with 'Mm hm?'. Declarative utterances like this are routinely treated as seeking answer type responses where they concern matters in the recipient's epistemic domain (Heritage, 2012). Mark knows better than Kim what Dennis was doing in writing the big letter and thus her turn is treated as seeking confirmation.

Another aspect of this turn's composition is its prosody. Consider Mark's continuing response to what Kim has said. Although he first confirms that indeed Dennis 'wrote a complaint' with 'Mm hm?', at lines 17 and 18 he adds '≠We:ll, (0.5) he has like: eight hundred dollars in: (.) rental cars fer (.) two weeks'. With the 'well' preface and the estimation of the amount Dennis has had to pay, this seems very much like a defense of what

Dennis has done or an explanation of why he did it. But what might lead Mark to respond to Kim's turn in this way? What makes such a response appropriate? Although it is apparently not initially heard this way, the turn 'So he wrote a complaint?' might be interpreted by Mark as conveying that Kim is surprised that this is what Dennis did. This hearing could be encouraged by an increase in the prosodic register of Kim's talk. The differences here are quantitative, meaning there is no unequivocal boundary between a simple, 'unmarked' prosody and one that conveys surprise and from there possibly conveying that what Dennis has done is an overreaction.

With that in mind, notice that Mark continues his response after a substantial 4 seconds of silence at line 16. While 'So he wrote a complaint?' may be heard initially as a simple formulation of upshot for confirmation, neutral with respect to the stance Kim is adopting, the 4 seconds of silence that follow Mark's confirmation provide a new context in which to see it and can lead him to revise his understanding of what Kim's move was doing.

Our analysis suggests that action in interaction is understood by participants via 'multiple drafts'.[21] As new information becomes available, a sign will take on a particular significance and the 'identity' of an action may be transformed.[22] Mark's response evolves in the light of a continuous flow of information from the agent whose action he is engaged with – Kim. Specifically, the silence that follows his confirmation may cast the turn being confirmed as challenging. The understanding of what was conveyed by the turn's design – a combination of lexicosemantics, grammatical construction and prosody – may change.

Let us summarize our discussion of this case. First, an action is constructed from components. It is compositional. Prosody, words, grammar and the body simultaneously contribute to the action import of a turn-at-talk. Second, an action is positioned in various ways. It is positioned in relation to any prior turn (overlapping, latched, after delay, etc.), in relation to a sequence of actions, in relation to some ongoing activity and overall structure. It is also positioned relative to what comes next (e.g. in relation to the silence that follows: 'So he wrote a complaint?'). Third, action emerges in time and has temporal structure. A particular component can project more to come. A component can be recognizable as just the first piece, and so on. Fourth, action is subject to multiple, successive interpretations or understandings or drafts. Fifth, action is subject to multiple possible descriptions simultaneously: under one description it continues the prior turn, under another it initiates repair, under another it affiliates with the displayed stance of another participant, and so on.

We argue that a hearer does not need to – consciously or otherwise – classify an action as being of a certain type in order to know how to deal with it. A hearer uses the components of a move to figure out, on the fly, what would be an appropriate token response to that token move. Certain token actions will share features with many others, and so are readily interpreted by means of heuristics that have been successfully applied before (Gigerenzer et al., 2011; Schelling, 1960), hence the intuition that there are familiar action categories such as request, complaint and noticing. But if actions can be dealt with at the token level, they need not be seen as tokens *of types* at all. The taxonomy of action types is fundamentally metasemiotic in nature. The idea that there is a list of possible actions is a construct.[23] A radical version of our

claim would be that there are no 'actions' in the singular, only 'action' as a non-enumerable phenomenon. To analyze interaction using action category labels might be a descriptive convenience, but there is no reason to think that this is what participants in interaction actually do.

Our account therefore contrasts with those which suggest that the orderliness of interaction depends on the capacity of participants to take a particular bit of conduct (e.g. the utterance 'That's a really nice jacket') and assign it to some particular action category – for example 'compliment'. For this binning approach, the central problem involves participants sorting the stream of interactional conduct into appropriate categories or bins. The mental operation might be translated as 'that's a compliment' (Levinson, 2013), or in subtler versions 'that's a possible compliment' (Schegloff, 2006). These accounts appear to involve a presumption about the psychological reality of action types that is akin to the psychological reality of phonemes (Dresher, 1995; Sapir, 1933). For the binning account to be correct, there must be an inventory of actions just as there is a limited set of phonemes in a language. Each token bit of conduct goes into an appropriate pre-existing action-type category. The binning approach also suggests that it would be reasonable to ask how many actions there are. But to ask how many actions there are is like asking how many sentences there are. Here lies the key to our account: action is generated out of constituent elements of context and code that we already need in order to infer utterance meanings, and to produce appropriate responses. There is no need to add another level, conscious or otherwise, at which action categorization is done.

Indeed, our approach predicts that it will seldom be easy to label an action, while it will usually be straightforward to know what an appropriate next move would be. These two things are not the same. Interactants are far from hopeless in knowing how to respond. Our point is that they can know how to respond quite reliably by working up from turn components and finding a token solution, and not by having to assign the whole turn to an action category.

To name an action is to thematize participants' accountability

We have argued that an interactant is able to produce an appropriate response to another's move in interaction by building a token understanding of what a person is doing by what they are saying, based on features of the move's design and positioning. We have further suggested that there is no analytical need to propose that interactants categorize actions by binning token moves as this or that action type. Now we want to argue that the explicit labeling of an action not only adds nothing that is needed, but that it would often add something that is not needed, namely, a specific claim about the accountability at hand. To name an action is to cast a move under a description (as an alternative to other possible descriptions), and this is a way of thematizing specific things for which a participant is claimed to be accountable (Sidnell, 2017).

Let us look at a case in which – unlike those in the previous section – the participants are concerned with categorizing moves in the interaction by explicitly labeling them in action terms. Our point will be that these descriptions do not identify or define the

actions; rather, they make claims about others' conduct in order to thematize a specific accountability for the participants in question.

In the following excerpt from a family dinner, a father, a mother, and their daughter aged around 10 are having an evening meal at home in Southern California:

[Field trip]
01 Dad: So Ci:n (0.2) tell me abou your day.
02 (0.5)
03 Cin: Uh:: .h
04 Dad: Wha' dju (d) learn.
05 (1.0)
06 Dad: [Oh^::H yeah we went to thuh- we went to uh: (.)
07 Cin: [Uh:m-
08 Cin: Claim Jumper.
09 Dad: Claim Jum[per oday.
10 Mom: [>Uh huh<
11 Mom: May I have a roll [please,
12 Cin: [For uh field rip.=
13 Dad: =Sure.
14 Dad: An' may I have thuh -butter please.
15 Mom: Yes.=.hh
16 (0.5)
17 Cin: Went to Claim Jumper for our field trip.=
18 Dad: Yiea:h, an'- an'- tell me about it.
19 (0.5)
20 Mom: °°Cindy wantuh roll, °°
21 (1.0)
22 Cin: Mm[::_
23 Mom: [It's soft, it's good,
24 Dad: Come on,
25 Cin: Ye[ah.
26 Dad: [Describe this thing to me,
27 Cin: Uh:m_ .h It was fu:n?,
28 (0.3)
29 Dad: No.
30 (.)
31 Dad: You're gonna h(h)afta do uh lot [better than that.]
32 Mom: [Well she's-] she'll s
33 she's-
34 Dad: Hey [shh::
35 Mom: [Just starting. [°She's just starting.°
36 Dad: [Don't' defend 'er:,
37 Dad: [I wanna hear this from °her.°
38 Mom: [((mumble something at high pitch))
39 Mom: °She'll do fine,°
40 (0.5)
41 Cin: Uhm (2.0) what's this?,
42 (1.0)
43 Mom: Uh garbanzo bean.

We begin when the father selects his daughter to 'tell (him) about' her day (line 01). By using the phrase 'tell me about', the father is naming a kind of action: 'to tell someone about something'. With this, the father thematizes a normative set of practices that would constitute an instance of the named action. He is not describing an action that has taken place, rather he is directing the daughter to carry out the action he is describing. By a combination of factors, the daughter, as a participant in the joint activity of conversation at dinner, is now obliged to do something that could be defended as an action of the kind described. For one thing, she is the recipient of a directive, linguistically formatted in the imperative. For another, the family is engaged in a social activity – that of having an evening meal together – which by virtue of the cultural context, identifies this action, specifically with reference to what happened during the day leading up to the evening meal, as one that the daughter should recognize. And third, he is her father, which has obvious implications for the power relationship between the two (e.g. such that the daughter is unlikely to be able to make the same directive to her father, or at least not in the same terms). So, with this line the father obliges the daughter to act in a certain way, thus prospectively holding her accountable for producing a swatch of conduct that defensibly fits his action description.

The father quickly follows up his directive 'tell me abou your day' with a question 'What'dju . . . learn' (line 04), narrowing the scope of what the daughter might say in her projected telling. After a second's silence the father continues to narrow the scope of what the daughter could do. With line 06, 'Oh, ::H yeah we went to thuh', the father 'does remembering' what the daughter was going to do that day, transposing himself as subject of this remembering. Before he produces the name of the place the daughter was going, she completes his turn (line 08), saying 'Claim Jumper', a restaurant that she visited on a class excursion that day.[24]

At this point, the mother inserts a request for a roll, immediately followed by the father's request for butter (see examples 1 and 2, earlier), and then the daughter picks up the topic, saying 'Went to Claim Jumper for our field trip' (line 17). What is the daughter doing with this assertion? It could be that she is fulfilling the obligation that the father gave her by saying 'tell me about your day' in line 01. The father immediately discounts this possibility by saying 'Yiea:h an' . . . tell me about it' in line 18, reacting to the assertion ('yeah') and then re-issuing the directive 'tell me about it'. Again, there is brief talk relating to passing rolls, and the father resumes with 'Describe this thing to me' (line 26). He has now reiterated his directive using two action descriptions: 'tell me about it' and 'describe it to me'. By doing so, he is showing that nothing that the daughter has produced so far is accepted as an instance of an action fitting the description. These descriptions are not merely seeking to categorize the daughter's talk, but are being used in the context of holding her accountable for the appropriateness and effectiveness of her conduct.

The daughter then moves from making an objective assertion about what happened to taking an evaluative stance, stating 'Uh:m . . . It was fun:?' in line 27. There are two notable aspects of her turn design here. First is the interjection *um* (Clark and Fox Tree, 2002; Schegloff, 2010). While interjections such as *um* and *uh* appear in many sequential contexts, this one resembles those cases in which such a form is used to preface an estimate or a guess. And although she is producing a subjective assessment of her own

experience, it is nevertheless delivered as a 'guess'. Another notable aspect is the try marking – a distinctive type of pitch rise in turn-final position, typically used for mobilizing a confirmation of recognition from the other (Sacks and Schegloff, 1979) – on 'fun'. She is not only stating that it was fun, but she is at that moment mobilizing a response from the father (Stivers and Rossano, 2010). A try-marker usually adds an implicit question, canonically occurring on a recognitional reference such as a name. The name assumes 'you know who/what/where I'm talking about', and the try-marker adds a question – 'Can you confirm that you recognize this reference?' – thereby making a confirming response relevant next. Line 27 – 'Uh:m . . . It was fu:n? – differs from the canonical pattern in lacking a recognitional reference. Here, the try marking appears to convey, together with the 'estimating' *um*, 'Is this adequate?'.

But this conflicts with the projected action of 'telling someone about something', which normatively requires a multi-unit turn, and this is what the daughter will be held accountable for. The daughter at this point has used a single turn construction unit (TCU), and is hearable as asking for confirmation that she is doing, or has done, the required thing. Giving evidence in support of the idea that the try-marker is asking a yes–no question, the father's first move (line 29) is to say 'No', responding to the implicit question in the try-marker. This is followed up by a clear rebuke in line 31: 'You're gonna h(h)afta do uh lot better than that'. This is a clear case in which a person is held accountable for their conduct in terms of its failure to fit with a verbal description of the desired behavior in action terms.[25]

To summarize, the father's rebuke makes sense if we interpret the daughter as offering 'Uh:m . . . It was fu:n?' as a possible instance of 'telling someone about something'. The daughter is held to account for her failure to produce an action of the described kind. This shows the power of an explicit action description and, we argue, it shows its very purpose in social interaction.

What happens next gives further support to this view. The mother says 'She's just starting' (beginning in overlap in line 32, and emerging in lines 33 and 35). This proposes an alternative action description of the daughter's turn from the father's original in line 01. By characterizing the line as 'just starting', the mother offers a charitable interpretation, implying that the daughter is in fact embarking on producing an action of just the requested kin.[26] If indeed the daughter is just starting, then the father's rebuke in line 31 is not justified. What can we say about the mother's action here?

There are two ways of answering. The father helps us with one way by naming the mother's action (line 36), attempting to hold her accountable for what she has just done: 'Don't *defend* her'. Was the mother defending the daughter? A binning analysis would suggest that there is a correct answer to this question, but in our view this question is not about defining action but about thematizing accountability. This is necessarily selective. Which is the right formulation to select? We might say, as the father does, that the mother was defending the daughter. Or we might say that she was rebuking the father for an uncharitable interpretation of the daughter's move 'Uh:m . . . It was fun?'. Indeed, the father's choice of the word 'defend' may imply that she had been under 'attack' from the father (an interesting collateral effect of the father's word choice).

But of course what the father *calls* the action does not tell us what the mother was doing. Our point is that by calling it something, he is making a claim to what the mother

can be held accountable for having done. This is what analysts are doing when they bin moves as nameable actions. Sometimes this may be justified (see the 'requests' in the second section), but there is seldom any basis for doing so – or for doing so in a way that claims to be definitive/exhaustive. In the previous section we concentrated on moves that have straightforward responses, but that are not straightforwardly nameable and do not imply straightforward accountability.

It is impossible to say, definitively, what singular action 'She's just starting' implements. Whatever action it accomplishes is bound up with the details of its context. No description can ever do more than gloss it. It is useful, however, to consider what the options are. Searle (1969), for instance, seems committed to describing it as an assertion (of belief). Schegloff (1996) might advise a 'praction' analysis in which a collection of like cases are brought together in order to reveal a contingent mapping between composition, position and function.[27] Levinson might conclude that whatever action this implements, it will be included in an inventory or 'bag' of actions.[28] Our own account insists that there are always alternate possibilities. There could have been other analyses than that which 'Don't defend her' implies. There is no single, unequivocal, definitive analysis for the father to arrive at. Rather, he chooses the analysis – that the mother was 'defending' the daughter in lines 32–35 – which fits his purposes. While this may seem obvious, it raises a less obvious challenge for analysts: What are our purposes, such that we may decide what aspects of a participant's accountability to thematize?

Conclusion

The conceptual problem of action in interaction has long exercised philosophers of language, and action is increasingly viewed in conversation analysis as a domain of order alongside turn taking, repair, and sequence organization. To define and understand action in interaction, a valid account must both (1) handle the empirical particulars of the languages and cultures of any given data and (2) explain how those local particulars are instances of a species-wide domain of order, with predictions for how variation in human interaction is constrained.[29] In conclusion, we want to comment on the implications of our account of (1), which has been the focus of this article, for our account of (2).[30]

One path to a universal theory of action begins with some global macro-categories for actions, which any action will be an instance of. The psychologist Michael Tomasello (2008) proposes three macro-actions in interaction, grounded in empirical findings of comparative behavioral psychology. These actions are requesting (to secure another's help for meeting your own goals), helping (to provide help for meeting another's goals), and sharing (for meeting common goals, and for the cohesion of social relationships). Tomasello argues that these macro-actions require distinct social motivations and underlying cognitive capacities that are both universal and unique in our species. He leaves implicit how these three categories are instantiated by token actions of the kind we discuss earlier.

The philosopher John Searle (1969, 2010) takes a different approach. He establishes macro-categories of action based on four distinct types of mental state, which he takes as prior: assertives (corresponding to beliefs), directives (corresponding to desires), commissives (corresponding to intentions) and expressives (corresponding to feelings). From

basic intentional states (Searle, 1983) grounded in individual human cognition, Searle derives general communicative actions.[31]

These two theories offer an exhaustive inventory of action types, or action elements, that are proposed to exist universally. They do not link the macro-action categories to actions at the more specific level of token cases in data from conversation. Accounts of action that begin with conversational data work with categories of action that are much more specific in nature, and greater in number, than the macro-action categories suggested by Tomasello and Searle. For example, the praction approach in conversation analysis (see Schegloff, 1996, 1997) begins with maximally specific characterizations of action types based solely on repeatedly observed occurrences of defined patterns of practices and their normative effects. Schegloff makes no claim that there is a universal or finite inventory of action types.[32] Levinson (2013, 2017) has, however, recently taken a step in this direction, asking what 'the inventory of action types' looks like. He states that with our current state of knowledge we have 'over a hundred in the bag' (Levinson, 2013: 122), and implies that the finite set required by a binning conception of action in interaction (at least for a putative first order of action) is unlikely to be orders greater than this.[33]

Our view is that no distinct set of action types, whether general or specific, is necessary or sufficient for thinking about action in interaction. We acknowledge that existing approaches have important virtues, methodological and otherwise. Action labels serve a useful role in analysis, and in scholarly communication. We have argued that a binning analysis – sorting the stream of conduct into discrete action types – has a role, but this role does not provide a general account for action in interaction. Categorization of actions is relevant to a categorizer's goal of thematizing accountability (see our examples 'tell me about your day', 'She's just starting' and 'Don't defend her'). This insight requires that we attend to the reflexive quality of language, something that Linguistic Anthropology has dealt with more than Conversation Analysis. Accountability is the glue that holds interaction together. For a theory of action to be built on this idea, we need to acknowledge that participants work on the assumption that people are pursuing goals, and they seek to ascribe reasons behind others' behavior that make sense in terms of those likely goals.

If there are no types of action independent from the observed tokens of goal-directed conduct, then is it possible to compare actions across cultures? One way to study action comparatively is to focus on the social problems that people need to solve (Schegloff, 2006a; Sidnell and Enfield, 2012). Some of those are likely to be universal, such as initiating repair (Dingemanse et al., 2015) or recruiting others' assistance (Floyd et al., 2014), and others are likely to be culture specific. These problems in need of interactional solutions are not, of course, actions themselves; rather, they are the ends that means are contrived to meet. Actions are not problems that need solving. They are the means-for-ends solutions to those problems. From an interpreter's perspective, the task is to work from the means – the practices used – to infer the ends. In addition, actions are bound up with the uptakes that they elicit and that in turn help consummate actions, given that action in interaction is, ultimately and necessarily, distributed and collaborative in nature (Enfield and Kockelman, 2017).

We conclude that participants in interaction do not need to recognize action types or categories in order to respond appropriately. Other theories that have addressed the problem of action ascription seem to assume that such ascription involves assigning each bit

of conduct to an action-type category, essentially a process of labeling. Our view is that when we are analyzing moves in the flow of interaction we are not labeling or otherwise identifying actions; rather, we are considering the details of turns-at-talk, under the assumption that the speaker's conduct is purposive and goal directed, for their relevance in deciding what to do next and how to do it. And all of their details are potentially relevant to what people will be understood to be doing. Action category labels are convenient heuristics, but they are ultimately neither necessary nor sufficient. Category labels cannot be a substitute for practice-based analysis of situated social action.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work is supported by a Faculty Collaborative Research Scheme grant (`Power and Accountability') from the University of Sydney.

Notes

1. 'Talk' includes semantic, grammatical and other potentially meaning-bearing features of language, as well as gesture and embodied conduct.
2. Axiom 1 is the starting point for speech act theory and much of the literature in philosophical pragmatics.
3. The response has to be 'allowed through' if it is going to help determine what the prior action was. In third position repair a speaker – post-response – indicates they were not doing the thing the recipient has taken them to be doing (Schegloff, 1992; Sidnell, 2017). This supports our point because it shows the power of the response to cast the prior as an action of a certain kind. This power needs constant monitoring throughout interaction.
4. Axiom 2 is the starting point for conversation analysis and much of the literature on social interaction.
5. This does not mean that a recipient always produces an 'appropriate' response, even when they know what this might be. In some cases they purposefully do not. That said, for a person to knowingly produce something other than an appropriate response, for a certain effect, entails that they act in accordance with their knowledge, at some level, of what an appropriate response would be.
6. When certain things happen in interaction, we are not surprised that they happen, but we do not want to say that we had actively *anticipated* or *expected* those things to happen. For this tacit awareness of what is normatively or culturally possible next, we use the term *subprehension* (Enfield, 2013: 222, fn. 28).
7. This is why we ultimately do not accept the practions model of action (Enfield, 2013: 100; Schegloff, 1992: 100), because it assumes a fixed connection between 'set of practices X' and 'action Y'. That said, we recognize that many form–action associations are frequent and entrenched.
8. Our semiotic account begins not with senders/speakers and thus not with intention, but with receivers/hearers and thus with interpretation. It is a recipient driven system (Enfield, 2013: 16–17; Krebs and Dawkins, 1984).
9. Overt labeling of actions by participants in interaction is not uncommon in cases of third position repair ('I wasn't criticizing, I was asking').

10. The distinction between 'treating as' and 'describing as' may point to two ends of a continuum, with a possible middle ground. Consider the case (Sidnell, 2017) in which person B's delivery truck is pointed in the wrong direction on a narrow one-way street. A: *You know this is one way right?* B: *I've got a permit so relax*. B's response seems to imply that A's turn was an accusation of sorts, but they do not call it anything or describe it as such. B does, however, implicate that A's initial turn was 'not relaxed', and in this way B is holding A accountable for a possible transgression. There might be action-implicating manners of formulation. Consider the insertion of curse words. *There's no soy milk in the fridge* might be a request or a complaint, or even a noticing (e.g. if we are looking for evidence as to whether a person known to drink soy milk is currently staying in the house), while *There's no fucking soy milk in the fridge* is less likely to be read as a noticing in the last context. Or consider also, A: *There's no soy milk in the fridge*, B: *That's true.* When B describes the prior turn's status as 'true', they effectively classify it as an assertion, thus not a request (a request cannot be true or false). While examples like these do not explicitly label a prior action, they nevertheless use words that imply a specific, nameable characterization.
11. If, for instance, someone says, 'this bread is dry' (thereby doing something describable as 'complaining'), another may pass the butter.
12. This has obvious and important consequences for our understanding of sequence organization (see, inter alia, Schegloff, 2007). Conversation analysts often describe sequences by reference to the first pair part of the base adjacency pair. So, for instance, we talk of 'invitation sequences', 'directive-response sequences', 'advice sequences', and so on. This is an entirely reasonable way of talking for heuristic purposes, but can easily be taken to suggest that the whole of sequence organization is anchored by the occurrence of discrete, binnable actions. Our proposal is that sequences are organized around participants' assumptions about the goal-directed, purposive character of one another's conduct. Action labels are convenient heuristic labels for talking about this purposive character of human conduct, but they do not, we argue, reflect anything about the underlying ontology of action. Of course, participants do sometimes actually name the action they are purportedly engaged in performing (e.g. 'I'd like to invite you to dinner') but, as it turns out, this does not guarantee that this is the action they will be understood to have done (see Enfield and Sidnell, 2017 for further discussion and examples).
13. See Heritage, 1984; Robinson, 2016; Sidnell, 2017.
14. Whether by 'recognizing' or 'ascribing'.
15. And issues can arise even with examples that are structurally similar to the ones above, such as: A: *May I have the butter please*, B: ((passes butter)). Was a question consummated here? Based on distributional evidence (i.e. numerous other cases in which a recipient says 'yes' or 'sure'), many analysts would treat this as two actions together (question+request), where the question is not overtly oriented to by the recipient. Other analysts have argued that people do not need to process the 'literal meaning' of conventional indirect requests like these (Gibbs, 1983). But what ultimately matters is what people can be held accountable for having done. This is relatively straightforward when you have a polar interrogative: it would be hard for the speaker to deny 'having asked a question' in addition to 'making a request'. But with other forms such as declaratives (see Note 10), accountability is more fluid and dependent on compositionality (language, body, context). (We thank Giovanni Rossi for raising these points with us.)
16. With respect, that is, to the problem of conceptualizing the nature of action in interaction. Such examples surely have much to tell us about other matters (such as e.g. the relevance of different formats to the design of action).
17. This concept of action (one in which the stream of conduct must be sorted into 'bins' in order to be understood at the level of 'action') is perhaps best illustrated by the work of Stephen C. Levinson, especially his article in the *Handbook of Conversation Analysis* (Levinson, 2013). There are, however, many other examples of analyses that presume such a model. The piece

by Levinson is a particularly useful point of reference in so far as he tends to articulate his underlying assumptions somewhat more explicitly than others. That said, we quite intentionally refrain from suggesting that the binning approach characterizes *all* work in conversation analysis (or any other approach to social interaction for that matter). Schegloff has made very similar points to the ones we are making in this article (in the course of his lectures, for instance) and has often pointed to the methodological implications by warning students against basing an analysis on a presumed inventory of action types (suggesting instead that analysis should begin with 'practices'). Indeed, Schegloff (2009) proposes an approach quite similar to ours in an important but as yet unpublished paper titled 'Prolegomena to an analysis of action'.

18. This is because action ascription by labeling is an act of categorization, and categorization of the same objective entity can be made in a large range of ways. This is true for any named phenomena. Take Fido: Is he a dog? A pet? A mammal? All are true descriptions.
19. See the discussion of 'appositionals' in Sacks et al. (1974: 719). See also Clark and Fox Tree, 2002; Local and Walker, 2012; Schegloff, 1981; Walker, 2012.
20. This points to an important conceptual problem with the binning approach. On a binning account, Kim's glance must of course be binned. But as what? If it does not bin, then it is not an action. If the machinery of action involves binning each possible action into one bin from the inventory, then how do we deal with very lean 'actions' like a gaze shift? Or even more problematic: What about silence? For example, after a request? Or a greeting? Is this also to be binned? If not, are we thereby committed to saying that it is not an action?
21. This notion of 'multiple drafts' is adapted from an argument by Dennett (1991) about perception and is akin to what Garfinkel called 'the documentary method of interpretation'.
22. Giovanni Rossi points out to us the parallel here with incrementality in language production/comprehension (e.g. Levelt, 1989).
23. The 'list' is composed partly of vernacular action labels. Ordinary people also categorize actions, but they do not typically compose lists/inventories and have little need for them in any case.
24. https://en.wikipedia.org/wiki/Claim_Jumper
25. This is the kind of rebuke that Gilbert (1992) describes in her account of joint action. Joint action entails mutual commitment, and this commitment has a moral dimension: if one person fails to keep up their part of the commitment, the other is entitled to 'rebuke' them. In Gilbert's example of going for a walk together, if one person walks faster than the other and draws ahead, the first person may be entitled to say something like 'I can't keep up, you are going to have to slow down!'. Conversation is joint action in this sense (Clark, 1996; Enfield, 2017).
26. Mother's 'She's just starting' also suggests an analysis of the daughter's use of try marking to check a recipient's understanding with minimal disruption to the progress of her telling. In this respect, it is perhaps important to note that 'try marking' can elicit recipient responses within a single TCU and before it reaches possible completion (see Sidnell, 2010).
27. On 'practions' in action analysis, see Enfield (2013: 100).
28. Following a binning approach, Levinson (2013: 122) raises the question of an inventory of actions, and asks how extensive it might be.
29. See Sidnell, 2007, 2009; Dingemanse and Floyd, 2014; Enfield, 2017; Sidnell and Enfield, 2012.
30. On action in relation to human diversity, see Part III of *The Concept of Action* (Enfield and Sidnell, 2017).
31. Searle (2010) has a fifth type of speech act – the declarative – that is not based on a basic psychological state, unless perhaps imagining.
32. While Schegloff (2006a) makes several proposals as to what is likely to universally underlie the organization of human interaction, he does not claim that any specific actions, or general action types, are universal, or otherwise serve as a universal comparative framework for understanding action in human interaction. What Schegloff lists are universal problems that

people need to solve in interaction (the turn-taking problem, the repair problem, etc.) – that is, interactional *ends* that people universally need to devise means for.

33. Levinson (2013) suggests that there is a distinction between 'primary' actions and more nuanced, 'optional' actions, implying that only the former are to be counted in the proposed inventory of actions, and that the latter are somehow derived in the context of overarching activities in which actions are situated.

References

Austin JL (1962) *How to Do Things with Words*. Cambridge, MA: Harvard University Press.

Bolden GB (2006) Little words that matter: Discourse markers 'So' and 'Oh' and the doing of other-attentiveness in social interaction. *Journal of Communication* 56: 661–688.

Clark HH (1996) *Using Language*. Cambridge: Cambridge University Press.

Clark HH and Fox Tree JE (2002) Using uh and um in spontaneous speaking. *Cognition* 84: 73–111.

Couper-Kuhlen E (2014) What does grammar tell us about action? *Pragmatics* 24(3): 623–647.

Curl TS and Drew P (2008) Contingency and action: A comparison of two forms of requesting. *Research on Language and Social Interaction* 41: 129–153.

Dingemanse M and Floyd S (2014) Conversation across cultures. In: Enfield NJ, Kockelman P and Sidnell J (eds) *Cambridge Handbook of Linguistic Anthropology*. Cambridge: Cambridge University Press, pp. 434–464.

Dennett DC (1991) *Consciousness Explained*. Boston, MA: Little, Brown and Company.

Dingemanse M, Roberts SG, Baranova J, et al. (2015) Universal principles in the repair of communication problems. *PLoS ONE* 10(9): e013610.

Dingemanse M, Rossi G and Floyd S (2017) Place reference in story beginnings: A cross-linguistic study of narrative and interactional affordances. *Language in Society* 46(2): 129–158.

Dresher E (1995) There's no reality like psychological reality. *Glot International 1*(1): 7.

Enfield NJ (2013) *Relationship Thinking: Agency, Enchrony, and Human Sociality*. New York: Oxford University Press.

Enfield NJ (2017) *How We Talk: The Inner Workings of Conversation*. New York: Basic Books.

Enfield NJ and Kockelman P (eds) (2017) *Distributed Agency*. New York: Oxford University Press.

Enfield NJ and Sidnell J (2017) *The Concept of Action*. Cambridge: Cambridge University Press.

Floyd S, Rossi G, Baranova J, et al. (2014) Recruitments across languages: A systematic comparison. Paper presented at the Fourth International Conference on Conversation Analysis, University of California at Los Angeles, CA.

Garfinkel H (1952) *The Perception of the Other: A Study in Social Order*. Cambridge, MA: Harvard University Press.

Gibbs RW (1983) Do people always process the literal meanings of indirect requests? *Journal of Experimental Psychology: Learning, Memory, and Cognition* 9(3): 524–533.

Gigerenzer G, Hertwig R and Pachur T (eds) (2011) *Heuristics: The Foundations of Adaptive Behavior*. New York: Oxford University Press.

Gilbert M (1992) *On Social Facts*. Princeton, NJ: Princeton University Press.

Goodwin C (1979) The interactive construction of a sentence in natural conversation. In: Psathas G (ed.) *Everyday Language: Studies in Ethnomethodology*. New York: Irvington Publishers, pp. 97–121.

Heinemann T (2006) "Will you or can't you?": Displaying entitlement in interrogative requests, *Journal of Pragmatics* 38: 1081–1104.

Heritage J (1984) *Garfinkel and Ethnomethodology*. Cambridge: Polity Press.

Heritage J (2012) Epistemics in action: Action formation and territories of knowledge. *Research on Language and Social Interaction* 45: 1–29.
Jakobson R (1957) Shifters, verbal categories, and the Russian verb. In: Jakobson R(ed.) *Selected Writings*, Vol II. The Hague: Mouton, pp. 130–147.
Jefferson G (1978) Sequential aspects of storytelling in conversation. In: Schenkein J (ed.) *Studies in the Organization of Conversational Interaction*. New York: Academic Press, pp. 219–248.
Local J and Walker G (2012) How phonetic features project more talk. *Journal of the International Phonetic Association* 42(3): 255–280.
Kockelman P (2005) The semiotic stance. *Semiotica* 157 (1–4): 233–304.
Krebs JR and Dawkins R (1984) Animal signals: Mind-reading and manipulation. In: Krebs JR and Davies NB (eds) *Behavioural Ecology: An Evolutionary Approach*, 2nd edn. London: Blackwell Publishing, pp. 380–405.
Levelt WJM (1989) *Speaking: From Intention to Articulation*. Cambridge, MA: The MIT Press.
Levinson SC (2013) Action formation and ascription. In: Sidnell J and Stivers T (eds) *The Handbook of Conversation Analysis*. Malden, MA: Wiley-Blackwell, pp. 103–130.
Levinson SC (2017) Speech acts. In: Huang Y (ed.) *The Oxford Handbook of Pragmatics*. New York: Oxford University Press, pp. 199–216.
Raymond G (2004) Prompting action: The stand-alone 'so' in sequences of talk-in-interaction. *Research on Language and Social Interaction* 37(2): 185–218.
Raymond G and Lerner GH (2014) A body and its involvements: Adjusting action for dual involvements. In: Haddington P, Keisanen T, Mondada L, et al. (eds) *Multiactivity in Social Interaction: Beyond Multitasking*. Amsterdam; Philadelphia, PA: John Benjamins, pp. 227–245.
Robinson J (2016) Accountability in social interaction. In: Robinson J (ed.) *Accountability in Social Interaction*. New York: Oxford University Press , pp. 1–44.
Rossano F (2012) *Gaze behavior in face-to-face interaction*. PhD Thesis, Radboud University, Nijmegen.
Rossi G (2015) *The request system in Italian interaction*. PhD Thesis, Radboud University, Nijmegen.
Rossi G and Zinken J (2016) Grammar and social agency: The pragmatics of impersonal deontic statements. *Language* 92(4): e296–e325.
Rumsey A (1990) Wording, meaning, and linguistic ideology. *American Anthropologist* 92(2): 346–361.
Sacks H (1995) *Lectures on Conversation*, vols 1 and 2. Oxford: Basil Blackwell.
Sacks H and Schegloff EA (1979) Two preferences in the organization of reference to persons and their interaction. In: Psathas G (ed.) *Everyday Language: Studies in Ethnomethodology*. New York: Irvington Publishers, pp. 15–21.
Sacks H, Schegloff EA and Jefferson G (1974) A simplest systematics for the organization of turn-taking for conversation. *Language* 50: 696–735.
Sapir E (1933) La Réalité Psychologique des Phonèmes [The Psychological Reality of Phonemes]. *Journal de Psychologie Normale et Pathalogique (Paris)* 30: 247–265.
Schegloff EA (1981) Discourse as an interactional achievement: Some uses of 'uh huh' and other things that come between sentences. *Georgetown University Round Table on Languages and Linguistics* 71: 71–93.
Schegloff EA (1982) Discourse as an interactional achievement: Some uses of "uh huh" and other things that come between sentences. In: Tannen D (ed.) *Georgetown University Round Table on Languages and Linguistics 1981: Analyzing Discourse: Text and Talk*. Washington, DC: Georgetown University Press, pp. 71–93.
Schegloff EA (1988) Description in the social sciences I: Talk-in-interaction. *IPRA Papers in Pragmatics* 2(1): 1–24.
Schegloff EA (1992) Repair after next turn: The last structurally provided defense of intersubjectivity in conversation. *American Journal of Sociology* 97(5): 1295–1345.
Schegloff EA (1996) Confirming allusions: Toward an empirical account of action. *American Journal of Sociology* 102(1): 161–216.

Schegloff EA (1997) Practices and actions: Boundary cases of other-initiated repair. *Discourse Processes* 23(3): 499–545.
Schegloff EA (2006a) Interaction: The infrastructure for social institutions, the natural ecological niche for language, and the arena in which culture is enacted. In: Enfield NJ and Levinson SC (eds) *Roots of Human Sociality: Culture, Cognition, and Interaction*. Oxford: Berg Publishers, pp. 70–96.
Schegloff EA (2006b) On possibles. *Discourse Studies* 8(1): 141–157.
Schegloff EA (2007) *Sequence Organization in Interaction: A Primer in Conversation Analysis*. Cambridge: Cambridge University Press.
Schegloff EA (2009) Prolegomena to the Analysis of Action(s) in Talk-in-interaction. In: *Proceedings of the 15th Annual Conference on Language, Interaction, and Social Organization*, University of California, Santa Barbara, CA, 14–16 May.
Schegloff EA (2010) Some other u(h)ms. *Discourse Processes* 47(2): 130–174.
Schegloff EA and Sacks H (1973) Opening up closings. *Semiotica* 8: 289–327.
Schelling TC (1960) *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.
Searle JR (1969) *Speech Acts*. Cambridge: Cambridge University Press.
Searle JR (1983) *Intentionality*. Cambridge: Cambridge University Press.
Searle JR (2010) *Making the Social World*. Oxford: Oxford University Press.
Sidnell J (2009) Comparative perspectives in conversation analysis. In: Sidnell J (ed.) *Conversation Analysis: Comparative Perspectives*. Cambridge: Cambridge University Press, pp. 3–28.
Sidnell J (2010) *Conversation Analysis: An Introduction*. Oxford: Wiley-Blackwell.
Sidnell J (2007) Comparative studies in conversation analysis. *Annual Review of Anthropology* 36(1): 229–244. doi:10.1146/annurev.anthro.36.081406.094313.
Sidnell J (2017) Action in interaction is conduct under a description. *Language in Society* 46(3): 313–337.
Sidnell J and Enfield NJ (2012) Language diversity and social action: A third locus of linguistic relativity. *Current Anthropology* 53: 302–333.
Silverstein M (1979) Language structure and linguistic ideology. In: Clyne PR, Hanks WF and Hofbauer CL (eds) *The Elements: A Parasession on Linguistic Units and Levels*. Chicago, IL: Chicago Linguistic Society, pp. 193–247.
Stivers T (2008) Stance, alignment, and affiliation during storytelling: When nodding is a token of affiliation. *Research on Language and Social Interaction* 41(1): 31–57.
Stivers T and Rossano F (2010) Mobilizing response. *Research on Language and Social Interaction* 43(1): 3–31.
Tomasello M (2008) *Origins of Human Communication*. Cambridge, MA: The MIT Press.
Walker G (2012) Coordination and interpretation of vocal and visible resources: 'Trail-off' conjunctions. *Language and Speech* 55(1): 141–163.

Author biographies

NJ Enfield is Professor of Linguistics at the University of Sydney, and director of the Sydney Social Sciences and Humanities Advanced Research Centre (SSSHARC). His work on language and human sociality is based on regular fieldwork in mainland Southeast Asia, especially Laos. Among his more recent books are *Relationship Thinking* (2013), *The Utility of Meaning* (2015), and *How We Talk* (2017).

Jack Sidnell is Professor of Anthropology at the University of Toronto. He has conducted field research in the Caribbean, Vietnam, India and North America. The structures of social interaction have been the object of long term study. Other research has focused on the anthropology of knowledge and the ontology of action and he is particularly interested in the possibilities for dialogue between anthropology and philosophy.