

# Part II Fundamental Structures of Conversation

---

UNCORRECTED PROOF

Su

UNCORRECTED PROOF

Su

# 6 Action Formation and Ascription\*

---

STEPHEN C. LEVINSON

*Max Planck Institute for Psycholinguistics, Radboud University, Nijmegen*

*Talk is constructed and is attended by its recipients for the action or actions it may be doing.*  
(Schegloff, 1996a: 5)

## 1 Introduction

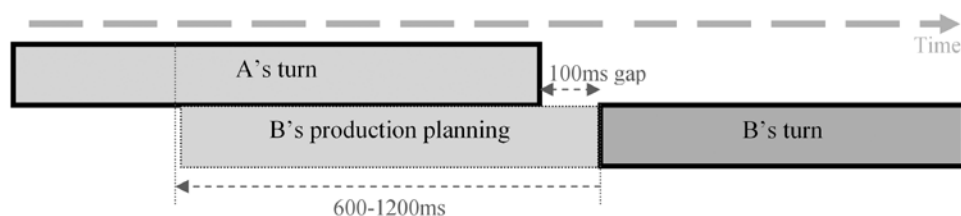
The ability to conduct conversation is a miracle in general, but particularly striking is the speed and apparent accuracy of action ascription. To see this, consider the following facts: it takes over 600ms to plan and execute the shortest turn-at-talk (see Levelt, 1989), while on average the gaps between turns are around 200ms, depending a bit on the language (de Ruiter, et al., 2006; Stivers, et al., 2009). As Figure 6.1 makes clear, this entails that B must plan his or her turn well before the prior speaker A's is finished. Of course, B's turn will mostly be tied to A's turn via sequence organization in particular: If A's turn was a question, B's turn is expectably an answer; if an offer, an acceptance or rejection is in order, and so forth (see Stivers, this volume, on sequence organization). So action ascription by B of A's turn is a prerequisite for the design of B's turn—the very 'proof procedure' that makes CA possible.<sup>1</sup> Although occasional misunderstandings of action content do occur, they are surprisingly rare.

The challenge for participants, then, is to assign at least one major action to a turn they have only heard part of so far. But to do this, they must have parsed what they have heard and understood its grammar well enough to predict both the content and its structure, so that they can predict when it will come to an end (otherwise their response may come too early or too late). Further, action

---

*The Handbook of Conversation Analysis*, First Edition. Edited by Jack Sidnell and Tanya Stivers.  
© 2012 Blackwell Publishing Ltd. Published 2012 by Blackwell Publishing Ltd.

Su



**Figure 6.1** Overlap of comprehension and production processes in conversation.

ascription involves numerous dimensions, as we will see, so it would seem to be a much more complex and indeterminate process than decoding the structure and content of the turn. That is the miracle, and trying to understand how it might be achieved is what this chapter is about.

First, a word about terminology. The standard terms in the CA literature are *action formation* and *action recognition*. *Action formation* is characterized by Schegloff (2007b: xiv) as follows:

The action formation problem: how are the resources of the language, the body, the environment of the interaction, and position *in* the interaction fashioned into conformations designed to be, and to be recognized by recipients as, particular actions—actions like requesting, inviting, granting, complaining, agreeing, telling, noticing, rejecting, and so on—in a class of unknown size?

The corresponding term for the recipient point of view, *action recognition* is potentially misleading, because it presupposes that actions have a correct identity, when actually the process of attributing an action to a turn is a fallible, negotiated, and even potentially ineffable process.<sup>2</sup> Consequently, the term used here is, for the most part, *action ascription*, by which is understood the assignment of an action to a turn as revealed by the response of a next speaker, which, if uncorrected in the following turn(s), becomes in some sense a joint ‘good enough’ understanding.<sup>3</sup>

Two factors in action ascription are clearly crucial, and they are the same factors that play a major role in much CA argumentation, namely turn design (see Drew, this volume, on turn design) and turn location (see Stivers, this volume, on sequence organization). There are other factors too, for example actions in other modalities, the context of ongoing activities (see Robinson, this volume, on overall structural organization), and the larger institutional framework and the social roles thus ascribed to participants (see Clayman, Gardner, and Komter, this volume, on interviewing, classrooms, and courtrooms, respectively). Here we will briefly review these other factors, concentrating on the two major factors in sections 4 and 6, but it will be useful first to sketch some of the background of the subject.

## 2 Background: Language As Action in Sociology and Other Disciplines

The topic of this chapter is in a number of ways clearly central to the CA enterprise, and there is a great deal of work on specific 'practices' for accomplishing interactional jobs which touches on it. Yet, direct empirical investigations in CA (e.g. under the rubric of 'action formation' or 'recognition') are few and far between. The result is unfortunate. For many of the other findings in CA rely on intuitive characterizations of the actions embodied in turns. For example, any application of the notions of adjacency pair, sequence or preference relies crucially on an identification of the character of, for example, a first turn and its second. But that identification is largely based on an appeal to our knowledge as societal 'members' or conversational practitioners. This loose hermeneutics is the soft underbelly of CA, and it is one of the reasons that other disciplines sometimes think of CA as a branch of the occult. This is not to deny there is some sterling work, reviewed in section 2; rather, this is merely to say that we do not have enough of it, and to exhort us collectively to further analysis.

Concern with *how* social actions are recognizable for what they are lies deep in the developments that led to Conversational Analysis, and can be traced from Weber through Parsons to Garfinkel (Heritage, 1984b; Heritage & Stivers, this volume; Parsons, 1937). In the context of his memorable analysis of the minimal story "The baby cried. The mommy picked it up,"<sup>4</sup> Sacks famously wrote (1995: 226) "a culture is an apparatus for generating *recognizable* actions." On many accounts, Sacks put forward two major frameworks of analysis: one, the membership categorization notion, and the other, sequential analysis, of which only the latter has been fully developed within CA (see Maynard, 2000, for example). The major development of sequential analysis at the expense of the other framework has been attributed to seeing how the larger frameworks may be constructed out of the "small parts" of sequential analysis (Schegloff, 1995b: xxxv). But this topic, the recognizability of actions, may reunite the two frameworks, in ways made clear below.

Meanwhile, other disciplines have quite a bit to say about action ascription and action typology. First, there is Philosophy. It is indeed a puzzle fit for a philosopher: causal ontology without physics. How on earth can wobbles in the air so fundamentally change the world, as when defendants are found guilty, wars declared, marital bonds created, institutions dissolved, boundaries agreed, banks declared bankrupt? It was Austin (1970a: 251) who said: "What we need besides the old doctrine about meanings is a new doctrine about all the possible forces of utterances," and he set about the lexicography of performative verbs (verbs that can appear in the frame *I hereby X*), thinking that would yield an initial characterization of action types (ordering, declaring, promising, etc.). Searle (1976) then systematized Austin's analysis and suggested, on the one hand, that speech acts could be individuated and characterized in terms of a set of four felicity conditions, and secondly that there were just five big families of types of speech

acts: representatives, directives, commissives, expressives, declarations.<sup>5</sup> Recently, he has gone on to suggest that the whole of Sociology can be recast in terms of an analysis of the latter (Searle, 2010). Linguistics rapidly got in on the act. Noting that the coding of action force was often indirect, one systematicity turned out to be that questioning or asserting a felicity condition on a speech act could work as a means of doing that act indirectly (Gordon & Lakoff, 1971), hence the *Could you* and *Would you* prefaces to indirect requests, not to mention the *I want* and *I'd like* versions. Motivations for deviating from direct expression could be spelled out in a theory of face-saving or politeness (Brown & Levinson, 1987). Some rapprochement of this line of work with CA was suggested in Levinson (1983), where, after a demolition of the discourse analytic treatment of actions as in Labov and Fanshel (1977), an analysis of indirect speech acts as *pre-s* was advanced. In addition, as CA moved (increasingly) into institutional settings, the sociology of activities began to be seen to constrain the attribution of actions in interesting ways (Atkinson & Drew, 1979; Levinson, 1979; Merritt, 1976). Institutional settings had the virtue of making plain the *action*-like component of language, as when language and nonlinguistic actions interdigitate seamlessly in service encounters or in the doctor's office.

Much of the follow-up work was done in the 1970s, when the topic of linguistic action was pursued in parallel in Linguistics, Psychology of Language, Developmental Psychology, and Artificial Intelligence or Natural Language Processing. In the latter case, interesting work suggested that conversation might be based on plan amalgamation—that is, that appropriate turns require an attempted reconstruction of the other's plans, so that plans could be assisted or resisted as the steps toward realizing them appeared. In that case, a turn comes not with a single paired action, but with a hierarchy of plans for which it is merely the tip of the iceberg (Allen & Litman, 1990; cf. also Labov & Fanshel, 1977).

There was then a lull before the subject reappeared in further thought about joint action (Bratman, 1987; Clark, 1996; Sebanz & Knoblich, 2008; Tomasello, 2008). Here actions are treated as contributions to a joint enterprise, requiring grounding in mutual assumptions and mental modeling of the other's expected contributions. Tomasello (2008) has also reinvigorated the questions about the typology of action types by suggesting that apes and infants share a primordial set of just three types: requesting, helping and sharing.

This (lamentably) brief survey suggests that the topic of action as expressed in language is of considerable interest to other disciplines, which have raised some quite fundamental questions of the kind: What is the underlying ontology of a verbal action? What is the inventory of social actions that can be performed through words? What is the origin of that inventory, and what constrains it? How does it fit in with the social institutions that mould less informal verbal interchange? Above all, how on earth are actions reliably attributed, namely actually recognized? Are we dealing with a mapping of action-to-utterance, or actually something much more complex like a reconstruction of the other's motives, with inevitable ineffability? Can there be a science or systematic investigation of this

crucial subject, a science which offers an account of the instrumentality of language?

### 3 Granularity of Action Description

First, a self-imposed restriction, without which I doubt there could be any systematic study of action. The particular sense of *action* being put central here is the ascription or assignment of a 'main job' that the turn is performing. The sense of 'main job' or primary action intended here is *what the response must deal with in order to count as an adequate next turn*. (A turn may perform two, perhaps three, 'main jobs' at once, but perhaps not often more—see section 12 below.) There is a much broader sense of 'how we do things with words' that has been explored by speech act theorists (e.g. Austin (1962), who distinguished locutionary, illocutionary and perlocutionary acts), sociologists (Goffman, 1976), and psychologists (Clark, 1996). These more elaborate schemes separate out the 'main overt business' (Austin's illocutionary level as opposed to the perlocutionary; Goffman's 'official business' as opposed to the 'ritual business'; Clark's 'track 1' vs. 'track 2') from the penumbra of less 'official' business that participants may intend to be recognized. For example, in answering you, I can hint that you should have known already (Stivers, 2011b), or in responding to an assessment indicate that actually I am the expert and not you (see Heritage & Raymond, 2005; Raymond & Heritage, 2006; and, more broadly, Heritage, this volume), or in being brusque in greetings indicate that I am annoyed with you. These are 'doings' which are usually 'off-record', and not easy to respond to directly without completely redirecting the talk. That does not mean that the design of the response will not reflect these other factors—it almost certainly will, producing an under-current of tit-for-tat, but these factors do not change the *nature of the sequential action type now due*, at least in the normal case. We will need to return to this issue after reviewing the role of sequential organization and turn format.

### 4 Sequence Organization and Action Ascription

A sequence is by definition "a course of action implemented through talk" (Schegloff, 2007b: 9). Sequences can have an elaborate many-turn course, but Schegloff argues that they can be thought of as built up out of elementary adjacency pairs, like question-answer, offer-acceptance, greeting-greeting, and so on. It is obvious that a first-pair part (colloquially, a 'first'), in projecting a matched second, maps onto the next turn an expectation of the action (or action alternatives) it ought to perform (Schegloff, 1968). Thus greetings project greetings in return, questions project answers, requests project compliances or rejections, invitations project acceptances or declinations, and so on. It is easy to imagine that the same utterance might have different actions mapped onto it by virtue of its location: *Well I have to be here till six* might, in principle, be an answer to a request

Su

for information (*When are you coming?*), a request refusal (after *Could you come over now?*), an invitation declining (*How about lunch together?*), and so forth. (In practice, there are likely to be small differences in turn design and timing, but the point is the words themselves are multi-action compatible.) One way of appreciating the power of firsts to map actions onto seconds is to consider how silence—a thing without any properties of its own—is differentially interpreted after different first-pair parts, as crudely glossed in italics below:<sup>6</sup>

## (1) Drew (1981: 249)

1 Mum: What's the time?  
 2 (3.0) ← *child does not know how to read the clock*  
 3 Now what number's that?

## (2) Schegloff (1979a: 37)

1 ((rings))  
 2 R: Hello  
 3 C: Hello  
 4 (0.2) ← *R cannot identify C's voice*  
 5 This is Yolk

## (3) Levinson (1983: 320)

1 C: I was wondering would you be in your office on Monday (.) by  
 2 any chance?  
 3 (2.0) ← *R will not answer positively*  
 4 C: Probably not

Given this projective power of adjacency pairs, the actions done in the second part slot are highly constrained. But even here, of course, options are available: apart from greetings and partings, alternate action types are usually available (e.g. request compliance vs. refusal). Here preference organization kicks in to aid recognition: usually any turn that is not designed to comply with the expectation of the first-pair part comes with an immediate delay, warning particles (*well, uh*), indirection, mitigation, excuses or other marking of deviation from the preferred alternate action (Pomerantz, 1984a and see Pomerantz & Heritage, this volume). In (4), for example, a question leading to a possible invitation or request is answered cagily, and the absence of a direct positive response is already discernible by the response's first syllable.

## (4) Schegloff (2007b: 68)

1 Bee: 'hmhh .hh So yih gonna be arou:n this weeken'?,  
 2 Ava: Uh::mm (0.3) Possibly.

Still, second-pair parts are not the only things that come after first-pair parts: insertion sequences (like requests for clarification) are always available, and



counters (Schegloff, 2007b: 16–19) may reverse the trajectory (A: *Where are you going?* B: *Well where are you going?*).

In contrast, first position turns, for example the first part of an adjacency pair, can in principle come without this clear projective advantage as a clue to the action being performed. Their vulnerability to misinterpretation may be one motivation for preambles of various kinds, including *pre*-s (illustrated below in (5)). Normally, of course, they will be located somewhere after preceding turns, which can cumulatively bias interpretation. (Even when they occur at or near the beginning of an interaction, that fact projects expectations of a small number of actions appropriate to that location, for example noticings about changes in appearance (*Wow!*) said on first meeting are much more likely to refer to the other's appearance than later on in the conversation; see Schegloff, 2007b: 86–7.)

The cumulative effect of sequential location can be clearly seen in examples like the following, where the arrowed line in 7 has an action ambiguity between a straight question on the one hand and a pre-announcement on the other (pre-announcements check whether the recipient already knows the news):

(5) Terasaki (1976: 45)

- 3 Rus: I know where yer goin,  
 4 Mom: Whe^re.  
 5 Rus: .h To that eh (eight grade)=  
 6 Mom: =Ye^ah. Ri^ght.  
 7 Mom: -> Do you know who's going to that meeting?  
 8 Rus: Who.  
 9 Mom: I don't kno:w.  
 10 (0.2)  
 11 Rus: .hh Oh::. Prob'ly .hh Missiz Mc Owen 'n Dad said  
 12 prob'ly Missiz Cadry and some of the teachers.

An utterance like “Do you know who's going to that meeting?” in line 7 might always be vulnerable to misunderstandings between two kinds of first actions: a question and a pre-announcement. But notice that in line 3 Rus produced something that might be characterized as a pre-guess, a wager taken up by Mom. So Mom, in producing line 7, could be read as doing a counterpart challenge: *Bet you don't know who's going to that meeting!* As it happens, that was not what was intended, but the example allows one to see how two structurally independent sequences can be stacked to predispose toward certain interpretations (see Schegloff, 1988e, 2007b: ch. 10). Many other observations about the stacking of actions have been made as well, for example that complaints tend to come in sequences (Drew & Walker, 2009: 2405), and that one request will tend to induce a counter-request (Schegloff, 2007b: 83).

Enough has been said to show that there is arguably no location in a conversation that comes without sequential expectations hedging in action attribution—there are overall structural constraints (e.g. things that have to be done right at the beginning, like greetings, followed by *How are you*s, etc.), and there are

cumulative sequential constraints throughout. Nevertheless, there are positions where the constraints of expected next actions are much less, as in the launch of new sequences. Here the major clue to the action type will lie in the turn format itself, to which we turn in the next section.

One clarification is in order: sequences contribute to action formation and recognition in two ways. First, as discussed at length, the location of a turn maps sequential expectations of action content onto it. Second, though, *the fact that an action is part of a larger sequence changes the very character of it*. A pre-invitation as in (6) is a double—a request for information and a broaching of an upcoming action. Responses are thus oriented to both, as the example illustrates. By the criteria mentioned at the beginning, the overt attention to both actions in the response recognizes the double character of the prior turn.

(6) SB (Schegloff 2007b: 31)

- 1 A: Ha you doin- <say what 'r you doing.  
 2 B: Well, we're just going out. Why.  
 3 A: Oh I was just going say come out and come over here . . .

## 5 Action Formation and Turn Design

We turn to consider how the details of turn design are used to signal or convey the intended action.

### 5.1 *A first pass: some early contributions from Linguistics*

The term *action formation* might be understood in a broader way to ask how all the properties of the interaction (linguistic formulation, position, gesture, setting, etc.) are enlisted by a speaker to make a turn recognizable as a specific action (Schegloff, 2007b: xiv), but here we will concentrate on the turn design, the linguistic signal itself (see also Drew, this volume, on turn construction).

Needless to say, the linguists have left a trail here, from which a number of helpful observations can be gleaned. First, it appears that all languages make distinctions between major 'sentence types', like declaratives, interrogatives and imperatives (most languages also have minor sentence types, as in the English *What a beauty!*, *designed*, *it seems*, for exclamations). Although there is no easy correlation with actions, these sentential forms do play a role in action ascription: imperatives for example (in English marked only by dropping the subject, but in many languages by morphology) are normally associated with urging or wishing the addressee to do or think something (*Shove over*, *Have a nice day*, *Don't speak with your mouth full*). Given that there are something of the order of a hundred or more distinct 'main' actions (see below), obviously a handful of grammatical forms will not alone solve the action recognition problem.

Another contribution from Linguistics (already mentioned above) is the recognition (under the rubric ‘indirect speech acts’) that stating or asking about the preconditions for an action might end up performing the action itself, as in *Is anyone using that cup?* being read as a (conditional) request for that cup. The claim is that questioning or stating (as appropriate) of any of the preconditions for an action can count as performing that action: thus *I want you to X*, *Has X already been done?*, *It'd be good if X*, etc., could all do requests in the right circumstances. These observations have their CA translation, most clearly perhaps into the language of pre-sequences. Checking a precondition for an action is the typical motivation for a pre-sequence (*Did you hear the news?* checks the appositeness of news-telling, *Are you going downtown?* checks the preconditions for a ride, etc.). When the action to which the pre- is a forerunner is entirely perspicuous given the pre- alone, as in *Are you using that pencil?* or *Can you reach that suitcase?*, then this invites sequence truncation as illustrated by the contrast between the following two examples (see Levinson, 1983: 356–64 for data citation and detail; earlier versions of the idea can be found in Goffman, 1976; Merritt, 1976).

## (7) Merritt (1976: 324)

- |               |   |
|---------------|---|
| 1. Pre-       | Hi. Do you have uh size C flashlight batteries? |
| 2. Go-ahead   | Yes sir   |
| 3. Request    | I'll have four please                           |
| 4. Compliance | ((turns to get))                                |

## (8) Sinclair (1976: 60) (truncated, nonappearing ‘reconstructed’ part in italics)

- |                |  |
|----------------|--|
| 1. Pre-        | Have you got Embassy Gold please?      |
| 2. Go-ahead    | ((Yes))                                |
| 3. First Part  | (( <i>Can I have a packet of 20</i> )) |
| 4. Second Part | Yes dear ((provides))                  |

This analysis gives an account of how and why specific formats may end up doing the things they do—if you like, a reconstructed evolutionary theory of practices. We will return to this analysis shortly.

Linguistics has contributed many studies of speech act design, for example of the way adverbs like *please* (Wichmann, 2004), *obviously*, or sentence-final particles, or evidentials in other languages, contribute to action design. This literature can be sampled through searches in, for example, the Bibliography of Pragmatics Online (<http://benjamins.com/online/bop/>), using rubrics like *speech acts*, *requests*, *questions*, and the like.

## 5.2 ‘Front-loading’, prosody, gaze and early cues to function

Given the point made by Figure 6.1, namely that action ascription must be made fast, even before the incoming utterance is complete, there is reason to think that

there may be cues to action type early in the turn. It is notable that many languages like English shift their question words right up to the front of the utterance, so that sentences with *wh*-words downstream (as in *You saw what?*) are likely to be read as repair initiators, rather than initial information-seeking actions. Even where, as in Japanese, *wh*- words are said by linguists to occur *in situ* (i.e. in the same position as a non-interrogative, as in *You saw what?*), in actual practice most of them are fronted (Hayashi, 2010; see also Yoon, 2010, on Korean), and where *in situ* order seems normal (as in Lao; Enfield, 2010c), it may be that the sentence is often rephrased to front the *wh*-word, or there are some other early indicators of function.

Interestingly, though, many languages (including these last three) have sentence-final particles for asking yes/no questions with variable import, so this cannot be the whole story. Even here, though, up to half of these are repair-initiators or confirmation requests, so not in first position in the base sequence (Enfield, 2010c); in addition there are very high proportions of 'declarative' yes/no questions, i.e. questions that are formally unmarked (40% in Lao and Japanese, 74% in Korean). Overall, then, despite the existence of sentence-final particles, there is some evidence for 'front-loading', or at least omission of 'back-loading', in these languages too. We treat it here as a likely universal bias given the nature of the turn-taking system and vulnerability of final items to overlap.

We have dwelled on questions because they are probably the most frequent high-level action type in all languages, and much researched in CA and neighboring disciplines (see Clayman, Gardner and Gill & Roberts, this volume, on news interviews, classrooms and medicine, respectively). *Wh*-words are ranked in the top 200 most frequent words in English and other languages, and they are consequently among the very most ancient, least borrowable of all forms. This centrality is partly because they are often, as Schegloff has put it, *vehicles* for other actions. We have just seen why this is the case—checking out whether a precondition for another action holds is usually done with a question, and in so doing the question may end up also doing the foreseeable next action it is a prelude to.

Questions are often discussed by linguists in connection with rising intonation, which would only become manifest at the end of the utterance. Rising intonation does indeed seem to be associated with *Wh*-questions, but they of course are often marked by question words appearing earlier. Rising intonation might be thought to play a special role in so-called 'declarative questions', that is yes/no questions lacking special marking (inversion in languages like English, particles or affixes in other languages), but corpus studies report that this may be absent or at best variable (see Enfield, 2010c; Geluykens, 1988; Levinson, 2010; Stivers, 2010; see also Couper-Kuhlen, *frth.*, for example). Again, a front-loading bias suggests looking elsewhere. Right at the beginning of a turn, it has to start on some particular pitch, and this pitch can deviate from the speaker's norm or from the immediately preceding pitch sample provided by a speaker. Instrumental measurement over a corpus of conversation in ten different languages shows that this cue is, in fact, used to mark questions that are *not* information seeking, for example, challenges, rhetorical questions, and the like (Sicoli, et al., in prep.).

Corpus results can be taken as indicators of phenomena that should be followed up by CA methods—a correlation between turn design and action type does not, by itself, tell us how or why a signal is deployed or recognized. There is in fact a nice analysis of pitch offset by Couper-Kuhlen (2001), who shows a parallel use of initial up-step: various preliminaries may be handled early in a phone call, making potentially problematic the location of the main reason for the call; these are, however, usually marked off by a high pitch offset right at the beginning of the turn.

CA studies of prosody in the service of action formation are relatively few (though see Couper-Kuhlen & Ford, 2004; Selting, 1996b; see also Walker, this volume, on phonetics more generally). Schegloff (1998b) provides a single case study which suggests a useful direction. He considers the following example (9), wherein two young women are discussing the visit (arranged by a match-making friend) of an eligible young man from Minneapolis to see Hyla. Hyla minimizes the possibilities (“gimme sumpn tuh do one night”, line 5), but Nancy foresees a mini-tragedy (“then yu’ll like him an hi’ll go back tuh Minneap’lis.”, lines 6–7). Schegloff suggests that the stressed pronoun “him” encodes an implicit compliment, because the contrastive stress suggests *It goes without saying, he’ll like you*. Some evidence for this is that Hyla produces the self-deprecating rejoinder typical of compliment receipts (the stress now suggesting *I’ll write to him, but I won’t hear back*).

(9) Schegloff (1998b: 248)

- 1 Nan: That sounds goo:d.  
 2 (0.2)  
 3 Hyl: Eh:::,=  
 4 Nan: =A’xi:::[gɪt,  
 5 Hyl: [Gimm[e sumpn [tɪh do [one night]  
 6 Nan: [Y e : h [except [then yu ll] like him en hill go  
 7 back [tuh Minneap’lis.=  
 8 Hyl: [hɦɦɦ hɦɦɦɦ]  
 9 Hyl: =.eh En ah’ll ne(h)ver hear fr’m him a[gai:n, ]  
 10 Nan: [nɪɦɦ hɦɦ] -heh

This suggests a not implausible model for how prosody plays into action formation, in contrast with an implausible model of a direct matching of prosody with action type (as in the idea of ‘the rising intonation of questions’)—implausible because of the multiplicity of action types and the graded nature of prosodic coding. Instead, the model suggests that participants make use of some much more low-level coding of meaning (here, stress to invoke a contrastive situation which implies an implicit action).

Contrastive stress is categorical—it contrasts with unmarked lexical stress. Many prosodic and phonetic features are gradual or continuous. For example, the pitch range, the amplitude and the degree of phonetic precision (hyperarticulation, lack of slurring) are gradient. These can nevertheless play a role in action

formation. Consider the following, where two students are talking about the difficulty of finding housing:

(10) Ogden (2006: 1766)

- 1 A: and it's (.) ?a l:ot of effort=  
 2 B: =it is quite a lot of hassle  
 3 unless you like cos sometimes it's just luck(y)  
 4 isn't it you just like walk in . . .

A second assessment, Ogden (2006) shows, typically displays phonetic upgrading (increased pitch range, increased phonetic precision) if it agrees with the first. But here the pitch range of B's second assessment is *reduced* compared to the first, the tempo is faster and the amplitude reduced—all harbingers of the partial disagreement following. (For other uses of relative pitch contrast, see Local, 2007.)

Another kind of early cue to action type may be kinesic. CA-informed studies of gaze show that initial and sustained speaker gaze is typical of question delivery (Rossano, et al., 2009). On the other hand, addressee gaze is sensitive to sequence closure and sequence type (Rossano, 2012)—for example, a bid for an extended turn-at-talk gets rapid addressee gaze. This affords the analyst insight into what exactly 'tips off' the recipient that an extended turn is forthcoming. Using gaze as an index of recipient recognition of a story in progress, Rossano shows that the first few words may be sufficient to cue that recognition (e.g. *Yesterday, I was . . .* where the past time and first person reference seem efficient cues to an upcoming story). Gestures are known to typically precede the linguistic units they are semantically connected to: thus a request to pass the water may well be slightly preceded by a gesture in that direction, again giving front-loaded cues to action recognition (see Kelly, 2001). Enfield (2009: 223ff.) suggests that this kind of rough and ready cuing might fall under the psychological notion of 'fast and frugal heuristics'. In the case of action recognition, the idea would be that if early cues suggest a specific action is in the offing, assume that is the correct identity until proven otherwise.

Finally, the limiting case of front-loading is where the cues are already provided in the context, either by virtue of the activity (see section 10 below) or by virtue of shared knowledge or epistemic states. Heritage (2012b) shows how, for the ascription of questioning, a skewed epistemic state (speaker does not know, addressee is assumed to) trumps any kind of formal marking on utterances. Declaratives that betray epistemic imbalance of this kind work fine as questions (*You're divorced currently* or *Yer lines been busy*) as Labov and Fanshel (1977) noted, but interrogatives in a context of equal epistemic access do other actions, for example assessments (*Isn't that good at long last* or *That Pat isn't she a doll?*).

### 5.3 *A second pass: detailed linguistic information and the problem of indirection*

It is clear that some actions (greetings, partings, *Yes* responses to yes-no questions) are coded by a limited set of dedicated formats. Requests are one class of first



actions that tend to have recurrent formatting.<sup>7</sup> In English conversation, they come overwhelmingly in *Can/Could/Would you* form, but there are nevertheless choices to be made. There are many studies of the coding of requests in specific languages and contexts (the Bibliography of Pragmatics Online lists over 200 references), but relatively few that meet CA criteria. Curl and Drew (2008), expanding on Brown and Levinson (1987), consider what seems to motivate the choice between two main kinds of prevalent request forms, the *Can/Could you* (modal) forms on the one hand, and the *I wonder if* form on the other. They find that the latter is used when there is doubt over whether the contingencies for the request are met (for example whether a shop can deliver, or whether a doctor's visit is warranted). In contrast, the modal forms are used when the speaker's entitlement to the requested service is high, and the contingencies are assumed to be known.<sup>8</sup> Thus low entitlement, high contingency motivates *I wonder if*, while high entitlement, low contingency motivates *Can/Could you* forms. One interest of the analysis is that it explains the greater prevalence of *I wonder if* forms in service exchanges and institutional contexts, where the speaker knows less about the constraints, without directly invoking the setting: the institutional setting is constituted via the mode of talk rather than the other way around. (Incidentally, similar findings concerning entitlement and contingency were found experimentally in the 1980s; see Levelt, 1989: 134–6.)

But requests, taken overall, are not such innocent beasts. By requiring actions from the other party, they inevitably raise issues about the nature of the social relation. Is this a relation in which the requestor has rights to presume compliance ('entitlement')? Are the expectations based on social status (kinship, institutional position) or on the more fickle exchange of friendly favors? There is often room for doubt about entitlement, and requests may then be approached circuitously or at least tentatively, in the hope that a pre-emptive offer may be forthcoming:

(11) Levinson (1983: 343)

- 1 C: Hullo I was just ringing up to ask if you were going to  
 2 Bertrand's party  
 3 R: Yes I thought you might be  
 4 C: Heh heh  
 5 R: Yes would you like a lift?  
 6 C: Oh I'd love one

Schegloff (2007b: 64, 81ff.) reviews the evidence for offers being preferred (in the technical sense) to requests, so that efforts may be made to 'fish' for an offer rather than request outright, and notes (a) that requests may be dealt with positively or negatively without ever actually being broached, and (b) that they may be disguised as other actions, for example an offer to help disguising a request for some joint activity. In these cases, the action of requesting is actually avoided but the effect otherwise achieved. Obviously, researching the unsaid is not an easy task: one needs to show that both participants are oriented to something not occurring

and deliberately kept off-stage as it were (see Pomerantz, 1980). The following is a transparent case of an unsaid request and its lightly adumbrated refusal:

## (12) Schegloff (2007b: 64)

- 1 D: Guess what.hh  
 2 M: What.  
 3 D: .hh My ca:r is sta::lled  
 4 (0.2)  
 5 ('n) I'm up here in the Glen?  
 6 M: Oh:.  
 7 {(0.4)}  
 8 {.hhh}  
 9 D: hhh  
 10 A:nd. Hh  
 11 (0.2)  
 12 I don' know if it's po:ssible, but {(0.2)/(hhh)} see I haveta open  
 13 up the ba:nk.hh  
 14 (0.3)  
 15 a:t uh: Brentwood?hh=  
 16 M: =Yeaḥ:- en I know you want- (.) en I whoa- (.) en I would, but-  
 17 except I've gotta leave in  
 18 aybout five min(h)utes. (hheh)

Offers are another action type where we know something about the distribution of action formatting. Curl (2006) reports a correlation of three main formats with three different kinds of occasioning. In the first, an offer is provided, after a preamble, as the reason for a call: these offers take the form *If X, then Y*, as in. *If there's anything we can do, let us know*. In the second context, it emerges slowly over the course of talk that the recipient may perhaps have some problem, and the offerer produces an offer of the form *Do you want me to X*. In the third context, the problem is made overt, and an immediate offer (apparently often in overlap) is made in response, but in diverse formats, for example *Can I bring some pies or something, I'll take you up Wednesday*, and so on.

To sum up this section, for a limited number of action types, we have detailed studies that show that the turns that embody these actions seem to be drawn mainly from a handful of limited formats. This might suggest that the problem of action formation and ascription has been exaggerated: there are well-worn practices for each type of situation. This is almost certainly a complacent conclusion.

This issue, the relative roles of what is coded in an utterance vs. what is 'read into' it by virtue of contextual (e.g. sequential) expectation, is a much-rehearsed controversy in the psychology of language. For example, do we identify words entirely through their phonetic character ('bottom up' from the signal), or do we partially guess their identity through the sense of what is being said ('top-down' through sequential expectation)? The cocktail party syndrome shows that we are certainly capable of the latter, but research seems to show that we use the signal where we can. Exactly the same issue arises in the study of action ascription: to what extent are actions finely coded in linguistic detail, or to what extent is rec-



ognition or ascription dependent on tracking complex sequential context? Just like psycholinguists are split on this issue, CA researchers are likely to make different initial assumptions here (with perhaps Curl and Drew (2008) and Couper-Kuhlen (2010) tending toward a 'bottom-up' analysis, and Schegloff (2007b) and Heritage (2012b) toward a 'top-down' one).

The fine-grained work on conversational actions of different kinds has mostly yet to be done. It is possible that no turns are, as it were, action interchangeable—that they carry in the detail of their linguistic formulation, prosodic delivery and phonetic articulation indelible cues to the action they are intended to deliver. That is a strong prediction, experimentally verifiable if true. But it is unlikely to be true in general. Some of the cases studied are action types that are often preceded by a lengthy negotiation, making unequivocal delivery finally relevant and relatively risk free. In many other circumstances, for example the actions leading up to this point, equivocal delivery and deniability will be desirable. In addition, as we will see below, actions get part of their character from the overall 'project' (see section 7) they are steps toward, and this will, in many cases, not be coded in their form, if only because projects are unlimited in kind.

## 6 Action Formatting vs. Content-Defined Practices

So far, the discussion of action formation has equivocated on the distinction between a practice defined in terms of turn format (e.g. *Could you, I wonder if, Why not . . .*) and a practice defined in terms of content, whatever the precise linguistic formulation.<sup>9</sup> Thus, we noted that stating or questioning a condition of an action may often be taken as performing that action (*Do you need that pen?*), where no precise restrictions on format are imposed. Some of the most interesting CA observations about action coding have been made in terms of content, not format.

Consider first pre-s. Pre-announcements have been noted to have recurrent turn formats (Levinson, 1983: 350–3; Schegloff, 2007b: 38; Terasaki, 1976)—they come in a frame with a variable (a linguistic item like *something* or *what*), where the third position announcement fills in the variable, often copying over the rest of the frame. *You know what?, Guess what?, You'll never believe what happened to me, Didju hear about Bill*, and so on, are all typical formats, but the range is greater (e.g. *Yer not in on what happen' .(hh)d?.*) and the formats could not be listed extensionally. The format has in fact to be stated in terms of constraints on content (roughly, a check on the recipient's existing knowledge, using a variable that leaves the news unspecified). But other pre-s may be much more various in shape—for example pre-offers may lack any specific format, as Schegloff (2007b: 35) notes.

Now consider an especially interesting case. Pomerantz (1980) noted a regularity of the following kind: if one states an externally observed aspect of what the recipient can be supposed to know in experiential terms, one can be seen to be 'fishing' for information about that experience. For example, in the following example (13), S notes that G's telephone line was busy (engaged), the sort of

remark that qualifies as ‘fishing’, but which fails to elicit a volunteered explanation, so that S follows up with a direct question:

(13) TC:1:1:2 (Pomerantz, 1980: 186)

- 1 G: . . . dju j’see me pull up?=  
 2 S: = hhh No:: I z trying you all day.en the line wz busy fer like  
 3 hours  
 4 G: Ohh:::::, ohh:::::, hhhhhh We::ll,hhh I’m g’nna c’m over in a  
 5 little while help yer  
 6 brother out  
 ((5 lines omitted))  
 7 S: hhh Uh:m, tlk hhh Who wih yih ta:lking to

Often though the ‘fishing’ will be successful:

(14) NB:II2.-1 (Pomerantz, 1980: 189)

- 1 A: Yer line’s been busy  
 2 B: Yeuh my fu(hh)! Hh my father’s wife called me . . .

Pomerantz characterizes ‘fishing’ in terms of telling one’s own side in hopes of eliciting the other side with the firsthand knowledge of what happened. She sees the practice as oriented to an acknowledgment of no rights to know, no rights to ask. Participants can ‘fish’ for invitations, offers, news, and so on. There is no fixed format—the practice has to be characterized in terms of stating one’s own side where the recipient can talk as the horse’s mouth, all against the background of relevant norms (rights to private information, rights to invite, etc.). ‘Fishing’ is an interesting borderline case for action defined in terms of required response—as the first case illustrates, the recipient can stonewall. In terms developed below, it is an implicit ‘project’.

## 7 Multiple Actions in One Turn

There is a trivial sense in which a turn can contain more than one action, as in the following, where in response to H’s *How are you?*, N answers *fine* and asks the same question reciprocally all in one turn. Turns can be composed of more than one turn-constructional unit, each performing a ‘main action’, so two units can clearly pack two actions into one turn.<sup>10</sup>

(15) HG 1:01–11

- 1 H: Hwaryuhh=  
 2 N: =Fi:ne how’r you,

But here we are interested in how a single unit can seem to do more than one action. Linguists have talked about indirect speech acts, presuming that in asking

a question like *Is that a spare pencil?* the speaker is also requesting. Schegloff (2007b: 9) has usefully coined the notion of one action being a *vehicle* for another, as mentioned earlier. Often questions are the vehicles for other actions, from complaints, to offers, or invitations. They are also the typical format of pre-s, which are clearly doing more than one thing—checking a precondition to a future action, and so, with varying degrees of transparency, leading up to that second action. Pre-s, we have seen, are thus doubles.<sup>11</sup>

It can go deeper: questions can be used to invite repair, and in inviting repair, they can be used as pre-disagreements. Such other initiations of repair may carry special stress or pitch (Selting, 1996b):

(16) Schegloff (2007b: 102)

- 1 A: Have you ever tried a clinic?
- 2 B: *What?*
- 3 A: Have you ever tried a clinic?
- 4 B: ((sigh)) No, I don't want to go to a clinic.

(17) Schegloff (2007b: 102)

- 1 S: We went to camp. Forget it. She wouldn't behave for anything.
- 2 A: W-when.
- 3 S: When we went to camp
- 4 A: She behaved OK

(There is also the case of *pre-pre-s* as in *Can I ask you a question?*, turns that overtly foreshadow a named action that then appears quite far downstream (Schegloff, 2007b: 44), but are pre-s to a preliminary to that action.)

Questions (as requests for information or confirmation) are such a common vehicle for other actions that some authors treat interrogative features as merely one of a number of turn design features (like gaze, prosody, epistemic asymmetry) calculated to elicit response (Stivers & Rossano, 2010), thus stripping such questions of their primary action force, and reducing them to mere packaging in the service of other actions. This would seem to make incomprehensible the double responses to such question vehicles as in (6). In any case, questions are not the only format for double actions: *I need some water* can do a request as well as *Can you pass the water?*

## 8 Actions vs. Projects

Although the actions reviewed, like offers and requests, are often delivered like silver bullets in a single turn, they do not for the most part come out of the blue. They may either, like Curl's (2006) responsive offers, be touched off by the other's immediate talk, or they may be carefully launched after the offerer has prepared the ground with extensive preamble. For example, offers of the *If X, then Y* kind

seem to be introduced by an extended telling whose pertinence seems to come clear to the recipient quite late as indicated in line 31 (Curl, 2006: 1260):

## (18) Holt: 2:3:1-2

- 5 Les: Hello, .tch.h I hh^ope you don't mind me getting in touch  
6 but uh- we metchor husband little while ago at a Liberal  
7 meeting.  
8 (0.3)  
9 Mar: Ye:is?  
10 Les: [.hh And he wz: (0.3) i-he told us something of what'd  
11 happen:ed,  
12 (0.5)  
13 Les: to him .hh An:' I wondered haa- (0.2) i-he said he m::ight  
14 have another position in vie:[w,  
15 Mar: [Mhm,  
16 Les: .hh (.) Uhm (0.3) .tch Well I don't know how that went, .h  
17 uh (.) It's just that I wondered if he hasn't (0.3) uh  
18 we have friends in: Bristol  
19 Mar: Ye:is?  
20 Les: who:- (.) uh: that u-had the same experience.  
21 Mar: Oh^i:.  
22 Les: And they uhm: .t (0.2) .hh He worked f'r a printing an:'  
23 paper (0.9) uh firm[u-  
24 Mar: [Ye:s,  
25 Les: uh[: - which ih puh- uh: part'v the Paige Group.  
26 Mar: [Yeh,  
27 (.)  
28 Les: .hh And he now has: u-a:: um (1.1) I don't think eez called  
29 it consultancy (0.2) They find positions for people: in the  
30 printing'n paper (0.4) indus[try:  
31 Mar: -> [Oh I see:[:.  
32 Les: [ hh An:d if: i-your  
33 husband would li:ke their addre:ss.  
34 Mar: [Y e :[:i s,  
35 Les: [<As they're  
36 specialists,  
37 Mar: Ye::is?  
38 (.)  
39 Les: Uhm: my husband w'd gladly give it to him.

Note here Leslie produces a preamble (whose lead-in status is recognized by the recipient's continuers) which is clearly going somewhere sensitive ("I hh^ope you don't mind..." line 5), but its precise identity is only apparently recognizable in line 31 ("Oh I see:").

Let us consider this distribution of the job over a preparatory and delivery phase a *project*. We need the notion of this larger 'project' entity for a number of reasons (see also Clark, 1996: 205ff.). First, the adumbration of a project may be sufficient for participants to see where it is heading and so gracefully abort it before more overt damage is done (see the covert request sequence (12) above,

adumbrated but never delivered). Extract (19) shows an offer blocked in the preparatory stage (line 11), then re-opened as the details of the offer become more attractive, the offer proper being delivered at the end of the excerpt (lines 22–6).

(19) Holt: S088:2:7:1 (Curl, 2006: 1261)

1 Les: [.hh Oh I'n (.) I'n sorry to bring you fr'm your  
2 wo:rk .hhh Serena you ^^know what you were telling me  
3 Les: about your father yesterday an' how he wz: eventually  
4 goin't'look for- a- .hh store manager job  
5 (.)  
6 Ser: Ye:is,=  
7 Les: =.hhhhhhh .t u-We:ll, where my husband works: they're  
8 actually goin' to adver|tise for one in the near fu|ture.  
9 (0.5)  
10 Les: .hhh|h But  
11 Ser: -> [Yes it won't be finished till July next yea: [r.  
12 Les: [.hhh  
13 ah: yes: I thought so  
13 Les: So I jus' thought I'd ring  
14 you an' ask .hhhh|h uh:m  
15 Ser: [Oh: (what a shame ^heh)=  
16 Les: =Okay?=  
17 Ser: =Yeah. No th- (0.2) that would be brilliant uh ac[tually  
18 Les: [It is  
19 actually a very well paid jo:b.  
20 (0.3)  
21 Ser: ^Oh:.  
22 Les: -> Yea:h, .t.hh Anyway eh:m .t.h ^tell im if he wants to  
23 have a word  
24 Ser: Yeah what's the=  
25 Les: -> =.hhh to ring my husband up when we c-come back from::n  
26 Newcastle

Second, because participants thus show that, at some point during the exchange, they orient to the tellings as 'leading to' (in this case) an offer, the preamble has to be characterized in terms of where it is heading. Much talk that looks desultory to the analyst may consist of projects launched but silently sunk on submerged interactional rocks.

Third, we need to be able to distinguish projects as courses of action from the sequences that may embody them. A clear example of this can be found in courtroom interaction (see Komter, this volume), where examination is conducted by means of question-answer sequences. Completely different courses of action are likely to be found in examination-in-chief, where the council for a client interrogates the client or his witness to extract a presentation favorable to their case, compared to cross-examination, where the other side's council interrogates the witness in order, for example to make a charge stick, or to show the witness is unreliable (Levinson, 1979). Exclusively in the latter case, questions may become the medium for sustained accusation (Atkinson & Drew, 1979), and the witness typically designs answers to resist this project. So we can have the same sequence

Su

with different projects. Likewise, we can have the same project instantiated in different sequences, as when a 'fishing' for information is abandoned in favor of a direct request (as in (13) above; see also Davidson, 1984). Thus projects cannot be reduced to sequences.

We need to firm up the notion of a project a little further (see here also Schegloff, 2007b: 244–50, who talks of "thematic threads", "courses of action", as well as "projects"). A project is not a sequence, for it may or may not be instantiated in a sequence: in (12) *My car is stalled*, the project of asking for rescue is evident, but never gets embodied in a request sequence (and in (19) the offer project very nearly got stillborn by an early blocking move). Pre-s and preambles as in (19) adumbrate projects, but if blocked, the sequence they are a prelude to may never get aired. The notion of project we need for action ascription is not 'thematic thread' but 'plan of action'—that is, a course of action that at least one participant is pursuing, which may at first be opaque to others then retrospectively discernible (cf. the 27 lines preceding "Oh I see:." in (18)), and then prospectively projectable. When the other buys into the project, it is likely to surface as a sequence; when not, its invisible hand will anyway have directed the talk.<sup>12</sup> Either way, projects play a crucial role in action ascription, because 'seeing' what project a turn adumbrates plays a crucial role in the response, either encouraging or discouraging the project.

## 9 Action Types: Issues of Level and Inventory—Granularity Revisited

We earlier raised the issue of granularity, but we are now in a better position to address it. Schegloff (1996b: 211), flushed with excitement over the discovery of what appears to be a new action type ('confirming allusions'), asked:

How many other such unknown jobs, functions, actions, practices—such as confirming allusions—might there turn out to be? Hundreds? Thousands? . . . We lack the social equivalent of the presumed 90% of the physical matter of the universe now unaccounted for and termed 'dark matter'.

Recollect that we earlier distinguished any and every way in which a practice might be said to do something from an action whose identity has strict sequential consequences. So what does the inventory of action types, *sensu stricto*, look like? It is certainly clear that it exceeds by far the vernacular metalinguistic terms like offering, complaining, requesting, offering, teasing, insulting, greeting, and so on (themselves particular to only one of the 7,000 languages on the planet). To that list, we need to add the technical terms for action types like assessments, pre-closings, self-identifications, pre-invitations, other-initiation of repairs, and so on. We certainly have over a hundred in the bag already.

But it is not clear that confirming allusions is one of them, and that is because it is not clear that confirmation by repetition vs. confirmation by other means has

differential sequential implications. 'Confirming allusions' is the practice of exact repetition of an utterance that seems to require confirmation, and the repetition (instead of a simple confirmation token like *Yeah*) alludes to the lead up to the offending utterance, and implies 'That was just what I was saying'. As Schegloff (1996b: 209) says:

As compared to the grosser types of action involved in the sequences we have examined—which we could formulate as 'confirmation' and 'disconfirmation', one or the other of which is virtually *mandated* by an interlocutor's offering of a candidate understanding to a speaker of what the speaker has just said—the form that the confirmation takes appears to be what might be called an 'optional action'.

It is the 'virtually mandated' action types that form the central structure of conversational interaction. Schegloff (1987a, 1996b: 165) has complained of Goffman's (1976) deflection of analytic attention from 'first-order actions' to the subliminal 'ritual actions'—let us then give the first-order, official business pride of place. The same danger may, perhaps, lie in the current analytic attention to the implicit struggles over epistemic territory, indubitably there, but often a second-order business. (However, as noted above in section 5.2, epistemicity can also play a central role in action formation, converting declaratives about things the addressee knows best into questions, as in *Your leg is hurting*. Many languages like Lao or Japanese have a range of particles expressing a range from speaker certainty to doubt, which change the nature of the expected response.)

## 10 Cross-Cultural Regularities in Action Types and the Productivity of Activity Types

So constrained by a relatively coarse level of granularity (but supplemented by the open-ended nature of projects), action types within a social group may be limited in type, thus constructing limited (or at least rule-driven) sequence-types out of action pairs. That would constitute the kind of thing a child growing up within a culture could learn, producing Sacks' apparatus for generating recognizable actions. Some of these actions—like requests for information, requests for action, offers—look universal or cross-culturally shared. Exactly which, and exactly why these action types, remains a mystery, understanding which might tell us a great deal about the phylogeny of the interaction style found in our species (see Canfield, 1993).

But other action types are almost certainly culturally circumscribed. Greeting sequences, for example, while sharing general functions, may look entirely different in different cultures (Duranti, 1997b; Irvine, 1974). The point is easily made by looking at exotic action types that lack any obvious English counterpart. Take, for example, Yéli Dnye father-in-law jokes made by very oblique allusion to some faux pas made by the recipient's father-in-law, for which the second is a return quip of the same kind (Levinson, 2005). Despite this proliferation of



culturally-specific action types, the local metalanguage for actions is likely to be a very poor guide (there is no Yélí Dnye word for father-in-law jokes). In fact, unwritten languages tend to have very little of the metalanguage we find in a language like English with a long literary tradition (which itself fails to cover all the base actions for adjacency pairs)—Yélí Dnye has, unlike English, no word for promises, offers, threats and the like.

One reason for the cultural proliferation of action types is the human predilection for inventing, defining and refining activities, in which special rules for allowable action inventories, action formats and sequence types may hold (Atkinson & Drew, 1979; Atkinson & Heritage, 1984; Levinson, 1979)—think of TV quizzes, umpires' pronouncements, auctioneers' calls, and so on. Our institutional imagination may be the only limit on possible action types.

Activities play a role in action ascription because they on the one hand amplify, and, on the other, constrain, the types of expectable action within them. You do not expect to buy a Picasso with a nod except in a fine art auction (Heath & Luff, 2007a). This is where the other less-developed half of Sacks' program fits in: the activity imposes a set of roles (*membership categorization* in Sacks' terms) and action types that go with them (unlike the buyer's nod, the auctioneer's nod may recognize a bid, but will not itself constitute a bid). Drew (1978) for example shows how a sequence of questions in a courtroom may add up to an accusation, and the answers construct a systematic denial. The view that the activity is constituted by the interaction, correct in itself, does not of course imply that the activity is *de novo* constructed on each occasion—it should be understood as an evolved set of practices that guides action formation and ascription within specific settings.

## 11 Action Streams and the Nonverbal

Linguists have traditionally given the verbal level of communication priority (the rest has sometimes been labeled 'paralanguage', the outer satellites of the verbal sun as it were). But recently the mood has changed, spurred by developments in gesture studies and sign language. Now the gestural and the verbal are increasingly seen as part of a single package (Enfield, 2009), with the emphasis merely shifted from mouth to hands in the case of sign languages. Similar issues have arisen in CA. C. Goodwin (1994, 2003e) takes the integrated perspective, but Schegloff (2007b: 11) holds on to the view that the structure of sequences, as made out of adjacency pairs, has no substantial counterpart in the nonlinguistic realm: "There is, therefore, no reliable empirical evidence for treating physically realized actions as being in principle organized as adjacency pairs."<sup>13</sup>

Schegloff's position seems hard to defend: I wave hello, you do so in response; I raise my eyebrows and gaze pointedly in the direction of the wine, you get it for me; I hold out my hand, you shake it.<sup>14</sup> Whole sequences of adjacency-pair-like



turns are possible: I put the book on the checkout counter, you ring it up, I hand out a bill, you give me the change (Clark, 2006). Where nonverbal procedures are highly developed as in surgery, putting out a hand may clearly request different things at different moments (Mondada, 2002).

Indeed, thinking about nonverbal action sequences may help us understand the verbal ones. In general, we seem to understand others' actions in terms of a grammar of motives (cf. Sacks' grammar of recognizable actions): if you are fiddling in your pockets in front of a door, I assume you are searching for your keys and will go on and open the door. If I have the same keys, you might put out your hand for them, or I might offer them. Actions come tiered: this one (finding keys) is a step toward that one (opening the door), which is jointly achieved by a sequence of others (putting the key in the lock, turning it, and so forth).<sup>15</sup> We can insert our actions into others' action streams (offering the keys) because we recognize the whole from one subpart (the fiddling in the pocket by the door) and thus can predict all the rest of the subparts. This ability to 'see' actions, to project coherent sequences from their subparts, and cooperatively offer to do part of them, is part of some special capacity for interaction that seems more or less confined to our species (Levinson, 2006b).

That capacity to 'see' the superstructure behind the implemented first action or two has however its distinct limits. I can 'see' you in the act of opening the door, and perhaps guess from the hour that you are just returning from work, but beyond that your plans may be ineffable. In the same way, regular practices give us enough insight into each other's projects that we can collaborate in them or avoid that. We may surmise beyond, but once we leave the *terra firma* of regular practices, our insight is hazy.

If we now turn back this perspective on the complexities of verbal interaction, it helps us to see how pre-s work: a pre- (*What are you doing tonight?*) projects a conditional next action, which is how recipients can make informed choices about encouraging (with a 'go ahead') or discouraging the next action. Its duality of function (request for information, request for permission to proceed) follows from its place in a hierarchy of actions: it is a request for information, but also a prelude to, for example, an invitation, and it is dealt with as both (*Nothing much*). Pre-s are just simple examples of turns that inherit part of their action potential from the larger project or action plan they are the initial steps for.

Returning to the streams of action, talk over the dinner table illustrates the possibility of maintaining more than one action stream simultaneously: animated talk about some unrelated topic can fly over the business of passing plates, circulating condiments, and so on (see Schegloff, 2007b: 10–11). The two streams may intrude upon one another, and in the case of many kinds of workplace or service interaction, they interlock in complex ways, so action ascription has to track both. One of the signal merits of thinking about conversational exchange in terms of action sequences is that it explains how speech and nonverbal interaction can be locked either into a single system (as in requests) or in a system of systems (as in conversation over dinner).

## 12 Bringing It All Together: the Distributed Nature of Action Coding

In this review, many issues have been traversed, and puzzles collected:

1. How many action types are there? Is there a finite inventory, or an indefinite range of possible actions?
2. How are actions 'recognized' or, as here preferred, ascribed?
3. How can we account for one turn doing more than one action? Is there an indefinite number of things a turn can do at once?
4. Turns can perform just one or more actions simultaneously, but the actions they perform fit into larger schema, only some of which seem captured by the notion of sequence.
5. The unsaid haunts the said, ghosts of actions apparently being visible to participants who navigate around them.
6. Sequential position maps expectable actions onto turns, but turns can also wear their actions overtly on the sleeve, indelible markers of identity as it were. Which of these processes (top-down, bottom-up) rules the game?
7. How do nonverbal action streams interact with verbal action streams?

This chapter suggests that many of these puzzles disappear if we take a slightly larger-scale perspective, namely that of the project. A project is an action plan, and like any plan of moderate complexity, it will have steps to be taken on its way to completion: to make the coffee, I have to, say, find a filter, fill the water up, find the coffee, ready a cup, and so on, observing which, you might step in and get the cups. A collaborative project will interdigitate my steps and yours. Clearly, in conversation, projects are interactionally negotiated, jointly launched, diverted or aborted. Actions then are in the service of projects, and projects are themselves actions to accomplish. That is why there is no simple answer to what action this turn is doing: it is doing something local, which governs its response types, but also part of something more global, which, as soon as it is recognizable, also plays a role in fashioning responses (as in the 'go ahead' or 'blocking' responses to pre-s). In short, there is a hierarchy of actions within a project.

At least three tiers of such embedded plans can be addressed at once. Consider the following bicker over a daughter's allowance or pocket money, where Mom's "Ten dollahs a week?" (line 15) is ostensibly locally a clarification question (per week or per month?) to a proposal for more money; but also, querying the amount produces an opportunity for daughter Virginia to justify the amount, where, if those justifications prove inadequate, grounds are thereby provided for rejecting the proposal. Virginia's project, asking for more pocket money (earlier more clothes), is countered by Mom's project of holding the status quo. Other-initiated repair, an information request, a challenge to produce reasons, a pre-accusation and thus likely refusal to grant the request, are all visible in the one turn.<sup>16</sup>

## (20) Virginia, p. 8.

12 Vir: But- you know, you have to have enough mo:ney?, I think  
 13 ten dollars'ud be good.  
 14 (0.4)  
 15 Mom: .hhh Ten dollahs a week?  
 16 Vir: Mm hm.  
 17 Mom: Just to throw away?  
 18 (0.5)  
 19 Vir: Not to throw away, to spe:nd.  
 20 (.)  
 21 Mom: ((shrilly)) On [WHAT? That's what I been tryin'a find=  
 22 Pr?: [eh hih hih  
 23 Mom: =out. besides McDo:nalds?,

If we understand that actions are frequently parts of projects, and action content partly inheres in the larger whole, many of the puzzles about action recognition and formation disappear. To return to our puzzles and allay them:

1. *How many action types are there? Is there a finite inventory, or an indefinite range of possible actions?*

Primary actions are limited by conventions of practice, but these can participate in indefinite projects, thus acquiring infinite flavors.

2. *How are actions 'recognized' or attributed?*

By a range of factors: format (linguistic shape), content (e.g. mentioning of conditions on another action), position in a sequence, the nature of the prior sequence, by detecting the underlying project from the current and preceding turns, and by tracking epistemic authority and other aspects of context.

3. *How can we account for one turn doing more than one action? Is there an indefinite number of things a turn can do at once?*

Primary actions can transparently be linked to next actions in projects, so acquiring double actionhood. The limits on such forward action-chaining are perhaps just set by project perspicacity—in ordinary conversation three or four may be a practical limit governing the depth of response type (beyond that, guessing may encounter general deniability).

4. *Turns perform just one or two actions simultaneously, but the actions they perform fit into larger schema, only some of which seem captured by the notion of sequence.*

Sequences are in the service of projects: they implement them, but projects are not reducible to sequences. Actions often form part of a larger project inheriting part of their import from the larger whole.

5. *The unsaid haunts the said, ghosts of actions being apparently visible to participants, who navigate around them.*

Sequences get played out, but projects are often abandoned halfway, sometimes left skillfully and deniably covert, at other times broached but diverted or actually blocked. In all of these cases, though, mutual orientation to the unrealized project may be discernable to interactants, and derivatively, to analysts.

6. *Sequential position maps expectable actions onto turns, but turns can also wear their actions overtly on the sleeve, indelible markers of identity as it were. Which of these processes (top-down, bottom-up) rules the game?*

Actions may be overt, but the projects they serve may as yet be covert: perhaps only in the case of the culminating actions in a project (the offer or request finally made after preambles) can coded content exhaust the action content—otherwise the role in the ongoing project contributes to action content.

7. *How do nonverbal action streams interact with verbal action streams?*

At least in two distinct ways: first, when the activity has a nonverbal base (e.g. shopping) but is facilitated by language, where actions in a single chain may be realized verbally or nonverbally; second, where two (or more) action chains are superimposed (e.g. talking over dinner) and they need to ‘time share’, where one (e.g. eating, serving a meal) may be given priority. Foreseeing the other’s project (e.g. wanting the water) may allow the two streams to run concurrently without overt interruption.

### 13 Future Directions

This chapter hopefully suggests a framework within which future work can be conducted. It has only indirectly indicated the directions that that research might take, for example, by drawing attention to ‘front-loading’ of cues to action, the tension between ‘top-down’ and ‘bottom-up’ bases for action ascription, and the crucial importance of discernible larger projects behind individual turns. Careful corpus work on specific practices and action types will clearly help to build up a broader picture. Work on service encounters and concurrent activities (as in talk over eating) will contribute to a picture over the handling of multiple action streams, which remains underdeveloped. Substantial insights are perhaps most likely to be gained by looking for cases where action ascription is overtly under scrutiny by participants. Understanding how misunderstandings come about, as in (5), will be crucial. Seeing how a turn in production may be repaired on the fly to redirect the action may also yield fundamental insight (see C. Goodwin, 1981), as in the snippet below where a proposal is re-crafted into an invitation by switching the pronoun from “we:” to “I” (see Drew, 2005a: 95):

(21) NB:VII:3

1 Emm: Wul why don't we: uh-m:=  
2 =why don't I take you'n Mo:m up there tuh Coco's.someday fer  
3 lu:nch

In general, our understanding of the processes of action ascription is still at such an elementary stage that almost any of the topics raised in this brief review offer broad avenues for future research.

## NOTES

\* Anyone working on a topic like this will find themselves inevitably ‘wrestling with Manny’. He has made so many pertinent remarks, but like any pioneer, left so many half-turned stones, that one can only follow the trail and try and look under a few of them. Schegloff (2007b) promises a book on the subject of this chapter, which therefore has a strict sell-by date. I have had much help on this chapter from Tanya Stivers, Nick Enfield, Paul Drew, Jack Sidnell and Penelope Brown for which I am most grateful.

1 See Sacks, Schegloff and Jefferson (1974: 728):

The turn-taking system has, as a by-product of its design, a proof procedure for the analysis of turns. When A addresses a first pair-part such as a ‘question’ or a ‘complaint’ to B, we have noted, A selects B as next speaker, and selects for B that he next perform a second part for the ‘adjacency pair’ A has started, i.e. an ‘answer’ or an ‘apology’ (among other possibilities) respectively. B, in so doing, not only performs that utterance-type, but thereby displays (in the first place to his co-participants) his understanding of the prior turn’s talk as a first part, as a ‘question’ or ‘complaint’.

- 2 But see Schegloff (1996a: 173, fn. 6), which makes the point that the possibility of clear-cut action misunderstandings shows that there are objective features of action coding: the action ascription may be wrong, but the reasons it went wrong can usually be detected.
- 3 The ‘proof procedure’ is thus actually a ‘disproof procedure’ (as Nick Enfield points out to me), an opportunity in third turn for a correction of a misunderstanding of a first turn displayed in a second.
- 4 At the ICCA meeting in Mannheim, July 8, 2010, Federico Rossano showed that bonobos enact exactly such a paired sequence of actions, giving us a glimpse of the deep phylogeny of our interactional structure.
- 5 Searle was here probably following Grice’s unpublished work on deriving ‘mood’ from classes of intention.
- 6 An additional point made by these examples is that withholding a turn or an overt action can also perform an action, as can minimal particles like *hmm*, *Oh*, *Well*, etc.
- 7 Many other actions have received CA attention. See, for example, Drew and Holt (1988) and Drew (1998) on the format of complaints.
- 8 Curl and Drew (2008) thus offer the new dimension of contingency to the account by Brown and Levinson (1987) which in effect sketches a theory of entitlement, factoring in the social relationship and the ‘weight’ of the action being performed.
- 9 The term *practice* is here used loosely. There seem to be two basic uses in CA: one that dwells purely on the format of turns, and another that links the format to an action so that a practice is partly defined in terms of the action executed by the deployment of the form (cf. Schegloff, 2007b: 231, who writes of “practices of implementing a course of action-in-interaction through talking”).
- 10 Schegloff (2007b: 197) notes that this can be characterized in terms of two overlapping adjacency pairs, with the second part of the first pair combining with the first part of the second, thus chaining sequences. Less trivial cases of complex turns are dispreferred responses, which may consist of multiple TCUs but perform only one main action, e.g. rejection.
- 11 Schegloff (in oral presentations on action formation) has pointed out that there is another way in which actions can be layered, e.g. in ironic or joke formulations. See Sidnell (2010a: 70–2) on irony.
- 12 If projects were only foreseeable sequences in the making, perhaps there would be no need for the concept. But the way in which, for example, the offering project in (18) unfolds, suggests something much more inchoate than a practice instantiated in a sequence.

- 13 Cf. Drew (2005a: 78):

But in important respects nonverbal conduct is subordinate to the verbal conduct with which it is intermeshed; it's probably true to say that none of the practices, devices or patterns identified in CA research are shaped or altered in any significant ways by accompanying nonverbal conduct.

- 14 The strong 'sequential relevance' of a proffered hand is reflected in the snubs recorded in the media (as when John McCain failed to shake Barack Obama's hand, <http://www.youtube.com/watch?v=LIUSa-ufwNY>), or equally when the second part is inappropriately produced (as when Prince Charles shook Robert Mugabe's hand without reflection at the Pope's funeral, [http://news.bbc.co.uk/2/hi/uk\\_news/4425385.stm](http://news.bbc.co.uk/2/hi/uk_news/4425385.stm)). Thanks to Paul Drew and Nick Enfield for the examples.
- 15 There is a great deal of scholarship on action tiers or hierarchies of plans from Philosophy, to Psychology, to Robotics. The way that AI programs assign actions to utterances using such hierarchies can be seen in Allen and Litman (1990). See Clark & Schaefer (1987) for an application to action recognition.
- 16 See Selting (1996b) for a study of the marked prosody of repair initiators that are doing more than just asking for repair (e.g. indicating surprise or astonishment or outrage).

UNCORRECTED PROOF