6

Conversational structure[1]

6.0 Introduction

In this Chapter we shall be centrally concerned with the organization of conversation. Definitions will emerge below, but for the present **conversation** may be taken to be that familiar predominant kind of talk in which two or more participants freely alternate in speaking, which generally occurs outside specific institutional settings like religious services, law courts, classrooms and the like.

It is not hard to see why one should look to conversation for insight into pragmatic phenomena, for conversation is clearly the prototypical kind of language usage, the form in which we are all first exposed to language – the matrix for language acquisition. Various aspects of pragmatic organization can be shown to be centrally organized around usage in conversation, including the aspects of deixis explored in Chapter 2 where it was shown that unmarked usages of grammatical encodings of temporal, spatial, social and discourse parameters are organized around an assumption of co-present conversational participants. Presupposition may also be seen as in some basic ways

[1] Illustrative data in this Chapter are drawn where possible from published sources so that readers can refer to them for additional context or further discussion; in those cases the source heads each extract. Where this has not been possible data have been drawn from transcripts circulated by workers in conversation analysis, such sources being indicated by customary identifying initials (e.g. US, DCD), a large proportion of which have been transcribed by Gail Jefferson; otherwise, data headed by a number (e.g. 176B) are drawn from the author's collection, some of which were transcribed by Marion Owen. Data without a heading are constructed for illustrative purposes unless otherwise indicated in the text. It has not been possible to check transcripts against the original recordings, so there may only be a modicum of consistency in the use of transcription conventions (see Appendix to this Chapter).

organized around a conversational setting: the phenomena involve constraints on the way in which information has to be presented if it is to be introduced to particular participants with specific shared assumptions and knowledge about the world. The issues touch closely on the distinction between **given** and **new** (see e.g. Clark & Haviland, 1977), and concern constraints on the **formulation** of information (that is, the choice of just one out of the indefinitely many possible descriptions of some entity – see Schegloff, 1972b), both of which are important issues in conversational organization. Similarly, implicatures derive from specific assumptions about conversational context: they do not always arise in the same way in all kinds of discourse – rather they are typical of conversation (yet, as we have seen, they have general grammatical reflexes, as in the constraints they impose on lexicalization). In the same way it may be argued that many kinds of speech act are built on the assumption of a conversational matrix – betting, for example, requires *uptake* to be effective, so that the utterance of *I bet you six pence* does not succeed without the interactional ratification typical of conversation. Indeed the conversational dependence of illocutionary force is such that the concept itself can be claimed to be substantially replaced by concepts of conversational function, as we shall see.

Nearly all the pragmatic concepts we have reviewed so far can thus be claimed to tie in closely with conversation as the central or most basic kind of language usage. Now if, as we shall argue, the proper way to study conversational organization is through empirical techniques, this suggests that the largely philosophical traditions that have given rise to pragmatics may have to yield in the future to more empirical kinds of investigation of language usage. Conceptual analysis using introspective data would then be replaced by careful inductive work based on observation. The issue raised here is whether pragmatics is in the long term an essentially empirical discipline or an essentially philosophical one, and whether the present lack of integration in the subject is due primarily to the absence of adequate theory and conceptual analysis or to the lack of adequate observational data, and indeed an empirical tradition. So far, in this book, we have reviewed the philosophically rooted traditions, but in this Chapter we turn to the outstanding empirical tradition in pragmatics. First, however, we should make clear the reasons for preferring this tradition to other approaches to the study of conversation.

6.1 Discourse analysis versus conversation analysis

In this section some different approaches to the study of conversation are assessed. At the risk of oversimplification, there can be considered to be two major approaches to the analysis of conversation, which we shall designate **discourse analysis** and **conversation analysis** (other distinctive approaches exist, of which the most important is probably the modelling of conversation using computer programs instead of human participants, as yet in its infancy – but see e.g. Power, 1979). Both approaches are centrally concerned with giving an account of how coherence and sequential organization in discourse is produced and understood. But the two approaches have distinctive and largely incompatible styles of analysis, which we may characterize as follows.

Discourse analysis (or DA) employs both the methodology and the kinds of theoretical principles and primitive concepts (e.g. *rule*, *well-formed formula*) typical of linguistics. It is essentially a series of attempts to extend the techniques so successful in linguistics, beyond the unit of the sentence. The procedures employed (often implicitly) are essentially the following: (a) the isolation of a set of basic categories or units of discourse, (b) the formulation of a set of concatenation rules stated over those categories, delimiting well-formed sequences of categories (coherent discourses) from ill-formed sequences (incoherent discourses). There are a number of other features that tend to go with these. There is typically an appeal to intuitions, about, for example, what is and what is not a coherent or well-formed discourse (see e.g. Van Dijk, 1972; Labov & Fanshel, 1977: 72). There is also a tendency to take one (or a few) texts (often constructed by the analyst) and to attempt to give an analysis in depth of all the interesting features of this limited domain (to find out, as some have put it, "what is really going on" – Labov & Fanshel, 1977: 59, 117). Into this broad avenue of work fall not only (and most obviously) the **text grammarians** (like Petöfi and Van Dijk – see de Beaugrande & Dressler, 1981: 24ff for review), but also the rather different work based on speech acts (or related notions) of researchers such as Sinclair & Coulthard (1975), Longacre (1976b), Labov & Fanshel (1977) and Coulthard & Brazil (1979).

In contrast, *conversation analysis* (or CA), as practised by Sacks, Schegloff, Jefferson, Pomerantz and others, is a rigorously empirical approach which avoids premature theory construction (see the

collections in Schenkein, 1978; Psathas, 1979; Atkinson & Heritage, in press). The methods are essentially *inductive*; search is made for recurring patterns across many records of naturally occurring conversations, in contrast to the immediate categorization of (usually) restricted data which is the typical first step in DA work. Secondly, in place of a theoretical ontology of *rules* as used in syntactic description, we have an emphasis on the interactional and inferential consequences of the choice between alternative utterances. Again in contrast to DA, there is as little appeal as possible to intuitive judgements – they may, willy-nilly, guide research, but they are not explanations and they certainly do not circumscribe the data; the emphasis is on what can actually be found to occur, not on what one would guess would be odd (or acceptable) if it were to do so. Intuition, it is claimed, is simply an unreliable guide in this area, as indeed it may be in other areas of linguistics (see e.g. Labov, 1972a). There is also a tendency to avoid analyses based on single texts. Instead, as many instances as possible of some particular phenomena are examined across texts, not primarily to illuminate "what is really going on" in some interaction (a goal judged impossible, such illuminations evading participants as well as analysts on many occasions), but rather to discover the systematic properties of the sequential organization of talk, and the ways in which utterances are designed to manage such sequences.

Which is the correct manner in which to proceed? The issue is a live one: DA theorists can accuse CA practitioners of being inexplicit, or worse, plain muddled, about the theories and conceptual categories they are actually employing in analysis (see e.g. Labov & Fanshel, 1977: 25; Coulthard & Brazil, 1979); CA practitioners can retort that DA theorists are so busy with premature formalization that they pay scant attention to the nature of the data. The main strength of the DA approach is that it promises to integrate linguistic findings about intra-sentential organization with discourse structure; while the strength of the CA position is that the procedures employed have already proved themselves capable of yielding by far the most substantial insights that have yet been gained into the organization of conversation.

There may well seem to be room for some kind of accommodation or even synthesis between the two positions; however there are some reasons to think that the DA approach as outlined is fundamentally

misconceived. We may start by noting that DA analysts can be divided into two basic categories – the text grammarians and the speech act (or interactional) theorists. The text grammarians believe, at least in the simplest formulations, that discourses can be viewed simply as sentences strung together in much the same way that clauses within sentences can be conjoined with connectives of various kinds. It follows that there are no problems for discourse analysis that are not problems for sentential analysis – "discourse can be treated as a single sentence in isolation by regarding sentence boundaries as sentential connectives" (Katz & Fodor, 1964: 490; see critique in Edmondson, 1978, 1979). However adequate such a view may be for written non-dialogic text, it is simply not feasible as a model for conversation where the links between speakers cannot be paraphrased as sentential connectives – for example (1) does not paraphrase as (2):

(1) A: How are you?
B: To hell with you

(2) How are you and to hell with you

(3) Anne said "How are you?" and Barry replied "To hell with you"

Even if (1) can be reported as (3), this shows nothing about the reducibility of (1) to (3), but merely that like all other kinds of events conversations are reportable (*contra* Katz & Fodor, 1964: 491).

The DA theorists that are therefore of interest to us are those who have been specifically concerned with conversation as a particular type of discourse, and we shall devote the rest of this section to a critique of their basic methods and assumptions. Here there is a remarkable underlying uniformity of views, a basic assumption (probably right as far as it goes) that the level at which coherence or order in conversation is to be found is not at the level of linguistic expressions, but at the level of the speech acts or the interactional moves that are made by the utterance of those expressions. Or, as Labov & Fanshel (1977: 70) put it: "obligatory sequencing is not to be found between utterances but between the actions that are being performed". It is thus possible to formulate the general properties of the whole class of models to which, in one guise or another, most DA theorists of conversation would subscribe (see e.g. Labov, 1972b; Sinclair & Coulthard, 1975; Longacre, 1976b; Labov & Fanshel, 1977; Coulthard & Brazil, 1979; Edmondson, 1981):

(4)
(i) There are unit acts – *speech acts* or *moves* – that are performed in speaking, which belong to a specifiable, delimited set
(ii) Utterances[2] are segmentable into unit parts – *utterance-units* – each of which corresponds to (at least) one unit act
(iii) There is a *specifiable function*, and hopefully a *procedure*, that will map utterance units into speech acts and vice versa
(iv) Conversational sequences are primarily regulated by a set of *sequencing rules* stated over speech act (or move) types

The kernel idea here is both simple and highly plausible: since sequential constraints are clearly not easily stated on the form or meaning of what is said, utterances have to be 'translated' into the underlying actions they perform, because on this deeper (or more abstract) level rules of sequencing will be straightforwardly describable. Such a model seems to capture the obvious regularities of the sort that answers generally follow questions, actions or excuses follow requests, acceptances or rejections follow offers, greetings follow greetings, and so on. The difficulties are thus generally considered to lie at the level of (iii) above, the translation from utterances into acts – "the rules of production and interpretation ... are quite complex; the sequencing rules are relatively simple" (Labov & Fanshel, 1977: 110) – and the various theories of indirect speech acts are therefore a focus of interest.

If the view is right then we can build up a model of conversation from a linguistic base by utilizing (while improving) the basic notions of speech act theory, merely adding a *syntax* for the concatenation of speech act categories that will capture the simple regularities noted above.

However there are some strong reasons to believe that such models are fundamentally inappropriate to the subject matter, and thus irremediably inadequate. Some of these have to do with the general problems that beset speech act theory, which we have already reviewed in Chapter 5. But in fact there are severe problems for each

[2] As noted in Chapter 1, considerable ambiguities attend the use of this term. We have generally used the term in prior Chapters to denote a sentence-context pair; however here, and generally elsewhere in this Chapter, it is being used in the sense of a product of an act of utterance, occurring within a *turn* (see below) at talk. On the notion *utterance-unit* see Lyons, 1977a: 633ff; Goodwin, 1981: 25ff.

of the basic assumptions in (4), which should be briefly indicated (see also Levinson, 1981a, 1981b).

First, there are a number of problems with assumption (4)(i). One of these is that some single-sentence utterances clearly perform more than one speech act at a time (if the notion of a speech act is to capture at least what utterances conventionally achieve) – consider, for example, the first utterance in the following exchange:

(5) A: Would you like another drink?
B: Yes I would, thank you, but make it a small one

The first utterance seems to be both a question and an offer, as indicated by the response. Now such multiple functions are not in principle problematic for assumptions (4)(i) and (4)(iii), but as they accumulate they do render the whole model considerably less attractive. How, for example, are the sequencing rules in (iv) to operate if more acts are being done than can feasibly be responded to directly? Moreover, as we shall see, the sources for multiple functions often lie outside the utterance in question, in the sequential environment in which it occurs; but such environments are not obviously restricted in kind, so that the existence of a well-defined and delimited set of speech act types, as required by the model, is quite dubious.

However, more problematic for the assumption in (4)(i) is the fact that conversational responses can be directed not just to the *illocutions* performed by utterances, but to their *perlocutions* too. Suppose, for example, that A and his companion B are at a party, and A being bored says to B:

(6) A: It's getting late, Mildred
B: a. But I'm having such a good time
b. Do you want to go?
c. Aren't you enjoying yourself, dear?

Then B might reply in any of the ways indicated, but none of these addresses the illocutionary force of A's utterance; rather they respond to a number of possible perlocutionary intents that A might have had. But this is highly problematic for the species of model in question: for perlocutions are unlimited in kind and number and any responses based on them will necessarily fall outside the scope of such a model.

There are serious problems too with (4)(ii), the requirement that there be identifiable utterance-units upon which speech acts or moves

can be mapped. Single sentences can be used to perform two or more speech acts in different clauses, and each clause (as we have seen) may perform more than one speech act. Further, there are many sub-sentential units that occur as utterances, and it is possible for non-linguistic vocalizations (e.g. laughter), non-vocal actions (like handing someone something requested), and sheer silence (e.g. after a loaded question) to perform appropriate responses to utterances. The problem is that in order for the function in property (4)(iii) to be well-formed, there must be an independently specifiable set of utterance-units onto which actions can be mapped. But in fact it is impossible to specify in advance what kinds of behavioural units will carry major interactional acts; rather the units in question seem to be functionally defined by the actions they can be seen to perform in context.

The requirement (4)(iii), therefore, inherits two problems: for a function to map actions onto utterance-units, there must be well-defined sets of (a) relevant actions and (b) relevant utterance-units. But we have seen that there are not. In addition, for this kind of model to have any real interest, we require not merely an abstract function, but an actual procedure or algorithm that will implement the function. But here we shall be even more disappointed, for as our discussion of speech act theory in Chapter 5 showed, there simply is no simple form-to-force correlation, and the attempts to bridge the gap (between what utterances 'literally' mean and 'actually' do in the way of actions) with theories of indirect speech acts have provided at best only partial solutions. For questions of context, both sequential (or discourse) context and extra-linguistic context, can play a crucial role in the assignment of utterance function. We can expect, therefore, no simple 'force conversion' rules to supply a general solution here, but rather some immensely complex inferential process that utilizes information of many different kinds. In the present state of knowledge, proponents of the kind of model outlined in (4) cannot expect to have even the general outlines of such an algorithm.

But this has an unfortunate consequence for such models, namely that they are unfalsifiable, and therefore essentially vacuous. The reasoning is this: suppose I claim (in accordance with the final assumption in (4)) that, given some set of speech act types or moves (let us call them X, Y and Z), only some sequences of these are *well-formed* or coherent sequences (say, XYZ, XZ, YXX) while all

others (like *ZXY, *XYX, *ZX, etc.) are *ill-formed* sequences. Then in order to be able to disprove this hypothesis it must be possible to test independently whether some sequence of utterances in fact corresponds to, say, the string XYZ. But such a test is only possible if there is an explicit procedure for assigning utterances to categories like X, Y and Z. And as there is no such procedure, there is no empirical content to the claim that strings of the form XYX do not or should not occur in discourse.

Finally we come to (4)(iv), the assumption that there is a set of sequencing rules, stated over speech act (or related) categories, which govern the sequential organization of conversation. This assumption is the motivating property of all such models, for the point of 'translating' utterances into the actions they perform is to reduce the problems of sequencing in conversation to a set of rules governing well-formed action sequences. The assumption embodies a strong claim about the 'syntactic' nature of sequential constraints in conversation, and essential to such a claim is that there should be clear cases of ill-formed sequences (like *XYX above) just as there are in sentence grammars (like **on cat the sat mat the*). Yet cases of such *impossible* discourses are hard if not impossible to find (see e.g. the successful contextualization by Edmondson, 1981: 12ff of the allegedly ill-formed discourses in Van Dijk, 1972). One reason for this is predicted by Grice's theory of implicature: any apparent conversational violation (e.g. a flouting of Relevance) is likely to be treated on the assumption that the utterances involved are in fact interpretable, if additional inferences are made (see Chapter 3 above). Another is that, as mentioned above, responses can be made to perlocutions, and perlocutions are not limited in kind and number and are not solely predictable from the utterances involved. A third is that our intuitions do not seem to be reliable guides in this area – sequences that we might judge 'ill-formed' in isolation do in fact frequently occur. Consider the following example (from Sacks, 1968, April 17):

(7) A: I have a fourteen year old son
B: Well that's all right
A: I also have a dog
B: Oh I'm sorry

which can seem in isolation quite bizarre, but when re-embedded in the actual conversation from which it is taken – in which A is raising a series of possible disqualifications for apartment rental with the

landlord B – it will seem natural and indeed quite unremarkable. So the fundamental basis for the postulation of general sequencing rules, namely the existence and predictability of ill-formed sequences, is seriously called into question.

It is an initial consideration of paired utterances like questions and answers, offers and acceptances (or rejections), greetings and greetings in response, and so on, that motivates the sequencing rules approach. But not only is conversation not basically constituted by such pairs (cf. Coulthard, 1977: 70) but the rules that bind them are not of a quasi-syntactic nature. For example, questions can be happily followed by partial answers, rejections of the presuppositions of the question, statements of ignorance, denials of the relevance of the question, and so on, as illustrated below:

(8) A: What does John do for a living?
B: a. Oh this and that
b. He doesn't
c. I've no idea
d. What's that got to do with it?

Rather we want to say that given a question an answer is relevant, and responses can be expected to deal with this relevance (see the explication of the notion of **conditional relevance** in 6.2.1.2 below). Such expectations are more like the maxims proposed by Grice, with their associated defeasible inferences, than like the rule-bound expectancy of an object after a transitive verb in English. This is made clear by, for example, the fact that in conversation inventive co-operative responses following questions may be preferable to answers:

(9) A: Is John there?
B: You can reach him at extension thirty-four sixty-two

Finally we should note that sequencing constraints in conversation could in any case never be captured fully in speech act terms. What makes some utterance after a question constitute an answer is not only the nature of the utterance itself but also the fact that it occurs after a question with a particular content – 'answerhood' is a complex property composed of sequential location and topical coherence across two utterances, amongst other things; significantly, there is no proposed illocutionary force of answering. But the model in question skirts the puzzling issue of constraints on topical coherence, despite

the fact that their relevance to issues of conversational sequencing is made clear by examples like (7). It seems, then, that it is doubtful that there are rules of a syntactic sort governing conversational sequencing, and that even if such rules could be found they would not give anything but a partial account of constraints on conversational sequences.

The conclusion that can be drawn is that all of the models that fall within the class having the general properties outlined in (4) are beset with fundamental difficulties. In addition, the actual analyses offered within theories of this kind are often quite superficial and disappointing, involving an intuitive mapping of unmotivated categories onto a restricted range of data. Even where this is not so (as in the major work by Labov & Fanshel, 1977), the analyses can often be shown to have obscured basic features of conversational organization (see e.g. the re-analysis of their data in (104) below).

It seems reasonable, then, to turn to CA as the approach that, at least at present, has most to offer in the way of substantial insight into the nature of conversation. It is important to see, though, that the basis for the rejection of DA is that the methods and theoretical tools advocated, namely those imported from mainstream theoretical linguistics, seem quite inappropriate to the domain of conversation. Conversation is not a structural product in the same way that a sentence is – it is rather the outcome of the interaction of two or more independent, goal-directed individuals, with often divergent interests. Moving from the study of sentences to the study of conversations is like moving from physics to biology: quite different analytical procedures and methods are appropriate even though conversations are (in part) composed of units that have some direct correspondence to sentences.

6.2 Conversation analysis[3]

Conversation analysis of the sort that will be described in the rest of this Chapter has been pioneered by a break-away group

[3] This Chapter, though relatively long because of the need to cite a considerable amount of data, is only a preliminary introduction. It may be supplemented with the introductory Chapters of Atkinson & Drew, 1979; the exemplary papers by Schegloff & Sacks (1973), Schegloff (1976); and the collections in Schenkein, 1978; Psathas, 1979; Atkinson & Heritage, in press. See also the introduction by Coulthard (1977). It should also be noted that for expositional purposes I have presented in a bold and simplified way a number of findings that are still treated as working hypotheses in conversation analysis.

of sociologists, often known as **ethnomethodologists**. The relevance of the sociological background to the pragmaticist is the methodological preferences that derive from it. The movement arose in reaction to the quantitative techniques, and the arbitrary imposition on the data of supposedly objective categories (upon which such techniques generally rely), that were typical of mainstream American sociology. In contrast, it was argued cogently, the proper object of sociological study is the set of techniques that the members of a society themselves utilize to interpret and act within their own social worlds – the sociologist's 'objective' methods perhaps not really being different in kind at all. Hence the use of the term **ethnomethodology**, the study of 'ethnic' (i.e. participants' own) methods of production and interpretation of social interaction (see Garfinkel, 1972; Turner, 1974a). Out of this background comes a healthy suspicion of premature theorizing and *ad hoc* analytical categories: as far as possible the categories of analysis should be those that participants themselves can be shown to utilize in making sense of interaction; unmotivated theoretical constructs and unsubstantiated intuitions are all to be avoided. In practice this results in a strict and parsimonious structuralism and a theoretical asceticism – the emphasis is on the data and the patterns recurrently displayed therein.

The data consist of tape-recordings and transcripts of naturally occurring conversation, with little attention paid to the nature of the context as that might be theoretically conceived within sociolinguistics or social psychology (e.g. whether the participants are friends or distant acquaintances, or belong to a certain social group, or whether the context is formal or informal, etc.).[4] As anyone who works on conversational data knows, heavy reliance inevitably comes to be placed on transcriptions and, as in phonetics, issues immediately arise here as to how **broad** or **narrow** such transcriptions should be, what notational systems should be used, and to what extent the exercise of transcription itself embodies theoretical decisions (see Ochs, 1979d). Excerpts from transcripts will here be given in the notation generally utilized in conversation analysis and listed in the Appendix

[4] It is not that the relevance of these factors is denied *a priori*, but simply that it is not assumed – if participants themselves can be rigorously shown to employ such categories in the production of conversation, then they would be of interest to CA. See e.g. Jefferson, 1974: 198.

to this Chapter: standard orthography will be used in some places where linguists might prefer phonetic transcription, and there is not, unfortunately, an adequate treatment of prosodic, and especially intonational, cues.[5]

In section 6.2.1 we shall present some of the most basic findings that have resulted from this kind of work. These findings are not in themselves, perhaps, of a very surprising sort, but we will then go on to show in later sections (especially, 6.2.2 and 6.2.3) that these apparently disparate little facts about conversation all fit together in a systematic way, and it is only then that one can begin to see that conversation has in fact an elaborate and detailed architecture.

One important *caveat* should be made immediately. The work here reviewed is based almost entirely on English data, especially telephone conversations and group talk, and we simply do not know at the present to what extent these findings extend to other languages and cultures. But although the findings here may be in part culturally specific, the methods employed should be of quite general application.

6.2.1 *Some basic findings*

6.2.1.1 Turn-taking

We may start with the obvious observation that conversation is characterized by **turn-taking**: one participant, A, talks, stops; another, B, starts, talks, stops; and so we obtain an A–B–A–B–A–B distribution of talk across two participants. But as soon as close attention is paid to this phenomenon, how such a distribution is actually achieved becomes anything but obvious. First there are the surprising facts that less (and often considerably less) than 5 per cent of the speech stream is delivered in **overlap** (two speakers speaking simultaneously), yet gaps between one person speaking and another

[5] Workers in CA have sometimes used *ad hoc* orthography to represent segmental features, to the irritation of linguists, although no serious theoretical issues seem to be involved (see Goodwin, 1977: 120, 1981: 47). I have taken the considerable liberty of standardizing the orthography of transcripts, but only where non-native speakers might otherwise have difficulty interpreting the text. Punctuation marks are also used by workers in CA to give some indication of intonation (see Appendix) and the original punctuation has therefore been reproduced in examples taken from these printed sources. One hopes that in future work a better system of prosodic transcription will be adopted (as used in e.g. the British tradition by Crystal (1969); O'Connor & Arnold (1973); Brazil, Coulthard & Johns (1980)).

starting are frequently measurable in just a few micro-seconds and they average amounts measured in a few tenths of a second (see Ervin-Tripp, 1979: 392 and references therein). How is this orderly transition from one speaker to another achieved with such precise timing and so little overlap? A second puzzle is that, whatever the mechanism responsible, it must be capable of operating in quite different circumstances: the number of parties may vary from two to twenty or more; persons may enter and exit the pool of participants; turns at speaking can vary from minimal utterances to many minutes of continous talk; and if there are more than two parties then provision is made for all parties to speak without there being any specified order or 'queue' of speakers. In addition the same system seems to operate equally well both in face-to-face interaction and in the absence of visual monitoring, as on the telephone.

Sacks, Schegloff & Jefferson (1974, 1978) suggest that the mechanism that governs turn-taking, and accounts for the properties noted, is a set of rules with ordered options which operates on a turn-by-turn basis, and can thus be termed a **local management system**. One way of looking at the rules is as a sharing device, an 'economy' operating over a scarce resource, namely control of the 'floor'. Such an allocational system will require minimal units (or 'shares') over which it will operate, such units being the units from which **turns** at talk are constructed. These units are, in this model, determined by various features of linguistic surface structure: they are syntactic units (sentences, clauses, noun phrases, and so on) identified as turn-units in part by prosodic, and especially intonational, means. A speaker will be assigned initially just one of these **turn-constructional units** (although the extent of the unit is largely within the speaker's control due to the flexibility of natural language syntax). The end of such a unit constitutes a point at which speakers may change – it is a **transition relevance place**, or TRP. At a TRP the rules that govern the transition of speakers then come into play, which does not mean that speakers will change at that point but simply that they may do so, as we shall see. The exact characterization of such units still requires a considerable amount of linguistic work (see Goodwin, 1981: 15ff), but whatever its final shape the characterization must allow for the **projectability** or predictability of each unit's end – for it is this alone that can account for the recurrent marvels of split-second speaker transition.

There is one other feature of turn-units that has to be mentioned before the rules can be presented, namely the possibility of specifically indicating within such a unit that at its end some particular other party is invited to speak next. Techniques for selecting next speakers in this way can be quite elaborate, but include such straightforward devices as the following: a question (offer, or request, etc.) plus an address term; a tagged assertion plus an address feature; and the various hearing and understanding checks (*Who?*, *You did what?*, *Pardon?*, *You mean tomorrow?*, etc.) which select prior speaker as next.

Operating on the turn-units are the following rules (slightly simplified from Sacks, Schegloff & Jefferson, 1978), where C is current speaker, N is next speaker, and TRP is the recognizable end of a turn-constructional unit:

(10) Rule 1 – applies initially at the first TRP of any turn
- (a) If C selects N in current turn, then C must stop speaking, and N must speak next, transition occurring at the first TRP after N-selection
- (b) If C does not select N, then any (other) party may self-select, first speaker gaining rights to the next turn
- (c) If C has not selected N, and no other party self-selects under option (b), then C may (but need not) continue (i.e. claim rights to a further turn-constructional unit)

Rule 2 – applies at all subsequent TRPs
When Rule 1(c) has been applied by C, then at the next TRP Rules 1 (a)–(c) apply, and recursively at the next TRP, until speaker change is effected

It may be asked whether Rule 1(c) is not just a special case of Rule 1(b), and therefore redundant. However there is some evidence that the self-selecting parties in Rule 1(b) should not properly include current speaker (C): for example, the delays between two turns by different speakers are statistically shorter than between two turn-constructional units produced by a single speaker, suggesting that opportunity for others to speak is specifically provided by Rule 1(b) (see Sacks, Schegloff & Jefferson, 1978: 54 n.30).

Careful consideration will show that the rules provide for the basic observations already noted. On the one hand they predict the following specific details. First, only one speaker will generally be speaking at any one time in a single conversation (although four or more speakers may often conduct more than one conversation

simultaneously). However where overlaps do occur, they can be predicted to be, at least in the great majority of cases, precisely placed: overlaps will either occur as competing first starts, as allowed by Rule 1(b) and illustrated in (11), or they will occur where TRPs have been misprojected for systematic reasons, e.g. where a tag or address term has been appended as illustrated in (12), in which case overlap will be predictably brief. The rules thus provide a basis for the discrimination (which we all employ) between inadvertent overlap as in (11) and (12) and violative interruption as in (13):

(11) *Sacks, Schegloff & Jefferson, 1978: 16*
J: Twelve pounds I think wasn't it.=
D: =//Can you bel*ie*ve it?
L: Twelve pounds on the Weight Watchers' scale.

(12) *Sacks, Schegloff & Jefferson, 1978: 17*
A: Uh *you* been down here before // havenche.
B: Yeah.

(13) *DCD:22*
C: We:ll I wrote what I thought was a a-a
rea:s'n//ble explan*ati*o:n
→ F: I: think it was a *very* rude *le*:tter

It is also predicted that when silence – the absence of vocalization – occurs, it will be differentially assigned, on the basis of the rules, as either (i) a **gap** before a subsequent application of Rules 1(b) or 1(c), or (ii) a **lapse** on the non-application of Rules 1(a), (b) and (c), or (iii) a selected next speaker's **significant** (or **attributable**) **silence**[6] after the application of Rule 1(a). Thus in (14) we have first a *gap* by delay of the Rule 1(b) option for just one second, then a lapse of sixteen seconds:

(14) *Sacks, Schegloff & Jefferson, 1978: 25*
C: Well no I'll drive (I don't mi//nd)
J: hhh
→ (1.0)
J: I meant to *off*er.
→ (16.0)
J: Those shoes look nice ...

While in (15) we have two clear cases of *attributable silence*, by virtue

[6] Henceforth the term *silence* is sometimes used in this technical sense, while the term *pause* is used as a general cover term for these various kinds of periods of non-speech. Other usages will be clear from the context.

of the fact that A's utterances select B as next speaker, and by Rule 1(a) B should then speak:

(15) *Atkinson & Drew, 1979: 52*
A: Is there something bothering you or not?
→ (1.0)
A: Yes or no
→ (1.5)
A: Eh?
B: No.

While making such specific predictions, the rules also allow for the observable variations in conversation: lapses may or may not occur; there is no strict limit to turn size given the extendable nature of syntactic turn-constructional units and the continuations allowed for by Rule 1(c); there is no exclusion of parties; the number of parties to a conversation can change. These diverse variations are allowed for basically because the system is **locally managed**, i.e. it operates on a turn-by-turn basis, organizing just the transition from current speaker to next, and is therefore indifferent to, for example, the pool of potential next speakers.[7]

An important consequence of the system is that it provides, independently of content or politeness considerations, an intrinsic motivation for participants to both listen and process what is said – for the transition rules require prior location of next speaker selection should it occur, and the projection of upcoming TRPs.

Where, despite the rules, overlapping talk occurs, detailed study has revealed the operation of a resolution system that is integrated into the main turn-taking system. First, if overlap occurs, one speaker generally drops out rapidly, as in (16):

(16) *Atkinson & Drew, 1979: 44 (simplified)*
D: ... he's got to *talk* to someone (very sor) supportive way towards you (.)
→ A: //Greg's (got wha-)*
G: Think you sh* – think you should have *one* to: hold him

Secondly, as soon as one speaker thus emerges into 'the clear', he typically recycles precisely the part of the turn obscured by the overlap, as in G's turn in (16). Finally, if one speaker does not immediately drop out, there is available a competitive allocation

[7] Although such factors do influence, for example, the details of techniques for next-speaker selection.

system which works roughly on a syllable-by-syllable basis, whereby the speaker who 'upgrades' most wins the floor, upgrading consisting of increased amplitude, slowing tempo, lengthened vowels and other features, as illustrated in (17):

(17) *US: 43*
→ J: But dis // person thet *DID* IT* IS GOT TO BE::
V: If I see the person
J: .hh taken care of

There is, then, quite an elaborate back-up machinery for resolving overlap if, despite the rules, it should occur (see Jefferson & Schegloff, 1975).

It is important to see that, although the phenomenon of turn-taking is obvious, the suggested mechanism organizing it is not.[8] For a start, things could be quite otherwise: for example, it is reported of the African people, the Burundi (see Albert, 1972: 81ff), that turn-taking (presumably in rather special settings) is pre-allocated by the rank of the participants, so that if A is of higher social status than B, and B than C, then the order in which the parties will talk is A–B–C. Of course in English-speaking cultures too there are special non-conversational turn-taking systems operative in, for example, classrooms, courtrooms, chaired meetings and other 'institutional' settings, where turns are (at least in part) pre-allocated rather than determined on a turn-by-turn basis, and these too emphasize that the rules in (10) are not the only possible or rational solution to the organization of the 'economy' of turns at talk. Nevertheless, there is good reason to think that like many aspects of conversational organization, the rules are valid for the most informal, ordinary kinds of talk across all the cultures of the world. There is even evidence of ethological roots for turn-taking and other related mechanisms, both from work on human neonates (see e.g. Trevarthen, 1974, 1979) and primate research (see e.g. Haimoff, in press).

Another indication that the suggested mechanism is far from obvious is that psychologists working on conversation have suggested a quite different solution to how turn-taking works. According to this

[8] It is also worth pointing out that the motivation for turn-taking is not as obvious as it may seem: as Miller has noted (1963: 418) turn-taking "is not a necessary consequence of any auditory or physiological inability to speak and hear simultaneously; one voice is poor masking for another" (cited in Goodwin, 1977: 5). The possibility of simultaneous translation bears witness to this (see Goldman-Eisler, 1980).

other view, turn-taking is regulated primarily by *signals*, and not by opportunity assignment rules at all (see e.g. Kendon, 1967; Jaffé & Feldstein, 1970; Duncan, 1974; Duncan & Fiske, 1977). On such a view a current speaker will signal when he intends to hand over the floor, and other participants may bid by recognized signals for rights to speak – a practice similar to the 'over' announcement on a field radio transmitter. One of the most plausible candidates for such signals is **gaze** : it seems roughly true, for example, that a speaker will break mutual gaze while speaking, returning gaze to the addressee upon turn completion (Kendon, 1967; Argyle, 1973: 109, 202; but see contrary findings in Beattie, 1978a; and see Goodwin, 1977, 1981 for a CA approach to gaze). The problem here is that if such signals formed the basis of our turn-taking ability, there would be a clear prediction that in the absence of visual cues there should either be much more gap and overlap or that the absence would require compensation by special audible cues. But work on telephone conversation shows that neither seems to be true – for example, there is actually less gap and shorter overlap on the telephone (see Butterworth, Hine & Brady, 1977; Ervin-Tripp, 1979: 392), and there is no evidence of special prosodic or intonational patterns at turn-boundaries on the telephone (although there is evidence that such cues are utilized both in the absence and presence of visual contact to indicate the boundaries of turn-constructional units – see e.g. Duncan & Fiske, 1977). In any case it is not clear how a signal-based system could provide for the observed properties of turn-taking anyway: for example, a system of intonational cues would not easily accomplish the observable *lapses* in conversation, or correctly predict the principled basis of overlaps where they occur, or account for how particular next speakers are selected (see Goodwin, 1979b, 1981: 23ff). Therefore the signalling view, plausible as it is, viewed as a complete account of turn-taking seems to be wrong: signals indicating the completion of turn-constructional units do indeed occur, but they are not the essential organizational basis for turn-taking in conversation. That organization seems rather to be based on an opportunity assignment of the sort specified by the rules in (10).

Another possible view that also seems to be incorrect is that, while turn-taking is indeed an option-based system, the options are organized not around surface-structural units, as suggested by Sacks,

Schegloff & Jefferson (1978), but rather around functional units – speech acts, moves, or perhaps ideational units (as in Butterworth, 1975). Such a view has an initial plausibility: as a participant one should wait until one sees what interactional contribution the other party is making, and then perform one's own. Again, however, such a view makes the wrong predictions – for example, since greetings, expressions like *How are you?*, etc., are generally precisely predictable, they ought to get regularly overlapped, but this is not the case. Similarly, where a speaker fails to make himself audible or comprehensible to a recipient, requests for **repair** ought to occur immediately after the 'repairable', whereas in fact the initiation of repair generally awaits the next TRP (see Sacks, Schegloff & Jefferson, 1978: 39, and section 6.3.2 below). And in general, given the apparent projectability of other persons' utterances, we should expect the majority of turns to be completed in overlap – and of course such is not the case. So despite its plausibility, this view too seems to be wrong: turn-taking is firmly anchored around the surface-structural definition of turn-units, over which rules of the sort in (10) operate to organize a systematic distribution of turns to participants.

6.2.1.2 Adjacency pairs

We now turn to another local management organization in conversation, namely **adjacency pairs** – the kind of paired utterances of which question–answer, greeting–greeting, offer–acceptance, apology–minimization, etc., are prototypical. We have already noted that these are deeply inter-related with the turn-taking system as techniques for selecting a next speaker (especially where an address term is included or the content of the first utterance of the pair clearly isolates a relevant next speaker). Once again, the existence of such paired utterances is obvious, but a precise specification of the underlying expectations upon which the regularities are based is not so easy. Schegloff & Sacks (1973) offer us a characterization along the following lines:

(18) *adjacency pairs* are sequences of two utterances that are:
- (i) adjacent
- (ii) produced by different speakers
- (iii) ordered as a **first part** and a **second part**
- (iv) typed, so that a particular first part requires a particular second (or range of second parts) – e.g. offers require

acceptances or rejections, greetings require greetings, and so on

and there is a rule governing the use of adjacency pairs, namely:

(19) Having produced a first part of some pair, current speaker must stop speaking, and next speaker must produce at that point a second part to the same pair

Adjacency pairs seem to be a fundamental unit of conversational organization – indeed it has been suggested that they are *the* fundamental unit (see e.g. Goffman, 1976; Coulthard, 1977: 70). Such a view seems to underlie the speech act models of conversation reviewed in section 6.1 above. However there are many other kinds of more complex sequential organizations operating in conversation, as we shall see, nor indeed can the constraints across such pairs be properly modelled by formation rules analogous to syntactic rules. It is therefore important to see that the characterization of adjacency pairs in (18) and (19) is only a first approximation, and is in fact inadequate in a number of important respects.

There are problems with each of the conditions in (18), but we shall focus on (i), adjacency, and (iv), the kinds of expectable second parts. First, strict adjacency is actually too strong a requirement: there frequently occur **insertion sequences** (Schegloff, 1972a) like the following in which one question–answer pair is embedded within another (where Q1 labels the first question, A1 its answer, and so on):

(20) *Merritt, 1976: 333*

A: May I have a bottle of Mich? ((Q1))
B: Are you twenty one ? ((Q2))
A: No ((A2))
B: No ((A1))

or like the following where a notification of temporary interactional exit and its acceptance are embedded within a question–answer pair:[9]

(21) *144/6*

B: U:hm (.) what's the price now eh with V.A.T. do you know eh ((Q1))
A: Er I'll just work that out for you= ((HOLD))

[9] *Hold* and *accept(ance)* are *ad hoc* terms for the parts of the adjacency pair that are used to initiate an interactional interlude or 'time out'. Interaction may then, but need not, be re-initiated by another adjacency pair (*Hello?*; *Hello*).

B: =thanks ((ACCEPT))
(10.0)
A: Three pounds nineteen a tube sir ((A1))

Indeed numerous levels of embedding are not at all infrequent, with the consequence that, say, a question and its answer may be many utterances apart; nevertheless the relevance of the answer is merely held in abeyance while preliminaries are sorted out, and insertion sequences are thus restricted in content to the sorting out of such preliminaries. In fact (21) is extracted from the larger sequence of nested adjacency pairs in (22) (here R labels a request first part, Q and A question and answer, respectively, and turns are numbered T1, T2, etc., for reference):

(22) *144/6*

T1 B: ... I ordered some paint from you uh a couple of weeks ago some vermilion
T2 A: Yuh
T3 B: And I wanted to order some more the name's Boyd ((R1))
T4 A: Yes // how many tubes would you like sir ((Q1))
T5 B: An-
T6 B: U:hm (.) what's the price now eh with V.A.T. do you know eh ((Q2))
T7 A: Er I'll just work that out for you= ((HOLD))
T8 B: =Thanks ((ACCEPT))
(10.0)
T9 A: Three pounds nineteen a tube sir ((A2))
T10 B: Three nineteen is it = ((Q3))
T11 A: =Yeah ((A3))
T12 B: E::h (1.0) yes u:hm ((dental click)) ((in parenthetical tone)) e:h jus-justa think, that's what three nineteen
That's for the large tube isn't it ((Q4))
T13 A: Well yeah it's the thirty seven c.c.s ((A4))
T14 B: Er, hh I'll tell you what I'll just eh eh ring you back I have to work out how many I'll need. Sorry I did- wasn't sure of the price you see ((ACCOUNT FOR NO A1))
T15 A: Okay

A number of points may be parenthetically made here. First, insertion sequences, which are of great interest in their own right, can

effectively structure considerable stretches of conversation. So what is strictly a local system, operating over just two turns – namely adjacency pair organization – can by means of the accumulation of first pair parts project a large sequence of expectable seconds, as in the structure schematized in (23):

(23) (Q1(Q2(Q3(Q4–A4)A3)A2)A1)

Secondly, we should note that in (22) neither the initial request (R1) nor the first question (Q1) ever receives its second part (an acceptance or rejection, and an answer, respectively). Nevertheless what takes place after these two turns, T3 and T4, takes place under the umbrella of the expectation that the relevant second parts will be forthcoming. Finally in T14 an explanation or **account** is provided for the failure to provide an A1 for Q1, showing that there is an orientation to the expected appropriate second part even though it never occurs. Further, note that the acknowledged failure to produce an A1 is sufficient to explain the absence of any response to R1: failure to resolve an insertion sequence regularly aborts the entire umbrella sequence too.

But the main point is that we need to replace the strict criterion of adjacency with the notion of **conditional relevance**, namely the criterion for adjacency pairs that, given a first part of a pair, a second part is immediately relevant and expectable (Schegloff, 1972a: 363ff). If such a second fails to occur, it is noticeably absent; and if some other first part occurs in its place then that will be heard where possible as some preliminary to the doing of the second part, the relevance of which is not lifted until it is either directly attended to or aborted by the announced failure to provide some preliminary action. What the notion of conditional relevance makes clear is that what binds the parts of adjacency pairs together is not a formation rule of the sort that would specify that a question must receive an answer if it is to count as a well-formed discourse, but the setting up of specific expectations which have to be attended to. Hence the non-occurrences of an R1 and an A1 in (22) do not result in an incoherent discourse because their absences are systematically provided for.

A second kind of problem that arises with the notion of an adjacency pair concerns the range of potential seconds to a first part. Unless for any given first part there is a small or at least delimited

set of seconds, the concept will cease, it seems, to describe the tight organization in conversation that is its principal attraction. But in fact there are, for example, a great many responses to questions other than answers which nevertheless count as acceptable seconds (rather than, say, beginnings of insertion sequences prior to answers) – including protestations of ignorance, 're-routes' (like *Better ask John*), refusals to provide an answer, and challenges to the presuppositions or sincerity of the question (and see (8) above). For example, we noted in (22) that in T14, the slot for an answer to Q1, we have not an answer but a promise to provide an answer at a later date, together with an account explaining the deferral. So while responses to, for example, questions may be restricted, they certainly do not constitute a small set, and this does seem to undermine the structural significance of the concept of an adjacency pair.

However the importance of the notion is revived by the concept of **preference organization**. The central insight here is that not all the potential second parts to a first part of an adjacency pair are of equal standing: there is a ranking operating over the alternatives such that there is at least one **preferred** and one **dispreferred** category of response. It must be pointed out immediately that the notion of *preference* here introduced is not a psychological one, in the sense that it does not refer to speakers' or hearers' individual preferences. Rather it is a structural notion that corresponds closely to the linguistic concept of **markedness**. In essence, preferred seconds are **unmarked** – they occur as structurally simpler turns; in contrast dispreferred seconds are **marked** by various kinds of structural complexity. Thus dispreferred seconds are typically delivered: (a) after some significant delay; (b) with some preface marking their dispreferred status, often the particle *well*; (c) with some account of why the preferred second cannot be performed. For the present (but see 6.3) a contrastive pair of examples will suffice to illustrate the notion:

(24) *Wootton, in press*

Child: Could you .hh could you put on the light for my .hh room

Father: Yep

[10] In examples from telephone calls, where the roles of caller and receiver may be relevant to the interpretation, caller is labelled C, receiver R.

(25) *176B*[10]

C: Um I wondered if there's any chance of seeing you tomorrow sometime (0.5) morning or before the seminar

→ (1.0)

R: Ah um (.) I doubt it

C: Uhm huh

R: The reason is I'm seeing Elizabeth

In (24) a granting of a request is done without significant delay and with a minimal granting component *Yep*. In contrast in (25), a rejection of a request for an appointment is done after a one second delay, and then, after further delay components (*ah um*, the micro-pause (.)), by a non-minimal turn (compare *I doubt it* with *No*), followed by an account or reason for the difficulty. In fact, rejections of requests are normally done in this marked way. Thus we can say grantings are preferred seconds (or **preferreds** for short) to requests, rejections are dispreferred seconds (or **dispreferreds**). This is a general pattern: in contrast to the simple and immediate nature of preferreds, dispreferreds are delayed and contain additional complex components; and certain kinds of seconds like request rejections, refusals of offers, disagreements after evaluative assessments, etc., are systematically marked as dispreferreds.

Preference organization is described in detail in section 6.3 below, but the relevance here is that by ordering seconds as preferreds and dispreferreds, the organization allows the notion of an adjacency pair to continue to describe a set of strict expectations despite the existence of many alternative seconds to most kinds of first parts.[11]

6.2.1.3 Overall organization

We have now described two kinds of local organization operating in conversation – local in the sense that turn-taking and adjacency pair organization operate in the first instance across just two turns, current and next. But there are quite different orders of organization in conversation: for example, there are certain recurrent kinds of sequence definable only over three or four or more turns, like those treated in following sections that deal with **repair** (6.3) or begin with **pre-sequences** (6.4). Further, there are some that can be called **overall organizations** in that they organize the totality

[11] The exceptions here include greetings, where return greetings are more or less the only kind of second.

of the exchanges within some specific kind of conversation, and it is these that we shall illustrate here.

One kind of conversation with a recognizable overall organization that has been much studied is the telephone call. But it is not by virtue of 'being on the telephone' that such conversations have most of the features of overall organization that they display: rather they belong fairly clearly to a class of verbal interchanges that share many features, namely those that are social activities effectively constituted by talk itself, like a chat on a chance meeting in the street, or a talk over the garden fence. These tend to have clear beginnings and carefully organized closings. Thus in telephone calls we can recognize the following typical components of an **opening section**: the telephone rings and, upon picking up the receiver, the person at the receiving end almost invariably speaks first, either with a *station identification* (name of a firm, a telephone number, etc.) or a plain *Hello*, whereupon the caller produces a *Hello*, often with a self-identification. If the call is between two friends or acquaintances we may expect an exchange of *How are you*s. Then at that point we expect some announcement from the caller of the reason for the call, and we thereby find ourselves projected into the substance of the call, and thus (as we shall see) into matters of topical organization.

To say this is to say little more than that telephone conversations have recognizable openings. But there is much elaborate structure here. For a start we may note that such openings are constructed largely from adjacency pairs: thus we typically get paired *Hello*s as an exchange of greetings, we may get self-identifications with paired recognitions, and an exchange of *How are you*s each with their paired responses (see Schegloff, 1972a, 1979a; Sacks, 1975, respectively, for each of these). There is, moreover, a puzzle about why the receiver, the person with the least information about the identity and purposes of the other, almost invariably talks first. The puzzle dissolves when we assimilate the openings of telephone conversations to **summons–answer** sequences. Such sequences in face-to-face interaction run typically in any of the following ways:

(26) *Terasaki, 1976: 12,13*

(a) A: Jim?
B: Yeah?

(b) A: Mo:m
B: What?

(c) A: ((knock knock knock))
B: Come in::

(27) *Atkinson & Drew, 1979: 46*

Ch: Mummy

M: Yes dear
(2.1)
Ch: I want a cloth to clean (the) windows

where the first utterance (or action) is a **summons**, the second an **answer** to the summons, the exchange establishing an open channel for talk. Schegloff (1972a) suggests that the ringing of the telephone is the summons component in such an adjacency pair, so that the first turn at talk (the receiver's *Hello*) is actually the second interactional move. This explains a number of features of telephone openings, including the strong compulsion to respond, and the reportable inference which motivates it – namely that (by conditional relevance) no response 'means' that 'no one is at home'. It even explains the mechanical ring–pause–ring, which imitates the recursive repetition of a verbal summons that is not attended to. That repetition is in turn the basis of the rare exceptions to the generalization that the receiver speaks first, for these occur where the receiver upon picking up the telephone after the first (mechanical) summons, fails to respond – we then get a repeated summons (now verbal) from the caller.

A moment's consideration will show too that summons–answer sequences are a little different from other adjacency pairs (like greetings–greetings, offers–acceptances/refusals) in that they are always a prelude to something. Moreover the something in question can be expected to be produced by the summoner as the reason for the summons. So summons–answer sequences are actually elements of (minimally) three-turn sequences, as illustrated below (and in (27) above):

(28)

T1	A: John?	((SUMMONS))
T2	B: Yeah?	((ANSWER))
T3	A: Pass the water wouldja?	((REASON FOR SUMMONS))

The three-part structure is evidenced by the common use of question components in T2 (like *What?*, *What is it?*, *Yeah?*), which, by simultaneously being the second part to the summons and a first part requesting reasons for the summons, provide for a three-turn structure constructed out of two adjacency pairs. One may also note the obligation that the summoner often feels, for example, in calling a store to find if it is open, to produce a T3 (e.g. *Oh I was just calling to see if you were open*) even though the presence of a T2 was sufficient

to make the third turn redundant. It is the three-turn structure of such sequences that establishes not only the obligation for the summoner to produce a T3, but an obligation for a recipient who has produced a T2 to attend to a T3. The sequence thereby serves to establish the co-participation necessary to conversation.

One important feature of opening sections in telephone conversations is the immediate relevance, and the potential problems, of *identification* and *recognition* (Schegloff, 1979a). Many telephone conversations have as their first three turns the following, or something closely similar:

(29)

	C: ((causes telephone to ring at R's location))
T1	R: Hello
T2	C: Hi
T3	R: Oh hi::

Such openings illustrate a basic finding of CA, namely that a single minimal utterance or turn can be the locus of a number of quite different overlapping constraints – it can thus perform, and can be carefully designed to perform, a number of quite different functions at once. Here for example, T1, despite being the first turn in the conversation, is not (as we have seen) the first move in the interaction: the ring is the summons, and T1 its answer. But T1 is also simultaneously a display for recognitional purposes of recipient's identity (in cases where recognition is relevant, as not always, e.g. in business calls), and it is notable that speakers tend to use a 'signatured' prosody or voice-quality in this turn (Schegloff, 1979a: 67). Despite the apparent greeting token in T1, greeting is not what the turn appears to do, as the discussion of T3 will make clear. T2 on the other hand is indeed a greeting token that does greetings, and greetings being adjacency-paired, T2 gets a return greeting in T3 (this showing that T1 is hardly a greeting after all, greetings being in general not reiterable kinds of things). But that is not all, indeed the least, of what is going on in T2 and T3.[12] T2, by virtue of its minimal greeting form, actually claims recognition of the recipient on the sole basis of the voice-quality sample offered in T1; and moreover T2 claims that the recipient should likewise be able to recognize the caller on the basis of the minimal voice-quality sample it provides. T3 then, in

[12] Note that the *Oh* in T3, normally a marker of receipt of new information, only makes sense if more than greetings are going on in T2 and T3 (see Heritage, in press).

performing return greetings, also claims to have recognized the caller. The overlapping organizations here are thus: (a) telephone (and other related) conversations begin with summons–answer pairs; (b) reciprocal greetings are relevant at the very beginning of calls; (c) also at the very beginning of calls, recognition (or identification) is a prime concern. Note that T2 is the slot for recognitions to be begun, recipient clearly not being able to do this in T1 in the absence of any evidence of who the caller might be. And despite the total absence in (29) of any *overt* recognitional devices (e.g. *Hi, Sam*), the expectation, based on overall organization, of the recognitional relevance of T2 is strong enough invariably to impose on *Hi, Hello*, and other minimal greeting components in T2, a claim that recognition of the recipient by the caller has been achieved (see discussion of (45), (46), and (81)–(85) below, and Schegloff, 1979a). We may summarize this as follows:

(30)	C: ((rings))	((SUMMONS))
T1	R: Hello	((ANSWER)) + ((DISPLAY FOR RECOGNITION))
T2	C: Hi	((GREETINGS 1ST PART))
		((CLAIM THAT C HAS RECOGNIZED R))
		((CLAIM THAT R CAN RECOGNIZE C))
T3	R: Oh hi::	((GREETINGS 2ND PART))
		((CLAIM THAT R HAS RECOGNIZED C))

We are introduced here to the richness of the communicational content that is mapped onto minimal utterances by virtue of **sequential location** – here a location whose specificity is due to the structure of opening sections of the overall organization of telephone calls.

The opening section of a telephone call is usually followed in what may be called **first topic slot** by an announcement by the caller of the reason for the call:

(31) *Schegloff, 1979a: 47*
R: Hello.
C: Hello Rob. This is Laurie. How's everything.
R: ((sniff)) Pretty good. How 'bout you.
→ C: Jus' fine. The reason I called was ta ask ...

The first topic slot immediately after the opening section is a privileged one: it is the only one that is likely to be almost entirely free from topical constraints arising from prior turns. The main body of a call is thus structured by topical constraints: the content of the first slot is likely to be understood as the main reason for the call (whether or not, of course, from the point of view of the caller, it 'really' is),[13] and after that topics should by preference be 'fitted' to prior ones – topics therefore often being withheld until such a 'natural' location for their mention turns up (Schegloff & Sacks, 1973: 300ff). Evidence for this preference for linked transitions from topic to topic can be found in the common experience of having things to say that one never manages to get in, and more demonstrably in the *marked* nature of the other main kind of transition, unlinked topic 'jumps'. Thus, for example, in the arrowed utterance in (32), a topic jump is signalled in a typical way by the features of increased amplitude, raised pitch, markers of self-editing and hesitancy (see Schegloff, 1979b) and a marker of discontinuity, *Hey*.

(32) *163*
R: It's o – it's okay we'll pop down tomorrow Gertrude
C: You sure you don't, it is an awful lot of it, you want to quickly nip down now for it
→ R: Okay I will. Er *HEY* you hmm that is have you been lighting a fire down there?

Sacks remarks (1971, April 5) that the relative frequency of marked topic shifts of this sort is a measure of a 'lousy' conversation. Instead, what seems to be preferred is that, if A has been talking about X, B should find a way to talk about Z (if Z is the subject he wants to introduce) such that X and Z can be found to be 'natural' fellow members of some category Y. However it should not be thought from this that such co-class membership is somehow antecedently given; rather it is something that is actually achieved in conversation.

This last point needs a little elaboration. It has been suggested, very plausibly, that topic can be characterized in terms of **reference**: A and B are talking about the same topic if they are talking about the same things or sets of referents (see Putnam, 1958; but see Keenan

[13] In some cultures there seems to be a preference for displacing the business of a conversation to later on – however, one needs to distinguish here an elaboration of openings to include conventional inquiries about health, family and so on, from a true difference in the use of the first free topic slot.

& Schieffelin, 1976). Alternatively, we can say that A and B are talking about the same topic if they are talking about the same or linked **concepts** (de Beaugrande & Dressler, 1981: 104). However it is easy to show that co-referentiality, or a set of shared concepts, is neither sufficient nor necessary to establish topical coherence. Consider, for example:

(33) *Sacks, 1968, April 17: 16*
A: God any more hair on muh chest an' I'd be a fuzz boy.
B: 'd be a *what*.
C: A // fuzz boy.
A: Fuzz boy.
B: What's that.
A: Fuzz mop.
C: Then you'd have t'start shaving.
(1.0)
→ B: Hey I shaved this morni- I mean last night for you.

Here the last two utterances both mention shaving, and share that concept, and also on the logical analysis of predicates (see Allwood, Andersson & Dahl, 1977: 72ff) would share some of their referents.[14] But, as Sacks (1968, April 17) points out, B's utterance is produced in such a way as to indicate that it is *not* topically tied to what has gone before. Rather the *Hey* marks (as it can be shown to do generally) the introduction of a new topic 'touched off' by the prior utterance, which is just evoked from memory by some chance association to the content of the prior turn.

But if shared reference, or a set of shared concepts, across turns is not sufficient to ensure shared topic, neither is it necessary for two turns to share some referents, or concepts, in order for topic to be

[14] It may be objected that the example indicates only that use of the same words, e.g. *shaving*, does not entail identity of reference. However, it is easy to show that identical referents may be picked out by terms either side of a topic break, here marked by *By the way* and increased amplitude following a pause:
Owen 8b
B: Probably *is* because of that I should think, yes, mm
A: Mm
(1.2)
A: ((louder)) By the way, do you want any lettuces
Here of course *I* and *you* both refer to the same entity, namely B, but neither topic is in any ordinary sense 'about B'. So the argument can be generalized: neither identical reference, nor the use of identical terms or concepts (with same or different reference) is sufficient to engender topical continuity.

preserved. For example, C's utterance below is topically tied to prior utterances:

(34) *Sacks, 1968, April 17*
A: If yer gonna be a politician, you better learn how to smoke cigars
B: Yeah that's an idea Rog
C: I heard a very astounding thing about pipes last night

but *pipes* and *cigars* are distinct concepts, and are terms with no overlapping sets of referents. Of course we can retreat and say: two utterances share the same topic or are at least topically tied only if there is some superordinate set which includes referents or concepts from both utterances (here, say, the set of 'smokables'). But then any two utterances share a topic (or at least are topically tied) because for *any* two sets of referents or concepts one can invent a superordinate set that includes them both – nor is this conversationally absurd (see e.g. (7) above where the shared class was 'apartment rental disqualifiers', hardly some 'natural' class).

The point is simply that topical coherence cannot be thought of as residing in some independently calculable procedure for ascertaining (for example) shared reference across utterances. Rather, topical coherence is something *constructed* across turns by the collaboration of participants. What needs then to be studied is how potential topics are introduced and collaboratively ratified, how they are marked as 'new', 'touched off', 'misplaced' and so on, how they are avoided or competed over and how they are collaboratively closed down.[15]

Now such collaborative procedures for opening, changing and closing down topics are not strictly part of the overall organization of telephone calls: they are local procedures that can operate throughout a call. But they interact in complex ways with matters of overall organization, hence their treatment here. For example, as we noted, later topical constraints give the first topic slot after the opening section a special importance, reinforced by the expectation that, after a summons and its answer, a reason for the summons will be presented. Further, the elaboration of *How are you*s provides a route into topical talk that can displace the reason for the call and its

[15] Relatively little work has been done here, but see Sacks, 1967–72 passim and summary in Coulthard, 1977: 78ff; Button & Casey, in press; Jefferson, in press; Owen, 1982.

first topic slot to later in the call, thereby providing a powerful motive for escaping from such elaborations (see Sacks, 1975). And techniques for topic closing are intimately connected to the introduction of the **closing section** shutting down the conversation: the closure of any topic after the privileged first one makes the introduction of the closing section potentially imminent, matters dealt with below. Finally, some kinds of telephone calls have an expectable overall organization that admits just one topic – such **monotopical** calls being typical of routine business calls or service inquiries. Interestingly, such calls are monotopical not in the sense that no more than one topic is ever addressed within them, but in the sense that the caller orients to the expectation of a single topic in the very introduction of further topics. Thus one finds, not only initial announcements in first topic slot that the caller has in fact more than one thing to say, but also careful tracking of the progress through the list of topics:

(35) *Birmingham Discourse Project TD.C1.2 (After initial inquiry)*
B: Yeah er two other things firstly do you know the eventual street number of plot 36
((several turns later))
Erm the other thing is erm ((ahem)) presumably be okay for somebody to have access to it before we move in to put carpets down and that

So matters of overall organization and of topical organization can be closely interlinked.

We come finally to the **closing sections** of the overall organization of telephone calls or similar kinds of conversation. Closings are a delicate matter both technically, in the sense that they must be so placed that no party is forced to exit while still having compelling things to say, and socially in the sense that both over-hasty and over-slow terminations can carry unwelcome inferences about the social relationships between the participants. The devices that organize closings are closely attuned to these problems. We find typically that conversations close in the following sort of manner:

(36) *172B(7)*
R: Why don't we all have lunch
C: Okay so that would be in St Jude's would it?
R: Yes
(0.7)
C: Okay so:::

R: One o'clock in the bar
C: Okay
R: Okay?
C: Okay then thanks very much indeed George=
R: =All right
C: //See you there
R: See you there
C: Okay
R: Okay // bye
R: Bye

The typical features here are the arrangements for a next meeting, a sequence of *Okay*s closing down the arrangements (or other topic), a *Thank you* produced by the caller, and a further sequence of *Okay*s just prior to a final exchange of *Good-bye*s. One very general schema for closing sections, of which (36) is merely one instantiation, might be represented thus:

(37)
(a) a closing down of some topic, typically a **closing implicative** topic; where closing implicative topics include the making of arrangements, the first topic in monotopical calls, the giving of regards to the other's family members, etc.
(b) one or more pairs of **passing turns** with **pre-closing items**, like *Okay*, *All right*, *So::*, etc.
(c) if appropriate, a **typing** of the call as e.g. a favour requested and done (hence *Thank you*), or as a checking up on recipient's state of health (*Well I just wanted to know how you were*), etc., followed by a further exchange of pre-closing items
(d) a final exchange of terminal elements: *Bye*, *Righteo*, *Cheers*, etc.

The crucial elements here (after (a) has been achieved) are (b) and (d). Essentially what the two components jointly achieve is a co-ordinated exit from the conversation: they do this by providing, in the form of the topic-less passing turns in (b), a mutual agreement to talk no more, this being a prelude to the exchange of the terminal adjacency pair in (d) that closes down the conversation. The mutual agreement is secured by one party producing a topic-less passing turn, indicating that he has no more to say, whereupon the other party – if he too has no more to say – may produce another such turn. The technical and social problems that closings raise are thus initially dealt with by providing that the closing section as a whole is placed in a

location that is interactively achieved: a pre-closing offer to close is issued in the form of *Okay, Right*, etc., and only if taken up do closings proceed. Further motivation for this pattern in closing sections will be provided below (but see Schegloff & Sacks, 1973).

A final point about closing sections that is of interest here is that components of the sort in (37)(c) indicate that the placement and content of closing sections is attuned to other aspects of overall organization. Thus, for example, the *Thanks* in (36) is oriented to the specific content of the first topic slot of that call, namely a request for a favour. Similarly one finds in closings reference to aspects of opening sections, as in *Sorry to have woken you up* referring back to *I hope I'm not calling too early*, or *Well I hope you feel better soon* referring back to responses to *How are you*s, and so on. Each aspect of overall organization, then, can be oriented to other aspects, as is exemplified in the attention paid in the opening sections of expectably monotopical calls to the imminence of closing immediately after the first topic is closed down (the attention revealed in the *Just two things* kind of bid for more than one topic).

We are now in a position to give a more technical characterization of what a conversation is. We must first distinguish the unit **a conversation** from **conversational activity**. The latter is something characterizable in terms of local organizations, and especially the operation of the turn-taking system in (10); there are many kinds of talk – e.g. sermons, lectures, etc. – that do not have these properties and which we would not want to consider conversational. Yet there are also many kinds of talk – e.g. courtroom or classroom interrogation – which exhibit features of conversational activity like turn-taking, but which are clearly not conversations. Conversation as a unit, on the other hand, is characterizable in terms of overall organizations of the sort sketched here in addition to the use of conversational activities like turn-taking (Schegloff & Sacks, 1973: 325; Sacks, Schegloff & Jefferson, 1974: 730–7).

6.2.2 *Some remarks on methodology*

The basic findings in the prior section have been presented (for the sake of brevity) in a way that CA workers would in fact be careful to avoid. The reason is that, for each substantial claim, the methodology employed in CA requires evidence not only that some aspect of conversation *can* be viewed in the way suggested, but that it

actually is so conceived by the participants producing it. That is, what conversation analysts are trying to model are the procedures and expectations actually employed by participants in producing and understanding conversation. In addition, for each conversational device we should like, by way of explanation, to elucidate the interactional problems that it is specifically designed to resolve – that is, to provide *functional* explanations, or expositions of rational design, for the existence of the device in question. There are, then, two basic methods to be employed in CA-style investigation:

(a) We should attempt to locate some particular conversational organization, and isolate its systematic features, by demonstrating participants' orientation to it
(b) We should ask, (i) what problems does this organization solve, and (ii) what problems does this organization raise – and therefore what implications does it have for the existence of further solutions to further problems?

These methods are important because they offer us a way of avoiding the indefinitely extendable and unverifiable categorization and speculation about actors' intents so typical of DA-style analysis. Let us therefore look at some illustrations of how the methods may be applied to yield and then confirm results of the kind we have reviewed.

We may start with the problem of demonstrating that some conversational organization is actually oriented to (i.e. implicitly recognized) by participants, rather than being an artefact of analysis. One key source of verification here is what happens when some 'hitch' occurs – i.e. when the hypothesized organization does not operate in the predicted way – since then participants (like the analyst) should address themselves to the problem thus produced. Specifically, we may expect them either to try to repair the hitch, or alternatively, to draw strong inferences of a quite specific kind from the absence of the expected behaviour, and to act accordingly.

Where hitches of these sorts are a recurrent possibility, there is likely to be a regularized repair procedure. Such occurs, we noted, in association with the turn-taking system, where a special set of procedures operates to reduce and resolve overlap, should this arise despite the rules assigning turns. But there are overlaps allowed (and thus their location and nature predicted) by the rules, and overlaps that contravene the rules (*interruptions*). When the latter occur, they

are subject not only to the standard resolution procedures, but also to overt reprimands and sanctions – and such overt attention to interruptions again indicates participants' orientation to the basic expectations provided by the rules:

(38) *DCD: 28*

Collins: Now // the be:lt is meh*

Fagan: is the sa:me mater*ial as // thi:s

→ Smythe: Wait a moment

Miss Fagan

Similarly, the conditional relevance of a second part of an adjacency pair given a first part is easily shown to be more than just an analyst's fancy. Consider for example what happens when, employing Rule 1(a) of the turn-taking system, a speaker addresses a recipient with the first part of a pair and receives no immediate response. Strong inferences are immediately drawn, either of the sort 'no response means no channel contact', or, if that is clearly not the case, then 'no response means there's a problem'. So, in the case of a failure to respond to a summons, the absence of a second part can, in the case of the telephone, be understood as 'recipient is not at home', or in face-to-face interaction as 'recipient is sulking or giving the cold shoulder' (Schegloff, 1972a: 368ff). Or, consider:

(39) *172B(7)*

T1 C: So I was wondering would you be in your office on Monday (.) by any chance?

T2 (2.0)

T3 C: Probably not

T4 R: Hmm yes=

T5 C: =You would?

T6 R: Ya

T7 C: So if we came by could you give us ten minutes of your time?

Here a two-second pause after the question in T1 is actually taken by C to indicate a (negative) answer to the question. How can this come about? Note first that (by Rule 1(a) of the turn-taking system) C has selected R to speak (a feature of address not being necessary as there are only two participants here). Therefore the two-second pause is not just anyone's pause or nobody's pause (i.e. a lapse): rather it is assigned by the system to R as R's silence. Then recollect that adjacency pairs can have dispreferred seconds, these in general being marked by delay (amongst other features). Therefore the pause can be heard as a preface to a dispreferred response. Now in full

sequential context it is clear that C's question is a prelude to a request for an appointment, and for such questions it turns out that negative answers (answers that block the request) are dispreferred (see 6.3 and 6.4 below). Hence C draws the inference from R's silence that he makes explicit in T3. (That he got it wrong, as indicated by R in T4, does not affect the point – such inferences are made, often correctly, though sometimes not.) Note here the remarkable power of the turn-taking system to assign the absence of any verbal activity to some particular participant as his turn: such a mechanism can then quite literally make something out of nothing, assigning to a silence or pause, itself devoid of interesting properties, the property of being A's, or B's, or neither A's nor B's, and further, through additional mechanisms, the kind of specific significance illustrated in (39) (a point taken up below).[16]

A fundamental methodological point can be made with respect to (39), and indeed most examples of conversation. Conversation, as opposed to monologue, offers the analyst an invaluable analytical resource: as each turn is responded to by a second, we find displayed in that second an *analysis* of the first by its recipient. Such an analysis is thus provided by participants not only for each other but for analysts too. Thus in (39) the turn in T3 displays how the pause in T2 was interpreted. Hence "the turn-taking system has, as a by-product of its design, a proof procedure for the analysis of turns" (Sacks, Schegloff & Jefferson, 1978: 44). A good case can therefore be made for the methodological priority of the study of conversation over the study of other kinds of talk or other kinds of text.

Having shown that participants themselves orient to the conditional relevance of, for example, an answer after a question, let us now briefly consider the kind of evidence that could be used to show that

[16] Examples of this sort provide a clue to the nature of conversational constraints. Participants are constrained to utilize the expected procedures not (or not only) because failure to do so would yield 'incoherent discourses', but because if they don't, they find themselves accountable for specific inferences that their behaviour will have generated. Thus defendants in political trials may hope that silence will count as rejection of the proceedings, only to find it read as admission of guilt. Or, in (39), R's disregard of the expectation that preferred responses will be immediate not only produces an unintended inference that has to be corrected, but, if sustained, may produce an inference of general reluctance to co-operate. Conversationalists are thus not so much constrained by rules or sanctions, as caught in a web of inferences.

the overall organizations we have claimed to be operative in conversation are actually oriented to by participants. As already noted, closing sections may refer back to opening sections and vice versa, indicating that "the unit 'a single conversation' is one to which participants orient throughout its course" (Schegloff & Sacks, 1973: 310). Further, if closing sections have the character suggested above, then a co-ordinated determination to close is mutually accepted by an exchange of pre-closings like *Okay*, and we expect thereupon the immediate exchange of terminal elements like *Bye*. But every now and then closings are in fact re-opened, and if these re-openings occur after the exchange of pre-closings, then they are typically *marked* as grossly misplaced, as in the extract below:

(40) *Schegloff & Sacks, 1973: 320*
C: Okay, thank you.
R: Okay dear.
→ C: OH BY THE WAY. I'd just like to say ...

Such misplacement markers demonstrate an orientation to the closing section as a unit not properly taking such interpolations, and thus once embarked on, properly final.

Let now consider the other basic methodological procedure, namely the search for the *raison d'être* of particular conversational organizations, and then the implications that the existence of one device has for the necessity for others. We may show in this way how all the structural facts we have reviewed (and indeed others too) are in fact closely integrated; and in doing so we may illustrate how in discovering one such organization the analyst is provided with a lever for prying up further levels of organization. So the assumption of functional inter-connection actually yields a powerful discovery technique.

Suppose we take the turn-taking system as the fundamental device, our initial discovery. What we then have is a system primarily designed to (a) organize the change of speakers and (b) keep only one speaker speaking at a time. But then we may ask: how is such a device 'cranked up', how is the machinery to be got rolling? Clearly we need some device that will establish (for the case of two parties) the A–B–A–B pattern of turns, while launching us into the business of the interaction. An adjacency pair, it would seem, would nicely do the job, setting up an initial A–B sequence. However, as the turn-taking rules permit a conversation to lapse, that is all such a pair might

achieve: A–B, finish. So we need an opening section that has at least a three-turn structure, wherein the first requests attendance from the other party, the second provides a slot for that other party to commit himself to an initiation of interaction, and the third turn is the slot for the initiating party to provide some initial business for the interaction. We then have the familiar structure, summons–answer–first topic, which establishes a co-ordinated co-participation, assigns speaking and receiving roles to the two parties for the first three turns, and thus cranks up the turn-taking machinery as minimally required. Small details of the design of such sequences reflect their adaptation to this task – e.g. the tendency (in face-to-face interaction) for the second turn to be an open question requiring by adjacency pair format the third turn necessary for the proper initiation of talk (as already noted). There is thus nothing *ad hoc* or arbitrary about the design of conversational sequences like summons sequences: they are rational solutions to particular organizational problems.

We now have the turn-taking machinery started up. But then the question arises: how do we suspend it? Consider: A and B are talking and A now wants, in response to B's remark, to tell an apposite story. But how is A to get such a substantial section of talk, when by the rules of the turn-taking system B is allowed at the very first TRP to compete by first start for the floor? Clearly obtaining such an extended turn at talk (by other than sheer listener apathy) requires special techniques. One such special device is a story announcement sequence of the stereotypical sort illustrated below:

(41) A: Have you heard the one about the pink Martian?
B: No
A: ((Story))

where a bid is specifically made for an extended space for the telling of a story, the telling being conditional on the acceptance of the bid. Or, from a recording:

(42) *Sacks, 1974: 338*

T1 K: You wanna hear muh- eh my sister told me a story last night.
T2 R: I don't wanna hear it. But if you must,
(1.0)
T3 A: What's purple an' an island. Grape-Britain. That's what iz sis//ter -
T4 K: No. To stun me she says uh there was these three girls ...((Story follows))

Here, in T2, R gives a reluctant go-ahead, while in T3 the other intended recipient produces a 'guess' at what kind of a story it is as a potential dismissal (T3), itself dismissed by the story teller in T4. Such sequences contain (minimally) in T1 an offer to tell, in T2 a 'go-ahead' or rejection, and then contingent on the 'go-ahead' the telling of the story in T3 (see 6.4 below). What such a structure achieves is the collaborative suspension of the turn-taking machinery, by joint agreement, for the duration of the story (there are of course other techniques for doing this – see Terasaki, 1976; Jefferson, 1978; Ryave, 1978).

But if we have achieved a suspension of speaker transition relevance over an extended period of talk, we now have yet another problem, namely how to start up the turn-taking machinery once again (or more strictly, since co-participation is still assured, re-invoke the relevance of TRPs). A solution here had better provide for the recognizability of *story endings* – for if they are recognizable then on such a completion the normal turn-taking machinery can once again automatically resume. So stories must be recognizable units if turn-taking is to be adjusted around them; and of course they are: stories, if of the 'funny' variety, typically have punchlines, whereupon laughter by listeners is immediately relevant (Sacks, 1974: 347ff); or if they are topically tied to the sequential locus in which they occur then endings are recognizable in part because they return participants to that particular topic (Jefferson, 1978); or other recognizable ending formats are used (Labov & Waletsky, 1966; Sacks, 1972).

Once again, then, we have the turn-taking machinery operating normally. But now let us suppose we want not merely to suspend it, but to close it down, i.e. to finish the conversation. Again some special device is needed which will provide a solution to the following problem: "how to organize the simultaneous arrival of co-conversationalists at a point where one speaker's completion will not occasion another speaker's talk and that will not be heard as some speaker's silence" (Schegloff & Sacks, 1973: 294–5). Again one basic ingredient suggests itself: an adjacency pair such that the first part announces imminent closure and the second part secures it. And we do indeed have the terminal exchange generally realized as A: *Bye*; B: *Bye*.

However there would be substantial problems for the use of the terminal exchange alone as a solution to the closing problem. For A

might have said all he wants to say, and therefore have issued a *Bye*, whereupon B, despite perhaps having important things to say (things perhaps that *must* be said in this conversation – see Sacks, 1975), would be constrained by the adjacency pair format to produce a second *Bye* that terminated the interaction. Therefore there needs to be some pre-terminal section where undelivered news and the like can be fitted in. This need is strongly reinforced by the topical organization we reviewed, since (a) one is constrained not to mention in first topic slot anything that one doesn't want to be taken as the main reason for engaging in interaction, and is therefore forced to hang on to these other 'mentionables', and (b) after first topic slot, mentionables should by preference be fitted to prior topics, requiring that one waits for a suitable slot for such deferred mentionables. However, such a slot may never come up, and there is therefore a need for some slot towards the end of a conversation specifically set aside as the place where such deferred mentionables can be unburdened.

What is needed for effective closings is therefore a device which (a) offers each party a turn for such deferred mentionables, (b) if such a turn is taken up, recycles the opportunity in (a), and (c), consequent upon no party taking up the opportunity in (a), makes the terminal exchange immediately relevant. And it is this that motivates the familiar four-turn closing section:

(43) A: Okay
B: Okay
A: Bye
B: Bye

where the first *Okay* yields the floor to the other party for any deferred mentionables that he may have, the second indicates that no such items have been withheld, and thus the exchange of topic-less passing turns may be taken as a mutual agreement that termination should now commence. The exchange of *Okay*s can thus be called *pre-closings* – producing the forewarning and collaborative co-ordination of closure, which the turn-taking system and topical organization independently but jointly require.

So in the way thus informally sketched, from one kind of conversational organization one can foresee the need for other kinds of organizations with specific properties, providing simultaneously both a search procedure for conversational organizations and explanations for their existence and design.

One further methodological preference is a growing tendency in CA to work with increasing numbers of instances of some phenomenon. Until one knows how, for example, certain kinds of sequence normally unfold, the analysis of individual complex cases will not yield up the rich texture they almost invariably conceal (see e.g. the analysis of (49) and (104) below).

In summary, then, CA methodology is based on three basic procedures: (a) collecting recurrent patterns in the data, and hypothesizing sequential expectations based on these; (b) showing that such sequential expectations actually are oriented to by participants; and (c) showing that, as a consequence of such expectations, while some organizational problems are resolved, others are actually created, for which further organizations will be required.

6.2.3 *Some applications*

In this section we illustrate how the observations above may be applied to yield insight into particular instances of talk. We will start by considering what is apparently just one phenomenon – silence, or a period of non-speech – and show how such pauses can be discriminated into many different kinds with quite different significances on the basis of their structural locations. Then we will summarize an analysis by Schegloff of an opaque little sequence rich in structural detail, showing that detailed analysis of individual segments of talk is made possible by the use of the general findings and techniques already reviewed. These examples should suffice to indicate how much organization there is to be discovered in the smallest extract of talk, and how powerful *sequential location* can be in the assignment of multiple functions to individual utterances.

There have been many theories about the significance of pauses and hesitations in conversation: some analysts, for example, have seen pauses as evidence of verbal planning, i.e. 'time out' for psychological processing either in the routine preparation of the fluent phases that often follow (Butterworth, 1975) or in the production of complex syntax (Goldman-Eisler, 1968; Bernstein, 1973). But the following observations show that any unitary account of pauses, and any account that does not take into consideration their role as potentially symbolic devices, will be fundamentally misguided.

The turn-taking system itself assigns different values to pauses within conversation. We have already described how the rules in (10)

discriminate between *gaps* (delays in the application of Rules 1(b) or 1(c)), *lapses* (non-application of the rules) and next speaker *silence* (after application of Rule 1(a)), as illustrated in (14) and (15) above. Where these rules assign a pause to some speaker as a silence, additional factors systematically play a role in its interpretation. For example, we have seen in (39) how a silence after a question of a special sort (a prelude to a request – a **pre-request** – see section 6.4 below) can be read, by virtue of preference organization, as indicating a negative answer. Or, consider the three-second silence in (44):

(44) *Drew, 1981 : 249*

M: What's the time- by the clock?
R: Uh
M: What's the time?
→ (3.0)
M: (Now) what number's that?
R: Number two
M: No it's not
What is it?
R: It's a one and a nought

Here, in the turn prior to the pause, a mother asks her child to try and tell the time. So, by Rule 1(a) in (10), the pause is a silence, attributable to the child R. But just because the question is an 'exam question' (and not, say, a pre-request), the silence here can be understood as 'answer unknown'. Such an analysis is made clear by the mother's next turn, where an easier question is asked that, if answered, might provide a partial solution to the first question.

Now in (45) we have a small pause after the second turn in the opening of a telephone call:

(45) *Schegloff, 1979a : 37*

C: ((rings))
T1 R: Hello?
T2 C: Hello Charles.
→ (0.2)
T3 C: This is Yolk.

As noted earlier, for a caller to provide a greeting in T2 (his first verbal turn) is to claim that the recipient should be able to recognize the caller on the basis of this sample of voice-quality alone. The second turn, we noted, is in fact the first part of an adjacency greeting pair; a second is therefore due. Once again, then, the delay (short though

it is) is R's delay and can be taken by C to indicate a problem for R. That the problem is here a problem in identification is shown by the repair C offers, after a significant pause has developed, namely an overt self-identification (*This is Yolk*). That the problem indicated to C by this small delay is not imaginary is shown by examples like the following, where in T3 R has to invite C to repair what C had taken to be an adequate self-identification (the *Hello* in T2):

(46) *Schegloff, 1979a : 39*

C: ((rings))
T1 R: Hello?
T2 C: Hello.
→ (1.5)
T3 R: Who's this.

Here a momentary pause is heard immediately as a problem with what is always underway in the first few turns of telephone conversations, namely the business of mutual identification. Therefore the significance of a pause here is determined by that set of overlapping organizations that converge on the first few turns of telephone calls, as indicated in (30) above; that set determines, via adjacency pair organization and the structure of opening sections, just how a pause in this location will be interpreted.

In (47) a pause, which can be analysed as somewhat similar to that in (45), occurs after an invitation. Once again, an invitation is a first part of an adjacency pair, and this assigns next turn to the other party:

(47) *Davidson, in press*

A: C'mon down *he*:re, =it's oka:y,
→ (0.2)
A: I got lotta stuff, =I got *be*:er en stuff

And, as in (45), a short pause occurs, hearable as the other party's silence, and clearly analysed in this (and many related examples) as some problem with A's invitation, which A consequently upgrades – i.e. an attempt is made to make the invitation more attractive (see Davidson, in press on the systematicity of this pattern).

Finally, the following example features the punchline of a dirty joke and the ensuing laughter. As we pointed out, after a story an appreciation is immediately relevant, and the temporary suspension of the turn-transition relevance is lifted. But here we have a two-second delay, and then instead of recipient laughter we have teller's laughter (with a further four-second delay interspersed). Only then does one

of the recipients (A) laugh, and then it has the careful syllabicity of mock laughter. The pauses here are assignable to story recipients as *their* silences, and the withholding of appreciation signals 'failed joke' (see Sacks, 1974).

(48) *Sacks, 1974 : 339*

K: ((tells dirty joke, ending thus:)) Third girl, walks up t'her – Why didn' ya *say* anything last night; W' *you* told me it was always impolite t'talk with my mouth full,
→ (2.0)
K: hh hyok hyok,
→ (1.0)
K: hyok,
→ (3.0)
A: HA-HA-HA-HA

Many further kinds of significant absences of speech can be found – see e.g. (66), (67), (76) and (77) below – and each kind draws the analyst's attention to the strong kinds of expectations that different conversational organizations, whether local, overall or intermediate in scope, impose on particular sequential slots. The demonstration is the more remarkable in that silence has no features of its own: all the different significances attributed to it must have their sources in the structural expectations engendered by the surrounding talk. So sequential expectations are not only capable of making something out of nothing, but also of constructing many different kinds of significance out of the sheer absence of talk. If conversational organization can map 'meaning' onto silence, it can also map situated significance onto utterances – and in fact can be shown to regularly do so.

Let us now turn to one short extract of conversation and show how the various findings and techniques we have reviewed can be applied to good effect. The argument is a brief résumé of Schegloff, 1976. The extract comes from a radio call-in programme broadcast in the United States, and in it B, who is a High School pupil, is reporting to the compère of the show, A, an argument that he has been having with his history teacher about American foreign policy. The teacher (T) holds that foreign policy should be based on morality, but B thinks it should be based on expediency – 'what is good for America'. It runs as follows:

(49) *Schegloff, 1976 : D9*

T1 B: An' s- an' ( ) we were discussing, it tur- , it comes down, he ((T)) s- he says, I-I-you've talked with thi- si- i- about

this many times. *I* ((B)) said, it came down t' this: =
= our main difference: *I* feel that a government, i- the main thing, is- th-the purpose of the goverment is, what is best for the country.

T2 A: *Mm*hmm

T3 B: *He* ((T)) says, governments, an' you know he keeps- he talks about governments, they sh- the thing that they sh'd do is what's right or wrong.

T4 → A: For *whom*.

T5 B: Well he says-//he-

T6 A: By what *stan*dard.

T7 B: That's what- that's exactly what I mean. He s- but he says ...

The particular interest of this extract is a crucial ambiguity associated with the utterance *For whom*. It is not, however, an ambiguity that lies in the linguistic structure of the utterance, nor has it to do with any lexical ambiguities of the words *for* and *whom*; and unlike linguistic ambiguities, which scarcely ever cause difficulties in context, this one demonstrably is (or becomes) ambiguous for the participants. The ambiguity is this: on one reading (R1) A, in uttering *For whom*, asks a question that we might paraphrase as 'What exactly did your teacher say – governments should do what's right *for whom*? Whom did he have in mind?' On the other reading (R2), A, in asking *For whom*, is actually trying to show that he agrees with B against B's teacher (T), and he is trying to show this by offering a potential piece of B's argument against T. To see this consider that B is reporting T as saying that foreign policy should be based on what is morally right – to which B might have retorted by saying *Yes, but right for whom?*, pointing out that ethical judgements of good or bad depend upon different parties' points of view. So on this reading, or interpretation, A in saying *For whom* is providing an utterance that B might have used against his teacher, thus showing agreement with B.

That both readings of the utterance become available to B is clear. First, in T5, he starts off responding to R1, the straightforward question interpretation, by beginning on a further specification of what the teacher says. But then A interrupts with a correction; we know this in part because only corrections of such sorts are priority items licensing violations of the turn-taking rules. But we also know that T6 is a correction because it utilizes a standard device for

correcting misunderstandings, namely **reformulation** that makes the same point in different words. In the following turn, B then displays understanding of the alternative reading, by acknowledging A's agreement with him, *that's exactly what I mean*. We can thus show that the ambiguity is a participant's (and not merely an analyst's) ambiguity: each party deals with each reading once – A by correcting B's interpretation, and then reformulating his own intended reading, and B by first beginning to respond to the non-intended reading, and then showing understanding of the second reading as an agreement with him against his teacher, by acknowledging A's agreement.

But how does the ambiguity arise? Since it is clearly not a matter of the grammatical or lexical ambiguity of *For whom*, the source of the ambiguity must lie outside the utterance itself in its sequential location in the conversation. We need now to show that the structural location itself predisposes us to both of the relevant interpretations.

Stories, we noted, require the suspension of the normal turn-taking system, which then requires resumption. This could be provided for, it was argued, if story endings are easily recognizable. One recognizable and recurrent story ending format is a summing up of the story, and that is what we find occurring in our extract – B says *It came down to this : our main difference is* ... and the summary follows. So the slot in which A says *For whom* is the first slot after a story ending. Such a slot is one where story recipients can be expected to do one of two things: they may ask for further details or clarifications of the story – and this is the sequential basis for the simple question interpretation, R1; or they may show understanding and appreciation of the story (as e.g. in the expectable laughter after a joke: see discussion of (48) above), and it is this possibility that forms the basis of the second, more complex, interpretation, R2. For one way of showing understanding is to express agreement in such a way that prior understanding must have taken place, and *For whom* does just this, by showing agreement through displaying understanding of the argument that B was having with his teacher.

But there's another element here: this agreement reading is reinforced by consideration of the kind of story that B's story is, namely an 'opposition story' or a reported argument. Such stories have as features not only an alternation of reported speakers, or an A–B–A–B structure of reported turns, but also, mapped onto the

alternation of turns, the alternation of positions, or sides in the argument. So when reported speakers change, the positions being argued for change. Such structural expectations lie behind our ability to understand some minimal story like *Pay the rent. I can't pay the rent* as being a reported argument where one party said *Pay the rent* and the other *I can't pay the rent*. Now it is just because B's story here is an opposition story that we can hear A's *For whom* as taking up B's position against the teacher. For B is reporting an argument in which the teacher (T) and he alternated in turns and positions in a T–B–T–B ... sequence. Further, we can see that it is just because in addition to being an opposition story, it is one which ends with a turn by T, that A can jump in and show story understanding by taking B's turn after T's. And for A to do this is an optimal way of displaying understanding, one of the expectable things to be going on in the first slot after a story.

Analyses of this sort, which show how surrounding conversational structure can impose rich interpretations on utterances, provide important lessons for linguistic and psychological theories of language understanding. First, they indicate that semantic interpretation is only a small and not perhaps the most complex aspect of the communicational significance of an utterance. Secondly, they show that speech act theory and allied theories of utterance function can only be considered crude and (at best) partial accounts of such situated significance (consider, for example, what little of interest speech act theory could say about *For whom*). Thirdly, such analyses suggest that while it is correct to look for the sources of such significance outside the utterance itself, it may be a mistake to look too far afield, and specifically that it can be premature to invoke the application of large quantities of background knowledge, as in the **frames** approach now popular in cognitive psychology and artificial intelligence approaches to language understanding (see e.g. Charniak, 1972).

6.3 Preference organization

6.3.1 *Preferred second turns*

As we have seen (6.2.1.2), alternative second parts to first parts of adjacency pairs are not generally of equal status; rather some second turns are **preferred** and others **dispreferred**. The notion of **preference**, it was noted, is not intended as a psychological claim

about speaker's or hearer's desires, but as a label for a structural phenomenon very close to the linguistic concept of **markedness**, especially as used in morphology:[17]

> The intuition behind the notion of markedness in linguistics is that, where we have an opposition between two or more members ..., it is often the case that one member is felt to be more usual, more normal, less specific than the other (in markedness terminology it is unmarked, the others marked). (Comrie, 1976a: 111)

Further, in morphology, "unmarked categories tend to have less morphological material than marked categories" and there is a "greater likelihood of morphological irregularity in unmarked forms" (Comrie, 1976a: 114). The parallel is therefore quite apt, because in a similar way **preferred** (and thus **unmarked**) seconds to different and unrelated adjacency pair first parts have less material than **dispreferreds** (**marked** seconds), but beyond that have little in common (cf. "irregular"). In contrast, dispreferred seconds of quite different and unrelated first parts (e.g. questions, offers, requests, summonses, etc.) have much in common, notably components of delay and parallel kinds of complexity. Some further examples will make this clear, but before proceeding we should point out that, in addition to the structural aspect of preference organization, we will need a rule for speech production, which can be stated roughly as follows: 'try to avoid the dispreferred action – the action that generally occurs in dispreferred or marked format'. (The two essential features of dispreferred actions are thus (a) they tend to occur in a marked format, and (b) they tend to be avoided.) Such a rule is non-circular if we already have an independent characterization of preferred or dispreferred alternatives on structural grounds.

So let us return to a characterization of dispreferred seconds – consider the following pair of invitations and their responses:

(50) *Atkinson & Drew, 1979: 58*
A: Why don't you come up and *see* me some//times
B: I would like to

(51) *Atkinson & Drew, 1979: 58*
A: Uh if you'd care to come and visit a little while this morning I'll give you a cup of *co*ffee

[17] The concept of *markedness* was originally developed by linguists of the Prague School; the classic references are Jakobson, 1932; Trubetzkoy, 1939: Chapter 3; see also Lyons, 1968: 79ff.

→ B: hehh Well that's awfully sweet of you,
((DELAY))((MARKER))((APPRECIATION))
I don't think I can make it this morning.
((REFUSAL or DECLINATION))
.hh uhm I'm running an ad in the paper and-and uh I have to stay near the phone.
((ACCOUNT))

Here (as Atkinson & Drew (1979: 58ff) point out) the invitation in the first example has an acceptance as a second part: the acceptance is of simple design and is delivered not only without delay but actually in partial overlap. In contrast, the invitation in the second example receives a refusal or declination as a second, and here we have all the typical features of dispreferreds, namely (as indicated by the glosses in capitals) delay, the particle *Well* which standardly prefaces and marks dispreferreds (and here we have a rival analysis to that offered in Chapter 3 in terms of implicature – see Owen, 1980: 68ff, 1981), an appreciation (notably absent from the acceptance in the prior example),[18] a qualified or mitigated refusal (*I don't think I can*), and an account or explanation for the dispreferred second. (Compare also the request examples in (24) and (25) above.)

The characteristics of dispreferred seconds can be further generalized (see Pomerantz, 1975: 42ff, 1978, in press; Atkinson & Drew, 1979: Chapter 2; Wootton, in press) – such turns typically exhibit at least a substantial number of the following features:

(a) *delays*: (i) by pause before delivery, (ii) by the use of a preface (see (b)), (iii) by displacement over a number of turns via use of *repair initiators*[19] or insertion sequences
(b) *prefaces*: (i) the use of markers or announcers of dispreferreds like *Uh* and *Well*, (ii) the production of token agreements before disagreements, (iii) the use of appreciations if relevant (for offers, invitations, suggestions, advice), (iv) the use of apologies if relevant (for requests, invitations, etc), (v) the use of qualifiers (e.g. *I don't know for sure, but* ...), (vi) hesitation in various forms, including self-editing
(c) *accounts*: carefully formulated explanations for why the (dispreferred) act is being done
(d) *declination component*: of a form suited to the nature of the first part of the pair, but characteristically indirect or mitigated

[18] Of course, appreciations can occur with invitation acceptances, but they typically occur *after* acceptances while they occur *before* rejections.

[19] This term is explained below.

There follow some examples of each of these (the headings indicate which features are especially notable in each extract):

(52) *Wootton, in press (Illustrating (a)(i))*
Ch: Can I go down an see 'im
(2.0)
( )
(1.8)
C'mo::n
(1.5)
Come'n te see 'im
(1.6)
C'mo::n
M: No:::

(53) *33A (Illustrating (a)(ii), (b)(iii), (c), (d))*
B: She says you might want that dress I bought, I don't know whether you do
→ A: Oh thanks (well), let me see I really have lots of dresses

(54) *Wootton, in press (Illustrating (a)(iii))*
Ch: I wan my ow:n tea .hh my*self*
→ M: (You) want what? =
Ch: = My tea my*se :lf*
M: No:w? We are all having tea together

(55) *176B (Illustrating (b)(i), (c))*
R: What about coming here on the way (.) or doesn't that give you enough time?
→ C: Well no I'm supervising here

(56) *176B (Illustrating (b)(v), (d))*
C: Um I wondered if there's any chance of seeing you tomorrow sometime (0.5) morning or before the seminar
(1.0)
→ R: Ahum (.) I doubt it

(57) *163 (Illustrating (b)(vi)) (R has been complaining that C's fire in the apartment below has filled R's apartment with smoke)*
C: ...is it-it's all right now- you don't want me to put it out?
→ R: E::r (1.5) well on the whole I wouldn't bother because er huhuh (2.0) well I mean what- what (0.5) would it involve putting it out
(0.5)
C: Hahaha ( ) hahah

(58) *163 (Illustrating (b)(iv))*
A: ((to operator)) Could I have Andrew Roper's extension please?
(9.0)

B: Robin Hardwick's telephone (1.0) hello
A: Andrew?
→ B: No I'm awfully sorry Andrew's away all week

Given a structural characterization of preferred and dispreferred turns we can then correlate the content and the sequential position of such turns with the tendency to produce them in a preferred or dispreferred format. And here we find recurrent and reliable patterns, e.g. refusals of requests or invitations are nearly always in dispreferred format, acceptances in preferred format. Table 6.1 indicates the sort of consistent match between format and content found across a number of adjacency pair seconds.

Table 6.1 *Correlations of content and format in adjacency pair seconds*

FIRST PARTS:	Request	Offer/Invite	Assessment	Question	Blame
SECOND PARTS:					
Preferred:	acceptance	acceptance	agreement	expected answer	denial[20]
Dispreferred:	refusal	refusal	disagreement	unexpected answer or non-answer	admission

Given such a correlation between the kind of action performed and the way in which it is done, we can then talk not only of preferred turns but also of preferred actions (namely, those normally performed in the preferred format).

Now given these patterns we can see that the analysis can in fact be pushed much further back into the structure, not only of the second part of a pair, but of the first part as well. Take again example (47), in which we noted that the delay of two-tenths of a second seems to be taken as evidence of a dispreferred action – namely an invitation rejection – coming up, so that A then adds further inducements (*lotta stuff* being intended to indicate that plenty of food and drink is available, including the specified beer). Similarly, reconsider (55), where again a small delay (marked (.)) is sufficient to indicate to R

[20] Note that blamings receive denials in simple preferred format, thus indicating again that preference cannot be identified with, for example, blamer's desire; see Atkinson & Drew, 1979: 80.

that there may be a problem with R's suggestion, namely the one R suggests. Or again:

(59) *144*
C: ...I wondered if you could phone the vicar so that we could ((in-breath)) do the final on Saturday (0.8) morning o:r (.) afternoon or
(3.0)
R: Yeah you see I'll I'll phone him up and see if there's any time free
(2.0)
C: Yeah
R: Uh they're normally booked Saturdays but I don't- it might not be

Here over the course of C's first turn there are a number of slots provided where R could have performed the preferred compliance with C's request (these include the prolonged in-breath, the eight-tenths of a second pause, the lengthened *o:r* and its following short pause, and of course the long three-second silence after the turn). Given that preferred actions are properly done without delay, the fact that R's compliance is systematically delayed indicates that significant problems are coming up.

What such examples illustrate is that over the course of a single turn's construction, interactional feedback is being systematically taken into consideration (see Davidson, in press). In this sense a single turn at talk by one speaker can itself be seen to be a joint production, here by virtue of the strong expectations for no gap between the transition of speakers provided by preference organization. There is also further evidence of quite different kinds which shows that a single speaker's turn is often a joint production, in that recipients' non-verbal responses are utilized to guide the turn's construction throughout the course of its production (see Goodwin, 1979a, 1981). Here, though, preference organization, in constraining the construction of second parts of adjacency pairs, can systematically affect the design of first parts – and as we shall see this can happen in more ways than one.

Preference organization, however, extends far beyond the confines of adjacency pairs. There are, for a start, kinds of turns paired less tightly than adjacency pairs, where a first part does not seem to require but rather makes apt some response or second – **action-chains** in Pomerantz's (1978) terminology. For example, after an

assessment (or assertion expressing a judgement) a second assessment is often due, as in:

(60) *Pomerantz, 1975: 1*
J: T's – it's a beautiful day out isn't it?
L: Yeah it's jus' gorgeous ...

(61) *Pomerantz, 1975: 1*
A: (It) was too depre//ssing
B: O::::h it is te::rrible

Given a first assessment there is a clear preference for agreement over disagreement. Disagreements here, and after assertions in general, typically have a *yes, but* kind of format (i.e. disagreement, prefaced with token agreement), or they are delayed, or prefaced with *well* like other dispreferreds:

(62) *Pomerantz, 1975: 66*
R: ... Well never mind. It's not important.
→ D: *Well*, it is important.

(63) *Pomerantz, 1975: 68*
R: ... You've really both basically honestly gone your own ways.
→ D: *Essentially*, except we've hadda good relationship at //home
→ R: .hhh *Ye:s*, but I mean it's a relationship where ...

We are now in a position to appreciate one kind of complexity that arises, where two different kinds of conversational expectations work in opposing directions. One such area is self-denigration: by the preference for agreement after assessments, if A self-denigrates, an agreement from B is preferred. But by an independent principle of a different order, namely a norm enjoining the avoidance of criticism, B should avoid such an agreement. The latter principle in fact generally takes precedence (if agreements occur at all after self-deprecations they are preceded by disagreements – see Pomerantz, 1975: 101):

(64) *Pomerantz, 1975: 93*
L: ... I'm so dumb I don't even know it. hhh! heh
→ W: Y-no, *y-you're not du:mb...*

(65) *Pomerantz, 1975: 94*
L: You're not bored (huh)?
→ S: Bored? No. We're fascinated.

It follows from this, and the nature of pauses as markers of dispreferred responses, that there is an asymmetry in the significance of a pause

after an ordinary assessment like (66) and after a self-deprecating assessment like (67):

(66) A: God isn't it dreary!
B: ((SILENCE = DISAGREEMENT))

(67) A: I'm gettin fat hh
B: ((SILENCE = AGREEMENT))

Further complexities arise in another special kind of assessment, namely compliments. Once again there are cross-cutting principles at work: a preference for agreement with the compliment, and a norm specifying the avoidance of self-praise. Compromise solutions employed here include down-graded agreements, shifts of praise to third parties, and plain disagreements (Pomerantz, 1978).

6.3.2 *Preferred sequences*

So far we have been concerned with how preference operates over a range of alternative seconds to some prior turn. We have, though, indicated that it can operate to structure that prior turn during the course of its production; we have also briefly indicated that the delay component of a dispreferred second can be realized by what may be called a **next turn repair initiator**, or NTRI, which invites repair of the prior turn in the next turn, as in (54) above, where M asks *You want what?*, or as in the arrowed turn below:

(68) *Pomerantz, 1975: 74*
A: Why *wha*t'sa *ma*tter with y-you sou//nd *HA*:PPY, hh
B: Nothing
→ B: I sound ha:p//py?
A: *Ye*:uh
(0.3)
B: No:,

A dispreferred 'second' turn can thereby become displaced into fourth turn, by the sequence: A:((ASSESSMENT)), B: ((NTRI)), A: ((RE-ASSESSMENT)), B: ((DISPREFERRED SECOND)). One motive here is that B thereby provides A with an opportunity to re-formulate the first turn in a more acceptable way. So preference organization can and often does spill over into a number of turns subsequent to a first turn.

One area where preference organization routinely operates within and across turns is a central conversational device, the organization of **repair** (Schegloff, Jefferson & Sacks, 1977). As was pointed out above, the tendency for an utterance to attend to those immediately

prior to it provides, for both analysts and participants, a 'proof procedure' for checking how those turns were understood. This would be of little use if there was no device for the correction of misunderstandings, mishearings or indeed non-hearings. There is of course such a device and it has the following properties. First, it provides a number of systematic slots across (at least) a three-turn sequence in which repair, or at least its prompting, can be done, as follows:

(69) T1 (includes repairable item) = first opportunity: here for self-initiated self-repair
Transition space[21] between T1 and T2 = second opportunity: here again for self-initiated self-repair
T2 = third opportunity: either for other-repair or for other-initiation of self-repair in T3
T3 = fourth opportunity: given other-initiation in T2, for other-initiated self-repair

There are two important distinctions here: first, **self-initiated** contrasted to **other-initiated** repair – i.e. repair by a speaker without prompting vs. repair after prompting; secondly, **self-repair**, repair done by the speaker of the problem or repairable item, contrasted to **other-repair**, done by another party. An example of repair in each opportunity should help to make the distinctions clear:

(70) *Schegloff, Jefferson & Sacks, 1977: 364*
(Illustrating self-initiated self-repair in opportunity 1)
N: She was givin' me a:ll the people that were go:ne
→ this yea:r I mean this quarter y'//know
J: Yeah

(71) *Schegloff, Jefferson & Sacks, 1977: 366*
(Illustrating repair in opportunity 2, again self-initiated self-repair)
L: An' 'en but all of the doors 'n things were taped up =
→ =I mean y'know they put up y'know that kinda paper 'r stuff, the brown paper.

(72) *Schegloff, Jefferson & Sacks, 1977: 378 (Illustrating other-initiated other-repair in opportunity 3)*
A: Lissena *pig*eons.
(0.7)
→ B: Quail, I think.

(73) *Schegloff, Jefferson & Sacks, 1977: 367 (Illustrating other-initiation of self-repair in opportunity 3)*
A: Have you ever tried a clinic?
→ B: *What?*
A: Have you ever tried a clinic?

(74) *Schegloff, Jefferson & Sacks, 1977: 368 (Illustrating self-repair in opportunity 4, following other-initiation by NTRI)*
B: .hhh Well I'm working through the Amfat Corporation.
→ A: The *who?*
→ B: Amfah Corporation. T's a holding company.

[21] *Transition space* labels "the beat that potentially follows the possible completion point of a turn" (Schegloff, Jefferson & Sacks, 1977: 366). A more detailed analysis here might be: first opportunity is immediately after the error, second is at end of turn, third is after recipient delay at end of turn, fourth at T2, fifth at T3, and sixth at still further remove (Schegloff, in prep. a).

The range of phenomena collected here under the concept of *repair* is wide, including word recovery problems, self-editings where no discernible 'error' occurred, corrections proper (i.e. error replacements) and much else besides. The claim (Schegloff, Jefferson & Sacks, 1977) is that the same system handles the repair of all these problems. The examples above are only illustrative: there are many ways, for example, in which self-repair within the turn is signalled (e.g. by glottal stops, lengthened vowels, long schwa, etc.), or other-initiation of self-repair is achieved (e.g. by *What?*, *'Scuse me?*, etc., or by echo-questions, or repetitions of problematic items with stress on problem syllables as in (74), (77) and (78)).

Now the second major component of the repair apparatus is a set of preferences setting up a rank ordering across the opportunity set above. Briefly, the preference ranking is as follows:

(75) *Preference 1* is for self-initiated self-repair in opportunity 1 (own turn)
Preference 2 is for self-initiated self-repair in opportunity 2 (transition space)
Preference 3 is for other-initiation, by NTRI in opportunity 3 (next turn), of self-repair (in the turn after that)
Preference 4 is for other-initiated other-repair in opportunity 3 (next turn)

The evidence for such a ranking is, first, that this corresponds to the ranking from the most frequently used to the least used resource, (other-repair, for example, being really quite rare in conversation). Secondly, the system is actually set up so that there will be a tendency for self-initiated self-repair, this being the type of repair relevant in the first two opportunities traversed. Thirdly, we have the typical delay by recipient following these two opportunities if they're not

immediately utilized, indicating a 'problem' and inviting self-initiated self-repair. Sometimes the invitation by the non-responder is successful as in (76), sometimes not, as in (77):

(76) *Schegloff, Jefferson & Sacks, 1977 : 364*
K: Did*ju* know the guy up there at-oh. What the hell is'z name use to work up't (Steeldinner) garage did their body work for 'em.
→ (1.5)
K: Uh::ah, (0.5) Oh:: he meh- uh, (0.5) His wife ran off with Jim McCa:nn.

(77) *Schegloff, Jefferson & Sacks, 1977 : 370*
A: Hey the first time they stopped me from sellin' cigarettes was this morning.
→ (1.0)
B: From *sell*ing cigarettes?
A: From buying cigarettes.

Fourthly, there is clear evidence that even where other parties can do the required repair, they produce an NTRI (i.e. other-initiation of a self-repair) instead of doing other-repair on many and probably most occasions. B's turn in (77) is a case in point, and the fourth turn in the following makes this explicit:

(78) *Schegloff, Jefferson & Sacks, 1977 : 377*
K: 'E likes that waiter over there,
A: Wait-*er*?
K: Waitress, sorry,
→ A: 'Ats better,

Finally in the rare event of other-repair occurring, it is followed by 'modulators' like the *I think* in (72) above, or prefaced by *y'mean*, or otherwise marked:

(79) *Schegloff, Jefferson & Sacks, 1977 : 378*
L: But y'know single *beds*'r *aw*fully *thin* to *sleep* on.
S: What?
L: Single beds. // They're-
→ E: Y'mean narrow?
L: They're awfully *na*rrow yeah.

So the repair apparatus as a whole is strongly biased both by a preference for self-initiation of repair and by a preference for self-repair over repair by others. As a consequence preference organization governs the unfolding of sequences concerned with repair.

We have now widened the scope of preference organization to cover not only rankings of alternative turns, but alternative solutions to problems (like the handling of repair), the solutions being either handled within a single turn or across a sequence of turns. However, preference also seems to operate across *sequence types* – for example, it seems that, if possible, prompting an offer is an action preferable to performing a request (Schegloff, 1979a: 49). Hence, as we shall see in 6.4.3, there is a special utility in a turn designed to prefigure or preface a request (a **pre-request**), for it provides for the possibility of recipient performing an offer instead, as below:

(80) *176*
→ C: Hullo I was just ringing up to ask if you were going to Bertrand's party
R: Yes I thought you might be
C: Heh heh
→ R: Yes would you like a lift?
C: Oh I'd love one

In a similar way, in the initial three turns of telephone calls recognition is done by, and totally submerged within, greetings (where this is possible) in preference to being achieved by a sequence involving overt self-identifications. Thus (81) is the preferred sequence, (82) the dispreferred:[22]

(81) *Schegloff, 1979a : 35*
C: ((rings))
T1 R: Hello
T2 C: Hello
T3 R: Hi

(82) *Schegloff, 1979a : 59*
C: ((rings))
T1 R: Hello,
T2 C: Hi. Susan?
T3 R: Ye:s,
T4 C: This's Judith (.) Rossman
T5 R: *Ju*dith!

The evidence that self-identification is in general dispreferred is this. Despite the fact that R can be assumed to have a strong interest in learning who is calling, in the first turn by callers (T2) self-identification is often absent; if, given this turn, identification is not

[22] This sort of preference may be quite culture-specific; see Godard, 1977 on French conventions.

immediately achieved by R, then C generally leaves a gap or pause for recognition to occur – hence, as we saw in 6.2.3, delay after a minimal T2 is understood as evidence of a problem in R's recognition of C. Only after such a delay does C offer an identification (as in (45) above) or does R request one (as in (46)).

Now given this dispreference for self-identifications, if a caller wishes to avoid overt self-identification but is uncertain that recipient can do identification on a minimal T2 (e.g. *Hello*), then he can produce a T2 that prefigures or prefaces a self-identification while withholding it, thus giving the recipient an opportunity to recognize the caller without actually claiming (as a plain *Hello* would) that recipient can do so. The caller can achieve this by the use of *Hello* plus the name of the recipient, with the characteristic low-rise intonation contour of a 'try' on the name, as in :

(83) *Schegloff, 1979a : 52*
R: Hello:,
→ C: Hello *Il*se?
R: Yes. *Be* :tty.

(Note that a high-rise contour would seem to signal genuine uncertainty about the recipient's identity, whereas a low-rise primarily conveys uncertainty about whether the recipient can recognize the caller – see Schegloff, 1979a: 50.) Now just as a pre-request invites the recipient to provide an offer, thus avoiding the dispreferred request sequence, so *Hello* plus name (with low-rise contour) invites recognition (as in (83)) in preference to the self-identification it prefigures, namely the dispreferred sequence instanced by (82). Hence such a dispreferred sequence, with overt self-identification, will generally only occur if in T3 recipient shows no evidence of recognition, like an address term or enthusiastic response (e.g. *Oh hi! How are you?*). And that it is dispreferred is further shown by the fact that, on the absence of recognition in T3, the caller sometimes provides no more identificatory material in T4 than a further sample of voice-quality:

(84) *Schegloff, 1979a : 55*
R: H'llo:?
C: Harriet?
R: Yeah?
→ C: Hi!
R: *Hi* :.

Finally, the dispreferred nature of self-identification is confirmed by the fact that, if it is after all required, it is often received with 'the big hello' – the upgraded recognition component in e.g. T5 in (82), regularly with an account of why recognition was not achieved earlier (e.g. *You sound different* – see Schegloff, 1979a: 48).

So for telephone recognitions between known parties the preference is for caller to provide the minimal cues he judges sufficient for recipient to recognize caller (note here the pause before the provision of the surname in (82)). And such a preference not only ranks T2 cues as in (85):

(85)
(i) Hi
(ii) Hello
(iii) Hello. It's me
(iv) Hello. It's Penny
(v) Hello. It's Penny Rankin

but it also ranks the two sequence types, (i) greetings alone, (ii) greetings followed by overt self-identifications and recognitions.

Preference organization thus extends not only across alternative seconds to first parts of adjacency pairs, but backwards into the construction of first parts, forwards into the organization of subsequent turns, and also across entire alternative sequences, ranking sets of sequence types.

6.4 Pre-sequences

6.4.1 *General remarks*

The term **pre-sequence** is used, with a systematic ambiguity, to refer both to a certain kind of turn and a certain kind of sequence containing that type of turn. We will, however, use the abbreviation **pre-s** for the turn-type, reserving *pre-sequence* for the sequence type. Some examples of pre-sequences and pre-s have already been introduced in passing. We have noted for example that a summons prefigures a turn which contains a reason for the summons, as in:

(86) *Atkinson & Drew, 1979 : 46*
Ch: Mummy
M: Yes dear
(2.1)
Ch: I want a cloth to clean (the) windows

Since such reasons can be various, summonses are 'generalized pre-s'; most pre-s, however, are built to prefigure the specific kind of action that they potentially precede. For example, pre-closings, often realized as tokens of *Okay*, are recognizable as potential initiations of closings, otherwise closings could not be co-ordinated. Pre-closings illustrate one major motivation for pre-s in general, namely that by prefiguring an upcoming action they invite collaboration in that action (as in pre-closings) or collaboration in avoiding that action (as in pre-self-identifications).

Some of the clearest kinds of pre-s are **pre-invitations**, like the following:

(87) *Atkinson & Drew, 1979 : 253*
→ A: Whatcha doin'?
B: Nothin'
A: Wanna drink?

(88) *Atkinson & Drew, 1979 : 143*
R: Hi John
C: How ya doin=
→ =say wh*at*'r you doing?
R: Well we're going out. Why?
C: Oh, I was just gonna say come out and come over here an' talk this evening, but if you're going out you can't very well do that

Notice that in both cases the pre-invitations are treated as transparent by the recipients – so that their responses are clearly attuned to the fact that an invitation (or related act) is potentially forthcoming in the next turn. Thus *Nothing* in (87) can be read as 'nothing that would make the offer of an evening's entertainment irrelevant' or the like, while the formulation of what R is doing in (88) is clearly attuned to the possibility of an upcoming invitation, which the *Why?* requests details of.

A pre-s is not just some turn that comes before some other kind of turn – most turns have that property; it is a turn that occupies a specific slot in a specific kind of sequence with distinctive properties. On the basis of examples like the pre-invitation ones above, we might attempt the following characterization of the structure of such sequences (although such a characterization requires generalization to other kinds of pre-sequence):

(89) (a) T1 (Position 1): a question checking whether some precondition obtains for the action to be performed in T3

T2 (Position 2): an answer indicating that the precondition obtains, often with a question or request to proceed to T3
T3 (Position 3): the prefigured action, conditional on the 'go ahead' in T2
T4 (Position 4): response to the action in T3

(b) *distribution rule*: one party, A, addresses T1 and T3 to another party, B, and B addresses T2 and T4 to A

Of course, a crucial part of the motivation for such a sequence is the conditional or contingent nature of T3 on the nature of T2; so, in the absence of an encouragement in T2, the sequence can be expected to abort along the following lines:

(90) T1: as in (89)
T2: answer indicates that precondition on action does not obtain – often so formulated as to specifically discourage the foreseeable action
T3: withholding of the prefigured action, usually with a report of what would have been done in T3, by way of explanation for T1

Such a sequence occurs in (88).

Given such a characterization, we then have no difficulty finding further kinds of pre-sequence, e.g. **pre-requests** like the following:

(91) *Merritt, 1976 : 337*
→ C: Do you have hot chocolate?
S: mmhmm
→ C: Can I have hot chocolate with whipped cream?
S: Sure ((leaves to get))

(92) *Merritt, 1976 : 324*
→ C: Do you have the blackberry jam?
S: Yes
→ C: Okay. Can I have half a pint then?
S: Sure ((turns to get))

(93) *172B(7)*
→ C: So um I was wondering would you be in your office on Monday (.) by any chance (2.0) probably not
R: Hmm yes=
C: =You would
R: Yes yes
(1.0)
→ C: So if we came by could you give us ten minutes or so?

Similarly one can recognize **pre-arrangements** for future contact, as in:

(94) *176B*
→ R: Erm (2.8) what what are you doing today?
C: Er well I'm supervising at quarter past
(1.6)
→ R: Er yuh why (don't) er (1.5) would you like to come by after that?
C: I can't I'm afraid no

Now there is a problem raised for our characterization of such sequences by examples like (93),[23] where the distribution of the characteristic actions is not exactly over the paradigmatic four-turn sequence proposed in (89). Such problems are more acute in examples like the following, where we have two **insertion sequences**, one in T2 and T3 concerned with repair, one in T4 and T5 concerned with establishing a temporary 'hold' in the turn-taking system, both these sequences being inserted between a pre-request on the one hand (in T1), and its response (in T6) with the follow-up request (in T7) on the other.

(95) *144(3)*
T1 → C: ... Do you have in stock please any L.T. one eight eight?
((POSITION 1))
T2 R: One eight eight ((HEARING CHECK))
T3 C: Yeah= ((CHECK OKAYED))
T4 R: =Can you hold on please ((HOLD))
T5 C: Thank you ((ACCEPT))
(1.5)
T6 R: Yes I have got the one ((POSITION 2))
T7 → C: Yes. Could I- you hold that for H.H.Q.G please
((POSITION 3))

Intuitively, then, we have two insertion sequences between what would be the paradigmatic turns T1 and T2 in (89). But if we have defined the notion of a pre-sequence in terms of just the sequence of turns in (89), how is this case to be assimilated? What we need here is the distinction introduced recently by Schegloff in the discussion of repair, between **turn location** – i.e. sheer sequential locus of a turn in a sequence by a count after some initial turn – and **position**, the response to some prior but not necessarily adjacent turn. Thus a second part of an adjacency pair separated from its first part by a two-turn insertion sequence will be in *fourth turn* but *second position*. What we want to claim, therefore, is that the structure in (89) actually

[23] And compare the different transcription of the same data in (39).

holds across a sequence of *positions* rather than *turns*, and that the utterance T6 in (95) is in *second position* despite being sixth in turn, and likewise T7 is in *third position* though seventh in turn.

But if this distinction between *turn* and *position* is not to render vacuous the claim that pre-sequences ordinarily have the structure in (89), then we must have an independent characterization of each *position*, so that it can be recognized wherever in a sequence of turns it actually shows up. This is not easy to do in a general way for all pre-sequences, but it can certainly be done for sub-classes of pre-sequences, like that including pre-requests, pre-invitations, pre-offers and the like. Indeed the glosses on the content of each turn (or, preferably now, position) in (89) indicate some typical recognizable features of each. Let us take up this problem of characterizing turns in particular positions with respect to one particular type of pre-sequence.

6.4.2 *Pre-announcements*

One class of pre-s of special interest are **pre-announcements** (see Terasaki, 1976, on whose work the following is based). We have already met the sub-class of these which are bids for story space (see examples (41) and (42) above), noting that they operate to gain ratified access to an extended turn at talk. But there are many other kinds of pre-announcements, like the following:

(96) *Terasaki, 1976 : 36*
D: .hh Oh guess what.
R: What.
D: Professor Deelies came in, 'n he- put another book on 'is order.

(97) *Terasaki, 1976 : 53*
D: I forgot to tell you the two best things that happen' to me today.
R: Oh super=what were they.
D: I got a B+ on my math test ... and I got an athletic award.

(98) *Terasaki, 1976 : 53*
D: Hey you'll never guess what your dad is lookih-is lookin' at.
R: What're you looking at.
D: A radar range.

Let us now attempt to characterize such sequences. One way of thinking about them (and perhaps pre-sequences in general) is that

they are made up of two superimposed adjacency pairs: a pre-pair (e.g. A: *Have you heard the news?*, B: *No*) and a second pair (e.g. B: *Tell me*, A: *John won the lottery*) – superimposed in that the second of the first pair and the first of the second pair occur in the same turn or position – namely position 2. Hence we often find in the turn occupying position 2 dual components of the kind in the second turn in (97), where *Oh super* looks backward to the prior turn, and *what were they* is a first part requiring the announcement as a second. We thus have the following structure for pre-announcements:

(99) *Position 1*: pre-sequence first part, generally checking on newsworthiness of potential announcement in position 3
Position 2: pre-sequence second, generally validating newsworthiness, *and* first part of second pair, namely a request to tell
Position 3: second part to second pair – the announcement delivered
Position 4: news receipt

As required by the distinction between *turn* and *position*, we should attempt to characterize the format of each position independently of its sheer sequential location (though the order of the positions must of course be maintained, so sequential considerations can still play an important part in the recognition of specific positions). Thus we can say of position 1 turns that, although they may be in any of the three main sentence-types (e.g. interrogative in (41) above, imperative in (96), declarative in (97) and (98)), they typically have at least one of the following elements: they *name* the kind of announcement (e.g. *what your dad is looking at* in (98), *the news* in the examples immediately below); and/or they evaluate it e.g. as *good news* in (100), and *terrible news* in (101):

(100) *Terasaki, 1976 : 33*
D: Hey we got good news.
R: What's the good news.

(101) *Terasaki, 1976 : 28*[24]
D: Didju hear the terrible news?
R: No. What.
D: Y'know your Grandpa Bill's brother Dan?
R: He died.
D: Yeah.

[24] This example is a principled exception to the schema in (99), as explained below in terms of a preference for *guessing* over *telling* in the case of bad news.

Further they often date the news (e.g. the specification *today* in (97)); and finally, and most importantly, such position 1 turns generally have some *variable*, a WH-word (as in *What?* in (96) and (98)) or an indefinite phrase (*a good thing happened*) or a definite but non-specific phrase (*the news*). It is of course this variable that these first turns in the sequence offer to instantiate in position 3.

Position 2 turns are generally characterized by (a), optionally, a response to position 1 taken as a question (e.g. *No* in (101)), and (b) almost invariably a question-like component. These question components are one-word questions like *What* (as in (96)) or echo-questions or questions like *What were they?* in (97), which copy parts of position 1's material. That is, they are built like NTRIs (next turn repair initiators), which include of course echo-questions. What they share with NTRIs is that they have the same double-directedness – they look back to prior turn (making possible the truncated format) and they look forward to next turn (hence the question format). In this way the format of position 2 turns is designed both as a second to position 1 turns and as first parts taking position 3 turns as seconds.

Turning to position 3 turns, the announcements themselves, we find a series of tight constraints on their format. For example, they sometimes retain the syntactic or case frame of their corresponding pre-announcements in position 1 (Terasaki, 1976):

(102) *Terasaki, 1976 : 26*
→ D: Oh. You know, Yuri did a terrible thing
((SUBJ)) ((VERB)) ((DIRECT OBJECT))
R: hhh! I kn*ow*.
D: You know?
→ She committed *sui*cide
((SUBJ)) ((VERB)) ((DIRECT OBJECT))

Alternatively, position 3 turns provide just the items that would fill the variable slot (here in bold face for recognitional purposes) typical in position 1 turns:

(103) *Terasaki, 1976 : 53*
→ D: Y'wanna know **who** I got stoned with a few w(hh)eeks ago? hh!
R: Who.
→ D: Mary Carter 'n her boy(hh) frie(hhh)nd. hh.

Note here too the tie back of position 3 to position 2, since position 3 provides just the information solicited in position 2 (and offered in

position 1). There are other variations, but the point here is that each position is indeed characterizable, independently of absolute location in a sequence of turns, as having certain kinds of (alternative) format.

Clearly, the design of the turn in position 1 is crucial: for it is on the basis of this that the recipient must decide whether or not he already knows the content of the announcement, and thus should abort the sequence. Hence the prefiguring of the syntactic frame of the announcement, as in (102), is a very useful clue to the recipient, as is the characterization of the announcement as 'news' or as a 'joke' or 'story', the dating of reportable events, and the evaluation of 'news' as 'good', 'terrible', etc. So we can appreciate that some phrase like *the two best things that happened to me today* (in (97)) is carefully formulated to prefigure what is coming up – namely two items, good things, and things that happened today.

It is instructive in this regard to reconsider an analysis by Labov & Fanshel (1977) of the very beginning of a psychiatric interview:

(104) *Labov & Fanshel, 1977 : 363 (transcription conventions converted to CA style)*

R: I don't (1.0) know, whether (1.5) I- I think I did- the right thing, jistalittle situation came up (4.5) an' I tried to uhm (3.0) well, try to (4.0) use what I- what I've learned here, see if it worked
(0.3)
T: Mhm
R: Now, I don't know if I did the right thing. Sunday (1.0) um- my mother went to my sister's again ... ((story continues))

In seventeen pages of painstaking analysis in DA style, Labov & Fanshel (1977: 113ff) analyse the patient's (R's) first turn here as containing various speech acts including questions, assertions and challenges. To achieve an understanding in such depth they look forward in the interaction to see what *the right thing* and *jistalittle situation* refer to; they then pack back into the gloss or "expansion" of the first turn these details gleaned from later on. Despite the obvious discrepancy between the information thus available to participants (who cannot look ahead in a transcript) and analysts (who can), the authors feel this procedure is justified by the analysts' relative lack of the knowledge available to participants about each other (ibid.: 120). They further argue that the various features here, including the glottalizations and hesitancy, and crucially the "vague

reference" in *thing* and *situation*, can be attributed to aspects of "interview style" (ibid.: 129).

Now contrast an analysis in CA style. R's first turn is a pre-announcement, formulated to prefigure (a) the telling of something she did (*I think I did the right thing*), and (b) the describing of the situation that led to the action (*jistalittle situation came up*). We are therefore warned to expect a story with two such components; moreover the point of the story and its relevance to the here and now is also prefigured (*use what I've learned here, see if it worked*). The alleged vagueness of *the right thing* and *jistalittle situation* is in fact the provision of just those variables typical of position 1 turns in pre-announcement sequences. That a pre-sequence analysis seems correct is reinforced by the fact that the recipient, the therapist, does indeed wait for each prefigured segment of the story, receiving the first with *Oh* (a typical news receipt item – see Heritage, in press), and the second with an agreement (*Yes I think you did (the right thing) too*), abstaining from any other substantial turns throughout the story. The point to be made here is that the original DA-style analysis, proceeding in an act-by-act fashion, is not attuned to the larger sequential structures that organize conversation; nor are such structures easily recognizable without a lot of comparative material.[25]

The recognition of pre-announcements can be problematic not only for analysts, but for participants too. Consider, for example, (105):

(105) *Terasaki, 1976 : 45*

T1		Kid:	I know where you're going.
T2		Mom:	Where.
T3		Kid:	To: that (meeting ...)
T4		Mom:	Right. Yah!
	→		Do you know who's going to that meeting?
T5		Kid:	Who.
T6	→	Mom:	I don't know!
T7		Kid:	Ou::h prob'ly: Mr Murphy an' Dad said prob'ly Mrs Timpte an' some o' the teachers.

Here in T4 we have a turn in question format: Kid takes it to be a

[25] Incidentally, CA can also provide rival analyses of other features of this first turn – for example, the hesitation and glottal stops attributed in the original analysis to "style" are also the typical markings of self-initiated self-repair, which is characteristic of the production of first topics (see Schegloff, 1979b), and is also used to request listener attention (Goodwin, 1981: Chapter 2).

pre-announcement, and for good reasons – the sequential locus is such that given his own use of a guess in T1, he may expect another 'riddle' in return. So in T5 Kid solicits the prefigured announcement, with *Who*. But it turns out that Mom intended T4 as a question, not a pre-announcement, as her response in T6 makes clear. Notice that Kid can then in fact produce an answer in T7, indicating that the *Who* in T5 was only intended as 'Go ahead and tell', not 'I don't know'. An ambiguity of this sort, which is shown here to be an ambiguity for participants like that in (49) above, is a good example of the kind of phenomenon that analysts using intuitions as data have failed to notice. Instead, such theorists might be concerned about another ambiguity, namely that between the 'direct' speech act interpretation of *Do you know p?* (for which *Yes* or *No* would be a complete and adequate response) and the 'indirect' speech act interpretation as a request to tell, which is not an issue for participants here (Schegloff, in prep. b).

What motivates the use of pre-s like these pre-announcements? There seem to be a number of motivations, sometimes working concurrently. We have already sketched, for pre-story turns like that in (42), a motivation based on the turn-taking system: if a speaker wishes to suspend temporarily the relevance of possible transition at each TRP, he may make a bid for ratification of an extended turn. This motivation, Sacks has pointed out, explains the frequent use of pre-announcements by those with restricted rights to speak – hence children's use of formulae like *Want to know something, Daddy?*

However, perhaps the most prominent motivation for pre-announcements is a keen concern with not telling people things that they already know. The concern ramifies through this and other conversational organizations, motivating a general tendency to "oversuppose and undertell" (Sacks & Schegloff, 1979), which we have already met in the domain of telephone identifications (Schegloff, 1979a: 50). Grice's maxim of Quantity, and its basis in rational co-operative efficiency, seems scant motive for the strong interactional aversion to self-repetition. In any case, the extraordinary fact is that everyone is expected to keep an account book, as it were, of every matter talked about with every other co-participant. If uncertainty does arise, or if there is reason to suppose that some other third party may have already imparted the 'news', then recourse can be had to the pre-announcement, which offers to tell contingently upon the

'news' not already being known. Thus solicitations to tell in position 2 will tend to commit the recipient to claiming that he has not previously heard the news (a commitment escapable – but not without loss of grace – by an *Oh that*, or the like, in position 4). A puzzle that then arises is how, from the sometimes fairly non-specific turns in position 1, participants can effectively judge whether what is prefigured is already known, as they confidently do:

(106) *Terasaki, 1976: 26*
D: ... Hey we got good news.
R: I know.

Here, besides the format of the pre-announcement itself, participants rely on features like sequential context (stories, for example, are often topic-tied to prior turns, providing a resource for guessing what story may now be told – see Jefferson, 1978) and the dating of news provided by the occasion of last meeting (whatever was 'news' then, should have been delivered then, so 'news' now must be 'news' since then – see Sacks, 1975).

We thus see in pre-announcements a concern, reflected in a sequential organization, with the distinction between given and new information that we have discussed elsewhere under the rubrics of presupposition and the maxim of Quantity. The concern runs deep in the usage of pre-announcements: the structure of position 1 turns is often so designed that it provides a frame that is given information, and a variable whose instantiation is thought to be new (as in (98), where the frame is *You'll never guesswhat your dad is looking at*, given by the situation, and what is delivered in position 3 is just what is new: *A radar range*). Further, in cases where a pre-announcement is delivered to a set of recipients in potentially different states of knowledge, one finds position 1 turns like *Some of you may not have heard the news*; and if one of these recipients is 'in the know' he may produce a 'collaborative' T2 like *Yeah tell 'em about it*, thus carefully excluding himself from those 'in the dark' (Terasaki, 1976: 20ff).

But there are other motivations for pre-announcements besides these. One important one involves preference organization, which we have noted can rank order not only alternative turns but also the choice between alternative entire sequences. Offer sequences can thus be preferred to request sequences (see example (80) and discussion above); and implicit recognition in greetings preferred to overt

self-identification sequences in the openings of telephone calls. So also there seems to be, in the case of the delivery of 'bad news', a preference for B guessing over A telling. Just as a pre-request can, in projecting an upcoming request, secure an offer, so a pre-announcement can obtain, and be specifically designed to obtain, a guess:

(107) *Terasaki, 1976 : 29*
D: I-I-I had something *terr*ible t'tell you.
So // uh
R: How terrible *is* it.
D: Uh, th- as worse it could *be*.
(0.8)
R: W- y'mean Edna?
D: Uh yah.
R: Whad she do, die?
D: Mm:hm,

Note here the delay after the third turn, that seems specifically to invite a guess. And note that in (101) above, the first pre-announcement is followed by a second, until recipient guesses. So another motivation for pre-announcements is that by prefiguring a dispreferred action, the telling of bad news, they can prompt a guess by the other party that obviates the need to do the dispreferred action at all.

In this discussion of pre-announcements we have shown that (a) a specific kind of sequence can be properly characterized as an ordered sequence of not necessarily contiguous turns of distinctive type, (b) on inception, it is recognizable to participants by virtue of characteristic position 1 turns, and (c) the usage of such sequences is strongly motivated by various principles of language usage.

6.4.3 *Pre-requests : a re-analysis of indirect speech acts*

We now have all the ingredients for a powerful re-analysis of the problem of *indirect speech acts*.[26] Strictly speaking, perhaps, we should say that on the CA view the alleged problem does not even arise; and in any case the terms of the two kinds of analyses are so starkly different that what is a problem for the philosophico-linguistic approaches is not for the CA approach, and perhaps vice versa.

[26] A number of authors should be credited here: Schegloff, in unpublished work; Goffman, 1976; Merritt, 1976; Coulthard, 1977: 71; Heringer, 1977. The argument also benefits from unpublished work by Paul Drew and John Heritage.

The problem, recollect, from the point of view of speech act theory, is that indirect speech acts do not have the 'literal force' (allegedly) associated by rule with their sentence-types, but rather some other force which a theory of indirect speech acts is concerned to explain. Thus the question is how, for example, sentences like *Is there any more?*, or *Can you reach that book?* or *Will you come here please?* can be effectively employed to perform requests. Let us restrict ourselves to requests as the variety of indirect speech act which has received the most attention.

A CA analysis might go roughly as follows. Pre-request sequences, we noted, properly have a four-position structure, providing the following kind of analysis:

(108) *Merritt, 1976 : 324*

Position 1:	A: Hi. Do you have uh size C flashlight batteries?	((PRE-REQUEST))
Position 2:	B: Yes sir	((GO AHEAD))
Position 3:	A: I'll have four please	((REQUEST))
Position 4:	B: ((turns to get))	((RESPONSE))

Now, as we have argued, it is possible to distinguish sheer sequential location in a sequence of turns, from position or location in a sequence of responses. So we need an independent characterization of, for example, position 1 turns in pre-request sequences, thus giving an account of how they can be recognized prior to position 3 turns being performed. One characteristic, we noted, for a wide range of pre-sequences, is that position 1 turns check that conditions for successful position 3 turns obtain. Why should this be so?

In the case of requests it seems clear that one prime motivation for employing pre-requests is provided by the preference ranking which organizes responses to requests themselves. Request refusals are dispreferred: therefore, by the accompanying rule for production, to be avoided if possible. One major reason for utilizing a pre-request is, then, that it allows the producer to check out whether a request is likely to succeed, and if not to avoid one in order to avoid its subsequent dispreferred response, namely a rejection. Given which, in cases of doubt, pre-requests are to be preferred to requests.

One kind of evidence for this is that not just any precondition on a request is generally usable in a pre-request: rather just those that are, in the particular circumstances, the usual grounds for *refusal* of that request (Labov & Fanshel, 1977: 86ff). It is no accident, for

example, that questions about recipient's abilities figure so largely in both indirect speech acts of requesting and pre-requests – they are also the favoured basis for request refusals in conversation:

(109) *164*
(*Context : A has asked a third party B for change, C standing by now joins in :*)
C: How much do you want?
A: Well a fiver, can you do five?
→ C: Oh sorry, I'd be able to do a couple of quid

(110) *170*
A: Hullo I was wondering whether you were intending to go to Popper's talk this afternoon
→ B: Not today I'm afraid I can't really make it to this one
A: Ah okay
B: You wanted me to record it didn't you heh!
A: Yeah heheh
B: Heheh no I'm sorry about that, ...

Note that in (110) B treats A's first utterance as a transparent pre-request, hence the apologies typical of dispreferreds and the guess at the request that would have been relevant had B been able to attend.

While in conversation inability seems to be the preferred grounds for request refusals (in preference to, say, unwillingness), in service encounters (in shops, offices, bars, etc.) it seems that requests for goods are usually rejected by explaining that the desired goods are not in stock (see Ervin-Tripp, 1976; Sinclair, 1976). Hence the symmetry across the pre-request format and the refusal format below is not accidental:

(111) *Merritt, 1976 : 325*
→ C: Do you have Marlboros?
→ S: Uh, no. We ran out
C: Okay. Thanks anyway
S: Sorry

What is checked in the pre-request is what is most likely to be the grounds for refusal; and if those grounds are present, then the request sequence is aborted.

We now have at least a partial characterization of position 1 turns in pre-requests: they check (and are therefore generally questions) the most likely grounds for rejection. We also have a *motivation* for this particular format – namely avoiding an action (the request) that would obtain a dispreferred second (a rejection), and hence the

checking of the most likely grounds for refusal in position 1 turns.

But there is a further motivation for using pre-requests, namely the possibility of avoiding requests altogether. As we noted, preference organization operates not only over alternative seconds but also over alternative sequences, so that offer sequences seem to be preferred to request sequences (Schegloff, 1979a: 49).[27] By producing a pre-request in turn 1, one participant can make it possible for another to provide an offer in turn (or position) 2, an offer of whatever it was that the pre-request prefigured, as below:

(112) *Merritt, 1976 : 324*
C: Do you have pecan Danish today?
→ S: Yes we do. Would you like one of those?
C: Yes please
S: Okay ((turns to get))

(113) (*as in (80)*)
C: Hullo I was just ringing up to ask if you were going to Bertrand's party
R: Yes I thought you might be
C: Heh heh
→ R: Yes would you like a lift?
C: Oh I'd love one
R: Right okay um I'll pick you up from there ...

There is a further possibility, beyond the preference for offers over overt requests. For it may be that, after a pre-request, it is preferred that neither a request nor an offer takes place at all. There are some parallels for such avoidance of overt actions altogether. In telephone recognitions, as we noted, there seems to be a preference not only for the use of minimal resources for mutual recognition, but for the work of recognition to be actually submerged and hidden altogether in the exchange of minimal greetings – i.e. a preference for recognition to appear to be no issue at all. There is another, rather different, parallel: in repair we noted that there is a preference for self-initiated self-repair over other-initiated repair either by self or other. But if self fails to initiate repair, and other is to initiate, then there is a way of doing this that avoids the use of an NTRI and the consequent three-turn

[27] In addition to the other cases in 6.3.2, we have also seen that, when telling bad news, there seems to be a preference for the sequence (a) T1: pre-announcement, T2: guess, over the sequence (b) T1: pre-announcement, T2: go-ahead, T3: announcement.

repair sequence: other can do such repair when his turn naturally comes up, simply by substituting (for example) a 'correct' term for the offending one, as in the following (the relevant terms are in bold face):

(114) *Jefferson, MS.*
(Hardware store: customer trying to match a pipe-fitting. C = customer, S = salesman)
C: Mm, the **whales** are wider apart than that.
S: Okay, let me see if I can find one with wider **threads**. ((looks through stock))
S: How's this.
C: Nope, the **threads** are even wider than that.

Here S, the expert, simply substitutes *threads* for *whales*, and C subsequently adopts the suggested usage. So here 'correction' is being effectively achieved without ever becoming the sort of interactional issue that it can become if done through the normal three-position sequence of other-initiated self-repair:

(115) *Jefferson, MS.*
A: ... had to put new gaskets on the oil pan to strop-stop the leak, an' then I put- an then-
R: That was a gas leak.
→ A: It was an oil leak buddy.
→ B: 'T's a *gas* leak.
→ A: It's an oil leak.
((dispute continues for many turns))

Jefferson (MS.) calls correction of the sort in (114) **embedded**,[28] and the sort in (115) **exposed**, and there is reason to think that embedded correction is preferred to exposed, in part perhaps because issues of competence are not overtly raised. So the avoidance of a sequence altogether in favour of a covert solution can sometimes be preferred in repair work also.

Returning to requests, what we may now suggest is that there may also be a preference for the avoidance of requests altogether. So, if you can see that someone wants something, and a pre-request may be an effective clue to that, then it may be most preferred to provide it without more ado, next most preferred to offer, and third in preference to simply solicit the request. If this is correct, then after a pre-request, we have the following preference ranking operating

[28] A misleading terminology for linguists – better would be *covert* or *implicit*.

over three kinds of sequences (ignoring those that are aborted when preconditions are not met):

(116)
(i) most preferred: Position 1: (pre-request)
Position 4: (response to non-overt request)
(ii) next preferred: Position 1: (pre-request)
Position 2′: (offer)
Position 3′: (acceptance of offer)
(iii) least preferred: Position 1: (pre-request)
Position 2: (go ahead)
Position 3: (request)
Position 4: (compliance)

Some examples of sequences of type (i) follow:

(117) *Sinclair, 1976: 60*
S: Have you got Embassy Gold please? ((POSITION 1))
H: Yes dear ((provides)) ((POSITION 4))

(118) *Merritt, 1976: 325*
C: Do you have Marlboros? ((POSITION 1))
S: Yeah. Hard or soft? ((INSERTION SEQUENCE))
C: Soft please
S: Okay ((POSITION 4))

It may be objected that the initial turns here are just indirect requests – but note that they are in the format of pre-requests, and the difference between, for example, (108) and (117) is simply that position 2 and position 3 turns are missing in the latter. Since (117) might very well have run off similarly to (108), it would be a post-hoc distinction to say of the first turn in (108) that it is a pre-request while to say of the first turn in (117) that it is an indirect request. Sometimes, though, the sequential pattern (i) in (116) is absolutely clear (see also (120) below):

(119) *178*
(Context: R and Sheila have delivered a joint lecture course, but failed to provide C, compiler of the relevant examination, with questions for it)
((immediately after greetings))
C: Um (1.5) you and Sheila have been doing some lectures for first year Microbiol//ogy
R: Right and oh my God it's the third of March or whatever- yes – fourth of March (1.0)
er we'll get them to you (1.0) toda:y ...

We want to say that C's turn here is treated as a pre-request or position 1 turn, which receives in response a position 4 turn, i.e. a response (compliance) as if the request had itself been delivered.

Now the great majority of sequences of this truncated type have pre-requests that are distinctive – namely, they appear to be *built to obtain a position 4 in second turn*. Consider, for example:

(120) *US: 24*
M: What're you doing wi' that big bow-puh tank.
Nothing.
(0.5)
V: ((cough)) uh-h-h (1.0) I'm not into selling it or giving it.
That's it.
M: Okay

Here a pre-request in the form of a question has tagged on to it a presumption of its answer (*Nothing*). But in supplying the answer to the question that checks the precondition on the request, M invites a position 4 response directly – and gets it (but it is the dispreferred second, a refusal). Note that M's *Okay* in accepting the refusal acquiesces in the request interpretation. Recollect a parallel in the case of pre-announcements, where position 1 turns are often carefully formulated (a) to provide enough information about the upcoming announcement for recipients to judge whether they have already heard it, (b) sometimes so formulated as to extract guesses in position 2 of what would otherwise appear in position 3. Similarly, then, pre-requests can be built *specifically* to invite position 4 responses.

One technique for this formulation is to provide in the pre-request all the information that would be required for recipient to comply with the request. Thus there is a systematic difference in the way the following two sequences run off, given the difference between their first turns:

(121) *Sinclair, 1976: 68*
S: Can I have two pints of Abbot and a grapefruit and whisky? ((POSITION 1))
H: Sure ((turns to get)) ((POSITION 4))
((later)) There you are ...

(122) *Sinclair, 1976: 54*
S: Do you have any glue? ((POSITION 1))
H: Yes. What kind do you want dear? I've got um, I got a jar or-

S: Do you have some tubes? ((another POSITION 1))
H: The tubes? Ah you're lucky, aren't you actually. That's twenty five ((POSITION 2))
S: Oh I'll take that ((POSITION 3))

In the first example the full specification in position 1 is built to take a position 4 turn in second turn, whereas in the second the lack of specification for kind and and quantity of goods does not make such a second turn possible. Now there are many other properties of position 1 turns that are designed to get position 4 responses. For example, there are the range of markers of **interactional pessimism** (Brown & Levinson, 1979: 320) which occur in position 1 turns like the following:

(123) *145B*
C: You don't have his number I don't suppose

(124) *151*
C: I wonder whether I could possibly have a copy of last year's tax return

In cataloguing such features of position 1 turns that obtain position 4 responses we would soon find ourselves in the business of noting all the features of 'indirect requests' – including that same troublesome pre-verbal *please* that causes such difficulties in theories of indirect speech acts.

We can now say that so-called indirect speech acts are position 1 turns – pre-requests – formulated so as to expect position 4 responses in second turn. Questions about whether they have 'literal' or 'indirect' (or both) forces or meanings simply do not, on this view, arise. Such position 1 turns mean whatever they mean; that they can be formulated so as to project certain conversational trajectories is something properly explored in the sequential analysis of successive turns.

Let us just review the ingredients of such an analysis: (i) we make a distinction between *position* and *turn*, which will allow us to claim that indirect speech acts are position 1 turns that get position 4 responses in second turn; (ii) we note that preference organization, in seeking to avoid request refusals, motivates the standard four-position pre-request sequence; (iii) we show that there's a motivation provided for the precise content of position 1 turns in such sequences – namely, in order to avoid request rejections, the material used to check whether a request is likely to succeed is drawn from the grounds

that would usually be used to reject the request (i.e. position 1 material is drawn from the dispreferred position 4 material), and therefore questions of ability and the existence of goods routinely crop up in pre-requests; (iv) we can find in preference organization a systematic preference for the avoidance of some sequences altogether, and this provides a motivation for the collapse of the four-position sequence into the two-position sequence consisting of a position 1 turn followed by a position 4 turn; (v), given (iv), and the general tendency for pre-s to be so formulated as to prefigure what is to come up next, we can expect position 1 turns to be expressly formulated to get position 4 turns in second turn – and hence for pre-requests of this sort to contain special markers (including *would, could, not, please*, etc.).

Careful comparison of this conversationally based account with the standard accounts of indirect speech acts will show that it renders many of the most problematic aspects of the indirect speech acts issue quite illusory. In actual fact the issue is not of great interest to conversation analysis, and we have covered it in detail only because it illustrates just one of a number of different ways in which CA insights can unravel linguistic problems.

6.5 Conclusions

6.5.1 *Conversation analysis and linguistics*

In this Chapter we have argued that conversation analysis has made important contributions to the understanding of utterance meaning, by showing how a large proportion of the situated significance of utterances can be traced to their surrounding sequential environments. Just as the problems of indirect speech acts can be re-analysed in CA terms, so many of the other central concepts in pragmatic theory may be amenable to CA (or other discourse-analytic) treatments. Grice's maxims are, of course, prime targets in this regard, but so are problems with presupposition (see Sacks, 1968, May 29; Prince, 1978a, 1978b), and even problems with the analysis of deixis (Watson, 1975; Sacks, 1976; Goodwin, 1977).

Less clearly, perhaps, CA has much to contribute too to the study of linguistic form: to prosodics, phonology, syntax and the description of the lexicon. It is worth spelling out here some of the observable relations between conversational and linguistic structure. Let us take some of the conversational organizations we have reviewed and ask

how each may be a functional source of, or explanation for, certain linguistic structures and expressions.

The turn-taking system, for example, directly motivates the prosodic and syntactic signalling of turn completion and incompletion. The signalling of incompletion then provides a motivation for syntactic subordination, and predicts a preference for left-branching structures, or traces to the left of right-branching structures that may be forthcoming. Thus the English relative clause in *I am reading the book which I gave you* is more vulnerable to overlap than the equivalent clause in the comparable Dravidian or Japanese sentence, glossable as 'The I to-you given book I am reading'; but the vulnerability is reduced by the location of the WH-word at the head of the clause. On the other hand, the possibility that after completion the speaker may be able to continue, so that a turn may extend over more than one turn-constructional unit, makes it desirable that syntactic structures allow open-ended conjunction or addition to the right. In addition to such very general functional pressures, the turn-taking system provides more specific demands on linguistic structure: for example, the provision in the rules for the selection of next speaker directly motivates tag-formation. There are also many particles in languages that seem to have a function only explicable in relation to the turn-taking system, whether they are floor-holders (like *uh* in English), floor-returners (like *hm* in English), or turn-enders (like tag-particles in many languages).

The repair system also motivates many aspects of linguistic or utterance structure (Schegloff, 1979b). Apart from the markers of self-repair (glottal stops, *I mean*, etc.) and simple understanding or hearing checks (*Pardon?*, etc.), there are special syntactic features of NTRIs or echo-questions (cf. *John went to the what?*). There are also some interesting interactions between segmental phonology and repair (Jefferson, 1974).

Adjacency pair organization likewise motivates aspects of linguistic structure. Indeed a general explanation for the cross-linguistic prevalence of the three basic sentence-types (declarative, interrogative and imperative) may lie in the basic distinction between, respectively, utterances that are not first-pair parts, utterances that are first parts to other utterances, and utterances that are first parts to actions. Adjacency pair organization also motivates other ways of typing first parts of pairs as requiring specific kinds of second part (e.g. *Yes/No*

vs. WH-questions; or the conventional sets of summons formats like *Hey!*, *Excuse me*, etc.). While preference organization, working across adjacency pairs, motivates the conventional announcers of dispreferred responses like *Well* and *Actually*.

Sequences of various sorts also have linguistic implications. As we noted for pre-sequences, it is possible to make a distinction between position and turn precisely because positions in a sequence are linguistically marked: thus pre-requests have those syntactic features (pre-verbal *please*, the polite past as in *I was wondering* ..., the *Could you* ... forms, etc.) previously associated with indirect illocutionary force. We have also noted the ways in which given vs. new information is packaged in the structure of pre-announcement sequences.

Topical organization is also an area that has direct linguistic implications, although it should not be thought that there is any direct connection between what has been talked of in linguistics under the rubric of **topic** or **theme** and the notion of discourse topic (see Keenan & Schieffelin, 1976). What is clear, however, is that certain syntactically marked constructions like left-dislocations (as in *John, I like him*) are used in attempts to control the flow of topic in the conversational sense (see e.g. Ochs & Duranti, 1979). Further, phrases like *By the way*, and interjections like *Hey*, mark introductions of new topics, while utterance-initial *Anyway* may mark return to prior topic (cf. Owen, 1982). Much work needs to be done here to clarify linguistic concepts of topic and their relation to discourse or conversational topic; given such clarification, it is possible that many syntactic constructions can be shown to be directly motivated by requirements of topical organization in conversation.

Finally, aspects of overall conversational organization also interact with linguistic structure, most noticeably in the linguistic formulae typical of openings and closings (Irvine, 1974; Ferguson, 1976), but also in the use of particles like *Well* and *Okay* in pre-closings, and the like.

In the present state of our knowledge, remarks of this sort can only be suggestive of the many, largely unexplored, ways in which conversational organization interacts with sentence and utterance structure.

6.5.2 *Some remaining questions*

It is perhaps no accident that the analyses produced by CA so far have a striking (if superficial) resemblance to the structuralist theories of linguistics that predominated before the 1960s. Both kinds of approach are concerned with corpora of recorded materials; both have as a central methodological tool the use of a 'slot and filler' heuristic – i.e. the investigation of how sequential (or *syntagmatic*) considerations restrict the class of items that may expectably follow, and of how items in that class contrast with one another (or stand in *paradigmatic* relations). The parallel is clearest perhaps in CA discussions of **formulation**, where the central issue is why one particular description is chosen from a set of paradigmatic alternatives (see Schegloff, 1972b; also Sacks, 1972 on **membership categorizations**). Just as structuralist analyses of linguistic structure have been shown to be theoretically inadequate as models of human competence, so in the long run CA analyses may perhaps be found deficient as rather simple reconstructions of the no doubt immensely complicated cognitive processes involved in conducting conversations. But meanwhile, at least, no other kind of investigation of conversational organization has yielded such a rich harvest of insights.

A possible puzzle that is raised here is whether, despite the remarks in section 6.1 above, CA is not after all a 'syntactic' model of conversation. For, it is largely concerned with constraints on sequential possibilities. However the differences are in fact substantial. First, some of the rules formulated in CA, e.g. the turn-taking rules described in section 6.2.1.1, are as much **regulative** as **constitutive**, to employ the distinction used by Searle (1969) to distinguish speech act rules (themselves constitutive of each kind of speech act) from, for example, traffic rules (merely regulative of independently existing traffic flow). Secondly, the CA rules describe unmarked expectations rather than the set of possible well-formed sequences or conversations; in this way such rules are much more like Grice's maxims than like linguistic rules. Consider, for example, the rule that given a first part of an adjacency pair, a second part should follow; as the notion of conditional relevance (which we introduced in 6.2.1) makes clear, failure to provide a second is itself a communicational resource that can be used to contribute effectively to conversation. So the adoption of the 'slot and filler' heuristic should not be construed as carrying with it the special sense of rule that is found in linguistics.

Another puzzle that arises is whether the unreflective use of categories like *request*, *invitation*, *greeting* and the like, does not embody an implicit theory of speech acts. Might it be that, while speech act theory has been attempting to provide an internal characterization of the function of turns, CA has been concerned with inter-turn relationships, so that some simple synthesis is possible? Workers in CA would reject such suggestions. First, they would point out that the terms *request*, *invitation*, *greeting*, and so on, are not the inventions of speech act theory, but rather part of a rich (if largely unexplored) natural language metalanguage (see e.g. Allwood, 1976; Verschueren, 1980). It does not follow from the existence of such terms either that there is a close connection between 'folk' metalanguage and the categories actually utilized in speech production, or that such categories are properly explicated by providing sets of necessary and sufficient conditions for speech act category membership (contrast: Searle, 1969; Levinson, 1979a). In any case, if pressed, CA workers would claim that intuitive use of categories like *request* should in fact be backed up by at least (i) a full sequential explication in terms of the range of expectable responses (like refusals, deferrals, compliances, etc.), (ii) an account of the way that requests are typically formulated in order to obtain the desired responses (see section 6.4.3 above). Secondly, it is incorrect to view CA as primarily concerned with inter-turn relationships: the discussion of examples like (59) or (102) above, as well as the re-analysis of indirect speech acts, should make clear that CA is specifically interested in the relationship between intra-turn structure and inter-turn organization or sequence. It is not clear, therefore, that CA has to yield any quarter at all to speech act theory (see Turner, 1974b).

This raises a final and central issue. To what extent are aspects of conversational organization universal? Or, to what extent are the features of that organization reviewed here restricted to English (or even a sub-variety thereof)? The issue is central for a number of reasons: if basic aspects of conversational organization are universal then: (a) linguists may be able to explain significant linguistic universals by pointing to universal functional pressures exerted by basic patterns of language use; (b) general patterns of child language acquisition may be explained by reference to a single basic learning situation, which is conversational; (c) pedagogical programmes in second language acquisition can take certain basic pragmatic parameters for granted; (d) there are distinct limits to the kind of social variation in the use of language that has been explored in the ethnography of speaking. Further, such universals would throw light on a basic facet of human nature – perhaps humans as a species are as much characterized by conversational activity as they are by differing cultures, complex social systems and tool-making.

At the present we simply do not know to what extent conversational organization is universal – very little comparative work has been done on this level for languages other than the familiar European ones (but see e.g. Moerman, 1977). But it seems safe to say that, of the features reviewed in this Chapter, those that have been described as *local management systems* – e.g. turn-taking, adjacency pair organization, repair systems – have a universal basis, even if the descriptions in this Chapter are culturally skewed in certain ways. Overall structural units – like the notion of a conversation – are much more likely to be culturally variable; indeed this is a significant theme in the ethnography of speaking (see e.g. Bauman & Sherzer, 1974). Intermediate organizations, like preference organization and pre-sequences, probably lie somewhere in between: such things probably exist in all cultures, although the kinds of actions they organize may be quite different (e.g. there is a difference even between American English and British English in the preferred response to a compliment). But these are speculations. Disentangling cultural particulars from universal tendencies can be a difficult job, and given the importance of the issue, this is likely to be a preoccupation of comparative pragmatics and sociolinguistics in the years to come.

Appendix: transcription conventions

The conventions used in this Chapter in all examples from cited sources (except those from Merritt, 1976, Sinclair, 1976 and Labov & Fanshel, 1977) are mostly those employed in Schenkein, 1978: xi–xvi and developed by Jefferson and others. The most important are:

//	point at which the current utterance is overlapped by that transcribed below
*	asterisks indicate the alignment of the points where overlap ceases
(0.0)	pauses or gaps in what is very approximately tenths of seconds

	(closer measurements often being irrelevant because the significance of pauses is linked to some sense of 'the beat' of any particular conversation – see Goodwin, 1981: 114)
(.)	micropause – potentially significant but very short pause, comparable perhaps to an average syllable duration or somewhere below 0.2 seconds' duration
CAPS	relatively high amplitude, or, in double parentheses, analytical labels
italics	syllables stressed by amplitude, pitch and duration
::	lengthened syllables
-	glottal-stop self-editing marker
= =	'latched' utterances, with no gap
?	not a punctuation mark, but a rising intonation contour
.	used to indicate falling intonation contour
,	used to indicate maintained ('continuing') intonation contour
(( ))	used to specify "some phenomenon that the transcriber does not want to wrestle with" or some non-vocal action, etc.
( )	uncertain passages of transcript
→	draws attention to location of phenomenon of direct interest to discussion
hh	indicates an audible out-breath, .hh an in-breath

7
Conclusions

> In the history of human inquiry, philosophy has the place of the central sun, seminal and tumultuous: from time to time it throws off some portion of itself to take station as a science, a planet, cool and well regulated, progressing steadily towards a distant final state ... Is it not possible that the next century may see the birth, through the joint labours of philosophers, grammarians, and numerous other students of language, of a true and comprehensive *science of language*? Then we shall have rid ourselves of one more part of philosophy (there will still be plenty left) in the only way we ever get rid of philosophy, by kicking it upstairs. (Austin, 1956: 131–2)

7.0 Introduction

In these conclusions we shall try to tie together some of the loose strands of thought that have run through this book, by considering the relation between pragmatics and other disciplines. One discipline will stand noticeably absent: philosophy, the 'prodigal provider', cannot easily re-absorb the empirical studies that it has spawned (but cf. Atlas, 1979). The general tenor of this book has been the description of how, from original, mostly philosophical concepts, a series of empirical modes of investigation have developed, which jointly form the climate of the Anglo-American tradition in pragmatics. As the quotation indicates, Austin foresaw, and indeed hoped for, just this development of a field that he, perhaps more than any other single individual, did most to promote.

In the sections below, we shall first consider the inter-relations between pragmatics and the other 'core' components of linguistic theory, then between pragmatics and the 'hyphenated' linguistic disciplines (sociolinguistics and psycholinguistics in particular), and finally between pragmatics and less closely related fields.