DISCOURSE PROCESSES 4, 93–116 (1981)

Some Pre-Observations on the Modelling of Dialogue

STEPHEN C. LEVINSON
Department of Linguistics
University of Cambridge

IS THERE A SYNTAX OF DIALOGUE?

What kind of a thing is dialogue, or more generally, conversation? We are led to ask what necessary and distinctive properties it exhibits, and specifically, whether there could be a *syntax* of conversational structure. If so, an important initial gambit would be to do a Chomsky on conversational structure, where by 'do a Chomsky' I mean ascertain where on the hierarchy of formal languages the supposed syntax of conversation is located.

Recollect that Chomsky (1963) established the familiar hierarchy in (1), with a correspondence between levels of automata and rewriting grammars and the kinds of formal language they can generate:

(1)

Type	*Automata*	*Grammars*
0	Turing Machine	Unrestricted rewriting
1	Linear-bounded automata	Context-sensitive phrase structure
2	Push-down automata	Context-free phrase structure
3	Finite automata	Finite-state

where the lowest numbered type is the most general mechanism (Turing machine being generally understood to be capable of any computable task), and the highest numbered, the most specific and restricted type. The hierarchy is an implicational scale in that if a language can be generated by Type 3, then it can be generated by Type 2, and so on, but not necessarily the reverse (see Bach, 1974, pp. 179–205; Hunt, 1975, pp. 20–43 for elementary expositions). Chomsky and others have shown that grammars for natural languages could not possibly be generated by Type 3 mechanisms, and probably require mechanisms even more general than Type 1.

If there is a syntax of conversation, where on this hierarchy would it lie? The answer would be of some importance: it would precisely delimit the class of appropriate models of conversation, and would thus be the first substantial step towards a formal treatment. In order to ask the question we must first locate or discover the syntactic units of conversation, or, more technically, we must ask: what is the alphabet over which the formal language is defined?

That is, what is the analogue of the morphemes (the terminal alphabet) and the grammatical categories (the non-terminal or auxiliary alphabet) of linguistic analysis on the plane of conversational structure? The terminal alphabet, seemingly at least, poses few problems: it consists for conversation of sentences, sentence fragments, particles that indicate channel state, and perhaps significant pauses and various non-linguistic vocalizations like laughter.

But what about the non-terminal alphabet? What parallels the grammatical categories of linguistic theory? A plausible answer, and indeed *the* only plausible answer that I can think of, is that the auxiliary alphabet of conversational structure is the set of *speech act categories*. Why? Consider the regularities that are expressed by saying that, for example, answers (inter alia) follow questions, offers are met with acceptances or rejections, bets require declinings or bet-uptakes, greetings should be followed by greetings, apologies by acceptances or rejections, requests for permission by grantings or refusals, and so on. Clearly, we could say, these are concatenations of elements that are restricted in specific ways that should be describable in terms of a formal language, and the elements are speech acts (or conversational *moves* in another terminology; see Goffman, 1976; Owen, 1978).

If we make the equation of non-terminal vocabulary in our theory of conversation with speech act categories (refined, perhaps, well beyond those currently available), various interesting possibilities open up. We may note, following Schegloff, that in addition to question-answer sequences, for example, one obtains question-question-answer-answer sequences, or even further levels of embedding as in (2):

(2) A. Excuse me, can you tell me where Woolworths is? (Q_1)
B. Ah, do you know Marks and Spencers? (Q_2)
A. Is that in Sydney Street (Q_3)
B. Yes (A_3)
A. Ah yes, then I know (A_2)
B. Well, Woolworths is just down from there (A_1)

where A_1 is the answer to Q_1, A_2, to Q_2, and A_3, to Q_3.

Now given that, in principle, there seems no limit to the possible number of embedded sequences except that imposed by memory limitations, we can express the regularities here most simply as the rewriting rule in (3):

(3) $Q_s \rightarrow Q\ (Q_s)\ A$

here Q_s stands for question sequence, a node that always dominates Q and A. If indeed we need this rule, then we know immediately by two well known

results that the syntax of conversation cannot be generated by the most specialized devices in the hierarchy in (1). For the rule is of the sort $A \rightarrow xAy$, that is of the essentially self-embedding kind, and generates formulae on the pattern $a^n b^n$, thus requiring an infinity of substitution classes, and both these properties cannot be captured by a finite state device (see Bach, 1974, pp. 181–9). Thus we are able to state theorem 1:

(4) Theorem 1: The syntax of dialogue is essentially self-embedding and requires an infinity of substitution classes; it cannot therefore be generated by a finite state device and requires a device at least of Type 2, and possibly even more general. And this would be, for such a theory of conversational structure, a significant result, constraining the relevant class of theories about that structure.

I hope to have painted a promising and appealing picture for such an approach. But I believe that this approach is in fact entirely inappropriate to the subject matter. Indeed I would like to state a counter hypothesis:

(5) Dialogue has no syntax; speech acts are not the relevant categories over which to define the regularities of conversation; there exists no other finite alphabet over which to define the regularities; and there are no concatenation rules of general application even if there were such an alphabet.

To see why I maintain the counter-hypothesis, it will be necessary to formulate the syntactic approach to models of dialogue in a more specific way, and it is to this that I now turn.

2. SPEECH ACT MODELS OF DIALOGUE

Syntactic models of conversational structure might be developed in various directions, yet they must all, it seems to me, share a basic set of properties that clearly define a specific class of possible modules. These properties, which we can treat as distinctive of the class, are those in (6):

(6) (i) There are unit acts - *speech acts* - that are performed in speaking.
(ii) Utterances are segmentable into unit parts - *utterance units* - each of which corresponds to a unit act.
(iii) There is a specifiable conventional *procedure that will assign unit acts to utterance units;* in other words, there is a function whose domain is the set of possible utterance units and whose range is the set of possible speech acts.
(iv) Sequences of acts are regulated by a set of conventional *sequencing rules stated over speech act types.*

The kernel idea behind such models is simple and attractive: since sequential constraints are clearly not easily stated on the form or meaning of what is said, we should *translate* utterances into the acts that they perform, because on this deeper level rules of sequencing will be straight-forwardly describable.

We can develop a hypothetical set of sequencing rules of the sort in (iv) either in terms of which individual kinds of speech acts can be concatenated with each other, or in terms of larger categories of speech act types, sub-categorized where necessary. Taking the latter line, we could attempt to employ the typology of speech act types suggested by Searle (1976), as in (7):

(7) Type I: Representatives (e.g. assertions)
II: Directives (e.g. orders, questions)
III: Commissives (e.g. promises, threats)
IV: Expressives (e.g. apologies, approvals)
V: Declarations (e.g. declaring war, making man and wife)

We might then suggest the following rules as attempts to capture certain sequential potentialities:

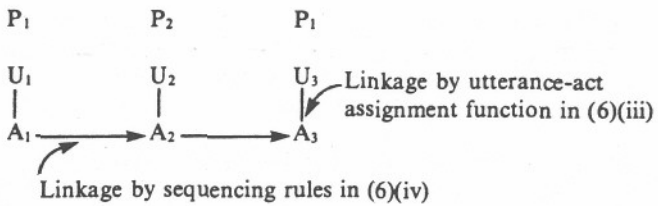
(8)

Type	is followable by	Type
I	→	I
		IV
II	→	I
		III
III	→	IV
IV	→	IV
V	→	IV

These rules state, *inter alia,* that assertions can be met with other assertions or expressives like 'How amazing!'; orders can be met with assertions ('I can't do it') or commissives ('I'll do it as soon as I can'); promises should be received with, for example, thanks; expressives can be met with expressives ('Good Lord,' 'Wow'); and declarations with expressives ('Behold your anointed king,' 'Hoorah!'). Of course these are simplistic, but allowing for sub-categorization where necessary, they might (it could be claimed) be rounded out in a way that captured both the details of individual speech act pairs (like question-answer) and the generalities statable at a more abstract level.

We are now in a position to construct a first-order approximation to a model of dialogue; the simplest plausible model along these lines can be represented as in (9), where P_1 and P_2 are the participants or interlocutors, U_n is an utterance unit, A_n the unit achieved by that utterance unit:

(9)



Although as a model (9) can only be described as crude, notice that closer approximations to reality can be achieved without essentially changing its nature. For example, we could define the notion of a *turn* at speaking, and allow each turn to contain more than one utterance unit, achieving more than one unit act, as in the response in (10):

(10) A: Well, how are you?
B: Oh very well thanks. And you?

In addition we could allow for embedded sequences of the sort we saw could occur with questions as in example (2) above. In fact, all sorts of additional frills could plausibly be added, while retaining as a central feature a translation procedure from utterances to acts (where the regularities of sequencing are held to lie) and indeed all the basic properties in (6) above.

Now I have set up a model of conversation that I wish to attack. The model is only a 'strawman' in the sense that it makes explicit what has generally been held implicitly, and lacks some of the frills that have been added in some accounts. Specific accounts that seem to me to fall within the general class of such models are numerous, and include Carbonell, 1978; Clarke, 1978; Goffman, 1976; Gazdar, 1978; Labov, 1972; Labov & Fanshel, 1978; Levinson, 1973; Owen, 1978; Perrault, Allen & Cohen, 1978; Power, 1977; Sinclair & Coulthard, 1975; and Woods (personal communication on work in progress). The list includes work by sociologists, linguists, social psychologists and workers in artificial intelligence – indeed, in general, work on machine question-answering systems and the like will fall into this class. Some remarks by ethnomethodologists, particularly on adjacency pair organization, seem to lie in the same sort of direction, but on the whole their approach is too flexible to fall within this framework. In general, then, this class of models represents one large vein of work on conversational structure, and as work increases on analysis by synthesis in this domain this class of models threatens to dominate the work in this area. It is therefore important to point out the sorts of difficulties that arise within this class, and it is to this that we may now turn.

3.0 PROBLEMS FOR SPEECH ACT MODELS OF DIALOGUE

If we examine the four basic assumptions that underlie speech act models of dialogue, namely those in (6) above, we will find that it is far from clear that actual conversation exhibits these properties. Let us take them one by one.

3.1. *The Existence and Identifiability of Unit Acts Corresponding to Specific Utterance Units.*

Two problems in particular arise with the isolation of speech acts in discourse: the first is that some unitary utterance units seem to involve more than one speech act in a number of different ways, and the second concerns the viability of a distinction between illocutionary and perlocutionary force in a theory of speech acts relevant to the analysis of conversation.

3.1.1. More Than One Speech Act at a Time Some utterances seem to achieve two or more speech acts simultaneously without one act being attributable to one utterance-part and another to another part. Simple examples are indirect speech acts like (11):

(11) Would you like another drink?

which is both a question and an offer (as shown by the possible response 'Yes I would, thank you' where the *yes* answers the question and the *thank you* acknowledges the offer).

Less obviously, perhaps, the first utterance in (12) is not just a question:

(12) A: What are you doing tonight?
B: Nothing, why?
C: I was thinking of going to a movie, wanna come?

It is also, as Sacks, Schegloff and associates have pointed out, a *pre-offer*. If we were to characterize A's first utterance as just a question, we would have to consider B's response ('nothing') palpably false; it isn't, of course, under the interpretation that it is a response to a question that is a pre-offer, and that it means essentially 'nothing that would make the offer of an evenings entertainment irrelevant'. Furthermore, A's interpretation of B's 'why?' relies on A's assumption that B saw A's first utterance as a pre-offer.

The question-pre-offer utterance in (12) could be an initial utterance after greetings, but there are other cases of apparently simultaneous speech acts where the dual interpretation of is due more strongly to location of the utterance in a sequence. Turner (1975) (discussed by Gazdar 1978) discusses for example the "double-duty utterance" of B's in example (13):

(13) A: What do you need
B: Four at thirty-three
A: Okay, will do

where the second utterance is both an answer to a question and a request (hence A's *Okay*). In other cases the double-duty nature of an utterance is only locatable with reference to aspects of the non-linguistic context; Owen (1978, p. 10) provides the following example:

(14) A: You've sold out of Listeners have you?
B: *Yes I'm terribly sorry dear, we have.* Is there something special in it this week?
A: *It's the Reith lectures.*

Here, apparently, B sees A's first utterance as both a question (hence the answer 'yes') and a potential complaint, or at least the raising of a matter worthy of apology (hence the 'I'm terribly sorry dear'). But the complaint reading is only licensed by the fact that B is a news agent who feels responsible for the reasonable provision of certain commodities, and by the fact that the day of enquiry is the day of publication of the Listener, a weekly publication, one of those commodities.

Another rather subtler kind of double-duty utterance is described by Schegloff (1976), where the dual interpretation arises from detailed structural properties of the kind of talk involved. In (15) B is describing to A how he (B) had been having an argument with his teacher (T):

(15) B_1: *He* (=T) says, governments, an' you know he keeps--he talks about governments, they shou--the thing that they should do is what's right or wrong.
A_1: For *whom*
B_2: Well he says-// he -
A_2: By what *stan*dard
B_3: That's what - that's exactly what I mean.

The interesting thing is the function of A_1: *For whom*. That it is a question is not reasonably in doubt; but that that is all it is is not so clear. B in B_2 provides a simple answer, whereupon A interrupts with a reformulation of A_1 that is appreciated in a different way by B. For what A seems to have intended is that A_1 should be considered an utterance in the *reported* conversation, in which case it would have been said by B to his teacher as a challenge: 'what is right or wrong *for whom?*' (see Schegloff, op.cit., for a closely argued justification of this analysis). A_1 is thus both a question and a challenge to provide a reasonable answer (in this way close to a rhetorical question), and moreover a

question spoken by A as if he were B to B as if he were T – it is an utterance in a conversation other than the present one, and thereby a contribution to the present one.

These examples raise a number of difficulties for the speech act modeller of dialogue. First, if there are pre-offers, may there not be pre-pre-offers, as well as pre-bets, pre-promises and so on? This threatens to expand any set of speech act categories indefinitely. Secondly, if there are double-duty utterances, may there not be triple-duty ones, and so on? And since there are quite different ways in which utterances can achieve double-duty status, there is no reason why these cannot be compounded to yield an indefinitely large number of act-interpretations.

Given that an utterance in context may achieve more than one 'act' simultaneously, should we not change the terminology? For an action is a composite formed of a chunk of behavior and a set of intentions. The multiplicity of interpretations, simultaneously assigned, is really an assignment of more than intention to the utterance, a chunk of behavior.

The change of terminology is apposite because it reminds us that intentions are not units, in the way that behavior chunks can be; intentions can have hierarchical organization (one being the source of another) and linear relations (one being a precondition to another), and for an interactant another interactant's intentions are only likely to be determinable up to a certain point. These facts surface in the next class of problems facing speech act unitization.

3.1.2. Chains of Perlocutionary Intent Ultimately more problematic, I think, than n-duty utterances, though related, is the fact that on occasions speakers seem to have great chains of motives that issue forth in an utterance. Take this case: I'm not enjoying the party that I have gone to with my companion Mildred, so I wish to leave, so I wish to suggest to Mildred that we both go, so I say to her, 'It's getting late, Mildred'. To which Mildred may felicitously reply with any of the following utterances (inter alia):

(16) A: It's only 11:15, darling.
B: Let's stay 'til Tony goes
C: Do you want to go?
D: Aren't you enjoying yourself dear?

where only the first one seems to respond directly to what is said. The others seem to respond to higher levels in the hierarchical chain of motives that led me to say 'It's getting late, Mildred'. Thus (16)b. is addressed to my desire that we both go, (16)c. to my desire that I go, and (16)d. to my ultimate motive in saying what I did.

There's no difficulty, I think, in showing that this is a very general phenomenon. Looking back at (12) we can see that the first utterance is a

question made with the intent of establishing whether the conditions for an offer obtain (hence the characterization as a 'pre-offer'), and indeed it has been suggested that most indirect speech acts are best treated as pre-sequences of this sort (see Heringer, 1977). Similarly for (13), (14) and (15), the multiplicity of function attributed to utterances comes about because a speaker has a hierarchy of intents behind an utterance: in (15), for example, A intended that, in asking a question as if he was B intending to challenge B's teacher, he would show his agreement with B's point of view.

Now these facts do definitely confront the speech act modeller of dialogue with substantial problems. First, we have shown that responses to utterances can be based on quite remote perlocutionary intents; yet speech act theory is founded on a basic distinction between illocutionary and perlocutionary acts (or intents, as I would prefer) and has nothing interesting to say about the latter. Secondly, the speech act theorist is not in a position to simply extend his theory to cover perlocutions; due to the probable infinity of possible perlocutions, his model does not and cannot have anything to say about them. Presumably, anything that anyone can possibly intend is, in the right circumstances, a possible perlocution. Thus it seems likely that there will also be an infinity of perlocutionary *types,* as well as an infinity of perlocutionary acts, so the original idea of reducing the varied and disorderly surface utterances to a limited set of speech act types loses most of its attraction.

3.2 *The Existence and Identifiability of Utterance Units Corresponding to Unit Acts.*

The second property that speech act models of dialogue have is that they presume the existence of *utterance units* (see (6) (ii) above), to be paired with act units, which are identifiable independently of the functions (the act units) they perform. This too proves to be a problematic assumption.

We have already seen that, as in (11) or (12) above, a single sentence can perform more than one act simultaneously. However, the sentence is not clearly the relevant unit to which to assign acts or intents. Some sentences perform more than one act because one sentence segment performs one act and another another. Thus (17) seems to be both an order and a threat, and (18) both a statement and a question (see Lakoff, 1974, for discussion):

(17) Shut up or I'll belt you
(18) Fillipino Lippis are delightful aren't they?

Further, just about any sentence part above a bound morpheme can operate as a full conversational contribution. Consider for example the response in (19):

(19) A: How do you want your coffee?
B: Light, with sugar, to go.

And note that, according to what went before, such 'sentence-fragments' (as Morgan, 1974, dubs them) can perform more than one act:

(20) A: *Who* did *what* to *whom?*!?
B: Stan, unmentionables, to Pam

In addition, what are we to say of an utterance like (21) (observed and reported to me by Andy Rogers):

(21) Bill, Mary; Mary, Bill

which was said by way of introduction. Here the relevant utterance units are either one, two or four, and it's hard to say which; we are also at a loss as to whether one, two or four acts took place. The coincidence of indeterminancy indicates clearly that we do not independently identify utterance-units and act-units.

This functional (or act-based) identification of the relevant utterance units to which acts are to be assigned is made especially clear by facts like attributable silence (see Sacks). For example, if a teacher reprimands a child with an utterance like (22), and the child remains mute, this can be 'heard' as an affirmative reply to the question:

(22) Johnny, did *you* smear Susie's face with the paint?

Or again, consider bidding in an auction: one finds exchanges like (23),

(23) A: Anymore bids? A delightful piece, going for a song. Going, going...
B: Twenty-five

but here almost any action on B's part, a wave of the hand, a scratch of the head, a grunt, could count as a successful bid, especially if B has been bidding hitherto. A large number of vocal and non-vocal units would function just as well as 'Twenty-five'.

In addition there are utterance units to which ordinary act-units should not be assigned, namely what have been called 'back-channel cues' (Duncan & Niederehe, 1974) – utterances consisting of items like *ah, oh, huhuh, mm, yeah.* But some of these, like *yeah, okay* and *I see* are homophonous with units that do count as turns, moves or speech acts (the ambiguity in function will remain sometimes even when intonation is taken into account). The distinction has to be made on functional grounds.

Utterance units then are very variable. They range from sets of sentences through sentence fragments to single lexemes, particles, or even silence or

non-verbal action. Which unit is the relevant unit for speech act assignment cannot be determined in advance, for utterance units seem to be identified on functional grounds. How then is speech act assignment to be achieved? It is to this problem that we now should turn.

3.3 The Existence of a Specifiable Function or Procedure That Will Assign Utterance Units to Unit Acts (Speech Acts).

We have discussed two of the basic properties of speech act models of dialogue, and we concluded that:

(a) speech acts are not unitary assignments to utterances from a well-defined act of speech act types, but rather an n-ary assignment of intents, where these are linked in specific ways, from an indefinitely large set of possible perlocutionary intents;
(b) utterance units are very varied in kind, must in the last resort be functionally defined (partly in terms of the speech acts to be assigned to them); and for any given functional role, there are likely to be an indefinitely large set of utterance units that will perform that function.

We now turn to the third basic property of speech act models of dialogue, namely the assumption that there is at least a specifiable function, and more ambitiously a specifiable procedure, that maps utterance units into speech act units.

Given conclusions (a) and (b), we see immediately that we are faced with problems. First, the domain of the function is defined in part in relation to the range – significant utterance-parts are not identifiable without reference to the speech acts that will be assigned to them. Secondly, and crucially, since neither the domain nor the range of the speech act assignment function are well defined, and indeed seem likely to resist attempts at precise specification, there is every reason to doubt the possibility of properly specifying such a function.

But let us suppose that, somehow, these problems can be overcome. We could for example, following Gazdar (1978) (in turn generalizing from work by Hamblin, 1971; Kamp, 1974; Stalnaker, 1978) treat speech acts as operations on, or modifications of, the knowledge and commitments held in a context – specifically, as functions taking contexts into (changed) contexts. (Note that this is a different, although probably more useful, notion than the one used hitherto in this paper, synonymous with illocutionary force.) And we could assume that further work on utterance units would find a way of characterizing them independently of speech act assignments. Even then, though, if the function was well defined, we would still be faced with overwhelming problems concerning the adequate specification of an actual

procedure that would adequately implement the abstract function – that is, provide us with a description of how assignment is actually achieved.

Perhaps the major problem here is the fundamental prevalence of *indirection* in human communication. The indirection no doubt comes about for various reasons: sometimes it is simply more efficient to say little and let one's interlocutor do some inference, sometimes people do not want to be held totally accountable for what they contextually imply, and most generally there are social rules and motives (like politeness) that require indirection (see Brown & Levinson, 1978).

The simplest answer to the assignment problem falls prey immediately to the fact of indirection. For this solution, often implicitly assumed in discussions in the philosophical and linguistic literature, would simply be to assign a unique illocutionary force to a sentence on the basis of its performative prefix (if overt) or its sentence type (often held to be a reflex of a covert performative prefix). Let us call this, following Gazdar (1978) rather loosely, the 'literal meaning hypothesis'.

It soon becomes clear, though, once attention is paid to the phenomenon of indirect speech acts (e.g., Gordon & Lakoff, 1971; Sadock, 1974; Searle, 1975) like (24), that there is no such simple relation between sentences and the speech acts they perform.

(24) Can you please pass the salt

Searle (1975) would like to maintain that sentences like (24) continue to perform questions, but happen in addition to serve as acts (possibly perlocutionary acts – Searle is not clear here) of requesting. However, sentences like (25) are counter-examples to such a claim:

(25) May I remind you that your account is overdue.

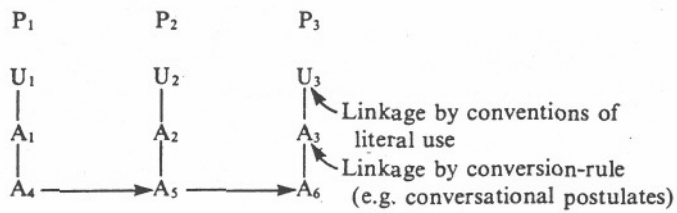
since it cannot possibly actually function as a request for permission to remind, since reminding is done in uttering the sentence without such permission being granted. Other counter-examples, and an illuminating discussion can be found in Gazdar (1978) who concludes that speech act assignment must be a mapping of sentence-context pairs, not sentences, into speech acts. (He is able to show too that Searle's attempt to equate the study of sentence-context pairs with the study of sentences that explicitly express their illocutionary force, via his 'principle of expressibility', is not a viable way out).

One possible solution to the indirect speech act problem might be to take the 'conversational postulate' view (Lakoff & Gordon, 1971) seriously, and consider that speech act assignment is based (a) on the literal force of what is said, unless, as in (24), this is ruled out (by context, and here by the particle

please which occurs only in requests), in which case (b) it is based on the application of a set of conventional force-conversion rules. Such rules would specify that, for example, *Can you X, Would you please X, I'd be most grateful if you'd please X* and so on would count as *I request that you do X* wherever the illocutionary forces based on their literal forms are contextually not viable. Let us call this the (speech act) conversion-rule hypothesis. The rules could be thought of as a set of functions that take an illocutionary force and a proposition and yield a different illocutionary force and a proposition and yield a different illocutionary force with the same (or part of the same) proposition. Proponents of such a view include Heringer (1972), Fraser (1975), Sinclair & Coulthard (1975), Labov (1972) and Labove & Fanshel (1978).

The speech act modeller of dialogue would then have to modify the model in (9) above to something more along the lines shown in (26):

(26)



The same sequencing rules would operate as before to give an account of sequential potentialities, but now on a deeper level, on acts more remotely linked to what has actually been said.

The problem with such a solution is that, while it might work for some cases, it is only at best partial. There are many kinds of indirect speech acts that are not based on any conventional conversion-rule, but rather on some complex inferential process. Consider, for example. (27):

(27) A: I could eat the whole of that cake
B: Thanks. It's quite easy to make actually.

Here B (correctly) interpreted A's remark as a compliment (as shown by the response 'thanks') on the cake that she had baked. But clearly there is no general rule of the sort that to say one could eat or do X counts as a compliment on X. Inferences of this kind depend on specific aspects of the context: (27) was said at a birthday party where the host had baked her own cake. There must of course be principles that underlie our understandings of such things, but if modelling conversation is to be a finite enterprise they had

better be a small but powerful set of general principles of inference to interlocutors' communicative intentions in specific contexts, rather than members of a huge and ad hoc set of conventional rules.

Consider too the cases of multiple-duty utterances in (12) through (15) above; the source of their speech act assignments has to be sought in the detailed facts about their sequential location in particular contexts between participant interlocutors. Thus, the conversion rule hypothesis fails for just the same reason as the literal meaning hypothesis fails: it attempts again to minimize the role of context in determining the acts or intents that are assigned to utterances.

We should point out in addition that responses can be tied (and often are) to the presuppositions or implicatures of what is said rather than to the literal content. Thus in (28), the first utterance only presupposes that B left her headlights on, although – as the response indicates – the conveying of this information is the main import of A's utterance in interrogative form:

(28) (A): Are you aware that you've left your headlights on?
B: Oh thank you so much.

Similarly in (29), the response is directed to the implicature, not the literal content of, the conditional sentence uttered by A:

(29) A: If John gets an A in chemistry I'll eat my hat
B: Come now, he's not so dumb.

These examples could be multiplied, but the point is that it makes little sense, as the main point of what is meant becomes more and more remote from what has actually been said, to talk of speech act assignment at all. In addition, these examples illustrate that the kinds of indirection available in conversation are legion; a review of some of these is available in Brown & Levinson (1978), where considerable attention is paid to the complexity of the resultant inferences and the ways in which they often rely on consideration of motives that the speaker may have for not saying what he actually intends to convey.

We have come far enough to present the speech act theorist with a fundamental dilemma. Either he retreats to his original position, in which illocutionary force is assigned on the basis of surface syntactic type, in which case assignment is a relatively simple affair but is largely irrelevant to how conversation proceeds; or he is faced with accounting for force and content that are only tenuously linked to what is actually said by mechanisms that are not simply a set of conventional rules, but rather a powerful set of little-understood inference principles.

3.4 The Assumption That Sequences of Utterances are Regulated by Conventional Sequencing Rules Stated Over Speech Act Types.

We come now to the final, and indeed the motivating, property of speech act models of dialogue. For the point of modelling a translation procedure from utterances to acts was to reduce the problems of sequencing in dialogue to a statement of regularities in sequences of acts.

3.4.1. Over-Emphasis on Adjacency Pairs. The initial attraction of the program probably stems from observation of adjacency-pair organization (e.g., Sacks & Schegloff, 1974, pp. 238–41), the way in which, for example, questions set up expectations for answers, greetings for greetings, offers for acceptances or declinings, and so on. But the bulk of conversation is not constructed from adjacency pairs, and it is easy to conceptually over-emphasize the constraints on trying between consecutive utterances. Compare, for example, what can follow an assertion: another assertion, a question, a bet, a promise, a back-channel cue, an offer and so on, perhaps indefinitely and without clear preference for one response rather than another. In this respect (as in many others), the hypothetical sequencing rules in (8) above are just plain wrong.

In any case, responses to first parts of adjacency pairs are a lot freer than the question-answer stereotype would suggest. Questions can be happily followed by questions, by partial answers, statements of ignorance, denials of the relevance of the question, details of the presuppositions of the question, and so on, as illustrated in (30):

(30) A: What does Joe do for a living?
B: i. Do you need to know?
ii. Oh this and that.
iii. I've no idea.
iv. What's that got to do with it?
v. He doesn't.

How limited is this list? To some extent, as I have shown elsewhere (Levinson, 1978), what a question is – and hence the nature of the set of relevant responses – is dependent on the peculiarities of the particular language game (social activity) it is embedded within. Suppose during a lecture I say (31):

(31) What would a reasonable model of dialogue look like?

I may intend no response, either from myself or from my audience, knowing for example that the answer is still an enigma. Or, in a classroom, truthful

answers to teacher's questions may be quite inadequate, either because some other child has already provided that answer or because the answer doesn't advance the teacher's line of argument or because of other special conditions pertinent to the particular language game involved (see Levinson, 1978, pp. E13–F6). The number of sets of adequate responses may thus be just as unlimited as the number of language games that employ devices recognizable as questions. Finally, let me just provide an example, taken from a shopping episode, of the kind of interactionally adequate response to a question that actually occurs but which the armchair analyst is unlikely to take into account:

(32) (A is a shopper in search of a pullover to go with a beige jacket; B is a shop assistant paired with A)
A: Yes um. Do you think that blue would stand a chance with um beige?
B: Well// // Try that one on (selecting a grey pullover).

where a suggestion in imperative form follows a question. In sum, I do not wish to deny that there are interesting facts about what may follow questions in discourse, but I have yet to see an interesting characterization of the constraints. Specifically, I doubt that the set of adequate responses can in general be characterized in speech act terms.

3.4.2 Preservation of Topic I assume that in attempting to give an account of sequencing constraints in conversation one is, at least in part, attempting to give an account of the preservation and coherence of topic in talk. Constraints on, say, second parts of adjacency pairs are partly topic constraints; an answer must be an answer to the relevant specific question posed prior to the answer.

The first point to note is that topic, and the related notion of relevance, is an elusive concept. Topical coherence is flexible: in conversational transcripts, and indeed in conversation, one can detect on-topic talk, topic-shift, and topic slide – that is, gradual movement across topics with some measure of coherency and connection (Goldberg, 1978, provides some nice examples). In addition, there are breaks in topic that are not perceived as rule-violations, as when new topics arise. Moreover, there is the phenomenon of *topical hierarchies:* consider, for example, the kind of licensed topic shift that can occur in work situations where chat must succumb to the communications demands of a priority task. At table, the same sort of thing can occur:

(33) A: Well, I had a lovely time in Rome
B: More wine? This is better than the last one and you'll never guess where it comes from.

In short, topic does not seem to exhibit the sort of inflexibility and precisely sanctionable nature that we associate with rule-bound activity.

A second, more fundamental, point to notice is that speech act models of dialogue have in and of themselves nothing whatsoever to say about topic (with the possible exception of question-answer topical coherencies). The reason of course is that speech act theory (as in Searle, 1969), in making the basic distinction between illocutionary force and propositional content, abandons further consideration of the latter (beyond whatever general constraints particular illocutionary forces put on propositional contents). So, to handle topical coherency, speech act models of dialogue would have to have an additional and independent theory of topical dependencies across propositional contents. We are provided with no clues about what such a theory might look like, but it would in any case be awkward, because topical coherency is not in fact independent of illocutionary force. For example, it is not the case that there is just an expectation that a bet will be met with a bet-uptake or declining: if Bill says "I bet you £ 5 I can get to London before you" and I say "You're on", then I've taken up the particular bet that was on offer.

Consideration of topic raises a very basic issue. In modelling dialogue we are attempting to give an account of what gives cohesion to conversations. Within speech act models of dialogue (at least of the sort outlined in (6)) this task is interpreted as giving an account of what is and what is not a well-formed sequence of utterances, by writing generative rules that will partition all possible such sequences into those two classes on the basis of abstract structural features. On this account, it should be possible to determine the acceptability of, or to generate, sequences without having available a full interpretation of them. In terms of a computational problem, the task is to construct a machine that, given (say) a large set of sentences, will generate acceptable dialogue, or alternatively, one that will recognize well-formed sequences essentially on the basis of relations across illocutionary forces.

An alternative definition of the task that I would prefer is to attempt to specify well-formedness of conversational sequences *relative* to interpretations of them. Sufficient motivation for this is provided by our exceptional ability to provide plausible interpretations for what at first look like bizarre sequences. As Grice (1975) pointed out, in examples like (34) what at first appears to be a radical nonsequitur can on reflection be seen to be implicating that, say, B is not free to talk about B's ills in the present company:

(34) A: What on earth is the matter with you?
B: What a lovely dress, where did you get it?

The problem, then, as I see it, is to attempt to specify what kind of knowledge above and beyond the 'literal meanings' of utterances (however that is thought of, for example as what is given by semantic representations

and their derivative inferences) is required to participate in and understand a conversation. If we want to think of this in terms of a computational task then we could think of it as a challenge to construct a machine that, having taken the text of a conversation as input, will output a set of interpretations of an intuitively correct sort, or alternatively as a challenge to construct two machines that will converse with one another while displaying appropriate interpretations of how the conversation is proceeding (in both cases, of course we have to make the outrageous assumption that the grammatical aspects of the problem have already been solved).

If this latter version of the problem is the one we really want to address, then 'syntactic' models of dialogue are built along the wrong lines, for the account of sequencing they can provide is based, at best, on a very partial interpretation—just the illocutionary component of what is said.

3.4.3. Responses Aimed at Prior Speaker's Higher-Level Goals. I have raised some general difficulties for rule-based approaches to sequencing constraints, but I now wish to consider a range of examples that seem to me to be knock-down counter-examples to such approaches, and in particular to property (iv) in (6) above. These examples are exchanges in dialogue where responses are aimed not at what has been said, but at the broader motive, or higher level goal, that is seen to lie behind what has been said.

A simple instance is (35):

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

where B's response is not an answer, and yet constitutes an eminently cooperative response on the understanding that the motive behind A's question is A's wanting to get in touch with John. Incidentally this kind of response to unstated higher level goals is closely related to what ethnomethodologists call 'formulation'. To illustrate consider (36):

(36) A: Where's John
B: He's at extension thirty-four sixty-two

Let us suppose (and in this case it is perhaps not quite clear) that B's response is a (truthful) answer to A's question; then we must admit that there is an indefinitely large set of alternative truthful answers to A's question, including, say: The northern hemisphere; England; Cambridge; the History Faculty Building; Roome 32; on the yellow chair; and so on. Which particular answer is the helpful response will depend on what the questioner's motive was in asking the question, this motive often being directly inferable from the contextual circumstances and in particular the kind of activity in which the

exchange is situated (for examples and an important discussion, see Schegloff, 1972).

Recent artificial intelligence work on dialogue has thrown up the same problem: Robinson & Hobbs (1978) discuss examples like (37) and conclude that these responses to higher-level goals form a basic category of responses in task-oriented dialogues (see also Grosz, 1977):

(37) A: What's the metric torque-wrench nipple-extractor look like?
B: It's on the bench in front of you.

Here again B's 'answer' is cooperative but only on the assumption that the reason for A's question is that he wishes to identify and find the wrench and that B reckons that a statement of its location will serve A's purpose better than a description of the instrument itself.

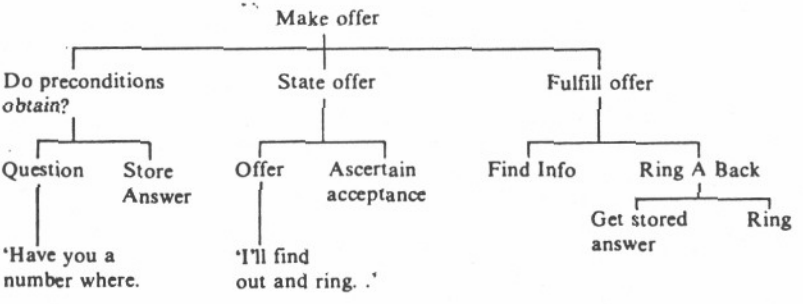
Take a final and slightly more involved example:

(38) A: Can you give me Mr. X's phone number?
B: Hmm. Have you a number where I can ring you back?
A: Thanks, but I'll be seeing him later anyway so it's alright.

Now note that this might equally well have gone a slightly different way, as in (39):

(39) A: Can you give me Mr. X's phone number?
B: Hmm. Have you a number where I can ring you back?
A: Yes, it's 60185
B: Good I'll find out and ring you back this afternoon
A: Thanks a lot

We can think of B's responses in (39) as being directed by a hierarchical plan organizing a set of sub-goals as in (40):



This structure of goals is implemented (up till 'Fulfill offer') in (39), but not fully in (38). But the interesting thing about the (naturally occurring) example in (39) is that A's response in the last line is not to the immediately preceding question by B, but rather to B's *unstated* higher-level goal ('make offer' in (40)) of offering to ring A back. We know this because A responds to B's question with 'thanks', *even though no offer ever took place.* This is not simply capturable in a speech acts model of sequencing, of a sort that say allowed embedded structures like (Request(question-answer)Offer to comply with request), because the offer in question never actually took place.

This in fact appears to be a very general phenomenon: given an utterance which is merely the first in a sequence predicted by a hierarchical structure of goals, one is free to respond to any of those higher level goals.

An important point about examples (33), (37) and (38) is that a sequencing-rules model would predict violations of rules in such instances, with consequent repair mechanisms, sanctions and so on. But in fact no such sense of violation is present: these are all eminently cooperative exchanges.

Finally, there are aspects of natural conversation that do seem to be based on conventional sequencing rules. Certain ritually prescribed sequences, for example, can be neatly described in a syntactic convention. An elegant study by Irvine (1974) describes the elaborate and variable greeting used by the Wolof of Senegal, using a phrase-structure grammar where the terminal vocabulary is a set of set-phrases, and optional and recursive devices capture the observable variations in greeting performances. But studies of such ritually prescribed sequences only serve to make clear how radically different such exchanges are from freely directed conversation.

3.5 Contextual Variation in Act Assignments and Sequencing

There is a final problem that undermines both the assumption ((iii) in (6)) that there is a precise procedure that assigns utterances to speech acts, and the assumption ((iv) in (6)) that there is a single set of conventional sequencing rules. The problem is that most properties of discourse, and especially these two, are context-sensitive. In general we may say that how an utterance will be taken, and in what way it will be responded to, are facts that are in part determined by the social activity within which the talk occurs. The whole phenomenon is discussed at length in Levinson (1978) and I merely wish to illustrate the nature of the problem here.

If I am in a vegetable shop being served by an assistant, and I point to a lettuce and say:

(41) That's a nice one

I am likely to find the lettuce removed from display, wrapped up and presented to me with an expectation of money in exchange. In other words, I will have (or be perceived to have) performed the act of selecting a particular lettuce for purchase. Contrast now the act performed if I say (41), peering through the window of a store I cannot afford to shop in, to an equally impecunious companion; we might say variously that it expressed a judgment or a wish (and if the companion was better off, a suggestion). In short, the actions that I perform by saying something are in part a function of the social situation that I find myself in. Note too that responses will be likewise adjusted; in the grocer's store (41) may be followed by 'Okay', while outside the expensive store the response may be 'I prefer that one', but not vice versa.

The sequential implications of the type of activity that talk occurs within are best explored by looking closely at rather special kinds of activities. Example (42) is taken from a judicial proceeding (The Watergate Hearings, New York Times 1973:577).

(42) Q_1: You saw all of the papers that were being reviewed, did you not?
A_1: Not all the working papers of the committee. I saw the recommendations that went to the President
Q_2: Did you read the recommendations that went to the President?
A_2: I am not sure I did or not. If I did it was not in any detail.

Now compare the following constructed exchange:

(43) Q_1: Did you see last week's Newsweek?
A_1: Part of it
Q_2: Did you read that part of it?
A_2: No

If (43) occurred in circumstances more cooperative than judicial proceedings, for example in a chat across the garden fence, we should find it bizarre. That is because, in such cooperative circumstances, *see* would implicate *read*, so Q_2 would be redundant, and A_2 an admission that A_1 was tantamount to a lie. So the only circumstances in which an exchange like (43) would be natural would be those where Grice's (1975) principle of cooperation was not fully in force, and thus the implicature in question was not generated.

If we now compare (43) to (42) we see that it has the same basic sequential structure, and we can conclude that in some social activities Grice's principle is indeed relaxed. The relevance to sequential considerations is obvious: Q_2 and A_2 would be a redundant and thus nonoccurring sequence in cooperative circumstances. (For a discussion of the implications see Levinson, 1978.)

The situational relativity of properties (iii) and (iv) of speech act models of dialogue make the prospects for such models look pretty grim: there would

have to be as many speech act assignment procedures and as many sets of sequencing rules as there are distinctive kinds of social activity in a culture (and that is not a static set; consider, for example, group therapy sessions).

4. CONCLUSION

My strategy has been to raise the possibility of treating the structure of dialogue in terms of some formal language, with all the obvious advantages that would thereby accrue. I then set up what I take to be the most plausible sub-class of such models, namely those based on a non-terminal vocabulary of speech act categories, and outlined the criterial properties of such a class. A great deal of recent work on conversation can be directly equated with some model in this class, as the references have indicated. Finally I have devoted considerable space to showing that the whole class of models is in principle incapable of modelling the actual properties of natural dialogue.

Speech act models of dialogue are thus not viable as general models of conversations, even though they may capture a number of observable regularities like those, for example, between adjacency pairs, or in ritualized sequences, like greetings. In limited domains, though, such models may have their utility; question-answering systems built along these lines would, I imagine, be quite successful and capable of interesting refinements. My objection to them though is more fundamental than that they are merely partial accounts; I do not believe that this is, fundamentally, how we participate in conversation at all.

Finally, I assume that in demolishing the possibility of this particular class of models I have also rendered the larger class of syntactic models of dialogue inappropriate. This assumption, obviously, only goes through on the grounds that speech act models are the only plausible members of the larger class, which I believe to be the case.

REFERENCES

Abbreviations:
CIS n = Papers from the nth regional meeting of the Chicago Linguistic Society
LS = Language in Society
L = Language
PM = Pragmatic Microfiche
IULC = Indiana Univ. Linguistics Club mimeo.

Atlas, J. & Levinson, S. The importance of practical reasoning in language usage. Part II. Report on research undertaken for the MSSB advanced workshop on formal pragmatics, Ann Arbor, 1973.

Bach, E. *Syntactic Theory*. New York, Holt Rinehart & Winston, 1974.

Bennett, J. *Linguistic Behaviour*. Cambridge, University Press, 1976.

Brown, P. & Levinson, S. Universals in language usage: politeness phenomena. *In E. Goody* (ed.), *Questions and Politeness*. Cambridge, University Press, 1978.

Carbonell, J. *Intentionality and* human conversations. In D. Waltz (ed.), 1978.

Chomsky, N. Formal properties of grammars. I. Luce, R., Bush, R. and Galanter, E. (eds.), *Handbook of Mathematical Psychology*, Volume II. New York and London, 1963.

Clarke, D. The syntax of action. (Lecture on work in progress delivered to Social Psychology seminar, Cambridge), 1978.

Cole, P. & Morgan, J. *Syntax and Semantics 3: Speech Acts*. New York, Academic Press, *1975*.

Duncan, S. & Niederehe, G. On signalling that it's your turn to speak. *Journal of Experimental Social Psychology*, *10*:234–47, *1974*.

Fraser, B. Hedged performatives. In Cole and Morgan, 1975.

Gazdar, G. *Implicature, Presupposition and Logical Form*. IULC, 1977.

Gazdar, G. Speech act assignment, University of Sussex, ms., 1978.

Giglioli, P. *Language and social context*. Penguin Books, Harmondsworth, 1972.

Goffman, E. *Frame Analysis*. Penguin Books, Harmondsworth, *1975*.

Goffman, E. Replies and Responses. LS 5: 257–313, 1976.

Goldberg, J. *On 'moves' in conversation*. Cambridge Univ., ms., 1978.

Gordon, D. & Lakoff, G. Conversational postulates. CLS 7, 63–84, 1971.

Grice, H. P. Meaning. Reprinted in Strawson, P. (ed.), *Philosophical Logic*, 1967, Oxford, University Press, 1957.

Grice, H. P. Utterer's meaning, sentence-meaning, and word-meaning. In Searle, J. (ed.), *The Philosophy of Language*. Oxford, University Press, *1971*.

Grice, H. P. Logic and conversation. In Cole and Morgan, 1975.

Grice, H. P. *Further notes on* logic and conversation. In P. Cole (ed.), *Syntax and Semantics 9: Pragmatics*. New York, Academic Press, 1978.

Grosz, B. The representation and use of focus in dialogue understanding. SRI technical note 151, Menlo Park, 1977.

Hamblin, C. L. The effect of when it's said. *Theoria* 36:249–263, 1970.

Hamblin, C. L. Mathematical models of dialogue. *Theoria 37:130–155, 1971*.

Heringer, J. *Some grammatical correlates of felicity conditions and presuppositions*. IULC, *1972*.

Heringer, J. Presequences and indirect speech acts. In E. O. Keenan and T. L. Bennett (eds.), *Discourse across time and space*. Southern California Occasional Papers in Linguistics, No. 5, Los Angeles, University of Southern California, 1977.

Hunt, E. *Artificial Intelligence*. New York, Academic Press, 1975.

Hymes, D. Towards ethnographies *of communication: the analysis of communicative events*. In P. Giglioli, 1972.

Isard, S. *Changing the* context. In E. L. Keenan, 1975.

Irvine, J. Strategies of status manipulation in the Wolof greeting. In Bauman, R. and Sherzer, J. (eds.), *Explorations in the Ethnography of Speaking*. Cambridge, University Press, 1974.

James, D. Some aspects of the syntax and semantics of interjections. CLS 8:162–172, 1972.

Jefferson, G. et al. Some notes on laughing together. PM 1.8:A2-D10, 1976.

Kamp, J. A. W. Free choice permission. *Aristotelian Society Proceedings* Supplementary volume 49:21–44, 1974.

Keenan, E. L. *The Formal Semantics of Natural Language*. Cambridge, University Press, 1975.

Lakoff, G. Syntactic amalgams. CLS 10:321–344, 1974.

Lakoff, G. Pragmatics in natural logic. In E. L. Keenan, 1975.

Lakoff, R. The logic of politeness; or minding your p's and q's. CLS 9:292–305, 1973.

Labov, W. Rules for ritual insults. In D. Sudnow (ed.), *Studies in social interaction*. New York, Free Press, *1972*.

Labov, W. & Fanshel, D. *Therapeutic Discourse*. New York, Academic Press, 1977.

Levinson, S. Models of discourse. University of California, Berkeley, ms., 1973.

Levinson, S. Activity types and language. PM 3.3 D1-G5, 1978.

Lewis, D. *Convention: a philosophical study*. Cambridge, Massachusetts, Harvard University Press, 1969.

Morgan, J. Sentence fragments and the notion 'sentence'. In B. B. Kachru et al. (eds.), *Issues in Linguistics*. Urbana, University of Illinois Press, 1973.

Morgan, J. Towards a rational model of discourse comprehension. In D. Waltz, 1978.

Sacks, H., Scheloff, E. & Jefferson, G. A Simplest Systematics for the Organization of Turn-Taking in Conversation. *Language* 50(4):696–735, 1974.

Sadock, J. *Toward a Linguistic Theory of Speech Acts*. New York, Academic Press, 1974.

Schank, R. & Nash-Webber, B. L. (eds.) *Theoretical Issues in Natural Language Processing*. Cambridge, Massachusetts, 1975.

Schegloff, E. On Some Questions and Ambiguities in Conversation. *Pragmatics Microfiche*. 2.2, D8-G12, 1976.

Searle, J. *Speech Acts*. Cambridge, Cambridge University Press, 1969.

Searle, J. Indirect Speech Acts. In Cole & Morgan (eds.), 1975.

Searle, J. The Classification of Illocutionary Acts. *Language in Society* 5(1), 1976.

Sinclair, J. & Coulthard, M. *Towards an analysis of discourse:* the language of teachers and pupils. Oxford University Press.

Turner, R. Some Formal Properties of Therapy Talk. In D. Sudow (ed.), *Studies in Social Interaction*. The Free Press, 1972.

Turner, R. *Ethnomethodology*. Harmondsworth, Penguin, 1974.