Acts in Discourse: From Monological Speech Acts to Dialogical Inter-Acts

PER LINELL AND IVANA MARKOVÁ

It is nowadays generally agreed that talk in dialogue (and also in monologue) is constituted by interactants' doing things, through discursive action and interaction. This kind of interest in speech activities was greatly promoted by Wittgenstein's and Austin's work on language, and the most influential approach has been speech act theory as developed by John Searle.

The natural habitat for speech acts is, of course, authentic situated discourse and conversational interaction. Yet, speech act theory was developed mainly in terms of intuitive analyses of made-up sentences in imagined contexts. Since the early 1970's, many attempts have been made to apply speech act theory to authentic discourse, and this research has indicated a number of shortcomings of the theory.

In this paper, we will try to do mainly two things. First, we will outline what we take to be the general epistemology underlying speech act theory, and demonstrate that the fact that many aspects of discursive action cannot be captured by speech act theory is connected with its epistemology of monologism. We will also outline an alternative epistemology for the study of communicative inter-acts, that of dialogism.

In the second part of the paper, we will consider some of the many existing proposals to apply speech act theory to authentic discourse. We will argue that these models are fraught with inadequacies and inconsistencies, and then suggest what a scheme for a dialogically based coding system of authentic discourse might look like.

1. THE MONOLOGISM OF SPEECH-ACT THEORY

Speech act theory (henceforth abbreviated SAT) is concerned with the supposed actions of people when they communicate through spoken inter-
action. It can hardly be disputed that SAT is the classical approach to utterances taken as actions, rather than as linguistic forms. SAT was initiated by Austin (1962) and developed by John Searle (1969, 1975), with whose name it is primarily associated. Searle (1975) presented an authoritative version of a supposedly exhaustive taxonomy of speech acts. Though several modifications have later been proposed, and some of these will be touched upon below, Searle's (1975) paper provides the most succinct and clearly articulated account of his own views on speech acts. Recently, Searle (1992) has commented fairly extensively on the relation of conversation to SAT.

SAT provides a theory which is clearly based on a perspective that views speakers as individual agents. Speech acts are performed by individual speakers, just as thoughts and intentions belong to individuals. John Searle himself gives the following, half self-ironical, account (1992:7):

Traditionally speech act theory has a very restricted subject matter. The speech act scenario is enacted by its two great heroes, "S" and "H"; and it works as follows: S goes up to H and cuts loose with an acoustic blast; if all goes well, if all the appropriate conditions are satisfied, if S's noise is infused with intentionality, and if all kinds of rules come into play, then the speech act is successful and non defective. After that, there is silence; nothing else happens. The speech act is concluded and S and H go their separate ways. Traditional speech act theory is thus largely confined to single speech acts.

SAT has roots in philosophy and linguistics. At the most fundamental level the philosophical assumptions of speech act theory are those of individualistic Cartesian philosophy. This individualistic or monological point of view portrays human agency, cognition and consciousness as faculties belonging solely to the individual rather than as processes resulting from interdependencies between individual and environment. It portrays the agent as an autonomous information-processing device (Rommetveit, 1991b), and it does not seem equipped to conceptualize the co-development of the mind and its cultural and social environment. Monologism is at least partly due to its non-developmental presuppositions typically conceiving facts as given, truth as being outside the individual, meaning as literal, and linguistic communication as based on truth-functional semantics.

With regard to linguistics, SAT was originally developed as an alternative, or rather a supplement, to traditional sentence-based semantics which was concerned with truth conditions and propositional content; as such, SAT argues for the universality of illocutionary content. However, SAT shares a host of properties with its intellectual ancestors within linguistics. In general, most of these are part of what Linell (1982) termed "the written language bias in linguistics". For one thing, it continues to work with sentences (or sentence equivalents) as autonomous units. Autonomy means that whatever the relevant contextual conditions are, they are treated as extrinsic with regard to the speech act. Secondly, classical SAT went on working with made-up examples (in imagined contexts and of course written on paper) rather than with authentic connected spoken discourse. Thirdly, SAT still accepts the idea that the task of semantics or pragmatics is to map onto each possible expression a unique semantic interpretation. (A weaker formulation is that this mapping should apply to each possible "reading" of the expression, which of course makes the argument circular.) Note that this view of linguistic semantics and pragmatics is characteristic of the work of both those who are concerned with sentences in abstract and some of those who deal with situated utterances. In other words, it is common to, for example, Chomskyan generative grammar, Grice's (1975) theory of conversational implicature, Sperber & Wilson's (1986) theory of relevance, and Labov & Fanshew's (1977: esp. ch. 9) model for analyzing real talk. Thus, while these theories naturally admit that many utterances are vague or ambiguous on the surface, they deny, in effect, that this also holds for a deeper level, when rules of implicature and interpretation have been applied (though Labov & Fanshel admit that "the problem of 'correct interpretation' can never be solved completely" (1977:73)).

However, our account of SAT in relation to actual discourse is not completely fair if left with the above-given quotation from Searle. In fact, the author goes on immediately to say the following (1992:7):

But, as we all know, in real life speech acts are often not like that at all. In real life, speech characteristically consists of longer sequences of speech acts, either on the part of one speaker, in a continuous discourse; or it consists, more interestingly, of sequences of exchange speech acts in a conversation, where alternatively S becomes H; and H, S.

Searle then proceeds to discuss what a theory of conversation might be like. He argues that there can be no theory of conversation, one reason being that there are no intentions at that level. We should not, any more than many of Searle's critics (see e.g. Searle et al., 1992), conclude that this shows that there can be no theory of dialogue, nor should we deny that there are plans and projects in discourse that extend far beyond single utterances. However, rather than continue along these lines, it seems more fruitful to go back to the substance of SAT and see where the theory goes wrong.

2. THE FOCUS OF SPEECH ACT THEORY: THE ASSIGNMENT OF RESPONSIBILITIES, RIGHTS AND OBLIGATIONS.

Every theory is subject to a particular perspective, and Searle (1992) clearly concedes that conversation involves much more than individual speech acts. In spite of that, while he (1975) discusses many properties and divisions among SA categories, these all pertain to individuals' actions, while interactions' aspects remain largely neglected. It is for this reason, we argue, that SAT cannot be
considered an exhaustive theory of "what people do with words". In his recent publica-
tion (1992), Searle admits that interaction involves shared or collective inten-
tionality (pp. 22, 138). But before proceeding to some of the riders, let us con-
consider what seems to be the primary focus and foremost concern of SAT, i.e.
the *illocutionary aspects* (illocutionary points and forces) of speech acts. We shall
look at what these aspects involve, and argue that illocutionary points, the very
heart of SAT, cannot be explained solely by reference to individual speakers' inten-
tions, i.e. in monologic terms.

Illocutionary points are the basic categories in terms of which SAT classifies
uses of language (Searle, 1975; 22; 1992:20). (There are eleven other
dimensions, which Searle (1975) considers relevant for the subclassification of
illocutionary acts, but the illocutionary point is definitely the basis.) We will
argue that illocutionary points are concerned with the *assignment of epistemic and
practical responsibilities, rights and obligations, to particular persons.* Epistemic
responsibilities are expressed in propositions and categorizations in discourse,
and speakers can therefore be held accountable for them. Associated with
epistemic responsibilities are rights and entitlements to introduce and maintain
topics and topical aspects (cf. Linell & Luckmann, 1991). (For a detailed notion
of epistemic responsibility, see Rommetveit, 1991b.) Practical responsibilities
are commitments to do certain things, and these are also often established
through communicative acts. The system of rights, duties and obligations (cf.
also Levinson, 1982:277–8) may, in the words of Muhlhäuser & Harré (1990:
29, 41), be characterized as a 'moral order' of dialogue and conversation. Searle
gives a distorted picture of this order through his persistence in dealing with it
only in terms of the reliability of individual intentions (Muhlhäuser & Harré,
1990:41).

We propose to interpret illocutionary points of speech acts *generally* in terms of
responsibilities, rights and obligations, even if Searle himself used the term
"obligation" only about commissives, e.g. promises; a promise assigns a
particular commitment (to do q) to the speaker. But with an appropriate
extension of the terms, we argue that the assignment and assumption of
responsibilities is a ubiquitous feature of spoken interaction. For example, an
assertion amounts to the speaker's taking on some (epistemic) responsibility for
the truth of what has been asserted (p). By the same token, the hearer is assigned
the right to expect the speaker to go on arguing in consistency with the assertion
(that p). Similarly, a request is initiated by a speaker, who will then be held
accountable for having requested something (q) from the hearer, and if the
request is duly understood and agreed to by the hearer, it assigns to him or her
the commitment or obligation to do q. In all cases, issues of intentionality and
sincerity are involved. The assignment of epistemic and practical responsi-
bilities is a collaborative undertaking involving reciprocities and mutualities
between interactants. Far from being individual businesses, the rights,
 obligations and commitments are distributed between, assumed by and

 assigned to the self and the other, and this has at least three different aspects,
 neither of which is properly accounted for by SAT.

The first important point is that the process of assigning rights and
responsibilities to individuals (speakers and/or recipients) is a collective
achievement which requires *sequential resources*. To be sure, a communicative
project of establishing epistemic and practical responsibilities (corresponding to a
'speech act') is typically *locally expressed*, but it is seldom something contained
in one individual's utterance only, but something *co-produced by interactants*
distributed and over a local sequence. Take for example a defendant's admission of
guilt in a court trial. Although in pleading guilty, the defendant takes
responsibility for the prior (criminal) actions, this is not something which can be
done and communicated by the defendant alone. It requires an appropriate
context: the defendant's contributions must come as a response to a question
asked by the judge in some special way and at the right stage of the trial.
Furthermore, the defendant's answer is regularly checked by the judge, e.g. by
requesting clarifications or confirmations or by issuing an authoritative
confirmation (declaring it as a valid admission). Linell et al. (1993), using an
extensive corpus of data, have demonstrated the different ways in which this
collaborative process may be carried out. Similar analyses can be made of other
'speech acts' such as promises or requests for action. Even simple assertions,
which are often given as prime examples of autonomous speech acts ("neither
constraining nor being constrained by surrounding speech acts", in Dascal's
(1992:37) words), are in fact seldom made in isolation; the whole point of the
assertion (that p) is that it is *spoken*, rather than something which is only tacitly
believed by the speaker; thus, the speaker tries to *communicate* that p and thereby
becomes socially responsible (in front of his or her audience) for the truth of p.3
Therefore, as a communicative project an assertion usually also involves
"illocutionary uptake" (to borrow Austin's (1962:116) term also employed by,
e.g., Sbisa, 1992); the hearer somehow signals his or her understanding and
reaction, and is then expected (if not obligated) to behave accordingly.
The project involves the 'co-operative completion' by the other (Muhlhäuser &
Harré, 1990:25). This is of course not to deny that there are communicative
situations in which recipients need not and do not provide overt responses.
However, even if an utterance, such as an assertion or a request, is met with
silence, that reaction, or absence of a reaction, will take on communicative
significance and retroactively influence the interactional and illocutionary
value of the utterance in question.

Secondly, although the discursive acts associated with the assignments of
rights and obligations are usually *not symmetrical*, i.e. only one speaker
contributes the 'core' utterance of the assertion, question, promise etc., and the
other typically responds to it, the assignments have *bilateral consequences*,
bounding the interactants under an implicit 'contract'. Thus, a promise of course commits
the speaker A to do q, but this, at the same time, means that, by accepting the
promise, the recipient B commits him- or herself to a certain attitude towards A (and q); since B has the right to expect the future action q, this will influence his or her own conduct. In other words, having performed the communicative project of promising, A and B are entitled to make legitimate, but different, demands on each other. Similarly, if A asserts that p and assumes the corresponding epistemic responsibility, then B, as we noted, is simultaneously given the right to believe that A believes that p and will act in accordance with this belief. To take the example of the defendant’s admission of guilt in court again, we argued that even if s/he clearly declares him- or herself responsible for certain actions, the court is at the same time assigned certain rights and obligations, e.g. those of making further inquiries into the circumstances and deciding on a sentence. Assuming that the defendant understands the situation (and if s/he does not, it should be explained by the judge or the attorneys), s/he in turn, by her/his admission, contributes to assigning to the court these rights and obligations. In most cases, the responsibilities assigned to the interlocutors in a dialogue are not identical. Instead, a web of partially different and complementary rights and commitments is assigned and assumed.

Thirdly, despite the complementarity just alluded to, the dialogues make responsibilities divided among, or (asymmetrically) shared by, interactants. The one partner in the dialogue, by adopting the perspective of the other, becomes partly co-responsible for that perspective; s/he then has to accept at least some of the categorizations and descriptions implied by it. For example, Rommetveit (1991a) shows this in his analysis of Nora’s dependence on Helmer in the first act of Ibsen’s A Doll’s House. Apart from this, no speaker is normally alone responsible for, and cannot fully intend, all the connotations and implications that can be associated with his or her utterances. As speakers, we “inherit” commitments from the culture to which we adhere and from the language we speak; words and utterances have built-in “histories” which we cannot fully escape (e.g. Bakhtin, 1986).

3. SOME FURTHER INADEQUACIES OF SPEECH ACT THEORY

Although the assignment of epistemic and practical rights and responsibilities is certainly an important aspect of communication, we have argued that it cannot be accounted for solely in terms of individual speech acts and intentions. Moreover, the building of coherent discourse involves many other aspects than these assignments, e.g. developing topics and sustaining coherence, and soliciting or inhibiting conversational activity on the part of the other, i.e. negotiating participation. All these aspects are thoroughly dialogical in nature and SAT has had very little to offer here. Accordingly, taken as a theory of spoken interaction in real communicative activities, SAT has been subjected to various criticisms. Of the various points discussed in the literature we shall mention four which are the most relevant for our purposes.

Searle (1975:22) notes that some speech act verbs mark or can take more than one illocutionary point, and more recently, Searle (1992: 143, et passim) admits that in some cases, e.g. indirect speech acts, the speaker performs more than one act per utterance. However, it is safe to say that, in principle, SAT assumes that each utterance has one illocutionary point. Yet, it has been made abundantly clear by many that most utterances fulfill several functions, and that, sometimes, the illocutionary function of an utterance is vague or ambiguous (e.g. Labov & Fanshel, 1977; Levinson, 1981, 1983). The point has often been made that communicative actions, as well as other actions, can be described as part of many kinds of overarching projects; Rommetveit (1990:86ff.) makes this point with his example of Mr. Smith, who, in mowing his lawn, can be seen as simultaneously doing many other things: keeping up property value, engaging in physical exercise and thus ‘working’, engaging in leisure-time activity and thus ‘not working’, avoiding his wife, to mention just a few. Similarly, Labov & Fanshel’s (1977) well-known analysis of a psychotherapeutic interview clearly shows the multifunctionality of single discourse contributions. For instance, when Rhoda, the young anorexic girl, asks her mother “Well, when do you plan t’ come home?” (cf. pp. 48ff, 155ff.), this question can be taken as a request for information, a request for action, a request for help, and a criticism of Rhoda’s mother. In addition, when Rhoda reports having asked this question, she may be seen as construing it as a status-supporting act with respect to both her mother and the therapist.

3.2. Activity Dependence

SAT takes illocutionary force to be inherent in the utterance act, treating various contextual conditions as extrinsic factors. Its critics (e.g. Levinson, 1979, 1983) have shown that some discourse contributions are fundamentally dependent on subordinate frames of activity types. For example, the defendant’s admission of guilt is indicative and partly constitutive of the court trial activity itself. That elementary communicative acts and global activities co-constitute each other was of course envisaged already in Wittgenstein’s (1958) concept of a language game. According to SAT, however, it is the elementary speech acts that build up discourse; it is not the more global activities and projects that motivate speech acts. (Searle, 1992, compromises this position somewhat.) Searle also tends to talk about “acts” rather than “speakers”, very much as if acts are independent not only of activity contexts and situation definitions but also of concrete speakers. However, activities are subject to human projects, concerns, commitments, and perspectives, and it is the interactants who set and take perspectives (e.g. Rommetveit, 1990; Graumann, 1993).
3.3. Sequentiality.

In SAT, the sequences surrounding a particular utterance are treated as external to that act and as simply consisting of other independent acts. Each speaker’s utterances are taken as if they were autonomous and not intrinsically contingent on the partner’s. For example, if a person A promises another person B to do q, then it is the individual speaker A’s utterance only which is the promise. If B says something too, e.g. taking up and confirming A’s utterance as a promise, then this utterance of B’s is another (kind of) independent speech act, and the two acts are not seen as part of a more comprehensive communicative project of establishing and mutually recognising a promise (cf. above). Hancher (1979), amending Searle (1975), admits speech act categories that require the collaboration of two or more individuals, such as cooperative commissives (e.g. contracts, marriages) and reciprocal cooperative declarations (e.g. gifts, sales, appointments) (cf. Stiles, 1981:241). Searle (1992:10) says basically the same with reference to offers, bets, and invitations. However, it is not that some speech acts require collaboration of both interactants while others do not. As we have pointed out, the allocation of epistemic and practical responsibilities is systematically dialogical in nature. Accordingly, it usually requires collaborative work accomplished over sequences by different interactants (Streeck, 1980; Levinson, 1983; Atkinson & Heritage, 1984:5-7; Linell et al., 1993; Sbisa, 1992).

3.4. Initiations versus Responses

All Searlean speech acts are discursive initiatives, e.g. assertions, requests for action, promises, accusations, as if dialogues may be construed as series of such active interventions. But discourse, monologue as well as dialogue, does not consist of disconnected utterances. Contributions are not only initiatives, but they are also responsive with regard to prior contributions. That discursive acts are double-sided, being both related to preceding acts and projecting and constraining next contributions, is a basic fact not captured by SAT. We shall return to this point later.

4. DIALOGICAL ANALYSIS: BASIC TENETS

In contrast to SAT, which is monologically based, we shall now introduce a dialogical conception of discursive interaction. This alternative conception comes from the epistemology of dialogism (Bakhtin, 1986; Holquist, 1990; Marková & Foppa, 1990). The term ‘dialogism’ was coined by neo-Kantians in the 1920’s who, for their purposes, assimilated Hegel’s social dialectic of the dyadic relationship between the self and other. According to dialogism, the socio-historical and ontogenetic developments of language, cognition and communication, as well as their maintenance in face-to-face interaction, are fundamentally built on the mutuality between self and other. Many developmental psychologists have expressed this dialogical principle of the development of the self; here are two examples:

the knowledge of the other and the knowledge of the self are parallel, and the task of understanding early social cognition must be in the task of studying self, other and their interaction (Lewis & Brooks-Gunn, 1979:240)

The mechanism of knowing oneself (self awareness) and the mechanism for knowing others are one and the same. (Vygotsky, 1978:29)

Mead (1934) has shown that in communication an action is never the individual’s action only. The individual merely starts an action, but it is completed by the other participant. Actions and intentions are dyadic. They must be conceived in terms of their counterparts, just as are figure and ground. Mead talks in this context about a ‘conversation of gestures’.

Dialogism is an epistemology of social processes and, more specifically, of humans interacting operating on a moment-to-moment basis in situated discourse. Here, in contrast to SAT, dialogism (Bakhtin, 1986; Marková & Foppa, 1990) portrays speakers (as well as listeners, writers, and readers) as being “in dialogue with” their interlocutors and various contexts, local (e.g. concrete situational) as well as global and abstract (cultural, sociohistorically constituted) ones. Since discourse and dialogue do not consist in individuals’ pronouncing utterances (performing individual speech acts) but in their attempting to establish shared and mutual understanding of things talked about, communicative acts are necessarily collective, involving both (or several) interactants, although of course the constitutive speaking and comprehending activities are characteristically asymmetrically distributed (Marková & Foppa, 1991). Single discourse contributions have a relational or Janus-like nature, and this can be explained in terms of a three-step analysis of communicative interactions.

4.1 The Three-Step Analysis

Some discourse analysts claim that dialogue can often be analysed in terms of three-part sequences. For example, Sinclair & Coulthard (1975) found frequent three-part exchanges in actual interaction: initiation-response-feedback (I-R-F, cf. section 5.1). In their model, an exchange, at least in educational settings, consists of three moves (and three discursive acts): person A (teacher) says something or asks a question, person B (student) responds, and
A then evaluates B's response. While Sinclair & Coulthard argue that a feedback or evaluation move is an important device for teachers' educational purposes, it has been claimed that the three-step sequence has a more general significance as a minimal unit of interaction (Sinclair, 1980:122).

However, in this section we shall turn to a fundamentally different way of looking at three-partedness of discourse employing the epistemology of dialogism (Marková, 1987, 1990a, 1990b; cf. also Foppa, 1990:170). The basic idea of a three-step analysis was implied by Mead back in 1934, in his analysis of the 'conversation of gestures' referred to above. Although superficially similar to Sinclair & Coulthard's analysis in terms of a three-part structure, there is a fundamental conceptual difference between them. In general, Mead's 'conversation of gestures' amounts to saying that for an agent A to have his or her act (or utterance) a communicated and mutually understood, it takes the interactants, A and B, at least three steps. First A makes an action (utterance) a, then B responds to A's action, usually through another action (gesture, utterance) b, and then A must show his or her reaction to B's response, again usually through another action (gesture, utterance) a'. Note that without the third step, while A has access to B's understanding (displayed, or at least partially displayed, in b), B has not yet received any reaction from A and hence cannot know whether b (and its presupposed understanding of a) fits in with A's ideas; hence, no mutual ground has been established (unless, of course, there are contextually established routines which make such checking procedures unnecessary).

So far, Sinclair & Coulthard's and Mead's analyses appear to be similar. However, with Sinclair & Coulthard, there is just a temporal and unidirectional sequence of acts from question to answer to evaluation. The last part is then A's evaluation of the communicative action of B. In contrast, in Mead's "conversation of gestures" A's action a', i.e. A's second action, is not an evaluation of B's action b only; rather, it is a response to both a and b. In the most general context of social communication (e.g. Luckmann, 1990), a' could be simply a changed behavioural attitude, say, of a threatening posture of one dog to the posture of the other (cf. Mead, 1934:14). In human communication, however, a' is a reflexive communicative action, based on understanding backwards or re-interpreting both interactants' previous steps. In other words, A understands his or her own communicative action on the basis of B's response to it. At the same time, human communicative actions are oriented forward. Speakers anticipate their interlocutors' messages and adjust their own messages to these anticipations.

Of course, it is not just the agent A but also B who interprets and re-interprets the other's and his or her own acts. Indeed, the three-step process gives rise to exchanges that, in most cases, are expanded far beyond just three units (cf. Jefferson & Schenkein, 1977). The reason for this is that second and third actions (defined with regard to a given first action) are seldom just responses; instead, agents use these opportunities to introduce new initiatives and new content into the interaction or conversation. Thus, second and third actions become first actions in their own right, projecting and anticipating new second and third actions. In this way, the three-step process is a generative principle that accounts for the inherent and dynamic "power" of dialogue to progress continuously (Linell, 1990; Foppa, 1990).

4.2 The Nature of the Discourse Contribution.

We can now apply these dialogical assumptions to a single discourse contribution. In this account, we will take the conversational turn5 to be the basic unit of dialogue, and show that it has a relational nature. However, an analogical argumentation is possible for sentence-sized utterances (the proto-typical form of "speech acts").

Taking one's own perspective and that of the other does not mean just re-interpreting past communicative actions. It also means projecting and anticipating one's interlocutor's point of view. Sense-making activities in producing a discourse contribution include a retrospective interpretation of a prior contribution and a prospective anticipation (projection) of what is going to follow, or of what would be a relevant continuation. Thus, a contribution is seen as the result of bidirectional forces; it is co-determined by the past context and by the commitment to the future; it is related both "backwards" to the prior (already given) contribution(s) and "forwards" to the potential next contribution(s) (to be given). In Heritage's words (1984: 242), a contribution to dialogue is both context-shaped and context-renewing. We propose to call these two relational aspects responsive and initiatory, the former dealing with coherence backwards and with respect to prior discourse, and the latter dealing with the introduction of new content, making discourse progress and creating relevance conditions for contributions to come.

The dialogical analysis of the single discourse contribution amounts to characterizing it in terms of both what it is and in terms of its relevant outside. This means that each contribution (turn, utterance) is characterized conceptually by three elements (cf. the three-step analysis):

- the relationship to prior contribution(s),
- the contribution itself (with its propositional and illocutionary content),
- the relationship to next contribution(s).

Therefore, an utterance, or a text, cannot be separated from its prior and projected contexts; texts and contexts co-develop and co-constitute one another. This theory proposes that discourse is made up of dialogical inter-acts, i.e. acts which are intrinsically dependent on their contextual relations, rather than of monological "speech acts".

© The Executive Management Committee/Basil Blackwell Ltd 1993.
As a brief summary, we can say that dialogical theory claims that communicators and their acts and projects, i.e. their discourse contributions, are “in dialogue with”, or dialectically related to, roughly two dimensions of contexts, the sequential environment (prior-subsequent relations) and the activity frame (and its associated background knowledge) (cf. section 3: sequentiality and activity-dependence). A more fine-grained analysis would of course complicate the picture considerably.

5. SPEECH ACTS IN ACTUAL DISCOURSE

As we have repeatedly stated, SAT was developed within philosophy and theoretical linguistics, without being confronted with empirical data from authentic discourse. Attempts have been made (by other scholars than Searle) to view connected discourse as built up of Searlean speech acts, i.e. to classify actual conversational contributions in terms of SA types. Vanderveken (1985) claimed that “any materially adequate theory of conversation must start with the recognition of the fact (sic!) that illocutionary acts are the constitutive units of conversation” (p. 181). And, indeed, many influential discourse analysts either explicitly develop SAT, or they draw at least implicitly on some of its crucial aspects. In order to show that an analysis of discourse which is based on SAT does not escape the monologism of SAT, i.e. the viewing of turns or utterances as independent units, we shall do two things. First, we shall examine one major theoretical approach to the study of discourse, that of the so-called Birmingham school of discourse analysis. Second, we shall examine some coding systems for authentic discourse.

5.1. Sinclair & Coulthard.

Sinclair & Coulthard (1975), and other works from the so-called Birmingham school of discourse analysis, notably Coulthard & Montgomery (1981), Edmondson (1981) and Stubbs (1983), have proposed a multi-level structural model of discourse, of which we will only briefly consider the levels (Sinclair & Coulthard call them “ranks”) of exchanges, moves, and acts. Furthermore, we will focus on the central exchange types, disregarding, e.g., boundary exchanges. The basic units organizing discourse (cf. Stubbs, 1983:129) are moves of three basic types (plus some subtypes): I (= initiations), R (= responses) and F (= feedback, follow-up). These make up sequences (“exchanges”) of various forms: I-R-F, I-R, R/I-R etc. (cf. Stubbs, 1983).

Superficially, this makes the model similar to the three-step analysis (Marková, 1990b, cf. section 4.1 above), and at least terminologically akin to initiative-response analysis (Linell et al., 1988; cf. section 6 below). On a closer scrutiny, however, the similarities are misleading, and we argue that the Birmingham models are strongly monologic.

Consider the following example of discourse, which is drawn from classroom interaction (see Coulthard et al., 1981:5):

(t) (T = teacher, P = pupil)
1. T: Can you tell me why you eat all that food? Yes.
2. P: To keep you strong.
3. T: To keep you strong. Yes. To keep you strong. Why do you want to be strong?
5. T: To make muscles. Yes. Well what would you want to use... what would you want to do with your muscles?
6. P: Sir, use them.

Coulthard et al. (op.cit.: 6) comment on this excerpt in the following terms: “Our first real break-through came when we realised that ‘utterance’ was not in fact a unit of interaction at all — there are significant breaks in the middle of utterances 3 and 5 in the example above of a kind repeated in many teacher utterances. The first half of the utterance links back to what the pupil has said, the second half links forward to what a pupil will say next.” The authors conclude that turns (their term is “utterance”) must be broken up into smaller units called “moves”. The turns in (t) would then be coded: I, R, F/I, R, F/I, R.

Each move contains, by definition, only one aspect (I or R or F). The same applied to the acts that fill the moves; these acts are basically speech acts, although the Sinclair & Coulthard model, as seen from a Searlean point-of-view, hosts a rather mixed bag of acts (examples are acts like accept, evaluate, comment, elicit, prompt). There are turns that are divided into two moves (acts) e.g., R/I or, as in (t), F/I, which are then taken as physical subdivisions of the turns involved, each part having its own value.

By contrast, in a dialogical analysis (cf. IR analysis as discussed below, section 6), each contribution contains both response and initiative, and these are then regarded as abstract relational aspects rather than as physical parts of turns. Accordingly, rather than denying that turns are units of interaction, such an analysis recognizes that which “links back” vs. that which “links forward” as relational aspects of each such unit. Of course, in some cases there is a rough correlation between aspect and physical segment, but this need not be the case. This can be easily demonstrated by slightly changing the turns in (t):

(t') (simplified of (t))
1. T: Can you tell me why you eat all that food? Yes.
2. P: To keep you strong.
3. T: Why do you want to be strong?
5.2. Some Monologically Based Coding Schemes

The Sinclair & Coulthard model was not, to the best of our knowledge, developed as a coding scheme to be used for a reliable and exhaustive categorization of units in actual discourse. In order to examine whether SAT can really be a theory of ‘what people do with words’ and whether it provides a theoretical basis for coding schemes, we have to turn our attention elsewhere. Indeed, a number of psychologists, psycholinguists and sociolinguists have developed coding schemes on the basis of SAT. These systems are interesting in that they make salient some of the inadequacies of SAT.

We noted (section 3.4) that all Searlean speech acts can be seen as initiatives (initiatives). It is therefore interesting to find that empirical researchers, just like Sinclair & Coulthard, have felt the need somehow to introduce the responsive dimension into their taxonomies. Dore (1977a,b), who was one of the first to make use of SAT in quantitative studies of interpersonal communication (in his case the discourse of young children and their interlocutors), introduced the category of responses, basically answers, compliances, agreements and acknowledgements (for a critique, see Stiles, 1981:234). Labov & Fanshel (1977), in their well-known case study of therapeutic discourse, tried to come to grips with responsive aspects in basically two rather different ways: on the one hand, they acknowledged the “meta-linguistic actions” of initiating, continuing and ending speech events, and, on the other hand, they introduced action subtypes in the other categories associated with first-slot (A’s initiative), second-slot (B’s response) and third-slot (A’s reaction to B’s response). In this way, they added several types of responses to their coding scheme. For them, responses are a minor family of acts, in effect subclassified in accordance with the initiatives to which they respond. Furthermore, responses and initiatives are separate individual acts (just like in SAT), rather than relational aspects of dialogical inter-acts.

d’Andrade & Wish (1985) have developed one of the more sophisticated coding schemes. They include the category of ‘reactions’, along with other variants of Searlean speech act types. They observe (p. 248), however, that the category of reaction is alien to SAT principles. While Searlean ‘initiatives’ can, to some extent (but cf. section 3.1), be construed as mutually exclusive categories, d’Andrade & Wish point out that this is no longer so with their amended system. Reactions are systematically also analyzable in terms of one or the other of the ‘initiations’. d’Andrade & Wish therefore allow double coding of utterances, i.e. the same utterance can be categorized in several ways. More specifically, the same utterance can be coded simultaneously as a reaction and an initiation, e.g. an answer which itself is an assertion. However, although they have departed to a considerable extent from other coding systems based on SAT, d’Andrade & Wish still operate, conceptually, within the framework of monologism. What they have shown is, nevertheless, that monologism can no longer be sustained in the study of authentic discourse.
More generally, Dore's, Labov & Fanshel's, and d'Andrade & Wish's coding schemes reveal the difficulties of any attempt to apply SAT to natural discourse. It has become necessary for all of them to introduce categories like responses or reactions. The way they cope with the problems involved in this was to mix the responsive-initiatory dimension (as described above) with the basic SAT concern, i.e. that of the allocation of epistemic and practical responsibilities. Thus they have developed coding schemes in which different kinds of categories are combined. They have not succeeded in dealing with multi-functionality in a logically consistent way. Instead, they have introduced multiple codings and have made arbitrary decisions concerning the number of different categories to be included in their systems.

6. TOWARDS A DIALOGICAL CODING SYSTEM

A coding system for interaction must aim at fulfilling a number of requirements. Some of these are general, having nothing to do with the interactional nature of the units:

a) it should be based on a logically consistent system, which does not mix categories of different sorts and hence does not allow for the introduction of arbitrary numbers of categories, and

b) it should permit an exhaustive coding of units (and this coding must of course reach reasonable standards of reliability).

A severe flaw in SAT-based coding systems is that they consist of categories that have been added to the schemes in ad-hoc ways, i.e. when new cases and new needs have appeared. The multi-ambiguities in speech act status is a major obstacle. Thus we have seen that Rhoda's question "Well, when do you plan t'come home?" in Labov & Fanshel's research was found to have at least half a dozen different functions, and there is no principled reason to assume that this sets the boundary.

An additional, and more specific, requirement on a coding scheme for interaction is that:

c) it should account for the dialogical properties of discourse contributions.

With respect to the last-mentioned point, recall that a dialogical conception of sequentiality and of the nature of single discourse contributions is different from that of recognizing constraints on sequences of self-contained units. It is a fundamental characteristic of human interaction that it consists of meaningful acts, and ultimately of meanings that are socially co-produced. Utterances, therefore, cannot be treated simply as (sequences of) physical behaviours, or

linguistic expressions with static interpretations. What makes dialogical theory appropriate for the study of interactions is that it conceives each contribution to be a relationally defined unit, i.e. an 'inter-act'. Its constituent features include its contextual relations to prior and projected (subsequent) contributions (see section 4.2).

In a series of papers (Linell, Gustavsson & Juvonen, 1988; Linell, 1990; Marková 1991) we have presented and applied a coding system for dyadic interaction called Initiator-Response (for short: IR) analysis, which tries to comply with the above-mentioned requirements and which is based on a systematic exploration of the discourse-building ties between turns in dialogue.

We mention this model here not because it deals with all problems of dialogue; it clearly does not, since any analysis of discourse is partial and subject to particular concerns and problem formulations. Moreover, there are a number of conceptual and methodological issues pertaining to a dialogical theory of coding. For example, a coding system must necessarily bracket some dialogical properties of the interaction. Space restrictions do not allow us here to specify how coding practices, in general, could be construed within a dialogical framework. Instead, we must be content merely to suggest that a dialogical coding system is not an unattainable goal.

Within IR analysis, each turn is analyzed both in terms of its retrospective ties (response aspects) and prospective ties (initiative aspects). Besides this principle of double-sidedness, the model includes a few other important dimensions, e.g. whether the responsive ties connect with a contribution of the speaker him- or herself or of the interlocutor (the self vs. other dimension, which is of fundamental significance in social psychology), and whether responsive ties are local or non-local. With a few further subcategorizations, all fully consistent with these three basic principles, the responsive and initiatory aspects of contributions are analysed as follows.

The response aspects of discourse contributions are subcategorized in terms of the following features:

- locally tied (to adjacent prior turn) vs. non-locally tied (tied further back in prior discourse) vs. no discourse-internal retroactive tie;
- tied to other's vs. speaker's own (adjacent prior) turn;
- focally tied (to major topical aspects) vs. non-focally tied (to functional, formal or topically peripheral aspects) of the other's prior turn (the latter category of non-focal ties comprises, among other things, metacommunicative and metalinguistic contributions);
- minimal vs. expanded responses (where "minimal" stands for what is minimally required as a response to a prior initiative);
- adequate (accepted) vs. non-adequate (or non-accepted) responses (this dichotomy only subcategorizing minimal responses).
The initiative aspects are subcategorized into only two types, soliciting and non-soliciting initiatives, which request vs. only invite (or possibly even discourage) a following contribution from the other.

In total, the system contains about twenty different turn types designed for exhaustive categorization of contributions to dyadic interaction. With some additional categories the system may be extended to cover multi-party dialogues as well. Basically, the system is designed to capture the features that build coherence, or cause incoherence, in dialogue. Since the various IR categories can be assigned numerical values on a scale of interactional strength, it also becomes possible to calculate measures of dominance (vs. submissiveness) in dialogue. 13

Coding systems for discourse presuppose reliable unitization and categorization, which always means that some properties of the stream of discourse must be disregarded. At least, this holds true if a major aim is to contribute to quantitative measures of variation. While we contend that all contributions to dialogue are both responsive and initiatory in character, it is, within a given finite coding system, possible to regard some contributions as purely initiatory or responsive. For example, with the IR system, a so-called free, i.e. untied, initiative is a turn which has no topical or other tie to anything else occurring prior to it in the same dialogue; yet, such an initiative is usually related to something in the context, or has an 'intertextual' relation to some other experience or communicative exchange in the biography of the interactant(s). Similarly, a so-called 'minimal response' is, within the IR coding scheme, defined as only responsive in the sense that it does not bring in anything beyond what has been requested through a prior contribution and it does no more than close a local topic; yet, in doing this, it does of course promote discourse and does bring about a new local context, thus being, however minimally, somewhat context-renewing or initiatory in nature. (The same holds for listener-support (backchannel) items and inadequate (non-accepted) responses, although these are even weaker as context-renewers.)

It is important that the recognition of these types does not in any way contradict the dialogical approach. Conceptually, these contributions are still to be conceived as both initiatory and responsive. Their being 'purely' initiatory or responsive simply means that one or the other component is so insignificant that, for the purposes of the specific coding, it is considered as missing. These 'purely' initiatory or responsive contributions are, therefore, extreme cases of dyadic relationships, with one component reduced to virtually nil. This is conceptually different from the monological approach in which the starting point is the individual's speech act with one function only. More generally, 'initiative' and 'response' remain for us aspects of discourse contributions defined in terms of (proactive vs. retroactive) relational properties that are co-constitutive of the given contribution and its relevant (local) outsides; they are not physical utterances filling first and second slots of exchange structures.

Speech act theory has been advanced as a major attempt to deal with "what people do with words" in spoken interaction. As such, the theory tries to analyze some important aspects of communicative actions, in particular people's assignments and assumptions of rights and obligations in dialogue, i.e. (in our terms) their epistemic and practical responsibilities. However, we argue that SAT provides a distorted picture even of these aspects which are the theory's primary concern. Ironically enough, Searle (1992:22) himself has explained why:

The phenomenon of shared collective behaviour is a genuine social phenomenon and underlies much social behavior. We are blinded to this fact by the traditional analytic devices that treat all intentionality as strictly a matter of the individual person. I believe that a recognition of shared intentionality and its implications is one of the basic concepts we need in order to understand how conversations work.

In this paper, we have tried to indicate that Searle could not "recognize" these implications, unless he abandoned his own theory on most points. Attempts to use SAT in coding authentic conversational interaction have been shown to be seriously flawed. We have also defined the main characteristics of a dialogical theory of interaction and suggested that coding schemes for authentic discourse can be devised on such a theoretical basis.

Per Linell
University of Linkoping

Ioana Marko6
University of Stirling

S-581 83 Linkoping
Stirling

Sweden
FK9 4LA

NOTES

1 An earlier version of this paper was presented at the Study Group on “The Dynamics of Dialogue” at the Werner Reimer-Stiftung, Bad Homburg, Germany, on March 6, 1992. We thank the group for valuable comments.

2 As is well-known and will be alluded to in this paper, one of Searle’s major interests concerns human intentionality. Here it appears that he has modified his stance in some respects over the years. He now seems to have left the position that intentions are (fully) behind or prior to action, and prefers to speak of intentionality in action (Searle, 1991). Furthermore, in addition to individual intentionality, he has now more clearly emphasized the importance of shared or collective intentionality (Searle, 1992:22, 138).

3 Bogen (1991) elaborates on a critique of the autonomous intentionalist semantics of SAT. For example, he says: “By beginning from the perspective of intentionalist semantics, the theory of speech acts commits itself to a monological account of meaning-
adequate expression, thereby ignoring entirely the constitutive role of contextual considerations in the situated production and recognition of intelligible utterances."

Searle (1975) makes a distinction between illocutionary point and illocutionary force, which is not always honoured.

Searle (e.g. 1992: 140ff.) argues that an assertion (and many other speech acts) can be made or performed even if the speaker does not succeed in communicating his/her meaning.

A talk at talk may be defined roughly as a continuous period in which a given speaker holds the floor. "Utterance" is sometimes used coterminously, though it is more common for this term to be used, somewhat loosely, about a sentence- or clause-sized segment of discourse. Therefore, "utterance" and "turn" are different units. However, we will not indulge in a discussion here. The arguments on monologism vs. dialogism may be applied to both kinds of units. Cf. Marková & Linell (forthc.)

This analysis applies at various levels, from single discourse contributions via topics etc. to whole "speech events" or texts. Cf. Bakhtin (1986).

We shall not consider Conversational Analysis (e.g. Atkinson & Heritage, 1984) in this paper. The reason is that this research paradigm does not set itself the task of coding discourse units. However, we would argue that Conversational Analysis is in some respects akin to dialogism (Linell, 1990). Another important theory, which we have to leave out, is that of Clark & Schafer (1987, 1986). Their theory of "contributions to discourse" takes care of the insight that illocutionary points can be established only in collaboration among interactants (cf. section 3). In some ways, however, Clark's theory of the "contribution" is similar to Sinclair & Coulthard's three-part structure and to Edmondson's (1981) above-mentioned plea for the "exchange" as the "minimal unit of social interaction". It recognises dialogical relations within such exchanges but fails to account for the retroactive and proactive relations of each such unit to its prior and subsequent units.

We will try to deal with these issues in a forthcoming paper (Marková & Linell, forthcoming).

In addition, IR codings may of course be combined with independent categorizations, for e.g. amount of talk, type of questions, support/challenge aspects etc. IR analysis has been successfully applied to discourses of many kinds, e.g. court trials (Adelswärd et al., 1987), language lessons (Gustavsson, 1988), social welfare interviews (Fredin, 1993), dialogues with mentally handicapped persons (Marková, 1991), fiction drama dialogues (Rommetveit, 1991a) etc. A full manual (Linell & Gustavsson, 1987) exists so far only in Swedish (but cf. Gustavsson, 1988).

For particular research purposes, it may be desirable to reduce the coding to fewer categories. One radical simplification is to work only with soliciting initiatives (comprising questions and requests, cf. Blank & Franklin, 1980: "obliges"), expanded responses (Blank & Franklin: "comments"), minimal responses, and, possibly, requests for clarification. Such a simplification brings the coding scheme close to, e.g. Blank & Franklin (op. cit.) and glosses over much of the dialogicality of contributions; for example, one easily slips back into a classification of whole turns into either initiations or responses.

REFERENCES


© The Executive Management Committee/Basil Blackwell Ltd 1993.