CONTEXT AND COGNITION: KNOWLEDGE FRAMES AND SPEECH ACT COMPREHENSION *

TEUN A. VAN DIJK

This paper is about the cognitive foundations of pragmatic theories. Besides the fact that the usual appropriateness conditions for speech acts, which are given in 'cognitive' terms, such as 'S knows/believes/wants ... (that) *p*', require empirical investigation, a sound theory of pragmatics must also explain how certain utterances in certain contexts are actually *understood as* certain speech acts. Speech act comprehension is based on rules and strategies for so-called 'context analysis', in which (epistemic) frames play an important role in the analysis of social context, social frames, and interaction type. Results of context analysis are then matched with those of pragmatic sentence analysis, viz. the illocutionary act indicating devices. Finally, some results from the cognitive analysis of discourse processing are applied in a brief account of the comprehension of speech act sequences and macro-speech acts.

1. The foundations of pragmatics

The philosophical and linguistic study of pragmatics requires an analysis of its *foundations.* This basis of pragmatic theories is on the one hand conceptual, e.g. in the analysis of action and interaction, and on the other hand *empirical*, viz. in the investigation of the psychological and social properties of language processing in communicative interaction.

One of the crucial components in an integrated theory of communicative inter-action by natural language is constituted by a *cognitive theory* of language use.

Such a cognitive theory will not only provide insight into the processes and structures involved in the actual production, comprehension, storage, reproduction

and other kinds of processing of sentences and discourses, but also in the ways *speech acts* are planned, executed and understood.

In particular such a cognitive theory of pragmatics will have to elucidate what the relations are between various cognitive (conceptual) *systems*, and the conditions for the appropriateness of speech acts in given contexts.

Besides the systems of beliefs, wants, wishes, preferences, norms and values, that of *conventional knowledge* plays an important role. It is so to speak the basic, and at the same time the social condition for the operation of the other systems in communication. Such systems are studied, e.g. in actual artificial intelligence, under the label of *frames*. A cognitive theory of pragmatics has as one of its tasks to specify how we are able to perform and understand acts of language, and how we are able to act 'upon' such understanding as it is related to cognitive frames.

2. The cognitive nature of pragmatic conditions

The appropriateness conditions of classical speech act theory are usually of a cognitive nature, and include such conditions as:

(i)   speaker knows that *p*
(ii) speaker believes that *p*
(iii) speaker wants *p*
(iv) speaker finds it good that *p* etc.

The concepts involved in such conditions are treated as primitives in pragmatics: they are not further analyzed, neither conceptually – e.g. in terms of epistemic logic – nor empirically. It is clear however that if a theory of pragmatics should claim to be empirically relevant, these various concepts should also be assessed at the proper psychological level, i.e. in terms of experimental or simulative (artificial) findings.

More specifically, it should be investigated which cognitive processes are underlying 'the assignment of 'appropriateness' in communicative contexts. In which respect are the actually formulated conditions 'idealizations', i.e. how 'far' are they from the actual acceptance/comprehension of speech acts?

Although pragmatic conditions have a cognitive basis, it should be borne in mind that the ultimate rationale for a pragmatic theory of language is to bridge the gap between utterances (and hence grammar) on the one hand, and interaction (and hence the social sciences) on the other hand. This may mean, for instance, that whatever a speaker *really* knows, thinks or wants during the accomplishment of a speech act, is irrelevant as long as his behaviour may be *interpreted as*, and hence socially *counts as*, exhibiting these various internal states.

This methodological caveat does not imply however that the cognitive analysis of pragmatic concepts is irrelevant. On the contrary, actual communication *does* involve *real* knowledge, beliefs and wants. There is even a general condition −viz. that of sincerity − which requires a direct link between what is thought, etc. and what is *shown*.

More generally it may even be said that the proper *social* conditions involved in the formulation of pragmatic rules, such as authority, power, role and politeness relations, operate on a cognitive basis: i.e. they are *relevant* only inasfar the speech participants *know* these rules, are able to *use* them, and are able to relate their *interpretations* of what is going on in communication with these 'social' properties of the context. [1]

3. Pragmatic comprehension

A pragmatic theory provides rules for *pragmatic interpretation*. That is, given certain utterances of natural language, it specifies the rules assigning a particular speech act or illocutionary force to this utterance, given a particular structure of the pragmatic context.

Now, this task is more a programme, at the moment, than actually carried out. We know something about various speech acts and their contextual conditions, but little about the systematic relationships with the (grammatical or other) structures of the utterance. These relationships are most conspicuous in the use of explicit performatives, certain particles, the syntactic form of sentences (indicative, inter-rogative and imperative), and in the propositional 'content' of sentences. [2]

Since we know not much more, we have little to offer for a more general theory of *pragmatic comprehension* – which is part of a cognitive theory of information processing. Pragmatic comprehension is the series of processes during which language users assign particular conventional acts, i.e. illocutionary forces, to each other's utterances. The problem thus is: how do hearer actually *know* that when a speaker utters such or such a sentence, that the speaker thereby makes a promise or a threat? What information must be available to the hearer in order to be able to make such assignments?

Obviously this information may come from various sources and through various 'channels':

---

[1] This kind of 'cognitive relativism' does not imply an 'idealistic' approach to language and interaction. As we will briefly mention below, conventions are involved, where conventions are taken as strictly objective properties of social structure. Methodological issues related to this issue will however be left undiscussed in this paper.

[2] The basic pragmatic concepts used in this paper have been discussed elsewhere; for a recent summary and for an analysis of the relations between the semantics and the pragmatics of dis-course, see van Dijk 1977a, also for further references.

A. properties of the structure of the utterance (as assigned on the basis of grammatical rules);
B.   para-linguistic properties, such as speed, stress, intonation, pitch, etc. on the one hand, and gestures, facial expression, bodily movements, etc. on the other hand;
C. actual observation/perception of the communicative context (presence and properties of objects, other persons, etc.);
D.   knowledge/beliefs in memory about the speaker and his properties, or about other properties of the actual situation;
E. more in particular: knowledge/beliefs with respect to the type of interaction going on, and the structures of preceding contexts of interaction;
F. knowledge/beliefs derived from previous speech acts c.q. previous parts of the discourse, both at the micro (or local) level and on the macro (or global) level;
G. general semantic, in particular conventional, knowledge about (inter)action, rules, etc. – especially those of pragmatics;
H. other kinds of general world-knowledge (frames).

The very fact that all these components may be involved in pragmatic comprehension justifies the well-known insight that often we are unable to assign a definite illocutionary force on the basis of a (semantic) comprehension of the utterance (sentence) alone:

(1) I'll come tomorrow!

may function as a promise, a threat, an announcement/assertion, etc. A specific force or function may be assigned only if the communicative context yields information about whether the speaker has certain obligations, the hearer certain wishes, the action a beneficiary role for the hearer, etc. All this information must however be 'detected' in complex processes of comprehension of previous acts and utterances, observation, stored or inferred assumptions, etc.

The question is, thus, how exactly are all these complex processes related?

4. Frames and speech acts

The actual comprehension of utterances as certain speech acts is based on a complex process involving the use of the various kinds of information mentioned above. The comprehension of particular observable indices, however, should be given in terms of more general knowledge: understanding involves general concepts, categories, rules and strategies.

This general 'knowledge' is not amorphous but *organized* in conceptual systems.

One of the ways to account for this organization is in terms of *frames*. [3] Frames are not arbitrary 'chunks' of knowledge. First of all they are knowledge units organized 'around' a certain concept. But, unlike a set of associations such units contain the essential, the typical and the possible information associated with such a concept. Finally, frames seem to have a more or less *conventional* nature, and thus should specify what in a certain culture is 'characteristic' or 'typical'. This criterion seems to single out especially certain 'episodes' of social interaction, such as going to the movies, travelling by train, or eating in a restaurant. Although we might also call a frame the set of epistemic units we have about books, balls and bananas, such units do not – as such – organize our understanding of the world in a way a conceptual frame – as we define it – does, viz. by also organizing our behaviour with respect to the world, and the ways we interpret other's behaviour, as in the frames we have about cashing a check or shopping. Although the distinction between mere concepts and the frame-like organization of conceptual knowledge is still vague – there may be fuzzy boundaries in the theory - we provisionally keep ourselves to the more restricted interpretation of the notion of frame.

The question which arises with respect to pragmatic theory is then: in what sense may we consider speech acts as 'frames'. Clearly, speech acts are acts and they also have a conventional nature. We know, intuitively, how to promise something, how to congratulate somebody, and this knowledge is clearly part of our world-knowledge. In which respect, however, would our 'promise'-frame *organize* our knowledge of the world in a different way than our knowledge of bananas or barking, i.e. in the sense that we 'know' one when we 'see' or 'hear' one? Could we say that there is a whole 'episode' during which a speaker is going about promising – e.g. by accomplishing all kinds of preparatory, component and auxiliary acts, as is the case in taking a train?

Although the appropriate accomplishment of speech act involves a number of conditions, and although there are ways of 'complex' promising or threatening, we would at first sight see no reason why a simple speech act would be a frame (and not, for instance, laughing or hitting). The only 'organizing' principle involved is that relating certain purposes, intentions and certain doings (utterances with certain properties) to contextual states and events.

---

[3] The current notion of 'frame' has been discussed mainly in artificial intelligence and cognitive psychology after Minksy's influential paper (Minsky 1975). For further elaboration of this concept, see the contributions in Bobrow and Collins, eds. 1975 also for further references. The notion, of which variants have become known as 'scripts', 'scenarios' or 'schemata', already appeared in artificial intelligence in Charniak's dissertation (1972), where the term 'demon' was used in order to denote conventional knowledge structures used in the interpretation of discourse. In fact, much of the actual discussion about knowledge representation has its roots in Bartlett's work on remembering (1932) in which the notion of a 'schema' played an important role.

For a critical discussion about the nature of frames and their function in discourse comprehension, see van Dijk 1977b.

Yet, speech acts may nevertheless be connected with frames. First of all, we have typical *speech act sequences* [4] of which the structure has a more or less conventional or 'ritual' character, such as giving lectures, preaching, making everyday conversation, or writing love letters. In such cases we clearly have a number of different (speech) acts, of which each may have a characteristic *function* in the performance of the episode: opening, introducing, greeting, giving arguments, defending, closing, etc. In such cases we may have different *strategies* for fully accomplishing our goals. Moreover, unlike (most) speech acts, they may be culture dependent.

Secondly, speech acts are interpreted on the basis of frame-like world knowledge, e.g. because they are part of such frames. Especially the institutional speech acts, such as baptizing, marrying, convicting or firing (taken as speech acts) are part of often highly conventionalized episodes. Without such frame-knowledge I would for instance be unable to differentiate the utterance "*I sentence you to ten years of prison*" when spoken to me by a judge, in a courtroom, at the end of a trial, etc. or as spoken by my friend being angry against me. We would know that the first speech act 'counts' and the second not, because only the first is part of an institutional frame.

Thirdly, the interpretation of speech acts requires knowledge of what might be called *meta-frames*: we know the general conditions under which actions are accomplished, when they are successful, etc. Thus, if somebody in his sleep tells me "*Can you open the window*?", I would hardly do so, because I have the general knowledge that only controlled, conscious and purposefully intended doings count as actions. [5]

Finally, the interpretation of speech acts involves world knowledge more in general. Speech acts often pertain to past or future activity of the speaker or the hearer: they are essentially functioning as expedient ways in which such activities are planned, controlled, commented upon, etc., or they are intended with the purpose to provide information for such actions. Hence, they basically require knowledge about what is necessary, plausible or possible in the real world. If somebody would tell us: "*I just jumped from the Eiffel Tower*", we would hardly take him seriously.

Similarly, when I congratulate somebody I should assume that something pleasant occurred to him, but our more general world knowledge will have to tell us *what* is pleasant, for *whom in what circumstances*. Pragmatics itself will not make explicit the latter conditions – which belong to a representation of our cognitive semantics.

In other words, whether the necessary conditions for the appropriateness of

[4] For an analysis of speech act sequences and their super-ordinate 'macro-speech acts', see van Dijk 1977a.

[5] For these general notions from the theory of action and their relevance for pragmatic theory, see van Dijk 1977a.

speech acts are actually satisfied must be decided by our knowledge of the world and its frame-like mental organization.

5. Context analysis

Before language users are able to match incoming information against the more general linguistic and other knowledge in memory, they must analyse the *context* with respect to which a certain speech act is performed.

One of the methodological principles which should be kept in mind is that the notion of 'context' is both a theoretical and a cognitive *abstraction*, viz. from the actual physical-biological-etc. *situation*. That is, a great number of features of the situation are not *relevant* for the correct comprehension of the illocutionary force of utterances. It will seldom be the case that, whether my speech participant has red hair or not, I will understand his utterances in different ways (with the possible exception of those cases where such situational details are thematized). Hence a speech understander will focus attention on specific properties of the situation which might be relevant for correct interpretation of both meaning/reference and pragmatic intentions/purposes. A next methodological point is that unlike pragmatics and (the rest of) grammar, a cognitive theory does not (only) have rules and concepts, but strategies and schemata, i.e. devices for a fast and functional processing of information.

Strategies and schemata are the basis of the normal processes of *hypothetical* interpretation: given a certain textual and contextual structure they allow fast assumptions about probable meaning and intention – even if the rules at a later moment may lead to a rejection of the hypotheses.

One of the most obvious examples of such cognitive processing devices is based on the typical structure of the sentence: if an interrogative structure is given, we may provisionally conclude that a question or request is made.

Similarly in the analysis of context. If a complete stranger is heading for us on the street, we may be pretty sure that (globally) he will ask a question or make a request – and not make some assertion about his love life, or a threat.

It seems to follow that in pragmatic comprehension we not only 'establish' a context out of the situation, but also have *schemata* for the analysis of such contexts. That is, if a context obviously satisfies a set of (ordered) *key* features it will be taken as characteristic for a specific set of *possible* speech acts.

It is clear that the analysis of context is a necessary but usually not a sufficient condition for the pragmatic comprehension of utterances. That is, a language user will merely have a certain 'set' towards the *possible* speech acts which may follow. The definite assignment of a speech act takes place, of course, after comprehension of the utterance itself, and after matching of the pragmatically relevant information from the utterance with the information from the context analysis. In this respect, pragmatic comprehension parallels the relative process of semantic comprehension,

in which previous discourse and knowledge of (semantic) context is important in the interpretation of individual sentences. Parallel to the notion of *presupposition*, then, we may introduce the notion of *pragmatic precondition*, defined as a necessary contextual property.

Pragmatic comprehension schemata pertain to the initial context of the verbal communication process, viz. the state which changes by the performance of an act of speech. Such an initial context is not only characterized by the events/actions immediately preceding the speech act, but possibly also by *cumulated* information from 'earlier' previous states and events.

Since however not *all* details from previous interaction states can be stored and retrieved, permanent procedures of *relevance assignment* must be at work to sort out the information which hypothetically will be important in further (inter-)action production/comprehension. It may be assumed that the procedures involved here are similar to those based on *macro-rules* in the processes of semantic comprehension of complex information (e.g. of discourse). 6 In other words, the initial context with respect to which a speech act is to be interpreted contains three kinds of information:

(i)   general semantic information (memory, frames);
(ii)   final state information from immediately preceding events/acts;
(iii)   global (macro -)information about the whole previous interaction structures/ processes.

Since (ii) and (iii) are relevant only for the processing of the actual context, these kinds of information are of the *episodic* kind.

From the remarks above we may conclude that pragmatic contexts are *structured*. More in particular we assume that fast cognitive processing requires that contexts are *hierarchially* structured - as is also the case for semantic (macro-)structures of discourse. This hierarchy is defined in terms of *social structure*: speech acts are integral part of social interaction (we do not usually perform them when being alone).

The hierachical structure of society allows us to determine which units (e.g. institutions, roles, actions) and relations are determined by those of higher *level*. In order to be able to determine whether a speech act is appropriate, we thus should first of all be aware of the most general social *setting* in which the interaction takes place, and then about the more specific or ad hoc particulars of this setting, e.g. actual properties of the speech participants.

Although the social context of speech acts is not the topic of this paper, it should be recalled that the relevant social structure should, as mentally represented,

---

6 The relevance of macro-structures and macro-rules in cognitive processing of complex information has been shown, both theoretically and experimentally, in van Dijk 1975, 1976, 1977a, van Dijk and Kintsch 1977 and Kintsch and van Dijk 1977.

be taken into account when analysing the processes of pragmatic comprehension. The important methodological correlate of this assumption is of course that it is, at this level of analysis, not so much the social situation itself, but the *interpretation/ construction* by the social members of that situation which counts in actual communication. Clearly, this kind of 'cognitive relativism' does not imply that those interpretations have no 'objective' basis. On the contrary, successful interaction requires that the interpretations of social structure are *conventionalized*.

Note that the social context is also an abstract construct with respect to actual social situations. In the first place all those properties are socially irrelevant which do not somehow condition the interaction of the social members: it is irrelevant what I actually think if I do not show my thought in my behaviour; nor is it relevant what I carry in my car if it has no interactional 'meaning'. Similarly for those interactions which are particular to one situation, such as sneezing or smiling to someone during a public lecture or a court session: they do not essentially determine the typical (inter-)actions defining such social settings in general.

Thus, social context analysis begins at the level of general *social context*. This general social context may be characterized by the following categories:

(i)   private
(ii) public
(iii) institutional/formal
(iv) informal

The precise definition of these concepts cannot be given here. Important is only that they define different kinds of social contexts, e.g. public institutions such as courts, traffic, or hospitals, public informal 'places' such as restaurants or buses, private institutions such as families and private informal 'settings' such as making love or beating somebody.

The different social contexts thus globally characterized are in turn defined by the following properties:

(i)   positions (e.g. roles, status, ect.)
(ii) properties (e.g. sex, age, etc.)
(iii) relations (e.g. dominance, authority)
(iv) functions (e.g. 'father', 'waitress', 'judge', etc.)

These properties of social contexts, and their *members*, are systematically related. They define the *possible actions* of the social members in the respective contexts. Social contexts may themselves be *organized*, e.g. by a certain structure of (social) *frames*. [7] Thus, within the general institutional context of a court, there

---

[7] We make a provisional difference between epistemic (cognitive) and social frames, as notions from cognitive psychology and sociology, respectively. Clearly, there are relations between the two notions, because the organization of social interaction has its influence on the organization of our knowledge about social structures. An analysis of social frames has been given by Goffman (1974).

are several frames, which are e.g. chronologically ordered, such as the charge-frame, the defense-frame and the judgement/conviction-frame. In these frames members are assigned specific functions/positions/properties and relations. More in particular these frames regulate which kinds of acts may be performed. In the private institution of a family the parent-function, being associated with a set of properties and relations (power, authority), defines a set of possible social actions, e.g. sending the children to bed, whereas the child function, conventionally, does not allow that a child sends one of his parents to bed. Besides the four defining categories of social contexts, then, we must have a set of *conventions* (rules, laws, principles, norms, values) defining which sets of actions are associated with which positions, functions, etc.

Thus, a convention will determine when in an informal public place a member with a certain property, relation, and function may 'greet' another member. I may not greet anybody in a crowded street, but may do so on a lonely mountain path, whereas I may greet any acquaintance or even people I have just interacted with in specific ways, etc. The complexity of these conventions and their constraints cannot be gone into here.

For our discussion about the contextual analysis taking place during pragmatic comprehension the (very fragmentary and informal) analysis of social context given above suggests that each language user should take into account the following information about this general social context: its specific type, the frame of the context now being relevant, the properties/relations of social positions, functions and the members filling these *categories*, as well as the conventions (rules, laws, principles, norms, values) determining the socially possible actions of the members involved.

Note that the analysis of a *particular* context in terms of the concepts mentioned above is possible only with respect to *general* knowledge of social structure. It is within this broader framework of the analysis of the social context in which the *specific* properties and relations (e.g. actions, interactions) of the speaker may be analyzed, viz. his previous behaviour (doings, actions, etc.) - e.g. the specific things he said before - as well as the *inferences* we make about the *internal structure* of the speaker, in terms of:

(i)   knowledge, beliefs
(ii) wants, desires, preferences
(iii) attitudes
(iv) feelings, emotions

both at 'he particular as well as on the more general level (norms, values, beliefs the speaker also exhibits in other situations). Note that part of the acts involved are instances of the conventional acts belonging to the action sets of some informal or institutional social context and its properties: in the institutional public context of traffic, a person with the function of policeman, has the right to give me the signal

that I should stop – whereby I am, by law, obliged to in fact stop, or else sanctions may follow; thus I interpret the particular signal as an instance of the institutionalized act as described. In other words, the specific signal is an *appropriate* act with respect to the context as defined.

Finally, context analysis as a component in speech act comprehension also involves *self-analysis* of the hearer. In order to understand that the particular speech act is appropriate to him, he must be aware of his own (previous) activities and the underlying knowledge, wishes, attitudes and emotions. Thus if in a particular context somebody offers me help, I must be aware of the fact that I am acting such that such an offer makes sense, and that the offer pertains to a goal which is probably (partly) identical with my own goal. More specifically, the speaker not only has information about the 'world' or the social structure in particular, but also about me, as the coparticipant in communication. The hearer, thus, will have to compare what the speaker apparently assumes about the hearer, with his own self-knowledge.

Although in this section we have been far from complete, and certainly not yet very explicit, we now seem to have a rough picture of the major components of a *schema for contextual analysis* applied by language users in order to evaluate which speech act is accomplished, and whether this speech act is appropriate relative to this context – or relative to this context analysis of the hearer.

It may be assumed that what has been remarked above also may be relevant in a theory of *speech act production*. That is, a speaker may only appropriately accomplish an intended speech act when he believes that the context satisfies the conditions of such a speech act.

6. Examples of context analysis

In order to illustrate informally our levels and categories operating in pragmatic context analysis of language users, we will give two examples of speech acts and a characterization of a context in which they are appropriate. The examples will be taken from both an institutional and a private kind of context.

First, consider the following utterance:

(2) "May I see your ticket, please?"

Although there is set of possible contexts with respect to which this utterance, taken as a request, may be appropriate, only one example will be given, viz. that of ticket inspection on trains, or means of public transport in general:

Social context type: Institutional. Public.
    Institution: Public Transport
    Frame: Ticket Inspection

A. *Frame structure*
    a. *Setting:*   train (during operation)
    b. *Functions:*   F(x): official ticket-inspector
                         G(y): passenger
    c. *Properties:* x has visible signs of being inspector of the (railroad) company; and/or x can identify himself as being an inspector of the (railroad) company;
                         x actually performs his task of ticket inspection;
                         y is the obvious ticket-holder (e.g. not a child travelling with his parent)
    d. *Relations:*   F(x) has authority over G(y);
    e. *Positions:*   (see functions): y is checked by x
                         x is checking y
B. *Frame conventions* (rules, norms, etc.)
    1. Each passenger must have a valid ticket when on means of public transport;
    2. Each passenger must show his ticket upon request to officials of the railroad company;
    3. A passenger which is not in the possession of a valid ticket will be fined $25.00.
    n. It belongs to the duties of ticket-inspectors to inspect tickets (...)

Contextual course of action:
    Macro-action: x takes train to Z.
        Previous acts: -x went to station
                            -x bought a ticket (=> has ticket)
                            -x entered carriage (=> is on train)
                            -x sees/looks at/hears inspector
                            -(x begins search for ticket)

It is roughly in this kind of context that an utterance as (2) may successfully be performed as an acceptable *request*. That is, the hearer will not only conclude from the form of the utterance itself (see below) that it is a request, but also that the request satisfies the specific request conditions as well as the *grounds* for these conditions, as defined by the social structure. Only those hearers who have the information available about the context as specified will be able to judge whether these grounds are sufficient for the acceptability of the act of requesting, and hence whether the request should be complied with or not.

Thus, if one of the contextual features does not obtain the request may become spurious and hence socially unacceptable (at least to some degree). If I am not on the train, no request for my train ticket may be made; similarly if the inspector has neither a uniform nor identification: I need not show my ticket to any passenger requesting so; etc. Note that the frame conventions are the basis from which speci-

fic knowledge, beliefs, duties/ obligations, etc. may be derived. Hence, passenger *x* (= hearer) knows that by regulation (law) B.1. he should also have a ticket, a knowledge which is a condition for actually buying the ticket, and for feeling the obligation to show it to an inspector, when requested (as by B.2).

As a second example we take a speech act from an informal private context:

(3) "Let me carry your suitcase!"

performed in a situation in which the speaker (*x*) takes the hearer (*y*) from the train, thereby *making an offer* to carry the hearer's luggage. Clearly, such an offer would be appropriate in many kinds of contexts, because there are no institutional restrictions on the interaction as in the first example.

Social context type: Informal. Private.
    Frame: Meet somebody (at arrival)
      A. Frame structure
          a. *Setting:*   Railway station. Platform.
          b. *Functions:* $F(x)$: host
                                 $G(y)$: guest
          c. *Properties:*   *x* is strong enough to carry suitcase
                                 y may be tired and/or
                                 y has difficulties wearing suitcase, and/or
                                 y is a woman or old person
                                 (. . .)
          d. *Relations:*   *x* knows y (and y knows *x*)
                                 (*y* knows that) *x* is there to meet *y*
          e. *Position:*   *x* is helper of *y*
      B. *Frame Conventions*
          1. If *x* meets y at arrival, *x* should be friendly to *y*;
          2. If *y* has difficulty (or may have difficulty) in accomplishing a task, *x* should help *y* (or offer to help *y*);
          3. Heavy luggage is a sufficient reason for offering help;
          4. Guests, and especially women, tired, ill or old persons should be offered help.
          (. . .)

Contextual course of action:
    Macro-action: Arrival of *y*. Y being met by *x*.
        Previous acts: - y prepared for arrival
                              - y got out of train
                              - y looked for x
                              -(...)
                              -*y* walked to *x*

-*y* smiled to *x*
-*y* greeted *x*
-(...)

This categorized list, then, is assumed to represent the major components of a schema necessary in the comprehension (and acceptance) of (3) as a request. Note that since we specify the knowledge which must be available for the adequate *comprehension* of speech acts, the acts are described, as in the previous example, from the *point of view of the hearer*.

In case the major features of this context should not be satisfied, the request may become inappropriate, or it may become not complied with. Thus, if I do not wait for somebody, does not know him/her, etc., I may not usually offer for help, according to the frame conventions. Similarly, a strong man will usually protest against the offer if this is made by a child or a weak person in general.

The categories used above are not definitive, and require further definition. Especially the difference between the *function* and the *position* categories are not always clearcut. By a function we understand a more or less fixed, often institutionalized, set of properties/ relations, e.g. professions, e.g. judge, ticket-inspector, etc. but also host, guest, etc. Positions on the other hand more specifically define the kind of relation operating in the frame, between the participants, i.e. they define their actual role in interaction: I may help, advise, prohibit, etc. somebody else, but may do so 'in' different functions, as a doctor, a friend, a mother.

7. The 'dynamics' of context

Since actions are accomplished in contexts, these contexts are not static, but dynamic: they change according to causal principles, conventions and other constraints on sequences of events and action. We have seen above that a speech act is usually embedded in such a course of (interaction-). Contextual analysis, thus, is a permanent *process* – of which we have considered the 'final state', being the initial state for the adequate comprehension of an utterance as a speech act. During (inter-)action a person will gradually *construct* the relevant features of the contexts and the changes to subsequent context states. He will not only be aware of the global characteristics of the social context, the frame and its components, but also of the actions actually performed in these contexts/ frames. He thus will construct the assumed purposes, intensions (and underlying knowledge, wishes) of the co-participants, in particular the speaker, and judge whether the purposes of the speech act are compatible with these initial conditions.

8. Utterance analysis

Although the analysis of the context not only provides the conditions with respect to which an utterance should be evaluated but also provides *expectations*

about plausible goals of participants and hence about possible speech acts they might perform in this context, it is clear that in most cases the final comprehension of an utterance as a specific speech act must be based on an analysis of the *utterance* itself. The question, then, is which typical properties of utterances indicate which speech acts has been performed – or may be counted as performed?

The answer to this issue should of course be based on the knowledge we have about the 'grammatical' processing of words, phrases, sentences and discourse. It is not the aim of this paper, however, to review the theoretical and experimental/simulative results with respect to the rules and strategies involved in phonological, morphological, syntactic and semantic processing of natural language. We are only interested in those features of these respective levels which may have the particular function of 'illocutionary act indicating devices'. Besides the information from the context, the text will provide the hearer with the clues about which act is to be assigned to the utterance.

Since there are some reasons to believe that the 'evaluation' of these indices in the utterance depends on the full interpretation of the utterance (e.g. in terms of clauses/propositions), we will start with the *semantic* level.

Semantics
    Reference: Denotation of speech participants (I, you);
                     Denotation of contextually present objects;
                     Denotation of properties of the context, and relations between paticipants;
                     Denotation of states, events, actions;
                     Denotation of modalities: time, possible world, obligations, etc. etc.

Thus, in our example "*Let me carry your suitcase*", the semantic structure of the utterance denotes both the speaker and the hearer, a (planned) action of the hearer, a (requested) permission, and the object of the action (the suitcase). Hence, semantically the hearer now knows that the utterance relates to an immediate action of the speaker with respect to an object of the hearer. Given the frame information (and further world knowledge), the hearer may interpret such an action as an helping action. Since there has not been request for help, permission must first be obtained, hence the reference to such a permission.

Syntax
    1. sentence forms (indicative, interrogative, imperative)
        E.g. *Let me* (. . .) in our example, being a 'formulaeic' expression for obtaining permission as part of offers (of own actions). In general indicating speech acts involving wanted (future) actions of hearers (orders, requests, etc.);
    2. word order; clausal structure of sentence;

Indicating presupposition—assertion or topic—comment structures, and hence what is supposed to be 'given' and what 'new' information; indication of possible topic of discourse;

3. syntactic functions (subject, indirect object, etc.)
    Indicating semantic functions and thereby roles of participants in the context;

4. tense
    Indicating when event or action takes place or will take place; present tense in performatives;

5. aspect
    Indicating the mood of the action, and thereby possibly the attitude involved in the speech act;
    E.g. I promise you to come tomorrow! vs. I am promising you to do it!

6. sentence/sequence structure
    Indicating the delimitation and ordering of speech acts, or whether one speech act is performed on the basis of a complex proposition, or two speech acts on single propositions.

Morphology/Lexicon (- semantics)

1 . word choice (in general)

     Use of certain words directly indicating the concepts involved and hence the possible referents of the expressions; in particular: deictic expressions, action predicates;

2. explicit performatives

3. fixed phrases (tag questions)
    E.g. *won't you, can't you; isn't it,* etc. indicating conditions (ability, truth, etc.) of speech acts;

4. words like *please* in our example, indicating requests, etc. 5. pragmatic particles (especially in Dutch, German, etc.) Indicating e.g. attitudes of the speaker with respect to propositional content and/or hearer, e.g. in accusation, reproach, defense, etc.

Phonology/phonetics

1. Intonation Together with sentence form and word order indicating e.g. stating or questioning;

2. Stress
    Indicating comment function, and thereby new information in sentence, or contrast, and thereby expectations of the hearer, or focus and thereby controlling the attention of the hearer;

3. Speed, pitch, loudness
    Indicating attitude, emotions, etc. of the speaker and thereby the respective speech acts connected with them, e.g. a warning vs. an assertion, as in "A stone is *falling*".

Paralinguistic activities
   1.   deictic movements
         E.g. pointing to the suitcase in our second example.
   2.   (other) gestures
         E.g. touching one's head in insults, accusations.
   3.   facial expressions
         E.g. smiling in greeting, angry face in accusations;
   4.   bodily movements
         E.g. drawing somebody away during a warning;
   5.   body-interaction (closeness, shaking hands, caressing)
         E.g. in greetings, declarations of love, fighting.

Although this list is certainly not complete, nor very explicit, it seems to show that at all linguistic and para-linguistic levels of the utterance we have a large amount of indications about certain features of the possible speech act involved. It is obvious that none of the indications is as such, in isolation, sufficient to establish certain speech acts. All levels must interpretatively be integrated, and connected with the contextual analysis. Thus, our second example yields further information about the possible (speech) act performed, viz. reference to a permission of the hearer, reference to speaker and hearer, reference to an immediate action of the hearer, reference to an object of action, a requesting sentence form (formulaeic), and possibly deictic movements (looking at, pointing at suitcase). Since the hearer knows that the action of the speaker is beneficiary to him, and knows from the context that all other conditions (including the referential ones) are satisfied, the hearer will by conventional knowledge interpret the utterance of (3) as an offer.

9. Comprehension of speech act sequences and macro-speech acts

In the analysis of the context given above it has been left open whether the initial context was brought about by the communicative participants, or by others, by events or by actions. In a similar way, it may be the case that the initial context is the final state and consequence of previous speech acts. Speech acts may occur in *speech act sequences*, typically so in conversation.

Thus, an assertion may bring about in the hearer the knowledge to understand why a threat or a promise is following it. Much in the same way as the comprehension of discourse as a sequence of connected and coherent sentences, requires an interpretation of these connections (e.g. all kinds of conditional relations between denoted facts), the comprehension of speech act sequences is based on the interpretation of the 'connection' between the subsequent speech acts. The most general connection condition has just been mentioned: previous acts establish the context with respect to which a subsequent speech act is evaluated. As for act sequences in general the conditional relation may be one of mere 'enabling', i.e. making a speech

act *possible* with respect to a previous speech act, or have a stronger nature, e.g. making a following speech act likely or even necessary. The latter case arises in all kinds of ritualized sequences, e.g. in such pairs as (congratulating, thanking).

Another kind of speech act relation is given when some subsequent speech act is intended as an *addition*, a *correction,* or *explanation* with respect to a previous speech act.

Speech acts occurring in sequences need not all be of the same 'level'. Just as in discourse or sentence, we might distinguish between 'subordinate' and 'superordinate' speech act relations, e.g. when some speech act is an *auxiliary* act with respect to another speech act, e.g. when I am stating that I am hungry in order to be able to appropriately accomplich the act of a request for food.

Finally, speech act sequences may also be analysed at a *global* level. This means that the sequence of speech acts is mapped *as a whole* onto one (or more) global speech acts of *macro-speech acts.* [8] Thus, a whole letter may globally function as a threat, a whole law as a prohibition. With respect to the individual speech acts of the sequence, such a pragmatic macro-structure is a kind of 'reduction': it defines what is the 'upshot' of an utterance, e.g. in terms of the global intention and purpose. The individual speech acts hereby function either as relatively 'irrelevant' aspects of the communication (e.g. greetings), or may be considered as normal conditions, components or consequences of the global speech act. *Macro-rules* specify how a speech act sequence is related with its global representation in terms of macro-speech acts.

Cognitively these assumptions imply that in contextual and sequential analysis a hearer will apply macro-rules to previous speech acts in order to be able to determine what the *relevant* pragmatic information is for the interpretation of subsequent speech acts. Thus, having understood a series of speech acts as a request, the hearer may himself plan the act of a refusal, whereby all kinds of details of the request may have been forgotten already. In this respect the global understanding of speech act sequences is not much different from the understanding of the other properties of the context (perception of objects, states, events, etc.)

One of the important functions of macro-comprehension is also that the speaker/hearer is able to establish connections between speech acts *with respect to* a macro-speech act. If not, we would not be able to plan and monitor a long discourse or conversation, nor would we in comprehension be able to understand what the speaker is 'doing' a11 the time. Thus, a series of greetings, assertions, questions, etc. may be interpreted as really 'belonging together', i.e. as a coherent sequence, only if we make the assumptions that *thereby* the speaker is making a global request, of which the respective 'micro-speech acts' establish the necessary preparations, conditions and components.

The specific properties of speech act sequences, and their relations with

[8] For an analysis of speech act sequences and macro-speech acts see the references given in note 4 of this article.

sequences of sentences in discourse will not further be analysed here. At the sequential (linear) level the comprehension process is not fundamentally different from that of the context in general. More interesting though is the fact that as soon as the speech act sequence becomes too complex, macro-processing is necessary.

10. Psychological process assumptions [9]

The account given above of speech act comprehension is not only theoretical, but also is an abtraction in another way. That is, we have enumerated a number of elements which are part of an assumed cognitive representation of the pragmatic context. It has been stressed however that both the context itself as its comprehension are 'dynamic', in the sense of requiring *processes*. In other words, we do not yet know *how* the representation of the context (and that of the text) is actually constructed during interaction. We do not know how exactly information from perception is combined with all kinds of inferences, the actualization of frames, or how all kinds of input information is organized, stored, combined with existing knowledge, or with wishes, emotions, attitudes, intensions or purposes. And finally, we do not know yet how all this 'external' and 'internal' information is mapped onto representations of the social context categories and structures.

Although all these questions are at the same time queries for future research on the cognitive basis of pragmatic structures, a certain number of assumptions may be derived from our discussion above.

First of all, the incoming information in communicative interaction is so amazingly complex that we need processes of different kinds which enable us to handle and control it. This means for instance that *macro-rules* must be applied both at the level of perception and on the level of event, act and language comprehension. That is, information must be *functionally organized and reduced*.

This organization, as we have seen, must be of some kind of *hierarchical* nature. Given some typical social context, it is determined what kind of frames may be involved, which in turn determine the possible positions, functions, properties and relations of the social members. The essentially conventional nature of social interaction, furhtermore, will allow the formation and hence the activiation and application of *epistemic frames*, as effective ways to organize our more general knowledge of the social frames involved.

At the same time the incoming information is not only organized but also reduced to macro-structures at several levels. A sequence of actions of a participant may thus be 'seen as' one global act, of which only the globally intended result may be relevant, as well as the global goal as 'purposed' by the participant. This macro-reduction is again possible due to the conventional frame knowledge, which specifies which global acts are associated with which conditions, components and consequences. The hierarchical structure of the epistemic frames, then, would allow a flexible and fast analysis of the levels and components necessary in the comprehend-

---

[9] For a detailed analysis of complex information processing, see Kintsch and van Dijk 1977.

sion of actually ongoing social interaction. The same principles hold in the comprehension of visual information ('images') and of complex event episodes. At a higher level, we also have general knowledge about the structure of action, so that each sequence of actions may be mapped onto a global action, abd we are able to focus attention on relevant results and consequences (goals) of action.

Finally, besides these various procedures in the analysis of complex information, language users apply all kinds of comprehension *strategies*, yielding fast hypotheses with respect to the most plausible structure of the actual and following context states. These strategies are based on 'indices' or 'key features' of text and context, e.g. question intonation in requests, fixed phrases, certain gestures, and of typical (ritualized) actions in certain frames, as in the case of the offer in our second example. Although some very general principles of speech act comprehension and of the cognitive processes underlying interaction have briefly been mentioned, it is obvious that insight into actual processing is virtually nonexistent at the moment. We may devise theoretical models for the schemata or procedures used by language users in the matching of context analysis with text analysis, and we may try to test such models empirically by systematic experiments and/or by artificial simulation of communicative interaction, but such models will at first be of the 'input-output' type, and will not yet provide insight into the actual processes.

Similarly, little is known at the moment of the cognitive (and emotional) 'underlying structures' in the *planning, execution* and *control* of complex speech interaction, and of the processes of storage, retrieval, recall, reconstruction, and re-application of speech act representations. That is, we do not know yet how wishes, wants or preferences interact with knowledge, beliefs and attitudes, and how all these systems lead to actions via complex procedures of decision making, purpose-formation and planning. In this respect the open problems of cognitive pragmatics are those of an empirical theory of action in general. The additional difficulty lies in the fact that in speech act production and comprehension we also need insight into the complex relations between the contextual analysis and the textual analysis. Clearly, we are only in the first stage of understanding the complex interdisciplinary domain between language, action, meaning, cognition, and social structures.

References

Bartlett, F.C. 1932. Remembering. London: Cambridge University Press.

Bobrow, Daniel, G. and Allan Collins, eds. 1975. Representation and understanding. New York, London: Academic Press.

Charniak, Eugene. 1972. Towards a model of children's story comprehension. Ph. D. Diss. MIT, Cambridge, Mass.

Dijk, Teun A. van. 1975. Recalling and summarizing complex discourse. University of Amsterdam, mimeo. In: W. Burghardt and K. Hölker, eds., Text processing. Papers on text analysis and text description. Berlin, New York: de Gruyter. In press.

Dijk, Teun A. van. 1976. Complex semantic information processing. Paper, International Workshop on Documentation and Linguistics, Stockholm. University of Amsterdam, mimeo. [The paper will appear in the Proceedings of the workshop.]

Dijk, Teun A. van. 1977. Text and context. Explorations in the semantics and pragmatics of discourse. London: Longmans.

Dijk, Teun A. van. 1977b. Macro-structures, knowledge frames and discourse comprehension Paper for the XIIth Carnegie-Mellon Symposium on Cognition, Pittsburgh, 1976. In: P. Carpenter and M. Just, eds., Cognitive processes in comprehension. Hilldale, N.J.: Erlbaum.

Dijk, Teun A. van. 1977c. 'Pragmatic macro-structures in discourse and cognition'. Paper, International Workshop on The Cognitive Viewpoint, Ghent. In: M. de Mey et al., eds., CC 77. International workshop on cognitive viewpoint, 99-113. Ghent: Communication & Cognition.

Dijk, Teun A. van, and Walter Kintsch. 1977. 'Cognitive psychology and discourse'. In: W.U. Dressier, ed., Current trends in text linguistics. Berlin, New York: de Gruyter.

Goffman, Erving, 1974. Frame analysis. New York: Harper & Row.

Kintsch, Walter, and Teun A. van Dijk. 1977. Toward a model of discourse comprehension and production. University of Colorado/University of Amsterdam, mimeo.

Minsky, Marvin. 1975. 'A framework for representing knowledge'. In: P. Winston, ed., The psychology of computer vision, 211-277. New York: McGraw Hill.

Teun A. van Dijk

Born 1943 studied I French language and literature, and then theory of literature, at the Free University and the University of Amsterdam, respectively. Did research, and published several books and many articles, on linguistic poetics, semantics and pragmatics, focussing in particular on the grammar and theory of discourse. Teaches discourse studies at the University of Amsterdam. Recently engaged him self in research on the psychology of discourse processing.

Publications

1972. Some aspects of text grammars. Ph.D. dissertation. The Hague: Mouton.

1975 (ed.). Pragmatics of language and literature. Amsterdam: North-Holland.

1977. Text and context. Explorations in the semantics and pragmatics of discourse. London: Longmans.