

# PARSING AND GENERATING THE PRAGMATICS OF NATURAL LANGUAGE UTTERANCES USING METACOMMUNICATION

David L. Sanford & J. W. Roach  
Department of Computer Science  
Virginia Tech  
Blacksburg, Virginia 24061

## Abstract

This paper reports a new theory of natural language processing and its implementation in a computer program, DIALS (for DIALOGue Structures). This represents a radical departure from the paradigmatic approach to natural language processing currently dominating the fields of artificial intelligence, linguistics, and language philosophy, among others. We use the theory of metacommunication to develop a "pragmatic grammar" for the structural analysis of dialogue. We are currently able to parse and generate over 5000 surface forms of a single underlying request content. We propose using this pragmatic information to manage the communication context, including inferring some of a speakers' goals and controlling status and politeness.

*Keywords:* discourse analysis, pragmatics, metacommunication, requests

## Introduction

Natural language processing has traditionally concentrated on syntax or semantics. Our new theory, called "Dialogue Structures," concentrates instead on pragmatic issues. This theory posits that indirect questions, emphasis, focus and speakers' goals, still problematic issues after years of research, can be determined, in part, by structural means. Every utterance consists of both a semantic, content portion and a communication management portion (Roach & Nickson, 1983, 1986). Roach and Nickson independently rediscovered the theory of metacommunication originally developed by Bateson (1951a, 1951b) and his successors (Watzlawick, Bavelas, & Jackson, 1967). Communication management, using metacommunication, guides the listener's attention, shows importance of topics, maintains a shared communication context and preserves social relationships (Sanford & Roach, in press) such as politeness, status, etc. Being able to process the pragmatic content of utterances leads, for the first time, to a system that can account for some of the tremendous range of expression in natural language. For example, the system we have built, called DIALS, can correctly generate and parse the request, "What time is it?" in more than 5000 different ways. DIALS uses a "pragmatic grammar," similar in concept to the well-known semantic grammar, to process natural language. Unlike systems that use semantic grammars, its domain is general rather than specific; i.e., DIALS' task domain is the management of communication, necessary in all dialogues. We do not claim that syntactic and semantic methods are wrong, only that one of the most important problems in parsing has been largely ignored. "Dialogue Structures" speaks, at least in part, to this problem. This paper will discuss the application of metacommunication to the understanding of the pragmatics of requests, an important cognitive task that humans learn to parse and generate correctly at a very early age.

Making requests is a common task in all types of communicative activities. Research on requests is often limited to question-answering (Lehnert, 1984) and often makes overly simplified assumptions about how communication works. Those studying language analysis from the viewpoint of linguistics, language philosophy, and related fields, seem to adopt a model of communication that says that language is explicit, that people say what they mean and mean what they say. Grice's (1975) well known Cooperative Principle, including the Manner maxims, express this "Transparent Model" of communication: speakers "avoid obscurity of expression," "avoid ambiguity," and try to "be brief."

Communication theory (Sanford & Roach, 1986) disagrees with the transparent model, arguing that people do not say exactly what they mean and seldom mean what they do say. Indeed, this “Guarded Model” of communication recognizes that speakers have hidden agendas; as Goffman (1959) says, one important goal in interactions is to save face, for which people use masks or façades. The transparent model to communication overlooks the existence of deception and face-saving. This paper will not discuss deception to any degree, but at least the guarded model recognizes its possibility. This paper will discuss how the theory of Dialogue Structures applies to requests, identifying the two major parts of a request, analyzing each part separately, and finally discussing DIALS, the computer program that implements the theory.

## Theory of Dialogue Structures

According to Bateson, there are at least two subcategories of metacommunication: “the propositions about codification [i.e., communication in which the content of the utterance is the process or mechanisms of communication] and the propositions about interpersonal relationship” (1951b, p. 214). Dialogue Structures deals with the subtle expression of interpersonal relationship from the metacommunicational cues of the form in which requests are phrased. That is, a speaker’s intentions cannot be directly identified but must be inferred from subtle metacommunicational cues and the interpreter’s knowledge of social norms and interpersonal relationships.

Dialogue Structures states that the surface form of a request is not simply an expression of pragmatic purpose, politeness, or clarity. The main thing expressed explicitly in a request is how demanding vs. how pleading a request is, i.e., the “imperative force” of the request. A request, therefore, can be represented in the following way:

$$\text{request} = (\text{content expressing a desire}) + (\text{structure expressing imperative force})$$

We shall define what we mean by “structure” in the next section. Dialogue Structures, then, must explain the two parts of a request: the structure expressing imperative force and the content expressing a desire. The following will examine separately these two parts of a request.

### *Structure Expressing Imperative Force*

“Imperative force” expresses how demanding vs. how pleading a request is. “Imperative force” was first used by Searle (1975) to express that a request is demanding; we expand it into a complete dimension of expressive power. That is, Searle said only two types of requests have imperative force: the explicit performative (e.g., “I order you to leave the room”) and the flat imperative (e.g., “Leave the room”); we say that all requests have imperative force, but it ranges through many gradations from strongly demanding to strongly pleading.

Most people use the term “structure” in this context to refer only to the syntax of an utterance. Certainly syntax is a structural component of utterances, but we use “structure” in a broader sense. We identify structural aspects to the pragmatics of utterances and use this structural level as well. Structural pragmatics involve the identification not only of the syntax of an utterance, but also of key word patterns; e.g., “I was considering . . .” is not only a declarative syntactic structure, but the key words identify it as a “Claim of Deliberation” in which any desire may be embedded. (Throughout this discussion, several request categories will be mentioned. There is not enough room in this report to give the full category system. See Sanford and Roach, in press.) For example, consider trying to identify the “transparent” purpose of a request to decide how to respond. We argue that a listener cannot tell the true pragmatic purpose of a request from the surface form. However, the request may structurally appear to be a request for information, for permission, or whatever. Therefore, although a request beginning, “Do you know . . .” sounds as if it is intended as a request for a yes/no response, we cannot tell without making inferences from subtle metacommunicational cues and the interpreter’s knowledge of social norms and interpersonal relationships. But we can identify it “structurally” as “Asking for Suggestion,” and

then investigate how demanding vs. pleading it is. A clear understanding of imperative force, therefore, is the first step in constructing an overall theory of requests.

This brings us to explaining how imperative force functions in communication. Imperative force may express politeness, status, or emphasis. Searle (1975) asserts that the choice of how to phrase a request is based solely on how polite an individual wants to be. Examining such transparent model approaches to requests would lead one to believe that politeness is most important, status is least important, and emphasis has no relation to the form of requests. Actually, the order is exactly the opposite, according to Dialogue Structures. We will first discuss the relation between politeness and status, then discuss the relation between status and emphasis, and finally present our analysis of how a hearer uses imperative force to interpret a speaker's utterances.

Status is more fundamental than politeness. Consider how the transparent model explains impolite behavior. Since people try to be clear, direct, unambiguous, and cooperative, the assumption that people would be impolite is untenable. Therefore, some other contrast to politeness is necessary. Lakoff (1973) says that the more direct the wording of a request, the more impolite but the clearer that request is; his contrast, therefore, is between politeness and clearness. Research shows this not to be the case. When Gibbs (1979) measured the time taken to interpret indirect requests embedded in a story context, it actually took longer to understand the literal than the indirect meanings. Also, research shows that children about two to three years of age have no more difficulty with indirect than with direct requests (Elrod, 1983; Shatz, 1978). Direct wordings of requests, therefore, are no more clear than indirect wordings.

Searle (1975), as mentioned, asserts that the request form is based solely on politeness. Research shows this to be an unjustified assertion:

The politeness of the directives used by a speaker appears to be affected directly by the status relationship between the speaker and the listener. Studies with adults have shown that polite request forms such as "May I please use your phone?" are more likely to be addressed to a listener whose age or professional position places him in a superior role. (James, 1978, p. 308)

This suggests that status is fundamental to determining politeness. Hill et al. (1986) refer to two components of politeness: "discernment," which involves "conforming to the expected norm," and "volition," which allows a speaker a "more active choice." We would say that behavior under the control of norms is deference, whereas politeness is always volitional. For example, a sergeant has more status or power than a private. If a private uses a pleading request to a sergeant, the private is not being polite but is expressing deference to the power of the sergeant. If the sergeant uses a pleading request to the private, the sergeant is being polite, since the sergeant can make a demand of the private and chose to be less demanding. It is inappropriate, therefore, to apply the terms "politeness" or "clearness" as the basic dimensions of comparison of requests. This evidence strongly supports the idea that "imperative force" is a metacommunicational cue that fulfills some other function(s) in the communication of requests. Indirect requests are patterns expressing metacommunicational information in which any type of desire may be embedded. This agrees with everyday observations: if someone asks, "Do you know the time?" and receives the response, "Yes," they assume the respondent is joking or being uncooperative. Requests are demanding vs. pleading; the status of the interactants and the emphasis being expressed determine whether a request is polite.

Emphasis is more fundamental than status. To return to our example, during battle a private can address an imperative request, such as "Pass the ammunition!" to a sergeant without being considered impolite or incurring the wrath of the sergeant. The emphasis on the importance of the desire in this context precludes any consideration of status or politeness. Indeed, it might be considered "impolite" to use a pleading form, since it takes more words to express pleading, and the extra time might be life or death.

To explain more fully this issue of the emphasis expressed with a request, consider the following two dialogue examples between a customer and an airline reservation agent:

### EXAMPLE 1

**CUSTOMER:** I need to go to L.A. Do you know if I can leave town today?  
**AGENT:** I'm sorry, but all flights to L.A. today are booked. I could get you on a flight tomorrow.

### EXAMPLE 2

**CUSTOMER:** I need to leave town today. Do you know if I can go to L.A.?  
**AGENT:** I'm sorry, but all flights to L.A. today are booked. I could get you on a flight to Dallas/Ft. Worth.

Using the analysis of transparent models, the first sentence of each customer is fairly clear, has no imperative force, and is less polite, whereas the second sentence is less clear, has no imperative force, and is more polite. Most people would neither interpret these requests as mixing impolite and polite forms nor say there is no imperative force to the requests. The wording puts emphasis on the information being expressed, allowing the hearer to infer the speaker's goals. Consider how inappropriate it would be for the agent to offer the first customer a flight to Dallas/Ft. Worth; consider how equally inappropriate it would be for the agent to offer the second customer a flight on the following day.

Determining the function of imperative force in a given request. Several types of information are needed to decompose the relative importance of the three determinants of imperative force: emphasis, status, and politeness. A participant needs to know standard status levels for established societal roles, such as teacher vs. student, boss vs. worker, etc. For an ongoing relationship, one needs to know the history of status negotiation within this given relationship. For a given dialogue, one needs to know the sequence of status negotiation moves across this interaction. For example, within an interaction, status is usually the first issue addressed in a sequence of dialogue moves. This is often shown by the phrasing of the pre-requests (Jacobs & Jackson, 1983), or what we term the "empty requests," by the requestor. Two excerpts taken from transcribed tape recordings of actual interactions between airline reservation customers and agents show what we mean.

### EXCERPT 1

**CUSTOMER:** I'm planning on a flight leaving April 5. I plan on leaving from Roanoke. I'd like to go to L. A. I was wondering if you could give me some flight information about that.

### EXCERPT 2

**CUSTOMER:** I need some information. I would like to get some information about taking a flight leaving April 5 from Roanoke to Los Angeles.

In the first excerpt the customer emphasizes the date and airport of departure with slightly less emphasis on the airport of arrival. The last sentence is empty of any information important or relevant to the request, and would probably be considered a pre-request if it were in the first position. In the second excerpt, the customer starts with an empty request, this time more easily defined as a pre-request. Then the customer shows that all the pieces of the request are of equal emphasis, expressed with a lower level of imperative force. The speaker is using a less demanding request form than status allows, expressing politeness and the willingness to allow any of the three factors in the request to be adjusted by the agent, as circumstances require.

The process for decomposing the purpose of the imperative force of an utterance now can be clearly stated. First, the hearer determines the relative level of imperative force being used. It is well known that some people habitually express everything as if it were vitally important, while others express everything they say as if it were unimportant. That is, some peoples' range of expression stays among the strongly and moderately demanding forms, while others stay among the strongly and moderately pleading. A hearer must identify the surface level imperative force for a particular utterance and compare it to the range of imperative force used by this speaker in the past. This is one reason we choose to call

our approach "Dialogue Structures," since imperative force is a structural aspect of utterances that must be examined over the course of dialogues and not just at a single point in time.

Second, the hearer checks the defined status between the speaker and hearer, originally by examining the roles of the two and the socially-defined status levels associated with those roles; then, by examining the history of status between the interactants, if there has been an on-going relationship; and finally, by examining the status negotiation, if any, during this interaction. At this stage, the speaker's referent power or the feelings of friendliness between the interactants is probably less important than the speaker's legitimate authority. Certainly there is a coloring effect from referent power and friendliness; one can easily imagine a policeman barking orders at a driver to help the driver avoid an accident (strongly demanding to express the emphasis on the importance of the information), while the driver interprets the imperative force as a status claim: "that stupid cop is yelling at me just to show how much power he has!" Here, again, we see the importance of watching this structural dimension across interactions rather than simply focusing on this particular instance of imperative force, i.e., watching dialogue structures. Policemen are trained to use their voice to achieve social control; most of their utterances will be demanding and therefore this particular utterance will not be extremely out of line with their dialogue patterns. After the hearer has determined the speaker's purpose in choosing a given level of imperative force, the speaker's referent power or the level of friendliness between the participants is much more important in helping to determine what the hearer will decide to do. One again can imagine a case where a person has made a pleading request, e.g., for the time, and the hearer correctly interprets the pleading as an attempt to be polite but decides not to answer based on the hearer's feelings of dislike for the speaker.

A hearer is ready to infer the purpose of the speaker now that the hearer has these three pieces of information: first, the imperative force of this particular request, second, the relative status of the participants, and third, the range of imperative force used by the speaker across a history of dialogue. Again, let us use an example:

STUDENT: I can't turn my paper in today.  
TEACHER: Would it be convenient for you to turn it in tomorrow?

How does the student interpret the teacher's request? The surface form of the request is an "Asking about Convenience," which is moderately pleading. The student knows that teachers have higher status than students, so it cannot be a status claim. Say the teacher has a history of using the entire range of imperative force in expressing requests, so this is nothing unusual in that regard. The student can refer to the semantic domain of discussion, realize that teachers consider class assignments important, so it cannot be emphasis; therefore, it is politeness.

Take the same interaction under another circumstance. Say the student had requested to turn in an early draft of the paper for preliminary review, but the paper is not really assigned to be turned in today or tomorrow. The student could rightly assume the teacher is expressing the relative unimportance of getting it in tomorrow, expressing emphasis. Or take the same interaction with another teacher. Say this teacher has a history of using mainly demanding request forms with students; that makes this pleading form unusual. It takes on a reverse effect and becomes a highly impolite, sarcastic demand.

We are just starting to investigate the rules for combining these three types of information to make the kind of inferences exemplified above. We already have the rules needed to identify the first type of information, the imperative force for a particular utterance. This is also the raw data for the third type of information, the imperative force used by a speaker across a history of dialogue. The second type of information, the relative status of given societal roles, will be stored in the database of the pragmatic grammar for use in these inferencing procedures.

### *Responding to the Content that Expresses a Desire*

The second issue needing explanation is the appropriate response to a request; whether to respond with information, action, confirmation, or whatever. Consider the method for making this determination

using the transparent model: identify the intention of the speaker and respond to that intention. From the perspective of the guarded model, the decision is determined by the information and situation of the hearer, since one cannot trust the speaker to identify clearly the speaker's intention. Take an example from an interaction between an airline agent and a customer. A request such as, "Can you book me for the 1:45 flight?" sounds as if it is intended as a request for a yes/no answer. If there are seats available at that time, most agents would respond not with "yes" but with the action of booking a flight. Information is needed only when the circumstances prohibit action. For example, Allen (1983) grapples with the issue of whether the following is a yes/no request or a request for information:

Do you know when the Windsor train leaves?

If the one being asked knows the information, then a sensible response might be:

7:14 in the morning.

Only if the person cannot provide the information that is indirectly requested should something else be given, for example:

I don't work here.

But in neither case is a "yes" or a "no" needed. Research on people shows that they often go ahead and answer the surface question, but add the information indirectly requested. That is, they do not always include the "yes" or "no," but do so quite often. This is not ruled out by our theory; certainly, we do not lose the information that the surface form is a request for a yes/no response. We simply assume that the surface form is not the speaker's main goal and go immediately to the embedded indirect request. There are cases in which people truly want a yes/no response, but these are cases that people have difficulty identifying unless that desire has been explicitly stated. One is reminded of the many times Perry Mason had to tell a witness, "Answer only 'Yes' or 'No.'" Otherwise, the witness invariably wanted to respond to the indirect question embedded in the surface form.

### *Computational Results*

We have constructed a parsing and generation system, called DIALS, that embodies aspects of the theory presented above. DIALS currently can handle individual sentences, separating the pragmatic from the semantic portions of a sentence and working with a caseframe of the pragmatic portion of a request. That is, it identifies the originator of the request, the proxy verbalizing the request, the receiver of the request, and the imperative force of the request, as well as the goal of the request, expressed in the semantic portion of the sentence. DIALS either starts with a surface form sentence and parses into this caseframe or starts with the caseframe and generates a surface form sentence. It handles indirect requests easily, without first determining a surface meaning then applying an inference engine to infer the indirect meaning, as suggested by Searle (1975), Allen (1983), and others assuming the transparent model. DIALS is independent of any semantic domain and therefore can be applied as a front end to any task, such as databases, operating systems, editors, etc. DIALS has been implemented in PROLOG; on a VAX 11/785, with an interpreter running at 1 klips, a pragmatic parse also producing a simplified sentence for analysis by a semantic parser requires between one and five CPU seconds.

### **Conclusion**

We realize our approach clashes directly with the paradigmatic artificial intelligence approach introduced by Charniak (1972) and exemplified by Allen (1983); this approach says that all pragmatic inferences must be made using an immense world knowledge base. But DIALS already can parse and generate over 5000 different wordings of a single underlying request content. It has been tested by comparing its coding of 547 sentences taken from transcribed recordings of conversations between airline customers and agents against human coding of the conversations; it was able to code the imperative force of the requests correctly over 95% of the time. It is still undergoing improvement aimed at achieving

100% correctness and at being able to parse and generate tens of thousands of surface forms of single underlying request contents. Also, it is being improved to allow it to keep track of the pragmatic structure of dialogue so the rest of the theory discussed in this paper can be implemented. But our preliminary successes make us feel that the "guarded model" of communication and metacommunication may have advantages in many pragmatic issues over the "transparent model" and its attendant need for an immense world knowledge base.

## References

- Allen, J. (1983). Recognizing intentions from natural language utterances. In M. Brady & R. C. Berwick (Eds.), *Computational models of discourse*. Cambridge, MA: M.I.T. Press, 107-166.
- Bateson, G. (1951a). Information and codification: A philosophical approach. In J. Ruesch & G. Bateson, *Communication: The social matrix of psychiatry*. New York: W. W. Norton, 168-211.
- Bateson, G. (1951b). Conventions of communication: Where validity depends upon belief. In J. Ruesch & G. Bateson, *Communication: The social matrix of psychiatry*. New York: W. W. Norton, 212-227.
- Charniak, E. (1972). *Toward a model of children's story comprehension* (Tech. Rep. AI TR-266). Cambridge, MA: M.I.T. AI Laboratory.
- Elrod, M. M. (1983). Young children's responses to direct and indirect directives. *The Journal of Genetic Psychology*, 143, 217-227.
- Gibbs, R. W., Jr. (1979). Contextual effects in understanding indirect requests. *Discourse Processes*, 2, 1-10.
- Goffman, E. (1959). *The presentation of self in everyday life*. Garden City, NY: Doubleday.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics, vol. 3: Speech acts*. New York: Academic Press, 41-58.
- Hill, B., Ide, S., Ikuta, S., Kawasaki, A., & Ogino, T. (1986). Universals of linguistic politeness: Quantitative evidence from Japanese and American English. *Journal of Pragmatics*, 10, 347-371.
- Jacobs, S. & Jackson, S. (1983). Strategy and structure in conversational influence attempts. *Communication Monographs*, 50, 285-304.
- James, S. L. (1978). Effect of listener age and situation on the politeness of children's directives. *Journal of Psycholinguistic Research*, 7, 307-317.
- Lakoff, R. (1973). The logic of politeness: Or, minding your p's and q's. *Papers from the ninth regional meeting, Chicago Linguistic Society*. Chicago: Chicago Linguistic Society, 292-305.
- Lehnert, W. (1984). Problems in question answering. In L. Vaina & J. Hintikka (Eds.), *Cognitive constraints on communication: Representation and processes*. Boston: Kluwer Academic Publishers, 137-159.
- Roach, J. W. & Nickson, M. M. (1983). Formal specifications for modeling and developing human/computer interfaces. In *Proceedings CHI'83 Human Factors in Computing Systems*. New York: A.C.M., 35-39.
- Roach, J. W. & Nickson, M. M. (1986). Software methodology: Rule-based approach to interface design. In R. W. Ehrich & R. C. Williges (Eds.), *Human-computer dialogue design*. New York: Elsevier, 51-59.
- Sanford, D. L. & Roach, J. W. (1986). The role of language in human-computer interfaces: Communication and language. In R. W. Ehrich & R. C. Williges (Eds.), *Human-computer dialogue design*. New York: Elsevier Publishers, 165-177.
- Sanford, D. L. & Roach, J. W. (in press). Representing and using metacommunication to control speakers' relationships in natural language dialogue. *International Journal of Man-Machine Studies*.
- Searle, J. R. (1975). Indirect speech acts. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics, vol. 3: Speech acts*. New York: Academic Press, 59-82.
- Shatz, M. (1978). Children's comprehension of their mothers' question-directives. *Journal of Child Language*, 5, 39-46.
- Watzlawick, P., Bavelas, J. B., & Jackson, D. D. (1967). *Pragmatics of human communication: A study of interactional patterns, pathologies, and paradoxes*. New York: W. W. Norton.