Parsing Meta-Communicational Statements in Dialogues:

A Computational Model

by

Ahmed T. Bouzid

Thesis submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Master of Science

in

Computer Science

APPROVED:

John W. Roach, Chairman

Edward A. Fox

Donald W. McKeon

January, 1991.

Blacksburg, Virginia

c.2

LD
5655
V855
1991
B679
c.2

Parsing Meta-Communicational Statements in Dialogues:

A Computational Model

by

Ahmed Tewfik Bouzid

Dr. John W. Roach, Chairman

Computer Science and Applications

(ABSTRACT)

Prevalent in dialogues are communications about the activity of talking itself. Examination of dyadic conversations readily shows the ubiquity of statements such as "Okay," "Well," "You know," indicating that in natural, spontaneous talk people are often concerned with the process of talking as well as with the content of their talk. In an analysis of a corpus of twenty seven transcribed telephone conversations (80 minutes of continuous talk), we found that such statements occur with an average frequency of 1.49 per conversational turn, and that a turn free of such statements is rare. A verbal expression focusing on the activity of talking we term a **meta-statement.**

In this thesis, we present a model of meta-statements. We propose that meta-statements are pragmatic units of talk used to regulate and maintain the conversational contexts shared by interactants. More specifically, we hypothesize that to determine the "meaning" of meta-statements, six aspects of a conversation's context must be modeled: the *rhetorical, conversational, interpersonal, emotional, cognitive,* and *communicational* aspects. We specify these models in what follows and outline the components of a conversational system that is able to parse the meaning of meta-statements.

Acknowledgements

I would like to thank first and foremost my advisor, Dr. John Roach, for his guidance and support. Without his patience and constant encouragements I would not have been able to complete this work. I would also like to thank the other members of my committee, Dr. Edward Fox and Dr. Donald Mackeon, for their constructive remarks concerning the content as well as the structure of the manuscript. Their comments and suggestions have been an immense help in improving this work.

Special thanks go to Dr. Charles Dudley for the big help he gave me with the material discussed in Chapter Six.

Many thanks go to my friends and fellow Natural-Language Group members: Dan Kilper, Russell Modlin, Ramachandra Venkateshwaran (Venky), Rajesh Virkar, and Rob Warner. I would also like to thank my friends, Bilal Hijazi, Carl Melder, Eskinder Sahle, and Abdurahman Yarali, for their help and friendship.

My deepest appreciation goes to my family for their constant emotional support. I thank my Mother and Father for their love and patience and ask their forgiveness for all my misdeeds. I also thank my brothers, Kamel and Reda, my sister, Souad, my Babelouad cousin, Amine,

my cousins in Belaabas, my uncle "Aamimou," and the rest of my family in Morocco. I also wish to thank my friends "back home," including Zohra and Abdurahman Boukhaf, Amina, Karima, Nadia, Mouh Znagui, Emir, Okba, and all the "hittists" of Algiers.

And last, but far from least, I wish to thank the Queen of Africa, whose beauty and serenity have taught me that there can still be grace and harmony in this world of chaos we live in.

To My Parents—with Maximum Love.

Table of Contents

Chapter 1. Introduction 1

Chapter 2. Literature Survey 6

2.1 Studies of Particles 7

1. And 7

A. Coordinating Ideas 7

B. Continuing an Action 8

2. Because 9

3. But 9

A. Contrasting Ideas 10

B. Contrasting Actions 10

4. Now 11

A. Progression of Ideas 11

B. Shifts in Participation Framework 11

5. Or 12

6. So 13

7. Then 13

8. Well 14
9. You Know 15
2.2 Attempts at Taxonomy 15
A. Reichman 16
1. Support 16
2. Restatement and/or Conclusion of Point Being Supported 16
3. Interruption 17
4. Return 17
5. Indirect Challenge 17
6. Direct Challenge 17
7. Subargument Concession 18
8. Support Challenges 18
9. Prior Logical Abstraction 18
10. Further Development 18
B. Keller 19
1. Semantic Framing 19
2. The Signalling of Social Context 21
3. State-of-consciousness Signals 21
4. Communication Control Signals 22
2.3 Summary 22

Chapter 3. Problem Definition 24
3.1 Scope of the Investigation 26
3.2 The Meta-Statement 27
3.3 Talk about Talk 28
3.4 Talk about Talkers 29
3.5 The Six-tuple Representation of Meta-statements 30
3.6 The Parsing Problem 34

On Attributing Meaning to Surface Forms 35
3.7 The Interpretive System 39
3.7.1 Action Models 40
3.7.2 Parser Components 41
1. State Parameters 41
2. Action Grammars 42
3. Surface Interpretations 43
3.8 Dynamics of the System 44
3.9 Aims of this Work--A Disclaimer 45
3.10 Summary 45

Chapter 4. The Rhetorical Model 48
4.1 The Unit of Talk 53
4.2 A Model of the Rhetorical Unit 55
4.2.1 The Component Unit 56
4.2.2 The Action Unit 59
4.3 A Classification of Rhetorical Connectors 60
4.3.1 Argumentative Connectors 60
Logical Connectors 61
Conceptual Connectors 62
Contextual Connectors 63
4.3.2 Expositive Connectors 63
Composite Connectors 64
Structural Connectors 65
Discursive Connectors 65
4.4 Rhetorical Actions Model 66
4.4.1 Actions 66
Component Actions 66

Connector Actions 67
4.4.2 Grammars 71
4.5 Summary 76

Chapter 5. The Conversational Model 85
5.1 Coordinating Dialogues 86
5.2 The Conversational Situation 90
5.3 Turns, Speakers, and Listeners 91
5.4 Turn Construction and Allocation 95
5.5 Overlaps, Interruptions, and Gaps 98
5.5.1 Overlaps 99
5.5.2 Interruptions 100
Resolution of Turn Disputes 102
5.5.3 Gaps 102
5.6 Conversation Openings 103
5.7 Conversation Closings 105
5.8 State Model 107
5.9 Action Model 109
5.9.1 Actions 109
5.9.2 Action Grammar 110
5.10 Summary 117

Chapter 6. Role Theory 118
6.1 The Role Theory Problem 119
6.2 Actions and Functions 122
6.3 Social Norms 122
6.4 Roles, Statuses, and Positions 124
6.5 Situs and Station 126

6.6 Internal Representations 127
6.7 Application of Role Theory to the Conversations 129
Norms, Functions, and Roles 129
Status, Position, and Situs 133
6.8 Summary 134

Chapter 7. Politeness 144
7.1 On Regularity in Interaction Behavior 148
7.2 The Nature of Interaction 151
7.3 The Notion of Face 153
7.4 Face Wants and Sensitivities 155
H's Negative Face 157
H's Positive Face 157
S's Negative Face 157
S's Positive Face 158
7.5 Status and Power 159
7.6 Distance and Affection 161
7.7 Politeness Strategies 163
On-record Politeness Strategies 164
Negative Politeness Strategies 165
Positive Politeness Strategies 166
Off-record Negative Politeness Strategies 167
7.8 Multiple Impact Analysis 169
7.9 Politeness Actions 170
7.10 Analysis of Transcripts 171
The Status Dimension 172
The Affect Dimension 175
Analysis of a Conversation 177

7.11 Summary 179

Chapter 8. Emotions 181

8.1 On What is to be Considered an Emotion 183

8.2 Emotional State Classes 186

8.3 Intentionality 189

8.4 Valence Dimensions 191

8.5 Modeling Valence Contexts 193

8.5.1 Modeling Goals 195

Goals Structure 195

Dynamics of Structure 196

Types of Goals 198

8.5.2 Modeling Standards 199

8.5.3 Modeling Tastes 201

8.6 A Classification of Emotion Types 202

Goals Based Emotions 203

Focus-on-Self 203

Focus-on-Others (Fortunes-of-Others) 204

Standards Based Emotions (Attribution) 205

Focus-on-Self 205

Focus-on-Others 206

Tastes Based Emotions (Attraction) 206

Compound Emotions 207

8.7 Global Parameters Influencing Valence Reactions 208

Sense of Reality 208

Psychological Proximity 209

Unexpectedness 211

8.8 Local Parameters Influencing Valence Reactions 212

Goals-based 212
Prospect-based 213
Focus-on-Others 214
Standards-based 215
Tastes-based 217
8.9 The Emotion State Model 217
8.10 The Verbal Expression of Emotions 218
8.10.1. Explicit Expressions 219
Explicit Reports 219
Expression Tokens 220
8.10.2. Implicit Expressions 221
I. Structural 221
II. Semantic 225
III. Conversational 227
8.11 Example 229
8.12 Summary 233

Chapter 9. The Cognitive Function 236
9.1 Referential 238
Definite References 239
1. Proper Names 239
2. Pronouns 239
3. Definite Descriptions 239
Indefinite References 240
1. Quantified Expressions 240
2. Indefinite Specifications 240
Indefinite Descriptions 240
9.2 Factual 241

9.3 Contextual 242
9.4 Presentational 244
1. Rhetorical Structuring 244
2. Complexity 245
3. Background Information Assumed 245
9.5 State Model 246
Knowledge Parameters 247
Context Parameters 248
Presentational Parameters 249
9.6 Action Model 250
Actions 250
Grammar 252
9.7 Examples 253
9.8 Summary 259

Chapter 10. The Communicational Function 260
10.1 The State Model 262
1. Receipt of Messages 263
2. State of Transfer 266
3. Converser Readiness 267
10.2 The Action Model 268
Actions 268
Signalling Actions 268
1. Receipt Actions 268
2. Readiness Actions 269
3. Transfer Actions 270
Checking Actions 270
1. Receipt Actions 271

2. Readiness Actions 271
3. Transfer Actions 271
Grammar 272
10.3 Example 275
10.4 Summary 277

Chapter 11. Toward an Integrated System 279
11.1 The Components of the System 280
A. Parameters 280
B. System Actions 283
C. Meta-communicational Actions 284
D. Databases 285
1. Surface Interpretations Database 285
2. FTA's Database 285
3. Politeness Strategies Database 285
4. Emotion Terms Database 286
5. Grammars Database 286
6. Roles Database 287
7. Actions' effects Database 287
11.2 The Structure of the Parsers 287
A. The Cognitive, Communicational, Conversational, and Rhetorical Parsers 288
B. The Emotions Parser 288
C. The Politeness Parser 290
11.3 Summary 291

Chapter 12. Results and Conclusions 305
12.1 What We Wanted To Do 305
12.2 What We Have Done 306

Design 306
Implementation 308
12.3 What Remains to Be Done 309

Appendix A. Surface Interpretations 312
Positional Classification of Tokens 312
Explicit Meta-statements 314

Vita 347

List of Illustrations

Figure 1. Token Action Transition 47

Figure 2. The Level-1 Decomposition of Conversation 77

Figure 3. The Level-2 Decomposition of Conversation 78

Figure 4. The Level-3 Decomposition of Conversation 79

Figure 5. The Level-4 Decomposition of Conversation 80

Figure 6. The Level-5 Decomposition of Conversation 81

Figure 7. The Units of the Level-1 Decomposition of Conversation 82

Figure 8. The Units of the Level-2 Decomposition of Conversation 83

Figure 9. The Units of the Level-3 Decomposition of Conversation 84

Figure 10. Family Group: Statuses and Positions 136

Figure 11. The Structure of Role, Status, and Position 137

Figure 12. The Structure of Situs and Station 138

Figure 13. The Agent-Customer Group 139

Figure 14. The Colleague-Colleague Group 140

Figure 15. The Agent-Supervisor Group 141

Figure 16. The Customer-Supervisor Group 142

Figure 17. The Customer-Agent-Supervisor Group 143

Figure 18. Goal Tree Structure for CLIMBING-HIMALAYAS goal 235

Figure 19. The Rhetorical Parser 293

Figure 20. The Conversational Parser 294

Figure 21. The Cognitive Parser 295

Figure 22. The Communication Parser 296
Figure 23. The Verbal Expression Parser 297
Figure 24. The Goals Evaluator 298
Figure 25. The Norms Evaluator 299
Figure 26. The Tastes Evaluator 300
Figure 27. The Emotions Parser 301
Figure 28. The FTA's Evaluator 302
Figure 29. The Politeness Strategies Parser 303
Figure 30. The Politeness Parser 304

List of Tables

Table 1. Face Impact of On-Record Negative Politeness Strategies 169

Table 2. Face Impact of On-Record Positive Politeness Strategies 170

Table 3. Politeness Strategies Used in Example 178

Table 4. Surface Interpretations (1) 315

Table 5. Surface Interpretations (2) 316

Table 6. Surface Interpretations (3) 317

Table 7. Surface Interpretations (4) 318

Table 8. Surface Interpretations (5) 319

Table 9. Surface Interpretations (6) 320

Table 10. Surface Interpretations (7) 321

Table 11. Surface Interpretations (8) 322

Table 12. Surface Interpretations (9) 323

Table 13. Surface Interpretations (10) 324

Table 14. Surface Interpretations (11) 325

Table 15. Surface Interpretations (12) 326

Table 16. Surface Interpretations (13) 327

Table 17. Surface Interpretations (14) 328

Table 18. Surface Interpretations (15) 329

Table 19. Surface Interpretations (16) 330

Table 20. Surface Interpretations (17) 331

Table 21. Surface Interpretations (18) 332

Chapter 1. Introduction

Almost invariably, science fiction's vision of tomorrow's world includes in some form or another a computer that is able to communicate with people through natural language. Whether it is Poul Anderson's *Muddlehead,* Karel Capeks' *Universal Robots,* or Arthur C. Clarke's *Hal,* tomorrow's "robot" is expected to be able to talk to humans in the language of humans.

Although today's computers are far more "communicative" than the computers of twenty or even ten years ago, the brunt of the communication burden, so to speak, is still on humans: it is humans who must learn the computer's language, and not the computer the human's. In the beginning--that is, when computers were first constructed and put into use--the burden on the human was total: programmers--i.e., people who wanted to "communicate" something to the computer--had to translate their communications into binary instructions: they had to completely transform what they wanted to say from their own language to the language of the

computer. Later on, assemblers were constructed and the burden on the human was somewhat lessened, so that humans did not have to translate directly what they needed to say to the level of the computer's language: they could use a more abstract language--the assembly language.

The trend of shifting the communication burden from the human to the computer continued, and high-level programmming languages became available. With high-level programming languages, people were afforded with a communication language that has, albeit only to a very limited extent, similarities with natural languages. Instead of code names and obscure operation codes, instructions and commands were now denoted by words that meant something to humans.

As the proliferation of computers into society accelerated, the "average" computer user was no longer a computer specialist or an engineer, but rather an "ordinary" person who had little or no training in the science of talking to computers--i.e., programming. Consequently, high-level languages as a means of communication with computers were no longer the answer: a language that placed an even smaller burden on the human side had to be devised.

The language that places all of the communicational burden on the computer and none on the human is, obviously, natural language. We are all of us fluent speakers of some natural language or another, and consequently, if there were a computer with which we could communicate in the language that we spoke, we would be able to do so without having to learn anything new. There would be no need for us to further train ourselves to be able to talk to the computer; we would communicate with it as though we were communicating with another human being.

* * * *

Today's computers are still a long way from reaching the ultimate goal of shouldering all of the communicational burden. When communicating with computers people are still required to follow a rigid format, use certain specific words, obey rules--in short, they are still burdened with restrictions. On the other hand, tomorrow's computers, science fiction writers predict,

will have little, if any trouble understanding people when people talk to them as though they (the computers) were themselves human. Moreover, computers would not only allow people to talk freely and *naturally,* without any restrictions on format or meaning, but would also *understand* what is said to them in the "same" way a human being would understand. In other words, tomorrow's computers, if we are to believe science fiction, will have the same sensitivities to language that human beings have.

Consequently, the first logical step in constructing such a computer system would be to determine the exact underlying laws and rules that people use when communicating with other people. Only after having identified what these rules are would it be possible to make a computer sensitive to the *meaning* of the communications it receives.

The general aim of this thesis is to identify some of these underlying rules. A more narrow goal--a goal the solution to which, as we will see in the following chapters, is intricately related to the general aim cited above--is to build a conversatinal model that takes into account the ubiquitous verbal particles that characterize natural, spontaneous talk--particles such as "well," "you know," "okay," "all right," etc. We shall use the term *meta-statements* to refer to these particles.

* * * *

Before we proceed with a detailed exposition of the problem this thesis will be addressing, we pause here to examine in what follows an extract from Arthur C. Clarke's science fiction tale, *2001: A Space Odyssey* (pp. 135, 136). The extract will demonstrate quite clearly that tomorrow's computers are expected to possess many if not all of the communication capabilities that people possess. The extract in question is a conversation between a human, David Bowman, and a robot, Hal. The conversation is taking place in a space ship, somewhere in space, and at the moment where we pick up the conversation, David Bowman, the astronaut, is questioning Hal in an attempt to determine the cause of an apparent on-board circuit fault. The excerpt follows:

"Have you any idea," [Bowman] said, "what's causing the fault?"
It was unusual for Hal to pause so long. Then he answered:
"Not really, Dave. As I reported earlier, I can't localize the trouble."
"You're *quite* certain," said Bowman cautiously, "that you
haven't made a mistake? You know we tested the other AE35 unit thoroughly,
and there was nothing wrong with it."
"Yes, I know that. But I can assure you that there is a fault. If
it's not in the unit, it may be in the entire subsystem."

....

"Well, I'll report it to Mission Control and we'll see what they
advise." He paused, but there was no reaction.
"Hal," he continued, "is something bothering you--something that might
account for this problem?"
Again there was that unusual delay. Then Hal answered, in his normal
tone of voice:
"Look, Dave, I know you're trying to be helpful. But the fault is either
in the antenna system--or in your test procedures. My information
processing is perfectly normal. If you check my record, you'll find it
completely free from error."
"I know all about your service record Hal--but that doesn't prove
that you're right this time. Anyone can make mistakes."
"I don't want to insist on it, Dave, but I am incappable of making
an error."

What is remarkable about the above exchange is how unlike a computer Hal "sounds." Hal sounds unmistakably human. After we have read the passage, we have no reservations concluding that there is a tension between Hal and Bowman, and that, overall, Hal is behaving defensively. In lines (2) and (13), for example, we are told that Hal pauses longer than usual before he gives his answer, and from this we are compelled to infer that Hal is *hesitating.* Moreover, we notice that Hal does not give his negative answers straighforwardly, instead he softens his talk. We see this in line (3), where he uses "Not really, Dave," instead of the more direct "No," in line (15), with "I know you are being Helpful. But...", and in line (21) with "I don't want to insist on it, Dave, but ...."

We notice also that Hal is aware of himself and his own state of mind. In line (4), for example, he refers to something he had said previously: "As I reported earlier," while in line (7) he declares that he knows and is aware of the information Bowman is communicating to him: "Yes, I know that." Hal is also aware of Bowman's lack of confidence in his competence:

"I can assure you..." (line (7)), "My information processing is perfectly normal" (lines (16) and (17)), and more explicitly, "I am incapable of making an error" (lines (21) and (22)).

What is also worth noting is that Bowman, a human, interacts with Hal almost exactly as though he were interacting with another human. In line (4), for example, he asks Hal to make sure that he (Hal) has not made any mistakes, prefacing his request with "you're *quite* certain;" in line (11) he goes so far as to ask about Hal's emotional state: "is something bothering you;" while in line (20), he tries to tell Hal as delicately as possible that "anyone can make mistakes."

Bowman, as we can see, communicates with Hal as though he were communicating with a human. And what is important to note is that, by communicating with a computer with the language he uses to communicate with humans, Bowman cannot help but attribute to Hal's actions--the words Hal says, the pauses, etc.--the same meaning he would attribute to them were he talking with a human being. Hal is no longer a machine for Bowman: he is a living entity with feelings, an ego, a personality, a self-consciousness--in short, an entity deserving the delicate subtleties of human communication.

Chapter 2. Literature Survey

A detailed, although by no means exhaustive, survey of the literature on meta-statements is given in [BOUZ91]. In this chapter, our aim is to briefly outline the major issues discussed in that paper without introducing too many details into the discussion. Illustrative examples as well as a more detailed examination of the concepts discussed may be found in [BOUZ91].

We examine two types of studies in this chapter: first are studies that concentrate on examining certain particular meta-statements, and second are studies that attempt to classify meta-statements. In section 2.1 we discuss the first type of studies, while in section 2.2 we concentrate on the second type.

2.1 Studies of Particles

The most detailed studies of the particles we shall be focusing on in this survey were conducted by Schiffrin [SCHI85], [SCHI87]. Important contributions were also made by van Dijk [DIJK79], [DIJK80], and Reichman [REIC85], and in what follows we shall mainly examine the theories presented by these three researchers. The particles studied are: *and, because, but, now, or, so, then, well,* and *you know.*

1. And

The following discussion begins with Schiffrin's examination of the role played by "and," and then turns to a discussion on van Dijk's analysis. Schiffrin identifies two roles that "and" plays in talk: first, it coordinates and connects idea units, and second it continues a speaker's action. She argues that although the roles are independent of one another, they are nonetheless simultaneously played.

A. Coordinating Ideas

Schiffrin found that "*and* is the most frequently used mode of connection at a local level of idea structure." Whereas 1002 "clause-sized idea units" in her corpus of analysis were prefaced by "and," only 440 were prefaced by "but," 206 by "so," and 53 by "or." She identifies two types of coordinating ideas functions where "and" is used: (1) defending a position, and (2) relating topic structures.

"And" is used when defending a position in the following four manners:

(1) Connecting events within the statement of a position.
(2) Connecting events within the statement of a support.
(3) Connecting two supports to a stated position.
(4) Prefacing the restatement of an already stated position.

When used to relate topic structures, "and" fulfills one of the following two functions:

(1) Connecting structures within a topic structure.
(2) Connecting structures with different topics.

B. Continuing an Action

The second function that "and" plays, Schiffrin argues, is that of continuing an action, or more specifically that of marking speaker-continuation. First, she notes that "and" is often used as a device for warding off talk that threatens to steer the conversation away from its current topic.

A second type of speaker continuation, Schiffrin argues, occurs when the speaker is addressing to the listener a series of questions, with the questions asked being often connected with "and."

A similar structure takes place when the speaker is requesting an elaboration of a response given by the listener, and the request for elaboration is marked with an "and." In such cases, the "and" marks not only that another question is going to be asked, but also that the question is somewhat related to the previous question.

The third function is that of turn transition. Speakers often offer the turn at talk by ending their turn with an "and."

Van Dijk [DIJK79] distinguishes between two types of connectives: those that are *semantic* and those that are *pragmatic*. Whereas he defines semantic connectives to be those that express relations between facts, he calls pragmatic those connectives that express relations between speech acts.

One use of "and" van Dijk points out is that of marking that the speaker wants to add something to the first speech act, and argues that "its function may be characterized in terms of concepts like 'addition' or 'continuation'" [DIJK79.450].

The second function that "and" can fulfill is to signal a contradiction or protest, as for instance when the speaker wants to prevent the listener from drawing false conclusions from a previous statement.

2. Because

Schiffrin identifies three uses of "because." First, she argues that "because" may be used to mark a fact-based cause, i.e., it may link two discourse units, where one unit contains information about the cause that brought about the consequences promulgated in the other unit.

In addition, she argues that "because" can play is that of prefacing a unit that contains background information. And third, "because" can also be used to preface accounts for requests.

3. But

Schiffrin notes that "but" is used mainly to contrast an upcoming unit with a previous unit. She makes the distinction between two types of contrasts: (1) contrasting ideas, and (2) contrasting actions.

A. Contrasting Ideas

The first use of "but" in contrasting ideas that Schiffrin mentions is the case where the contrast marked is between a hypothetical situation and an actual situation.

The second use of "but" occurs when the content of what is going to be said may lead the listener to draw conclusions about the speaker which the speaker does not want the listener to draw. "But" in these cases is used after a disclaimer statement is explicitly stated by the speaker in the attempt to avoid giving the "wrong" impression.

In other instances, "but" is used when the speaker wants to return to the main subject of the talk after going into what Schiffrin calls a "socially more cooperative" portion of talk.

B. Contrasting Actions

Five major uses of "but" are identified under the category of "contrasting actions."

First, "but" may be used to return to a general statement after going into a description of a particular which in some way contradicts the general statement. "But" may also be used when marking a return to the point of the discussion. And third, "but" may occur when a speaker's turn is threatened by an interruption, and more specifically, by an interruption that may be caused by the other speaker or speakers agreeing with the current speaker. "But" in such instances is used to retain the turn.

The fourth use of "but" is the case where the speaker is defending a point against a challenge. And fifth is the case when "but" is used when prefacing a disagreement.

Van Dijk identifies four pragmatic functions that "but" may fulfill. First, he notes that a speaker may use "but" in order to mark that an essential condition for the fulfillment of a directive given by another speaker does not hold.

"But" can also be used, van Dijk notes, to indicate that the speaker does not accept the other speaker's previous speech act by questioning one of its conditions.

The third pragmatic function fulfilled by "but" is one where the speaker is expressing surprise, as in: "But, you had your hair cut!"

The fourth function identified is when the speaker concatenates his own speech acts (as opposed to connecting another's to one's speech act), as for example when marking the satisfaction or non-satisfaction of "illucutionary conditions, or to emphasize that a previous speech act only holds under specific circumstances."

In addition, van Dijk notes that "weaker forms may be used as conversational starters, e.g., when a speaker tries to interrupt another speaker, even in those cases where no contradiction or semantic contrast is involved."

4. Now

Schiffrin identifies two types of functions that "now" can fulfill: (1) Marking the progression of the ideas treated in the discourse, and (2) Marking shifts in the participation framework.

A. Progression of Ideas

According to Schiffrin, "'Now' occurs in discourse in which the speaker progresses through a cumulative series of subordinate units." In addition, she argues that "now" can also occur in a discourse where a subordinate unit is to be interpreted in relation to a larger structure.

B. Shifts in Participation Framework

Parts of a discourse may also be differentiated by a shift in the mode through which a speaker is related to the information presented. "Now" is often used to signal this shift in mood.

The first type of shifts is when the speaker changes his mood from the declarative to the interrogative. The second type is when the speaker ends the presentation of a sequence of facts with a rhetorical question. The third type is when the speaker goes from the narrative mood to the evaluative mood. Finally, according to Schiffrin, "now" can also be used to shift from a general position to a more specific example that supports that position.

In addition to shifts, "now" can also be used when the speaker is negotiating the right to control what will happen next in talk.

Reichman also examines the function of "now," but identifies only one particular function: "the clue word 'Now' often signals that a new active context space is about to be developed that will contain elaborative comments on the topic and elements of the preceding active context space" [REIC85.44].

5. Or

According to Schiffrin, the main function of "or" is to mark a "speaker's provision of options to a hearer," and more specifically, it marks the two-way choice offered by speaker to hearer of either accepting only one member of the disjunct or all members of the disjunct.

Van Dijk argues that "or" is often used to introduce questions, and more specifically, questions where the speaker tries to determine whether the conditions necessary for a previously performed speech act hold. In addition, a speaker may precede an explanation justifying the performance of a speech act.

Van Dijk notes that "or" may fulfill a politeness function when the speech act that is put in question is a request made by the speaker to the listener. In such a situation, the "or" may be used by the speaker to give a "way out" to the listener, if the latter is not able or does not want to fulfill the speaker's request.

6. So

Schiffrin distinguishes between two functions of "so." The first function is concerned with connecting what she calls idea units, while the second function is concerned with the organization of transitions in talk.

Three types of idea unit connection uses of "so" are identified by Schiffrin. First, she argues that "so" can be used to mark a "fact-based" result relation at both local and global levels of discourse. Second, she argues that "so" can mark "the use of idea units as a basis for additional knowledge. And third, she points to the use of "so" with requests, to mark that an action was motivated by prior discourse, with the action being usually a request.

In addition to its use with requests, "so" can also fulfill a function in the management of transitions in the "participation framework."

Van Dijk notes that "so" is pragmatically used to link "two speech acts of which the second functions as 'conclusion' with respect to the first speech act." A variant of this case is when one speaker draws a conclusion from another speaker's speech act.

In addition to drawing conclusions from facts, van Dijk notes, a speaker may use "so" to draw conclusions from *the force* of the speech act.

Reichman suggests that "so" is often used to preface a restatement of a point already stated, after presenting information supporting that point.

7. Then

Two functions of "then" are identified by Schiffrin: (1) marking successive ideas in discourse time, and (2) marking successive actions in discourse time. In the first case "then" is used to mark temporally successive events, whereas in the second case, "then" marks the succession of actions in discourse, and in particular, marking *warranted requests.*

8. Well

Schiffrin found that often "well" is used to preface an answer to a question. She distinguishes between two types of such prefaces: (1) those that preface the introductory phrases, and (2) those that preface the main portion of the answer.

In addition to prefacing answers, Schiffrin holds, "well" can also occur within request/compliance pairs. She begins by arguing that "non-compliance with a request is more likely to be marked with *well* than is compliance." In this case, the "well" functions to mark the hesitation, or reluctance of the speaker to fail to comply with the other speaker's request.

In general, Schiffrin argues that "well" is often used with disagreements, denials, and in situations where the speaker is unable to provide sufficient or satisfactory answers.

Next, she argues that "well" is often used to preface requests. She identifies four types of uses of "well" with requests. First, "well" may be used to preface a request for clarification. Second, she notes that a series of requests marked with "well" can occur when questions on a speaker's agenda are not answered because of a divergence from the topic. "Well" in such situations can be used to preface a return to the requests or questions. Third is the case when the response given to a request does not satisfy some need for information; the request in such case is often rephrased with a "well" prefacing it.

And the fourth and last use of "well" with requests that Schiffrin identifies "is when a request is based on the speaker's use of information from prior talk to draw an inference from which confirmation is requested."

In addition to its use with requests, "well," Schiffrin found, can also be used to respond to a referent/response pair.

Also, "well" is often used when reporting speech, i.e., when quoting talk "whose original time, place, and possibly author, is not concurrent with the on-going conversation."

A related use of "well" is marking a shift in speaker orientation, when going, for instance, from a description of events to an evaluation of those events.

9. You Know

"You know" is found by Schiffrin to fulfill three types of functions. First, it may be used with relation to information states. One such use is the assessment by one speaker of another speaker's knowledge state.

The second type of use of "you know" is in arguments, when the speaker uses "you know" to appeal "to shared knowledge as a way of converting an opponent to one's own side dispute."

And the third use of "you know" is in narratives, "when a hearer is invited to share in the information transfer being accomplished through narrative discourse."

Having briefly examined the functions identified for some of the most studied meta-statements, we now turn to examining two particular studies where a taxonomy of meta-statements is presented.

2.2 Attempts at Taxonomy

In this section, we examine two attempts at classifying meta-statements. First, we examine the taxonomy suggested by Reichman, then we focus on Keller's taxonomy.

A. Reichman

Reichman identifies ten types of what she calls *conversational moves.* Conversational moves she defines to be "the various kinds of semantic and logical relations that can hold between utterances of discourse." To characterize a type of conversational move, Reichman first specifies "the components that explicitly or implicitly are required for the completion of particular moves," and then identifies certain *clue words* that "usually accompany different types of conversational moves in a discourse." These "clue words" are what we call meta-statements.

In what follows, we give Reichman's ten conversational moves. For each move we first briefly describe the move, then we give the various clue words that, according to Reichman, apply to the move in question.

1. Support

Description: The speaker moves to support a claim.

Clue Words: "Because," "Like," "Like when."

2. Restatement and/or Conclusion of Point Being Supported

Description: Speaker restates a previously made point, either in terms of an elaboration, rewording, or summary, or draws a conclusion from what has already been said.

Clue Word: "So."

3. *Interruption*

Description: The speaker digresses from the subject before closing it, with the intention of returning to that subject after the digression has been closed.

Clue Words: "Incidentally," "By the way."

4. *Return*

Description: The speaker closes a digression and returns to the main topic of the talk.

Clue Words: "Anyway," "In any case."

5. *Indirect Challenge*

Description: The speaker signals to the listener that although he is not going to directly attack the listener, he is nonetheless going to challenge the latter's argument at least indirectly.

Clue Words: "Yes, but," "Right, but," "Except, however."

6. *Direct Challenge*

Description: The speaker directly challenges the listener's preceding utterances.

Clue Words: "No," "No, but."

7. Subargument Concession

Description: The speaker, although accepting the validity of the listener's preceding statement, decides to argue another related point which has not yet been resolved.

Clue Words: "All right, but," "Okay, but."

8. Support Challenges

Description: The speaker challenges the listener's point. Three methods of challenge are identified: (1) rejecting the relevance of the listener's statement as supporting the latter's point (e.g., "so what?"), (2) denying the truth of the listener's statement (e.g., "that's false"), or (3) "claiming that the domain specified [by the listener] in the support is inappropriately matched to the domain of the claim."

9. Prior Logical Abstraction

Description: The speaker closes a current context space and shifts the discourse "into a context space whose topic is 'logically prior' to that of the closed space."

Clue Words: "But look," "But listen," "But you see."

10. Further Development

Description: The speaker signals that what follows will contain a further elaboration of what has been stated previously.

Clue Words: "Now."

B. Keller

We now examine the taxonomy of meta-statements suggested by Keller. Keller approaches the analysis of conversational speech by examining what he calls "the psycholinguistic strategies used by the participant in a conversation" to structure the content of that conversation as well as the procedure undertaken to carry out the conversation. These strategies, he argues, often take the form of explicit, "semi-fixed" verbal expressions he calls *gambits.* These gambits are what we call meta-statements.

Keller's taxonomy, we found, is by far the most comprehensive and the most structured of studies on meta-statements. Five hundred gambits commonly used in North American English are identified from an analysis of over 10,000 sentences of Canadian English speech, and a four way classification of these gambits is given.

Keller distinguishes between four functions that gambits have: semantic framing, social signalling, state-of-consciousness signalling, and communicative signalling. In what follows, we briefly discuss each of these functions.

1. Semantic Framing

The semantic framing function of gambits designates that function which refers to the semantic information presented by the speaker. For example, a speaker may want to qualify some stretch of talk he is going to give or a stretch of talk he has already given in some particular way, e.g., "I have some bad news," qualifies the upcoming stretch of talk as "bad news," while "that's only my opinion," qualfies what has been said as "an opinion." Seven subcategories of semantic framing functions are identified. We discuss each in what follows.

(a) Major Semantic Field Indicators:

The functions identified in this category delimit the semantic fields, i.e. the semantic contexts that are the focus of the discourse, marking transitions from one context to the next. In all, three functions are identified:

(i) Congruent Major Semantic Field: when the upcoming context is a continuation of the current context, e.g., "I have a question on that."

(ii) Incongruent Major Semantic Field: when the current context and the next context are not related. Two subfunctions are identified: (1) that of *initiating a digression,* e.g., "This reminds me," and (2) that of *returning to main topic,* e.g., "In any case."

(b) Various Aspects of a Topic:

The function of the gambits in this category is to relate information about an aspect of the upcoming topic. Five aspects are identified:

(i) A List: e.g., "First," "To begin with."
(ii) A Main Aspect: e.g., "The main thing is," "Most of all."
(iii) A Surprising Aspect: e.g., "Believe it or not," "Strangely enough."
(iv) An Unpleasant Aspect: e.g., "To be realistic," "Let's face it."
(v) An Emphasized Aspect: e.g., "The main thing is," "It bears emphasizing."

3. Opinion:

In this category, the gambits function to qualify the speaker's talk as an opinion.
(i) Guessing: e.g., "My guess is."
(ii) An Opinion: e.g., "I'm pretty sure that."
(iii) A Conviction: e.g., "To my mind."
(iv) A Personal Viewpoint: e.g., "As far as I can tell."
(v) Personal Circumstances: e.g., "In my case."
(vi) Confidential Information: e.g., "Just between you and me."

4. Action Strategy:

The gambits in this category serve as a preface to the performance of some action.
(i) A Suggestion: e.g., "Here's what you can do."
(ii) A Plan: e.g., "What we have in mind is..."

5. Subject-expansion:

The gambits in this category are used by speakers to inform the listeners that what follows is an expansion of the current subject:

(i) Expanding a Point: e.g., "As far as that is concerned."
(ii) Adding Items: e.g., "Another thing."
(iii) Giving a Reason: e.g., "The reason why."
(iv) Explaining a Result: e.g., "Consequently."

(v) Positive Contingency: e.g., "In case of."
(vi) Negative Contingency: e.g., "Unless."
(vii) Restatement: e.g., "What I mean is."
(viii) Appearance and Reality: e.g., "You may think... but in fact..."

6. Subject-evaluation:

The gambits in this category are used by speakers to express how they evaluate a point or a subject they are about to discuss:

(i) Reservations: e.g., "Yes, but consider."
(ii) Taking Into Account: e.g., "Seeing as how...," "keeping in mind..."
(iii) Seeing the Other Side: e.g., "Yet on the other hand," "Mind you though."

7. Argumentation:

The gambits here are used by speakers during arguments:
(i) Generalization: e.g., "Most of the time."
(ii) Exceptions: e.g., "As an exception."
(iii) Examples: e.g., "As an example."
(iv) Summerizing: e.g., "In short."

2. *The Signalling of Social Context*

The second main category of gambit functions is comprised of those gambits which signal the speaker's involvement in the social context created by the conversational exchange. Six subfunctions are identified.[1]

1. ***I want to have a turn:*** e.g., "May I interrupt for a moment."
2. ***I want to keep my turn:*** e.g., "Wait a second."
3. ***I want to abandon my turn:*** e.g., "That's about it."
4. ***I don't want to take a turn:*** e.g., "I'll pass on that."
5. ***Why don't you take a turn:*** e.g., "And what about you?"
6. ***I want to leave the conversational group:*** e.g., "It's been nice talking to you."

3. *State-of-consciousness Signals*

The third category of gambits includes those verbals statements that "indicate a person's state of consciousness concerning information, opinions, or emotions." Eleven subfunctions are identified (again the semantics of the functions are paraphrased):

---

[1] The subfunctions are identified by a paraphrase of their "meaning:"

1. ***I am ready to receive information:*** e.g., "I'd like to hear all about it."
2. ***I am not ready to receive information:*** e.g., "I'm not really interested in that."
3. ***I am ready to provide information or opinion:*** e.g., "I've got something to tell you."
4. ***I am not ready to provide information or opinion:*** e.g., "I don't want to get into that."
5. ***I am sharing some knowledge/an opinion/an emotion with you:*** e.g., "I knew that."
6. ***I am not sharing your knowledge/opinion/emotion:*** e.g., "I didn't know that."
7. ***I am not (entirely) sharing your opinion:*** e.g., "I don't think so."
8. ***I would like you to share my opinion:*** e.g., "What do you say about that."
9. ***I yield to your opinion:*** e.g., "On second thought."
10. ***I am ready for some particular action:*** e.g., "I'd love to."
11. ***I am not ready for some particular action:*** e.g., "I'd rather not."

4. Communication Control Signals

The fourth and final category includes gambits that are used by speakers to request or relate information about states of understanding or hearing:

1. ***Do you understand/hear me?:*** e.g., "Are you following me?"
2. ***I understood/heard you:*** e.g., "Okay."
3. ***I did not understand/hear you:*** e.g., "Pardon me?"
4. ***You must have misunderstood me/not heard me right:*** e.g., "That's not what I meant."

2.3 Summary

We presented in this chapter an overview of work done on meta-statements. The most detailed work on particular meta-statements has been done by Schiffrin, and our previous discussion concentrated on her contributions. Van Dijk's work, as well as work done by Reichman, was also presented.

In addition, we gave a survey of two particular schemes for classifying meta-statements. Reichman's classification identifies ten categories of meta-statements, while Keller's classification subdivides meta-statements into four categories.

In this thesis, our goal is to give a comprehensive, structured model of the functions that meta-statements fulfill. Our study will not examine meta-statements by examining a set of particular verbal expressions or words, nor will it concentrate on one particular function of meta-statements. Instead, we will proceed by first identifying the various functions that meta-statements fulfill, and then examine how these functions are fulfilled by meta-statements. In the chapter that follows, we examine more closely the problem we are confronting and give an outline of our solution.

Chapter 3. Problem Definition

Our aim in this thesis is to build a computational system that can parse meta-statements, i.e., a parser that can map a meta-statement's surface form to its *meaning.* Obviously, then, our first step in building this parser is to establish what a meta-statement means. In other words, we must establish the semantic domain of meta-statements.

In order to solve this problem, we must first proceed by defining exactly what we mean by the term "meta-statement." Only with such a definition will we be able to define the semantic domain of meta-statements. After establishing the domain to which we are to map the surface form of meta-statements, our next task is to establish *how* we are to map a surface form to its corresponding meaning in the semantic domain. That is, we must construct a device that takes as input lexical items--the meta-statements--and gives as output a structure describing the meaning of those lexical items.

We begin this chapter with a discussion on the scope of our investigation. Following this, in section 3.2, we give a precise definition of what we call meta-statements and propose that a meta-statement is a report on two types of objects: *talk* and *talkers.* In section 3.3, we examine the first type of reports, i.e., talk about talk, while in section 3.4, we discuss talk about talkers. In section 3.5, we identify two types of "talk about talk" reports--those pertaining to the structure of talk and those pertaining to the process of talking--and four types of "talk about talkers"--those reports describing speakers' emotional states, those pertaining to politeness, thoses describing speakers' ability or willingness to communicate, and those relating information about speakers' knowledge and understanding. To illustrate the six aspects of meta-statements, we examine the meta-statements contained in a conversation fragment.

In section 3.6, we discuss some issues involved in solving the parsing problem in general. Our argument in this thesis is that the solution to the parsing problem is not to be found through an examination of structures and forms of talk (i.e., syntax) or through the modeling of word meanings (i.e., semantics). Instead, we propose that the solution lies in the construction of a model that takes into consideration aspects such as who the speakers are, what relationship exists between them, how they feel, what they understand, etc., i.e., we propose the construction of the *pragmatics* of communicational situations.

Section 3.7 contains an overview of the major components of this pragmatic model, while section 3.8 gives a brief description of how the parts of this model interact with one another to produce a dynamic parser.

But before we begin our discussion on meta-statements, we first establish the scope of our investigation in this study.

3.1 Scope of the Investigation

There are three major restrictions on the scope of our investigation. First, the exchange situation we are examining in this analysis is the *telephone dialogue.* Obviously, this is only one of many other possible communicational settings, such as direct face-to-face interactions, letter correspondences, communications via teletype, CB radios, etc. Second, we are restricting our examination to dialogues where the interactants are strangers to one another, and where one of the interactants is a seller (viz. a travel agent) while the other is a buyer (viz. a customer). Other possible relationships may be that of friend/friend, mother/child, teacher/pupil, etc. And third, we are examining only verbal expressions; our field of investigation does not involve other channels of communication, e.g., body gestures, facial expressions, etc.

Although we recognize that we have greatly narrowed the scope of our investigation with these three restrictions, we argue that the model of discourse we have developed can be easily extended to model settings that involve interactants besides buyers and sellers. Moreover, according to [BIRD70], [HALL63], and [HASL87], a model developed for the verbal exchange can be used to analyze nonverbal communication in general. Therefore, we argue that although we have used verbal messages in our investigation, the system we have developed is sufficiently powerful to function regardless of how messages are sent.

The data we have used in this investigation comes from a collection of twenty-seven telephone conversations between an airline reservation agent and customers.

We now proceed to define what we mean by the term meta-statement.

3.2 The Meta-Statement

By *meta-statement* we mean any utterance that pertains to the *discursive system.* That is, a meta-statement is an utterance that relates information about *the activity of talking.* More generally, *meta-messages* refer to any *message* that pertains to the discursive system. Thus, a meta-statement is a particular type of meta-message, namely a *verbal* meta-message.

We identify two sets of objects as the most basic elements that make up the discursive system. First, we have the utterances exchanged (in general, the messages sent) during the interaction; and second we have the interactants involved in the exchange, i.e., the conversers. Communications dealing with the former objects we term "talk about talk," those with the latter "talk about talkers." An example of a "talk about talk" meta-statement would be the utterance, "What was that question again?" where the object talked about is explicitly the utterance, viz. the question uttered previously by the other speaker. An example of a "talk about talkers" meta-statement would be the utterance "I don't understand," where the statement uttered by the speaker is a description of the speaker's own state of understanding.

We will at the outset make the point that this dichotomy is vague at best and is by no means a clear-cut classification. In the first example cited above we can clearly see that the utterance is not exclusively of the "talk about talk" type; it also has a "talk about talkers" aspect to it. The speaker may very well have been saying at *the same time,* "I don't understand what you said," or perhaps, "I couldn't hear you," both cases being "talk about talkers" meta-statements. And similarly for the second utterance, although the statement is primarily a report on the speaker's own cognitive state, it is implicitly "talk about talk" as well since the speaker is referring to something the other interactant had said previously. So then, as we can see already, a meta-statement can relate information about both objects of the discursive system simultaneously.

The "talk about talk" and "talk about talkers" dichotomy, however, serves the purpose of identifying the major aspects of the meta-statement, and so we shall retain it for now.

Subsequently we will expand it as we progress towards a more exact and fine grained characterization of the functions of the meta-statement.

3.3 Talk about Talk

At its most basic level, "talk about talk" deals with individual utterances, and almost all meta-statements can be said to be "talk about talk" in this sense. After all, it is through verbal communication that the meta-statement's messages are sent, and even when the meta-statement is primarily of the "talk about talkers" kind, it is also in the sense of talkers with respect to talk, and thus indirectly with respect to talk itself.

Meta-statements along the "talk about talk" dimension, however, pertain to more than simply individual utterances. Utterances stand in relation to one another, and a meta-statement can relate information about these relationships.

Moreover, a meta-statement can relate information about not only one utterance, but also about a *group* of utterances that together form a larger unit. And just as a meta-statement can relate information about relationships between utterances, it can also refer to the relationships between groups of utterances.

Another aspect of "talk about talk" meta-statements relates not to talk itself, *per se,* but rather to the *process of talking.* In our case, the rules used for regulating this process are dictated by the dialogue's communication protocol. One way these rules are observed (or violated) is through the meta-statement. An example of this type of meta-statements would be the utterance "may I say something?" where the speaker explicitly brings to the foreground his concern for the process of talking by asking for the conversation turn.

The function of meta-statements dealing with individual utterances or groupings of utterances and the relations that exist between them we will term the **rhetorical** function of the meta-statement. That function which deals with the process of talking, on the other hand, we

will term the **conversational** function. In other words, the rhetorical aspect of a meta-statement is a report on the *structure of talk,* whereas the conversational aspect of a meta-statement is a report on the *process of talking.* We shall give a more thorough exposition of these two aspects in chapters Four and Five, respectively. For now, let us investigate the other dimension of meta-statements, that is, the "talk about talkers" dimension.

3.4 Talk about Talkers

Whereas "talk about talk" focuses on the structures and processes of talk, "talk about talkers" is concerned with the interactants engaged in the encounter. We distinguish between four functions that meta-statements can fulfill along the "talk about talk" dimension.

First, we have the **communicational** function. An example of a communicational meta-statement would be the utterance "Can you hear me?" where the speaker is trying to ascertain the integrity of the communicational lines that link the two interactants. Most of the time, the communicational aspect is not explicitly pronounced. It is only when communication is physically threatened, e.g., a bad telephone connection, that the communicational function comes to the foreground.

The second function of "talk about talkers" meta-statements relates information about the **cognitive** states of the interactants. Here, it is the understanding of the interactants that is the concern. An example of a cognitive meta-statement would be "Do you understand what I said?" where the interactant is focusing on whether or not his partner has understood something said previously. Like the communicational aspect, in most conversational settings the cognitive aspect is rarely articulated. However, in encounters where explanations and descriptions are frequently given (e.g., the tutor/student encounter), the cognitive aspect may not only be frequently addressed, but may even be the aspect focused on the most often.

The third "talk about talkers" function of meta-statements pertains to the **emotional** state of the interactants. Here, it is the emotional state of the conversers to what is being said that is addressed. An utterance of the sort "Oh, that's nice," for example, would be classified as primarily emotional. The speaker of the utterance is telling us about his feelings. He is, in a sense, giving us a report on his emotional state.

The fourth function fulfilled is the one dealing with the interpersonal relationship that exists between the interactants. This function we call the **interpersonal** function of the meta-statement. An example of a primarily interpersonal meta-statement would be the utterance "I'm sorry about that, sir," where the interactant seems to be explicitly tending to what we call *the face* of his partner, and thus, in a sense, addressing the relationship that exists between him and his partner.

3.5 The Six-tuple Representation of Meta-statements

We have therefore identified *six* functions that a meta-statement can fulfill. Along the "talk about talk" dimension, we have identified the *rhetorical* function and the *conversational* function, whereas along the "talk about talkers" dimension we have determined four distinct functions: the *communicational* function, the *cognitive* function, the *emotion* function, and the *interpersonal* function. We propose that a meta-statement is an action that may fulfill any combination of these functions simultaneously. A meta-statement is therefore not one single action but many, and up to six, distinct parallel actions.

To clarify the notion of multi-functionality of meta-statements, let us take an example.

The following is an extract from a conversation contained in our corpus of twenty seven telephone conversations. The exchange occurs between a travel agency representative, Mrs.

Crackenthorpe,[2] and a male customer. The phrases enclosed between square brackets indicate what we consider to be meta-statements. The exchange follows.

Example 3.1:[3]

1 -AGENT.
2 [American Airlines, Mrs. Crackenthorpe.]
3 -CUSTOMER.
4 [Yes,] [um] I'm planning a trip from
5 [ah] Roanoke to L.A. I'd like some
6 [ah] information.

Let us examine the meta-statements contained in this opening sequence. The agent begins the conversation by giving the name of the company she is representing, that is, American Airlines, and then her last name, which she qualifies with the title "Mrs.," indicating thus both gender and marital status. This is a standard opening move for business establishments receiving telephone calls from customers. The agent is in effect, by this opening line, advertising the *role* she is willing to play in this incipient stage of the encounter, namely, the role of "travel agency representative." At the same time, the agent refers to herself as "Mrs. Crackenthorpe," a more formal address term than, for example, the first name term, and indicates that she is married through the use of the token "Mrs." Clearly then, this opening move contains many *interpersonal* messages. Paraphrasing the utterance, we might say that the utterance is a condensed way of saying: "I represent American Airlines. My job is that of a travel agent. I am a woman. I am married. I am willing to engage only on a purely transactional level with you at this point. Given the roles we are playing, I will strive to be friendly to you, but I do not plan on our developing a personal relationship with each other. This is why, among other things, I wish that you refer to me as Mrs. Crackenthorpe...."

---

[2] The real names have been replaced here and throughout the Thesis with ficticious ones.

[3] Conversation Phone.03, lines 1-6

A second function of the very same meta-statement runs along the *conversational* dimension; namely, the agent is opening a conversation. She is saying that she is willing to engage in a one-to-one conversational encounter with the customer and that she is ready to proceed with the exchange. This is the *conversational* function.

Third, the agent is also, as it were, testing the channels of communication. The physical message, i.e., the sound she sends, will serve to test whether the other interactant is able to hear her or not. At the same time, the agent is signalling that her own channels of communication are ready to receive signals, i.e., that she has decided to dedicate the line of exchange to the customer in question. This function served by the meta-statement is what we call the *communicational* function.

Fourth, the agent is testing whether or not the other person is aware of the situation he is involved in. The message is in a way a warning to all those who dialed the wrong number that this is not the number they were trying to contact. This is the *cognitive* function of the meta-statement.

Fifth, the statement is implicitly setting the stage for the topic to be discussed. That is, the agent is saying by the meta-statement: "Let's talk about the problem you want to solve." Topic introduction is one type of what we term *rhetorical* actions.

Finally, the statement is a report on the emotional state of the interactant, although in this case, the state is apparently "neutral" and no outward manifestation of a specific emotion is made.

Now let us analyze the customer's response. The customer begins with the meta-statement "yes," and goes on to announce the purpose of his call, twice using the meta-statement "ah" and once the meta-statement "um." The "yes" is a rich meta-statement in almost all of the six dimensions. First, it addresses the communicational dimension, announcing that, yes, he can indeed hear fine (whereas if he could not hear the agent, he might have started with a statement such as "Hello? I can't hear you..."). Second, it announces that he has understood the situation, that is, that he knows he is in contact with an airline reservation agency, and that that is what he wants (whereas if he had dialed the wrong number,

he probably would have answered with something indicating that he had not meant to call the agency). This is the cognitive function. Third, it indicates that he is willing to engage in the conversational encounter with the agent. This is the conversational function. Fourth, that he accepts the topic announced by the agent, namely, the "let's solve your problem" topic, implicitly suggested in the previous statement by the agent. This is the rhetorical function. Fifth, that he accepts the role taken by the agent and the role assigned to him, i.e., he accepts the relationship established between them, namely that of agent and customer (whereas if the caller were not playing the role of customer, e.g., the caller is the agent's husband, then his first utterance would probably address the roles being played, informing the agent that he is not playing the role of customer but that of husband, e.g., "Hi, honey. It's me."). This is the interpersonal function. And finally, he signals his own mood, apparently here along the lines of showing caution and concentration on presenting the problem. This is the emotion function.

The two "ah" tokens and the "um" token serve to indicate that the lines of communication are still working (whereas the silence of a pause might lead the listener to think that a breakdown in the physical lines of communication has occurred); emotionally they signal the speaker's hesitation; conversationally they serve to signal that the turn is not yielded yet; rhetorically they serve to indicate that what follows is an integral part of what preceeded; interpersonally they signal that the speaker regards the relationship such that he feels it is not inappropriate to hedge in his words (compare this with the less relaxed speaking protocol of a private to a general during a round of inspection, for example); and cognitively they serve to indicate that the speaker is searching for the appropriate words to express himself.

As we can see from this brief example, a meta-statement performs not only one single action but many different actions simultaneously. Our problem, then, is to determine what these actions are, i.e., we need to establish how we are to go from the meta-statement's surface form to the meta-statement's six-tuple representation. Determining how a word or a sequence of words is to be mapped to their meaning is generally called *the parsing problem.*

3.6 The Parsing Problem

In general, the parsing problem is concerned with mapping a sequence of words to a meaning structure representing the "meaning" of those words. Parsing a string of words then consists of mapping the string from the lexical space to the semantic space. As we noted above, the semantic space corresponding to meta-statements are actions along the six identified dimensions. Thus the meaning of a meta-statement is a tuple denoting the actions performed by uttering the meta-statement being parsed. However, as in any semantic domain of discourse, the surface form of a meta-statement is never exclusively associated with one single meaning. That is, there is no one-to-one mapping from the set of meta-statements, i.e., surface forms, to the set of corresponding actions.[4] Any one given surface form can be used to express much more than one single, particular action. As an example, take the meta-statement "yes, sir." Consider this meta-statement's "meaning" when uttered by a private to a general in response to a latter's order. In this case, the statement "yes, sir," would probably be mapped along the interpersonal axis to mean: "I acknowledge your order, and I consider it to be my duty to carry it out...." Now consider the same statement when uttered by an adult to a child. Depending on the situation, it may very well be the case that the adult is reminding the child who is the adult and who is the child [SANF88]. The "yes, sir" in other words, may very well be used in this case to impart a sense of irony. As another example, take the often used meta-statement "well". If analysed along the emotion axis, we find that, among other functions, it can be used to express hesitation, as in "Well... I don't know if I should;" surprise, as in "Well! I never expected that;" indignation, as in "Well! I've never been so insulted in my whole life;" or perhaps a sense of expectancy, as in "Well?... don't you have something to tell me?" And there are even more possible usages of the "well" meta-statement (see Chapter

[4] We shall use the term "meta-action" to refer to these actions.

Two). The point, however, is that the meta-statement's surface form in no way gives us what meaning we are to tag to it.

The question we must address then is: how are we to determine the meaning of a meta-statement? From what elements should we construct our interpretive system, and what procedures should we employ to identify the meta-actions that correspond to the meta-statements we are parsing?

On Attributing Meaning to Surface Forms

The classical paradigm for modeling language has roots that date as far back as the third century B.C., when a theory of signs and language was advanced by the *Stoics.* The Stoics argued that philosophy is divided into three domains: physical, logical, and ethical [RUSS45.258]. The most important consequence of this division is the creation of a school of thought that "separates sensation and perception as part of how we know the truth from the five senses as part of the psyche" [YNGV86]. Consequently, a theory of language based on this dichotomy would view language *independently of the psychological and social contexts in which it is used.*

The opposing view, which emerged with the advent of the scientific method, rejects in principle the Stoic dichotomy and argues that the study of language, and thus its modeling, must be based on observable data--data gathered from actual language use, taking into account empirically supported psychological and social theories:

> We can use the data obtained from people to understand people not indirectly through language but directly from the point of view of how they communicate. The unjustified assumptions of the traditional objects of language could then be given up.... It would be possible to develop explanations of linguistic phenomena of the sort we have been striving for, in terms of people, directly on the basis of the evidence from people, not indirectly through language. Such linguistics would be human linguistics rather than a linguistics of language. It would seek a scientific understanding of how people communicate [YNGV86.38].

Belonging to the former school are theorists such as Chomsky [CHOM65] who treated language (at least initially) as "an abstract device ... dissociable from the uses, users and

functions of language" [LEVI83.33]. When the notion of generative grammars was introduced [CHOM65], and with the advent of the computer in the 1940's, it became commonly believed that the parsing problem could be solved with the application of the theories developed in the field of structural linguistics, and that, given the power and speed the digital computer afforded, it was only a matter of time before problems such as converting speech into writing or machine translation could be completely solved. All that needed to be done, it was held, was to supply a computer with a body of rules (e.g., an automata description of a language) and the solution would follow easily [CHOM65.3].

Needless to say that it was not long before it dawned on those working on the problems of language that the state of the art of the theories as well as that of the technologies of the day fell short of providing a solution. As Chomsky later noted:

> It has, I believe, become quite clear that if we are ever to understand how language is used or acquired, then we must abstract for separate and independent study a cognitive system, a system of knowledge and belief, that develops in early childhood and that interacts with many other factors to determine the kinds of behavior that we observe; to introduce a technical term, we must isolate and study the system of *linguistic competence* that underlies behavior but that is not realized in any direct or simple way in behavior. *And this system of linguistic competence is qualitatively different from anything that can be described in terms of the taxonomic methods of structural linguistics, the concepts of S-R psychology, or the notions developed within the mathematical theory of communication or the theory of simple automata [CHOM68].*

To illustrate what Chomsky means, let us take an example. The sentence "A little boy kicked the big blue ball" is a sentence that is said to be, using Chomsky's terminology, *well-formed.* That is, a *competent* speaker of the English language would consider such a sentence "good English" and would agree that "it makes sense." If we examined the sentence's grammatical structure, we would find that it translates to: *[Article] [Adjective] [Noun] [Verb] [Article] [Adjective] [Adjective] [Noun].* The question now is: are *all* sentences with a grammatical structure such as this one well-formed sentences? Consider the sentence: "The spiteful pillow bought a stupid yellow adjective." This sentence has the same grammatical structure as the sentence about the little boy, and yet, it is obvious that it is not a well-formed sentence, i.e., pillows are not endowed with the faculty of feeling spite, or of feeling anything; pillows do not

have the ability to buy things; adjectives are not commodities bought or sold; and adjectives cannot be described as stupid or as yellow. In other words, the sentence "does not make sense." Consequently, in answer to our question, we conclude that *not* all sentences that have the same structure as the sentence about the little boy are well-formed. The restrictions imposed by the syntax do not give us enough information about what sentences are well-formed and what sentences are not well-formed. We must also have information about the *meaning* of words. For example, we must know that pillows are inanimate objects, and that inanimate objects do not feel [VIRK91]. With this information, we would then know that "The spiteful pillow" is not an acceptable sequence of words, i.e., it is not well-formed.

Consequently, the notion of solving the parsing problem by developing a syntactic grammar (no matter how elaborate) was dispelled, and purely syntactic parsers were replaced by parsers that took into account not only the lexical type but also the meaning of the words parsed. Such parsers became to be known as *semantic* parsers.

Semantic parsers that receive information from a grammar as well as information from the meaning of the words being parsed will therefore accept only a proper subset of the sentences a purely syntactic parser would accept: with its knowledge about the meaning of words, a semantic parser can detect cases where a sentence does not make sense because the combination of the words violates some semantic law, e.g., bying adjectives. Of course, what set of sentences the parser will accept as well-formed will depend on how word meanings are encoded. The more information about the semantics of the words, the more restrictive the parser will become.

An important characteristic that syntactic and semantic parsers have in common is that they are both *context free.* That is, the mapping from a linguistic form to a meaning structure does not take into consideration the context within which the mapping is done. Instead, meaning is tagged to a form by taking into consideration only information that is resident in the parser: i.e., in the case of syntax the body of syntax rules, and in the case of semantics the syntax rules along with the semantics attached to the words being parsed.

This brings the discussion to the third "branch" of language, *pragmatics.* As Morris states in [MORR38.6], there are three separate branches in the inquiry about signs [LEVI83.1]: first is the study of "the formal relation of signs to one another"--i.e., syntax; second is the study of "the relations of signs to the objects to which the signs are applicable"--i.e., semantics; and third is the study of the "the relation of signs to interpreters"--pragmatics. That is, to study language, Morris holds that all of these three aspects must be taken into consideration.

To illustrate the need for the pragmatic aspect of language, let us take an example. Consider the meta-statement "I understand" uttered by some speaker A in response to speaker B's meta-question "Do you understand?" How is speaker B to interpret this meta-statement? What "meaning" is he to attribute to it? Is he necessarily to take it as though A had meant: "I have fully understood the meaning of what you said, and no further explanation is necessary at this point...."? That is, is he to take the meaning of the meta-statement to be the *surface meaning* of that statement, i.e., the "semantic" meaning provided by the meaning of the individual words that make up the statement? But what if B has a strong suspicion that A has not really understood what was said? Would he still take the utterance to mean the same thing? Obviously not. The "I understand" statement would then take on a completely different meaning for B. B may very well take A's utterance to mean: "I don't really understand, but I just have no time to chat at this point," or as "I don't really understand, but I don't want to talk about this particular subject," or perhaps as "I don't understand, but I don't want you to think I am dumb." In none these cases would the *surface meaning* of "I understand"--i.e., the meaning inherited from the semantics of the words--correspond to the "true" meaning of the statement--"true" in the sense of how the interpreter sees it. Obviously then, word meanings do not suffice to attribute the true meaning of an utterance. As Grice points out [GRIC57], there is a difference to be made between *speaker-meaning* and *sentence-meaning.* Irony, for example, exploits precisley this difference. The sentence "Linguistics is fascinating" has the sentence-meaning "Linguistics is fascinating" but may have as speaker-meaning "Linguitics is deadly boring" [LEVI83.17].

Consequently, our conclusion is that syntax and semantics do not provide us with the solution to the parsing problem. Linguistic structures and meanings of words alone do not contain all the information that is expressed during communication. The context within which communication is taking place must be brought into consideration. That is, we must model pragmatic aspects such as who the speaker is, who the listener is, what relationship exists between the speaker and the listener, between the speaker and his utterances, where the talking is taking place, when it is taking place, etc.

In our paradigm, the pragmatic context corresponds to a model of how an interactant views the conversational situation. In other words, *we propose to build a model of the internal state of the person making the interpretations*, viz., the person attributing meaning to the utterances received by that person. Meaning for us is defined only in terms of how the interpreter defines it. There are no meanings existing independantly of perceivers: meaning is defined by perception. And since perception is in function of the perceiver's state, we must then construct the perceiver's state in order to construct meaning.

In short, in order to address the parsing problem, we propose to construct a model of the world as it is perceived by the interpreter. In the next section, we give an overview description of this model.

3.7 The Interpretive System

The task of the interpretive system is to determine the meaning of meta-statements. As we argued already, in our paradigm we hypothesize that the meaning of a meta-statement is a tuple of actions. That is, our parser receives as input meta-statements, and produces as output a tuple of actions.

To determine to what actions a meta-statement is to be mapped, our parser receives information from the following three sources: (1) state parameters, (2) action grammars, and (3) surface interpretations.

In what follows, we begin by first describing how we model actions in our system, and then we proceed with a description of the components needed during the parsing of a meta-statement.

3.7.1 Action Models

For the action model, a list of basic actions is specified for each of the six dimensions of the system. These lists of actions correspond, in a sense, to our lexicon of meanings that meta-statements can take. In the case of the communicative, cognitive, and emotion dimensions, we identify two types of actions: *signalling state actions,* and *checking state actions.* For example, the utterance "I hear you" when interpreted along the communicational axis would be the action Signal-Own-Ability-to-Receive, while "Do you understand?" along the cognitive axis of interpretation would be the action Check-Other's-Understanding.

With regard to interpersonal actions, an action is characterized in terms of its face impact--that is, in terms of its *politeness.* In order to determine the face impact of an action, we make use of the concepts of *face threatening acts* (FTA's) and that of *politeness strategies,* presented in [BROW87]. An example of a face threatening act would be the action Insult-Other, while a politeness strategy action would be the action Admit-Impingement-On-Other.

Along the rhetorical dimension, a set of actions is associated with each of the rhetorical connectors we shall identify. For example, with the category 'Digression' we have actions such as Open-Digression, Close-Digression, Abort-Digression, etc.

With the conversational dimension, we also associate a set of actions. The actions Open-Conversation, Close-Conversation, Hold-Conversation are examples of such actions.

3.7.2 Parser Components

While parsing the meaning of a meta-statement, the parser receives information from three sources. We discuss each in what follows.

1. State Parameters

As we argued in section 3.6, in order to model what an utterance means, we must first model the state of the interpreter. What an utterance means is not attached to or inherent in the lexical form of the utterance. What an utterance means depends on the conditions within which the utterance is used. Therefore, to "understand" what an utterance means, we must model the context where it is used.

We have argued that there are six axes of interpretation for meta-statements. Therefore, to determine the meaning of a meta-statement, we must build a model of the conditions of the discursive system along each of the six identified axes. For instance, to understand what a meta-statement means communicationally, we build a model the communicational state of the conversers. For this, we use parameters to register the physical communicational state of the interactants, such as whether they received a message clearly, whether they are able to give or receive a message, whether they have finished transferring their message, etc. Similarly, for the cognitive state we use parameters to register the cognitive state of the interactants with repect to what is being said in the conversation. What is being said is represented in our system by *meaning structures,* i.e., semantic units representing the information exchanged between the conversers [VIRK91], and the cognitive state of an interactant is measured with respect to these meaning structures--what meaning structures the interactant thinks he understands, and which ones he feels he does not understand.

When modeling the emotional state of an interactant, we need to associate with each meaning structure not the state of understanding, but the emotion evoked by that structure.

Therefore, a model of emotions is needed. For this we have adopted the model developed in [ORTO88].

The interpersonal state is modeled in our system in terms of the *roles* played by the interactants. An interactant's behavior depends on the behavior expected of him and on the way he is perceived [BALE70], [BALE50]. In our model a role is played relative to four entities: the other interactant *(status role),* the group within which the interactant exists *(position role),* the organization to which the role player belongs *(situs role),* and society at large *(station role)* [BATE75]. We then establish the norms of behavior expected of the interactants. These norms will then be the basis upon which are defined what are called *face wants and sensitivities* [GOFF55], [BROW87]. An interactant's face wants and sensitivities play a crucial role in the way the interactant perceives his partner's actions, i.e., whether he will consider the actions performed by his partner to be polite or impolite.

In the case of the rhetorical state, we are modeling the *structure* of the conversation. For this, we define the concept of a *rhetorical unit.* A rhetorical unit is any stretch of talk where a unit of information is communicated. Rhetorical units are connected by *rhetorical connectors,* and the role a unit plays in the overall structure is a function of what the connector connects to other units. The elements of a list, for example, are connected by the 'list' connector, and together they form a larger unit--a list.

Finally, the conversational state of a conversation models the state of the discourse exchange, e.g., whether the conversation is at the opening stage, or the closing stage, whether it is "on hold," or whether it has been closed, etc.

2. *Action Grammars*

A second source of information used by the parser is what we call *action grammars.* An action grammar specifies what sequences of actions are possible along a specific dimension. For example, after a speaker proposes to a listener to engage in a conversation (a meta-action

along the conversational dimension), only two actions from the listener *along the conversational dimension* are possible: (1) the listener accepts the proposal, or (2) the listener rejects the proposal. By consulting the conversational action grammar, the parser in this example would receive information regarding what actions may follow the previously performed action. In this case, the the meta-statement being parsed may be mapped to one of the two conversational actions: "accept open move" or "reject open move."

Action grammars are in effect a means of determining *the context* in which the utterances being parsed occur. Action grammars, in other words, help us limit the possible spheres of interpretation by limiting what can follow what in terms of actions.

3. *Surface Interpretations*

The third source of information used by the parser is the *surface interpretation* of meta-statements. With each meta-statement we associate a set of surface interpretations. These surface interpretations correspond to the meaning a purely semantic parser would attribute to the meta-statements. For example, the surface interpretation of "I understand" would be the cognitive action: Signal-Own-Understanding, while that of "I don't understand" would be: Signal-Own-Non-Understanding. Of course, as we have already argued, surface interpretations do not necessarily correspond to either what the speaker means or what the listener perceives. Often, completely different interpretations are made. However, in situations where no information is present indicating that the meta-statement we are parsing has a meaning different from its surface meaning (i.e., in situations where the relevant state parameters have no values suggesting a contrary interpretation), the surface meaning may quite possibly correspond to the actual meaning.

3.8 Dynamics of the System

As we said in section 3.6, our approach to solving the parsing problem is based on modeling the interpreter by taking into account *human* characteristics rather than language factors. In section 3.7, we described the major parts that make up the interpretive system. We saw that two models need to be constructed: a model of the state of affairs--i.e., the elements of the conversation--and a model of the actions performed by the interactants involved in the interaction. In this section, we give an overview discussion of the overall dynamics of the interpreter, i.e., how the parser works as an integrated system.

A fundamental characteristic of our system is its time based nature. Interpreter states hold only for intervals of time and then change, and meta-actions are performed at certain points in time. The system's mode of operation and nature, therefore, are inherently transitional. That is, the interpreter is perpertually changing as it reacts to the environment with which it interacts. Unlike the traditional parser, our system is context-sensitive and therefore time dependent.

As we said, the interpretation of an utterance is totally a function of the interpreter's state. We can talk only about 'perceived' actions and never about 'intended' actions. In our dialogue situation, we determined the existence of six sub-states that need to be modeled. Consequently, the interpretation of a surface utterance will take into account these six states. As we also said, a meta-statement's interpretation usually consists of more than one action, and possibly of six simultaneous actions: an action along each of our six axes of interpretation. Figure 1 schematizes how an utterance T is converted to a six-tuple interpretation.

Once the six interpretations are produced, the system will then react to the actions perceived. That is, its state model will be changed by what it understood. The result, then, is a modified--a new--interpreter.

Our parser, therefore, is dynamic in nature. Unlike a static parser, which is built and preestablished once and for all (e.g., a syntax parser), our parser changes as it reacts to the world it communicates with.

3.9 Aims of this Work--A Disclaimer

Our aim in this thesis is to build a framework within which a conversational system sensitive to pragmatic aspects of talk can be constructed. The system we shall specify in what follows is by no means a "final product." On the contrary, we expect that many extensions and modifications will need to be made as work on the system progresses. At this incipient stage of the development, our goal is to establish the structure of the system and to identify the various engines needed to make the system a dynamic parser.

3.10 Summary

In building an interpretive system for handling the parsing of meta-statements, we have determined the existence of six axes of interpretation. We hold that the meta-action corresponding to a meta-statement is a tuple of actions, each action belonging to one of the six identified axes. The problem we face then is determining what actions to attribute to a given meta-statement. For this we have found that we need to create a model of the conversational state. That is, in order for the interpretive system to determine what functions a meta-statement is fulfilling, it must have a model of the state of the world it is concerned with, i.e.,

a model of the conversation. A conversation's most basic elements are (1) the utterances exchanged, i.e., talk, and (2) the interactants involved in the exchange, i.e., talkers.

In order to model talk we need to model two aspects of the conversation: (1) the rhetorical aspect, and (2) the conversational aspect. To model the talkers we need to model four aspects of the interactants: (1) their communicational state, (2) their cognitive state, (3) their emotional state, and (4) their interpersonal state.

In chapters Four through ten that follow, we examine each of these aspects more closely. Chapters Four and Five will present the models we are using for the rhetorical and the conversational aspects, respectively. In Chapter Six, we examine the issue of modeling roles. In Chapter Seven, we present a model of politeness. Chapter Eight contains a discussion on emotions, while chapters Nine and Ten focus on the cognitive and communicational aspects of meta-statements, respectively. Finally, in Chapter Eleven, we descibe how the elements of the model specified are to be integrated to create a dynamic parser.

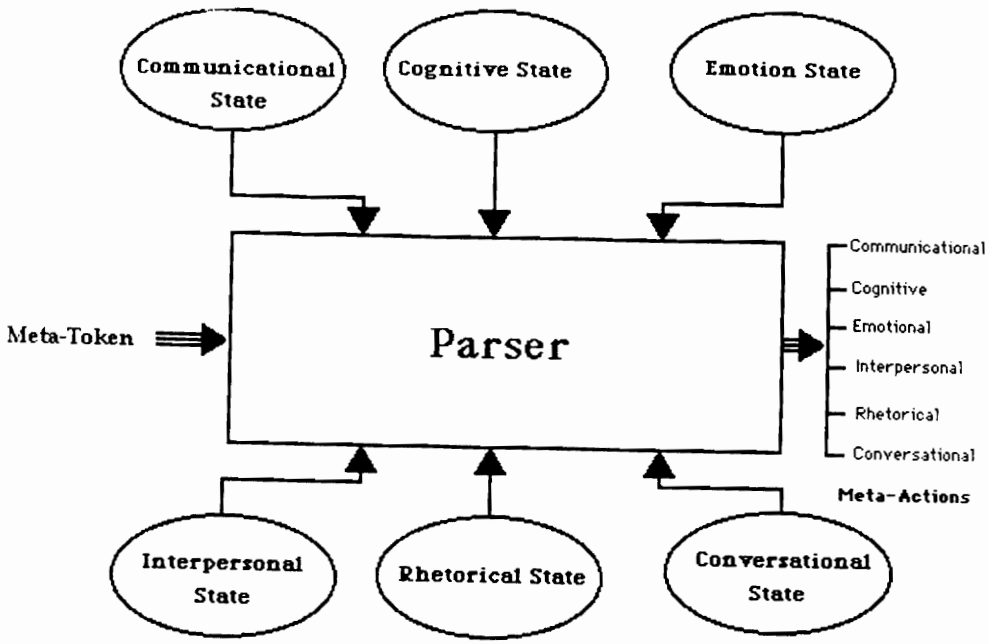


Figure 1. Token Action Transition

Chapter 4. The Rhetorical Model

Analysis of the briefest conversational exchange will readily reveal that dialogue is far from being a simple game of action and reaction. The notion that each conversational move depends only on the preceding move, or that a conversation is linear in structure is overly simplistic. Even in the most casual conversations, where the interactants are talking "for the sake of talking," and where the exchange is "spontaneous," we find that the structure of the conversation is far from being simple.

Among the functions fulfilled by the meta-statement is the *structuring of talk.* We call this "the rhetorical function" of meta-statements. When raising an issue, for example, or switching from one topic to a new topic, a speaker will need to signal to his partner that an issue has been raised or that a new topic has been introduced. Without this signal the listener might become confused and perhaps even unable to follow his partner's flow of ideas. One way to signal transition is through meta-statements. The statements "on another matter" or "concerning something else," for instance, are examples of meta-statements that might be used to signal transition from one topic to another.

In this chapter, we examine the rhetorical function of the meta-statement: how is the meta-statement used to connect parts of discourse? As we will see in the examples of this chapter, very often the presence of a meta-statement can make the difference between a perfectly intelligible sequence of utterances and a totally incoherent piece of dialogue. When the relationship between two parts of the dialogue is not immediately deducible from what is said, a speaker will often need to explicitly mark how the parts are to be connected with one another. This explicit connection is often achieved by using a meta-statement.

In the following analysis, our goal is to build a model to account for this connecting function of meta-statements. With this model we will be able both to parse the meta-statement's rhetorical function--i.e., understand the role it plays in the structure--as well as to build and organize discourse structures using meta-statements.

The first step in building our model is to address the problem of how to subdivide talk; that is, we must determine what *the unit of talk* is. The second step is to build a model of the rhetorical unit, which in turn requires a *rhetorical state* model and a *rhetorical action* model of the conversation. The state model necessitates a classification of rhetorical connector types along with the parameters each type requires for its definition. Whereas for the action model we will need first to specify the rhetorical actions and their parameters and then proceed to modeling sequences of actions through *action grammars.*

But before going any further in our analysis, let us at this point take an example to illustrate some of the concepts we shall be discussing.

An Example

The following conversational excerpt, although brief, will illustrate the high degree of structure one almost always finds in conversations. As in previous examples, we have enclosed meta-statements between square brackets. We have also partitioned the transcribed excerpt so as to facilitate the demarcation of the rhetorical units. Spans of lines will be denoted by square brackets, e.g., [2,7] indicates the stretch of lines from line (2) to line (7), inclusive. An ordered set of lines is enclosed between parentheses, e.g., (2,3,5) means lines (2), (3), and (5).

A rhetorical unit can be composed of spans of lines and of ordered sets of lines. So then, the rhetorical unit ((2,3,5),[7,9],(11,13)) consists of lines (2), (3), and (5), lines (7) through (9), and lines (11) and (13).

Let us now examine the following conversation excerpt:[5]

1 -CUSTOMER.
2 [Um]
3 in April,
4 April 5th
5 I'm gonna be going
6 [um]
7 to Los Angeles ...
8 -AGENT.
9 [Uh-huh.]
10 -CUSTOMER.
11 [and]
12 I would like
13 [ah]
14 some information on
15 [ah]
16 flights out that way.
17 -AGENT.
18 [Okay.]
19 [All right.]
20 Do you know how long you'll be staying if you leave on the 5th?
21 When you'll be returning?
22 -CUSTOMER.
23 [When returning?]
24 [Um]
25 [let's see,]
26 it'd probably be for,
27 it's not certain right now
28 [but]
29 it'd probably be for
30 [like]
31 10 days to two weeks.
32 -AGENT.
33 [Okay.]
34 [So]
35 you would be staying at least one week
36 [then,]
37 [okay.]
38 -CUSTOMER.
39 [Yes, ma'am.]

---

[5] Lines [9-32] of conversation phone.01

The fragment of dialogue presented above can be partitioned into three units. In the first unit, the customer furnishes the agent with some information (lines [2,31]). In the second unit, the agent signals that she has received enough information and that she is now about to draw an inference (lines [33,34]). And in the third unit (lines ([35,39]), the agent draws her inference based on the information the customer gave, with the customer confirming the validity of the inference drawn. Figure 2 gives a schematic view of the highest level structure of this exchange. We call this decomposition of the dialogue a level-1 decomposition. The first unit (lines [2,31]), which we shall call "the data-request unit", is basically the unit where the data about the departure date, the return date, and the destination location are requested and given. The second unit, which we shall call "the connector unit," connects units one and three. The third unit (lines [35,39])--"the inference unit"--is where some conclusions and inferences are drawn based on the data given in the first unit.

Each of these units is in turn made up of smaller units. The data-request unit is also composed of three units. In the first unit, the customer gives the agent some information and then makes his request (lines [2,16]). In the second unit (lines [17,21]), the agent requests an elaboration. And in the third unit (lines [23,31]) the customer furnishes the requested elaboration.

The inference unit is also decomposed into two smaller units: unit one (line (35)) is where the agent draws her inference, while in unit two (lines ([36,39]) the inference is closed. Figure 3 shows the level-2 decomposition of the dialogue.

The units of the level-2 decomposition are themselves composed of even smaller units. The first unit of the data-request unit is composed of four smaller units. The first unit (line (2)) is a meta-statement used to open the second unit (lines [3,7]), while the third unit (lines (9,11)) is used to connect the second unit with the fourth unit (lines [12,16]). Semantically, the second unit contains background information that sets the stage for the request that is about to be made in the fourth unit--i.e., the request for information resulting from the fact that the customer is planning to take a trip.

Examining the elaboration unit we can distinguish between two smaller units. First (line (23)), the customer opens the elaboration by echoing, or rewording, the agent's request for an elaboration (line (21)). This is unit one. The second unit (lines [24,31]) contains the elaboration. Figure 4 gives the level-3 decomposition.

If we go one more level down, we find that the unit composed of lines [3,7] is in fact the combination of four units. Unit one is lines [3,4], unit two line (5), unit three line (6), and unit four line (7). Unit one and two are implicitly connected by a conjunction. The third unit is however an explicit conjunction, and it connects units two and four. We notice that unit one contains two smaller units--the second unit (line (4)) is an elaboration of line (3). Similarly, the unit in lines [12,16] is composed of five smaller units: line (12), line (13), line (14), line (15), and line (16). Lines (13) and (15) are conjunctive units, and they function to connect the first, third, and fifth units.

The elaboration (lines [24,31]) also is composed of even smaller units: the first (lines [24,25]) opens the elaboration, which begins only at line (26) and is interrupted with a digression (line (27)). The digression is closed at line (28) and the elaboration resumes from line (29) to line (31). Thus we have four units: unit one is lines [24,25], unit two lines ((26),[29,31]), unit three line (27), and unit four line (28). Figures 5 and 6 give us the level-4 and level-5 decompositions, respectively.

As we can see, brief as it is, this stretch of dialogue requires five levels of abstraction to model its structure. The conclusion to be drawn from this example, then, is that in order to model a discourse, the model that should be adopted must allow for building structures.

Our first step in developing this model is to describe how we will model a rhetorical unit.

4.1 The Unit of Talk

In this section we examine the question of determining what discourse unit should be adopted for the development of our rhetorical model. It is obvious that the utterance level of analysis is inadequate for our purposes. Utterances are seldom units in themselves, and very often we find that an utterance is in fact itself an elaborately structured piece of discourse. The problem then is determining what sort of an entity is the rhetorical unit. Should it be a grammatically correct sentence, a phrase, a conversational turn? How are we to demarcate the boundaries of a unit? Should we use syntactic delimiters, or perhaps should we take the propositional content of what is being said into consideration?

It is now generally accepted in the field of Discourse Analysis that any adequate model of discourse must account for the existence of structures in talk. Speech act theorists, for example, have started to see the need for observing speech act sequences on a range larger than that of the traditional adjacency pair view. Van Dijk suggests the use of *macro speech acts* [DIJK77], whereas Reichman suggests using what she calls *context spaces.* And according to Halliday, "The speaker is obliged to chunk his speech into information units. He has to present his message in a series of packages" [YULE83].

Thus, whether we are dealing with van Dijk's macrostructures, Reichman's context spaces, or Halliday's information units, what is clear is that a paradigm for grouping must be adopted. The difficulty, however, is determining what the unit of grouping is.

Some argue that the sentence is the unit on which the analysis of discourse should be based. Chafe holds that: "Syntactically there is a tendency for idea units to consist of a single clause" [CHAF70].

Harris, on the other hand, suggests the term "utterance" for a unit of speech. According to his definition, an utterance is "any stretch of talk by one person, before and after which there is silence on the part of that person" [HARR]. While Halliday maintains that: "the tone

group is a phonological unit that functions as realisation of information structure" [HALL67.203].

In our opinion, the syntactic approach to partitioning natural speech is wrong for a very simple reason: the language of conversation is quite often syntactically faulty. People have a tendency to stop in the middle of sentences, start new sentences before finishing ones they had initiated, and sometimes even finish sentences someone else had started. In any case, the syntactically based paradigm of partition would simply not be able to handle such cases.

Harris's utterance suggestion is also inadequate. Between two silences a speaker can build quite complex structures of discourse. We would hardly be justified in considering that such structures could be candidates for a unit of talk.

On the other hand, Halliday's suggestion of demarcating units of information by tonal changes comes close to our own view. The only difference is that Halliday makes use of a paralinguistic[6] means of communication--tonality of voice--whereas we are concentrating on the verbal statement. This is not to say that there is any incompatibility between the two views. On the contrary, we believe the two views to be complementary. In any case, we believe that a meta-message, whether verbal or non-verbal, is usually used to delimit the boundaries of an information unit.

Thus, the unit of discourse we shall adopt for our model is the *information unit.* That is, a rhetorical unit is any stretch of discourse, whether it is one word long or whether it is made up of many conversational turns, where a unit of information is expressed. The information unit, we hold, is demarcated by meta-messages. These meta-messages are used by the speaker to signal to his listener how to subdivide and structure stretches of talk. In our case, we are analyzing one type of meta-message, the verbal meta-message. Verbal meta-messages are quite often used in natural dialogue, and one of the functions they fulfill is the demarcation of information units.

---

[6] Goodwin [GOOD81] also suggests that paralinguistic signals are used in the demarcation of information unit boundaries, especially through the exchange of gazes.

In addition to the physical structuring of talk--i.e., at what points in the talk the units begin and end--meta-statements are also used to define the *semantic* role played by the identified units. That is, meta-statements give information as to how to relate and connect the *meaning* of information units to one another. In order to account for this function of the meta-statement, we must therefore identify the semantic roles that information units can play.

For this we set up a classification of how information units can be connected. It is with respect to these connections that the semantics of the units is defined. That is, an information unit--i.e., a rhetorical unit--is defined in terms of its rhetorical function: in terms of the role it plays in the structure.

In our conversational set up--the telephone dialogue--we can therefore expect at least two types of rhetorical delimiters: the tonal delimiter and the verbal delimiter. Since we are dealing only with text--transcribed conversations--we will therefore not examine the tonal delimiter. We will focus only on the verbal rhetorical delimiter.

In what follows, we examine the model we have adopted for representing rhetorical units.

4.2 A Model of the Rhetorical Unit

The level-1 decomposition of the conversation given in section 4.1 gave us a structure composed of three rhetorical units: the first unit being connected to the third unit by the second unit. We saw that in the third unit the agent drew an inference based on the information presented in the first unit. The second unit's function, we saw, was to signal that the third unit was to be interpreted as an inference. Each of these units, then, played a role in the structure of the dialogue. The first and the third units in this example are said to be *component units.* The second unit, on the other hand, is an *action unit,* since through it the agent performs a rhetorical action, i.e., she connects the component units.

In this section, we describe the rhetorical model we use for each of these types of units. We begin the discussion with the component unit model.

4.2.1 The Component Unit

There are seven characteristics of a component unit that we need to model. They are its *identifier, coordinates, class, connector type, connector parameters, furnisher,* and *aborted-status.*

Identifier: Each component unit is assigned a unique identification tag. This tag is called the UNIT-ID.

Coordinates: The coordinates of a rhetorical unit identify the unit's *physical* address in the discourse. That is, a unit's coordinates locate the unit's position within the sequence of talk. We adopt the simple addressing scheme where a unit is demarcated by two points: the *start* and *end* points. A point in our system is a pair of integers (T,W), where T indicates the *conversational turn sequence number,* and W is the *word sequence number* within the turn in question. Thus, the first unit of the level-1 decomposition of the example in section 4.1 has coordinates ((1,1),(5,29)), since it begins with word 1 of turn 1 and ends with word 29 of turn 5.

In most cases the rhetorical units encountered are stretches of talk, but in many instances we may need to model "empty units," that is, structures that do not contain any stretches of talk. The need for a construct that allows for empty units becomes obvious if problems such as speech planning, conversational interruptions, and implicit connections are to be handled. The customer of our conversation, for example, most probably did not plan the exact wording of his utterances. But he must have had at least an idea, a schema, vague as it may have been, as to the general structure of how he was going to begin his conversation. If our system is to simulate the converser's behavior, it must have a way of "thinking" about

structures abstractly and independently of the actual utterances from which those structures may (or may not) be ultimately composed.

We achieve this by allowing the coordinate parameters to be unspecified. That is, we give the system the ability to model a unit without specifying the values of the start and stop parameters. Thus, we allow three types of coordinate states: (1) both the start and the end parameters are specified (the unit is said to be *closed);* (2) the start is specified, but the end is not (the unit is said to be *open);* and (3) neither the start nor the end is specified (the unit is said to be *new).*

The first case occurs when the unit in question has already been encountered during the dialogue: we know exactly where it starts in the discourse and where it ends. In the second case, the unit's start boundary is known, but not where it will end, or whether it will have additional structures added to the ones that already compose it. Needless to say, at any point during a dialogue we will always have at least one rhetorical unit open: the conversation. Until it is formally closed by both interactants, a conversation will remain open.

For the third case, the only thing known about the unit is how it is connected to other units. Physically, the unit does not exist. For example, when a speaker says "in other words," by uttering this meta-statement he has created, for the moment at least, a unit about which we know only one thing: it is supposed to contain a rewording of previous statements. Before the speaker begins his actual rewording, the unit exists only hypothetically--i.e., it does not have any discourse coordinates.

We must note, however, that a closed unit may at any point during the conversation be opened again, and thus once more be made into an open unit. Often, a speaker will tell a story, finish it, and go on with the conversation, only to return to that story later in the talk in order to add some detail that was initially left out. While the speaker is adding his detail, the story, which had been previously totally defined, becomes once again partially defined, and remains so until it is once again closed.

Class: A component rhetorical unit belongs to one of two classes: the *nucleus* class or the *satellite* class [MANN87]. The class of a unit is determined by the connector that connects it

to other units. For example, the component units of a list structure are all nuclei, since the list connector is defined to connect only nuclei. On the other hand, the third unit of the example in 4.1 is a satellite since it is the unit where an inference is drawn, and given the nature of the connector in this case--the inference connector--the first unit is classified as a nucleus, since it contains the information from which an inference is drawn. On the other hand, the third unit, where the inference is drawn, is classified as a satellite (see section 4.4 for a detailed specification of nucleus and satellite requirements).

Connector Type: As we said already, units are defined in terms of the role they play within a structure. A structure's nature is defined by the connector that operates on the constituent components. Thus, to define a component unit we must also specify what connector is applied to it. For example, in the case of a list item unit, the connector applied to the unit is the list connector. In the case of the two component units of our example, both units, the satellite and the nucleus, are connected by an inference connector.

Connector Parameters: With some connectors, in addition to the connector's type, we must also specify other parameters. For example, the list connector has as parameters the length of the list and the sequence number of the item currently being furnished.

Furnisher: This parameter specifies which speaker is furnishing the unit.

Aborted-status: In many cases, the specification of a unit may be cut short before the speaker is able to finish its specification. To model such cases we have the aborted-status parameter. This parameter can have one of two values: *yes* if the unit was aborted, and *no* if it was not. A unit is said to be *aborted* if the parameter has the value 'yes'.

Thus, a component unit is modeled in our system with the following frame:

```
(COMPONENT-UNIT
  (UNIT-ID          ?unit-id)
  (COORDINATES
                    (START ?start)
                    (END   ?end))
  (CLASS            ?class)
  (CONNECTOR-TYPE   ?connector-type)
  (CONNECTOR-PARAS  ?connector-paras)
  (FURNISHER        ?furnisher)
  (ABORTED          ?aborted))
```

4.2.2 The Action Unit

We now turn to the second type of rhetorical unit: the *action unit.* In an action unit, the speaker performs a rhetorical action. For example, in the second unit of the level-1 decomposition of the example in 4.1, the agent signals to the customer that in the third unit an inference is to be drawn based on the information given in the first unit. The second unit, is thus called an action unit. (The rhetorical actions identified in our model are given later in the chapter.)

Five parameters need to be specified in order to model an action unit. In addition to a unique identification tag (UNIT-ID), we must also specify the *coordinates* of the unit, the *rhetorical action* performed in the unit, the *actor* performing that action, and the *components* related to the action (specified as a list of component UNIT-ID's).

As with component units, we may also need to specify additional parameters about an action in order to model that action fully. For example, the rhetorical action performed by the statement "My three most favorite composers are:" is a rhetorical action unit where the length of a list is specified. The action is then a "length of list specification" action, while the parameter it specifies--the length--is three.

Thus, an action unit is modeled internally by the following frame structure:

```
(ACTION-UNIT
  (UNIT-ID              ?unit-id)
  (COORDINATES
                        (START ?start)
                        (END   ?end))
  (ACTION               ?action)
  (ACTOR                ?actor)
  (COMPONENTS           ?list-components))
```

As we can see, then, in the case of the component unit, the specification of a rhetorical unit depends on the type of connector that specifies the unit's status and parameters, while in the case of the action unit, the unit's specification depends on the action performed in the

unit. Therefore, in order to specify component units we must have a detailed classification of rhetorical connectors, along with the parameters they require for their specification, while for modeling an action unit, we need a classification of rhetorical actions and their parameters.

4.3 A Classification of Rhetorical Connectors

A two-part classification of rhetorical connectors is presented in this section. Rhetorical units may be connected by an *argumentative* connector or by an *expositive* connector. If the rhetorical units connected are all nuclei, then the connector is said to be *multi-nuclei.* If the connector connects one nucleus with satellites, then the connector is called a *nucleus-satellite* connector.

In what follows, the model's rhetorical connectors are given. With each connector we specify the connector parameters--if any--that need to be specified.

4.3.1 Argumentative Connectors

Rhetorical units are said to be bound by an argumentative connector if the connector that connects the constituent units is based on the *informational content* of the units in question. In the example of section 4.1, the customer connects unit [2,7] with unit [12,16] using an argumentative connector (11). The "and" in line (11) is classified as an argumentative connector because it functions to signal that the information content of the first unit ([2,7]) is to be regarded as *backgound* information with respect to the third unit ([12,16]).

We distinguish between three types of argumentative connectors: *logical, conceptual,* and *contextual.*

Logical Connectors

The logical connector functions to signal the *semantic dependency* that exists between events and actions. A speaker may, for example, describe how the occurrence of some action brought into effect some new state of affairs, or how a state of affairs was the consequence of some action. In both cases, he is pointing to the *logical relationship* between the events.

Seven logical connectors are identified in our model: *causality, conditional, conflict, consequence, enablement, inference* and *violated expectation.* In what follows, we briefly paraphrase the semantics of each connector, illustrating each connector with a brief example. The connector signals in the example are enclosed between square brackets, while the units are enclosed between curly brackets. A unit followed by an S is a satellite, while one followed by an N is a nucleus.

Causality: (nucleus-satellite) The action or state described in the nucleus unit is caused by the action or state described in the satellite unit.

Example:

- {There was a blackout} N
  [because]
- {the storm knocked down the power line.} S

Conditional: (nucleus-satellite) The occurrence of the action or the existence of the state described in the satellite is prerequisite for the occurrence of the action or the existence of the state described in the nucleus.

Example:

- {I will gladly go} N
  [but only if]
- {you pay me.} S

Consequence: (nucleus-satellite) The state of affairs described in the satellite is the result of the state of affairs or action described in the nucleus.

Example:

- {The storm knocked down the power line.} N
  [So,]
- {there was a blackout.} S

Enablement: (nucleus-satellite) The action or state of affairs described in the satellite brings into effect a new state of affairs which makes the action described in the nucleus possible.

Example:

- {Yesterday I won five dollars in the lottery} S
  [so,]
- {I bought myself a new pair of socks.} N

Inference: (nucleus-satellite) The action or state of affairs described in the satellite is inferred from the information given in nucleus (see lines [33,37] of example in section 4.1).

Violated Expectation: (nucleus-satellite) The state of affairs or actions described in the satellite were not expected in light of the expected consequences of the actions or state of affairs described in the nucleus.

Example:

- {He was twice bitten by the killer snake} N
  [and yet]
- {nothing happened to him.} S

Conflict: (multi-nuclei) The actions or state of affairs described in the two nuclei are incompatible and cannot both hold at the same time.

Example:

- You say {you have no money} N
  [But how is it that]
- {you can afford a new dress.} N

Conceptual Connectors

The second type of argumentative connectors is the *conceptual* connector. The conceptual connector is used to compare aspects of the actions or states of affairs described in the nuclei it connects. There are three conceptual connectors:

Similarity: (multi-nuclei) The similarity between the actions or states of affairs described in the nuclei is pointed out.

Example:

- {The bomb blast sounded} N
  [just like]
- {a thunder clap.} N

Difference: (multi-nuclei) The difference between similar aspects of the actions or states of affairs described in the nuclei is pointed out.

Example:

- {A bicycle has two wheels} N
  [whereas]
- {a car has four.} N

Contrast: (multi-nuclei) Different aspects of comparable actions or state of affairs described in the nuclei are pointed out.

Example:

- {A car is faster than a bicycle}
  [but]
- {a bicycle costs less.}

Contextual Connectors

The third type of argumentative connectors is the *contextual* connector. The contextual connector is used to connect a satellite that has information that creates a context in which the interpretation of the nucleus is to be made. There are two contextual connectors:

Background: (nucleus-satellite) The action or state of affairs already described in the satellite is to be used as background context information for the interpretation of information about to be furnished in the nucleus (see lines [2,16] of example).

Justification: (nucleus-satellite) The information already furnished in the nucleus is to be interpreted in light of information about to be furnished in the satellite.

Example:

- {Can you tell me what time it is.} N
  [You see,]
- {my watch isn't working.} S

4.3.2 Expositive Connectors

We now turn to the second type of rhetorical connectors: *expositive* connectors. Whereas in the case of argumentative connectors the information content of the units is the basis on which the connections are made, in the case of expositive connectors the connection is based on the speaker's intentions as to how he wants to present his ideas and how he wants to organize

them. We distinguish between three types of expositive connectors: *composite*, *structural*, and *discursive* connectors.

Composite Connectors

Composite connectors connect units that together form a conceptual whole. We have four composite connectors: *list*, *plan*, *story*, and *entity*.

List: (multi-nuclei) The nuclei connected by the connector form a list structure. Each nucleus contains a list item.

Parameters:

* *List-length:* the length of the list.
* *Item-number:* the sequence number of the current list item being specified.

Example:

- {My three favorite composers are:}
- {Chopin} N
- {Mozart} N
  [and]
- {Bach.} N

Plan: (multi-nuclei) The nuclei connected form a plan. Each nucleus contains a plan goal.

Parameters:

* *Plan-goal:* the goal the plan is to achieve.

Story: (multi-nuclei) The nuclei connected form a story. Each nucleus contains a story episode.

Entity: (multi-nuclei) The nuclei connected describe some entity.

Parameters:

* *Entity:* the entity described.

Structural Connectors

The structural connector is used to organise the form of talk into discourse structures. We have two structural connectors.

Topic: (multi-nuclei) Each of the nuclei connected treat some aspect of the same topic.

Parameters:

* *Topic:* the topic treated.

Digression: (nucleus-satellite) The satellite unit's subject matter is to be regarded as a digression from the main topic, which is the nucleus's subject matter (see lines [26,29] of example).

Discursive Connectors

In the case of discursive connectors, the connector connects two units where the actions or states of affairs described in one unit are described differently in the other unit. There are five discursive connectors.

Summary: (nucleus-satellite) The satellite unit contains a more succinct presentation of the subject matter presented in the nucleus. Meta-statements such as "in short" and "to make a long story short" are used to signal the upcoming summary satellite.

Elaboration: (nucleus-satellite) The satellite unit contains a more detailed presentation of the subject matter presented in the nucleus (see lines [3,4] of example).

Rewording: (nucleus-satellite) A different presentation of the subject matter presented in the nucleus is presented in the satellite (see lines [20,21] of example).

Examples: (nucleus-satellite) The satellite contains information that illustrates a point made in the nucleus. Usually, example satellites are announced with meta-statements such as "for example," and "for instance."

Generalization: (nucleus-satellite) The nucleus contains particular information upon which is based the generalization presented in the satellite. Generalization satellites are usually prefaced by statements such as "so, in general," and "as a rule," etc.

4.4 Rhetorical Actions Model

Having identified the ways rhetorical units are defined and connected, we proceed next to specifying the rhetorical actions that may be performed on those units. We identify with each rhetorical connector a set of rhetorical actions and a rhetorical grammar. In the first section, we discuss the model's rhetorical actions, while in the second section we examine the rhetorical grammars.

4.4.1 Actions

We begin the specification of rhetorical actions by examining those actions that are performed to furnish component units. These actions are necessary independently of the way the units in question are connected. After this we proceed with the specification of further actions that are needed for each particular rhetorical connector category.

Component Actions

As we saw in section 4.3.1, a component unit can be in one of four states: new, opened, closed, and aborted. The actions specified in this section are those that change the states of the units they affect.

We saw that each component unit has a main furnisher. Consequently, actions are distinguished depending on the main--or intended--furnisher of the component unit in question. Thus, if the speaker furnishing some component unit were to abort the specification of the unit, we would then say that he had performed action three--abort unit for self. On the other hand, if he were to interrupt his partner, then he would have performed action seven, namely the 'abort unit for other' action. Actions one through eight are called *furnish component actions.*

These are the actions the speaker uses to furnish components. In the section that follows, we identify the actions needed for each rhetorical category. In each category we will include the action *furnish-nucleus* or *furnish-satellite* (as appropriate, given the connector type). These actions are in fact sequences of furnish-component actions. The acceptable sequences will be given later in the form of the *furnish unit grammar.*

The eight actions listed below are those actions used to furnish component units. To the right of each action we specify the parameters that the action affects. We use 'cs' to indicate that the start coordinate parameter of the component unit is affected, whereas 'ce' indicates that the end coordinate parameter is affected. The 'a' refers to the 'aborted-status' parameter of the unit, the 'f' to the 'furnisher' parameter, the 't' to the 'connector type', and the 'l' to the 'class'. We now give the actions:

1 - Open new unit for self (cs,a,f).
2 - Open aborted unit for self (cs,a,f).
3 - Abort unit for self (ce,a,f).
4 - Close unit for self (ce,a,f).
5 - Open new unit for other (cs,a,f).
6 - Open aborted unit for other (cs,a,f).
7 - Abort unit for other (ce,a,f).
8 - Close unit for other (ce,a,f).

Connector Actions

We make the distinction between three modes of signalling the connector's type. *Preconnecting* refers to the case where the connector type is signalled before the elements connected are given. For example, "the following is a list of people who have won the lottery,"

defines a list structure before the identification of structure's elements has begun, i.e., it pre-connects the nuclei.

On the other hand, *inter-connecting* refers to the case where the connection is performed after the specification of at least one unit. Most of the time, the connector is of this type. The meta-statement "in short," for example, is usually used after the nucleus and before the satellite.

In the case of post-connecting, the connector's type is signalled *after* all the components have been specified. For example, in the sentence "I need something sharp; a knife, for instance," the connector that links the nucleus "I need something sharp" and the satellite "a knife" is the example connector. In this case, the connector is signalled with post-connector "for instance."

Actions nine through fourteen below are called *connecting* functions:

9 - Pre-connect for self (cs,t,l).
10 - Inter-connect for self (cs,ce,t,l).
11 - Post-connect for self (ce,t,l).
12 - Pre-connect for other (cs,t,l).
13 - Inter-connect for other (cs,ce,t,l).
14 - Post-connect for other (ce,t,l).

The following connectors require only furnishing and connecting actions: *Causality, conditional, consequence, enablement, inference, violated expectation, conflict, similarity, difference, contrast, background, justification, summary, elaboration, rewording, examples, generalization,* and *digression.*

With the other connector types, however, in addition to the furnishing unit actions there are also specific actions. The connector types that require additional actions are: *topic, list, plan, story,* and *entity.* The bold faced Furnish action, as we said, refers to a sequence of component furnishing actions (this sequence corresonds to strings that can be parsed by the Component Furnishing grammar).

We specify the actions for each in what follows.

List:

1 - Open own list (cs,f).
2 - Open other's list (cs,f).
3 - Close own list (ce).

4 - Close other's list (ce).
5 - Abort own list specification (ce,a).
6 - Abort other's list specification (ce,a).
7 - Increase length of own list (sp-length).
8 - Increase length of other's list (sp-length).
9 - Decrease length of own list (sp-length).
10 - Decrease length of other's list (sp-length).
11 - Reopen own self-aborted list (ce,a).
12 - Reopen own other-aborted list (ce,a).
13 - Reopen other's self-aborted list (ce,a).
14 - Reopen other's other-aborted list (ce,a).
15 - Reopen own closed list (ce).
16 - Reopen other's closed list (ce).
17 - Add own list item (sp-itnumb).
18 - Add other's list item (sp-itnumb).
19 - **Furnish** own nucleus.
20 - **Furnish** other's nucleus.

Plan:

1 - Open own plan specification (cs,f,sp-goal).
2 - Open other's plan specification (cs,f,sp-goal).
3 - Close own plan specification (ce).
4 - Close other's plan specification (ce).
5 - Abort own plan specification (ce,a).
6 - Abort other's plan specification (ce,a).
7 - Retract own plan goal.
8 - Retract other's plan goal.
9 - Rearrange own goals sequence.
10 - Rearrange other's goals sequence.
11 - Reopen own plan specification after natural close (ce).
12 - Reopen other's plan specification after natural close (ce).
13 - Reopen other own-aborted plan specification (ce,a).
14 - Reopen other other-aborted plan specification (ce,a).
15 - Reopen own closed plan specification (ce,a).
16 - Reopen other closed plan specification (ce,a).
17 - Add goal to own plan.
18 - Add goal to other's plan.
19 - **Furnish** own nucleus.
20 - **Furnish** other's nucleus.

Story:

1 - Open own story specification (cs,f,sp-episode).
2 - Open other's story specification (cs,f,sp-episode).
3 - Close own story specification (ce).
4 - Close other's story specification (ce).
5 - Abort own story specification (ce,a).
6 - Abort other's story specification (ce,a).
7 - Retract own story episode.
8 - Retract other's story episode.
9 - Rearrange own episodes sequence.
10 - Rearrange other's episodes sequence.
11 - Reopen own story specification after natural close (ce).
12 - Reopen other's story specification after natural close (ce).
13 - Reopen other own-aborted story specification (ce,a).

14 - Reopen other other-aborted story specification (ce,a).
15 - Reopen own closed story specification (ce,a).
16 - Reopen other closed story specification (ce,a).
17 - Add episode to own story.
18 - Add episode to other's story.
19 - **Furnish** own nucleus.
20 - **Furnish** other's nucleus.

Entity:

1 - Open own entity specification (cs,f,sp-part).
2 - Open other's entity specification (cs,f,sp-part).
3 - Close own entity specification (ce).
4 - Close other's entity specification (ce).
5 - Abort own entity specification (ce,a).
6 - Abort other's entity specification (ce,a).
7 - Retract own entity part.
8 - Retract other's entity part.
9 - Rearrange own parts sequence.
10 - Rearrange other's parts sequence.
11 - Reopen own entity specification after natural close (ce).
12 - Reopen other's entity specification after natural close (ce).
13 - Reopen other own-aborted entity specification (ce,a).
14 - Reopen other other-aborted entity specification (ce,a).
15 - Reopen own closed entity specification (ce,a).
16 - Reopen other closed entity specification (ce,a).
17 - Add part to own entity.
18 - Add part to other's entity.
19 - **Furnish** own nucleus.
20 - **Furnish** other's nucleus.

Topic:

1 - Signal topic switch (cs,sp-topic).
2 - Accept topic switch.
3 - Reject topic switch (ce,topic).
4 - Abort own topic switch (ce,a).
5 - Retry own self-aborted switch (ce,a).
6 - Retry own other-aborted switch (ce,a).
7 - Retry other's self-aborted switch (ce,a).
8 - Retry other's other-aborted switch (ce,a).
9 - **Furnish** nucleus.

Figures 7, 8, and 9 show the rhetorical units of the level-1, level-2, and level-3 decompositions of the example in section 4.1, respectively. CU indicates that the unit is a component unit. The unit's class (nucieus or satellite) is also given. AU indicates an action unit. The action performed in the unit is given hyphenated to the 'AU'. In the case where the action is

a pre-, inter-, or post-connector, the connector type is given (e.g., 'Inference' for the level-1 decomposition).[7]

4.4.2 Grammars

Having specified the model's actions, we now proceed to the specification of the action grammars.

In our model, we distinguish between seven rhetorical grammars: the *component furnishing* grammar, *the connector* grammar, and the *topic, list, plan, story,* and *entity* grammars.

The component furnishing grammar specifies the allowed sequences of component furnishing actions. The grammar is given in what follows. The connector grammar specifies the allowed sequences of connecting actions, where the units are connected by one of the following connectors: *causality, conditional, consequence, enablement, inference, violated expectation, conflict, similarity, difference, contrast, background, justification, summary, elaboration, rewording, examples, generalization,* and *digression.* The grammars for the topic, list, plans, story, and entity connectors, being different from the other connector grammars, are given separately.

The grammars follow:

Component Furnishing Grammar

F_A --->

A_1F_{A1}
| B_5F_{A1}

F_{A1} --->

A_4
| B_8
| A_3F_{A2}

---

[7] For the line numbers demarcating the units, refer to the corresponding diagrams; e.g., for Figure 7 see Figure 2.

$| B_7F_{A2}$

F_{A2} --->

A_2F_{A1}
$| B_6F_{A1}$

F_B --->

B_1F_{B1}
$| A_5F_{B1}$

F_{B1} --->

B_4
$| A_8$
$| B_3F_{B2}$
$| A_7F_{B2}$

F_{B2} --->

B_2F_{B1}
$| A_6F_{B1}$

Connector Grammar

The actions used in the following grammar are the connecting actions. The F refers to the F of the above grammar. This grammar is used for the following connectors: **causality, conditional, consequence, enablement, inference, violated expectation, conflict, similarity, difference, contrast, background, justification, summary, elaboration, rewording, examples, generalization,** and **digression.**

The grammar is as follows:

S --->

$A_9F_AS_1$
$| B_{12}F_AS_1$
$| F_AS_1$
$| F_BS_1$

S_1 --->

$A_{10}F_AS_2$

| $A_{10}F_A$
| $B_{13}F_BS_2$
| $B_{13}F_B$
| F_AS_2
| F_BS_2
| F_A
| F_B

S_2 --->

A_{11}
| B_{14}

In addition to this grammar, we also specify the action grammar for the *topic*, *list*, *plan*, *story*, and *entity*.

Topic Grammar

S --->

A_1S_{A1}
| B_1S_{B1}

S_{A1} --->

B_2F_A
| B_2F_B
| B_2F_AS
| B_2F_BS
| B_3S_{A2}
| A_4S_{A3}

S_{A2} --->

A_6S_{A1}
| A_1S_{A1}
| B_7S_{B1}

S_{A3} --->

A_5S_{A1}
| A_1S_{A1}
| B_8S_{B1}

S_{B1} --->

A_2F_B
| A_2F_A
| A_2F_BS
| A_2F_AS
| A_3S_{B2}
| B_4S_{B3}

S_{B2} --->

B_6S_{B1}
| B_1S_{B1}
| A_7S_{A1}

S_{B3} --->

B_5S_{B1}
| B_1S_{B1}
| A_8S_{A1}

List, Plan, Story, and Entity Grammar

S --->

S_A
| S_B

S_A --->

A_1S_{A1}
| A_2S_{B1}
| $A_{15}S_{A1}$
| $B_{16}S_{A1}$

S_B --->

B_1S_{B1}
| B_2S_{A1}
| $B_{15}S_{B1}$
| $A_{16}S_{B1}$

S_{A1} --->

A_5S_{A2}
| B_6S_{B2}
| A_7S_{A1}
| B_8S_{A1}
| A_9S_{A1}
| $B_{10}S_{A1}$
| $A_{17}F_AS_{A1}$
| $B_{18}F_BS_{A1}$
| A_3S
| B_4S

S_{A2} --->

$A_{11}S_{A1}$
| $B_{14}S_{A1}$
| S_A
| S_B

S_{B1} --->

B_5S_{B2}
| A_6S_{A2}
| B_7S_{B1}
| A_8S_{B1}
| B_9S_{B1}
| $A_{10}S_{B1}$
| $B_{17}F_BS_{B1}$
| $A_{18}F_AS_{B1}$
| B_3S
| A_4S

S_{B2} --->

$B_{11}S_{B1}$
| $A_{14}S_{B1}$
| S_B
| S_A

If we refer back to the example given in section 4.1, we can say that at level-1, the sequence of actions is: "open new unit for self" (line (2)), "close unit for other" (line (33)), "inter-connect for other" (line (34)), "open new unit for self" (line (35)), and "close unit for self" (line (36)). This sequence of actions would then be denoted by the string (taking *A* to be the customer and *B* to be the agent): $A_1B_8B_{13}B_1B_4$. The string then is parsed by the connector grammar as follows:

S ==> F_AS_1
==> $A_1F_{A1}S_1$
==> $A_1B_8S_1$
==> $A_1B_8B_{13}F_B$
==> $A_1B_8B_{13}B_1F_{B1}$
==> $A_1B_8B_{13}B_1B_4$

4.5 Summary

As we have seen in the examples given in this chapter, dialogues often have quite complex discourse structures. A brief dialogue fragment we analyzed required the use of a structure five levels deep. A model that accounts for highly structured discourse is therefore needed.

One function of meta-statements is to give structure to talk. Speakers very often use meta-statements to delimit the boundaries of the components used to build their structures. Thus, in order to account for this function of the meta-statement, we must provide a model of the units that are being delimited.

The units from which a conversation is composed are *rhetorical units.* We distinguished between two types of rhetorical units: *component units* and *action units.* Component units are characterized by the way they are connected to other units. That is, the type of connector used to build the structure defines the component units from which the structure is built. Thus, a classification of rhetorical connector types is needed for building the component unit model.

The action unit, on the other hand, needs for its model the specification of rhetorical actions. Thus, we identified the rhetorical actions accounted for in our system. In addition, we furnished *action grammars* through which we encoded the possible sequences of rhetorical actions.

With the context created by the action grammars and the component unit models built, we can thus model the rhetorical structure of our conversations. As a result, the identification of the rhetorical function of meta-statements is greatly facilitated.

In addition to being able to interpret the function of meta-statements accurately--i.e., parse the rhetorical meaning of meta-satements--the system will also be able to structure the information it communicates so as to make it easier for the listener to understand what is being said. Consequently, our system will be more efficient and more coherent as a conversational partner.

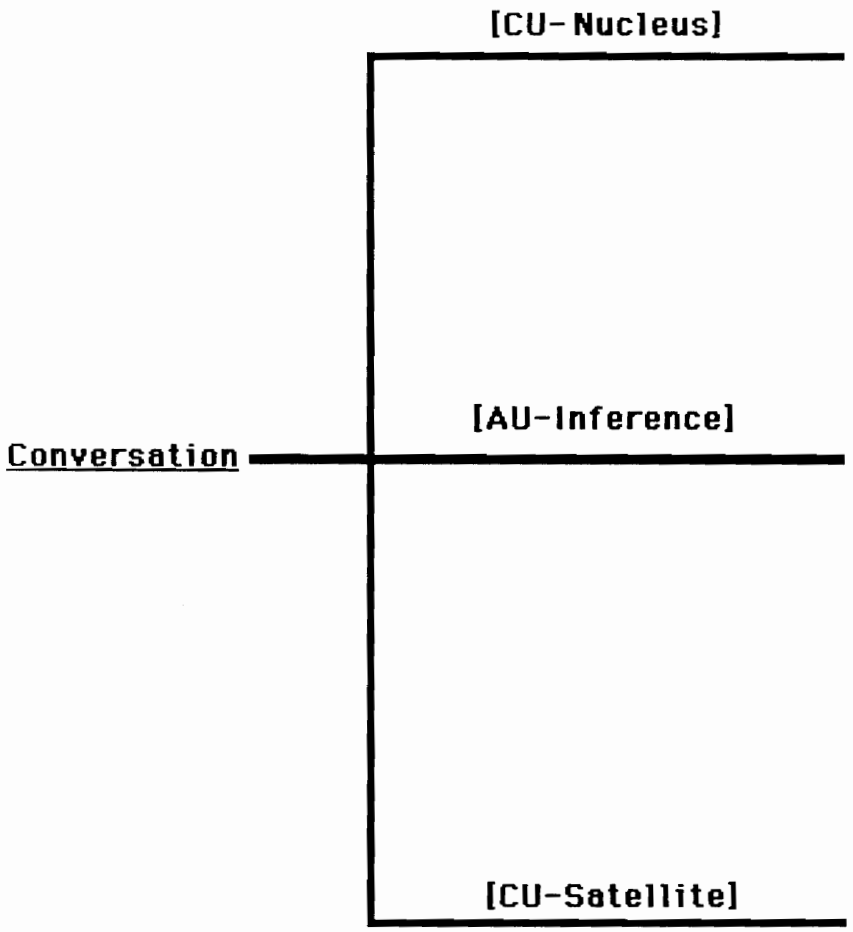


Figure 2. The Level-1 Decomposition of Conversation

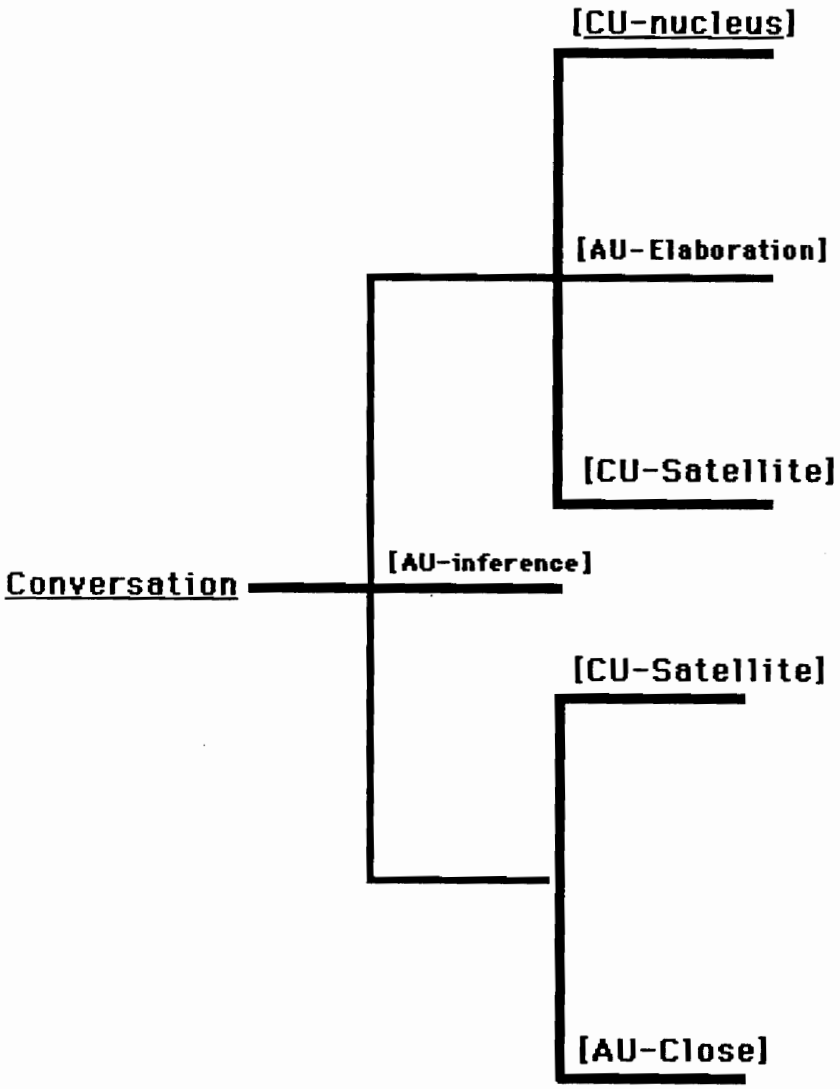


Figure 3. The Level-2 Decomposition of Conversation

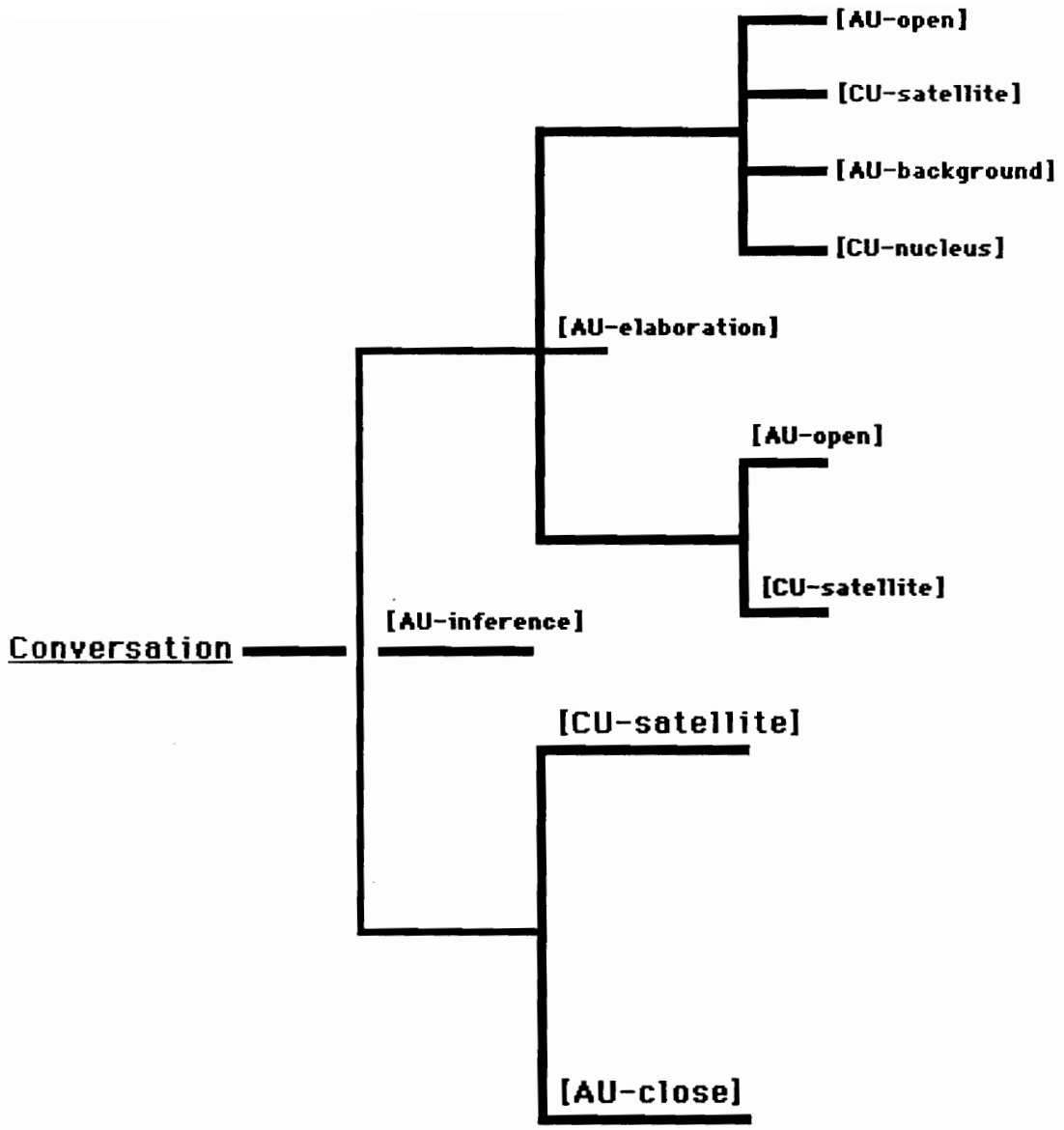


Figure 4. The Level-3 Decomposition of Conversation

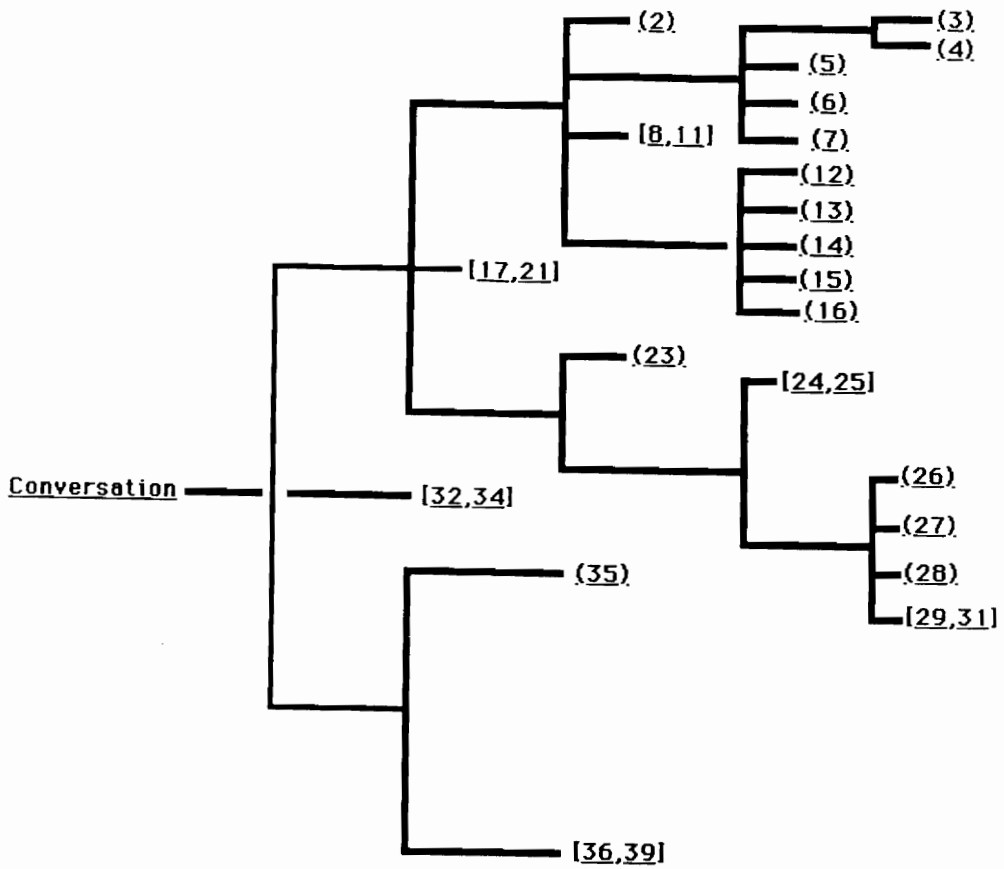


Figure 5. The Level-4 Decomposition of Conversation

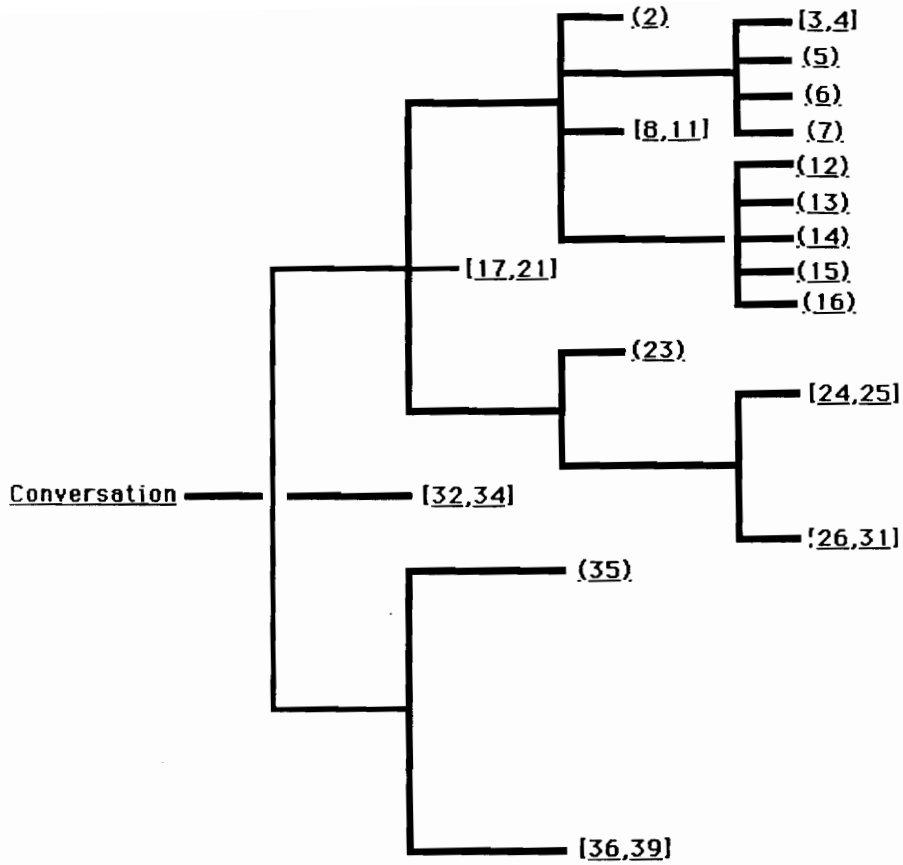


Figure 6. The Level-5 Decomposition of Conversation

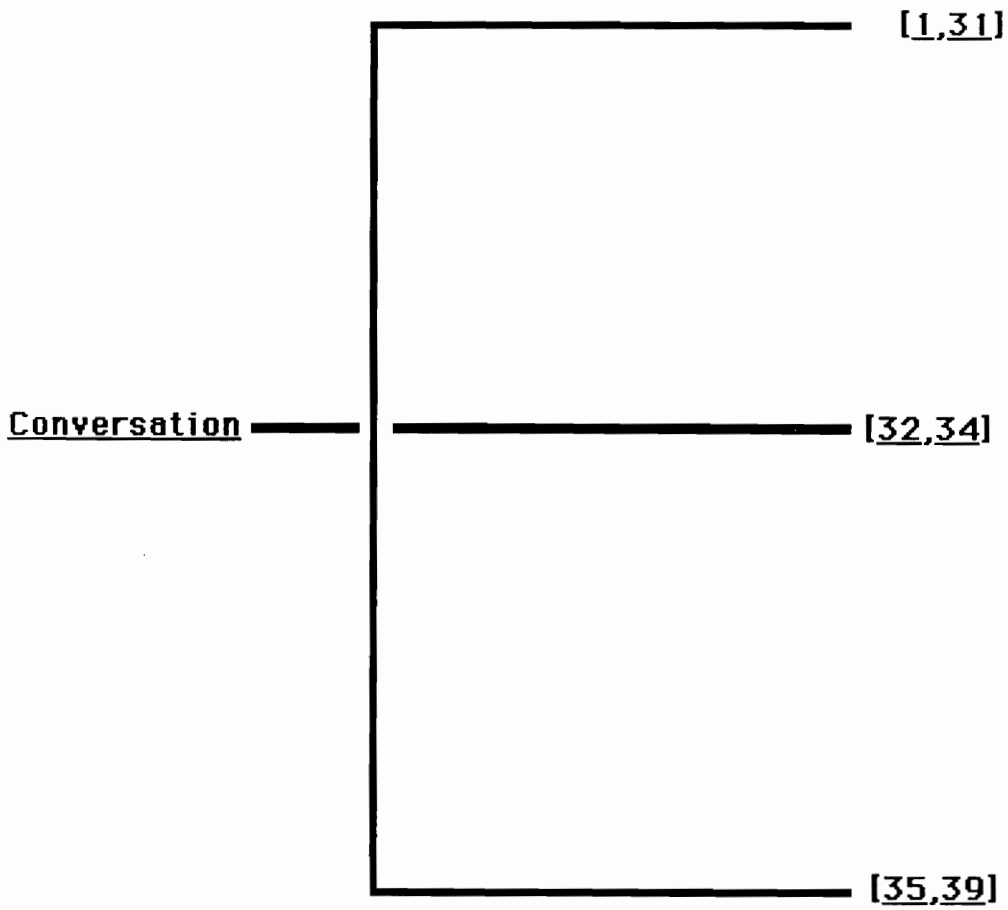


Figure 7. The Units of the Level-1 Decomposition of Conversation

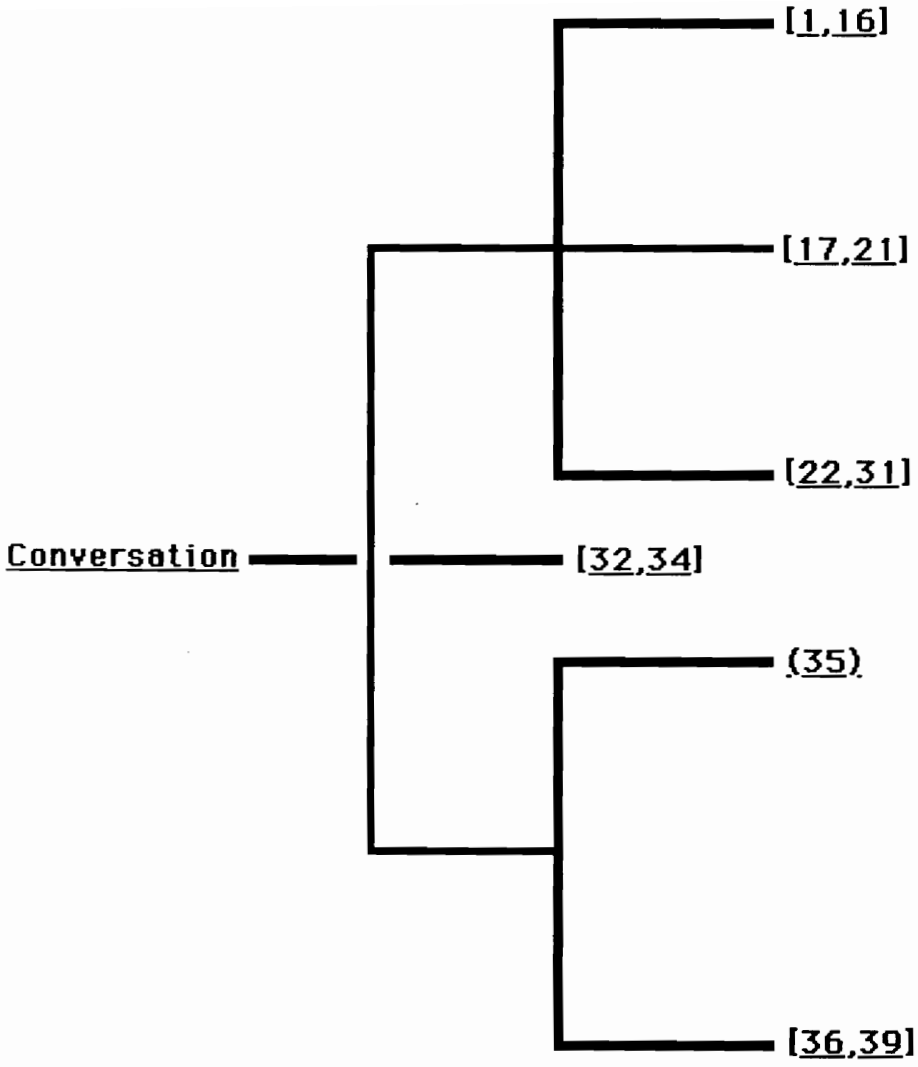


Figure 8. The Units of the Level-2 Decomposition of Conversation

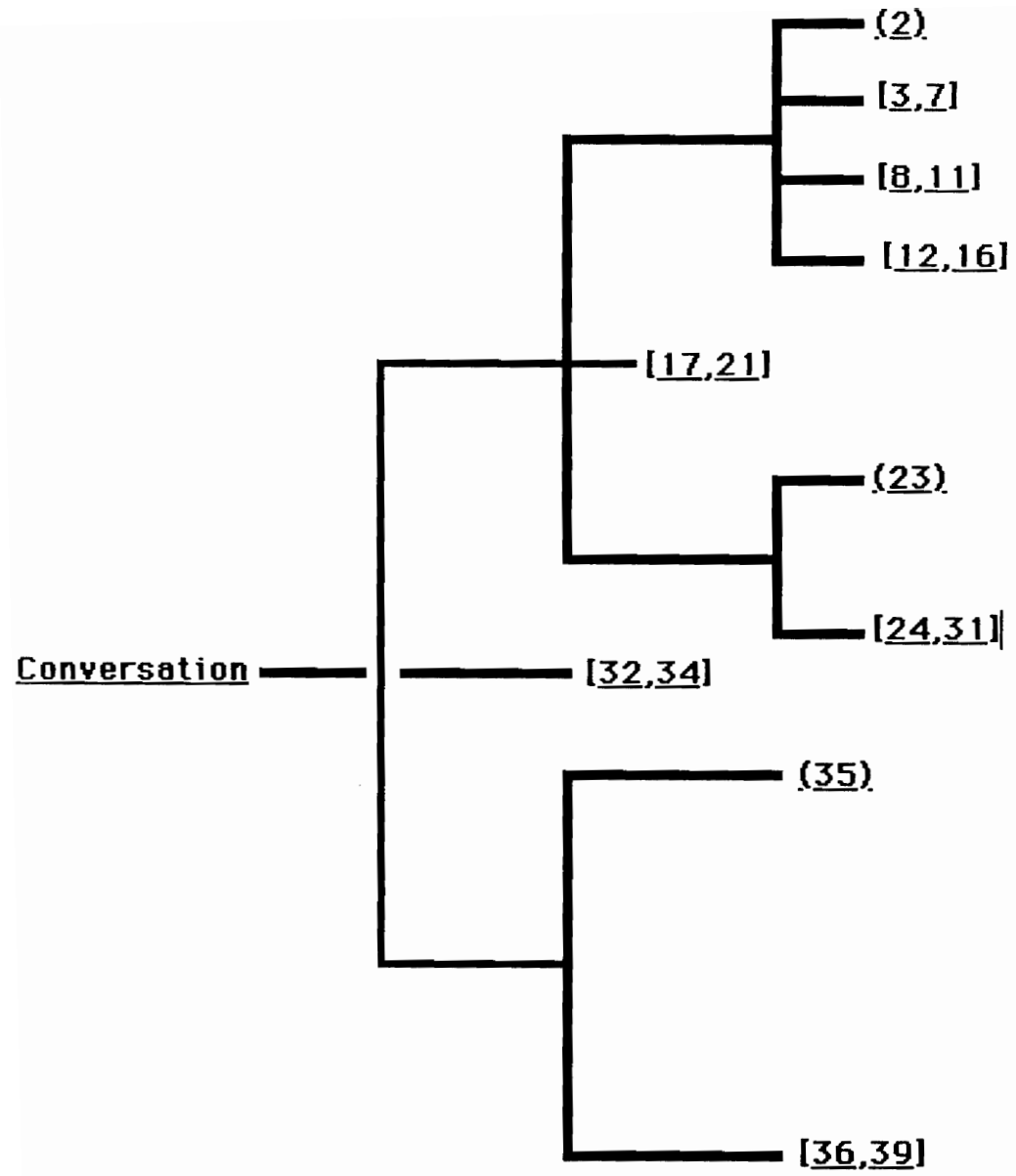


Figure 9. The Units of the Level-3 Decomposition of Conversation

Chapter 5. The Conversational Model

As we said in Chapter Three, a meta-statement fulfills two functions along the "talk about talk" dimension: a rhetorical function, where the problem of structuring talk is addressed, and a conversational function, where the process of the dyadic exchange is managed. In Chapter Four, we described the model needed to account for the rhetorical function of meta-statements. We specified both the state model describing the rhetorical state of the conversation, and the action model for taking into account the rhetorical actions that may be performed to effect the rhetorical state. In this chapter we concentrate on the second dimension of the "talk about talk" aspect of meta-statements, the conversational dimension.

We begin the chapter with a general discussion on the need for coordinating dialogues, motivating this need with some examples that illustrate the crucial role meta-statements play in the synchronization of talk. In the second section, we specify the nature of the exchange situation we are analyzing and contrast it to other situations. Following this, we treat the fundamental issues of turn taking, speaker role, and listener role. In the fourth section we concentrate on the mechanism for turn constructions and allocations, while in the fifth section we examine the issue of overlaps, interruptions, and gaps. In sections six and seven we deal

with conversational openings and closings, respectively. In the last two sections, we specify the models we have adopted for the system. The conversational state model is first presented in section eight, and in section nine the action model is specified.

5.1 Coordinating Dialogues

As in any other situation where a scarce resource is to be alternately allocated for the exclusive use of competing parties, conversations also need a mechanism for regulating the assignment of their resources. The shared resource in the case of conversations is the speaking turn, while the competing parties are, obviously, the conversers involved in the exchange. The task of the regulating mechanism, then, is to determine to whom talk time is to be allotted, and in what way such an allottment is to be effected.

In addition to handling the exchange of turns, the conversational protocol also coordinates the initiation of conversations. The initiation phase of a conversation is critical, since it is then that communication is established--i.e., contact is made. Before an exchange can be brought into effect, both parties involved in the exchange must first agree in some way that they do in fact want to engage in dialogue. The initiation phase is where this agreement is negotiated.

Conversation endings also fulfill an important function. For example, before an exchange can be terminated, the interactants must first make sure that their partner does not have anything else to say. Thus, signals to the effect that an interactant has nothing to add to what has already been said need to be sent so that the closing of the conversation can be carried out.

And finally, the protocol mechanism must also be able to account for situations where a conversation is temporarily put on hold. On many occasions, an event external to the conversation may force the momentary suspension of talk. The converser putting the conversa-

tion on hold will then need to produce the appropriate signal so the other interactant is made aware of the situation. And similarly, once an interruption is over and both interactants are ready to resume talk, an appropriate sequence of signals must be exchanged to restore the flow of talk.

In the sections that follow, we shall analyze in more detail the coordination problems that arise in conversations. But let us first take a few brief examples to illustrate in general terms the conversational function of meta-statements.

We begin with an example showing how a conversation is typically opened. As discussed in Chapter Three (see section 3.1), the conversational situation we are studying is the telephone dialogue between a travel agent and a customer. The customer is the interactant making the call, and the agent is invariably the initiator of the conversation, i.e., the agent is forcibly the first to speak since he or she is the one answering the telephone. As usual, meta-statements are enclosed in square brackets. The opening sequence follows:

Example 5.1:[8]

```
1 AGENT.
2    [Good morning, this is United Airlines
3    Reservations, Mrs. McGillicuddy.]
4 CUSTOMER.
5    [Yes,]
6    I need some information, [please.]
     ....
```

Lines (2) and (3) form a typical opening utterance from agents. Agents very often begin with a salutation ("Good morning"), then give the name of the company they represent ("United Airlines") and their name ("Mrs McGillicuddy"). In line (5) the customer aknowledges the opening move with the "yes," and then proceeds in line (6) with the conversation.

Next, we give an example illustrating how turns are retained and relinquished using meta-statements:

---

[8] Conversation phone.04, lines 1-6.

Example 5.2:[9]

1 AGENT.
2 [Now] that one you'd arrive Charlotte at 7:42 and
3 then leave again at 8:25 [...]
4 CUSTOMER.
5 [Uh-huh.]
6 AGENT.
7 arriving 9:45 in Dallas/Fort Worth.
8 CUSTOMER.
9 [Okay.]
10 AGENT.
11 [And then] the Braniff leaves at 11:15 [...]
12 CUSTOMER.
13 [Right.]
14 AGENT.
15 arriving 12:28 p.m. Los Angeles time.
16 CUSTOMER.
17 [About] 12:30. [Okay.]
18 AGENT.
19 [All right.]
20 [Okay.]
21 CUSTOMER.
22 [Um] are there any, [let's see,] [I guess so,]
23 breakfast and [...]
24 AGENT.
25 [Right,] [yeah,] you'll have full meal service on
26 that, [ah let's see,] [I believe,] from Charlotte to
27 Dallas/Fort Worth

In this exchange, between lines (1) and (20) the turn is said to belong to the agent. As we shall see later in the chapter, ownership status is to be distinguished from speaking status, and as we can see in this example, although the turn is said to be owned by the agent, the customer speaks, namely at lines (5), (9), (13), and (17). Turn transfer is executed at line (19), (20), and (22). The agent relinquishes the turn with two meta-statements in succession (lines (19,20)), while the customer accepts the turn with the meta-statement "um" in line (22). The agent reclaims the turn in line (25) after the customer hesitates.

The fragment below illustrates the need for managing conversation suspensions:

Example 5.3:[10]

1 AGENT.
2 Are you departing from Roanoke?
3 CUSTOMER.

---

[9] Conversation phone.02, lines 140-167.

[10] Conversation phone.02, lines 50-62.

4 [Ah] yes [ma'am.]
5 AGENT.
6 [Okay.]
7 [Hold just one moment,] and I'll check that, [sir.]
8 CUSTOMER.
9 [Okay.]
10 AGENT.
11 [Sir, I appreciate your waiting.]
12 CUSTOMER.
13 [Okay.]

At line (7), the agent informs the customer about the need for a momentary pause. The customer accepts the pause at line (9) with an "okay," and at line (11), the agent returns to the conversation, thanking the customer for waiting. The customer acknowledges the thanks and expresses his willingness to continue with the conversation (line (13)) with another "okay."

Finally, the fragment below is a typical sequence of utterances used for closing conversations:

Example 5.4:[11]

1 CUSTOMER.
2 [Okay] [well,] [thank you very much.]
3 AGENT.
4 [Okay?]
5 [And] [thanks for calling American Airlines.]
6 [Have a good day.]
7 CUSTOMER.
8 [Okay,] [you too.]
9 AGENT.
10 [Bye-bye.]
11 CUSTOMER.
12 [Bye-bye.]

The closing sequence in this conversation is initiated by the customer (line (2)) with meta-statements "okay" and "well," followed by a thanking expression. The agent responds with a questioning meta-statement "okay?", a thanking expression, and a parting salutation (lines (4,5,6)). In line (8), the customer responds with another "okay" and a parting salutation. It is only at lines (10) and (12) that the conversation is finally closed. As we shall see later in this chapter, the closing of conversations is usually a two step process, as it is here.

---

[11] Conversation phone.03, lines 629-640.

From the brief excerpts above, it is clear, then, that conversations are generally regulated by a turn taking protocol system. And it is equally clear that meta-statements are quite often used for following the rules of this protocol. The task we must therefore undertake is the construction of a model that accounts for all the conversational functions that a meta-statement can fulfill. But before we begin the description of our model, let us specify the basic conversational characteristics of our exchange situation.

5.2 *The Conversational Situation*

Our conversational situation has a few characteristics that distinguish it from other forms of dyadic exchanges. First of all, as we already said, given the nature of the situation, it is always the customer who calls the agent, and it is invariably the agent who begins the exchange. This establishes a rather fixed pattern of opening sequences, usually consisting of two phases: the agent opens the conversation with a salutation, and the customer answers with a salutation and begins the conversation proper. In other situations, the opening sequence is usually longer than the minimal two phase exchange found here, as in the case when the interactants are friends.

The second characteristic of our dialogues is that the length of the conversation is not specified at the outset. Although the exchange is expected to terminate within a "reasonable" period of time, no time limit is really set when the conversation begins.[12] Consequently, the closing of a conversation needs to be negotiated locally. That is, since no prior agreement is set on when to close the exchange, the time of termination must be determined within the conversation.[13]

---

[12] Interviews or therapy sessions [CLAY89], on the other hand, are almost always constrained by time considerations.

[13] In some situations, however, a converser may introduce time constraints as the conversation

Third, in our situation, there are no restrictions set on the length of a turn. Although again there is the expectation that conversers hold the turn for "reasonable" spans of time, no limit is really specified.[14] Since the conversations we are studying are transactional in nature--i.e., information is exchanged, tickets are bought, etc.--the tendency is for turns to be relatively short, although occasionally, especially when explanations are given, they do become rather long.

And finally, which converser is to manage the exchange of turns is not specified in any way. Turn exchanges are also negotiated by both parties [CLAY89]. In some cases, the turn is offered by the speaker; in other cases it is requested by the listener. Neither interactant is really given the role of turn allocator. This is not always true of dyadic exhanges. In some instances where the social status differential between the conversers is very high, the interactant with the lower status may never initiate a turn, leaving it to the higher status interactant to determine who is to speak when.

The above mentioned characteristics have important consequences on the nature of our protocol system. In what follows we discuss the salient features of this system.

5.3 *Turns, Speakers, and Listeners*

In this section, we discuss the basic notions of *turn at talk, speaker,* and *listener.* Since the definition of each of these concepts is inherently dependent on each of the other two concepts, we shall discuss them in conjunction with one another.

---

progresses. Although this has the effect of establishing a tentative time limit, the closing must still be locally negotiated. In none of the twenty seven conversations studied here did we encounter time constraints introduced within an exchange.

[14] In debates, on the other hand, time limits are set on the turn of talk.

As we noted previously, conversation is an inherently temporal phenomenon. It has a starting point in time, progresses through a span of time, and has an ending point in time. This is not necessarily true of all dyadic communication situations: the reader-writer setting,[15] for example, is not temporally constricted; that is, the time that the message formulation begins, the time it takes the writer to formulate the message, the time the reader begins reading the message, and the time it takes the reader to read the message--all these issues are not immediately consequential. The message is formulated at some point and is received at some later point, and that is all that matters.

Between the conversation and the reader-writer setting, however, there is an endless variation on the degree of temporal restrictions. For example, although the letter exchange is to a certain extent temporally open, temporal restrictions sometimes are introduced, e.g., "please write soon."

Conversations, then, are marked by a certain degree of immediacy, and therefore conversers must allocate time to speak, time to listen, time to pause, if necessary, and even time to end their conversations.

Another basic characteristic of conversations is their interactional nature. A message is sent to a specific listener, and a reaction to the message is expected to be manifested directly by the listener to the sender. The reader-writer situation, for example, lacks this characteristic. The writer may write with no specific reader in mind, and although he may have some idea about the reaction of his readers, he usually does not experience it directly (barring correspondence between writer and readers).

The most frequent task to be performed by interactants involved in a conversation is the determination of who is to speak and who is to listen. Given the immediate and interactional nature of conversations, the allocation of turns to speak is one of the most important concerns to be addressed in order for the conversation to progress. That is, given that a reaction must be exhibited within a certain period of time, both interactants must give each other the op-

---

[15] By the reader-writer setting we refer to communications between a writer and a readership at large.

portunity to give their messages (i.e., act and react), and receive messages (i.e., be in a position to react).

As we said previously, turn allocations must be effected in a cooperative way. Both interactants must enter in agreement about who is to take the turn at talk. If this agreement cannot be reached, it may very well be impossible to continue with the conversation. This suggests then that the notion of conversational turn has inherent in it the notion of ownership. The interactants in fact *own* turns at talk, and with this ownership they are entitled to certain basic "rights." Among these rights is the expectation that one ought to be asked for permission to use the turn when one owns the turn, and the right to refuse relinquishing the turn when one is in possession of it.

Turn disputes result when the ownership statuses of interactants, as perceived by each interactant, are not compatible. As we said in Chapter Three, all aspects of a conversation are modeled in terms of how they are perceived by the respective interactants. In our context, for example, we model how an interactant perceives the turn ownership status. Does interactant A feel that he owns the turn, or does he feel that the turn is owned by interactant B. And similarly for interactant B. We can never talk about a turn being owned by some interactant; we can only talk about whether or not the interactants feel they own a turn. Only with this paradigm will we be able to model the four possible ownership statuses that may exist between interactants A and B, namely, (1) A feels he owns the turn, while B does not feel he owns it; (2) B feels he owns the turn, while A feels he does not; (3) neither A nor B feel they own the turn; and (4) both A and B feel that the turn is theirs. Thus, whether A (or B) considers himself to be a speaker or a listener will depend on whether he feels he owns the talking turn, or whether he feels that it is B's (A's) property.

Along with the rights that come with owning a turn, there are also obligations that need to be fulfilled. Among these is the obligation to relinquish the turn at some point during the exchange. By definition, ownership in *multi-party* conversations cannot be exclusive, and each interactant is expected to relinquish the turn after a certain, though not specified, period of possession. Thus, it is incumbent upon each converser to show at least a minimal concern

for his partner's desire to speak, and, for example, offer the turn if the latter seems desirous to take it.

The listener also has obligations. Since there is an implicit agreement between the interactants that each is to tend to the other's turn taking needs, the listener must periodically inform the speaker that he (the listener) does not want the turn (if he does not want to take it) while the speaker is speaking. In this way, the speaker can talk without having to stop unnecessarily and offer the turn.

An important observation to be made at this point is that the concepts of speaker and turn owner are different. The speaker is not necessarily the turn owner, and the turn owner is not necessarily the person speaking. In the example below, for instance, although the customer speaks at lines (5), (10), and (15), the turn is nevertheless said to be owned by the agent. Conversationally, the customer's utterances serve only one purpose, and that is to indicate that the turn is not desired and that the ownership of the turn is not challenged yet:

Example 5.5:[16]

1 AGENT.
2 [Now,] what you could at least do for yourself,
3 is set up a tentative reservation.
4 CUSTOMER.
5 [Uh-huh.]
6 AGENT.
7 [And] then [you know] buy the ticket as soon as
8 you can.
9 CUSTOMER.
10 [Right.]
11 AGENT.
12 [And] thataway at least you know the minute I'm able
13 to buy my ticket that that fare is frozen.
14 CUSTOMER.
15 [Uh-huh].

Having discussed in general terms some of the issues raised by the turn taking aspect of conversations, we proceed next to an analysis of how turns are constructed and allocated.

---

[16] Conversation phone.03, lines 110-124.

5.4 Turn Construction and Allocation

By nature, dialogues are negotiative. In order for a dialogue to progress, the parties involved must work cooperatively. Since it is expected that the exchange progresses in a series of alternating turns, the conversers must work together to effect turn transfers from one interactant to another. Unlike the mediated exchange, where a third party is assigned the responsibility of deciding when a turn starts, how long it is to last, and when it ends, in dialogue these issues must be resolved by the conversers themselves.

As we said in the previous section, the most fundamental issue to be addressed in the exchange of turns is that of turn ownership: how is the turn acquired for ownership, how is it protected, how is it declined when offered, how is it relinquished, and what signals do conversers send to inform each other of their conversational turn intent?

We distinguish between two types of turn exchange strategies: (1) the "current speaker selects next" strategy, where the owner offers the turn to the speaker, and (2) the "next speaker selects next" strategy, where the listener prompts the owner for the turn.

In the case where the owner offers the turn, the offer is usually signalled without a verbal preface. Most of the time the signal is accomplished by a tonal inflection [HALL67]. In some instances, however, verbal signals are in fact used, as with the meta-statement "Right?" in line (4) of the example below, where the agent explicitly offers the turn to the customer:

Example 5.6:[17]

1 AGENT.
2 I don't intend you calling American Airlines on
3 the first of January.
4 [Right?]
5 CUSTOMER.
6 [Right.]

When verbally signalled, in our transcripts the expressions used are "okay," "all right," "right?" and "you know." Expressions such as "sir" and "please" are almost always used only

---

[17] Conversation phone.03, lines 336-341.

at the end of turns, and their presence often signals a turn offer. Questions, it must be noted, are a strong indicator that the turn is offered by the converser asking the question.

The second type of turn allocation requires that the *listener* move to acquire the turn. In the following example, the agent takes the turn in line (6) by anticipating the customer's utterance:

Example 5.7:[18]

```
1 CUSTOMER.
2   [Uh-huh,] [I see,] [um] how about, [let's see,] what
3   would the price be, I don't know if you can do
4   this, for my nephew from [ah] [...]
5 AGENT.
6   [From Seattle to Los Angeles only?]
7 CUSTOMER.
8   [From Seattle down to Los Angeles, yeah.]
```

While in the example below, the agent moves to acquire the turn (line (7)) with a request for permission through the "let me" meta-statement, in a way showing politeness and a desire to avoid a confrontation:

Example 5.8:[19]

```
1 CUSTOMER.
2   [Okay.]
3   [And also] [ah] about my flight from, [ah] to
4   Seattle, is there [ah] a lot of layovers there [also,]
5   or [...]
6 AGENT.
7   [Let me] check the times for you on that schedule,
8   [let's see how that works.]
```

Notice that in both of these cases the situation is not as clear-cut as that of example 5.6. Here the customer seems to offer the turn, though not straightforwardly, by hesitating (line (4) of example 5.7, and (5) of example 5.8), with the agent seizing the moment, as it were, and moving to take the turn. Using the terminology of [SACK74], moments such as this are called *transitional relevance places* (TRP's). It is at such places that listeners have the most opportune occasion to make their move for acquiring the turn.

---

[18] Conversaton phone.02, lines 261-268.

[19] Conversaton phone.02, lines 277-284.

Obviously, there are two exceptions to the rule of moving to acquire a turn during TRP's. First, the listener may wish to take a turn *before* a TRP is encountered, either out of necessity (the urgency of the message, for example), or because he feels that not enough TRP's are provided by the speaker. In either case, the listener is said to *interrupt* the speaker.

The second exception occurs when the listener does not wish to claim the turn ownership at all, even though a TRP is encountered. In such a case the listener is expected to display his non-challenge explicitly. In telephone conversations, such displays are invariably performed verbally, whereas in face to face encounters they are usually both verbal and paralinguistic, (e.g., nods, face expressions, etc.). In example 5.5, for instance, the customer in lines (5), (10), and (15) is displaying his non-challenge for the turn. Each of the points at the end of the respective previous stretches of talk is a TRP. Semantically, the stretches of talk between the customer's utterances are complete, and the customer could have safely moved to claim the turn. Since he chose not to move to acquire the turn, and since the agent did provide him with the opportunity to make a move, the listener was therefore obliged in some sense to decline the implicit offer explicitly.

We found that in our transcripts, the most frequently used meta-statements of this type are "uh-huh" utterances, and variations thereof. Also among the formulations used are "okay," "yeah," and "all right." In addition, we also found many instances where exclamations such as "oh, boy," "oh, I see," and "God, really" are used.

Turn ownerships are often initiated with "okay" and "well," as well as "now," "all right," "yes," and connectors such as "so." In addition, turns are frequently started with a hesitation marker, such as "ah" and "um."

Resuming turns is usually performed with a connector, with "and" used most often. "But" is also used on occasion, usually when returning from a digression. Also among the formulations used to resume a turn are "then," "and then," and "okay."

The last important issue to discuss is that of turn protection: how do turn owners "ward off" listeners' attempts at taking the turn, and how do they signal their desire to keep the turn?

Duncan calls such messages "attempt-suppressing" signals, and observes that rarely do auditors try to take a conversational turn when such a signal is displayed [DUNC72.287].

He observes that one of the signals used for suppressing an attempt on the part of the listener is using hand gesticulations while speaking. In the case of telephone exchanges, where the conversers are not able to see one another, hand gesticulations are obviously not used for this purpose. Instead, turns are very often 'protected' using meta-statements. In the example below, the customer uses the "um," "okay," "let's see," and "uh" to protect his turn before he really begins his question:

Example 5.9:[20]

1 CUSTOMER.
2 [Um] [okay,] [let's see] [uh,] what is the [uh,] how long,
3 you say there's, I have to stop out from Roanoke
4 to Charlotte?

Suppressing turn attempts is mostly performed with statements such as "ah" and "uh," or statements of the like. Also used are expressions such as "you know," and "let's see," and "see." Occasionally, "okay" is also used for turn protection.

Having discussed the major issues involved in modeling turn allocations, we proceed to a discussion of the problems that may arise when turn exchanges do not proceed as expected.

5.5 *Overlaps, Interruptions, and Gaps*

In the ideal turn exchange system, the allocation of turns would take place smoothly and without any overlaps or gaps in the ownership of the turn.[21] In such a system, a speaker would assume the turn ownership as soon as his partner has relinquished it, would continue talking

---

[20] Conversation phone.02, lines 100-103.

[21] Note that we say overlaps or gaps in *the ownership* of the turn and not overlaps or gaps in *talk*. Simultaneous talk and silences do not necessarily indicate that a turn dispute is in effect. As we will see shortly, overlaps frequently occur in natural talk, and most of the time, such simultaneous talk

for the duration of his turn, and as soon as his claim over the turn expires he would relinquish the turn, with the other converser at once reclaiming the ownership of the turn. However, this is hardly the case in real life situations. Overlaps in speaking turns, speakers talking at the same time, interruptions, gaps and silences between turns are very common. In fact, conversations without such phenomena would be considered "unnatural" and would sound "artificial."

As we already said in section 5.3, simultaneous talk does not necessarily indicate a dispute over ownership of the turn. In fact, we saw that very often a listener *is supposed* to interject occasionally while the speaker is talking, so as to signal his non-challenging status. Such instances of simultaneous talk must, therefore, be distinguished from the case where a listener interjects to claim the turn while the owner of the turn does not intend to relinquish the turn. The former we call *overlaps*, the latter *interruptions.*

5.5.1 Overlaps

Overlaps are those conversational situations where both conversers are talking at the same time without the existence of a turn ownership dispute. We distinguish between two types of overlaps: the *back-channel* overlap and the *trail-off* overlap. Back-channels were introduced in the previous section (see example 5.5), and it was determined that their main function is to indicate non-challenge for the turn.

Trail-off overlaps, on the other hand, are those instances where the listener begins assuming ownership of the turn *while* the speaker is concluding the talk within his turn.

In our data, we found two types of trail-off overlaps. First is the *simple trail-off,* where the listener perceives the speaker to be in the last stages of his turn. The second type is the

---

is not an instance of turn disputes. Similarly, a long silence does not necessarily mean that neither converser wants the turn. It may very well be that the silence is a case of the turn owner reflecting on what to say next.

anticipated trail-off, where the listener finishes what the speaker was saying and does it so that he (the listener) puts a claim on the turn. Example 5.10 is an instance of a simple trail-off, whereas example 5.11 is a case of anticipated trail-off:

Example 5.10:[22]

1 AGENT.
2 Robert.
3 [And] the telephone number at home would be area
4 code ...
5 CUSTOMER.
6 [Okay] the area code 703.

Example 5.11:[23]

1 CUSTOMER.
2 [and ah] I have a nephew up there and I might have
3 to pick him up and ah bring him down to me with,
4 [ah] to Los Angeles.
5 Is there a ...
6 AGENT.
7 If you went Roanoke to Seattle and then to Los
8 Angeles?
9 CUSTOMER.
10 [Yeah.]

5.5.2 Interruptions

Interruptions, on the other hand, occur without the existence of a TRP. Unlike the overlap, an interruption occurs while the owner is still desirous of holding the turn.

In our paradigm, where we must always speak in terms of perceptions, an interruption must therefore be analyzed in terms of conversers' perceptions. Thus, for example, whether the speaker perceives that he is being interrupted will depend on whether or not he thinks he

---

[22] Conversation phone.03, lines 372-376.

[23] Conversation phone.02, lines 181-190.

sent a cue to signal a TRP. Similarly, whether or not a listener feels he is interrupting will depend on his perception of TRP signals. If the listener thinks he received a TRP signal, then he will consider the simultaneous talk to be an instance of trail-off overlaps, whereas if he starts talking without perceiving that a TRP message was sent, then he will consider himself to be interrupting his partner. Four listener/owner TRP state perceptions cases are possible. The first case is when both owner and listener consider that a TRP is in effect and the listener starts talking before the speaker has stopped. This situation would be viewed by both interactants as an instance of trail-off overlaps. The second case is when the owner does not signal a TRP, but the listener for some reason does perceive one (either because the speaker inadvertently sent a TRP message or because the listener "misinterpreted" a signal). In this case we have the state where the speaker views the listener's talk as an interruption, while the listener thinks that it is a case of trail-off overlapping.

The third case is when the speaker thinks that a TRP is in effect, while the listener does not perceive the TRP (not having perceived the TRP signal sent by speaker). In this case, the listener views the simultaneity as an interruption, while the speaker feels that a transition of turns is occurring. The fourth case is when both speaker and listener perceive that no TRP is in effect; that is, the listener and the speaker perceive the former to be interrupting the latter.

The distinction between the four cases is important in at least one respect: the interpersonal management of faces (see Chapter Seven). In the second case, where the listener sees a TRP while the speaker does not see it, the speaker might become offended by the listener's lack of politeness: the speaker perceives the listener to be interrupting, and thus expects an apology of some sort, while the latter, perceiving that he has been offered the turn, sees no need to apologize since, according to his perception, he is only following the rules. Similarly, in the third case, where the speaker perceives a trail-off while the listener perceives an interruption, the listener may be overly polite (or perceived to be so by the speaker) since he (the listener) is apologizing for an interruption the speaker does not perceive.

Resolution of Turn Disputes

In case an interrupted speaker decides to dispute the ownership of the turn, he has at his disposal *a repair mechanism.* Through this mechanism both sides of the dispute move towards *the resolution* of the contest [DUNC74]. The speaker who emerges from the dispute with the turn is called *the winner.*

Among the techniques afforded by this mechanism is the use of *voice volume differential.* According to Meltzer, Morris, and Hayes, the *relative difference* in voice amplitude can be effective in the retention of turn: "'Soft-Spoken' persons may win interruptions via the surprise tactic of being loud only when they need to be" [MELT71.393].

Another method of contention is the explicit complaint, such as "let me finish," or "I'm not done yet." Complaints, however, in order to be effective must be formulated as soon as the turn ownership dispute begins [ZIMM75]. If not, the speaker complaining may lose ground for his protest (i.e., the turn becomes the property of the other speaker).

Still another option available to an interrupted speaker is to ignore the interruption altogether and resume speaking as though no challenge were forwarded. With the complaint technique, the speaker complaining needs to change topics: he needs to stop talking about what he was discussing before the interruption occurred and initiate his complaint. This may occasion too much effort in some instances, as when the speaker is concentrating on his train of thoughts [ZIMM75].

5.5.3 Gaps

In addition to overlaps and interruptions, natural conversations "deviate" from the ideal turn exchange model by their inclusion of gaps and silences between turns.

The existence of these gaps can be explained by observing that although a turn is normally a sought resource, it is not necessarily so. That is, a listener usually does take the

turn when it is offered him, but he may not necessarily take it. It may be by choice: he has nothing to say, and thus does not see the need to take the turn, or it may be because he does not perceive the TRP signal sent by the speaker and still perceives the speaker to claim ownership over the turn.

In the first instance, the gap may ensue from the listener's failure to emit the proper back-channel signal. As we said, when a TRP is in effect, speakers expect some type of back-channel signalling from the listener in order to determine whether or not the turn ownership is being challenged. A silence, or a conspicuous lack of back-channel messages may cause the speaker to pause and perhaps even explicitly offer the turn to the listener.

The second case is an instance of incompatible perceptions; namely, the speaker has emitted a yield turn signal, but the listener did not perceive it.

In addition, a silence gap may simply be an instance of the turn owner pausing to reflect on what to say next, with both interactants agreeing on who among the conversers if the owner of the turn.

Having examined the issues of interruptions and gaps, we next proceed to an analysis of conversation openings.

5.6 Conversation Openings

Conversations are usually initiated with a highly regular exchange sequence. Before conversers can really begin the conversation, they first proceed through a sequence of moves called *an opening.* Openings serve an important function: they are the stage of the conversation where the nature of the dialogue and the identity of the interactants are specified [SCHE68]. It is during openings that the conversers negotiate what sort of an exchange they will have, or even whether the exchange will take place or not.

The conversational situation we are examining--the telephone dialogue--places some restrictions on the types of opening sequences usually encountered. Among these is what Schegloff [SCHE68.1076] calls *the distribution rule.* The distibution rule maintains that the answerer (in our case, the agent) is the one expected to answer first.

In our interviews, the first receiver's utterance is invariably a self-identification. Agents always give their name (usually the full name, sometimes only last name, and occasionally the name and a title, e.g., "Mrs.") and the name of the company they represent (e.g., "Delta Airlines"). The caller's responses are usually simple greetings, followed by a brief statement of the problem.

Example 5.12:[24]

1 AGENT.
2 [Piedmont reservations, Mrs. Eyelesbarrow.]
3 CUSTOMER.
4 [Hi,] I was wondering if you could help me with some
5 information.

Very often, we found that customer responses begins with a simple "yes." This is typical of situations where the caller is contacting a service establishment, where the identity of the answerer, beyond the role they play (i.e., agent) is not relevant to the situation. Since agents usually answer with an identification specification, the customers, having nothing to add to the opening exchange, often use the *minimal "yes."* The minimal "yes" mainly functions to signal that the customer is aware of who he is dealing with and that he is ready to start the conversation. The exchange below is a typical opening:

Example 5.13:[25]

1 AGENT.
2 [American Airlines, Mrs. Crackenthorpe.]
3 CUSTOMER.
4 [Yes,] [um] I'm planning a trip from [ah] Roanoke to
5 L.A. I'd like some [ah] information.

---

[24] Conversation phone.03, lines 1-5.

[25] Conversation phone.02, lines 1-5.

5.7 Conversation Closings

As we noted in our previous discussion, in the conversational setting we are analyzing, the length of the dialogue exchange is not specified beforehand. This, as we remarked, puts the burden of determining when the exchange is to end on the interactants. It is they who must negotiate (as the conversation is progressing) when to bring the interaction to a close.

According to Schegloff and Sacks [SCHE73], in dialogues such as the ones we are analyzing, conversers need to address two problems before a conversation can be closed successfully. First, they must signal their desire to close, and do it such that their signal is recognized by the other interactant. Halting with a silence, for example, is not a sufficient signal in itself for closing a conversation, since it may be interpreted as a gap in the flow.

The second issue that needs to be resolved is the determination by the speakers that their partner has nothing more to say, or at least nothing that must be said under the circumstances. Since the time the exchange is to end is not specified at the outset, speakers can never be sure that their partner has nothing more to add to what has already been said. Therefore, in addition to signalling their intentions, they must also determine in some way the intentions of their partner.

One way speakers can determine whether or not further discussion is still planned by the other speaker is through the use of a *preclosing exchange* [CLAY89.667].[26] The simplest type of preclosing sequences is the technique of *passing turns* [CLAY89.667]. The interactants in this case simply pass the turn one after the other, thus establishing not only that they wish to bring the exchange to a stop but also that their partner has nothing more to say. We found many instances of preclosing sequences in our corpus of conversations. The following is a typical closing sequence:

---

[26] Schegloff and Sacks [SCHE73.303] use the term "unmentioned mentionables" for describing such discussion.

Example 5.14:[27]

1 AGENT.
2 [and again,] weekdays generally on the super-saver
3 are always cheaper.
4 CUSTOMER.
5 [Yeah,] [okay.]
6 AGENT.
7 [Okay?]
8 CUSTOMER.
9 [All right.]
10 [Hey, I appreciate your ah ...]
11 AGENT.
12 [Sure!]
13 CUSTOMER.
14 [all your ah punching up there.]
15 AGENT.
16 [You're quite welcome.]
17 [Thank you.]
18 [Bye-bye.]
19 CUSTOMER.
20 [Bye-bye.]

At line (7) the agent inititates the preclosing sequence with a questioning "okay?" The customer responds in line (9) with an "all right." Lines (7) and (9) form a passing turns exchange. The customer resumes the preclosing sequence with a thanking utterance, and the exchange of thanks continues through line (17). It is only in line (18) that the closing is undertaken with the parting salutation "Bye-bye," which is used once by each converser.

Another technique for preclosing is the *promise for further contact.* We found in our transcripts that when used, the promiser is always the customer, as in the example below:

Example 5.15:[28]

1 CUSTOMER.
2 [All right.]
3 That's all the information I need for now,
4 and then I'll get back with you when I get
5 my arrangements made.
6 AGENT.
7 [Okay.]

Also used often by customers is the *expression of satisfaction.* In the example above, the customer initiates the preclose move with the statement "that's all the information I need for

---

[27] Conversation phone.05, lines 421-440.

[28] Conversation agency.03, lines 64-69.

now" (line (3)). In addition, often used for preclosing is the exchange of thanks, usually in the form of "thanks," "thank you," and occasionally "I thank you."

5.8 State Model

We now begin the specification of the conversational state model our system will use. The specification will consist in determining the conversational parameters needed for the model.

From the above discussion,the following parameters are considered necessary for the specification of the state of a conversation at any point in time:

Conversation State (CS). At any point during an exchange, the conversation is said to be in one of seven states:

1 - Opening Stage State: The conversation is initiated, but no "substantive" discussion has been undertaken yet.

2 - Main Stage State: The opening has been concluded, and "substantive" matters are being discussed.

3 - Suspended State: During a pause, with both interactants planning to resume the exchange after the pause.

4 - Preclosing Stage State: After main discussion, but before the exchange of parting salutations.

5 - Closing Stage State: During the exchange of parting salutations.

6 - Closed State: After the exchange has been terminated.

7 - Aborted State: After the closing of the conversation without going through a closing sequence.

Internally, this parameter is represented as follows:

```
(CONV-STATE      ?state)
```

with ?state taking one of the following values: OPEN, MAIN, SUSPEND, PRECLOSE, CLOSING, CLOSED, and, ABORTED.

Turn Owner (TO). The perceiver will at any point in time regard the turn to be owned by one of the two interactants:

1 - Self: the perceiver sees the turn as belonging to him.

2 - Other: the perceiver sees the turn as belonging to his partner.

Internally, this parameter is represented as follows:

```
(TURN-OWNER      ?owner)
```

with ?owner being either SELF or OTHER.

Current Speaker (SP). As we said, speakers are not necessarily owners. This parameter identifies the interactant perceived to be speaking:

1 - Self: perceiver is speaking.

2 - Other: other is speaking.

3 - Both: both perceiver and other are speaking simultaneously.

```
(SPEAKER         ?speaker)
```

with ?speaker being SELF, OTHER, or BOTH.

TRP flag (TRP). Transitional relevance places (TRP's) are points in the conversation where it is "appropriate" for interactants to move toward the acquisition of the turn. At any point in time, a TRP is either in effect or not:

1 - Yes: TRP is in effect.

2 - No: TRP is not in effect.

```
(TRP     ?flag)
```

with ?flag being either YES or NO.

Interruption Flag (IF). In the case the perceiver moves to speak without the perceived (by the interrupter) existence of a TRP, then the interruption flag becomes *on.* Only after the restoration of the turn to speaker (interactant interrupted) or the transfer of turns from speaker to listener is the value of the flag *off* once again:

1 - On: interruption perceived to be in effect.

2 - Off: no interruption perceived to be in effect.

```
(INTERRUPT-FLAG          ?flag)
```

with ?flag being either ON or OFF.

The state of a conversation at any point in time, as perceived by a converser, is thus the specification of the value of each of the above parameters.

5.9 Action Model

We now specify the conversational action model. We first specify the conversational actions, and then discuss the conversational action grammar.

5.9.1 Actions

Below is the list of conversational actions. To the right of each action we specify the parameter(s) that may be affected by the action:

Opening Actions:

1 - Open conversation (CS-1)
2 - Accept open conversation move (CS-1).
3 - Reject open conversation move (CS-1,CS-7).

Preclosing actions:

4 - Signal preclose move (CS-4).
5 - Accept preclose move (CS-5).
6 - Reject preclose move (CS-2).

Exchanging turn actions:

7 - Signal desire to acquire turn (TO,IF).
8 - Yield turn (TO).
9 - Refuse to yield turn (TO).
10 - Signal TRP (TRP,TO).
11 - Explicitly offer turn (TO,TRP).
12 - Accept offered turn (TO).
13 - Decline offered turn (TO).
14 - Signal desire to retain turn (TO).

Conversation pause actions:

15 - Signal conversational pause (CS-3).
16 - Accept conversational pause request (CS-3).
17 - Reject conversational pause request (CS-3).
18 - Signal return to flow after pause (CS-2).

19 - Signal desire to resume conversation after closing is initiated (CS-2).
20 - Abort conversation (CS-7).
21 - Sustain Flow (SP).
22 - Signal parting salutation (CS-5).

5.9.2 Action Grammar

The grammar for the conversational model is specified in full in what follows. To illustrate the grammar, we apply it to a conversation fragment.

Conversational Grammar

S --->

S^A
| S^B

S^A --->

$OPEN - OK^A\ S_1^B$
| $OPEN - NO^A\ S_2$
| A_{20}
| B_{20}

S_1^A --->

$PAUSE - OK^A\ S_1$
| $PAUSE - NO^A\ S_1$
| $EXCHANGE - OK^A\ S_1^B$
| $EXCHANGE - NO^A\ S_1^A$
| $PRECLOSE - OK^A\ S_2^A$
| $PRECLOSE - NO^A\ S_1^B$
| $A_{14}S_1^A$
| A_{20}
| B_{20}

S_2^A --->

$CLOSE - OK^A$
| $CLOSE - NO^A\ S_1^B$
| A_{20}
| B_{20}

$OPEN - OK^A$ ---> A_1B_2

$OPEN - NO^A$ ---> A_1B_3

$PAUSE - OK^A$ ---> $A_{15}B_{16}PAUSE - OK_1^A$

$PAUSE - OK_1^A$ --->

A_{18}
| B_{18}

$PAUSE - NO^A$ ---> $A_{15}B_{17}$

$EXCHANGE - OK^A$ --->

$OFFER - OK^A$
| $REQUEST - OK^A$
| $TRP - OK^A$

$EXCHANGE - NO^A$ --->

$OFFER\text{-}NO^A$
| $REQUEST\text{-}NO^A$
| $TRP\text{-}NO^A$

$OFFER\text{-}OK^A$ ---> $A_{11}B_{12}$

$REQUEST\text{-}OK^A$ ---> B_7A_8

$TRP\text{-}OK^A$ ---> $A_{10}B_{12}$

$OFFER\text{-}NO^A$ ---> $A_{11}B_{13}$

$REQUEST\text{-}NO^A$ ---> B_7A_9

$TRP\text{-}NO^A$ ---> $A_{10}B_{21}$

$PRECLOSE\text{-}OK^A$ ---> A_4B_5

$PRECLOSE\text{-}NO^A$ ---> A_4B_6

$CLOSE\text{-}OK^A$ --->

$A_{22}B_{22}$
| A_{22}

$CLOSE\text{-}NO^A$ --->

$A_{22}B_{19}$
| A_{20}
| B_{20}

S^B --->

$OPEN\text{-}OK^B\ S_1^A$
| $OPEN\text{-}NO^B\ S_2$
| B_{20}
| A_{20}

S_1^B --->

$PAUSE\text{-}OK^B\ S_1$
| $PAUSE\text{-}NO^B\ S_1$
| $EXCHANGE\text{-}OK^B\ S_1^A$
| $EXCHANGE\text{-}NO^B\ S_1^B$
| $PRECLOSE\text{-}OK^B\ S_2^B$
| $PRECLOSE\text{-}NO^B\ S_1^A$
| $B_{14}S_1^B$
| B_{20}
| A_{20}

S_2^B --->

$CLOSE\text{-}OK^B$
| $CLOSE\text{-}NO^B\ S_1^A$
| B_{20}

| A_{20}

$OPEN - OK^B$ ---> B_1A_2

$OPEN - NO^B$ ---> B_1A_3

$PAUSE - OK^B$ ---> $B_{15}A_{16}PAUSE - OK_1^B$

$PAUSE - OK_1^B$ --->

B_{18}
| A_{18}

$PAUSE - NO^B$ ---> $B_{15}A_{17}$

$EXCHANGE - OK^B$ --->

$OFFER - OK^B$
| $REQUEST - OK^B$
| $TRP - OK^B$

$EXCHANGE - NO^B$ --->

$OFFER - NO^B$
| $REQUEST - NO^B$
| $TRP - NO^B$

$OFFER - OK^B$ ---> $B_{11}A_{12}$

$REQUEST - OK^B$ ---> A_7B_8

$TRP - OK^B$ ---> $B_{10}A_{12}$

$OFFER - NO^B$ ---> $B_{11}A_{13}$

$REQUEST - NO^B$ ---> A_7B_9

$TRP - NO^B$ ---> $B_{10}A_{21}$

$PRECLOSE - OK^B$ ---> B_4A_5

$PRECLOSE - NO^B$ ---> B_4A_6

$CLOSE - OK^B$ --->

$B_{22}A_{22}$
| B_{22}

$CLOSE - NO^B$ --->

$B_{22}A_{19}$
| B_{20}
| A_{20}

Example 5.16: [29]

AGENT.
[Piedmont reservations, Mrs. Eyelesbarrow.]
CUSTOMER.
[Hi,] I was wondering if you could help me with some information.
AGENT.
[Yes, sir?]
CUSTOMER.
[Um] in April, April 5th I'm gonna be going [um] to Los Angeles ...
AGENT.
[Uh-huh.]
CUSTOMER.
[and] I would like [ah] some information on [ah] flights out that way.
AGENT.
[Okay.]
[All right.]
Do you know how long you'll be staying if you leave on the 5th?
When you'll be returning?
CUSTOMER.
When returning?
[Um] [let's see,] it'd probably be for, it's not certain right now [but] it'd probably be for like 10 days to two weeks.
AGENT.
[Okay.]
[So] you would be staying at least one week [then,]
[okay.]
CUSTOMER.
[Yes, ma'am.]
AGENT.
[All right,] it probably will be cheaper for you if you can fly week days both ways.
CUSTOMER.
[Uh-huh.]
AGENT.
[Since] you picked a Monday, [I believe,] to leave that would be a good idea if you could return say on a Monday through Thursday.
CUSTOMER.
[Uh-huh.]

Analyzing the above conversation fragment in terms of conversational actions, we obtain the following transcription:

AGENT.
[OPEN-CONVERSATION]

---

[29] Conversation phone.02, lines 1-43.

3 CUSTOMER.
4 [ACCEPT-OPEN-CONVERSATION-MOVE] {...}[30]
5 {...} [SIGNAL-TRP]
6 AGENT.
7 [SUSTAIN-FLOW]
8 CUSTOMER.
9 [SIGNAL-DESIRE-TO-RETAIN-TURN] {...} [SIGNAL-DESIRE-TO-RETAIN-TURN]
10 {...} [SIGNAL-TRP]
11 AGENT.
12 [SUSTAIN-FLOW]
13 CUSTOMER.
14 [SIGNAL-DESIRE-TO-RETAIN-TURN] {...} [SIGNAL-DESIRE-TO-RETAIN-TURN]
15 {...} [SIGNAL-TRP]
16 AGENT.
17 [ACCEPT-OFFERED-TURN]
18 [SIGNAL-DESIRE-TO-RETAIN-TURN]
19 {...}
20 {...}
21 {...} [SIGNAL-TRP]
22 CUSTOMER.
23 [ACCEPT-OFFERED-TURN]
24 [SIGNAL-DESIRE-TO-RETAIN-TURN] [SIGNAL-DESIRE-TO-RETAIN-TURN] {...}
25 {...}
26 {...} [SIGNAL-TRP]
27 AGENT.
28 [ACCEPT-OFFERED-TURN]
29 {...}
30 [SIGNAL-TRP]
31 CUSTOMER.
32 [SUSTAIN-FLOW]
33 AGENT.
34 [SIGNAL-DESIRE-TO-RETAIN-TURN] {...}
35 {...} [SIGNAL-TRP]
36 CUSTOMER.
37 [SUSTAIN-FLOW]
38 AGENT.
39 [SIGNAL-DESIRE-TO-RETAIN-TURN] {...}
40 {...}
41 {...} [SIGNAL-TRP]
42 CUSTOMER.
43 [SUSTAIN-FLOW]

Using *A* for the agent and *B* for the customer, the sequence of conversational actions can be denoted as follows:[31]

A_1
B_2B_{10}
A_{21}
$B_{14}B_{14}B_{10}$
A_{21}

---

[30] {...} designates a stretch of talk.

[31] Each line represents a stretch of talk by a speaker (not necessarily a turn).

$B_{14}B_{14}B_{10}$
$A_{12}A_{14}A_{10}$
$B_{12}B_{14}B_{14}B_{10}$
$A_{12}A_{10}$
B_{21}
$A_{14}A_{10}$
B_{21}
$A_{14}A_{10}$
B_{21}

Using the grammar, the above sequence can be parsed, top down, as follows:

S $==>$ S^{A}
$==>$ $OPEN - OK^{A}S^{B}$
$==>$ $A_1B_2S^{B}$
$==>$ $A_1B_2EXCHANGE - NO^{B}S^{B}$
$==>$ $A_1B_2TRP - NO^{B}S^{B}$
$==>$ $A_1B_2B_{10}A_{21}S^{B}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}S^{B}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}S^{B}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}EXCHANGE - NO^{B}S^{B}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}TRP - NO^{B}S^{B}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}S^{B}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}S^{B}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}S^{B}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}EXCHANGE - OK^{B}S^{A}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}TRP - OK^{B}S^{A}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}B_{10}A_{12}S^{A}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}B_{10}A_{12}A_{14}S^{A}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}S^{B}B_{10}A_{12}A_{14}EXCHANGE - OK^{A}S^{B}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}B_{10}A_{12}A_{14}TRP - OK^{A}S^{B}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}B_{10}A_{12}A_{14}A_{10}B_{12}S^{B}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}B_{10}A_{12}A_{14}A_{10}B_{12}B_{14}S^{B}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}B_{10}A_{12}A_{14}A_{10}B_{12}B_{14}B_{14}S^{B}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}B_{10}A_{12}A_{14}A_{10}B_{12}B_{14}B_{14}EXCHANGE - OK^{B}S^{A}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}B_{10}A_{12}A_{14}A_{10}B_{12}B_{14}B_{14}TRP - OK^{B}S^{A}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}B_{10}A_{12}A_{14}A_{10}B_{12}B_{14}B_{14}B_{10}A_{12}S^{A}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}B_{10}A_{12}A_{14}A_{10}B_{12}B_{14}B_{14}B_{10}A_{12}S^{A}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}B_{10}A_{12}A_{14}A_{10}B_{12}B_{14}B_{14}B_{10}A_{12}EXCHANGE - OK^{A}S^{B}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}B_{10}A_{12}A_{14}A_{10}B_{12}B_{14}B_{14}B_{10}A_{12}TRP - OK^{B}S^{A}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}B_{10}A_{12}A_{14}A_{10}B_{12}B_{14}B_{14}B_{10}A_{12}S^{A}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}B_{10}A_{12}A_{14}A_{10}B_{12}B_{14}B_{14}B_{10}A_{12}EXCHANGE - NO^{A}S^{A}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}B_{10}A_{12}A_{14}A_{10}B_{12}B_{14}B_{14}B_{10}A_{12}TRP - NO^{A}S^{A}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}B_{10}A_{12}A_{14}A_{10}B_{12}B_{14}B_{14}B_{10}A_{12}A_{10}B_{21}S^{A}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}B_{10}A_{12}A_{14}A_{10}B_{12}B_{14}B_{14}B_{10}A_{12}A_{10}B_{21}A_{14}S^{A}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}B_{10}A_{12}A_{14}A_{10}B_{12}B_{14}B_{14}B_{10}A_{12}A_{10}B_{21}A_{14}EXCHANGE - NO^{A}S^{A}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}B_{10}A_{12}A_{14}A_{10}B_{12}B_{14}B_{14}B_{10}A_{12}A_{10}B_{21}A_{14}TRP - NO^{A}S^{A}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}B_{10}A_{12}A_{14}A_{10}B_{12}B_{14}B_{14}B_{10}A_{12}A_{10}B_{21}A_{14}A_{10}B_{21}S^{A}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}B_{10}A_{12}A_{14}A_{10}B_{12}B_{14}B_{14}B_{10}A_{12}A_{10}B_{21}A_{14}A_{10}B_{21}A_{14}S^{A}$
$==>$
$A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}B_{10}A_{12}A_{14}A_{10}B_{12}B_{14}B_{14}B_{10}A_{12}A_{10}B_{21}A_{14}A_{10}B_{21}A_{14}EXCHANGE - NO^{A}S^{A}$
$==>$ $A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}B_{10}A_{12}A_{14}A_{10}B_{12}B_{14}B_{14}B_{10}A_{12}A_{10}B_{21}A_{14}A_{10}B_{21}A_{14}TRP - NO^{A}S^{A}$

$$= = > A_1B_2B_{10}A_{21}B_{14}B_{14}B_{10}A_{21}B_{14}B_{14}B_{10}A_{12}A_{14}A_{10}B_{12}B_{14}B_{14}B_{10}A_{12}A_{10}B_{21}A_{14}A_{10}B_{21}A_{14}A_{10}B_{21}S_1^A$$

5.10 Summary

In this chapter, we analyzed the conversational function that meta-statements fulfill. As we have seen, interactants involved in a conversational exchange observe a well-defined set of protocol rules. These rules, we argue, are of vital importance since without them the exchange would be impossible to carry out.

The major functions that the conversational protocol fulfills are: the opening of a conversation, the exchange of turns between speakers, the resolution of turn disputes and interruptions, and the closing of a conversation. We discussed the major issues each of these functions raises, and we presented some solutions the protocol offers.

The many examples presented in this chapter illustrate that meta-statements play an important role in the management of the conversational aspect of conversations. To model this function of meta-statements, we have proceeded as we did in Chapter Four by specifying of the conversational state model and the conversational action model. But whereas a rhetorically sensitive interpreter offers a more coherent system, an interpreter sensitive to the protocol of conversing offers a system that makes for a better conversational partner.

Chapter 6. Role Theory

Chapters Seven and Eight of this thesis will focus on two "talk about talkers" functions of meta-statements, the politeness function and the emotion function, respectively. As we will see in those chapters, to be able to model these two functions, we must first have a model of the social context within which the interactants we are modeling are acting. That is, we must have an idea about what behavior is expected from the interactants and how the interactants stand in relation to the other interactants. A model of the *social role* enacted by interactants is such a social context.

In the case of politeness, it is easy to see the importance of modeling enacted roles. An act performed by person A to person B may be polite if person A is enacting role 1 and person B is enacting role 2, but may be impolite if person A is enacting role 3 and person B

role 4. For example, the action of patting a friend on the back is polite, whereas the action of patting a total stranger is most probably not.

Similarly, what emotional reaction a person will have will depend on the role that person is playing. For example, an order we receive from our boss will probably not hurt our ego as much as an order from a peer or an underling.

Therefore, in order to be able to parse the politeness and emotive meanings of meta-statements, we must first have a way of modeling the roles being enacted by interactants. In the sections that follow, we specify such a model. First, however, we examine the problem of modeling roles in general.

6.1 The Role Theory Problem

There is no lack of confusion regarding *what* the term "role" precisely refers to. A quick survey of the literature on role theory readily reveals that there are almost as many definitions of the concept of role as there are people trying to define that concept. Parsons defines a role to be "what the actor does in his relations with others" [PARS51]; Cottrell sees it as "an internally consistent series of conditional responses;" while Goffman regards a role to consist of "the activity the incumbent would engage in were he to act solely in terms of the normative demands upon someone in his position" [GOFF81]. Confusion also arises on the issue of whether the term "role" refers to the actions *expected* from a person playing the role, or whether it refers to the actions *performed* by the role player. Some insist that the distinction between what is expected and what is performed is fundamental; others dismiss it as superfluous and irrelevant. And yet others argue that roles should not be regarded simply as patterns of behavior, but rather as the underlying causes that bring into effect those patterns. And instead of the term "role," some use the term "status," while others use the term "position." In short, what Neiman and Hughes [NEIM51] remarked forty years ago still holds quite

true today: "the concept of role is at present still rather vague, nebulous, and non-definitive." The problem of defining precisely what the concept of role refers to we shall term *the role theory problem.*

The notion of role is intuitively appealing. Long before sociologists recognized its value as an analytical tool, the layman already talked and thought in terms of roles: he bought his bread from the "baker," his meat from the "butcher," his milk was brought by the "milkman," while his letters were delivered by the "postman," etc. But when we approach the role concept analytically we are at once confronted with many theoretical complications. First, what *is* a role? On what is the notion of role based? Is "wife," for instance, a role? Is "mayor" a role? Surely, we do often say that so-and-so is playing the role of "wife," or that so-and-so is playing the role of "mayor." But are both "wife" and "mayor" roles in the same sense? A wife, for instance, is wife only to one person (at least in the western culture), her husband. A mayor, on the other hand, is mayor of the entire city he governs. Can we, then, say that both wife and mayor are roles in the same sense? And let us take an even more disturbing case. Let us say that a certain Mr. Jones is an employee in a factory that manufactures shoes, and let us say that his job in that factory is to monitor the quality of the shoes produced there. Let us also suppose that he heads a team of specialists, and that at the end of each month he writes a report which he submits to his supervisor and in which he discusses the findings he and his team have gathered. Now, the question is: can we say that Mr. Jones plays the *roles* of "employee," "head of the quality production team," and "quality report writer"? How can "employee" and "head of the quality production team" be both roles if part of being an employee (i.e., playing the "role" of employee) consists of being the "head of the quality production team"? And how can "head of the quality production team" and "quality report writer" both be "roles" if writing the quality report--playing the "role" of "quality report writer"--is part of heading the quality production team--playing the "role" of "head of the quality production team"? How can a whole and part of that whole both be entities of the same type?

Obviously, then, there is a definitional problem. Some "roles" are defined in terms of certain other more "basic" "roles": our employee has a supervisor, and thus plays the "role"

of "subordinate" with his supervisor, but he is also himself a supervisor, and thus plays "supervisor" with his subordinates. "Subordinate" and "supervisor," in a sense, are more "basic" than "employee." Both, along with other "roles," form the "role" of "employee." But then, if a "role" is composed of other "roles," how can we define the concept of "role" without using the concept itself in the definition? That is, how can we avoid a circular definition?

The first step toward finding a solution to this problem will consist in identifying the most basic concept on which the notion of what is generally called "role" is based. Having identified this basic concept, our subsequent steps in solving the problem will require the construction of a hierarchical structure for modeling human activity, where the entities in each level of the hierarchy are constituted from elements in the level immediately below it. Thus, instead of one level of abstraction, where *any* activity is said to constitute a role, we would be dealing with many levels of abstractions, starting from the most basic social activities and ascending to levels of more and more complex structures of behavior.

The works of Bates and Harvey[32] in the area of modeling human social systems are particularly relevant to our purposes [BATE75]. Although the immediate aim in their works is "to build a structural model of society by defining, or, more properly, redefining a set of concepts that will allow us to conceive in the abstract what societies are like as systems" [BATE75.17], the structural model they do present offers a powerful paradigm for solving the role theory problem. The appeal of their approach stems mainly from its "bottom up" nature. They start by identifying the most basic elements on which their theory is to be based, and then proceed in a constructive fashion to define new concepts to account for increasingly complex social structures. The discussion that follows, and the terminology used in that discussion, are based on their theory.

---

[32] I am indebted to Dr. Jack Dudley for his help in this area, and particularly for having brought to my attention the works of Bates and Harvey.

6.2 Actions and Functions

The most basic element in social behavior is the *action.* It is single, distinct actions that people perform, and individual acts are the smallest identifiable and *observable* units of social behavior with which we must start our analysis. As we already said in the introduction, however, social behavior cannot be explained only in terms of individual actions. Social actions do not occur in a social vacuum; they occur in a social context, and only within that context do they acquire their social meaning. Thus, in order to explain behavior, a level of abstraction higher than that of action must be adopted.

This level of abstraction is the *function* level. Actions are performed to fulfill a function, and their meaning is directly inherited from the functions they fulfill. For example, the actions of a Moslem praying: bowing the head, kneeling, whispering, etc., are actions directly observable by anyone, but only those who know or are told in advance that the person performing these activities is praying could understand what the actions *mean,* i.e., what function they fulfill. Without an idea about the function fulfilled, the actions would be meaningless to the observer.

6.3 Social Norms

The notion of function by itself does not, however, explain the presence of regularity in social behavior. Why do *all* moslems, for example, pray in almost the exact same fashion? Why do they all perform not only the same actions, but even follow the same sequence of actions? And in general, how can the prevalence of order in social behavior be explained?

Two explanations are possible. We could argue that regularity comes from *personality,* and that people who act alike have *similar* personalities. Or we could argue that *culture* is the

source of regularity in social behavior, and that people act alike because they have the same culture.

The first suggestion can be easily discarded by observing that not all people who engage in some pattern of behavior have similar personalities. Moslems who pray are not all alike. "Moslem" is not a personality attribute, but a statement about the person's religious beliefs. Thus, the prayer routine of a moslem is not due to his personal disposition. A moslem does not pray as he or she does because of some inherent personal characteristic. Moslems pray as they do because they share the same culture. That is, it is from the shared culture that the regularity comes.

Culture, then, is the source of regularity. But as a monolothic entity, the concept of culture cannot be very useful to our "bottom up" approach. As we said, people behave through individual actions, and to introduce the notion that people's behavior is affected by culture, we must bring the concept of culture to the level of action. In other words, we must break down the idea of culture as an encompassing whole into its most basic elements. The most basic elements of a culture are its *norms.*

A norm is defined by Bates and Harvey [BATE75.64] as a "program or blueprint for action." A norm is a guideline for how one is *expected* to act, and thus has inherent in it the notion of "should" or "ought." To be a member of a culture, one has an *obligation* to observe the norms dictated by that culture. To ignore those norms, one may risk punishment and perhaps even expulsion from society.

Norms become actions through individual behavior, and when an individual is acting a norm we say that the norm is being *enacted.* Of course, norm enactments are not identical from one individual to the next. For instance, although all moslems follow closely the norms for praying, nonetheless each moslem prays somewhat differently from others. For instance, one with an easy-going personality might perhaps perform the prayer movements smoothly and unhurriedly, while another, with a more agitated character, might perform his movements more bruskly. If in a hurry, one might perform the prayer very rapidly, whereas if not pressed by time one might take longer to complete his prayer. In short, norms interact with other

factors, such as personality and situational circumstances, to determine the exact actions to be performed.

6.4 Roles, Statuses, and Positions

The notion of role we shall be discussing in this section differs fundamentally from the one we discussed in section 6.2. As we noted there, the general term "role" is ambiguous and problematic. It is used to refer to many different types of social behavior. We also saw that the problem inherent in the notion of "role" was that it referred to social structures of varying complexity, so that in many instances a "role" was composed of other "roles." This posed a definitional problem, since the definition of "role" had to include the notion of role in it, and thus the task of defining "role" was rendered impossible.

To avoid the "lumping together" of structurally different concepts by using one ambiguous term, we will make a clear distinction between those concepts. The first step in making this distinction is to use an unambiguous terminology: we must use different terms for different concepts. In this section we discuss three such concepts: *role, status,* and *position.*

The definition of the three concepts of role, status, and position can be greatly simplified if the concepts are illustrated through an example. Suppose the Jones family is composed of Mr. Jones, Mrs. Jones, Jones Jr., and Miss Jones. These four individuals together form a *social group,* viz. the Jones family. Each of the members of the Jones family is expected to behave according to certain norms--norms that the culture (viz. the American culture) deems appropriate for family behavior. Each of the individuals also stands in a relationship with the other members of the family: Mr. Jones is the father of Junior and Miss Jones, Mrs. Jones is the mother of Junior and Miss Jones, Miss Jones is the sister of Junior, who is, therefore, her brother, while Mr. Jones is also the husband of Mrs. Jones, who thus is his wife, etc. Figure

10 shows a scheme encoding some of the possible relationships between the members of this social group.

Mr. Jones, who is both father and husband in his family, is said to occupy the father-husband *position.* He is father to Junior, father to Miss Jones, and husband to Mrs. Jones. The relationship he holds with each of the members of his family defines one of his *statuses.* That is, Mr. Jones holds three statuses in the family: he is father to Junior--"father-to-son" status; father to Miss Jones--"father-to-daughter" status; and husband to Mrs. Jones--"husband-to-wife" status. Each of Mr. Jones's relationships comes with a set of expectations about Mr. Jones's behavior. With his children, for instance, he is expected to behave as a father should behave. Fatherly behavior includes being a "disciplinarian," a "provider," a "protector," etc. With his wife, on the other hand, he is expected to be a "sex partner," a "provider," as well as a "protector." Each of Mr. Jones's behavior expectations we call a *role.* Thus a status is a collection of roles.

Each of Mr. Jones's roles fulfills some particular function. The disciplinarian role's function, for example, is to punish a subordinate if the subordinate "misbehaves;" a protector's function is to protect, while a provider is supposed to provide for those who depend on him. Each of these functions is expected to be fulfilled in a certain way, through certain types of actions. Punishing, for instance, may be done by Mr. Jones by grounding his children, or perhaps by spanking them. If noises are heard late at night in the house, Mr. Jones is expected to be the one to get up and survey the premises for any prowlers. Mr. Jones is also expected to work in order to provide for his wife and children. These actions are the *norms* of behavior Mr. Jones is supposed to follow when aiming to fulfill the functions of his roles.

In summary, then, a role is a set of norms, a status is a set of roles, while a position is a set of statuses. Figure 11 depicts a structural scheme for encoding the hierarchy of these three concepts.

6.5 *Situs and Station*

We have hitherto focused on human behavior within *groups* of actors. We saw that each member of any given group occupies a position within that group, and we defined a position as a set of statuses. We saw that the positions forming a group were inter-related with one another through complementarity of role functions, and we argued that for each role there existed a complementary role. The abstraction of human behavior patterns, however, does not end with the group level. The group is not the largest social structure in which human behavior can take place. Mr. Jones, for instance, belongs to at least two groups: the group he heads, and where he is the supervisor, and the group where he is subordinate and which his supervisor heads. When we want to talk about Mr. Jones's behavior, without specifying whether it is occurring in one group or another, we must abstract our description so that the distinction between the two groups is eliminated. When we say that Mr. Jones is an employee, we are in fact doing precisely that; i.e., we are abstracting our observation. As an employee, Mr. Jones is both a supervisor and a subordinate. He is supervisor in the "quality control team," and he is subordinate when interacting with his superiors.

The collection of groups to which Mr. Jones, belongs, as an "employee," is termed *an organization.* Mr. Jones is said to be a member of his work organization, and whereas he occupies positions in the groups forming that organization, we say that he occupies a *situs* in the organization structure. A situs, then, is a collection of positions.

But Mr. Jones does not belong only to his work organization. He also belongs to at least one other organization: the one that includes his family and his friends and that may very well be totally different from his work organization. Mr. Jones may also belong to the local church, or the local golf club. He may in addition be involved in a local crime watch group, or he may be a volunteer to do some community service. In short, Mr. Jones is likely to be involved in many organizations within his society. Abstracting to the level of society, we say that the collection of activities that Mr. Jones is involved in, as a member of society, constitutes Mr.

Jones's *station.* The collection of organizations to which Mr. Jones belongs constitute his *community.* Thus, a station is a set of situses.

With the notions of situs and station, we can talk about Mr. Jones in abstract ways, ways that would not have been possible if we had restricted our analysis only to groups, i.e., only to sets of individuals. The abstraction of situs enables us to talk about Mr. Jones as an "employee." "Employee" is not a role, since it is not based on one function. It is not a status, since an employee is not a person-to-person relationship, and it is not a position, since employees do not form a group of complementary statuses. Instead, it is a situs: Mr. Jones is member of an organization, and within that organization he belongs to many groups.

In addition, we can talk about Mr. Jones the "middle-class male." "Middle-class male" also, like "employee," is not a role. Nor is it a status or a position. And it is not a situs either, since a middle-class male can belong to many organizations, and being a middle-class male does not imply belonging to any particular organization. Instead it is a station. Figure 12 depicts a possible station structure for a middle-class male.

With the definition of station we have completed the theory specification of the hierarchical structure on which we will base our behavioral analysis. We began the specification of this structure by identifying the most basic elements of human behavior, the social norm. We then proceeded with the identification of increasingly complex structures of human behavior, where at each level of complexity behavioral patterns were defined in terms of patterns at the level below it. We thus saw that there exist five levels of human behavior abstractions: the role, the status, the position, the situs, and the station.

6.6 *Internal Representations*

We now turn to a description of the data structures we will use to model norms, roles, and statuses.

Internally, the enactment of a role is represented as follows:

```
(ROLE-ENACTED
        (ROLE-TAG ?role-tag)
        (ROLE     ?role))
```

with ROLE-TAG being a unique tag identification number, and ROLE being a frame with the following structure:

```
(ROLE
        (ROLE-ID  ?role-id)
        (ACTOR    ?actor)
        (FUNCTION ?function)
        (NORMS    ?norms))
```

with, ROLE-ID being the identification number identifying the role enacted, i.e., the first column of Table 5 in Appendix B given in [BOUZ91]; ACTOR being the speaker enacting the role in question (given in column two of Table 5); FUNCTION being the function on which the role is based (given in column three); and NORMS being the list of norms that make up the role (given in column four).

While enacting a role, an actor will not necessarily enact all the norms expected or acceptable for that role; a norm he does enact is modeled internally as follows:

```
(NORM-ENACTED
        (NORM-TAG    ?norm-id)
        (NORM-SPECS  ?specs))
```

with NORM-TAG being a unique tag used to represent the instance of an enactment by a speaker of a norm, while NORM-SPECS is a frame of the following form:

```
(NORM-SPECS
        (NORM-ID       ?norm-id)
        (NORM-CATEGORY ?cat)
        (NORM-DATA     ?data)
        (NORM-TYPE     ?type))
```

with NORM-ID being the identification number of the norm that is being enacted (i.e., ?norm-id is one of the identification numbers given in Tables 3 and 4 of Appendix B given in [BOUZ91]);

NORM-CATEGORY being the category to which the norm belongs; NORM-DATA being the semantic data needed to specify the norm; and NORM-TYPE being the norm's type (i.e., OB for obligatory, NE for normally expected, and OP for optional). Tables 3 and 4 in Appendix B in [BOUZ91] give the specifications of the ninety one norms used in our system.

A status is modeled as follows:

```
(STATUS
        (STATUS-ID ?status-id)
        (ROLE-IDS  ?role-ids))
```

with STATUS-ID being the identification number used for the status in question (column 1 of Table 6 in Appendix B in [BOUZ91]), and ROLE-IDS indicating the identification numbers of the roles that constitute the status (column 2 of Table 6).

In the following section, some of the concepts we have been discussing in the previous sections are illustrated through analysis of transcribed conversations.

6.7 *Application of Role Theory to the Conversations*

We give in what follows a brief illustration of how we will use the model we presented in the previous section to represent the social context of interactions.

Norms, Functions, and Roles

The social group we are focusing on in our analysis is the agent-customer group. Since the size of the group is two, the maximum number of statuses composing the position of each group member is one.

We identify at least three roles for each of the two statuses. For the agent status, the roles are: information provider, information seeker, and seller. For the customer status, the roles are: information seeker, information provider, and buyer. For each of these roles there is a function. A total of six functions are thus identified. The function of an information provider is to provide information appropriate in the situation; the function of an information seeker is to seek relevant information; the function of a seller is to convince a potential buyer to buy what he has to sell, while the function of a buyer is to buy what he wants to buy. Each of these functions is surrounded by norms of behavior. Norms, as we said, are expectations about behavior. The norms around the information provider dictate what sort of information the provider is expected to provide. The norms around the information seeker function dictate what sort of information a seeker is expected to seek.

In what follows we have compiled the norms enacted in our transcripts. The list obviously does not exhaust all possible norms. The norms are given in table form in tables 3 and 4 of Appendix B in [BOUZ91].

1. A requests to be told: [33]

(about)
1 - C's last name.
2 - C's first name.
3 - C's home telephone number.
4 - C's office telephone number.
5 - the departure time.
6 - the arrival location.
7 - the departure date.
8 - the type of the dog C is taking along.

(whether)
9 - C wants to cancel previous reservation.
10 - C has any preference for a departure time.
11 - C has any preference for an arrival time.
12 - C has kennel for dog.
13 - C is buying ticket now.
14 - C can drive to departure location.
15 - C wishes to have tickets mailed.
16 - C wants reservations to be cancelled.
17 - C needs clarifications.

---

[33] The A stands for agent, while the C stands for customer.

2. C requests to be told:

(about)
1 - the arrival time.
2 - the departure time.
3 - the stops involved.
4 - altervative flight routes.
5 - the duration of flight.
6 - the plane type.
7 - ticket price.
8 - length of layovers.
9 - the number of layovers.

(whether)
10 - A had already made a reservation.
11 - any special arrangements are made for handicaps.
12 - help for handicaps will cost more.
13 - a kosher meal can be provided.
14 - a kosher meal will increase price.
15 - types of meals offered depend on the departure time.
16 - costs of alternatives differ.
17 - dog will be walked in stop overs.
18 - dog will be fed and given water in stop overs.
19 - there would be enought time between flights to walk the dog.
20 - the airline assumes any liability if dog is injured.

3. C Informs A:

(about)
1 - departure date.
2 - departure location.
3 - arrival location.
4 - C's name.
5 - C's home telephone number.
6 - C's office telephone number.
7 - the weight and the type of dog.
8 - C's fears of airplane model type.

(whether)
9 - C does not want to make tentative reservation.
10 - C wants to cancel previous reservation.
11 - C is making a reservation.
12 - C is buying a ticket now.
13 - trip is round trip.

(that)
14 - C is handicapped.
15 - it is C's first time flying.
16 - C taking a dog along.
17 - C fears flying.

4. A Informs C:

(about)
1 - availability of kosher meal.
2 - availabilty of help for handicapped persons.
3 - existence of direct flights betwen two locations.
4 - existence additional charges for handicapped persons.
5 - the least expensive flight option.
6 - special rates.
7 - airline name.
8 - flight number.
9 - snacks offered.
10 - type of airplane.
11 - kennel sizes available.
12 - price.
13 - cost of each kennel size.
14 - additional cost for kosher meal.
15 - time constraints on flight itinerary changes.
16 - arrival time to stop location.
17 - arrival time.
18 - the arrival airport.
19 - departure time from stop location.
20 - departure time from original location.
21 - length of flight.
22 - the number of stops.

(whether)
23 - kennel is for C to keep.
24 - dog is not fed or given water during stop.
25 - dog is to be walked by owner during stops.
26 - company takes responsibility.
27 - ticket prices vary with the day of the week they are purchased.
28 - buying now is better.
29 - taxes are included.

(that)
30 - C's name is needed for holding seats.
31 - meal types depend on the departure time.
32 - buying now is better.
33 - the price given is the cheapest fare.

5. A suggests to C:

(that)
1 - C make a tentative reservation.
2 - C call back to confirm reservation.
3 - C be at airport ahead of time.
4 - C inform himself about airplanes

(How)
5 - C can reduce the layover times by taking alternate flights.

The above norms are then clustered around the following six roles:

Information Provider (Agent): [4.1-36].
Information Seeker (Agent): [1.1-17].
Seller (Agent): [1.1-4,9,13,15,16],[4.5,6,12,13,14,27,28,29,32,33],[5.1,5].

Information Provider (Customer): [3.1-17].
Information Seeker (Customer): [2.1-20].
Buyer (Customer): [2.7,12,14,16],[3.4-6,8-11].

The roles are given in table form in Table 5 of Appendix B in [BOUZ91].

Status, Position, and Situs

Let us examine our customer-agent interaction situation in terms of the statuses, positions, and situses it involves.

As we have said already, since the size of our group is two, then the number of statuses is also two, one for each group member. Consequently, each group member's position is composed of one status. Figure 13 shows the positions in the agent-customer group.

However, agents do not interact only with customers. They also interact regularly with their colleagues, for example, whom they might occasionally seek for assistance when facing a problem they are unable to solve by themselves. In addition, agents interact with their supervisors, or with their subordinates (if they have any), etc. The customer also can interact with someone other than the agent. For instance, he might request to speak to the agent's supervisor.

Figures 14, 15, and 16 respectively show the colleagues group, the agent-supervisor group, and the customer-supervisor group. In each group, each member of that group is said to hold a position. The agent holds a position in the customer-agent group and a position in the customer-supervisor group. The agent also has a position in the customer-agent group and a position in the agent-supervisor group. The customer's two positions form his situs in the "travel agency" organization (assuming, for the sake of simplicity that a customer has only these two positions in the organization). The agent also occupies a situs in the "travel

agency" organization, a situs composed of the positions he occupies in the customer-agent group and the agent-supervisor group.

There was one instance in our transcribed dialogues[34] where a customer requested to speak with the agent's supervisor. The customer in that particular instance was in a hurry and felt that the agent he was dealing with was not diligent enough in fulfilling his requests. With the introduction of the supervisor in the interaction, three social groups were thus brought together: customer-agent, customer-supervisor, and agent-supervisor. Figure 17 shows the situation schematically.

We end our role theoretical analysis of the transcripts by briefly discussing some distancing behavior exhibited in our transcripts.

6.8 Summary

Among the functions fulfilled by meta-statements are the politeness and emotion functions. To model these two functions, we argued that we need to build a model of the social context within which talk (and therefore meta-statements) is taking place. Modeling the roles played by interactants is such a model.

We began this chapter with a general inquiry about the meaning of the term "role." Intuitively, one can easily grasp what is meant when it is said that someone is playing a role, but how is one to model the notion of role precisely?

We saw that the problem inherent in defining what "role" means comes from the use of one term to refer to social structures of varying complexity. The term role, for example, is used for "mother" as well as for "middle aged suburban woman." The behavior of a mother

---

[34] Conversation roleplay.04, lines 184-185.

is very much restricted to a particular type of behavior, whereas the behavior of a middle aged suburban woman is much broader and may even include the behavior of a mother. Thus a distinction between behaviors in terms of complexity needed to be drawn.

A paradigm for modeling behavior is presented by Bates and Harvey. We saw that they differentiate between five levels of behavioral complexity: role, status, position, situs, and station. The most basic notions on which a role is based are those of function and norm. A role is composed of a function and a cluster of norms. A status is a cluster of roles, a position a cluster of statuses, a situs a cluster of positions, and a station a cluster of situses.

In addition to investigating the concepts involved in the definition of the five levels of behavioral complexity, we also discussed the issues of behavioral conflict and distancing. We ended the discussion with an analysis of a conversation in terms of its role theoretical aspect.

As we said at the outset of this discussion, people do not behave in a social vacuum. Each action is part of a pattern of actions. Thus, to understand an action, it will not suffice for an observer to observe that action by itself. The observer must also have an idea about the pattern to which the action belongs. Without an understanding of this pattern, the action will have no meaning. Since one of our basic goals is to model human understanding and to interpret human actions, a mechanism for modeling patterns of human social behavior is essential. We presented in this chapter a paradigm for constructing such a model.

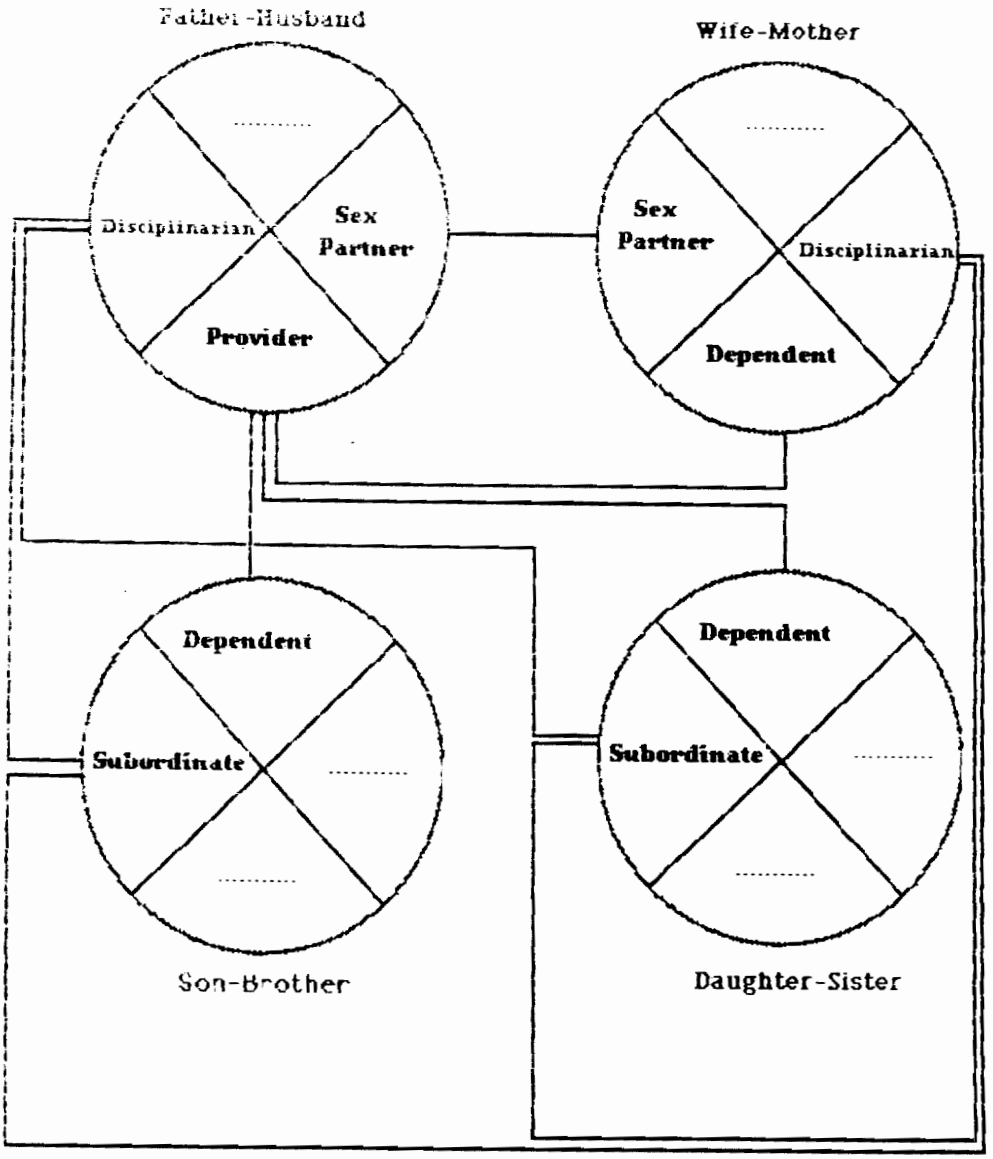


Figure 10. Family Group: Statuses and Positions

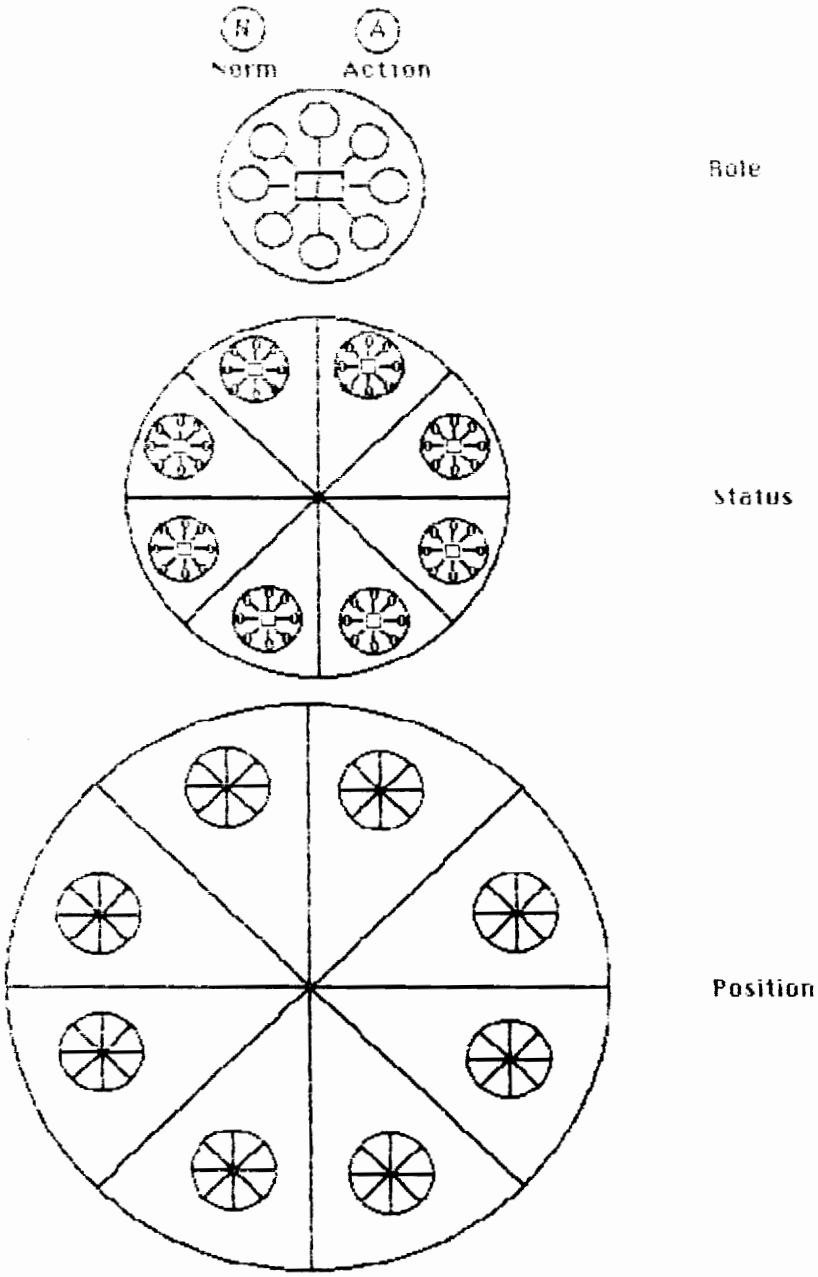
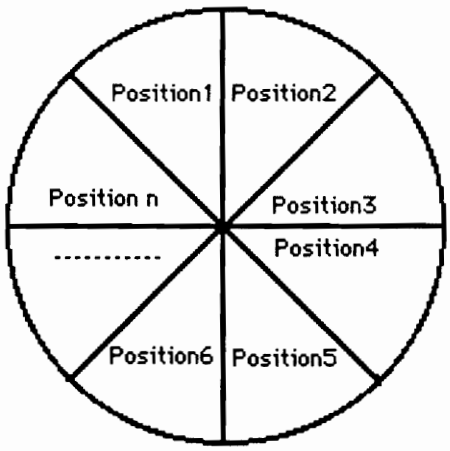
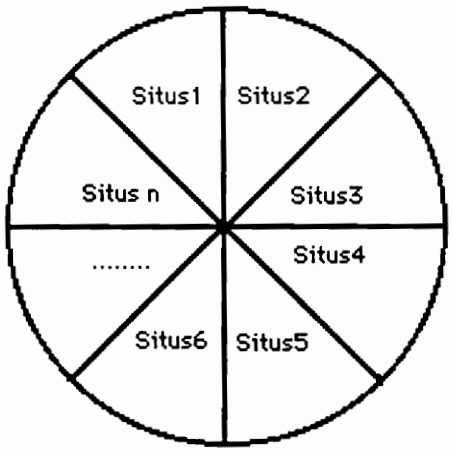


Figure 11. The Structure of Role, Status, and Position



Situs={position1 ,...,positionn}.



Station={Situs1 ,...,Situsn}.

Figure 12. The Structure of Situs and Station

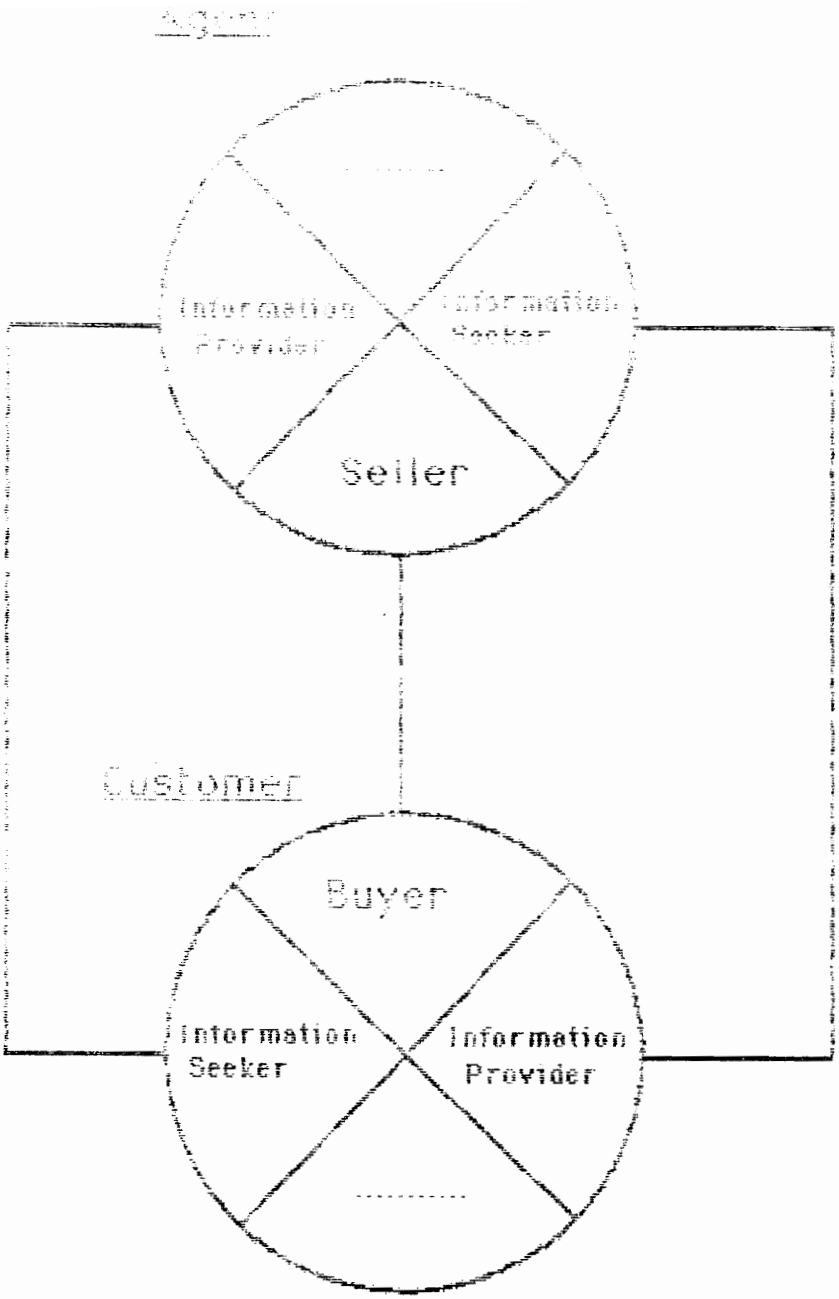


Figure 13. The Agent-Customer Group

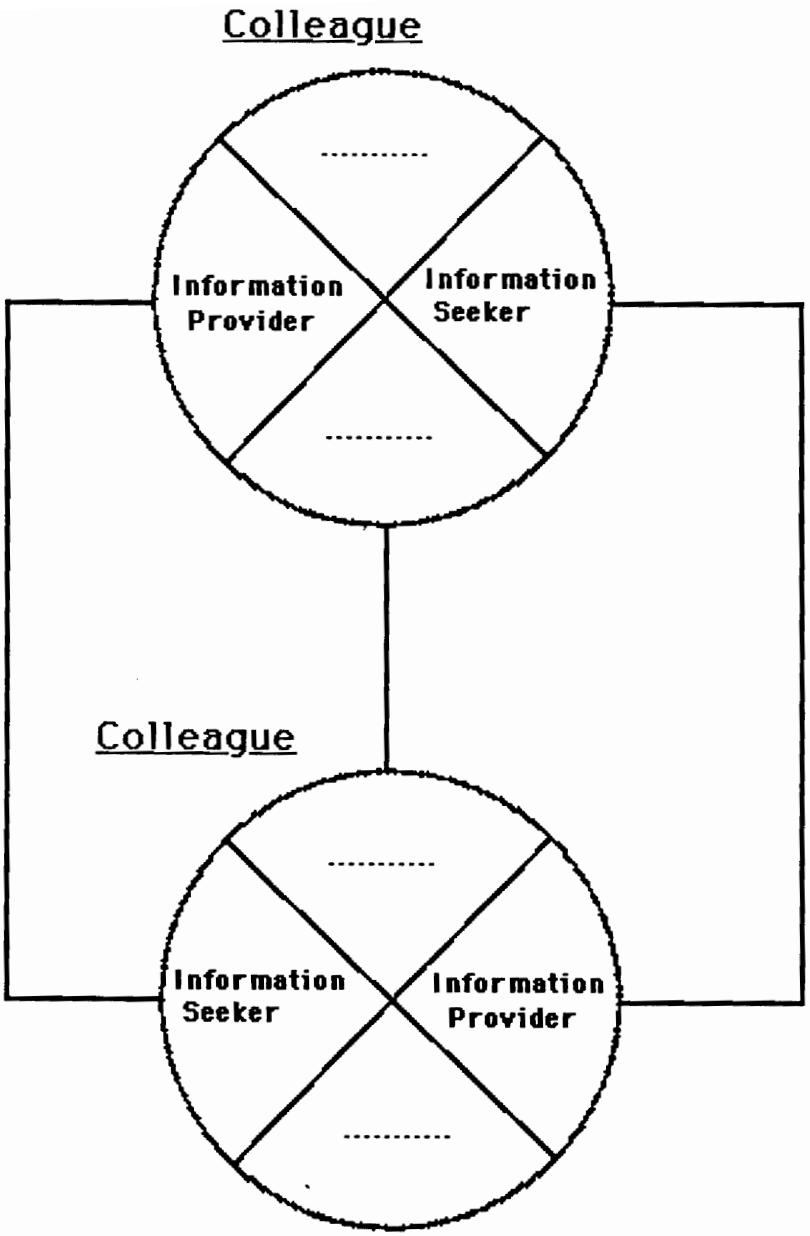


Figure 14. The Colleague-Colleague Group

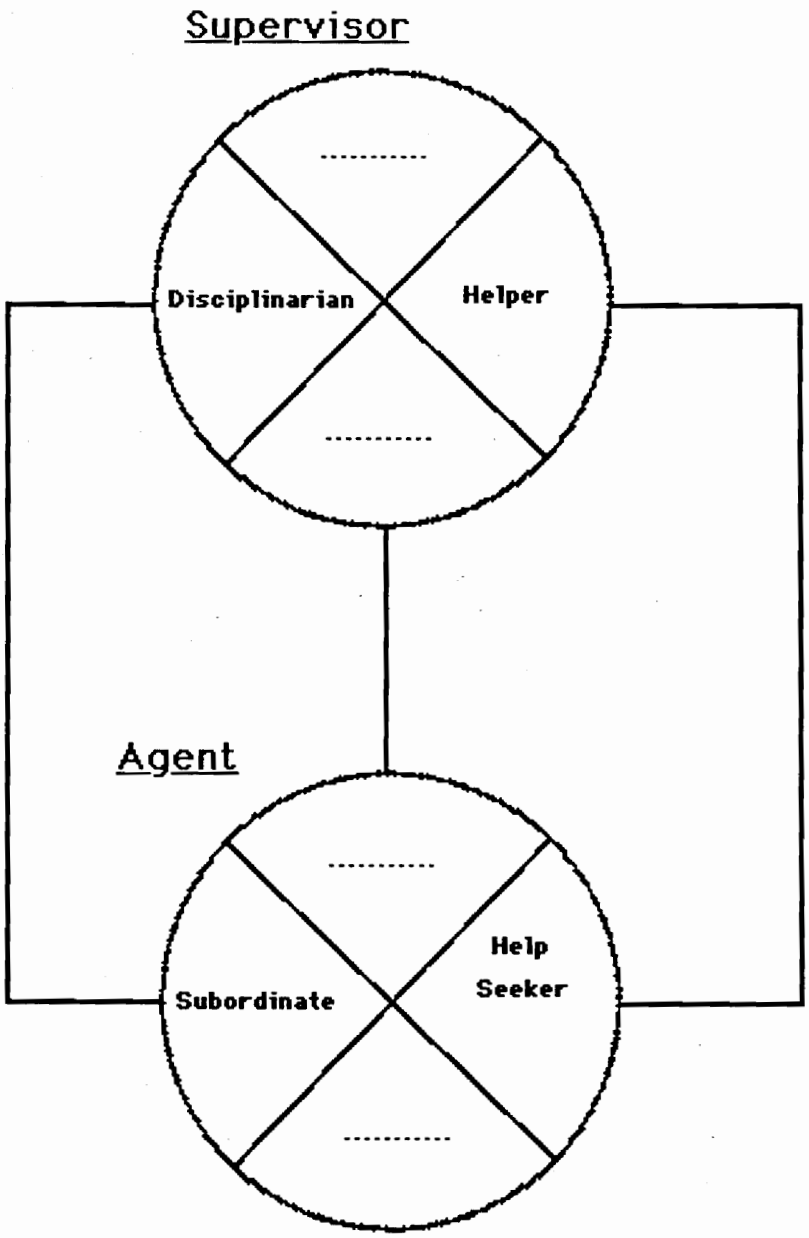


Figure 15. The Agent-Supervisor Group

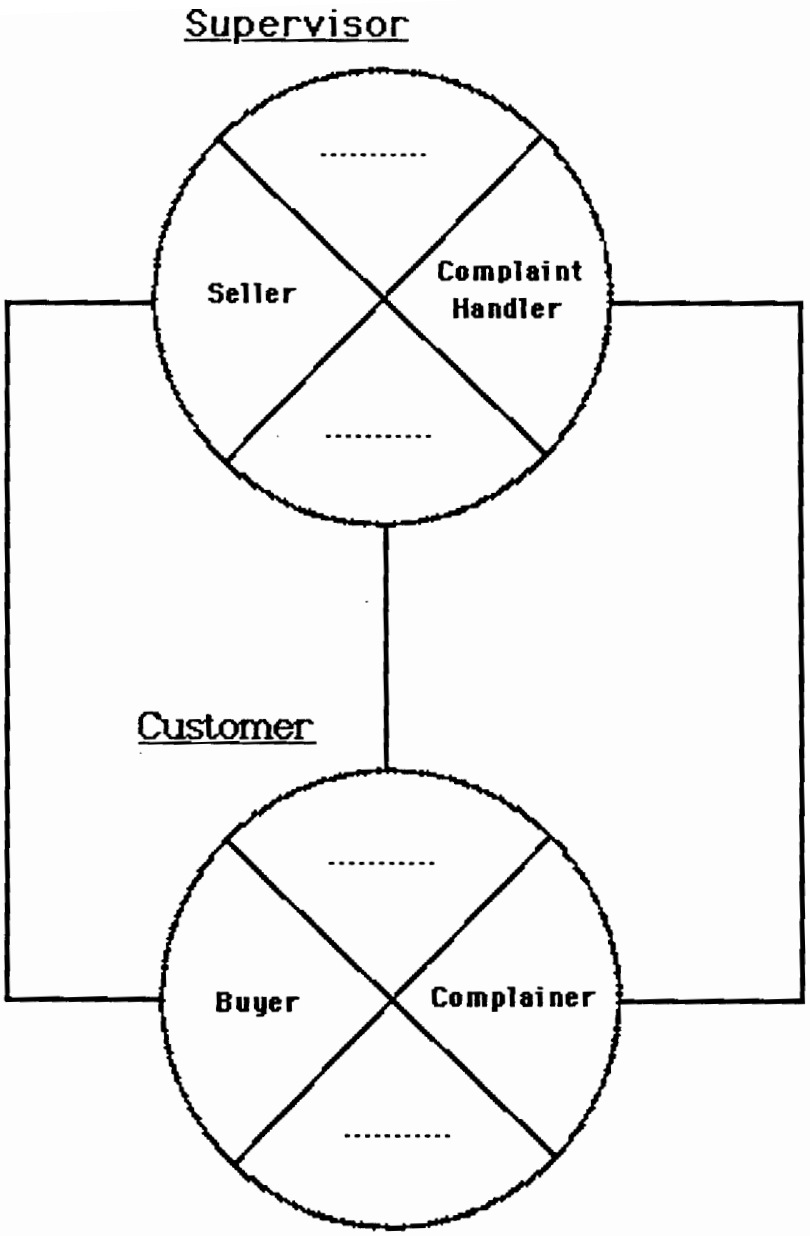


Figure 16. The Customer-Supervisor Group

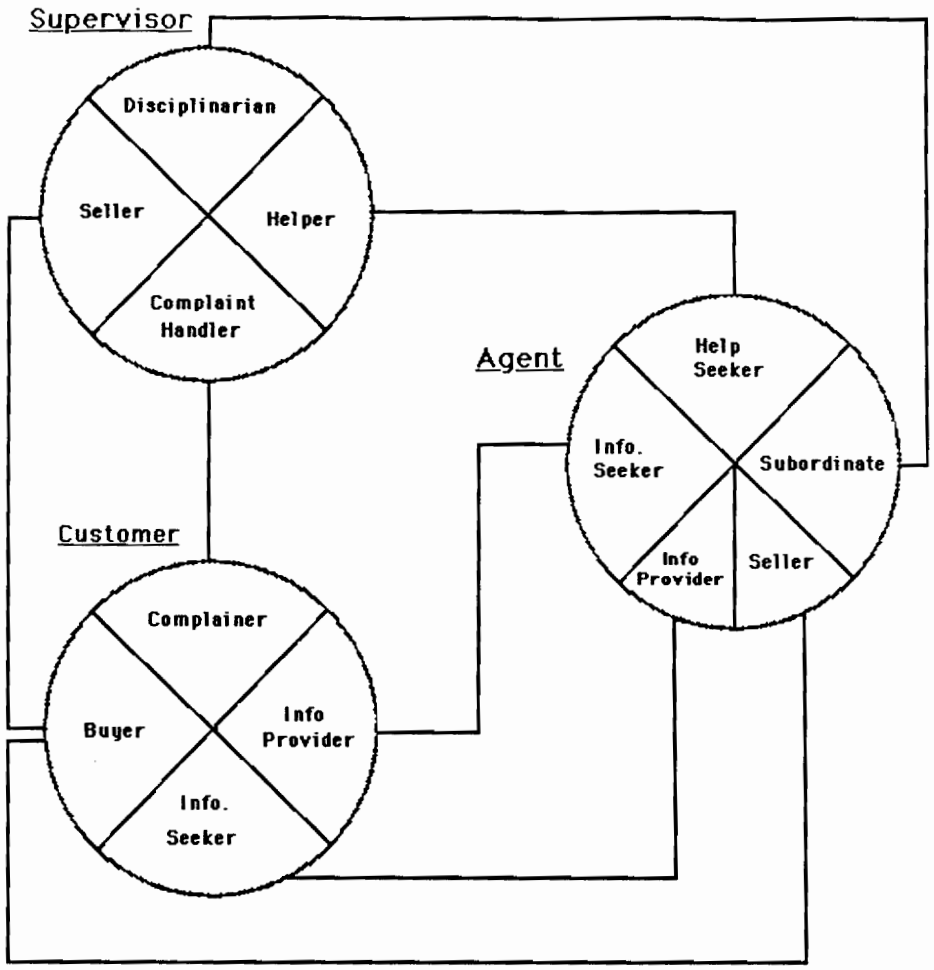


Figure 17. The Customer-Agent-Supervisor Group

Chapter 7. Politeness

We discussed in the previous chapter some issues concerning human behavior in social context. We argued that the "meaning" of an action is not inherent in the action itself, but that it is instead inherited from the function on which the action is based and the social structure within which the action is performed. We thus concluded that in order to understand the behavior that an actor exhibited, one had to be aware of two things: (1) the task the actor was solving and (2) the social structure within which the actor was solving that task.

As we will see in the examples given in this chapter, meta-statements are often used to convey messages pertaining to the relationship that exists between interactants. Statements such as "sir," "please," "thank you," "I'm sorry," etc., are messages that fulfill what we call a *politeness* function.

In this thesis, we are focusing on one particular type of human behavior: the telephone dialogue between an airline reservation agent and a customer. At the end of the previous chapter, we discussed the task aspect of such telephone conversations. We enumerated the various functions that agents and customers fulfill, and we identified the actions performed to fulfill those functions. However, behavior cannot be accurately modeled using only tasks and functions. Certain actions cannot be explained if we limited ourselves only to tasks specific to the functions being fulfilled. If we were to cast all actions into a function specific paradigm, how would we then explain actions such as greeting, thanking, apologizing, bidding farewell, etc? Ostensibly, these actions have nothing at all to do with setting up flights or asking for information. That is, they are devoid of any "functionality" per se, in the sense that they fulfill one of the specific tasks agents and customers strive to fulfill. Moreover, greeting, apologizing, bidding farewell, etc., are actions performed in almost all types of interactions, regardless of the tasks being fulfilled. If we are to analyze behavior only in terms of functions and tasks, how are we then to account for these task independent actions?

An important characteristic of the behavior we are focusing on (viz. the telephone dialogue) is its interactive aspect: we are analyzing behavior that involves people interacting with other people. Of course, not all social behavior is necessarily interactive. Some types of behavior involve one person solving some problem alone, e.g., a man shaving himself in the morning. Often, however, people engage in behavior involving them with other people. The difference between these two types of behavior--i.e., individual behavior versus group behavior--is that whereas in the case of the individual acting by himself all the actions are exclusively oriented toward the specific tasks being fulfilled, in the case of group behavior there are also actions performed to manage the *interaction* between the members of the group. Problems such as "ego clashes" and "personality conflicts" are not minor and cannot be ignored, for a failure to resolve them may very well render the group's task of fulfilling its functions impossible. Actions to avoid and repair such problems are thus vital not only for the group's performance, but also for its survival as a group.

Therefore, in order to model human behavior in groups more accurately, we need to account for more than just task specific actions: we need also to model those actions that address the ego concerns of interactants. Disregarding these concerns would leave much of the behavior exhibited by people totally unexplainable. How would we explain, for instance, the existence of so many different ways of formulating the same request? [SANF87]. Why do we say in some instances "Can you give me your name, please?" while in other instances we say "What's your name?" or "Let's have your name." Why should there be so many different ways for saying the "exact same thing"?

But is it in fact the "exact same thing" that each of these utterances expresses? The first request, for instance--"Can you give me your name, please?"--is rather formal and somewhat overly polite, the sort of thing one might hear from a salesman, for example. "What's your name?" on the other hand, is markedly less formal and more direct. Such a formulation one might say while being introduced to a peer. On the other hand, "Let's have your name" has an unmistakably threatening connotation about it. One might hear such a formulation from the lips of a disgruntled teacher planning to punish the name bearer in question. Moreover, it would probably be highly inappropriate to say, "Let's have your name" while being introduced to another person. It would be too rude, perhaps, or too impolite, and would most probably do little good furthering the initiated relationship. "Can you give me your name, please?" also would sound odd and somewhat inappropriate during an introduction. The formulation is too formal, too businesslike, too polite. In short, a formulation appropriate in one situation is not necessarily fitting for another situation. Each situation calls for certain types of formulations to be used.

This implies, then, that there is much more at play in a situation than the semantics of what is being said and done. A dimension different and independent from semantics must be introduced. Only by including this new dimension will we be able to explain, for instance, why people greet one another when they meet, why they bid farewell to one another when they part, why they apologize, why they thank, why they hedge and hesitate, and why they choose to say things one way and not another. That is, we must not only consider the se-

mantics of the situation to understand the actions performed in that situation, we must also take into consideration the *pragmatics* of that situation. This chapter addresses some of the issues involved in modeling this added dimension of behavior.

In what follows, we will concern ourselves with the following questions: How do people behave when interacting with other people? What types of actions do they perform, and what are the motives that drive them to perform those actions? In the first section, we discuss some general issues related to behavior exhibited during interaction and will argue that interaction behavior is marked by high degree of regularity. In section two, we examine the nature of this regularity and argue that its origins are due to the pragmatics of communication. In section three we discuss the notion of *face,* while in section four the notion of *face wants* is examined. Section five focuses on the impact that differences in status between interactants have on an interaction, and in section six the impact of distance and affection between interactants is examined. In section seven we focus on strategies of politeness that interactants use, and a taxonomy of politeness strategies is given. We end the chapter with an analysis of conversations to illustrate some of the concepts treated in this chapter.

But before begin our discussion, let us first examine a brief conversational excerpt to illustrate the politeness aspect of meta-statements.

The following analysis is taken from an analysis of an entire conversation, given in [BOUZ91]. As that analysis revealed, politeness is indeed a major concern during interaction, and meta-statements are used extensively to effect the various strategies employed. The analysis follows.

Example 7.1:

1 AGENT.
2 [Hello,] Eastern Reservations.
3 Patsy Frisbey, may I help you?
4 CUSTOMER.
5 [Yes.]

6 [Um] I'm planning on a flight leaving April 5th.
7 It's Easter week and I plan on leaving from
8 Roanoke and I'd like to go to L.A. and
9 I was wondering if you could give me some
10 flight information about that.
11 AGENT.
12 That's April 5th?
13 CUSTOMER.
14 Yes [Ma'am.]

The agent begins with a salutation, "hello" (line (2)), and in line (3) offers help. In line (5) the customer answers with a "yes," hedges in line (6), and from line (6) to line (10) extensively uses the strategy of being conventionally indirect (i.e., declaratives, e.g., "I'm planning," "I plan," and "I was wondering..."). In line (12) the agent echoes the departure date given by the customer, uses a contraction, and an impersonalization of H with a point of view distancing in space through the use of "that." The customer shows agreement and gives deference with "ma'am."

As we can see already, meta-statements are indeed used for politeness purposes during conversations. In what follows we discuss some of the basic notions related to the issue of politeness in general.

7.1 On Regularity in Interaction Behavior

Although each person is a unique individual, with his own personality and character, it is nevertheless still true that people's behavior with other people is marked by a high degree of regularity. One has only to casually observe people conversing to see quite clearly how little that is original or singular occurs during interactions. Conversations are almost always started with greetings and terminated with farewells; when someone pays us a compliment we usually say "thank you," or maybe pay them back the compliment; when we inconvenience someone we apologize or ask for forgiveness; we bow or curtsy to a queen; we salute a

general; we exchange a "high five" with a buddy while we kiss a lady's hand. In short, the way we communicate and interact with other people seems to be regulated by a very well defined and almost always obeyed body of behavior rules, and it is clear that when we behave we base our actions and our interpretations of other people's actions on this body of rules and regulations.

The questions we must face now then are: Why is there regularity in human interaction behavior? What is at the source of this regularity? Why do people behave so similarly in similar situations?

One might argue that this regularity comes from the inherent similarity that exists between people. People are alike, and so they act alike. We all greet one another when we meet because we are all basically the same, one might argue. And we all apologize under certain circumstances--when we feel we are inconveniencing someone, for instance--because we all have the same psychological dispositions to do so. But this explanation does not carry us far in answering our inquiry about what is at the source of regularity in human interaction. That we are to a great extent psychologically similar to one another is a fact beyond question. But this by itself is not an answer. All we may conclude from the observation that we are "basically the same" is that regularity in behavior is possible, and that the conditions for regularity exist. Our similarities, by themselves, do not cause any regularity in behavior.

We argue instead that this regularity comes from pragmatic considerations of communication. To communicate with one another we need a common language. To relate our positions and feelings toward other people, we need to have a way of expressing ourselves in such a way that others will understand us. And to understand how others stand with respect to us and what they feel toward us, we must be able to understand what they tell us. We must know, for instance, what to say and how to say what we want to say if we want others to see that we are appreciative of their services and that we feel indebted to them--i.e., we must know how to *thank*. If we have been taught properly, we will say "thank you," or "thank you very much," or "thanks a lot," or perhaps "I thank you." By saying this we hope the other person will understand how we stand with respect to them. That is, they will understand that

what we did by saying "thank you" is thank them, and that when thanking, one is expressing that he stands in a certain position with respect to the person he is thanking. In other words, they will understand us, and we them, because we share a common language, the language of interaction.

The language of interaction, like all other languages, is a means of performing actions and describing states of affairs with accuracy and economy. Take, for example, the act of greeting: what does it mean when an acquaintance of mine and I exchange "hello's" when we come across one another in the street. Let us say that "hello" is all that we say to one another, and that after the exchange each goes his way. What did those "hello's" mean? Semantically speaking, what are they? What did I mean by saying "hello," and what did my acquaintance's "hello" mean to me?

Greeting is one of the most basic actions performed in groups, whether the group is a human society or a pack of baboons. Its function seems to consist of, among other things, signalling recognition, conveying peaceful intentions, and communicating a desire to maintain the existing relationship between the greeters at status quo. All at once, these considerations are addressed with one simple "hello." Imagine the alternative; that is, imagine the non-existence of this interactional language. How would we then go about telling our acquaintance the way we stand with respect to them, without needing to go into an absurdly long explanation, that is, without having to stop and say to our acquaintance: "I know you and I recognize you; I have no evil intent upon you, and I mean to do you no harm; I plan on remaining friendly with you for the moment, and I hope you also plan on remaining friendly with me"? Obviously, such an explanation would be simply too wasteful and would require much effort and exertion. If we did not have this interactional language we would most probably spend the greater part of our waking moments explaining ourselves and clarifying our positions to others.

The argument, then, is that regularity in interactional behavior comes from the existence of a language of interaction. The function of this language is to enable those who use it to convey how they stand with respect to other interactants, and do so efficiently and effectively.

Having argued that interactional behavior is marked by a certain degree of regularity, and having advanced the notion that this regularity is due to the existence of a language of interaction, our next step is to examine the functions fulfilled by the language of interaction. We begin this examination with a general discussion on the nature of interaction, while in the subsequent sections a more detailed presentation of the interaction language functions is given.

7.2 The Nature of Interaction

It goes without saying that man needs the help of others to survive in this world. Consequently, it is essential that he be able to interact with those he is living with, so that he may communicate to them what help he wants from them.

However, man's need to interact with other people cannot be completely argued functionally. That is, it would not be accurate to argue that man interacts with others only because he needs others to fulfill certain functions he is not able or willing to fulfill by himself. There is strong evidence to suggest that interaction is to a certain extent an end in itself and not exclusively a means to an objective end. It was found, for instance, that depriving infants of handling over extended periods of time led to a deterioration of the infants' health, making them vulnerable to various types of diseases [SPIT45]. Berne [BERN64] also argues that intrinsic in the human being is a *hunger for stimulus,* noting that "indeed, solitary confinement is one of the punishments most dreaded even by prisoners hardened to physical brutality..." [BERN64.13] Moreover, he argues that "if the reticular activating system of the brain stem is not sufficiently stimulated, degenerative changes in the nerve cells follow..." [BERN64.14]. The point is, then, that people seek interaction with others to fulfill a basic human need, a need independent of and different from the functional needs that interactions may fulfill.

Interaction is then necessary for fulfilling at least two types of needs: (1) functional needs, in the sense that man cannot fulfill certain functions without interacting with other people, and (2) psychological needs, in the sense that man instinctively needs to interact with others to maintain his psychological health.

But interaction has its hazards, and instead of the help sought, an interactant might suffer injury. Consequently, an interactant finds himself facing a dilemma: on one hand he has a need to communicate with others (for functional and psychological reasons), while at the same time he fears that interacting with others will threaten him and lead him to suffer the evil doings of those who wish him evil. In short, he finds himself pulled by two opposing impulses: the impulse to seek and approach others, and the impulse to avoid and flee from others.

Obviously, it would hardly be in his interest to succumb totally to either one of the two impulses. If he trusted each and every person with whom he came in contact, and considered all men good and saintly, he would very soon find himself victim to some schemer's designs, and it would probably not take very long before he would acquire the notoriety of being "foolish" and "idiotic." In fact, it seems that a minimum degree of guardedness and aloofness are prerequisite for that much sought social prize of "respectability."

Similarly, it would hardly be any better were he to succumb totally to the other impulse, trusting no one and suspecting everybody. He would live his days in paranoid fear and isolation. Instead, it would be to his best interest not to fall to either extreme, but instead to maintain a balance between total trust and total suspicion: he must find a middle ground.

The language of interaction is the means by which man tells others about his fears and his desires, and through which he learns about the fears and the desires of those who come his way.

The implication then is that the communal man--at least one who has not fallen to either of the above mentioned extremes and has instead settled to a "normal" middle--behaves as though bound by an *interactional contract* drawn between him and the members of his society. That is, not only does he act as though he were aware of others' hopes and fears but also as

though others were aware of his hopes and fears, and as though they were aware that he is aware of their hopes and fears. Only with this mutual "recognition" is interaction possible. If he believed that others did not have hopes similar to his-- i.e., they did not want to approach other people--and if he believed that they did not have the same fears he has, then how can he interact with them?

7.3 The Notion of Face

The principle of mutual recognition implies that when interacting an actor should be aware (at least theoretically) of two things: (1) the power that he has over others and (2) the power that others have over him. The principle does not, of course, imply that the power an actor perceives himself to possess and the power he perceives others to possess over him are necessarily of equal strength. All that may be inferred from the principle is that when interacting, a person is aware of both his strengths as well as his vulnerabilities, i.e., he is to a certain extent aware that his actions will affect others, and that the actions of others will affect him. Only by recognizing the *consequentiality* of his actions and that of those with whom he interacts is he be able to truly interact with others.

With the recognition of the consequentiality of his actions, and with the awareness that others are similar to him, an actor is (again, at least theoretically) able to infer what actions he should pursue and what actions he should avoid. Those actions he "knows" will please others (since others are like him) he should seek to perform (if the actions are not harmful to him), while those actions he "knows" will harm others he should avoid.

The term we shall use to refer to the sensitivities incurred as a consequence of expectations about behavior toward a self is *face* [GOFF55]. We will say then that each interactant will come to an interaction with a face, and that each will have *face wants and sensitivities.* An interactant who is comfortable with the interaction situation and is satisfied with the way

others are treating him we will say is *in face.* In the opposite situation where he feels inadequate in the situation or feels that others are not treating him according to his expectations--whether they are not fulfilling the minimum obligations claimed by the interactant or whether they are exceeding them--we will say that he is *out of face* [GOFF67].

The face metaphor is quite appropriate and captures the very idea of vulnerability in interaction. It is through our faces that we usually communicate: with our eyes we see others, with our ears we hear others, and with our mouths we speak to others. And whether we intend to or not, we express many of our feelings through our faces. Our face is, as it were, our window to the world and the front with which we *face* it. Consequently, it is only natural that those who seek to harm us should do so through our face. When we come to an interaction we come with a face, a face that we lay bare for all present to see. We may be fortunate and have our face accepted, or we may be unfortunate and see our face rejected. With acceptance we *radiate* pride; with rejection we *blush* and hide our face away--thus the theatrical gesture of hiding one's face with one's hand when ashamed.

An important face related emotion people experience is the emotion of *embarrassment,* that feeling of inadequacy usually accompannied by physical manifestations showing restlessness and discomfort, e.g., blushing or fidgeting [GOFF67]. Embarrassment seems to have at least two sources: one psychological and the other structural.

Psychological embarrassment results from the existence of an egocentric view of the events taking place, i.e., the embarrassed interactant considers himself the center of the interaction situation. Consequently, the interactant will have expectations about face that will most certainly not be recognized by others, resulting in the interactant in question feeling embarrassed, e.g., the situation where something one says is ignored [GOFF67].

Structural embarrassment, on the other hand, results from the interactant finding himself in a situation where he must face two audiences each of which expects him to fulfill a role not entirely compatible with the role the other audience expects [GOFF67].

People do not, of course, all react to embarrassing situations in the same manner. Some will lose total control, and even perhaps take flight, whereas others will react as though

not embarrassed at all. The ability to keep one's composure while embarrassed is called *poise.*

Embarrassment is not exclusively an ego emotion. That is, one may feel embarrassment because one feels his situation awkward, but one may also, as a consequence of the mutual recognition principle, feel embarrassment for others. One who does not "feel for others" is said to have a heart of stone and is little admired for his heartlessness. Not only is one expected to protect his face, he is also expected to protect the face of others. And this he is expected to do not simply for contractual reasons--i.e., protect me and I'll protect you--but also out of a feeling of empathy and identification due to the acceptance of another as a similar self.

Consequently, on many occasions an interactant will find himself facing yet another dilemma: the one pitting his desires and impulses to protect and promote his face against his desires and impulses to protect and promote the face of others. On the one hand we must protect ourselve and fight for our "dignity," if fight we must, but we must also try to protect the "dignity" of others and rarely do we find enjoyment in the defacement of others. Thus, we sometimes give up a little of our face so that we may save the face of another, and sometimes others give up a little of their face so that ours may be saved. Without this giving and taking, and the small and big sacrifices that we make, interaction would not be possible.

We proceed in the next section to a detailed analysis of the various types of face wants and sensitivities claimed by interactants.

7.4 Face Wants and Sensitivities

In section 7.2, we identified two types of conflicting impulses that characterize human interaction behavior: the impulse to seek others out, and the impulse to avoid others. We argued

that the seeking impulse comes from an instinctive psychological need that man has to be among other people, whereas the impulse to avoid others, we argued, comes from fears and suspicions that man has about the intentions of other people. This dichotomy of concerns translates in the terminology used in [BROW87] to the distinction between two types of faces: *positive face* and *negative face.* Positive face refers to "the positive consistent self-image or 'personality (crucially including the desire that this self-image be appreciated and approved of) claimed by interactants," while negative face refers to "the basic claim to territories, personal preserves, rights to non-distraction--i.e. to freedom of action and freedom from imposition" [BROW87.61]. An action that threatens either the positive or the negative face of an interactant they call a *face threatening act* (FTA). In addition to contextual FTA's, they argue that there are acts that are *intrinsically* threatening, in the sense that "by their nature [they] run contrary to the face wants of the addressee and/or of the speaker" [BROW87.65].

FTA's are not performed only on others; an interactant may also perform an FTA on himself, thus imposing on one or both of his own faces, e.g., hitting himself on the head or calling himself "stupid" for having forgotten to do something. In addition, an FTA may simultaneously be threatening not only to one face or to one interactant, but also to the faces of other or all other interactants simultaneously. For example, a man breaking down in tears is threatening not only his own positive face but also the negative face of those around him, since he is in a very real sense forcing them to console him or inquire about his state, thus imposing on their freedom to choose what to do.

In what follows we present the classification of FTA's suggested in [BROW87.65]. FTA's are here classified in terms of the types of faces they threaten, i.e., whether they threaten positive or negative faces, and whether they threaten the actor (i.e., speaker S), or another interactant (i.e., Hearer H). After the classification, we give the internal presentation we are using in our system.

H's Negative Face

The FTA's in this category are impositions on hearers' freedom of action and choice. Among the FTA's identified in this category are the following:

- *1 - Expressions where S indicates that he likes or would like something of H's.*
- *2 - Expressions of strong (negative) emotions.*
- *3 - Giving advice or suggestions,* thus pressuring the hearer to follow a certain course of actions.
- *4 - Making offers or promises.*
- *5 - Warning, daring, making threats or giving orders.*
- *6 - Requesting, and thus pressuring H to do something S wants done.*
- *7 - Reminding, and thus pressuring H to remember to do something.*

H's Positive Face

The FTA's in this category are damaging to the hearer's positive face.

- *1 - Disagreeing or contradicting, and thus rejecting an opinion held by H.*
- *2 - Showing blatant non-cooperation, and thus not sharing H's goals to accomplish task.*
- *3 - Raising of controversial topics.*
- *4 - Bringing of good news about self or bringing bad news about H.*
- *5 - Showing irreverence (mention of taboo topics).*
- *6 - Criticising H.*
- *7 - Complaining about H.*
- *8 - Accusing H.*
- *9 - Reprimanding H.*
- *10 - Expressing contempt to H.*
- *11 - Ridiculing H.*
- *12 - Insulting H.*

S's Negative Face

The following FTA's threaten the speaker's own negative face. The first and second FTA's are threatening because they constrict the speaker's future freedom. FTA's three to six, on the other hand, are impositions because of the effort itself made to perform them.

- *1 - Accepting H's offers, and thus incurring a debt.*
- *2 - Making unwilling promises or offers to H.*
- *3 - Giving excuses.*
- *4 - Responding to H's faux pas.*
- *5 - Expressing thanks.*
- *6 - Accepting H's thanks or apologies.*

S's Positive Face

The following FTA's threaten the speaker's own positive face:

- *1 - Accepting a compliment,* and thus possibly appearing to be conceited.
- *2 - Apologizing.*
- *3 - Non-control of emotion.*
- *4 - Confessing to guilt or responsibility.*
- *5 - Self humiliating action or self contradiction.*

Internally, a face threatening action is modeled with the following frame:

```
(FTA
     (TAG        ?tag)
     (TYPE       ?type)
     (UTTERANCE  ?utterance)
     (FTA        ?fta))
```

where, TYPE indicates what type of face the FTA is threatening (i.e., H-NEG/H-POS/S-NEG/S-POS); UTTERANCE refers to the utterance in which the FTA is formulated; and FTA refers to the specific FTA performed (e.g., FTA #2 of type S-NEG). TAG simply contains a tag uniquely identifying the particular FTA in question.

Face sensitivities and wants are not, however, completely measured in terms of the intrinsic imposition of FTA's. In addition to intrinsic impositions, there are at least two contextual factors that come into play: the relative status differences between the interactants, i.e., which of the interactants is at a position of power over the other, and the emotional distance between the interactants, i.e., whether the interactants are emotionally close (they like one another) or emotionally distant (they do not like one another). We treat each of these dimensions in the following two sections.

7.5 Status and Power

Each interactant comes to an interaction with a certain amount of power to influence the course of events during that interaction. However, not all interactants have equivalent amounts of power. Instead, some will possess greater power than others to direct what will take place and occur during the interaction. Those who have the greatest power in the group are called *leaders,* and they are said to occupy the highest status in the group.

A group is characterized by a structure of statuses, where each group member occupies a certain location relative to other members of that group.[35] Thus a member may be subordinate to some members, equal to certain other members, while superior to yet other group members. With each of these three types of group members, he will have to adopt a certain pattern of behavior such that this difference in status is recognized. For instance, he might talk condescendingly to his "inferiors," informally with his peers, whereas when speaking to his "superiors" he might perhaps choose to employ a more respectful tone of voice.

Power to control is in fact based on the contractual nature of interaction. Each of the interactants involved in an interaction has a certain degree of freedom to terminate his involvment in that interaction and thus leave the group. However, not all of the interactants have the same degree of freedom to terminate their participation. Some will be more reluctant than others to leave the group; they may feel a strong need to remain within the group, because of functional or psychological reasons (or both). Consequently not all members will have the same power. Those who will lose the least from the termination of the interaction will obviously have greater control, since they hold in their hands a powerful weapon, that with which they may eliminate the group as a group. Those who will lose the most from the termination, however, (before they have fulfilled their goals, for example) will have less power, since they are at the mercy of those who may want to effect it. Thus a salesman is at a con-

---

[35] The notion of status in this chapter is to be distinguished from the one presented in chapter six. Status here refers to a relative degree of power.

siderably lower status with his potential buyer before the latter makes a purchase, since the buyer usually will not suffer considerably from breaking his interaction with the salesman before buying (i.e., there are other salesmen and other stores), whereas the salesman will lose a commission, and customers are not always easy to attract. Once the sale is made, however, the customer will lose much of his high status, and in fact may come very close to being equal in status to the salesman. A situation where the interactants are peers is a chat between two friends, where neither of the two friends is particularly anxious to maintain the chat nor particularly anxious to terminate it. In such a situation the interactants are on equal footing, have equal power to either terminate or sustain the interaction, and thus occupy the same status.

In short, we distinguish between three types of statuses that any interactant may occupy relative to another interactant: *superior, inferior,* and *equal.*

Status differences play an important role in determining face wants and sensitivities. Thus a superior will have a more sensitive face than a peer or an equal regarding formality of the interaction. A peer, on the other hand, has sensitivities for being treated as a peer and neither as a superior nor as an inferior. While an inferior will expect to be treated as such, i.e., as an inferior, and will not take offense if he is treated as he expects.[36]

Internally, the following parameter is used:

```
(REL-STAT
        (CONV1 ?conv1)
        (CONV2 ?conv2)
        (DIFF  ?diff))
```

with DIFF containing the relative status differential existing between converser 1 (CONV1) and converser 2 (CONV2). DIFF has the value HIGHER, LOWER, EQUAL if ?conv1 has higher, lower, or equal status, respectively, relative to ?conv2.

---

[36] Whereas he might take offense if he were treated with exceeding respectability, which he might perceive as mockery.

7.6 Distance and Affection

In addition to the intrinsic impositions of face threatening acts and the differences in the statuses of interactants, face sensitivities are also influenced by a third important factor: the affection that exists between the interactants. The imposition of a friend we will accept, but the imposition of a stranger we will not. Just as we will accept the imposition of someone who has power over us, we will also accept the imposition of someone for whom we hold an affection. And just as we will be reluctant to accept the imposition of an "inferior," so we will be reluctant to accept the imposition of someone for whom we hold little or no affection.

We identify three factors that may influence the degree of affection that interactants feel for one another. First is the *solidarity* of the interactants. We distinguish between three types of solidarity:

(1) *Structural solidarity* is said to exist between two interactants if the interactants belong to the same social group or structure, e.g., the solidarity of belonging to the same family, race, nationality, gender, religion, labor union, team, etc. Hostility may also result from belonging to social groups that are antagonistic to one another.

(2) *Ideological solidarity* is said to exist between people sharing the same opinions or beliefs. On the other hand, antagonism may result from holding opposing opinions.

(3) *Situational solidarity* is said to be in effect between people who are facing or have faced similar situations or problems, e.g., "misery loves company," and "we are all in the same boat together."

According excessive importance to solidarity, at the expense to other considerations, produces what we call "prejudiced" behavior, e.g., jingoism, race, age, or gender discrimination, etc. On the other hand, behavior according little importance to solidarity is called "treacherous," and the actor a "traitor."

Internally, the following parameter is used:

```
(SOLIDARITY
          (DIR  ?dir)
          (TYPE ?type))
```

with DIR taking the value POS when the interactants are in a situation of solidarity, NEG when the interactants oppose one another, and NIL in the case where no solidarity exists. TYPE takes the value STRUCT, IDEOL, SITUAT, respectively for structural, ideological, and situational solidarity.

The second source of affection is *personality.* We may perhaps not feel solidarity with another person, but because of their personality we may feel a certain affection for them. We distinguish between two types of sources for personality affection: *character* and *performance.*

We may like a person for a certain trait in their character, e.g., for the dignity of their conduct, for the consideration they display for others, for their self-possession, etc.[37]

In addition, we may like people for what they have accomplished, or what they are able to accomplish, i.e., for their performance. This affection is usually displayed by a feeling of *respect.*

Internally, the following parameter is used:

```
(PERSONALITY
        (DIR  ?dir)
        (TYPE ?type))
```

with DIR taking the value POS when the personality of the other converser (i.e., the partner) reduces the affection distance, NEG if it increases it, and NIL if the personality is not considered at all. TYPE takes the value CHAR, if the interactant's character is the aspect in focus, and PERF, if his performance is the focus.

The third and last source of affection is *familiarity.* We distinguish between two factors that may influence familiarity. First is *time,* i.e., how long the interactants have known one

---

[37] This dimension is to be distinguished from the ideological dimension of solidarity: we may hate someone for his opinions, but we may still respect his dignity or his self-sacrifice.

another. And second is *experience,* i.e., how much do the interactants know about one another. Familiarity cannot be measured exclusively along one of the two dimensions, i.e., people who have known one another for an extended period of time are not necessarily more familiar with one another than other people who have known one another for a shorter period of time. It may be that in one encounter two interactants have learned a great deal about one another, whereas in many previous encounters little was revealed.

Internally, the following parameter is used:

```
(FAMILIARITY
        (DIR  ?dir)
        (TYPE ?type))
```

with DIR taking the value POS when the interactant's familiarity reduces the affection distance, NEG if it increases it, and NIL if the familiarity is not considered at all. TYPE takes the value TIME, if the familiarity is based on time, and EXP, if it is based on experience.

7.7 Politeness Strategies

We have argued up to this point that there are at least three important factors that influence the face wants and sensitivities claimed by interactants. First are basic needs felt by interactants, needs that are instinctively manifested. Second are the status and power differentials that exist between interactants. And third are the affectional distances existing between the interactants. We have argued that the face impact of an action--i.e., its imposition--is a function of these three factors, and that the impact of an action will be reduced or accentuated depending on how the interactants stand with respect to one another in terms of status and affection.

As we argued in section 7.3, it is in the best interest of everyone involved in an interaction to protect not only their own respective faces but also the faces of the other interactants

involved in the interaction. Certain actions, as we saw in section 7.4, are intrinsically threatening to the face of a hearer. An action especially performed to reduce the imposition caused by an FTA (face threatening act) is called *a politeness strategy.*

According to Brown and Levinson [BROW87], there are two types of politeness strategies: strategies that are said to be *on-record* and strategies said to be *off-record.* An actor is said to perform an action on-record if the action he performs is not ambigous. "Can you lend me five bucks?" is a clear and unambigous request to borrow money. An action is said to be off-record if its meaning is ambiguous, or at least such that more than one interpretation is possible. A speaker may, for instance, "hint" to a hearer to lend him some money with, "Gee, I forgot to go to the bank," but the "request" (if there was one intended) is said to be off-record, since the request was not explicitly stated. Should the hearer respond to the above utterance by saying, "I can't lend you any money," the original speaker may legitimately claim that he never requested to borrow any money, i.e., he never went on-record with the request.

Politeness strategies that reduce the imposition of an FTA with an on-record action recognizing the imposition are said to be on-record. The "can you" formulation in the above loan request is an on record politeness strategy because the imposition is performed on record and recognized on-record. The "hint," on the other hand, is an off-record strategy, because the imposition is *not* performed on-record.

On-record Politeness Strategies

The following are politeness strategies performed by the speaker with explicit recognition of the imposition being made by S on H.

Negative Politeness Strategies

The following strategies of politeness are directed at reducing the face impact of an action on the negative face of another interactant. The actor in this case perceives that he is encroaching on the freedom of another interactant and that it is incumbent upon him to signal his recognition of the other's claim to freedom from imposition.

1. Be Conventionally Indirect

(1) *Inquiry about Ability.e.g.*
(2) *Generalize.*
(3) *Use Exclamations.*

2. Hedges

(1) *Weakening Hedges.*
(2) *Strengthening Hedges.*

3. Be Pessimistic

(1) *Use of Subjunctive.*
(2) *Use of Negative (with Tag).*
(3) *Use of Remote Possibility Markers.*

4. Thanking

5. Give Deference

(1) *Addressee Honorifics.*
(2) *Referent Honorifics.*

6. Apologize

(1) *Admit the Impingement.*
(2) *Indicate Reluctance.*
(3) *Give Overwhelming Reasons.*
(4) *Beg Forgiveness.*

7. Impersonalize S & H

(1) *Performatives.*
(2) *Imperatives.*
(3) *Impersonal Verbs.*
(4) *Replacement of the Pronouns 'I' & 'You' by indefinites.*

(5) *Reference Terms like 'I' Avoidance.*
(6) *Point-of-view Distancing:*
    (a) *Distancing in Time.*
    (b) *Distancing in Space.*

8. State the FTA as a General Rule

9. Nominalize

10. Go On-record as Incurring a Debt

11. Opening and Closing Salutations

12. Use the Token "Please"

(1) *The post-please token.*
(2) *The pre-please token.*

Positive Politeness Strategies

Positive politeness strategies are actions which have the impact of redressing the positive face of H. The impact $i(P)$ of a positive politeness strategy P contains at least the element r_H^+. Brown and Levinson identify the following strategies:

1. Notice, Attend to H's (interests, wants, needs, goods).

2. Exaggerate (interest, approval, sympathy with H).

3. Intensify Interest in H:
(1) *Using the vivid present.*
(2) *Using directly quoted speech.*
(3) *Using tag questions.*
(4) *Overstatement.*

4. Use In-Group Identity Markers:
(1) *Address forms:* e.g., terms such as *man, brother, honey, pal,* etc.
(2) *Use of in-Group language or dialect.*
(3) *Use jargon or slang.*
(4) *Contraction and ellipsis.*

5. Seek Agreement:
(1) *Raising of safe topics.*
(2) *Use repetition.*

6. Avoid Disagreement:

(1) *Token Agreement.*
(2) *Pseudo-agreement.*
(3) *"White" lies.*
(4) *Hedging opinions.*

7. Presuppose/Raise/Assert Common Ground:

(1) *Gossip, small talk.*
(2) *Point-of-view operations:* S's
    (a) *Personal-centre switch (S to H).*
    (b) *Time switch.*
    (c) *Place switch.*
(3) *Presuppostion Manipulations:*
    (a) *Presuppose knowledge of H's wants and attitudes.*
    (b) *Presuppose H's values are the same as S's values.*
    (c) *Presupposes familiarity in S-H relationship.*
    (d) *Presuppose H's knowledge.*

8. Joke.

9. Assert or Presuppose S's Knowledge of and Concern for H's Wants.

10. Offer, Promise.

11. Be Optimistic.

12. Include both S and H in the Activity.

13. Ask for (or Give) Reasons.

14. Assume or Assert Reciprocity.

15. Give Gifts (goods, sympathy, understanding).

Off-record Negative Politeness Strategies

The following are politeness strategies performed by the speaker without explicit performance of the imposition [BROW87].

1. Give Hints

2. Give Association Rules

3. Presuppose

4. Understate

5. Overstate

6. Use Tautologies

7. Use Contradictions

8. Be Ironic

9. Use Metaphors

10. Use Rhetorical Questions

11. Be Ambiguous

12. Be Vague

13. Over-Generalize

14. Displace H

15. Be Incomplete, Use Ellipsis

Internally, a politeness strategy is modeled with the following frame:

```
(POL-STRAT
        (TAG        ?tag)
        (TYPE       ?type)
        (UTTERANCE  ?utterance)
        (STRATEGY   ?strategy))
```

with UTTERANCE being the utterance where the strategy is employed, TYPE indicating what type of politeness strategy is used (i.e., ON-REC-NEG/ON-REC-POS/OFF-REC) and STRATEGY referring to the particular strategy used (e.g., OFF-REC #15). TAG simply contains a tag uniquely identifying the particular politeness strategy in question.

7.8 Multiple Impact Analysis

As we have already said, an action usually impacts more than one face at a time. A strategy performed to redress an imposition on one face may at the same time be damaging to another face. Tables 1 and 2 below give the impact of the on-record strategies identified. S^+ and S^- refer to the positive and negative faces of speaker S, respectively, while H^+ and H^- refer to the positive and negative faces of hearer H, respectively. *D* indicates that a strategy may be damaging to the designated face, while an R means that the strategy may be redressive to that face.

For example, positive politeness strategy #10 (offering or promising) has possibly three simultaneous impacts: (1) a redressive impact on the positive face of the hearer, (2) a damaging impact on the negative face of the hearer (see FTA #3, in FTA's on H's negative face), and (3) a damaging impact on the negative face of the speaker (FTA #2, in FTA's on S's negative face).

Table 1. Face Impact of On-Record Negative Politeness Strategies

Strategy	S+	S-	H+	H-
N-1	D	nil	R	R
N-2	D	nil	R	R
N-3	D	nil	R	R
N-4	D	nil	R	R
N-5	D	nil	D	R
N-6	D	nil	R	R
N-7	nil	nil	D	R
N-8	nil	nil	D	R
N-9	nil	nil	D	R
N-10	D	D	R	R
N-11	nil	nil	R	R
N-12	nil	nil	R	R

Table 2. Face Impact of On-Record Positive Politeness Strategies

Strategy	S+	S-	H+	H-
N+1	nil	nil	R	D
N+2	nil	nil	R	nil
N+3	nil	nil	R	nil
N+4	nil	nil	R	D
N+5	nil	D	R	nil
N+6	nil	D	R	nil
N+7	D	nil	R	nil
N+8	D	nil	R	nil
N+9	nil	nil	R	D
N+10	nil	D	R	D
N+11	nil	nil	R	D
N+12	D	nil	R	D
N+13	D	D	R	D
N+14	nil	nil	R	nil
N+15	nil	nil	R	D

7.9 Politeness Actions

We distinguish between two types of politeness actions: threatening actions and redressing actions. An action is said to be threatening if its execution threatens the face of one of the interactants. An utterance is thus said to be threatening if its promulgation results in an FTA. Similarly, an action is said to be redressing if its execution redresses the face of one of the interactants. An utterance is said to be redressing if its promulgation results in a politeness strategy. In our system, a meta-statement is mapped along the politeness dimension to either a threatening action--i.e., an FTA--or a redressive action--i.e., a politeness strategy.

We argued in the previous section that an FTA may threaten more than one face, and that a politeness strategy may redress more than one face. In the case of a threatening

action, the action may be threatening to not just one face but to all four faces: S-NEG, S-POS, H-NEG, and H-POS; likewise with politeness strategies.

Thus, the two types of politeness actions identified in our system have the following internal representations:

```
(POL-ACT-FTAS
        (LIST-FTAS   ?list-ftas))
```

while a politeness strategy action has the following frame:

```
(POL-ACT-STRATS
        (LIST-STRATS ?list-strats))
```

with LIST-FTAS being a list of FTA frames of length at least one (an FTA must threaten at least one face) and at most four (there are only four faces to be threatened); and LIST-STRATS being a list of POL-STRAT frames also of minimum length of one and maximum length of four.

7.10 Analysis of Transcripts

We illustrate in this section some of the concepts introduced in the previous sections. Our discussion begins with an analysis of the status dimension of the conversations we are studying. We follow this with an analysis of the affect dimension. The third section contains an analysis of a conversation in terms of the politeness strategies used.

The Status Dimension

Our dialogue situation defines rather specific status positions for each interactant involved in the conversation. The situation involves a customer and a travel agent, and as we argued in the previous chapter, the relationship existing between customers and agents involves three major roles for each of the two interactants. First is the role of information seeker played by the customer (complemented by the agent's information provider role); second is the role of information provider played by the customer (complemented by the agent's information seeker role); and third is the customer's buyer role (complemented by the seller role the agent plays). Each of these roles stands with a certain status relative to its complementary role. The role of information seeker played by the customer puts the customer in a status position lower than the status position played by the agent. It is the agent who has the information, and thus the agent has power to either give it or refuse giving it. This power is of course not absolute, i.e., the agent is also a seller, and before he can sell he must furnish the customer with information on what he is selling. Nevertheless, we argue that possessing information that others want is in itself advantageous to the person possessing the sought information, and thus raises his relative status.

Let us examine the following excerpt to illustrate the influence of status on the strategies of interaction used.

Example 7.2: [38]

AGENT.
[Good morning, this is United Airlines
Reservations, Mrs. McGillicuddy.]
CUSTOMER.
[Yes,]
[I need some information,] [please.]
I would like to [um] ask, get some information about
taking a flight leaving April 5th from Roanoke to
Los Angeles.
AGENT.
[Okay,]

[38] Conversation Phone.04, lines [1-27].

12 [and] returning?
13 CUSTOMER.
14 [Um] a week later.
15 The following week, [I believe] that would be the
16 12th.
17 AGENT.
18 [The 12th.]
19 [And] how many are traveling?
20 CUSTOMER.
21 Just 1.
22 AGENT.
23 [Okay,] I'll be glad to check and see what's
24 available.
25 Are you able to depart anytime on those days?
26 CUSTOMER.
27 [Ah] yes.

We notice in lines [5,9] and lines [14,15] the customer's lower status position: in line (6) he uses two negative polietness strategies: (1) conventional indirectness--"I need some information"--and the post-tag "please," and in line (7) he uses "would like" and a hedge particle--"um." The agent reponds with minimal effort in lines [11,12], while the customer resumes with a hedge in line (14)--"um"--and another hedge in line (15)--"I believe."

On the other hand, starting from line (23), the roles are reversed, and it is the agent who becomes the information seeker. With this reversal we notice also a shift in statuses, reflected in the politeness strategies used. The agent begins with a statement that she does not consider the request the customer made an imposition--"Okay, I'll be glad to check and see what's available"--a positive politeness strategy, while in line (25) she uses two additional negative politeness strategies: (1) the strategy of being conventionally indirect--"are you able"-- and (2) the strategy of referent honorofics with the use of "depart" instead of "leave" or "go." The customer, on the other hand, answers with minimal politeness, using only a hedge--"ah"--in line 27, before answering in the affirmative.

The buyer/seller relationship in our situation is even more clear cut. The buyer clearly has the higher status. This status difference is most pronounced when a move is made by the agent to effect a sell, a move usually initiated with the agent's asking for the customer's name. The following two instances in conversation (phone.03) lines [299-300] and [359-360] illustrate this:

Example 7.3:

1 AGENT.
2 [Okay,] [and] what is your name, [please?]

Example 7.4:

1 AGENT.
2 [Okay,] may I have your last name, [sir?]

In the first example the agent uses two negative politeness strategies. First a hedge with "okay," and then the post-please tag "please." In the second example, she uses two strategies also, starting with an "okay," and then an addressee honorific, "sir."

The following example illustrates even more clearly the "inferior" status the agent inherits from her seller role. Five different politeness strategies are used in fourteen instances:

Example 7.5: [39]

169 AGENT.
170 [And] [if you like] we can go ahead and [just]
171 tentatively hold it for you.
172 We can hold it until Saturday ['cause] that would
173 be the last day and [if you like] we can mail them
174 out on that date, we can put them on a credit card
175 or you can pay by cash or check.

In line (170) the agent uses three strategies: (1) the hedge "if you like," (2) negative politeness strategy #7.5, impersonalize S & H with term 'I' avoidance, using "we" instead, and (3) negative politeness strategy #4, with the particle "just." In line 172, the "we" is again used, along with positive politeness strategy #4.4, by using the contraction "'cause." In line 173 the hedge "if you like" is used, and once more "we" instead of "I" is employed. In 174, "we" is again used instead of "I." In addition, negative politeness strategy #1 is used throughout with the auxiliary "can" used on five separate instances and "would" used once in line (172).

---

[39] Phone.04 lines [169-175].

The Affect Dimension

We turn now to the second dimension of analysis, the *affect* dimension.

We argued in section 7.6 for the existence of three axes of affection: solidarity, personality, and familiarity. The solidarity axis, we proposed, has three sub-axes: an ideological axis, a structural axis, and a situational axis. Ideological solidarity exists between people who share the same beliefs or opinions; structural solidarity between people who belong to the same social structure; while situational solidarity exists between people who are facing (or have faced) some common problem raised by the situation they are facing. We illustrate each of these dimensions briefly with the following examples.

In our conversations, the ideological dimension of solidarity was rarely manifested. This is to be expected since giving or asking for another's beliefs is usually not done in transactions such as the ones we are studying. In one instance, however, the agent, discovering that the customer is graduating in "communications," informed the latter that she thought that communications is "such a good field to be in," while the customer answered with "I know."[40] With this exchange, the affectional distance between the interactants is somewhat reduced, the agent and the customer "discovering" that they share the same opinion about the value of graduating in communications.

Structural solidarity was also rarely manifested. One instance was the agent revealing to the customer, who was going to look for a job in L.A., that she, the agent, had lived in L.A. The solidarity in this case is manifested through the agent and customer's being (potentially) both "people who lived in L.A."

Situational solidarity, on the other hand, was more prevalent. A pronounced instance was the long conversational fragment in conversation phone.03, stretching from line (466) to line (619). Learning that the customer feared flying the DC10 airplane (line 466) and later that it was the customer's first time flying (line 496), the agent tells the customer about her expe-

---

[40] Conversation Roleplay.08, lines [63-73].

riences while she was trying to learn how to drive (line 564), drawing a parallel between the customer's fears of flying and her own fears of learning how to drive. Obviously, her actions were motivated by situational solidarity, i.e., she identified with the customer's fears, having herself faced a situation where she also was anxious and nervous.

The second dimension of affect is personality. We identified two personality axes: character and performance. The influence of character was difficult to evaluate in our conversations. There was at least one instance, however, where character seemed to have an impact. In the above example of the agent trying to reassure her customer that the latter had nothing to fear from flying an airplane, the agent explicitly informed the customer of her appreciation for the latter's openness in informing her that it was his first time flying (line 496). That is, she felt affection for the customer because of the latter's openness of character. The example of the agent's praising the customer's field of study is one instance where performance was a factor in affection distancing. That is, discovering that the customer was graduating in communications, the agent, feeling that communications "is such a good field to be in," reduces her affection distance from the customer.

Familiarity is the third affectional dimension. We argued that the age of a relationship has an impact on the affection existing between interactants. This is not to say that the longer one knows another person the better one will like them. Obviously, there are instances where the more one knows about another person the less they will like them. Nonetheless, we argue that the mere fact of knowing someone over an extended period of time reduces the affectional distance. For our situations, thus, the longer the conversation, the more the time familiarity factor reduces the distance.

The second factor influencing familiarity is experience, i.e, the degree to which the interactants know one another. Time alone is not a sufficient indicator of the degree of familiarity. In conversations (e.g., phone.05, phone.06) nothing aside from immediate concerns about flight scheduling and related issues is discussed. On the other hand, in other conversations (e.g., phone.03, phone.04) both the agent and the customer reveal something about themselves. In conversation phone.03 the customer reveals to the agent that it is his first time

flying (lines 547-548); the agent in turn tells him a story about her own fears when she was learning how to drive (lines 564-582); while in conversation phone.04 (lines 182-185) the customer reveals that he is handicapped and needs a wheelchair. In each of these instances, the familiarity is increased.

A point we must stress is that we are not implying either an increase or a decrease in affection as a result of the increase or decrease in familiarity. That is, it is not always the case that the more one learns about another, the more one will like him. It may very well be that the converse is true.

Analysis of a Conversation

Conversation phone.01 was examined in terms of the politeness strategies and FTA's employed by the interactants involved in that conversation, and the results obtained from that analysis are given below in Table 3. The analysis reveals that, indeed, politeness is a major concern during interaction, and that meta-statements are used extensively to effect the various strategies employed. The following notation is used. An on-record strategy is denoted by an *N,* while an off-record strategy by *F.* A positive strategy is denoted by a "+" appended to the N or the F, while the digits that follow the sign designate the strategy number. Thus, on-record positive strategy *joke* is denoted by N+8.

Analysis of the transcription gives us the following results:

Table 3. Politeness Strategies Used in Example

Strategy	Customer	Agent
N+1	1	2
N+4.4	2	0
N+5	18	9
N+5.2	6	1
N+6.2	2	0
N+6.4	0	1
N+7.2.a	1	0
N+10	0	1
N+12	1	0
N+15	1	1
N-1	3	2
N-2	19	12
N-3.1	13	7
N-5	1	0
N-5.2	1	2
N-6	0	1
N-7.6	0	1
N-7.6.B	1	0
N-11	0	2
F-15	7	2

The totals follow: the customer used 8 different positive politeness strategies for a total of 32 positive politeness strategy applications, while the agent used 6 different positive politeness strategies for a total of 15 application; the customer used 6 different negative politeness strategies for a total of 38 negative politeness strategy applications, while the agent used 7 different negative politeness strategies for a total of 27 applications. As Table 3 shows, the customer used 7 instances of off-record politeness strategies, while the agent used 2. In all, 20 different strategies were used, for a total of 121 instances of politeness strategy applications. A surprising result is that the customer generated 63.64 percent (77) of these applications, while the agent generated 36.36 percent (44). One would have expected the agent to be far more polite, but in this instance, the customer was decidedly the one who used politeness strategies more often.

7.11 Summary

In this chapter, we have discussed the topic of politeness. We have argued that there exists a language of interaction that people use when interacting with one another, and that this language addresses both the fears and the desires inherent in human beings. Each person has sensitivities that pertain to basic wants intrinsic in any human being: (1) a want to be free from the impositions of others, and (2) a want to be accepted by others. The sensitivities ensuing from the first want constitute what is called the interactant's negative face, while those ensuing from the second constitute the interactant's positive face.

The argument we followed is based on four assumptions: (1) each interactant comes to an interaction with face sensitivities; (2) each interactant is aware that the other interactants also have face sensitivities of their own; (3) each interactant is aware that he has the power to violate the face sensitivities of the other interactants; and (4) each interactant is aware that the other interactants have the power to violate his own face sensitivities. Consequently, given this mutual vulnerability, we argued that it is in the interest of each interactant to seek to fulfill not only his own face sensitivities, but also those of the other interactants.

We have also argued that face sensitivities and wants are influenced by three factors: (1) the intrinsic imposition of acts, (2) the status difference between the interactants, and (3) the affection distance existing between the interactants. In section 7.4, we gave a classification of intrinsically face threatening actions; in section 7.5 we discussed the influence of status and power on politeness; while in section 7.6 the impact of distance and affection.

Section 7.7 contained a classification of politeness strategies. The classification scheme is based on the paradigm suggested in [Brown&Levinson], with extensions made along two dimensions: (1) a formalisation for the model is suggested, and (2) a model for multiple face impact FTA's is outlined. The chapter ends with a discussion on results obtained after examination of a conversation in terms of its politeness aspect.

The issues raised in this chapter lead us directly to concerns that will be the focus of the chapter that follows. Human interaction cannot be adequately analyzed and modeled without at least a basic understanding of the emotional impulses that drive human beings to act and react as they do. The next chapter therefore discusses issues related to modeling the human emotional system.

Chapter 8. Emotions

In the previous chapter, the first "talk about talkers" function of meta-statements was discussed. We saw that often, meta-statements are used in politeness strategies, and that a meta-statement can function as a message pertaining to the relationship existing between the interactants. In this chapter, the second "talk about talkers" function of meta-statements, the *emotion* function, is discussed. As we shall see, when people talk with other people, the language they use carries in it messages that communicate their emotional state. People do not constrain themselves to stating facts; they also have a strong tendency to convey how they feel and what emotions they have about those facts. Meta-statements are often used for just such a purpose, and in this chapter our goal is to account for the use of meta-statements to convey messages about emotional states.

Before we outline the chapter, let us first define what exactly we want to accomplish. As we will see in the examples given in this chapter, meta-statements are often used to convey information about speaker's emotional states. Talk is seldom purely informational, and even in interactional settings where one would not expect interactants to behave emotionally, communication about conversers' emotional states does take place. And since, as we shall see, meta-statements are frequently used as messages about emotional states, we must therefore devise a way of modeling the emotional meaning that meta-staments have. The first step in achieving this is to determine how emotional states of conversers can be modeled. That is, we must address questions such as, what does it mean that a person is in emotional state X, how is emotional state X characterized, what are the factors and the parameters that need to be taken into consideration when modeling emotional state X, etc.

Second, since our aim is to create a system that is *sensitive* to emotions, we must therefore determine how emotional states are manifested behaviorally. Since the only behavior to which our system is sensitive is verbal, our task is therefore to make the system infer emotions from verbal behavior, and more specifically from meta-statements. With a system able to infer emotions from behavior on one hand, and internally represent the inferred emotions on the other, we can then claim that we have, at least to a certain extent, captured the emotional dimension of meta-statements.

In section One of the chapter, we examine the question: what is to be considered an emotion? It is agreed that love is an emotion, but is irascibility, for example, an emotion? Do emotions have to be accompanied with physiological disturbances, or can an emotion be purely intellectual?[41] In section Two, we make the distinction between three types of emotional states: episodes, moods, and dispositions. In section Three, the problem of intentionality, i.e., what an emotion is *about,* is examined. In section Four, we outline what we call the *valence dimensions* of emotions. A valence dimension is a context within which a situation or an event is evaluated. We distinguish between three valence dimensions: goals, standards, and

---

[41] The model presented here is based on the cognitive model described in [ORTO88].

tastes. We hold that an emotion is a reaction to the impact an event or a state of affairs has on a valence context. Section Five contains a detailed presentation of how valence contexts are modeled in our system.

In section Six, we give a classification of emotion types. The classification is based on the discrimination between the various types of impacts an event or a state of affairs has on the valence contexts. In section Seven, we present the global parameters that influence the evaluation of an event's or state of affair's impact on a valence context. Global prameters influence all of the identified types of emotions. Local parameters, which we discuss in section Eight, on the other hand, have an influence only on a restricted set of types of emotions. In section Nine, the internal representation of emotional states is given.

In section Ten, we discuss how emotions are verbally expressed. We distinguish between two types of expressions: explicit expressions and implicit expressions. Each type is examined in detail. We conclude in section Eleven with a few examples illustrating the concepts we have been discussing.

8.1 On What is to be Considered an Emotion

Although it is generally held that terms such as "love," "anger," and "fear" refer to what is called an emotion, agreement is not unanimous when we talk of things such as "respect," "admiration," and "patriotism." Is respect an emotion? Certainly we do say that one *feels* respect for someone else. But are the "feelings" involved with respect of the same nature as those involved when an emotion such as anger is in effect? For one thing, anger seems to be more immediate, more intense, and certainly more physiologically pronounced than respect. One pounds a table when one is anger, or yells, or becomes red in the face, whereas when one feels respect for someone else one hardly any physiological manifestations take place. In a sense, feeling respect is more abstract than feeling angry. Moreover, one is in a

state of anger for a few minutes, or a few hours, whereas respect seems to be an atemporal phenomena: one respects someone else not for one minute or one hour or one day; one simply respects, and there is no notion of time involved. Should we therefore call emotions only those manifestations which are of a violent and immediate nature, such as anger or should our definition include "milder" emotions, such as respect and admiration?

Another related problem is that of moods. Should we include moods in what we consider to be emotions? Let us take a mood such as irascibility. Is irascibility an emotion? If we compared it to anger, we can see that it differs from it in at least two important respects: first, one *becomes* angry, whereas one *is* irascible. That is, the emotion of anger has an aspect of urgency and intensity to it, whereas that of irascibility is marked by a certain degree of continuity and stability. Second, one is angry at or over something, whereas one is irascible over nothing in particular. That is, anger has an object whereas irascibility does not. Should we then exclude moods from what we consider to be emotions because moods do not have a particular object to which they refer? Or perhaps should we require that emotions be intense and immediate and thus exclude moods in this way, since a mood is by definition phenomenon protracted over time?

The solution offered to this problem by Ryle [RYLE49], a solution we will adopt, is to distinguish between what he calls an *episode* and a *disposition.* He bases this distinction on a more basic differentiation which contrasts *feelings* to *inclinations.* Feelings, he says, are those "sorts of things which people often describe as thrills, twinges, pangs, throbs, wrenches, itches, prickings, chills, glows, loads, qualms, hankerings, curdlings, sinkings, tensions, gnawings and shocks" [RYLE49]. Feelings he characterizes to be immediate, temporally located and marked by physiological disturbances. An inclination, on the other hand, he defines to be a *tendency* to exhibit certain feelings under certain circumstances. Thus, he argues that saying that a man is vain is to say that this vain man has an inclination to have certain feelings and behave in a certain way when he is facing certain situations: "when thwarted, he feels acute dudgeon and when unexpectedly successful, he feels buoyant." That is, the vain man has a *disposition* to experience certain *episodes.* An emotion, Ryle then concludes, is simply

a disposition to behave in a certain characteristic way. Thus for him, respect would be an emotion, since there are certain characteristic behaviors that are exhibited under "appropriate" circumstances when one feels respect for someone else: one may use "respectful" language, feel flattered or proud when the respected person praises one, or defend the respected person's reputation when that reputation is unfairly attacked. Anger also is an emotion according to Ryle, but only in the sense that it is a disposition to behave and feel in a certain way when "made angry." And love, in its various shades and degrees of intensity, is also an emotion. Ryle holds that romantic adolescent love, with its characteristic intensity and urgency, is an emotion as much as the abstract love for justice. Both involve dispositions to behave in some characteristic way under "appropriate" conditions, and both of these dispositions are actualised through episodes, e.g., the passionate kissing between two lovers, or the self-sacrifice of a lover of justice.

An observation worth making at this point is that Ryle approaches the emotion issue linguistically, in the sense that he seems to be making the assumption that the terminology of ordinary language contains all terms that describe all possible human emotions. Whether this position is justified may be subject to debate. Are there certain clinical cases, special anxieties, for instance, which have no ordinary terminology that refers to them? Certainly the vocabulary of a psychiatrist is much more technical and exact on certain types of anxieties than ordinary language is. And could not one argue that there are some emotions that do not even have a name, whether in ordinary language or wheter in the language of the psychiatrist? And if this is so, why should we then follow Ryle's suggestion that we study emotions by studying the language that describes them?

Although it may be the case that there are emotions which do not have a name, and that ordinary language is not rich enough to cover all types and shades of emotions, we nevertheless believe that Ryle's suggestion is sound and should be followed for the following reasons. First, ordinary language covers if not the whole set of possible emotions, at least certainly a subset, and most probably an important subset, and it is a legitimate approach to study this subset through the ways the terms referring to them are used. Second, those

emotions which are explicitly identified by language, ordinary or specialized, are the *only* emotions which can be analyzed--i.e., a phenomenon needs to have a name and be identified before it can be referred to, let alone studied. That there may be certain "emotions" not covered by language should not in the least deter us from studying emotions through language, because after all we can only *talk* about those emotions for which we have a name. And in fact, we may have little choice in this matter, since we can study only what we can point to.

8.2 Emotional State Classes

A problem we mentioned in the previous section concerned the issue of whether we should consider moods to be emotions or whether we should treat them as a separate phenomenon. It is obvious that moods involve emotional experiences, but it is equally obvious that a statement such as "Jack has been irascible all day" is quite different from a statement such as "Jack became angry when Mary slighted him." To say that Jack was irascible all day is to describe Jack's *general* emotional tendencies during the day in question. The statement is not to be taken to mean that during every moment of the day Jack was angry at something. Instead, the statement is telling us that Jack had a *tendency* to be easily angred, and that this tendency lasted "all day." In contrast, the statement, "Jack became angry when Mary slighted him" refers to a particular emotional experience, an *episode,* that Jack experienced during which Jack did feel anger, and unlike the mood statement, this statement stresses the fact that anger was in effect "when Mary slighted him." Emotional states referring to moods we will refer to as *mood* states, whereas those referring to an episode as *episodic* states.

There are two basic differences to note between mood and episodic states. The first is *temporal.* Episodes span short intervals of time, whereas moods span longer intervals. Thus, "Jack jumped with joy upon learning that he passed the exam," is an episode because it describes Jack's emotional state during the brief moment "upon learning that he passed the

exam." On the other hand, "Jack has been living in ecstacy ever since he learned that he passed the exam," is a mood statement because it describes Jack's emotional state during the whole interval between the time he learned that he passed the exam and the narrative present.

The second difference between an episode and a mood is that during an episodic emotional state, the emotion described is *effectively prevalent*, whereas in the case of a mood it is *frequently prevalent.* That is, to use our two previous examples, upon learning that he passed the exam--i.e., during that period of time when he learned the news--Jack *did feel* the emotion of Joy, i.e., the emotion was *in effect.* In the second case, that is, in the statement "Jack has been living in ecstacy ever since he learned that he passed the exam," Jack's experiencing ecstacy was not in effect during *every moment* ever since he learned the good news. It is more likely that although he was ecstatic quite frequently ever since he learned of his passing the exam, during other moments he had other completely different emotions, and his ecstacy was not in effect.

Distinct from moods and episodes are emotional states that form the third category of emotional classes, *dispositions.* The main characteristic that distinguishes dispositions from moods and episodes is that whereas moods and episodes are temporal, dispositions are not. In the statement, "Jack hates blueberries," Jack's emotion of hatred is a reference neither to one of Jack's past or present emotional experiences, nor to one of Jack's moods. Instead, it is a statement about one of Jack's dispositions to experience certain feelings concerning blueberries. Phobias are one type of emotional states that fall under the category of dispositions. Being afraid of heights, for example, is a statement about one's disposition to feel fear when standing on elevated surfaces.

It is worth noting that episodes may form the basis for certain moods or dispositions. In our previous example of fear of heights, an episodic experience of being afraid of heights may have its source the phobia, i.e., the disposition to fear heights. Similarly for moods. A statement such as "Jack feels awful when the sky is gray," is a disposition about Jack's experiencing a certain mood under certain conditions.

Formally, to model an episode emotional state we need three parameters: (1) the time the emotion was triggered, (2) the length of time the emotion was in effect, and (3) the emotion in effect. A mood, on the other hand, needs only (2) and (3) for its definition, whereas dispositions needs only (3).

Thus the three types of *emotion-states* are represented internally as follows:

```
(EMO-STATE
        (PERSON     ?person)
        (STATE-ID   ?state-id)
        (STATE      ?state-type)
        (EMOTION-TYPE ?TYPE))
```

The state-id is simply a unique tag used to distinguish between emotion-states. STATE is one of the following three structures. If emotion state is an episode, the structure is as follows:

```
(EPISODE
        (TRIGGER-TIME ?trigger-time)
        (LENGTH       ?length)
        (EMOTION      ?emotion))
```

If the state-type is a mood, then frame is as follows:

```
(MOOD
     (LENGTH       ?length)
     (EMOTION      ?emotion))
```

If the state-type is a disposition, then frame is as follows:

```
(DISPOSITION
            (EMOTION ?emotion))
```

To summerise, we distinguish between three types of emotional states: (1) episodes, (2) moods, and (3) dispositions. Whereas episodes and moods are characterized by a temporal instantiation--i.e., they are in effect in actuality--dispositions are potential emotional states that may become actualized either as episodes or moods.

8.3 Intentionality

Hume originally pointed out that what differentiates an emotion such as anger or fear from a pure sensation such as a headache is the fact that emotions refer to something whereas pure sensations have no referential object [HUME78]. Thus one is angry about something, is afraid of something, whereas one suffers a headache or a toothache not with relation to something external but simply because of physiological "disturbances." Brentano also singled out this referential characteristic that emotions seem to have, pointing out that emotions, and thinking phenomena in general, "are uniquely characterized by the fact that they have an object upon which, as we might say, they are directed" [BREN89]. This reference to an object Brentano termed *intentionality.*

The problem that immediately arises is that not all emotions are equally intentional. With many emotions the object of the emotion can be easily referred to and pointed out: I feel love for Mary, and so Mary is the object of my love; I hate Joe, and so Joe is the object of my hatred; I am afraid of the dog that is snarling at me, and so the dog is the object that I fear. With many other emotions, however, the "object" of the emotion is far from being specific. This is especially true of moods. Irascibility is a case in point: one simply *is* irascible, and there is no notion of being irascible *at* something or someone. One may display his irascibility through characteristic behavior at certain objects, but these objects are not the object of the emotion of irascibility, as Mary is the object of my Love, but rather objects through which the emotion is instantiated. Whereas with my love for Mary, the *focus* of my emotion is the single "object" Mary, with irascibility my object is, as it were, the whole world.

In many instances, an emotion can be directly related to an event that in some way triggered that emotion. In our previous example where "Jack became angry when Mary slighted him," the event "Mary slighted him" is the object of Jack's anger. On the other hand, with other emotions, the emotion's intention is not an event but instead a *state of affairs.* Suppose I learn that "Jack is a rich man," and suppose that upon learning this the emotion

triggered in me is that of envy. Obviously, in this case the emotion of envy is related to the state of affairs of Jack's being rich. That is, the object of my envy is Jack's state of being rich. Emotions such as love, liking, hatred, envy, are usually associated with states of affairs. In the statement "Jack loves Mary," the object of Jack's love is who Mary *is,* or more specifically, the person Jack sees Mary to be, i.e., her attributes, including her physical features, her character, her beliefs, her deeds, her affection for him, etc. Jack may at a certain point discover new and crucial information about Mary, e.g., that she is in fact affectionate towards him only because he is rich, and may, as a consequence, have a change of heart. That is, he may perceive a completely new state of affairs with respect to Mary, and as a consequence may not feel love for her any longer. In either case, whether loving or not, the emotion's object is a state of affairs.

With many other emotional states, however, the emotions felt during those states have no identifiable objects, whether in the form of events or states of affairs. Moods are usually "objectless" emotional manifestations, e.g., "I am feeling low today, but I don't know why." However, there may also be emotional episodes during which no object may be identified, as for example with "unjustifiable" feelings of anxiety or foreboding that last only for a few moments. And dispositions may also be objectless, as in "Jack is a cheerful fellow." Jack's disposition of being cheerful is related neither to some specific event or events nor to some specific states of affairs, but is instead related to one of his character traits. The "object" of objectless emotions--or rather, the objectless character of objectless emotions--is denoted by the value *nil.*

To summarize, we distinguish between three types of intentional objects: (1) events, (2) states of affairs, and (3) nil objects. An emotional state, whether it is an episode, a mood, or a disposition, is always realted to one of these three intentional objects. An intentional object is internally represented as follows:

```
(EMO-OBJECT   ?object)
```

8.4 Valence Dimensions

In addition to characterizing emotions in terms of the intentional objects to which they are related, we also characterize emotions in terms of what we call their *valence dimension.* We define a valence dimension to be the semantic context with which an emotion's intentional object is related. We distinguish between three types of valence dimensions: *goals, standards,* and *tastes.* To clarify the definition, let us take some examples.

Let us suppose that Jack's goal is to graduate by the end of this semester, and let us suppose that between him and graduation stands only one exam: if he passes the exam, he will pass the course and be able to graduate, but if he fails the exam, he will also fail the course and consequently not be able to graduate. Let us say that Jack in fact does pass the exam, and that, as a result, he feels ecstatic. In this instance, we say that the valence dimension of Jack's ecstacy is *goal based.* Jack had a goal--to graduate--the fulfillment of which he valued greatly, and one of the necessary subgoals of this goal was to pass all courses, a subgoal of which was to pass the course in which he took the exam, a subgoal of which was to pass the exam in question. Fulfilling the lowest level goal, i.e., passing the exam, fulfilled the higher goal of passing the course, which in turn enabled the fulfillment of the higher level goal of graduating. That is, the evaluation of the event of passing an exam is in this case based on the various goals and plans that Jack has.

Not all emotional reactions are necessarily related to goals, however. Some emotions are instead related to *standards.* By "standards" we mean those events or states of affairs one believes *ought* to obtain. The keyword in the definition is "ought." The object of a standard emotion is not related to a goal or a plan, but rather to a *norm,* a code of conduct. For example, emotions such as disdain, indignation, guilt, are all related to what one *believes* ought not to obtain. Emotions such as pride, admiration, respect, on the other hand, are related to what one believes ought to obtain, and thus one has a positive emotion as a consequence of their having obtained. No notion of functionality--i.e., in terms of promoting or

undermining goals and plans--is present. Instead, standards emotions are evaluated relative to *moral* structure.

The third valence dimension refers to what we call *tastes.* Whereas in the case of plans an event or a state of affairs is evaluated with respect to a goal structure, and in the case of standards with respect to a moral structure, in the case of tastes the evaluation is completely subjective. Emotions such as liking, loving, hating, are often totally a matter of taste. Thus, in "Jack hates blueberries," Jack's hating blueberries may be explained neither in terms of goals nor in terms of moral beliefs. It may simply be a matter of *taste* that Jack hates blueberries, and instead of a goal or a moral structure, the reaction may be based solely on Jack's subjective likes and dislikes.

Along each of the three valence dimensions identified we distinguish between two types of reactions: those that are *positive* and those that are *negative.* For the goal based dimension, reacting positively is referred to by the term *pleased,* whereas the negative reaction is referred to by the term *displeased.* The terms used for standards are *approving* and *disapproving,* respectively for positive and negative reactions, while those used to tastes are *liking* and *disliking.* These terms are not to be taken as emotional descriptions; we are using them here in a purely technical sense.

Before we go any further, let us make clear the following point: it is not our intention to imply that any situation is evaluated in terms of only one valence dimension, at the exclusion of the other two dimensions. In the example of Jack's passing the exam, for instance, Jack may have felt a positive emotion along all three dimensions: i.e., in addition to his being pleased by his passing the exam (because it enabled the fulfillment of a valued goal), he may also have felt approval--e.g., he believes that passing exams is something gqod--or perhaps also liking--he likes to succeed for the sake of succeeding. In such an instance, not only one but possibly all three types of emotions may be simultaneously in effect. That is, the emotional reaction elicited by any intentional object depends on the valence dimension along which the object is evaluated, and since an object may be evaluated along more than one dimension, an intentional object may elicit more than one emotional reaction.

8.5 *Modeling Valence Contexts*

As we said in the previous section, there are three context dimensions upon which an emotion is based: goals, standards, and tastes. We saw in section 8.3 that the intentional object of an emotion can be either an event, a state of affairs, or a nil entity. In this section we propose to specify how intentional objects and valence contexts interact and how that interaction influences emotional reactions and states.

To clarify what we mean let us take the valence context of goals, and let us see how emotions, intentional objects and contexts are related.

Emotions based on goals are those emotional reactions where events or states of affairs are evaluated in terms of whether those events and states of affairs promote or undermine goals. An event or state of affairs is evaluated negatively, i.e., is deemed undesirable, if it undermines a goal, and positively, i.e., desirable, if it promotes the goal. In addition, the degree to which an emotion is deemed desirable or undesirable depends on how much the goal accomplished or hindered is sought. An event or state of affairs that facilitates the accomplishment of a long sought and greatly valued goal will produce an intensely positive reaction such as joy or ecstacy, whereas one that facilitates only a mildly desirable goal will produce a lukewarm positive reaction such as contentment.

And similarly, an event or state of affairs that hinders the accomplishment of some goal will produce a negative reaction. If the goal is greatly valued then the emotion will be intense, e.g., shock, perplexity, etc., whereas if the goal is only mildly valued then the emotion also will be mild in its intensity, e.g., regret, displeasure, etc.

We will call the parameter measuring the valuability of an intentional object along the goals dimension the *desirability* of that intentional object

The second valence dimension we identified in the previous section had norms instead of goals as a basis upon which events or states of affairs are evaluated. Therefore, in order to be able to model emotional reactions along this dimension we must build a model of norms.

We saw in Chapter Six that social behavior is largely based on norms of conduct, and that human behavior is regulated to a great extent by expectations to conform to those norms. An emotional reaction can result as a consequence of violating or conforming to expected norms of conduct.

In addition to conduct, which was the focus of Chapter Six, evaluations along the standards valence dimension can also have as intentional objects states of affairs. Thus reacting with shock upon seeing a naked woman is a reaction to a violation of a norm that is related to a state of affairs rather than an action or an event.

The parameter measuring the valuability of an intentional object along the standards dimension we shall call the *praiseworthiness* of that object.

The third valence dimension has as its semantic context the subjective tastes of the person experiencing the emotion. To be able to model this semantic domain, we must create a knowledge base about the concerned person's likes and dislikes. This knowledge base, however, must be built in such a way that the reactions along the taste dimension can be inferred when no explicit information exists in the knowledge base concerning the object in question. For instance, knowing that Jack hates blueberry pies, we might tentatively infer that he also hates blueberry yogurt, blueberry juice, blueberry chewing gum, i.e., that he hates blueberries. And also, if we knew that Jack loves strawberries, then we might infer that he loves strawberry pies, strawberry ice cream, strawberry juice, and in general anything containing strawberries.

In short, our model must be capable of comparing two "similar" intentional objects and infer, knowing the like/dislike reaction with respect to one object, the reaction with respect to the other object. In addition, the system should also be able to abstract from one set of particular objects to a category of objects which will elicit a certain type of like/dislike reactions.

Internally, the three parameters are represented as follows:

```
(EMO-PARAM DESIRABILITY
           (OBJECT ?object)
           (VALUE  ?value))
```

with ?value being TRUE, if ?object is desirable, and FALSE, if ?object is not desirable.

```
(EMO-PARAM PRAISEWORTHINESS
           (OBJECT ?object)
           (VALUE  ?value))
```

with ?value being TRUE, if ?object is praiseworthy, and FALSE, if object is not praiseworthy.

```
(EMO-PARAM APPEAL
           (OBJECT ?object)
           (VALUE  ?value))
```

with ?value being TRUE, if ?object is appealing, and FALSE, if object is not appealing.

8.5.1 Modeling Goals

In this subsection, we discuss some issues concerning goals and plans that will be crucial in our modeling emotions and emotional states.

Goals Structure

Figure 18 shows the goal structure for the CLIMB-HIMALAYAS goal. Although the structure is obviosuly not detailed enough and certainly not complete, it will nonetheless do for our purposes.

The structure is hierarchical: a node in one level has the nodes connected to it from a lower level as subgoals. Thus, the goal CLIMB-HIMALAYAS has as subgoals TRAIN and GO-TO-HIMALAYAS. Each of the links that link the structure's nodes are labeled in the following way. A link is labeled with an **N** if the subgoal connected to the higher goal is *necessary* to the higher goal in question. Thus, in order to accomplish goal CLIMB-HIMALAYAS, two

subgoals are necessary: TRAIN and GO-TO-HIMALAYAS. A link is labeled with an **S,** on the other hand, if the subgoal it connects to the higher goal is by itself sufficient for the fulfillment of the higher goal. Thus, for example, the TRAIN goal can be filled *either* by TRAIN-SELF, *or* by GET-SPONSOR, *or* by JOIN-NATIONAL-CLIMBING-TEAM.

A link labeled with an **F** links a subgoal that *facilitates* the accomplishment of the higher goal. Thus, achieving CLIMB-HIMALAYAS will facilitate the goal MAKE-DAD-PROUD. The former goal is neither sufficient nor necessary to fulfill the latter goal, it only facilitates it. And finally, a link labeled with an **I** indicates that the subgoal linked to the higher goal is *inhibitive* to the higher goal in question. The goal CLIMB-HIMALAYAS, for instance, will inhibit the goal KEEP-MOM'S-PEACE-OF-MIND, since climbing a mountain will probably worry and disturb Mom's peace of mind.

Internally, an object's evaluation in terms of its goal aspect is represented as follows:

```
(EMO-VALUE GOAL
           (OBJECT     ?object)
           (SUPER-GOAL ?super-goal)
           (TYPE       ?type))
```

with ?super-goal being the goal under which the goal to which ?object is related falls, and ?type taking one of the following three values: N, if the goal is necessary, S, if the goal is satisfactory, I, if the goal is inhibitive, and F if the goal is facilitative.

Dynamics of Structure

Pruning vs. Severing

When a goal is fulfilled, the consequence of this fulfillment is reflected in the plan structure by *pruning* the node representing that goal. For example, referring to Figure 18, if the national climbing team accepts me, then the JOIN-NATIONAL-CLIMBING-TEAM node will be pruned. On the other hand, if the fulfillment of a goal becomes impossible, then the structure will be

modified by *severing* the node representing that goal. If in Figure 18 the national team rejects me, then the JOIN-NATIONAL-CLIMBING-TEAM will be severed.

In what follows, we specify the six operations associated with pruning and severing nodes from a goal structure.

(1) If an S subgoal is acheived,[42] then prune away *all* the other subgoal nodes that are related to the pruned subgoal's higher goal.
(2) If an N subgoal is achieved, then prune *only* that subgoal's node.
(3) If all of a node's subgoal nodes are pruned, then prune that node.
(4) If an S goal is not achievable, then sever only the node associated with that goal. If only one S node is left after severing an S node, then relabel that remaining node with an N label.
(5) If an N goal is not achievable, then sever the respective node along with all nodes linked to the severed node higher node.
(6) If all subgoal nodes are severed, then sever the higher node.

To illustrate, suppose the goal GET-SPONSOR is fulfilled. Since the goal's node is linked to its higher node TRAIN by an S link, then, according to (1), all other subgoal nodes--TRAIN-SELF and JOIN-NATIONAL-TEAM--are to be pruned, and consequently, according to (3), node TRAIN is to be pruned. In addition, according to (2) only the TRAIN subgoal node is to be pruned, with the other subgoal, GET-TO-THE-HIMALAYAS remaining to be achieved.

Disturbance Propagation

We shall use the term *pruning sequence* to refer to a sequence of pruning operations resulting from the pruning of nodes whose goals have been fulfilled, and the term *severing sequence* for a sequence of severing operations resulting from the severing of nodes whose goals cannot be fulfilled.

The impact of an intentional object i is evaluated in terms of how far into the goals tree its disturbance is propagated. We hypothesize the following:

(1) If i creates a severing sequence, then the longer the sequence the more intense the undesirability.

---

[42] An S subgoal is a goal linked by an S to its higher goal. Similarly for N.

(2) If i creates a pruning sequence, then the longer the sequence the more intense the desirability.

Types of Goals

We distinguish between goals along two different dimensions. First along the *pursuit type* dimension, and second along the *fulfillment type* dimension.

Pursuit Type Distinction:

Goals differ from one another with respect to the type of pursuit on which they are based. *Active-pursuit* goals--A-goals--are the types of goals where one wants to achieve something, e.g., wanting to buy a house, or to climb a mountain, etc. A-goals are to be distinguished from *interest* goals--I-goals--in that the achievement of A-goals may be promoted by exerting effort to realize the goal--e.g., working hard and saving money in order to buy a house--whereas the achievement of I-goals cannot be influenced by expenditure of effort. An example of an I-goal is the desire that the hometown team to wina game. This goal is an I-goal because one has no influence over one's favorite team's winning or losing (and if one had such an influence, then the goal would become an A-goal).

The third category of goals includes those we call *replenishment* goals--R-goals. R-goals are those wants which are never satisfied permanently. For instance, the want to eat or drink is satisfied only temporarily, and five or six hours after eating, the desire to eat surfaces again. R-goals are not to be confused with A-goals. An R-goal is an abstraction that captures the *repeatability* of certain A-goals--A-goals which are individual instantiations of the R-goal in question. For example the desire to eat in general is an R-goal, whereas my wanting to eat now is an A-goal.

Fulfillment Type Distinction:

Another dimension along which goal types are distinguished is the extent to which goals *can* be fulfilled. Some goals can be fulfilled either completely or not at all, e.g., passing an exam or failing. One cannot pass an exam halfway or only partially; one either passes or fails. These we call *all-or-none* goals. Another type of goals are those goals that have gradations of fulfillment, e.g., sweeping the floor. The floor can be swept completely, or only half or a third of the floor can be swept. These goals are said to be *partially-attainable.* In addition, there are goals that are said to be *only-partially-attainable.* For instance, the goal of attaining knowledge can never be fully fulfilled, but only partially so.

8.5.2 Modeling Standards

We now turn to the second valence dimension, standards. As we said, standards are beliefs one holds about what events or situations *ought* to obtain, given one's standards, such as norms, moral beliefs and precepts.

We saw in Chapter Six that human social conduct is heavily based on expectations borne from the conventions built around a role's norms and functions. The equivalent of function in this chapter's terminology is goal, and a behavior associated with a role is partially concerned with achieving certain goals and fulfilling certain tasks. However, there is also a normative aspect to behavior which we designated in Chapter Six by the term "norm," and we argued that behavior was evaluated in terms of its conformity to norms of conduct. Emotional reactions are often based on evaluations of conduct with respect to the normative code dictated on the actor by the social structure within which the actor is behaving. Thus, in a one-to-one relationship, the normative code dictated by the status that is assumed by each of the parties involved in the relationship in question is called forth--e.g., the husband and wife codes--whereas in a group the position normative code is evoked. Similarly, in an organiza-

tion it is the situs code which is called forth, and in a community the code pertaining to the station is evoked.

In addition to behavior, we must also model states of affairs that either conform or deviate from expected states of affairs. When we enter a McDonald's restaurant, we expect to see a counter, people dressed in McDonald's uniforms, the smell of food, seats and tables, etc., i.e., a state of affairs. Not all McDonald's are the same, but they all more or less conform to our expectations about what a McDonald's should look like. Our system must therefore be able to evaluate the degree of conformity of one state of affairs compared to an expected state of affairs. And if there is deviation, our system must be able to evaluate the degree of deviation as well.

In addition to evaluating the degree of conformity of an intentional object, our system must also be able to differentiate between three types of norms. First are those norms we call *obligatory*. A norm is said to be obligatory if it is in the form "one must do X," or "things must be X." "One *must* pay for his marchandise before leaving the store" is an event obligatory norm, whereas "the size of the class must be at least ten," is a state of affairs obligatory norm. Second are those norms we call *optional.* An event optional norm is the statement "one *should* donate to charity," whereas a states of affairs optional norm is the statement "one should be generous." Between obligatory and optional norms are *normally expected* norms. To conform to an optional norm would bring about praise, whereas to conform to an obligatory norm would bring little or no praise to the person conforming to the norm. Also, to violate an optional norm would bring about little or no negative reactions, whereas the violation of an obligatory norm would most probably cause rather intense criticism. A normally expected norm combines obligatory and optional norms in the following way: violations will bring about little criticism and conformity little praise. One who wears blue pants should not expect to be patted on the back for conforming to the norms dictated by the dress code, and one who wears yellow pants with green stripes and pink dots should not expect to be harshly criticised for his "deviation."

The distinction between these three types of norms is important in determining the intensity of an emotional reaction. Deviations from normally expected norms will usually elicit mild reactions, e.g., reproach, while those deviations that violate an obligatory norm will usually be accompannied by more intense reactions, e.g., outrage. And similarly, conformity to a prohibitive norm--e.g., one does not steal--will not elicit reactions that are as intense as those elicited by conformity to an optional obligatory norm--e.g., one should give to charity.

Thus, the six influences on the praiseworthiness parameter are as follows (*i* indicates the intentional object):

(1) The approval value of i is high if there is conformity to an optional norm.
(2) The disapproval value of i is nil if there is violation of an optional norm.
(3) The approval value of i is low if there is conformity to a normally expected norm.
(4) The disapproval value of i is low if there is violation of a normally expected norm.
(5) The approval value of i is nil if there is conformity to an obligatory norm.
(6) The disapproval value of i is high if there is violation of an obligatory norm.

Internally, an object's evaluation in terms of its standards aspect is represented as follows:

```
(EMO-VALUE STANDARD
           (OBJECT     ?object)
           (NORM-SPECS ?norm-specs)
           (CONFORMITY ?conformity))
```

with ?conformity being TRUE if the norm in question is followed, and FALSE if it is violated, and ?norm-specs being a NORM-SPECS frame, as specified in Chapter Six.

8.5.3 Modeling Tastes

This third valence dimension captures those situations where the emotional reaction is based on personal tastes and biases. As we said in section 8.4, in order to be able to model tastes, we must have not only a static knowledge base containing information about likes and dislikes, but also a mechanism for dynamically inferring likes and dislikes about a situation's intentional object. This mechanism must obviously have at least two components: (1) current

data about established likes and dislikes with respect to intentional objects previously evaluated and (2) a mechanism for evaluating the similarity between a situation's intentional object and an evaluatively tagged intentional object present in the knowledge base. For instance, suppose we have in our knowledge base the statement "Jack hates blueberries," and suppose we were to ask ourselves, "how would Jack like blueberry juice?" A reasonable answer might be "Jack would hate blueberry juice, because blueberry juice comes from blueberries." To be able to reach such a conclusion, our system must have a way of comparing the two intentional objects "blueberries" and "blueberry juice" and determine that these two objects are quite similar within the context in which they are evoked--i.e., that they are both evaluated in terms of their edibility.

Internally, an object's evaluation in terms of its tastes aspect is represented as follows:

```
(EMO-VALUE  TASTES
            (OBJECT ?object)
            (VALUE  ?value))
```

with value being TRUE if ?object is liked, and FALSE, if ?object is not liked.

In the following we proceed with the presentation of the classification of emotion types we will use in our model.

8.6 A Classification of Emotion Types

Not surprisingly, the classification of emotion types is based at the highest level on valence contexts. Thus we distinguish between three categories of emotions: those based on goals, those based on standards, and those based on tastes.

Goals Based Emotions

We saw in the previous section examples where an emotional reaction was directly related to the impact events and states of affairs had on goals entertained by the person experiencing the emotional reaction. Emotions falling under this category belong to one of two types of subcategories: (1) those focusing on self, and (2) those focusing on others.

Focus-on-Self

This subcategory further divides into two subcategories: (1) those whose intentional object is non-prospective in natures, and (2) those whose intentional object is prospective.

Non-Prospective (Well-being)

Non-prospective emotions are those whose intentional object is an event or state of affairs that has obtained or is in the process of obtaining. Thus, joy upon learning that one passed all his classes is a non-prospective emotion, since passing the exams has already taken place. This category is also reffered to as the well-being category.

Prospective (Prospect-based)

Prospective emotions are those emotions whose intentional object is an event or a state of affairs that is characterized by its being viewed as a future prospect. We make the distinction between two types of prospective emotions: (1) those whose intentional object has not yet obtained, vs. (2) those whose intentional object did obtain. Emotions in this category are said to be prospect-based.

Unrealized

Unrealized prospetive emotions have as an intentional object events or states of affairs that have not obtained yet. While waiting for the results of his examination, Jack might at times experience the emotion of hope at the prospect of his success, and fear at the prospect of his failure. The intentional object of each of these two emotions is said to be prospective since it has not yet taken place but still belongs to the future.

Realized

On the other hand, emotions whose object was a previous prospect that has now obtained are said to be realized prospective emotions. If Jack should learn that he did pass all his classes he will feel satisfaction. Should he fail, he will feel disappointment. We further make the distinction between two types of realized prospective emotions: (1) those whose object confirms a fear or a hope, which we will call *confirming realized prospective,* and (2) those whose object disconfirms a fear or a hope, and these we will call *disconfirming realized prospective* emotions. Satisfaction thus is a confirmed emotion, since it results from the obtaining of an expected good thing, while the feeling of relief is a disconfirmed emotion, since it results from the non-obtaining of a feared bad thing.

Focus-on-Others (Fortunes-of-Others)

In this category, the emotional reactions involve another person distinct from the person feeling the emotion. That is, the intentional object of the emotions in this category involve in some way an other person. We distinguish between two types of focus on others emotions: (1) those that are desirable for others, and (2) those that are undesirable for others. The term fortunes-of-others is also used for the emotions forming this category.

Desirable for Others

Desirable for others emotions are those emotions whose intentional object is an event or a state of affairs that another person evaluates as good for themselves--i.e., good for the other person. Jack winning the lottery most probably will be evaluated as good by Jack. The emotion I or any one else (besides Jack) might feel with respect to the event "Jack won the lottery" is then a Desirable for Others emotion. If I were a good friend of Jack, I might feel happy for him. If, on the other hand, I hated Jack I, might very likely react with resentment.

Undesirable for Others

Undesirable for Others emotions are those emotions whose intentional object is an event or a state of affairs that another person evaluates as bad for themselves. Jack failing an exam most probably will be evaluated by Jack as something bad. If I liked Jack I might feel pity for him, while if I hated him I might gloat over his failure.

Standards Based Emotions (Attribution)

We now turn to the second type of emotions, those that are reactions to conformity or deviation to norms and standards. We distinguish between two types of standards based emotions: (1) those focusing on the self, and (2) those focusing on others. The emotions of this category are refered to as attribution emotions.

Focus-on-Self

Standard based focusing on the self are emotions such as feeling ashamed, embarrassed about an event or a state of affairs one is in some way related to. Pride, for instance, is usually an emotion felt after one has accomplished something that is deemed good, e.g., donating to charity. Shame, on the other hand, is an emotion usually felt by one after one does something deemed to be bad: e.g., stealing. Notice that one may also be proud not of an event but

of simply a state of affairs: e.g., one is proud of the beauty of one's face, or one is ashamed of the ugliness of one's face.

Focus-on-Others

Focus on Others emotions, on the other hand, are emotional reactions that evaluate the conformity or deviation of another person or persons from norms and standards. For example, admiring someone is an emotion felt for one who does something that is considered to be good: e.g., one admires generous people. On the other hand, if someone does something bad, then one will have a tendency to feel reproach for that person, e.g., somone who does not give his seat on the bus to an elderly person who is standing up.

States of affairs can also be intentional objects for this category of emotions. One may admire another person's generosity (and not simply a particular act of generosity), or may reproach another person's miserliness.

Tastes Based Emotions (Attraction)

The third category of emotions includes those emotions that are reactions to violations or adherences to preset likes and dislikes. Most prominent among the emotions belonging to this category are the emotions of love and hate. The emotions in this category are also referred to as attraction emotions.

Compound Emotions

In addition to the tree major types of emotions, we also identify a fourth category of emotions, that which includes emotions that are the result of combining well-being and attribution emotions.

These emotions are the result of evaluating an intentional object along two dimensions simultaneously. For example, an intentional object that is evaluated as praiseworthy, involves another person (who is responsible for the event or states of affairs), and is desirable for self brings forth an emotion such as gratitude, while an object that is blameworthy, involves another person (who is responsible for the event or states of affairs), and is displeasing to self brings forth an emotion such as anger. On the other hand, an intentional object that is evaluated as praiseworthy, is desirable, and one is responsible for the praiseworthy and desirable event or state of affairs brings forth an emotion such as gratification, while an intentional object that is evaluated as blameworthy and undesirable, and for which one is responsible brings forth an emotion such as remorse.

We have presented in this section a classification of emotions based on the semantic context with respect to which events or states of affairs are evaluated. In what follows we identify precisely those parameters that influence how events or states of affairs are evaluated along each of our three dimensions. In section 8.7, we describe factors that influence all types of reactions--global variables--while in section 8.8 we specify those factors that are particular to each valence dimension.

8.7 Global Parameters Influencing Valence Reactions

We discussed in section 8.5 some factors that influence how an event or a state of affairs is evaluated given the valence context within which the event or state of affairs occur. For instance, we gave the example of an event enabling a goal to become fulfilled, and we argued that along the goals dimension, the emotion elicited by such an event will probably be something akin to satisfaction or joy, depending on how desirable the fulfilled goal is.

However, in addition to the parameters presented in section 8.5--i.e., desirability, praiseworthiness, and appealingness--there are also other factors that have influence on the type and intensity of emotional reactions. In this section we discuss those parameters that have influence over all types of emotional reactions, while in the section that follows we discuss parameters that are particular to only specific emotion categories.

Sense of Reality

The sense of reality parameter is a measure of how cognitively *real* the intentional object under evaluation is perceived to be. The simple hypothesis we shall hold is this: *the more real an object is evaluated, the more intense is the emotional reaction felt towards that object.* Fear is the emotion that readily illustrates the influence of the sense of reality factor: the more real the danger, the more fearful one becomes. Hope also is another case where this correlation can be illustrated. One cannot hope for something that is absolutely not feasible: there must be at least a germ of a chance for the object to realise, otherwise the emotion felt would not be hope.

An interesting issue related to the notion of reality concerns those situations where the events and states of affairs are known to be absolutely fictitious and yet nevertheless elict

emotional reactions. What we have in mind are emotional reactions to movies or fictional tales. People do react with fear, anxiety, disgust, delight, or even joy when watching movies, even though they are perfectly aware that what they are watching is nothing more than fiction. How does the sense of reality correlate with the intensity of emotions for *these* cases?

In answer to this, we make the following observation: A movie, or any fictional message, is effective only to the degree to which it succeeds in emulating real life situations. Should its plot be weak, its dialogue artificial, its characters badly drawn, or the acting badly executed, then the believability of a movie will weaken, and with this weakening the emotional reactions elicited by the situation depicted in the movie will be few and not as intense as those a more believable depiction would have elicited.

Internally, this parameter is represented as follows:

```
(EMO-GLOB-PARAM
                    SENSE-REAL
                    (OBJECT ?object)
                    (VALUE ?value))
```

with ?value taking one of the following values: TRUE, if the ?object is perceived true, and FALSE otherwise.

Psychological Proximity

Related to, but distinct from the sense of reality parameter is the psychological proximity of the intentional object. This parameter is an evaluation of the degree of proximity between the intentional object and the person evaluating the object. We distinguish between two types of proximities: *spatial* and *conceptual.* By spatial proximity we refer to the time and space distance between the person evaluating the object and the object evaluated. For example, one might feel greater and greater anxiety as a dreaded event approaches, or less and less distress as a traumatic event recedes into the past. Similarly, one might feel more disgust at the

sight of a mutilated body a few feet away than if one were simply told about the mutilated body.[43]

Conceptual proximity, on the other hand, is an evaluation of an object's proximity to the observer along non-physical dimensions (by physical we mean, of course, time and space). The news of the death of someone we knew has greater impact on us than the news about a stranger's death. The person we knew was closer to us than the stranger not in terms of time or time or space, but rather in terms of familiarity. Solidarity, which restricts the intentional objects to persons, is another important type of conceptual proximity. One feels closer, as it were, to people of one's own kind.

As we said at the beginning of our discussion, proximity is to be distinguished from sense of reality. The two parameters are related, but distinct. Usually they are positively correlated with one another, i.e., the more proximate an object the more real the object in question is evaluated, but this is not necessarily the case for all situation. Cases such as mental blocks of shocking situations are highly proximate, but, because they are blocked, and this as a consequence of their being *too* proximate, their sense of reality is reduced to zero from the point of view of the shocked observer. As time goes by, and the event recedes into the past, i.e., the object becomes less proximate temporally, the reality of the shocking event may slowly start to "sink in" and as a consequence a whole sequence of emotional reactions subsequently follow.

Internally, this parameter is represented as follows:

```
(EMO-GLOB-PARAM PROXIMITY
                (OBJECT ?object)
                (TYPE   ?type)
                (VALUE  ?value))
```

with ?type being SPATIAL or CONCEPTUAL, and ?value being either NEAR, or FAR.

---

[43] Notice the relationship between the sense of reality and proximity here. A mutilated body we do not see is not as "real" as one we do see. But technically speaking, the existence of the mutilated body is real in both situations. That is, it is not a matter of denying that a body exists that the disgust reaction is milder in the case where the body is not in sight, but rather a matter of not being physically near it.

Unexpectedness

The third and last global parameter is the one that measures the unexpectedness of the intentional object under evaluation. We propose that evaluations are correlated to the unexpectedness parameter in the following manner:

(1) - Unexpected positive objects are evaluated more positively than expected ones.
(2) - Unexpected negative objects are evaluated more negatively than expected ones.

Thus, one is more delighted at finding ten dollars, when one did not at all expect to find the money, than one who is given ten dollars from a person who had borrowed the money and was expected to return it any day, even though in both cases one is ten dollars richer. An similarly, one who expects to fail an exam will feel less disappointment than a person who expected to pass, even though both might have desired to pass with equal intensity.

An issue that is of importance for modeling expectations is the distinction between two types of unexpectedness. First is the unexpectedness based on a violation of expectations: I expect my alarm to sound off at eight o'clock every day, and if it does not then I evaluate that event as unexpected. Second is the unexpectedness that is not based on any expectations, but somehow "springs out of the blue," e.g., finding ten dollars. The distinction is important because it is crucial in the construction of the mechanism for evaluating intentional objects. Those intentional objects that violate an explicit expectation can be easily tagged as being unexpected, and the degree of their unexpectedness positively correlated to the degree of expectedness of the violated expectation. But those objects that did not violate any specific expectations are not so easily amenable to evaluation, and more crucially to identification as unexpected objects. For our current needs, we shall *not* deal with these types of expectations. To solve the "out of the blue" unexpectedness problem, the problem of relevance in context must be solved,[44] i.e., how relevant is an event or a state of affairs given the current context. This involves issues such as cognitive focus management and classification and identification

[44] See [SPER86] for a relevant discussion on this issue.

of events and situations in terms of context, among other non trivial problems, issues that are certainly not treatable within the space of our current work.

Internally, this parameter is represented as follows:

```
(EMO-GLOB-PARAM UNEXPECTEDNESS
                (OBJECT ?object)
                (VALUE  ?value))
```

with ?value being TRUE if ?object was unexpected, and FALSE, if ?object was expected.

8.8 Local Parameters Influencing Valence Reactions

We now turn to a specification of those parameters that are particular and apply only to some types of emotions and not others.

Goals-based

We saw in section 8.5 how the local parameter of an object's desirability can be measured in terms of that object's impact on the fulfillment of goals. We argued that the desirability of an object is partially influenced by the types of goals promoted or undermined and the extent to which goals are fulfilled or inhibited. In addition to the desirability parameter, we specify in what follows parameters that are specifically relevant to two subcategories of emotion, (1) those that are focused on the self and prospect based, and (2) those that are focused on others.

Prospect-based

We identify three prospect based parameters:

(1) *Likelihood:* which measures the degree the object is considered likely to obtain.[45] Relief is greater if the feared object was certain to obtain, but did not, and disppointment is greater if the object was certain to obtain but did not.

Internally, this parameter is represented as follows:

```
(EMO-LOC-PARAM
                LIKELIHOOD
                (OBJECT ?object)
                (VALUE  ?value))
```

with ?value being TRUE if the object's prospective realisation is likely, and FALSE otherwise.

(2) *Effort:* which measures the effort expended to realize or prevent the prospective object. Thus, the more one endeavored to prevent an evil the more relieved one is that it did not occur, and the more one endeavored to obtain some intended object one deemed desirable and failed to obtain that object, the more disappointed one is.

Internally, this parameter is represented as follows:

```
(EMO-LOC-PARAM
                EFFORT
                (OBJECT ?object)
                (VALUE  ?value))
```

with ?value being GREAT, if great effort was expended in bringing into effect the object, and LOW, if little effort were expended.

(3) *Realization:* which refers to the degree to which the realized prospective event is realized. Obviously, only partially-attainable and only-partially-attainable goals may be involved in this case. Getting a "good" grade will probably bring forth a joy emotion, but how "good" the grade is will have an influence, probably directly proportional, on the intensity of the joy felt.

---

[45] Likelihood is to be collapsed with the global parameter *expectedness*. It is explicitly mentioned for this category of emotions because it is crucial in the evaluation of the emotion's intensity.

Internally, this parameter is represented as follows:

```
(EMO-LOC-PARAM
                REALIZATION
                (OBJECT ?object)
                (VALUE  ?value))
```

with ?value taking the value COMPLETE if the ?object was realized completely or nearly completely, HALF, if the object was realised halfway, and LITTLE, if the object was realised only to a small extent.

Focus-on-Others

For focus on others we identify three parameters:

(1) *Desirability-for-Other:* which refers to the degree the other person, i.e., the person who is in focus, considers the object desirable or undesirable. Thus, if something that a friend considers desirable happens to my friend I will feel good for my friend, whereas the thing is considered bad by my friend, then I will feel pity or sorrow.

Internally, this parameter is represented as follows:

```
(EMO-LOC-PARAM
                DESIRE-OTHERS
                (OBJECT ?object)
                (PERSON ?person)
                (VALUE  ?value))
```

with ?value taking the value TRUE if ?person considers ?object desirable, and FALSE otherwise.

(2) *Liking:* which measures the degree the person in focus is liked by the person experiencing the emotion. The more one likes a person the more they will wish them good, and vice versa.

Internally, this parameter is represented as follows:

```
(EMO-LOC-PARAM
                LIKING
                (PERSON ?person)
                (VALUE  ?value))
```

with ?value taking the value TRUE if ?person is liked, and FALSE otherwise.

(3) *Deservingness:* which refers to the degree the person in focus deserves the benefits of a desirable object or the losses caused by an undesirable object. For instance, one might feel great resentment towards a person who received a perfect score in an examination, knowing that that person cheated and so did not deserve the perfect score. Similarly, one may feel pity for a person who received a bad score on the exam, knowing that that person knew the material covered in the exam but failed to perform well because they lost their composure.

Internally, this parameter is represented as follows:

```
(EMO-LOC-PARAM
                DESERVE
                (OBJECT ?object)
                (PERSON ?person)
                (VALUE  ?value))
```

with ?value taking the value TRUE if ?person is seen as deserving the benefits or adverse consequences of ?object, and FALSE otherwise.

Standards-based

We now turn to parameters that pertain locally to objects evaluated relative to norms. In section 8.5, those factors influencing *praisworthiness* were discussed. The following two parameters are also relevant to the evaluation of the intensity of standards-based emotions.

(1) *Identification:* which refers to the identification of a praised or blamed object with another object that is deemed praiseworthy or blameworthy. For instance, one may praise another person on the beauty of that person's daughter, and the other person might feel proud as a

result. The direct object praised is the daughter, but the parent felt pride because they identified with the praised object, i.e., their daughter, and thus directed the praise towards themselves. The same holds for blame. Of might feel guilty not because one did anything, but because one is somehow associated with something or someone that did something "bad," e.g., feeling guilty when meeting a person whose father was killed in a war between that person's country and one's own country.

Internally, this parameter is represented as follows:

```
(EMO-LOC-PARAM
                IDENT
                (OBJECT ?object)
                (VALUE  ?value))
```

with ?value taking the value TRUE if there is an identification with ?object, and FALSE otherwise.

(2) *Deviation:* which refers to the degree the event or state of affairs deviates from the expected norms. Notice that this is a measure of *degree* and is different from the deviations discussed with regards to the evaluation of praiseworthiness. In the evaluation of praiseworthiness, an object's conformity is categorized along one of the three dimensions: obligatory, normally expected, and optional, and a deviation is qualitatively tagged as a violation of one of these three types of conformities. The deviation parameter is a measure of the degree of deviation once the deviation type is established. Wearing a pink tie is a deviation from normally expected norms, but wearing a tie two feet wide is more "devious" than wearing a pink tie, and wearing a tie that flashes lights is even more "devious"--except if one were a clown, in which case the norms of evaluations would be reversed.

Internally, this parameter is represented as follows:

```
(EMO-LOC-PARAM
                DEVIATION
                (OBJECT ?object)
                (VALUE  ?value))
```

with ?value taking the value HIGH, if the deviation is high, and LOW, if the deviation is low.

Tastes-based

We discussed in section 8.5 the parameter of appealingness which measured the degree an object conforms to one's likes and dislikes. An additional parameter is particularly relevant for this category, that of *familiarity*. We introduced in the previous section the global parameter of proximity, and more specifically that of *conceptual* proximity. Familiarity is one particular type of conceptual proximity parameter in that it measures the extent to which an object is recognizable as a previously encountered object. The hypothesis is here that the more familiar the object is, the more appealing it will be.

Internally, this parameter is represented as follows:

```
(EMO-LOC-PARAM
                FAMILIARITY
                (OBJECT ?object)
                (VALUE  ?value))
```

with ?value taking the value TRUE of ?object is familiar, and FALSE otherwise.

8.9 The Emotion State Model

We are now in a position to describe how we model an emotion state in our system. As we said in section 8.2, we distinguish between three types of emotion states: episodes, moods, and dispositions. In modeling an emotion state we first begin by determining whether we are talking about a particular instance of an emotional manifestation, whether the manifestation is a tendency over a period of time, or whether it is a general tendency of the person in question. The description of the emotional state is contained in the *emotion-state* slot of the *emotion-state-description*. In section 8.2 we gave a description of the model used for specifying emotion states.

In addition, an emotion state is also characterized by the emotion type in question. In the previous section we gave a list of the emotion types.

The third parameter in the specification of the emotion-state-description is the intensity of the emotion. We distinguish between three grades of emotion intensity: *intense, neutrally intense,* and *weakly intense.* Thus "fear" would be neutrally intense, "alarm" intense, whereas "apprehension" would be weakly intense.

The specific emotion is indicated by an English word and is specified in the *token* entry.

Thus the specification:

EMOTION-STATE-DESCRIPTION

<EMO-STATE> : {emotion-state}
<EMO-TYPE> : {emotion-type}
<INTENSITY> : {graded-relative-value}
<TOKEN> : {English word}

8.10 The Verbal Expression of Emotions

In this section we examine the issue of how emotional states are verbally expressed. We distinguish between two types of verbal emotional expressions: (1) those that are *explicit,* e.g., "I'm overjoyed," and (2) those that are *implicit,* i.e., not stated verbally by the person experiencing the emotion. Within explicit expressions we make the further distinction between (1) *explicit reports* and (2) *tokens of expression.* Implicit expressions are also divided into two subcategories: (1) those that are *structural,* e.g., distancing in time and space, and (2) those that are *semantic,* such as repetitions and slips.

Before we begin our discussion, however, we feel the need to make the following point clear: although we use the term "expression" to refer to verbal emotional behavior, we do not

imply that there is a *conscious* effort by the talker to express his emotional state. To imply that there is a conscious intention we would have to make the following assumptions: (1) the speaker is aware of his emotional state and knows *exactly* how he feels and (2) the speaker knows *exactly* how to express his emotional state verbally. These assumptions, as we stated in our previous discussions, we do not make. That is we allow for the speaker to be totally unaware of his true emotional state, and we allow for the speaker to be awkward in his expression, even when he knows "exactly" how he feels. This is, however, not to say that we preclude thoses instances where a speaker formulates an emotional expression intentionally, nor that we exclude instances of deceitful behavior. We allow an emotional expression to be either intentional or "impulsive."

8.10.1. Explicit Expressions

Explicit Reports

By "explicit reports" we mean those verbal statements where the person speaking describes his or another person's emotional state, e.g., "I feel great," "you seem to be angry," etc. The surface meaning of emotional state descriptions, like that of any other verbal statement, is not taken by our system to be necessarily an accurate description of the actual emotional state. That is, a speaker may state that he is feeling great, but in fact be feeling otherwise. Whether it is accurate in its depiction of the extant emotional state or wheter it is innacurate, nevertheless an explicit report in itself is relevant. In the case where the depiction is accurate, and nothing indicates that it is otherwise, the report is informationally relevant, i.e., we are told what the emotional state is. And in the case where the depiction is not accurate, the incongruity between the actual state and the described state may itself be a source of information, e.g., a person saying "I'm very pleased to hear that!" receiving bad news is displaying

not pleasure--the statement's surface meaning--but irritation, and the report of this irritation is partly based on the contrast between the nature of the news (bad) and the person's statement about his reaction to the news (as though the news were good).

Internally, an explicit report is represented as follows:

```
(EMO-EXP
        EXPLICIT
        EXPL-REP
        (REPORT     ?report)
        (UTTERANCE ?utterance))
```

with ?report being the report in question, and ?utterance the utterance in which the report occurs.

Expression Tokens

Expression tokens are those verbal statements which, although they do not explicitly depict the emotional state in effect, are nonetheless emotionally expressive, e.g., "Oh my God!" when expressing fright, "Oh-Oh" when expressing apprehension, "Wow!" when expressing awe, etc. As with explicit reports, a distinction must be drawn between the surface meaning of a token and its actual meaning. As we shall see in the example of the next section, expression tokens are frequently used by speakers.

Internally, an expression token is represented as follows:

```
(EMO-EXP
        EXPLICIT
        EXP-REP
        (EXPRESSION ?expression)
        (UTTERANCE  ?utterance))
```

8.10.2. Implicit Expressions

The second category of emotional expressions are those we call *implicit* expressions.[46] Whereas in an explicit expression a statement about the speaker's emotional state is explicitly formulated (whether the statement is accurate or not is another issue) in the case of implicit expressions, the emotional state is inferred indirectly from the statement. We distinguish between three types of inferences: (1) those that are based on the *structure* of statements, (2) those that are based on the *meaning* of statements, and (3) those that are based on the ongoing conversational process. We examine each of these three types of emotional expressions in what follows.

I. Structural

By structural expressions we mean those verbal statements which, through their structure, are expressive of the emotional state of the speaker. We identify nine structural devices.

1. Spatial Separation

A speaker may express his positive or negative feelings towards the object of his emotion by talking about the object as though it were spatially near or far from the speaker, regardless of its actual spatial proximity. It is generally believed that a positive emotion will induce a near reference to the object, whereas when a negative emotion is felt the object is referred to as though it were far. Two types of spatial separation devices are identitfied:

- a - *Demonstratives:* using "this" instead of "that" is usually indicative of positive feelings towards the object indicated.
- b - *Place Reference:* using "here" is considered more positively expressive than using "there."

---

[46] The Classification given here is based on the taxonomies given in [COLL85] and [MEHR81].

Internally, the representation is as follows:

```
(EMO-EXP
        IMPLICIT
        SPAT-SEP
        (TYPE       ?type)
        (UTTERANCE ?utterance))
```

with type being either DEMONSTRATIVE or PLACE-REF.

2. Temporal Separation

Emotional states can also be inferred from the tense used by the speaker. It is held that a speaker will use the past tense, when it is not necessary for him to do so, instead of the present tense in order to distance himself from something he evaluates negatively. At the lower end of the evaluation scale is the past perfect, preceded by the past, preceded by the perfect, preceded by the present. That is, "I am playing cards with Mary" is less negative than "I have been playing cards with Mary," which is less negative "I played cards with Mary," which is less negative than "I had been playing cards with Mary."

Internally, the representation is as follows:

```
(EMO-EXP
        IMPLICIT
        TEMP-SEP
        (TYPE       ?type)
        (UTTERANCE ?utterance))
```

with temp taking one of the following values: PRESENT/PAST/PAST-PERF.

3. Over-inclusion

A statement is said to be over-inclusive if the speaker uttering the statement attributes an event or a state of affairs to an agent that is more inclusive than the actual agent. For example, Jack may say, "everybody hates Larry," when in fact he means that *he* hates Larry.

Internally, the representation is as follows:

```
(EMO-EXP
        IMPLICIT
        OVER-INCL
        (UTTERANCE ?utterance))
```

4. Under-inclusion

In many instances, a speaker may express his negative evaluation of an object by restricting his positive evaluation of that object to a specific feature or aspect the object in question has. For example, Jack may say "I liked the scenery" when asked "did you like the movie?" His being over-specific in his statement may be indicative of his negative evaluation of the movie in general. Notice that the statement "I liked the scenery" is simultaneously explicit and implicit. It is explicit since the speaker reports his liking of the "scenery," and implicit since the restriction to the movie's scenery may indicate that the movie overall was evaluated negatively.

Internally, the representation is as follows:

```
(EMO-EXP
        IMPLICIT
        UNDER-INCL
        (UTTERANCE ?utterance))
```

5. Negation

In many instances, the presence of a negative statement may be used to infer the speaker's emotional state. For example, Jack may express his lack of enthusiasm about the movie by saying "it wasn't bad" when asked whether he liked the movie or not. By using the negative "it wasn't bad" he is describing his evaluation in a much more general way than with a more specific statement such as "it was good."

Internally, the representation is as follows:

```
(EMO-EXP
        IMPLICIT
        NEGATION
        (UTTERANCE ?utterance))
```

6. Linguistic Positioning:

In general, it is considered that objects evaluated positively are mentioned earlier than objects evaluated negatively [SANF88], [SANF87].

Internally, the representation is as follows:

```
(EMO-EXP
        IMPLICIT
        POSITIONING
        (POSITION  ?position)
        (UTTERANCE ?utterance))
```

with ?position taking one of the following three values: BEGIN/MIDDLE/END, depending on whether the mention of the object evaluated is at the beginning, middle, or end of ?utterance, respectively.

7. Speaker-Object Relationships

An evaluation can also be inferred by examining how the object of focus is related to the speaker. We distinguish between two types of Speaker-Object relationships.

- a - *Attribution to an External Agent:* for example, "I have to go" instead of "I want to go."
- b - *Using the Passive:* for example, "She congratulated me," instead of "I was congratulated by her."

Internally, the representation is as follows:

```
(EMO-EXP
        IMPLICIT
        SPK-OBJ-REL
        (TYPE      ?type)
        (UTTERANCE ?utterance))
```

with ?type taking one of the following two values: EXT, for attribution to an external agent case, and PASS, when the passive is used.

8. Modification

Feelings of uncertainty and confidence can be expressed by introducing in the verbal formulations certain verbal particles that modify an unqualified statement and stress the speaker's level of confidence in his evaluations. We distinguish between two types of modifications:

- a - *Qualification:* Uncertainty can be conveyed with statements such as "I guess," "I believe," "I think," "it could be that," etc.

- b - *Objectification:* Certainty, or at least a desire by the speaker to convey it, can be expressed by statements such as "it is obvious," "it is evident," "it's a fact that," etc.

- c - *Linguistically Unnecessary Insertions* Frequently used are intra-statement particles such as "just," "you know," "you understand," and the frequently used "uh"'s and "ah"'s, which seem to convey a feeling of doubt and hesitation.

Internally, the representation is as follows:

```
(EMO-EXP
        IMPLICIT
        MODIF
        (TYPE      ?type)
        (UTTERANCE ?utterance))
```

with ?type taking one of the following two values: QUAL, OBJ, LING, for cases a, b, c above, respectively.

II. Semantic

We now turn to the second category of implicit emotional statements, those that are semantically based. Whereas structural statements were inferred from the form of the uttered

statements, in the case of semantic statements, the inferences are based on the *meaning* of what is said.

1. Frequency: Whether it is evaluated positively or negatively, an object mentioned frequently is one towards which the speaker feels an intense emotion. Thus, one may infer that Jack loves (or hates) Mary from the tendency the former has to bring up the subject of Mary every time he can.

Internally, the representation is as follows:

```
(EMO-EXP
        IMPLICIT
        FREQ
        (VALUE      ?value)
        (UTTERANCE ?utterance))
```

with ?value taking one of the following three values: OVER, for overly frequent mentionings, UNDER, for less than usual frequency of mentioning, and AVG, for average frequency of mentioning.

2. Retention: Remembering details about certain objects while forgetting important information about other objects is an indication that the former are considered important while the latter are not.

Internally, the representation is as follows:

```
(EMO-EXP
        IMPLICIT
        RETENTION
        (VALUE      ?value)
        (UTTERANCE ?utterance))
```

with ?value taking one of the following three values: HIGH, for the case where the retention exhibited is higher than expected, LOW, for low retention, and NORMAL, for retention considered "normal."

3. Freudian Slips: Not infrequently, people use one word when they mean to use another word, and do this without even realizing that they have done so. The "true" emotions that one

has about the object in focus, or the object brought into focus by the slip, may sometimes be accurately inferred from the slip.

Internally, the representation is as follows:

```
(EMO-EXP
        IMPLICIT
        FREUD
        (USED       ?used)
        (INTENDED   ?intended)
        (UTTERANCE  ?utterance))
```

with ?used being the word used and ?intended the word that was intended.

III. Conversational

The third source for inferring emotional states is conversational behavior. We saw in Chapter Five that a conversation is a highly regulated activity, and that a protocol of exchange is closely followed by conversants. Emotional states can be inferred from behavior that does not conform to the expectations established by the protocol. We identify five types of conversational statements:

1. ***Silences:*** Silences may be indicative of mild reactions such as hesitation or reluctance, or strong reactions such as shock or amazement.

Internally, the representation is as follows:

```
(EMO-EXP
        IMPLICIT
        SILENCE
        (VALUE ?value))
```

with ?value taking the value LONG, SHORT, or NORMAL, depending on whether the silence is considered to be long, short, or "normal."

2. Interruptions: An interruption may indicate that the speaker interrupting is eager to say something, whether it is a case of the speaker being unable to contain his joy or his anger, or whether the speaker feels that the information he has to convey is of relevance or urgency.

Internally, the representation is as follows:

```
(EMO-EXP
          IMPLICIT
          INTERRUPT
          (VALUE    ?value))
```

with ?value being TRUE if an interruption is in effect, and FALSE otherwise.

3. Turn Retention: When a turn is retained over intervals of time longer than expected of the speaker in question, then it may be inferred that the speaker is eager to talk about the object in focus. Similarly, turn lengths shorter than expected of the speaker may indicate the speaker's desire to avoid the object of focus.

Internally, the representation is as follows:

```
(EMO-EXP
          IMPLICIT
          TURN-RET
          (VALUE ?value))
```

with ?value taking the value LONG, SHORT, or NORMAL, depending on whether the turn retention is considered to be long, short, or "normal."

4. Openings: We saw in Chapter Five that opening sequences for conversations are highly regular patterns. Deviations from those patterns may be a source for drawing inferences about speaker emotional states. An opening without any salutations may signal, depending on the context, hositility, anger, or amicability.

Internally, the representation is as follows:

```
(EMO-EXP
          IMPLICIT
          OPEN
          (VALUE ?value))
```

with ?value taking the value NORMAL or ABRUPT, depending on whether a "normal" opening sequence is used in the opening or whether the opening is effected without an opening sequence.

*5. **Closings:*** Closings also may be a source for drawing inferences. A speaker abruptly closing a conversation, without the presence of any external interference forcing the termination, may be displaying anger or agitation.

Internally, the representation is as follows:

```
(EMO-EXP
        IMPLICIT
        CLOSE
        (VALUE ?value))
```

with ?value taking the value NORMAL or ABRUPT, depending on whether a "normal" closing sequence is used in the closing or whether the conversation is closed without a closing sequence.

8.11 Example

The interactive situation we have been studying throughout is the travel agent customer telephone dialogue, where a customer calls a travel agent to arrange a flight or request some information. At the end of Chapter Six, we gave a detailed description of the functions upon which the respective roles of "customer" and "agent" are based. We also dicussed the norms that constituted each of these roles. In addition, we argued in Chapter Seven that each interactant involved in an interaction has face wants and sensitivities, and that many of the norms of conduct tend directly to those wants and sensitivities. The functions of the various roles an interactant takes on during an interaction are the basis upon which the goal tree will be built. The norms dictated by the role, on the other hand, including those pertaining to face

wants and sensitivities inherited from the role, form the standards base. Thus, even before the conversation begins, an interactant already possesses a well defined goal tree and a substantial standards base.

On the other hand, the tastes base is not available before the interaction begins--unless the interactants know something about one another. As the conversation progresses, however, and more is revealed to each interactant about the other interactant, the taste bases will progressively expand.

With the three valence contexts established, events and states of affairs can be evaluated in terms of their effects on the states of the valence contexts. With this, a model of the system's emotional state as well as a model of the system's partner (as perceived by the system) can be constructed.

Chapter Six, and more specifically section 6.9, contains a detailed discussion on the various goals each of the interactants has as he or she plays their respective roles. And Chapter Seven, section 7.8, contains a discussion on face oriented expectations (i.e., standards about interactional behavior).

In what follows we analyze a few conversation fragments along the emotion dimension.

The following are segments from conversation phone.03. The segments are within the stretch of conversation where the agent and the customer are discussing the latter's apprehension about flying the DC10, and flying in general (lines [454,619]).

```
453     .CUSTOMER.
454        [Okay,] I have a few questions.
455     .AGENT.
456        [Sure].
457     .CUSTOMER.
458        The [ah] flight from Chicago to L.A.
459     .AGENT.
460        [Uh-huh].
461     .CUSTOMER.
462        What kind of plane is that [do you know?]
463     .AGENT.
464        That's a DC10 at the moment.
465     .CUSTOMER.
466        [Oh].
467        I hate to say this.
468     .AGENT.
469        That's [about] the safest plane around after all the
470        checkups they've had.
```

471 .CUSTOMER.

The customer begins with a hesitant question, qualifying the question with the "ah" particle in line 458. In line 462, the question is further qualified with the post-tag "do you know." In addition, the customer's apprehension about the plane type is expressed with the "that" demonstrative, which refers to the plane type in question. The agent also distances herself from the plane type with the "that" demonstrative in line 464. In addition, she uses the "at the moment" qualifier, further indicating her hesitation.

In line 466, we have an instance of an expression token with "Oh," which expresses the customer's surprise. The surprise is qualified in line 467 with the explicit report "I hate to say this." The agent interrupts with a defensive statement about the reliability of the DC10, but nevertheless, still uses the "that" demonstrative, and furthemore qualifies her statement with "about."

It may be inferred that in this segment of the conversation, the customer's episodic emotional state is dominated by the emotion of apprehension, a mild form of fear. The fear in question has as object the plane make, and more specifically, the reliability of the DC10.

The agent's response to the agent's apprehension is defensive, and we can clearly see that she was expecting a negative reaction from the customer (i.e., she was dreading) after she told him that the plane type is DC10.

Further in the conversation, the customer reveals that this is the first time he is flying.

494 .CUSTOMER.
495 [Right].
496 [Well] this is [like] the first time I'm flying.
497 .AGENT.
498 [Oh] it is?
499 .CUSTOMER.
500 [Yeah].
501 I'm hesitant about the DC10 since I've heard so
502 much about them.
503 .AGENT.
504 [Well] it's quite possible that the equipment has
505 changed again ...

The disclosure is stated in a hesitant manner, with "well" and "like" used before the statement proper is formulated in lines 495 and 496. The agent responds with surprise, with the "oh"

expression token in line 498. The customer continues with an explicit report, stating that he feels "hesitant about the DC10." The agent insists on defending the DC10's reliability, but nevertheless still does qualify her statements with "well" and "it's quite possible."

Apprehension still dominates the customer's episodic emotional state along the goals valence dimension. In addition, however, the customer also displays his concern for not wanting to be perceived as being unnecessarily fearful. That is, the customer displays a desire to justify his fear. This, we argue, he does for two reasons: (1) a desire to protect his positive face, i.e., not behave like a coward, and (2) a desire to protect the agent's positive face by taking the trouble to explain the basis of the apprehension. The first concern thus springs from a mild form of shame, i.e., being ashamed of one's fears. The second concern is based on the customer's desire to alleviate the agent's feeling shame from perceiving the customer to be reproaching him. This the customer does by decreasing the responsibility of the agent (who is also, technically speaking, the agent of the blame) and shifting the blame to the object, i.e., the DC10.

In the third segment below, the agent tells the customer about an incident she had when she was learning to drive, aiming by this to alleviate the customer's fears.

563 .AGENT.
564 [You know] when I first, when I first started to
565 learn how to drive it was a disaster.
566 I was all trembled ...
567 .CUSTOMER.
568 [Yeah].
569 .AGENT.
570 and fidgety and my instructor nearly hit the
571 ceiling, poor guy.

The agent begins with an explicit report, "I was all trembled" in line 566. The emotion described may safely be inferred to be "fear." The agent continues her statement in line 570, describing in addition another person's emotional state, i.e., that of her instructor. The instructor's state is not clearly described, but we may infer from the description given about his behavior that it must have been an "anger" type emotion, e.g., irritation. The anger episodic state in question has the following parameters: (1) the undesirable object to the instructor is

the "non-cooperative" behavior--i.e., fearful behavior--of the student, (2) the degree of the blameworthiness of the student's fears, (3) the degree of identification between the student with the fearful behavior and the blamed object (here identical), and (4) the degree of deviation of the student's fears from expected fearful reactions students usually have.

In addition, notice the "poor guy" used by the agent in line 571. The insertion may have been motivated by the agent's desire to avoid giving the customer the impression that she considers the instructor's reaction blameworthy, or that she harbors any feelings of resentment or indignation.

8.12 Summary

We presented in this chapter a model of human emotions. A model of emotional states is needed if the emotion function that meta-statements fulfill is to be taken into account. The model we have adopted is cognitive. That is, we hold that an emotion is based on beliefs and evaluations. More specifically, we consider an emotion to be an evaluation of its intentional object. An emotion's intentional object, we hold, is either an event, a state of affairs, or a nil object. The evaluation of an event or state of affairs is based on the impact that event or a state of affairs has on what we call valence contexts. We identify three valence contexts: goals, standards, and tastes. An emotion along the goals valence context is an evaluation of an event or state of affairs in terms of whether the event or state of affairs furthers or hinders the various goals that are in effect. An emotion along the standards dimension is an evaluation of an event or state of affairs in terms of whether the event or state of affairs conforms to or deviates from expected norms and standards. And similarly, along the tastes dimension, an event or state of affairs is evaluated in terms of whether the event or state of affairs conforms to or deviates from the tastes and likes of the person making the evaluation.

In addition, we make the distinction between three classes of emotional states: episodes, moods, and dispositions. Thus, an emotional state will be characterized by the following characteristics: (1) its class, (2) its object, and (3) its valence dimension.

With this model, we are able to model at least in part emotional states and behaviors during conversations. Sensitivity to emotional manifestations will enable our system to better understand utterances in general and put the meanings conveyed by other conversers in an affective context. In addition, our system can exhibit emotional behavior and in this way behave in a manner that is similar to that of a human converser. This makes the system not only more "natural sounding," but also more interesting to interact with.

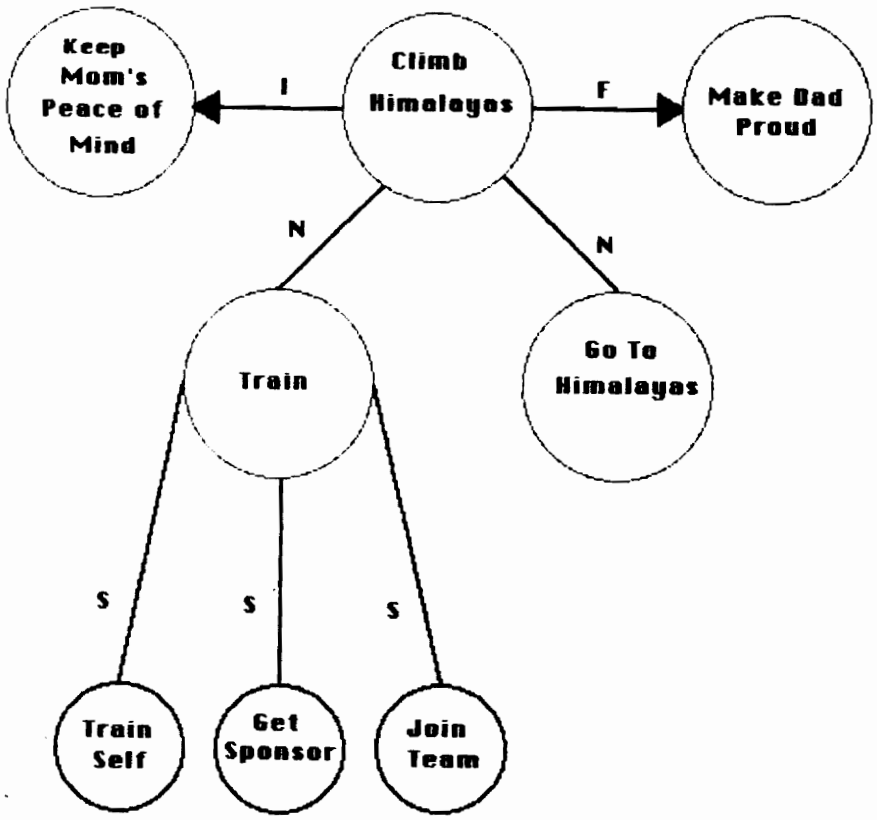


Figure 18. Goal Tree Structure for CLIMBING-HIMALAYAS goal

Chapter 9. The Cognitive Function

We now turn to the third "talk about talkers" function of meta-statements, the *cognitive* function. Two important aspects of talkers that we must model are their state of knowledge and their state of understanding: what do the conversers know, and to what extent do they understand what is being said to them. This chapter focuses on how meta-statements are used by talkers to describe and inquire about states of knowledge and understanding during conversations.

The first two sections focus on modeling states of knowledge. In section One, we discuss the problem of *reference,* i.e., what a speaker is referring to. We make the distinction between two types of references: (1) definite references and (2) indefinite references. Each type is discussed in detail.

In section Two, we discuss modeling the *factual* state of conversers, i.e., the knowledge base that they possess. We distinguish between two types of facts that are of interest to us: (1) facts about the vocabulary used in the talk, and (2) facts concerning the semantic domain talked about.

Sections Three and Four focus on modeling states of understanding. Section three contains a discussion on the relationship between context and understanding. For example, in many cases, a person may fail to understand an utterance not because the person does not possess the necessary vocabulary or because the utterance is ambiguous, but instead because the person is not able to relate the utterance to the established context. The issue of relevance is thus examined in this section.

The fourth section contains a discussion on the relationship between understanding and *how* information is presented. We will discuss the issues of rhetorical structuring, complexity, and assumed background information. Section Five contains a specification of the state model, while section Six contains a specification of the action model. Section Seven contains examples illustrating the model adopted.

But before we begin our discussion, let us first examine the following conversation fragment in terms of the cognitive function of the meta-statements used.

Example 9.1: [47]

CUSTOMER.
Would I be served dinner on there?
AGENT.
[Um] a dinner and snack, [right.]
CUSTOMER.
[Dinner and snack,] [okay.]
[Ah] what time would they be leaving?
AGENT.
[Well,] the night flight out of Kennedy leaves at 11
p.m. arriving L.A. at 1:35 a.m.
CUSTOMER.
1:35 a.m. arrival.
[Okay,] in [um] L.A. or?
AGENT.
1:35 a.m. in L.A.
CUSTOMER.

[47] From conversation phone.01, lines 119-135.

135 In L.A., [okay.]

....

The customer begins with a question on whether a dinner will be served during the flight. In terms of its cognitive aspect, the utterance is a request for confirmation or denial: the customer is trying to assess his own cognitive state, i.e., whether the information he possesses is true or false. The agent answers with a confirmation, specifying that in addition to dinner a snack will be served. The agent's confirmation is further emphasized by the meta-statement "right." The customer replies by echoing the agent's utterance, post-fixing his utterance with an "okay," indicating thus his own understanding. This is then followed by another question on the departure time, to which the agent replies in lines (127) and (128) by furnishing the necessary information. The customer again echoes the agent's utterance in line (130), indicating his understanding of the agent's utterance and post-fixing his utterance with "okay." In line (131), the customer again requests the agent's confirmation, this time asking whether L.A. is the arrival location the agent was referring to. In line (133), the agent confirms the customer's statements, and in line (135), the customer echoes the agent's confirmation, post-fixing the confirmation with "okay."

From this brief excerpt, we can see that conversers are indeed concerned with their own and their partner's states of knowing and understanding. In what follows, we discuss in more detail the issues involved in modeling the cognitive state of speakers.

9.1 Referential

In this section, we examine the first aspect of conversers' knowledge state, their knowledge of what is being *referred to.* We make the distinction between two types of references: (1)

definite references and (2) *indefinite* references. In what follows we examine each of these two types.[48]

Definite References

By definite references we mean those references whose object is some individual referent. In "London is a big city," for example, the object referred to is a particular city, viz London. We distinguish between three types of definite references:

1. Proper Names

As in the example given above, a term that "uniquely" identifies the referent is used.

2. Pronouns

Pronouns such as "I," "this," "he," etc., may also be used to refer to some particular referent. Obviously, reference with pronouns is heavily context dependent.

3. Definite Descriptions

Expressions such as "the man standing to your right," may also be used to refer to a particular referent. In this case, the identity of the referent is indicated through a description of the referent.

---

[48] The following discussion is based on [THRA80].

Indefinite References

Indefinite references, on the other hand, have no particular referent. We distinguish between three types of indefinite references:

1. Quantified Expressions

Expressions such as "every man," "any woman," "a man," are expressions that refer to sets of entities rather than any particular entity.

2. Indefinite Specifications

Pronouns such as "something," "many," "few," are also indefinite references.

Indefinite Descriptions

Expressions such as "three dogs," "a blue bird," "man" are also indefinite in their reference. In this case, reference is made to a generic object, or many generic objects, rather than a specific referent.

9.2 Factual

The second aspect of speakers' knowledge states that we need to model is their state of factual knowledge: what do they know. We distinguish between two types of knowledge important for our model: (1) *vocabulary knowledge* and (2) *semantic knowledge.*

By vocabulary knowledge we mean knowledge about the words and phrases used during talk. A listener may, for example, fail to understand the person he is talking with because a word or an expression the latter used does not belong to the listener's vocabulary. In such cases, the listener might ask the speaker to define the unknown word or expression for him.

Semantic knowledge, on the other hand, refers to knowledge about facts. In conversational situations where information is exchanged between the conversers, such as our agent-customer dialogues, the evaluation of semantic knowledge states is by far the main cognitive concern of the conversers. Information requests, offers, confirmations, and denials are all actions performed with the aim of assessing semantic knowledge states, and, as we will see in the examples discussed below, the conversers in our transcripts spend considerable effort describing their own knowledge state and evaluating the state of their partner.

We make the distinction between two types of semantic knowledge: *propositional* knowledge and *non propositional* knowledge.

By propositional knowledge we mean information that can be qualified as being either true or false. The statement "the capital of England is London," for example, is said to be propositional, since we can evaluate its truth value to be true.

Non propositional knowledge, on the other hand, is knowledge that cannot be evaluated as either true or false. Statements that begin with "suppose," for example, are said to be non propositional. The supposition, "suppose I went with you," for example, cannot be qualified as either true or false. Suggestions also are non propositional.

9.3 Contextual

In addition to modeling the knowledge state of speakers, we also need to model speakers' state of understanding. Knowledge (at least in our use of the term in this context) refers to static information, e.g., propositions. Understanding, on the other hand, refers to the *dynamic processing of information:* how information is related to current information.

Incoming information is evaluated in our system with respect to a pre-existing context in terms of its *relevance:* how relevant is the incoming information. In cases where the incoming information cannot be related to the existing context, then the information is said to be *irrelevant.* In the case where the information can be related to the context, then the information is said to be *relevant.* [49]

Incoming information is said to be irrelevant if one of the following three cases holds:

1 - *The incoming information cannot be readily connected to any information existing within the current context,* i.e., a converser says something "out of the blue."[50]

2 - *The incoming information already exists in the context.* (Again, a distinction must be made between the irrelevance of the information and the irrelevance of the act of giving the irrelevant information. It may be, for example, that the speaker wants to stress a point (see Chapter Eight), and thus, although the information given is irrelevant, the act of giving that information is relevant.)

3 - *The incoming information is not consistent with the context, but has little support to override the present, contradicting information.*

Information that impacts and modifies the context into which it is introduced is on the other hand said to be relevant. Relevance, however, is a matter of degree. That is, not all

---

[49] The following discussion is based on [SPER86].

[50] Note that although the information itself may not be relevant, its utterance need not be irrelevant, i.e., the utterance of something "out of the blue" may be significative, e.g., the speaker wants to change the subject.

statements are equally relevant within a given context. Two factors influence the evaluation of the degree of relevance of a statement: (1) the contextual *effects* resulting from the introduction of the information contained in the statement, and (2) the *processing effort* required for processing the information. To illustrate, let us give a few examples.

Suppose we have the following context:

(1) Jack is married to Mary.

(2) Mary wants to raise children.

(3) Because of illness, Mary cannot bear children.

(4) Couples who want to raise children but can't bear them should adopt a child.

Now suppose we introduce the information:

(5) Jack wants to raise children.

Given the context (1)-(4) and the new information contained in (5), the context can be modified to further include:

(6) Jack and Mary should adopt a child.

Statement (6) is obtained from (1) (i.e., that Jack and Mary form a couple), the conjunction of (2) and (5) (thus "Jack and Mary both want to raise a child), and an instantiation of (4), with "couples" subsitituted with "Jack and Mary."

Suppose now that instead of (5) above, we have the statement:

(5a) Mary has black hair.

The statement is obviously less relevant than the original statement (5). Although it is not totally irrelevant (i.e., it is about Mary, who is mentioned in the context), it has nonetheless little effect on the context. The information contained in the statement is merely added to the context.

Now suppose that instead of (5) above, we have the statement "Jack wants to raise children, and today is the twenty fourth of May." Intuitively, the statement, "Jack wants to raise children" by itself is more relevant than the conjunction "Jack wants to raise children, and today is the twenty fourth of May." Although both statements modify the context in the

same way, the effort for processing the second statement is greater than the effort needed for processing the first statement, thus making the first statement more relevant than the second.

In general, we hold the following:

1 - The greater the contextual effects of incoming information, the more relevant the information is.

2 - The less the effort for processing incoming information, the more relevant the information is.

9.4 Presentational

In addition to modeling incoming information relative to an established context, we must also model *how* that incoming information is presented, since the manner with which meaning is presented is also important in determining whether or not a listener understands what is being said to him. For example, if the speaker structures his information badly, if he overwhelms his listener with facts, or assumes that the listener knows more about the subject he is discussing than the listener actually knows, then the listener may as a result have difficulty understanding what the speaker is telling him.

We distinguish between three areas of concern: (1) rhetorical structuring, (2) complexity, and (3) background information assumed.

1. Rhetorical Structuring

By rhetorical structuring we mean the use of explicit meta-statements in demarcating rhetorical structures in talk. Rhetorical meta-statements such as "in other words," "on the other hand," "first of all," serve to clarify how one part of the talk is to be related to another. In

certain cases, the relationship between parts of talk is obvious, so that such meta-statements are not necessary for understanding. In other cases, however, these meta-statements are so crucial that their absence may cause a loss in coherence, consequently making understanding more difficult.[51]

2. Complexity

The complexity of a presentation is also an important factor in determining whether or not a listener understands what the speaker says. We distinguish between two factors for evaluating complexity. *Retention requirements* refers to the amount of information the listener is required to keep in mind in order to understand, e.g., long lists, deep digressions, etc. The greater the amount, obviously, the harder it will be for a listener to understand.

The second factor influencing the complexity of a presentation is *the grammatical form* of the sentences uttered by the speaker. Long sentences with many internal digressions and sidetracks, for example, will strain the listener and make it hard for him to understand.

3. Background Information Assumed

Obviously, whether or not a listener understands what is being said depends greatly on what the listener knows, i.e., what background information he possesses. If the speaker assumes that the listener knows more than what the listener actually knows, then the listener may have difficulty understanding the speaker. In such a case the listener might request that the speaker speak more simply or explain more fully what they are talking about.

---

[51] See Chapter Four for a discussion on rhetorical structures.

There is also the case where the speaker assumes the listener to know less than he actually knows. In such a case, the listener might request that the speaker talk without going into unnecessary explanations.

We now turn to the specification of the cognitive state model we will use for our system.

9.5 State Model

As we said previously, we distinguish between two types of cognitive states: (1) the state of knowledge and (2) the state of understanding. In evaluating a converser's state of knowledge we are determining whether or not the converser possesses facts, i.e., knowledge. In evaluating a state of understanding, on the other hand, we are examining the process through which the converser relates incoming information to a body of pre-existing knowledge. Thus, in our model we use two parameters to model a converser's cognitive state: (1) the *knowledge* parameter and (2) the *understanding* parameter. Each of these two parameters can take one of two values: true or false. The value true for the knowledge parameter means that the converser in question *knows* the fact in focus, whereas the value *false* means that the converser does not know the fact. The true value for the understanding parameter means that the converser in question is able to relate the incoming information with the present information, whereas the value false means that the converser is not able to relate the information with the context. In what follows we first specify the parameters that determine the value of the knowledge state of a converser, and then we specify the parameters that determine the value of a converser's understanding state.

Knowledge Parameters

There are two types of knowledge parameters: (1) *reference parameters* and (2) *factual parameters.*

Reference parameters model a converser's knowledge of what a term refers to. Internally, the reference parameter is modeled with the following frame:

```
(COG-PARAM
      (TYPE      REF-DEF)
      (PERSON    ?person)
      (STATE     ?state)
      (REF-TERM  ?ref-term)
      (OBJECT    ?object))
```

with ?state taking one of four values. The following three values represent the case where converser ?person is not able to relate the reference term used in the description of ?object (i.e., ?object in such a case is NIL): PROP-NAME, if the reference term is a proper name; PRONOUN if the reference term is a pronoun; and DEF-DESC, if the reference term is a definite description. The fourth value, SAT, is for the case where ?person is able to relate the reference term to what is being referred to (i.e., ?object has a non nil value).

In the case where the reference term is indefinite, then the frame is as follows:

```
(COG-PARAM
      (TYPE      REF-IND)
      (PERSON    ?person)
      (STATE     ?state)
      (REF-TERM  ?ref-term)
      (OBJECT    ?object))
```

with ?state taking one of four values. Again, the first three values represent the case where converser ?person is not able to relate the reference term used in the description of ?object: QUANT-EXP, if the reference term is a quantified expression; PRONOUN, if the reference term is a pronoun; and INDEF-DESC, if the reference term is an indefinite description. The fourth

value, SAT, is for the case where ?person is able to relate the reference term to what is being referred to.

Factual parameters, the second type of knowledge parameters, are evaluations of information with respect to the knowledge base.

To model the vocabulary knowledge, we must consult the converser's knowledge base of words. When an incoming word is absent from the converser's knowledge base we raise a flag signalling the absence. Internally, the absence is represented by:

```
(COG-PARAM
      (TYPE   WORD-MISSING)
      (WORD   ?word)
      (PERSON ?person))
```

Context Parameters

Incoming information that cannot be related to the context is flagged as irrelevant, and the type of irrelevence indicated. As we said in section 9.3, we distinguish between three types of irrelevances: (1) those due to a lack of connection between the incoming information and the context, (2) those due to the pre-existence of the incoming information in the context, and (3) those due to a lack of support to the incoming information, when that information contradicts information existing in the context. If the incoming data is considered not relevant, the following frame is used to model its non relevance:

```
(COG-PARAM
       (TYPE     NON-RELEVANCE)
       (PERSON   ?person)
       (STATE    ?state)
       (DATA     ?data))
```

with ?state taking the value NO-CONNECT, if the data is not relevant because it cannot be related to the current context; PRE-EXISTS, if the incoming data is already present in the knowledge base; and NO-SUPPORT, if the incoming data has nothing to support its validity.

In the case where the information is relevant, we make the distinction between two types of relevances: (1) *weak* relevances, where the incoming information is simply added to the knowledge base, with no further modifications to the knowledge base, and (2) *strong* relevances, where the incoming information causes, in addition to adding the incoming information, further modifications to the context. Internally, the types are represented as follows:

```
(COG-PARAM
      (TYPE     RELEVANCE)
      (PERSON   ?person)
      (STATE    ?state)
      (DATA     ?data))
```

with ?state taking the value WEAK if the data is weakly relevant; and STRONG, if it is strongly relevant.

Presentational Parameters

Three parameters measure the evaluation of the presentation. *Rhetorical structuring* is a measure of whether or not the presentation is *coherent*, i.e., whether the rhetorical structures can be easily identified and related to one another. Internally, a parameter is used to indicate the state of coherence:

```
(COG-PARAM
      (TYPE   COHERENCE-STATE)
      (PERSON ?person)
      (STATE  ?state))
```

The second measure is that of *Complexity*. We distinguish between two types of complexities: (1) retention requirements and (2) grammatical forms. Internally we have the following representation:

```
(COG-PARAM
      (TYPE       COMPLEXITY)
      (PERSON     ?person)
      (STATE      ?state)
      (UTTERANCE ?utterance))
```

with ?state taking the value RETENTION if the complexity is due to excessive demands on retention by ?person, and GRAMM-FORM, if the grammatical forms used cause ?person difficulty to understand. ?utterance is the utterance that is focused on.

Finally, the *background information assumed* parameter is a measure of how much background information is assumed by a converser. We distinguish between two cases: (1) the case where too much information is assumed, and (2) the case where too little information is assumed. Internally, we use the following representations:

```
(COG-PARAM
      (TYPE   BACKGROUND)
      (PERSON ?person)
      (STATE  ?state))
```

with ?state taking the value TOO-LITTLE if ?person is assumed (by other speaker) to have too little background information (i.e., ?person knows more than other speaker thinks he knows), and value TOO-MUCH if ?person is assumed to have too much background information (i.e., ?person knows less than other speaker thinks he knows).

9.6 *Action Model*

We specify in this section the action model. We begin first with a specification of the cognitive actions, and then specify the cognitive grammar.

Actions

The following are the cognitive actions included in our system.

1 - State Own Satisfactory Cognitive State: [52] The speaker informs the listener that he (the speaker) knows or understands what his partner is talking about. This action is usually performed with meta-statements such as "uh-huh" and "right" while sending a back-channel signal (see Chapter Five). The action is also often performed by restating or echoing something the other person had said, thus implicitly checking own understanding.

2 - State Own Non Satisfactory Cognitive State: The speaker informs the listener that he (the speaker) does not know or does not understand what his partner is talking about. The action is usually performed with a request for clarification or repetition.

3 - Check Other's Satisfactory Cognitive State: The speaker checks whether the listener's cognitive state is satisfactory. The speaker may "perform" the action implicitly, and instead of explicitly requesting that the listener describe his state (e.g., "Do you understand?"), the speaker may infer from what the speaker says whether the latter understood or not. When descriptions of cognitive states are explicitly requested they are usually performed during a TRP (see Chapter Five), either verbally with questions such as "right?" or "okay?" or non verbally with trail offs. The action may also be performed with an offer to restate or re-explain something said previously, thus indirectly requesting from the listener a statement about their cognitive state.

4 - Check Own Satisfactory Cognitive State: The speaker checks whether his cognitive state is satisfactory. Usually, this action is performed by echoing the listener in the interrogative, e.g., restating in the form of a question a listener's statement, or rewording a speaker's previous statement.

5 - State Other's Satisfactory Cognitive State: The speaker informs the listener that the listener knows or understands what the speaker is talking about. Usually, the action is performed with meta-statements such as "right," "that's right," or "that's correct."

---

[52] "Cognitive state" refers to either a state of "knowing" or a state of "understanding." That is, the meta-statement: "I know that the capital of England is London" is a "State Own Satisfactory Cognitive State" along the kowledge dimension, while the meta-statement: "I understand your explanation" is a "State Own Satisfactory Cognitive State" along the understanding dimension.

6 - State Other's Non Satisfactory Cognitive State: The speaker informs the listener that the listener does not know or does not understand what the speaker is talking about. The action may be performed after an explicit request by the listener for a confirmation, or after the speaker infers from the listener's previous statements that the latter does not understand the speaker.

7 - Provide meaning: The speaker gives information or requests information.

The actions are thus:

1 - State Own Satisfactory Cognitive State.
2 - State Own Non Satisfactory Cognitive State.
3 - Check Other's Satisfactory Cognitive State.
4 - Check Own Satisfactory Cognitive State.
5 - State Other's Satisfactory Cognitive State.
6 - State Other's Non Satisfactory Cognitive State.
7 - Provide meaning.

Next, we give the grammar for parsing sequences of the actions given above. As usual, we use *A* and *B* to designate the speaker performing the action, with the action performed indicated by the subscript to the right of the letter.

Grammar

The Cognitive grammar follows:

S --->

A_7S_1
| B_7S_4

S_1 --->

A_1S_1
| A_3S_2
| A_7S_1
| B_1S_1
| B_2S_1
| B_4S_3
| B_5S_1

$| B_6S_1$
$| B_7S_4$

S_2 --->

B_1S_3
$| B_2S_1$
$| B_4S_3$

S_3 --->

A_5S_1
$| A_6S_1$
$| A_7S_1$

S_4 --->

B_1S_4
$| B_3S_5$
$| B_7S_7$
$| A_1S_4$
$| A_2S_1$
$| A_4S_6$
$| A_5S_4$
$| A_6S_1$
$| A_7S_1$

S_5 --->

A_1S_6
$| A_2S_1$
$| A_4S_6$

S_6 --->

B_5S_4
$| B_6S_4$
$| B_7S_4$

9.7 Examples

We begin this section by examining a fragment of conversation in terms of the cognitive actions perfomed within it. First we describe the actions, and then we give a symbolic transcription of the actions and show how the sequence of the actions is parsed by the grammar.

Example 9.2: [53]

AGENT.
[Good morning, this is United Airlines
Reservations, Mrs. Stark.]
CUSTOMER.
[Yes,]
[I need some information,] [please.]
I would like to [um] ask, get some information about
taking a flight leaving April 5th from Roanoke to
Los Angeles.
AGENT.
[Okay,]
[and] returning?
CUSTOMER.
[Um] a week later.
The following week, [I believe] that would be the
12th.
AGENT.
[The 12th.]
[And] how many are traveling?
CUSTOMER.
[Just] 1.
AGENT.
[Okay,] [I'll be glad to check and see what's
available.]
Are you able to depart anytime on those days?
CUSTOMER.
[Ah] yes.
AGENT.
[Okay.]
CUSTOMER.
[Preferably] the morning.
AGENT.
[Okay,] [fine.]
I'll check and see what we have available as well
as check the rates for you.
[Please hold.]
CUSTOMER.
[Thank you.]
AGENT.
[Thank you for holding now.]
CUSTOMER.
[Thank you.]
AGENT.
[Okay,] [I'll be glad] to go over our schedule if
you'd like to jot it down and tell you what we
have available.
CUSTOMER.
[All right.]
AGENT.
[Okay,] on the 5th of April [...]

---

[53] Conversation Phone.04 lines 1-65.

51 CUSTOMER.
52 [Uh-huh.]
53 AGENT.
54 from Roanoke to Chicago we have Piedmont flight
55 301 out of Roanoke.
56 CUSTOMER.
57 [That's Piedmont] [...]
58 AGENT.
59 [Piedmont 301] [...]
60 CUSTOMER.
61 [Uh-huh.]
62 AGENT.
63 which departs at 7:15 a.m.
64 CUSTOMER.
65 [Uh-huh.]
.....

We have in lines [2,3] a provide-meaning action, more specifically, a provide-information action. The customer responds with another provide-information action in lines [5,9]. In line (11) the agent responds with a statement of own understanding, and in line (12), perform another provide-meaning action, this time requesting information. In lines [14,15] the agent provides information, while the agent in line (18) echoes the customer's information, by this indicating her own understanding. She then again asks for information in line (19), which the customer provides in line (21).

In line (23) the agent signals that she has received enough information to be able to "see what's available." She then requests that the customer confirm or disconfirm her information in line (25). The customer confirms the agent's statement in line (27), and the agent responds with an "okay" in line 29, indicating her own understanding. The customer then provides additional information in line (31), to which the agent responds with "Okay, fine," indicating her own understanding.

A conversational pause follows from line (36) to line (44). In line (50) the agent finally begins giving the customer the information the latter is seeking. In line (52) the customer responds to the agent's information with an "uh'huh," indicating own understanding. The agent continues giving information in lines (54) and (55). The customer responds in line 57 by requesting a confirmation "That's Piedmont," which is signalled by a pause (indicated by "..."). The agent confirms in line (59) the information, and the customer responds in line (61) that he

understood. The agent then gives more in formation in line (63) and the customer responds with a confirmation in line (65).

The sequence of actions is thus:

```
1       AGENT.
2-3         (GIVE-INFO) [A7]
4       CUSTOMER.
5           (STATE-OWN-SAT) [B1]
6-9         (GIVE-INFO) [B7]
10      AGENT.
11          (STATE-OWN-SAT) [A1]
12          (REQUEST-INFO) [A7]
13      CUSTOMER.
14-16       (GIVE-INFO) [B7]
17      AGENT.
18          (STATE-OWN-SAT) [A1]
19          (REQUEST-INFO) [A7]
20      CUSTOMER.
21          (GIVE-INFO) [B7]
22      AGENT.
23-24       (STATE-OWN-SAT) [A1]
25          (CHECK-OWN-SAT) [A4]
26      CUSTOMER.
27          (STATE-OTHER-SAT) [B5]
28      AGENT.
29          (STATE-OWN-SAT) [A1]
30      CUSTOMER.
31          (GIVE-INFO) [B7]
32      AGENT.
33          (STATE-OWN-SAT) [A1]
34-36       (GIVE-INFO) [A7]
37      CUSTOMER.
38          (STATE-OWN-SAT) [B1]
39      AGENT.
40          (GIVE-INFO) [A7]
41      CUSTOMER.
42          (STATE-OWN-SAT) [B1]
43      AGENT.
44-46       (GIVE-INFO) [A7]
47      CUSTOMER.
48          (STATE-OWN-SAT) [B1]
49      AGENT.
50          (GIVE-INFO) [A7]
51      CUSTOMER.
52          (STATE-OWN-SAT) [B1]
53      AGENT.
54-55       (GIVE-INFO) [A7]
56      CUSTOMER.
57          (CHECK-OWN-SAT) [B4]
58      AGENT.
59          (STATE-OTHER-SAT) [A5]
60      CUSTOMER.
61          (STATE-OWN-SAT) [B1]
62      AGENT.
```

63 (GIVE-INFO) [A7]
64 CUSTOMER.
65 (STATE-OWN-SAT) [B1]

The sequence of actions is thus: [A7 B1 B7 A1 A7 B7 A1 A7 B7 A1 A4 B5 A1 B7 A1 A7 B1 A7 B1 A7 B1 A7 B1 A7 B4 A5 B1 A7 B1]. The grammar given above parses it as follows:

$S ==> A_7S_1$
$S ==> A_7B_1S_1$
$S ==> A_7B_1B_7S_4$
$S ==> A_7B_1B_7A_1S_4$
$S ==> A_7B_1B_7A_1A_7S_1$
$S ==> A_7B_1B_7A_1A_7B_7S_4$
$S ==> A_7B_1B_7A_1A_7B_7A_1S_4$
$S ==> A_7B_1B_7A_1A_7B_7A_1A_7S_1$
$S ==> A_7B_1B_7A_1A_7B_7A_1A_7B_7S_4$
$S ==> A_7B_1B_7A_1A_7B_7A_1A_7B_7A_1S_4$
$S ==> A_7B_1B_7A_1A_7B_7A_1A_7B_7A_1A_4S_1$
$S ==> A_7B_1B_7A_1A_7B_7A_1A_7B_7A_1A_4B_5S_1$
$S ==> A_7B_1B_7A_1A_7B_7A_1A_7B_7A_1A_4B_5A_1S_1$
$S ==> A_7B_1B_7A_1A_7B_7A_1A_7B_7A_1A_4B_5A_1B_7S_4$
$S ==> A_7B_1B_7A_1A_7B_7A_1A_7B_7A_1A_4B_5A_1B_7A_1S_4$
$S ==> A_7B_1B_7A_1A_7B_7A_1A_7B_7A_1A_4B_5A_1B_7A_1A_7S_1$
$S ==> A_7B_1B_7A_1A_7B_7A_1A_7B_7A_1A_4B_5A_1B_7A_1A_7B_1S_1$
$S ==> A_7B_1B_7A_1A_7B_7A_1A_7B_7A_1A_4B_5A_1B_7A_1A_7B_1A_7S_1$
$S ==> A_7B_1B_7A_1A_7B_7A_1A_7B_7A_1A_4B_5A_1B_7A_1A_7B_1A_7B_1S_1$
$S ==> A_7B_1B_7A_1A_7B_7A_1A_7B_7A_1A_4B_5A_1B_7A_1A_7B_1A_7B_1A_7S_1$
$S ==> A_7B_1B_7A_1A_7B_7A_1A_7B_7A_1A_4B_5A_1B_7A_1A_7B_1A_7B_1A_7B_1S_1$
$S ==> A_7B_1B_7A_1A_7B_7A_1A_7B_7A_1A_4B_5A_1B_7A_1A_7B_1A_7B_1A_7B_1A_7S_1$
$S ==> A_7B_1B_7A_1A_7B_7A_1A_7B_7A_1A_4B_5A_1B_7A_1A_7B_1A_7B_1A_7B_1A_7B_1S_1$
$S ==> A_7B_1B_7A_1A_7B_7A_1A_7B_7A_1A_4B_5A_1B_7A_1A_7B_1A_7B_1A_7B_1A_7B_1A_7S_1$
$S ==> A_7B_1B_7A_1A_7B_7A_1A_7B_7A_1A_4B_5A_1B_7A_1A_7B_1A_7B_1A_7B_1A_7B_4S_3$
$S ==> A_7B_1B_7A_1A_7B_7A_1A_7B_7A_1A_4B_5A_1B_7A_1A_7B_1A_7B_1A_7B_1A_7B_4A_5S_1$
$S ==> A_7B_1B_7A_1A_7B_7A_1A_7B_7A_1A_4B_5A_1B_7A_1A_7B_1A_7B_1A_7B_1A_7B_4A_5B_1S_1$
$S ==> A_7B_1B_7A_1A_7B_7A_1A_7B_7A_1A_4B_5A_1B_7A_1A_7B_1A_7B_1A_7B_1A_7B_4A_5B_1A_7S_1$
$S ==> A_7B_1B_7A_1A_7B_7A_1A_7B_7A_1A_4B_5A_1B_7A_1A_7B_1A_7B_1A_7B_1A_7B_4A_5B_1A_7B_1S_1$

We now turn to the issue of contextual relevance. Relevance in our context can be easily evaluated, since the semantic context is quite precise. We gave at the end of Chapter Six a description of the functions and norms of the roles enacted by agents and customers. It is these functions and norms that form the basis upon which relevance is evaluated.

We did not encounter any instances in our transcripts where a converser introduced in the talk information that was totally irrelevant, or irrelevant to the point where the other converser was unable to relate the incoming information to the context. There were many instances of "weak" relevance, however, where the information given was semantically irrelevant. For example, there were cases where a converser reworded something obvious that

the other converser said, or inferred something that was "trivial" or "inconsequential." However, such semantically "irrelevant" statements were most often of relevance along another functional dimension, e.g., politeness, rhetorical structuring, conversational organization, etc.

We also did not encounter many instances of reference ambiguity. Only seldom did either the agent or the customer explicitly request that their partner disambiguate a reference. The fragment below shows such an instance:

Example 9.3: [54]

66 CUSTOMER.
67 [Okay.]
68 [Um] can you tell me if there is a meal served on
69 that flight?
70 AGENT.
71 A what?
72 CUSTOMER.
73 A meal?
74 AGENT.
75 Which one of the two?
76 CUSTOMER.
77 [Um] the flight leaving Roanoke at 11, at 1:15
78 arriving in New York at 2:18 and leaving at 4?

In this instance, confusion arose because the agent and the customer had been discussing more than one flight, and the customer did not use reference terms that were specific enough to enable the agent to identify which of the two flights he was referring to. Thus in line (75), the agent requests that a clarification be made on the identity of the flight. The customer responds by providing a description of the flight in terms of its departure location (Roanoke), its arrival location (New York), and its departure and arrival times.

[54] Lines 66-78 of conversation Phone.06.

9.8 Summary

We discussed in this chapter the cognitive function of meta-statements. We argued that meta-statements are used by conversers to describe their own cognitive state as well as inquire about the cognitive state of their partner. We distinguished between two aspects of conversers' cognitive states relevant to our model: (1) the state of knowing, and (2) the state of understanding. For our purposes, two types of knowledge are important: (1) knowledge about the words used in talk, and (2) knowledge about the semantic domain talked about. In the case of understanding, we made the distinction between two aspects of talk that influence the degree of understanding of conversers: (1) the relevance of the incoming information to a context, and (2) the manner with which that information is presented.

In addition, we specified the cognitive state model, where we identified the various internal parameters, and the cognitive action model, where the cognitive actions and the cognitive grammar were presented.

In the next chapter we specify the fourth and last "talk about talkers" function, the communicational function.

Chapter 10. The Communicational Function

We now turn to the fourth and last "talk about talkers" function of meta-statements: the *communicational* function. We saw in the previous chapter that during conversations, conversers often describe and inquire about each other's cognitive states. Another important concern that conversers have is one about the *physical* state of the exchange. During telephone conversations, it is not infrequent when a speaker cannot hear what their partner said. In such cases, the speaker may ask his partner to repeat what he had said using formulations such as "what?" "I can't hear you," "can you repeat that?" etc. These formulations are meta-statements, and their function is mainly to bring into attention the communicational state of the exchange. In this chapter, we focus on how conversers use meta-statements to describe and assess their own as well as their partner's communicational state.

We begin by examining a fragment of conversation illustrating the communicational function of meta-statements. In section one we specify the state model, while in section two the action model is described. We end the chapter with an example illustrating how the model is used.

Example 10.1: [55]

58 AGENT.
59 Piedmont 301 [...]
60 CUSTOMER.
61 [Uh-huh.]
62 AGENT.
63 which departs at 7:15 a.m.
64 CUSTOMER.
65 [Uh-huh.]
66 AGENT.
67 [and] arrives in Chicago with breakfast at 8:50,
68 8:43 p.m., a.m., [excuse me.]
69 CUSTOMER.
70 [8:43 a.m.?]
71 AGENT.
72 [Uh-huh.]
73 CUSTOMER.
74 [And] that's with breakfast?
75 AGENT.
76 [Pardon me?]
77 CUSTOMER.
78 Is that with breakfast?
79 AGENT.
80 With breakfast.
81 CUSTOMER.
82 [All right.]

In addition to its back channel function[56] the "uh-huh" messages in lines (61) and (65) also fulfill a communicational function. In both cases, the customer is informing the agent that he is able to receive messages satisfactorily. In line (70), the agent restates the arrival time in the form of a question ("8:43 a.m.?"). Cognitively, as we argued in the previous chapter, the statement is a "check own understanding" action. Communicationally, however, the customer is trying to determine whether he "heard right," i.e., whether the message he received is the message the agent sent. The agent responds that the message received by the customer is

---

[55] Lines 59-82 of Phone.04

[56] See Chapter Five for a discussion on back channel messages.

the message originally sent with the statement "uh-huh" in line (72). In line (74), the customer asks whether a breakfast will be served on the flight, but the agent apparently does not receive the message satisfactorily, and in line (76) she asks the customer to repeat what he had said. The customer repeats his question in line (78), and the agent confirms the information in line (80).

In addition to the statements identified above, evaluation of communicational states is also done implicitly. We argued in the previous chapter that conversers evaluate their partners' state of knowledge and understanding through the statements the latter make. If a question is asked, for example, and the listener's response is not "appropriate" (from the speaker's point of view), then a possible reason for the listener's inappropriate response is that he did not *understand* what was said. In the case of the communicative aspect of the exchange, implicit evaluations are even more prevalent. The mere fact that a listener does not stop the speaker and ask them to repeat their message is in itself a message, i.e., "Go ahead, I can hear you fine so far."

In what follows, we discuss some issues related to modeling the communicational state of conversers.

10.1 The State Model

To model the communicational state of conversers, we need to evaluate three parameters: (1) the state of the messages received, (2) the state of the transfer, and (3) the readiness of the conversers to give and receive messages.

1. Receipt of Messages

Conversers' main communicational concern during an exchange is to ensure that the messages they receive are the messages sent by their partners, and that the messages they send are the messages received by their partners. However, conversers do not follow every message they send with a request from their partners to repeat their message. Instead, they *assume* that the messages they send are received without distortion, and only when listeners request that messages be repeated (e.g., "I can't hear you"), or when they have reasons to believe that there was distortion do speakers become explicitly concerned with the integrity of the message sent. Similarly, listeners do not follow every receipt of a message by echoing the message received and requesting that the speaker confirm that the message they echoed corresponds with the message originally sent by the speaker. Instead, listeners assume that the messages they receive are the messages sent to them, and only when they are unable to receive a message (e.g., they can't hear), or when they have reason to believe that there was distortion of the message during the transfer do they explicitly examine the validity of the messages they received.[57]

There are two reasons why conversers communicate with *the assumption of integrity*, instead of using explicit evaluations. First, requesting explicit evaluations requires more effort than implying integrity. For example, a two message implicit exchange between speakers A and B would cost each speaker one statement:

A: **Send Message1.**
B: **Send Message2.**

---

[57] Note that we are limiting our discussion here and throughout the chapter to the communicational dimension. Therefore, a listener may very well request that the speaker repeat something the latter said, even though the listener has no doubts about the validity of the message received. Such actions, although on the surface communicational, are in fact actions along other metacommunicational dimensions, e.g., cognitive, as when the speaker says something "out of the blue," emotional, as when the listener reacts with incredulity, etc.

That is, speaker A sends his message and assumes that speaker B will receive it without distortion, and speaker B receives message1, assumes that it is the message sent by A intact, and sends his message, also assuming that A will receive it without distortions.

If, however, each speaker did not implicitly assume that their message would be received without distortion, then the exchange will be as follows:

A: **Send Message1.**
B: **Echo Message1.**
A: **Confirm Message1.**
B: **Send Message2.**
A: **Echo Message2.**
B: **Confirm Message2.**

That is, instead of two messages, the exchange now costs six messages.

The second reason why communication integrity is implicitly assumed, until indications to the contrary, is that a paradigm where speakers must be "certain" about the integrity of messages may cause communicationally intractable problems due to the nature of interactional communication. More specifically: if for each message a listener receives he must send a message indicating that the message was received, then communications will stall after the first message sent, i.e., A sends message M1, B sends response R1 (itself a message), indicating that M1 was received, to which A responds with R11 (another message), indicating that R11 was received, to which B responds with R111, etc. But this scenario is obviously avoided in natural conversations, and this because speakers assume the integrity of the communicational lines until they have reason to believe otherwise.

Thus we hold that during natural communication, each converser holds the following assumptions:

1 - *Each message sent is received by the listener without distortion.*
2 - *If a sender believes that a message he sent was distorted on its way to the listener, he will signal to the listener that the message was distorted.*
3 - *If a receiver believes that a message he received was distorted on its way to the receiver, he will signal to the sender that the message was distorted.*

Note that the above are *working assumptions* and *not* a model of beliefs or guidelines for behavior. That is, when two conversers communicate, their exchange protocol is based on the above three assumptions. If A sends a message to B, he will assume that the message he

sent will be received without distortion by B, and B will assume that the message he received is A's message undistorted. Should the message arrive to B distorted, it is up to B to signal the distortion, if he believes that there was one. If B does not perceive a distortion, or does not signal to A that there was one, then A will assume that B received the message A sent, or the message A belives B received. Similarly, if A believes that there was distortion in the transfer, then it is up to A to point to the distortion, since B will assume that the message he received is undistorted until he has reason to believe otherwise. Note that each converser is free to choose how he will respond to a message distortion, i.e., whether he will correct it or whether he will "overlook" it. In the following exchange, B decides to overlook (or rather, take advantage of) A's hearing "seventy" instead of the original "seventeen:"

A: How much do you want for that watch there?
B: Seventeen dollars, sir.
A: I'll give you sixty for it, and that's my final offer.
B: Sixty five and it's yours.

Now, should some other converser point to A that B originally said "seventeen" and not "seventy," then A can accuse B of communicational dishonesty, and more specifically of having violated assumption 2 of the exchange protocol. However, if no speaker C intervenes, and if following B's last utterance above, A responds by saying "Aha! I've caught you, you liar. You said 'seventeen,'" then B can also accuse A of having been communicationally dishonest, i.e., of having violated assumption 1 of the exchange protocol. That is, A on purpose acted as though he had received a distorted message, and in line (4) implicitly signalled to B that the message he received is not the one sent by B originally. This is a violation of assumption 1, and B may defend himself by pointing out that he had originally said "seventeen," and that it was A who intentionally distorted his message for his own entrapment purposes.

In our model, we keep a parameter that measures the receipt state of a converser with respect to a message. Internally, we have the following representation:

```
(RECEIPT-MESSAGE
        (MESSAGE   ?message)
        (STATE     ?state)
        (PERSON    ?person))
```

with ?message referring to the message in question, and ?state taking the value TRUE if the message is believed to have been received satisfactorily by ?person, and FALSE otherwise.

The state of a message is internally modeled as follows:

```
(MESSAGE-ENDED
        (STATE  ?state)
        (PERSON ?person))
```

with ?state taking the value TRUE if the message transfer has been completed, and FALSE otherwise. Note that we are talking here about a *message,* and not the whole transfer. This parameter therefore should not be confused with parameter TRANSFER-STATE.

2. State of Transfer

In addition to modeling the state of messages, we also need to model the *state of the transfer.* In telephone conversations, and more so in face-to-face exchanges, the state of the transfer is only occasionally verbally stated. But in communications such as those between two CB radio operators, messages are almost always begun and terminated with verbal statements, e.g., "calling station XYZ," or "over and out." Such statements are used by conversers to punctuate the beginning and end of their message transfers so that the transfer of messages can be effected without confusion with respect to whether or not the transfer of a message was completed.

Internally, we use the following representation:

```
(TRANSFER-STATE
          (STATE  ?state)
          (PERSON ?person))
```

The variable ?value can take one of the following three values: PRE-BEGIN, if the transfer is of messages is about to start (i.e., in the opening phase before the messages

proper are being sent); IN-PROGRESS, if the messages proper are currently being sent; and ENDED, if the exchange of messages is completed.

3. Converser Readiness

The third set of parameters relate information about conversers' readiness to give and receive information. We distinguish between three parameters: the READY-RECEIVE, the READY-GIVE, and the READY-BEGIN parameters. The READY-RECEIVE parameter models a converser's readiness to receive messages; the READY-GIVE parameter models a converser's readiness to give messages; while READY-BEGIN models a converser's readiness to begin the transfer of messages. Internally, we have the following representations:

```
(READY-RECEIVE
        (STATE  ?state)
        (PERSON ?person))
```

with the frame modeling the readiness of ?person to receive messages, with ?state being TRUE if ?person is ready, and FALSE otherwise.

An similarly for giving,

```
(READY-GIVE
        (STATE  ?state)
        (PERSON ?person))
```

where ?state taking the value TRUE if ?person is ready to give messages, and FALSE otherwise.

For readiness to begin the transfer, the following frame is used:

```
(READY-BEGIN
        (STATE  ?state)
        (PERSON ?person))
```

where ?state taking the value TRUE if ?person is ready to begin the transfer, and FALSE otherwise.

10.2 The Action Model

We now turn to a specification of the action model. We begin with the actions, and follow this with a specification of the grammar.

Actions

We distinguish between two types of actions: (1) signalling actions and (2) checking actions. By signalling actions we mean those actions where the converser describes his own communicative state. Checking actions, on the other hand, are those actions where the converser tries to assess the communicational state of his partner. Within each category, we identify three types of actions, each corresponding to one of the three parameters described in the previous section.

Signalling Actions

1. Receipt Actions

Receipt actions are actions where the converser states whether the message was received satisfactorily or non-satisfactorily. The actions are as follows:

1 - *Signal satisfactory receipt of message:* the converser signals that he has received the message sent to him. Most often, the signal is implicit, i.e., not having stated that the message was not received satisfactorily is a "statement" that the message was received satisfactorily. When explicitly stated, the signal is almost always in the form of a back channel. In its most verbally explicit form, the action may be performed with statements such as "I hear you," "I hear you loud and clear," "I read you," etc.

2 - *Signal non-satisfactory receipt of message:* the converser signals that he did not receive the message properly. Most often the signal is sent explicitly, as in the statement "I can't hear you," or in the form of a request, as in, "Can you repeat that?"

3 - *Signal other's satisfactory receipt of message:* the speaker signals to the listener that the message the latter received is the message intended by the speaker, e.g., "that's right," or "you heard right."

4 - *Signal other's non satisfactory receipt of message:* the speaker signals to the listener that the message the latter received is not the message intended by the speaker, e.g., "that's not what I said."

2. Readiness Actions

Readiness actions are those actions where the speaker describes whether or not he is ready to send or receive a message. The receiving actions are:

5 - *Signal readiness to receive message:* the speaker informs the listener that he is ready to receive the latter's message, e.g., "go ahead," "tell me about it," etc.

6 - *Signal non-readiness to receive message:* the speaker informs the listener that he is not ready to receive the latter's message, e.g., "can you hold," "one moment," etc.

And the sending actions are:

7 - *Signal readiness to give message:* the speaker informs the listener that he is ready to transfer his message to the listener, e.g., "I have something to tell you."

8 - *Signal non-readiness to give message:* the speaker informs the listener that he is not ready to transfer his message to the listener, e.g., "I'll tell you about it later."

3. *Transfer Actions*

Transfer actions are those actions where the speaker informs the listener of the state of the transfer. The actions are:

9 - *Signal readiness to begin transfer:* the speaker signals his readiness to begin transferring messages. This action is performed when conversations are opened or when returning from a pause.

10 - *Signal non readiness to begin transfer:* the speaker signals to his partner his non readiness to begin transferring messages. This action is performed in situations such as conversational pauses.

11 - *Signal that message transfer is done:* the speaker signals to his partner that the transfer of his message is completed. When verbally performed, statements such as "Right?" "Okay?" are used.

12 - *Signal that message transfer is not done:* the speaker signals to his partner that the transfer of his message is not completed. When verbally performed, statements such as "I'm not done yet," "Let me finish," are used.

Checking Actions

The second type of actions are actions where the speaker is trying to assess the communicational state of his partner. Again, we distinguish between three types of actions:

1. Receipt Actions

Receipt actions are those actions where the speaker is trying to determine whether he or his partner satisfactorily received the incoming messages. We identify two actions:

13 - *Check own satisfactory receipt of message:* the speaker tries to determine whether or not the message he received is the message sent to him by his partner. Most often, the action is performed by echoing the received information.

14 - *Check other's satisfactory receipt of message:* the speaker tries to determine whether or not the listener received or is able to receive his message, e.g., "can you hear me?"

2. Readiness Actions

Readiness actions are those actions where the speaker is trying to determine his partner's readiness to give or receive messages. The actions are:

15 - *Check other's readiness to receive message:* the speaker tries to determine whether or not the listener is ready to receive a message, e.g., "should I begin?"

16 - *Check other's readiness to give message:* the speaker tries to determine whether or not the listener is ready to give a message, e.g., "You want to tell me about it?"

3. Transfer Actions

Transfer actions are those actions where the speaker is trying to determine his partner's transfer state. We identify one action in this category:

17 - *Check whether other's transfer of message is done:* the speaker tries to determine whether his partner has finished transferring his message, "are you through?"

18 - *Check other's readiness to begin transfer:* the speaker tries to determine whether his partner is ready to begin the transfer of messages. The actions are thus:

1 - Signal satisfactory receipt of message.
2 - Signal non-satisfactory receipt of message.
3 - Signal readiness to receive message.
4 - Signal non-readiness to receive message.
5 - Signal readiness to give message.
6 - Signal non-readiness to give message.
7 - Signal readiness to begin transfer.
8 - Signal non readiness to begin transfer.
9 - Signal that message transfer is done.
10 - Signal that message transfer is not done.
11 - Signal other's satisfactory receipt of message.
12 - Signal other's non satisfactory receipt of message.
13 - Check own satisfactory receipt of message.
14 - Check other's satisfactory receipt of message.
15 - Check other's readiness to receive message.
16 - Check other's readiness to give message.
17 - Check whether other's transfer of message is done.
18 - Check other's readiness to begin transfer.

In addition, the following action is also identified, so as to facilitate the specification of the grammar. The action is not necessarily meta-communicational.

19 - Transfer message.

Grammar

The communicational grammar follows:

S --->

$A_{18}S_1$
| $B_{18}S_{18}$
| A_3S_2
| B_3S_2

S_1 --->

B_9S_2
| $B_{10}S$

S_2 --->

$A_{15}S_3$
| $A_{16}S_4$
| $A_{19}S_5$
| $B_{15}S_{10}$
| $B_{16}S_{11}$
| $B_{19}S_{12}$

S_3 --->

B_5S_8
| B_6S_2

S_4 --->

B_7S_{12}
| B_8S_2

S_5 --->

$A_{11}S_2$
| $A_{12}S_5$
| $A_{14}S_9$
| $A_{19}S_5$
| $B_{13}S_8$
| $B_{17}S_7$
| $B_{19}S_{12}$
| A_9S_6
| B_9S_{13}

S_6 --->

B_9S_{13}
| $B_{10}S_2$
| $A_{19}S_2$

S_7 --->

$A_{11}S_{17}$
| $A_{12}S_5$

S_8 --->

A_3S_5
| A_4S_5
| $A_{19}S_5$
| $B_{19}S_{12}$

S_9 --->

B_1S_5
| B_2S_5
| $A_{19}S_5$
| $B_{19}S_{12}$

S_{10} --->

A_5S_{15}
| A_6S_2

S_{11} --->

A_7S_5
| A_8S_2

S_{12} --->

$B_{11}S_2$
| $B_{12}S_{12}$
| $B_{14}S_{16}$
| $A_{13}S_{15}$
| $A_{17}S_{14}$
| $A_{19}S_5$
| $B_{19}S_{12}$

S_{13} --->

A_9S_6
| $A_{10}S_2$
| $A_{19}S_2$

S_{14} --->

$B_{11}S_{17}$
| $B_{12}S_{12}$

S_{15} --->

B_3S_{12}
| B_4S_{12}
| $A_{19}S_5$
| $B_{19}S_{12}$

S_{16} --->

A_1S_{12}
| A_2S_{12}
| $A_{19}S_5$
| $B_{19}S_{12}$

S_{17} --->

$A_{11}S_{17}$
| $A_{12}S_5$
| $A_{18}S_1$
| $B_{11}S_{17}$
| $B_{12}S_{12}$
| $B_{18}S_{18}$

S_{18} ---> $B_{19}S_{12}$

10.3 Example

We now examine the following short conversation in terms of the communicational function of its meta-communicational statements. We follow this by giving a full parse of the sequence of communicational actions performed within the conversation:

Example 10.2: [58]

AGENT.
[What can I help you with?]
CUSTOMER.
I would like to find out how much it cost to fly
from Roanoke to Pittsburgh.
AGENT.
When are you going?
CUSTOMER.
The 22nd.
AGENT.
When are you coming back?
CUSTOMER.
23rd.
AGENT.
Full fare - 180 round trip.
CUSTOMER.
$180.
Suppose I want to come back on Saturday
which is 24th.
AGENT.
If you buy your ticket 7 days before you leave,
it will be $130 round trip.
CUSTOMER.
$130.
If I buy my ticket one week ahead of time
without including Saturday that would make
any difference?
[Okay.]
[I call you back.]
Let me consult with my department here.
AGENT.
[Okay.]
[Good-bye.]
CUSTOMER.
[Bye.]

---

[58] Conversation Agency.14.

The conversation begins with an offer of help from the agent, line (2), which is also a communicational statement to the customer that the agent is ready to receive messages. The customer responds by directly stating their request, thus implicitly signalling readiness to give messages (lines [4-5]). The agent responds with a question, again implicitly acknowledging that the message was received (line (7)). From lines (9) to (15), the agent and the customer exchange messages with the implicit assumption by each that the other is receiving their messages. In line (17), the customer echoes the price "$180," and in lines (18) and (19) suggests an alternate return date. The agent's response in lines [21,22] is an implicit acknowledgment to the echoed $180 (i.e., since the agent did not correct the customer about the price, the customer will assume that the agent is confirming the $180 price).

Again the customer in line (24) echoes the quoted price, restates the agent's utterance, and procedes in lines [28,30] to end the message transfer. From lines (32) to (35), the agent and the customer exchange end of message transfer signals, and the transfer is closed.

The main point to note in the exchange is how the conversers exchange their messages with the implicit assumption that the other converser is able to receive their messages, and that they themselves are receiving undistorted messages. In two instances, the customer restated information given to him by the agent (lines (17) and (24)), but went on to make further statements (lines [18,19] after the echo in line (17), and lines [25,26] after the echo in line (24)), thus not altogether requiring the agent to confirm the customer's echoed information explicitly. The agent in neither case explicitly confirms the information, but goes on to address the customer's subsequent utterances. Given the implicit assumptions made by both conversers, however, the agent in both cases confirms the prices by not disconfirming them.

Below we give the conversation transcribed into the communicational actions it contains:

```
1       AGENT.
2         [SIGNAL READINESS TO RECEIVE]
3       CUSTOMER.
4-5       [TRANSFER MESSAGE]
6       AGENT.
7         [TRANSFER MESSAGE]
8       CUSTOMER.
9         [TRANSFER MESSAGE]
```

10 AGENT.
11 [TRANSFER MESSAGE]
12 CUSTOMER.
13 [TRANSFER MESSAGE]
14 AGENT.
15 [TRANSFER MESSAGE]
16 CUSTOMER.
17 [CHECK OWN SATISFACTORY RECEIPT OF MESSAGE]
18-19 [TRANSFER MESSAGE]
20 AGENT.
21-22 [TRANSFER MESSAGE]
23 CUSTOMER.
24 [CHECK OWN SATISFACTORY RECEIPT OF MESSAGE]
25-27 [TRANSFER MESSAGE]
28-30 [SIGNAL END OF MESSAGE TRANSFER]
31 AGENT.
32-33 [SIGNAL END OF MESSAGE TRANSFER]
34 CUSTOMER.
35 [SIGNAL END OF MESSAGE TRANSFER]

The above sequence is translated into: *[A3 B19 A19 B19 A19 B19 A19 B13 B19 A19 B13 B19 B9 A9 B9].*

And the grammar given above parses it as follows:

$S ==> A_3S_2$
$S ==> A_3B_{19}S_{12}$
$S ==> A_3B_{19}A_{19}S_5$
$S ==> A_3B_{19}A_{19}B_{19}S_{12}$
$S ==> A_3B_{19}A_{19}B_{19}A_{19}S_5$
$S ==> A_3B_{19}A_{19}B_{19}A_{19}B_{19}S_{12}$
$S ==> A_3B_{19}A_{19}B_{19}A_{19}B_{19}A_{19}S_5$
$S ==> A_3B_{19}A_{19}B_{19}A_{19}B_{19}A_{19}B_{13}S_8$
$S ==> A_3B_{19}A_{19}B_{19}A_{19}B_{19}A_{19}B_{13}B_{19}S_{12}$
$S ==> A_3B_{19}A_{19}B_{19}A_{19}B_{19}A_{19}B_{13}B_{19}A_{19}S_5$
$S ==> A_3B_{19}A_{19}B_{19}A_{19}B_{19}A_{19}B_{13}B_{19}A_{19}B_{13}S_8$
$S ==> A_3B_{19}A_{19}B_{19}A_{19}B_{19}A_{19}B_{13}B_{19}A_{19}B_{13}B_{19}S_{12}$
$S ==> A_3B_{19}A_{19}B_{19}A_{19}B_{19}A_{19}B_{13}B_{19}A_{19}B_{13}B_{19}B_9S_{13}$
$S ==> A_3B_{19}A_{19}B_{19}A_{19}B_{19}A_{19}B_{13}B_{19}A_{19}B_{13}B_{19}B_9A_9S_6$
$S ==> A_3B_{19}A_{19}B_{19}A_{19}B_{19}A_{19}B_{13}B_{19}A_{19}B_{13}B_{19}B_9A_9B_9S_{13}$

10.4 Summary

We examined in this chapter the fourth "talk about talkers" function of meta-statements, the communicational function. We argued that one of the concerns that conversers have when

they communicate is the physical aspect of the exchange. We identified three types of such concerns: (1) concerns about the receipt of information, e.g., "I can't hear you," (2) concerns about the readiness to receive or give information, e.g., "I want to tell you something," and (3) concerns about the transfer of messages, e.g., "are you finished talking?" We further distinguished between those actions where the speaker describes his own communicational state and those actions where the speaker checks his own or his partner's communicational state. We identified thus the actions and specified the grammar. We ended the chapter with an example illustrating the model.

Having at this point examined all six functions of meta-statements, in the next chapter we specify how in our system these six functions are integrated.

Chapter 11. Toward an Integrated System

Having discussed the separate models needed to specify the six functions of meta-statements, we proceed in this chapter with the specification of how these models are integrated into a dynamic conversational system. We begin first with a general overview of the components of the system and discuss the role that parameters, system actions, meta-communicational actions and grammars play in the system. In section 11.2, we examine the structure of the parsers and processing engines used by the system. The cognitive, communicational, rhetorical, and conversational parsers, having parsers with the same structure, are discussed first, while the emotions and politeness parsers, being composed of lower level engines, are discussed separately.

We will attempt to present the system's structure and dynamics without going into unnecessary details. The details of the system are fully given in the appendices in [BOUZ91].

This chapter is meant to give, on the one hand, a high level overview of the system, and at the same time serve as an explanation of the appendices. We begin with a discussion of the elements of the system.

11.1 The Components of the System

A. Parameters

As we argued in previous chapters, in order to model each of the six meta-communicational functions of meta-statements, we need to have present within the system a state model of the context relevant to each of the six functions. Thus, for example, in order to model the communicational function of a meta-statement, we must have a mechanism for modeling the communicational state of the converser: the meta-action "I can't hear you," for instance, is a statement describing the communicational state of the person uttering the statement, and therefore, to model what the statement *means* communicationally, we must have a model of the internal state of the converser. On hearing this statement, the listener might modify his communicational model of the speaker so that now in this model he believes the speaker not to have heard what was said to him. It is this modification of the internal model which we refer to as the action of "understanding," or more specifically in this case, understanding along the communicational dimension.

In Appendix A in [BOUZ91], we specify the parameters needed to model each of the six meta-communicational contexts. For the rhetorical contexts, in addition to the parameters we also specify the rhetorical connectors needed. Table 1 in the appendix specifies twenty three connectors. The second column of the table specifies the component class types. As we said in Chapter Four, there are two types of components classes: the nucleus-statellites

class, and the multi-nuclei class. NS is used to refer to the former, while MN refers to the latter. The third column specifies the parameters needed for the connectors. Only four connectors need parameters. The list connector is specified with the LENGTH and ITEM-NUMB parameters. The LENGTH parameter specifies the length of the list rhetorical unit opened, while the ITEM-NUMB gives the item number of the list nucleus being specified, e.g., in "fourth, ..." the nucleus unit following "fourth" is the fourth list nucleus being specified, and so the list parameter ITEM-NUMB is dynamically updated upon parsing the statement "fourth" to four.

In addition, the connector PLAN has as parameter PLAN-GOAL. PLAN-GOAL contains the goal for which the plan that is being specified was designed. Similarly, for the connector ENTITY, the entity under scrutiny is specified in the DESC-ENTITY parameter. The TOPIC connector has as parameter DESC-TOPIC. DESC-TOPIC contains information about the topic of the rhetorical unit currently open. The other connectors have no parameters.

The parameters for the conversation, politeness, emotion, cognitive, and communication state model are then specified. A parameter is specified as a frame. For example, the conversational parameter CON-PAR#1 is a frame with one entry, CONV-STATE, which has as its domain the following seven values: OPEN, MAIN, SUSPEND, PRECLOSE, CLOSING, CLOSED, and ABORTED (see Appendix A). At any point during the conversation, this parameter has its CONV-STATE field, its only field, set to one of these seven values.

For the politeness parameters, parameter POL-PAR#1 has as its third field the frame FTA. The values that FTA field can take are specified in Table 31 of Appendix F. For parameter POL-PAR#9, the values for its STRATEGY field are specified in Tables 32 and 33.

As we explained in Chapter Eight, the emotional state of a converser is modeled with parameter EMO-STATE. This parameter is a frame with four fields: PERSON refers to the person experiencing the emotion, STATE-ID is a tag used to identify uniquely the state being modeled, STATE contains information characteristic of the type of state the emotion is, while EMOTION-TYPE specifies one of the twenty two frames for modeling the emotion category. Appendix J contains a specification of the emotion types.

The STATE field is itself a frame. Three types of frames are used to model the state: the EPISODE frame, the MOOD frame, and the DISPOSITION frame. The field TRIGGER-TIME has as its domain points in time, LENGTH contains an interval in time, while EMOTION contains an English term for describing the emotion. The emotion terms used by our system are given in Appendix D.

In addition to the six types of parameters, we also use in our system parameters to model the roles enacted by the conversers. A role is modeled with the ROLE frame. As we argued, a role is specified by a function and a cluster of norms. A role's function is specified in the FUNCTION-ID field of the ROLE frame in question, while the norms are modeled in the NORM-IDS field, which contains a list the enacted norms' ids. An enacted norm is specified by a unique identification tag in NORM-ID, while a NORM-SPECS structure specifies the norm. The NORM-ID field specifies which of the ninety one norms the norm is. The CATEGORY field specified which of the five categories the norm belongs to. Five categories are specifies: SEEK-INFO-A refers to a norm that applies to the agent and which consists of the agent seeking information from the customer; SEEK-INFO-C is an information seeking norm that applies to the customer; GIVE-INFO-A is an information giving norm that applies to the agent; GIVE-INFO-C is an information giving norm that applies to the customer; while SUGGEST-INFO-A is a norm that applies to the agent and which consists in the agent making a suggestion to the customer.

The NORM-TYPE field specifies one of the three norm types: OB, stands for obligatory, NE, for normally expected, and OP, for optional.

A function is specified similarly by an identification tag in the FUNCTION-ID field, while in FUNCTION-TYPE the type of the function is specified. In Table 2 of Appendix B, the four functions used in our system are given.

In addition to modeling roles, the status of each customer is modeled with the frame STATUS. A status, as we argued in Chapter Six, is a cluster of roles enacted within a social group with respect to another actor. We have two statuses in our model. The first is the one

assumed by the agent, and which is composed of roles 1, 2, and 3. Status 2, on the other hand, is assumed by the customer, and it is composed of roles 4, 5, and 6.

B. System Actions

The parameters specified in Appendix A are subject to modifications as information about conversers and the conversation enters the system. Once incoming data has been examined and its impact on the various models determined, the relevant parameters reflecting the internal modifications must be updated. These updates are effected by *system actions.*

Along each of the six meta-communicational dimensions we specify a list of system actions. For example, system action 1 along the rhetorical dimension is: Create component unit. The effect of this action, once performed by the system, is to create a frame for modeling a component unit. This unit will have as an identification tag the "id" furnished by rhetorical system action 1.

Action 2 in the rhetorical dimension is "Set parameter in component unit." The action has as its parameters "id," "field," and "value." The effect of action 2 is thus to set the value of field "field" of the component with UNIT-ID "id" to "value." Thus, action 2(id,3,N) has the effect of setting the value of the CLASS field (the third field in the frame) of the component unit with identification tag id to "N." Actions 3 and 4 are analogous to actions 1 and 2, except that they apply to action units.

Action 5, on the other hand, applies to parameters RHE-PAR#1 and RHE-PAR#2. Action 5(1,1,id) sets the value of parameter RHE-PAR#1's first field to "id," while action 5(2,1,id) sets the value of parameter RHE-PAR#2's first field to "id."

The remaining system actions for the other meta-communicational dimensions function analogously. The only action that differs is the system action along the cognitive dimension. Here, the parameter affected by the action is not identified by its identification tag, but rather by the value contained in the TYPE field of the parameter frame. Thus the cognitive system

action 1(NON-RELEVANCE,3,NO-CONNECT) has the effect of setting the STATE field of parameter COG-PAR#4 (the parameter with type NON-RELEVANCE) to NO-CONNECT.

In Appendix G, we specify the effects that meta-communicational actions have on the internal state models by specifying the system-actions performed when the meta-actions are parsed. For example, in Table 37, the effect of conversational action 1 is 1(1,1,OPEN). That is, when it is determined by the parser that along the conversational dimension action 1 was performed, conversational system action 1 is performed, such that the first field of parameter 1 is set to the value "OPEN."

C. Meta-communicational Actions

As we argued in Chapter Three, a meta-statement may be parsed to more than one meta-communicational action. Therefore, when meta-communicationally parsing an utterance the output produced is not one action, but a list of actions. Appendix C gives a specification of how a meta-action is modeled. The third field of the frame modeling a meta-action contains a list of identification tags referring to the individual action frames of which the meta-action is composed.

A meta-communicational action is identified by one of eleven categories: six categories for rhetorical actions, and one for each of the remaining five meta-communicational categories. Tables 7 to 14 in Appendix C contain the actions identified by our system. In the third column of Tables 7, 8, 9, 11, 12, and 13 those parameters which are parsed along with the action are specified. For example, action PO 1, perform FTAs, has as a parameter the fta's performed. That is, when parsing, in addition to recognizing that action PO 1 is performed, the FTAs performed are also identified.

The conversational and communicational actions are without parameters.

D. Databases

We use within our system seven databases. We describe each in what follows.

1. Surface Interpretations Database

We have identified a list of *explicit* meta-communicational tokens. These tokens are characterized in two ways. First their *position* in the talk is identified. For example, the token "okay" when used at the beginning of a conversational turn has a meaning different from the meaning it has when it is in the middle of a turn or a sentence, or at the end of a turn. In Appendix E, the notation used for encoding a token's position is explained. Tables 24 to 30 in Appendix E contain the tokens actually identified in our corpus of twenty seven conversations.

2. FTA's Database

In addition, a database containing the types of FTA's recognized by our system is specified in Table 31 of Appendix A. The second column of the table contains a "decription" of the FTA identified in the first column. These descriptions obviously do not suffice by themselves to enable the characterization of an utterance in terms of the FTA's it contains. An engine with its own databases and structure must be built for this purpose. The FTA evaluator is this engine, and we will discuss it when we turn to the specification of the politeness parser.

3. Politeness Strategies Database

We also include a database of politeness strategies in our system. The strategies are given in Tables 32 and 33 of Appendix F. Similarly to the FTA's database, the "description" field is

not enough by itself. A politeness strategies evaluator must be built. We will discuss this evaluator when examining the politeness parser.

4. Emotion Terms Database

A list of emotion terms has been compiled in Tables 15 through 23 of Appendix D. The first column contains the English term, the second the category to which the emotion belongs (see Appendix J for a full specification of the categories), while the third column contains the intensity of the emotion.

5. Grammars Database

The grammars used for each of the six meta-communicational dimensions are specified in Appendix K. As we explained in Chapter Three, these grammars help us reduce the alternatives to which an action is to be mapped along one of the six dimensions. For example, after the action of opening a conversation, the *only* actions possible along the conversational dimension are: "accept open conversation move" or "reject open conversation move." Thus, when parsing, the conversational parser--i.e., the parser that is focusing on the conversational function of meta-statements--will need only to choose which of the two actions, CO 1 or CO 2, to map the action to, instead of having to search through the whole space of twenty two conversational actions. The notation used is explained in Appendix K.

6. Roles Database

A database containing the norms, functions, roles and statuses is specified in Appendix B. Table 2 contains the functions, Tables 3 and 4 the norms, Table 5 the roles, and Table 6 the statuses.

7. Actions' effects Database

The effects that meta-actions have on the system's internal parameters are given in Appendix G. The effects of the rhetorical actions are specified in Tables 34, 35, and 36; those of conversational actions in Table 37; the effects of the emotion actions in Table 38; while the effects of politeness, cognitive and communicational actions are specified in Tables 39, 40, and 41 respectively.

The effects of meta-actions are encoded as system actions and are given in the second column of the tables cited above (see the discussion on systems actions given in section B above). Certain actions have no effect, and so the second column in these cases are left blank.

11.2 The Structure of the Parsers

In this section, we discuss the structure of the parsers used in our system. First the rhetorical, conversational, cognitive, and communicational parsers are specified. In section two the emotions parser is discussed, while in section three the politeness parser is presented.

A. The Cognitive, Communicational, Conversational, and Rhetorical Parsers

The four parsers described in this section all have the same structure. The engines receive input from four sources: the utterance to be parsed, the action grammar for the function in question (e.g., the cognitive action grammar for the cognitive parser), the parameters modeling the function's state model, and the surface interpretations database (the latter being the same for all functions). The output of the parsers is the action along the function in question. Thus, the output of the cognitive parser is a cognitive function, that of the communicational parser the communicational action, etc. Figures 19 to 22 show the structure of the rhetorical, conversational, cognitive, and communicational parsers respectively. The internals of the parsers are discussed in more details in Appendix L.

B. The Emotions Parser

The emotions parser has a more complicated structure. The utterance that is to be parsed is not directly fed into the parser, but instead is first evaluated by four separate engines:

1. The Verbal Expression Parser:

The first engine is the *verbal expression parser,* which analyzes the structure and content of the utterances input to it in terms of their emotional content. This parser has as input the emotion terms database, the expression tokens database, the output of the implicit expressions evaluator and the conversational expressions evaluator. The output of the verbal expression parser consists in setting the appropriate values for emotion parameters 21 to 38 (i.e., the parameter used for modeling verbal emotional expressions). Figure 23 show the structure of this engine.

2. Goals Evaluator:

The second engine is the *goals evaluator.* This engine evaluates utterances purely in terms of their domain semantic content--i.e., in terms of data relating information about the goals and the plans of the person in focus. As we argued in Chapter Eight, one of the three valence contexts on which our model of emotions is based is the goals context. The goals evaluator is the engine the system uses to determine the impact that information contained in an utterance has on the goals, thus enabeling us to model emotions based on fulfillment or hinderance of goals. Figure 24 gives the structure of the goals evaluator.

3. Norms Evaluator:

Analogously, in order to model the second valence context, standards, a *norms evaluator* is used. This engine evaluates the information contained in an utterance in terms of its impact on expectations of behavior, i.e., norms. Obviously, the engine needs as input the norms that the conversational situation requires, and the content of the utterances parsed is compared with these norms. The engine is specified in more detail in Figure 25.

4. Tastes Evaluator:

Finally, the impact of an utterance on the third valence context, the tastes context, is computed by the *tastes evaluator.* Figure 26 gives the structure of this engine.

The output of these four engines consists in setting values in the appropriate emotions parameters. The parameters modified by each engine, along with a more detailed examination of these engines are give in Appendix M.

The emotions parser proper--that is, the engine that yields as output an emotion action--receives input from three sources: (1) the emotion parameters (once they have been set by the four engines described above), (2) the emotion actions grammar (used as discussed

previously to narrow the search space), and (3) the surface interpretation database. In case the emotion action output is either EM 1 or EM 2, then in addition to the action tag (i.e., EM 1 or EM 2), the system also determines the emotion in question (described by one of the English terms provided in Appendix D) and the state (e.g., one of the three frames EPISODE, MOOD, or DISPOSITION). Figure 27 shows the structure of the emotions parser.

C. *The Politeness Parser*

We now turn to the politeness parser. Similarly to the emotions parser, the politeness parser also does not receive the utterance directly as input. Instead, the utterance is first evaluated by two lower level engines: the FTA evaluator and the politeness strategy evaluator.

1. The FTA Evaluator:

This engine evaluates an utterance in terms of its FTA impacts. That is, the content and form of the utterance the engine receives as input are analyzed to determine the face impacts of the utterance. In addition to the utterance, the FTA evaluator has the following inputs. First is the database of FTA's specified in Appendix F, Table 31. It is to one of these FTA's that the face threats (if any) of the utterance are mapped. Second are the parameters used for modeling the roles involved in the situation. As we argued in Chapter Seven, the role an interactant enacts establishes (at least to a certain extent) negative as well as positive face wants. Thus, whether or not an action is an FTA depends on the roles being played.

Third are the politeness parameters, and more specifically, parameters 3, 4, and 5. In addition, the surface interpretation database is also another source of input for the FTA evaluator. Figure 28 shows the structure of this engine.

2. Politeness Strategies Evaluator:

The structure of the politeness strategies evaluator is similar to that of the FTA evaluator, except that instead of receiving input from the FTA database, the evaluator receives input from the politeness strategies database. This engine evaluates an utterance in terms of the politeness strategies (if any) it effects. Figure 29 shows the structure of this engine.

The politeness parser proper thus receives input from three separate sources (in addition to the utterance): (1) the politeness parameters affected by the FTA evaluator and the politeness strategies evaluator, (2) the politeness grammar, and (3) the surface interpretations database. The output of the utterance is one of the four politeness actions specified in Table 12. In case action PO 1 is performed, the face impacts are also parsed and determined (passed along from the FTA evaluator). In case the action is PO 2, PO 3, or PO 4, then the politeness strategy (passed along from the politeness strategies evaluator) is parsed and determined. The impact of the output politeness actions are given in Table 39, Appendix G. Figure 30 shows the structure of the politeness parser.

11.3 Summary

We gave in this chapter an overview presentation of the system's components and structure. The system's components are: (1) the parameters used for modeling the six meta-communicational state models, (2) the system actions which when performed modify the system parameters, (3) the meta-communicational actions to which verbal statements are parsed, and (4) the databases which contain information needed by the parsers to carry out the parsing.

In addition, we outlined the structure of the various engines needed by the system, and we specified their inputs and the outputs the system expects from them. We will not undertake to build all the various engines needed for each of the parsers outlined in this work, since their

development is beyond the scope of this work. Our task, as we said in Chapter Three, is to create a framework within which the pragmatic aspects of talk (in our case dialogue) can be modeled and used during interaction between computers and humans.

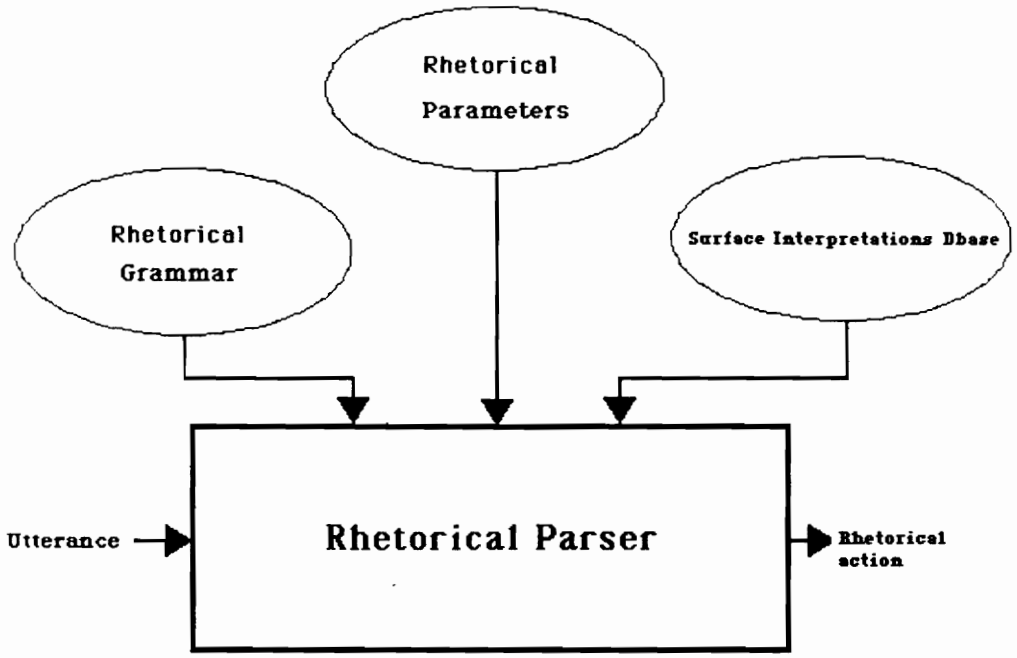


Figure 19. The Rhetorical Parser

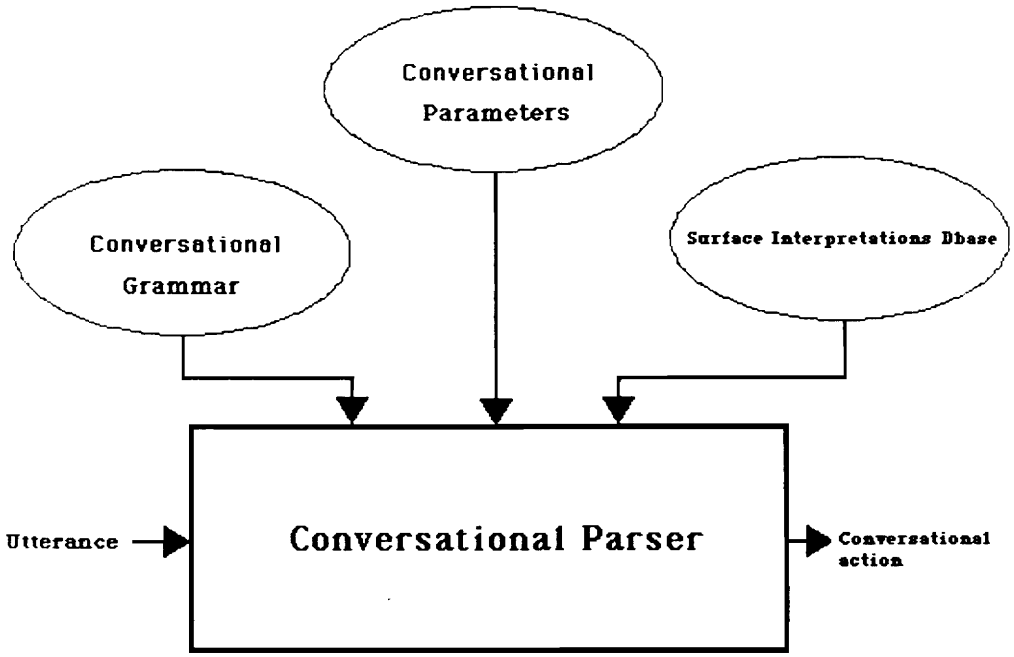


Figure 20. The Conversational Parser

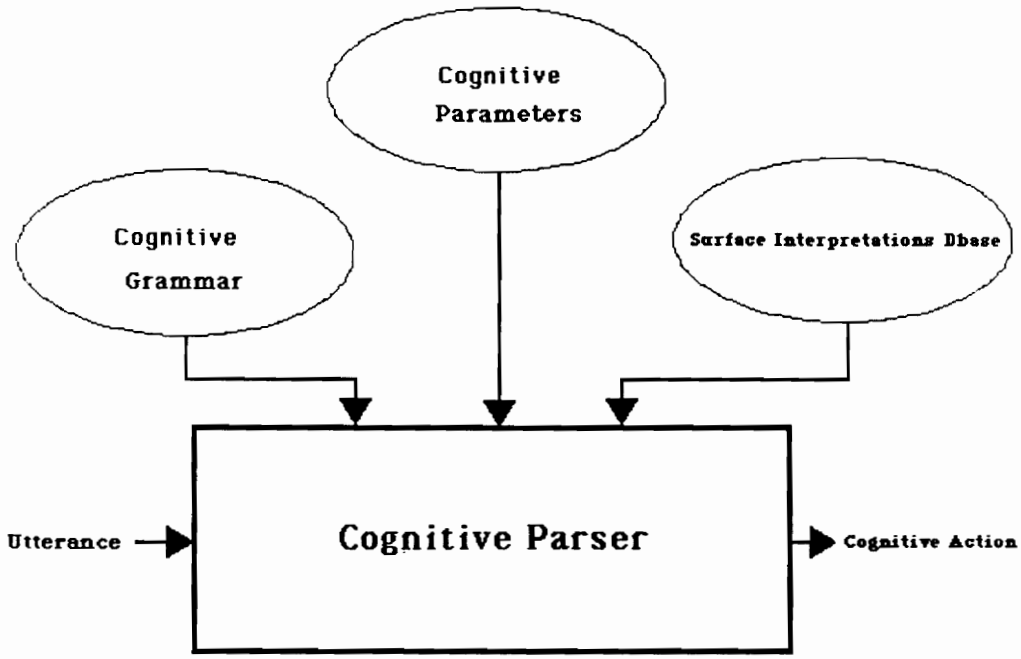


Figure 21. The Cognitive Parser

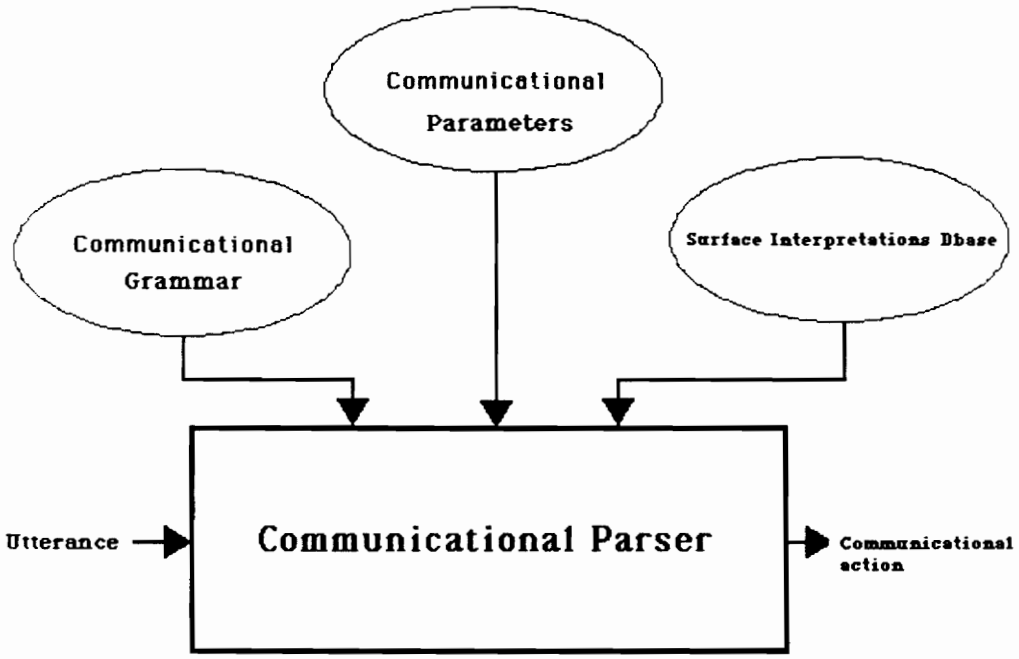


Figure 22. The Communication Parser

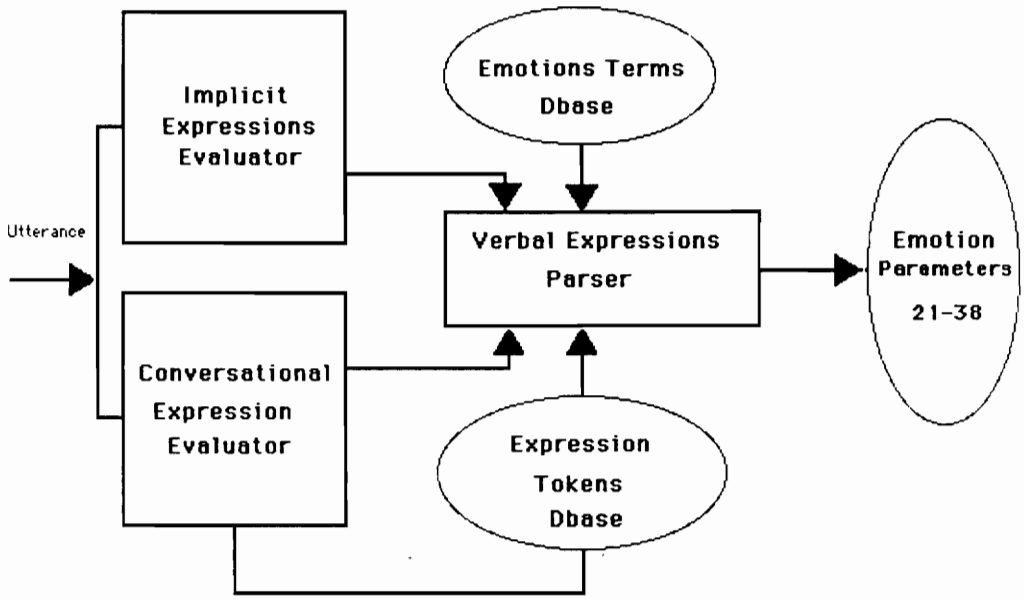


Figure 23. The Verbal Expression Parser

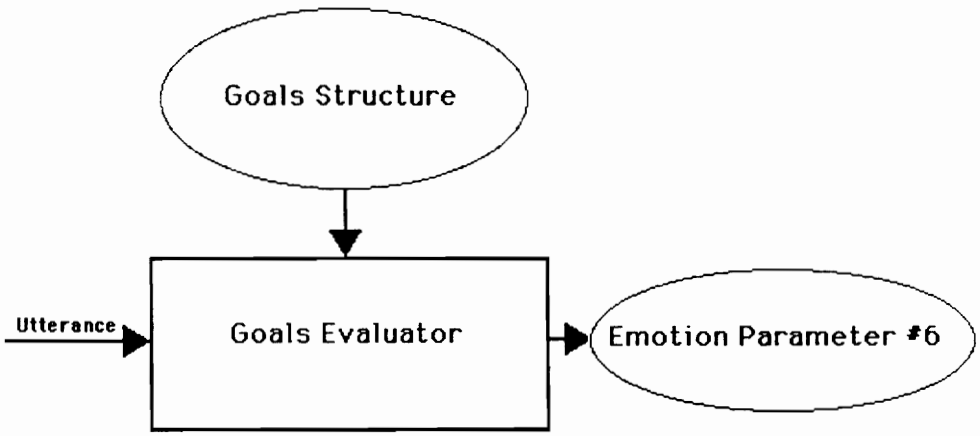


Figure 24. The Goals Evaluator

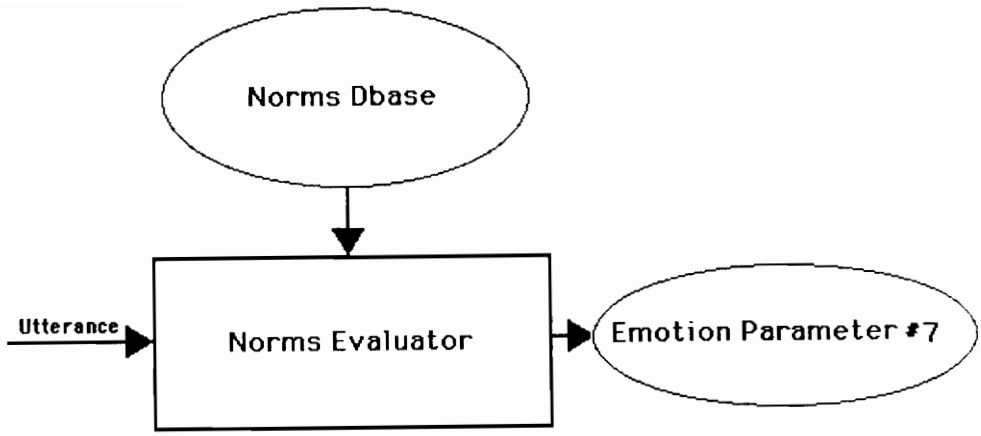


Figure 25. The Norms Evaluator

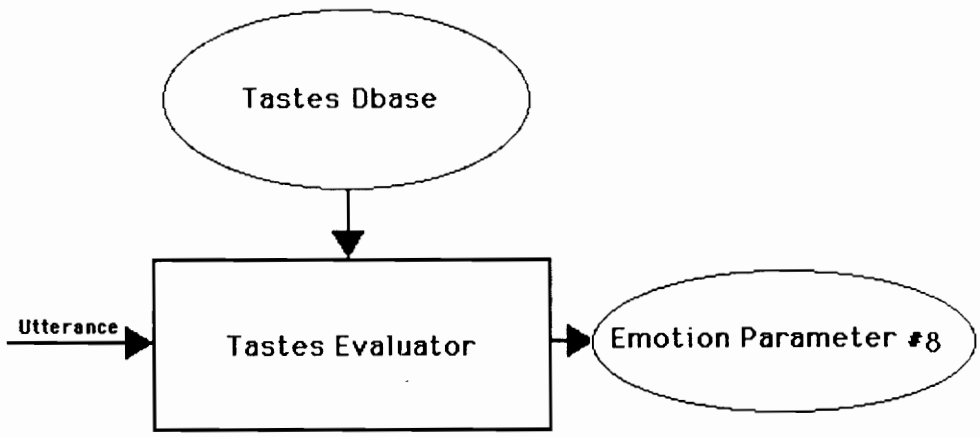


Figure 26. The Tastes Evaluator

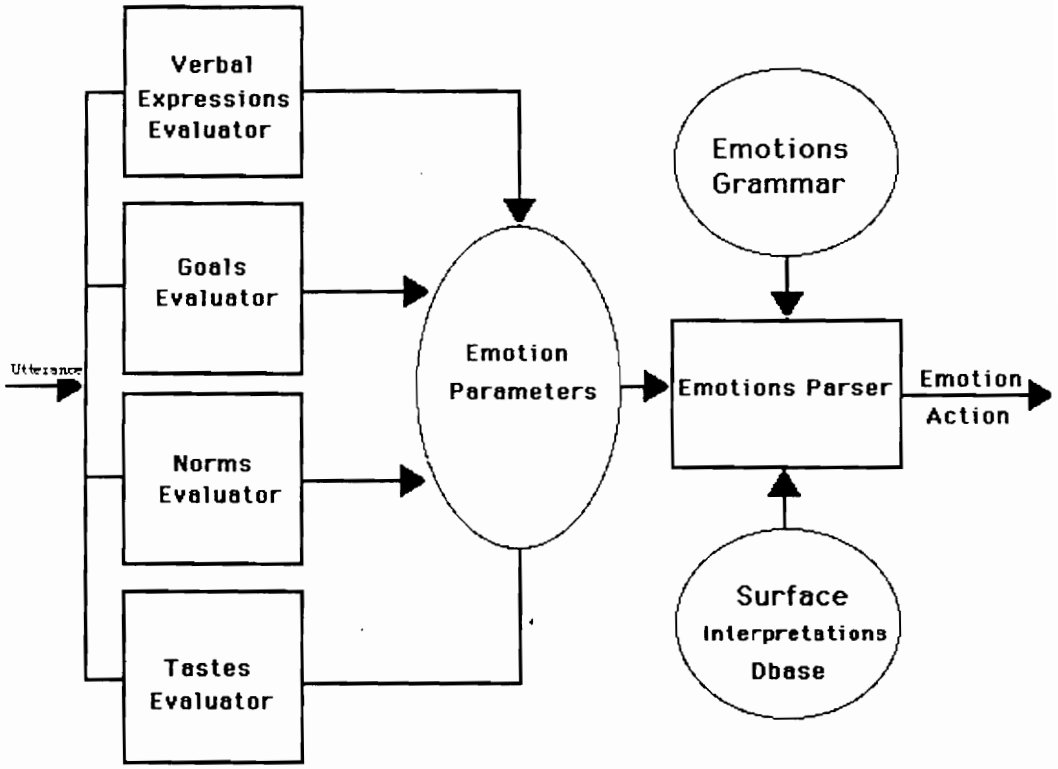


Figure 27. The Emotions Parser

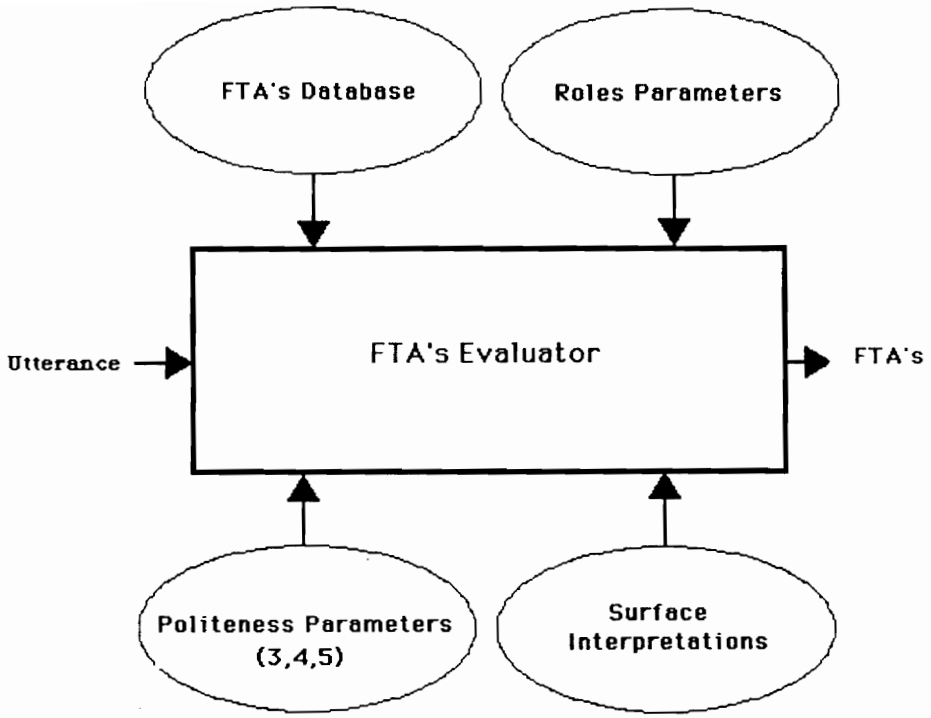


Figure 28. The FTA's Evaluator

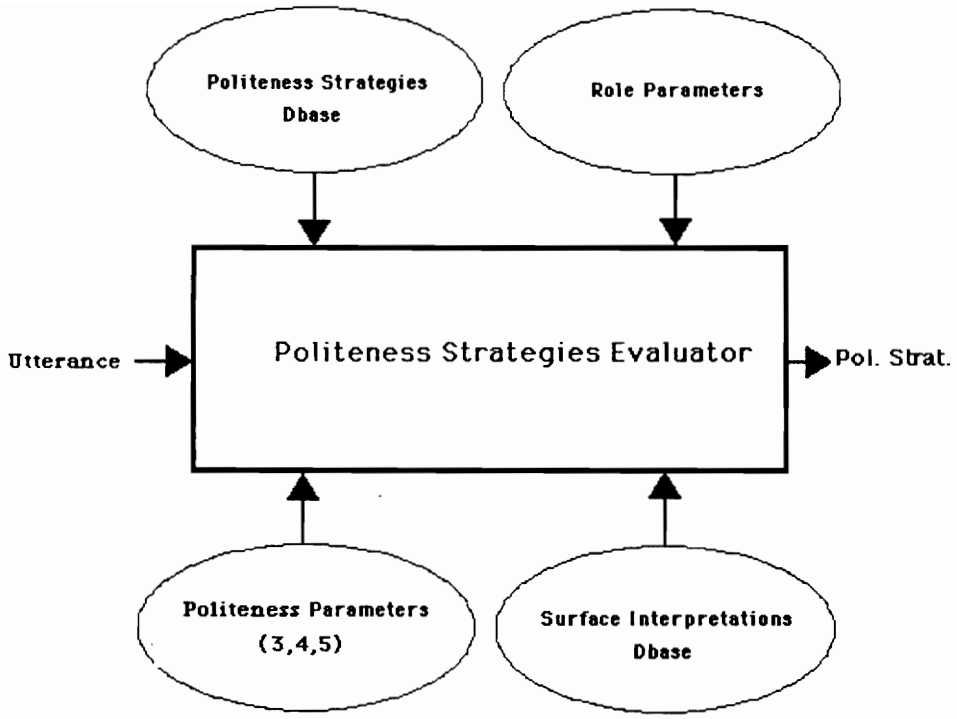


Figure 29. The Politeness Strategies Parser

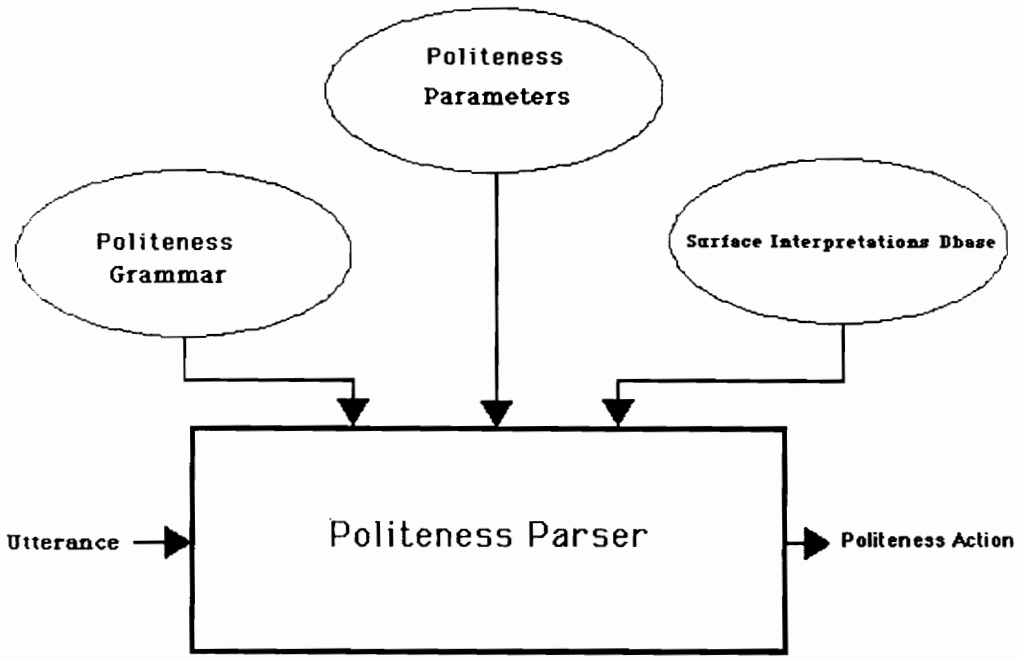


Figure 30. The Politeness Parser

Chapter 12. Results and Conclusions

We give in this chapter some concluding remarks about the aims of our work, the goals we have accomplished, and the work that still remains to be done.

12.1 What We Wanted To Do

The immediate problem this effort addressed was the construction of a system able to parse the meaning of meta-statements. After examination of real speech produced by people talking over the telephone, it became clear to us that no model of natural, spontaneous talk could be

satisfactory if meta-statements were ignored. Meta-statements are not "unnecessary" linguistic particles. Instead, they are messages rich with information about the context of the communicational setting where the exchange is taking place. This context is what is called the pragmatic aspect of communication.

12.2 What We Have Done

In what follows, we briefly discuss what we have accomplished in this work. We first present an outline of the work we have done in terms of design, and then we discuss what we have implemented, i.e., what we have coded.

Design

To solve the problem of parsing meta-statements, we had to answer two questions. First, we had to define what a meta-statement can mean, and second, we had to determine how a meta-statement is mapped to its corresponding meaning.

The first "problem" is definitional: that is, what a meta-statement can mean (i.e., its semantic domain) depends on how we delimit the domain of reference of meta-statements. Since our task is to build the pragmatic aspect of communication, i.e., the communicational context within which the conversations we want to model are taking place, the domain of meaning that is relevant to us is the domain of information about the relationship between speakers and the activity of talking.

With respect to the second problem, our task was to determine the elements and the procedures needed to map meta-statements to their corresponding meanings. To solve this

problem, we proposed the construction of a model of the context within which the meta-statements being parsed occur.

Concerning the semantic domain of meta-statements, our hypothesis was that there are two types of functions that a meta-statement fulfills. First, a meta-statement may relate information about the talk exchanged during an interaction. We called this function the "talk-about-talk" function of meta-statements, and we distinguished between two functions that are said to be "talk-about-talk." First is the rhetorical function, through which a meta-statement relates information about the structure of talk, and second is the conversational function, through which a meta-statement relates information about the process of talking.

The second type of meta-statement functions we termed "talk-about-talkers." Here, the meta-stament's function relates information about the conversers involved in the exchange. Four "talk-about-talkers" functions were identified: (1) the politeness function, (2) the emotion function, (3) the cognitive function, and (4) the communicational function.

In chapters Four and Five the rhetorical and conversational functions were examined and models for encoding the rhetorical and conversational states of talk were specified. Chapter Seven contains a discussion on issues of politeness and face threats, while in Chapter Eight a model of emotions as well as a discussion on how emotions are verbally expressed were given. In Chapter Nine we examined the cognitive function of meta-statements, while in Chapter Ten the communicational function was discussed.

In addition to examining the individual functions of meta-statements, we also discussed the important problem of modeling roles. We focused on this problem in Chapter Six, where we outlined a model for describing social behavior of varying degrees of complexity.

In Chapter Eleven, an outline of the system incorporating the various sub-systems identified in chapters Four through Ten was given. We specified the engines needed, the databases that must be furnished, and the structure of the system as a whole.

The appendices given in [BOUZ91] contain the details of the system. In Appendix A the parameters needed to model the six states of a conversation are specified along with their internal representation. Appendix B contains information about the roles that we need to

model. Appendix C contains a specification of the meta-actions recognized by our system as well as a specification of the internal representation of meta-actions.

In Appendix D of [BOUZ91] (Appendinx A of this work), we give a list of the meta-statements encountered in our transcripts. To each meta-statement, when appropriate, we tag a surface meaning corresponding to the context-free meaning of the words that constitute the meta-statement.

In Appendix E we give a list of the FTA's and politeness strategies that we need to incorporate in our system; Appendix F specifies the effects that meta-actions have on the internal parameters; while in Appendix G we specify the emotion categories used in our model of emotions. In addition, we specify in Appendix H a list of emotion words we will be using to refer to specific emotion states. For each word we specify the emotion category under which the emotion designated by that word falls, as well as the intensity of the emotion described by that word.

Implementation

At the time of writing, the following system elements have been implemented:

- A database containing information about norms, roles, and statuses (Appendix B).
- A database containing meta-statement surface interpretations (Appendix D).
- A database containing information about FTA's and politeness strategies (Appendix E).
- A list of emotion words and the emotion categories to which the emotions described by the words belong (Appendix H).
- A database containing information about the effects of meta-actions on internal parameters (Appendix F).
- The six grammars (Appendix I).
- The parameters needed for modeling each of the six identified dimensions (Appendix A).

- A driver for effecting changes on system parameters given the meta-actions parsed--i.e., an engine to process the information given in Appendix F.

12.3 What Remains to Be Done

Needless to say, there is much that still remains to be done before we can claim to have built a working parser. At this point in the development, the only legitimate claim we can make is that we have identified the *static* elements of our system.

What still remains to be accomplished is the construction of the *dynamic* part of the parser: i.e., the various engines identified in Chapter Eleven. The structure of the parser, i.e., what inputs are needed, what outputs are expected, what parameters and databases each engine needs, etc., are at this point developed to a satisfactory degree. What remains to be done is the development of the internals of the engines needed.

For example, we know that the cognitive parser receives input (in addition to the meta-statement it is parsing) from the cognitive grammar, the cognitive parameters, and the surface interpretations database, and that the output it produces is the cognitive action corresponding to the meta-statement being parsed. The problem that remains to be solved is: how is the parser to use all that information and what algorithm should it follow to choose among the many possible cognitive actions *the* appropriate cognitive action, i.e., the cognitive action that corresponds to the meta-statement's meaning? That is, having now created a model of the context, how is this model to be used to parse meta-statements?

Work also still remains to be done with respect to modeling the emotional state of interactants. We outlined in Chapter Eight a paradigm for modeling emotions, and we gave a specification of the various elements that, according to the paradigm we have adopted, need

to be modeled. Our feeling, however, is that the model described here, and the level of detail at which it is described, will need many further extensions.[59]

Similarly, work still remains to be done in the field of politeness. More specifically, engines for detecting FTA's and politeness strategies must be further developed. At this point in the development, the system is capable of identifying various politeness strategies that are encoded in the verbal formulations of utterances, and more specifically, what Sanford and Roach [SANF88] term "the imperative force" of utterances. However, the parser is still unable to identify many other FTA formulations and formulations of politeness. Additional work still remains to be done in that field.

Moreover, the model of roles presented here is quite far from being as fully developed as it should be. Many issues still need to be resolved.

The most important of these issues, in our opinion, is the question of relating roles to politeness and emotion. We have suggested in various points of the work that roles are a key element in modeling social contexts. For example, the face wants of an interactant are strongly determined by the roles being played by the interactants, and consequently, what impositions or politeness strategies the system can identify will depend on what roles are being enacted.

In addition, given the face wants established by a role, the range of emotional reactions that an interactant is susceptible to experiencing will also be in function of the roles being played. For instance, we gave in Chapter Six the example of one not being as upset by an order issued to one by one's boss as one would be if the order were issued by a peer or an underling.

In short, we hypothesize that a possible important extension will be the specification of how roles, politeness, and emotions relate and interact with one another, and how the parameters and actions specific to one of these three fields can be used to modify the parameters and actions of another field.

---

[59] Work on these extensions is currently conducted in work done in [WARN91].

The third and last (at least for now) possible extension of the roles model is to develop a model that can describe more than just group behavior. A model of behavior along the organization and society levels may prove to be beneficial in situations where models used for group roles are not applicable.

Before we end our discussion, we need to make the following final observation. The model we have built will most probably be useful not only in parsing the meaning of meta-statements, but also in the parsing of general utterances. How helpful the model will be, however, is a question whose answer we will discover as the system is further developped.

Appendix A. Surface Interpretations

The following appendix contains a list of the explicit meta-statements we have found in our twenty seven transcribed conversations. With each of the meta-statements we associate what we call a *surface interpretation* (see Chapter 3). In addition, with each meta-statement we associate *the position* of that meta-statement in the talk. The scheme we have adopted for encoding meta-statement positions is explained below.

Positional Classification of Tokens

A token is tagged in accordance with its position among turn uterrances. The tag is pair or a triple of digit numbers, of the form (F1,F2) or (F1,F2,F3).

The convention is as follows:

Field-1: Position of the token in the utterance:

1 = Begins utterance.

2 = In the midst of utterance.
3 = At the end of utterance.
4 = Token forms only utterance.in turn.
5 = Token forms an utterance (and at least one other utterance exists in turn).

Field-2: Position of the utterance in turn:

If F-1 = 1,
1 = First.
2 = Other.

If F-1 = 2, None.

If F-1 = 3,
1 = Last.
2 = Other.

If F-1 = 4, None.

If F-1 = 5,
1 = First.
2 = Last.
3 = Other.

Field-3: Token Markings:

1 = Question mark follows.
2 = Exclamation mark follows.
3 = Taper off follows.
4 = Pause precedes.
5 = Pause follows.
6 = Unmarked.

For example, 313 means that the token occurs at the *end* (3) of the *last* utterance (1) and is followed by a *tapering off* (3); whereas 115 means that the token occurs at the *beginning* (1) of the *first* utterance in the turn (1) and is followed by *a pause* (5).

Explicit Meta-statements

In what follows, we give the surface interpretations of explicit meta-statements found in our corpus of transcribed conversations. In Column Two of the table we give the position of the token, while columns Three to Eight each give the exlpicit surface actions performed along the corresponding dimension. The format for actions is as follows: single numbers enclosed in parentheses (e.g., (10)) indicate meta-actions. Pairs of numbers enclosed in double parentheses indicate actions and their parameter value: e.g., ((10 2)) means that action number ten is performed and that the value of its parameter is 2 (see corresponding tables for mapping from numbers to values). In the case where more than one action is possible, the disjunct of actions is indicated by the specification of a list longer than one, e.g., (1 2) means that actions 1 and 2 are possible. A "nil" denotes that no explicit surface action exists.

The asterisk to the left of some meta-statements indicates that the meta-statement in question is an explicit emotion expression token. See Chapter Eight for a discussion on expression tokens.

Table 4. Surface Interpretations (1)

TOKEN	POS	RHE	COV	POL	EMO	COG	COM
(about)	116	((1))	(12)	((1 38))	(1)	nil	(1)
(about)	25	nil	nil	nil	(1)	nil	nil
(about)	26	nil	nil	nil	(1)	nil	nil
(actually)	116	((10 20))	(12)	nil	nil	nil	(1 10)
(again)	115	((10 21))	nil	nil	nil	nil	(1 10)
(ah)	115	((1))	(12)	((1 44))	(1)	nil	(10)
(ah)	116	((1))	(12)	((1 44))	(1)	nil	(10)
(ah)	126	nil	nil	((1 44))	(1)	nil	(10)
(ah)	26	nil	nil	((1 44))	1	nil	(10)
(ah)	313	((4))	(10)	((1 44))	(1)	nil	(10)
(ah)	322	nil	nil	((1 44))	(1)	nil	(10)
(ah)	326	nil	nil	((1 44))	(1)	nil	(10)
(ah okay)	115	((1))	(12)	((1 44))	(1)	(1)	(1 10)
(ah yes)	115	((1))	(12)	((1 44))	(1)	(1)	(1 10)
(all right)	115	((8))	(12)	nil	(1)	(1)	(1)
(all right)	125	nil	nil	nil	(1)	(1)	(1)
(all right)	46	nil	(21)	nil	(1)	(1)	(1)
(all right)	516	nil	nil	nil	(1)	(1)	nil
(all right)	536	nil	nil	nil	(1)	(1)	nil
(also)	126	((10 17))	nil	nil	nil	nil	(1 10)
(also)	26	nil	nil	nil	nil	nil	(10)
(also)	516	nil	nil	nil	nil	nil	(10)
(although)	116	((10 10))	(12)	nil	nil	nil	(1)
(and)	115	((10 ?))	(12)	nil	nil	nil	(1 10)

Table 5. Surface Interpretations (2)

TOKEN	POS	RHE	COV	POL	EMO	COG	COM
(and)	116	((10 ?))	(12)	nil	nil	nil	(1 10)
(and)	126	((10 ?))	nil	nil	nil	nil	(1 10)
(and)	24	nil	nil	nil	nil	nil	(10)
(and)	26	nil	nil	nil	nil	nil	(10)
(and again)	115	((10 21))	(12)	nil	nil	nil	(1 10)
(and also)	115	((10 20))	(12)	nil	nil	nil	(1 10)
(and also)	125	((10 20))	nil	nil	nil	nil	(10)
(and then)	115	((49) (67))	(12)	nil	nil	nil	(10)
(and then)	116	((49) (67))	(12)	nil	nil	nil	(10)
(and then)	126	((49) (67))	nil	nil	nil	nil	(10)
(and then)	26	((49) (67))	nil	nil	nil	nil	(10)
(are you sure)	511	nil	(10)	nil	(3)	(5)	nil
(around)	114	(1)	(12)	nil	(1)	nil	(1)
(basically)	114	((10 19))	(12)	nil	(1)	nil	(1)
(basically)	26	((10 19))	nil	nil	(1)	nil	nil
(because)	126	((10 1))	nil	nil	nil	nil	nil
(because)	26	((10 1))	nil	nil	nil	nil	nil
(but)	115	((10 7) (10 10))	(12)	nil	nil	nil	(1)
(but)	116	((10 7) (10 10))	(12)	nil	nil	nil	(1)
(but)	125	((10 7) (10 10))	nil	ni	nil	nil	(10)
(but)	24	((10 7) (10 10))	nil	nil	nil	nil	(10)
(but)	26	((10 7) (10 10))	nil	nil	nil	nil	(10)
(but yes sure)	116	((10 18))	(12)	nil	nil	(1)	(1)
(by the way)	116	((10 17))	(12)	nil	nil	nil	(1 10)

Table 6. Surface Interpretations (3)

TOKEN	POS	RHE	COV	POL	EMO	COG	COM
(bye)	46	((4))	(22)	((1 26))	nil	nil	(1 9)
(bye-bye)	26	((4))	(22)	((1 26))	nil	nil	(9)
(bye-bye)	526	((4))	(22)	((1 26))	nil	nil	(9)
(can I help you)	111	((5))	(10)	((2 4))	nil	nil	(7)
(can I help you with something)	521	((5))	(10)	((1 29) (2 4))	nil	(3)	(7)
(can I help you with something)	531	((5))	(10)	((1 29) (2 4))	nil	(3)	(7)
(certainly)	116	((1))	(12)	nil	(1)	nil	(1)
(come back and see us)	526	((5))	(22)	((1 26))	nil	nil	(9)
(could you hold on just a second)	531	((10 18))	(15)	nil	nil	nil	(10)
(do you think)	26	nil	nil	nil	nil	(3)	nil
(er)	113	((1))	(12)	((1 44))	(1)	nil	(10)
(exactly)	516	nil	nil	nil	nil	(5)	nil
(excuse me)	316	nil	(10)	((1 15))	(1)	nil	nil
(fine)	326	nil	nil	nil	(1)	(1 5)	nil
(fine)	46	nil	(21)	nil	nil	(1 5)	(1)
(first of all)	125	((10 13))	nil	nil	nil	nil	(1)
(for example)	24	((10 22))	nil	nil	nil	nil	nil
(go ahead)	26	((5))	nil	nil	nil	nil	(3)
(go ahead)	526	((5))	(11)	nil	nil	nil	(3)
(*god)	115	((1))	(12)	((1 3))	(1)	nil	(1)
(*god really)	46	nil	(21)	((1 3))	(1)	nil	(1)
(good)	46	((8))	(21)	((1 3))	(1)	nil	(1)
(good)	516	((1))	nil	((1 3))	(1)	nil	nil
(good enough)	536	nil	nil	((1 3))	(1)	nil	nil

Table 7. Surface Interpretations (4)

TOKEN	POS	RHE	COV	POL	EMO	COG	COM
(good morning)	115	((1))	(1)	((1 26))	nil	nil	(7)
(*gosh)	116	((1))	(12)	((1 3))	(1)	nil	(1)
(got it)	516	nil	nil	((1 38))	nil	(1)	nil
(have a good day)	526	((4))	(22)	nil	nil	nil	(9)
(hello)	115	((1))	(1)	((1 26))	nil	nil	(7)
(here we go)	326	((10 18))	nil	nil	nil	nil	(10)
(here we go)	526	((10 18))	nil	nil	nil	nil	(10)
(hey)	125	nil	nil	((1 3))	nil	nil	nil
(hi)	115	((1))	(1)	((1 26))	nil	nil	(7)
(hold just one moment)	126	((10 18))	(15)	nil	nil	nil	(10)
(hold on one moment)	44	((10 18))	(15)	nil	nil	nil	(10)
(hopefully)	26	nil	nil	nil	(1)	nil	nil
(how are you)	521	((5))	(1 10)	((1 26))	(3)	nil	(7)
(how are you doing)	311	((5))	(1 10)	((1 26))	(3)	nil	(7)
(how are you today)	116	((5))	(1 10)	((1 26))	(3)	nil	(7)
(how can I help you)	41	((5))	(10)	((1 26))	nil	nil	(7)
(huh)	536	nil	nil	((1 44))	(1)	nil	(10)
(I appreciate)	26	nil	(4)	((2 24))	(1)	nil	nil
(I appreciate it)	316	((4))	(4)	((2 24))	(1)	nil	nil
(I appreciate it)	526	((4))	(4)	((2 24))	(1)	nil	nil
(I appreciate your waiting)	316	((10 18))	(18)	((2 24))	(1)	nil	nil
(I beg your pardon)	41	nil	(10)	((2 24))	(1)	(2)	(2)
(I believe)	26	nil	nil	((1 44))	(1)	nil	nil
(I believe so)	316	nil	nil	((1 44))	(1)	(1 5)	nil

Table 8. Surface Interpretations (5)

TOKEN	POS	RHE	COV	POL	EMO	COG	COM
(I can't believe this)	526	nil	nil	((1 38))	(1)	(2)	nil
(I certainly do thank you)	46	((4))	(4)	((2 24))	(1)	nil	(9)
(I couldn't say)	526	nil	nil	((1 38))	nil	(2)	nil
(I don't blame you)	516	nil	nil	((1 12) (1 38))	(1)	nil	nil
(I don't know)	125	nil	nil	((1 38))	nil	(2)	(1)
(I don't know)	126	nil	nil	((1 38))	nil	(2)	(1)
(I don't know if I want to)	116	((1))	(12)	((1 38))	(1)	(2)	(1)
(I guess)	115	((1))	(12)	nil	(1)	(1 5)	(1)
(I guess)	25	nil	nil	nil	(1)	(1 5)	nil
(I guess)	26	nil	nil	nil	(1)	(1 5)	nil
(I guess)	326	nil	nil	nil	(1)	(1 5)	nil
(I guess not)	316	nil	nil	nil	(1)	(2 6)	nil
(I guess so)	26	nil	nil	nil	(1)	(1 5)	nil
(I have a few questions)	316	((10 17))	nil	nil	nil	(2)	(5)
(I hate to say this)	526	nil	nil	((1 13))	(1)	nil	(5)
(I know)	26	nil	nil	nil	nil	(1)	nil
(I know)	314	nil	nil	nil	nil	(1)	nil
(I know)	516	nil	nil	nil	nil	(1)	nil
(I know)	536	nil	nil	nil	nil	(1)	nil
(I know the feeling)	326	nil	nil	((1 63))	(1)	(1)	nil
(I mean)	26	((10 21))	nil	nil	nil	(1)	(5)
(I need some information)	25	((10 17))	nil	nil	nil	(2)	(3)
(I need some information)	26	((10 17))	nil	nil	nil	(2)	(3)
(I see)	26	nil	nil	nil	nil	(1)	(1)

Table 9. Surface Interpretations (6)

TOKEN	POS	RHE	COV	POL	EMO	COG	COM
(I see)	516	nil	nil	nil	nil	(1)	(1)
(I tell you)	526	((10 17))	nil	nil	nil	nil	(5)
(I tell you what)	126	((10 17))	nil	nil	nil	nil	(5)
(I thank you very much)	536	((4))	(4)	((1 9) (2 24))	nil	nil	(9)
(I think)	116	((1))	(12)	nil	(1)	(4)	(1)
(I think)	126	nil	nil	nil	(1)	(4)	(1)
(I think)	24	nil	nil	nil	(1)	(4)	nil
(I think)	26	nil	nil	nil	(1)	(4)	nil
(I think)	314	nil	nil	nil	(1)	(4)	nil
(I think so)	326	nil	nil	nil	(1)	nil	nil
(I understand that)	116	((1))	(12)	nil	nil	(1)	(1)
(I wasn't sure about that)	326	nil	nil	((1 38))	(1)	(2)	nil
(I will be honest)	116	nil	(12)	((1 13))	(1)	nil	(1)
(if at all possible)	26	nil	nil	((1 13))	(1)	nil	nil
(if you don't mind holding)	316	((10 18))	(15)	((1 29))	(1)	nil	(10)
(in other words)	116	((10 21))	(12)	nil	nil	nil	(1 10)
(in other words)	26	((10 21))	nil	nil	nil	nil	(1 10)
(is that all)	41	nil	(4)	nil	nil	(3)	(17)
(is that all right with you)	311	nil	(10)	((1 29))	(3)	(3)	nil
(is that all right with you)	521	nil	(10)	((1 29))	(3)	(3)	nil
(is that fine)	521	nil	(10)	((1 29))	(3)	(3)	nil
(is that fine so far)	311	nil	(10)	((1 29))	(3)	(3)	nil
(is that it)	521	nil	(4 10)	((1 29))	(3)	(3)	(17)
(is that satisfactory)	521	nil	(10)	((1 29))	(3)	(3)	nil

Table 10. Surface Interpretations (7)

TOKEN	POS	RHE	COV	POL	EMO	COG	COM
(it is neat)	526	nil	nil	nil	(1)	nil	nil
(it might be good though)	526	nil	nil	nil	(1)	nil	nil
(it was unreal)	126	nil	nil	((1 30))	(1)	nil	(1)
(it's not certain right now)	25	nil	nil	((1 38))	(1)	nil	nil
(I'll be right back with you)	536	((4))	(15)	((1 38))	nil	nil	(9)
(i'll get back to you)	316	((4))	(4)	((1 38))	nil	nil	(9)
(i'll get back to you)	326	((4))	(4)	((1 38))	nil	nil	(9)
(i'll get back with you)	46	((4))	(4)	((1 38))	nil	nil	(9)
(i'll tell you what)	25	((10 17))	nil	((1 38))	nil	nil	(5)
(i'm not sure)	516	nil	nil	((1 38))	(1)	(2)	nil
(i'm really pleased)	126	nil	nil	((1 38))	(1)	nil	nil
(i'm sorry)	115	((1))	(12)	((1 15) (1 38))	(1)	nil	nil
(i'm sorry)	126	nil	nil	((1 15) (1 38))	(1)	nil	nil
(i'm sorry)	516	nil	nil	((1 15) (1 38))	(1)	nil	nil
(just)	114	((1))	(12)	nil	(1)	nil	nil
(just)	26	nil	nil	nil	(1)	nil	nil
(just a moment)	324	((10 18))	(15)	((1 38))	nil	nil	(10)
(just a moment)	526	((10 18))	(15)	((1 38))	nil	nil	(10)
(just a moment)	536	((10 18))	(15)	((1 38))	nil	nil	(10)
(just a second)	516	((10 18))	(15)	((1 38))	nil	nil	(10)
(kinda)	26	nil	nil	((1 37))	(1)	nil	(5)
(let me)	126	nil	nil	((1 56))	nil	nil	(5)
(let me)	26	nil	nil	((1 56))	nil	nil	(5)

Please note the entry given in the footnote:[60]

Table 11. Surface Interpretations (8)

TOKEN	POS	RHE	COV	POL	EMO	COG	COM
(let me see)	25	nil	nil	((1 56))	nil	nil	(5)
(let me see)	326	nil	nil	((1 56))	nil	nil	(5)
(let me see)	516	nil	nil	((1 56))	nil	nil	(5)
(let me see here)	326	nil	nil	((1 56))	nil	nil	(5)
[see footnote below]	516	((10 19))	nil	((1 56) (1 29))	nil	nil	(5)
(let me try)	46	nil	nil	((1 56))	nil	nil	(1 5)
(let me try something here)	515	((10 14))	nil	((1 56))	nil	nil	(5)
(let's assume)	24	nil	nil	((1 56) (1 38))	nil	nil	(5)
(let's check that)	326	nil	nil	((1 56) (1 38))	nil	nil	(5)
(let's imagine)	126	nil	nil	((1 56) (1 38))	nil	nil	(5)
(let's say)	114	((10 14))	(12)	((1 56) (1 38))	nil	nil	(1 5)
(let's say)	126	((10 14))	nil	((1 56) (1 38))	nil	nil	(5)
(let's say)	26	((10 14))	nil	((1 56) (1 38))	nil	nil	(5)
(let's see)	115	((1))	(12)	((1 56) (1 38))	nil	nil	(1 5)
("")	25	nil	nil	((1 56) (1 38))	nil	nil	(5)
("")	26	nil	nil	((1 56) (1 38))	nil	nil	(5)
("")	326	nil	nil	((1 56) (1 38))	nil	nil	(5)
("")	516	nil	nil	((1 56) (1 38))	nil	nil	(5)
("" here)	26	nil	nil	((1 56) (1 38))	nil	nil	(5)
("" here)	326	nil	nil	((1 56) (1 38))	nil	nil	(5)
("" how that works)	526	nil	nil	((1 56) (1 38))	nil	nil	(5)
("" what we can do here)	46	((10 17))	nil	((1 56) (1 38))	nil	nil	(1 5)
(let's try)	116	((10 14))	(12)	((1 56) (1 38))	nil	nil	(1 5)
(let's try)	26	((10 14))	nil	((1 56) (1 38))	nil	nil	(5)

[60] Due to formatting restrictions, the token for this entry is given in this footnote. The token is: "Let me show you what we have so far."

Table 12. Surface Interpretations (9)

TOKEN	POS	RHE	COV	POL	EMO	COG	COM
(like)	116	((1))	(12)	((1 37))	nil	nil	(1)
(may I help you)	311	((5))	(10)	((1 29))	nil	nil	(7)
(maybe)	116	((1))	(12)	nil	(1)	nil	(1)
(maybe)	26	nil	nil	nil	(1)	nil	nil
(no)	114	((1))	(12)	((2 8))	nil	nil	(1)
(no)	115	((1))	(12)	((2 8))	nil	nil	(1)
(no)	116	((1))	(12)	((2 8))	nil	nil	(1)
(no)	46	nil	nil	((2 8))	nil	nil	(1)
(no)	511	nil	(10)	((2 8))	nil	nil	nil
(no)	516	nil	nil	((2 8))	nil	nil	nil
(no but)	116	((10 7))	nil	((2 8))	nil	nil	(1)
(no problem)	316	nil	nil	((2 25) (1 9))	(1)	nil	nil
(no way)	316	nil	nil	nil	(1)	nil	nil
(not at all)	536	nil	nil	((2 25) (1 9))	nil	nil	nil
(now)	115	((10 17))	(12)	nil	nil	nil	(1)
(now)	116	((10 17))	(12)	nil	nil	nil	(1)
(now)	124	((10 17))	(12)	nil	nil	nil	nil
(now)	126	((10 17))	nil	nil	nil	nil	nil
(now will that be all)	41	((4))	(4)	((1 30))	(3)	nil	(1 17)
(*of course)	26	nil	nil	nil	(1)	nil	nil
(*oh)	115	((1))	(12)	((1 3))	(1)	nil	(1)
(*oh)	116	((1))	(12)	((1 3))	(1)	nil	(1)
(*oh)	126	nil	nil	((1 3))	(1)	nil	nil
(*oh)	24	nil	nil	((1 3))	(1)	nil	nil

Table 13. Surface Interpretations (10)

TOKEN	POS	RHE	COV	POL	EMO	COG	COM
(*oh)	43	nil	(21)	((1 3))	(1)	nil	nil
(*oh)	516	nil	nil	((1 3))	(1)	nil	nil
(*oh boy)	126	nil	nil	((1 3))	(1)	nil	nil
(*oh boy)	46	nil	(21)	((1 3))	(1)	nil	(1)
(*oh good)	516	nil	nil	((1 3))	(1)	nil	nil
(*oh I know)	125	nil	nil	((1 3))	(1)	nil	nil
(*oh I see)	516	nil	nil	((1 3))	(1)	nil	nil
(*oh it is)	41	((5))	(21)	((1 3))	(1)	nil	(1)
(*oh no)	115	((1))	(12)	((1 3))	(1)	nil	(1)
(*oh no)	116	((1))	(12)	((1 3))	(1)	nil	(1)
(*oh yeah)	46	nil	(21)	((1 3))	nil	nil	(1)
(oh okay)	46	nil	(21)	((1 3))	nil	nil	(1)
(oh okay)	516	((1))	nil	((1 3))	nil	nil	nil
(*oh that is great)	516	((1))	nil	((1 3))	nil	nil	nil
(okay)	115	((1))	(12)	((1 42))	nil	nil	(1)
(okay)	125	nil	nil	((1 42))	nil	nil	nil
(okay)	126	nil	nil	((1 42))	nil	nil	nil
(okay)	25	nil	nil	((1 42))	nil	nil	nil
(okay)	314	nil	nil	((1 42))	nil	nil	nil
(okay)	324	nil	nil	((1 42))	nil	nil	nil
(okay)	41	((5))	(21)	((1 42) (1 33))	(3)	(1 5)	(1)
(okay)	46	nil	(21)	((1 42) (1 33))	nil	(1)	(1)
(okay)	511	((5))	(10)	((1 42) (1 33))	(3)	(1 5)	(17)
(okay)	516	((1))	nil	((1 42) (1 33))	nil	(1)	nil

Table 14. Surface Interpretations (11)

TOKEN	POS	RHE	COV	POL	EMO	COG	COM
(okay)	526	nil	nil	nil	nil	(1)	nil
(one moment please)	516	((10 18))	(15)	((1 27))	nil	nil	(10)
(or)	116	((10 ?))	nil	nil	nil	nil	(1)
(pardon me)	41	((5))	(10)	((1 15))	nil	nil	(2)
(please)	326	nil	nil	((1 27))	nil	nil	nil
(please)	311	nil	nil	((1 27))	nil	nil	nil
(possibly)	26	nil	nil	nil	(1)	nil	nil
(pretty good)	326	nil	nil	nil	(1)	nil	nil
(pretty good)	46	nil	(21)	nil	(1)	nil	nil
(pretty good)	516	((1))	nil	nil	(1)	nil	nil
(probably)	114	((1))	(12)	nil	(1)	nil	(1)
(probably)	116	((1))	(12)	nil	(1)	nil	(1)
(probably)	126	nil	nil	nil	(1)	nil	nil
(probably)	26	nil	nil	nil	(1)	nil	nil
(probably)	516	((1))	nil	nil	(1)	nil	nil
(probably not)	46	nil	nil	nil	(1)	nil	nil
(quite possible)	26	nil	nil	nil	(1)	nil	nil
(really)	26	nil	nil	nil	(1)	nil	nil
(really)	311	((5))	(10)	nil	(1)	(1)	(1)
(really)	316	nil	nil	nil	(1)	nil	nil
(really)	41	nil	(21)	nil	(1)	(1)	(1 3)
(really)	516	((1))	nil	nil	(1)	nil	nil
(really)	536	nil	nil	nil	nil	nil	nil
(right)	115	((1))	(12)	((1 42))	(1)	(1 5)	nil

Table 15. Surface Interpretations (12)

TOKEN	POS	RHE	COV	POL	EMO	COG	COM
(right)	311	((5))	(10)	((1 42))	nil	(3 4)	(1)
(right)	316	nil	nil	((1 42))	nil	(1 5)	nil
(right)	324	nil	nil	((1 42))	nil	(1 5)	nil
(right)	41	((5))	(21)	((1 42) (1 33))	nil	(3 4)	(1)
(right)	46	nil	(21)	((1 42) (1 33))	nil	(1 5)	(1)
(right)	516	nil	nil	((1 42) (1 33))	nil	(1 5)	nil
(right)	521	nil	(10)	((1 42) (1 33))	nil	(3 4)	nil
(roughly)	114	((1))	(12)	nil	(1)	nil	nil
(see)	((1))	(12)	nil	((1 56))	nil	nil	(1)
(see)	116	((1))	(12)	((1 56))	nil	nil	(1)
(see)	26	nil	nil	((1 56))	nil	nil	nil
(since)	116	((10 ?))	(12)	nil	nil	nil	(1)
(since)	26	nil	nil	nil	nil	nil	nil
(sir)	115	((1))	(12)	((1 10))	nil	nil	(1)
(sir)	311	((5))	(10)	((1 10))	nil	nil	nil
(sir)	316	nil	nil	((1 10))	nil	nil	nil
(so)	115	((1))	(12)	nil	nil	nil	(1)
(so)	116	((1))	(12)	nil	nil	nil	(1)
(so)	125	((1))	nil	nil	nil	nil	nil
(so)	126	((1))	nil	nil	nil	nil	nil
(so)	26	nil	nil	nil	nil	nil	nil
(so)	313	((11) (12))	nil	nil	nil	nil	nil
(so that)	26	nil	nil	nil	nil	nil	nil
(so-called)	26	nil	nil	nil	(1)	nil	nil

Table 16. Surface Interpretations (13)

TOKEN	POS	RHE	COV	POL	EMO	COG	COM
(sorry I couldn't help you)	326	nil	(4)	((1 15) (1 38))	(1)	nil	(9)
(sorry that this takes so long)	46	nil	nil	((1 15) (1 38))	(1)	nil	(110)
(sure)	115	((1))	(12)	((1 41))	nil	nil	nil
(sure)	314	nil	nil	((1 41))	nil	nil	nil
(sure)	41	((5))	nil	((1 41))	nil	nil	nil
(sure)	46	nil	nil	((1 41))	nil	nil	(1)
(sure)	516	((1))	nil	((1 41))	nil	nil	(1)
(thank you)	316	nil	nil	((1 9) (2 24))	(1)	nil	(9)
(thank you)	46	nil	nil	((1 9) (2 24))	(1)	nil	(9)
(thank you)	526	nil	nil	((1 9) (2 24))	(1)	nil	(9)
(thank you)	536	nil	nil	((1 9) (2 24))	(1)	nil	(9)
(thank you anyway)	46	nil	(4)	((1 9) (2 24))	(1)	nil	(9)
(thank you for calling)	26	((4))	(4)	((1 9) (2 24))	(1)	nil	(9)
(thank you for calling)	46	((4))	(4)	((1 9) (2 24))	(1)	nil	(9)
(thank you for calling)	526	((4))	(4)	((1 9) (2 24))	(1)	nil	(9)
(thank you for holding now)	46	((1))	(18)	((1 9) (2 24))	(1)	nil	(9)
(thank you for your help)	314	((4))	(4)	((1 9) (2 24))	(1)	nil	(9)
(thank you very much)	314	((4))	(4)	((1 9) (2 24))	(1)	nil	(9)
(thank you very much)	46	nil	(4)	((1 9) (2 24))	(1)	nil	(9)
(thank you very much)	526	nil	(4)	((1 9) (2 24))	(1)	nil	(9)
("" for your help)	46	((4))	(4)	((1 9) (2 24))	(1)	nil	(9)
("" for you information)	316	((4))	(4)	((1 9) (2 24))	(1)	nil	(9)
("" for waiting)	316	((1))	(18)	((1 9) (2 24))	(1)	nil	(9)

Table 17. Surface Interpretations (14)

TOKEN	POS	RHE	COV	POL	EMO	COG	COM
(thanks)	46	nil	(4)	((1 9) (2 24))	(1)	nil	(9)
(thanks a lot)	46	nil	(4)	((1 9) (2 24))	(1)	nil	(9)
(thanks for calling)	26	((4))	(4)	((1 9) (2 24))	(1)	nil	(9)
(thanks for waiting)	516	((1))	(18)	((1 9) (2 24))	(1)	nil	(9)
(that is a problem)	516	((1))	nil	nil	(1)	nil	nil
(*that is excellent)	536	nil	nil	((1 42))	(1)	nil	nil
(*that is fine)	316	nil	nil	((1 42))	(1)	nil	nil
(*that is fine)	326	nil	nil	((1 42))	(1)	nil	nil
(*that is good)	316	nil	nil	((1 42))	(1)	nil	nil
(*that is good)	516	((1))	nil	((1 42))	(1)	nil	nil
(that is it)	46	((4))	nil	nil	nil	nil	(1)
(that is right)	516	((1))	nil	((1 42))	nil	nil	nil
(that is true)	46	nil	(21)	((1 42))	nil	nil	(1)
(that is what I was thinking)	26	nil	nil	((1 42))	nil	nil	nil
(*that is wonderful)	46	nil	(21)	((1 30))	nil	nil	(1)
(*that sounds fine)	46	nil	(21)	nil	(1)	nil	(1)
(*that sounds good)	316	nil	nil	nil	(1)	nil	nil
(*that sounds good)	326	nil	nil	nil	(1)	nil	nil
(*that will be fine)	46	nil	(21)	nil	(1)	nil	(1)
(*that would be not too good)	516	((1))	nil	((1 6) (2 8))	(1)	nil	nil
(*that wouldn't do at all)	316	nil	nil	((1 6) (1 38) (2 8))	(1)	nil	nil
(thataway)	26	nil	nil	((1 37))	nil	nil	nil
(*that'll be great)	326	nil	nil	((1 38) (1 42))	(1)	nil	nil
(*that's fine)	326	nil	nil	((1 38) (1 42))	(1)	nil	nil

Table 18. Surface Interpretations (15)

TOKEN	POS	RHE	COV	POL	EMO	COG	COM
(*that's fine)	516	((1))	nil	((1 38) (1 42))	(1)	nil	nil
(*that's good)	26	nil	nil	((1 38) (1 42))	(1)	nil	nil
(*that's good)	46	nil	(21)	((1 38) (1 42))	(1)	nil	(1)
(that's it)	516	((4))	nil	((1 38))	nil	nil	nil
(*that's no good)	46	nil	nil	((1 38) (2 8))	(1)	nil	(1)
(*that's perfect)	316	nil	nil	((1 38) (1 42) (1 30))	(1)	nil	nil
(that's quite all right)	46	nil	nil	((2 25))	nil	nil	(1)
(that's the way)	316	nil	nil	((1 38) (1 42))	nil	nil	nil
(that's true)	46	nil	(21)	((1 38) (1 42))	nil	nil	(1)
(then)	115	((49) (67))	(12)	nil	nil	nil	(1)
(then)	116	((49) (67))	(12)	nil	nil	nil	(1)
(then)	126	((49) (67))	nil	nil	nil	nil	nil
(then)	24	((49) (67))	nil	nil	nil	nil	nil
(therefore)	26	((10 3))	nil	nil	nil	nil	nil
(though)	326	((10 ?))	nil	nil	nil	nil	nil
(uh)	115	((1))	nil	((1 44))	nil	nil	nil
(uh)	116	((1))	(12)	((1 44))	nil	nil	nil
(uh)	25	nil	nil	((1 44))	nil	nil	nil
(uh-huh)	115	((1))	(12)	((1 37))	nil	nil	(1)
(uh-huh)	314	nil	nil	((1 37))	nil	nil	nil
(uh-huh)	326	nil	nil	((1 37))	nil	nil	nil
(uh-huh)	46	nil	(21)	((1 37))	nil	nil	(1)
(uh-huh)	516	((1))	(21)	((1 37))	nil	nil	nil
(um)	116	((1))	(12)	((1 44))	nil	nil	nil

Table 19. Surface Interpretations (16)

TOKEN	POS	RHE	COV	POL	EMO	COG	COM
(um)	126	nil	nil	((1 44))	nil	nil	nil
(um)	25	nil	nil	((1 44))	nil	nil	nil
(um)	41	((5))	(21)	((1 44))	nil	nil	(2)
(umm)	516	((1))	nil	((1 44))	nil	nil	nil
(*very good)	316	nil	nil	((1 42))	(1)	nil	nil
(*very good)	516	((1))	nil	((1 42))	(1)	nil	nil
(we can try)	46	nil	nil	((1 56))	nil	nil	(1)
(well)	114	((1))	(12)	((1 44))	(1)	nil	(1)
(well)	115	((1))	(12)	((1 44))	(1)	nil	(1)
(well)	116	((1))	(12)	((1 44))	(1)	nil	(1)
(well)	125	nil	nil	((1 44))	(1)	nil	nil
(well)	126	nil	nil	((1 44))	(1)	nil	nil
(well)	25	nil	nil	((1 44))	(1)	nil	nil
(well)	26	nil	nil	((1 44))	(1)	nil	nil
(well)	313	((5))	nil	((1 44))	(1)	nil	nil
(well)	43	((5))	nil	((1 44))	(1)	nil	(1)
(what can I do for you)	311	((5))	(10)	((1 29))	nil	nil	(7)
(what can I help you with)	41	((5))	(10)	((1 29))	nil	nil	(7)
(what is the whole thing together)	316	((13 19))	(10)	((1 29))	nil	nil	(3)
(what is this for)	41	((13 20))	(10)	((1 29))	nil	(2)	(3)
(what seems to be the problem)	521	((13 20))	(10)	((1 29))	nil	(3)	(3)
(what seems to be the problem)	531	((13 20))	(10)	((1 29))	nil	(3)	(3)
(what's this mean)	116	((13 20))	(12)	((1 38))	nil	(2)	(3)
(while)	116	((49) (67))	(12)	nil	nil	nil	(1)

Table 20. Surface Interpretations (17)

TOKEN	POS	RHE	COV	POL	EMO	COG	COM
(will that be all)	311	((13 20))	(10)	((1 29))	nil	(3)	(17)
(yeah)	115	((1))	(12)	((1 37) (1 42))	nil	nil	(1)
(yeah)	116	((1))	(12)	((1 37) (1 42))	nil	nil	(1)
(yeah)	25	nil	nil	((1 37) (1 42))	nil	nil	nil
(yeah)	311	((5))	(10)	((1 37) (1 42))	nil	nil	nil
(yeah)	314	nil	nil	((1 37) (1 42))	nil	nil	nil
(yeah)	316	nil	nil	((1 37) (1 42))	nil	nil	nil
(yeah)	46	nil	(21)	((1 37) (1 42))	nil	nil	(1)
(yeah)	516	((1))	nil	((1 37) (1 42))	nil	nil	nil
(yeh)	116	((1))	(12)	((1 37) (1 42))	nil	nil	(1)
(yes)	114	((1))	(12)	((1 42))	nil	nil	(1)
(yes)	115	((1))	(12)	((1 42))	nil	(1 5)	(1)
(yes)	26	nil	nil	((1 42))	nil	(1 5)	nil
(yes)	316	nil	nil	((1 42))	nil	(1 5)	nil
(yes)	326	nil	nil	((1 42))	nil	(1 5)	nil
(yes)	46	nil	(10)	((1 42))	nil	(1 5)	(1)
(yes)	516	((1))	nil	((1 42))	nil	(1 5)	nil
(yes)	526	nil	nil	((1 42))	nil	(1 5)	nil
(yes sure)	46	nil	(10)	((1 42))	nil	(1 5)	(1)
(yes ma'am)	316	nil	nil	((1 35) (1 42) (1 38))	nil	(1 5)	nil
(yes ma'am)	46	nil	nil	((1 35) (1 42) (1 38))	nil	(1 5)	(1)
(yes sir)	41	((5))	(10)	((1 35) (1 42))	nil	(1 5)	(3)
(yes yes)	516	((1))	nil	((1 42) (1 40))	(1)	(1 5)	nil
(you are quite welcome)	46	((4))	(5)	((2 25) (1 9))	(1)	nil	(9)

Table 21. Surface Interpretations (18)

TOKEN	POS	RHE	COV	POL	EMO	COG	COM
(you are welcome)	46	((4))	(5)	((1 9) (2 25))	nil	nil	(9)
(you know)	116	((1))	(12)	((1 51))	nil	nil	(1)
(you know)	125	((1))	nil	((1 51))	nil	nil	nil
(you know)	25	nil	nil	((1 51))	nil	nil	nil
(you know)	26	nil	nil	((1 51))	nil	nil	nil
(you know)	311	((5))	(10)	((1 51) (1 33))	nil	(3)	nil
(you know)	((4))	nil	((1 51))	nil	nil	nil	nil
(you know)	41	((5))	nil	((1 33))	nil	(3)	(1)
(you mean)	116	((1))	nil	nil	nil	(4)	(1)
(you think)	311	((5))	(10)	nil	nil	(4)	nil
(you too)	314	nil	nil	nil	nil	nil	nil
(you're quite welcome)	516	((4))	(5)	((1 38) (1 9) (2 25))	nil	nil	(9)
(you're welcome)	46	((4))	(5)	((1 9) (1 38) (2 25))	1	1	(9)
(...)	26	nil	nil	nil	(1)	(1)	nil
(...)	313	nil	nil	nil	(1)	(1)	nil
('cause)	26	((10 1))	nil	((1 38))	nil	nil	nil
(*(laugh))	312	nil	nil	((1 53))	(1)	nil	nil
(*(laugh))	516	nil	nil	((1 53))	(1)	nil	nil

[ACKR73] Ackrill, J. L. (1973). Aristotle's Ethics. Edited by Warnock, M. Robert Maclehose & Co. Ltd.

[ALLE83] Allen, J. F. (1983). Recognizing Intentions from Natural Language Utterances In Brady, M. and Berwick, R. C., *Computational Models of Discourse*, Cambridge, MA: MIT Press, 107-166.

[ARIS41] Aristotle (1941). Rhetoric. In McKeon, R. (ed.). *The Basic Works of Aristotle.* Random House.

[ARIS68] Aristotle (1968). De Anima; Books II and III (with certain passages from Book I). Trans. Hamlyn, D. W. Clarendon Press.

[ARNO60] Arnold, M. B. (1960). Emotion and Personality (2 vols.). Columbia University Press.

[ARNO70] Arnold, M. B. (1970). Feelings and Emotions: The Loyola Symposium. Academic Press.

[ATKI79] Atkinson, J. M. (1979). The Organization of Verbal Interaction in Judicial Settings. Atlantic Highlands, NJ: Humanities Press.

[BACH79] Bach, K. & Harnish, R. M. (1979). Linguistic Communication and Speech Acts. Cambridge, Mass.: MIT Press.

[BALE50] Bales, R. F. (1950). Interaction Process Analysis: A Method for the Study of Small Groups. Reading, Mass.: Addison-Wesley.

[BALE70] Bales, R. F. (1970). Personality and Interpersonal Behavior. NY: Holt, Rinehart and Winston.

[BALL81] Ballmer, T. T. & Brennenstuhl, W. (1981). Speech Act Classification: A Study of the Lexical Analysis of English Speech Activity Verbs. NY: Springer-Verlag.

[BATE75] Bates, F. L., Harvey C. C. (1975). The Structure of Social Systems. NY: Gardner Press, Inc.

[BEDF56] Bedford, E. (1956-7). Emotions. *Proceedings of the Aristotelean Society,* vol. 57.

[BERN64] Berne, E. (1964). Games People Play. NY: Random House.

[BIDD79] Biddle, B. J. (1979). Role Theory: Expectations, Identities, and Behavior. Academic Press.

[BIRD70] Birdwhistell, R. L. (1970). Essays on Body Motion Communication. University of Pennsylvania Press.

[BOBR79] Bobrow, D. G. (1979). Studies in Cognitive Science. Academic Press.

[BODE77] Boden, M. A. (1977). Artificial Intelligence and Natural Man. NY: Basic Books.

[BREN89] Brentano, C. (1889). On the Origin of our Knowledge of Right and Wrong.

[BROW87] Brown, P. & Levinson, S. C. (1987). Politeness: Some Universals in Language Usage. NY: Cambridge University Press.

[BOUZ91] Bouzid, A. T. (1991). Parsing Meta-Communicational Statements in Dialogues: A Computational Model. Virginia Polytechnic Institute, Department of Computer Science, Technical Report (in progress).

[CALH84] Calhoun, C. & Solomon, R. C. (1984). What Is an Emotion. Oxford University Press.

[CERC83] Cercone, N. J. (1983). Computational Linguistics. NY: Pergamon Press.

[CHAF70] Chafe, W. L. (1970). Meaning and the Structure of Language. University of Chicago Press.

[CHEE88] Cheepen, C. (1988). The Predictability of Informal Conversation. London: Printer.

[CHOM65] Chomsky, N. (1965). Aspects of the Theory of Syntax. Cambridge, Mass.: MIT Press.

[CHOM68] Chomsky, N. (1968). Language and Mind. MIT: Hartcourt, Brace, & World Inc.

[CLAR83] Clarke, D. D. (1983). Language and Action: A Structural Model of Behavior. NY: Pergamon Press.

[CLAY89] Steven, E. C. (1989). The Production of Punctuality: Social Interaction, Temporal Organization, and Social Structure. *American Journal of Sociology,* 95, vol. 3, 659-91.

[COHE81] Cohen, H. (1981). Understanding Social Relationships. Ames: Iowa State University Press.

[COLL85] Collier, G. (1985). Emotional Expression. L. Erlbaum Associates.

[COMR76] Comrie, B. (1976). Linguistic Politeness axes: speaker-addressee, speaker-referent, speaker-bystander. *Pragmatics Microfiche* 1.7: A3. Dept. of Linguistics, University of Cambridge.

[DARW99] Darwin, C. (1899). Expression of the Emotions in Man and Animals. Appleton-century-crofts.

[DESC11] Descartes, R. (1911). The Passions of the Soul. In Haldane, E. L. & Ross, G. R. (ed. & trans.). *The Philosophical Works of Descartes.* Cambridge University Press.

[DEWE94] Dewey, J. (1894). The Theory of Emotion. *Psychological Review,* vol. 1.

[DIJK77] Van Dijk, T. A. (1977). Text and Context: Explorations in the Semantics and Pragmatics of Discourse. NY: Longman.

[DIJK79] Van Dijk, T. A. (1979). Pragmatics Connectives. *Journal of Pragmatics,* 3, 446-456.

[DIJK80] Van Dijk, T. A. (1980). Macrostructures: An Interdisciplinary Study of Global Structures in Discourse, Interaction, and Cognition. Hillsdale, NJ: L. Erlbaum Associates.

[DIJK83] Van Dijk, T. A. & Kintsch, W. (1983). Strategies of Discourse Comprehension. NY: Academic Press.

[DIJK85] Van Dijk, T. A. (ed.) (1985). Handbook of Discourse Analysis. Orlandon: Academic Press.

[DUNC72] Duncan, S. Jr., (1972). Some Signals and Rules for Taking Speaking Turns in Conversations. *Journal of Personality and Social Psychology,* vol. 23, no. 2, 283-292.

[DUNC74] Duncan, S. Jr. & Niederehe, G. (1974). On Signalling that It's Your Turn to Speak. *Journal of Experimental Social Psychology,* 10, 234-247.

[DUNC77] Duncan, S. & Fiske D. W. (1977). Face-to-face Interaction. Hillsdale, NJ: L. Erlbaum Associates.

[EDMO81] Edmonson, W. (1981). Spoken Discourse: A Model for Analysis. NY: Longman.

[ERVI72] Ervin-Tripp, S. (1972). On Sociolinguistic rules: alternation and co-occurence. In Gumprez, J. J. & Hymes, D. (eds.) (1972). *The Ethnography of Communication,* 213-50. American Anthropologist, 66.vi, pt2, Special publ. Washington, D.C.

[FILH84] de Souza Filho, D. M. (1984). Language and Action. Philadelphia: Benjamins.

[FIND79] Findler, N. (ed.) (1979). Associative Networks: Representation and Use of Knowledge. NY: Academic Press.

[FREU53] Freud, S. (1953). Complete Psychological Works of Sigmund Freud, Strachey, J., (ed.). Bantam Books.

[GARF67] Garfinkel, H. (1967). Studies in Ethnomethodology. Englewood Cliffs, NJ: Princeton Hall.

[GOFF55] Goffman, E. (1955). On Face Work: An analysis of Ritual Elements in Social Interaction. *Psychiatry,* 18, 213-231.

[GOFF61] Goffman, E. (1961). Encounters: Two Studies in the Sociology of Interaction. Indianapolis, Bobbs-Merrill.

[GOFF67] Goffman, E. (1967). Interaction Ritual. Anchor Books.

[GOFF69] Goffman, E. (1969). Strategic Interaction. University of Pennsylvania Press.

[GOFF71] Goffman, E. (1971). Relations in Public: Microstudies of the Public Order. Basic Books.

[GOFF74] Goffman, E. (1974). Frame Analysis: An Analysis on the Organization of Experience. Harper & Row.

[GOFF81] Goffman, E. (1981). Forms of Talk. Philadelphia: University of Pennsylvania Press.

[GOOD81] Goodwin, C. (1981). Conversational Interaction. NY: Academic Press.

[GREE86] Greene, J. (1986). Language Understanding: A Cognitive Approach. Philadelphia: Open University Press.

[GRIM75] Grimes, J. E. (1975). The Thread of Discourse. The Hague: Mouton.

[GRIS86] Grishman, R. (1986). Computational Linguistics. NY: Cambridge University Press.

[GROS86] Grosz, B. J. & Sidner C. L. (1986). Attention, Intentions and the Structure of Discourse. *Computational Linguistics,* 12, vol. 3, 175-204.

[GUMP82] Gumprez, J. J. (1982). Discourse Strategies. NY: Cambridge University Press.

[HALL63] Hall, E. T. (1963). The Silent Language. Garden City, NY: Grune and Stratton.

[HALL67] Halliday, M. A. K. (1967). Notes on Transitivity and Theme in English: Part 2. *Journal of Linguistics,* 3, 199-244.

[HASL87] Haslett, B. (1987). Communication: Strategic Action in Context. Hillsdale, NJ: L. Erlbaum Associates.

[HEGE07] Hegel, G. W. F. (1807). Phenomenology of Spirit. Trans. Miller, A. V. (1977). Oxford: Clarendon Press.

[HOBB85] Hobbs, J. R. (1985). On the Coherence and Structure of Discourse. Center for the Study of Language and Information, Stanford University, Technical Report CSLI-85-37.

[HUME78] Hume, D. (1978). A Treatise of Human Nature. Selby-Bigge, L. A. (ed.). Clarendon Press.

[JAFF70] Jaffe, J. & Feldstein, S. (1970). Rhythms of Dialogue. NY: Academic Press.

[JAME84] James, W. (1884). What Is an Emotion. *Mind,* vol. 9.

[JAME90] James, W. (1890). The Principles of Psychology. (2 vols.) MacMillan.

[KELL79] Keller, E. (1979). Gambits: Conversational Strategy Signals. *Journal of Pragmatics,* 3, 219-238.

[KEND67] Kendon, A. (1967). Some Functioins of Gaze-direction in Social Interaction. *Acta Psychologica,* 26, 22-63.

[LABO72] Labov, W. (1972). Sociolinguistic Patterns. University of Pennsylvania Press.

[LEHN78] Lehnert, G. W. (1978). The Process of Question Answering. Hillsdale, NJ: L. Erlbaum Associates.

[LINT36] Linton, R. (1936). The Study of Man. NY: Appleton-Century.

[LIPP22] Lippman, W. (1922). Public Opinion. NY: Harcourt, Brace.

[LITT83] Littlejohn, S. W. (1983). Theories of Human Communication. Belmont, California: Wadsworth Pub. Co.

[LYON80] Lyons, W. (1980). Emotions. Cambridge University Press.

[MANN87] Mann, W. C. & Thompson, S. A. (1987). Rhetorical Structure Theory: A Theory of Text Organization. ISI Reprint series, ISI/RS-87-190.

[MCLA84] McLaughlin, M. L. (1984). Conversation: How Talk is Organized. Beverly Hills: Sage Publications.

[MCKE85] McKeown, K. R. (1985). Studies in Natural Language Processing, volume 2: Text generation using discourse strategies and focus constraints to generate natural language text. Cambridge: Cambridge University Press.

[MEAD34] Mead, G. H. (1934). Mind, Self, and Society. Chicago: University of Chicago Press.

[MEHR81] Mehrabian, A. (1981). Silent Messages: Implicit Communication of Emotions and Attitudes. Wadsworth, Inc.

[MELT71] Meltzer, L., Morris, W. N., & Hayes, D. P. (1971). Interruption Outcomes and Vocal Amplitudes: Explorations in Social Psychophysics. *Journal of Personality and Social Psychology,* vol. 18, 3, 392-402.

[MEYE84] Meyer, B. J. F. (1984). Prose Analysis: Purposes, Procedures, and Problems. In Britton, B. K. & Black J. B. *Understanding Expository Text: A Theoretical and Practical Handbook for Analyzing Explanatory Text,* chapter 2. Hillsdale, NJ: L. Erlbaum Associates.

[MILL63]

Miller, G. A. (1963). Review of Universals of Language, ed. by Greenberg, J. *Contemporary Psychology,* 8, 417-18.

[MILL76] Miller, G. A. & Johnson-Laird, P. N. (1976). Language and Perception. Cambridge, Mass.: Belknap Press of Harvard University Press.

[MORE60] Moreno, J. L., (ed.) (1960). The Sociometric Reader. Glencoe, Ill.: The Free Press.

[MUNN61] Munn, N. L. (1961). Psychology: The Fundamentals of Human Adjustment. Harrap.

[NEIM51] Neiman, L. J. & Hughes, J. W. (1951). The Problem of the concept of role--A re-survey of the literature. *Social Forces,* 30, 141-149.

[ORTO88] Ortony, A., Clore, G. L., & Collins, A. (1988). The Cognitive Structure of Emotions. NY: Cambridge University Press.

[PARE35] Pareto, V. (1935). Mind and Society. Harcourt Brace Jovanovich.

[PARS51] Parsons, T. (1951). The Social System. Glencoe, Ill.: The Free Press.

[PARS53] Parsons, T., Bales, R. F., & Shils, E. (1953). Working Papers in the Theory of Action. Glencoe, Ill.: The Free Press.

[REIC85] Reichman, R. (1985). Getting Computers to Talk Like You and Me: Discourse Context, Focus, and Semantics (an ATN Model). Cambridge, Mass.: The MIT Press.

[ROGE74] Rogers, W. T. & Jones, S. E. (1974-1975). Effects of Dominance Tendencies on Floor Holding and Interruption Behavior in Dyadic Interaction. *Human Communication Research,* 1, 113-22.

[RUES51] Ruesch, J. & Bateson, G. (1951). Communication: The Social Matrix of Psychiatry. NY: Norton.

[RUSS45] Russell, B. (1945). A History of Western Philosophy and Its Connection with Political and Social Circumstances from the Earliset Times to the Present Day. NY: Simon and Schuster.

[RYLE49] Ryle, G. (1949). The Concept of Mind. Hutchinson.

[SACK74] Sacks, H., Schegloff, E. A., Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language,* vol. 50, no. 4.

[SAND87] Sanders, R. E. (1987). Cognitive Foundations of Calculated Speech: Controlling Understanding in Conversation and Persuasion. Albany: State University of New York Press.

[SANF87] Sanford, D. L. & Roach, J. W. (1987). Representing and Using Metacommunication to Control Speakers' Relationship in Natural Language Dialogue. *International Journal of Man-Machine Studies,* 26, 301-319.

[SANF88] Sanford, D. L., and Roach, J. W. (1988). A Theory of Dialogue Structures to Help Manage Human Computer Interaction. *IEEE Transactions on Systems, Man, and Cybernetics,* 18, 567-574.

[SART71] Sartre, J. (1971). Sketch for a Theory of Emotions. Translation by Mairet, Meuthuen.

[SCHA81] Schank, R. C. & Riesbeck, C. K. (1981). Inside Computer Understanding: Five Programs Plus Miniatures. Hillsdale, NJ: L. Erlbaum Associates.

[SCHA82] Schank, R. C. (1982). Dynamic Memory: A Theory of Reminding an Learning in Computers and People. NY: Cambridge University Press.

[SCHA84] Schank, R. C. (1984). The Cognitive Computer: On Language, Learning, and AI.

[SCHE68] Schegloff, E. A. (1968). Sequencing in Conversational Openings. *American Anthropologist,* 70, 1075-1095.

[SCHE73] Schegloff, E. A., and Sacks, H. (1973). Opening up Closings. *Semiotica,* 6, 289-327.

[SCHE77] Schegloff, E. A., Jefferson, G., Sacks, H. (1977). The Preference for Self-Correction in the Organization of Repair in Conversation. *Language,* vol. 53, no. 2, 361-382.

[SCHE79] Schegloff, E. A. (1979). Identification and Recognition in Telephone Conversation Openings. In Psathas, G. (ed.): *Everyday Language: Studies in Ethnomethodology.* NY: Irvington Publishersm Inc.

[SCHI85] Schiffrin, D. (1985). Conversational Coherence. *LANGUAGE,* vol. 61, no. 3.

[SCHI87] Schiffrin, D. (1987). Discourse Markers. NY: Cambridge University Press.

[SEAR83] Searle, J. R. (1983). Intentionality: An Essay in the Philosophy of Mind. NY: Cambridge University Press.

[SHAN14] Shand, A. (1914). The Foundations of Character. Macmillan.

[SKIN61] Skinner, B. F. & Holland, J. G. (1961). The Analysis of Behavior. McGraw-Hill

[SPER86] Sperber, D. & Wilson, D. (1986). Relevance: Communication and Cognition. Cambridge, Mass.: Harvard University Press.

[SPIN83] Spinoza, B. (1883) The Ethics. In Elwes, R. H. M. (trans.). *The Chief Works of Benedict de Spinoza,* vol. II. George Bell & Sons.

[SPIT45] Spitz, R. (1945). Hospitalism: Genesis of Psychiatric Conditions in Early Childhood. *Personality Study of the Child,* 1, 53-74.

[TANN84] Tannen, D. (1984). Conversational Style. Norwood, NJ: Ablex Pub. Corp.

[TAYL87] Taylor, T. J. & Cameron, D. (1987). Analysing Conversation: Rules and Units in the Structure of Talk. NY: Pergaman Press.

[THRA80] Thrane, T. (1980). Referential Semantic Analysis: Aspects of a Theory of Linguistic Reference. Cambridge University Press.

[VIRK91] Virkar, R. S. (1991). PhD. Dissertation (in progress), Virginia Polytechnic Institute, Computer Science Department.

[WALK78] Walker, D. E. (1978). Understanding Spoken Language. North Holland.

[WATZ67] Watzlawick, P., Beavin, J. B., & Jackson D. D. (1967). Pragmatics of Human Communication. Norton.

[WARN91] Warner, R. (1991). MS Thesis (in progress), Virginia Polytechnic Institute, Computer Science Department.

[WEBE47] Weber, M. (1947). Theory of Social and Economic Organization. Trans. A.M. Henderson and Talcott Parsons. NY: Oxford University Press.

[WEST75] West, C. & Zimmerman, D. H. (1975). Small Insults: A Study of Interruptions in Cross-Sex Conversations Between Unacquainted Persons. In Thorne B., Kramarae C., and Henley N. (eds.). *Language, Gender, and Society.* Rowley, Mass.: Newbury House Publishers.

[WILL76] Willis, F. N. & Williams, S. J. (1976). Simultaneous Talking in Conversations and Sex of Speakers. *Perceptual and Motor Skills,* 47, 1067-1070.

[WILS86] Wilson, T. P. & Zimmerman, D. H. (1986). The Structure of Silence between Turns in Two Party Conversation. *Discourse Processes,* 9, 375-390.

[YULE85] Yule, G. (1985). Study of Language. NY: Cambridge University Press.

[YULE83] Yule, G. & Brown, G. (1983). Discourse Analysis. Cambridge University Press.

[YNGV86] Yngve, V. H. (1986). Linguistics as a Science. Bloomington: Indiana University Press.

[YNGV70] Yngve, V. H. (1970). On Getting a Word in Edgewise. Papers from the sixth regional meeting of the Chicago Linguistic Society, Chicago: Chicago Linguistic Society.

[ZIMM75] Zimmerman, D. H., & West, C. (1975). Sex Roles, Interruptions and Silences in Conversation. In Thorne B., and Henley, N. Language and Sex: Difference and Dominance. Rowley, Mass.: Newbury House Publishers.

Vita

Ahmed Tewfik Bouzid was born in Casablanca, Morocco on the 5th of October, 1963, of Touria Alami and Abdelkader Bouzid. He began his education at the primary school of "Hadikatou l'Hourria" in Algiers, received his secondary education from "The Gharmoul Polytechnic Lycee," and his high school education from the "Lycee Rochambo," in Bethesda, Maryland.

In 1983, Ahmed entered the George Washington University to pursue a Bachelor degree in Computer Science. Four years later he graduated with the sought degree and very proudly joined the illustrious Computer Science Department at Virginia Tech.

After many false starts and close calls, Ahmed somehow managed to get through the requirements for his MS degree and, God knows how, wrote up something that was consequently accepted as a thesis (and which you now have in your hands). Glowing and flushed by his unexpected "success," Ahmed dared destiny (and common sense) and decided to join the PhD program. He is currently, as we speak, working on his PhD degree, and already he bitterly and deeply regrets his decision.