Corpus Collection and Analysis
Preliminary Results

PLUS
A Pragmatics-Based Language Understanding System

ESPRIT P5254

November 18, 1991

CAP Gemini Innovation CAP Gemini SCS BeCom
ITK LIMSI Omega Generation UMIST
University of Bristol University of Göteborg

Corpus Collection and Analysis

Jens Allwood, Philippe Gobinet, Christophe Godin, Björn Haglund, Kristiina Jokinen, Torbjörn Lager, Didier Pernel, Violaine Prince, Harriet Dahlgren (ed) and Joakim Nivre (ed)

Abstract:
The present volume is a collection of papers devoted to the collection and analysis of dialogue corpora in PLUS.

DOCUMENT HISTORY

Date	*Version*	*Author*	*Comments*
25-10-1991	1.0	Joakim Nivre, Harriet Dahlgren	
15-11-1991	1.1	Joakim Nivre, Harriet Dahlgren	Revisions in chapters 3 and 4, and in appendices A and C. Minor corrections in chapters 1 and 6. Bibliography added.

Document ID : W1-Corpus Collection and Analysis
Issue date : November 18, 1991
Version : 1.1
Status : Final
WP number : 1
Distribution : Corpus
Location : UG /plus/WP1/corpus/corpus.tex
Keywords : Corpus Collection, Corpus Analysis, Dialogue, Wizard of Oz

Contents

Chapter 1

Introduction

1.1 Corpus Studies in PLUS

The goal of the corpus studies, as described in the PLUS technical annex, is to provide the project with relevant information for the design of system components. This concerns all the three main components of the system: NL engine, dialogue manager, and knowledge bases. In addition, the corpus studies will provide useful information for the test and evaluation tasks.

1.1.1 Corpus Collection

In designing the corpus collection task, the following parameters were considered:

- **Languages:** Since portability to new languages is a major concern of the PLUS project, it is important to try to separate, as far as possible, language-dependent and language-independent aspects of dialogues. Hence, dialogues were collected in three different languages: *French* (by CAP and LIMSI), *English* (by UMIST), and *Swedish* (by UG).

- **Dialogue types:** By *dialogue types* we mean different configurations of participants and communication channels. First, we have so-called *Wizard-of-Oz terminal dialogues* (often abbreviated WOZ below), where the subject (a human being) interacts through a terminal with what he believes to be a computer system. This is the most important type for the PLUS project, since it is the one which most closely resembles the dialogues that will be possible with the final system. However, for comparison, we have also collected *human-human terminal dialogues* and (only in Swedish) *human-human telephone dialogues*. The comparisons are important *inter alia* because many of the theoretical approaches (e. g.

in pragmatics) that will be exploited in PLUS have been developed mainly with human-human (especially spoken) communication in view.

- **Scenarios:** The term *scenario* refers here to the specific dialogue task that the subject is confronted with. The scenarios have been chosen from the same subdomains of the Yellow Pages that have been selected for the demonstrator, namely: *car hire*, *restaurants*, and *personal insurance*. (The scenarios and instructions that were given to the subjects can be found in appendix A.)

- **Subjects:** For the selection of subjects it was not considered necessary to use any special sampling technique in order to ensure representativity. Instead, each site was instructed to use whatever subjects were available and considered suitable. All subjects were asked to fill out a form with personal characteristics and provide an evaluation of the system (The questionnaires are exemplified in appendix B).

On the basis of these considerations, it was decided that for each combination of *language*, *dialogue type*, and *scenario*, dialogues from at least 10 subjects should be collected during the first phase of the project.

Software tools

As none of the software tools available for Wizard-of-Oz and human-human terminal dialogues was considered wholly adequate for the purposes of the project, it was decided that software tools had to be developed as part of the corpus collection task. Since, at the outset, the different sites involved had rather different conditions in terms of equipment, databases to be used as support, etc., it was decided that it would be most convenient for each site to develop their own tools, suited to their specific situation.

Collected Corpora

Below we present, for each language involved, the corpora collected during the first phase of the project.

French In total, 75 dialogues have been collected. 43 were Wizard-of-Oz terminal dialogues (16 car hire, 15 restaurant, 12 insurance), and 32 were human-human terminal dialogues (10 car hire, 12 restaurant, 10 insurance). The total number of subjects was 56. They belonged to three main socio-professional categories: students (24), researchers/engineers (16), and administrative employees (16).

English 98 dialogues have been collected altogether. Of these, the first 10 were Wizard-of-Oz dialogues collected as part of a pilot study (5 car hire scenario, 4 restaurant scenario, 1 insurance scenario) and another 10 were dialogues collected on other scenarios than car hire, restaurant and insurance.The remaining 78 dialogues consist of 46 Wizard-of-Oz (17 car hire, 17 restaurant, 12 insurance) and 32 human-human terminal dialogues (11 car hire, 11 restaurant, 10 insurance). The subjects were 1st and 2nd year students, postgraduates, staff from administration, bank clerks and radio representer, aged 20-59. The total number of subjects was 30, of which 5 took part in the pilot study.

Swedish A total number of 106 dialogues have been collected. Of these, 12 were collected as part of a pilot study (12 human-human telephone, 4 human-human terminal). The remaining 90 dialogues consist of 30 Wizard-of-Oz dialogues, 30 human-human terminal dialogues, and 30 human-human telephone dialogues. In each type, the three scenarios (car hire, restaurant, insurance) are represented by 10 dialogues each. The total number of subjects (excepting the pilot study) was 20. 19 were students, and 1 was an administrative employee.

1.1.2 Corpus Analysis

As stated in the PLUS technical annex, the corpus task in the first phase of the project was primarily concerned with the *collection* of dialogue corpora, whereas the major part of the *analysis* are to be carried out within tasks devoted to more specific topics in the areas of NL, pragmatics, and knowledge bases.

However, some preliminary work on analysis was also carried out in the early phase of the project. It is on these preliminary explorations of the PLUS dialogue corpora that we report in this volume.

1.2 The Present Volume

The present volume is a collection of papers devoted to the collection and analysis of dialogue corpora in PLUS. Most of these papers are early contributions dating from the first phase of the project. They have only been minimally edited for inclusion in this volume, and no real attempt has been made to turn the collection into a homogenous report.

It should therefore be emphasized that the present volume is *not* a first version of the corpus deliverable, but rather a collection of working papers reporting on preliminary results from different stages of the project. It should also be noted that further

analysis of the dialogue corpora is currently being performed within different tasks of the project, notably tasks dealing with pragmatic rules and principles, world and application models, and lexicons and grammars for the NL engine.

Let us now briefly describe the different contributions to this volume:

- Chapter 2, written by Philippe Gobinet (CAP), Christophe Godin (CAP), and Didier Pernel (LIMSI), presents the French corpus collected at CAP and LIMSI. The emphasis is on collection rather than analysis, although some general remarks about analysis are provided towards the end.
- Chapter 3, by Violaine Prince (LIMSI), reports on a preliminary analysis of the French WOZ corpus. The results are mainly descriptive statistics on a large number of variables concerning dialogue structure and dialogue strategy.
- Chapter 4 is written by Kristiina Jokinen at UMIST and contains preliminary results from the study of English WOZ dialogues. The aim is to provide a general characterization of the dialogues and to produce illustrative examples of interesting phenomena for further analysis.
- Chapter 5, by Björn Haglund and Torbjörn Lager at UG, presents a preliminary analysis of the Swedish WOZ corpus. The approach is similar to that of Jokinen in chapter 4.
- Chapter 6, finally, is a study by Jens Allwood and Björn Haglund at UG. It is an attempt to apply Allwood's model of communicative activity analysis[1] to WOZ dialogues (using data mainly from the English corpus). It is the most recent contribution in the volume and it is clearly focused on the analysis of pragmatic aspects of dialogues.

The volume has been edited by Joakim Nivre and Harriet Dahlgren at UG. Joakim Nivre has also written the introductory chapter.

1.3 A Note on Notation

In the presentation of dialogue excerpts, contributions from the system (wizard or human) are generally preceded by an S (for System) while contributions from a subject in the experiment is preceded by a U (for User). Isolated utterances or parts of utterances without a preceding S or U are from subjects/users unless otherwise indicated. In addition to the examples cited in the text, longer dialogue excerpts from different corpora can be found in appendix C.

[1]This model is decribed, for example, in Allwood [3]

Chapter 2

Collection of the French Corpus

2.1 Software Description

A special software has been developed by CAP for the collection of the French corpus carried out at CAP and LIMSI. This software is intended to provide two things:

- a basic interface for the subject (typically connected to a VT220 or any equivalent simple terminal),
- a graphic interface for the wizard which helps him to access the real-size YP database through the Minitel while communicating with the subject.

The whole software is thus divided into two parts corresponding to two different processes. The subject process runs on one machine and is driven by the "server-process" of the wizard which runs on another machine of the network.

A session is defined as a sequence :

- subject-login into the system,
- sequence of typed-interactions between the subject and the system,
- subject-logout.

Each session, possibly composed of one or more dialogues, is recorded in a "spyfile" whose name summarizes the basic session features (location of the recording, date, session identifier, etc.)

A spyfile contains a header which provides general information on the session and a body which consists of a sequence of textual interventions, each labelled by either

"subject" or "system". Time information is also provided at each step : absolute time counts time from the login of the subject while relative time corresponds to the period between two consecutive interventions.

2.1.1 Subject Interface

The subject interface has been designed in the simplest way. The subject process is run by the command"SP" which first tries to connect to the server process. If it fails, one must try again. If not, a welcoming message appears on the subject's screen and invites the subject to start querying the system. Before any character is typed on the keyboard by the subject, the system (wizard) can send a message to the subject, but as soon as the subject has typed the first character, the system can no longer interrupt him until the he hits the special key "ESC" which stands for "end of subject message". Any action attempted by the wizard during the subject's turn is simply discarded. Each character from the subject is sent to the wizard as soon as it is typed. The subject also has the possibility to send multi- line queries since a hit of the carriage return character "CR" only breaks the current line and starts a new one. The subject HAS to hit "ESC" to send his question to the system.

2.1.2 Wizard Interface

We have developed a complete graphic interface for the wizard's needs. It runs under the X-window environment on a SUN Sparc station and uses the standard MOTIF graphic library facilities. Let us briefly describe the graphic interface. It is composed of 4 text-windows :

- a window which displays the YP-querying system output (Read-Only window)
- a window that enables the wizard to query the YP-server (Read-Write window)
- a window which displays what the subject actually sees (Read-Only window)
- a window that enables the wizard to prepare his future response to the subject (Read-Write window)

The main advantage of this graphic interface is that the cut/copy/paste facilities are available between these windows. Hence, the wizard can query the YP-server while preparing his answer. As soon as the desired information is available in the YP window, it can be copied to the response-window, modified, and then sent to the subject as a coherent block.

2.1.3 Differences with the UMIST software

- We wanted to be able to really access to the YP data-base commonly used in France thanks to the "minitel" (phone number 11). This YP-service is queried by the wizard while the dialogue is proceeding with the subject. The graphics facilities help the wizard to manage simultaneously a lot of information which reduces the mean response time of the wizard to the subject. This appears to be of crucial importance for WOZ experiments as a real-size standard YP data-base is actually used. Still, the system is evaluated quite slow in general by the majority of the subjects.
- The output of the YP-server is transferred by cut and paste to the subject.
- After each subject *or* wizard intervention, any of them can react and continue the dialogue (i.e. *either* the subject *or* the wizard). In the UMIST software the switching between wizard and subject is necessary and the potential conflicts are bufferized until the one who has the turn decides to send his message. The bufferized data are then immediately flushed to the screen.
- The data sent by the wizard are discarded in case of conflict (the subject has started to type something but has not yet hit the "end of message" key).

2.2 Corpus Collection

The actual corpus collection was mainly performed at LIMSI. 56 subjects (47 at LIMSI and 9 at CAP) participated, and 75 dialogues were collected (some subjects participated in several simulations).

The subjects belong to three main socio-professional categories:

- 24 subjects (25 dialogues) are students,
- 16 subjects (20 dialogues) are either researchers or engineers (this group is referred to below as "researchers"),
- 16 subjects (30 dialogues) are administrative employees (secretaries, accountants, etc.).

The collection concerned both Human-Human dialogues (20 subjects, 32 dialogues) and so called Wizard of Oz simulations (36 subjects, 43 dialogues). For these two kinds of simulations, the experimental conditions were identical, only the oral instructions given to the subjects differed. The subjects were asked to query the system/human in order to perform one or several simple tasks described in written scenarios. These scenarios dealt

with the three agreed domains, namely car hire, restaurants and personal insurance (see appendix A for further details). Once the simulation was finished, the subjects were asked to fill in an identification form and to roughly evaluate the simulation. In the case of Wizard of Oz simulations, a more detailed evaluation questionnaire was proposed (see appendix B for further details).

The following tables give a more precise description of the distribution across dialogue types, tasks, and subject groups:

- Wizard of Oz simulations:

	students	administratives	researchers
car hire	8	6	2
restaurant	8	7	0
insurance	5	4	3

- Human-Human simulations:

	students	administratives	researchers
car hire	1	4	5
restaurant	1	5	6
insurance	2	4	4

2.3 Corpus Analysis

2.3.1 Preliminary Remarks

Here are some general remarks related to the corpus collection:

- All the subjects at CAP who were involved in a H-H dialogue, actually believed that they were in front of a prototype system and not communicating with a human operator. This phenomenon occurred even when the instructions given to the subject where particularly explicit.
- There is only a limited number of eccentric dialogues or dialogues trying to evaluate the limits of the system.
- It appears to be impossible to tell whether a dialogue has been collected within a H-H or a WOZ context by looking only at the output dialogue.
- LIMSI justified their long response-time by referring to long delays due to network problems in accessing the database, while CAP invoked the prototype status of the system. Both explanations were equally accepted by the subjects. In any

case, the subjects tended to attribute the long response-time to the navigation of the system inside the data-base and *never* to the parsing or to the understanding of their messages.

- The number of dialogues is not sufficient to extract data from it with statistical validity.
- The subjects all agree on two points. The system is rather convivial but its speed should be augmented notably in future versions. They all expected naturally from a convivial system to be able to give more information than the (name, address, phone number) that is provided by the current system.

2.3.2 Corpus Analysis

At the time of writing, the analysis task has only just started, but we list below a set of interesting topics that we have decided to deal with in the corpus analysis (see also chapter 3 of this volume). One of our main concerns will consist in identifying the use of pragmatic, contextual and world knowledge in the dialogues.

This is a list of topics we have selected a priori for the corpus analysis:

- Pragmatics
- What kinds of events and parameters does pragmatics depend on?
- Ergonomy
- Meta-dialogue
- Strategy: identify subgoals, number of strategy changes, beginning and end of the dialogue, conflicts with system knowledge
- Subject's beliefs about the system's capacity or behaviour
- Structural organization of the dialogues
- Nature of the subject's questions (strongly focused or evasive)
- Vocabulary
- Syntax
- Comparison of the H-H and WOZ collections
- Adaptation of the subject to the system

The kind of information that can be extracted from the corpus should be discussed and clearly defined in conjunction with UG and UMIST.

2.3.3 Open Questions and Future Work

The experience of the corpus collection gives rise to a few questions for future work:

- With our current collection tool, the wizard cannot interrupt the subject after he has started to write. It would be interesting to test the ability of the system to actually interrupt the subject at any time.
- Because of the corpus size (interesting enough but still rather limited) we wonder whether a full lexicon can be extracted from this corpus. If yes, should it be completed in another way to have a sufficiently large coverage of the application vocabulary?
- The same question arises for the grammar design.
- Will we need to collect more data? The confrontation of the collected corpora with the aims of the PLUS project may reveal insufficiencies in the information contained in the corpora (for instance lack of pragmatic inferences). If we had to collect corpora again in an equivalent context (PLUS), would we carry out the collection in the same way? What would change? Does this modification have important consequences for the PLUS project? All this should be discussed after the first corpus analysis.

Chapter 3

Preliminary Analysis of the French WOZ Corpus

3.1 Introduction

We have run a preliminary survey of 16 dialogues of the WOZ type, taken from the corpus collected at LIMSI. In this analysis, we have classified dialogues according to some general criteria specified in chapter 2 of this volume and also according to more specific criteria listed below in this chapter.

The aim of this first step in the analysis was to investigate, in the collected dialogues, some dialogue features which are considered in dialogue models such as Vilnat [24] and Luzati [15], based on the work of Roulet and Moeschler. These features could be classified as:

- Specific to the *interactive part* of the dialogue (see Luzati's model). We were interested in recognizing and modelling interventions that could be termed "meta-dialogue".
- Specific to the *intentional and functional part* of the dialogue (see Vilnat's model). User's goals and user's strategies are of importance here, especially as they could provoke modifications in the dialogue strategy.
- Finally, some features that are often assumed to be part of human conversation and which concern structural information about dialogues.

The existence or absence of these features has been considered. Moreover the pervasiveness of some of them seemed to us a significant point worth investigating. Our study was motivated by a desire to determine relevant aspects of the application dialogues before setting up our models. We believed this was a proper experimental attitude.

Therefore the importance of the figures given hereafter is only qualitative. It is a way of expressing the existence, the absence or the relative impact of some dialogue aspects. We have restricted our survey to dialogue phenomena regardless of particular linguistic phenomena (such as ellipsis, misspelling, anaphora, metonomy, etc.).

3.2 General Remarks

These remarks concern the totality or the majority of the analysed dialogues.

- Dialogue structure: The dialogues are generally structurally simple. That is, they do not contain many embedded structures. When a question arises, the exchange is generally closed after a few (two or three) interventions: no heavy backtracking in the heap of unanswered questions.
- Users' demands: It seems that users are very keen on extra information such as the price of the service, a cost/quality comparison All these demands we have considered as not belonging to a "pure" YP information system. It is as if they feel that a computerized system, being necessarily more sophisticated than a manual one, has also the task to improve the service to the customer.
- Eccentric dialogues: The report says that only a few dialogues could be labelled as eccentric. The survey of the 16 dialogues shows that although none of them is eccentric, most of them contain some eccentric parts: a sentence or two, concerning the evaluation of the system, or expressing a will to track the limits of the system.

3.3 General Criteria

3.3.1 Meta-dialogue

We distinguish two types of meta-dialogues: an ergonomical meta-dialogue and a general meta-dialogue.

Ergonomical Meta-dialogue

The ergonomical meta-dialogue is about the system's properties. We have determined the following cases:

- Speed/wait characteristics. We distinguish two types:

  - user's remarks: sentences issued from the user such as: "you are slow", "please hurry", or "wait till I finish noting down" or "I am noting these" etc. (SU)
  - system waiting sentences or justifications (SS)

- Morphological features. If issued from the system, they concern the form with which the user has to formulate a query. if issued from the user they are spelling correction justifications. (MF)
- System capacity. The system indicates to the user how many queries it might answer or when the user may state his/her request. Or the user indicates to the system what are the relevant queries he/she wants answered. (SC)
- System's behaviour. Remarks from the user concerning the apparent behaviour of the system, such as "make an effort", "I want an answer", etc. (SB)
- User's behaviour. Remarks from the user making an "opening" towards a change of goal or of strategy. System's encouragement for precision. (UB)

General Meta-dialogue

We have put into this category all requests about spelling ("how do you spell such and such?" or "is 'insurance' with an a or an e?", etc.) and the variable is noted (SP). We have also found requests for and expressions of extraneous information, such as curiosity about the capabilities of the system, expression of discouragement or happiness about the performance of the system, noted (CU).

3.3.2 External Dialogue Structure

We have tried to recognize the following operational variables: dialogue/exchange opening, dialogue closing. We also give a general variable about dialogue length (number of utterances).

Dialogue Opening

Two major types of opening:

- Formal human explicit opening, such as "good morning" or any polite form preceded by "good morning" (familiar expressions such as "hi" also belong to this category) (FH)
- Implicit opening. This type is subdivided into:

- direct polite opening, such as "I would like to know I" (DP)
- direct laconic opening, such as " What is I" , "do you have I"(DL)
- key word opening, such as "Insurance" (KW)
- "story" opening: when a subject expresses his/her needs, reasons, or justifications, such as "I am looking for", "I need to go..." (S)

Dialogue Closing

In the same way, one may define the following divisions.

- Formal human explicit closing, such as "goodbye", "so long" or a polite form + "goodbye".(FH)
- Implicit closing. This type is subdivided into:
  - direct polite closing: "thank you" (without goodbye) (DP)
  - direct laconic closing: such as "no", "that's all" (DL)
  - brutal stop: by system shutdown or no answer (BS)
  - "story" closing: when a subject expresses his/her needs, reasons, or justifications, such as "I have already had what I wanted", etc. (S)

Dialogue Length

We indicate the mean number of utterances for these 16 dialogues, the maximum range, and the minimum.

3.3.3 Dialogue Strategy

We have determined three main operational variables: changes in goals, changes in strategy, and conflict with the system knowledge.Subgoal identification is a variable we intend to study but which has not been examined so far.

Changes in Goals

Both user's and system's goals are identified that are issued from dialogue decomposition. We study the apparent changes in goals.

- User's goals : these involve brutal changes (not the same topic) as well as slight changes (goals compatible with or analogical to the preceding one) without any introducing expression that will indicate a preference or a position. (UG)
- System's subgoals: whenever they are needed. Precision about user's preferences, retrieval goals and so on. (SG)

Changes in Strategy

In this category we put utterances expressing user's change of mind. The user expresses his/her preference or his/her insatisfaction within a context, he/she changes his/her goal by explicitly giving the corresponding subgoal. We consider it as shifting from a general information retrieval strategy to a criterion-based specific strategy. (CM)

Conflict with System/User Knowledge

In this category we put utterances expressing:

- Direct contradiction with the system knowledge (issued by the user), such as "this adress is false", "this is not true", "your information is out of date", " I have been told otherwise", etc. (DC)
- Insistence on a particular request. The user repeats the request with another form or in extenso (UI)
- Insistence on the system's behalf about an ambiguous request (SI)

3.3.4 Nature of Subject's Questions

In this category we are interested in the following features:

- Precision/vagueness in user's request. Too general, or too abstract. (VD)
- Type of utterance. This concerns the request which cannot be specified as a direct request. In this category, we find two subtypes:
  - Justification from the user for asking a question. Telling a story about his/her request. This variable concerns utterances which are not openings for dialogues. (JU)
  - Explanation requests from the user. He/she does not understand the system's replies. (ER)

3.3.5 Other

Many other domains are just being investigated. Among them inconsistent utterances, subject's beliefs about the system capacity and behaviour, vocabulary, syntax, etc.

3.4 Results

3.4.1 Meta-dialogue

Ergonomical Meta-dialogue

SU	SS	MF	SC	SB	UB
1	4	2	3	3	3

Remark: These values have been extracted from 10 dialogues out of 16. Not all dialogues contain ergonomical meta-dialogues but it nevertheless seems frequent.

General Meta-dialogue

SP	CU
1	4

Remark: These values concern 3 dialogues out of 16. But a quick look on other dialogues have made us notice some sentences that will enhance these values. Total for meta-dialogue expressions: 21. Total number of dialogues concerned: 11 out of 16.

3.4.2 External Dialogue Structure

Opening/Closing

OPENING					CLOSING				
FH	DP	DL	KW	S	FH	DP	DL	BS	S
3	2	2	3	6	6	7	1	2	0

Remarks:

- We notice a "story" like opening dominating slightly. Didier Pernel remarks that most openings, even when not directly "story-like" are rather vague or general in term of their content.
- A truly dominant FH/DP for closing.

Dialogue Length

MEAN VALUE	MAXIMUM	MINIMUM
28	50	10

Remarks:

- These figures correspond to the number of utterances in the dialogue.
- Two dialogues with high figures (40 and 41 utterances) correspond to chaining two scenarios.
- Nevertheless, the longest dialogues (48 and 50 utterances) are single subject (scenario).
- We find these figures relatively high.

3.4.3 Dialogue Strategy

Changes in goals

UG	SG
15	1

Remarks:

- These figures concern 4 dialogues out of 16.
- One of these dialogues involves 4 changes in user's goals. Another involves 7 changes.
- One of these dialogues involve only brutal changes.

Changes in strategy

CM	
2	

Remarks:

- We intend to define more operational variables for the "changes in strategy" field.
- These changes of mind concern 2 dialogues out of 16.

Conflict with System/User Knowledge

DC	UI	SI
1	7	3

Remarks:

- These figures concern 8 dialogues out of 16.
- Some dialogues involve many utterances corresponding to one or many of these three variables.

Total for dialogue strategy traces: 29. Total number of different dialogues involved: 10 out of 16.

3.4.4 Nature of Subject's Questions

VD	JU	ER
8	5	5

Remarks:

- For VD we have processed 6 different dialogues (/16).
- For ER 5 different dialgoues (/16).
- For JU 3 different dialogues (/16).
- Some dialogues instantiate many variables.

- Some utterances instantiate variables in this category and variables in other categories: ER and SP.
- Some utterances instantiate two variables in this category, e. g. VD and JU.

3.5 Conclusion

The conclusion of this first "horisontal" analysis can be stated as follows. First, the interactive part of the communicative action, expressed by the dialogue, seemed to be very present in this WOZ experience. Its salient features, described by what we called operational variables, are compatible with the assumptions made by Luzati as requirements for his model. Therefore, we believe that this model is applicable to account for and predict this dimension of dialogue. We think it embodies a portion of dialogue pragmatics which is related to the Gricean maxim of cooperation. Second, the intentional part of the dialogue is also active in the analysed data. Its features and their relative importance have made us recognize the necessity of a *planning mechanism* to adapt to changes in goals and strategies, a *topic and focus search mechanism* to determine the relationship between new goals and the previous ones, and a *belief and expectations model* which is able to account for shared or non-shared knowledge between the system and the user, and for the user's intentions. These look like requirements for the PLUS system, based on empirical studies. Finally, the attempt to generalize some ideas about a recurring syntax of the dialogues (the recognition of openings and closings, the nature of exchanges whether topical or ergonomical, the possibility of having many exchanges as confirmed by long dialogues) which could be formalized in the PLUS system as *dialogue structure rules.*

Chapter 4

Preliminary Analysis of the English WOZ Corpus

4.1 Corpus Collection

The English corpus collected at UMIST contains 46 Wizard-of-Oz dialogues (17 on car hire and on restaurant scenarios and 12 on insurance scenarios) and also 32 human-human terminal dialogues (11 on car hire and restaurant scenario and 10 on insurance scenario). Comparision to other type of information seeking dialogues is offered by ten dialogues made on scenarios outside the YP scope (three deal with a scenario where the subjects were asked to find a suitable course to continue their studies in the educational institutions in Manchester, and seven with a conference scenario where subjects had to try out an automated conference booking service to find out relevant information for participating a conference in Paris). All the dialogues are written terminal dialogues.

The dialogues were collected using a computer communication system written by Afsaneh Haddadi at UMIST. This Software is a prototype allowing human-human communication between two different Sun workstations, or between a Sun and a dumb terminal or terminal emulator, and it is described in Haddadi (1991). The software enables the input from the user to be shown on the wizard's screen as it is being typed, thus giving the wizard time to plan responses. The responses by the wizard, on the other hand, are buffered, so that the user cannot see the process of message composition. This gives the user the illusion of a computer being at the other end. The partners cannot interrupt each other, i.e. after having sent the message by hitting the ESC-key, the participant has to wait until the other has responded and given back the turn. On this point the software was similar to the Swedish one, but differs from the software used in the French corpus collection.

The dialogues were recorded into a file. The file contains the user and the wizard

turns, and it is headed by the start time. The user's mistypings and corrections are also recorded. The file also contains time information about the length of each turn.

The aim of the UMIST corpus collection was to collect data rather than to model a system, the properties of which were at that time somewhat unspecified. Consequently, the restrictions put on the wizard's side in the WOZ-studies were general. The wizard was asked to be as helpful as possible given the limitations of the Yellow Pages database.

A set of responses for the wizard was stored in a file, but the wizard could also freely phrase her replies, if the ready-made replies were unsuitable in the context. The wizard was to use language that were natural, yet not too human-like in order to maintain the illusion about the computer responses.

The wizard also had a small database that contained the appropriate information about car hire firms, restaurants and insurance companies needed in the dialogues. The database was compiled on the basis of the Greater Manchester YP, concentrating on the information in the areas that were relevant for the scenarios. As the YP does not divide restaurants into type categories, a local restaurant guide was used to help in this categorization.

Our purpose was to have as different subjects as possible, and especially subjects that were would-be users of the future system. Half of the subjects thus comprised staff from the administration, bank clerks, radio representers, electrical engineers, and a nurse. The other half consisted of 1st and 2nd year students and postgraduates, who were easily available in the academic environment. The age range of the subjects was 20–59.

Statistics of the UMIST corpus studies:

		Dialogues	People
PLUS	Pilot (WOZ)	10	5
	WOZ	46	19
	Human	32	11
	Total	88	35
OTHER	Pilot (WOZ)	2	1
	WOZ	7	6
	Human	1	1
	Total	10	8
TOTAL		98	36

The setup of the dialogue session was as follows: the subjects were given the instruction sheet that contained general information about the experiment and how to use the system, as well as the scenario sheet, and then left to work on their own. A supervisor was available to assist, if problems arised, but did not otherwise intervene in the experiment. After the experiment session, the subjects were asked to fill in a questionnaire

(see appendix B).

The screen that faced the subjects in the very beginning of the session contained information about the YP service system, about how to quit, and a prompt to type the first question as follows:

```
**********************************************************
*                                                        *
*          Welcome to the YELLOW PAGES Directory         *
*                                                        *
**********************************************************

What can I do for you? (To quit the system type quit.)

YP  1:
```

The question on the screen was understood as the very first system turn and consequently, the user turn as a response to this. The dialogue files, however, did not record the first system turn (see the sample example in appendix C). It should also be mentioned that the screen did not contain greetings, and this may be one of the reasons why the dialogues did not contain greetings that normally start and end conversations: greeting was not invited by the system and obviously it was also considered unnecessary by the users who thought they were talking to a machine that does not understand — or require — such subtleties of human communication. We will come back to this when studying the structure of the dialogues.

4.2 Nature and Type of the Dialogues

4.2.1 Simulation of the Future System's Capabilities vs. Data Collection

We have primarily aimed at collecting data and providing information especially on the user's behavior. There were three reasons for this. First, in order to determine external and internal requirements for the PLUS system design, we considered it important that attention is also paid to the requirements that potential users of an electronic YP service impose on the system's robustness and user-friendliness.

Second, the simulation of the future system's capabilities was considered somewhat difficult a task, the reason being the vagueness in general requirements. Problematic issues dealt with the kind of behavior that the system should exhibit:

- the amount of initiativeness that the system should take,
- the type and nature of inferences that the system should be able to make, i.e. how deeply it should use the world knowledge and reasoning in understanding user utterances,
- the type and nature of information w r t the understood goal expressed in the first user turn, e.g. how well the system knows about the different issues that concern car hiring or personal insurance.

We decided that the wizard should not take initiative except in asking specification questions, that the wizard should use world knowledge that is considered general to the domain, and that the plan that the system is able to follow should mainly be based on the YP database organization.

Third, from the methodological point of view, the simulation may also easily end up in a wicked circle. The modelling of the future system's behavior is an important part of the experiment design, but it will also constrain the data that is to be obtained, and consequently, the dialogues may not provide independent data about possible real dialogues that may take place, and on the basis of which the future system's design is to be built, but something that fits to our predefined conception about how the dialogues *should* go.

4.2.2 The Subjects' Evaluation of the Dialogues

All the subjects believed that they were talking to a computer. Even though the wizard sometimes used very sophisticated and human-like language, the users accepted this as being an indication of the new advanced prototype system that can understand natural language. On the other hand, in a couple of cases, when the wizard made obvious mistakes like sending the wrong type of information, the users simply regarded these as system errors that commonly occur when one is dealing with computers. Thus, features that professional computer scientists may consider too human-like, actually supported the subjects in their belief that it was a computer that they were talking to.

Interestingly, the users seemed to be more concerned about quick and exact replies than fluent dialogues. Obviously, because fluent dialogues were something that they did not expect a computer to master; usually they were surprised of the capabilities of a computer to understand and produce such a human-like language. However, even if the subjects acknowledged the naturalness of the system responses, they regarded the long response time as the main factor in distinguishing their computer conversations from real human-human conversations. Thus, in building a robust system, considarable amount of attention is to be paid to fast and efficient algorithms (cf. similar results also in the French corpus collection).

4.3 Evaluation of the Dialogues

4.3.1 General Aspects

The wizard responses may not always be the most fortunate ones, and one might think of other alternatives as well. We are aware of the impact that the system's responses have on the subject's queries. Especially the beginning of the conversation is essential: "explicit, non-polite conditions elicit the most concise requests, while inexplicit but polite conditions elicit polite responses most commonly".[1]. However, in many cases, the dialogues produced seem to depend more on the subject's imagination than the system responses, and some of the subjects indeed understood the experiment as a situation where they could try out the system's limits.

4.3.2 The beginning and the end

As mentioned, the dialogues did not contain greetings, not even the human-human ones. Only one subject used *bye* before quitting her WOZ-scenarios. The lack of conventionalized start and end signals of a dialogue may be due to the software, but one should notice that greetings were not common in the French and Swedish corpora either, even though the wizard explicitly greeted the user. Greetings must thus be considered optional at most in this context, and an explanation for this can be found in the fact that the experiment was not considered a genuine conversational situation.

How to start

The beginning of the dialogues was usually strightforward, and the dialogue types did not seem to differ in this respect: both WOZ-dialogues and human-human-dialogues started directly with the specification of the query.

The start was formulated in four different ways:

1. An (indirect) conventionalized request
   a. Interrogative *Can you tell me of any car hire firms in Entwistle?;*
   Can you give me a list of indian restaurants as near as possible to Levenshulme;
   could you tell me where I can find a restaurant in central Manchester which serves indian or thai food;
   could you please give me a selection of telephone numbers for Mexican and Indian Restaurants located in the centre of Manchester?
   b. Declarative *I would like the number of a local insurance agent;*
   i would like to find out about car hire services based around Bolton

[1]This is shown in a study by Richards and Underwood [20], quoted by Fraser and Gilbert [11]

2. An indirect (non-conventionalized) request
   a. Interrogative: *are there any restaurants near my home;*
   are there Indian restaurants in central Manchester?;
   Do you know any Indian restaurants in Manchester?;
   How many restaurants are there in Manchester that serve spicy food?
   b. Declarative: *I am looking for a restaurant;*
   I am looking for the names of Mexican restaurants in Chorlton cum Hardy;
   I need to hire a car;
   I live in Entwistle outside Bolton and want to hire a car;
   I want some insurance details.First of all, can I have some assessments for life insurance

3. a direct request
   a. Imperative: *please give me information on the indian restaurants in central Manchester;*
   please supply details of small restaurants in the didsbury area;
   Please tell me the phone numbers of car hire firms around Bolton
   b. Noun Phrase: *indian restaurants in the centre of Manchester;*
   car hire firms please

4. A direct question
   What categories of restaurants do you list?

The three first types seem to be equally common, but only one direct question was found (and this dealt with meta-knowledge: the user wanted to know the classification and coverage of the database).

Notice that we have divided direct requests into two groups: imperative sentences and noun phrases *indian restaurants in the centre of Manchester*. NPs can also be interpreted as elliptical declarative sentences, elliptical conventionalized requests etc., but the exact syntactic form of the sentence that the NP belongs to is not important here: all we need to know is that the utterance is a request, and this can be inferred on the basis of the context it appears in. We must emphasize, however, that not all NP-utterances function as requests: most often they appear as answers to questions, and thus contextual information is of crucial importance. Direct requests are also commonly accompanied by the word *please*, which seems to indicate request rather than politeness in these contexts.

How to end

The closing of the dialogues is more problematic. The subjects seemed to use the system not as an equal dialogue partner, but as a machine, and they did not consider it relevant to go through the long closing ceremonies, or indicate that they have got

all the information they wanted. Of the dialogues, 24 % contained thanking at the end of the dialogue (about two thirds (63 %) of the subjects used *thank you*, but not necessarily at the end of the dialogue). When the task was accomplished, the conversation was considered finished, too.[2]

In human-human dialogues, however, the subjects' behaviour was different. Even though the setup was the same, the subjects tended to be more polite and cooperative: 70 % of the dialogues ended with thanking, and of these, 3/4 also contained acceptation or indication of future plans in a form of the phrase: *Thank you, that is all I want to know.* and *Thanks, I'll give Hertz a ring.* All, but one of the subjects thanked at the end.

From the user's point of view, the conversation is over when the required information is obtained. For the system, the accomplishment of a task means that its expectations of the continuation of the dialogue return to the very beginning: it is waiting for the user to take the initiative and either continue with a new request or quit.[3] From the system's point of view, the lack of explicit closing markers is thus problematic, as a long pause after a completed task may mean either that the user is thinking what to ask next or has gone home.

There are three solutions in this problematic situation. First, after having listed the required information, the system may take the initiative and, in the same turn, ask the user to make a decision about the continuation: *Is this all you wanted to know, Would you like to continue* etc. The negative side of this alternative is that it automatically makes the dialogue awkward and non-human-like, especially if the question is always phrased in the same way. The second alternative is to use mechanical time calculation: after a given waiting time, the system may decide that the conversation is over and close the connection. The third alternative is a combination of these two: after a certain waiting time, the system may take a more subtle initiative and ask if the user wants to continue the dialogue, and then, if no reply comes to this question, close the connection.

4.3.3 Pragmatic Aspects

The corpus collection tasks shows clearly that the collected dialogues are simpler and more straightforward as regards their conversational implicatures and dialogue structure than was expected: usually people made direct questions, which could be quite directly answered. Some even stated all the relevant search parameters in the very first utterance:

[2]This was also the most common case when the supervisor's help was needed: "I've got the information that I need, what shall I do next?".

[3]Of course, the beginning of dialogue in the beginning of the whole conversation or in the middle of a conversation differ on one crucial point: in the latter case, the system already has a model of a discourse built on the basis of the previous dialogues, while in the first case, no such model exists.

(1) U : Could you please give me a selection of telephone numbers for Mexican and Indian Restaurants located in the centre of Manchester.

However, some examples of such turn takings that can be regarded as schoolbook data for the speech act theory or Gricean maxims can be found. Some uncontroversial examples are the following (the two last ones are not from YP scenarios):

(2) U : Can I have a list of Curry houses in the Rusholme area that serve vegetarian food?
S : I'm sorry, the list of restaurants I have does not include such details; I can provide a list of names and phone numbers if that would help.
U : Yes, please.
S : Please wait...

(3) S : In which area?
U : well I live in Entwistle but need to travel to Rochdale and the airport

(4) U : Well, I don't know if I've got the right qualifications to do that

S : 3 GCE 'A' Levels with grades BBC. You must at least have a Maths 'A' level.

(5) U : Is it more vocational course than the last one?

S : You can take a year out in industry in this course.

It should also be noticed that many users started their dialogues with a general statement which in this context is to be understood as an indirect request for information, not as a description of the person's activities. However, indirection in these cases can be understood on the basis of the dialogue context: the very first user turn is a request to the system to supply certain kind of information. Examples:

(6) U : I am looking for a restaurant

(7) U : I need a hire car for tomorrow afternoon

(8) U : i need some addresses of insurance companies

One can find at least four reasons for the straightforwardness of the dialogues. Firstly, as Bill Black has pointed out, the information seeking dialogue is quite different from that of a problem solving one. We did not have subjects struggling with decisions and comparisons, since the scenarios were so simple that the tasks to be performed by the subjects were almost trivial: the subjects did not see any problem in hiring a car and leaving it in another place. On the other hand, if the person was familiar with car hiring, then professional questions were asked.

It must be noticed that even though producing "good" dialogue depends to a large extent on the wizard's capability to invite the subject to produce pragmatically interesting pieces of conversation, the wizard cannot push forward a natural dialogue, if the subject does not find the hidden inferences:

(9) U : car hire companies in bolton area
S : What would be best for you? Perhaps you could explain the type of journey you will be making.
U : hire for the day to make a short trip and leave the car at airport
S : Some companies allow you to collect the car from one branch and return it to another. However, they don't usually give unlimited mileage with this option.
U : yes unlimited
S : Did you consider that a company that is local to Bolton might not be able to collect if from the airport at reasonable cost?
U : can the car be delivered to my house
S : Yes do you want a chauffeur-driven service?
U : no thank you
S : Do you want the Bolton branch number or the airport numbers?
U : bolton
S : Please wait...

Conversational richness of the information seeking dialogues was thus more dependent on the persons' creative imagination than their linguistic or problem solving capabil-

ities. Some subjects (mostly students) were ready to accommodate themselves to the scenario and invent different possibilities to try out the limits of the computer, others found talking to a 'blank screen' somewhat odd.

Secondly, the Yellow Pages as it now shapes itself, does not contain enough information to stimulate vivid conversations. Hence, dialogue situations were somehow misleading from the naturalness view point. When subjects were told to use natural language in a way they would use it in an ordinary human communication, they generalized this to mean that the computer also had human database capabilities. In other words, they forgot they were looking for information from the YP, and made questions that would be natural when discussing such things face-to-face. For instance, in restaurant situations, questions like "what is the cheapest restaurant", "which one is in the good food guide", "which has car parking spaces", "is it a walking distance from the bus station" were frequent questions. Unfortunately, the YP does not contain such information, and thus, after several unsuccessful questions, the question-answer-conversation just died out.

Thirdly, there are psychological issues that had an impact on the experiment. The study situation in general was unknown and something special to the subjects. Some were nervous, because this was their first contact with a computer, some considered the situation a kind of test on their capabilities and were thus tense.[4]

Fourthly, also experience in computing and typing affected the behavior of the subjects. Many subjects had minimal acquaintance with typing, and this made their behaviour slow and unsure. However, usually they were able to use the keyboard and produce written language, but we should bear in mind that some potential YP Service users *are* unexperienced in computers and minimally acquainted with typing, and thus the information seeking situation as such should not be complicated with extra problems of how to "talk" to the computer with too sophisticated operations.[5]

One may be tempted to say that in plain information seeking situations like these, there is no need to exploit the whole repertoire of human pragmatic inference capabilities: the search task itself is well-defined and constrained, and it will terminate, when the information can be given or when no requested information is available. However, the straightforwardness of the collected dialogues does not necessarily give the right picture of the whole domain of the information seeking dialogues and their simplicity may be spurious: if the requested information cannot be given, or the reason for the search

[4]An example of an unintended problem solving situation is an oldish man who explained his strategy to have been "just to think what would be the best way to put the questions so that it would understand me; would it understand 'restaurants indian' or 'indian restaurants'." This implies that there is work to be done also in teaching a new generation which is *not* used to a keyword computer language, but would consider natural language as a normal way to communicate with computers, too.

[5]One of our subjects were quite unfamiliar with both typing and production of written language, and each request was a big effort by him. The requests were simple, sometimes misspelled, one or two word utterances like *carhire; car s; hertz; whole list.* Even though this subject represents an extreme case, the system should, of course, be prepared to handle telegraphic utterances as well.

turns out to be more complex, a richer problem solving situation may arise. The former case can be illustrated by the following example:

(10) U : could you tell me where i can find a restaurant in central manchester which serves indian or thai food
S : There are two Thai restaurants:
Siam Orchid 236-1388
Royal Orchid 236-5183
There are many Indian restaurants in Manchester. Can you be more specific?
U : king street
S : Please wait ...
The YP has no details of Indian restaurants on King Street
U : any indian restaurants in withington
S : Sarah Indian Cuisine
489 Wilmslow Road
Withington
446 2565

An instructive example of the latter (given by one of our dialogues) is the situation, where the tacit presupposition of the car hire scenario is actually broken: the subject cannot drive. Another problem behind the original information seeking task thus becomes visible, and the user naturally proceeds to attempt to solve this one: how to get from one place to another. This situation is analogous to the I-wish-to-redo-my-bathroom-and-what-shall-I-do example described in the PLUS Technical Annex.

It should also be mentioned that despite the scarce occurences of uncontroversially pragmatic turn-takings and conversational implicatures, the dialogues provide good data for exhibiting the use of the huge amount of contextual and textual information that is needed in natural language understanding (see examples in the next section).

4.4 Some Observations for Further Studies

Already a preliminary analysis of these dialogues calls for a high quality DM. The Cognitive Analyzer should be able to cope with various kinds of corrupt input and infer the 'real' meaning of an utterance, the Goal Formulator should be capable of deciding what is the best and most natural way to continue, and Response Planner be capable of producing flexible utterances as responses to the user queries. The dialogues also place the NLE in a new light: since the utterances are often far from the grammatical English sentences, the rules of grammar cannot be rigid S → NP VP-type, but a flexible

lexicalist view is required. What would be the correct NLE output for *Yes, please*; *yes unlimited*, *carhire*[6], *what about Rusholme*, or *only Bolton?*[7]

In order to produce robust dialogues, the system should be capable of handling two types of problematic input. First, formal, or syntactic-semantic-phonological issues like coordination, ellipsis, and anaphora, as well as punctuation oddities and different types of misspellings. Second, content, or pragmatic-textual-world-knowledge issues like metonymy, synonymy, metatextual anaphora, pragmatic particles, etc. This division does not correspond to the division of labour between DM and NLE.

Below we list some issues that are to be paid attention to:

- Punctuation: commas and full stops are readily left out, which makes the segmentation of the input complicated: *no thats all i need thank you ....* One should also notice a very special use of a comma referring to ' in contexts like *i,d definitely like...*
- Capitals: most users like to use small letters throughout the question, also for *I* and proper nouns.
- Different types of misspellings:[8]
  - two words written together: *ihave, carhire, thats*
  - mixed letters:[9] *purchance, somw, restraunt*
- Morphology: how to decide the correct singular form

  (11) U : do they serve for veggies
  S : what is a veggie? (veggy?)

- Peculiarities of English syntax. The following question was unintelligble to me at first, even on the basis of the previous context of the Wizard's reply:

  (12) S : A minority of consultants place large adverts in the YP; however, most entries are name/location only
  U : who place large adverts

[6]The user was surprised that the system could not understand his simple typing *carhire*, but still used the words *car hire firms* in one of its replies.

[7]Notice that in this very utterance, punctuation is the only cue for the intended speech act.

[8]Most misspellings were actually noticed by the user and corrected immediately.

[9]Should the DM somehow be able to use the phonological information of a misspelled word in its attempts to understand the utterance? In both types of misspellings, it is still the phonological/graphic form of the word that guides the humans towards the correct interpretation, and in some cases, misspellings are passed unnoticed. In the DM, this kind of exploitation of phonological/graphic cues may facilitate the inference making: a misspelled word need not be a totally unknown element in the input, but a familiar word, even though a little bit distorted. Of course, the system must still take care of the compatibility of the "familiar" word in its context, but the system is now only required to do simple checking, not abduction about what would be the compatible filling in this context.

- Pragmatic particles: *so, perhaps, roughly, then*
- Tags: *it is a French taxi, isn't it*
- Ellipses:
  - in answers to questions, i.e. the answer fills a 'slot', and we can argue that there is no ellipsis in these cases:

    (13) S : In which area is entwistle?
    U : bolton
  - occurring in questions, and thus, to understand the question, the context in which the question functions as a 'slot-filler' must be recreated:

    (14) U : which is the cheaper service
    S : *Of the 2 chauffeur firms?*
  - "long-distance ellipsis":

    (15) U : Will my employer pay expenses for the hotel?
    S : I'm sorry, you'll have to ask your employer that.
    U : *Assuming that he will*, can the Conference office arrange accommodation at the Conf. centre.
- Anaphora and reference resolution (choose your favourite example from the dialogues)
- Metatextual reference:

  (16) U : Please supply a list of telephone numbers for self-drive car hire companies in or around the area known as Entwistle (near Bolton).
  S : Please be more specific in your requirements; for instance, the YP can supply lists of those firms which offer unlimited mileage, delivery/collection, one way rentals, cheap local rates.
  U : Please supply telephone numbers as in question number 7, but only those which operate a delivery/collection service.
- Metonymy: *hot spicy food* for *restaurant*
- Paraphrase and synonymy: *reasonable rates* is interpreted as *cheaper*, *low cost* etc.
- Antonymy: *poor* vs. *better quality printers*

- User asks for definition or queries about the database directly: *what do you mean by 'proceesings'* (sic!), *What more information do you have about these companies? , is that all* (i.e. is the information given exhaustive w r t the database), *What can it tell me about them*
- Different levels of background knowledge: *four wheels* in the context of daisy wheel printers or in general

Chapter 5

Preliminary Analysis of the Swedish WOZ Corpus

5.1 Introduction

Let us depict the PLUS target system, in relation to its intended user and background database, in the following way:

user ←→ system ←→ YP data base

As we had no clear idea of what the target system will look like, and as we needed to simulate *some* system (otherwise it wouldn't be Wizard of Oz), we had to 'design' a system of our own. Now, let us depict the setting of the simulations in the following way:

subject ←→ wizard ←→ YP database subset

The subjects, numbering ten, were all university students or staff, aged 19–63. Eight of them are native Swedish speakers, but there are also one Spanish and one Hungarian native speaker. Typing ability is ranging from bad to medium and aquaintance with computers from none to medium.

The YP database subsets choosen were car hire, restaurants and personal insurance.

Thirty dialogues were collected using the Wizard of Oz method. This is a report on the preliminary analysis of these dialogues. The investigation of the material has been systematic, although not formal. We have tried to look at it from a 'pre-theoretical'

point of view. In general, we have started out with *all* examples of what seemed to be of one phenomenon, and then deleted duplicates, leaving what we hope is illustrative examples of the phenomenon in question. We have reflected on them, classified, and in other ways systematized them.

5.2 The Simulated System

In the following we will try to describe the simulated system from an external point of view. This system should be compared with the PLUS target system, as described in the PLUS external requirements. We certainly do not want to claim that we have been able to simulate the target PLUS system. This will be a later task.

It should also be noted that this description is an idealization, which the wizard has not always been able to act in accordance with.

5.2.1 The System's Knowledge

In terms of lexical and grammatical knowledge, the system is just as well equipped as any reasonably well educated Swedish speaking person. In terms of general common sense knowledge the system knows, for instance, that hot and spicy food is served in oriental and mexican restaurants.

The system's geographic knowledge is only mediocre. It knows the difference between the city as a whole, the centre, the eastern part, the western part etc. as well as the area where the airport is located. It can often (but not always) associate the name of a street with an area (e.g. it knows that Avenyen is located in the centre.) However, it doesn't possess the concept of 'distance' or the concept of 'nearest' (in terms of distance), i.e. for three locations A, B and C, it doesn't know if A is nearer to B than to C, or vice versa.

In terms of application knowledge, the system knows basically what is explicitly listed in the Swedish Yellow Pages, which is to say that it knows about addresses and telephone numbers. It also knows which company is located in which area (cf. above). However, it doesn't know about opening hours, prices or price levels, public transportations etc.

The system does, to a certain extent, possess meta-knowledge, i.e. it knows what it knows and what it doesn't know.

5.2.2 The System's Language Understanding Capabilities

The user's input to the system is textual only (no pointing, clicking, dragging). The system is very clever when it comes to understanding language. It can handle spelling errors (sometimes it does not even detect them), severe syntactic anomalies (relative to some 'standard school grammar'), abbreviations, etc. The system does easily resolve anaphoric references. However, the system doesn't understand linguistic meta-talk like in the following:

(17) U : is there any type of insurance that covers everything?
S : No.
U : do you know what "everything" refers to?
S : I don't understand. Please rephrase!

In short, the system understands written language nearly as well as a human.

5.2.3 The System's Language Generation Capabilities

For output the system delivers text only; no menus, nor maps. The system always expresses itself in complete and grammatically correct sentences (often more than one per turn). The sentences produced by the system contain no spelling errors. Ellipsis is not commonly used. The sentences contain almost no pronoun anaphora. Rather than using a pronoun to refer to e.g. an earlier introduced car rental service, the system reuses the name of it:

(18) U : I would like to have information on Budget Rent a Car, please
S : Budget Rent a Car is at Odinsg. 8 and has telephone number 800970.

But here's one common type of anaphoric reference performed by the system:

(19) U : Is Taj Mahal known as a good restaurant?
S : Information about that is not available.

Below we list some standard system responses (in the restaurants domain): Note that '*' is used as a list type constructor.

- Questions:

(20) S: Which area do you prefer?
S: Which type of place do you prefer?
S: Which type of food do you prefer?

- Enumeration of classes:

(21) S: Choose between the following types of places: <type-of-restaurant>*
S: Choose between the following types of food: <type-of-food>*
S: Choose between the following kitchens: <nationality>*
S: You may choose between the following areas: <part-of-city>*

- Counting and enumeration of instances of classes:

(22) S: In <part-of-city> there are <number> <type-of-restaurant>.
S: The following places in <part-of-city> are <type-of-restaurant>: <instance>*
S: In <part-of-city> there are <number> restaurants that serves <type-of-food>.
S: The following places in <part-of-city> serves <type-of-food> : <instance>*
S: In <part-of-city> there are <number> restaurants that serves <nationality> food.
S: The following restaurants in <part-of-city> serves <nationality> mat: <instance>*

- Values of properties associated with instances:

(23) S: <restaurant> has telephone number <phone-number>.
S: <restaurant> is at <address>.
S: <restaurant> is at <address> and has telephone number <phone-number>.

Similarly for the car hire and insurance domains.

5.2.4 The Generation of Helpful Responses

If the user starts with a question that will result in too many hits when searching the database (i.e. too many to present to the user), the system will respond by asking questions in order to obtain information that will further constrain the search.

If the user doesn't answer a question in a way that the system can handle, the viable alternatives are presented in a menu-like format. The rationale behind this is to present the user with the system's limitations in terms of vocabulary and world knowledge, and at the same time provide a way of solving the problem.

If the user is asking for information that the system cannot provide (because it was not designed that way), the system informs the user about this. At the same time it also informs the user about what it positively *can* provide. This is the way information on limitations in terms of application knowledge is given.

The system may also inform the user that he may contact the companies in question and in this way obtain more information.

Unfortunately, the wizard has not always been able to correctly implement these strategies. On the other hand, other interesting pragmatic features of the system are probably unconsciously implemented by the wizard.

5.2.5 General Performance

The speed of the system is not great. It takes up to 100 seconds (actually, the extreme worst was 166 seconds) to generate an answer, from the moment when the system has been given the 'floor'. In order not to frustrate the user, the system generates 'Please wait..'-responses before doing anything else:

(24) S: Please wait ...
The following small insurance companies offer home insurances:
- Holmia
- Försäkringsbolaget Royal

S: Just a moment ...
No, about that I have no information.

To save space, and to increase readability, we have choosen to delete these responses from the examples given in this document.

5.3 The Structure of the Dialogues

In this section we will simply present fragments taken from the collected material.[1] But we will also connect the fragments into bigger chunks (phases) and in this way begin to impose some structure upon the dialogues. We will distinguish between the opening phase, the user's first request, a phase of specification of this request, a phase of presentation of the result of the inquiry, a phase which is characterised by the user showing ignorance of the limitations of the system, and finally, a phase constituting the closing of the dialogue. The giving and taking of initiative will also be indicated.

5.3.1 The Opening of the Dialogues

The user is first met by a screen with a "Yellow Pages Information System" -logo and a "hit space to begin"-message. When the user hits the space key, the system clears the screen and produces the following:

(25) S: Welcome to the Yellow Pages Information Service!
What can we do for you?
GSI>

Now, the first sentence obviously is a greeting. The second sentence looks like a question, and the prompter is how 'the floor' is given to the user. The user is thus given the initiative.

5.3.2 The User's First Utterance

The first utterance from the user has a somewhat special status. This is one of the places where the user shows politeness (if he shows it), this is where he takes (or refrains from taking) the initiative, and most importantly, this is where the user informs the system of what he needs and wants. The user's main reason for contacting the system is likely to be revealed here. It is also important to note that this utterance has no significant prior context, no topical constraints arising from prior turns.

We have found only two utterances (same subject) longer than one sentence:

[1] In this preliminary version of this document we will, for your convenience, present translations of data rather than the actual data. It is of course impossible to give completely faithful translations. We think, however, that the gist of the examples will go through.

(26) U : Hello. I would like to have some information about where I can go to have a really good dinner. I will ask a friend out, and he is very fond of hot, spicy food. Where shall we go?.

(27) U : Hello. I would like to contact different insurance companies. Do you think you can help me?

More common is that the user utters either a full sentence or some other phrase. Among the full sentences we find declarative ones:

(28) U : I would like to find a car rental service located in Partille

(29) U : I want to know where I can find a restaurant with hot and spicy food

(30) U : I want facts about insurance companies

(31) U : I would like to know what oriental restaurant there are

(32) U : I want to know what car rental companies are available in Göteborg.

(33) U : Hello, I want to know which restaurants in Göteborg city serve hot and spicy food

(34) U : I am interested in insurance companies and their services

(35) U : I want to hire a car!

It is interesting to note that all the declarative sentences says something quite direct about the 'mental state' of the user (what he wants, what he wants to know, his interests). There are also sentences in the interrogative and among them we find both yes/no-questions (surface grammatically, that is),

(36) U : are there any Thai restaurants in the neigbourhood of Kungsportsplatsen?

(37) U : can you give me the phone number to one insurance company specialising in family insurances and one that has good car insurances or a combination of both

and WH-questions:

(38) U : which car rental companies are there in the centre

(39) U : which insurance companies are there in göteborg

(40) U : which companies have cars for rent ?

(41) U : which insurance companies have insurances for both home and car?

In the WOZ-corpus none of the initial sentences are in the imperative, although imperative sentences can be found in the human-human dialogue corpus:

(42) U : give me car rental services in central Göteborg

(43) U : give me the restaurants located in the Landala area

Now, if we look at phrases other than sentences, the first thing to note is that all but one are noun phrases:

The following can be regarded as 'elliptic' answers to the system's initial question "What can we do for you?"

(44) U : information about car rental in göteborg

(45) U : information about restaurants

(46) U : addresses to non-Swedish restaurants, for instance Lebanese or Chinese, located in the centre

In other cases "information on" can be thought of as implicitly understood:

(47) U : possibilities for renting cars in Göteborg

(48) U : car hire

(49) U : Restaurants

(50) U : Insurance consulting

(51) U : Restaurants with hot and spicy food, located in the centre!

(52) U : insurance companies

(53) U : insurance companies concerning financial security

The following cases are problematic:

(54) U : insurance companies?

(55) U : insur.companies?

The question marks indicate that they are questions, rather than answers to the system's question.

We also have one case of a verb phrase:[2]

(56) U : Rent a car!

Most users seem to treat the second sentence in the system's first utterance as a question, and try to answer it as such. But some users apparently do not; they may very well use an interrogative construction and ask a question instead of 'formally' answering.

[2]The Swedish original is not ambiguous between an infinitive verb phrase and a sentence in the imperative.

5.3.3 The Specification Phase

The system may now in principle respond to the user's request by giving a list of addresses and telephone numbers, as in the following made up dialogue fragment:

(57) U : I want addresses and telephone numbers to car hire companies, located in the eastern part of the city, which can hire me an ordinary small car.
S : The following companies in the eastern parts of the city rent private cars:
- Hertz, Stampg 16 A, tel. 803730
- Statoil Olskrokens Bensin o. Biluthyrning AB, L. Olskroksg. 2, tel. 258515

However, we did in fact *not* record any dialogue with a structure like that. The user never specified his request in sufficient detail for the system to be able to immediately supply this information. There is always what we may call a 'specification phase', where the system asks questions, and the user answers, in order to determine the values of parameters like preferred location, size of company or preferred type of food/car/insurance (depending on the domain).

In this phase the system initially has the initiative.

The answers to the system's questions are usually short, often one word only:

(58) S : Which type of car are you interested in?
U : private car

or short phrases:

(59) S : Which type of place do you prefer?
U : Restaurant with hot and spicy food in the neigbourhood of Avenyn (Göteborg)

(60) S : Which type of insurance are you interested in?
U : Everything! Car, home, life etc

(61) S : Which type of car are you interested in?
U : a car with normal comfort

The specifications often comes in sequences:

(62) S : Which type of car are you interested in?
U : private car
S : Which area do you prefer?
U : Kålltorp

Sometimes the system gives a reason for asking a question:

(63) S : In Göteborg there are 18 companies which hire private cars. Which area do you prefer?
U : the centre

In this case, the first sentence should be understood as the way the system tells the user why the companies are not listed; the list simply wouldn't fit on the screen.

The user does not have to answer the question, and sometimes (well, once) he didn't:

(64) S : In Göteborg there are 8 insurance companies offering home-, life- and car insurances. Do you prefer a big or a small company?
U : I want quotations from, for instance, Folksam och Trygghansa

It is interesting to note that the system got as many as four different answers (from four different users) to the type of disjunctive question known as 'binary choice question' :

(65) S : Do you prefer a big or a small insurance company?
U : a small

(66) U : a big

(67) U : it doesn't matter. the one that is most advantageous for students

(68) U : a middle sized

The system's question does not always result in a satisfactory answer. Sometimes the system doesn't understand the user's answer, or find it too general, and will therefore ask for a clarification.

(69) S : Which area do you prefer?
U : hisings backa
S : You may choose between the following areas:
- the centre
- the western parts of the city
- the eastern parts of the city
- Landvetter
U : the centre

(70) S : Which type of car are you interested in?
U : car
S : Choose between the following types of cars:
- private cars
- trucks
- buses
- trailers
- vans
- mobile homes
U : private cars

5.3.4 The Presentation Phase

The system presents the user with a list of addresses and telephone numbers.

(71) S : The following companies in the eastern parts of the city rent private cars:
- Hertz, Stampg 16 A, tel. 803730
- Statoil Olskrokens Bensin o. Biluthyrning AB, L. Olskroksg. 2, tel. 258515

The initiative has now been given to the user. Some users are closing the dialogue here (see 'The Closing of the Dialogue'). Some users start a new task. Others start 'squeezing' the system for more information on the elements in the list (see 'The Phase of Ignorance'). Some users, when they see the result of the interaction, change their minds about the parameter values given:

(72) S : Do you prefer a big or a small insurance company?
U : a small
S : The following small insurance companies offer home insurances:
- Holmia
- Försäkringsbolaget Royal
U : I rather choose a big, then
S : The following big insurance companies offer home insurances:
- Ansvar
- Folksam
- WASA FÖRSÄKRING
- Länsförsäkringar
- Skandia Försäkrings AB
- Trygg-Hansa

5.3.5 The Phase of Ignorance

Quite often the dialogue points to a clear discrepancy between what the user thinks the system knows and what it actually knows.

(73) U : Is Taj Mahal known as a good restaurant?
S : We have no information about that.
U : Is it expensive?
S : Information is lacking.
U : Are there any public transportations in the neighbourhood?
S : Information is lacking.
U : Taxi?
S : The only available information concerning the restaurants is addresses and telephone numbers

This is not very helpful. Fortunately, the system is often more helpful than this, in that it reveals the scope of its knowledge much earlier:

(74) U : I want price information from, for instance, Folksam and Trygghansa
S : Information about prices is lacking. You can have addresses and telephone numbers to the insurance companies.

(75) U : I would like to have facts about the opening hours of the Mexican restaurants
S : We don't have information about opening hours. We have information about addresses and telephone numbers, only.

(76) U : Which kitchen cooks the best food?
S : We can not inform you about this. We suggest that you consult a restaurant guide.

In the examples listed above, the user 'presupposes' the availability of a certain kind of information. In other cases the user explicitly asks about this availability:

(77) U : is there any information on how to find, for instance, Postgatan
S : No, there is no such information.

(78) U : would you be able to recommend one of these restaurants that serves spicy food?
S : No, I cannot recommend specific restaurants. But Mexican restaurants are well known for their spicy food.

(79) S : [...] there is a Malaysian restaurant in the centre. Restaurant Malaysia Sate is at Ö. Hamng. 19 and has telephone number 115710.
U : do you have any quotations?
S : Information about prices or opening hours is not available, only addresses and telephone numbers.

It even happens that the system responds that it doesn't have information of a certain kind, but the user insists that perhaps it can get it from somewhere...

(80) U : do you have any further information about cars for hire
S : No, only addresses och telephone numbers to car rental services.
U : wouldn't you be able, in some way, to find information about cars for hire concerning, not only addresses and telephone numbers, but something else
S : No, I'm sorry. You should contact the car rental services yourself, e.g. by calling them.

The system may also inform the user that he may contact the companies in question and in this way obtain more information:

(81) U : Can I fetch the car at the Central station and leave it at Landvetter airport?
S : I have no information about that. I suggest that you turn to Avis for information.

5.3.6 The Closing of the Dialogues

Many users apparently are satisfied when the system has listed addresses and phone numbers to a set of companies, so they just say 'thank you' and expects the system to end the session.

(82) S : Ansvar is at N. Alleg. 6. Folksam is at Storg. 53. WASA is at Stampgatan 14. Länsförsäkringar is at Ö. Larmgatan 16. Skandia is at Sten Stureg. 12. Trygg-Hansa is at Vasag. 45.
U : thank you
S : Thank you and goodbye from the Yellow Page Information Service.

(83) S : The following companies at Landvetter airport rent private cars:
- Hertz, tel. 946020
- Avis, tel. 946030
- InterRent, tel. 946210

U : thank you for your help and goodbye!
S : Thank you and welcome back!

Some users are even telling the system about their choice.

(84) U : I have decided to go for Inter Rent Biluthyrning AB
S : I understand. Thank you and welcome back!

(85) S : The following restaurants in the centre serves Mexican food:
- O'Learys Bar and Restaurant, Ö. Hamng. 37, tel. 115519
- CARAMBA TEX MEX GRILL and BAR, Ö. Hamng. 37, tel. 116796
- Taste Five, Postg. 26, tel. 802440

Information about prices or opening hours is not available, only information about addresses and telephone numbers.
U : caramba tex sounds good. thanks for your help
S : Thank you and goodbye from the Yellow Page Information Service.

(86) U : Then I will contact Folksam. Thanks for your help.

(87) S : Thank you and welcome back!

Some users, when the system has listed addresses and phone numbers to a set of companies, apparently feel that there must be more information available (see above), but when they realize that this is not the case, they will end the session.

(88) S : No, there is information about telephone numbers and addresses only.
U : Thank you anyway
S : Thank you and goodbye from the Yellow Page Information Service.

Some users write 'thank you' (or something equivalent), some writes 'bye' (or something equivalent), and others writes both 'thank you' and 'bye' (but then always in the same sentence).

(89) U : thank you

(90) U : unbelieveble! you are clever. Bye!

(91) U : thank you for your help and bye!

One user explicitly indicated the end of the dialogue, using a kind of 'computerese' expression:[3]

(92) U : thank you for the information, quit informing.

The last turn is on the system (the user is not given a prompter):

(93) S : Thank you and goodbye from the Yellow Page Information Service.
Welcome back!

(94) S : Thank you and welcome back!

[3]The Swedish expression used here was *tack för informationen, informationen avslutas*

5.3.7 The Structure of the Dialogues - A Summary

Let us summarize some observations of the structure of the dialogues by giving the following 'dialogue grammar'. This grammar is not observationally adequate (there are dialogues in the corpus that it will not generate), and even less descriptively adequate[4] (the correct structure is not always assigned to the dialogues), but it gives a rough picture of the structure of most of the dialogues. A dialogue can be described as an opening followed by a sequence of one or more tasks, followed by the closing of the dialogue.

dialogue → opening, tasks, closing.

The opening has the following structure:

opening → sys-greeting, sys-'question', opt-user-greeting.

The sequence of tasks can be described in the following way,

tasks → task.

tasks → task, tasks.

A task may be described as the user's initial specification of a 'desire' for information followed by an optional specification phase, and ended by the system giving some information to the user.

task → user-main-request, opt-specifications, sys-main-answer.

opt-specifications → [].

opt-specifications → specifications.

The fact that specifications often comes in sequences may be described as follows:

specification → [].

specifications → specification, specifications.

[4]In Chomsky's terminology

A specification phase begins with the system asking the user a question. If the user gives an appropriate answer, this phase is over, but it may be that there is need for a phase of clarification in between.

specification → sys-asks-question, opt-clarification, user-answer.

opt-clarification → [].

opt-clarification → clarification.

The fact that the system's question may actually be preceded by the system giving a reson for asking is caught in the following rule:

sys-asks-question → opt-sys-gives-reason-for-asking, sys-asks-question2.

A clarification subdialogue can be described as the user providing information which the system doesn't understand or takes as irrelevant in the current context. The system reacts by listing a set of relevant answers from which the user may pick one.

clarification → user-gives-unknown-answer, sys-lists-viable-answers.

The closing phase of the dialoges consists of the user's and the system's closing utterances, and may thus be described as follows:

closing → user-closing, sys-closing.

We feel that some phenomena in the dialogues should not be accounted for in the grammar, but are best described as deviant in relation to it.

This 'grammatical exercise' may not be of value when writing a 'real' dialogue grammar for the PLUS target system (if we decide to do that). But at the very least, it could be of value when discussing possible improvements of the simulations.

5.4 Some Notes on Language Use

In this section we will suggest some generalisations having to do with the way the dialogue partners use language to communicate. We will touch briefly upon spelling and related issues, we will consider the use of referential expressions and 'ellipsis' and we will look at some cases of ambiguity found in the dialogues.

5.4.1 Spelling and Related Issues

As was also noted by the UMIST team:

- commas and full stops are often missing
- users very often do not capitalize in the proper way, e.g. they use small letters throughout
- misspellings are common, but the user seems to be eager to correct himself when he detect an error.
- two words are often written together

As examples of these phenomena are fairly dependent on a specific language, in this case Swedish, we do not exemplify them.

5.4.2 The Use of Referential Expressions

Proper Names

In the collected dialogues, proper names are used mainly for two purposes: to refer to companies, and to refer to locations like streets, places or areas:

Folksam (insurance company)

Storgatan (street)

Centralstationen (railway station)

Landvetter (airport)

Proper names do more often than not contain more than one word:

InterRent Europcar Biluthyrning AB [5]

Restaurant Taj Mahal

But on the other hand, they are often used in abbreviated forms:

Storg.

InterRent

Taj Mahal

Anaphoric Expressions

Use of pronoun anaphora is very limited,[6] but definite noun phrase anaphora is quite common. Some other phenomena which might be classified as examples of anaphora are mentioned in 4.2.5.

- Intrasentential pronoun anaphora:

  (95) U : I choose to contact WASA FÖRSÄKRING, Stampgatan 14 and get myself information about *their* rules of insurances and offers

  (96) U : Can I fetch the car at the Central station and leave *it* at Landvetter?

- Intersentential pronoun anaphora:

  (97) S : The following insurance companies offers home- and car insurances:
  - Ansvar, N. Alleg. 6, tel. 177900
  - Folksam, Storg. 53, tel. 600500
  - WASA FÖRSÄKRING, Stampgatan 14, tel. 614500
  - Länsförsäkringar, Ö. Larmgatan 16, tel. 638000
  - Skandia Försäkrings AB, Sten Stureg. 12, tel. 816000
  - Trygg-Hansa, Vasag. 45, tel. 819000

  (98) U : which of *these* have student discounts?

[5] 'AB' is a Swedish correspondent to 'Ltd' or 'Inc'.
[6] As is also shown by Dahlbäck and Jönsson [10]

(99) U : Is Taj Mahal known as a good restaurant?
S : Information about *this* is lacking.
U : Is *it* expensive?
S : Information is lacking.

- Definite noun phrase anaphora:

(100) S : Which type of car are you interested in?
U : private car
...
U : I would also like to know *the cost per mile and hour.* And also if it's possible to leave *the car* at Landvetter airport .

(101) U : Can I fetch *the car* at the Central station and leave it at Landvetter?

(102) S : The following companies in the western parts of the city rent private cars: - Statoil Biluthyrning Järntorget, Järntorget, tel. 117896 - Masthuggsgaraget, Andra Långg. 48, tel. 170210
U : is there anything in Askim?
S : No, we have no information about that.
U : I want to know the price for renting a car for five hours from *both companies*

(103) S : The following companies in the eastern parts of the city rent private cars: - Hertz, Stampg 16 A, tel. 803730 - Statoil Olskrokens Bensin o. Biluthyrning AB, L. Olskroksg. 2, tel. 258515
U : If I hire a car from *one of the listed companies* will I then be able to leave the car at Landvetter airport?

Expressions Referring to Location

The category of location tend to be important in these exchanges, mainly because we are dealing with companies, and one aspect of a company is that it is physically located in space, i.e. in a specific area (part of city), and at a specific address. The user is also

located in space, and the distance relation between these two locations is important, as is the relation between distances (comparision of distances).

In the following, we have divided the expressions referring to location into corresponding categories.

- addresses (exact location)

  (104) S : Restaurant Taj Mahal *is at Järntorget 4* and has telephone number 121922. [...]

  (105) U : I want to know *where* I can find a restaurant serving hot and spicy food

- areas (part of town)

  (106) S : Which *part of the city* do you prefer?

  (107) S : You may choose between the following areas:
  - the centre
  - the western parts of the city
  - the eastern parts of the city
  - the northern parts of the city
  - the southern parts of the city

  (108) U : I would like to find a car rental service *in Partille*

- distance

  (109) S : Which type of place do you prefer?
  U : restaurant with hot and spicy food *in the neighbourhood of Avenyn (Göteborg)*

  (110) S : Restaurant Taj Mahal is at Järntorget 4 and has telephone number 121922. [...]
  U : Are there public transportations *in the neigbourhood*?

(111) U : which of these restaurants *are located in the neigbourhood of kungsportsplatsen*?

What does 'in the neigbourhood of' mean? Walking distance? One kilometer or two?

- relation between distances

(112) U : [referring to a set of car hire companies] which is *closest to södra gubberogatan*

- other

Some companies are 'distributed' over different locations, as for example is the case with the car rental companies with head office in town and a local branch at the airport.

(113) U : some company that might have premises both at Landvetter and in the centre of Göteborg
S : [...] You should turn to a company with a local branch at Landvetter airport.

Expressions Referring to Time

The dialogues are all in the present tense.

The system knows nothing about opening hours (and time is not introduced in any other way), and therefore expressions related to the concept of time are rare. But here is one:

(114) U : I would like to have facts about the opening hours for the Mexican restaurants
S : Information about opening hours is lacking. You can have information about addresses and phone numbers.

If the system could deal with opening hours, expressions related to time would become more important.

'Ellipsis' and Related Phenomena

Below we have collected examples where utterances are, or appear, 'incomplete' in relation to some standard conception of grammaticality, and we have also made an attemt to classify them. This tentative classification is of course *not* exclusive. It is furthermore far from clear whether the majority of the cases exemplify 'genuine ellipsis', at least not until a clear definition is given.

In connection with some of the examples we indicate *one* way of 'expanding' it, so as to make it less difficult to understand in isolation.

In some cases the user exploit the immediate context, e.g. the present topic, or a recent listing or something like that:

(115) U : is there more information?

Meaning: Is there more information concerning the current topic?

(116) S : The following companies in the centre rent private cars:
- Kungsgaragets Biluthyrning, Ekelundsg. 5-7, tel. 112877
- Avis Biluthyrning Licensee/Riksbilar AB, Centralstation, tel. 170410
- Göteborgs Biluthyrning AB, Teaterg. 22-24, tel. 184700
- Heden Biluthyrning AB, Engelbrektsg. 55, tel. 200035
- Vasagaraget, Teaterg. 22-24, tel. 200770
- Aveny Biluthyrning AB, Teaterg. 22-24, tel. 203790
- Budget Rent a Car, Odinsg. 8, tel. 800970
- Hyrbilen, Skåneg. 1, tel. 804780
- InterRent Europcar Biluthyrning AB, Odinsg. 1, tel. 805390
- Wagnefors Biluthyrning AB, Södra v. 58, tel. 810090
- Quick Rent a Car, Karl Gustavsg. 29, tel. 819735

U : the cheapest alternative?

Meaning: For the companies listed, which is the cheapest alternative?

(117) S : The following big insurance companies offer home insurances:
- Ansvar
- Folksam
- WASA FÖRSÄKRING
- Länsförsäkringar
- Skandia Försäkrings AB
- Trygg-Hansa

U : which offers car insurance, as well?

Meaning: Which of these offers car insurance as well?

(118) S : The following companies in the western parts of the city rent private cars:
- Statoil Biluthyrning Järntorget, Järntorget, tel. 117896
- Masthuggsgaraget, Andra Långg. 48, tel. 170210

U : information on prices ?

Meaning: For the companies listed, do you have information on prices for renting a car?

In other cases it is the *wide context* that is exploited (like the fact that the user terminal is located in Göteborg):

(119) U : I would like to know which oriental restaurants there are

Meaning: I would like to know which oriental restaurants there are in Göteborg

In some cases no, or only rudimentary, indications of intended 'illucutionary force' or 'mode' are given. (Functionally it is often irrelevant whether the utterance is construed as a question, a request for information, or a declaration of intent, goals etc.):

(120) S : The following companies in the centre rent private cars:
- Kungsgaragets Biluthyrning, Ekelundsg. 5-7, tel. 112877
- Avis Biluthyrning Licensee/Riksbilar AB, Centralstation, tel. 170410
- Göteborgs Biluthyrning AB, Teaterg. 22-24, tel. 184700
- Heden Biluthyrning AB, Engelbrektsg. 55, tel. 200035
- Vasagaraget, Teaterg. 22-24, tel. 200770
- Aveny Biluthyrning AB, Teaterg. 22-24, tel. 203790
- Budget Rent a Car, Odinsg. 8, tel. 800970
- Hyrbilen, Skåneg. 1, tel. 804780
- InterRent Europcar Biluthyrning AB, Odinsg. 1, tel. 805390
- Wagnefors Biluthyrning AB, Södra v. 58, tel. 810090
- Quick Rent a Car, Karl Gustavsg. 29, tel. 819735

U : the cheapest alternative?

Meaning: Which is the cheapest alternative of the ones listed?

(121) U : restaurants with spicy food, located in the centre!

Meaning: Give me information on restaurants with spicy food, located in the centre!

(122) S : Skandia has telephone number 816000. Folksam has telephone number 600500. Trygg-Hansa has telephone number 819000. WASA has telephone number 614500. Länsförsäkringar has telephone number 638000. Ansvar has telephone number 177900.

U : Addresses as well

Meaning: Give me the addresses to all the listed companies as well

Some cases could be construed *either* as elliptical *or* as exploiting some anaphoric use of e.g. quantifier expressions:

(123) S : All of the listed companies offer car insurance.

U : can I have the phone numbers to everyone

Meaning: Can I have the phone numbers to everyone of the companies listed?

(124) U : is there any type of insurance that covers everything?

Meaning: Is there any type of insurance that covers every type of insurance I have asked about?

(125) S : There are no insurance companies that offer home insurance only.
U : is there any with only car insurance.?

Meaning: Is there any insurance company that offers car insurance only

(126) S : Which type of car are you interested in?
U : one to a reasonable price

Meaning: The type of car which has a reasonable price

Sometimes the system presents something that may best be regarded as a 'menu' (e.g. a multiple choice question or a listing of alternatives). The user then often produces something that looks like an ellipsis but could just as well be classified as simply as a 'choice indication':

(127) S : There are 8 insurance companies that offer home- and personal insurances. Do you prefer a big or a small company?
U : a big

That the above example also can be analysed as elliptic can be seen by comparing it to the following example from another dialogue:

(128) S : Do you prefer a big or a small company?
U : I prefer a big company

Related to choice indications are 'keywords':

(129) S : Welcome to the Yellow Pages Information Service!
What can we do for you?
U : insurance companies

The difference between a keyword and an 'elliptic' one word phrase might be that the latter is dependent on the (linguistic) context for its interpretation, whereas the former is dependent more on a wide, general context, like the fact that the user is interacting with an information providing, Minitel-like, system.

Ambiguity

In this section we will list some cases of ambiguity.

- Modifier scope ambiguity

(130) U : Arabic or Indian in the neigbourhood of järntorget
S : The following restaurant in the western parts of the city serves Arabic food:
- Kaninen, Järntorget 6, tel. 242324
The following restaurants in the western parts of the city serve Indian food:
- Restaurant Taj Mahal, Järntorget 4, tel. 121922
- Restaurant Tandoor, Bang. 21, tel. 147881

In this case there probably is a preferred reading (Arabic in the neigbourhood of järntorget or Indian in the neigbourhood of järntorget). How, in general, to find the preferred reading in cases like this is a difficult issue. But it turns out that, in this particular example, the ambiguity doesn't matter! There is only one arabic restaurant in Göteborg, located at Järntorget, so obviously, the system didn't have to resolve the ambiguity.

- Referential ambiguity

(131) S : The following restaurants in Göteborg serves Indian food:
- Restaurant Taj Mahal, Järntorget 4, tel. 121922
- Restaurant Tandoor, Bang. 21, tel. 147881
- Restaurant Taj India, Odinsg. 6, tel. 150569
- Old Town Restaurant, Artillerig. 8, tel. 843172
U : Is there any other information about *the restaurants*

This use of the definite NP *the restaurants* is referentially ambiguous between the restaurants as a class (the set of all restaurants) and the listed instances of restaurants. There is no preferred reading, we would say, but this ambiguity probably doesn't matter anyway.

- Quantifier scope ambiguity

We have found no clear-cut cases of quantifier scope ambiguity.

Politeness

One would think that the system's way of greeting the user is likely to influence the politeness of the user's response.

But only in one case does the user greet the system:

(132) U : Hello, I want to know which restaurants in the centre of Göteborg serve hot and spicy food

The wordings are often quite polite:

(133) U : I would like to find a car rental service located in Partille

U : I would like to know which oriental restaurants there are

U : can you give me the phone number to one insurance company specialising in family insurances and one that has good car insurances or a combination of both

Only if he really want to close the session, does the user produce 'thank you' or something similar:

(134) U : Thank you anyway
U : Thank you
U : thank you for the information, quit information.
U : thank you for the information.
U : unbelievable! you are clever. Bye !
U : thank you for your help
U : caramba tex sounds good, thank you for your help
U : thank you and goodbye
U : thank you for your help and good bye!
U : Thank you so much and goodbye .

5.5 The user's view of the system

In this section we will attempt a characterization of how the system appears to the user. There are two sources for information about this:

- the dialogues
- the questionnaires

5.5.1 The User's View of the Speed of the System

Nearly all subjects complained that it took long time for the system to respond:
"Too much waiting for an answer.."
"too slow and sluggish"

5.5.2 The User's Expectations of the System's Knowledge

We have found that, clearly, the knowledge of the system does not match the expectations of the users. Below we give a fairly complete list of what, judging from the dialogues, the users expect the system to know :

- information on prices, price level, milage cost
- information on discounts (for students)
- information on how to find the way to a specific address
- information on opening hours

- information on accessability by means of public transportations (bus, tram, taxi)
- information on which place is 'the best'

That users want this kind of information can be backed up by the questionnaires:

"If I would want to use a computer intead of the YP book, I would demand more services form the computer, and the possibility to connect different kinds of services."

The users seem to experience their lack of knowledge of the system's range of knowledge in a very negative way:

"It would be better if one knew what kind of information the computer had to offer, instead of having to ask for it in vain. Some kind of menu suits me better."

5.5.3 The User's View of the System as a Dialogue Partner

The users appreciated the way the computer could understand what they were writing:

(135) U : unbelievable! you are clever, Bye!

"The computer was clever understanding what I wrote, even if I spelled something wrong or used abbrevations."

5.6 Suggestions for Improvements of the Simulator

We may not know exactly what the PLUS target system will look like, but we know for sure that it will not understand natural language as well as a human. So when it comes to understanding language, we do actually want to simulate a less clever system. This is not an easy task for the wizard, though, as it involves surpressing selected parts of his skill using language, without really having time to think about it.

On the other hand, when it comes to knowing things about restaurants, car hire services and insurance companies, we do want our simulated system to behave better than it

currently does. In terms of dialogue structure we want 'the phase of ignorance' to be reduced to a minimum. (However, there will always be things that a user expects the system to know, but which it doesn't). One way of achieving this is to simply add more information to the database subset. Consider the following categories of information, derived form the user's wish lists:

- information on opening hours and telephone hours
- information on how to find the way to a specific address
- information on accessability by means of public transportations (bus, tram, taxi)

Information about opening hours and telephone hours is relatively simple to add. And this is an important category, especially if we want to try out different methods of reasoning about time (like the event calculus). If we add this kind of information, and make some changes in the instructions to the subjects, perhaps we could expect exchanges like the following:

(136) U : I called them yesterday at this time, but they didn't answer. Are they open now?
S : No, but they will open again tomorrow at nine a clock.
U : Thank you, I will call them then.

Information on how (public) transportations can be used to get to the company in question may be much harder to get.

Now consider the following category of information:

- information on prices, price level, milage cost, information on discounts (students)
- information on which place is 'the best' (has the best food, best service etc)

An argument against adding knowledge of the first category to the knowledge base is that this kind of information changes frequently and it is impossible for the system to guarantee the correctness of such information at every time. An argument against adding information of the second category (evaluations) is that the information service cannot take responsibility for it. It's simply too dependent on subjective measures of values. But we think there is a way around this. If we take the information broker metaphor (i.e. the view of the system as something that acts as an intelligent intermediary between the information in the Yellow pages data base and the user) more seriously, we can let the broker 'quote' from the yellow pages database. In this way, the system does not *make* judgements, it merely *conveys* judgements. That's a totally different matter, as the system doesn't have to take responsibility for what it conveys.

Finally, we would like to propose two simple additions to the system and the scenario, which will make the dialogues more interesting. First, a small printer could be connected to the subject's machine and instructions to ask the system to produce a hardcopy of the result of the interaction could be given to him. Second, it could also be interesting if the subject could ask the system to connect him to a certain telephone number. These amendments are realistic, they are easy to add, they help to motivate the subjects, and they will generate more interesting dialogues.

Chapter 6

Communicative Activity Analysis of a Wizard of Oz Experiment

6.1 Introduction

6.1.1 The Experiment

In this chapter we will describe and analyze a simulation of a factual information seeking system for a subsection of the "yellow pages". The subsection contains information about (i) car hire (ii) restaurants, and (iii) personal insurance. The simulation was accomplished by use of the so-called Wizard of Oz technique.[1] Using this technique, subjects who believe they are communicating with a prototype of an automatic information system are, in fact, communicating with another person via a computer terminal. The person who is simulating the system (and is called the wizard with an allusion to the American Wizard of Oz story) has at his/her disposal an information base and a set of preprogrammed "canned" response patterns. The actual simulation was preceded by an introductory session where subjects were given a set of instructio

The output of the simulation sessions were, on the one hand, a set of logged written dialogues between "user" and "system" and, on the other hand, a completed set of questionnaires. In the present study, we will only report on an analysis of the logged written dialogues.

[1] Cf Dahlbäck and Jönsson [9]

6.1.2 Communicative Activity Analysis

The analysis has been guided by the "communicative activity analysis" described, for example, in Allwood [3] and Allwood, Nivre and Ahlsén [5]. The superordinate perspective of this analysis is that language and communication are seen as instruments for individual and collective information processing in the service of social activities, the main purpose of which does not have to be communicative.[2]

Further, it is suggested that in studying the instrumental aspect of language and communication for social activity, it is often useful to distinguish 4 analytical dimensions pertaining to the phenomena observed:

1. A determining vs. determined dimension
2. A general vs. specific dimension
3. A collective vs. individual dimension
4. A global vs. local dimension

All dimensions must be related to a perspective and a level of abstraction which means that a given phenomenon can be seen as determined or general in relation to one phenomenon without being so in relation to another phenomenon. The dimensions are not dichotomous but rather ordinal and continuous which means that one phenomenon can be more or less determining, general, collective or global than another.

With regard to the application of the four dimensions to communicative aspects of a social activity, features of the communicative behavior are, broadly speaking, regarded as determined and features of the social activity, context and participating individuals are regarded as determining. In particular, general determining features are given by the natural, cultural (including language system) and social institutional environment, while specific determining factors are given by particular activities or particular individuals pursuing the activities.

General features of communicative behavior are features like the general distinguishability of a phonological, morphological and syntactic aspect and the restrictions of a general nature that pertain to these aspects. Specific features of communicative behavior arise, for example, from its employment in a particular social activity by particular persons.

With regard to the dimension collective-individual; collectively determining phenomena can affect all participants in communication equally. Some of these are general such

[2]For a classical illustration see, for example, Malinowski's [16] analysis of the instrumental role of language and communication for a Trobriand fishing trip.

as laws of nature or certain cultural conventions. Others are more specific and dependent on particular activities. Individually determining phenomena are phenomena that differentially affect communication through the participants in the interaction. Such phenomena are mainly of three types: biological (e.g. a physical handicap affecting communication), psychological (e.g. beliefs or other attitudes affecting communication) and social (e.g. sex role, age role, social class). If individual features are general, i.e. are shared by all individuals, they, by this very fact, also become collective (e.g. certain expectations concerning motivation, rationality and agency of the individuals pertaining to communication; cf. Grice [12] and Allwood [1]).

General collective features of the (determined) communicative behavior are such features as turntaking, sequences and feedback (cf Allwood [3] and Allwood, Nivre and Ahlsén [5]). Specific collective features, for example, arise as a result of the employment of communication in a particular activity. Thus, turntaking and feedback are very different in a formal lecture and in an informal conversation. In the former case, we have extremely asymmetric turntaking and very restricted opportunities for overt verbal feedback-giving from the listeners, while in the latter case we have symmetrical turntaking and quite unrestricted feedback-giving from the listeners.

Finally, we turn to the dimension global-local. Global determining phenomena are phenomena which affect communication throughout the course of an activity. Such phenomena can then, in accordance with what has been indicated above, be subclassified as general or specific and collective or individual. Local determining phenomena are phenomena which only determine communication during a phase or a part of an activity. Like global determining phenomena, local determinants can be more or less general and more or less collectively shared. We may further note, that since the difference between local and global, among other other things, involves a difference in level of abstraction, communicative behavior which on a global level can be regarded as, largely speaking, determined, with general and specific aspects and individual and collective aspects, on a local level,becomes equally determining and determined. The preceding discourse, in particular the preceding utterance, always has a very strong determining influence on the succeeding utterance. Every utterance has the dual property of being determined by preceding discourse and being determining for succeeding discourse. The consequence is that local communicative behavior is determined, not only by global determining factors, but also by local determining factors including preceding communicative behavior. The result can be seen with regard to the choice of local communicative exchange patterns to continue interaction, e.g. an unclear question might lead to a clarification sequence as well as in the way each new utterance both is bound by and continues preceding discourse.

In table 6.1 below, an overview is given of the dimensions in communicative analysis and their intersection.

In what follows, we will mainly consider the two rightmost (specific) columns of the diagram, i.e. we will be concerned with the analysis of a particular (specific) activity—an

Table 6.1: Dimensions of Communicative Activity Analysis and their Interaction

		General		Specific	
		Collective	Individual	Collective	Individual
Determining	Global	Nat. env. Culture Language Social inst.	Human nature -rat. agent -biol. psych. Social characteristics	Activities -purposes -roles -artifacts -environ.	Ind. identity -biol. -psych. -social
	Local	Activated aspects of global, coll.	Activated aspects of global, ind. e. g. psych. moods	Activated aspects of spec. coll. global + preceding discourse	Activated aspects of spec. ind. global
Determined	Global	Dim. of coll. comm. beh. e. g. turntaking, feedback, sequencing	Dim. of ind. comm. beh. e. g. phonology, morphology syntax	Coll. comm. behavior relevant to activity	Ind. comm. behavior relevant to activity and coll. comm. behavior
	Local	Activated relevant aspects of coll. comm. behavior	Activated relevant aspects of ind. comm. behavior	Activated relevant coll. comm. beh. + obligated by prec. discourse	Activated required ind. comm. beh. + optional behavior

experimental simulation of a computer assisted system for factual information seeking. In section 6.2, we give an account of the global determining factors of this activity, concentrating on specific collective factors such as acitivity purpose and roles, and on general individual factors such as are encoded in the Gricean maxims (cf Grice [12]). In section 6.3, we discuss local determining factors, and in section 6.4 we analyze the functional structure of individual communicative expressions and the determining role of preceding discourse. Section 6.5 contains an analysis of a specific WOZ dialogue, taken from the English corpus. In section 6.6, we discuss the possibilities of formalizing regularities and functional dependencies found ing dialogues, and we conclude in section 6.7 by discussing the relation of the proposed analysis to the architecture of the PLUS system.

6.2 The Wizard of Oz Experiment—Global Determining Factors

6.2.1 Specific, Collective, Global Factors

Activity Purpose (Function)

The purpose or function of the activity was explicitly given to the subjects through an oral briefing and written scenarios. The given purpose had a multi-layered character as follows:

1. Participation in a scientific experiment of WOZ character (only known to the research staff)
2. Participation in a scientific study testing a prototype for an NL information system concerning the yellow pages.
3. Participation in three simulated factual information seeking tasks, where subjects are told to "play-act" the role of the user. The factual information tasks concern:
   (a) Car hire
   (b) Restaurants
   (c) Personal insurance

The three tasks were described in "scenarios" (see appendix A).

Roles

Persons participate in activities by occupying or playing roles which have been constituted by the activities. Such roles can often be characterized in terms of socially stereotyped rights and obligations. For example, if you are a customer in a shop, you normally have the right to receive information about the quality and price of the goods for sale and you have the obligation to pay for the goods should you wish to acquire them.

Similarly to the fashion in which activities can be simultaneous, and superposed, roles too an be simultaneous and superposed. Imagine in the previous example that the customer was accompanied by his child in the shop. His customer role would then be superposed on his father role. The roles would still be separate, i.e., the rights and obligations of both roles would still be separately valid, even if the resulting behavior might be influenced by both roles. In general, relations between roles mirror relations between the activities they are constituted by.

Logically, an activity or a role can be a specification of, a part of or conjoined with another activity or role. Thus, the Wizard of Oz experiment and the wizard role are specifications of a scientific experiment and the researcher role, respectively and communication and the communication roles (sending and receiving) are parts of the WOZ research activity and wizard role, respectively.

More generally, two or more roles may be conjoined, e.g., to form (parts of) a third role. The communication role, the data-base managing role,. the role of observing and keeping a record (log-file) of the interactions are parts or sub-roles in relation to the wizard role. Thus, the wizard role may be viewed as a conjunction of these sub-roles.

Roles are instantiated through instances of action and behavior. In particular, the following holds:

> If R1 is a specification of R2 and A instantiates R1 then A instantiates R2 as well.
>
> If R1 is a part of R2 and A instantiates R1 then A instantiates R2.

For instance, the communication role that forms a part of the wizard role is also a specification of a (general) communication role. Thus, any action instantiating the wizard-communication role also instantiates the wizard role and the (general) communication role.

Further, we may associate an extension to a role consisting of all possible actions that instantiate the role (role potential).

From the above considerations, it follows, for example, that the extension or role potential of the WOZ-communication-role is the intersection of the extensions of the WOZ-role and the (general) communication role.

Returning to the relations between persons and roles, we have distinguished two such relations (i) occupying (having) a role and (ii) playing a role. A person may occupy a role without being aware of this fact if he meets the requirements for occupying the role. He may also, however, play a role without meeting such requirements if he exhibits the expected stereotypical role behavior or intends to exhibit such behavior.

A consequence of this way of characterizing a role is that one may be assigned a role without playing it or play a role without having been assigned to it.

In what follows we will characterize roles by the rights (permissions) and obligations which are connected with the role. We will also, when the need arises, attempt to characterize the requirements which have to be met in order to occupy a role. Such requirements can be behavioral or substantive and can concern communicative ability or certain types of knowledge and skills, such as knowledge of computers in the wizard role.

Over and above requirements, desiderata for playing a role well can be added. Examples of such additional desiderata might be speed of typing, intelligence in problem solving or perseverance, etc. Possibly, after more careful study, it might be possible to divide requirements relating to a role into necessary, sufficient and facilitating requirements for the role.

The Roles of the WOZ Experiment

The main roles of each instance of the experiment are:

1. Experiment manager
2. System
3. User

The roles of the interactions (wizard and user) could also be described as role positions. This is so, since the person who occupies one of these role positions, in fact, "plays several roles" or phrased alternatively participates in several "role levels".

1. The Experiment Manager

   The role of the experiment manager is limited to the initial briefing and instruction phase and to a final phase where subjects are asked to complete a questionnaire including questions asking them to evaluate the system. For the study of the corpus this material is relevant only as a contextual background factor.

2. The System

   The person occupying the "system position" is through the activity connected with 3 role levels where the main level is the "system level".

   (a) *Researcher role:* He or she is a researcher in computational linguistics involved in the PLUS project, more specifically, in the present situation he/she is the "wizard".

   (b) *The system role:* He or she is a "wizard", i.e. trying to simulate a factual information-giving system for the 3 selected areas of the yellow pages.

   (c) *Communication roles:* The person taking the system role alternates between being in a sender role, a receiver role and a data base manager role.

   Let us know try to characterize the system role in terms of the rights and obligations which connect it to the communication roles of receiving and sending information

(a) *Receiving and processing:* No restrictions—the wizard is only limited by his humanly given capacity for understanding. This entailed, for example. full understanding of spelling errors, abbreviations, ellipses, cross reference, etc. (To make future studies more realistic, appropriate restrictions could well be introduced here).

Since the wizard is incrementally interpreting screen messages, he or she has increased time for data base management and choice of appropriate reply. Given the role of information provider, the system will have a tendency to interpret each message from the user as a request for information or acknowledgement of such information given. Since the wizard is acquainted with the scenarios, he/she will often be able to predict what the next move from the user will be and can use this information to interpret the users message. The data base manager role obligates the wizard to find relevant information in the data base and to decide whether this can be transmitted by means of available canned message forms.

(b) *Sending (production):* As a sender, the system is also obligated by the restrictions mentioned above, however, in short form the general restriction might be expressed as "being as computer-like as possible". In the experiments carried out, the wizard, therefore:

- used canned answers as often as possible.
- used canned answers as patterns also for other responses which more explicitly meant that complete sentences according to "school grammar" were used with the exception of lists, and that restricted vocabulary in the direction of official bureaucratic language was used.
- took no initiatives except those motivated by the task of providing the user with information within the three scenarios. The initiatives that were taken would, therefore, typically be questions for specification and clarification.
- did not provide information which went beyond the limitations given by the available Yellow Pages data base. In a few cases, common sense considerations were employed. For example, the wizard knew that Mexican cuisine is well-known for spicy food.
- never evaluated any products in the yellow pages data base, even when evaluations were asked for.
- never gave information on geographical distance. This was not available in the database for the application domain.
- Restricted helpfulness to the following aspects:
  - full comprehension, e.g. the system could always identify locations (limited by WOZ personal knowledge), even though no information over and above address was available in the data base
  - suggestions for where (phone No.) the user could turn for more information

- information given was always by the wizard believed to be correct and relevant for the user
- the information was presented as clearly and perspicuously as possible, e.g., attempts were made to reduce "large amounts of information" presented by asking for specifications from the user.

3. The User

(a) *Subject role:* On a first level, the user occupies the role of "subject-in-scientific-experiment". This carries with it the right to be treated courteously and the obligation to follow instructions given. A consequence of this could be, for example, an urge to test the system's robustness by typing in abbreviations or misspelled words. However, this last urge is probably checked by an opposite urge to spell correctly when observed by one's teachers. The subjects being students might have identified the situation as one in which they themselves in some sense were tested.

(b) *The user role:* On a second level—the main level—the subjects were asked to play the role of "user" or "information seeker" in three different scenarios. The obligations of a user are relatively restricted and are confined to communicating clearly. Since the user believes he is communicating with a machine, he/she probably believes he/she has to spell correctly and write grammatically. This is, however, counteracted by the user's information that he/she is communicating with a new prototype with more human-like qualities, which perhaps, therefore, could be expected to handle the lack of such features. Communicating with a machine further produces fewer obligations concerning politeness and manners. Obligations involving giving information to the system about textual relevance and coherence relations might also be diminished. Since, furthermore, the interaction is preserved on the screen, 40 lines back (about 10 messages) and a prompter ends and starts each new contribution, there is almost no need for requests for repetitions or feedback information.

The rights of the user are "to get his/her money's worth" of correct, adequate helpful and courteous factual information. Since these rights are fairly vague and far reaching, it is to be expected that the user's expectations might be more far reaching than the system's capacity for providing information.

The resulting behavior of the persons in the "user role position" is probably to be seen as a compromise between the rights and obligations connected with the role of "research subject" and rights and obligations connected with the role of "user" proper. Details of the user role are given in the scenarios above.

Another perspective on the roles of "user" and "system" might be gained by comparing them to, or maybe even taking them to the instances of the "master" and "slave" roles,

respectively. The use is the master and has great freedom of action - he/she can do more or less what they want. The system is the slave who has put up with the users whims, always interpreting them charitably and cooperatively.

6.2.2 General, Individual, Global Factors

From a general perspective, individual communication requires both action and interaction. Communication as well as action and interaction, in their turn, are based on certain features of human nature. These features can, for example, be divided into biological psychological and social features. They include such abilities as the ability to send, receive and process acoustic or optical waves in a psychologically and socially relevant way—an activity in which the three types of features are integrated with each other.

One might also say that communication, action and interaction rely on at least three types of constraints and enablement conditions:

1. *Causal*—which pertains perhaps primarily to the interaction of the human organism with the physical environment
2. *Conventional*—which pertains perhaps primarily to the psychological and social sides of human nature.
3. *Rational-functional*—which pertains primarily also to psychological and social sides of human nature.

We will here only remark briefly on constraints and enablement conditions of the third rational-functional type. The extent to which there are constraints and enablement conditions of this type depends on the extent to which communication, action and interaction are rational and functional.

Starting with Grice [12], who formulates four maxims (quality, quantity, relation and manner) and an overriding principle of cooperation in order to capture certain preconditions (in the Kantian sense) of conversation, there have been a number of attempts at finding such constraints and enablement conditions for communication. Sometimes these attempts have been made with explicit consideration of other types of action and interaction than communication. Sometimes they have not. In Allwood [1], the connection is made with other types of action and interaction and six maxims of motivated, rational action are given. In addition, an analysis is given of the role of ethical motivation and cooperation in communication. Horn [13] presents an analysis based only on the Gricean maxim of quantity and Sperber and Wilson [23] present an analysis based on the notion of relevance. In what follows, we will make use of a combination of the principles found in Grice [12] and Allwood [1].

Another set of general constraints and enablement conditions are connected more specifically with communication. There can be no communication unless the parties are willing and able to:

1. start and continue communicating,
2. perceive each other's messages,
3. understand each other's message,
4. react to each other's messages.

These basic communicative functions have led to the development, in all languages, of a set of elementary feedback mechanisms (cf Allwood [4] and Allwood, Nivre and Ahlsén [6]), which enable communicators to on-line interactively manage the functions in question. Other examples of interactive communication management functions are turntaking mechanisms (cf Sacks et al [21]) and sequencing mechanisms (cf Sacks and Schegloff [22]).

Over and above interactive communication management, there is also a need for own communication management. A sender or a speaker needs, on-line, to be able to plan and change his message in direct response to his own ideas and the interlocutors reactions. A set of such mechanisms is described in Allwood, Nivre & Ahlsén [5].

Finally, it should perhaps be repeated that features which generally are present in the communication of all individuals, ipso facto, become collectively relevant. If there is general awareness of the features, as there is in the case of many of the features of rational functionality, they can become collectively relevant features of the awareness and intentionality which are connected with communication, action and interaction.

6.3 Local Determining Factors

6.3.1 Local Determinants

Communicative behavior on a local level is influenced by preceding discourse, especially the immediately preceding utterance, and by activated relevant aspects of the general and specific globally determining factors. In the simulated computer supported factual information dialogues we are considering here, this could involve such things as specific linguistic and cultural constraints on the asking and answering of questions. It could also involve activation of specific competence requirements by the individuals playing the typical roles of the activity. In this case, being the wizard (i.e., the system) requires a number or rather specialized competencies and abilities, while the user role requires somewhat less specialized competence.

6.3.2 Functional Specification of Activities

When it comes to the local activation of the function and purpose of the activity itself, this can often be captured by functionally specifying the activity into subactivities which have a means-ends relation to the activity as a whole. Each subactivity has its own function or goal which provides a local functional context within the activity as a whole. A subactivity can then be broken down into further subactivities, if this should be desirable for some purpose. As a baseline for the subactivities, we will assume a set of functions or purposes which can be realized through individual communicative behavior. When appropriate, we will refer to these as speech acts or, more generally, as communicative acts. It should here be noted that there is no one-to-one relationship between an instance of individual comunicative behavior and the comunicative functions, purposes, acts that can be realized through the behavior. In general, we assume that the relationship is one-to-many, since a statement like *it's raining*, in a given context, easily, simultaneously to being a statement, can function as a request (e.g. to close the window), a warning and a reminder.[3] It might even be argued that the relationship is many-to-many, since one can imagine individuals jointly making statements, by two separate utterances, like in the following example

A: Bill is going out
B: at six oclock

If one feels uncomfortable with the notion of collective speech acts, one might alternatively say that A and B have both produced partial speech acts.

Functional specification of the WOZ experiment and the factual information seeking dialogue

The overriding purposes of the WOZ experiment and the factual information seeking dialogue might be expressed as 1 and 2, respectively:

1. To carry out an experiment simulating a factual information seeking dialogue with a robust and cooperative computer concerning three areas of the yellow pages.

2. To provide factual information about three selected areas of the yellow pages in a way which is robust (tolerates spelling errors, nonconventional grammar, lack of spacing, lack of explicit information) and cooperative (considers the probable information needs of the user and gives quick, competent, sufficient, relevant and perspicuous information).

[3]Cf Allwood [2].

The purpose of the WOZ expreiment can then be broken down into subgoals and subactivities like finding willing subjects, deciding how to simulate a computer, deciding what should be meant by helpful, robust and factual information, etc. Since a functional specification of the actual experiment as an activity is not our main concern in this paper, let us instead consider the purposes of the factual information seeking dialogue itself.

The overriding purpose of the information seeking dialogue can be connected with a number of more or less necessary subgoals and subactivities which serve as a means to realize the purpose.

Subactivities and Subgoals of Factual Information Seeking

1. Opening interation
   - indicating ability and willingness to start (user and/or system)
   - opening channel (e.g., by greeting) (user and/or system)
   - identification of participants (user and/or system)
2. Presenting query
   - eliciting query (system)
   - presenting query (user)
3. Evaluating query
   - interpret query (system)
   - check if information is available and if there is a need for specification (system)
   - express evaluation of query (system)
4. Specifying query
   - elicit specification (system)
   - provide specification (user)
5. Answering query
   - report inability to provide information (system)
   - accept inability and terminate (user)
   - give desired information (system)
   - accept information and confirm receipt (user)

6. Evaluating answer
    - interpret answer (user)
    - check if information is adequate, sufficient, etc. and if there is a need for specification or supplementation (user)
    - express evaluation of answer (user)
7. Specifying answer
    - elicit specification (user)
    - provide specification (system)
8. Supplementary query
    - present supplementary query (user)
    - evaluate supplementary query (system)
9. Expression of satisfaction
    - elicit information regarding whether the users information need is satisfied (system)
    - indicate that information need is satisfied (user)
10. Closing
    - summarize the information which has been given (system and/or user)
    - express gratitude (system and/or user)
    - indicate ability and wish to end (e.g., by greeting /system and/or user)

In shorter form the list of subactivities might be given as follows:

1. Opening
2. Query (main or supplementary)
3. Evaluation
4. Specification
5. Answer
6. Evaluation
7. Specification
8. Supplementary query

9. Expression of satisfaction
10. Closing

A further simplification would be to say that information seeking dialogues consist of:

1. An opening
2. Queries (main and supplementary)
3. Answers
4. A closing

These four functions are realized by one or more utterances, where each utterance elicits an evaluation of its own contribution and the next utterance either implicitly or explicitly gives this evaluation. If the evaluation is negative (or unclear), this can be directly reported by writing e.g, "I don't understand" or " I don't want to answer" etc., or more indirectly by writing "what", "pardon", etc., or by asking for a specification or a clarification. Often the direct and the indirect negative report are combined as in "I don't understand, could you please clarify".

If the evaluation is positive, this can also be directly reported by saying "I understand" etc., or indirectly by carrying out the evocative intention of the preceding contribution. As in the negative case, positive direct and indirect reactions can be combined as in "I understand. The answer is 2 m".

All utterances, e.g. queries and answers, but also those which realize openings and closings are evaluated and, if need be, elicited, specified and clarified. Evaluation is, thus, a necessary ingredient of communication, while elicitation, specification and clarificiation are useful ancillary communicative moves.

6.4 The Functional Structure of Individual Communicative Expressions and the Determining Role of Preceding Discourse

6.4.1 Obligations and Options

Let us now take a closer look at the functional structure of an individual communicative expression (we will use the term communicative expression as the general correspondent to what in spoken language is called an utterance). With the possible exception of the first and the last communicative expression of an interaction, the functional structure of a communicative expression can be characterized as follows:

1. Obligated (functions obligated by the (immediately) preceding discourse)
2. Optional (functions which are not obligated by preceding discourse)
3. Obligating (functions which are obligating for the (immediately) succeeding discourse

Above we have argued that each communicative expression with all its functions is relevantly influenced by certain global and local determining factors. Among the local factors, we have mentioned that an important part is played by the discourse which precedes a given expression (in particular the immediately preceding expression).

The obligated functions of an expression are, thus, those functions of an expression which, in a given context, meet certain obligating functions of a preceding expression. Two main types of obligating and obligated functions may be distinguished (cf Allwood, Nivre and Ahlsén [5]):

1. Obligations arising from requirements of interactive communication management
2. Obligations arising from requirements generated primarily by the interaction of the evocative dimensions of (non-management directed) communicative acts with the embedding (activity) context

6.4.2 Obligations and Options Arising from Requirements of Interactive Communication Management

Interactive communication management includes the following three types:

1. Basic communication feedback, i.e. whether sender and receiver are willing and able to
   - continue (start or end) communicating,
   - perceive message,
   - understand message,
   - react to message (+ an indication of the nature of this reaction).
2. Turntaking, i.e., mechanisms for distribution of the right to communicate (more properly, for distributing the right to be sender and possibly also the obligation to be receiver). Turntaking functions include: having—not having turn, taking (interrupting)—refusing turn, accepting—giving up turn, keeping—losing turn and assigning a turn. As we can see, turntaking is, thus, conceptually related to the ability and willingness to continue communication.

3. Sequential structuring (of communicative acts, subactivities, topics, etc.). Since this last type of management function concerns, for example, entering or leaving subactivities, it is more often globally than locally determined.

Each interactive communicative expression is assumed to realize one or more positive or negative feedback and turntaking function. The realization can be explicit by special linguistic mechanisms or implicit through communicative acts which presuppose particular feedback and turntaking functions. Some interactive communicative expressions, besides feedback and turntaking, realize global sequential structuring functions.

The requirements of feedback and turntaking are met through the pairing of obligating and obligated aspects of adjacent expressions with regard to the basic communicative functions of feedback, i.e., continuation, perception, understanding and reaction, as well as turntaking.

The options of feedback and turntaking concern:

- whether the information is implicit or explicit,
- choice of specific kind of feedback and turntaking information.

The corresponding obligations could be restated as the single requirement that each contribution gives and elicits information about feedback and turntaking. Note that a consequence of this view is that also a refusal to communicate will count as giving information, albeit negative information, on feedback and turntaking.

6.4.3 Obligations and Options Arising from Interactive Requirements of Communicative Acts

Besides obligations related to the interactive management of basic communicative functions, there are obligations related to particular communicative acts. Perhaps the best way of illustrating this is to present a table of these obligations as they are related to the stereotypical expressive and evocative functions of the four moods declarative, interrogative, imperative and exclamative (cf. also Allwood, Nivre & Ahlsén [6]).

The claim made by the table is that the declarative, *the door is open*, stereotypically is used to express a judgement by the speaker with the evocative intention of getting the listener to share the judgement. Similarly, an interrogative like *is the door open*, stereotypically is used to express an act of wonder, connected to a desire for information, with the evocative intention of receiving the desired information from the listener, etc. When communicative expressions realizing the stereotypical combinations of the expressive and evocative functions of the four moods are used for communication in

Table 6.2: Mood Functions

Mood	Expressive	Evocative
Declarative The door is open	Judgement (Belief)	Shared judgement (Belief)
Interrogative Is the door open?	Wonder Desire for info	Answer Giving desired info
Imperative Open the door!	Desire	Satisfaction of desire
Exclamative An open door!	Arbitrary other attitudes	Attention

some social activity, certain stereotypical obligations or with a weaker formulation, expectancies are generated.

These obligations are the result of the interaction of mood functions with (i) considerations of rational motivated action, ethics and cooperation, (ii) language specification conventions and (iii) activity requirements and conventions.

The obligations are of two types:

1. obligations contracted by the sender through the use of the mood (sender commitments or sender obligations)
2. obligations which the sender attempts to put on the receiver(s) through the use of the mood. These obligations directly correspond to:
   (a) an obligating function generated by the mood carrying expression in a particular context (potential receiver obligations) and
   (b) the corresponding obligated functions found in an actual response to this expression (response obligations or actual assumed receiver obligations).

If we connect the mood functions with mood obligations that might arise in some particular context, the result generated might, for example, be the one shown in table 6.3.

In the table, CPU indicates contact, perception and understanding, the three basic feedback functions of IA management. Let us illustrate the way the table is to be read by analyzing an example of a factual information seeking dialogue.

Table 6.3: Moods: Functions and Obligations.

Mood	Expressive	Evocative	Sender related obligations	Receiver related obligations	
				Potential receiver obligations	Response obligations or Actual resumed receiver obligations
Decl	Belief	Belief	1. Actual belief 2. Evidence 3. Belief that R can understand and evaluate expressed judgement	1. Accept CPU, evaluate, report 2.Report on evaluation	1. Report on acceptance 2. Report on evaluation
Int	Wonder	Answer	1. Actual wonder 2. Need of info 3. Belief that R can provide info	1. Accept CPU, evaluate if possible to provide info and report 2. Find info 3. Report	1.Report on acceptance 2. Answer (report on found info)
Imp	Desire	Satisfaction of desire	1.Actual desire 2. Need 3. Belief that R can satisfy desire without risk to himself	1. Accept CPU, evaluate if possible to satisfy desire and report 2. Satisfy desire 3. Result or report on result	1. Report on acceptance 2. Desired action (or report on desired action)
Excl	Any attitude	Attention	1. Actual attitude	1. Accept CPU	Optional report on CPU

(137) U : I would like some information about Korean restaurants
S : Unfortunately, there is no information available. We have no Korean restaurants listed in Göteborg. Would you like information about any other kind of restaurant?

Since we do not have the discourse preceding the user's contribution, we will abstain from commenting on it's obligated aspects. The remainder of the analysis could look like this.

User: I would like some information about Korean restaurants

- Mood: Indicative:Declarative
- Optional:
  - Speech Act: Request
  - Expressive: Judgement (belief) concerning own likes and desires to the effect that sender has a desire for information
  - Evocative: That receiver shares this judgement (belief) and, consequently, tries to satisfy desire for information
- Obligations:
  - Sender related obligations:
    1. Sender has the belief expressed
    2. Sender has evidence for this belief, i.e., that sender has a desire for information
    3. Sender has a reasonable need for the information
  - Potential receiver obligations:
    1. Accept CPU and evaluate and report
    2. Evaluate whether judgement can be shared and report
    3. Evaluate whether desired information can be provided and report
    4. If possible, provide information

System: Unfortunately, there is no information available. We have no Korean restaurants listed in Göteborg. Would you like information about any other kind of restaurant?

- Mood: Indicative:Indicative:Interrogative:YNQ
- Speech Act: Statement (excuse):Statement (explanation):Question (compensating offer)

- Obligated:
  - IA:
    1. Implied CPU and turn acceptance
  - CA:
    1. Implied sharing of judgement
    2. Explicit negative evaluation of whether information can be provided (cancelling (4) above)
- Optional:
  - Speech Act: Explanatory statement:question giving offer
  - Expressive:
    1. judgement
    2. desire for information
    3. willingness to provide beneficial action for receiver
  - Evocative:
    1. that receiver shares judgement and accepts excuse
    2. that receiver provides desired information
    3. that receiver accepts beneficial action
- Obligations:
  - Sender related obligations:
    1. Sender has beliefs expressed
    2. Sender has evidence for beliefs, i.e. has searched data base
    3. Sender has the desire for information and the willingness to provide beneficial action that he has expressed
    4. Sender has a reason for this desire and willingness, i.e. that it has the role of helpful information provider in a factual information seeking system
  - Potential receiver related obligations:
    1. Accept CPU and evaluate and report
    2. Evaluate own willingness and ability to accept statement containing excuse as well as question containing offer and report
    3. If evaluation is positive provide information

In section 6.5 below, we will use roughly the categories presented here to analyze an English WOZ dialogue.

After having somewhat lengthily considered the obligations generated by communicative actions, let us now turn to options. The options connected with communicative action concern:

1. whether to engage in communicative action over and above IA management obligations,
2. choice of type of communicative action compatible with restrictions given by general and specific, global and local, collective and individual determining factors.

The choice of communicative act is, thus, always regarded as optional, within the degrees of freedom given by the determining factors. This is the reason why the label optional, in the example discussed above, has been used to characterize those of the user's and system's communicative acts which are not bound by the preceding contributions.

6.4.4 Obligations, Options, Requirements, Assumptions, Preferences

Instead of the concepts "obligated", "obligating" and "optional" which make a rather strong assumption about the relations which can hold between utterances in discourse, we could have chosen concepts which would have characterized these relations in a different and often somewhat weaker way. More specifically, we could have spoken about preferences, assumptions, expectancies or requirements instead of obligations. For mainly methodological reasons, we have, however, chosen obligations. The methodological reason is that the use of this concept to characterize interutterance relations makes a strong claim about their contractual, glue-like character. This claim can subsequently be modified in favor of concepts which characterize the relations in different or weaker terms.

The choice between the concepts mentioned above is, however, not mutually exclusive. Because of their differences, it is relatively easy to point to cases where several of them might be needed, as in the following:

A: could you come here
B: yes (moves towards A)

A possible analysis of the relations between A's and B's utterances is to say that A's utterance is a request couched in the form of a yes/no question, which has an obligated response, a yes or a no, and as a preferred response a yes. If analyzed in this way, we could, thus, for example, make use of the distinction between obligations and preferences.

6.4.5 Obligations, Options, Backward-Forward Orientation, Boundedness, Novelty of Information and Thematisation

The functional trichotomy obligating-optional-obligated is also related to the distinctions between backward-looking and forwarding-looking, bound and free, response and initiative, old and new, topic and comment which have been proposed by different authors to capture various aspects of the way in which each new utterance in a dialogue both is connected to earlier discourse and unconnected in the sense that it contributes something new. Just as above, we suspect that the mentioned pairs of terms do not cover the same aspects of the interrelatedness of discourse and that, therefore, a very fine-grained analysis could make use of all of them in slightly different ways.

The concepts chosen by us are related to the mentioned concepts roughly in the following way:

With regard to forward and backward orientation, all parts of an utterance can be potentially relevant in both directions, however, the obligated functions are mainly backward-looking and the obligating functions are mainly forward-looking. The optional functions have a lesser bias in either direction but are, if anything, slightly more forward-looking. Concerning the distinction bound-free, the obligated functions are bound by preceding discourse. Similarly, the obligated functions are most responsive and least initiating while the optional and obligating functions are least responsive and most initiating. Finally, the obligated functions of an utterance are perhaps to a greater extent than the optional and obligating functions related to information which is old or in topic or theme position, while the inverse holds for information which is new or in comment or rheme position.

6.4.6 Possible Rules and Regularities

Let us now turn to a consideration of what kind of regularities have been noted so far and to the question of whether rules could be formulated for these regularities and the supplementary question of the nature of these rules. One way of dividing the regularities observed would be the following:

1. Regularities depending on relations within communicative contributions
2. Regularities depending on relations between communicative contributions
3. Regularities depending on relations between general and specific, global determining factors (e.g. Gricean-like maxims and activity function) and local collective or individual features of communicative contributions.

Rules corresponding to regularities of the first type would be rules of phonology, morphology, syntax and semantic composition applying within and/or between the syntactically and semantically maximal units of a communicative contribution. Rules of this type could also concern the sequencing of pragmatic, semantic or syntactic maximal units within a communicative contribution. Finally, rules of this type regulate the placement of interactive (IA) and own communication (OC) management expressions in relation to non-management oriented expressions. For IA management, what here is involved, is mainly rules of addition and insertion, while OC management, in addition to these, also requires rules of deletion and reordering (cf Allwood, Nivre & Ahlsén [5]).

Rules corresponding to regularities of the second type, which are the main concern of this paper, could be called rules of expectable exchange obligations and are of two main types:

1. CIAM exchange obligations
2. CA exchange obligations

Rules of the first type—rules of communicative interaction management (CIAM) exchange obligations–are rules which, in a context sensitive way, state correspondences between CIAM functions. Rules of the second type—rules of communicative action (CA) related exchange obligations—are rules which, in a context sensitive way, derive potential receiver obligations from the performed communicative actions and state expectable correspondences between potential receiver obligations and actual assumed receiver obligations.

Rules corresponding to the third type of regularities are also of several kinds:

1. Rules of functional activity specification - these are statements which specify a given function or activity into subfunctions and/or subactivities (cf above). The functional specification rules, thus, feed directly into the rules for CA exchange obligations.
2. Rules for supplementing local sender and receiver obligations by deriving obligations from activity roles and general considerations of rational, motivated (ethical, cooperative) action.
3. Rules stating restrictions and/or requirements on topic development, inference drawing, etc. We will not deal with this type of rules in this paper.
4. Rules stating restrictions and/or requirements on interactive task (communicative or otherwise) accomplishment (cf below).
5. Rules stating restrictions and/or requirements on information update, e.g. coreference restrictions, nonmonotonicity, etc.

In the next section, we will now proceed to an empirical illustration of phenomena connected with rules and regularities of type 2 and types 3.1, 3.2 and 3.4. In section 6.6, we will discuss some possibilities of formalizing such rules and regularities.

6.5 Analysis of an English WOZ Dialogue

6.5.1 Introduction

We will first make some further remarks on the background assumptions for the analysis to be pursued below. The model we are using is roughly depicted in figure 6.1.

The states, which are affected by the discourse, contain the participants' "current belief-sets", as well as their "current task structures". (These are regarded as dynamic structures, and kept separate from some presupposed static structures.) Especially important are the participants' beliefs about each others beliefs and goals, since these will be instrumental in their attempts to cooperate, and to their manner of applying Gricean-like maxims. (Being relevant means, among other things, adapting to the other's interests; abstaining from telling what is already known presupposes some opinion on what is known by the other, and so on.)

The task structures consist of a number of tasks related to the activity of which the interaction is a part, as well as to tasks (obligations, reactions to expectancies) deriving from the communicative interaction itself. In the preliminary analysis we have explicitly listed only tasks related to the main activity, and given only a rather course-grained analysis of them at that. As the interaction goes on, the participants generate tasks for each other (and possibly for them selves). A task generated may be refused, but normally accepted. In order to solve a task, a participant may break it down into sub-tasks, some of which may be presented to the other party, who may in turn do likewise, and so on.

The set of tasks and their interrelations form the task structure, and we have given a picture of how this is changing during the interactions, as well as by what "general linguistic means" it is affected.

The *effects* of utterances, thus, correspond to the change of state that they cause/promote. It should, however, be noted that state-changes are affected also by other things than utterances (like, e.g., data base searches, internal planning and reasoning etc.) Thus the solution, and resulting canceling, of a task is normally not brought about by utterances alone.

The *properties*, both internal and relational, of utterances are also recorded in the analyses. As we have seen above, the model used for an analysis of interutterance relations looks roughly like figure 6.2.

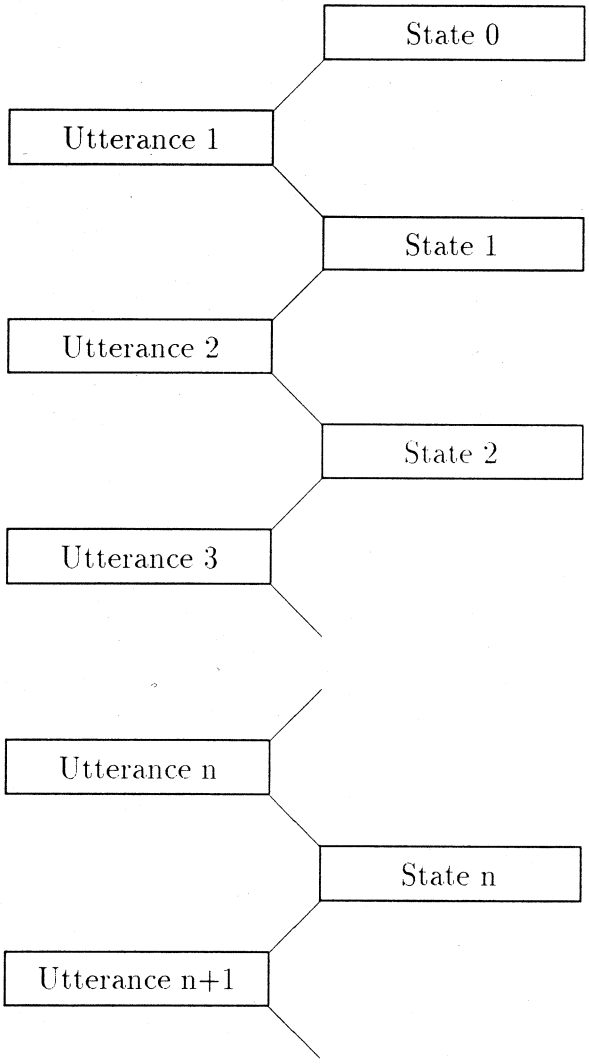

Figure 6.1: Interutterance Relations and State Changes

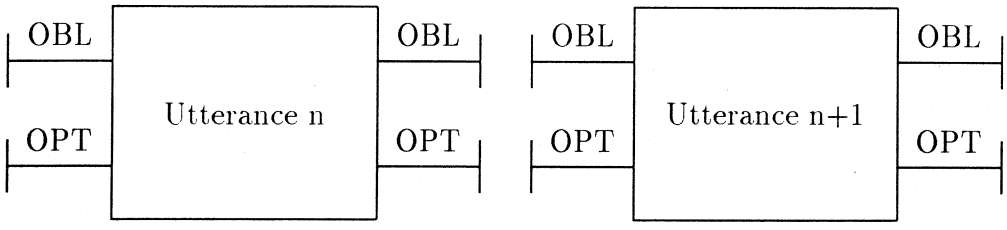

Figure 6.2: Interutterance Relations: Obligations and Options

To the left we have "in-coming" obligations and options, and to the right we have "out-going" obligations/expectancies and options. A sequence of utterances form a "locally coherent" part of a discourse if the "in-coming" parts of each utterance in the sequence "match" the "out-going" parts of the (immediately) preceding one. Such a sequence forms, so to speak, an interlocked chain, where the locks (or glue) are (is) formed by the matching OBL and OPT parts of the utterances.

Amongst the properties of utterances recorded in the analyses some are "inherited" from the expression type of the token used in the utterance. (Syntactic category and structure, as well as mode, belong here.) These properties are thus rather context-independent. Other properties are context-dependent, meaning that different tokens of the same expression type could differ regarding these properties.

6.5.2 Conventions of Analysis and Coding

In what follows we will "utterance-by-utterance" analyze one English WOZ dialogue, attempting in the analysis to make use of the analytical categories introduced above (mainly in section 6.4).

More specifically, the format will be that each utterance (communicative expression) is listed on a separate page with a partially coded analysis following it. The coding categories are, as far as has been possible, structured in the form of attributes and values. Below we now explain the categories with reference to the first "utterance" of the English dialogue. We will later attempt to analyze this dialogue in total.

Example 1

U1: can i have a list of car hire companies in the Entwistle area.

Mood: Inter:aux.inv.Q:Y/N

<OBL

IA Man: {TM :Accept turn
FB:CPU Seq: }

CA Dep: Follow instruction/accept offer*
Cont dep: classification of SA type of U1 partly given by screen instruction

<>OPT

OC Man:

CA: (Question), Request
Expr: desire for info, wish
Evoc: satisfy desire and wish

>OBL

CA Dep:	Sender(U): Have desire, have wish, have motive Receiver(S): Evaluate question and request, carry out request Cont dep: Answer should contain list satisfying predication of U1
IA Man:	{FB: CPU} TM: Accept turn
Activity Status:	Opening—Query 1 <opening interaction, presenting Query 1 >elicit answer to Query 1
System tasks:	ST1 = find and present list
User tasks:	[UT1 = Enter request]*

*) This is based on a guess about how the interaction started. We presume that the screen contained a text which could be interpreted alternatively as an instruction to subjects participating in the experiment or an offer to the users, e.g., something like: Please enter your request.

The Contributed Communicative Expression (Utterance)

The contributed communicative expression of example 1 is:

U1: can i have a list of car hire companies in the Entwistle area.

U1 indicates the name of the user's first "utterance". The utterance is reproduced without "cleaning up", i.e. uneven spacing, lack of capitals, spelling errors, etc. are included.

Mood

Mood: Inter:aux.inv.Q:Y/N

The first coding which is done of an utterance is to assign a mood to it. So far, we have made use of the following moods, where ":" indicates successive values specifying attributes:

Indicative (Indic)	: Declarative (Decl) : Phrase (Phr)—includes sentences, phrase categories and word class categories

: List

Interogative (Inter) : Auxilliary inversion (aux.inv.) : Yes-No-question (Y/N)
: Do-supported question (DQ) : Yes-No-question (Y/N)
: Tag question (TQ) : Yes-No-question (Y/N)
: Question word question (WhQ) : Wh-word
: Disjunctive question (DisjQ) :Yes-No-question (Y/N)
: Alternative question (Alt)

Imperative (Imp) :

Exclamative (Excl) :

Obligated Functions

<OBL

IA Man: {TM :Accept turn
FB:CPU Seq: }

CA Dep: Follow instruction/accept offer*
Cont dep: classification of SA type of U1 partly given by screen instruction

This part of the analysis attempts to capture the obligated (<OBL) functions of a contribution (or "utterance", cf above). First we code those functions which are connected with interaction management (IA Man). These are of three types: (i) turn management or turntaking (TM), (ii) feed-back (FB) and (iii) global sequential structure (Seq). As values of turn management we use turn taking—turn rejecting, turn accepting—turn giving, turn keeping—turn loosing—turn returning. As values of feedback we use contact (C), perception (P), understanding (U) and reaction (R). Usually R is left out since its most important aspects are elaborated by the communicative act dependent (CA Dep) functions. Values of global sequential structuring (Seq) are, for example, opening, closing, summarization (of subactivities or topics). Seq is only used when there is an explicit linguistic indication of sequential structuring. When the structuring is implicit and concerns the relation between activity and subactivities, rel

"CA Dep" and "Cont dep", as already mentioned, concern obligated functions and content features dependent on the content and type of communicative act of the preceding "utterance". Examples of values of Cont dep are coreferential repetition (Co-ref:Rep), Coreferential anaphor (Co-ref:Anaph), coreferential ellipsis (Co-ref:Ellips). For both CA Dep and Cont dep it is still an open question what a more extensive list of values might look like. In our analysis, as will be seen, we experiment with many different values.

Curly brackets ("{" and "}") are used to indicate that functions are present (or filled) implicitly. Thus in example 1 the TM, FB and Seq obligations are met implicitly.

Optional Functions

<>OPT

Own C Man:
CA: (Question), Request
Expr: desire for info, wish
Evoc: satisfy desire and wish

Two types of functions are coded as optional:

1. Own communication management (OC Man)
2. Communicative act (CA)

These functions are regarded as optional in the sense that they are not obligated by the preceding utterance. They are, however, as has been mentioned above, constrained by global determining factors.

OC Man refers to mechanisms for planning and changing one's own contribution on-line, while communicating. Values of this attribute are, for example, change, choice and process.

CA refers to the communicative acts performed in the contribution. If there are several acts, they are listed as concatenated, e.g. statement-statement-request. As values of this attribute we allow most verbs and nouns that could be used to describe aspects of the act of communicating, including, of course, basic communicative acts such as statement, question, request and exclamation. Sometimes we use ordinary brackets (i.e."(" and ")") to surround communicative act labels we believe to be less important for the following interaction. Thus we believe that, from an interactive point of view, it is more important to classify U1 as a request than as a question.

Each communicative act is characterized by the attributes expressive (Expr) and evocative (Evoc). The values of "Expr" are the attitudes which are expressed by the communicative act, e.g. *statements* express belief, *questions* express desire for information, *requests* express wish and *exclamations* express any attitude. The values of "Evoc" are the evocative intentions which are connected to a communicative act, e.g. *statement*—evoke belief, *question*—evoke answer (providing desired information), *request*—evoke a desired action and *exclamation*—evoke attention. For concatenated communicative acts expressed attitudes and evocative intentions are enumerated in the sequential order of the respective communicative acts.

Obligating Functions

>OBL

CA Dep:	Sender(U): Have desire, have wish, have motive Receiver(S): Evaluate question and request, carry out request Cont dep: Answer should contain list satisfying predication of U1
IA Man:	{FB: CPU} TM: Accept turn

Just as we have asumed that each contribution has two kinds of obligated functions, we asume that it also has two kinds of obligating functions:

1. Communicative act dependent (CA Dep)
2. Interactive management dependent (IA Man)

"CA Dep" has two values, sender related obligations, i.e. commitments made by the sender through the use of a certain communicative act, and receiver related obligations, which the sender tries to get the receiver to assume, and which, in fact, need the receivers acceptance in order to become operational. It is not the sender or receiver role per se, which is obligated through the use of a particular communicative act, but rather the persons taking these roles at a given moment. We have indicated this by putting U for user and S for system in brackets after sender and receiver. This enables us, in a cumulative way, to keep track of the ways in which user and system become obligated through the interaction.

We have assumed two basic types of sender incurred obligations: (i) sincerity related obligations, whereby the sender commits him/herself to have the attitude expressed by the communicative act, and (ii) reasonableness related obligations, whereby a sender is committed to have reasons for his/her expressed attitude; e.g. a belief requires evidence, etc.

With regard to potential receiver obligations, we have also asumed that:

1. The receiver should always evaluate willingness and ability to attend to the evocative intentions connected to the preceding communicative act and report if willingness or ability is not present.
2. If willingness and ability are present, he/she should attempt to realize the goals of the evocative intentions (i.e. believe statements, carry out requests, attend to exclamations etc) and if the result of such attempts is not evident through performed actions, report on the realization and

3. evaluate whether the communicative act activates other obligations he/she might have through activity role or as a motivated rational agent.

The values of the attribute "Cont dep" are supposed to be obligating features which arise as the result of an interaction between the expressive and evokative aspects of the present communicative act and its content. For example, a which-question does not only expect an answer but, by content dependency, it expects an answer which gives an enumeration (which could contain 0 or more elements) of items satisfying the predication used in the question. This is also true of example 1, where there is even an explicit question about a list.

The attributes "IA Man", "FB", "TM" and "Seq" and their values "CPU", "Accept turn", etc. have the same meanings as discussed above. The only difference is that they are seen as obligating rather than obligated here. In our analysis we have left out "Seq" as an obligating attribute, except in the last contribution, since it is only relevant there.

The curly brackets ("{" and "}"), indicating that functions are implicit, have in our analysis been used to show that FB functions but not TM functions are implicit. The reason for this is that we assume that each contribution is terminated by, eg., pressing the 'return key' which is assumed to show up as, e.g. a prompt appearing on the receivers screen. This would then mean that turn giving obligating turn acceptance is explicit in the system, at hand, even if this can not be seen in the log.

Activity Status

Activity Status: Opening—Query 1
<opening interaction, presenting Query 1
>elicit answer to Query 1

This coding concerns the function of a contribution in relation to activity and local subactivity. "Activity Status" has as values the local subactivity or subactivities. (If several subactivities are superposed the names of the subactivities are given separated by dashes ("—") as in example 1. Each contribution is then specified as to whether it is an obligated (<) or obligating (>) part of a subactivity and a further attempt is made of specifying what type of obligated or obligating activity/subactivity function it is. In example 1, we have guessed that U1, simultaneously, concludes a request to open interaction and a request to present a query, both of which have been elicited by the screen text. These two functions, in turn, create an obligatiion for the system to answer the query.

Task Structure

System tasks: ST1 = find and present list

User tasks: [UT1 = Enter request]*

These codings represent an attempt to combine CA dependent obligations with local (sub)activity obligations, separately for system and user. The tasks represent a kind of "discourse model" for system and user. When a task is first introduced in the analysis, it is, for convenience, given an informal description. When a task has been completed its designating expression (e.g. first user task: UT1) is enclosed in square brackets (e.g. [UT1]). If a task is abandoned or left uncompleted, its designating expression is enclosed in ordinary brackets (e.g. (UT1)). The tasks are numbered according to the order in which they are introduced in the interaction. We can also see that there is a correspondence between the task structure and the activity status of an expression. U1 concludes subactivities 1 and 2, which have both been introduced by the system. Simultaneously U1, thus, provides the actual query which it is the system's main task to answer.

It should also be noted that the record of task structures for user and system form part of the description of the "state-boxes" of figure 6.1, rather than of the utterances. Thus what is coded is an aspect of the state immediately "after" the utterance under which the code is found.

6.5.3 An English Factual Information Seeking Dialogue

On the pages that follow, we now present an "utterance-by-utterance" analysis of an English WOZ dialogue from the UMIST corpus.

U1: can i have a list of car hire companies in the Entwistle area.

Mood: Inter:aux.inv.Q:Y/N

<OBL

IA Man: {TM :Accept turn
FB:CPU Seq: }

CA Dep: Follow instruction/accept offer*
Cont dep: classification of SA type of U1 partly given by screen instruction

<>OPT

OC Man:

CA: (Question), Request
Expr: desire for info, wish
Evoc: satisfy desire and wish

>OBL

CA Dep: Sender(U): Have desire, have wish, have motive
Receiver(S): Evaluate question and request, carry out request
Cont dep: Answer should contain list satisfying predication of U1

IA Man: {FB: CPU}
TM: Accept turn

Activity Status: Opening—Query 1
<opening interaction, presenting Query 1
>elicit answer to Query 1

System tasks: ST1 = find and present list

User tasks: [UT1 = Enter request]*

S1: Where's Entwistle?

Mood: Inter:WhQ:where

<OBL

IA Man: {FB: CPU
TM: Accept turn
Seq: }

CA Dep: {Neg eval of ability to carry out request}
Cont dep: co-ref repetition of name in U1

<>OPT

OC Man:

CA: Question of clarification/specification
Expr: desire for info
Evoc: satisfy desire by presenting info

>OBL

CA Dep: Sender (S): have desire, have motive*
Receiver (U): evaluate, satisfy desire
Cont dep: Answer must concern location of named referent

IA Man: {FB: CPU}
TM: Accept turn

Activity Status: Specification (Query)
<Evaluate ability to answer query, Elicit specification
>Give specification

System tasks: ST1

User tasks: [UT1], UT2 = provide system with info about location of Entwistle

*) An acceptable motive for the user but not for the system is plain curiosity. Acceptable motives for the system are only such that are related to being robust and cooperative.

U2: Entwistle is outside Bolton.

Mood: Indic:Decl

<OBL
IA Man: {FB: CPU
TM: Accept turn
Seq: }
CA Dep: satisfy desire by answering question
Cont dep: co-ref repetition of name in S1, Answer concerns location
<>OPT
OC Man:
CA: Statement: answer
Expr: belief
Evoc: belief
>OBL
CA Dep: Sender(U): have belief and evidence
Receiver(S): evaluate, make use of IC for carrying out request in U1
Cont dep: Answer to U1 included in list for Bolton area
IA Man: {FB: CPU}
TM: Accept turn

Activity Status: Specification (Query)
<Presentation of specification
>Evaluate specification and answer query

System tasks: ST1

User tasks: [UT1], [UT2]

S2: There are a fair number of companies in Bolton - about three pages worth.

Mood: Indic:Decl

<OBL
IA Man: {FB: CPU
TM: Accept turn
Seq: }
CA Dep: {Pos eval of statement in U2,
Neg eval of ability to carry out request in U1}
Cont dep: co-ref rep of name in U2
<>OPT
OC Man:
CA: Statement {Explanation of why request in U1 is not carried out}
{Indirect request for further specification}
Expr: belief
Evoc: belief. inference that system needs further spec
>OBL
CA dep: Sender(S): have belief. have evidence
Receiver(U): evaluate and react
Cont dep:
IA Man: {FB: CPU}
TM: Accept turn

Activity Status: specification 2 (specification 1 (query))
<Negative evaluation of sufficiency of specification in U2.
Indirect elicitation of further specification.
>Give specification 2

System tasks: ST1

User tasks: [UT1], [UT2], UT3 = react adequately on info given

U3:

<OBL
IA Man: TM: Accept turn
>OBL
IA Man: TM: Loose turn

S3:

<OBL
IA Man: TM: Accept turn
>OBL
IA Man: TM: Accept turn

(Probably un-intended turngiving by user, and intended giving back of turn by system. Since mainly turn management is relevant no other codes have been given.)

U4: Which companies do not have to have the car returned to the same as

S4:

<OBL

IA Man: TM: Accept turn—Return turn

>OBL

IA Man: TM: Accept turn

U5: address as hired from.

Mood: Inter:WhQ:which

<OBL

IA Man: {FB: CPU
TM: Accept turn (twice)
Seq: }

CA Dep evaluation of and reaction to S2 (which serves as further spec of info task)
Cont dep: co-ref with common noun (topic of query 1), presupposed: location of referent of queried topic the same as in U1 and U2

<>OPT

OC Man: user has probably intended to correct message, but instead given turn over to system, who immediately has given it back ??
Probably incomplete deletion of "as".

CA: Question,
Expr: desire for info
Evoc: find and present info

>OBL

CA Dep: Sender(U): have desire for info, have motive
Receiver(S): evaluate and satisfy desire for info
Cont dep: Answer must provide equivalent of listing satisfying pred in U4-5

IA Man: {FB: CPU}
TM: (Return mistakenly given turn). Accept turn

Activity Status: Specification 2 (Specification 1 (Query 1))
<Give Specification 2 (replace Query 1 by more specific Query 2)
>Initiate answer of Query 2

System tasks: (ST1), replaced by the more confined ST2 = find and present list of companies in Bolton with one way rental service

User tasks: [UT1], [UT2], [UT3]

S5:

Budget Rent-a-Car
Bradford House
287-289 Manchester Road
Bolton
0204 391611
One way rentals available
EuroDollar Rent-a-Car
Bridgeman Street
Bolton
0204 365373
Unlimited mileage, one way rentals available, all cars have radios
Hertz
Manchester North branch
061-273 8884
One way rentals, delivery/collection service, cheap local rates, unlimited mileage

Just those three are in my list here.

Mood: Indic:List-Indic:Decl

<OBL
IA Man: {FB: CPU
TM: Accept turn
Seq: }
CA Dep: satisfy desire by answering question in U4-5
Cont dep: enumeration in form of list satisfying pred in U4-5
<>OPT
OC Man:
CA: Statement(answer) of list-statement (comment on previous statement)
Expr: belief-belief
Evoc: belief-belief
>OBL
CA Dep: Sender(S): have beliefs, have evidence
Receiver(U): evaluate (esp. first statement)
Cont dep:
IA Man: {FB: CPU}
TM: Accept turn

Activity Status: Query 2

<Answer Query 2
>Evaluate answer and terminate or continue with new query

System tasks: (ST1), [ST2]

User tasks: [UT1], [UT2], [UT3]

U6: Can I have a price list for the cheapest cars for hire for 1 day from the listed companies?

Mood: Inter:aux.inv.Q:Y/N

<OBL

IA Man: {FB: CPU
TM: Accept turn
Seq: }

CA Dep:
Cont dep: presupposed: query topic restricted by
U1, U2 and U4-5

<>OPT

OC Man:

CA: request for info
Expr; wish for info
Evoc: find and present info

>OBL

CA Dep: Sender(U): wish for info
Receiver(S): evaluate and satisfy wish for info (answer question)
Cont dep: Answer should contain list of referents of query topic restricted by predication and properties given
in U1, U2 and U4-5

IA Man: {FB: CPU}
TM: Accept turn

Activity Status: Supplementary specifying query 3
<Evaluation and specifying query
>Evaluate and answer query 3

System tasks: (ST1), [ST2], ST3 = find info on prices and present cheapest cars from listed companies

User tasks: [UT1], [UT2], [UT3]

S6: I'm sorry I've only got the Yellow pages here, they don't give any prices.

Mood: Indic:Decl-Indic:Decl-Indic:Decl

<OBL
IA Man: {FB: CPU
TM: Accept turn
Seq: }
CA Dep: explanation why answer can't be given
Cont dep: YP presupposed as source of info for query topic
<>OPT
OC Man:
CA: Exclamation and Excuse-Explanation-Explanation
Expr: regret, belief
Evoc: belif, excuse
>OBL
CA Dep: Sender(S): has attitude of regret, has belief, has evidence
Receiver(U): evaluate and react
Cont dep: reaction should accept link between
presupposed info source and query topic
IA Man: {FB: CPU}
TM: Accept turn

Activity Status: Query 3
<Neg evaluation of ability to answer Query 3
>Evaluate response and terminate or continue

System tasks: (ST1), [ST2], (ST3)

User tasks: [UT1], [UT2], [UT3], UT4 = react on info by closing
or by presenting new task for system

U7: Can the car be delivered to me home?

Mood: Inter:aux.inv.Q:Y/N

<OBL

IA Man: {FB: CPU
TM: Accept turn
Seq: }

CA Dep: (Neg eval) suppl. spec. question
Cont dep: presupposed ownership link between car and company,
co-ref. me = user

<>OPT

OC Man:

CA: Question
Expr: want of info
Evoc: find and present info

>OBL

Sender(U): has want of info, has motive for want
Receiver(S): evaluate and carry out request by more than minimal **info**

CA Dep: (answer question), meeting request
Cont dep: ref of car owned by company restricted by properties
in U1, U2, U4-5, U6. Answer also restricted by pred in U7.
Answer should explicitly or implicitly affirm (yes)
or reject (no) the queried proposition

IA Man: {FB: CPU}
TM: Accept turn

Activity Staus: Query 4
<(Neg eval of answer), suppl spec of Query 4
>evaluate and answer Query 4

System tasks: (ST1), [ST2], (ST3), ST4 = find and present info on delivery at **home**

User tasks: [UT1], [UT2], [UT3], [UT4]

S7:

Hertz
Manchester North branch
061-273 8884
One way rentals, delivery/collection service, cheap local rates,
unlimited mileage
say they do that, most other big companies do it if you ask.

Mood: Indic:Decl

<OBL
IA Man: {FB: CPU
TM: Accept turn
Seq: }
CA Dep: Answer to S7
Cont dep: Hertz instance of company owning car restricted to pred of U7.
Pro-predicates"do that","do it" co-ref pred of U7.
<>OPT
OC Man:
CA: Statement
Expr: belief
Evoc: belief
>OBL
CA Dep: Sender(S): has belief, has evidence
Receiver(U): evaluate and react
Cont dep:
IA Man: {FB: CPU}
TM: Accept turn

Activity Status: Query 4
<Answer to Query 4
>Evaluate answer to Query 4 and terminate or continue

System tasks: (ST1), [ST2], (ST3), [ST4]

User tasks: [UT1], [UT2], [UT3], [UT4], UT5 = react on info by closing or by presenting new task for system

U8: Does Hertz have a 24 Hr. service/

Mood: Inter:DQ:Y/N

<OBL

IA Man: {FB: CPU
TM: Accept turn
Seq: }

CA Dep: {Neg eval of sufficiency of answer}
Cont dep: Co-ref repetition of "Hertz"

<>OPT

OC Man:

CA: Question
Expr: want of info
Evoc: find and present info

>OBL

CA Dep: Sender(U): has want of info, has motive for want
Receiver(S): evaluate and satisfy want of info
Cont dep: Answer should affirm or reject queried proposition

IA Man: {FB: CPU}
TM: Accept turn

Activity Status: Query 5
<Neg eval of sufficience of answer to Query 4
>Answer Query 5

System tasks: (ST1), [ST2], (ST3), [ST4], ST5 = find and present info on Hertz 24 h service

User tasks: [UT1], [UT2], [UT3], [UT4], [UT5]

S8: I don't know I'm afraid. It's not specified in their advert, but then it isn't in anyone else's either.

Mood: Indic:Decl-Indic:Decl-Indic:Decl

<OBL
IA Man: {FB: CPU
TM: Accept turn
Seq: }
CA Dep: Neg eval of ability to answer-Excuse-Explanation
Cont dep: Knowledge object given by implied proposition in U8. Pron "it" bound by abstraction over pred in this proposition. "Their" co-ref with Hertz in U8.

<>OPT
OC Man:
CA: Admission-Explanation
Expr: belief (and regret)
Evoc: belief (and excuse)

>OBL
CA Dep: Sender(S): has belief, has evidence
Receiver(U): evaluate and react
Cont dep:
IA Man: {FB: CPU}
TM: Accept turn

Activity Status: Response to Query 5
<Neg eval of ability to answer
>Eval answer, terminate or continue

System tasks: (ST1), [ST2], (ST3), [ST4], (ST5)

User tasks: [UT1], [UT2], [UT3], [UT4], [UT5], UT6 = react on info by closing or by presenting new task for system

U9: Thank you. quit

Mood: Excl:VP-Indic:V alt. Excl:VP-Imp:V

<OBL

IA Man: {FB: CPU
TM: Accept turn}
Seq: closing

CA Dep: (Evaluate response) and thank
Cont dep: Object of thank S1-8, (esp S8),
"quit" linked to screen instruction?

<>OPT

OC Man:

CA: Thank and request
Expr: gratitude-belief motivated by wish (to end interaction)
alt. gratitude-wish (to end interaction)
Evoc: share belief and help realize wish,
alt. satisfy wish by ending interaction

>OBL

CA Dep: Sender(U): be satisfied, have belief, have wish to end,
have motive for these
Receiver(S): evaluate and satisfy wish to end

\+ Cont dep: "quit" bound to interaction as a whole

IA Man: {FB: CPU}
TM: terminate

Activity Status: Closing
<Evaluation of answer and dialogue
>Elicit closing

System tasks: (ST1), [ST2], (ST3), [ST4], (ST5), ST6 = end interaction

User tasks: [UT1], [UT2], [UT3], [UT4], [UT5], [UT6]

S9: You'll need to type quit on a separate line to get off.

Mood: Indic:Decl:Cond
<OBL
IA Man: {FB: CPU
TM: Accept turn
Seq: }
CA Dep: {Neg eval of wish to satisfy users wish (to end)}
Cont dep: presupposed link between typing "quit" and termination instruction given on screen
<>OPT
OC Man:
CA: Statement (instruction)
Expr: belief
Evoc: shared belief
>OBL
CA Dep: Sender(S): have belief
Receiver(U): evaluate and react
Cont dep:
IA Man: {FB: CPU}
TM: Accept turn

Activity Status: Closing
<Neg eval of closing attempt
>Evaluate and terminate or continue

System tasks: (ST1), [ST2], (ST3), [ST4], (ST5), (ST6)

User tasks: [UT1], [UT2], [UT3], [UT4], [UT5], [UT6], UT7 = type "quit" on a separate line

U10: quit

Mood:	Indic:V alt. Imp:V	
<OBL		
	IA Man:	{FB: CPU
		TM: Accept turn}
		Seq: closing
	CA Dep:	Accept and follow instruction
		Cont dep: effect of typing "quit" given in S9,
		(and screen message ?)
<>OPT		
	OC Man:	
	CA:	Request (instruction)
		Expr: belief, wish to stop alt. desire to stop
		Evoc: evaluate and satisfy wish alt. follow instruction
>OBL		
	CA Dep:	Sender(U): have belief, have wish to stop and motive for wish
		Receiver(S): satisfy wish*
		Cont dep: responsive action should be the object of wish
		(i.e. stop interaction)
	IA Man:	{FB: CPU}
		TM: terminate
		seq: closing of interaction
Activity Status:		Closing
		<Acceptance of closing instruction
		>Evaluate and terminate
System tasks:		(ST1), [ST2], (ST3), [ST4], (ST5), (ST6)
User tasks:		[UT1], [UT2], [UT3], [UT4], [UT5], [UT6], [UT7]

*) The systems obligation and/or right to evaluate the request to quit is diminished by the asymmetric relation between the user and system roles, in this type of dialogue.

6.6 On the Possibilities of Formalizing the Regularities and Functional Dependencies Found in the Dialogues

6.6.1 What should be formalized?

It is not necessarily desirable to formalize the Communicative Activity Theory (CAT), at least not as a whole, since it is not the case that the system, as opposed to its designers, should use this theory. The system should behave in a certain way (helpful, robust, cooperative etc), but not necessarily by having access to, and reason from, an explicit theory of helpful, robust, cooperative etc. behaviour.

A similar remark can be made concerning some of the factors which, according to CAT, determines certain features of the dialogues. Thus that the system's purpose is to help the user to find information that the system has access to in some database is a global factor that is constant and therefore not one that necessarily must be explicitly stored in the system. The balance between what should be explicitly stored and used by means of inference processes, and what should be implemented as procedural features of the system is a difficult design issue.

Roughly the same goes for Gricean type maxims. Whether these should be formalized and explicitly stored in full generality, or if one should settle for the "relevant special cases" depends on how general a system one is aiming for. To take an example; how should the system treat the opening

U1: "I need a hire car for tomorrow afternoon"?

If uttered to an employee of a car hire company or to a psychiatrist or to a YP-information system, it would, in all probability, require quite different responses. A fully general, pragmatically competent, system should of course be able to decide on an appropriate response in these (and inumerably many other) cases. It could then look in some internal storage to find that it should enact a YP-information service system, and at some other place that it should be helpful etc. A less general, but still pragmatically competent, system could be designed to "automatically" exploit constraints valid in all its intended applications.

Exactly what level of generality should be aimed for has, of course, to be decided sooner or later (preferably sooner), and then it will be easier to tell precisely what aspects of conversation rules, global constraints etc that will have to be explicitly coded. Still, it is important to have a general theory, amongst other things in order to know what the special cases are to be special cases of, and to be able to show that they really are special cases.

6.6.2 Rules of Communicative Analysis for the Analysed Dialogues

In this section we will discuss the possibilties of formulating rules that cover the communicative activity occurring in factual information seeking dialogues of the analyzed type. We will concentrate on rules covering the activity from the systems point of view, and we will also here exploit the fact that we are dealing with a rather specialiced type of dialogue, in which one participant (the system) has the purpose of helping the other participant to obtain information from the Yellow Pages data base (YPDB). This means, inter alia, that the systems "interests" are subordinated to those of the user, and, thus, that the systems initiatives are restricted to those motivated by its purpose to serve the user's (expressed or infered) interests. The resulting asymmetry in rights and obligations of the participants is an important source of simplifying constraints in comparison to those that might hold in an arbitrary type of dialogue.

The perspective chosen in this section is the following: the cooperative activity is carried out by means of a series (or, more accurately, a partially time-ordered set of) actions performed by the system and user. These actions gradually change, among other things, the informational and motivational states of the participants. The general goal of the system is, if possible, to provide the user with information (from the YPDB) that satisfies his/her desires for knowledge. In the paper we have so far concentrated on those phenomena of the activity that concern its communicative aspects. But besides overt communicative expressions, we will in this section also consider other types of actions, especially those performed by the system, like data base searches, acts of interpretation which are implied by or required for the use of communicative expressions.

A few words on activity status and task structure. These notions belong to different dimensions of analysis. Activity status pertains to the activity (in the case at hand a species of factual information seeking dialogue between a "client" and a "server") as such, whereas the task structure pertains to the participants in the activity. Especially the task structure is rather superficially described in the manuscript, and only partially coded in the analysis. So let's say something more. First, a task structure consists of more than a (list of the) tasks for the participants, namely relations between and properties of these. Such relations are, e.g., "replaces", "overrides", "interrupts", "is a sub-task of", "has precedence over" and so on, and some properties are "is completed", "is in progress", "is temporarily stacked" etc. (Such properties and relations are not recorded in the presented analysis.) The idea is, of course, that the verbal exchange between the participants is partly aimed at regulating, controlling and reporting on the task structure and the work done on the tasks. (In cooperative activities a task of one participant may well be a sub-task of a task of another participant. In our type of activity, the ultimate control over the general task structure lies, of course, in the hands of the user.) The carrying out of tasks is however only partly done on the dialogue level. (To complete its tasks the system has, e.g., to make database searches etc.) The

point of including at least some aspects of task structures in analyses is to make explicit what effects certain features of the dialogue has on some important nonverbal aspects of the activity and how these, conversely, influence aspects of the dialogue. Thus all aspects of a dialogue can't be captured by a description of the verbal intercourse alone. Not all regularities in a set of dialogues (corpus) can, consequently, be be captured in a dialogue grammar dealing with only verbal aspects of the dialogues.

A "typical sequence of actions" performed by the system in a phase of a dialogue might, at a rather coarse level of description, look something like this:

1. receive a string from the user
2. interpret the string as a "message"
3. if the message is a request for information about some matter
   then: store the information that the user wants information about the particular matter
   transform the message to a YPDB search query
   search the YPDB for the desired information
   if the information is found
   then: store this information internally
   evaluate whether the info matches the users wants
   evaluate whether the info is suitable to present to user
   if so
   then: transform it into a NL answer
   send the answer to user
   else: (e.g. if there are too many YPDB-hits)
   decide on a suitable request for specification
   transform request into NL expression
   send the expression to the user and "stack" the present task
   else: construct an NL statement with the content that no information on the matter is available
   send statement to user
   else if the message is a statement about the users wants or desires...

and so on.

Obviously the above example is both grossly oversimplified and incomplete, but it points to a method for pursuing the goal of formulating action rules (including rules pertaining to a "dialogue grammar", etc.) and deciding on what aspects of the dialogue history and the informational and motivational states of system and user that have to be represented.

As a matter of fact, we believe that all functional aspects of utterances covered in this paper have to be represented if the system is to be able to pursue a reasonably natural dialogue with a user. Thus actions like "interpret string" and "transform into

NL response" have to be broken down into much more detailed action sequences, and given explicit preconditions and result requirements.

6.6.3 Production systems

Production systems, or rewrite systems, are computational devices, well known from the theory of automata, mathematical linguistics, and recursion theory. Their mathematical properties have been extensively studied and are in many respects well known. Their computational power is, e.g., that of Turing machines (cf Post [19]). Production systems have also been used in modelling of psychological and cognitive systems (cf. Newell and Simon [18]; Buchanan and Feigenbaum [8]), and they are nowadays a standard tool in expert systems technology (cf McDermott and Forgy [17]; Brownston et al [7]; Jackson [14]). A traditional production system has three components: a set of productions, a work-space and a control system. The productions consist of condition-action-pairs. The workspace contains symbolic expressions which can satisfy (or fail to satisfy) conditions in productions. If a condition in a production is satisfied by an expression in the workspace, the control system may activate the production, which means that the action specified in the action part of the production is carried out. Usually this means that the content of the workspace is altered, possibly to the effect that the conditions of other productions become satisfied (or cease to be satisfied). The most important function of the controll system is to handle so called conflict resolution in cases where two or more productions have satisfied conditions, and where the respective actions are incompatible, or where the carrying out of the action of one of the productions has the effect that the condition of another no longer is satisfied.

Production systems form a very flexible tool for modelling monotonic as well as non-monotonic inferences, deterministic as well as indeterministic dynamic systems, sequential as well as parallel calculations, and so on. They are also fairly easy to implement in, e.g., logic programming languages, which makes it possible to test rule systems at an early stage of construction. For many types of production systems, there exist well developed theories. This is the case for systems with monotonically growing workspace content, systems with strong restrictions on the expressions used in the workspace (and the productions), and systems which use fixed linear precedence orderings of productions as conflict resolution strategy. Also for other types of system there are theories and knowledge gained from numerous applications.

Some of the popularity of production systems as a theoretical tool in many areas no doubt derives from the rather intuitive character of the production rule form, which makes it possible to "understand" each rule in isolation, as it were. This, rather than the universality of the format, is the main reason for choosing a production rule format for early stage specification and development of complex systems.

We thus propose that, at least for a start, rules are given in the form of production rules,

i.e., as pairs of conditions and actions (to be obligated or permitted if the condition is satisfied). The conditions concern primarily states of the dynamic records, like the user model, the task structure, the activity status, the discourse history (and especially the preceding utterance) and the actions pertain to the behaviour of the system. (It would be making our task too simple to already at this stage try to externally "regulate" as opposed to "predict" the behaviour of the user.)

Some examples of rather generic contents of rules might be:

1. if the user has expressed a desire for information about a particular matter, then the system should construct a task structure containing at least the tasks to search the YPDB for that information, and to report to the user on the result of this search in a way that respects all the obligating functions of the utterance by which the user expressed the desire.

2. if the user has expressed a belief with the content that the user has a desire for information about a particular matter, then the system should conclude that the user has thereby also expressed that desire, and

3. if the user has expressed a desire for information on whether the system has information about a particular matter, then the system should conclude that the user has thereby expressed a desire for information about that particular matter.

Together these rules should make user utterances of the types:

> "Which Indian restaurants are there in Gothenburg?",
> "I want to know which Indian restaurants there are in Gothenburg.", and
> "Do you know what Indian restaurants there are in Gothenburg?"

on the whole functionally equivalent with regard to the task of finding and reporting on YPDB information. However, a response beginning like "Yes, there are ..." would be odd as a reply to the first (and maybe the second) type of user utterance, and one beginning "Certainly, there are ..." would be even more odd as a response to the first type of user utterance, which shows that communicative functions over and above those pertaining to communicative act and content are important for, e.g., response planning. In this case, we see that the specified mood of the information request is important. A "yes" or a "no" (and equivalents like "certainly") are expected (or weakly obligated) after a Y/N-Q, tolerated after a Declarative but odd after a Wh-Q.

6.7 Relation of the Proposed Analysis to a PLUS-like System

6.7.1 The Natural Language Engine

In terms of the analyses we have presented above and the architecture so far given for the PLUS system, the NL engine could be supposed to deliver a parsed analysis of a communicative expression which had as prominent features

1. A structure separating OCM, IACM and MM parts of an expression (cf. Allwood, Nivre, Ahlsén [5])
2. A mode (mainly for the MM parts)
3. A syncategorematic structure, especially a predication structure (mainly for the MM parts)

Scope and reference resolution as well as determination of categorematic predicates are deliberately left out. This kind of structure is then supposed to serve as an input to the cognitive analyzer and as an output from the response planner.

6.7.2 The Cognitive Analyzer

The cognitive analyzer takes as input a string of the kind described in 6.7.1 and, in the terminology introduced above, it produces the following six types of analysis (we will mainly be concerned with the perspective of the system):

1. An analysis of OCM and IACM parts. Roughly speaking, in the dialogues at hand, unless there are explicit OCM and IACM parts, it can be assumed that there is no OCM and that IACM functions are implicitly positive, i.e., each contribution except the first and last, implies obligated and obligating "turn accept".
2. An analysis of the obligated communicative act dependent functions of the (user) expression. These can probably only be got through a matching procedure between (i) the obligating CA dependent (including content dependent) requirements of the preceding (system) contribution and (ii) the presently relevant obligations (on the user) arising from taking part in rational, ethical cooperative interaction, having a particular role in a particular subactivity of an activity, and (iii) the features of the parsed utterances which are compatible with these obligations.

3. An analysis of the optional functions of the expression, in particular what (optional) communicative acts it is used for. The mood of the expression, in combination with a consideration of requirements derived from a consideration of MRA (motivated rational action), role, activity and subactivity are probably very relevant for this kind of analysis. The analysis of communicative acts should also include an analysis of the expressed attitudes and evocative intentions of the communicative act. This analysis can probably most profitably be done once the communicative acts have been determined.

4. An analysis of the epistemic and attitudinal commitments made in using the expression. These commitments are, in fact, the expression relevant expressive and evocative sender obligations discussed above, and are derived, more or less directly, from the attitudes given by 3, above, in combination with a consideration of the user's sender obligations (sincerity and reasonableness commitments).

5. An analysis of the obligating functions of the user's contribution. These can, together with considerations of the system's obligations as an ethically motivated rational cooperating agent, in a particular subactivity of the activity provide a major part of the input to the goal formulator.

6. An analysis of the presently relevant activity status and task structure. The results of such an analysis, possibly in combination with an analysis of topic structure and content dependencies, could then provide a major part of what has been called the "discourse model".

6.7.3 The Goal Formulator

Given an input to the goal formulator as described in 6.7.1 above, the output of the goal formulator could be expected to specify the following functions of the planned expression (functional goal plan):

1. Obligated functions
   (a) IA management
      i. FB
      ii. TM
      iii. Seq
   (b) CA dependent
      i. Report on evaluation of preceding CA
      ii. Carrying out functions
2. Optional functions

(a) Expressive and evocative

(b) Communicative acts (compatible with 2a, 1a and 1b above, and other determining constraints)

3. Obligating functions

   (a) IA management

      i. FB

      ii. TM

      iii. Seq

   (b) CA dependent (compatible with 2a, 2b and other determining constraints)

4. Sender commitments

   (a) Sincerity

   (b) Rational motivation (reasonableness)

6.7.4 The Response Planner

In the response planner, the functions given by the goal planner have to be explicitly or implicitly, sequentially or simultaneously realized in a plan for a communicative expression specified down to roughly the following level:

1. OCM—IA—MM structure
2. Mood specification of mainly MM
3. A determinate syncategorematic structure, especially a predication structure of mainly MM
4. Determinate scope and coreference relations
5. Determinate aspects of categorematic predicates

Bibliography

[1] Allwood, J. (1976) Linguistic Communication as Action and Cooperation. *Gothenburg Monographs in Linguistics* 2. University of Göteborg: Department of Linguistics.

[2] Allwood, J. (1977) A critical look at speech act theory. In Dahl Ö. (ed.) *Logic, Pragmatics and Grammar.* University of Göteborg: Department of Linguistics.

[3] Allwood, J. (1984) On Relevance in Spoken Interaction. In Kjellmer, G. and Bäckman, S. (eds.) *Papers on Language and Literature. Gothenburg Studies in English* 60. University of Göteborg: Department of English.

[4] Allwood, J. (1988) Om det svenska systemet för språklig återkoppling. In *Svenskans beskrivning* 16. University of Linköping: Department of Communication Studies.

[5] Allwood, J., Nivre, J. and Ahlsén, E. (1990) Speech management—on the non-written life of speech. *Nordic Journal of Linguistics* 13.

[6] Allwood, J., Nivre, J. and Ahlsén, E. (forthcoming) On the Semantics and Pragmatics of Lingustic Feedback. *Journal of Semantics.*

[7] Brownston, Farrell, Kant and Martin (1985) *Programming Expert Systems in OPS5.* Reading, Mass.: Addison-Wesley.

[8] Buchanan and Feigenbaum (1978) DENDRAL and Meta-DENDRAL: Their Application Dimension. *Artificial Intelligence 11.*

[9] Dahlbäck, N. and Jönsson, A. (1986) A System for Studying Human-Computer Dialogues in Natural Language. Research Report LiTH-IDA-R-86-42, NLPLAB, IDA Linköping.

[10] Dahlbäck, N. and Jönsson, A. (1989) Empirical Studies of Discourse Representations for Natural Language Interfaces. *Proc. of the 4th Conf. of the European Chapter of the Association for Computational Linguistics*, Manchester, April, 1989.

[11] Fraser, N.M. and Gilbert, G.N. (1991) Simulating speech systems. *Computer Speech and Language* 5, 81-99.

[12] Grice, H. P. (1975) Logic and Conversation. In Cole and Morgan (eds) *Syntax and Semantics* III. New York: Academic Press.

[13] Horn, L. R. (1973) Greek Grice. *Proceedings of the Ninth Regional Meeting of the Chicago Linguistic Society*, 205–214.

[14] Jackson, P. (1986) *Introduction to Expert Systems.* Reading, Mass.: Addison-Wesley.

[15] Luzzati, D. (1989). Recherches sur le dialogue homme-machine: modèles linguistiques et traitements automatiques. Thèse de doctorat d'état ès lettres, Université de Paris III.

[16] Malinowski (1923) The Problem of Meaning in Primitive Languages. In Ogden, C. K. and Richards, I. A. (1923) *The Meaning of Meaning.* New York: Harcourt, Brace and World.

[17] McDermott and Forgy (1978) Productions Systems Conflict Resolution Strategies. In Waterman and Hays-Roth (1978) *Pattern Directed Inference Systems.* New York: Academic Press.

[18] Newell and Simon (1972) *Human Problem Solving.* Engelwood Cliffs, NJ: Prentice-Hall.

[19] Post, E. L. (1936) Finite Combinatory Processes—Formulation I. *The Journal of Symbolic Logic 1.*

[20] Richards, M.A. and Underwood, K.M. (1984). How should people and computers speak to each other? *Interact '84*, 33-36.

[21] Sacks, H., Schegloff, E. and Jefferson, G. (1974) A simplest systematics for the organization of turn-taking for conversation. *Language* 50, 696–735.

[22] Sacks, H. and Schegloff, E. (1973) Opening up closings. *Semiotica* 8, 289–327.

[23] Sperber, D. and Wilson, D. (1986) *Relevance: Communication and Cognition.* Oxford: Blackwell.

[24] Vilnat, A. (1984). L'élaboration d'interventions pertinentes dans une conversation homme-machine. Thèse de 3e cycle, Université de Paris VI.

Appendix A

Instructions and Scenarios

Below we reproduce the instructions and scenarios given to the subjects before the WOZ dialogues and human-human dialogues, respectively. The scenarios contain variables (such as <UNIVERSITY> and <TOWN>) which were given appropriate values locally by each team.

A.1 Instructions to Subjects: WOZ

We have developed a prototype for a computerised system for Yellow pages information service. This new service is an advance on existing services, because it can interpret your request for information expressed in ordinary <LANGUAGE>

We now want to test if the system can manage a natural dialogue.

The computer is meant to be equally capable of understanding you whether you are careful to express what you want in a complex question, or whether you let it work out how best to provide you with the information you need as elaborated in the ongoing dialogue. In fact, you are encouraged to try to use your language in as natural a way as possible, as if you were speaking to another person. In this way, once the dialogue has started, you should not find yourself having to resort to long, complete sentences, for example.

Below you will find three tasks which we want you to solve using the system.

You should not expect the information system to be able to tell you everything you need to know. However, you may find that the system has som background information about locations and which companies have several branches, etc., so do try to get as much information as you can before spending time and money on the phone. Also, bear in mind that you will only have one session with the system, so it will be necessary to compile a list of all the plausible alternatives whilst using it.

A.1.1 Car Hire

You are a postgraduate student at <UNIVERSITY> university, and you live in the village of <VILLAGE>, outside <TOWN>. (Look at the map supplied and familiarise yourself with the area.)

You are booked on a flight from <CITY> airport tomorrow evening for your holidays, departing at 7.00 pm. Just today, you have been invited for a job interview also tomorrow, in the afternoon, with a small company based in a village <DISTANCE> north of <TOWN>.

You desperately want this job, and want to attend the interview, but you can't change your holiday travel arrangements, so instead of going to the airport by bus, you have decided to hire a car to go first to the interview and then to the airport. You will get some expenses from the company, but probably not the whole cost of the car hire, so it is important to be able to check out a few companies to minimise the cost to you. Also, if it's not too expensive, you would like to pick up the car close to home or have it delivered, but leave it at the airport.

Now, we want you to try to find this information from our prototype of a new interactive Yellow Pages database, using the computer terminal.

A.1.2 Restaurant

You live in <CITY>. A dear friend of yours, who lives in the countryside, far away from the plesaures of a big city, pays you a short visit. Unfortunately, you are not much of a cook. Therefore you decide to take him to a restaurant. As you know your friend has a taste for hot and spicy food, you decide to take him to a restaurant serving this kind of food. Find a suitable one, not too far away from your home, using the computerized information service.

Find a suitable restaurant, not too far away from your home, using the information service.

A.1.3 Personal Insurance

You and your partner have just decided to settle down permanently and raise a family. You have already found a suitable flat, but you need to buy some new furniture and household equipment. And in order to be able to do this you have borrowed some money.

You and your partner realize that in your new life situation with financial undertakings and responsibilities for each other (and maybe children before long) you ought to look

over your insurance situation. There is the home and its equipment to protect, there is the financial security of the family to think of, and by the way your partner has that old Renault that is probably not adequately insured.

So what you need right now is to contact some insurance companies to see what alternatives there are, what the cost would be for various alternatives, and perhaps to get some useful pieces of advice.

Use the computerised information service to find out how to get the information you need.

A.2 Instructions to Subjects: Human-Human

We are developing a computerised system for Yellow pages information service. This new service will be an advance on existing services (like the French Minitel), because it will be able to interpret your request for information expressed in ordinary <LANGUAGE>

We want to model the computer-generated responses as closely as possible on human dialogue capabilities, and so through the computer terminal, you are actually connected to another person, who has access to the Yellow pages database for the area in question, and will interpret your question intelligently, and produce responses consistent with the information we intend to be in the computer's "knowledge base".

Below you will find three tasks which we want you to solve using the simulated system. You are encouraged to try to use your language in as natural a way as possible. Ask your questions, make your comments and so on, reflecting the way you would like to interact with a system like this.

You should not expect the information service to be able to tell you everything you need to know. However, you may find that the service has som background information about locations and which companies have several branches, etc., so do try to get as much information as you can before spending time and money on the phone. Also, bear in mind that you will only have one session with the service, so it will be necessary to compile a list of all the plausible alternatives whilst using it.

A.2.1 Car Hire

You are a postgraduate student at <UNIVERSITY> university, and you live in the village of <VILLAGE>, outside <TOWN>. (Look at the map supplied and familiarise yourself with the area.)

You are booked on a flight from <CITY> airport tomorrow evening for your holidays,

departing at 7.00 pm. Just today, you have been invited for a job interview also tomorrow, in the afternoon, with a small company based in a village <DISTANCE> north of <TOWN>.

You desperately want this job, and want to attend the interview, but you can't change your holiday travel arrangements, so instead of going to the airport by bus, you have decided to hire a car to go first to the interview and then to the airport. You will get some expenses from the company, but probably not the whole cost of the car hire, so it is important to be able to check out a few companies to minimise the cost to you. Also, if it's not too expensive, you would like to pick up the car close to home or have it delivered, but leave it at the airport.

Now, we want you to try to get this information from our simulated system, using the computer terminal.

A.2.2 Restaurant

You live in <CITY>. A dear friend of yours, who lives in the countryside, far away from the plesaures of a big city, pays you a short visit. Unfortunately, you are not much of a cock. Therefore you decide to take him to a restaurant. As you know your friend has a taste for hot and spicy food, you decide to take him to a restaurant serving this kind of food.

Find a suitable restaurant, not too far away from your home, using the information service.

A.2.3 Personal Insurance

You and your partner have just decided to settle down permanently and raise a family. You have already found a suitable flat, but you need to buy some new furniture and household equipment. And in order to be able to do this you have borrowed some money.

You and your partner realize that in your new life situation with financial undertakings and responsibilities for each other (and maybe children before long) you ought to look over your insurance situation. There is the home and its equipment to protect, there is the financial security of the family to think of, and by the way your partner has that old Renault that is probably not adequately insured.

So what you need right now is to contact some insurance companies to see what alternatives there are, what the cost would be for various alternatives, and perhaps to get some useful pieces of advice.

Use the simulated information service to find out how to get the information you need.

Appendix B

Questionnaire

All subjects participating in the collection of dialogues were asked to supply information both about their sociobiographic characteristics and their evaluation of the system/experiment. The questionnaires used to elicit this information differed slightly between the teams. Below we reproduce the questionnaire used by the English team.

1. Age?
2. Sex?
3. Occupation?
4. Your experience (none, some, a lot, an expert) on
   (a) typing?
   (b) computing?
   (c) yellow pages?
5. Give your opinion of the responses given by the system. Were they natural and human-like or did they sound artificial? Were they logical and informative? Please explain why/why not.
6. How would you evaluate your own performance? Was it different from having a face-to-face conversation with a person? Please explain why/why not.
7. Give any comments on the dialogue in general.

Appendix C

Dialogue Excerpts

Below we provide three illustrative excerpts from the PLUS dialogue corpus. The first is an English WOZ dialogue on the restaurant scenario, the second is an English translation of a Swedish WOZ dialogue exemplifying the personal insurance scenario, and the third is the beginning of a French Human-Human car hire dialogue, translated into English.

For readability reasons, we have not included time stamps in the excerpts below. We have also omitted all information about characters that were deleted by the user before sending the message to the system.

C.1 English WOZ Restaurant Dialogue

U: can you give me a list of indian and chinese restaurants in the
Fallowfield ,Withington, and Rusholme areas please
S: Please wait ...
Indian:
The Indian Cottage, Rusholme 224-0446
Sanam Restaurant, Rusholme 224-8570
Shezan, Rusholme 224-0895
Chinese:
Happy Garden, Didsbury 445 6979
U: are they all licenced to sell alcohol
S: The YP cannot provide this information
U: which ones are open after midnight
S: The YP cannot be specific; however, in general such restaurants are open until 2 am
U: do any of them have a special student night
S: The YP cannot provide this information; please contact the restaurants concerned

U: are there any italian restaurants in fallowfield
S: Please wait ...
Conca D'Orco
220 Wilmslow Road
224 2373
Mamma Rosa
317 Wilmslow Road
257 2971
U: what about french
S: Please wait ...
Still in Fallowfield?
U: yes, or in the area
S: The Mock Turtle
256 Wilmslow Road
Fallowfield
224 2340
The Lime Tree Cafe
9-11 Wilmslow Road
Rusholme
225 7108
U: can you give me a list of all the adresses that you have found for me
S: Indian:
The Indian Cottage, Rusholme 224-0446
Sanam Restaurant, Rusholme 224-8570
Shezan, Rusholme 224-0895
Chinese:
Happy Garden, Didsbury 445 6979
Conca D'Orco
220 Wilmslow Road
224 2373
Mamma Rosa
317 Wilmslow Road
257 2971
The Mock Turtle
256 Wilmslow Road
Fallowfield
224 2340
The Lime Tree Cafe
9-11 Wilmslow Road
U: is there another chinese restaurant in town
S: There are many Chinese restaurants in town; please be more specific
U: in the town center
S: Would you like a list?

U: yes please
S: Please wait ...
Hong Kong, 236 7003
Hopewell City, 236 0091
Kai's, 236 2041
Kwok Man, 236 9298
Golden Rice Bowl, 832 9033
Little Yang Sing, 228 7722
MR Kuk's, 236 3539
Pearl City, 236 2574
Win Wah, 236 2214
Woo Sang, 236 4170
Yang Sing, 236 2200
Quan Ju De, 236 5236

C.2 Swedish WOZ Insurance Dialogue

S: Welcome to the Yellow Pages Information Service!
What can we do for you?
U: insurance companies?
S: Just a moment, please ...
What type of insurance are you interested in?
U: Everything! Car, home, life etc
S: Just a moment, please ...
Would you prefer a large or a small company?
U: small
S: Just a moment, please ...
The following small companies offer home insurances:
- Holmia
- Försäkringsbolaget Royal

U: I think I prefer a large company after all
S: Just a moment, please ...
The following large companies offer home insurances:
- Ansvar
- Folksam
- Wasa Försäkring
- Länsförsäkringar
- Skandia Försäkrings AB
- Trygg-Hansa

U: which offer car insurances as well?

S: All the listed companies offer car insurances.
U: may i have phone no. to all
S: Just a moment, please ...
Skandia has telephone number 816000.
Folksam has telephone number 600500.
Trygg-Hansa has telephone number 819000.
Wasa has telephone number 614500.
Länsforsäkringar has telephone number 638000.
Ansvar has telephone number 177900.
U: Addresses as well
S: Just a moment, please ...
Ansvar is located at N. Alleg. 6.
Folksam is located at Storg. 53.
Wasa is located at Stampgatan 14.
Länsforsäkringar is located at Ö. Larmgatan 16.
Skandia is located at Sten Stureg. 12.
Trygg-Hansa is located at Vasag. 45.
U: thanks
S: Thank you and goodbye from the Yellow Pages Information Service.
Welcome back!

C.3 French Human-Human Car Hire Dialogue

S: Welcome to the yellow pages access system.
Hit the <esc> key to send your text to the system.
Hello, please formulate your query ...
U: I want to hire a car of category A for 4 days I want to pick it up
at Saint Maur and to return it at the Roissy airport
S: Sorry, I haven't any information on the different car categories proposed by
the companies. Do you want a chauffeur-driven or a self-drive car?
U: Chauffeur-driven
S: Please wait ...
S: There are several towns named "Saint Maur" in France.
Please indicate the one you are referring to in the following list:
1. Saint Maur (Cher)
2. Saint Maur (Gers)
3. Saint Maur (Indre)
4. Saint Maur (Jura)
5. Saint Maur (Oise)
6. Bac Saint Maur (Pas de Calais)

7. Saint Maur des Bois (Manche)
8. Saint Maur des Fosses (Val de Marne)

U: 8
S: Sorry, you can't find a chauffeur-driven car hiring company in Saint Maur des Fosses. Do you want to try a search in neighbouring towns?
U: No I will take self-drive
S: Still in Saint Maur des Fosses?
U: Yes
S: Please wait ...
S: Two companies (namely "Avis" and "Hertz") have both branches in Saint Maur des Fosses and Roissy. Do you want to know their whereabouts?
U: Yes
S: Here they are:

Avis:	95 Av Foch	Saint Maur des Fosses	(1) 48 89 57 00
	Terminal 2	Charles de Gaulle airport	(1) 48 62 59 59
Hertz:	3 Rue Albert de Mun	Saint Maur des Fosses	(1) 42 83 10 14
	Terminal 1	Charles de Gaulle airport	(1) 48 62 29 00

(The dialogue continues.)