Corpus–Based Information Presentation for a Spoken Public Transport Information System*

M.M.M. Rats and **R.J. van Vark** and **J.P.M. de Vreught**
Knowledge Based Systems
Technical Computer Science
Delft University of Technology
alparon@kgs.twi.tudelft.nl

Abstract

The Alparon project aims to improve VIOS, Openbaar Vervoer Reisinformatie's (OVR) automated speech processing system for public transport information, by using a corpus–based approach. The shortcomings of the current system have been investigated, and a study is made of how dialogues in the OVR domain usually occur between a human operator and a client. While centering our attention on the presentation of information by the Automated Speech Processing (ASP) system, we describe the implications of this corpus–based approach on the implementation of our prototype.

1 Introduction

The Alparon research group in Delft aims to improve Automated Speech Processing (ASP) systems for information retrieval and information storing dialogues. The current focus is on dialogue management for a research project of Openbaar Vervoer Reisinformatie (OVR) (Vark et al., 1996). OVR provides information about Dutch public transport systems, ranging from local bus services to long distance trains. They are developing an ASP system, called VIOS, to automate part of its dialogues held at its call centres.

Although the second version of VIOS is implemented, it is far from perfect. After conducting user–appreciation studies it was clear that the next version of VIOS should act more human like.

This is the reason why we will take the human operator as the source of inspiration for improvement. The proposals that we will put foreward in this paper will be based on a study of about 100 human–human OVR-dialogues selected from a corpus of 5205 telephone conversations recorded at the call centers of OVR. The pilot corpus consists of dialogues that concern the exchange of train information only.

OVR dialogues proceed in a specific way: first greetings are exchanged, then the client formulates his query, next the operator gives the desired information, and finally both parties say goodbye to each other. In this paper, we will concentrate on the improving the presentation of the travel scheme.

We can roughly distinguish three different situations: the standard situation where everything runs smoothly, the situation where there is a repair operation by either the client or the operator, and the situation where a topic shift occurs. In our prototype, the dialogue manager and the text generator will collaborate to handle these situations. Since the dialogue manager has knowledge concerning known and new information, it can instruct the text generator to present the text in a natural way.

In the next section, we will discuss related work. Then, we will describe in detail how information is presented in VIOS and in the OVR corpus. Next, we will present the strategies an ASP system has to follow if such a system is to present information in a sensible manner. In the last section, we will describe future research.

2 Related Work

VIOS is a Dutch version of the train timetable information system developed by Philips Aachen for the Deutsche Bundesbahn (Aust and Oerder, 1995; Aust et al., 1995). It was developed within the European LE–MLAP project MAIS (Strik et al., 1996), the aim of which was to make a Dutch and French version of the Philips system.

A related European research project was the LE–MLAP project RAILTEL (Bennacef et al., 1996) which strived to develop prototype telephone services for access to train travel information. Prototype systems are being developed for France

*This work is funded by OVR and Senter.

Figure 1: A VIOS presentation of a train connection

V: Ik heb de volgende verbinding gevonden. Vertrek vanuit Delft om twintig uur tweeënveertig, aankomst in Rotterdam CS om twintig uur zesenvijftig, daar overstappen naar Utrecht CS, vertrek om eenentwintig uur zeven, aankomst in Utrecht CS om eenentwintig uur drieënveertig. Wilt u dat ik de verbinding nog eens herhaal?	V: I have found the following connection. Departure from Delft at twenty hours forty-two, arrival at Rotterdam CS at twenty hours fifty-six, there change to Utrecht CS, departure at twenty-one hours seven, arrival in Utrecht CS at twenty-one hours forty-three. Do you want me to repeat the connection again?

(LIMSI), Italy (CSELT), and the U.K. (CCIR).

The successor of these projects is the European project ARISE, which aims to improve the previous versions of the different systems. OVR is a partner in both MAIS and ARISE.

Besides Alparon several other universities and companies in the Netherlands are working to improve VIOS. Most of them are working in the OVIS project (Boves et al., 1995), which aims to develop the next version of VIOS for the Dutch context. In the Alparon project, we are allowed to try and test out new ideas beyond this next version.

Our work is also inspired by other related research projects. One of them is the TRAINS project (Allen et al., 1995), that tries to build an interactive planning assistant that helps a user to construct and monitor plans concerning a railroad freight system. Another related project is described in (Carenini et al., 1994). They have developed a system that uses text planning and user modelling techniques to generate natural language descriptions of migraine, its symptoms, triggering factors and prescriptions. The system is capable of handling follow–up questions requesting further information, and generating responses in context of previously supplied information. In (Cawsey, 1993) the EDGE system is described, which is able to plan explanations which allow interactions with the user.

For text generation, we found useful ideas in (Prevost, 1996), where an architecture is described for the generation of spoken text with contextually appropriate intonation. The schema approach (McKeown, 1985), and the systems DYD (Deemter and Odijk, 1995), and GoalGetter (Klabbers et al., 1997) were inspiring too.

3 Information presentation in VIOS

In VIOS, the travel plan is presented using templates filled with specific stations and times. The station names and times have a disambiguated form, always resulting in a full and uniquely identifying description. The templates and the words for stations and times are prerecorded and their accoustic representations are concatenated to form complete sentences. Together these sentences form a monologue. No interruptions by the caller are possible (see figure 1).

The architecture of VIOS is such that the dialogue management component and the text generating component work in sequential order. When the travel scheme has been determined, the dialogue manager sends the entire scheme to the text generator. This component communicates the plan to the caller. During the presentation process, the system stops listening. So, no feedback from the user to the system is possible.

Appreciation studies have shown that the VIOS presentation of information is far from ideal for the following reasons (Blasband, 1997; Manintveld and Rothkrantz, 1997; Peters and Kock, 1997):

- All information is given at once. It would be more convenient for the caller if the system would provide the information in smaller chunks.

- The caller is not able to interrupt when he does not comprehend the given information.

- The use of full names for stations is very unnatural and confusing, especially when the caller has used other descriptions to introduce them.

- The system makes the caller feel hunted in processing and copying down the information since it speaks too fast.

- The information is spoken in an unnatural way. The concatenated phrases exhibit differences in speech tempo and loudness. Since they miss a natural intonation pattern, they are awkward to understand.

In this paper, we will only work towards a solution for the first two problems. The next three items will be left for further research. We will take the human–human dialogues as an example. This choice is dictated by our comprehensive appreciation study, where 500 respondents were asked to call different Wizard of Oz simulation systems, VIOS, and a

Figure 2: An information phase of an OVR dialogue

C: ik wilde graag weten als ik morgen om: morgenochtend om negen uur [u:h] op de Lelylaan in Amsterdam moet zijn hoe laat moet ik vanuit Hilversum vertrekken dan? ⋮	C: I would like to know if I have to be tomorrow at: tomorrow morning at nine o'clock [u:h] at the Lelylaan in Amsterdam at what time do I have to leave from Hilversum? ⋮
I: dat wordt[u:h] acht uur negentien vertrekken uit Hilversum	I: that will be[u:h] depart from Hilversum at eight hours nineteen
C: acht uur ja	C: eight hours yes
I: en dan overstappen Amsterdam	I: and then change Amsterdam
C: [u:h] ja	C: [u:h] yes
I: vertrek Amsterdam acht vijfenvijftig	I: departure Amsterdam eight fifty-five
C: acht vijfenvijftig	C: eight fifty-five
I: komt u negen nul drie aan Amsterdam Lelylaan	I: will arrive at Amsterdam Lelylaan at nine o three
C: oké dus acht uur negentien vertrekken oké hartstikke bedankt	C: okay so depart at eight hours nineteen okay thank you very much

human operator. One of its most important conclusions was that callers appreciate the human operator over all kinds of automated systems.

4 Information presentation in human–human OVR dialogues

A study of a sample of 100 information dialogues out of a corpus of over 5000 dialogues shows that the presentation of a travel plan in a human–human dialogue involves more than just a monologue that presents the entire plan at once. The information presentation has a more interactive form. A human information service presents the travel plan in a stepwise way, generally giving at least one piece of new information with each turn. This presentation follows the temporal order of the different stages in the travel plan. After each step in the information presentation, the caller shows that he has processed the step by an acknowledgement. The dialogue fragment displayed in figure 2 shows an example of an information presentation in an OVR dialogue. We see that each step in the plan is acknowledged, before the next one is given.

This stepwise presentation and acceptance of the travel plan is one of the most important characteristics of the information phase of a naturally occurring OVR dialogue. In this way, the operator is able to communicate the information as clear as possible and the caller can relate the new information to information already known.

Table 1 shows the quantitative facts that underly our description. It shows the amount of utterances the information service applies in one turn. We see that 87% of the turns contain only one utterance and 10% contain two utterances. The maximum of utterances per turn is 5, which is in only 0.4% of the cases. Only in this case with five utterances, a whole travel plan is given at once. This confirms our view that in human–human OVR dialogues, the travel plan is given in steps.

Table 1: The amount of utterances in each turn of the information service

Amount of utterances	percentage
1	87%
2	10%
$\geq$3	3%

A closer look at the exact information transfer in an OVR dialogue reveals even more about the exact information structure of the individual utterances of the information service. The information transfer in an OVR dialogue consists of three phases:

1. a query phase,
2. a search phase, and
3. an information phase.

During the query phase the caller poses his query, and the information service tries to understand this query as clear as possible. In the search phase (often indicated by a silence since the operator is searching), the information service applies the database query, and chooses the right travel plan. In the information phase, the information service communicates the travel plan to the caller and the caller tries to get the plan as straight as possible.

In each of these phases, different information elements play a crucial role. During the query phase, these elements are the departure place, the arrival place, a global indication of the departure or arrival time, the day of travel, and if the caller wants a direct connection. The information service needs these information elements to compose an appropriate database query and to choose the most suitable travel plan. For the presentation of the travel plan, other information elements become important: the departure time, the arrival time, the places where to change, the directions of trains, the departure and arrival times at the places of change. The caller needs to know these information elements to carry out this plan.

In the information phase, the information service uses these information elements to compose her presentation. In general, she will use the elements that are already known from the query phase or from previous utterances within the information phase, as a point of attachment for presenting the unknown elements. The order of the steps in the travel plan are the guiding principle behind the order in which the elements are presented. The dialogue fragment in figure 2 illustrates this. In the query phase of this dialogue, the speakers have established the day, a global arrival time, the arrival place, and the departure place. The information service has found a proper travel plan and starts her presentation. In consecutive turns, she gives the departure time (new) at the departure place (given), then the place where to change (new), then the departure time (new) at the place of change (given in the previous utterance), then the arrival time (new) at the arrival place (given).

Tables 2, 3, and 4 show the quantitative facts that underly our description. Table 2 shows the amount of information elements for each utterance of the information service. We see that 69% of the utterances contain 2 information elements, that 18% contains only one element, and 12% contains 3 elements. This confirms our view that mainly two information elements per utterance are given.

Table 2: The amount of information elements in each utterance of the information service

Information elements	Percentage of utterances
1	18%
2	69%
3	12%
4	1%

Table 3 shows the given–new divisions in utterances with one, two, three, and four information elements respectively. We see that utterances with one element contain in most of the cases (68%) a new element. In the other cases (32%), they either refer to a crucial step in the information exchange, or give a alternative description of a previously introduced element. We see that most utterances with more than one information element contain at least one new element. The given-new distinctions in the table reflect the order in which the elements occur. One can observe that in utterances with two elements, the speakers prefer to mention the new element first, while in case of three elements, speakers prefer to mention a given element first.

Table 3: The given-new division in utterances with one, two, three, and four information elements respectively.

Information elements	Given-new division	Percentage
1	new	68%
	given	32%
2	given-new	28%
	given-given	4%
	new-given	46%
	new-new	22%
3	new-given-given	12%
	new-new-given	4%
	given-new-given	15%
	given-new-new	54%
	new-given-new	15%
4	new-new-new-new	100%

The table confirms our view that speakers tend to present at least one piece of new information per utterance. Usually, they will relate this new information with an entity introduced in the preceding context. This communicative rule was found in other corpora of spoken discourse as well (Chafe, 1987; Rats, 1996).

Table 4 shows which specific information elements generally serve as given and which serve as new information. We see that most departure and arrival times (97% in both cases) are new for the caller, while all departure and arrival places (100% in both cases) are known from the previous context. For the places of change and the directions, the result is roughly fifty–fifty.

The table confirms our observation that departure and arrival places are generally introduced during the query phase and serve as given information during the information phase. On the contrary, places of change and the directions of train are mostly introduced in the information phase, and become given information after introduction. The departure and

Table 4: The information status of the information elements in the information phase.

Information element	Given	New
Departure time	3%	97%
Arrival time	3%	97%
Departure place	100%	-
Arrival place	100%	-
Place of change	54%	46%
Direction	42%	58%

arrival times commonly serve as new information.

After the presentation and acceptance of a whole travel plan, a caller may ask for new travel plans. This happens in 26% of the dialogues. Most of these plans (84%) concern earlier or later connections. In 16% of the cases, the callers ask for another travel plan, for instance, for a connection from the station where the previous trip ended or another connection from the same departure place.

5 Repair in the presentation phase

Figure 2 shows an information phase where the caller has no problems in processing the presentation of the travel plan. Nevertheless, during communication problems may arise. In such cases, the caller will interrupt the presentation by starting a repair sequence to solve the problem:

- The caller will start a reconfirmation sequence, if he is not sure that he has heard the operator's utterances well and he wants the information service to repeat, to complete, or to confirm previously given information. A reconfirmation sequence consists of a reconfirm by the caller and an appropriate answer by the information service.

- The caller will apply a clarification sequence, if he wants extra information about the current plan. A clarification sequence consists of a wh–question of the caller and an appropriate answer by the information service.

- The caller will apply a checking sequence, if he wants to check extra information about the current plan that he suspects to be true. A checking sequence consists of a check by the caller and an appropriate answer by the information service.

- The caller will start a correcting sequence, if he notices that the information service gives inappropriate information. A correcting sequence consists of a correction and possibly a negative acknowledgement by the caller and an appropriate answer by the information service.

In case the caller does not notice problems himself, the information service may infer from the caller's responses that the caller did not process her utterances as intended. She can infer this from his acknowledgements, but also from his reconfirmations, checks, and wh–questions. In such cases, she will interrupt the information presentation by a correcting sequence.

Table 5 shows the frequency of the caller's repair sequences compaired with the presence of positive acknowledgements. We see that repair sequences do not occur as frequently as positive acknowledgements. Most of the caller's reactions (63%) are positive acknowledgements. Considerably less of the caller's reactions are wh–questions (4%), checks (10%), or reconfirmations (17%), and very few reactions are corrections. In 6% of the cases, the reactions concern questions for other or related travel plans.

Table 5: Dialogue acts expressed by the caller in the information phase of an OVR dialogue

Dialogue acts	Frequency
Positive acknowledgements	63%
wh–questions	4%
Checks	10%
Reconfirmations	17%
Corrections	0.3%
Other (travel plans)	6%

Repair sequences appear at different places in the information exchange. They may appear directly after the utterance to which they react. However, they may also occur after the acknowledgement of the utterance to which they react, and at the end of the complete presentation and acceptance of the travel plan. Table 6 shows the frequencies of these three possibilities for each repair act.

We see that wh–questions, checks, and questions for an extra travel plan mainly occur after the complete presentation of the travel plan. By contrast, reconfirmations and corrections mainly occur directly after the problematic utterance. This difference is of course understandable. Reconfirmations and corrections directly concern problems in processing the previous utterance, while wh–questions and checks mainly ask or check extra information about the travel plan. It is more polite to keep these last kind of questions until the information service is ready with her presentation.

Table 6: The place of the repair sequences in the presentation

Dialogue act	after the problematic utterance	after acknowledgement of the problematic utterance	after the complete presentation
wh–question	-	18%	82%
Check	10%	3%	86%
Reconfirmation	90%	-	10%
Correction	100%	-	-
Other(travel plan)	-	-	100%

Nevertheless, the table shows that speakers may violate these habits, since they may utter reconfirmations after the whole presentation (10%) —although it seemed that they had understood and accepted it— and wh–questions and checks directly after an informing utterance or the acknowledgement of that utterance.

6 Towards a new strategy of information presentation

Analysis of information presentation in VIOS and OVR dialogues shows an important difference in strategy. The VIOS system presents complete travel plans as a whole, while human operators give the information in several chunks. Future releases of VIOS will have to follow this latter strategy as much as possible, because it highly influences the appreciation of clients.

As a result, the travel plan will have to be divided in manageable chunks of information which follow the temporal order of the travel schedule. Each piece of information corresponds to a turn in the dialogue. Such a turn will introduce exactly one new information element as happens in most of the OVR dialogues. This new information element will often be accompanied by an already given element.

We have taken a first step in this direction, by extracting presentation scenarios for different dialogues situations from our sample corpus. The choice for a certain scenario will depend on two types of information:

1. the information elements that the system has gathered during the query phase, and
2. the information that the system has received from the database query.

The information acquired from the database will influence the choice for a certain scenario most, since a travel scheme with two changes will result in another presentation than a direct connection. Nevertheless, the scenario must also contain the right given-new combinations for the individual utterances. The information elements that the system has gathered during the query phase will have to serve as the given information, while the new information that the system has received from the database query will have to function as the new information.

Table 7 gives an example of such a scenario. It shows which scenario should be used, given certain information elements gathered during the query phase and the information elements brought up by the database query. Each line in this scenario refers to a separate chunk of information. It also shows the order in which the elements should be uttered. The table also gives a possible linguistic form of the separate lines in the scenario. After each line, an acknowledgement or a short repair sequence may follow.

The dialogue manager will incorporate the lines into separate statements and will send them one by one the text generator, awaiting the user's reaction before to decide to go on. The dialogue manager will proceed with the next chunk if the user has acknowledged the presented information. In this case, a relatively long period of silence can also be taken as a positive acknowledgement. If the user reacts by a wh-question, a check, or a reconfirmation, the appropriate response will be given before it will continue the presentation. If a misinterpretation is detected the system will first start a correction sequence.

When the complete presentation is finished and thus acknowledged by the caller, he may either finish the conversation or pose a new query. In the last case, the whole dialogue management process will be started again: the representation of the query will be updated, a new database query will be posed, and an appropriate scenario will be chosen. The same will happen when the user corrects the system because it does not give the plan he wants.

To improve information presentation further, dialogue management and text generation will have to collaborate intensively. The text generator will have to choose the right linguistic form, following

Table 7: An example of a scenario

Information given by the query phase:	Departure place, arrival place, and a global indication of the departure or arrival time
Information given by the database query:	One train connection with one change, exact departure times, exact arrival times, place of change which is the same as the direction of the train, the direction of the second train is the same as the arrival place
Scenario:	Departure_Time(new)–Departure_Place(given), Arrival_Time(new)–Place_of_Change(new), Place_of_Change(given)–Departure_Time(new)–Arrival_Place(given), Arrival_Place(given)–Arrival_Time(new).
Possible linguistic form:	Ik heb een trein om [DeT] uit [DeP]. Die komt om [ArT] aan in [PoC]. Daar vertrekt om [DeT] de trein naar [ArP] en dan bent u daar om [ArT] (I have a train at [DeT] from [DeP]. It will arrive at [ArT] in [PoC]. There will leave at [DeT] the train to [ArP] and then you will be there at [ArT].)

the prescribed dialogue act and given-new division. The corpus work described in Rats and Bunt (Rats and Bunt, 1997) on information packaging in Dutch information dialogues may be used for this. The dialogue manager will probably have to incorporate extra contextual information into its instructions, in case several repair sequences will appear between two information chunks. Since in such a case, the system may have to use extra linguistic devices to show the user that he is going to continue the presentation of the travel plan.

7 Future Work

We are currently working on a precursor for the next version of Vios in which our ideas on automated speech processing are incorporated (Vark et al., 1997). Dialogue management will have a predominant role in this precursor as our study has shown dialogue management to be the significant difference between current ASP systems and human-human dialogues. The strategies described in the previous section will serve as an important system guideline to present information. We hope the next version will increase user's acceptancy of automated speech processing systems.

References

J. Allen, L. Schubert, G. Ferguson, P. Heeman, C. Hee Hwang, T. Kato, M. Light, N. Martin, B. Miller, M. Poesio, and D. Traum (1995), The TRAINS Project: A Case Study in Building a Conversational Planning Agent, *Journal of Experimental and Theoretical AI*, 7:7–48.

H. Aust and M. Oerder (1995), Dialogue Control in Automatic Inquiry Systems, in *Proceedings of the ESCA Workshop on Spoken Dialogue Systems*, pp. 121–124.

H. Aust, M. Oerder, F. Seide, and V. Steinbiss (1995), The Philips Automatic Train Timetable Information System, *Speech Communication*, 17:249–262.

S. Bennacef, L. Devillers, S. Rosset, and L. Lamel (1996), Dialog in the RailTel Telephone–Based System, in *International Conference on Spoken Language Processing*, pp. 550–553, Philadelphia.

M. Blasband (1997), Initial Evaluation of the Dutch Environment, Technical report, ARISE Deliverable D 3.10.

L. Boves, J. Landsbergen, R. Scha, and G.–J. van Noord (1995), Language and Speech Technology, Technical report, NWO Priority Programme.

G. Carenini, O. Mittal, and J. Moore (1994), Generating Patient–Specific Interactive Natural Language Explanations, in *Proceedings of the eighteenth Annual Meeting on Computer Applications in Medical Care*, McGraw–Hill Inc.

A. Cawsey (1993), Planning Interactive Explanations, *International Journal of Man-Machine Studies*, 38:169–199.

W. Chafe (1987), Cognitive Constraints on Information Flow, in R. Toulmin, editor, *Coherence and Grounding in Discourse*, pp. 21–51, John Benjamins Publishing Company.

K. van Deemter and J. Odijk (1995), Context Modeling and the Generation of Spoken Discourse, Technical report, IPO manuscript 1125, To appear in Speech Communication 21 (1/2).

E. Klabbers, J. Odijk, J.R. de Pijper, and M. Theune (1997), From data to speech: a generic approach, Technical report, IPO manuscript 1202.

W.A.Th. Manintveld and L.J.M. Rothkrantz (1997), The OVR–WOz experiment: Setup and Analysis, Technical Report 97–04, Alparon, Delft University of Technology.

K. McKeown (1985), *Text generation. Using discourse strategies and focus constraints to generate natural language text*, Studies in Natural Language Processing, Cambridge University Press.

Alexandra Peters and Jens Kock (1997), The digital phone operator, User appreciation and evaluation of the VIOS speech recognition system, Technical Report 97–05, Alparon, Delft University of Technology.

S. Prevost (1996), An Information Structural Approach to Spoken Language Generation, in *Proceedings of the 34th Annual Meeting of the ACL.*

M.M.M. Rats (1996), *Topic Management in Information Dialogues*, Ph.D. thesis, Tilburg University.

M.M.M. Rats and H.J. Bunt (1997), Information Packaging in Dutch Information Dialogues, in *To appear in the Proceedings of the HCM-workshop Discourse and Spoken Dialogue.*

H. Strik, A. Russel, H. van den Heuvel, C. Cucchiarini, and L. Boves (1996), A Spoken Dialogue System for Public Transport Information, in *Proceedings of the Department of Language and Speech*, University of Nijmegen.

R.J. van Vark, J.P.M. de Vreught, and L.J.M. Rothkrantz (1996), Classification of Public Transport Information Dialogues using an Information Based Coding Scheme, in *Workshop Dialogue Processing in Spoken Language Systems*, European Conference on Artificial Intelligence, pp. 92–99, revised paper has also appeared in Lecture Notes in Artificial Intelligence 1236, pp. 55–69, 1997.

R.J. van Vark, J.P.M. de Vreught, and L.J.M. Rothkrantz (1997), An Automated Speech Processing System for Public Transport Information Services, in *Third International Congress on Information Engineering*, pp. 212–221.