

Grounding information in route explanation dialogues.

Philippe Muller & Laurent Prévot

IRIT, Université Paul Sabatier, Toulouse (France)

LOA, ISTC-CNR, Trento (Italy)

Abstract

We have studied the role of several factors influencing the grounding of information in a French corpus of direction-giving dialogues. We focused on how speakers agree on specific locations that are parts of a route explanation. Spatial knowledge is thus seen here as a means of a precise semantic analysis of a certain type of dialog act (feedback responses). We show here the different kinds of feedbacks that can be accounted for by our data, looking at the type of lexical cues in utterances, the type of dialog act targeted by the feedback and the span of the conversation that falls under the scope of such a feedback. Dialog acts also depend on the type of spatial descriptions that are used (motion description, introduction of landmarks or description of landmarks). We then propose an interpretation of this as different mechanisms for speakers' agreement.

1 Introduction

The subject of our study is how speakers in a conversation can agree on the information they exchange when that information concerns spatial or spatio-temporal objects. We have collected route explanation dialogues in a city environment where speakers have to agree (i) on a set of spatial landmarks and (ii) on how they can get from one point in the city to another. Of particular interest to us was *positive feedback* and how it relates to the dynamics of settled information in a conversation (or Conversational Common Ground as defined in (Clark, 1996)), and to the type of spatial knowledge that is involved. There have been several studies detailing the roles that assertions, questions and answers can take in a conversation. Coding schemes for dialogue such as Dialog Act Markup in

Several Layers or DAMSL (Core & Allen, 1997; Carletta & al., 1997) take great care in distinguishing these functions. Following the DAMSL (Core & Allen, 1997) terminology, dialogue acts are divided between those having mainly a *forward-looking function* and those having mainly a *backward-looking function*. A lot of attention has been given to the question/answer pair, its semantics and pragmatic interpretation (Ginzburg, 1996; Asher & Lascarides, 1998). Less emphasis has been put on the role of all speech turns ensuring that information exchanged is properly interpreted (*feedback*). The important work of (Traum, 1994) has studied in some detail how these utterances play a role in deciding the status of information exchanged during a dialogue (mutually accepted or under discussion). He emphasizes that different levels of acknowledgment exist as described by (Clark, 1996) or (Allwood, 1995). We want to show here how distinguishing between such turns is central to the establishment of information, along with question/answer pairs; how they can be accounted for in a structural theory for representing dialogue and how they interact with spatial information.

We have thus studied the role and influence of several discourse markers on acknowledgments in a French corpus of direction-giving dialogues. We focused on how speakers agree on specific locations that are part of an explanation. Spatial knowledge is thus seen here as a means of a precise semantic analysis of the meaning of a certain type of dialog act.

2 Route explanation and dialogue

Route explanations are a specific type of communication about spatial information that has been studied in psychology and psycholinguistics as a good entry in human spatial cognition and the relation between spatial representation and language (Taylor & Tversky, 1992). This is probably because a route is a highly structured and complex piece of information, which is not obvious to communicate or understand. It is common to see a route as a series of temporally ordered actions defined with respect to a set of spatial landmarks. (Denis, 1997) claims that actions are of two main types: change of orientation or plain progressions. In this respect, landmarks are seen as ways of locating actions and other relevant landmarks, or ways of checking the description. When the explanation is given through an interaction between two speakers (as opposed to a monologue as in (Denis, 1997)), specific

communicative issues appear: the receiver of the information will react to the explanation by asking about places she does not know, the giver of the explanation might ask for confirmations of understanding, etc. The structure of the interaction will be associated with the structure of the route, in a manner common to task-oriented dialogues. When studying interaction of this kind, we wanted to study the interplay of communicative actions (questions, answers, feedback) with the type of information exchanged: landmarks descriptions, motion instructions, etc, as a way of studying how information is grounded between participants.

3 A corpus of explanation dialogues and its annotation

To collect empirical data about communication of spatial knowledge we designed the following experiment: we recorded conversations between two speakers on both sides of a telephone line (to restrict communication to verbal communication). We had two sets of subjects. Each one in the first set was supposed to give an explanation on how to get to his/her location (an apartment in the city center) to a subject in the other group, which was located at another apartment not too far from the first one, so that it could be reached on foot.

A “giver” was always associated with a “receiver” she/he didn't know to avoid the use of personal mutual knowledge. To ensure that all conversations were realistic and all subjects motivated by the task, everybody in the second group was invited for a drink at the first apartment a few days after the experiment was conducted. We collected 21 dialogues, for a total of 747 speech turns that were divided in 1235 segments/ dialog acts. All the examples of the paper are extracted from the corpus and presented with their location in the corpus (Prévot, 2004).

We chose to annotate dialogues with the following principles: for each communicative act, we distinguish the form (e.g. assertions, questions) and the function of the act in the context. Moreover we wanted a precise description of the type of task-related information that was involved. If landmarks are central to the route structure, we want to see at what levels they appear and in what types of dialog acts. That's why we needed a more precise annotation scheme than general ones such as DAMSL or Switchboard-DAMSL (Jurafsky et al., 1997). We did not want to use the scheme from the Maptask corpus (Carletta et al, 1997) because there was the most confusion

between Maptask annotators on dialog acts relating to acknowledgment, as explained in (Kowtko, 1996), and acknowledgments constitute the very type we are interested in. The surface part of the annotation was just a distinction between *assertions*, *questions* (*yes-no*, *wh-*, or *disjunctive questions*), *requests* and *disfluencies*. This annotation was carried on the based of the syntactic surface form and the intonation. The function-related part of the annotation scheme deals with task-based dialog acts and communicative acts. Communicative acts were divided between *answer*, *acknowledgment*, *recap*, *reject*, *correction*, *request for feedback*, *request for clarification*, *social act*, and *indeterminate*.

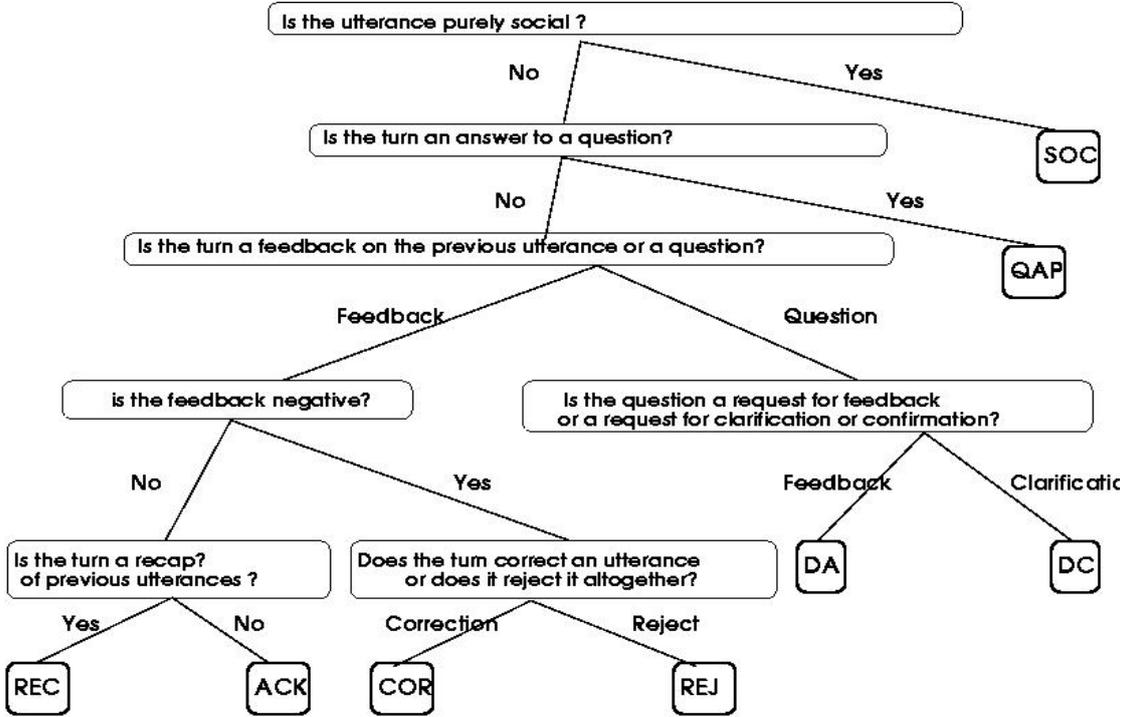


Figure 1: Decision tree for the annotation of dialog acts with a communicative function.

Figure 1 shows the annotation decision tree, which is a subpart of the global tree when it is clear that the relevant turn has a communicative function.

4 Dialog acts and types of spatial descriptions

Denis and Briffault (Denis, 1997, Denis et Briffault, 1997), in their study of purely verbal route descriptions (and without interaction), offer the following structure for explanations of a route:

1. Locate the receiver at the starting point. This means definition and anchoring of the starting point by the receiver and orienting of the receiver.
2. Start the progression ("do that").
3. Introduce a landmark. The landmark indicate the end of a step and help orienting the receiver again.
4. Oriente the receiver.
5. Repeat 2,3,4 until the end point is reached.

They assume that landmarks are introduced in the order of the progression along the route. In contrast divide orientation dialogues in three phases: opening, orientation and closing. Golding and colleagues (Golding et al., 1996) claim that a question about orientation has actually two parts: how to proceed to a place and how to identify the place (if it is not known in advance). This last part is important, as the first one generally leads to a zone around the end point (in our corpus, it was a flat located in a small, not well-known street). Another difference in our corpus is that the starting point of the receiver might not be known by the giver, and there is also a need for explanation. In some cases, speakers use a strategy different from these: they just try to find a landmark near the endpoint and that the receiver might know, and build a simple route from there.

We will now have a more precise look at the basic elements of a description that makes the semantic side of our dialogues. We take the classification of (Denis, 1997) as a basis (itself a refining of (Klein, 1982) and (Riesbeck, 1980)): *prescriptions without landmark*, *prescriptions with landmarks*, *introductions of landmarks*, *description of landmarks* and *comments*, and we complete it with two types of dialog acts: *positioning* and *precision of a description*. Percentages are given relative to the number of speech acts that are task-related (about half of the total number of 1235 speech acts). In the following, a "/" in examples indicates a speaker change.

- prescriptions without landmark (IOL): 5% of task-related speech acts in our corpus (17% in Denis'). They can be either simple instructions (e.g. *tu continues* "you go on") or re-orientations (e.g. *tu tournes à gauche* "you turn on the left").

- prescriptions with landmarks (IWL) 16% of task related speech acts (33% in Denis'). They can make use of previously introduced landmarks, described with definite descriptions (proper names, demonstratives, etc) or introducing new landmarks in the process. (e.g. *tu traverses la rue de Metz* "you cross Metz Street"). Finer semantic distinctions can be made by using a typology of actions or motion predicates. Denis has five types: go-to/go-towards X, take X (a street e.g.), go out of X, cross X, go past X, turn at X.

- introductions of landmarks (IL): 20% in proportion (36 in Denis'). They are static descriptions of the form "*il y a X*" (there is [a] X), "*tu verras*" (you will see), etc. (e.g. *il y a un hôtel de police à cet endroit là* "there is a police station at this place"). Indefinite or definite descriptions can be used, when the speaker assumes the landmark should be known by the other speaker.

- descriptions of landmarks (DR): 47%, the most important class in our corpus (vs. 11% in Denis). The interactive nature of our corpus is the most straightforward explanation for this: speakers spend a lot of time making sure the landmarks they used are identified and produce a lot of utterances which come back on previously introduced entities (e.g. *c'est celle qui va des boulevards vers Ranguel*. "it's the one that is going from the Boulevards to Ranguel").

- comments (COM): 6% of speech acts, mostly used to describe the task itself: "*tu verras c'est facile*" (it will be easy).

- positioning (LOC) 4%: they correspond to isolated phrases that could be part of a larger turn that would for example introduce a landmark or a prescription, but that are interrupted by another speaker: "*après le carrefour [/ oui / tu verras X]*" (after that junction (yes) you will see X), but they can be also be dislocations that are not finished "*quand tu es à X / je ne connais pas X / quand tu es à Y tu tournes à gauche*" (**when you are at X** / i don't know X / **when you are at Y** you turn left).

- precision of prescriptions (PRE, 2%): they bring some information about a preceding prescription; an example is "tu remontes la rue X / mmh / vers Y en fait" (you go up the street / mmh / **towards Y actually**"). Remaining task-related acts were labelled TA. These are constituted by prescriptions or description related to the task but not to the route, such as the characteristics of the button for the doorbell, or the explanation of where is the flat once you entered the building.

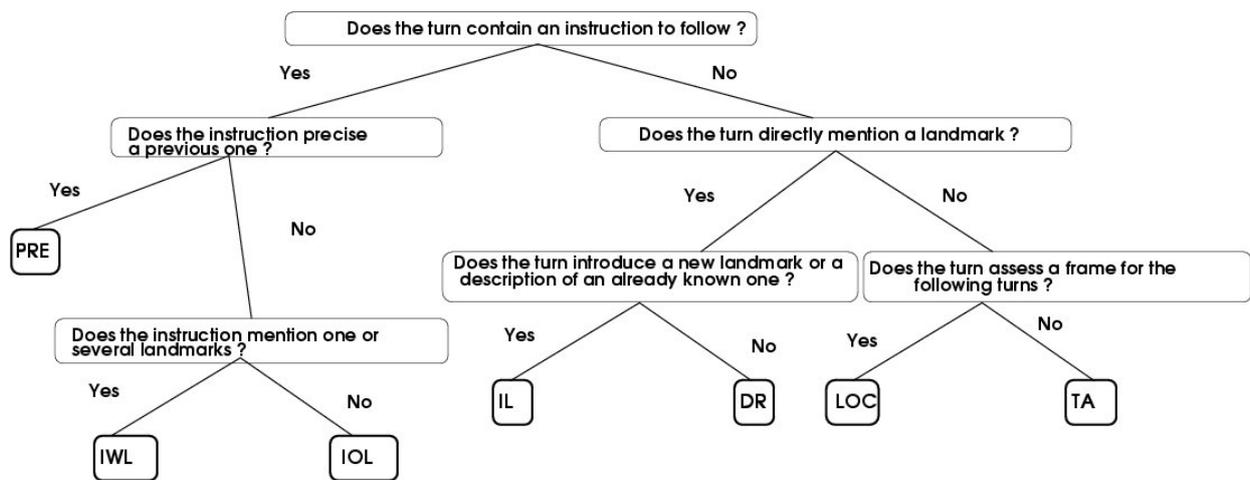


Figure 2: Decision tree for task-related dialog at involving spatial descriptions

The most relevant factor for our study, from a semantic point of view, is the presence or absence of a spatial entity, a landmark, because grounding of information depends almost exclusively on agreeing on these landmarks. People know or don't know some referents introduced, they usually have no problem following or planning an action once this is clear. So we expect to observe different ways of grounding on the classes that may involve new referents (introduction of landmarks, prescriptions with landmarks, most descriptions of landmarks), from acts that don't involve new referents (prescriptions without landmarks, positioning, precision) but are part of the route, and from acts that just describe the task itself (comments).

Figure 2 shows the decision tree for dialog act that are task-related and involve spatial descriptions.

Label	Function	kappa	annotator 1	annotator 2
IWL	Instruction with landmark	0.65	103	92
IOL	Instruction without landmark	0.65	23	31
DR	Description of landmark	0.59	239	195
IL	Introduction of new landmark	0.52	126	138
LOC	Localisation	0.51	64	57
PRE	Precision on a segment	-	26	8
COM	Comment	0.55	43	45
TA	other task-related utterance	-	26	9
ACK	Acknowledge	0.88	346	364
QAP	answer to a question	0.62	117	87
REC	Recap	-	6	21
REJ	Reject	0.54	18	16
COR	Correction	0.18	11	16
RF	Request for feedback	0.51	21	37
RC	Request for clarification	-	13	4
SOC	Social	-	9	1
IND	Indeterminate	-	7	50

Table 1: Agreement on dialog act labels

Table 1 sums up the number of each type of act annotated by our two annotators, with a measure of agreement between them, for the most relevant categories for our study. The kappa measure, advocated by (Carletta, 1996) was used to balance for chance agreement. A good agreement is usually considered at 0.8, while a kappa between 0.6 and 0.8 is considered acceptable. There was the most confusion between the annotators for IWL and IL, because it was not clear in some cases if a turn was an instruction or not. Another source of discrepancies were the unclear status of a lot of turns as polar questions or mere assertions (because the syntax is often the same in modern oral French), leading to different labels for the following turns (either answers or acknowledgments). This prompted us to start a study on this distinction, taking also intonation into account.

5 Different types of acknowledgment

Feedback effects range from full rejection to full acceptance. Here, we will focus on positive feedback (or acknowledgment). Acknowledgments are sometimes mixed with back-channels. Back-channels are speech units

that are uttered “by the listener” or to be more precise uttered by the participant who does not have the ground. Acknowledgments are often back-channelled, but in theory any kind of contribution may be. In consequence, we will focus on the positive feedback aspect without bothering whether it is back-channelled or not. The speaker uttering the response might have heard, understood or agreed on the target of the acknowledgment. It is not often clear what factors take part in defining this level of acceptance, nor how they interact with the other functions of feedback. Therefore, during the annotation task we had only one kind of acknowledgment, but we wanted to be able to distinguish different levels *a posteriori*. Feedback is associated with the establishment of different kind of objects: propositions and/or their truth on the one hand and referents and their identity on the other hand. We can further refine this in four phenomena:

- **grounding** (understanding, settling a proposition) is not related to the truth of utterances or the acceptance of an order. It is just a coordination between the speakers on what has been said, not necessarily its truth nor to acceptance of the information as exemplified in (1). In these examples, G stands for the Giver of the explanation, R for the receiver.

(1) — **Dialogue 2.9**—

Original transcription	<i>Translation</i>
G21: et au coin y’a un restau de tapas.	<i>and in the corner there is a tapas restaurant.</i>
R22.1: ouais.	<i>yeah.</i>
R22.2: je vois pas du tout.	<i>I don’t see at all.</i>

- **accepting** (agreement), as opposed to grounding, leads to the acceptance of the truth about the information agreed on or at least it leads to the acceptance of the information regarding the current purpose. See example (2).

(2) — **Dialogue 1.3** —

Original transcription	Translation
G21: tu vois la première rue à droite avant d'arriver sur la place en fait.	<i>you see the first street on the right before arriving on the square in fact.</i>
R22.1: ouais d'accord.	<i>yeah ok</i>

- **anchoring** (establishing a referent) is finding an internal anchor in one's beliefs for what has been said by the other speaker (e.g. a common referent has been found). See example (3).

(3) — **Dialogue 2.3** —

Original transcription	Translation
G41.3: c'est au 27 rue des Polinaires.	<i>it is at the 27 Polinaires street.</i>
R42.1: ouais je vois.	<i>yeah I see.</i>
R42.2: en fait c'est rue des Polinaires.	<i>in fact it is Polinaires street</i>
R42.3: d'accord	<i>ok</i>
F43: voilà	<i>that's it</i>

- **closing** is terminating a span of discourse as a sub-dialogue successfully or not (See F43 in (3)). We decided to include it in our list because (i) such kind of closing -- after successful exchange or not -- are still agreed among the speakers and (ii) they are concerning mainly positive closures.

However systematically finding the function of a given feedback utterance is not an easy task. There is not direct entailment between them except maybe that *acceptance* requires *grounding*. Accepting an utterance also requires, most of the time, to have anchored it beforehand. Anchoring a sentence means to establish all the referents it contains.

Count	French markers	Produced by Giver(alone)	by Receiver(alone)	English equivalents
141	oui, ouais	34(28)	107(85)	yes/yeah
67	d'accord	18(13)	49(33)	ok, I see
47	voilà	38(31)	9(9)	exactly, that's it
37	ok	9(6)	28(18)	ok
29	mhmm	8(8)	21(20)	mmmh
18	bon	10(5)	8(2)	now, ok, well
14	je vois	1(1)	13(2)	I see
12	repeat	8(8)	4(4)	-
22	other	-	-	-

Table 2: Break-up of markers use, by speaker (alone means there was no other marker in the given utterance)

We have isolated a set of discourse markers (DM) associated with positive feedback in French. They are listed in table 3, along with approximate English equivalents. We add to the DM analysis the observation of informational redundant utterances that help participants infer acceptance as shown in (Walker, 1992). The asymmetry between the two participants allowed us to investigate which markers were correlated with speaker roles. These roles can be associated with initiative or with the competence about the issue currently discussed.

Markers	ouais	oui	ok	d'accord	voilà	mhmm	bon	je vois	others	Tot
Landmark	83	26	30	64	14	9	9	17	4	256
Instruction	38	10	3	18	2	14	0	2	0	87
Feedback	23	6	18	18	34	6	19	1	3	128
Task	21	12	7	23	2	9	3	2	0	79
Total	165	54	58	123	52	38	31	22	7	550

Table 3: Targets of acknowledgment markers by general dialog act types

We have also looked at the function of the previous utterance with the context (its relational function). The data presented in table 4 distinguishes coarsely between task-related assertions describing an itinerary: introduction of landmarks or description of landmarks, instructions, communication management turns that are not related to the task (mainly feedback). We think these distinctions are important with respect to the difference between acceptance and anchoring, since anchoring is mainly about landmark management. The utterance

concerned by the management of landmarks mainly aims to anchor the referents they include, while instructions need to be accepted. Anchoring underlies the establishment of "managing referent utterances" and grounding the establishment of "instruction utterances".

For a more precise breakup of the relationship between the type of semantic act targetted by a given feedback marker, we can have a look at Table 4.

marker transl.	ouais yeah	oui yes	ok	d'accord ok/i see	voilà that's it	mhmm	bon Well	je vois I see	other	total	
IL		23	12	4	17	3	3	2	5	1	70
DR		60	14	26	47	11	6	7	12	3	186
IWL		29	8	1	14	1	11	0	2	0	66
IOL		7	1	1	2	0	2	0	0	0	13
LOC		13	8	3	13	0	9	1	1	0	48
PRE		2	1	1	2	1	1	0	0	0	8
TA		3	3	4	6	1	0	0	1	0	18
COM		5	1	0	4	1	0	2	0	0	13
ACK		12	3	11	7	27	3	7	0	2	72
REC		1	0	0	0	2	0	0	0	1	4
QAP		8	1	7	9	5	2	12	1	0	45
COR		2	2	0	2	0	1	0	0	0	7

Table 4: Precise relation between feedback markers and dialog act types (kappa=0.74).

In order to evaluate the functions of the different markers, we have statistically tested the correlation between each marker and the different role that seemed relevant. We only present here what is more interesting with respect to spatial information, that is the nature of the target of the feedback utterance. We have made a different test for each marker with respect to the following possible types of targets: instructions (IWL, IOL), landmark introduction or description (IL, DR), other task related acts and feedback turns. In each case we used an exact Fisher test on the

contingency table (marker is present in feedback, marker is not present in feedback) against types (landmark, instruction, other task, feedback). Table 5 indicates for each case, the level of significance p and when it was under 0.05, we put the difference with respect to an expected random distribution (marginals), to show what type was more likely to have influenced the test. In other words, a positive difference indicates a correlation between a marker and the role of the target, and the higher the difference, the further the presence of the marker is from a random distribution.

Marker	p	Landmark	instruction	task	feedback
ouais	0,0004	6,2	11,9	-2,7	-15,4
oui	0,0651				
ok	0,0550				
d'accord	0,0346	6,75	-1,46	5,33	-10,63
voilà	0,0000	-10,2	-6,23	-5,47	21,9
mhm	0,0004	-8,69	7,99	3,54	-2,84
bon	0,0000	-5,43	-4,9	-1,45	11,79
je vois	0,0276	6,76	-1,48	-1,16	-4,12
Others	0,3976				

Table 5: Correlation marker vs target of feedback

This confirms our observations: some markers provide feedbacks more often on semantic acts involving referents (*ouais, d'accord, je vois*) or instructions (*ouais, mmhm*) and others are more typically used on feedbacks (*voilà, bon*).

6 Conclusion

The goal of our study was to isolate the different kinds of positive feedback that can be expressed in a conversation and how they can be related to a spatial task, namely direction-giving. We have identified the different kinds of feedbacks that are supported by our data, looking at various lexical cues in utterances, the type of dialog act targeted by the feedback and the span of the conversation that falls under the scope of such a feedback. Refining dialog acts with respect to the specificities of direction-giving show the importance of the distinction between

managing referents in the dialogue history, giving instructions and managing the coordination in the interaction. It provides a good empirical basis for the characterisation of different feedbacks acts (grounding, accepting, anchoring).

References

Allwood, J. (1995). An Activity Based Approach to Pragmatics, *Gothenburg Papers in Theoretical Linguistics*, 76.

Asher, N. and Lascarides, A (1998). Questions in dialogue, *Linguistics and Philosophy*, 21:237-309.

Carletta, J. (1996). Assessing agreement on classification tasks: the kappa statistic, *Computational Linguistics*. 22(2):249-254.

Carletta J., Isard A., Isard S., Kowtko J., Doherty-Sneddon G. and Anderson A. (1997). The reliability of a dialogue structure coding scheme *Computational Linguistics*. 23(1):13-31.

Clark, H. H. (1996). *Using Language*. Cambridge University Press.

Core, M. and Allen, J. (1997). Coding dialogs with the DAMSL annotation scheme. In *Working Notes of the AAAI Fall Symposium on Communicative actions in Humans and Machines*, p 28-35, Cambridge, MA.

Denis, M. (1997). The description of routes: a cognitive approach to the production of spatial discourse. *Current Psychology of Cognition*, 4(16):409-458.

Denis, M. and Briffault, X. (1997). Les aides verbales la navigation. In Denis, M., editor, *Langage et cognition spatiale*, *Sciences Cognitives*, 127–154. Masson, Paris.

Ginzburg, J. (1996). Interrogatives: Questions, Facts and Dialogue. In Lappin S., editor, *Handbook of Contemporary Semantic Theory*, Blackwell, Oxford.

Golding J.M., Graesser A.C and Hauselt J. (1996). The Process of Answering Direction-Giving Questions When Someone is Lost on a University Campus: The Role of Pragmatics, *Applied Cognitive Psychology*, 10:23-39.

Jurafsky, D., Bates, R., Coccaro, N., Martin, R., Meteer, M., Ries, K., Shriberg, E., Stolcke, A., Taylor, P., and Ess-Dykema, C. V. (1997). Switchboard discourse language modeling project final report. *Technical Report Summer Research Workshop Technical Reports 30*, Johns Hopkins University, Baltimore, MD.

Kowtko, J. (1996), *The function of intonation in task-oriented dialogues*, PhD Thesis, University of Edimburgh.

Klein, W. (1982) Local deixis in route directions. In Jarvella R. and Klein W., *Speech, place and Action*, 161-182.

Prévot, L. (2004). *Structures sémantiques et pragmatiques pour la modélisation de la cohérence dans des dialogues finalisés*. PhD, Université Paul Sabatier, Toulouse.

Riesbeck, C. K. (1980). "You can't miss it!" : Judging clarity of directions. *Cognitive Science*, 4:285-303.

Taylor, H. and Tversky, B. (1992). Spatial mental models derived from survey and route descriptions. *Journal of Memory and Language*, (31):261-292.

Traum, D. (1994). *A computational theory of grounding in natural language conversation*. PhD thesis, University of Rochester.

Walker, M.A. (1992). Redundancy in Collaborative Dialogue. *Proceedings of COLING'92*.