Interpretation of Yes/No Questions as Metaphor Recognition

Tarmo Truu **Haldur Õim** **Mare Koit**
University of Tartu
Liivi 2, Tartu 50409
Estonia
{tarmo.truu, haldur.oim, mare.koit}@ut.ee

Abstract

Estonian institutional phone calls are analyzed with the further aim to develop a human-computer dialogue system. The analysis is based on the Estonian Dialogue Corpus. Linguistic cues of yes/no questions are found out that can be used for their automatic recognition.

1 Introduction

Automatic recognition of user questions is one of the main tasks of a dialogue system (DS) which interacts with a user in a natural language. An analysis of human-human conversations is needed in order to find out how do speakers formulate their requests and how hearers understand them.

When a speaker wants the hearer to perform an action, he can express his request directly, using an imperative form (*pass me the salt*); however, it is more polite to use an indirect request (such as *would you pass me the salt?*), which doesn't presuppose any hearer's attitude towards the requested action (in fact, she is questioned about that). Various methods for modulating the strength of utterances are chosen according to the degree of familiarity, respect, relative social roles of the participants of communication, and the impact that the contents of the acts might have on them (Brown and Levinson, 1987). Indirect speech acts may also be considered as allowing more than one characterization. On the standard view, an indirect speech act occurs when a speaker uses an utterance to perform an additional speech act to the one that is 'directly' associated with the utterance in view of its appearance, as illustrated by *Do you know what time it is?* (as a request to tell what time it is) or *What time do you think it is?* as a reproach for being late (Bunt and Girard, 2005).

The idea that constructing new meanings from explicitly given ones forms an inherent part of text understanding process is well known in (cognitive) linguistics. The situation where in order to understand a text one should, proceeding from what is explicitly said in the text, carry out certain operations to reach the 'real' (intended) meaning of the text is in fact much more common in natural communication. Processing indirect speech acts (such as a request in the form of a question) constitutes just one – but very common – case/example of quite analogous processes. Would it be possible to establish some more general mechanisms in human communication which underlie and unite different meaning construction processes? This is the problem we will approach from the point of view of modeling the process of recognizing (understanding) indirect speech acts.

Let's take two examples, one of which is a typical example in the cognitive theory of metaphor and the other – a typical example in the treatment of indirect speech acts. The sentence *My dentist is a real robber* represents a typical use of metaphor. On the other hand, such sentences express indirect speech acts, in the given case e.g. an accusation. The indirect meaning is recognized through interpreting the sentence as metaphorical. Such cases apparently will be outside the abilities of a DS in the near future. On the other hand, in the treatment of indirect speech acts one popular type is request in the form of a question concerning some aspect – as a rule, a pre-requirement of the requested action, e.g. *Can you tell me the arrival time of the bus?* From the point of view of cognitive semantics we can treat such uses as the above question as cases

Joakim Nivre, Heiki-Jaan Kaalep, Kadri Muischnek and Mare Koit (Eds.)
NODALIDA 2007 Conference Proceedings, pp. 398–401

of metonymy: being able to do some D is just one part (prerequisite) of doing D. For the computational analysis of dialogue and recognition of indirect acts this offers a much more clear possibility, especially in the context of institutional information-seeking phone calls. One of the (hypothetical) rules here could be: if the customer is asking whether a prerequisite for an action expected from the information operator does hold then he in fact intends to get the action performed (can you tell me => tell me). The types of actions performed by information agencies can be delimited and their structures where their prerequisites are explicitly formulated can be realized in the corresponding DS (e.g. in the form of frames of actions they are expected to carry out).

In this paper, we will analyze yes/no questions. There are two subtypes of such questions which could be called as direct and indirect ones: (1) (direct) yes/no questions which expect a simple answer *yes* or *no* (e.g. asking *Can you tell phone numbers of private persons?* a speaker intends to get the answer *yes* or *no*), (2) (indirect) yes/no questions which expect giving information (e.g. by asking *Can you tell me the arrival time of the bus?* a speaker intends to receive the arrival time of the bus, the answer *yes* would be insufficient). Let us call these two subtypes as closed (CYN) and open (OYN) yes/no questions, respectively (Gerassimenko et al., 2004). Our aim is to find out (1) how to recognize yes/no questions, (2) how to differentiate these two subtypes, (3) how to model it in a DS.

The rest of the paper is organized as follows: the second section describes the corpus and tools used for the analysis; the third gives an overview of the results of analysis – some linguistic cues which have been found out for recognition of yes/no questions; the fourth section represents some ideas how to model the interpretation of speech acts in a DS. Finally, some brief conclusions are presented.

2 Corpus and Tools Used

Our current study is based on the Estonian Dialogue Corpus (EDiC). The corpus contains over 900 authentic human-human spoken dialogues, including over 800 phone calls. Dialogue acts are annotated in the corpus using a DAMSL-like typology of dialogue acts (Gerassimenko et al., 2004).

Dialogue acts used for requesting information form a certain act group which is differently classified in different typologies. In the typology used by us, questions are determined as the utterances which have a specific form in Estonian: interrogatives, a specific word order and/or intonation. Questions are differentiated from directives. For example, *Can you tell me the arrival time of the bus?* is considered as a question (indirect request, OYN) but *Tell me the arrival time of the bus* is a directive (request). OYN and CYN have similar form in Estonian but they expect different reactions from the partner. A CYN is a direct dialogue act and expects the answer *yes* or *no* (e.g. *Are you open in winter? – Yes.*) while an OYN expects giving information (e.g. by asking the question *Is there a bus that arrives after 8 p.m.?* the customer intends to learn the departure times of buses). An OYN is an indirect dialogue act – the speaker forms his actual request as a question (an act of another type, Hennoste et al., 2005).

For the analysis, a sub-corpus of EDiC was chosen consisting of 312 directory inquiries. Customers ask phone numbers, addresses, opening hours of institutions, etc. The Workbench of EDiC was used for calculations and analyses.[1]

3 Corpus Analysis

Our first aim is to find out some linguistic cues which can be used for recognition of yes/no questions. Let us consider two examples from EDiC, the first one is annotated as a closed and the second one as an open yes/no question[2]:

```
'kas te mulle 'saate firma 'nime
vaadata 'numbri järgi,=
```

can you give me the name of a firm on the basis of a phone number? CYN

```
.hh kas teie käest saaks informat-
siooni kui palju võiks maksta sõit
'Inglismaale.
```

could you give information about how much does a trip to England cost? OYN

The linguistic form of both utterances is the same but the expected responses are different.

Customers asked 76 and operators 67 CYNs. The number of OYNs is 163 and 19, respectively. The reason of the significant difference in numbers

[1] http://math.ut.ee/~treumuth/
EDiC is accessible via the Workbench, but it is password-protected.

[2] Transcription of conversation analysis is used in examples (Schegloff, 1986).

of OYN is that the goal of a customer is to get information, and it is reasonable for him to expect a longer answer to a yes/no question than simply *yes*. An indirect question includes both direct and indirect meanings (Clark, 1991). In our case, an OYN includes two meanings – a direct and an indirect wish of the speaker.

We analyzed only customers' yes/no questions having an aim to find out how the computer performing the role of an operator could recognize users' dialogue acts. The analysis was carried out in two parts. The first part examined most important cue words, and the second one considered the same cue words together with the interrogative *kas* (whether) which is a significant key of yes/no questions in Estonian. The Table represents the results of our analysis: the numbers of cues found in OYNs, CYNs and in the remaining part of the sub-corpus, and possible recognition percents which one can expect to achieve using these cues.

The most interesting cue includes the word *saama* (to be able). Using only this word alone as a cue one can achieve the recognition percents 3.9 and 12.3 for CYN and OYN, respectively. But using this word together with the interrogative *kas* (whether) the percents increase to 22.5 and 52.5, respectively. It is not surprising because the question *kas sa saad teha D (are you able to do D)* includes a prerequisite of doing D.

The amount of the analyzed sub-corpus is too small to make some general conclusions. Still, it is clear that there exist linguistic cues which can be used for automatic recognition of yes/no questions.

4 Computational Model of Interpretation of Yes/No Questions

Communication is the intentional exchange of information. The speaker S wants to inform the hearer H about proposition p. H infers that S intended to convey q (where q=p in ideal case). S is not intentionally ambiguous but most utterances have several interpretations. H infers the most probable interpretation of p (speaker's interpretation or metaphorical interpretation). Such framework can be implemented both for dialogue analysis and recognition of metaphors (or metonymes) where the task of a hearer (reader) is to understand the actual intention of the speaker (author).

Table. Cues, their numbers and percents in open and closed yes/no questions

Cue	# OYN	# CYN	# other acts	% OYN	% CYN
kas (whether)	92	53	75	41.8	24.0
mingi (some, a certain, a kind of)	38	1	52	41.7	1.1
vä (or)	3	6	0	33.3	66.6
mõni (some, any, a few)	5	0	2	71.4	0.0
midagi (something, anything)	21	0	24	46.6	0.0
ütlema (to tell)	27	1	43	38.0	1.4
võimalik (possible)	9	4	24	24.3	10.8
näiteks (for example)	10	3	48	16.3	4.9
saama (to (be able)	31	10	210	12.3	3.9
ega (nor)	8	1	9	44.4	5.5
kas + another interrogative	7	0	0	100.0	0.0
kas + mingi	17	0	6	73.9	0.0
kas + mõni	3	0	0	100.0	0.0
kas + midagi	9	0	1	90.0	0.0
kas + võimalik	5	2	0	71.4	28.5
kas + ütlema (to tell)	13	0	11	54.1	0.0
kas + näiteks	4	3	1	50.0	37.5
kas + saama	21	9	10	52.5	22.5
tahtsin + küsida (I wanted to ask)	6	0	1	85.7	0.0

Similar methods have been used for solving of both tasks. The computational models of both dialogue act interpretation and metaphor recognition can be divided into two classes. The first class has been called cue-based or probabilistic. The idea is that the hearer (or reader) uses different linguistic cues of the utterance to build its non-literal meaning. Sometimes, an utterance can be considered as an idiom. The second class of models implements the inferential approach. Such models are based on belief logics and use logical inference to reason about the speaker's intentions (Jurafsky and Martin, 2000).

We are working on a computer model of information seeking dialogues in Estonian and experi-

menting with different approaches to recognizing communicative intentions, including indirect ones, and expressing these intentions (by the computer) in a human manner. Recognizing and using metaphorical and metonymic expressions is one of the methods investigated, the general and uniting key-concepts being 'meaning construction' and 'communication through reasoning'. Our work is based on the EDiC and our general framework is one kind of BDI model worked out in Artificial Intelligence (Koit and Õim, 2004).

How to differentiate the OYNs and CYNs? We propose the following analysis cycle: first, linguistic cues are used to recognize the type of a dialogue act (a yes/no question), and secondly, frame representations of dialogue acts are used to interpret the act. We have built frames of questions, and a frame of an OYN is a combination of the frames of the CYN and wh-question. On the ground of the hypothetical rule formulated in Section 1 a frame of request can be inferred and constructed. Therefore, we try to combine the two kinds of computational models to interpret questions as dialogue acts – cue-based and inferential-based.

Solving of metaphors, metonymes and indirect speech acts can be considered as meaning disambiguation: the task is to choose one of several possible meanings which suits with a context. The hearer has to find out the most probable intention of the speaker.

5 Conclusion and Future Work

Estonian directory inquiries were analysed with the further aim to develop a DS. Yes/no questions were considered in order to find out some cues which can be used for their automatic recognition. It turned out that the most important linguistic cues of yes/no questions are (1) the interrogative *kas* (whether), (2) the pronouns *mingi* (some, a certain, a kind of), *mõni* (some, any, a few), *midagi* (something, anything), 3) the verb *saama* (to be able). By combining of these cues, the recognition accuracy can be increased. For example, *kas* + *saama* is a significant cue.

The first task of the computer is to recognize a yes/no question. After that, a closed and an open yes/no question can be differentiated as direct and indirect dialogue acts. An indirect dialogue act includes at least two meanings while the partner has to react to the most important one.

A simple DS is implemented which gives information about flights leaving from the Tallinn Airport. Our future work concerns implementation of the described method of recognition of yes/no questions in the DS.

Acknowledgement

This work is supported by Estonian Science Foundation (grant No 5685).

References

Penelope Brown and Stephen C. Levinson. 1987. Politeness: some universals on language usage. *Studies in Interactional Sociolinguistics* 4. Cambridge University Press.

Harry Bunt and Yann Girard. 2005. Designing an Open, Multidimensional Dialogue Act Taxonomy. Claire Gardent and Bertrand Gaiffe (eds). *Dialor'05. Proc. of the ninth workshop on the semantics and pragmatics of dialogue (SEMDIAL)*. Loria, 37-44.

Herbert H. Clark. 1991. Responding to Indirect Speech Acts. *Pragmatics.* Ed S. Davis. Oxford: Oxford University Press.

Olga Gerassimenko, Tiit Hennoste, Mare Koit, Andriela Rääbis, Krista Strandson, Maret Valdisoo and Evely Vutt. 2004. Annotated dialogue corpus as a language resource: an experience of building the Estonian dialogue corpus. *The first Baltic conference "Human language technologies. The Baltic perspective"*. Riga, 150–155.

Tiit Hennoste, Olga Gerassimenko, Riina Kasterpalu, Mare Koit, Andriela Rääbis, Krista Strandson and Maret Valdisoo. 2005. Questions in Estonian Information Dialogues: Form and Functions. *Text, Speech and Dialogue. 6th International Conference TSD 2005*. Ed. V. Matousek, P. Mautner. Springer, 420-427.

Daniel Jurafsky and James H. Martin. 2000. *An introduction to natural language processing, computational linguistics, and speech recognition.* Prentice Hall.

Mare Koit and Haldur Õim. 2004. Argumentation in the Agreement Negotiation Process: A Model that Involves Natural Reasoning. *Proc. of the Workshop W12 on Computational Models of Natural Argument. 16th European Conference on Artificial Intelligence.* Valencia, Spain, 53-56.

Emanuel Schegloff. 1986. The routine as achievement. - *Human Studies*, 9, 111-152.