Determining Dialogue Acts in Estonian DialogueCorpus

Tiit Hennoste[1], Mare Koit[2], Maret Kullasaar[2],
Andriela Rääbis[1], Krista Strandson[1], Evely Vutt[2]

[1] University of Tartu, Department of General Linguistics, Tiigi 78,
50410 Tartu, Estonia
{thenno, araabis, kstrand}@psych.ut.ee
http://www.cl.ut.ee/
[2] University of Tartu, Institute of Computer Science, Liivi 2,
50409 Tartu, Estonia
{koit, maret, nurm}@ut.ee
http://www.cs.ut.ee/

Abstract. The paper gives an overview of Estonian dialogue corpus. The corpus currently includes dialogues of two kinds – recordings of spoken conversations, and dialogues collected by Wizard of Oz technique. We have worked out a typology of dialogue acts and are annotating them in our corpus. Some problems of determining of dialogue acts are considered, among them different kinds of questions. First of all, we are interested in question-answer dialogues, and our further aim is to develop a dialogue system which can interact with a user in natural language following norms and rules of natural human communication.

1 Introduction

The paper discusses some problems of annotating Estonian dialogue corpus.
Our study is done in the framework of a research project. The aim is to work out an experimental dialogue system which can interact with a user in Estonian and give him/her information about trips, bus, train etc. schedules.
We started our research with analysis of some dialogue models and corpora for the purpose to work out a model suitable for our implementation. Simultaneously, we collected a dialogue corpus and started to annotate dialogue acts in the corpus. The annotated corpus is necessary for developing and testing a dialogue system.
Estonian dialogue corpus consists of two kinds of texts.
1) (Transliterated) dialogues from corpus of spoken Estonian [1].
2) Dialogues collected in computer simulations using the Wizard of Oz (WOZ) method.
There are currently 232 transliterated texts in our corpus with total length of 64,600 word forms – 120 telephone calls and 112 face to face conversations.

All the recordings were transliterated using transcription of conversational analysis [2]; for Estonian cf. [3]. Dialogue acts are annotated in 17 dialogues, each by 2 annotators. 10 from annotated dialogues were discussed and unified by the whole team.

21 dialogues are collected by WOZ technique. The participants were asked to test a program which gives adequate and correct answers to their questions in Estonian but actually a person was playing the role of the computer through a network [4]. We wrote a program - an analogue of telnet-like chatrooms for implementing and saving dialogues in our experiments. A text field on a www page was used for putting in participants' utterances and getting the Wizard ones. Dialogue acts are tagged in all of the WOZ dialogues.

To simplify the annotating process, we wrote an annotation environment [5]. The program provides a menu for an annotator where he/she can choose tags – dialogue acts. We have taken the demonstration system of the project MATE [6] as an example. Several annotators are currently testing our annotation tool.

2 Dialogue Acts

When interacting, the people are doing something with help a language – they are asking, answering, greeting, etc. Such acts (realised by a language) are called as speech acts, conversational acts, dialogue acts, speech actions, etc.

The first well known typology of conversational acts which based on real dialogues was worked out by John Sinclair and Malcholm Coulthard [7]. This typology further was developed by Anna-Brita Stenström [8]. Several researchers are considering practical problems of determining conversational acts in the last decade – corpus linguists, discourse and conversation analysts, language technologists [9], [10], [11].

We are interested in certain kind of dialogues – information seekings. We are examining dialogue acts and their linguistic realisation in such dialogues. Our goal is to determine linguistic features characterising different dialogue acts for that a dialogue system could automatically recognise them and respond adequately to user queries.

Several theoretical problems connected with peculiarity of natural speech and oral dialogue must be solved for this purpose. We decided to use the method of conversation analysis [12], [13], [14]. The departure point of this method is the fact that a human reacts always and certainly to the previous turn in conversation. Another reactions *can* be but this reaction *must* be.

There are three basic organisations in conversation: turn-taking, adjacency pairs, repair organisation.

Participants are **taking turns**. Turns are ordered sequentially and linked together into definite sequences. The hearer must recognise when the turn of the speaker is finished and when she may start to speak. There are two sub-problems here.

1. What is the carrying entity of a dialogue act? A turn is a continuing speech of one speaker. Turns consist of turn constructional units, TCU, which boundaries are intonational and grammatical-pragmatical. In ideal case, the carrying entity of a

dialogue act must be the least linguistic unit used for performance an action. The main carrying unit of an act is TCU but it does not hold as rule. A part of TCU, or many TCUs together can realise an act. Sometimes, a phrase or a word can be separate dialogue acts. Therefore, no certain linguistic construction exist which could be classified as a carrying unity of dialogue acts. The study of empirical material is necessary for determinig different dialogue acts and their linguistic realisations.

2. Can one unit have only one or more functions? In our system, one dialogue act may perform many functions. It means that acts can be classified on several grounds. One classification critery is formal (for example, open questions, yes/no questions, etc.). The another critery is essential - a function of an act in conversation (adjustable questions, re-questions, etc.). Therefore, any turn can get more than one tag in annotated dialogue corpus.

A problem is depth of the classification of dialogue acts. Our analysis shows that it is necessary to classify different kinds of dialogue acts with different depth. The acts must be brought out whereafter a partner must choose between different reaction types, i.e. acts which change the course of conversation. Such acts can be various – commands, wishes, asks. It is necessary for a partner to choose his/her reacting act from this same group of acts.

Conversation is based on a system of **adjacency pairs**. This fact is most important for information seeking dialogues. An adjacency pair is a pair of turns where the first part requires the certain second part (so as a question requires an answer). Thus the speaker is expecting hearer's answer or explanation if it is impossible to answer.

The second parts can be classified as expected and non-expected ones. Expected second parts are, for example, fulfilling a command, giving the requested information, etc. They have different linguistic realisations [14], [15].

It is necessary to find basic types of adjacency pairs and their possible variants. For example, question–answer is an adjacency pair, and questions requiring different answering ways are different variants.

Any adjacency pair can contain insertion sequences, pre-expansions, post-expansions. Insertion sequences often are used if a responder does not have sufficient information to answer the question. For example, a caller asks for good restaurants but the data base of the information office is classified using other features: location, type of food. In this case, a responder must explain his/her possibilities, or suggest an another data classification before answering.

Participants must understand, was it the first part of an adjacency pair or not, and is the next turn the second part or an insertion sequence.

Conversation runs smoothly only in ideal case. Real conversations are full of false starts, repetitions, pauses, corrections of mistakes, non-understandings etc. This fact is weakly estimated in dialogue models. Thus, a permanent machinery is needed for solving communication problems – **repair organisation**. It can be classified according-ingly 1) who initiates and who makes the repair, and 2) is the repair made to solve problems in prior text or is it used to avoid problems before-hand [14], [16].

It is important for the computer to understand a self-repair of a user. If the user says: *please give me all buses from Tartu to Tallinn, (.) to Pärnu* then the computer must understand that the user is interested only in buses to Pärnu – it is a self-repair. Currently, we do not annotate self-repairs in our corpus.

3. Macrostructure of Information-Seeking Dialogues

Information seeking dialogue consists of four different parts with different functions typically (cf. Example 1).
- a ritual begin – greeting, introducing (rows 1-6 in the next example)
- solving communication problems in cooperation – non-understanding, incredibility of information (rows 8-20). Often this part follows after the first question and forms an insertion sequence within the first adjacency pair. Solving communication problems can be repeated within the next adjacency pairs.
- giving the requested information – answering questions, accounting phone numbers, etc. (rows 22-25)
- a ritual end – thanking, leavetaking (rows 27-28).

Example
```
1. ((kutsung))
((call))  CALL
2.V:   ´Estmar=´info
'Estmar= 'information,  RESPONDING THE CALL | INTRODUCING
3. ´Leenu=kuuleb
Leenu is hearing  INTRODUCING
4. tere
Good afternoon GREETING
5. H:  tere ´õhtust,
Good afternoon  RESPONDING THE GREETING
6. ma paluks ´Maarjamõisa ´kööki.
Please give me Maarjamoisa's kitchen.  REQUEST
7. (2.2)
8. V:  ´haigla juures või poli´kliinikus.
By hospital or by ambulance? ALTERNATIVE QUESTION | ADJUSTMENT
OF CONDITONS OF ANSWER
9. H:  ´haigla.
Hospital  ALTERNATIVE ANSWER: ONE | PERFORMANCE OF REPAIR
10. (0.5)
11. V: e ´köögi numbrit ei ole meil ´antud.
mm we do not have kitchen's number FULFILLING OF DIRECTIVE:
MISSING INFORMATON
12. (0.5)
13. H: sääl neil ´on telefon peal.
They have a phone there        ASSERTION
14. V: jah nendel võib ´olla
Yes they can have             AGREEMENT WITH RESERVATION
15. aga meil ei ole ´antud köögi ´numbrit.
but we do not have kitchen's number     ADDING  INFORMATION:
STRESSING
16. (0.5) ma saan teile ´anda ´üldinfo ´numbri.
I can give you a general number     OFFER
17. (0.8) Maarjamõisa ´haiglas.
of the hospital        ADDING INFORMATION: ADJUSTING
18. (1.0)
19. H: ahah.
OK.   SIGNAL OF NEW INFORMATION
20. (1.0) jah, (.) ´olge pai, ´andke.
yes, give please        YOU ARE WELCOME | REQUEST
```

```
21. (.)
22. V: neli neli kaheksa,
four four eight   FULFILLING OF DIRECTIVE: GIVING INFORMATION
23. (0.8)
24. H: kaheksa
eight          MEMORISING REPETITION
25. V: üks üks viis
one one five FULFILLING OF DIRECTIVE: GIVING INFORMATION
26. (.)
27. H: ma tänan.
thank you      THANKING
28. V: ja palun?
you are welcome   YOU ARE WELCOME
```

Transition from one part to another always is a result of bilateral contract. One side proposes to finish a part, and another must accept it. If partner does not accept the proposal then the same part is continuing. Thus we must select the dialogue acts with help of which transitions between different parts of conversation are possible. The computer must understand if a user signals to change a topic or finish interaction.

We have classified our system of dialogue acts taking in account the problems considered above.

There are 8 goups of acts in our system.

1) Rituals – greeting, thanking, introducing, finishing.

2) Acts for re-structuring conversation, with help of which the speaker starts a new topic or changes the type of conversation.

3) Acts for exchanging of turn-takings, with help of which the speaker is asked to continue, or the existence of contact is checked.

4) Repairing acts, with help of which partners are solving communication problems.

5) Directive acts for giving and receiving of commands, requests, etc.

6) Questions and answers – pairs of acts, with help of which one partner asks a question and another answers them.

7) Acts for taking up of attitudes, with help of which one partner represents an attitude (belief, evaluation, charge) and another responds them.

8) The last group contains remaining acts (giving information, argument, conclusion, promise, etc.).

All the groups, except of the last, can form adjacency pairs. For that reason, they are divided to two sub-groups: the first and second parts. The first parts are used to give commands, ask questions, etc. The second parts express reactions to commands, answers to questions. Acts from the 8th group can supplement both the first and second parts. A simplified formal grammar defining our dialogue act system is given in [17].

4 Some problems of determining dialogue acts

4.1 Questions

We are using two different classifications simultaneously: formal and functional. By form, linguistic realisation, questions can be classified as open questions, closed questions yes-no, declarative questions, alternative questions. From the other side, questions can have different functions in conversation – adjustable questions, re-questions. If necessary, questions get two different tags in an annotated dialogue.

Example: *What's the name of this place more exactly?* OPEN QUESTION | ADJUSTABLE QUESTION

Two groups of problems can be mentioned: how to differentiate 1) closed questions yes-no from declarative ones, and 2) adjustable questions from adjustments of conditions of answer?

Formally, **closed and declarative questions** are similar. In Estonian, words *kas* or *vä* are used typically. The difference is in expected answers. A closed question expects an answer yes or no (*Are you open in winter? – Yes/no*). Using a declarative question, the speaker wants to get some information. For example, if a speaker asks *Do some bus arrive at 8 p.m. in Tallinn?* then he expects that the hearer lists the departure times of buses.

The situation will be more complicate if the hearer interpretes wrong the question. Example: a caller wants to get some information about a hostel and asks *Do you have any conference rooms?* He expects to get a description of conference rooms. A pause after the question indicates that the responder has interpreted the questions as a closed question and beliefs that answer *yes* is sufficient. The caller is expecting a longer answer and does not take over the turn.

Formally, both of **adjustable questions and adjustments of conditions of answer** are questions. Their function is to get an additional information from a partner. There are two distincting factors: 1) who asks the question – the asking or the answering participant, and 2) does the question start a repair or not.

The next question about the same thing is an adjustable question if a speaker has asked a question before them. Asking the question, the speaker will adjust some details. It does not start a repair. Example: a caller wants to know the location of a hostel. Answer *in West country* is too general, and he asks an adjustable question *is it near the sea*?

If the answering participant needs an additional information to compile the answer then he/she starts an inserted sequence. We annotate such kind of questions as adjustment of conditions of answer. The question starts a repair sequence and the partner performs the repair in the next turn – gives information. Now the answering participant can implement the second part of an adjacency pair – he/she answers the first question. Thus, an adjustment of conditions of answer is a responder's act. Example: a caller asks for concert trips. The responder's act *what kind of concerts are you interested in?* is an adjustment of conditions of answer.

Re-question is similar with the questions considered so far. It is a way to start a repair for both participants. Re-question distincts from adjustable question and adjustment of conditions of answer – the known information is asked once more. Re-question often expresses astonishment if an answer was non-expected.
Example: a caller asks for the phone number of teachers room of a school.
```
Responder: I can give only director, office and head
Caller: you don't have a teachers room?     RE-QUESTION
```

4.2 Continuer and acknowledgement

The speaker must understand when the hearer wants to start her own turn and when she gives a feedback.

The central dialogue particles for expressing reaction to previous turn are *mhmh, ja* and *ahah* in Estonian (correspondingly, *hem, yes, oh*). They are playing different roles in conversation: *mhmh* shows that a participant is hearing and distancing, *jah* shows agreement, *ahah* shows that previous information was new for the hearer.

A dialogue particle can form a turn alone or start a new more longer turn.

If a hearer uses particle (often *mhmh*) inside a partner's long turn then she signalises that the speaker may continue. The dialogue act is continuer in this case. It is more difficult to interprete particles if they are used as reactions to partner's previous finished turns. In this case, the hearer may take over the turn. If she does not want it then the particle signalises that turn goes back to the speaker and dialogue act is a continuer (rows 4,7,9,12 in the next example). If, on the contrary, the hearer wants to take over the turn then she goes forward immediately after the particle. We annotate such dialogue acts as acknowledgements (row 14). An acknowledgement finishes a sequence and a question, request, thank is following - the first part of the next adjacency pair.

Example
```
1. H: millal lähevad täna: bussid ´Rakvere umbes=e ´kaheteist
ja ´kolme vahel.
when do the buses depart to rakvere between 12 and 3  p.m. OPEN
QUESTION
2. (1.5)
3. V: läheb (.) kaksteist ´viisteist.
Twelve fifteen  OPEN ANSWER: GIVING INFORMATION
4. H: mhmh        CONTINUER
5. V: sis läb kell ´üks.
Then at one o'clock   OPEN ANSWER: GIVING INFORMATION
6. (0.5)
7. H: jaa        CONTINUER
8. V: ja pool ´kaks, kolmteist kolmkümend.
And half past one OPEN ANSWER: GIVING INFORMATION
9. H: mhmh        CONTINUER
10. V: kell ´kaks,     | AVATUD VASTUS: INFO ANDMINE |
At two o'clock        OPEN ANSWER: GIVING INFORMATION
11. (0.5)
12. H: jah        CONTINUER
13. V: ja pool ´neli on siis.
And then half past three   OPEN ANSWER: GIVING INFORMATION
```

```
14. H: mhmh.    ACKNOWLEDGEMENT
15..hh aga: milline neist kõige ´kiiremini=nagu ´sõidab.
which of them is going more quickly   OPEN QUESTION
```

5 Future work

Our further work is concentrated on some other problems of (formal) determining dialogue acts. Our aim is to refine the annotating software for automatic tagging of some kinds of dialogue acts. Simultaneously, we are adjusting our typology.

References

1. Tartu University Corpus of Spoken Estonian.   http://sys130.psych.ut.ee/~linds/english/
2. Psathas, G.; Anderson, T. The "practices" of transcription in conversation analysis. – Semiotica 78 (1990) 75–99
3. Hennoste, T. Sissejuhatus suulisesse eesti keelde I–IX. Akadeemia (2000) 5, 1117–1150; 6, 1343–1374; 7, 1553–1582; 8, 1773–1806; 9, 2011–2038; 10, 2223–2254; 11, 2465–2486; 12, 2689–2710; (2001) 1, 179–206
4. Kullasaar, M. Eestikeelse dialoogikorpuse arendamine "võlur Ozi" tehnikaga. TÜ (2001)
5. Nurmsalu, E. Eestikeelse dialoogikorpuse märgendustarkvara. TÜ (2001)
6. Dybkjaer, L. MATE Deliverable D6.2. Final Report.
   http://mate.nis.sdu.dk/about/deliverables.html (used 19. 02. 2001)
7. Sinclair, J.M.; Coulthard, R.M. Towards of Analysis of Discourse: The English used by Teachers and Pupils. London: Oxford UP (1975)
8. Stenström, A.-B. An Introduction to Spoken Interaction. London and New York: Longman (1994)
9. Allwood, J.; Ahlsen, E.; Björnberg, M.; Nivre, J. Social activity and communication act-related coding. – Cothenburg Papers in Theoretical Linguistics 85. Dialog Coding – Function and Grammar. Göteborg Coding Schemas. Ed by Jens Allwood. Goteburg (2001) 1– 28
10. Stolcke, A.; Coccaro, N.; Bates, R.; Taylor, P.; Van Ess-Dykema, C.; Ries, K.; Shriberg, E.; Jurafsky, D.; Martin, R.; Meteer, M. Dialogue act modeling for automatic tagging and recognition of conversational speech. – Computational Linguistics 26:3 (2000) 339–373
11. Jokinen, K.; Hurtig, T.; Hynnä, K.; Kanto, K.; Kaipainen, M.; Kermanen, A. Self-Organizing Dialogue Management. Proc. of NLPRS Workshop, Tokyo (2001)
12. Sacks, H. Lectures on conversation. Ed. by Gail Jefferson. Cambridge & Oxford: Blackwell (1992)
13. Hutchby, I., Wooffitt, R. Conversation Analysis. Principles, Practices and Applications. Cambridge, Polity Press (1998)
14. Ten Have, P. Doing Conversational Analysis. Sage (1999)
15. Levinson, S.C. Pragmatics. Cambridge UP (1982)
16. Schegloff, E. A., Jefferson, G., Sacks, H.. The preference for self-correction in the organization of repair in conversation. – Language 53 (1977) 361-382
18. Koit, M. Annotating dialogue acts in Estonian Dialogue Corpus. Proc. of RANLP'01. Ed. G. Angelova, K. Bontcheva, R. Mitkov, Nicolas Nicolov, Nikolai Nikolov. Tzigov Chark, Bulgaria (2001) 273-275