Human Language Technologies – The Baltic Perspective A. Tavast et al. (Eds.) © 2012 The Authors and IOS Press. This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License. doi:10.3233/978-1-61499-133-5-120

Towards Automatic Recognition of the Structure of Estonian Directory Inquiries

Mare KOIT^{a,1}

^a University of Tartu, Estonia

Abstract. We analyse dialogues in order to determine the dialogue structure formed by micro-level units – dialogue acts. The empirical material of the study is a sub-corpus of Estonian directory inquiries. Dialogue recordings are transliterated by using the transcription of conversation analysis. Dialogue acts are annotated in the corpus. Rules for identification of different dialogue parts will be formulated which use sequences of dialogue acts and their position in dialogue. Our further aim is to implement software for automatic pragmatic analysis of dialogues in order to recognize their linear structure as well as sub-dialogues.

Keywords. dialogue, structure, automatic recognition

Introduction

Studies of human dialogue behaviour indicate that natural dialogue utterances are very often multifunctional. This observation has inspired the development of multidimensional approaches to dialogue annotation and analysis. The most frequently used annotation scheme is DAMSL [1], which allows multiple labels to be assigned to utterances in four layers. Both symbolic and statistical learning methods have been used for dialogue processing [2-4], e.g. a mechanism is presented to learn repeated substructures of dialogue from a dialogue corpus [5].

We are implementing software for automatic pragmatic analysis of Estonian dialogues. Our current aim is to determine the formal structure of dialogue formed by micro-level units – dialogue acts. Empirical material of our study is a sub-corpus of Estonian directory inquiries, recorded in authentic situations. People call information services asking for different information – phone numbers, addresses, opening hours of institutions, etc. We are looking for cues that enable to identify the linear structure of a directory inquiry. We also aim to identify sub-dialogues of a dialogue – information sharing and (other-initiated) repair (non-understanding, checking, and reformulation) [6]. Information-sharing can be initiated by both dialogue participants – by the information provider in order to adjust the conditions of answer, and by the caller if

¹ Corresponding Author: University of Tartu, J. Liivi 2, 50409 Tartu, Estonia; E-mail: mare.koit@ut.ee.

the answer needs to be adjusted. Similarly, repairs can be initiated by both participants whenever during a dialogue.

The paper is organized as follows. Section 1 gives an overview of the Estonian Dialogue Corpus and the dialogue act typology used for annotation of the corpus. Section 2 is dedicated to the corpus analysis. Section 3 presents the typical structure of a directory inquiry found as a result of the analysis. Section 4 makes conclusions.

1. Corpus and Dialogue Act Typology

Our current study is based on the Estonian Dialogue Corpus (EDiC) [7] that consists of three parts. The first part is formed by human-human dialogues taken from the Corpus of Spoken Estonian – it includes at present 1137 telephone calls (directory inquiries, calls to travel agencies, to outpatients' offices, telemarketing calls, etc) and 45 face-to-face conversations, altogether 1182 transcribed texts with total size of 250,000 running words. The second part of the EDiC contains almost 200 written dialogues collected in computer simulations using the Wizard of Oz method. The third part – human-computer interactions (more than 1000 written dialogues) is being collected with two simple dialogue systems which interact with a user in Estonian. One of them gives information about cinema programs in Tartu, and another – dental treatment information.

Dialogue acts (DA) are annotated in the corpus by using a special annotation typology founded in the University of Tartu in 2001 [7,8]. The primary aim of creating the typology was study of Estonian human-human conversations. The typology is based on the principles of organization of conversation borrowed from conversation analysis (CA) [9]. The departing point of the CA is that a partner always must react to the previous turn regardless of his/her own plans and strategies. Some classes of DAs conventionally form pairs where the utterance of the first act makes the second act relevant. Such related act pairs are called *adjacency pairs*, AP (e.g. question – answer). In our typology, the acts are divided into AP acts and non-AP acts. Every AP act has the first pair part and the second pair part. All the DAs are divided into two groups: information acts (questions, directives, etc.) and communication managing acts (conventional acts, feedback, etc.). The communication managing acts are divided into two sub-groups: the acts managing fluent conversation, and the acts for solving communication problems (so-called repair acts). The overall number of DAs is 126, the full typology can be found in [10].

Names of DAs consist of two parts: the first two letters give abbreviation of the name of act-group, e.g. DI directives; the third letter is used only for AP acts – the first (F) or second (S) part of an AP act; 2) full name of the act. In the following dialogue example (1), the participants are client (A) who calls information service, and information provider (B) who answers the call. Transcription of CA is used; DAs are annotated by using our typology.

(1) A: ((summons)) RIF: Summons
B: ülikooli=info RIS: Answer | RS: Introduce
university information
A: .hh tere, RIF: GREETING
good morning
ma paluks eesti keel võõrkeelena õppetooli [`telefoni.] DIF: Request please give me the phone number of the Chair of Estonian as foreign language
B: [{-}] seitse `viis kaks=kaks seitse DIS: Giving information
seven five two two seven
A: * kaks kaks seitse, * VR: Acknowledgment: Neutral
two two seven
[ai]täh RIF: Thanking
thank you
B: palun? RIS: Please
you are welcome

Our typology does not provide multiple annotation levels like DAMSL [1] does. Anyway, the levels can be differentiated indirectly. For example, the first pair parts of APs correspond to forward-looking functions and the second pair parts to backward looking functions in DAMSL. On the other hand, our typology is more detailed in some categories, e.g. it includes all together 29 conventional acts (two in DAMSL – opening and closing) and six questioning acts (info-request in DAMSL).

Dialogue acts have been annotated in human-human dialogues several years ago. Two linguists annotated the same texts separately, by using software that simplifies the selection of texts from a corpus and DAs from a menu, and then a third person disambiguated the annotations. The kappa coefficient has been 0.73 [11]. At present, software of semi-automatic annotation of DAs can be used. It takes as input the transcript of a recording, divides it into turns and utterances, assigns up to five DA tags to every utterance using Naïve Bayes classifier, and then a human annotator makes a choice between the tags or assigns (one or more) new tags to an utterance [12].

The EDiC contains more than 500 directory inquiries. For this study, we have occasionally chosen 50 directory inquiries (consisting of 1131 utterances, 6134 words, 1202 dialogue act tags). The Workbench of Estonian Dialogues [13] was used for calculations. We analyse the dialogues in order to find out cues and to formulate rules for automatic recognition of different parts of directory inquiries.

2. Corpus Analysis

2.1. Directory Inquiries

A typical call consists of three parts: a conventional beginning, main information part, and a conventional ending. The kernel of the information part is an AP directive – grant or question – answer. Sub-dialogues can occur after a request (question) and/or answer: an adjusting/specifying question is asked and answered, or a repair for solving a communication problem is initiated. The kernel can be repeated – more than one question can be asked and answered.

The corpus analysis suggests use adjacency pairs of DAs as the main cues for recognition of different parts and sub-dialogues of a directory inquiry.

The conventional opening and closing parts can be recognized looking for APs of rituals (and the single conventional act RS: Introduce) at the beginning or at the end of a dialogue, respectively (Ex. 1).

The main part of a call begins with a caller's (participant A) request (in 80% of cases, Ex. 1) or question (Ex. 2).

(2)

A: palun `öelge mulle kas `teie ütlete ka era`isikute korteri`telefone.

 QUF: Closed Yes/No

 please tell me do you give private persons' phone numbers as well

 B: ei,
 QUS: No

 no

 era'isikuid annab ainult tasuline `infoliin.
 AI: Adjusting

 only information service for fee gives private persons

Different question types are used by A to start the main part of a call: closed yes/no

(Ex. 2), wh-question, open yes/no, or alternative question. In the simplest case, the main part of a dialogue consists of only one AP: request – simple information (Translated dialogue consists).

giving information (Ex. 1); 16% of the analysed dialogues correspond to the simplest pattern.

However, the main part often includes sub-dialogues. Two types of sub-dialogues occur which can be initiated both by A or B – information sharing and (other-initiated) repair.

2.2. Information Sharing

Client's (participant A) request can be adjusted by the information provider (B) before answering. In such a case, an information sharing sub-dialogue will be initiated by B (Ex. 3; sub-dialogue is marked by --> and <--).

```
(3)
A: tere RIS: Greeting
hi
ma `paluksin Vanemuise: kassa `telefon. DIF: Request
I'd like to ask the phone number of the Vanemuine ticket office
(...)
---> B: suure või `väikese maja. QUF: Alternative | ACF: Adjusting the
conditions of answer
of the large or small building
<--- A: suure. QUS: Alternative: One
of the large one</pre>
```

Information-sharing sub-dialogues initiated by B were found in 30% of the analysed dialogues (Figure 1). The greatest number of sub-dialogues is three (in one dialogue, where A asks for the phone number of a military base but B offers him the number of a person on duty instead). An information sharing sub-dialogue can be recognized on the basis of the AP formed by the DAs ACF: Adjusting the conditions of answer and ACS: Adjusting the conditions of answer.



Figure 1. Sub-dialogues: information sharing initiated by B

Similarly, A can initiate an information sharing sub-dialogue if the answer was insufficient. No clear cues have been found for recognition of these sub-dialogues but typically, such a sub-dialogue begins with A's question and ends with B's answer (Ex. 4; sub-dialogue is marked by --> and <--).

(4)

--> A: aga kui `tasulise infoliinil ei `ole seda numbrit, kust ma sis võiksin veel seda `teada saada. QUF: Wh-question

```
but if the service for fee does not have the number then where could I get it from <-- B: no siis ilmselt `rohkem võimalusi ei `ole. QUS: Giving information obviously, there are no more options then
```

A initiates much less information sharing sub-dialogues than B – in 14% of the analysed dialogues. This means that A has got sufficient information in most cases. A needs to adjust B's answer mainly in such a case if it is negative, i.e. if B is not able to give the needed information.

2.3. Other-Initiated Repair

Conversation never runs smoothly. Troubles are usual (non-hearing, non-understanding, misspeaking, etc.). The hearer, having a problem with the last turn of the speaker, initiates a repair sequence. The partner carries out a repair. Our typology differentiates between three types of repair initiations: checking, non-understanding and reformulation (Ex. 5; B initiates three sub-dialogues, which are marked by --> and <--).

(5)

A: palun 'öelge mulle Tallinasu Tallinas asuva: aktsiaselts (.) 'Meriner (.) ee 'telefoninumber. DIF: Request please tell me the phone number of the company Meriner in Tallinn --> B: firma=nimi=oli QUF: Wh-question | RPF: Nonunderstanding the company's name was

```
<-- A: `Meriner.
                       QUS: Giving information
                                                       RPS: Repair
Meriner
--> B: Meriner.
                      QUF: Offering answer
                                                    RPF: Checking
Meriner
<-- A: jah.
                       QUS: Yes | RPS: Repair
yes
(2.0)
--> B: kas kõik on eesti `tähed seal. OUF: Closed Yes/No | RPF:
Reformulation
are Estonian letters all
<-- A: jaa.
                       QUS: Yes | RPS: Repair
yes
```

68% of the dialogues include repairs. The biggest number of repair sub-dialogues is seven (in one dialogue, where *B* initiates repairs because *A* can not formulate his request exactly, and *A* initiates repairs due the poor audibility). In most cases, a dialogue includes one single other-initiated repair (Figure 2).



Figure 2. Sub-dialogues: other-initiated repair

Regardless both of A and B can initiate repair, the repairs are mostly initiated by A considering the information given by B. The most frequent repair initiation is checking (Figure 3). A checks the received phone numbers and addresses, which must be exact.



Figure 3. Repair initiation by A and B

Repair sequences can be recognized on the basis of APs of DAs RPF: Checking/ Non-understanding/ Reformulation and RIS: Repair. In limited cases (17% of dialogues), repair evaluation (non-AP act VR: REPAIR EVALUATION) follows to the second pair part RIS: Repair. Sub-dialogues are marked by --> and <--.

84% of the analysed dialogues contain sub-dialogues. 14% of the dialogues include both information sharing and repair.

3. The Dialogue Structure

After distilling [14] some DAs we get the following general structure of a directory inquiry. Here the DA tags correspond to the used typology; '/' is used for alternatives, and (...) for the DAs that can be missed. Sub-dialogues are marked by --> and <--.

```
[Opening]
A: ((summons))
                   RIF: Summons
  RIS: Answer
                     RS: Introduce [B introduces the service company]
R·
   (RS: Introduce [B introduces himself/herself])
   ( RIF: Greeting
                      )
A: RIS: Greeting
[Main part]
A:
    DIF: Request / QUF: Wh-question/ Open yes-no
    ( [information sharing initiated by B]
--> B: ACF: Adjusting the conditions of answer
<--A: ACS: Adjusting the conditions of answer
    ( [other-initiated repair]
--> B/A: RPF: Reformulation/ Checking/ Non-understanding
   A/B: RPS: Repair
```

```
<-- (B/A: VR: Repair evaluation )
      )
          VR: Acknowledgment: neutral QUS/DIS: Deferral )
   B: (
DIS/QUS: Giving information
       ( [information sharing initiated by A]
   --> A: QUF: Closed Yes-No/ Offering answer/ Wh-question
   <-- B: QUS: Yes/ Giving information
       )
       ( [other-initiated repair]
   --> A/B: RPF: Checking/ Non-understanding/ Reformulation
      B/A: RPS: Repair
   <-- ( A/B: VR: Repair evaluation )
   ( A: VR: Acknowledgment: neutral/ Bounder: neutral/ Change of
state: neutral
               )
```

[Closing] A: RIF: Thanking (RIF: Greeting) (B: RIS: Please RIS: Greeting)

The general structure also visualizes (in form of sequences of DAs) the strategies used by participants for requesting, giving, receiving and specifying information.

4. Conclusion

The dialogue structure is needed when implementing human-computer systems, which interact with a user in natural language following rules of human-human communication. The dialogue manager of a dialogue system has to be informed in the structure of communication in order to understand the user's utterances and to determine its own responses. On the other hand, automatic recognition of the dialogue structure can be considered as a separate task of automatic analysis of a coherent text and can be implemented in various applications, e.g. text summarization, information retrieval, etc.

Implementation of the results of the current analysis is in progress.

Acknowledgements

This work is supported by the European Regional Development Fund through the Estonian Centre of Excellence in Computer Science (EXCS), the Estonian Science Foundation (grants 9124 and 8558), and the Estonian Ministry of Education and Research (grants SF0180078s08 and EKT11005).

References

- J. Allen, M. Core, DAMSL: Dialogue Act Markup in Several Layers (Draft 2.1). Technical Report. University of Rochester, Rochester, NY, 1997.
- [2] J. Carletta, A. Isard, S. Isard, J. Kowtko, G. Doherty-Sneddon, and A.Anderson, The reliability of a dialogue structure coding scheme. *Computational Linguistics*, 23 (1997), 13–31.
- [3] H. Bunt, Dynamic Interpretation and Dialogue *Theory. The Structure of Multimodal Dialogue II.* Ed. by M.M. Taylor, F. Néel, and D.G. Bouwhuis. John Benjamins Publishing Company, Philadelphia/Amsterdam, 1999, 139–166.
- [4] H. Bunt, J. Alexandersson, J. Carletta, J.-W. Choe, A. Chengyu Fang, K. Hasida, K. Lee, V. Petukhova, A. Popescu-Belis, L. Romary, C. Soria, and D.R. Traum, ISO 24617-2: A semantically-based standard for dialogue annotation. *Proc. of LREC-2012*, European Language Resources Association (ELRA), Istanbul, Turkey, 2012, 430–437.
- [5] D. Field, S. Worgan, N. Webb, M. Hepple, and Y. Wilks, Automatic Induction of Dialogue Structure from the Companions Dialogue Corpus, *Proc. of the 4th International Workshop on Human-Computer Conversation*, Bellagio, Italy, October 2008.
- [6] T. Hennoste, O. Gerassimenko, R. Kasterpalu, M. Koit, A. Rääbis, K. Strandson, and M. Valdisoo, Information-Sharing and Correction in Estonian Information Dialogues: Corpus Analysis. Proc. of Second Baltic Conference on Human Language Technologies, Tallinn, 2005, 249–254.
- [7] O. Gerassimenko, T. Hennoste, M. Koit, A. Rääbis, K. Strandson, M. Valdisoo, and E. Vutt, 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"*. Commission of the Official Language at the Chancellery of the President of Latvia, Riga, 2004, 150–155.
- [8] T. Hennoste, O. Gerassimenko, R. Kasterpalu, M. Koit, A. Rääbis, and K. Strandson, From Human Communication to Intelligent User Interfaces: Corpora of Spoken Estonian, *Proc. of LREC-2008*, European Language Resources Association (ELRA), Marrakech, Morocco, 2008, 2025–2032.
- [9] I. Hutchby, and R. Wooffitt, *Conversation Analysis*. Polity Press, Cambridge, 1998.
- [10] M. Koit, Experiments on Automatic Recognition of Dialogue Acts. Proc. of SPECOM'2009: 13th International Conference Speech and Computer, St. Petersburg, Russia, 2009, 533–538.
- [11] M. Fišel, Machine learning techniques in dialogue act recognition, *Estonian Papers in Applied Linguistics* 3 (2007), 117–134 http://www.rakenduslingvistika.ee/ajakirjad/index.php/aastaraamat/-article/view/ERYa3.08/139
- [12] S. Aller, Dialoogiaktide märgendamine Eesti dialoogikorpuses: ülevaade ressurssidest ja tarkvaraarendus. Magistritöö. Tartu Ülikool, arvutiteaduse instituut, 2012 http://comserv.cs.ut.ee/forms/ati_report/
- [13] M. Treumuth, Eesti dialoogikorpus ja selle töötlemise tarkvara. Magistritöö. Tartu Ülikool, arvutiteaduse instituut, 2004 http://dspace.utlib.ee/dspace/handle/10062/1172
- [14] A. Jönsson, and N. Dahlbäck, Distilling dialogues A method using natural dialogue corpora for dialogue systems development. Proc. of the 6th Applied Natural Language Processing Conference, Seattle, 2000, 44–51.