From Human Communication to Intelligent User Interfaces: Corpora of Spoken Estonian

Tiit Hennoste[*]**, Olga Gerassimenko, Riina Kasterpalu, Mare Koit, Andriela Rääbis, Krista Strandson**

University of Helsinki[*], University of Tartu

J. Liivi 2, 50409 Tartu, Estonia

E-mail: tiit.hennoste@helsinki.fi, {olga.gerassimenko, riina.kasterpalu, mare.koit, andriela.raabis, krista.strandson}@ut.ee

Abstract

We argue for the necessity of studying human-human spoken conversations of various kinds in order to create user interfaces to databases. An efficient user interface benefits from a well-organized corpus that can be used for investigating the strategies people use in conversations in order to be efficient and to handle the spoken communication problems. For modeling the natural behaviour and testing the model we need a dialogue corpus where the roles of participants are close to the roles of the dialogue system and its user. For that reason, we collect and investigate the Corpus of the Spoken Estonian and the Estonian Dialogue Corpus as the sources for human-human interaction investigation. The transcription conventions and annotation typology of spoken human-human dialogues in Estonian are introduced. For creating a user interface the corpus of one institutional conversation type is insufficient, since we need to know what phenomena are inherent for the spoken language in general, what means are used only in certain types of the conversations and what are the differences.

1. Introduction

The most natural way of language use is spoken communication. As the technology is developing, intelligent user interfaces to different databases that enable spoken language input become more sophisticated and more popular. An efficient user interface benefits a lot from a well-organized corpus of spoken language that can be used for investigating the strategies people use in conversations in order to be efficient and to handle the spoken communication problems.

Thus the analysis and modeling of the spoken language requires a corpus. The corpora used for investigating human-human communication and for user interface creation are generally restricted to interaction in the framework of certain tasks which the user interface is designed for. For example, the COCONUT [1] corpus includes computer-mediated human-human dialogues in which two subjects cooperate on buying furniture for a house. The VERBMOBIL [2] corpus includes bilingual situational dialogues recorded with a role-playing manner (schedule arrangement, hotel, sight-seeing). The TRAINS [3] corpus includes problem solving dialogues where one participant plays the role of a user and has a certain task to accomplish, and the other plays the role of the system by acting as a planning assistant.

Our goal is to model natural dialogue on the computer, i.e. the computer as a dialogue participant must follow the norms and rules of human-human communication as much as possible. For modeling the natural behaviour and testing the model we need a dialogue corpus where the roles of participants are close to the roles of the dialogue system and its user.

Studies of spoken language have shown that language usage is variable and some of the variants are rare, they appear only in large corpora. At the same time, the variants are used in certain domains. Therefore, such variants can not be excluded when implementing a user interface. Secondly, the grammar of spoken language is interactional which means that different grammatical constructions are connected with certain communicative functions.

We would like to show that both the interaction studies and interaction modeling could benefit from a large corpus. Such corpus comprises of different varieties of the spoken interaction, so that it is possible to distinguish between the traits inherent for the spoken language in general (comparing it to written language) and the specific traits of varieties on all language layers (lexis, morphology, syntax, semantics, pragmatics). For that reason, we collect and investigate the Corpus of the Spoken Estonian and the Estonian Dialogue Corpus as the sources for human-human communication investigation. We analyse human-human dialogues in order to find out how Estonians communicate with each other, how they express their intentions and understand each other, how the language use depends on Estonian cultural space. A large national spoken language corpus makes it possible to provide such background studies.

We investigate the corpora using quantitative analysis and especially qualitative microanalysis in order to find out the conditions that supervise principles of language use.

---

[1] http://www.pitt.edu/~coconut/

[2] http://verbmobil.dfki.de/

[3] http://www.cs.rochester.edu/research/trains/

2. The Corpus of Spoken Estonian

The corpus of spoken Estonian is planned as an open corpus. Our corpus is divided by the five dimensions that influence the language use (Hennoste, 2000):

- social and dialectical background of the interactants
- dialogue *vs* monologue
- the degree of spontaneity of speech
- the closeness of contact between participants (immediate, telephone or mass-media)
- the degree of casuality (institutionality) with four sub-boundaries: relations between participants (familiar *vs* unknown); roles of the participants in the situation (private person *vs* representative of an institution); physical setting (private room *vs* official room); main purpose of the interaction (participation *vs* information).

The corpus includes mainly audio recordings. Currently the corpus comprises of 1777 transcribed texts, or 1,171,817 text units (words, vocal units, e.g. pause fillers or UM-s in our terminology, pauses) in total. Typical stretches of transcribed everyday conversation and longer institutional dialogues are about five to fifteen minutes long. Shorter institutional dialogues and phone conversations have been transcribed fully. The proportions of the corpus are as follows:

- telephone conversations 63% (1116):
  159 private calls
  953 institutional calls: 466 directory inquiries, 93 travel agency information requests, 87 outpatients' department, 66 services (post office, car workshops etc.), 45 telesales conversations, 32 colleagues dialogues, 24 shopping information, 23 taxi calls, 16 bus transport information, 109 other conversations (incl. 4 false connections)

- face-to-face conversations 29% (521):
  164 everyday conversations
  357 institutional dialogues or monologues: 101 shop dialogues, 29 service dialogues (post office, library, shoemaker, hairdresser etc.)**,** 20 conversations between strangers on the streets**,** 17 doctor-patient encounters**,** 24 interviews**,** 15 travel agency dialogues, 12 classroom interactions**,** 11 meetings**,** 25 conference presentations**,** 16 lectures**,** 10 sermons**,** 77 other conversations

- media broadcasts (TV and radio) 8% (140): 42 TV, 98 radio.

The corpus is transcribed using a transcription system of the conversational analysis (CA, see Hutcby & Fooffitt, 1998: 77-92) so that the categories crucial from the interactional point of view are used in the transcription (Hennoste, 2000). The central categories are (transcription symbols see Appendix):

- turn-construction units (TCU, utterances in our terminology), which end in a potential transition relevance places. The main criterion for boundaries of the utterance is intonation (Hennoste & Rääbis 2004: 27-30; cf. Ford & Thompson, 1996)
- words and different vocal units (e.g. UM-s) which have different functions in interaction
- pauses/intervals between the words, utterances or turns
- prosodic and paralinguistic features of the speech (intonation, stress, tempo, drawling etc.)
- overlapping speech and latching.

Example 1 demonstrates the use of transcription symbols (C – caller, A – answerer).

```
(1)
(ring)
A:  jaa?
    Yes?
C:  halloo? (.) ´magasid=vä.
    Hello? Were you sleeping?
A:  ei ´maganud. natuke ´tegin ´tööd.
    No, I wasn't. I was working a little.
C:  aa. (0.5) ee kuida teil sis ´läheb, ma
    ´emmet ´läbi akna ´nägin, aga mai=saand
    mai=saand  tema  juurde  ´minna  sest
    ma=n=´üksinda=hh.
    Oh. How are you then, I saw mummy through
    the window, but I couldn't go to see her
    because I was alone.
A:  .hh nt-t ta ´läheb. (.) ´issi on jälle
    ´käele haiget teind=ja
    So-so. Daddy has hurt his hand again.
    (0.5)
C:  ´maal jälle=vä.=
    In the country again?
A:  =jaa.
    Yes.
```

Each transcription is provided with a header that lists 44 situational factors that have been found to affect language use in the analysis of various languages (Hennoste, 2000). The main domains in the description are:

0. technical information about recording
1. situation (time, place, sphere of communication, norms of interaction, etc.)
2. participants (names, social characteristics, relations between participants in the situation, etc.)
3. topic and subject of conversation
4. text type and genre, planned or unplanned text
5. language, dialect or register
6. additional information.

Example 2 represents the header of the conversation (1).

```
(2)
0. Technical information about recording
0.1. Tape recorder: Grundig BB 390
```

```
Tape: BASF FE I ferro extra 90 min
0.2. Recorder: Liia Kusmin
Transcriber: Andriela Rääbis 1998, 2000.
Transcription grade: 3
0.3. Part of the tape transcribed: whole
conversation
1. Situation
Telephone conversation
1.1. Time: 16.03.1998
Caller
Town: Tartu
home/office
table phone/mobile phone

Answerer
Town: Tartu
home/office
table phone/mobile phone

1.6. Specification of situation: private
(answerer)/ public (caller) place, private/
institutional conversation

2. Participants
2.1. Names: Anni (caller), Liia (answerer)
bystanders: –
2.2. Characteristics of participants:
Anni:
gender: female
age or date of birth: 37
education: higher
nationality: Estonian
home place: Tartu
social status (worker, farmer, serviceman,
intellectual, businessman, retired, housewife,
unemployed, official, student etc.): lawyer
disabilities, speech defects: –

Liia:
gender: female
age or date of birth: 23
education: higher
nationality: Estonian
home place: Tartu
social status (worker, farmer, serviceman,
intellectual, businessman, retired, housewife,
unemployed, official, student etc.): teacher
disabilities, speech defects: –

2.5 Characteristics of participants at the moment:
Anni is tired.

2.7. Relations between participants: strangers/
acquaintances/ intimates
status: equal/ lower/ higher

3. Theme and subject of conversation: daddy has
hurt his hand; birthday; renovation of the
bathroom.

4. Text type and genre
dialogue/ monologue/ polylogue
planned/ unplanned text
```

3. Estonian Dialogue Corpus

The biggest part of the Estonian Dialogue Corpus (EDiC)[4] is formed by dialogues taken from the Corpus of Spoken Estonian – currently 945 calls for information (directory inquiries, calls to travel agencies, etc.) and 116 face-to-face conversations, altogether 1061 transcribed texts with a total length of 178,100 running words.

The second part of the EDiC contains 22 written information dialogues (2500 running words) which have been collected in computer simulations using the Wizard of Oz method.

The third part (human-computer interactions) is collected with two simple dialogue systems – "Travel agent" that gives information about flights departing from Tallinn Airport, and "Theatre agent"[5] that gives information about theatre programmes in Estonia.

The EDiC is annotated on three levels: morphological, syntactic (both partially) and dialogue acts level.

3.1. Morphological Annotation

For morphological analysis, the analyzer ESTMORF (Kaalep, 1997) created for written Estonian was adapted for spoken Estonian[6]. The Estonian language has rich morphology, syntactic and pragmatic relations in utterances are often expressed by means of morphology (Fig. 1, cf. Gerassimenko et al., 2004).

```
A
<s>
teated   information
    teade+d //_S_ com pl nom //
tere good morning
    - tere+0 //_B_ //
    tere+0 //_S_ com sg gen //
    tere+0 //_S_ com sg nom //
</s>

C
<s>
tere good morning
  - tere+0 //_B_ //
    tere+0 //_S_ com sg gen //
    tere+0 //_S_ com sg nom //
ma    I
    mina+0 //_P_ pers ps1 sg nom //
sooviksin   should like
    soovi+ksin //_V_ main cond pres ps1 sg ps
af //
teada  to know
    tead+a //_V_ main inf //
```

Figure 1: Morphological analysis of the dialogue
```
A: teated /information, tere / good morning
C: tere / good morning, ma sooviksin teada ... /I'd
like to know ...
```

3.2. Syntactic Annotation

For syntactic analysis the Constraint Grammar analyzer created for written Estonian was adapted for spoken language (Fig. 2, cf. Müürisep & Nigol, 2007).

```
Se # this
    see+0 //_P_ dem sg nom // **CLB @NN>
veranda # veranda
    veranda+0 //_S_ com sg nom // @SUBJ
on # is
    ole+0 //_V_ main indic pres ps3 sg //@+FMV
minu # my
    mina+0 //_P_ pers ps1 sg gen // @P>
meelest # opinion
    meelest+0 //_K_ post #gen // @ADVL
maailma # world's
    maa_ilm+0 //_S_ com sg gen // @NN>
kihvtim # coolest
    kihvti=m+0 //_A_ comp sg nom // @AN>
asi # thing
    asi+0 //_S_ com sg nom // @PRD
$.
. //_Z_ Fst //
```

Figure 2: Syntactic analysis of the utterance
```
Se veranda on minu meelest maailma kihvtim asi. /
In my opinion, this veranda is the coolest thing in
the world.
```

3.3. Dialogue Act Annotation

3.3.1. Basics of Dialogue Act Typology

When communicating, people perform actions using language – they are asking, answering, etc. Such actions have been called *communication acts*, *dialogue acts* etc., and they have been determined as functions which are connected with a speaker's intentions. Several researchers have considered practical problems of determining dialogue acts in the last decade – corpus linguists, discourse analysts, language technologists (e.g. Allwood et al., 2001; Stolcke et al., 2000; Jokinen et al., 2001).

In our corpus, dialogue acts are annotated by using a special annotation typology elaborated at the University of Tartu (Ex 3, Gerassimenko et al., 2004). Our typology is general, not domain- or problem-oriented and can be used for annotating both spoken human-human and human-machine dialogues. Still, the major part of our typology coincidences with well-known typologies (e.g. DAMSL, SWBD-DAMSL).

Our dialogue act typology is based on the principles of organization of conversation borrowed from CA. CA is based on an empirical, inductive microanalysis of conversation data (see e.g. Hutchby & Wooffitt, 1998). The main idea behind the analysis is that conversation is the collaboration of participants based on three mechanisms: turn taking organization (turn construction and turn distribution), repair organization, and sequence organization (preference organization and adjacency pairs). CA considers turns and does not consider dialogue acts or dialogue act typology. However, a turn can consist of different acts. For that reason, we prefer to consider dialogue acts. The principles of our dialogue act typology are as follows (Hennoste & Rääbis, 2004: 15-37).

1. We differentiate acceptable units (which can be analysed) from the non-acceptable units (as DAMSL does). Non-acceptable are only technically un-interpretable utterances in our typology (and not e.g. unfinished utterances).

2. 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. Therefore the analysis of relations between two turns is central. Some classes of dialogue acts conventionally form pairs where the first act makes the second act relevant. Such related act pairs are called *adjacency pairs,* AP (e.g. greeting – greeting, question – answer). The computer must be able to differentiate the first part of an AP (which is expecting a reaction) from the acts that do not expect any reaction as well as understand whether the following act is expected or not.

In our typology, the acts are divided into AP acts and non-AP acts. Every AP act has the first pair part and second pair part (cf. forward-looking and backward-looking functions in DAMSL).

3. The acts used in dialogue are typically divided into two groups – information acts (questions etc.) and communication managing acts (rituals, feedback etc.). The studies have shown that human-human communication is never completely fluent. Therefore, a mechanism is needed to signal and solve different interactional, grammatical and semantic problems. All dialogue act typologies include problem solving acts but typically they belong to different dialogue management or feedback acts and do not form a whole sub-system (cf. Allwood et al., 1992; Bunt, 1999). For example, in DAMSL one repair act represents communication status (*Abandoned*) but most of the them belong to the group of backward-looking acts (*Signal-non-understanding, Completion, Correct-misspeaking, Repeat-rephrase*).

CA assumes that there exists a problem solving mechanism with its own rules, which is called repair organization. CA brings out four types of the repair mechanism (Schegloff, 1979): *self-initiated self-repair*, *other-initiated other-repair, other-initiated self-repair*, *self-initiated other-repair*. The first two are initiated and made by the same person, the others by different persons. It is important to note that in many cases similar means are used in making up of information acts as well as repair acts. For example, most of the other-initiated self-repairs are questions. The computer has to understand when the question is about information and when it initiates repair.

Therefore, we have divided the communication managing acts into two sub-groups: the acts managing fluent conversation, or dialogue managing acts, and the acts for solving communication problems, or repair acts.

4. A turn can consist of several acts. Acts of the turn can be related with the acts in previous or the following turn but acts of one turn can be related one with another as well. We bring separately out such acts that are related with previous acts of the same speaker in the same turn.

5. Our typology is empirical and open. It is based on the presupposition that dialogue acts are empirical phenomena and it is impossible to predefine theoretically all the acts. For that reason, every act group includes a sub-group 'other'. This sub-group includes the acts that are not determined in the current typology. If needed, a new act group will be defined on the basis of this sub-group.

Summing up, we can say that there are two main classification principles of dialogue acts in our typology. First, acts are divided into single acts and adjacency pair acts where the first pair part makes the second pair part relevant. Secondly, the acts are divided into three groups: information acts, the acts managing fluent conversation, and the acts for solving communication problems.

There are 12 groups of acts in our typology, the overall number of dialogue acts is 127.

3.3.2. Overview of Typology

I. AP acts
DIALOGUE MANAGING ACTS
1. Conventional acts (greeting, thanking, etc.) are linguistically formulaic expressions which can be presented as lists.
2. Topic change acts are used to start a new topic or sub-topic.

REPAIR ACTS
3. Repairs initiated and made by different participants.
4. Contact control acts. The speaker checks the functioning of the communication channel (*do you hear, hallo*). These acts typically occur in phone conversations and are formulaic expressions which can be presented as lists.

INFORMATION ACTS
5. Directives and reactions (request, proposal, offer, etc.).
6. Questions and answers.
7. Opinions and reactions (assertion, etc.).

II. Non-AP acts
DIALOGUE MANAGING ACTS
1. Conventional (contact, call, etc.).
REPAIR ACTS
2. Repairs initiated and made by the same person.

INFORMATION ACTS
3. Primary single acts (narration, promise, rhetorical question, etc.).

4. Additional information (specification, softening, etc.) – by such acts a speaker adds some new information to the information act in the same turn, e.g. (s)he answers a question but gives some additional information which was not requested.
5. Responses (continuer, acknowledgement, etc.) – acts that traditionally are considered as narrow feedback. Using such acts, the hearer reacts voluntarily to the partner's previous turn.

Names of the dialogue acts consist of two parts separated by a colon: the first two letters give abbreviation of the name of an act-group, e.g. QU – questions, VR – voluntary responses. The third letter is used only for AP acts – the first (F) or the second (S) part of an AP act; 2) full name of the act, e.g. QUF: WH-QUESTION, QUS: GIVING INFORMATION, VR: CONTINUER (Ex 3).

Dialogue acts have been annotated manually using software that simplifies the selection of texts from a corpus and dialogue acts from a list. Two linguists annotate the same text separately and then a third person disambiguates the annotations. For each utterance, a qualitative microanalysis is performed. The analysis is based on CA and interactional linguistics.

```
(3)
C:  mt=.hh tere,                RIS: GREETING
    Hello.
    öelge=palun: `pensioniameti `telefoni
    (.) .h `number (.) `Tartus. DIF: REQUEST
    Please give me the phone number of the
    pension department in Tartu.
    (...)
A:  ee `number on `seitse=neli=`neli?
                        DIS: GIVING INFORMATION
    The number is seven four four.
    (0.5)
C:  jah?=               VR: NEUTRAL CONTINUER
    Yes?
A:  =seitse ´neli ´kolm ´kuus.
                        DIS: GIVING INFORMATION
    Seven four three six.
    (2.5)
C:  aitäh?                      RIF: THANKING
    Thanks.
A:  palun                       RIS: PLEASE
    You're welcome.
```

4. Implementation: an Example

In this section, we consider an example which demonstrates that directives and questions are differently used in different conditions. Our analysis is based on the EDiC.

4.1. Directives and Questions

Directives and questions are strongly related act groups. Some of dialogue act typologies consider them as a single group. For example, DAMSL has an act category *Info-requests* which includes the acts that set an

obligation for the hearer to give information. In our opinion, such approach is too general and does not differentiate the different uses of the dialogue acts. Sometimes directives and questions are differentiated on the basis whether the user does need some information (question) or (s)he wants to influence the hearer's future non-communicative actions (directive). We claim that it is not important for dialogue continuation whether the hearer must to do something outside of current dialogue or not. (S)he must react both to a question and a directive because both are the first pair parts of APs.

We differentiate directives and questions on the basis of their form. Questions are the info-requests which have specific formal features (interrogatives, specific word order etc.). Questions can be further classified on the basis of the expected reaction: 1) wh-questions, 2) alternative questions, 3) closed yes/no questions, 4) open yes/no questions, 5) yes/no questions offering answer.

Closed yes/no question expects an answer *yes* or *no*. Open yes/no question is expressed in Estonian using the same means as closed yes/no but the expected answer is giving information like in case of wh-questions (*kas te saaksite mulle öelda X telefoni?/ could you give me the phone of X?*). Yes/no question offering answer, includes a presumption of a true answer (*pluss maksud, jah?/ taxes added, yes?*). Therefore, both directives and questions are the first parts of APs but they have different linguistic

forms (Hennoste et al., 2005; Hennoste et al., 2006).

4.2. Analysis Using Corpus

For this paper, we have analysed initial info-requests of clients who are calling a service provider. Three situation types are represented in selected sub-corpus: directory inquiries, calls to outpatients' offices and taxi requests (Table 1).

In our data, requests are formulated in four possible linguistic ways: directive (almost exclusively with verb in conditional), wh-question, open yes/no question and closed yes/no question. Closed yes/no questions expect an answer *yes* or *no*. The other utterances expect giving information or an action of a service provider.

Table 1 shows that most of requests are directives and that directives and questions are used with different frequency in different situations. The qualitative microanalysis demonstrates that directives and questions are used in different preconditions. Let us compare the use of directives (DIF: REQUEST) and open yes/no questions (QUF: OPEN YES/NO).

The **taxi requests** (ordering a taxi) are formulated mainly as directives (Ex 4).

Situation type	# dialogues	Caller's initial requests (%)				
		Directives	Open yes/no questions	Wh-questions	Closed yes/no questions	Other dialogue acts
Calls to outpatients' office	26	50%	31%	4%	4%	11%
Taxi requests	22	77%	13%	–	5%	5%
Directory inquiries	60	62%	17%	21%	–	–
Total	108	62%	19%	13%	2%	4%

Table 1: Overview of the sub-corpus analysed.

(4) ma palun `taksot `Ringtee `kuuskend kaheksa `bee.   DIF: REQUEST
I would like a taxi to Ringtee sixty eight B.

Open yes/no questions were used only in three cases. In all those cases a caller was not sure whether the request can be granted or not because it was untypical (two cars or an untypical car were requested, Ex 5). We can say that callers claim their entitlement to having their request granted (cf. Curl & Drew forthcoming).

(5) 'on teil 'kahte autot 'Lossi 'kolmteist saata   QUF: OPEN YES/NO
Can you send two cars to Lossi 13.

The **calls to outpatients' offices** include different

requests. In addition to booking an appointment with a doctor (which is the most frequent one), callers also make inquiries of possible benefits (Ex 6).

(6) ma sooviks doktor 'Vaheri juurde 'aega.
                    DIF: REQUEST
I'd like to book an appointment with doctor Vaher.

Our analysis shows that a caller uses mainly a directive for booking an appointment with a doctor. But unlike in taxi calls, there may be some contingencies associated with the request being granted (the patient may not belong to the doctor's list etc.). Still, callers use mainly a directive as they claim their entitlement to make the request. Contrary to that, questions are used when a caller

is not sure whether his request can be granted (e.g. because of the limits of chargeless care funded by sick fund may be already reached the end, Ex 7).

(7) kas `teie=juurde  `lapsi saab ka
regist`reerida=vel `vana aasta sees=hh.
                                QUF: OPEN YES/NO
Do you book appointments for children before the end of the year.

Therefore, if a caller doubts the receptionist's ability to grant the request then (s)he chooses the format of a question and does not use a directive.

The calls to the **directory inquiries** are different in two ways both from taxi calls and calls to outpatients' offices. First, the caller wants to get information, not to elicit an action of an official. Secondly, there are different types of information to inquire of. In our data, we have inquiries mostly of phone numbers, which are asked in 45 cases (75%), there are fewer inquiries of addresses, opening hours of institutions, fields of activity of firms, etc.

Almost all the inquiries of phone numbers are clearly formulated, an institution is specified exactly. The linguistic format is a directive (Ex 8). Only in few cases the caller was uncertain about the exact address of the company or was not sure about some other facts.

(8)   palun   'Tallinna   'Tõnismäe
'hambapolikliinik.   DIF: REQUEST
 Tõnismäe dentist office in Tallinn please.

The second group of inquiries is formed as open yes/no questions. Here we find both requests of phone numbers and of other data. First, there are inquiries of general data (do you have any data about, are there any hours when they answer to the phone calls, where to call if someone has lost his job, etc.). A caller is not sure whether there is any information available, or whether an answer can be given because of request is too vague. Secondly, there are requests for certain special information.

General data and the special information are requested also in a form of a directive. Why a caller uses questions instead of directives? The choice has been typically explained as an act of politeness. However, there are examples to show that there may be other reasons either additionally to politeness influencing the choice or there may be rather different reasons to use a question instead of directive. In the example 9 the caller does not know whether such a number is included in the data base at all.

(9) palun kas teil 'on: 'Vesseli kaupluse
'numbrit 'Elvas.   QUF: OPEN YES/NO
Please do you have the number of the Vessel shop in Elva.

Or the caller may admit that (s)he does not know what

exactly (s)he is about to make inquiry. Or (s)he may lack of the personal experience and expertise in calling to directory inquiries.

Summing up, our analysis demonstrates that there exists no difference between formulating the requests which expect filling a gap in knowledge (e.g. with a phone number) and the requests that expect an action of a receptionist or an operator (e.g. sending a taxi). The analysis has shown that almost all of usages of open yes/no questions can be explained with the uncertainty of a caller about the possibility of having the request granted (the data base may not contain the data; background information given by the caller is too general to perform the search; the inquiry is formulated too vaguely).

There is another significant feature to point at. In all cases, a directive is used for frequent and typical requests (ordering a usual taxi, booking an appointment with a doctor, making inquiry of a phone number).

As mentioned before, a decision in favour of a question sometimes has been explained with politeness. There are different means for expressing politeness in languages, and using a question is only one of them. Our analysis does not confirm the claim that a question form is used for a polite request. There are other means used in our data, e.g. conditional mood, some polite words (e.g. *palun 'please'*). Only 7 requests (5 directory inquiries, 2 taxi calls) do not have these markers.

We claim therefore that a speaker is certain in requesting frequent and typical things and more hesitant in rare cases. Being certain in granting his/her request the caller chooses a directive, and contrary being uncertain the caller chooses the format of a question.

5. Conclusion

We argue for the necessity of studying human-human spoken conversations of various kinds in order to create different user interfaces to databases. We have collected a corpus of spoken Estonian and an Estonian dialogue corpus. We are investigating and comparing various human-human conversations with the aim to create intelligent user interfaces which can respond to a user in a way a human official does.

For creating a user interface the corpus of one institutional conversation type is insufficient, since we need to know what phenomena are inherent for the spoken language in general, what means are used only in certain types of the conversations and what are the differences. One needs to analyse large corpora for different registers in order to explain how and why people use different language means in different situations and for different purposes. We certainly need restricted (sub-)corpora for certain tasks or research areas and they can be easily created from a large corpus.

6. Acknowledgements

7. References

Allwood, J., Ahlsen, E., Björnberg, M., Nivre, J. (2001). Social activity and communication act-related coding. In J. Allwood (Ed.), *Cothenburg Papers in Theoretical Linguistics 85. Dialog Coding – Function and Grammar. Göteborg Coding Schemas*. Goteburg, pp. 1--28.

Bunt, H. (1999). Dynamic Interpretation and Dialogue Theory. In M.M. Taylor, F. Neel & D.G. Bouwhuis (Eds.), *The Structure of Multimodal Dialogue II.* Philadelphia/ Amsterdam: John Benjamins Publishing Company, pp. 139--166.

Curl, T., Drew, P. (forthcoming). Contingency and action: a comparison of two forms of requesting. In *Research on Language and Social Interaction.*

Ford, C., Thompson, S. (1996). Interactional Units in Conversation: Syntactic, Intonational, and Pragmatic Resources for the Management of Turns. In E. Ochs, E. A. Schegloff & S. A. Thompson (Eds.), *Interaction and Grammar,* Cambridge: Cambridge University Press, pp. 134—184.

Gerassimenko, O., Hennoste, T., Koit, M. Rääbis, A., Strandson, K., Valdisoo, M.,Vutt, E. (2004). Annotated Dialogue Corpus as a Language Resource: An Experience of Building the Estonian Dialogue Corpus. In *Proceedings of the First Baltic Conference "Human Language Technologies. The Baltic Perspective".* Riga: Commission of the Official Language at the Chancellery of the President of Latvia, pp. 150--155.

Hennoste, T. (2000). Eesti suulise kõne uurimine: transkriptsioon, taust ja korpus. *Keel ja Kirjandus,* 2, pp. 91--106.

Hennoste, T., Gerassimenko, O., Kasterpalu, R., Koit, M., Rääbis, A., Strandson, K., Valdisoo, M. (2005). Questions in Estonian Information Dialogues: Form and Functions. In V. Matousek, P. Mautner (Eds.), *Proceedings of the 8th International Conference TSD 2005: Text, Speech and Dialogue*. Berlin: Springer, pp 420--427.

Hennoste, T., Gerassimenko, O., Kasterpalu, R., Koit, M., Rääbis, A., Strandson, K., Valdisoo, M. (2006). Cue-based Interpretation of Customer's Requests: Analysis of Estonian Dialogue Corpus. In T. Salakoski, F. Ginter, S. Pyysalo & Tapio Pahikkala (Eds.), *Proceedings of FinTAL 2006: Advances in Natural Language Processing. 5th International Conference on Natural Language Processing.* Turku, Finland: Springer, pp. 206--213.

Hennoste, T.; Rääbis, A. (2004). *Dialoogiaktid eesti infodialoogides: tüpoloogia ja analüüs.* Tartu: Tartu Ülikooli Kirjastus.

Hutchby, I., Wooffitt, R. (1998). *Conversation Analysis. Principles, Practices and Applications.* Polity Press.

Jokinen, K., Hurtig, T., Hynna, K., Kanto, K., Kaipanen, M., Kerminen, A. (2001). Selforganizing dialogue management. In: *Proceedings of the Natural Language Pacific Rim Symposium (NLPRS) Workshop "Neural Networks and Natural Language Processing",* Tokyo, Japan.

Kaalep, H.-J. (1997). An Estonian Morphological Analyser and the Impact of a Corpus on Its Development. In *Computers and Humanities*. Vol. 31 № 2, pp. 115--133.

Müürisep, K., Nigol, H. (2007). Disfluency Detection and Parsing of Transcribed Speech of Estonian. In Z. Vetulani (Ed.), *Proceedings of Human Language Technologies as a Challenge for Computer Science and Linguistics. 3rd Language & Technology Conference.* Poznan, Poland: Fundacja Uniwersitetu im. A. Mickiewicza, pp. 483--487.

Schegloff, E. (1979). The Relevance of Repair to Syntax-for-Conversation. In *Syntax and Semantics 12. Discourse and Syntax.* New York: Academic Press, pp. 261--288.

Stolcke, A., Coccaro, N. Bates, R., Taylor, P., Van Ess-Dykema, C., Ries, K., Shriberg, E., Jurafsky, D., Martin, R., Meteer, M. (2000). Dialogue act modeling for automatic tagging and recognition of conversational speech. In *Computational Linguistics* 26:3, pp. 339--373.

Appendix: Transcription Symbols

Utterances (TCU-s)

.	falling (final) intonation
?	rising intonation
,	half-falling intonation

Pauses/silences

(.)	micro-pause: 0.2 sec or smaller
(…)	longer pause than a micro-pause
(0.8)	timed pause in seconds

Prosodic and paralinguistic phenomena

`	stressed word or syllable
>… <	faster segment
<… >	slower segment
*… *	softer segment
AHA	louder segment
khm	hem
hehe	laughing with open mouth
mhemhe	laughing with closed mouth
s(h)õna	word is voiced laughingly
$...$	segment uttered in a laughing voice
@...@	change in a tone of voice, e.g. imitation
-	cut-off, truncation (word is not finished)
:	drawling , lengthening of a sound
.hhh	audible inbreath
.jaa	word is voiced during inhalation
=h	voiced exhalation (at the end of word)

Overlapping speech and latching

=	latching (no silence between two items)
[	beginning of the overlapping speech
]	end of the overlapping speech

Obscurities and comments

{--}	impossible to hear what was said
(( ))	transcribers comments