

## К ВЗАИМОДЕЙСТВИЮ КОМПЬЮТЕРА И ЧЕЛОВЕКА НА ЕСТЕСТВЕННОМ ЯЗЫКЕ\*

### TOWARDS HUMAN-COMPUTER INTERACTION IN NATURAL LANGUAGE

*Страндсон К. (krista.strandson@ut.ee), Герасименко О. (olga.gerassimenko@ut.ee),  
Кастерпалу Р. (riina.kasterpalu@ut.ee), Койт М. (mare.koit@ut.ee), Рязбис А. (andriela.raabis@ut.ee)*  
Тартуский университет, Эстония

В статье рассмотрена структура естественных справочных диалогов для разработки эстонской диалоговой системы, общающейся с пользователем на естественном языке. На материале Эстонского диалогового корпуса проанализированы языковые характеристики информационных запросов клиента и ответов информатора. Результаты исследования будут использованы в двух диалоговых системах, находящихся в стадии разработки.

#### 1. Введение

Диалоговые корпуса, состоящие из записей естественных разговоров, представляют собой хорошую основу для изучения культурно обусловленных стратегий и методов, используемых собеседниками для достижения целей в разговоре. Наиболее естественным и распространенным способом общения является (устный) диалог. Компьютерная система, претендующая на роль участника диалога, т.е. интеллектуального агента, должна понимать и генерировать единицы естественного языка, а также управлять диалогом наравне с собеседником. Хотя и существует предположение, что люди общаются с компьютерной системой иначе, чем с другим человеком [1], мы убеждены, что основополагающие принципы ведения диалога на естественном языке нужно искать в естественном диалоговом поведении, и лишь затем анализировать специфику общения человека и компьютера.

Существует ряд диалоговых систем на естественном языке, выполняющих информационные запросы пользователя или определенные практические задания [2], [3]. В моделировании диалогового поведения важную роль играют устные диалоговые корпуса (например, в проектах Switchboard, Verbmobil, BNC). Телефонные службы, основанные на устных диалоговых системах, широко внедряются в областях информационного поиска и практических операций, включая системы транспортных расписаний, справочные и технические службы, системы управления интеллектуальным домом и системы навигации.

Наша цель – разработать диалоговую систему (ДС) для устной коммуникации на эстонском языке. Мы изучаем естественные справочные диалоги, чтобы выяснить, как клиент формулирует запрос, какие языковые средства используют участники диалога, какова структура устного институционального диалога и что должна учитывать система, следующая правилам естественной коммуникации. Исследование естественных диалогов важно и для автоматического распознавания речи: описание частотных шаблонов облегчает расшифровку беглой речи. Схожим образом, описание речевого поведения информатора помогает смоделировать естественное, лаконичное и практичное речевое поведение ДС.

На материале русских справочных диалогов структуру и динамику изменения текущего сознания коммуникантов исследовал А.Е. Кибрик [11]; наше исследование сосредотачивается на эксплицитных деталях речевого взаимодействия, однако мы используем близкий подход.

В разделе 2 мы представляем обзор эмпирического материала – Эстонского диалогового корпуса и типологии диалоговых актов, используемой для разметки корпуса. Разделы 3 и 4 посвящены анализу корпуса и соответственно рассмотрению запросов клиента и ответов информатора. В разделе 5 рассмотрены результаты корпусного анализа, которые будут использованы в двух экспериментальных ДС. Раздел 6 содержит выводы.

#### 2. Эмпирический материал

Наше исследование основано на Эстонском диалоговом корпусе (EDiC, <http://math.ut.ee/~koit/Dialog/EDiC.html>). Корпус состоит из записей и расшифрованных текстов более 1100 аутентичных устных диалогов, в т.ч. около 1000 телефонных диалогов. В корпусе размечены диалоговые акты

\* Работу поддерживает Эстонский научный фонд (грант № 7503) и Министерство просвещения Эстонии.

(функции диалогических реплик). Используемая нами типология диалоговых актов основана на конверсационном анализе (КА) [4], который исследует техники, используемые людьми в социальном взаимодействии. Согласно КА, некоторые диалоговые акты образуют смежные пары, в которых появление первого элемента обуславливает появление второго элемента (так, информационный запрос требует ответа). Если второй элемент не следует непосредственно за первым (например, вместо ответа собеседник задает уточняющий вопрос), его ожидание сохраняется на протяжении вставной секвенции. Компьютер как участник диалога должен быть способен отличить первый элемент смежной пары (ожидающий ответа) от второго элемента, а также от реплик, не ожидающих специфических ответных действий (например, реплик обратной связи: *так/ясно*).

Для описания поведения клиента и информатора в информационных диалогах из корпуса были отобраны 60 случайных звонков в справочную службу. Информационные запросы клиента имеют форму вопроса (специального, напр. *Когда идет последний поезд на Тарту?* или общего, напр. *Вы не подскажете мне последний поезд на Тарту?*) или директива (напр., *Подскажите мне последний поезд на Тарту*). Случаи использования общих и специальных вопросов мы рассматривали в ряде предыдущих статей [5], [6]. В этой статье мы рассмотрим смежные пары директивов и реакций на них («директив – выполнение директива»). Директивный запрос выражает желание или потребность говорящего получить некую информацию. Типичными реакциями на запрос являются предоставление информации или сообщение об отсутствии информации (прим. 1, С – клиент, I – информатор; транскрипция КА).

(1)

C: õelge=palun: `pensioniameti `telefoni (.) .h `number (.) `Tartus.

*скажите, пожалуйста, номер телефона пенсионного департамента в Тарту ЗАПРОС*

I: ee `number on `seitse=neli=`neli?

*ээ, номер семь четыре четыре.. ИНФОРМАЦИЯ*

Рассмотрим, какие диалоговые акты используются для оформления запросов и ответов (табл.1).

ЗАПРОСЫ КЛИЕНТА	# ЗАПРОСЫ	ОТВЕТЫ ИНФОРМАТОРА	# ОТВЕТЫ
Запрос	52	Предоставление информации Отсрочка Отсутствие информации	129 31 4
Запрос + Смена темы	2	Предоставление инф.	2
Предоставление инф. + Запрос	1	Предоставление инф.	1
Согласие + Запрос	2	Предоставление инф.	2
Утвердительный ответ + Запрос	1	Предоставление инф.	1
Всего	58		170

Таблица 1. Запросы и ответы в справочных диалогах

Ответных реплик в три раза больше, чем реплик с запросами, так как первой реакцией на запрос обычно является сообщение об отсрочке (информатору нужно время, чтобы свериться с базой данных). Кроме того, информатор часто разбивает ответ на блоки и произносит в несколько реплик (репликой в традиции КА считается поток речи между сменами говорящих (прим. 2), могущий содержать несколько диалоговых актов).

(2)

I: üks hetk.

*один момент (минуточку) ОТСРОЧКА*

(7.5) I: `üld`info, (.) &gt; kinnitamata andmetel &lt; neli kolm viis?

*общая информация, по неподтвержденным данным, четыре три пять.. ИНФОРМАЦИЯ*

(.)

C: jaa?

*да (так?) ОБРАТНАЯ СВЯЗЬ*

I: viis kaheksa,

пять восемь.. ИНФОРМАЦИЯ

C: jaa?

да (так?) ОБРАТНАЯ СВЯЗЬ

I: null null.

ноль ноль. ИНФОРМАЦИЯ

C: < null null. >

ноль ноль. ОБРАТНАЯ СВЯЗЬ

### 3. Запросы клиентов

#### А. Начальный запрос

Задачей информатора является распознать цель пользователя (получить информацию) и помочь ее достичь. Обычно диалог содержит один запрос (2/3 запросов в подкорпусе – начальные), однако возможны и несколько запросов в одном диалоге.

Запрашиваемая информация разнообразна – телефонные номера (75%), адреса, приемные часы учреждений, сфера деятельности фирм. В эстонском языке запрос формулируется при помощи следующих шаблонов:

*palun X* 'пожалуйста (будьте добры), X' или  
*palun öelge (mulle) X* 'пожалуйста, скажите (мне) X' или  
*ma sooviks X* 'я хотел(а) бы X'

где X обозначает пробел в знаниях клиента, который он желает заполнить.

#### В. Последующие запросы

К этой группе относятся 1/3 запросов: как правило, они встречаются в диалогах, в которых информатор не смог удовлетворить первичный запрос клиента (запрошенная информация отсутствует или не может быть представлена сразу) и предложил другую релевантную информацию. Такому предложению обычно предшествует уточняющий поддиалог (3).

(3)

I: .hh ma 'pakuksin selle 'üldise 'info['laua, jah .hh]

я бы предложила общий информационный отдел, да ОТСУТСТВИЕ ИНФ + ПРЕДЛОЖЕНИЕ

C: [jah, {oleks 'hea. jah}]

да, было бы хорошо. да. СОГЛАСИЕ + ЗАПРОС

I: 'number kinnitamata andmetel 'seitse 'kolm, 'neli,

номер, по неподтвержденным данным, семь три четыре.. ИНФОРМАЦИЯ

После ответа на запрос клиент может инициировать уточняющий поддиалог, чтобы запросить добавочную информацию, относящуюся к теме запроса (4): номера телефонов, адреса, имена, сферы деятельности учреждений и т.д.

(4)

C: sooviks=seda (.) 'T:oode numbrit, a- akselts 'Toode.

мне нужен номер Тооде, акционерного общества Тооде ЗАПРОС

/—/

I: 'kaheksa null='null (1.0) 'seitse 'null,

восемь ноль ноль .. семь ноль.. ИНФОРМАЦИЯ

(0.5)

C: \* mhmh \*

угу ОБРАТНАЯ СВЯЗЬ

I: null='null.

ноль ноль ИНФОРМАЦИЯ

C: null='null.

ноль ноль ОБРАТНАЯ СВЯЗЬ

C: ja 'kus ta 'asub.

и где они находятся? УТОЧНЯЮЩИЙ ВОПРОС

I: .hh 'Tähe 'sada kuus 'teist,

Тяхе сто шестнадцать ИНФОРМАЦИЯ

С. Лингвистические характеристики запроса

Для автоматического распознавания диалоговых актов в нашем корпусе использовались статистические методы [7]. Точность распознавания была невысока (в среднем, около 50%), что объясняется неоднородностью корпуса, содержащего диалоги различных ситуационных типов (звонки в регистратуру, турагентства, транспортные и страховые службы), подробностью типологии, содержащей 127 диалоговых актов, и полифункциональностью речевых шаблонов. Разнообразный и подробно размеченный материал ценен для исследования естественного человеческого общения в различных ситуациях, но неудобен для тренировки статистических методов. Точность распознавания можно улучшить добавлением лингвистических правил, лексических и синтаксических ключевых сигналов, характеризующих директивные запросы.

Для формулирования запросов используется ограниченное число шаблонов. Большинство запросов включают глагол-сказуемое (12 различных глаголов): наиболее распространены глагол *paluma* (*просить*, в 1 л. ед.ч. *пожалуйста*), *üttelema* (*сказать*) и *soovima* (*желать*, табл. 2).

ГЛАГОЛ	НАКЛОНЕНИЕ (#)			ВСЕГО
	инд.	конд.	имп.	
<i>paluma</i> 'просить' (в 1 л. ед.ч. «пожалуйста»)	6	8		14
<i>üttelema</i> 'сказать'			8	8
<i>soovima</i> 'желать' («я хочу/хотел бы»)	2	5		7
<i>tahtma</i> 'хотеть'		5		5
<i>võtma</i> 'брать' («дайте»)		5		5

Таблица 2. Наиболее частотные глаголы в запросах клиентов.

Глаголы-сказуемые часто используются в определенном наклонении и лице [5], как ключевые слова. Глаголы можно разделить на две группы: использующиеся в императиве (*öelge* 'скажите', *andke* 'дайте') и использующиеся в первом лице индикатива или кондиционалиса (*palun* 'прошу' «пожалуйста», *soovin/tahan* 'желаю/хочу' «я хотел бы», *võtaksin* 'я бы взял(а)' «дайте мне»). Индикатив, универсальная форма декларативных актов, используется в нашем корпусе лишь в случае устойчивого выражения *palun* ('прошу' «пожалуйста») для смягчения прямых директивных высказываний. Глагол *просить* может также использоваться в 1 лице кондиционалиса для усиления вежливости [8]. Признак кондиционалиса *-ks-* может служить надежным сигналом для автоматического распознавания.

Другим сигналом служит позиция глагола в реплике. Глагол начинает реплику в 22 случаях (38%), в других случаях реплика начинается с местоимения *mina/ma* 'я' (16 случаев, 28%) или с дискурсивного маркера (6 случаев): *et* 'что' «так, то есть», *aga* 'а', *äkki* 'вдруг' «может быть». Глагол в этих случаях занимает вторую позицию.

Последующие запросы, в отличие от начальных, обычно начинаются с союза *aga* 'а' или *ja* 'и' или включают частицу *ka* 'тоже' или *veel* 'еще'. Эти слова, указывающие на продолжение, не могут быть использованы в начальных запросах (5) (6). Глагол *võtma* 'брать' в кондиционалисе встречается только в последующих запросах.

(5)

С: *aga* 'võtaks Lelula 'numbri.

С: а (я) взяла бы («дайте-ка») номер Лелула

(6)

С: .hh *võtaks selle 'mängumaa 'ka, Anni 'mängumaa.*

С: (я) взяла бы («дайте-ка») эту игровую комнату тоже, игровую комнату Анни

Запросы, выполняющие более одной функции (напр., согласие + запрос, см. табл. 1), нетипичны, и их форма полностью соответствует добавочной функции высказывания (прим. 3, реплика 2).

Сигналами для распознавания директивных запросов могут служить: 1) глаголы, 2) связанные с семантикой глагола морфологические формы, 3) порядок слов.

#### 4. Ответы информатора

После запроса диалог может развиваться двояко: 1) информатор выполняет запрос немедленно; 2) информатор инициирует поддиалог уточнения, проясняя запрос. Запрошенная информация предоставляется немедленно в 60% случаев (1).

В поддиалогах уточнения (40%) информатор уточняет название или расположение учреждения, номер телефона, дату запрашиваемого события и т.д. (7).

(7)

I: .hhh siis te mõtlete 'Epiritari.

*так вы имеете в виду Эпитар? УТОЧНЯЮЩИЙ ВОПРОС*

C: 'j:ah, 'Epiritari.

*да, Эпитар. УТВЕРДИТЕЛЬНЫЙ ОТВЕТ*

I: jah,

*да (ага) ОБРАТНАЯ СВЯЗЬ*

I: üks='hetk

*один момент ОТСРОЧКА*

I: (...) .hh 'number on 'üldinfo,=

*номер - общая информация.. ИНФОРМАЦИЯ*

Вторые элементы смежной пары запрос-ответ представляют собой предоставление информации, сообщение об отсутствии информации и сообщение об отсрочке (табл. 1).

Для информационного агента важно, как представить информацию наиболее естественным и в разумной степени лаконичным путем: как построить ответ, как разделить информацию на удобные для восприятия блоки.

##### *А. Лингвистические характеристики ответов*

Простейший способ сформулировать ответ – использовать в ДС готовые шаблоны возможных ответов, например:

*<Учреждение> - телефонный номер <X>.*

Однако информаторы в естественных диалогах имеют дело с более сложными запросами и ведут себя иначе. В 1/3 случаев используются экономичные эллиптические конструкции. Самая распространенная цель клиента – узнать телефонный номер. Информаторы разделяют телефонные номера, состоящие из 7 цифр, на две или три реплики (3, 2 и 2 цифры или 3 и 4 цифры), делая между ними паузы, в которое клиент подтверждает получение информации или переспрашивает недослышанное (2). Интонация реплик постепенно понижается, позволяя клиенту спрогнозировать конец номера; после него клиент обычно повторяет цифры последней реплики или номер целиком. В этом случае молчание информатора является подтверждением правильности понимания. Если клиент хочет эксплицитного подтверждения правильности, он оформляет повтор как вопрос: «<Номер>, да?» [9]. ДС должна учитывать, желает ли клиент подтверждения повтора.

Если ответу на запрос клиента предшествует уточняющий поддиалог, он, как правило, оформляется полным предложением: информатор маркирует возвращение к основной линии диалога, связывая реплику с запросом.

Сообщение об отсутствии информации также более подробно: информатор поясняет, почему не может предоставить информацию (8) или предлагает замещающую информацию (3).

(8)

I: ma vaatan et 'selle nimega nagu ei 'näita.

*я смотрю, что с этим названием вроде бы не показывает.*

Для сообщения об отсрочке (нерелевантной для ДС, которая оперирует базами данных во много быстрее человека) используется ограниченный список устойчивых выражений:

*üks hetk 'один момент' (минуточку) или*

*üks hetk, palun 'один момент, пожалуйста' (7).*

*В. Запросы, не получившие ответа*

Первичный запрос не получает ответа в тех случаях, если он неясен информатору и в процессе его уточнения клиент формулирует новый запрос (3 случая).

**5. Структура диалога**

Институциональный телефонный диалог четко структурирован. В конвенциональной вводной части информатор представляется, и собеседники обмениваются приветствиями. Отвечая на звонок, информатор придерживается шаблона:

<название учреждения> <имя информатора> <приветствие>

Клиент, как правило, ограничивается ответным приветствием (9) и переходит к главному информационному запросу, открывающему главную часть диалога. Информатор может инициировать уточняющие поддиалоги после запроса, клиент – после получения ответа на запрос. Диалог заканчивается конвенциональной частью, в которой клиент благодарит информатора, вместе с тем инициируя завершение разговора, а информатор принимает благодарность и тем самым завершает разговор. Если диалог длится дольше обычного, то клиент не ограничивается благодарностью, но и эксплицитно прощается (10).

(9)

I: info`telefon=  
*справочная служба ИДЕНТИФИКАЦИЯ*  
I: Kersti=  
*Керсти ИДЕНТИФИКАЦИЯ*  
I: tere  
*здравствуйте ПРИВЕТСТВИЕ*  
C: tere päevast.  
*добрый день ПРИВЕТСТВИЕ*

(10)

C: suur `tänu teile.=  
*большое спасибо вам БЛАГОДАРНОСТЬ*  
I: =jaa palun.  
*да, пожалуйста ПРИНЯТИЕ*  
C: head päeva.  
*до свидания. ПРОЩАНИЕ*  
I: <кладет трубку>

Для ДС важно, что 1) при выполнении запроса предпочтительны короткие реплики; 2) информация делится на короткие блоки, разделенные паузами – это позволяет клиенту записать или запомнить ее и подтвердить ее получение; 3) информатор предлагает замещающую информацию, если не может выполнить запрос клиента; 4) реплики обратной связи указывают, что информация успешно получена и партнер готов слушать дальше; 5) для приветствия, прощания, выражения согласия и обратной связи используется ограниченный набор устойчивых выражений.

Работая над ДС, мы пытаемся смоделировать «идеального» информатора – четко произносящего реплики, терпеливого и вежливого. В естественных диалогах информатор не всегда отвечает этому идеалу, однако люди-информаторы пользуются удобными стратегиями, присущими устной речи и позволяющими информатору быть лаконичным и кооперативным.

Типичная структура звонка в справочную службу представлена формальным описанием (схема 1).

диалог ::= конвенциональное\_введение (запрос (уточнение)\* ответ (уточнение)\*)+  
          конвенциональное\_окончание  
ответ ::= предоставление\_информации | отсутствие\_информации  
конвенциональное\_введение ::= (звонок ответ) (введение-приветствие приветствие)  
конвенциональное\_окончание ::= благодарность принятие [прощание прощание]  
уточнение ::= запрос ответ

*Схема 1. Грамматика информационного диалога*

Грамматика представляет звено в справочную службу как последовательность смежных пар диалоговых актов, в которых первый элемент производится одним участником диалога, а второй – его собеседником.

Смежные пары диалоговых актов предполагают использование магазинной структуры. Начальный запрос клиента задает цель, помещаемую в магазин и остающуюся там, пока запрос не будет выполнен. Каждый последующий запрос задает новую (под)цель, которая должна быть выполнена, чтобы достичь начальной цели. Цели помещаются в магазин в хронологическом, а извлекаются в обратном хронологическом порядке.

В несколько ином членении структура информационного диалога описана на материале русских справочных диалогов А.Е. Кибриком, выделяющим фазу запроса (формулирование и прием запроса), фазу нормализации запроса (промежуточные уточняющие вопросы и вспомогательные ответы) и фазу ответа (ответ оператора на запрос и снятие запроса пользователем) [11].

Для общения с пользователем на эстонском языке разработаны две простые ДС [10]. Одна из них (Транспортный агент) предоставляет информацию о вылетах из Таллиннского аэропорта. Другая ДС (Театральный агент) отвечает на вопросы о программах эстонских театров. В обе программы интегрирован синтез текст-речь. Обе системы пользуются ключевыми словами для распознавания запросов клиента и шаблонами предложений для генерирования ответов. Для ведения диалога используется регулярная грамматика.

## **6. Выводы**

Мы проанализировали эстонские естественные справочные диалоги с дальнейшей целью разработки диалоговой системы.

Диалог состоит из трех частей: 1) конвенционального введения в диалог, содержащего идентификацию информатора и взаимные приветствия, 2) основной части, в которой информатор отвечает на запрос(ы) клиента, и 3) конвенционального завершения диалога.

Начальный запрос клиента задает цель, которая должна быть достигнута в сотрудничестве с оператором. Собеседники могут инициировать поддиалоги для уточнения запроса или ответа. Автоматическому распознаванию запросов способствуют ключевые сигналы, выявляемые в репликах запросов (определенные глагольные формы и порядок слов). Поведение информаторов в естественных диалогах экономично и лаконично, при возможности используются эллиптические конструкции. В конвенциональных фазах и репликах диалога используются устойчивые выражения. Нужно учитывать это в разработке ДС, общающейся на естественном языке. Результаты работы будут использоваться в «оестествлении» существующих диалоговых систем.

## **Список литературы**

1. Lutz H.-D. Software-ergonomische Entwicklung – eine Herausforderung für die Computerlinguistik // SDV – Sprache und Datenverarbeitung, Vol. 25, 1, 2001, pp. 5-20.
2. McTear M.F. Spoken Dialogue Technology: Toward the Conversational User Interface. London: Springer Verlag, 2004.
3. Minker W. and Bennacef S. Speech and Human-Machine Dialogues Boston/Dordrecht/London: Kluwer Academic Publishers, 2004.
4. Hutchby I. and Wooffitt R. Conversation Analysis. Principles, Practices and Applications. Polity Press, 1998.
5. Koit M., Valdisoo M., Gerassimenko O., Hennoste T., Kasterpalu R., Rääbis A., and Strandson K. Processing of requests in Estonian institutional dialogues: corpus analysis // Text, Speech and Dialogue, Proceedings. (Eds) V. Matousek, P. Mautner, Springer, 2006, pp. 621-628.
6. Gerassimenko O., Kasterpalu R., Koit M., Rääbis A., and Strandson K. Initial requests in institutional calls: corpus study // International Conference Recent Advances in Natural Language Processing. Proceedings: Recent Advances in Natural Language Processing, (Eds) G. Angelova, K. Bontceva, R. Mitkov, N. Nicolov, N. Nikolov. Shoumen, 2007, pp. 230-234.
7. Fishel M. Complex taxonomy dialogue act recognition with a Bayesian classifier // Proc. of DECALOG Workshop on the Semantics and Pragmatics of Dialogue. Rovereto, Italy, 2007, pp. 161-162.
8. Ereht M. (Ed.) Estonian language // Linguistica Uralica Supplementary Series, vol 1. Estonian Academy Publishers, Tallinn, 2003.
9. Õim H. and Koit M. Developing a dialogue system that interacts with a user in Estonian // Inquiries into Words, Constraints, and Contexts. Festschrift for Kimmo Koskenniemi on his 60th Birthday (2005). Antti Arppe, Lauri Carlson, Krister Lindén, Jussi Piitulainen, Mickael Suominen, Martti Vainio, Hanna Westerlund and Anssi Yli-Jyrä (Editors). CSLI Studies in Computational Linguistics ONLINE. Copestake, Ann (Series Editor). CSLI Publications, Stanford, California. Available on-line at: <http://csli-publications.stanford.edu/site/SCLO.html> pp. 278-288.

10. Treumuth M., Alumäe T., and Meister E. A natural language interface to a theater information database // Language Technologies, IS-LTC 2006: Proceedings of 5th Slovenian and 1st International Conference, 9th - 10 October, Ljubljana, Slovenia. (Eds) T. Erjavec, J. Žganec Gros. Ljubljana, 2006, pp. 27 - 30.

11. Кибрик А.Е. Динамика информационного диалога (на материале взаимодействия со справочно-информационной службой 09) // Очерки по общим и прикладным вопросам языкознания. Москва: УРСС, 2001. стр. 301-313.