Параметризация типов предложений предметной области для системы устного фразаря-переводчика

Яценко В.В., *младший научный сотрудник*

В статье рассматриваются подходы построения системы перевода устного сигнала в рамках предметных областей. Блок интерпретации получает произнесённое предложение в виде последовательности слов, распознанной декодером. На выходе системы принимается решение о принадлежности распознанной последовательности слов типу предложения, задающего тип смысла. Распознавание выполняется с учётом параметров, которые описывают множество возможных вариантов высказываний. Проанализированы альтернативные подходы моделирования ограничений на допустимые последовательности слов. Надёжность распознавания HMM-декодера в условиях сформированных акустической и лингвистической моделей позволила получить приемлемую интерпретацию распознанного сигнала. Это легло в основу разработки демонстрационной системы устного фразаря-переводчика.

• *распознавание и интерпретация речи* • *словарь-переводчик* • *предметная область* • *тип смысла* • *тип предложения* • *украинская разговорная речь*.

In this paper we describe approaches to build the spoken translation system within a subject area. The decoded sequence of words enters to the interpretation subsystem, which finally makes decision concerning the sentence type and the respective meaning type for the pronounced sentence. Possible variations of date, time, place etc. that may occur in sentences are parameterized and integrated to the language model. Strict, free and phonetic word based word grammars for speech decoder are analyzed. Acoustic and language models created for the HMM-based decoder shows such performance that allows for understanding response accuracy sufficient for practical application. The demonstration version of the spoken interpreter has been developed and presented.

• *apeech recognition and understanding* • *spoken phrasebook* • *subject area* • *meaning type* • *sentence type* • *ukrainian spoken language*.

Среди важных практических задач, связанных с распознаванием речи, к которым относятся системы надиктовывания текстов, справочные системы, системы речевого управления оборудованием, системы речевого диалога и т.д., мы выделяем систему устного перевода. Актуальность задачи, в частности, отображается востребованностью автоматизации всем известного бумажного разговорника, в котором пользователь вынужден искать необходимую фразу и озвучивать её перевод. Вместо этого пользователю предлагается

только произнести фразу на родном языке в выбранной теме. Далее система делает всё самостоятельно. Дополнительного внимания требует вопрос моделирования параметров в предложениях, т.е. необходимо предусмотреть все возможные варианты значений для определённого предложения. Например, в вопросе о путешествии в конкретный город параметром будет название города.

Такие системы актуальны в свете использования их при разговоре с носителем другого языка. Пользователь не только получает перевод фразы, но и её озвучивание, что существенно облегчает общение в неродной языковой среде.

При построении систем устного перевода в рамках предметных областей возникает ряд проблем, общих с проблемами задачи понимания речевого сигнала. Необходимо построить модели всех возможных предложений языка диалога, которые выражают один и тот же смысл, смоделировать параметры слов в типах предложений, сгенерировать и найти наиболее правдоподобные эталонные сигналы, учитывая параметры.

Для исследования и спецификации ограничений на допустимые последовательности слов во фразах использовались LISP-структуры [1, 2]. На основе этих структур генерируется большое количество предложений, которые имеют одинаковый смысл с точностью до параметров. Впрочем, существует ряд ограничений на использование этой технологии, связанных как с субъективным фактором при построении LISP-структур, так и с увеличением количества вычислений, обусловленных существенным усложнением графа распознавания.

В качестве альтернативы LISP-структур предлагается способ оценивания принадлежности последовательности слов типам предложений, которые характеризуют смысл [3]. Этот подход требует развития, в частности, с целью учёта возможных ошибок распознавания.

Для моделирования ограничений на порядок следования слов использовались грамматические знания [2]. Для моделирования параметров в типах предложений использовались базы данных и базы знаний. Была сформулирована лингвистическая модель интерпретации распознанного сигнала с учётом параметров.

Общая структура системы устного перевода в пределах предметных областей

Распознавание и смысловая интерпретация слитной речи выполняются во взаимосвязанном процессе, конечная цель которого — перевод смысла сообщения на другой язык.

Рассмотрим задачи распознавания и интерпретации слитной речи [1, 2] и их взаимосвязь. Распознавание речи — процесс автоматической обработки сигнала с целью определения последовательности слов, которые передаются этим сигналом. Смысловая интерпретация языка — процесс автоматической обработки речевого сигнала с целью выявления смысла, передаваемого сигналом, и представление этого смысла в определённой канонической форме, удобной для дальнейшего использования в системе устного перевода.

Очевидно, что смысловая интерпретация языка является более высокой степенью обобщения информации, чем распознавание. Поскольку каждую мысль можно выразить различными предложениями в языке диалога без изменения содержания, то следует определить некоторые ограничения на допустимые последовательности слов в предложениях. Поэтому, при интерпретации смысла речи различные предложения, которые передают одну и ту же мысль, должны отражаться в один и тот же результат, т.е. ответ распознавания не должен противоречить синтаксису, семантике и прагматике предметной области.

Ввиду этого, предлагается рассмотреть структуру системы устного перевода в рамках предметных областей (рис.1). Задача смысловой интерпретации слитной речи с целью дальнейшего перевода основывается на том, что сначала пользователь должен задать предметную область (далее ПО), с которой он хочет работать. Для этого нужно назвать эту ПО. Вообще рассматривается 15 ПО, с которыми может работать пользователь.

Активатор выбирает названную ПО и загружает подсловари ПО с соответствующими этой области типами предложений и грамматику, по которой моделируются допустимые ограничения на последовательности слов в предложениях.

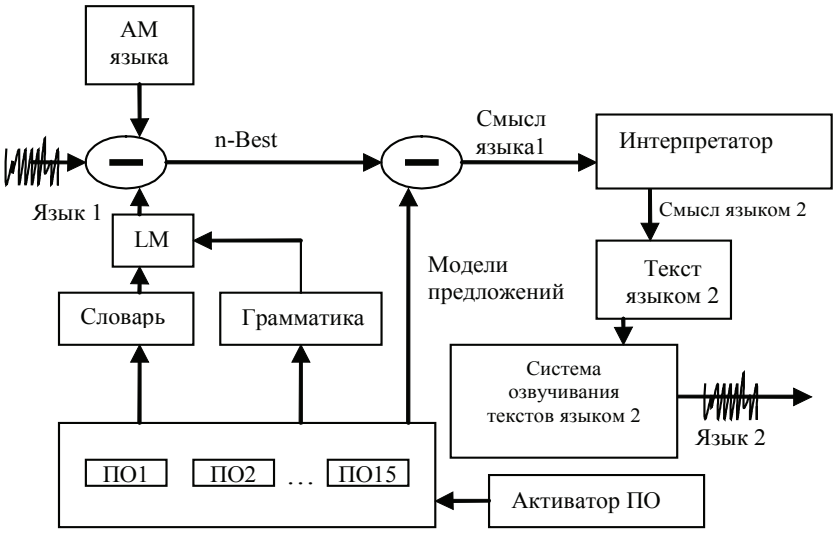


Рис.1. Структура системы устного перевода в рамках предметных областей

Диктор произносит на языке 1 предложение, которое распознаётся с учётом акустической модели и, построенной согласно словарю соответствующей ПО и грамматики, лингвистической модели (LM). Затем выбирается *n* лучших последовательностей слов и сравнивается с нагенерированными моделями предложений, которые могут задавать соответствующий тип предложения (далее ТП). Используя вероятностное оценивание, принимается решение о принадлежности распознанной последовательности слов к ТП. По этому ТП определяется тип смысла (далее ТС) и интерпретатор находит соответствующий ТС на другом языке. На выходе мы должны получить текст на языке 2, который озвучивается соответствующей системой озвучивания текстов на языке 2.

В описанной структуре перевода остаётся достаточно сложная задача интерпретации распознанного сигнала. Пути решения этой задачи, описанные в [2], основываются на том, чтобы научиться экономно задавать все допустимые предложения в языке диалога.

Таким образом, автоматический перевод фразы, произнесённой на языке 1, на язык 2 с озвучиванием результата, с помощью предлагаемой структуры устного перевода будет заключаться в том, чтобы сначала для сигнала, который произносится диктором, найти наиболее правдоподобный, возможно, параметризированный, ТП среди всех ТП, задающих ТС. Затем определить сам ТС произнесённого содержательного высказывания и найти для него подходящий ТС в языке 2 с учётом параметров. Наконец, предложение, полученное на языке 2, озвучивается.

Моделирование типов предложений с учётом параметров

Поскольку структура перевода должна работать в рамках ПО, то предлагается рассмотреть определённую иерархию речевых сигналов [2]. Подразумевается, что вся деятельность человека разбивается на ПО по аналогии с бумажным разговорником. Каждая ПО состоит из конечного множества ТС.

В каждый ТС входит множество эквивалентно содержательных ТП. ТП — конструкция, экономно задающая множество предложений, полученных из одного предложения независимыми допустимыми заменами и допустимыми перестановками или выпадением слов и словосочетаний.

В рамках задачи распознавания, интерпретации и перевода речевого сигнала немаловажен вопрос описания параметров слов во фразах, где могут быть разные варианты имён собственных, времени, адресов и т.д. Значение термина «параметр» может иметь разную интерпретацию в зависимости от контекста. В общем параметром называют величину, значения которой служат для различия элементов некоторого множества между собой.

Рассмотрим пример ТП «просьба разбудить человека в определённое время», из ПО «Гостиница». Базовая структура будет иметь вид:

$$\left((\text{разбудите}) \begin{pmatrix} \text{меня} \\ \text{нас} \\ * \end{pmatrix} \begin{pmatrix} \text{пожалуйста} \\ * \end{pmatrix} \begin{pmatrix} \textit{в}\ \ \$\text{time} : \text{app} \\ \begin{bmatrix} \text{в} \\ \text{через} \end{bmatrix}\ \ \$\text{time} \end{pmatrix} \right)$$

В круглых скобках () указаны подсловари, которые можно переставлять местами, а в квадратных [] — которые нельзя переставлять. Символ * означает пустое слово.

Этой структурой можно сгенерировать много предложений с учётом параметров. Среди этих предложений будут, например, и такие:
Разбудите меня, пожалуйста, в семь часов.
Пожалуйста, нас разбудите в пять утра.
В семь тридцать разбудите меня.
Разбудите нас в шесть.
Разбудите меня через шесть часов.

Стоит отметить, что предложения разговорной речи тоже необходимо учитывать.

В этом примере параметр — «временно́е предназначение»: $time:app, $time. Рассмотрим первый параметр $time:app. Он описывает любое время с точностью до, например, минут, в контексте определённого события. Чтобы предусмотреть все варианты и значения этих параметров вводится специально разработанная и описанная параметрическая грамматика словаря на основе формы Бекуса — Наура (BNF). Такую грамматику можно подать в развёрнутом виде (таблицы 1–2).

В приведённом примере базовая структура задаёт 4!·1·3·2·3=432 параметризированных предложений, допустимых в языке диалога. Если учесть, что каждый параметр содержит большое количество вариантов, то количество предложений значительно увеличится.

Таблица 1

Базовые структуры параметров временно́го предназначения

Обозначения	Пример	Параметризация для русского языка
$time:app	в шесть	$hour:nadj-at
	в шесть тридцать	$hour:nadj-at [$teen:n | $dec:5max]
	в шесть часов	$hour:nadj-at $hour-i
	в шесть часов утра	$hour:nadj-at $hour-i $time:post
	в шесть тридцать утра	$hour:nadj-at [$teen:n | $dec:6max] $time:post
	в шесть час. тридцать мин.	$hour:nadj-at $hour-i $min:n-u
	в шесть часов тридцать минут утра	$hour:nadj-at $hour-i $min:n-u $time:post

Таблица 1 (окончание)

$min:n-u	одна минута	$min:n1 $min1
	две минуты; 53 минуты	$min:n2 $min2
	5 минут; 37 минут	$min:n5 $min5
$min:n5	20; 45	$dec:5max [$digit5]
	5; 7	$digit5
	12; 15	$teen:n

Все предложения языка диалога можно задавать с помощью ТС и соответствующих им ТП, используя структуру, приведённую в примере. С помощью LISP-структур генерируется огромное количество предложений, имеющих одинаковый смысл. Поскольку построение LISP-структур довольно громоздкое, требует много ручной работы, то был разработан автоматизированный спецификатор ПО.

Таблица 2

Описание значений параметров для временно́го предназначения

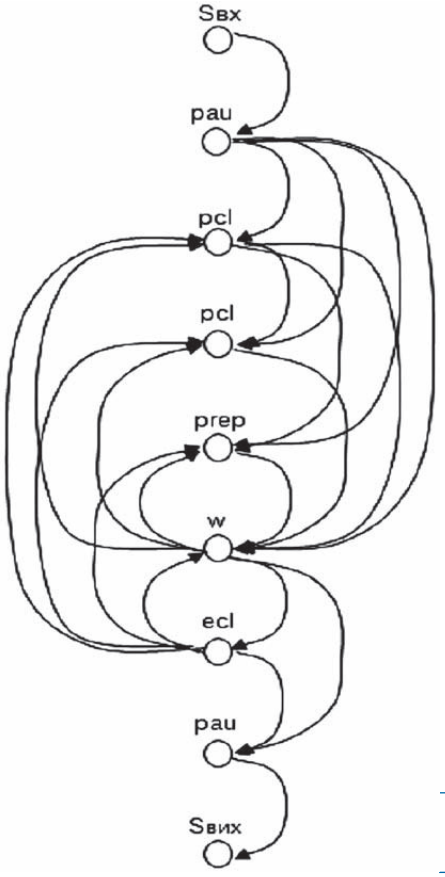
$hour:nadj-at	первый	one
	второй	two
	…	
	двадцать третий	twenty three
$hour-i	часа	o'clock
$min:n1	одна	one
	…	
	пятьдесят одна	fifty one
$min:n2	две	two
	три	three
	…	
	пятьдесят три	fifty three
$teen:n	десять	ten
	…	…
	девятнадцать	nineteen

$dec:50max	двадцать	twenty
	тридцать	thirty
	сорок	forty
	пятьдесят	fifty
$digit5	пять	five
	…	
	девять	nine
$min1	минута	minute
$min2	минуты	minute
$min5	минут	minute
$time:post	утра	a.m.
	дня	p.m.
	вечера	p.m.
	ночи	a.m.

Для построения всех возможных предложений языка устного диалога можно использовать так называемую ориентированную семантическую сеть (далее ОСС) [1, 2], одновременно задающую ограниченную грамматику порядка следования слов.

Альтернатива этой грамматики — грамматика свободного порядка следования слов. Между этими противоположными, по сути, грамматиками может быть построено множество других относительно свободных или относительно ограниченных грамматик. Мы предлагаем несколько ограничить свободную грамматику за счёт лингвистического понятия о фонетическом слове [2].

Под «фонетическим словом» понимаем слово с неотделимыми от него сопутствующими словами. Например, неотделимыми являются предлог от существительного или прилагательного, частица «не» перед глаголом и частица «б» после него. Предлагаемая нами, относительно свободная грамматика, представлена в виде графа (рис. 2), где *pau* — слово-пауза в начале и в кон-



це фразы, *pcl* — проклитик, *prep* — предлог, *w* — нейтральное слово, *ecl* — энклитик.

Впрочем, при такой грамматике принятие решений относительно смысла предложения неочевидно.

Рис. 2. Граф относительно свободной грамматики на основе понятия про лингвистическое слово

Статистическое оценивание принадлежности последовательности слов к типу предложения

При распознавании в условиях грамматики, которая не задаёт строгих ограничений на последовательности слов, очевидно, могут быть получены ответы распознавания, не входящие во множество предложений, которые сгенерированы определённым ТП. Это может быть обусловлено как ошибками при распознавании, так и при формировании ТП экспертом. Кроме того, сам пользователь может произнести предложение с различного рода отклонениями или аграмматизмами, например, повторить некоторое слово дважды.

Поэтому предлагается оценивать вероятность типа предложения *ST* с ОСС при распознанной последовательности слов $(w_1, w_2, ..., w_n)$ и объявлять ответом интерпретации тот тип предложений *ST**, для которого эта вероятность является наибольшей:

$$ST^* = \underset{ST}{\operatorname{argmax}}\, P(ST / w_1, w_2, ..., w_n), \qquad (1)$$

Вероятность в левой части (1) может быть записана также по формуле Байеса в следующем виде:

$$P(ST / w_1, w_2, ..., w_n) = \frac{P(ST)}{P(w_1, w_2, ..., w_n)} P(w_1, w_2, ..., w_n / ST), \qquad (2)$$

Рассматривая последовательность $(w_1, w_2, ..., w_n)$ как Марковский процесс, отображаем каждый из множителей условной вероятности в правой части (2) в виде:

$$P(w_1, w_2, ..., w_n / ST) = \prod_{k=1}^{n} P(w_k / ST, w_{k-m}, ..., w_{k-1}) \ , \quad (3)$$

$$P(w_1, w_2, ..., w_n) = \prod_{k=1}^{n} P(w_k / w_{k-m}, ..., w_{k-1}) \ , \quad (4)$$

где $m \geq 0$ — порядок процесса.

Оценивание каждого из множителей правой части выражений (3) и (4) может производиться различными способами в зависимости от выбранного порядка процесса.

Мы рассматривали наиболее простой случай, когда $m = 0$. Тогда, учитывая формулу Байеса, выражение (2):

$$P(ST / w_1, w_2, ..., w_n) = P(ST) \prod_{k=1}^{n} P(ST / w_k) \ . \quad (5)$$

Логично сделать предположение относительно равной вероятности всех типов предложений. В действительности, некоторые смыслы встречаются чаще других. Это зависит от предыдущего смысла (контекста). Остаётся рассчитать выражение вида $P(ST / w_k)$. Для этого рассмотрим $ST(w_k)$ — множество типов предложений, в которых встречается слово w_k. Тогда:

$$P(ST/w_k) = \begin{cases} |ST(w_k)|^{-1}, & \text{если } ST(w_k) \cap ST \neq \varnothing, \\ \alpha(ST, w_k), & \text{иначе}. \end{cases} \quad (6)$$

Выражение $\alpha(ST, w_k)$ отображает смысл вероятности того, что слово w_k распознано ошибочно вместо некоторого слова w: $ST(w) \neq \varnothing$. Эту вероятность можно оценить на основе некоторой меры минимальной редакторской правки $d(w_k, w)$, например, расстоянии Левенштейна. При вычислении этой меры штрафуются вставки, удаления и замены символов фонемного текста сравниваемых слов. Таким образом, выражение $\alpha(ST, w_k)$ предлагается оценивать как:

$$\alpha(ST, w_k) = \max_{ST(w) \neq \varnothing} \left(\max\left\{ 1 - \frac{d(w_k, w)}{L(w)}, 0 \right\} \times |ST(w)|^{-1} \right), \quad (7)$$

где $L(w)$ — количество фонем в слове w.

Решение относительно принадлежности распознанной последовательности слов некоторому ТП принимается на основании (1) — (7).

В случае, когда распознанная последовательность слов при таком оценивании совпадает с определённым ТП, принятие решения очевидно. Но может быть так, что некоторые слова распознались ошибочно. Такое предложение можно отбросить, не найдя для него соответствующий ТП. А можно попробовать оценить, к какому ТП ближе распознанная последовательность слов. И определить гипотетический ТП, т.е. который можно объявить ответом интерпретации.

Рассмотрим это на примере. Допустим последовательность распознанных слов.

$$(w_1, w_2, w_3) = Допоможіть \ \ було \ \ маска$$

Обозначим ST_1 _ ТП, к которому эта последовательность будет ближе всего.

$$ST_1 = Допоможіть \ мені \ будь \ ласка$$

Оценим вероятность принадлежности распознанной фразы (w_1, w_2, w_3) к ТП ST_1. Воспользуемся формулами (5) — (6).

$$P(ST_1 / \underset{w_1}{допоможіть}, \underset{w_2}{було}, \underset{w_3}{маска}) = P(ST_1)P(ST_1 / w_1)P(ST_1 / w_2)P(ST_1 / w_3) \cong$$

$$\cong 1 \cdot \frac{1}{|ST(w_1)|} \cdot \alpha(ST, w_2) \cdot \alpha(ST, w_3).$$

Поскольку мы предположили, что все ТП имеют одинаковую вероятность, то $P(ST_1) = 1$. Далее нужно подсчитать значение каждого множителя. Слово $w_1 =$ «допоможіть» распозналось правильно, поэтому

$$P(ST_1 / w_1) = \frac{1}{|ST(w_1)|} = 1.$$

Слова «було» и «маска» явно не принадлежат ST_1. Мы предполагаем, что эти слова являются гипотетическими ошибками распознавания.

Чтобы подсчитать вероятность принадлежности распознанной последовательности слов (w_1, w_2, w_3) ТП ST_1, сформируем множество слов из ST_1, которые там остались без учёта правильно распознанного слова w_1 =«допоможіть». Получится следующее множество: $w(ST_1) \setminus w_1 = \{мені, будь, ласка\}$.

Как было упомянуто выше, оценивать вероятности ошибочно распознанных слов w_k будем на основе меры Левенштейна — d, а именно используя формулу (7). Нам нужно посимвольно сравнить фонемный текст каждого гипотетически ошибочно распознанного слова $(w_2, w_3) = \{було, маска\}$ с множеством слов $w(ST_1) \setminus w_1 = \{мені, будь, ласка\}$. Вводится следующая система штрафования: вставки, удаления, замены символов фонемного текста штрафуются одним балом, а совпадение фонем не штрафуется. На рис. 3 представлены результаты сравнения фонемного текста и подсчитаны минимальные расстояния меры Левенштейна.

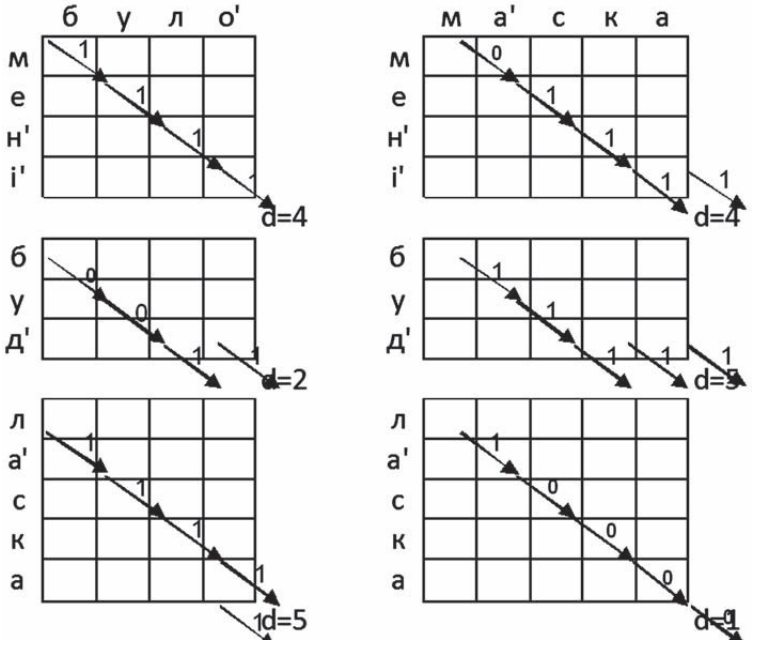


Рис. 3. Сравнение символов фонемного текста

Далее для слова $w_2 =$ «було» в развёрнутом виде описан подсчёт вероятности по формуле (7):

$$P(ST_1 / w_2) = \max_{ST(w)\neq\varnothing}\left(\max\left\{1-\frac{d(w_2,w)}{L(w)},0\right\}\times|ST(w)|^{-1}\right)=$$

$$= \max_{ST(w)\neq\varnothing}\left\{\max\left\{1-\frac{d(w_2,'мені')}{L(мені)},0\right\}\times|ST(мені)|^{-1},\ \max\left\{1-\frac{d(w_2,'будь')}{L(будь)},0\right\}\times|ST(будь)|^{-1},\right.$$

$$\left.\max\left\{1-\frac{d(w_2,'ласка')}{L(ласка)},0\right\}\times|ST(ласка)|^{-1}\right\}=$$

$$= \max_{ST(w)\neq\varnothing}\left\{0,\frac{1}{3}\times|ST(будь)|^{-1},0\right\}.$$

Аналогично для слова $w_3 =$ «маска»:

$$P(ST_1 / w_3) = \max_{ST(w)\neq\varnothing}\left\{0,\ 0,\ \frac{4}{5}\times|ST(ласка)|^{-1}\right\}.$$

Мы видим, что для каждого из этих слов существует ненулевая вероятность того, что они могут принадлежать ТП ST_1.

Таким образом, вероятность того, что распознанная фраза $(w_1, w_2, w_3) = Допоможіть\ було\ маска$ принадлежит ТП ST_1 будет:

$$P(ST_1 / допоможіть, було, маска) \cong 1\cdot\frac{1}{3|ST(будь)|}\cdot\frac{4}{5|ST(ласка)|}\le\frac{4}{15}.$$

Подсчитав окончательно по формуле (5) вероятность принадлежности распознанной фразы к предполагаемому ТП, мы видим, что гипотеза данного ТП не отбрасывается и при отсутствии других гипотез может быть ответом интерпретации.

Экспериментальные результаты

Предложенные в работе методы оценивания принадлежности последовательности слов к ТП были экспериментально проверены на фразах из обычного разговорника. В работе для примера рассматривались три ПО: «Повседневные фразы», «Путешествие», «Гостиница». Эти ПО содержат 47 + 102 + 68 = 217 ТС. В среднем на ТС приходится 4,17 базовых предложения.

Акустические модели для декодера разработаны на основе речевого корпуса отдельно произнесённых слов, в котором принимали участие 60 дикторов [2]. Средствами [3] проведено обучение 55 скрытых Марковских моделей фонем. Максимальное количество нормальных законов в смеси — 20.

Для эксперимента произвольным образом было выбрано 500 фраз. Смысловая интерпретация проводилась на основе результата пофонемного распознавания речевых сигналов [4] в условиях свободной и относительно свободной (на основе фонетических слов) грамматик относительно слов [2]. Из результатов проведённого эксперимента (таблица 3) следует, что для двух типов грамматик отклонение смысловой интерпретации не превышает 5%, что является приемлемым для прикладной системы.

В условиях ограниченной грамматики скорость распознавания в 10 раз превышает реальное время, а в условиях свободной и относительно свободной грамматик распознавание происходит быстрее реального времени на ресурсах нетбука.

Таблица 3

Результаты распознавания и смысловой интерпретации 500 предложений из двух предметных областей

Тип грамматики	Надёжность распознавания (%)		
Ограниченная	96,7	94,1	98,3
Свободная пословная	53,4	4,2	86,1
Относительно свободная	79,1	20,8	96,2

На основе проведённых исследований разработана демонстрационная программная модель для перевода произнесённых предложений с русского языка на английский (рис. 4). При этом последовательность слов в русском предложении может быть любой из допустимых. Предложению, произнесённому на русском языке, ставится в соответствие англоязычный ТС или ТП, а первое предложение этого ТС объявляется результатом перевода.

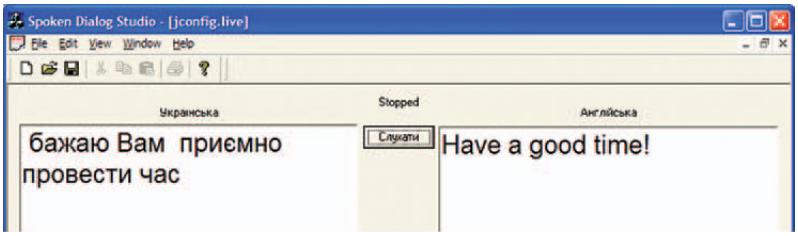


Рис. 4. Демонстрационное программное обеспечение модели устного фразаря-переводчика

Выводы

Рассматриваемая в работе система устного перевода является электронным аналогом бумажного разговорника, взаимодействие с которым происходит наиболее естественным способом — голосом. При распознавании произнесённой пользователем фразы используются лингвистические и семантические знания по выбранной ПО. Введённые при этом «мягкие» ограничения на порядок следования слов позволяют повысить надёжность распознавания, не повышая требований к вычислительным ресурсам. Разработанное программное обеспечение даёт возможность формировать грамматики следования слов для распознавания слитной речи как на основе ТП, так и на основе лингвистического понятия «фонетическое слово».

Использование параметрических моделей ТП даёт возможность пользователю более свободно и разнообразно общаться, что расширяет сферу использования системы. Предположение, что наблюдаемые последовательности слов — Марковский процесс, дало возможность сформулировать более гибкий способ формирования результата смысловой интерпретации.

На основе экспериментальной модели разработана программная модель устного словаря-переводчика, для перевода с русского языка на английский в рамках предметной области, которая работает в режиме реального времени на ограниченных вычислительных ресурсах.

Одни и те же фразы, произнесённые с различной интонацией, могут выражать как вопросительное предложение, так и повествовательное. Итак, в дальнейшей работе следует исследовать возможность распознавания интонации и ритма (просодики) с целью автоматической расстановки знаков препинания в распознанных фразах.

В дальнейшем также планируется ставить в соответствие русскоязычной фразе более точный англоязычный аналог среди ТП по ТС.

Литература

1. *Vintsiuk T.K*. Analysis, Recognition and Understanding of Speech Signals, Kyiv: Naukova Dumka, 1987.

2. *Sazhok M., Yatsenko V*. Spoken translation system based on speech understanding in subject area // All-Ukrainian Int. Conference on Signal/Image Processing and Pattern Recognition UkrObraz'2010. Kyiv, 2010. P. 103–106.

3. *Lee, Kawahara T*. and *Shikano K*. Julius — an open source real-time large vocabulary recognition engine. — In Proc. European Conference on Speech Communication and Technology (EUROSPEECH). 2001.P.1691–1694.

4. *Young S.J*. et al. HTK Book, version 3.1, Cambridge University. 2002.

Сведения об авторах

Яценко Валентина Витальевна —
работает в Международном научно-обучающем центре информационных технологий и систем в отделе распознавания и синтеза речевых сигналов. Занимается формированием словарей, фраз для словарей-разговорников и интерпритацией распознанных фраз и переводом их на другой язык. Киев. val-yatsenko@yandex.ru