

В.В. Яценко, Н.Н. Сажок

## Система устного перевода спонтанных высказываний в рамках предметных областей

Статья посвящена развитию системы устного перевода в рамках предметных областей путем моделирования однородных словосочетаний и обработки гипотетических ошибок декодера, что типично для распознавания спонтанной речи. Приемлемая надежность интерпретации спонтанных высказываний позволила разработать демонстрационную систему устного перевода.

The article is devoted to the further development of the spoken translation in subject areas by homogeneous phrase modeling and processing hypothetical errors of the decoder that is typical for spontaneous speech recognition. Acceptable reliability of spontaneous utterance understanding allowed for developing a demonstration system for spoken interpretation.

Статтю присвячено розвитку системи усного перекладу в межах предметних областей шляхом моделювання однорідних словосполучень та обробки гіпотетичних помилок декодера, що є типовим для розпізнавання спонтанної мови. Прийнятна надійність інтерпретації спонтанних висловлювань дозволила розробити демонстраційну систему усного перекладу.

**Введение.** Ежедневное общение людей базируется на постоянном обмене информацией, который происходит по-разному: либо в устной форме, либо в письменной, либо какой-то другой. При устном общении на родном языке человек не очень задумывается над построением спонтанно произносимой фразы, что ведет к искажению грамматических конструкций, возникновению вставок, заполненных пауз, служащих источником ошибок распознавания.

При рассмотрении задачи автоматического распознавания произнесенной фразы или речевого сигнала все еще недостаточно внимания уделяется вопросу смысловой интерпретации речи, что служит ключом к восстановлению содержания высказывания.

Так, задачи распознавания и интерпретации речевого сигнала следует рассматривать в совокупности при формировании ответа распознавания. С одной стороны, это дает возможность формировать более адекватный ответ распознавания. С другой, – имея возможность оперировать смыслами, легко перейти к такому же смыслу, но уже на другом языке, получая таким образом систему устного перевода.

Сегодня система распознавания речи востребована и применима во многих практических задачах: в системах надиктовывания текстов, справочных системах, системах речевого управления различным оборудованием, системах речевого диалога (например, посредством телефонной связи). Среди этих задач выделим *систему устного перевода*. Актуальность задачи, в частности, подтверждается востребованно-

стью автоматизации известного всем бумажного разговорника. Бумажный разговорник – это книжка, разбитая на темы и имеющая множество страниц. Пользователь вынужден искать необходимую фразу и озвучивать ее перевод другим языком самостоятельно. Вместо этого предлагается только произнести фразу на родном языке в выбранной теме. Далее система все выполняет самостоятельно.

Для исследования и спецификации ограничений на допустимые последовательности слов во фразах использовались *LISP*-структуры [1, 2]. На основе этих структур строятся порождающие грамматики, генерирующие большое количество предложений, имеющих один и тот же смысл с точностью до значений параметров. Выяснилось, 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 динамического программирования.

В качестве альтернативы *LISP*-структур предложен способ оценивания принадлежности последовательности слов типам предложений, характеризующих смысл [3]. Этот подход требует развития, в частности, с целью учета возможных ошибок распознавания и экономного задания однородных слов.

Сначала рассматривается общая структура системы устного перевода, последующий раздел посвящен описанию моделирования типов предложений с введением метаслов, далее опи-

сывается способ оценивания принадлежности последовательности слов к типу предложения с учетом гипотетических ошибок декодера, в конце описаны эксперименты и демонстрационная система.

### Общая структура системы устного перевода в пределах предметных областей

Суть задачи распознавания речевого сигнала заключается в том, что это – процесс автоматического анализа сигнала с целью определения последовательности слов, передаваемых этим сигналом. А задача смысловой интерпретации – это процесс автоматической обработки речевого сигнала с целью выявления смысла, передаваемого сигналом, и представление этого смысла в определенной канонической форме, удобной для дальнейшего использования, например, в системе устного перевода [1, 2].

Для системы устного перевода задача смысловой интерпретации есть ключевой, поскольку для множества предложений, которые могут передавать одну и ту же мысль, нужно отобразить однозначный результат, по которому потом получить перевод на другой язык.

Поэтому 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 не выходит за рамки одной или нескольких предметных областей (ПО). При выборе ПО конструируется лингвистическая модель (ЛМ), включающая в себя соответствующие словари, списки однородных слов и фраз (метаслова и метафразы), а также грамматику, по которой моделируются допустимые ограничения на последовательности слов в предложениях. Акустическая модель (АМ), позволяющая генерировать эталонные сигналы предложений, остается неизменной для определенного языка.

Диктор произносит на языке 1 некоторое предложение, сигнал которого преобразуется в пространство первичных признаков и поступает в декодер на акустический компаратор, сравнивающий предъявленный сигнал с эталонами,

которые генерируются вследствие применения порождающих грамматик из АМ и ЛМ.

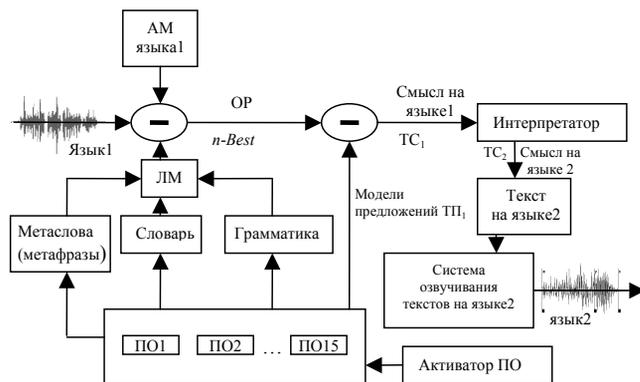


Рис. 1. Структура системы устного перевода в рамках предметных областей

Многозначный ответ распознавания (OP) в виде  $N$ -лучших ( $N$ -Best) последовательностей слов сравнивается с допустимыми типами предложений (ТП), которые моделируют предложения, имеющие один и тот же смысл. По наилучшему ТП определяется тип смысла (ТС), а Интерпретатор находит соответствующий ТС на другом языке. На выходе должны получить текст на языке 2, который может озвучиваться соответствующей системой.

В описанной структуре наименее исследована задача перехода от ответа распознавания к типу смысла. Пути решения этой задачи [2] основываются на том, чтобы научиться экономно задавать все возможные допустимые предложения в языке диалога.

Далее будут рассмотрены способы задания текстов всех возможных предложений, выражающих один и тот же смысл, а также введения ограничений на последовательности слов в декодере.

### Моделирование типов предложений с учетом однородных слов

Поскольку структура перевода должна работать в рамках предметных областей, предлагается рассмотреть определенную иерархию речевых сигналов [2]. Подразумевается, что вся деятельность человека разбивается на предметные области (ПО) по аналогии с бумажным разговорником для иностранных языков. Каждая ПО состоит из конечного множества ТС. В каждый ТС входит множество эквивалентно со-

держательных типов предложений. Тип предложения – это 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 немаловажным есть вопрос описания параметров слов во фразах, где могут быть разные варианты имен собственных, времени, адресов и другого, т.е. в каком-то смысле однородных слов.

В данной статье предложено использовать понятие *метаслова* как параметра, описывающего множество возможных семантически и синтаксически гомогенных (однородных) слов или словосочетаний во фразах при общении. Также вводится понятие *метафраза*, возникающее вследствие введения сочетаний метаслов во фразы.

Рассмотрим пример ТП просьба разбудить человека в определенное время, из ПО «Гостиница». Базовая конструкция имеет вид:

$$\left( \text{разбудите} \left( \begin{matrix} \text{меня} \\ \text{нас} \\ * \end{matrix} \right) \left( \begin{matrix} \text{пожалуйста} \\ * \end{matrix} \right) \left( \begin{matrix} \text{в } \$time:app \\ \text{через } \$time \\ \left[ \begin{matrix} \text{около} \\ \text{после} \end{matrix} \right] \$time:prox \end{matrix} \right) \right)$$

В круглых скобках ( ) указаны подсловари, которые можно менять местами, а в квадратных [ ] – которые нельзя переставлять. Переставлять подсловари можно только внутри старших скобок. Символ \* (звездочка) означает пустое слово.

Представленной конструкцией можно сгенерировать достаточно много предложений с учетом метаслов, вводимых мнемоникой через символ \$. Среди порожденных предложений будут, например, и такие:

- Разбудите меня, пожалуйста, в семь часов.*
- Пожалуйста, нас разбудите около пяти утра.*
- В семь тридцать разбудите меня.*
- Сегодня ночью разбудите нас, пожалуйста.*
- Разбудите меня через шесть часов.*

В этом примере метасловом есть *\$time:app* – момент времени и *\$time* – промежуток времени, *\$time:prox* – приближение времени. Рас-

смотрим метаслово *\$time:app*, которое будет составным. Оно описывает любой момент времени, например, с точностью до минут в контексте определенного события. Это метаслово порождает громадное количество словосочетаний, поскольку человек может выразить момент времени многими способами: указывать или нет единицы времени (часы, минуты), период суток, употреблять относительные выражения (например, «завтра»). Чтобы описать все варианты и значения метафраз вводится параметризованная грамматика на основе формы Бекуса–Наура (*BNF*). Такую грамматику можно подать в развернутом виде (табл. 1, 2).

В приведенном примере базовая структура задает  $4! \cdot 1 \cdot 3 \cdot 2 \cdot 4 = 576$  метафраз, допустимых в языке диалога. Если учесть варианты однородных слов, то количество допустимых предложений будет значительно больше.

Т а б л и ц а 1. Фрагмент спецификации составного метаслова *момент времени*

Мнемоника	Пример	Языкозависимые представления метафразы в <i>BNF</i>
<i>\$time:app</i>	в шесть	<i>\$hour:nadj-at</i>
	в шесть тридцать	<i>\$hour:nadj-at [\$steen:n   \$dec:5max]</i>
	в шесть часов	<i>\$hour:nadj-at \$hour-i</i>
	в шесть часов утра	<i>\$hour:nadj-at \$hour-i \$time:post</i>
	в шесть тридцать утра	<i>\$hour:nadj-at [\$steen:n   \$dec:6max] \$time:post</i>
	в шесть часов пять минут	<i>\$hour:nadj-at \$hour-i \$min:n-u</i>
	завтра утром	<i>\$day:rel [\$time:app-daypart]</i>
<i>\$min:n-u</i>	одна минута	<i>\$min:n1 \$min1</i>
	две минуты; 53 минуты	<i>\$min:n2 \$min2</i>
	5 минут; 37 минут	<i>\$min:n5 \$min5</i>
<i>\$min:n5</i>	20; 45	<i>\$dec:5max [\$digit5]</i>
	5; 7	<i>\$digit5</i>
	12; 15	<i>\$steen:n</i>

Для построения всех возможных предложений языка устного диалога можно использовать так называемую ориентированную семантическую сеть (ОСС) [1, 2]. Эта сеть задает порождающую грамматику, которая вводит ограничения на порядок следования слов при распознавании. Альтернативой этой грамматики служит грамматика свободного порядка следования слов. Между этими противоположными,

по сути, грамматиками может быть построено множество других относительно свободных или относительно ограниченных грамматик. Авторы предложили несколько ограничить свободную грамматику, используя лингвистическое понятие о фонетическом слове [4].

Таблица 2. Двухязычное описание значений метаслов

Обозначения	Рус. язык	Англ. язык
\$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.

Под фонетическим словом понимаем слово с неотделимыми от него сопутствующими словами. Например, неотделим предлог от существительного или прилагательного, частица *не* перед глаголом и частица *б* после него. Предлагаемая относительно свободная грамматика представлена в виде графа (рис. 2), где *pau* – слово-пауза в начале и в конце фразы, *pcl* – условный проклитик, *prep* – предлог, *w* – нейтральное слово, *ecl* – условный энклитик.

Следует отметить, что при такой грамматике принятие решения относительно ответа ин-

терпретации не очевидно, кроме того, необходимо учесть ошибки распознавания, что и будет рассмотрено далее.

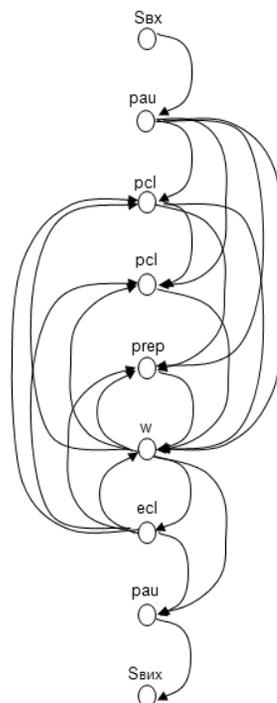


Рис. 2. Граф относительно свободной грамматики на основе понятия о лингвистическом слове

### Соотнесение ответа распознавания к типу предложения

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

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

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

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

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

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

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

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

где  $m \geq 0$  – порядок процесса.

Оценивание каждого множителя из правой части выражений (3) и (4) можно выполнять различными способами в зависимости от выбранного порядка процесса.

Рассмотрен наиболее простой случай, когда  $m = 0$ . Тогда, учитывая формулу Байеса, выражение (2) подадим в виде

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

Здесь логично предположить априорно равную вероятность всех типов предложений. В действительности некоторые смыслы встречаются чаще других, к тому же это зависит от предыдущего смысла (контекста), но рассмотрение контекста смыслов отложим на будущее. Остается рассчитать выражение вида  $P(ST / w_k)$ . Для этого рассмотрим  $ST(w_k)$  – множество типов предложений, в которых встречается слово  $w_k$ . Тогда

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

Выражение  $\alpha(ST, w_k)$  отражает смысл вероятности того, что слово  $w_k$  распознано ошибочно вместо некоторого слова  $w$ :  $ST(w) \neq \emptyset$ .

Эту вероятность можно оценить на основе некоторой меры минимальной редакторской правки или некоторого «расстояния» между словами  $d(w_k, w)$ . При вычислении этой меры штрафуются вставки, удаления и замены символов фонемного текста сравниваемых слов. Выражение  $\alpha(ST, w_k)$  предлагается оценивать так:

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

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

Решение относительно принадлежности распознанной последовательности слов некоторому типу предложения принимается на основании выражений (1) – (7).

Рассмотрим процесс принятия решения на примере. Допустим, ответом распознавания будет последовательность слов:

$(w_1, w_2, w_3) = (\text{«Допоможіть»}, \text{«було»}, \text{«маска»})$ .

Обозначим  $ST_1$  тот ТП, которому эта последовательность в действительности соответствует:

$ST_1 = \text{«Допоможіть мені будь ласка»}$

Оценим вероятность принадлежности распознанной последовательности слов  $(w_1, w_2, w_3)$  типу предложения  $ST_1$ . Воспользуемся выражениями (5) и (6):

$$\begin{aligned} P(ST_1 / \text{«Допоможіть, було, маска»}) &= P(ST_1) P(ST_1 / w_1) P(ST_1 / w_2) P(ST_1 / w_3) \cong \\ &\cong 1 \cdot \frac{1}{|ST(w_1)|} \cdot \alpha(ST, w_2) \cdot \alpha(ST, w_3). \end{aligned}$$

Поскольку принято предположение, что все типы предложения имеют одинаковую вероятность, то можно принять условно  $P(ST_1) = 1$ . Затем следует подсчитать значение каждого множителя. Слово  $w_1 = \text{«Допоможіть»}$  относится к правильному  $ST_1$  (и только к нему), поэтому

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

Слова *було* и *маска* явно не принадлежат  $ST_1$ . Предположим, что эти слова есть гипотетическими ошибками распознавания.

Чтобы подсчитать вероятность принадлежности распознанной последовательности слов  $(w_1, w_2, w_3)$  типу предложения  $ST_1$ , сформируем множество слов из  $ST_1$ , которые там остались без учета правильно распознанного слова  $w_1 = \text{допоможить}$ . Это будет следующее множество  $w(ST_1) \setminus w_1 = \{\text{мени, будь, ласка}\}$ .

Как было упомянуто, оценивать вероятности ошибочно распознанных слов  $w_k$  будем на основе расстояния между словами –  $d(\cdot, \cdot)$ , используя формулу (7). Необходимо посимвольно сравнить фонемный текст каждого гипотетически ошибочно распознанного слова  $(w_2, w_3) = \{\text{було, маска}\}$  с множеством слов  $w(ST_1) \setminus w_1 = \{\text{мени, будь, ласка}\}$ . Допустим, что используется следующая система оценки расстояния между словами: вставки, удаления, замены символов фонемного текста слов увеличивают расстояние на единицу, а совпадение фонем не изменяют расстояния. Результаты подсчета расстояний между гипотетически ошибочно распознанными словами и словами из  $ST_1$  представлены на рис. 1.

Далее для слова  $w_2 = \text{було}$  в развернутом виде выполняется подсчет вероятности по формуле (7):

$$\begin{aligned} P(ST_1 / w_2) &= \max_{ST(w) \neq \emptyset} \left( \max \left\{ 1 - \frac{d(w_2, w)}{L(w)}, 0 \right\} \times |ST(w)|^{-1} \right) = \\ &= \max_{ST(w) \neq \emptyset} \left\{ \max \left\{ 1 - \frac{d(w_2, \text{'мени'})}{L(\text{'мени'})}, 0 \right\} \times |ST(\text{'мени'})|^{-1}, \right. \\ &\quad \max \left\{ 1 - \frac{d(w_2, \text{'будь'})}{L(\text{'будь'})}, 0 \right\} \times |ST(\text{'будь'})|^{-1}, \\ &\quad \left. \max \left\{ 1 - \frac{d(w_2, \text{'ласка'})}{L(\text{'ласка'})}, 0 \right\} \times |ST(\text{'ласка'})|^{-1} \right\} = \\ &= \max_{ST(w) \neq \emptyset} \left\{ 0, \frac{1}{3} \times |ST(\text{'будь'})|^{-1}, 0 \right\}. \end{aligned}$$

Аналогично для слова  $w_3 = \text{маска}$ :

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

Очевидно, что для каждого из слов  $w_2$  и  $w_3$  существует ненулевая вероятность того, что они могут ошибочно наблюдаться в типе предложений  $ST_1$  в результате действия декодера.

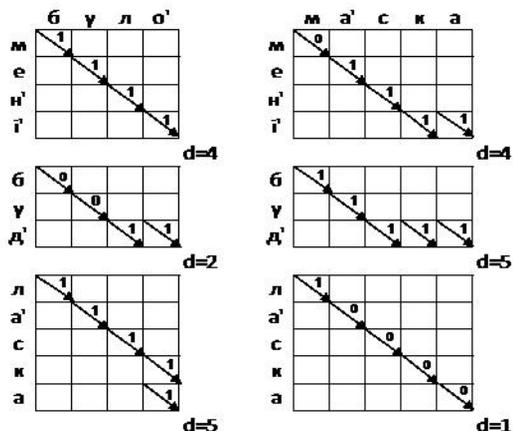


Рис. 3. Определение меры минимальной правки на фонемном тексте

Таким образом, вероятность того, что распознанная фраза  $(w_1, w_2, w_3) = \text{Допоможить було маска}$  принадлежит типу предложений  $ST_1$  оценивается как:

$$\begin{aligned} P(ST_1 / \text{допоможить, було, маска}) &\cong \\ &\cong 1 \cdot \frac{1}{3|ST(\text{'будь'})|} \cdot \frac{4}{5|ST(\text{'ласка'})|} \leq \frac{4}{15}. \end{aligned} \quad (8)$$

Значения  $|ST(\text{'будь'})|$  и  $|ST(\text{'ласка'})|$  определяются количеством типов предложения, в которых встретились соответствующие слова. Следовательно, гипотеза рассматриваемого ТП не отбрасывается, она имеет право на существование и при отсутствии других более правдоподобных гипотез может быть объявлена ответом интерпретации.

Для апробации предложенного метода была проведена серия экспериментальных исследований.

### Результаты экспериментов

Для исследования разработанной компьютерной модели были созданы две системы: экспериментальная и демонстрационная.

Экспериментальная система состоит из комплекса программных модулей. В рамках этой системы, распознавание фраз (предложений) осуществлялось декодером в условиях ограниченной (на основе LISP-структур), свободной и

относительно свободной (на основе фонетических слов) пословных грамматик [2, 4].

Для оценки параметров акустических моделей фонем использовался украиноязычный речевой корпус УкрЕко, содержащий записи 50 тыс. слов и фраз от более 150 дикторов. Этот корпус создан благодаря Гранту Президента Украины для одаренной молодежи № 58 по Распоряжению Президента Украины от 14 декабря 2005 г. № 1276/2006 рп. Параметры акустических моделей для декодера оценены на основе произнесенных 60 дикторами изолированных слов – подмножеств корпуса УкрЕко. Средствами [4] проведено обучение 55 акустических моделей фонем. Словари произношения сформированы для украинского и русского языков с аппроксимацией произношения слов украинскими фонемами.

Разработанная экспериментальная система используется для оценки надежности смысловой интерпретации. Предложенные в статье методы оценивания принадлежности последовательности слов к типу предложения были экспериментально проверены на фразах, взятых из разговорника, и произнесенных с внесением допустимых искажений в тексте фразы. Как примеры рассмотрены три ПО: Повседневные фразы, Путешествие, Гостиница. Эти ПО содержат  $47 + 102 + 68 = 217$  типов смысла. В среднем на тип смысла приходилось 4,17 типа предложения.

Для эксперимента использовался речевой сигнал 500 фраз, к которым были применены алгоритмы пофонемного распознавания речевых сигналов [3, 5] в условиях ограниченной, свободной и относительно свободной (на основе фонетических слов) пословных грамматик. Принятие решений по смысловой интерпретации для двух последних грамматик осуществлялось согласно статистическому оцениванию (1) – (7). Результаты экспериментов приводятся в табл. 3.

Из приведенных в табл. 3 результатов экспериментов можно сделать выводы, что для двух типов грамматик (ограниченной и относительно свободной) погрешность содержательной ин-

терпретации ТС не превышает четырех процентов, что приемлемо для прикладной системы.

Таблица 3. Результаты распознавания и смысловой интерпретации 500 предложений из двух предметных областей

Тип грамматики	Надежность распознавания (%)		
	слов	предложений	ТС
Ограниченная	97	94,1	98,3
Свободная пословная	54,3	4,2	86,1
Относительно свободная	79,1	20,8	96,2

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

Для демонстрации работы программных модулей на основе проведенных исследований была разработана демонстрационная программная модель с графическим интерфейсом пользователя (GUI) для перевода предложений в рамках предметных областей (рис. 4).

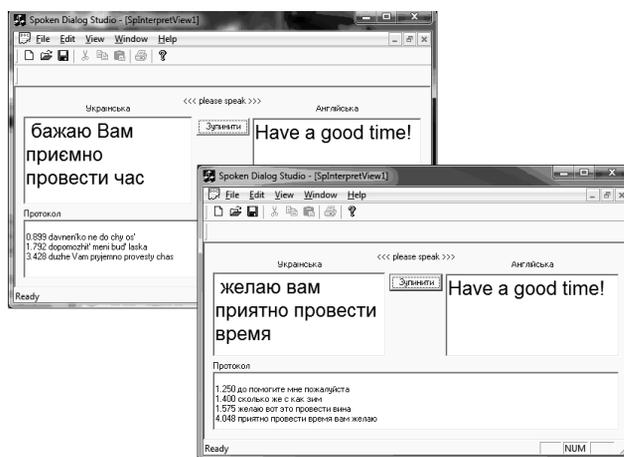


Рис. 4. Демонстрационная система устного переводчика с GUI

Предложенная демонстрационная система в настоящее время оперирует тремя языками (украинский, русский и английский) и 15 предметными областями, общий словарь для всех ПО составляет 2400 слов. Ввод речевого сигнала функционирует для украинского и русского языков. При этом порядок следования слов может быть любым из допустимых. Для произнесенного предложения находится наиболее подходящий для него тип предложения и соот-

ветствующий тип смысла, а первое предложение этого типа смысла объявляется результатом перевода на любой из языков вывода.

**Заключение.** Рассматриваемая система устного перевода представляет собой электронный аналог бумажного разговорника, взаимодействие с которым происходит наиболее естественным способом – голосом.

Разработанная концепция метаслов при моделировании типов предложений позволяет эксперту более эффективно формировать базы данных и знаний, охватывая практически все разнообразие выражения смысла. Предложенная обработка гипотетических ошибок декодера повышает достоверность принятия решения при смысловой интерпретации, особенно в условиях повышенных признаков спонтанности речи и в шумах. Реализованные усовершенствования бесспорно расширят сферу использования разработанной системы.

Одни и те же фразы с различной интонацией могут выражать как вопросительное предложение, так и повествовательное. Поэтому в дальнейшем предстоит исследовать возможность распознавания интонации и ритма (просодики) с целью автоматической расстановки

знаков препинания в распознанных фразах. Планируется также ставить в соответствие фразе исходного языка более конкретный переводной аналог среди типов предложений согласно типу смысла.

1. Винцюк Т.К. Анализ, распознавание и смысловая интерпретация речевых сигналов. – Киев. Наук. думка, 1987. – 264 с.
2. Сажок М., Яценко В. Система усного перекладу на основі інтерпретації мовленнєвого сигналу в межах предметних областей // Всеукр. міжнар. конф. з оброблення сигналів і зображень та розпізнавання образів УкрОбраз'2010 – Київ, 2010. – С. 103–106.
3. Lee T., Kawahara, Shikano K., Julius – an open source real-time large vocabulary recognition engine // Proc. Europ. Conf. on Speech Communication and Technology (EUROSPEECH). – 2001. – P. 1691–1694.
4. Сажок М., Яценко В. Порівняння послівних граматики для усного перекладу в межах предметної області // Всеукр. міжнар. конф. з оброблення сигналів, зображень та розпізнавання образів УкрОбраз'2008 – Київ, 2008. – С. 55–58.
5. *HTK Book*, v. 3.4 / S. Young, G. Evermann, M. Gales et al. // Cambridge University. – 2006. – 368 p.

Тел. для справок: +38 (044) 502-6333 (Киев)  
E-mail: yatsenko.valya@gmail.com, sazhok@gmail.com  
© В.В. Яценко, Н.Н. Сажок, 2013