

УДК 004.522

Учреждение Российской академии наук Санкт-Петербургский институт информатики и автоматизации РАН

Санкт-Петербург, 199178, 14 линия, 39.

<http://www.spiiras.nw.ru/speech>

А64 Анализ разговорной русской речи» (АР³-2010): Труды четвертого междисциплинарного семинара – СПб.: ГУАП, 2010.– 72 с.
ISBN 978-5-8088-0542-2

Издание представляет собой сборник докладов, сделанных на заседаниях четвертого междисциплинарного семинара «Анализ разговорной русской речи» (АР³-2010), проходившего 26 – 27 августа 2010 года в Учреждении Российской академии наук Санкт-Петербургском институте информатики и автоматизации РАН. Семинар посвящен обсуждению особенностей разговорной речи и возможных подходов к автоматическому анализу русской речи. Междисциплинарный подход к изучению речи позволит скорее продвинуться в моделировании речевой деятельности и решить фундаментальную проблему человеко-машинного диалога.

УДК 004.522

Статьи печатаются в авторской редакции.

ISBN 978-5-8088-0542-2

© СПИИРАН, 2010

© Коллектив авторов, 2010

© ГУАП, оформление 2010

А.В. Венцов, Е.И. Риехакайнен

Восприятие омофонов и структура ментального лексикона*

*Санкт-Петербургский государственный университет,
Санкт-Петербург, Россия,
av.ventsov@gmail.com, reha20@rambler.ru*

1. Введение

Ранее было показано [1, 2], что основным фактором, обеспечивающим надёжное распознавание редуцированных словоформ, является контекст. Однако в ряде случаев обращение к нему является невозможным или не позволяет выбрать единственную интерпретацию неоднозначного сигнала. В условиях естественной коммуникации это происходит тогда, когда неоднозначной оказывается самая первая словоформа в высказывании (и опора на контекст невозможна в принципе) и когда левый контекст не позволяет снять неоднозначность. В таких случаях, по-видимому, все единицы перцептивного словаря, активированные входным речевым сигналом, должны сохраняться в памяти слушающего до тех пор, пока контекст не позволит выбрать одну из них. Однако даже при отсутствии достаточного контекста одни словоформы являются более вероятными кандидатами на распознавание, чем другие, т.е. получают более высокую степень активации. Результаты распознавания редуцированных словоформ позволяют считать частотность тем фактором, который влияет на степень активации единиц при отсутствии достаточного контекста [3]. Тем не менее, выдвинутое предположение требует дальнейшей проверки, поскольку словоформы, которые выбирались в качестве ответов в экспериментах по распознаванию редуцированных словоформ, различались не только частотностью, но и степенью соответствия входному речевому сигналу.

В статье будут рассмотрены некоторые результаты исследования, направленного на проверку следующей нулевой гипотезы: при наличии нескольких единиц в перцептивном словаре, в равной степени соответствующих входному сигналу и допускаемых существующим контекстом, преимущество в процессе конкуренции получает та из них, которая является наиболее частотной.

2. Результаты проведенного эксперимента

В качестве экспериментального материала были использованы омофоны — «разные слова, совпавшие по звучанию, при различии морфонологического состава» [4: 288]. В словаре слушающего омофоничным отрезкам должны соответствовать разные по лексическому и/или грамматическому значению единицы, имеющие одинаковую фонемную структуру (т.е. в одинаковой степени соответствующие входному речевому сигналу). Предполагается, что в перцептивном словаре омофоны представлены одной единицей, обнаружение которой в акустическом сигнале приводит к немедленной активации информации следующих уровней (грамматических характеристик, частоты встречаемости) об обоих омофонах [5: 65]. При этом, на наш взгляд, в условиях отсутствия контекста, снимающего неоднозначность, единица, имеющая большую частотность, может быть активирована сильнее.

В качестве экспериментального материала были использованы 4 типа омофонов, одним из которых были возвратные глаголы прошедшего времени единственного числа с безударным окончанием /a/: [sl*ma+l*s']¹ (сломалось-сломалась), [stuča+l*s'] (стучалось-стучалась) и т.п. В рамках статьи будут рассмотрены только результаты, полученные на данном типе омофонов.

* Работа выполнена при частичной поддержке гранта РФФИ № 09-06-00244-а.

¹ Символ «+» в транскрипции указывает на ударность предшествующего гласного. Безударные аллофоны фонемы /a/ обозначаются в фонетической транскрипции знаком «*».

Отбор стимулов для тестовой последовательности осуществлялся с учётом относительной частотности омофонов в паре. Частотность словоформ определялась по полным данным основного корпуса Национального корпуса русского языка [6] (далее — НКРЯ; 162 132 407 словоупотреблений). Безусловно, соотношение текстов в данном корпусе (прежде всего то, что значительная его часть — это художественные тексты, а доля устной речи составляет лишь около 4%), время их создания (с середины XVIII века), а также наличие значительного количества ошибок не позволяют считать, что НКРЯ отражает реальную частоту встречаемости словоформ в русской речи. Однако поскольку некоторые из омофоничных словоформ являются очень редкими, определение их частотности по какому-либо из подкорпусов НКРЯ или по другим корпусам (например, по более сбалансированному Корпусу русского литературного языка, далее — КРЛЯ [7]) не представлялось возможным.

В условиях устно-письменного эксперимента на ответы испытуемых могла повлиять степень соответствия между звучанием словоформы и её написанием, поэтому для каждого из выбранных типов омофонов были подобраны следующие пары:

- омофон, орфографическая запись которого ближе к произношению, является более частотным, чем второй омофон данной пары;
- омофон, орфографическая запись которого ближе к произношению, является менее частотным, чем второй омофон данной пары.

Всего было отобрано 40 пар омофонов, 6 из них относились к типу, рассматриваемому в статье.

К сожалению, необходимый для проведения эксперимента материал невозможно было получить из записей спонтанной речи в силу отсутствия представительных корпусов естественной звучащей русской речи большого объёма. Поэтому тестовая последовательность была составлена из изолированных омофонов, которые были извлечены из предложений, прочитанных двумя дикторами (мужчиной и женщиной). Стимулы предъявлялись испытуемым для прослушивания через наушники. Каждый омофон звучал только один раз.

Испытуемые были разделены на две группы, первой из которых предлагалось просто записать услышанное. При такой инструкции межстимульный интервал составлял 5 секунд, эксперимент продолжался 4 минуты 25 секунд. Подобное задание достаточно близко к диктанту и, следовательно, при его выполнении роль степени соответствия между написанием и произнесением может оказаться значительной.

Задание, предложенное второй группе испытуемых, должно было уменьшить влияние данного фактора: участникам эксперимента предлагалось составить любую фразу с каждым стимулом. Если же они не могли этого сделать, им предлагалось записать хотя бы услышанное слово. Предполагалось, что, выполняя подобную инструкцию, испытуемые будут не просто механически записывать услышанное, а обращать внимание прежде всего на план содержания предъявляемых словоформ.

В данном случае пауза между стимулами составила 20 секунд, а продолжительность эксперимента — 14 минут 25 секунд.

В эксперименте приняли участие 85 человек в возрасте от 15 до 54 лет: 50 в первой группе и 35 — во второй. Испытуемыми были 43 женщины и 42 мужчины (соотношение для первой части 26/24, для второй — 17/18 соответственно).

Все испытуемые были наивными носителями языка: ни один из них не имел лингвистического/филологического образования и не являлся студентом филологического факультета.

Стимулы, которые представляли собой оканчивающиеся на безударный /а/ формы прошедшего времени единственного числа возвратных глаголов, были восприняты испытуемыми очень хорошо: в бланках ответов не встретилось ни одного отказа. При подсчёте результатов был вычислен процент форм женского и среднего рода в ответах испытуемых, а также доверительные интервалы на 5%-ном уровне значимости для всех полученных процентов.

Предположим, что в ментальном лексиконе носителя языка один из разделов образован возвратными глаголами прошедшего времени, а внутри последнего существуют подразделы, объединяющие словоформы, которые совпадают по числу и роду. Тогда можно было ожидать, что оценки этой группы стимулов испытуемыми 1-ой группы примерно поровну будут распределяться между словоформами женского и среднего рода, поскольку в корпусах разного объема количество соответствующих словоформ оказывается примерно одинаковым (см. Таблицу 1).

Таблица 1. Частотность глагольных форм на **-лась/-лось** в корпусах разного объема

	КРЛЯ		НКРЯ	
	<i>0,2 млн</i>	<i>1,05 млн</i>	<i>6 млн</i>	<i>162 млн</i>
Частота глаголов на -лась	459	2907	19943	447141
Частота глаголов на -лось	524	2908	19447	454092
Отношение частот	0,8759	0,9996	1,0255	0,9847

Однако, как следует из Таблицы 2, в 1-ой группе испытуемых по мере увеличения в НКРЯ отношения частоты словоформ женского рода к таковой для среднего рода процент форм женского рода в ответах систематически увеличивается, а процент форм среднего рода — уменьшается. Можно предположить, что в условиях данного эксперимента при оценке данного класса омофонов испытуемые ориентируются на частоту, с которой представлены в ментальном лексиконе соответствующие словоформы.

Таблица 2. Распознавание омофонов на **-лась/-лось**.

Стимул	Соотношение частот форм ж.р. и ср.р.		1-я группа, % ответов		2-я группа, % ответов	
	<i>абсолютн.</i>	<i>относит.</i>	<i>ж.р.</i>	<i>ср.р.</i>	<i>ж.р.</i>	<i>ср.р.</i>
[slučī+I*s']	1310 / 16743	0,078	6	94	5,7	94,3
[p*luči+I*s']	1544 / 6784	0,227	2	98	5,7	94,3
[pr*d*lža+I*s']	1793 / 2252	0,796	28	72	51,4*	48,6*
[sl*ma+I*s']	263 / 151	1,742	36*	64*	85,7	14,3
[z*kry+I*s']	618 / 99	6,242	44*	56*	77,1	22,9
[stuča+I*s']	63 / 6	10,5	60*	40*	71,4	28,6

*Звёздочкой отмечены те пары значений, которые достоверно не различаются.

Эти данные говорят также о том, что информация о частоте сопутствует представлению каждой словоформы в словаре и что носители языка при необходимости умеют ею пользоваться.

Наличие влияния суммарной частотности классов на **-лась** и на **-лось** можно предположить только для стимулов [sl*ma+I*s'], [z*kry+I*s'] и [stuča+I*s'], для которых количество форм женского и среднего рода в ответах 1-ой группы испытуемых достоверно не различается. Большая значимость данного фактора по сравнению с частотностью отдельных словоформ именно для данных стимулов может объясняться сравнительно низкой абсолютной частотностью омофонов, входящих в эти пары (см., например, [8]).

В ответах же 2-ой группы испытуемых не наблюдается столь строгой зависимости результатов распознавания от частотности словоформ. Видимо, при необходимости сконструировать фразу с услышанным омофоном испытуемый может опираться не только на частоту встречаемости соответствующей глагольной словоформы, но и на другие факторы.

Так, можно предположить, что в данном случае выбор формы глагола во многом определяется тем существительным или местоимением, которое испытуемый использует в качестве подлежащего при глаголе-стимуле. Например, большее количество форм женского рода в ответах 2-ой группы испытуемых по сравнению с 1-ой может объясняться большей частот-

ностью существительных женского рода по сравнению с существительными среднего рода (как в целом, так и только в именительном падеже): в подкорпусе НКРЯ со снятой омонимией количество существительных женского рода превосходит количество существительных среднего рода примерно в три раза.

Кроме того, значимым фактором может оказаться количество различных существительных и местоимений женского и среднего рода, которые могут употребляться в функции подлежащего при той или иной глагольной форме, а также соотношение их суммарных частотностей. В некоторых случаях, по-видимому, можно говорить о влиянии на выбор формы глагола частоты встречаемости конкретных существительных с конкретными глаголами в пределах одной фразы (например, *дверь закрылась, машина сломалась, это случилось, всё получилось*). Однако все сделанные предположения носят предварительный характер и требуют более детального изучения.

3. Заключение

Проведённый эксперимент в очередной раз свидетельствует о том, что ментальный лексикон слушающего представляет собой сложную многомерную структуру, элементы которой соединены связями, ведущую роль в формировании которых играет частотность.

При этом, несмотря на то, что полученные результаты позволяют сделать ряд интересных выводов и наблюдений, необходимо отдавать себе отчёт в том, что представления о частотности единиц в речи складываются на основе субъективного опыта каждого человека. Следовательно, соотношения частотностей единиц в ментальных лексиконах различных носителей языка могут различаться, и данные, извлечённые из любого, даже представительного и хорошо сбалансированного корпуса, позволяют получить лишь приблизительные представления о той частотности, которой пользуется конкретный носитель языка в процессе распознавания речи.

Литература

1. Риехакайнен Е.И. Стратегии восприятия редуцированных словоформ (на материале русской спонтанной речи) // Научные чтения — 2006. Материалы конференции. 13-14 ноября 2006. С.-Петербург. СПб, 2008. С. 52-59.
2. Венцов А.В., Риехакайнен Е.И., Слепокурова Н.А. Ментальный лексикон и восприятие редуцированных словоформ // Труды первого междисциплинарного семинара «Анализ разговорной русской речи» (АР³ – 2007). СПб, 2007. С. 21-24.
3. Риехакайнен Е.И. Роль фактора частотности в процессах порождения и восприятия редуцированных словоформ // VII выездная школа-семинар «Проблемы порождения и восприятия речи»: Материалы. Череповец, 2008. С. 53-60.
4. Ахманова О.С. Словарь лингвистических терминов. М., 2007.
5. Венцов А.В. Восприятие устной речи и ментальный лексикон // Русская языковая личность: Материалы шестой выездной школы-семинара. Череповец, 2007. С. 63-69.
6. Национальный корпус русского языка. URL: <http://www.ruscorpora.ru>
7. Корпус русского литературного языка. URL: <http://www.narusco.ru>
8. Alegre, M., Gordon, P. Frequency Effects and the Representational Status of Regular Inflections // Journal of Memory and Language 40 (1999). P. 41-61.