

УДК 81'42

О ВЛИЯНИИ ЧАСТОТНОСТИ КОЛЛОКАЦИИ ЛЕКСЕМ НА ВЗАИМОСВЯЗИ ЕДИНИЦ МЕНТАЛЬНОГО ЛЕКСИКОНА

Ирина Германовна Овчинникова
профессор кафедры речевой коммуникации
Пермского государственного университета
614990, Пермь, ул. Букирева, 15. iri-ovchinniko@yandex.ru

В статье рассматривается соотношение различных оценок частотности лексемы (в языке и в текстах) в аспекте значимости частотности лексем для формирования связей единиц ментального лексикона. Ментальный лексикон трактуется как сеть отношений, активируемых различными факторами. Исследование выполнено на материале коллокаций лексемы ЧЕЛОВЕК в Национальном Корпусе Русского Языка и данных различных ассоциативных словарей и экспериментов. Установлены обусловленные возрастом различия во входящих и исходящих связях лексемы в сознании носителей русского языка и определена мера соответствия силы взаимосвязей единиц ментального лексикона и частотности коллокаций слов в текстах.

Ключевые слова: корпус; ментальный лексикон; речевой онтогенез; вербальные ассоциации.

1. Введение¹

1.1. Актуальность исследования обусловлена необходимостью прояснить представление о ментальном лексиконе, весьма противоречивое в современной когнитивной науке. Между тем воспроизводящая модель ментального лексикона необходима для компьютерного моделирования речевой деятельности.

1.2. **Цель** данного исследования – установить влияние частотности коллокаций лексем в речи на предсказуемость их связей в ментальном лексиконе. Соответственно, свои задачи мы видим в характеристике современных моделей ментального лексикона, обсуждении фактора «частотность слова» (субъективная и объективная), в установлении корреляции между речевой частотностью (частотой употребления) и предсказуемостью взаимосвязей с другими единицами в ментальном лексиконе.

1.3. **Исходная гипотеза исследования.** Мы полагаем, что СМИ и разговорная речь оказывают непосредственное влияние на взаимосвязи единиц ментального лексикона, отраженные в ассоциативных связях, и именно частотные контексты закрепляются в ассоциативных реакциях. В свою очередь, частотные ассоциации «всплывают» при планировании речевого высказывания, актуализируясь в процессе лексического выбора и заполняя слоты синтаксического фрейма высказывания. Поскольку контекст развертывается слева направо, то выбор последующих лексем в речевой цепи отчасти обусловлен ассо-

циациями предшествующей лексемы. Наличие такого влияния установлено на отрицательном языковом материале [Ахутина 2002] и учитывается в вероятностных моделях речевой деятельности, прежде всего в коннекционистских [MacWhinney 2005]. Можно предположить, что, задавая в ассоциативном эксперименте в качестве стимулов лексемы «левого» контекста, мы получим ассоциативные реакции, совпадающие с «правым» контекстом. Наличие среди ассоциаций слов из «правого» контекста можно рассматривать в качестве отражения влияния текстов на формирование взаимосвязей единиц ментального лексикона. Между тем пока не выявлена значимость такого влияния в зависимости от целого ряда лингвистических и экстралингвистических факторов.

1.4. **Материал исследования** представляют данные ассоциативных экспериментов (более 3000 ассоциативных реакций) и около 700 контекстов из Национального корпуса русского языка (www.ruscorpora.ru/) и Национального корпуса русского литературного языка (www.narusko.ru/). Мы ограничились анализом ассоциативного поля лексемы ЧЕЛОВЕК, поскольку семантическое поле ЧЕЛОВЕК актуально для языкового сознания детей и взрослых носителей русского языка. Для носителей русского языка ЧЕЛОВЕК – это центральное понятие языкового сознания [Овчинникова и др. 2000]. Источником ассоциативных реакций послужили данные письменных ассоциативных экспериментов, опубликованные в

русских ассоциативных словарях [Русский... 1994-2002].

Процедура исследования заключалась в сопоставлении ассоциативных реакций детей и взрослых на слово-стимул ЧЕЛОВЕК (исходящие связи), в определении стимулов, на которые возникает частотная реакция *человек* (входящие связи), в создании подкорпуса СМИ и разговорной речи на основе корпусов русского языка, в сопоставлении контекстов лексемы человек с данными ассоциативных экспериментов, отражающими входящие и исходящие связи лексемы в ментальном лексиконе. Таким образом, мы использовали семантический анализ, дискурсивный анализ, дистрибутивный анализ и, для установления статистической достоверности полученных результатов, корреляционный анализ.

2. Характеристика современных моделей ментального лексикона

2.1. В современных моделях речевой деятельности и языковой способности ментальный лексикон чаще всего рассматривают в качестве хранилища декларативных знаний, как аналог семантической памяти. В таком случае последовательно проводится разграничение декларативных и процедурных знаний, причем лексикон ассоциируется с декларативными знаниями, а грамматика – с процедурными знаниями. Между тем экспериментальные данные свидетельствуют о том, что морфологическая и синтаксическая (т.е. грамматическая) информация о слове представлена в ментальном лексиконе. В частности, можно обнаружить сведения о специальном доступе к корневым морфемам освоенных лексем, что позволяет носителю языка не только морфологически идентифицировать слово, но и распознавать случаи омонимии и определять семантическое сходство слов [Pylkkanen et al. 2006]. «Синтаксический портрет» слова фиксируют аффиксы (суффиксы и окончания), синтаксические характеристики лексемы проявляются также в ассоциациях на основе контекстных связей. Таким образом, в ментальном лексиконе, трактуемом как хранилище декларативных знаний, все же доступны грамматические характеристики слова.

2.2. Обычно пишут о двух основных способах моделирования ментального лексикона. Во-первых, представляют ментальный лексикон в качестве иерархической сети; во-вторых – в качестве сети с распространяемой активацией, в которой таксономические связи отнюдь не выполняют главенствующую роль, а расстояния между узлами зависят от многих факторов [Elman 2004]. Слово хранится в ментальном лексиконе как сеть взаимосвязей с другими единицами. Последняя модель оказывается более адек-

ватной и позволяет интерпретировать разнообразные наблюдения над нормой и патологией, а также разнородные экспериментальные данные. Несмотря на взаимосвязь и взаимозависимость организации ментального лексикона и способа извлечения единиц из него, проблему способа доступа к этому хранилищу декларативных знаний решают отдельно.

В ментальном лексиконе представлена семантическая сеть. Наряду с ней в некоторых моделях выделяют фонологические и орфографические сети, каждая из которых связана с узлами семантической сети. Фонетическое или орфографическое сходство слов активует несколько узлов семантической сети. Благодаря исследованию *tip of the tongue phenomena* («слово на кончике языка») доказана автономность хранения плана выражения от плана содержания лексической единицы [Brown, McNeill 1966]. Вероятно, хранилище значений – одно, востребованное при любой форме коммуникации (устной и письменной). Однако активация связей варьирует в зависимости от типа речевой деятельности (порождения или восприятия речи).

Экспериментальное изучение распознавания многозначных слов позволило установить, что план выражения (звучание или написание слова) является ключом к одновременной активации всех значений многозначного слова при его восприятии вне контекста [Pylkkanen et al. 2006]. Иначе говоря, узел фонологической (орфографической) сети активует паутину в сети семантической. Такой доступ к разнообразной семантической информации обеспечивает возможность выбора необходимого варианта в неопределенных условиях, с одной стороны, и может быть истолкован в качестве доказательства отдельности семантической сети, самостоятельности лексического модуля языковой способности, с другой.

На активацию семантической сети оказывает влияние фактор частотности совместной встречаемости слов в речи: у часто встречающихся слов обнаружены сильные взаимные связи [Elman 2004]. В частности, разграничение ассоциативных и семантических связей между словами проводят на основе частотности коллокаций членов ассоциативной пары: ассоциации обеспечены высокой частотой совместной встречаемости, семантические связи – наличием общих сем, принадлежностью к одной категории.

2.3. В коннекционистских работах ментальный лексикон рассматривается в качестве сети отношений единиц, обеспечивающих обработку событий и поведения вне режима реального времени [Tabor & Tanenhaus 2001]. Такая сеть способна

«обучаться» прогнозированию, т.е. предсказанию последующего элемента речевой цепи, учитывая не только контекстные характеристики; в основе обучения сети – способность сравнивать текущую информацию с воспринятой ранее и сохраненной. В сущности, в данном случае обсуждается не лексикон. Дж.Элман полагает, что в сети нет такого компонента, как лексикон, т.е. хранилище знаков, в классическом смысле [Elman 2004]. В сети лексическое знание – знание слов и значений – скрыто в том влиянии, которое слова оказывают на ментальные репрезентации (внутренние состояния), возникающие в ответ на слово. Причем слова представляют собой очень важные, но не единственные детерминанты ментальных репрезентаций. Таким образом, слово рассматривается не в качестве носителя значения, а в качестве средства доступа к ментальным репрезентациям непосредственно [Ibid].

3. Взаимосвязи лексики ЧЕЛОВЕК в ментальном лексиконе носителя русского языка

Обратимся к обсуждению взаимосвязей лексики ЧЕЛОВЕК, установленных на основе данных письменного ассоциативного эксперимента. Стимулы для эксперимента были получены в результате обработки коллокаций лексики ЧЕЛОВЕК в подкорпусе, выбранном нами из национальных корпусов русского языка.

В результате ассоциативного эксперимента с восьмилетними детьми и данных детского ассоциативного словаря [Овчинникова и др. 2000] обнаружены следующие характеристики ЧЕЛОВЕКА, актуальные для детского сознания. В ходе анализа было выявлено 38 стимулов, на которые *человек* появляется в качестве частотной реакции. По результатам ассоциативного эксперимента можно сказать, что реакция *человек* в соответствии с закономерностями свободного ассоциирования появляется прежде всего на стимулы-прилагательные (АККУРАТНЫЙ, БЕЗГРАМОТНЫЙ, ВЕЖЛИВЫЙ, ЖАДНЫЙ, ЛЮБОПЫТНЫЙ, НАДЕЖНЫЙ, НЕВОСПИТАННЫЙ, СЛАБЫЙ, СПОКОЙНЫЙ, УМНЫЙ, ХИТРЫЙ, ЦЕННЫЙ), немногим реже – на стимулы-существительные (ВРАГ, ГОРДОСТЬ, ДОБРОТА, СУДЬБА, ТАЛАНТ, ЧЕСТЬ) и совсем нечасто – на стимулы-глаголы (ДРУЖИТЬ, ЛЕНИТЬСЯ, ОБИЖАТЬСЯ, ОБМАНЫВАТЬ). Таким образом, из 38 стимулов, на которые встретилась реакция ЧЕЛОВЕК, прилагательные составляют 42,9%, существительные – 39,2%, а глаголы только 17,9%.

В ментальном лексиконе ребенка представлена прежде всего положительная оценка ЧЕЛОВЕКА по различным параметрам (физиче-

ские качества, характер, внешние характеристики).

Обратимся к анализу материалов ассоциативных экспериментов с взрослыми носителями языка (см. Русский ассоциативный словарь, 1994-2002). Насколько мы можем судить, в ментальном лексиконе детей и взрослых носителей русского языка представлены сходные характеристики ЧЕЛОВЕКА (*умный, друг, гордость, верность, надежный, любопытный, мечтать*). На основе обработки данных ассоциативных экспериментов с взрослыми носителями языка мы установили актуальность оценки ЧЕЛОВЕКА по параметрам: внешность, социальные характеристики, действия, возраст. В качестве существенного различия отметим многомерность образа ЧЕЛОВЕКА в ментальном лексиконе взрослых. Многомерность отражает, во-первых, разнообразие семантических связей: здесь можно увидеть и крылатые фразы, и устойчивые выражения, например, *человек – это звучит гордо, человек – амфибия, человек – загадка* и др., и трансформированные выражения, например, *человек – человеку друг* от лат. *Homini lupus est* – человек человеку волк. Во-вторых, заметна большая вариативность грамматических связей: у взрослых в качестве реакций присутствуют не только существительные, прилагательные и глаголы, но и местоимения (*он, у него, наши*), числительные (*первый, последний*) и наречия (*всегда, хорошо, уже*).

Как видим, лексема ЧЕЛОВЕК обладает разнообразными входящими связями и может быть отнесена к ядру вербально-ассоциативной сети. Исходящие связи лексики довольно предсказуемы, поскольку в ассоциативном поле присутствуют сравнительно частотные реакции.

4. Сопоставление данных ассоциативного эксперимента с контекстами из Корпуса: оценка влияния фактора «частотность коллокации»

Мы оценивали меру взаимозависимости частоты коллокаций и частотности реакции в ассоциативном эксперименте с детьми и взрослыми методом корреляционного анализа. Значения коэффициентов корреляции представлены в табл.1.

Величина коэффициента корреляции варьирует от -1 до +1. При положительных значениях коэффициента говорят о положительной взаимосвязи: с увеличением или уменьшением одного признака происходит соответствующее изменение другого. При отрицательных значениях – об отрицательной: с увеличением одного признака другой уменьшается, и наоборот. Сильную корреляцию показывает коэффициент не ниже 0,7; среднюю – от 0,69 до 0,5; умеренную – от 0,49 до

0,3; слабую – от 0,29 до 0,2; ниже 0,2 – очень слабую, практически отсутствующую. Значения коэффициента корреляции частоты реакции ЧЕЛОВЕК в эксперименте и данных Корпуса показывают высокую корреляцию для взрослых носителей языка и среднюю – для восьмилетних детей. Такие результаты кажутся весьма достоверными. Взрослые носители языка обладают большим коммуникативным опытом, в том числе – в восприятии речи в СМИ и обиходном общении. Опыт речевой перцепции детей много скромнее, соответственно, и корреляция частот ниже. Слабая корреляция частотных характеристик совпадающих ассоциативных связей слова ЧЕЛОВЕК у взрослых и детей отражает существенность различий в структуре ментального лексикона, обусловленную накопленным коммуникативным и когнитивным опытом, а также нейропсихологическими особенностями детского возраста.

**Таблица 1
Значения коэффициентов корреляции**

Значения коэффициента корреляции	Корпус	Взрослые	Дети
Корпус		0,842096	0,627694
Взрослые	0,842096		0,391692
Дети	0,627694	0,391692	

5. Обсуждение результатов исследования

5.1. Полагаем, что можем интерпретировать пересечения характеристик ЧЕЛОВЕКА в сознании детей и взрослых как принципиальное совпадение базовых компонентов картины мира. Устойчивые характеристики обеспечивают принципиальную возможность успешной коммуникации между членами языкового коллектива, принадлежащими к разным поколениям.

5.2. Очевидно, фактор частотности коллокаций стоит признать существенным для формирования входящих и исходящих взаимосвязей в ментальном лексиконе. Между тем частотность совместной встречаемости лексем в тексте отчасти зависит от собственной частоты каждой из лексем. Чем выше частотность лексем, тем выше теоретическая вероятность ее совместной встречаемости с любой другой лексемой. Рассмотрим влияние объективной и субъективной частотности лексем на ее функцию в ментальном лексиконе.

5.2.1. Мы не обнаружили достаточно достоверных данных о влиянии объективной частотности слова на его репрезентацию в ментальном лексиконе ребенка или взрослого носителя язы-

ка. Установлено, что частотность существенна при восприятии и идентификации слова для семилетних детей, для взрослых испытуемых этот фактор статистически не значим [Smith et al. 2006]. Вероятно, семилетние дети чувствительны к объективной частотности в силу ограниченности их актуального лексикона по сравнению с более старшими носителями языка: чем меньше единиц освоено, тем существеннее характеристика каждой из единиц. Семилетние дети достоверно хуже распознают редкие слова по сравнению с частыми, более старшие носители языка при распознавании слова опираются на сформированный механизм вероятностного прогнозирования, что позволяет снизить влияние фактора «объективная частотность слова».

5.2.2. Субъективная частотность слова, по видимому, влияет на его взаимосвязи с другими единицами в индивидуальном лексиконе. Между тем, в исследовании И.А.Углановой не обнаружено статистически значимой зависимости между типом ассоциативной реакции и временем реакции от частотной характеристики слова. Исследователь замечает, что «фактор частоты проявляет себя лишь в тех категориях, которые содержат незнакомые или малознакомые слова, т.е. тогда, когда за словом не стоит сеть отношений, тогда и обнаруживаются типологические (тип реакции) флуктуации. Относительно времени реакции: частота, безусловно, должна на него влиять. Отсутствие корреляции, возможно, вызвано ситуацией эксперимента» [Угланова 2004: 20]. Таким образом, у малочастотных, т.е. малознакомых и незнакомых, слов тип ассоциативной реакции отражает колебание выбора направления поиска актуальных взаимосвязей. Очевидно, для малочастотных слов более актуальны нетипичные способы реагирования, отступления от поиска семантических, ситуативных или контекстных связей. Иначе говоря, малочастотные слова не обладают надежными исходящими связями.

5.2.3. Влияние субъективной частотности слова на количество входящих или исходящих связей неоднозначно. В целом для высокочастотных слов характерно большое количество входящих связей. Однако статистическая значимость взаимосвязи переменных «субъективная частотность слова» и «количество стимулов, активирующих данное слово в качестве ассоциативной реакции» пока не доказана.

5.2.4. Таким образом, влияние фактора частотности слова на его позицию в вербальной сети, на количество входящих или исходящих связей нельзя описать в виде простой непосредственной

зависимости. В целом, в качестве тенденции отметим:

- субъективная частотность слова в большей мере значима для структуры его связей в ментальном лексиконе, чем объективная;
- слова, которые носитель полагает субъективно частотными, обладают большим количеством входящих взаимосвязей;
- слова, которые носитель языка считает малочастотными, характеризуются нетипичным направлением поиска исходящих связей.

5.3. Частотность как характеристика лексемы, очевидно, не является достаточной для определения ее места в сети отношений, в ассоциативно-вербальной сети. Частотность, особенно объективная частотность, – слишком «глобальная» характеристика, не проявляющаяся непосредственно в организации взаимосвязей в ментальном лексиконе. Влияние субъективной частотности слова на количество входящих и исходящих связей, по-видимому, необходимо исследовать более целенаправленно. Частота совместной встречаемости, в отличие от частоты отдельной лексемы, играет достаточно заметную роль в формировании вербально-ассоциативных отношений. Это фактор оценивают в качестве весьма существенного, поскольку частотность совместной активации является основой для создания прочных нейронных связей.

¹Исследование выполняется при поддержке РГНФ10-04-00139а «Вариативность новостных текстов в аспекте информационного анализа: разнообразие речевых жанров, выражение функционально-грамматических категорий, формальные критерии выделения и обработки».

Список литературы

Ахутина Т.В. Типология ассоциативных реакций в норме и патологии: материалы конференции, посвященной 100-летию

А.Р.Лурия «А.Р.Лурия и психология XXI века». М., 2002. <http://www.auditorium.ru/>

Овчинникова И.Г., Береснева Н.И., Дубровская Л.А., Пенягина Е.Б. Лексикон младшего школьника. Пермь, 2000.

Русский ассоциативный словарь / под ред. Ю.Н. Караулова. М., 1994-2002.

Угланова И.А. Субъективная оценка частоты слова и ее категоризация: онтолингвистический аспект: автореф. дис. ... канд. филол. наук. Пермь, 2004. 24 с.

Bate E., Goodman J.C. On the Insuperability of Grammar and the Lexicon: Evidence from Acquisition. In: Tomasello, M. & Bates, E. (eds), *Essential Readings in Developmental Psychology*. Oxford, 2001. P.134-162.

Brown R., McNeill D. The "tip of the tongue" phenomenon // *Journal of Verbal Learning and Verbal Behavior*, 1966, №5. P.325-337.

Elman J. An alternative view of the mental lexicon. *Trends in Cognitive Science // Trends in Cognitive Sciences*. 2004. Vol.8. №7. P.301-306.

MacWhinney B. *New Directions in the Competition Model // Beyond Nature-Nurture: Essays in Honor of Elizabeth Bates*. Mahwah, New Jersey, London, 2005. P.81-110.

Pylkkanen L., Llinas R., Murphy G. The Representation of Polysemy: MEY Evidence // *Journal of Cognitive Neuroscience*, 2006, Vol.18, №1. P.97-109.

Smith Ph., Turner J., Brown P., Henry L. The distinct contributions of age of acquisition and word frequency in auditory word perception // *The Quarterly Journal of Experimental Psychology*, 2006, №59 (12). P.2121-2134.

Tabor W., Tanenhaus M.K. Dynamical system for sentence processing // *Connectionist Psycholinguistics* (Christiansen, M.H. & Chater, N., eds.). Ablex Publishing, 2001. P.177-211.

HOW FREQUENCY OF COLLOCATIONS INFLUENCES ON RELATIONS OF UNITS IN THE MENTAL LEXICON

Irina G. Ovchinnikova

Professor of Speech Communication Department
Perm State University

The article is concerned with the correlation between lexemes frequency and attractive peculiarities of their representations in the mental lexicon. Based on Corpora data and results of free associative tests the author evaluates the correlations between frequency of lexemes and the probability of activation associative link between them. The difference in the organization of associative links of the lexeme MAN in the mental lexicon of Russian speaking children and adults is discussed. The author shows that only collocations in texts have essential influence of the strength of associative links for children and adults as well.

Key words: Corpora; mental lexicon; association; correlations; speech ontogenesis.