

**СИНХРОННОЕ И ДИАХРОННОЕ
В
СРАВНИТЕЛЬНО-ИСТОРИЧЕСКОМ
ЯЗЫКОЗНАНИИ**



STUDIA ACADEMICA

**МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ М. В. ЛОМОНОСОВА**

Филологический факультет

**СИНХРОННОЕ И ДИАХРОННОЕ
В
СРАВНИТЕЛЬНО-ИСТОРИЧЕСКОМ
ЯЗЫКОЗНАНИИ**

*Материалы VII Международной научной конференции
по сравнительно-историческому языкознанию
(Москва, 31 января – 2 февраля 2011 г.)*

Под общей редакцией В.А. Кочергиной

«studia academica»

ДОБРОСВЕТ
Издательство «КДУ»
Москва
2011

УДК 801
ББК 81
К75

*Печатается по постановлению Редакционно-издательского совета
филологического факультета МГУ имени М. В. Ломоносова*

Рецензенты:

д.ф.н., проф. А. А. Волков к.ф.н. доцент Е. А. Брызгунова

Редакционная коллегия:

В. А. Кочергина (главный редактор), И. И. Богатырева,
О. А. Волошина, В. К. Казарян

К75 Синхронное и диахронное в сравнительно-
историческом языкознании — М.: «Добросвет», «Изда-
тельство „КДУ“», 2011. — 280 с.

ISBN 978-5-98227-791-6

В сборник включены статьи, представляющие собой материалы VII Международной научной конференции по сравнительно-историческому языкознанию «Синхронное и диахронное в сравнительно-историческом языкознании», проводимой кафедрой общего и сравнительно-исторического языкознания филологического факультета МГУ имени М.В. Ломоносова (Москва, 31 января — 2 февраля 2011). Сборник предназначен для компаративистов, а также для специалистов в различных областях языкознания.

УДК 801
ББК 81

ISBN 978-5-98227-791-6

© Авторы, 2011.
© «Добросвет», 2011.

**Вывод закона синхронного полисемического
распределения языковых знаков на основе диссипа-
тивной стохастической динамической модели эволю-
ции знаковых ансамблей**

ПОДДУБНЫЙ В.В. (ТОМСК)
ПОЛИКАРПОВ А.А. (МОСКВА)

В работах [1, 2] представлена эволюционная модель жизненного цикла языкового знака – от момента его зарождения до момента выхода из употребления. Жизненный цикл знака определяется двумя основными процессами: процессом роста полисемии знака (приобретения знаком новых, как правило, всё более абстрактных значений) и процессом постепенного выхода из употребления ранее приобретённых значений, начиная с наименее абстрактных. Способность знака к порождению новых значений была названа его ассоциативно-семантическим потенциалом (АСП), она измеряется максимальным количеством значений знака, которые он способен породить в ходе его жизненного цикла. Первый процесс, процесс приобретения знаком новых значений, постепенно замедляется по мере уменьшения («растрачивания») АСП. Второй процесс, процесс потери ранее приобретённых значений, начинается с некоторым запаздыванием по отношению к первому и протекает аналогично, но более медленно. Разность между количеством приобретённых знаком значений и количеством значений, вышедших из употребления к данному моменту времени, составляет размер актуальной полисемии знака, т.е. количество живущих в этот момент времени значений знака. Кривая развития этого процесса во времени для каждого знака – унимодальная кривая с максимумом, смещённым к началу процесса.

Языковые знаки в общем ансамбле знаковых единиц данного уровня (например, лексического) различаются величинами их АСП и, соответственно, параметрами кривой развития их полисемии. Более того, есть основания допустить, что в любом знаковом ансамбле того или иного национального языка знаки по их АСП находятся не в хаотическом, а во вполне закономерном статистическом соотношении. То есть, АСП знаков распределены сообразно некоторому закону, а именно – согласно экспоненциальному (показательному) закону. Если это так, то и конфигурация пучка полисемических траекторий разных знаков, появляющихся в данном языке в некоторый период времени и последовательно развивающихся (реализующих их АСП), тоже должна определяться этим законом.

Рассмотрение одномоментного сечения всех полисемических траекторий всех знаков, продолжающих жить (т.е. продолжающих приобретать и/или терять ранее приобретённые значения) даёт картину синхронного соотношения знаков по величине актуальной полисемии – числу присущих им (в данный момент) значений. Эмпирическая статистика синхронного (одномоментного) распределения полисемии всего ансамбля лексических знаков того или иного языка на сегодняшний день содержится в словарях или корпусах текстов.

Возникает вопрос, может ли некоторая математическая модель процесса эволюции языковых знаков прогнозировать синхронное (одномоментное) распределение полисемии, адекватное эмпирическим распределениям, полученным, например, из представительных словарей разных языков?

Наш ответ заключается в том, что если в эволюционной (диахронической) модели учтены действительно принципиальные особенности, наиболее фундаментальные закономерности реальной жизни,

реального соотношения микро-объектов (лексем) по их динамическим свойствам в изучаемом макро-объекте (словаре), то и теоретически реконструируемые на этой основе синхронные срезы того или иного состояния словарного ансамбля (при условии стационарности моделируемого процесса, стабильности его внешних условий) должны быть близки тому, что есть в реальности.

Для того чтобы осуществить указанный прогноз (а также для математического моделирования и других существенных характеристик процессов и состояний в лексике и в других знаковых ансамблях языка – морфемных и фразеологических) нами предложена [3] диссипативная стохастическая динамическая модель развития языковых знаков, удовлетворяющая принципу «наименьшего действия», одному из фундаментальных вариационных принципов природы. Модель предполагает пуассоновский характер потока рождения языковых знаков, экспоненциальное (показательное) распределение знаков по их АСП и оперирует разностными стохастическими уравнениями специального вида, вытекающими из принципа наименьшего действия для диссипативных процессов.

Выпишем эти уравнения для i -го знака ансамбля, снабдив процесс рождения новых значений верхним индексом 1, а процесс выпадения значений из употребления – верхним индексом 2:

$$\begin{aligned} t_{i,k+1}^{(1,2)} &= t_{i,k}^{(1,2)} + \tau_i^{(1,2)} / (G_i - k) + \xi^{(1,2)}, & k &= \overline{1, G_i - 1}, \\ t_{i,1}^{(1)} &= t_i, \quad t_{i,1}^{(2)} = \tau_{0i} + t_i, \quad \tau_i^{(2)} > \tau_i^{(1)}, & i &= \overline{1, N}. \end{aligned} \quad (1)$$

Здесь $t_{i,k}^{(1,2)}$ – момент рождения (индекс 1) или выхода из употребления (индекс 2) k -го значения i -го знака, G_i – случайный АСП i -го знака (распределён по показательному закону со средним значением $\langle G \rangle$), t_i – случайный момент появления в языке i -го знака, так что $t_{i+1} = t_i + \tau$, где τ – случайный интервал времени между появлением соседних знаков

(распределён по экспоненциальному закону со средним значением $\langle \tau \rangle$), τ_{0i} – случайное запаздывание начала процесса выхода из употребления значений i -го знака по отношению к моменту появления знака в языке (распределено по экспоненциальному закону со средним $\langle \tau_0 \rangle$), $\tau_i^{(1,2)} = c^{(1,2)} \langle G \rangle / G_i$ – постоянные времени процессов рождения новых значений знака и выхода их из употребления (обратно пропорциональны АСП, так что они случайны и их распределение определяется распределением АСП), $\xi^{(1,2)}$ – случайные флуктуации моментов рождения и выхода из употребления значений знака (распределены по равномерному закону с нулевыми средними и полуширинами интервалов распределения $\tau_i^{(1,2)} / (G_i - k)$), N – число знаков (слов) в языке. Очевидно, $L_{i,k} = t_{i,k}^{(2)} - t_{i,k}^{(1)}$ – длительность жизни k -го значения i -го знака. Нетрудно видеть, что эта величина подчиняется рекуррентному соотношению:

$$\begin{aligned} L_{i,k+1} &= L_{i,k} + (\tau_i^{(2)} - \tau_i^{(1)}) / (G_i - k), \\ L_{i,1} &= \tau_{0i}, \quad k = \overline{1, G_i - 1}, \quad i = \overline{1, N}, \end{aligned} \quad (2)$$

так что $L_{i,k+1} > L_{i,k}$ т.е. длительность жизни каждого значения любого i -го знака увеличивается с ростом k . Возраст k -го значения i -го знака в момент времени t равен $A_{i,k}(t) = t - t_{i,k}^{(1)}$, если в момент времени t это значение существует (уже появилось, но ещё не вышло из употребления, т.е. если $t_i < t < t_{i,G_i}^{(2)}$).

Полисемия i -го знака развивается с момента $t_{i,1}^{(1)} = t_i$ появления i -го знака в языке до момента $t_{i,G_i}^{(2)} = t_i + \tau_{0i} + L_{i,G_i}$ выхода из употребления последнего (G_i -го) значения i -го знака. Интервал времени длиной $\tau_{0i} + L_{i,G_i}$ от $t = t_{i,1}^{(1)}$ до $t = t_{i,G_i}^{(2)}$ есть интервал жизненного цикла i -го знака.

Модель (1) является диссипативной стохастической динамической моделью эволюции ансамбля языковых знаков. В этой модели 5 параметров:

- интенсивность потока новых знаков $1/\langle\tau\rangle$ (или средний интервал времени $\langle\tau\rangle$ между появлением соседних знаков в потоке),
- среднее запаздывание начала процесса выхода значений знака из употребления по отношению к моменту появления знака в потоке $\langle\tau_0\rangle$,
- среднее значение АСП $\langle G\rangle$,
- коэффициенты $c^{(1)}$, $c^{(2)}$ ($c^{(1)} \ll c^{(2)}$) обратной пропорциональной зависимости постоянных времени процессов рождения и выхода из употребления значений знака от его АСП (приблизительно равны средним значениям $\langle\tau^{(1)}\rangle$, $\langle\tau^{(2)}\rangle$ этих постоянных времени, причём $\langle\tau^{(1)}\rangle \ll \langle\tau^{(2)}\rangle$).

Для проверки адекватности модели (1) на материале лексических знаков использовалось 5 представительных словарей русского и английского языков (3 русского и 2 английского) разных типов (типология толковых словарей представлена в работе [4]):

ССРЛЯ – "Словарь современного русского литературного языка" в 17-ти тт. (1948–1965), (большой по типу словарь);

МАС – "Словарь русского языка" в 4-х тт. под ред. А. П. Евгеньевой (1957–1961), (средний по типу словарь);

СО – "Словарь русского языка" С. И. Ожегова (1972, 9-е издание), (краткий по типу словарь);

Shorter – "Shorter Oxford English Dictionary" (1962), (средний по типу словарь);

Hornby – A.S. Hornby. "Oxford Advanced Learner's Dictionary of Current English" (1982), (краткий по типу словарь).

На рис. 1 представлено семейство эмпирических кривых распределения полисемии, полученных из указанных словарей.

На рис. 2–6 для каждой из эмпирических кривых распределения полисемии, полученных из словарей, приведены кривые распределения

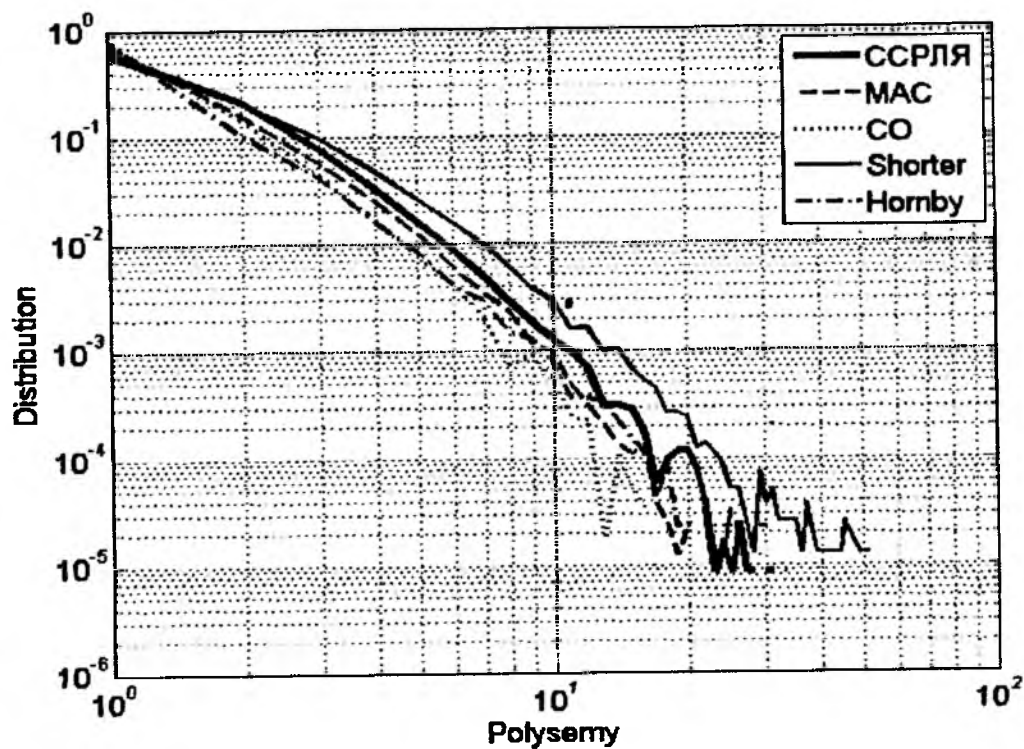


Рис. 1. Семейство эмпирических распределений лексической полисемии по 5 словарям русского и английского языков

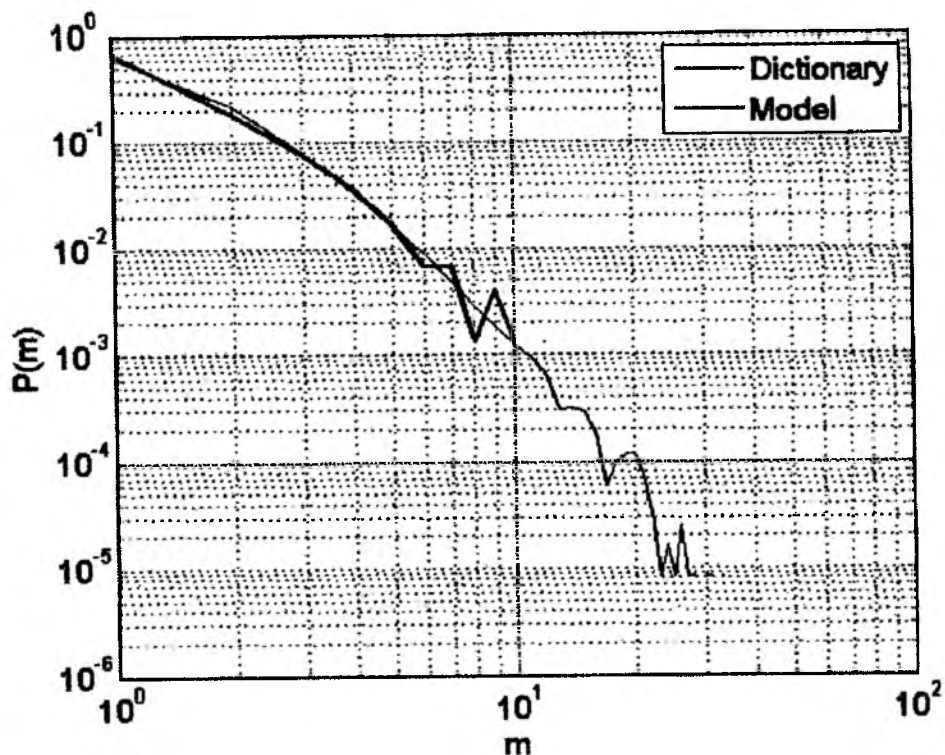


Рис. 2. Распределения лексической полисемии по словарю ССРЛЯ и по модели
 $m_M = 1,6388$, $m_D = 1,6973$, $\sigma_M = 1,286$, $\sigma_D = 1,3998$, $H = H_0$: $p = 0,99133$

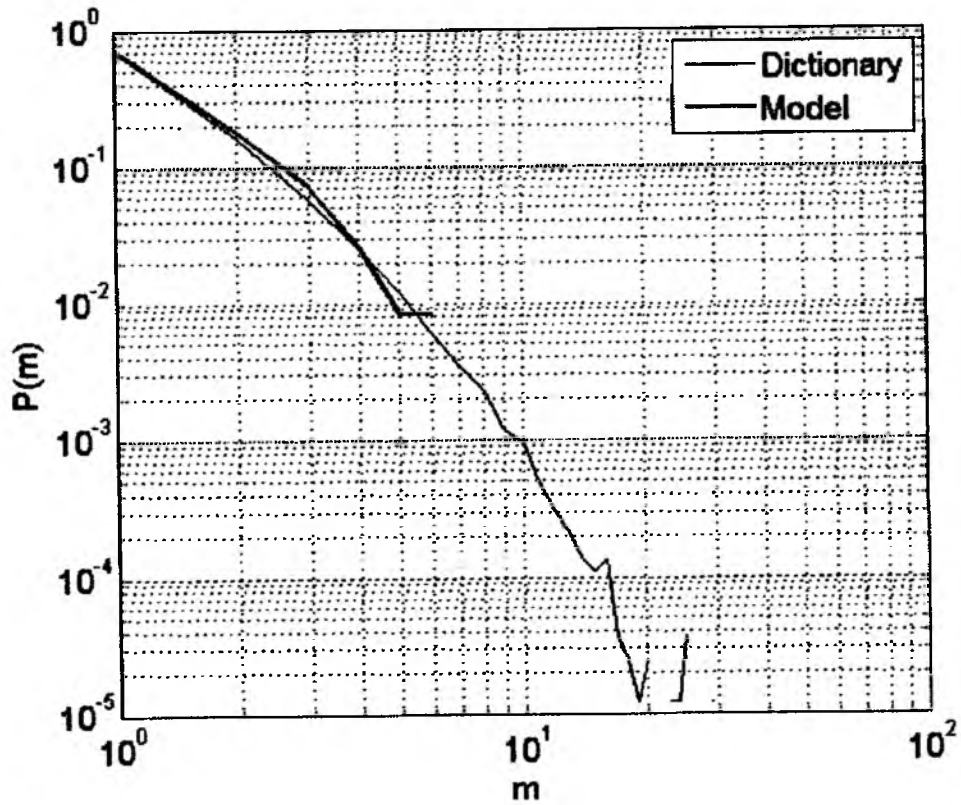


Рис. 3. Распределения лексической полисемии по словарю МАС и по модели
 $m_M = 1,4697$, $m_D = 1,503$, $\sigma_M = 0,89921$, $\sigma_D = 1,1651$, $H = H_0$: $p = 0,99996$

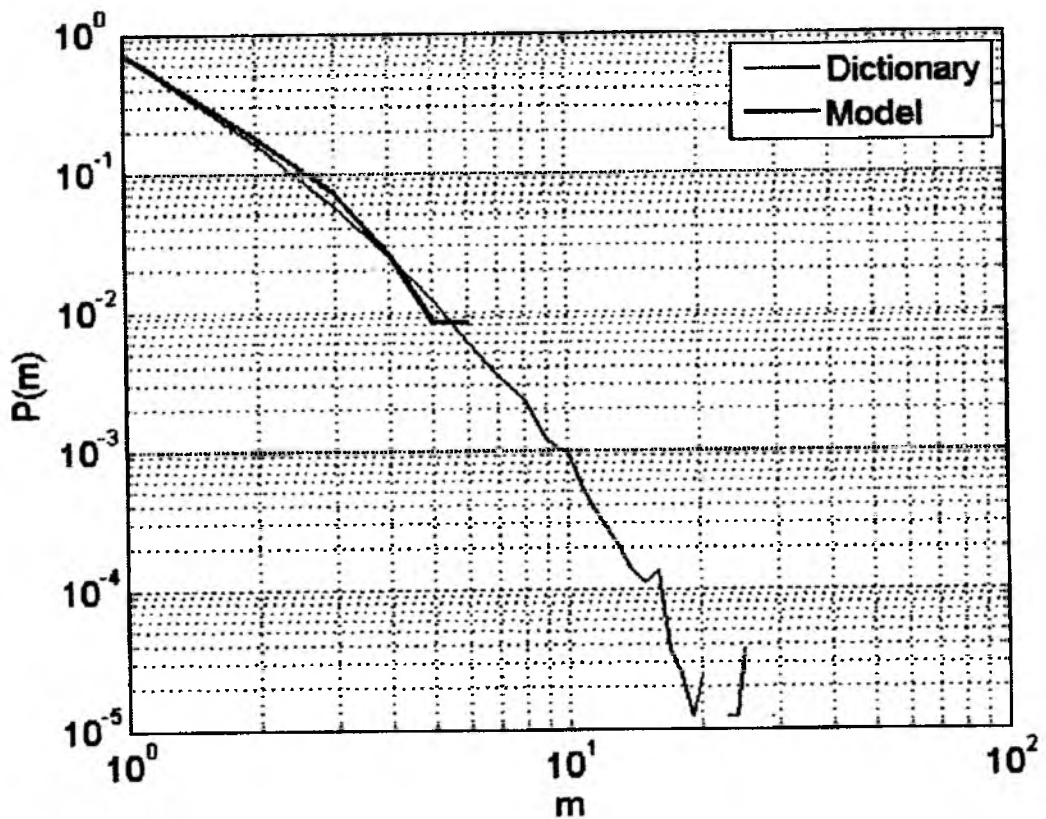


Рис. 4. Распределения лексической полисемии по словарю СО и по модели
 $m_M = 1,4697$, $m_D = 1,6973$, $\sigma_M = 0,89921$, $\sigma_D = 1,1651$, $H = H_0$: $p = 0,99996$

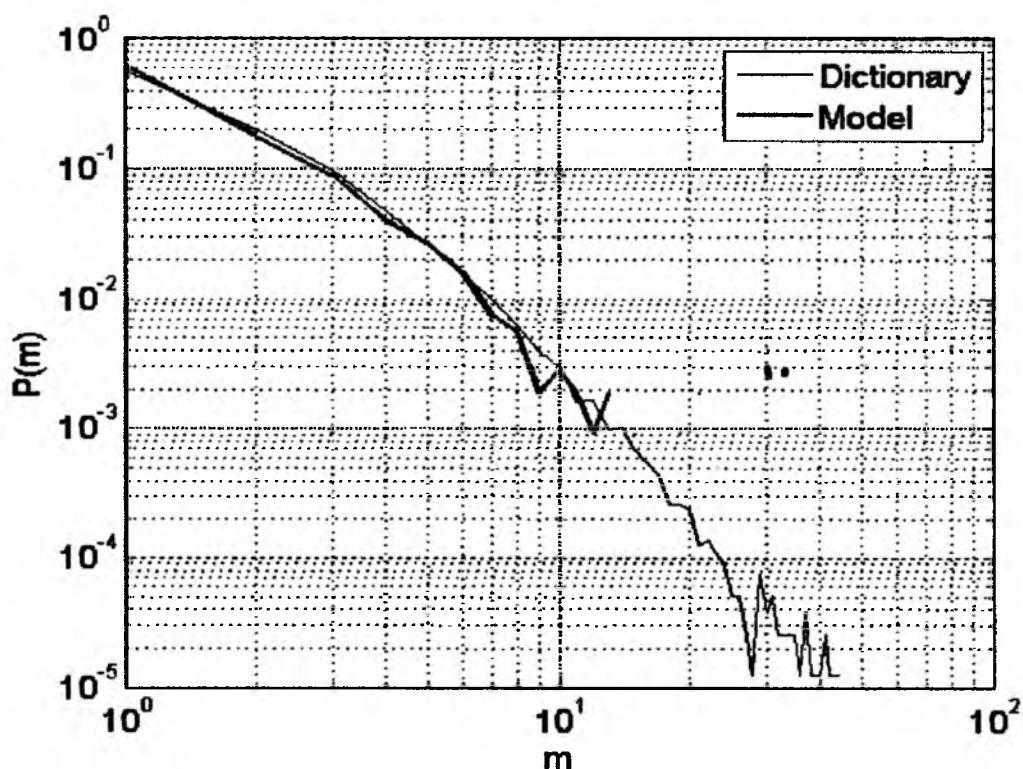


Рис. 5. Распределения лексической полисемии по словарю Shorter и по модели $m_M = 1,9027$, $m_D = 2,0114$, $\sigma_M = 1,8938$, $\sigma_D = 2,004$, $H = H_0$: $p = 0,53091$

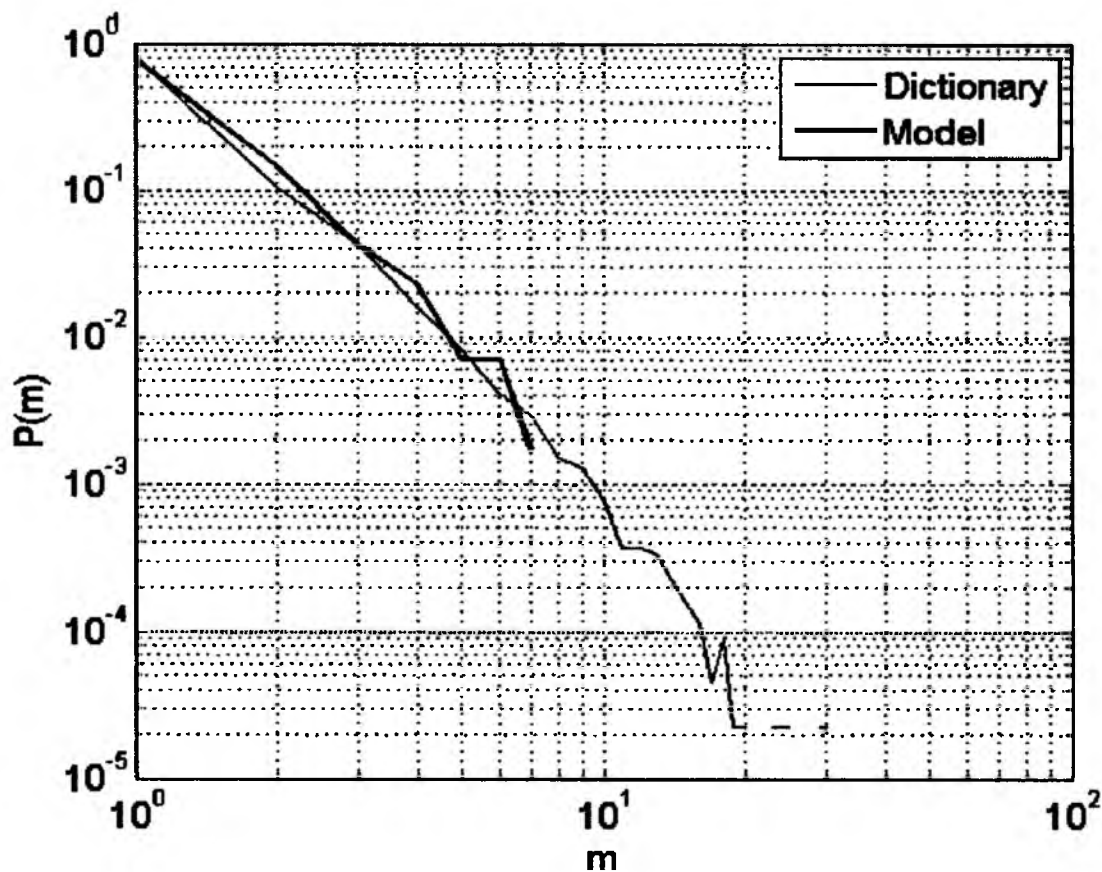


Рис. 6. Распределения лексической полисемии по словарю Hornby и по модели $m_M = 1,3654$, $m_D = 1,3597$, $\sigma_M = 0,84794$, $\sigma_D = 1,0909$, $H = H_0$: $p = 1$

полисемии, выведенные по синхронному (одномоментному) срезу диссипативной стохастической модели эволюции ансамбля из нескольких десятков тысяч языковых знаков. Модель идентифицировалась для каждого словаря соответствующим подбором указанных выше 5 параметров.

Как видно по рис. 2–6, выведенные из нашей модели распределения полисемии (жирные линии) статистически значимо (по критерию Колмогорова-Смирнова) не отличаются от эмпирических распределений, полученных на основе данных из указанных выше представительных словарей русского и английского языков (тонкие линии). Более того, эволюционно-полисемические параметры $\langle \tau \rangle$, $\langle \tau_0 \rangle$, $\langle \tau^{(1)} \rangle$, $\langle \tau^{(2)} \rangle$, $\langle G \rangle$, (т.е. те параметры, которые были исходно заложены в модель и отвечают за скорость смены значений и слов в языке) оказались имеющими и синхронный, лингво-типологический смысл. А именно, словарь английского языка, как языка явно относительно более аналитического, более ограничен по набору лексических единиц, чем словарь русского языка. Вследствие этого максимальная, а также средняя полисемическая, а поэтому и максимальная и средняя частотная нагрузка на каждую английскую лексическую единицу должна быть выше, а «изнашивание» каждой из них в английском языке вследствие этого должно происходить быстрее. То есть, в целом, языковой «метаболизм» (смена знаков и значений) в английском языке должен осуществляться интенсивнее, чем в русском. Это, в частности, должно также проявляться в относительно меньшей сохранности лексического материала английского в сравнении с русским за один и тот же период времени, например, за тысячелетие. Данные по сохранности лексических единиц русского и английского языков за последнее тысячелетие по глоттохронологическому списку подтверждают этот вывод. Проверка вывода на более широком словарном материале сейчас проводится.

Под общей редакцией *В. А. Кочергиной*

**СИНХРОННОЕ И ДИАХРОННОЕ В
СРАВНИТЕЛЬНО-ИСТОРИЧЕСКОМ ЯЗЫКОЗНАНИИ**

Оригинал-макет подготовлен издательством «Добросвет»

Формат 60x84 1/16.

Объем 17 печ. л.

Тираж 500 экз. Заказ Т-004.

Издательство «Добросвет».

109147, Москва, Ж-14, Марксистская ул.,
9 — 505.

«Издательство „КДУ“», 119234, Москва а/я 587.

тел./факс: (495) 939-44-91, 939-57-32.

e-mail: kdu@kdu.ru, <http://www.kdu.ru>

Отпечатано в типографии «КДУ»

тел./факс: (495) 939-44-91 e-mail: press@kdu.ru