ВОПРОСЫ ЯЗЫКОЗНАНИЯ

№ 5 2004

© 2004 г. Е.С. МАСЛОВА

ДИНАМИКА ТИПОЛОГИЧЕСКИХ РАСПРЕДЕЛЕНИЙ И СТАБИЛЬНОСТЬ ЯЗЫКОВЫХ ТИПОВ

Памяти Джозефа Гринберга

1. ВВЕДЕНИЕ

Один из эмпирических результатов типологических исследований последних десятилетий, начиная со знаменитой статьи Дж Гринберга [Greenberg 1963], состоит в наблюдении, что распределение языков мира по многим типологическим параметрам очень неравномерно: если представить множество теоретически возможных языков в виде многомерного пространства **W**, окажется, что существующие языки концентрируются в определенных областях **W**, а при удалении от таких областей степень "заселенности" **W** постепенно убывает М Драйер [Dryer 1998] иллюстрирует это наблюдение диаграммой, аналогичной рис 1а, представляющей одну из типичных проекций **W**, где точки соответствуют существующим языкам. При описании конкретного параметра (или нескольких взаимосвязанных параметров) наблюдения такого рода часто формулируются в качестве так называемых "статистических универсалий", т.е. утверждений вида "большинство языков мира обладает свойством α", которым соответствует условная граница "тесно заселенной области", внутри которой находятся языки, обладающие свойством α (см. рис. 1б).

Строго говоря, утверждения этого типа описывают свойства множества существующих языков – языковой популяции **E**. Их значимость для теоретического языкознания – понимаемого как наука о свойствах любого возможного естественного языка, или, в терминах введенной выше метафоры, как изучение свойств пространства **W**, – является весьма спорным вопросом. Существующие взгляды на эту проблему колеблются от практически полного отрицания теоретической ценности статистических универсалий до столь же полного отрицания теоретической релевантности "редких исключений" из общего правила, при котором различие между абсолютными и статистическими универсалиями практически не принимается во внимание

С одной стороны, языковая популяция земного шара – объект скорее исторический, чем лингвистический; распределение ее элементов по пространству **W** зависит от истории ее возникновения и развития и не могло не меняться на протяжении этой истории – очевидно, что в период непосредственно после возникновения первого языка (или языков) это распределение было иным С другой стороны, как аргумент против лингвистической значимости статистических универсалий это наблюдение сохраняет силу и для универсалий абсолютных, за исключением, возможно, очень небольшого множества тривиальных утверждений. Наши знания о том, какие языки возможны, могут опираться только на данные наблюдаемой популяции **E**. Если все языки в **E** обладают некоторым свойством, это также может оказаться "случайным" – никак не связанным со свойствами пространства **W** – свойством данной популяции. В самом деле, гносеологически, многие статистические универсалии –

это "бывшие" абсолютные универсалии ("Любой естественный язык обладает свойством α"), к которым нашлись исключения; хорошо известный пример – универсалия базового порядка составляющих SO. С другой стороны, некоторые статистические универсалии могли "оказаться" абсолютными: например, представляется вполне возможным, что щелкающие согласные возникли всего один раз в истории человечества. Если бы тот язык, в котором они возникли, исчез, то запрет на такие согласные мог бы предстать фонологам этого вполне возможного мира как абсолютная универсалия. Можно предположить, что мы находимся в точно такой же ситуации по отношению к целому ряду других маловероятных языковых явлений, а значит резкое теоретическое противопоставление статистических универсалий абсолютным не имеет эмпирических оснований.

В практике лингвистической типологии принято рассматривать абсолютные универсалии как "крайний случай" универсалий статистических [Comrie 1989]: любое отклонение от равномерного распределения признается лингвистически значимым, если его нельзя считать чисто случайным, т.е. если оно является статистически значимым [Bell 1978]. Проблема состоит в том, что статистически значимые отклонения типологических распределений от равномерного могли с большой вероятностью возникнуть "случайно" с лингвистической точки зрения. Основные источники таких отклонений – лингвистические свойства праязыка (или праязыков) [Maddieson 1991: 352] и возникновение и рост языковых семей в ранний период существования языковой популяции, когда она была достаточно мала [Maslova 2000]. Сохранение таких отклонений в современной языковой популяции возможно благодаря относительной стабильности языковых состояний: свойства любого языка существенно зависят от того, каким этот язык был на предыдущем этапе своей истории, и эта зависимость может, в принципе, сохраняться очень долго (или даже бесконечно, ср. [Lass 1997: 302])[1]. Следовательно, оценка лингвистической значимости наблюдаемых распределений возможна только на основании хотя бы приблизительных оценок стабильности изучаемых состояний. Основная задача этой статьи – описать и обосновать один из возможных методов получения таких оценок.

Статья построена следующим образом. В разделе 2 излагается простая вероятностная модель динамики типологических распределений. Описанию и примерам применения предлагаемого метода посвящены разделы 3–4. В рамках принятой модели, этот метод позволяет оценить стабильность языковых состояний и степень лингвистической значимости наблюдаемых статистических тенденций. В разделах 5–7 рассматриваются три основные проблемы, которые обычно занимают центральное место при обсуждении методологии типологических исследований, но не учтены в принятой в разделах 2–4 модели: влияние исторических случайностей на распределения лингвистических переменных в популяции (раздел 5), существование зависимостей между параметрами типологической вариации (раздел 6) и роль языковых контактов в процессе языковых изменений (раздел 7).

[1] На достаточно коротких интервалах времени, утверждение об относительной стабильности свойств языка тривиально: за исключением случаев резкого перехода сообщества на другой язык и возникновения креольских языков, язык каждого следующего поколения совпадает с языком предыдущего по подавляющему большинству грамматических параметров, а значит, находится в очень сильной зависимости от него. Однородность большинства генетических групп языков по многим грамматическим параметрам позволяет предположить, что такого рода статистическая зависимость сохраняется достаточно долго. Это предположение, строго говоря, не сводится к гипотезе сохранения языкового состояния: даже если состояние изменилось, статистическая зависимость может сохраняться.

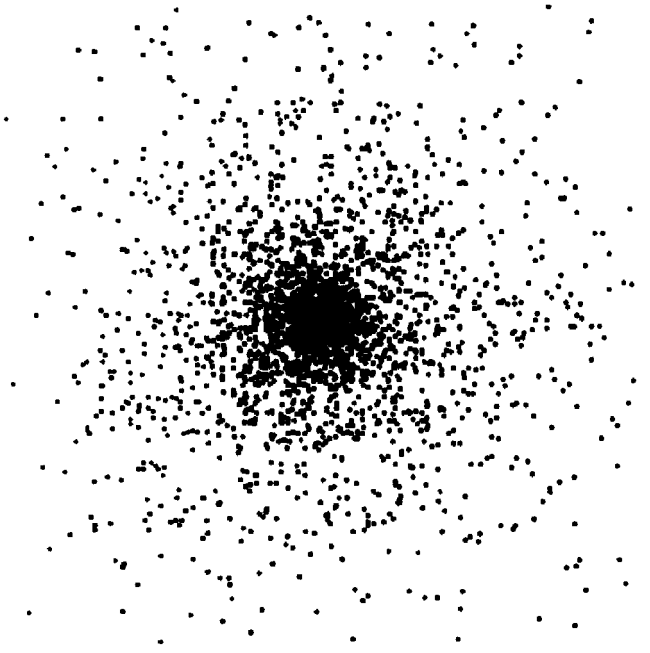

Рис. 1а. Распределение существующих языков по пространству возможных языков

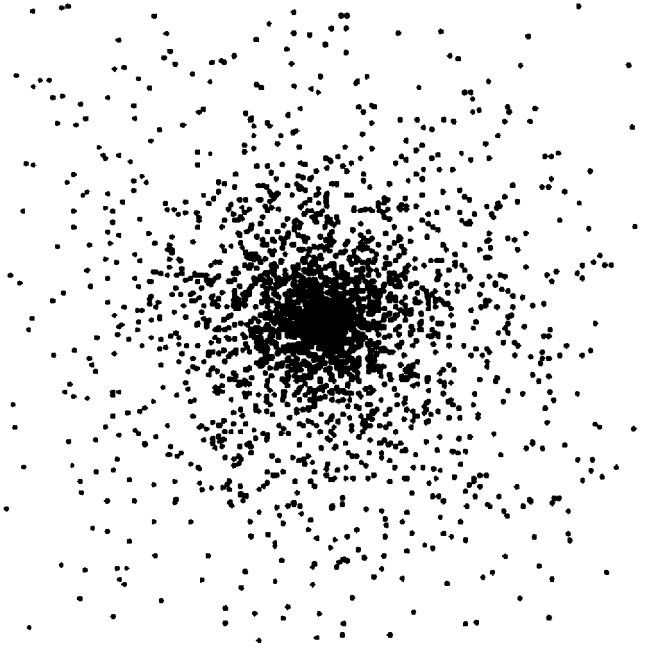

Рис. 1б. Условное изображение статистической универсалии

2. МИНИМАЛЬНАЯ ВЕРОЯТНОСТНАЯ МОДЕЛЬ ЛИНГВИСТИЧЕСКИХ ПЕРЕМЕННЫХ

В этой статье понятие статистической универсалии рассматривается для простейшего случая бинарных дискретных типологических параметров. В дальнейшем такие параметры обозначаются как [X], где X – одно из значений параметра, выделенное как "положительное"; например, обозначение [SO] соответствует параметру "доминантный (наиболее частотный) порядок следования субъекта и объекта в независимом переходном предложении". Каждый такой параметр можно представить как лингвистическую переменную l, которая может принимать значения 1 и 0. Для удобства, положительное значение (1) всегда приписывается тому значению, которое a priori предполагается более вероятным.

В модели, предложенной Дж. Гринбергом [Greenberg 1995], свойства таких переменных описываются двумя параметрами: стабильность (stability) и частота возникновения (frequency) положительного значения, определяемыми через вероятности перехода p_{ij} от одного значения к другому за некоторую единицу времени Δt, которая в дальнейшем называется временным шагом, т.е. вероятности того, что язык окажется в состоянии $l = j$ при условии, что он находился в состоянии $l = i$ на предыдущем шаге ($p_{i0} + p_{i1} = 1$)[2]. При этом предполагается, что сами эти вероятности представляют собой своего рода языковые универсалии, т.е. свойства любого естественного языка.

Если язык существует достаточно долго, то вероятности $s_i = P(l = i)$ значений переменной l в этом языке не зависят от исходного значения этой переменной и полностью определяются вероятностями переходов: их отношение $\omega(l)$ равно отношению вероятностей перехода:

(1) $\omega(l) = s_1/s_0 = p_{01}/p_{10} (s_1 + s_0 = 1)$

Такое распределение $\{s_i\}$ называется стационарным, так как оно соответствует состоянию равновесия между синхронными вероятностями и вероятностями перехода: если в некоторой совокупности языков наблюдается стационарное распределение, то общее количество переходов из нулевого состояния в положительное за любой интервал времени приблизительно равно количеству обратных переходов, так что распределение $\{s_i\}$ не меняется.

Поскольку предполагается, что вероятности перехода универсальны, то определяемое этими вероятностями стационарное распределение также является языковой универсалией, т.е. это распределение можно считать свойством пространства возможных языков W, а не конкретной языковой популяции. Следовательно, коэффициент $\omega(l)$ можно считать количественной мерой "силы" универсального ограничения (constraint) на нулевое значение переменной l. Абсолютной универсалии соответствует значение $\omega(l) = \infty$, т.е. нулевая вероятность перехода к запрещенному типу; при равномерном стационарном распределении (в традиционных терминах, при отсутствии универсальных ограничений на значения переменной) $\omega(l) = 1$; в целом, чем меньше стационарная вероятность нулевого значения (т.е. чем сильнее предполагаемая универсалия), тем больше значение $\omega(l)$. Этот коэффициент можно назвать "весом" универсалии.

Время конвергенции к стационарному распределению определяется суммой вероятностей изменения значения переменной, $\mu(l) = p_{01} + p_{10}$. Это свойство ти-

---

[2] Формально это модель простейшего марковского процесса с двумя состояниями и дискретным временем. Все математические факты в этом и следующем разделах относятся к любому такому процессу и могут быть найдены практически в любом учебнике по теории вероятностей. При этом предполагается, что "длина" временного шага выбирается таким образом, чтобы вероятности перехода были достаточно малы, так что можно пренебречь вероятностью более одного перехода на протяжении одного временного шага.

пологических параметров обычно называют (не)стабильностью, однако при этом возникает некоторая терминологическая путаница. по Гринбергу, стабильность – это свойство языкового типа, т.е. одного из значений лингвистической переменной (определяемое только вероятностью ухода из этого типа), а не свойство параметра вариации В этой статье, величина $\mu(l)$ называется мобильностью параметра (ср. [Hawkins 1983]): чем более мобилен параметр, тем меньше время конвергенции. Термин "стабильность" используется только для описания свойств значений параметра

Для практических применений важно соотношение времени конвергенции и временной "глубины" существующих в популяции языковых семей. Точнее, если современный язык L отдален от языка-предка A приблизительно на n временных шагов (другими словами, их разделяет временной интервал $n\Delta t$), то вероятность любого значения переменной l в L отклоняется от стационарной не более, чем на величину $\varepsilon = (1 - \mu(l))^n$ независимо от значения этой переменной в A. Следовательно, те отклонения от равномерного распределения, которые существенно превышают ε, отражают свойства стационарного распределения. Только такие отклонения считаются в рамках данного подхода "лингвистически значимыми". В этой статье наблюдаемые статистические свойства языковой популяции называются статистическими тенденциями, а универсальные свойства языка, которые могут приводить к возникновению таких тенденций, – стохастическими универсалиями. Статистическая тенденция лингвистически значима, если она отражает стохастическую универсалию.

3. МЕТОД ОЦЕНКИ ВЕРОЯТНОСТЕЙ ПЕРЕХОДА

Практическое применение описанной в разделе 2 модели для оценки лингвистической значимости статистических тенденций предполагает нахождение оценок вероятностей перехода p_{ij}. Может показаться, что для этого необходима информация о значениях изучаемой переменной l на предыдущем временном шаге для достаточно большого числа существующих языков, что представляет собой, очевидно, невыполнимое условие. Гипотеза Гринберга состояла в том, что качественные оценки вероятностей перехода можно получить, сравнивая распределения значений l в популяции в целом и внутри языковых семей [Greenberg 1978; 1995]. Предлагаемый здесь метод можно рассматривать как уточнение и развитие этой идеи.

Базовое наблюдение состоит в том, что, хотя мы не знаем, каковы были значения интересующих нас переменных на предыдущем временном шаге, для многих пар близкородственных языков мы можем с уверенностью предположить, что эти значения совпадали – постольку, поскольку эти языки представляли собой один и тот же язык. Для обоснования предлагаемого метода необходимо описать несколько подробнее, что происходит в произвольном множестве языков за один временной шаг. Пусть f – частота положительного значения переменной в рассматриваемом множестве. Вероятность q этого значения на следующем шаге задается следующим уравнением:

$$(2)\quad q = fp_{11} + (1 - f)p_{01}.$$

Назовем коэффициентом дивергенции d условную вероятность того, что значения переменной l для двух случайно выбранных языков после этого шага различны при условии, что эти значения совпадали на предыдущем шаге. Легко показать, что

$$(3)\quad d = 2fp_{11}p_{10} + 2(1 - f)p_{00}p_{01}.$$

Сравнивая эти уравнения, легко обнаружить, что вероятность положительного значения q и коэффициент дивергенции d связаны следующим линейным уравнени-

ем с коэффициентами a и b, зависящими только от вероятностей перехода (но *не* от исходной частоты f)

(4) $d = aq + b$, где $a = 2(p_{10} - p_{01})$, $b = 2p_{01}p_{11}$

Рассмотрим теперь некоторое достаточно большое множество **M** пар родственных языков; пусть **M*** – множество языков, входящих в **M**; **A** – множество предков этих языков, существовавшее некоторое время (Δt) назад, f – неизвестная нам частота положительного значения l в **A**, тогда частота положительного значения l в **M*** является наилучшей оценкой вероятности q, а частота пар с различными значениями этой переменной в **M** – наилучшей оценкой коэффициента дивергенции d при исходной частоте f.

Оценки вероятностей перехода можно получить с помощью сравнения двух или более множеств $\mathbf{M}_i$ с разными распределениями переменной l. Если имеется несколько разных пар значений (q_i; d_i), метод линейной регрессии дает оценки коэффициентов уравнения (4), т.е. величин $2(p_{10} - p_{01})$ и $2p_{01}p_{11}$. Зная значения этих величин, легко вычислить вероятности перехода, решив простое квадратное уравнение Несколько различных множеств $\mathbf{M}_i$ можно получить, разбив языковую популяцию на частичные популяции, соответствующие группам языковых семей. В следующем параграфе эта процедура описывается более детально на примере анализа трех типологических параметров

4. ПРИМЕР ПРИМЕНЕНИЯ МЕТОДА: БАЗОВЫЙ ПОРЯДОК СЛОВ

База данных по базовому порядку составляющих (S, O, V), опубликованная Р. Томлином [Tomlin 1986], представляет собой случайную выборку **S** (более 1000 языков) из языковой популяции **E** Полное множество **M** пар наиболее близкородственных языков в этой выборке было построено с помощью следующей процедуры.

Выберем некоторую многоуровневую классификацию **G**, приблизительно отражающую генеалогическое древо языков (и диалектов); в данном случае, использовалась классификация Ethnologue [Grimes 1997]. Для каждого языка L из **S** найдем множество **L** языков из **S** таких, что (а) все языки из **L** принадлежат к некоторому классу из **G** и (б) не существует более узкого класса в **G**, включающего одновременно L и хотя бы один другой язык из **S**. Если **L** состоит из единственного элемента L, этот язык исключается из рассмотрения. Если **L** включает два языка, оно включается в множество пар **M**. Если **L** состоит из трех или более языков, то из них случайным образом выбирается пара языков для включения в множество **M**. Полученное таким образом множество **M** содержит 325 пар родственных языков. Так как исходная выборка **S** очень велика и случайна, можно с достаточной долей надежности считать, что ее "плотность" (степень родства между наиболее близкородственными языками) постоянна, т.е. полученное множество **M** более или менее однородно по времени расхождения языков в паре (около 1000 лет); эта величина и принимается в дальнейшем за 'временной шаг"[3].

Для целей этого исследования, данные о базовом порядке слов были представлены в терминах трех бинарных параметров, [SO], [SV] и [VO]. Для оценки вероятностей перехода, множество пар **M** необходимо разбить на несколько подмножеств $\mathbf{M}_i$ таким образом, чтобы частоты положительного значения изучаемого параметра по

[3] Теоретически можно предположить, что этот временной шаг слишком велик для моделирования изменений порядка слов, т е. порядок слов в среднем меняется так быстро, что существует значительная вероятность двух или более переходов на протяжении одного тысячелетия (см. сноску 2). Представляется, однако, что такое предположение находится в противоречии с эмпирически наблюдаемой однородностью генетических языковых групп с меньшим временем дивергенции по признаку базового порядка слов.

возможности значительно отличались друг от друга. С этой целью, языковая популяция разбивалась на несколько частичных популяций, соответствующих множествам семей в Ethnologue; эта классификация содержит 100 макросемей, 73 из которых представлены в S. Для каждого параметра семьи упорядочивались по частоте положительного значения данного параметра внутри семьи, и полученный список разбивался на несколько множеств семей приблизительно равного размера, из которых и выбирались множества M_j.

Для параметров [SV] и [VO] с помощью этой процедуры удалось выделить по четыре достаточно больших множества пар M_j. Данные о распределении пар в этих множествах приведены в табл. 1–2, а соответствующие оценки параметров процессов языковых изменений суммированы в табл. 4.

Таблица 1

Распределение пар по параметру [SV]

	SV/SV	VS/VS	SV/VS	*q*	*d*
M_1	2	13	8	0.261	0.348
M_2	62	20	15	0.157	0.154
M_3	61	2	2	0 015	0 031
M_4	92	0	0	1 0	0.0

Таблица 2

Распределение пар по параметру [VO]

	VO/VO	OV/OV	VO/OV	*q*	*d*
M_1	2	76	10	0.080	0.114
M_2	13	34	18	0.338	0.277
M_3	51	22	9	0 677	0.110
M_4	81	15	11	0.808	0.103

Предварительно отметим, что полученные оценки для переменной [SV] хорошо согласуются с данными синхронных наблюдений в том смысле, что вероятность перехода в состояние SV выше вероятности обратного перехода, причем обнаруженная универсалия оказывается сильнее, чем можно было бы предположить на основе синхронных наблюдений: предсказанная стационарная вероятность SV (0.98) выше наблюдаемой в популяции частоты этого состояния (0.86). Необходимо подчеркнуть, что эта оценка получена без учета распределения переменной в популяции; напротив, метод включает построение частичных популяций, для которых распределения переменной существенно отличаются от наблюдаемого в популяции в целом. Иными словами, стохастическая универсалия подтверждается независимо от ее "*проявления*" *в синхронном распределении*. Различие между вероятностями перехода по параметру [VO] существенно менее значительно, что также согласуется с данными синхронного распределения; более того, полученные оценки подтверждают, что этот параметр значительно менее мобилен, чем [SV] [Dryer 1997]. С другой стороны, полученные оценки вероятностей перехода для [VO] предполагают, что стационарная вероятность VO (0.635) больше, чем стационарная вероятность OV, тогда как синхронные частоты различаются противоположным образом, что указывает

на существование стохастической универсалии, не проявившейся в синхронной статистической тенденции Более подробно этот результат обсуждается в разделе 6.

Разбиение на частичные популяции с разными распределениями переменной *l* может оказаться невозможным, если ее распределения среди языков одной семьи приблизительно о д и н а к о в ы для всех сколько-нибудь больших семей. Такая ситуация может возникнуть в двух принципиально разных случаях Во-первых, переменная может быть настолько мобильна, что ее распределение внутри каждой семьи приближается к стационарному В этом случае, сам факт сходства наблюдаемых распределений демонстрирует их близость к стационарному распределению, следовательно, лингвистическую значимость наблюдаемой тенденции. С другой стороны, если мобильность переменной очень мала и в с е языки – предки достаточно больших языковых семей находились в одном и том же состоянии, то частота этого состояния в любой семье близка к 100%. Наиболее очевидный частный случай такой ситуации – абсолютная универсалия, в этом случае единственный возможный вывод состоит в том что наблюдаемое состояние очень стабильно, т е вероятность ухода из этого состояния близка к нулю даже для очень большого временного шага в несколько тысячелетий.

Если вернуться к стохастическим универсалиям, то возникает вопрос, как различать (а) мобильные переменные с распределением, приближающимся к стационарному, и (б) немобильные переменные с распределением, отражающим случайное совпадение значений переменной во всех языках-предках. Прежде всего, если наблюдаемое распределение не очень сильно отклоняется от равномерного, то интерпретация (б) отпадает она возможна только при условии близости к нулю частоты редкого типа в каждой семье С другой стороны, если это условие выполняется, предлагаемый метод ограниченно применим, поскольку полное совпадение распределений внутри семей практически невозможно. Пример такой ситуации – параметр [SO] Хотя этот параметр принимает только положительное значение во многих языковых семьях, удается выбрать три достаточно больших множества пар M_j с несколько разными распределениями этой переменной Распределения типов пар в этих выборках и соответствующие значения q_j и d_j приведены в табл. 3.

Таблица 3

Распределение пар по параметру [SO]

	SO/SO	OS/OS	SO/OS	*q*	*d*
M_1	39	6	18	0 762	0 286
M_2	92	1	10	0 942	0 097
M_3	92	0	0	1 0	0 0

Эти данные позволяют получить оценки вероятностей перехода (p_{10} = 0.004, p_{01} = = 0 586). В соответствии с этими оценками, переход из OS в SO на рассматриваемом временном шаге был почти в 150 раз более вероятен, чем обратный переход, что означает, что предполагаемая стационарная вероятность SO (0.99) несколько выше наблюдаемой частоты этого состояния Таким образом, наблюдаемое распределение довольно точно отражает свойства стационарного распределения. Действительно, вероятность ухода из OS настолько велика, что параметр [SO] очень мобилен, т.е. сходится к стационарному распределению за относительно небольшое количество шагов В частности, для языка типа OS вероятность с о х р а н и т ь этот тип после шести временных шагов (что приблизительно соответствует минимальной временной глубине семьи) составляет 0 012. Таким образом, мы имеем дело с очень мобильным параметром, синхронное распределение которого приближается к стационарному.

5. О РОЛИ ИСТОРИЧЕСКИХ СЛУЧАЙНОСТЕЙ

Предлагавшиеся в типологической литературе способы решения задачи различения 'случайного" и "лингвистически значимого" в наблюдаемых распределениях типологических переменных сводятся, в значительной степени, к одной простой и, на первый взгляд, весьма разумной идее Эта идея состоит в том, что из языковой популяции можно целенаправленно выбрать некоторое подмножество I таким образом, чтобы исключить – или, по крайней мере, минимизировать – возможное влияние исторических случайностей на распределение лингвистических переменных Поскольку основным потенциальным источником статистически значимых (но лингвистически "случайных') сдвигов в таких распределениях является возникновение и распространение языковых семей, можно предположить, что если включить в I один случайно выбранный язык из каждой генетической группы некоторой временной глубины, то распределение лингвистических переменных в I будет "избавлено ' от влияния исторических событий, которые произошли за соответствующий период времени [Bell 1978]. В дальнейшем я буду обозначать распределение переменной *l* в множестве типа I как **L(I)**, а ее распределение в популяции – или в случайной выборке из популяции – как **L(E)**.

Несмотря на несомненную интуитивную привлекательность, этот метод – особенно в его наиболее сильном' варианте, когда в выборку I входит не более одного языка из каждой языковой семьи – не решает поставленной задачи, а в некоторых случаях может ухудшить результаты, т е **L(I)** может х у ж е отражать свойства стационарного распределения, чем **L(E)** Представим себе множество $\mathbf{E_0}$ языков – предков современных языковых семей Распределение лингвистических переменных в $\mathbf{E_0}$ возникло также, как и распределение в **E**, т е явилось результатом одновременного воздействия двух типов стохастических процессов (исторических процессов возникновения языковых семей и языковых изменений в истории каждого языка) на праязык или "мини-популяцию праязыков Как показано в [Maslova 2000], в н е б о л ь ш о й популяции языков исторические процессы с б о л ь ш о й в е р о я т н о с т ь ю порождают неравномерные распределения лингвистических переменных, дополнительным источником "случайности" являются собственно свойства праязыка (или праязыков) Следовательно, распределение $\mathbf{L(E_0)}$ можно считать, с лингвистической точки зрения, случайным.

Если переменная *l* достаточно мобильна, то распределение **L(I)** в множестве, содержащем один язык-потомок каждого языка из $\mathbf{E_0}$, незначительно зависит от начального состояния (т.е от свойств $\mathbf{E_0}$) и, с большой вероятностью, отражает свойства стационарного распределения. Но в этом случае этими же свойствами обладает и языковая популяция в целом, прежде всего потому, что в больших популяциях исторические процессы практически не могут привести к значительным изменениям в распределениях лингвистических переменных [Maslova 2000]. На основе достаточно большой случайной выборки можно оценить влияние таких процессов на **L(E)** Например, параметр [SV] принимает положительное значение для 86% языков в полной случайной выборке Томлина [Tomlin 1986] В случайной подвыборке типа I эта переменная имеет в т о ч н о с т и такое же распределение Тот же результат наблюдается для параметра [SO] – и в полной выборке, и в случайной подвыборке типа I параметр принимает положительное значение приблизительно в 96% языков. Таким образом, для мобильных переменных построение выборки типа I не меняет результатов по сравнению со случайной выборкой из популяции.

Если переменная *l* изменяется медленно, то распределение **L(I)** сильно зависит от начального состояния Это означает, что метод не решает поставленной задачи – свойства **L(I)** в значительной степени определяются теми случайностями, которые привели к появлению $\mathbf{E_0}$ В принципе, это относится и к популяции **E** в целом. Сравним, однако, распределения "медленного" параметра [VO] в популяции и в выборках типа I. В популяции распределение по этому параметру близко к равномерному – пе-

ременная принимает положительное значение для 48% языков, тогда как в L(I) частота этого значения заметно меньше, около 40%. Близкие результаты дает выборка Драйера для так называемых "родов" языков (genera), т.е. языковых групп типа германской или романской подгрупп индоевропейской семьи – около 41% родов содержат языки VO [Dryer 1998]. Драйер [Dryer 1989] объясняет это расхождение совпадением двух исторических случайностей – широким распространением австронезийских языков и языков банту, из чего вроде бы следует, что преобладание OV в L(I) точнее отражает универсальные свойства языка, чем равномерное распределение L(E) Это объяснение оставляет открытым вопрос, почему то же совпадение никак не повлияло на распределения [SV] и [SO].

С другой стороны, в соответствии с данными полученными в разделе 4, стационарная вероятность VO составляет около 64% (см. табл. 4), т е. переход от OV к VO несколько более вероятен, чем обратный переход.

Таблица 4

Оценки динамики изменений базового порядка слов

	Вероятности переходов		Вес универсалии	Мобильность параметра	Стационарная вероятность
	p_{01}	p_{10}	$\omega(I) = p_{01}/p_{10}$	$\mu(I) = p_{01} + p_{10}$	$s_1 = p_{01}/(p_{01} + p_{10})$
[SO]	0.586	0.004	146 5	0 590	0 993
[SV]	0.236	0.004	59	0 240	0 983
[VO]	0.096	0.055	1 74	0 151	0 635

Эти оценки хорошо согласуются с имеющимися данными о таких переходах [Givón 1979, Newmeyer 2000]. Но это означает, что L(E) оказалось б л и ж е к стационарному распределению, чем L(I). Иными словами, выборки типа I лучше сохраняют случайные свойства исходной популяции $\mathbf{E}_0$ (в данном случае, преобладание OV), а значит, дают менее лингвистически значимые результаты, чем случайные выборки. Этот на первый взгляд неожиданный эффект имеет простое качественное объяснение Если оба значения параметра очень стабильны, то в большинстве семей п р е о б л а д а е т исходное значение, поэтому при случайном выборе одного языка из семьи он с большой вероятностью отражает именно исходное значение В результате, выборка типа I довольно точно отражает распределение параметра в исходной популяции $\mathbf{E}_0$ С другой стороны, если одно значение (в нашем примере, OV) несколько менее стабильно, чем другое, семьи с менее стабильным исходным значением содержат, в среднем, несколько бблшее число языков, и з м е н и в ш и х значение параметра. В результате, в популяции в целом частота более стабильного состояния (VO) возросла по сравнению с $\mathbf{E}_0$, т е приблизилась к его стационарной вероятности, и это произошло не "случайно", а именно благодаря небольшому различию в вероятностях перехода. Иными словами, в результате попытки исключить влияние размеров семей на распределение переменной, мы полностью теряем информацию о различиях в стабильности между значениями этой переменной, которая содержится в распределениях этой переменной в н у т р и языковых семей с разными исходными значениями

Эти наблюдения позволяют сделать два вывода. Во-первых, попытки элиминировать эффекты исторических случайностей с помощью построения выборок типа I не приводят к ожидаемому улучшению оценок вероятностей языковых типов. для мобильных параметров, те же оценки можно получить на основе случайной выбор-

ки, а для медленных параметров такая выборка дает м е н е е лингвистически значимые результаты С другой стороны, приведенные выше соображения позволяют предложить относительно простую эвристику для оценки лингвистической значимости наблюдаемых тенденций если распределения некоторой переменной в популяции в целом и в выборке типа I приблизительно одинаковы, то, с большой вероятностью, это достаточно мобильный параметр, распределение которого близко к стационарному, а следовательно, лингвистически значимо. Именно эта ситуация наблюдается для параметров [SO] и [SV]. Если же распределения L(I) и L(E) значительно различаются, то переменная имеет низкую мобильность, и распределение L(I) отражает прежде всего свойства исходной популяции $\mathbf{E}_0$. Тогда само различие между распределениями L(I) и L(E) отражает универсальную тенденцию языковых изменений, а именно L(E) ближе к стационарному распределению, чем L(I). Как показывает анализ параметра [VO] в разделе 4, это не значит, что по L(E) можно судить о свойствах стационарного распределения, т. е. о "весе" стохастической универсалии. Именно в случаях расхождения между L(I) и L(E) предложенный здесь метод оценки вероятностей переходов наиболее полезен, так как он дает возможность обнаружить стохастические универсалии, не проявившиеся в синхронных статистических тенденциях.

6. ДИНАМИЧЕСКИЕ ЗАВИСИМОСТИ МЕЖДУ ПАРАМЕТРАМИ

Поиск лингвистически мотивированных статистических з а в и с и м о с т е й между параметрами находится в центре внимания типологических исследований. В рамках рассматриваемой здесь модели, лингвистически значимая зависимость между двумя параметрами – это различие в е р о я т н о с т е й п е р е х о д а по одному из этих параметров в зависимости от значения другого параметра Описанный в разделе 3 метод можно использовать для обнаружения таких зависимостей.

Предположим, что мы хотим определить наличие зависимости вероятностей перехода для переменной l_1 от значения другой лингвистической переменной, l_2. Выделим два множества пар $\mathbf{M}_i$, в которых переменная l_2 принимает одинаковые значения, например, $\mathbf{M}_0$([VO]) содержит все пары, в которых оба языка имеют базовый порядок слов OV, а $\mathbf{M}_1$([VO]) – все пары языков VO Хотя мы не можем с уверенностью предполагать, что параметр [VO] принимал это же значение в и с х о д н о м языке каждой пары, ясно, что р а с п р е д е л е н и я значений этого параметра в соответствующих множествах исходных языков $\mathbf{A}_i$ различны, так как сохранение исходного значения в каждом языке намного более вероятно, чем его изменение (см. табл 4). Следовательно, если вероятности изменений l_1 зависят от значения l_2, то они различны и для множеств $\mathbf{A}_i$. Для каждого множества $\mathbf{M}_i$ легко оценить значение вероятности q_i положительного значения l_1 и с помощью данных, полученных применением основного метода оценки вероятностей перехода, вычислить о ж и д а е м о е значение коэффициента дивергенции d_i, т.е. вероятность появления дивергентной пары при условии независимости l_1 от l_2. Например, в множестве $\mathbf{M}_0$([VO]) частота положительного значения параметра [SV] составляет $q_0 = 0.699$. Подставляя оценки вероятностей перехода из табл. 4 в уравнение (4) получаем ожидаемое значение $d_0 = 0.149$ При помощи критерия χ^2 с одной степенью свободы можно оценить, соответствует ли н а б л ю д а е м о е количество дивергентных пар ожидаемому. В данном случае, множество $\mathbf{M}_0$([VO]) состоит из 141 пары, 17 из которых разошлись по параметру [SV]. Соответствующее значение $\chi^2 = 0\,91$, что означает, что наблюдаемое значение очень близко к ожидаемому при условии независимости Аналогичная ситуация наблюдается для множества $\mathbf{M}_1$([VO]), т.е. вероятности изменения значений [SV] можно считать независимыми от значения [VO]. Данные сопоставления трех пар рассматриваемых параметров приведены в табл. 5. Эти данные показывают, что процессы перехода по этим параметрам можно считать независимыми друг от друга.

Таблица 5

Проверка динамической независимости параметров

l_1	l_2		Общее число пар	Число дивергентных пар	q	d	χ^2
[SV]	[VO]	M_0	141	17	0.698	0.149	0 91
		M_1	146	0	0.986	0.015	2 25
[SO]	[VO]	M_0	141	16	0.924	0 104	0 07
		M_1	146	1	0.982	0.031	2.68
[SO]	[SV]	M_0	269	2	0.996	0.013	0.73
		M_1	36	8	0 778	0 286	0 73

С другой стороны, сама возможность существования динамических зависимостей до некоторой степени ограничивает применимость описанной в разделе 2 модели бинарных параметров если переменная *l* зависит от значений множества D других лингвистических параметров, то вероятности изменения значений *l* на *n*-ом шаге зависят от распределения $L_D(n-1)$ параметров D на предыдущем шаге. Поскольку само распределение $L_D(n)$ может быть нестационарным, то полученные для одного шага оценки вероятностей перехода для *l* нельзя считать универсальными свойствами этого параметра, т.е. минимальная вероятностная модель бинарных переменных оказывается, в общем случае, неприменимой. Тем не менее, как показано выше, предложенный метод оценки вероятностей перехода можно использовать для обнаружения таких зависимостей, т е , по сути дела, более сложных стохастических универсалий. Применимость метода за пределами области применимости модели обусловлена тем, что предположение об универсальности вероятностей перехода никак не используется при получении оценок этих вероятностей. Более того, если предполагается, что обнаруженная зависимость сводится к универсальному стохастическому запрету на определенное сочетание значений переменных (т.е. к импликационной стохастической универсалии), эту гипотезу можно проверить с помощью того же метода, выбрав новую переменную *l**, нулевое значение которой определяется как наличие ' запрещенного" сочетания свойств, а положительное – как отсутствие хотя бы одного из этих свойств

7. О РОЛИ ЯЗЫКОВЫХ КОНТАКТОВ

Краеугольный камень описанной в разделе 2 динамической модели составляет предположение, что вероятность изменения состояния языка определяется универсальными, собственно лингвистическими, свойствами этого состояния, его внутренней стабильностью Альтернативная стохастическая модель языковых изменений может быть основана на гипотезе, что эта вероятность определяется прежде всего контактной ситуацией, в которой находится языковое сообщество. Представим себе такую простейшую модель. Пусть *c* – доля языков в популяции, которые находятся под достаточно сильным влиянием языковых контактов; поскольку вероятность оказаться в такой ситуации, очевидно, не зависит от лингвистических свойств языка, то коэффициент *c* можно считать постоянным для языков с разными значениями изучаемого параметра *l* Далее, пусть *p* – вероятность изменения значения данного параметра при условии, что его значения в контактирующих языках различны. В рамках контактной модели *p* не зависит от текущего значения параметра. Тогда вероятность перехода из состояния $l = i$ в состояние $l = j$ определяется тем,

с какой вероятностью второй контактный язык окажется в состоянии $l = j$, которая в свою очередь, зависит прежде всего от частоты таких языков в популяции. Таким образом, эта модель предсказывает следующие вероятности переходов

(5) $p_{01} = cpf, \quad p_{10} = cp(1-f),$

где f, как и раньше, обозначает частоту положительного значения l. Иными словами, предсказание состоит в том, что вероятности перехода соотносятся так же, как и частоты значений в популяции. Легко заметить, что в точности этим же свойством обладает стационарное распределение [см. уравнение (1)] Иными словами, в этой модели любое синхронное распределение является стационарным, а значит остается неизменным сколь угодно долго. Даже очень сильные статистические тенденции – такие, как преобладание языков SO – могут, однажды случайно возникнув, "само-возобновляться" за счет языковых контактов при полном отсутствии различий в стабильности языковых состояний.

Это означает, что высокую степень "согласованности" вероятностей перехода с наблюдаемой статистической тенденцией, которая в разделах 2–4 интерпретировалась как критерий лингвистической значимости последней, можно объяснить и в рамках контактной модели: например, вероятность перехода от SO к OS может быть намного меньше вероятности обратного перехода просто потому, что в популяции намного меньше языков OS, и, следовательно, меньше вероятность оказаться под влиянием такого языка Другими словами, проблема состоит в том, что языковые контакты способны создавать "тенденции" языковых изменений, т.е. различия в вероятностях перехода, которые, в отличие от стохастических универсалий, являются не свойством языка, а свойством языковой популяции с неравномерным распределением лингвистической переменной. Можно предположить, что наблюдаемые различия в вероятностях перехода представляют собой некую комбинацию этих явлений. Задача, таким образом, состоит в том, чтобы определить относительный "вклад" внешних (популяционных) и внутренних (универсальных) факторов в эту комбинацию.

Прежде всего, если обнаруженное различие вероятностей перехода не согласуется с наблюдаемым синхронным распределением (как, например, для параметра [VO]), то его можно объяснить только наличием стохастической универсалии, т.е. различием во внутренней стабильности состояний. Более того, такая универсалия, скорее всего, имеет бóльший вес, чем можно было бы предположить на основе полученных оценок, так как ей противодействует тенденция, создаваемая языковыми контактами. В самом деле, многие переходы из VO в OV объясняются именно влиянием языковых контактов [Newmeyer 2000]. Общий метод оценки относительной роли языковых контактов и стохастических универсалий может быть основан на применении описанной в разделе 6 процедуры обнаружения динамических зависимостей к языковым ареалам с разными синхронными распределениями изучаемого параметра (ср. [Dryer 1989]); более детальная разработка этой идеи представляет собой одно из необходимых направлений дальнейших исследований динамики типологических распределений.

СПИСОК ЛИТЕРАТУРЫ

Bell 1978 – *A Bell* Language sampling // J H Greenberg ed. Universals of human languages V. I Method and theory Stanford, 1978

Comrie 1989 – *B Comrie*. Language universals and linguistic typology 2-nd ed Oxford, 1989.

Dryer 1989 – *M S Dryer* Large linguistic areas and language sampling // Studies in language 13 1989

Dryer 1997 – *M S Dryer*. On the 6-way word order typology // Studies in language 21 1997.

Dryer 1998 – *M S Dryer*. Why statistical universals are better than absolute universals // Papers from the 33rd annual meeting of the Chicago linguistic society. 1998

Greenberg 1963 – *J.H Greenberg* Some universals of grammar with particular reference to the order of meaningful elements // J.H Greenberg ed. Universals of language. Cambridge (Mass.), 1963.
Greenberg 1978 – *J.H Greenberg* Diachrony, synchrony and language universals // J.H. Greenberg ed Universals of human languages. V. I. Method and Theory. Stanford (California), 1978.
Greenberg 1995 – *J.H Greenberg*. The diachronic typological approach to language // M. Shibatani, T Bynon eds. Approaches to language typology. Oxford, 1995.
Givón 1979 – *T. Givón*. On understanding grammar. New York, 1979.
Grimes ed 1997 – *B Grimes* ed. Ethnologue: languages of the world (plus supplement· ethnologue index). 13th edition Dallas, 1997.
Hawkins 1983 – *J.A Hawkins*. Word order universals. Quantitative analysis of linguistic structure. New York. 1983
Lass 1997 – *R Lass*. Historical linguistics and language change Cambridge, 1997
Maddieson 1991 – *I Maddieson* Investigating linguistic universals Proceedings of the XIIth International congress of phonetic sciences V I 1991
Maslova 2000 – *E. Maslova* A dynamic approach to the verification of distributional universals. Linguistic typology 4–3 2000
Newmeyer 2000 – *F.J Newmeyer*. On the reconstruction of "Proto-World" word order // C. Knight et al. eds The evolutionary emergence of language Cambridge, 2000.
Tomlin 1986 – *R S. Tomlin*. Basic word order: Functional principles London, 1986.