



A. В. СУПРУН

О СИСТЕМНОМ И СТАТИСТИЧЕСКОМ ПРЕДСТАВЛЕНИИ ГРАММАТИЧЕСКИХ ИЗОГЛОСС СЛАВЯНСКОГО МИРА

Понятие системы прочно закрепилось в современной науке о языке. Правда, понятие это толкуется по-разному, но теперь уже невозможно подходить к явлениям языка без учета так или иначе понимаемой их системности. В частности, понятие системы из характеристики явлений внутриязыковых — грамматическая система, лексическая система, звуковая система — стало переноситься на сферу использования этих явлений, на сферу pragmatики в широком смысле слова. Теперь речь идет не только о системе фонем или о системе глагольных времен, но и о системе стилей или о системе диалектов.

Системный принцип изучения диалектов с разных сторон обосновывается за последнее время в славянской диалектологической литературе. Попимание диалектов как системы может быть интерпретировано по крайней мере двояко. Во-первых, диалекты могут рассматриваться как различные отражения относительно единой системы, как результат некоторых изменений в этой системе; поскольку сами изменения носят системный характер, следует ожидать, что системными будут и результаты изменений. Во-вторых, группу диалектов мы можем рассматривать как диалекты некоего единого языка в том случае, если между системами диалектов устанавливается некоторая система корреспонденций; эта система корреспонденций и будет по существу системой диалектов языка. Вторая интерпретация системности диалектов более обща, поскольку в ней не принимается в расчет генезис диалектов, относительно которого могут быть различные мнения.

Такое понимание системности диалектов предполагает многомерность системы. Система диалектов является, если угодно, не суммой систем диалектов, а их произведением, ибо это не система отношений между отдельно взятыми фактами диалектов, а система отношений между системами диалектов. Вообще говоря, система диалекта может включать в себя несколько стилевых систем, причем вполне возможно использование в качестве одного из стилей данного диалекта некоторого общего для нескольких диалектов под- или наддиалекта (ср., например, язык фольклора или использование литературного языка как некоторого официального стиля носителями различных диалектов современных языков). Однако при рассмотрении территориального размещения языковых явлений целесообразно так или иначе отвлечься от фактов социально-стилистического плана и сосредоточить внимание только на тех явлениях, варьирование которых в пане время определяется лишь географически.

Но при этом условии многомерность диалектной системы остается. Она, естественно, не облегчает исследования, по в свете си-

стемного подхода к фактам языка учет этой многомерности совершенно необходим, хотя практическое осуществление такого учета требует особых приемов.

Чтобы избежать многомерности диалектной системы, в порядке первого приближения можно прибегнуть к некоторым упрощениям. Наиболее простое из них состоит в том, что строится географическая система некоторого факта *a*, затем строится такая же система факта *b*, факта *c* и т. д.; эти географические системы накладываются друг на друга, и то, что получается в результате таких наложений, рассматривается как система диалектов. Приближенность такого способа рассмотрения диалектной системы вытекает из того, что при его использовании исчезает по сути дела система каждого из диалектов. Если для наглядности представить географическую систему как некоторую горизонтальную сетку, а систему каждого из диалектов как вертикальный столбец, то при описанной процедуре построения диалектной системы языка мы получим лишь ряд горизонтальных срезов, совмещение которых не всегда правомерно, так как явления, попадающие в срез в одном диалекте, в системе другого диалекта могут занимать совершенно иное место в отношении других явлений системы данного диалекта. Системы отдельных диалектов, существующие в реальности, при такой процедуре разрушаются, а взамен них конструируются некоторые искусственные вертикальные системы. Аналогией такого подхода к построению диалектной системы языка может служить попытка построения прайзыка без учета относительной хронологии, без учета взаимоотношения элементов прайзыка; при такой попытке прайзык получался бессистемным набором, а не системой явлений. Современная же диалектология стремится к построению системы, ее не может удовлетворить просто сбор материала, его накопление (хотя сбор и накопление фактов продолжают оставаться в числе важнейших задач науки).

Для диалектологии важно знать не только, какую географическую систему составляют диалектные факты, но и то, как они взаимодействуют в системе отдельных диалектов. Это можно показать на примере систем родственных языков. Можно, например, допустить, что какое-то окончание сохранилось в разных языках. Так, в русском и чешском языках в родительном падеже у числительных типа *пять* выступает одинаковое окончание (русск.-*и*, чешск.-*i*), но этого еще недостаточно для утверждения того, что слово *пять* в русском и чешском языках склоняется одинаково. С другой стороны, слово *кость* в творительном падеже в русском имеет окончание *-ью*, а в чешском *-í*, но этого слова недостаточно для того, чтобы сказать, что слово это в русском и чешском языках в единственном числе имеет различные модели склонения. Рассмотрение небольшой подсистемы, которую можно было бы назвать микросистемой склонения этих типов слов, показывает, что типологически склонение слова *кость* в русском и чешском языках почти тождественно (различие состоит лишь в том, что в чешском сохраняется звательная форма); и в русском и в чешском языках слова этого типа имеют по три формы, одна из которых используется в иминительном и винительном падежах единственного числа, другая — в творительном, а третья — для остальных падежей. Слово же *пять* имеет в русском и чешском языках разные типы склонения: в русском — трехформенную парадигму, аналогичную рассмотренной только что парадигме слова *кость*, а в чешском — двухформенную парадигму с противопоставлением винительного-иминительного косвенным падежам. Существенно подчеркнуть, что рассмотрение даже всех падежей *по парно* в русском и чешском языках не могло бы дать того результата, который получается при сопоставлении целых

микросистем, ибо попарное совпадение или несовпадение окончаний падежей еще не дает основания говорить о тождестве или нетождестве парадигм в типологическом плане. Более того, различие русского *-ью* и чешского *-í* в склонении слова *кость* должно быть объяснено не на морфологическом, а на фонетическом или фонологическом уровне, а это значит, что для получения полной информации о соотношении русской и чешской парадигм склонения недостаточно самих лишь этих парадигм. А между тем, процедура построения системы диалектов как суммы географических систем отдельных фактов без учета роли этих фактов в системах отдельных диалектов таит в себе опасность того, что факты, получающие объяснение в системе своего диалекта, могут оказаться разрозненными и потерявшими все системы диалекта. Разрешающая способность лингвогеографического описания при этом понижается.

Система каждого языка и диалекта распадается на некоторые подсистемы, разумеется, связанные между собой, но все же относительно обособленные. Не забывая о связи между подсистемами, для удобства изучения иногда приходится прибегать к анализу именно этих подсистем (микросистем). Видимо, и в лингвогеографии было бы целесообразно обратиться не только к картографированию отдельно взятых фактов, но и таких подсистем (микросистем). Так, например, вполне реально провести изогlossenу не (или не только) отдельных падежных форм, но и изогlossenу типов (моделей, микросистем) склонения тех или иных групп слов.

Если взять для примера микросистемы склонения слов типа *пять* в современных славянских языках, то обнаруживается четыре типа таких микросистем, включающих соответственно одну, две, три или четыре формы. Распространенность этих типов микросистем варьируется по литературным языкам и по диалектам. Распространение этих типов микросистем (парадигм) по языкам в целом таково:

одноформенная: (болгарский, македонский), сербско-хорватский, словенский, словацкий, серболужицкие;

двухформенная: чешский, верхнелужицкий;

трехформенная: польский, русский, белорусский, украинский;

четырехформенная: украинский, словацкий, словенский, серболужицкие.

Следует отметить, что имеются вполне определенные связи этой микросистемы с другими микросистемами (парадигмами) склонения. Так, в частности, четырехформенная микросистема — большей частью уже не система собственно числительных типа *пять*, а система более широкого круга числительных (она включает также *три*, *четыре*, иногда и *два*). А типологически тождественная русской микросистеме трехформенная микросистема склонения украинских числительных типа *п'ять*, существующая с четырехформенной, имеет в творительном падеже окончание *-ма*, связывающее ее с четырехформенной парадигмой. С другой стороны, необходимо учитывать и частично уже отмеченные случаи сходства или несходства микросистемы склонения числительных типа *пять* с микросистемами склонения не только числительных, но и других частей речи (ср. сходство в ряде языков трехформенной микросистемы числительного *пять* с микросистемой склонения существительных типа *кость* в единственном числе или сходство четырехформенной микросистемы таких числительных с четырехформенной системой склонения прилагательных во множественном числе).

К сожалению, данные, приводимые даже в монографических описаниях диалектов, нередко очень бедны сведениями о системе склонения числительных; обычно упоминается, например, форма *пятёх* или *пятих*,

иногда даже без указания на то, имеется ли соответствующая форма для других падежей и как соотносится данная парадигма (если она существует) с другими парадигмами. Поэтому имеющиеся в диалектных описаниях сведения лишь несколько дополняют наши сведения о распространении тех или иных микросистем склонения числительных типа *пять*, по пока не могут быть достаточными ориентирами для уточнения границ ареалов. Так, одноформенная микросистема (несклоняемость хотя бы в некоторых случаях) представлена в некоторых украинских, белорусских, польских говорах; двухформенная микросистема известна в русских говорах; четырехформенная микросистема встречается в ряде говоров русского, белорусского, польского языков.

Таким образом, оказывается, что наибольшее распространение в современных славянских языках в территориальном отношении имеют две «полюсные» микросистемы склонения числительных типа *пять* — одноформенная и четырехформенная, причем последняя, несмотря на свою распространенность, по крайней мере в литературных языках не является единственной: с ней сосуществуют другие микросистемы склонения. С другой стороны, двухформенная и трехформенная микросистемы встречаются в более ограниченных ареалах. Уже эти факты, сопоставленные с микросистемами склонения других числительных, а также других частей речи, позволяют высказать мысль о некоторых особенностях экспансии одноформенной и четырехформенной микросистем склонения числительных типа *пять* в славянских языках. Что-то, видимо, мешает полному принятию четырехформенной парадигмы, «выгодной» тем, что она сближает различные типы числительных — *пять*, с одной стороны, и (*два*), *три*, *четыре*, с другой. Возможно это «что-то» заключается в том, что принятие такой парадигмы для числительных привело бы к поглощению их прилагательными. Одноформенная парадигма, если не учитывать в данном случае болгарского и македонского языков, вообще утративших склонение, стала единственной только в сербско-хорватском языке; во всех других славянских языках она используется как факультативная в определенных случаях. К сожалению, неполные данные об использовании такой парадигмы в говорах не дают возможности сделать заключения о некоторых возможных условиях принятия одноформенной парадигмы как единственной. Но сравнительно большая распространенность одноформенной парадигмы позволяет думать, что хотя бы в некоторых отношениях она соответствует тенденциям в эволюции славянских числительных. Полные диалектные данные, быть может, дали бы возможность развить эти соображения, в частности установить «эпицентры» экспансии одно- и четырехформенной микросистем.

Нельзя, конечно, полностью отрицать значения изоглосс отдельных, даже изолированных, фактов. Но следует думать, что именно изоглосса микросистемы дает больше возможностей для обобщений, для вскрытия особенностей языковых явлений. Нужно сказать, что иногда изоглосса отдельной формы может рассматриваться как представитель (репрезентант) некоторой микросистемы. Так, есть некоторые основания допустить, что формы *пятёх*, *пятих* и подобные являются репрезентантом четырехформенной микросистемы. В этом случае изоглосса формы *пят-x* будет как бы знаком изоглоссы всей микросистемы и может быть в качестве такой использована при различных интерпретациях данных лингвистической географии. При этом, однако, весьма существенно иметь твердое убеждение, что та или иная форма действительно является репрезентантом той или иной микросистемы. А убедиться в этом — по существу равносильно проведению системной изоглоссы.

В рассмотренном примере системная изоглосса проводится для небольшой микросистемы. Однако можно провести изоглоссы, характеризующие более крупные языковые подсистемы. Правда, в этом случае, по-видимому, нельзя будет ограничиться непосредственными наблюдениями или сравнительно простыми обобщениями, подобными тем, которые производятся при построении парадигмы одного типа склонения. Здесь нужны будут более генерализованные данные. Так, если бы в качестве объекта для картографирования было избрано склонение существительных в славянских языках, можно было бы воспользоваться обобщающими данными о склонении, которые приведены, например, в коллективном докладе группы советских славистов «Об одном подходе к типологии славянских языков»¹ на таблицах XIV и XV (количество типоформ, неразложимых сочетаний граммем, морфов, число типов нейтрализации; степени отклонения от синтетически простого языка, от аналитически простого языка, от идеального языка).

Можно было бы попытаться построить изоглоссы, характеризующие склонение числительных в целом. Для склонения числительных показательно количество типоформ. По убыванию этого количества в склонении числительных славянские языки располагаются в следующем порядке: украинский, белорусский, русский, польский, чешский, нижнелужицкий, верхнелужицкий, словенский, словацкий, сербско-хорватский, (македонский, болгарский). Учитывая значительные колебания в использовании форм склонения, характерные для славянских языков, можно думать, что наличие более полных сведений о склонении числительных в славянских диалектах могло бы дать очень интересные материалы для воссоздания картины истории и современного состояния славянских числительных. К сожалению, такие сведения имеются лишь по считанному числу диалектов.

Надо сказать, что пример с сосуществованием микросистем склонения одного и того же числительного в одном и том же языке на одном и том же этапе его развития, иногда даже в одинаковых окружениях, не единичен. В сосуществовании языковых явлений можно видеть проявление языковой динамики. Поскольку история языка творится в синхронии, можно полагать, что именно в процессе выбора одной из сосуществующих форм осуществляется тенденция к эволюции употребления этих форм. В связи с этим немалый интерес представляет использование статистического анализа при характеристике тех или иных существующих явлений в диалекте или языке. Модель системы диалекта или системы систем диалектов будет неполной без указания относительного веса, вероятности тех или иных существующих (и тем самым конкурирующих) явлений. В диалектологии славянских языков имеется некоторый опыт использования статистических данных. Можно указать на наблюдения В. Дорошевского над носовыми гласными в мазовецких говорах, по материалам которых еще в тридцатых годах он поставил вопрос о статистическом представлении изоглосс². Из работ последнего времени с более близким к теме этой статьи материалом следует отметить работу П. Янчака о возможностях статистической разработки колеблющихся диалектных явлений на лингвистических картах³, вы-

¹ З. М. Волоцкая, Т. Н. Молошная, Т. М. Николаева, И. И. Ревзин, Т. В. Цивьян. Об одном подходе к типологии славянских языков. В сб.: «Славянское языкознание», М., 1963.

² W. Dorożewski. Pour une représentation statistique des isoglosses.—«Bulletin de la Société de linguistique de Paris», XXXV, 1, 1935; ср. также: В. Дорожевский. Диалектология и сравнительно-исторический метод в языкознании.—«Вопросы языкознания», 1956, № 5.

³ Р. Янчак. Možnosti statistického zpracování, kolísavých nářečních jevů na jazykových mapách.—«Prace filologiczne», t. XVIII, cz. 1, Warszawa, 1963.

полненную на материале форм родительного падежка *uhlí/uhlího* в чешских говорах. По-видимому, возможно не только несколько иное графическое представление статистических данных на картах (например, при помощи секторных диаграмм или при помощи линий, соответствующих некоторым статистическим показателям в тех или иных местностях), чем то, которое предлагается Янчаком, но и использование в лингвистической географии некоторых более сложных статистических параметров.

Такие более сложные параметры целесообразно использовать, видимо, в том случае, когда имеется необходимость, с одной стороны, дать некоторые статистические сведения, учесть вероятности появления некоторых языковых явлений, а с другой стороны, учесть системный характер этих явлений. Так, например, можно воспользоваться коэффициентом корреляции между рядами цифр, характеризующих вероятность появления тех или иных форм в различных диалектах или языках⁴. Конечно, коэффициент корреляции не отражает всего богатства материала, в частности статистического материала, об употребительности тех или иных форм в говорах; но он в обобщенном виде позволяет представить отношения между диалектными системами с колеблющимися данными о частоте отдельных элементов этих систем. Таким образом, картографирование коэффициента корреляции в известной мере аналогично картографированию, например, степени отклонения от идеального языка, в рассмотренном примере тоже, конечно, не отражающей всего богатства склонения славянских существительных, но зато дающей некоторое сравнительное представление о соотношении особенностей систем склонения существительных в исследуемых языках или диалектах.

Одним из интересных колеблющихся во всех славянских языках грамматических явлений, которое вполне сравнимо и потому удобно в качестве примера, является согласование в числе сказуемого с количественным подлежащим. Речь идет, таким образом, о колебаниях формы числа сказуемого в конструкциях типа: *Пять человек идет/идут по улице*. Такие колебания в большей или в меньшей мере присущи всем славянским языкам, причем, что существенно, варьируются в зависимости от того, какое числительное употреблено в подлежащем. Поэтому можно говорить о некоторой статистической системе согласования в числе сказуемого с количественным (числительным) подлежащим в каждом из славянских языков (и диалектов) и путем использования коэффициентов корреляции между этими системами показать отношения между ними в географическом аспекте.

Одна из возникающих при этом проблем — выбор эталона, стандарта, с которым целесообразно производить сравнение. В качестве такого эталона может быть избрано некоторое отвлечение от реальных свойств явлений, как, например, «идеальный язык» в упоминавшемся выше коллективном докладе группы советских славистов или некоторая генетически предшествующая модель, сравнение с которой позволило бы в наибольшей степени выявить имеющиеся в возникших на ее базе языковых или диалектных системах тенденции. В диалектологических разысканиях, быть может, более правомерно сравнение с реальными языками или диалектами, воплощающими в себе те или иные особенности вариации данной системы. Каждый из этих эталонов обладает своими достоинствами и своими недостатками; поиск эталона вместе с тем было бы целесообразно вести не только с учетом особенностей данной микросистемы, но и всей языковой системы в целом: всеобщий эталон мог бы быть очень

⁴ Об использовании в языковедческих исследованиях коэффициентов корреляции см., например, Р. М. Фрумкин. Статистические методы изучения лексики. М., 1964.

полезен для обеспечения сравнимости различных данных. В этом смысле в поисках такого эталона много общего с поисками языка-посредника для машинного перевода и поисками языковых универсалий.

Коэффициенты корреляции доли единственного числа в согласовании сказуемого с количественным подлежащим были рассчитаны по отношению к чешскому языку, как языку с наименьшими колебаниями в употреблении форм числа сказуемого при количественном подлежащем, языку, в котором это согласование весьма регламентировано в зависимости от того, какое числительное выступает в подлежащем, к болгарскому языку, где при относительно строгой регламентации решительно преобладает множественное число сказуемого, а также к русскому языку, в котором согласование сказуемого с количественным подлежащим наименее регламентировано. В качестве типов подлежащих, при которых варьируется вероятность появления единственного числа сказуемого, избраны подлежащие, включающие числительные *один*, *два*, *три — четыре*, *двоев*, *пять, много — несколько*. Более дробное членение типов подлежащих для проверки их согласования со сказуемыми в восточнославянских языках, вследствие особенностей числительных в ряде языков, было бы затруднительно и существенно не изменило бы полученных коэффициентов. Полученные коэффициенты приводятся в таблице; следует иметь в виду, что второй десятичный знак в таблице приводится условно, ибо в ряде случаев погрешность превышает 0,1 (в известной степени допустимо пренебречь такой погрешностью, поскольку все коэффициенты получены одинаковым образом, но иметь в виду эту погрешность, так же как и оценочный характер некоторых исходных данных, необходимо). К сожалению, мы не располагаем соответствующими данными по славянским диалектам, которые могли бы осветить многое из того, что сейчас не ясно; но смысл данных, приводимых в таблице и представленных на рис. 1—3⁵, состоит в нашем примере не только

Таблица

По отн. к яз.	Бол	М	СХ	Слн	СЛ	П	Ч	Слц	Р	У	Бел
Р	0,87	0,88	0,95	0,89	0,97	0,91	0,66	0,59	1,00	0,98	0,99
Бол	1,00	1,00	0,81	0,73	0,84	0,71	0,71	0,84	0,87	0,88	0,83
Ч	0,71	0,73	0,80	0,99	0,98	0,86	1,00	0,98	0,66	0,75	0,64

и, пожалуй, не столько в географической интерпретации конкретного языкового явления (согласования сказуемого с количественным подлежащим), сколько в утверждении возможности представления при помощи коэффициента корреляции географической системы языковых систем, элементы которых колеблются, варьируются.

Бессспорно, установление процентного отношения тех или иных явлений в диалектах не может считаться легкой задачей, по без ее решения нам не удастся с необходимой глубиной проникнуть в понимание закономерностей функционирования и эволюции некоторых фактов. Можно допустить, что при сборе материала для описания говора будет спровоцировано повторение некоторых необходимых конструкций, ана-

⁵ Сокращения в таблице и на рисунках означают: Бел — белорусский, Бол — болгарский, М — македонский, П — польский, Р — русский, СЛ — серболужицкие, Слн — словенский, Слц — словацкий, СХ — сербско-хорватский, У — украинский, Ч — чешский языки.

лиз употребления которых с известным приближением позволит говорить об их варьировании хотя бы в некоторых узловых пунктах.

Получение системных и статистических данных по говорам, по-видимому, даст возможность вскрыть некоторые межъязыковые изоглоссы, уточнить ареалы распространения языковых явлений, раскрывая тем самым в ряде случаев механизм действия языковых тенденций, в частности в области языковых инноваций, которые изучаются еще недостаточно. Необходимость получения таких данных диктуется, однако, не только внутренними потребностями диалектологии, но и запросами тех отраслей науки, которые являются «потребителями» диалектологических материалов, в частности запросами истории языка, которая обнаруживает все большую тягу к системному и вероятностному анализу фактов. С другой стороны, нельзя не учитывать, что в современном мире диалекты в чистом виде постепенно отмирают, причем процесс этот происходит со все возрастающей скоростью. Разрушение, отмирание «чистых» говоров сказывается прежде всего именно на системе. Если отдельные факты еще и сохраняются в виде больших или меньших островков в море наступающих литературных языков (такими островками в большинстве современных славянских языков являются, например, реликты системы двойственного числа), то цельность системы диалектов, а тем самым и системы диалектных систем будет нарушена. Система диалектных систем уже претерпела

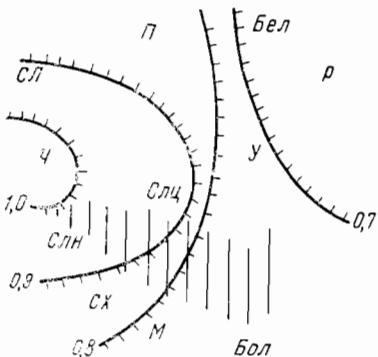


Рис. 1

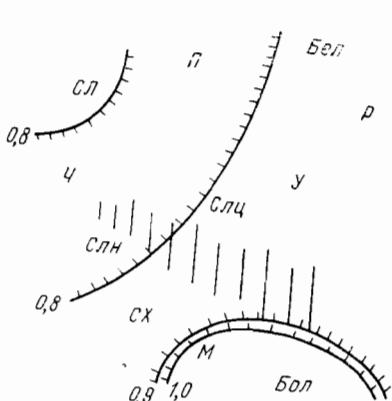


Рис. 2

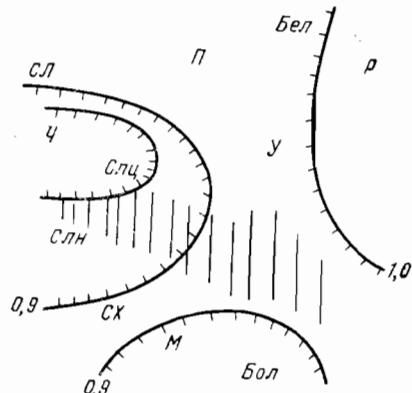


Рис. 3

немало преобразований, и на пути к ним стоят еще большие преобразования, в общем, видимо, направленные на стирание географических различий в языках. Поэтому то, что не будет записано в ближайшие годы, спустя два, три или пять десятков лет будет потеряно для науки навсегда.

Вот почему представляется разумным при сборе диалектологических материалов в наше время совмещать достаточно подробные, монографического типа описания определенной сетки пунктов, производимые по строгому плану, с более поверхностными характеристиками более

густой сетки точек. Это и предусматривается в «Вопроснике Общеславянского лингвистического атласа» для сбора материалов, по которому предполагается использовать полный и сокращенный варианты вопросника. В «Вопроснике» совершенно справедливо отмечается, что экспраторы «обязаны в каждом населенном пункте установить систему вокализма, консонантизма и просодии»⁶. Наряду с этим было бы, видимо, целесообразно устанавливать хотя бы основные системы склонения и спряжения. С другой стороны, эти системы должны быть отражены не только «в томах пояснений к картам атласа», но и непосредственно на самих картах. Хотелось бы, чтобы такой фундаментальный научный труд, каким является Общеславянский лингвистический атлас, отражал бы системный подход к языку не только на фонологическом уровне и отражал бы его в основной своей части (на картах).

За лингвистическими картами, за лингвистическими атласами, конечно, может и должно сохраняться значение капитальных и географически систематизированных собраний диалектных материалов. Но первичный материал может и должен также получить на картах определенное обобщение. Это вполне соответствует мысли о том, что «лингвистическим картам в полной мере присуще неотъемлемое свойство любой карты — генерализованность изображения»⁷. Построение системных, а также системно-статистических изоглосс как раз и является тем обобщением явлений и контуров их распространения, о котором говорят картографы, обобщением, вытекающим из системного устройства самого картографируемого объекта — языка. Можно думать, что генерализованное представление некоторых явлений языка на лингвистических картах отнюдь не понизит научного уровня карт, а, напротив, повысит его, поднимет их до уровня необходимых в науке обобщений. Компрессия информации о лингвистических явлениях в говорах, которая будет достигаться на системно-статистических картах, сделает их еще более ценным орудием в руках лингвистов, так как системное и статистическое представление изоглосс помогает проникнуть в глубины системы диалектных систем.

⁶ «Вопросник Общеславянского лингвистического атласа». М., 1965, стр. 10.

⁷ Е. М. Пospelov. Некоторые картографические вопросы лингвистической географии. — «Советское славяноведение», 1966, № 1, стр. 61.