

**В. В. Дьячков (ИЯз РАН)**

**V. V. Dyachkov (Institute of Linguistics, Russian Academy of Sciences)**

**О некоторых проблемах терминологии, принятой в базе данных**

**«Языки мира»**

**Towards some terminological problems of the database “Languages of the World”**

**Аннотация**

В данной статье рассматривается проблема представления признаков в базе данных «Языки мира», разрабатываемой на основе одноименной энциклопедии в секторе прикладного языкознания Института языкознания РАН. В статье кратко описываются теоретические проблемы, связанные с представлением некоторых признаков в базе. На примере представлений категории аспекта в базе предлагается решение проблем подобного типа, базирующееся на достижениях современной типологии.

This article considers the problem of representation of features in the database "Languages of the World", developed at the department of applied linguistics in the Institute of Linguistics (Russian Academy of Science) on the basis of the encyclopedia of the same name. The article briefly describes the theoretical problems associated with the representation of certain features in the database. A solution to problems of this type based on the achievements of modern typology is proposed illustrated by the example of the category of aspect presented in the database.

**Ключевые слова**

Типологические базы данных, аспектология, типология

Typology-oriented databases, theory of aspect, typology

Данная статья посвящена проблеме признакового пространства базы данных «Языки мира», а именно — вопросу его оптимального представления в новой версии. Эта проблема связана с работой по проверке признакового пространства, которая проводи-

лась в 2013—2018 гг. В ходе этой работы автором данной статьи было предложено большое количество изменений в названиях признаков, обусловленных как фактически-ми ошибками в признаковом пространстве, так и необходимостью смены многих терминологических установок, принятых в предыдущей версии базы данных (далее — БД).

Кратко опишем основные типы ошибок, встречающихся в базе данных. К ним мы относим: 1) терминологические ошибки; 2) признаки, о которых в статье, посвященной конкретному языку, нет никакой информации (или соответствующая статья не найдена в БД); 3) устаревшие термины, имеющие аналоги в самой БД; 4) опечатки. Данная классификация основана на нашем собственном исследовании языкового пространства (о проблеме корректности признаков в БД «Языки мира» см. также [1]).

Мы не будем касаться подробно вопроса опечаток, поскольку он не представляет собой теоретически нагруженной проблемы, в отличие от всех остальных. Точно так же не является интересной проблема признаков типа 2. Наличие признаков, о которых нет никакой информации в печатной версии энциклопедии, связано с тем, что информация о ряде языков вносилась вручную, на основании рукописей, и некоторые редкие признаки оказались связанными исключительно с этими языками. По нашему мнению, информация о таких языках должна быть либо удалена из базы, либо перепроверена вручную.

Что касается терминологических ошибок, их наличие может установить специалист, основываясь на понимании корректности соотнесения некоторого термина с обозначаемым им явлением. Ошибки подобного типа связаны с тем, что термины употребляются по отношению к явлению, которое нельзя описать данным термином. Приведем два примера. В базе данных было обнаружен признак «твердые/мягкие гласные» (с частотностью  $F = 1$ ). Некорректность данного признака обусловлена, очевидно, тем, что гласные не могут быть твердыми или мягкими, и если в какой-то языковедческой традиции гласные делятся на два подкласса с такими названиями, эти названия не должны фигурировать в типологически ориентированной базе данных.

Второй пример связан с тем, что один из признаков содержит следующие противопоставления по долготе: «долгие/краткие/сверхкраткие/беглые гласные» (что характерно, частотность это признака  $F = 0$ ). Признак также является некорректным, поскольку понятие беглого гласного, действительно принятое в некоторых языковедческих традициях, тем не менее, в общем случае не связано с понятием долготы: гласные, как известно, могут быть краткими или долгими, либо, в отдельных случаях, сверхкраткими или сверхдолгими.

Наибольшую сложность представляют собой признаки типа 3, который включает в себя термины, не используемые в современной типологии. О них мы поговорим ниже.

Стоит отметить, что полное удаление из базы данных «неправильных» признаков противоречило бы общей установке на максимальное сохранение информации в базе данных. Даже в том случае, если термин признается неверно употребленным и/или находящимся вне теоретических установок современной типологической науки, в ряде случаев сохранение этого термина может быть оправдано: так, термин «порода», употребляемый в арабистике по отношению к типам спряжения глагола, может быть переформулирован как более общий термин, например, как «тип спряжения». Вместе с тем, как представляется, оставить этот признак в базе данных представляется более целесообразным решением: во-первых, терминология, принятая в конкретных языковедческих традициях, может оказаться более точной и подходящей для данного языка, а во-вторых, проблема решается путем введения в глоссарий базы термина «порода» с его описанием.

Вместе с тем, как представляется, отдельные признаки и их значения, присутствующие в базе, нуждаются в значительной корректировке. В ходе нашей работы с признаковым пространством наиболее обширное расхождение терминологии базы с общепринятой типологической практикой мы выявили на примере представления категории аспекта. Более того, именно этот раздел базы, как нам представляется, является единственным разделом, нуждающимся в радикальной перестройке. Коротко резюмируя, представим набор основных значений признака «Основные видовые противопоставления» в том виде, в котором они сейчас представлены в базе:

- 1 Длительный/ недлительный
- 2 Законченный/ длительный
- 3 Однократный/ многократный
- 4 Однократный/ многократный/ итеративный/ результативный
- 5 Предельный/ непредельный
- 6 Пунктуальный/ курсивный
- 7 Результативный/ нерезультативный
- 8 Совершенный/ несовершенный
- 9 Совершенный/ несовершенный/ нейтральный
- 10 Совершенный/ несовершенный/ перфектный/ стативный
- 11 Длительный — недлительный/ результативный — нерезультативный/ хабитуальный — нехабитуальный
- 12 Кратный — некрatный/ процессуальный - непроцессуальный/ совершенный — несовершенный
- 13 Однократный — постоянный — прерывистый/ результативный — нерезультативный/ совершенный — несовершенный
- 14 Конкретный — неконкретный/ совершенный — несовершенный
- 15 Длительный — недлительный/ однократный — многократный/ совершенный — несовершенный
- 16 Длительный — недлительный/ результативный —- нерезультативный
- 17 Длительный — недлительный/ совершенный — несовершенный
- 18 Законченный — длительный/ однократный — многократный/ совершенный — несовершенный
- 19 Законченный — длительный/ результативный — нерезультативный/ совершенный — несовершенный
- 20 Однократный — многократный/ результативный — нерезультативный/ совершен-

ный — несовершенный

21 Однократный — многократный/ совершенный — несовершенный

22 Результативный - нерезультативный/ совершенный - несовершенный

Таким образом, инвентарь принятых на данный момент терминов выглядит следующим образом: *(не)длительный, законченный, однократный, многократный, итеративный, результативный, (не)предельный, (не)результативный, пунктуальный, курсивный, (не)совершенный, нейтральный, перфектный, стативный, (не)длительный, (не)хабитуальный, (не)кратный, (не)процессуальный, (не)совершенный, постоянный, прерывистый, (не)конкретный*. Нетрудно увидеть, что подобные термины не соответствуют ни одному из принятых в современной аспектологии терминов (см., например, обобщающие исследования [5] и [6]). Из приведенного списка видно, что существующими значениями признака покрывается также семантическая зона множественности (см. классификацию ее типов в [3] и [4]). С учетом того факта, что указанные выше исследования носят обобщающий фундаментальный характер, каталогизируют факты многих языков мира и предлагают основания для межъязыкового сравнения аспекта и глагольной множественности, возникает вопрос об интеграции научных достижений типологии в теоретические основания базы данных «Языки мира».

Вместе с тем, как мы уже указали выше, есть некоторая общая необходимость сохранять информацию, представленную в базе, поэтому процесс изменения названия признаков не должен уменьшать ее количество. В связи с этим возникает вопрос, как увязать имеющиеся в базе признаки с предполагаемыми изменениями.

Самое простое решение вопроса заключается в том, чтобы априорно задать некоторую классификацию имеющихся в языках мира аспектуальных значений в терминах межкатегориальных типов Э. Даля (см. [6]). Такая классификация предполагает, что любая аспектуальная граммема в любом языке мира может быть сведена к некоторому межкатегориальному типу — перфекту, имперфекту, плюсквамперфекту и т.д. Наблюдение Даля состоит в том, что в языках мира множества выражаемых граммемами значений

имеют тенденцию к группировке, и эти тенденции повторяются из языка в язык: скажем, результативное значение часто группируется со значением недавно произошедшего события, экспериенциальным значением, и такие пучки значений можно выделять на большом множестве языков. Именно эти пучки называются межкатегориальными типами.

Если принять подход Даля, для каждого языка, имеющегося в базе, необходимо выделить набор аспектуальных граммем и понять, к какому межкатегориальному типу они относятся. При этом мы можем подвергнуть процессу классификации только т. н. внутренний аспект, в то время как другой вид аспекта, внешний, который мы здесь именуем множественностью, окажется неохваченным. Для присвоения ярлыков мы должны, таким образом, построить некоторую новую априорную классификацию показателей множественности. Подход обладает еще одним существенным недостатком: он предполагает, что в каждый из более трехсот рефератов, описывающих языки, требуется (в общей сложности 300 раз) вносить изменения, что является трудоемким процессом.

Стоит отметить, что в существующем варианте описания признака «Основные видовые противопоставления» принята некоторая характеристика аспектуальной системы языка в бинарных терминах, что обуславливает наличие в описаниях противопоставлений типа «длительный — недлительный», «кратный — некратный» и т. д. К сожалению, необходимо признать, что подобные бинарные оппозиции не используются в современной типологии, и нельзя сказать, что они исчерпывающим образом описывают имеющиеся в языках видовые противопоставления.

Кроме того, среди имеющихся терминов встречаются такие, которые следует, по видимому, признать устаревшими или полностью/частично соответствующим другим терминам из списка. Так, термин *законченный* с большой долей вероятности соответствует термину *совершенный*, хотя, как представляется, термин *законченный* может обозначать и граммему комплетива. Термин *многократный* по определению должен быть отождествлен с термином *итеративный*, поскольку итеративный аспект по определению является способом выражения многократного события (см. [3]). Сомнительным представля-

ется также термин *кратный*, который, судя по его внутренней форме, также призван обозначать неоднократно происходящие события.

Не представляется оправданным употребление термина *курсивный*, аналога которому в современной типологии, по-видимому, нет. Описательными терминами, не соответствующими никаким сравнительным концептам современной типологии, являются также, по-видимому, термины *нейтральный*, *постоянный*, *прерывистый*, *(не)конкретный*.

Употребление некоторых терминов противоречит принятому в настоящее время разделению аспекта и акциональности, принятому в двухкомпонентных теориях вида (см. обсуждение данного вопроса в [2]). Исходя из этого понимания вида, следует отнести такие термины, как *(не)предельный*, к сфере акциональности и исключить из нашего списка. Это решение продиктовано соображением о том, что, хотя в языках мира есть граммы, вводящие в структуру глагола именно предельность (предельность вводят, например, русские префиксы в словах типа *написать*, *зарезать* или *ужалить*), неверно считать, что именно противопоставление по предельности/непредельности глагола является центральным видовым противопоставлением. В русском языке таковым является противопоставление перфективного и имперфективного аспекта. Эта система является экзотической в сопоставлении с другими языками мира, однако в целом хорошо вписывается в теорию межкатегориальных типов (см. [6]).

Итак, если вычеркнуть из нашего списка все названные выше признаки, которые мы сочли некорректными или не укладывающимися в современные типологические представления, то список оставшихся признаков будет выглядеть следующим образом (мы некоторым образом классифицируем признаки в соответствии с нашим представлением об организации семантической зоны аспекта):

- 1) *(не)длительный, (не)процессуальный*
- 2) *(не)результативный, (не)совершенный*
- 3) *однократный, итеративный, пунктуальный*
- 4) *перфектный, стативный*

### 5) (не)хабитуальный

В данном списке вполне употребительными являются термины *хабитуальный*, *перфектный*, *итеративный* — они являются общепринятыми и могут быть оставлены в базе данных. Корректным является также термин *стативный*, поскольку в языках мира действительно встречается особый класс стативных глаголов, который по своим свойствам сильно отличается от нестативных глаголов.

Что касается остальных терминов, то они по-прежнему являются спорными. Термины класса 1, по-видимому, относятся к классу понятий, выражаемых межъязыковым категориальным типом Прогрессив (см. [6]), однако точное решение на этапе построения априорной классификации принято быть не может — необходимо дополнительное исследование языков, имеющих данный признак. Термины класса 2 имеют отношение к понятиям, описываемым межъязыковым категориальным типом Перфектив (манифестацией которого является в том числе и русский совершенный вид), а также к употребительному понятию результата.

Термины класса 3 *однократный* и *пунктуальный* относятся, возможно, к сфере глагольной множественности, а точнее, к понятию семельфактива, который выделяет некоторый квант действия, обозначаемого т. н. мультипликативным глаголом (см. [6, 56—67]). Кроме того, под термином *пунктуальный* могут подразумеваться и т. н. пунктивные глаголы, то есть глаголы, обозначающие мгновенные события (*рухнуть*, *лопнуть*). Представляется, впрочем, что такие глаголы имеются в любом языке, и поэтому их выделение представляется неоправданным, в отличие от выделения специальных средств выражения семельфактивных значений, которые имеются не во всех языках.

Суммируя сказанное, мы предложим таблицу, в которой мы для каждого признака, имеющегося в базе, представим вариант его замещения на другой признак, соответствующий употребительному в области аспектологии термину. В качестве дополнения к нашей априорной классификации мы также введем в нее оставшиеся термины, используемые в указанных нами работах по аспекту и глагольной множественности и обозначаю-



щие другие сравнительные концепты. Для простоты мы указываем в таблице только признаки без префикса *не-*.

термин, используемый в базе данных	межъязыковой тип/сравнительный концепт, соответствующий термину
<i>длительный</i>	прогрессив/имперфектив
<i>процессуальный</i>	
<i>законченный</i>	перфектив/комплетив
<i>предельный</i>	
<i>совершенный</i>	
<i>результативный</i>	результатив
<i>перфектный</i>	перфект
<i>однократный</i>	семельфактив?
<i>пунктуальный</i>	семельфактив?
<i>курсивный</i>	-
<i>кратный</i>	итератив
<i>многократный</i>	
<i>итеративный</i>	
<i>прерывистый</i>	-
<i>постоянный</i>	хабитуалис?
<i>конкретный</i>	-
<i>хабитуальный</i>	хабитуалис
<i>стативный</i>	статив
<i>нейтральный</i>	-
-	конклюдив
-	экспериенциальный перфект
-	плюсквамперфект
-	мультипликатив

Получившийся список признаков, таким образом, является компромиссным вариантом, который, с одной стороны, позволяет сохранить основные имеющиеся в базе данных термины, а, с другой стороны, предлагает исчерпывающую классификацию типов явления. Стоит понимать, что наше предложение базируется на некотором наиболее общепринятом варианте классификации аспектуальных морфем и может не учитывать некоторые граммы, имеющиеся в языках мира. Вместе с тем, оно учитывает общепринятое на данный момент выделение межкатегориальных типов, которые являются основанием для сравнения аспектуальных грамм в разноструктурных языках. Предполагается, что в новом варианте базы признак «Основные видовые противопоставления» будет иметь значения из правого столбца приведенной выше таблицы, и при этом часть противопоставлений будет конвертирована в новые признаки автоматически, в то время как другая часть должна быть заполнена специалистами с учетом данных по конкретным языкам.

Итак, в данной статье мы рассмотрели на примере аспектуальных противопоставлений проблемы, возникшие в ходе нашей работы над базой данных «Языки мира», и предложили решение для самой главной, на наш взгляд, проблемы — проблемы представления категории аспекта, которое представляется архаичным для актуального состояния типологической науки. Безусловно, выводы и рекомендации, сделанные на основе исследования, имеют предварительный характер, однако, на наш взгляд, они должны быть учтены в процессе модернизации признакового пространства базы данных.

### **Литература**

1. Соловьев В. Д., Балалыкина Э. А., Байрашева В. Р. Анализ структуры базы данных «Языки мира». I. Словообразование // Труды школы по компьютерной и когнитивной лингвистике. Казань: Отечество. 2007.
2. Татевосов С. Г. Акциональность в лексике и грамматике. Глагол и структура события. М., Языки славянской культуры, 2015.

3. Храковский В. С. Семантические типы множества ситуаций и их естественная классификация // Храковский В. С. (ред.). Типология итеративных конструкций. Л., 1989. С. 5—53.

4. Шлуинский А. Б. Типология предикатной множественности: количественные аспектуальные значения. Дисс... канд. филол. наук. М., 2005.

5. Bybee J., Perkins R., Pagliuca W. The Evolution of Grammar (Tense, Aspect, and Modality in the Languages of the World). Chicago and London, The University of Chicago Press, 2004.

6. Dahl, Ö. Tense and Aspect Systems. Oxford: Blackwell, 1985.