

*Л. К. Бабурин*

## **СТЕПЕНЬ ВЗАИМНОЙ СИСТЕМНО-СИНЕРГЕТИЧЕСКОЙ БЛИЗОСТИ ИСПАНСКОГО, ФРАНЦУЗСКОГО И ИТАЛЬЯНСКОГО ЯЗЫКОВ**

*Работа представлена кафедрой романской филологии.*

*Научный руководитель – доктор филологических наук, профессор Р. Г. Пиотровский*

**Статья посвящена определению степени системно-синергетической близости испанского, французского и итальянского языков на основании лексико-статистических данных. Результаты работы можно использовать для внутрилингвистической таксономии романских языков, которая отличается от традиционных классификаций, построенных на внешних критериях или частных лингвистических индикаторах.**

**The aim of the article consists in determining the degree of the system-synergic proximity among Spanish, French and Italian on the basis of the lexicostatistical data. The results of this research can be used for the immanent (in sense of L. Hjelmslev) linguistic taxonomy of the Romance languages, which differs from traditional classifications that use also extra-linguistic criteria.**

В нашей предыдущей работе<sup>1</sup> рассматривалась степень близости испанского языка с французским и итальянским. Здесь с применением тех же критериев рассматри-

вается также соотношение французского и итальянского языков. Не повторяя подробного описания процедуры исследования, напомним его основные принципы.

Работа проводилась в синхронии. Это означает, что в ней учитываются формальные и семантические сходства лексических единиц (ЛЕ) лишь на современном этапе их существования, вне прямой связи с условиями их происхождения. Вместе с тем при определенных затруднениях в данном исследовании использовался этимологический словарь романских языков.

**Методика исследования** выглядит следующим образом. Для каждой пары языков ( $L_i, L_j$ ) были построены сопоставительные таблицы. Одни из этих таблиц содержат номера ЛЕ, указанные в частотных словарях (ЧС), другие – те номера, которые получаются в каждой таблице после исключения из нее служебных слов. Столбцы в таблицах распределяются следующим образом:

- 1) нумерация ЛЕ языка  $L_i$ ;
- 2) ЛЕ языка  $L_i$  с их грамматическими пометами;
- 3) частоты ЛЕ языка  $L_i$ ;
- 4) нумерация ЛЕ языка  $L_j$ ;
- 5) ЛЕ языка  $L_j$  с их грамматическими пометами;
- 6) частоты ЛЕ языка  $L_j$ ;
- 7) разность между номером ЛЕ языка  $L_i$  и номером соответствующей ей ЛЕ языка  $L_j$  на данной строке;
- 8) квадрат разности по каждой строке; внизу – сумма всех квадратов разностей в столбце.

По полученным результатам построена сводная таблица, в которой к каждой рассматриваемой паре языков приводятся коэффициенты их корреляции.

Из нечеткого множества степеней сходства при помощи операции сгущения мы получили три:

- 1) полное сходство; сюда входят случаи полного соответствия одновременно по форме и по основным значениям, причем допускается формальное расхождение по фонетическим и/или морфологическим законам (так, в паре исп. *estar* – ит. *stare* фонетически закономерно соответствие *es* – *s* в начале слова перед согласным и морфо-

логически закономерно соответствие *-r* – *-re* в инфинитиве), а также отнесение одной и той же ЛЕ в ЧС разных языков к разным частям речи без изменения ее значения (ср.: исп. *roso* ај. – фр. *peu* ав.);

- 2) частичное сходство (с поправочным коэффициентом 0,5); сюда входят случаи расхождения, лишь отчасти обусловленного фонетическими и/или морфологическими законами (ср.: исп. *cada* – фр. *chaque*), префиксального и постфиксального расхождения (ср.: исп. *aparecer* – ит. *comparire*; исп. *ақин* – ит. *qui*; исп. *сумо* – фр. *comment*), а также вхождения слова в одном языке в состав эквивалентного ему сложного слова в другом (ср.: исп. *hoy* – фр. *aujourd'hui*);

- 3) несходство (с поправочным коэффициентом –1); сюда входят случаи полного расхождения по форме при полном или частичном совпадении основных значений.

В случаях супплетивизма работа велась следующим образом. В алфавитных списках ЧС приводятся частоты каждой формы того или иного слова. Это позволяет разбить супплетивные формы таким образом, чтобы сопоставить их сходные в языке  $L_i$  и языке  $L_j$  пары по отдельности. Возьмем, например, испанский глагол *ser* и эквивалентный ему французский глагол *être*. Для начала следует выделить ряды их форм по общим основам. В списке форм исп. *ser* выделяется четыре таких ряда: *er-*, *es* (одна форма), *fu-*, *s-*; сходным образом – у фр. *être*: *es* + *est*, *ût-* + *kt-*, *fu-*, *s-*. При выделении каждого такого ряда производится сложение частот тех форм, которые к нему относятся. Выделив эти ряды, мы получаем возможность сопоставить их по отдельности. Так, в полные соответствия включаются: исп. *es* и фр. *es* + *est*; исп. *fu-* и фр. *fu-*; исп. *s-* и фр. *s-*. Что касается фр. *ût-* + *kt-*, то этот ряд форм по фонетическим ключам входит в полное соответствие с исп. *estar*. При всех этих операциях происходит соответствующее перераспределение номеров в частотном списке как языка  $L_i$ , так и языка  $L_j$ , но лишь относительно каждой данной пары языков.

Определенные проблемы в приведенном примере вызывает исп. *er-*, не входящее в сопоставление ни с одним рядом французских форм. Ввиду этого исп. *er-* исключается из рассмотрения. Однако в том случае, если такие «лишние» основы выявляются и в языке  $L_i$ , и в языке  $L_j$ , они приводятся в таблице полных несоответствий. При этом, чтобы избежать излишней разбивки ЛЕ, основы, относящиеся к одной и той же ЛЕ, даются в таком случае на одной строке вне зависимости от того, однородны они или разнородны. Например, у испанского глагола *ir* выделяются следующие группы форм: *fu-*, *i-* + *yendo*, *v-*. Последняя входит в полное соответствие с основой *v-* эквивалентного итальянского глагола *andare*. В то же время две другие группы форм исп. *ir*, так же как основа *and-* ит. *andare*, остаются «незадействованными». Поэтому они заносятся в одну строку таблицы с полными несоответствиями, что выглядит следующим образом:

55.	<i>ir v.</i> ( <i>fu-</i> + <i>i-</i> + <i>yendo</i> )	815	49.	<i>andare v.</i> ( <i>and-</i> )	1115
-----	-----------------------------------------------------------	-----	-----	-------------------------------------	------

Второстепенные значения в данной работе не учитываются. Ведущим критерием выделения основных и второстепенных значений является частотность: если какой-либо ЛЕ языка  $L_i$  в пределах области полных и частичных соответствий эквивалентны две (и более) ЛЕ языка  $L_j$ , одна из которых по своей частотности меньше данной ЛЕ языка  $L_i$  в десять раз и более, то такая ЛЕ языка  $L_j$  исключается из рассмотрения (операция сгущения). Ср.: исп. *verdad* n. 326 – фр. *vérité* n. 97 + *vrai* n. 27. В данном случае исключается фр. *vrai*, частотность которого более чем в десять раз меньше частотности исп. *verdad*. В то же время если соответствие с подобным соотношением частот является единственным, то оно учитывается (ср.: исп. *siendo* av. 803 — фр. *quand* av. 19).

В тех случаях, когда ЛЕ языка  $L_i$  имеет в языке  $L_j$  полные и/или неполные соответствия, а также соответствия с полным рас-

хождением по форме, последние исключаются из рассмотрения. Например, исп. *palabra* имеет французские эквиваленты *parole* и *mot*. С первым из них оно обнаруживает полное сходство, со вторым – полное несходство. Соответствие исп. *palabra* – фр. *mot* не рассматривается. Вместе с тем действует правило основного значения (см. выше). Ср.: исп. *ciudad* n. 318 — фр. *ville* n. 184 + *citù* n. 24. Со словом *ville* испанское *ciudad* обнаруживает полное несходство, со словом *citù* – полное сходство. Однако *citù* имеет недостаточно высокую частотность по сравнению с *ciudad*, в отличие от *ville*. Поэтому соответствие исп. *ciudad* – фр. *ville* учитывается, а соответствие исп. *ciudad* – фр. *citù* исключается из рассмотрения.

Сходства, восходящие не к латинско-романскому, а к индоевропейскому периоду, в данной работе считаются несходствами. Так, исп. *luz* эквивалентно фр. *lumiure*. Их латинскими этимонами являются соответственно *lux* (основа *luc-*) и *luminaria* (< *lumen*). Слово *lux*, в свою очередь, происходит от и.-е. *\*louk-s*, а *lumen* – от родственного ему и.-е. *\*leuqs-men* (или *\*louqs-men*)<sup>2</sup>. В латинский период эти два слова не воспринимались как однокоренные, так как имели уже разные основы. Поэтому указанная испанско-французская лексическая пара заносится в таблицу несходств.

Обработка результатов, представленных в таблицах, была ориентирована на коэффициент корреляции по Спирмэну:

$$\rho = 1 - \frac{6 \sum_{i=1}^{i=n} d^2}{(n^3 - n)} = 1 - \frac{6 \sum_{i=1}^{i=n} d^2}{n(n^2 - 1)},$$

где  $d$  – разность между номером ЛЕ языка  $L_i$  и номером соответствующей ЛЕ языка  $L_j$ ,  $n$  – количество строк с лексическими сопоставлениями в данной таблице,  $i$  – номер в ЧС.

С учетом трех перечисленных выше степеней сходства приходилось усложнять

математическую обработку, используя формулу Спирмэна в измененном виде:

$$\rho = \left( 1 - \frac{6 \left( \sum_{i=1}^{i=n} d_1^2 + 0,5 \sum_{i=1}^{i=n} d_2^2 \right)}{(n_1 + n_2) ((n_1 + n_2)^2 - 1)} \right) \cdot \frac{n_1 + n_2}{n_1 + n_2 + n_3},$$

где  $d_1$  – разность между номером ЛЕ языка  $L_i$  и номером соответствующей ЛЕ языка  $L_j$  в таблице с полными сходствами,  $d_2$  – разность между номером ЛЕ языка  $L_i$  и номером соответствующей ЛЕ языка  $L_j$  в таблице с частичными сходствами,  $n_1$  – количество строк с лексическими сопоставлениями в таблице с полными сходствами,  $n_2$  – количество строк с лексическими сопоставлениями в таблице с частичными сходствами,  $n_3$  – количество строк с лексическими сопоставлениями в таблице с несходствами.

$$-1 = \rho = 1,$$

где 1 обозначает большую близость языка  $L_i$  и языка  $L_j$  по частотности (положительную связь),  $-1$  – их взаимную противопоставленность по частотности (отрицательную связь), а 0 – отсутствие связи по частотности между ними (внеположенность). Между родственными языками возможен показатель только выше нуля.

Приведем пример вычисления по результатам таблиц при помощи формулы Спирмэна.

Исп. – фр. (по номерам ЧС знаменательных ЛЕ)<sup>3</sup>:

$$\rho = (1 - 6(147287 + 0,5 \cdot 31115) : (103 + 22)((103 + 22)I - 1))((103 + 22) : (103 + 22 + 55)) \approx (1 - 6 \cdot 162844,5 : 1953000)0,7 \approx 0,5 \cdot 0,7 = 0,35.$$

Таблица 1

Результаты

Пара	Коэффициент корреляции	
	По номерам общего ЧС	По номерам ЧС знаменательных ЛЕ
Испанский – французский	0,15	0,35
Испанский – итальянский	0,21	0,4
Французский – итальянский	0,3	0,56

Такие результаты (табл. 1) требуют интерпретации, отличающейся от предложенной в нашей предыдущей статье<sup>4</sup>. В табл. 1 по номерам как общего ЧС, так и ЧС знаменательных ЛЕ прослеживается следующая иерархия исследуемых пар языков по убывающей степени близости: французский – итальянский, испанский – итальянский, испанский – французский. В таком случае итальянский язык представляется центральным (каковым он, собственно говоря, и является – по крайней мере, с историко-географической точки зрения) с «иррадиацией» в ареалы других романских языков. Эти ареалы при этом более удалены друг от друга, чем относительно центра, с которым они в языковом плане обнаруживают различную степень сходства. Такая схема соотносится с результатами исследования Дж. Э. Ри<sup>5</sup>, также использовавшего лексический критерий.

ПРИМЕЧАНИЯ

<sup>1</sup> Бабурип Л. К. Степень близости испанского языка с французским и с итальянским // Известия Российского государственного педагогического университета им. А. И. Герцена: Аспирантские тетради. 2007. № 15 (39). С. 42–50.

<sup>2</sup> Walde A. Lateinisches etymologisches Wörterbuch. 3., neubearbeitete Aufl. von J. B. Hofmann. Heidelberg: Carl Winter's Universitdsbuchhandlung. 1. Bd. (A-L), 1938. S. 824, 832.

<sup>3</sup> Конкретные числовые данные в приведенном математическом выражении, а также коэффициент корреляции по номерам общего ЧС в паре испанский – итальянский незначительно изменены при перерасчете. Просим извинения за неточности, допущенные нами в предыдущей публикации.

<sup>4</sup> Напомним, что в статье, указанной в первом примечании, речь шла о противопоставлении испанского и итальянского языков как принадлежащих непрерывной Романии (Romania continua – термин А. Алонсо) французскому как представителю прерывной Романии (Romania discontinua).

<sup>5</sup> Ср.: Kroeber A. L. Three Quantitative Classifications of Romance // Romance Philology. Berkeley, Los Angeles: University of California Press, Vol. XIV. N 3. February. 1961. P. 192–194.