

## О ЗАДАЧЕ ВЫБОРА ОФИЦИАЛЬНЫХ ЯЗЫКОВ ЕВРОСОЮЗА<sup>1</sup>

А.Ю. Филатов

*Институт систем энергетики им. Л.А. Мелентьева, Иркутск*  
*e-mail: fial@irlan.ru*

**Аннотация.** Работа посвящена одному из аспектов гетерогенности – проблеме меж-  
страновых коммуникаций. Изучаются подходы к решению задачи определения опти-  
мального числа и выбора состава официальных языков Евросоюза. Показано, что изме-  
нение характеристик ущемления существенно влияет на порядок включения языков в  
состав официальных.

**Ключевые слова:** теория игр, лингвистические модели, официальные языки Евросоюза.

### **Введение**

Миром правит гетерогенность. Различаются отдельные люди – у каждого из нас  
свои индивидуальные ценности и интересы, доходы и потребности. Людей можно объ-  
единить в различающиеся между собой социальные страты. Наконец, различаются  
страны – у каждой своя история, идеология, культура, доминирующая религия, титуль-  
ная нация, экономика, лингвистика.

Возникает вопрос: гетерогенность – это хорошо или плохо? С одной стороны,  
можно привести множество примеров (одними из самых ярких, наверное, являются Си-  
ликоновая долина, ставшая домом для людей со всего света, и институт Макса Планка  
в 30-х годах XX века), когда гетерогенность вела к всплеску культурного или техниче-  
ского прогресса. С другой стороны, повышенный уровень гетерогенности может спо-  
собствовать непониманию между людьми, дестабилизировать политическую обстанов-  
ку и даже вести к национальным и религиозным конфликтам.

Особенно внимательно данную проблему необходимо рассматривать при при-  
соединении к более-менее однородной группе нового существенно отличающегося уча-  
стника. В качестве примера можно привести проблему формирования межстрановых  
союзов, таких, как СНГ или Евросоюз – в частности, стоит ли включать в последний  
Турцию или, скажем, Украину, давать ли новым участникам некоторый особый статус  
и не станут ли данные действия причиной чьего-то ущемления в правах.

### **Постановка задачи выбора официальных языков Евросоюза**

Рассмотрим статистику. Евросоюз к 2004 году насчитывал 15 государств: Авст-  
рию, Бельгию, Великобританию, Германию, Грецию, Данию, Ирландию, Испанию,  
Италию, Люксембург, Нидерланды, Португалию, Финляндию, Францию и Швецию. В  
2004 году к Евросоюзу присоединились Венгрия, Кипр, Латвия, Литва, Мальта, Польша,  
Словакия, Словения, Чехия и Эстония. В 2007 году после вхождения Болгарии и  
Румынии состав Евросоюза расширился до 27 членов, и это не предел. Однако уже сей-  
час 23 языка являются государственными в этих странах. Среди них есть широко рас-  
пространенные: немецкий язык является родным для 85 миллионов жителей Евросою-  
за, английский – для 62, французский – для 61 миллиона. Есть малые языки: на маль-  
тийском и ирландском говорит всего 300–600 тыс. человек. Есть достаточно распро-  
страненные языки, не являющиеся государственными: русский (4,2 млн), каталонский  
(4,1 млн), турецкий (2,2 млн), арабский (1,6 млн).

---

<sup>1</sup> Исследования выполнены при финансовой поддержке РГНФ (проект 06-02-00266а)

Возникает проблема международных коммуникаций: если сократить число официальных языков Евросоюза до 1–2, то многие страны, включая крупные, оказываются в ущемленном положении. В то же время очевидна невозможность придать официальный статус не только всем 23 языкам, но и значимой их части. Затраты на перевод документов на все языки (62% изначально готовятся на английском, 26% – на французском и только 10 с небольшим процентов на остальных языках), а также связанные с этим задержки и возможные ошибки и разночтения, затраты на синхронный перевод официальных встреч и прочие издержки (притом что уже сейчас они составляют 17% общих расходов Евросоюза) многократно перекроют все выгоды сокращения ущемления. Необходим компромисс. Задача состоит в выборе количества официальных языков и в определении их состава.

В конце 2005 года были собраны данные, для скольких граждан Евросоюза какой из языков является родным, и сколько человек какой из языков знают на разговорном уровне. Результаты сведены в таблице 1. В частности, видим, что даже на английском языке не говорит и 50% населения Евросоюза. Говорящих на других языках (включая немецкий, являющийся родным для наибольшего числа людей) еще меньше.

**Табл. 1.** Число граждан Евросоюза, говорящих на различных языках

Язык	Родной, млн.чел.	Говорят, млн.чел.	Говорят / родной
Английский	62,4	182,6	2,93
Немецкий	85,3	121,7	1,43
Французский	60,7	97,2	1,60
Итальянский	57,7	64,8	1,12
Испанский	39,7	54,1	1,36
Польский	39,2	40,9	1,04
Румынский	21,0	22,2	1,06
Голландский	21,9	24,0	1,10
Русский	4,2	22,4	5,33

#### Учет близости языков

Важным при определении состава языков является их близость друг к другу. Очевидно, что среднестатистическому болгарину проще выучить другой славянский язык, нежели любой из латинских, для испанца португальский ближе английского, а выучить финский или греческий для кого бы то ни было – достаточно сложная задача.

В 1992 году была построена [1] матрица расстояний языков. Было отобрано 200 основных понятий, и для каждой пары из 95 индоевропейских языков определялось, сколько из этих понятий имеет общие корни. Если обозначить  $n_{lm}^0$  – число несовпадений корней, а  $n_{lm}^1$  – число совпадений, то расстояние между языками находится по формуле

$$y(l, m) = n_{lm}^0 / (n_{lm}^0 + n_{lm}^1).$$

Здесь сумма величин  $n_{lm}^0$  и  $n_{lm}^1$  не всегда равнялась двумстам, поскольку были понятия, для которых не удавалось в точности определить общность или различие корней. В то же время, если  $y(l, m) = 0$ , можно считать, что языки абсолютно близки, а если  $y(l, m) = 1$  – абсолютно различные.

В соответствии с полученными Дайеном и соавторами результатами 23 государственных языка, представленные в Евросоюзе, можно разделить на 8 групп:

1. Английский язык.
2. Германская группа (немецкий, голландский, шведский, датский).
3. Итальянская группа (французский, итальянский, испанский, португальский, румынский).

4. Славянская группа (словенский, чешский, словацкий, польский, болгарский).
5. Балтийская группа (литовский, латышский).
6. Греческий язык.
7. Ирландский язык.
8. Неиндоевропейские языки (финский, эстонский, венгерский, мальтийский).

Фрагмент матрицы расстояний языков для Евросоюза представлен в таблице 2. В таблице отсутствуют неиндоевропейские языки, а также 3 языка, получивших официальный статус в 2007 году: болгарский, румынский и ирландский. В свою очередь, в матрицу включены русский и украинский языки. Введены следующие обозначения:

IT – итальянский,	SW – шведский	CZ – чешский
FR – французский	DA – датский	SK – словацкий
SP – испанский	EN – английский	PL – польский
PT – португальский	LI – литовский	GR – греческий
GE – немецкий	LA – латвийский	RU – русский
NL – голландский	SV – словенский	UA – украинский

**Табл. 2.** Матрица расстояний между языками

	IT	FR	SP	PT	GE	NL	SW	DA	EN	LI	LA	SV	CZ	SK	PL	GR	RU	UA
<b>IT</b>	0	0,20	0,21	0,23	0,73	0,74	0,74	0,74	0,75	0,76	0,78	0,76	0,75	0,75	0,76	0,82	0,76	0,77
<b>FR</b>	0,20	0	0,27	0,29	0,76	0,76	0,76	0,76	0,76	0,78	0,79	0,78	0,77	0,76	0,78	0,84	0,77	0,78
<b>SP</b>	0,21	0,27	0	0,13	0,75	0,74	0,75	0,75	0,76	0,77	0,79	0,77	0,76	0,75	0,77	0,83	0,77	0,78
<b>PT</b>	0,23	0,29	0,13	0	0,75	0,75	0,74	0,75	0,76	0,78	0,80	0,78	0,76	0,76	0,78	0,83	0,77	0,78
<b>GE</b>	0,73	0,76	0,75	0,75	0	0,16	0,30	0,29	0,42	0,78	0,80	0,73	0,74	0,74	0,75	0,81	0,76	0,76
<b>NL</b>	0,74	0,76	0,74	0,75	0,16	0	0,31	0,34	0,39	0,79	0,80	0,75	0,76	0,75	0,77	0,81	0,78	0,79
<b>SW</b>	0,74	0,76	0,75	0,74	0,30	0,31	0	0,13	0,41	0,78	0,79	0,75	0,75	0,74	0,76	0,82	0,75	0,76
<b>DA</b>	0,74	0,76	0,75	0,75	0,29	0,34	0,13	0	0,41	0,78	0,80	0,73	0,75	0,73	0,75	0,82	0,74	0,76
<b>EN</b>	0,75	0,76	0,76	0,76	0,42	0,39	0,41	0,41	0	0,78	0,80	0,75	0,76	0,75	0,76	0,84	0,76	0,78
<b>LI</b>	0,76	0,78	0,77	0,78	0,78	0,79	0,78	0,78	0,78	0	0,39	0,66	0,62	0,60	0,64	0,83	0,62	0,63
<b>LA</b>	0,78	0,79	0,79	0,80	0,80	0,80	0,79	0,80	0,80	0,39	0	0,68	0,67	0,64	0,67	0,85	0,64	0,64
<b>SV</b>	0,76	0,78	0,77	0,78	0,73	0,75	0,75	0,73	0,75	0,66	0,68	0	0,34	0,31	0,37	0,82	0,39	0,36
<b>CZ</b>	0,75	0,77	0,76	0,76	0,74	0,76	0,75	0,75	0,76	0,62	0,67	0,34	0	0,09	0,23	0,84	0,26	0,24
<b>SK</b>	0,75	0,76	0,75	0,76	0,74	0,75	0,74	0,73	0,75	0,60	0,64	0,31	0,09	0	0,22	0,83	0,26	0,19
<b>PL</b>	0,76	0,78	0,77	0,78	0,75	0,77	0,76	0,75	0,76	0,64	0,67	0,37	0,23	0,22	0	0,84	0,27	0,20
<b>GR</b>	0,82	0,84	0,83	0,83	0,81	0,81	0,82	0,82	0,84	0,83	0,85	0,82	0,84	0,83	0,84	0	0,83	0,77
<b>RU</b>	0,76	0,77	0,77	0,77	0,76	0,78	0,75	0,74	0,76	0,62	0,64	0,39	0,26	0,26	0,27	0,83	0	0,22
<b>UA</b>	0,77	0,78	0,78	0,78	0,76	0,79	0,76	0,76	0,78	0,63	0,64	0,36	0,24	0,19	0,20	0,77	0,22	0

Пусть для каждого  $i$ -человека известен его родной язык  $n(i)$  и множество  $L(i)$  (для некоторых состоящее только из  $n(i)$ ) языков, которые этот человек знает. Обозначим за  $T$  множество официальных языков Евросоюза. В [2], [3] Вебером и соавторами введены 4 расстояния:

$$1. \delta_i^{(d,p)} = \begin{cases} 0, & n(i) \in T, \\ 1, & n(i) \notin T. \end{cases}$$

Данная величина равна нулю, если родной язык человека является официальным языком Евросоюза, и единице в противном случае.

$$2. \delta_i^{(d,s)} = \begin{cases} 0, & L(i) \cap T \neq \emptyset, \\ 1, & L(i) \cap T = \emptyset. \end{cases}$$

Данная величина равна нулю, если человек говорит на одном из официальных языков Евросоюза, и единице в противном случае.

$$3. \delta_i^{(y,p)} = \min_{t \in T} y(n(i), t).$$

Данная величина равна минимальному расстоянию между родным языком данного человека и одним из официальных языков.

$$4. \delta_i^{(y,s)} = \min_{l \in L(i), t \in T} y(l, t).$$

Данная величина равна минимальному расстоянию между одним из языков, на котором говорит человек, и одним из официальных языков.

Посчитав для любого набора официальных языков  $T$  среднее по всем гражданам Евросоюза значение  $\Delta^*(T)$  одной из этих величин, получим характеристику ущемления. Этот «индекс ущемления» необходимо минимизировать, учитывая, однако, существенно увеличивающиеся издержки при включении в набор новых языков. Таким образом, получаем двухкритериальную задачу, которую, в частности, можно решать с помощью следующей свертки:

$$\alpha \Delta^*(T) + |T|^\beta \rightarrow \min.$$

Есть и другие способы: условная минимизация, линейная свертка, максиминная свертка, метод идеальной точки и т.д.

### Наилучшие наборы официальных языков

Для начала попробуем отыскать наилучшие наборы официальных языков для любого их количества:  $|T| = 1, 2, 3, \dots$  Эти наборы будут различаться в зависимости от выбранной характеристики ущемления и возраста. Влияние возраста на знание языков продемонстрировано в таблице 3.

Табл. 3. Доля граждан Евросоюза, не говорящих на соответствующих языках, %

	Всего	15-29 лет	30-44 лет	45-60 лет	> 60 лет
Английский	63	45	59	68	76
Немецкий	75	74	75	76	75
Французский	80	78	81	80	81
Итальянский	87	87	87	87	87
Испанский	89	87	89	90	89
Польский	92	92	92	92	92
Голландский	95	95	95	95	95
Русский	95	96	95	95	96

При этом единственным языком, знание которого существенно меняется со временем, является английский. Его действительно молодежь в возрасте до 30 лет (а они определяют тенденции!) знает лучше, чем старшие поколения. Однако последовательность включения языков в набор официальных претерпевает некоторые изменения.

Приведем на схемах 1–3 соответствующие последовательности для индексов ущемления  $\Delta^{(d,s)}$  и  $\Delta^{(y,s)}$ . Обе характеристики не включают в число ущемленных людей, разговаривающих на одном из официальных языков, даже если это не их родной язык. Последовательности соответствуют всему населению, однако для первого из индексов дана также последовательность для молодежи в возрасте до 30 лет.

Приведены соответствующие значения индекса ущемления. К 23 государственным языкам добавлен русский язык. Ранее не рассматривавшиеся языки имеют следующие обозначения: RO (румынский), HU (венгерский), BG (болгарский), FI (финский). Если несколько вариантов последовательности включения языков в набор официальных принципиально не отличаются друг от друга, они обозначены как варианты а, б, с и т.д. Заметим, что для того, чтобы доля ущемляемого населения стала менее од-

ного процента, необходимо включение в состав официальных 19 языков для индекса  $\Delta^{(d,s)}$  и 14 языков для индекса  $\Delta^{(y,s)}$ .

**Схема 1.** Последовательность для индекса ущемления  $\Delta^{(d,s)}$ . Все население.

1	2	3	4	5	6	7	8	9	10a	10b	10c
<b>EN</b>	<b>1 + GE</b>	<b>2 + FR</b>	<b>3 + IT</b>	<b>4 + SP</b>	<b>5 + PL</b>	<b>6 + RO</b>	<b>7 + HU</b>	<b>8 + PT</b>	<b>9 + CZ</b>	<b>9 + GR</b>	<b>9 + RU</b>
62,6	49,3	37,8	29,5	22,4	16,4	12,9	10,9	9,2	7,7	7,7	7,7
11	12	13	14a	14b	15	16a	16b	17	18a	18b	19
<b>10a+ GR</b>	<b>11 + BG</b>	<b>12 + NL</b>	<b>13 + FI</b>	<b>13 + SW</b>	<b>14a+ SW</b>	<b>15 + LT</b>	<b>15 + SK</b>	<b>16a+ SK</b>	<b>17 + LV</b>	<b>17 + DA</b>	<b>18a+ DA</b>
6,2	5,0	4,0	3,3	3,3	2,7	2,2	2,2	1,7	1,3	1,3	1,0

**Схема 2.** Последовательность для индекса ущемления  $\Delta^{(d,s)}$ . Молодежь до 30 лет.

1	2	3	4	5	6	7	8	9	10	11a	11b
<b>EN</b>	<b>1 + FR</b>	<b>2 + GE</b>	<b>3 + IT</b>	<b>4 + SP</b>	<b>5 + PL</b>	<b>6 + RO</b>	<b>7 + HU</b>	<b>8 + PT</b>	<b>9 + CZ</b>	<b>10 + GR</b>	<b>10 + BG</b>
44,6	34,5	25,8	19,9	14,4	10,4	7,8	6,3	5,1	3,9	3,1	3,1
12	13	14a	14b	14c	14d	14e	18				
<b>11a+ BG</b>	<b>12 + NL</b>	<b>13 + RU</b>	<b>13 + FI</b>	<b>13 + SL</b>	<b>13 + LT</b>	<b>13 + LV</b>	<b>13 + FI / SK / LT / LV</b>				
2,3	1,8	1,5	1,5	1,5	1,5	1,5	0,7				

**Схема 3.** Последовательность для индекса ущемления  $\Delta^{(y,s)}$ . Все население.

1	2	3	4	5	6	7	8a	8b	9
<b>EN</b>	<b>1 + FR</b>	<b>2 + PL</b>	<b>3 + GE</b>	<b>4 + IT</b>	<b>5 + HU</b>	<b>6 + SP</b>	<b>7 + GR</b>	<b>7 + RO</b>	<b>8a + RO</b>
43,1	24,0	16,6	11,4	9,0	6,9	5,2	4,0	4,1	2,9
10a	10b	10c	12	13a	13b	14			
<b>9 + CZ</b>	<b>9 + FI</b>	<b>9 + BG</b>	<b>10a + FI/BG</b>	<b>12 + SW</b>	<b>12 + PT</b>	<b>13a+ PT</b>			
2,1	2,1	2,2	1,3	1,1	1,1	0,9			

Наиболее существенные изменения наблюдаются, если учитывать матрицу расстояний между языками. Неожиданно польский язык (как наиболее распространенный из славянских) вошел в тройку главных языков Евросоюза, немецкий же, являющийся родным для самого большого числа жителей Европы, оказался лишь четвертым. Вообще при учете близости языков на более высокие позиции попадают языки, не имеющие аналогов: значимо повысили свой статус венгерский, греческий и финский. В то же время португальский язык, близкий к испанскому, передвинулся с 9-го места на 14-е, а близкий к немецкому голландский (7-й по числу говорящих на нем в Европе) и так находившийся на 13-м месте, вообще ушел из списка.

Интересно также посмотреть, что произойдет, если ориентироваться исключительно на родные языки, используя в качестве меры ущемления индекс  $\Delta^{(y,p)}$ . Последовательность будет выглядеть следующим образом: GE  $\rightarrow$  GE+IT  $\rightarrow$  GE+IT+PL. Видим, что в тройку не входят самые распространенные языки: английский оказывается на 4-й

позиции, а французский – лишь на 5-й. Следовательно, выбор характеристики ущемления может существенно изменить набор официальных языков.

Вхождение новых стран (особенно крупных по численности населения) также существенно меняет картину. Например, вхождение Украины в Евросоюз может сделать представителем славянских языков словацкий. Польский же откатывается при этом далеко назад. В связи с этим можно рассматривать модели стратегического поведения стран – будут ли они поддерживать расширение Евросоюза или противиться этому.

### Оптимальное число официальных языков

Вернемся к проблеме нахождения оптимального числа официальных языков. Графически продемонстрируем зависимость индексов ущемления от числа языков на рис.1. Графики даны для всех 3 рассмотренных ситуаций: «все население», «молодежь» и «все население, с учетом матрицы расстояний между языками».

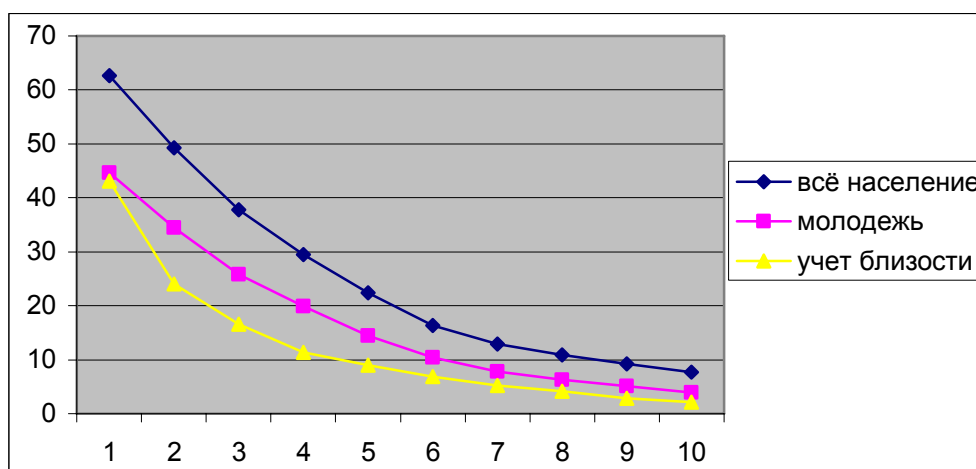


Рис.1. Зависимость индексов ущемления от числа официальных языков

Проблема осложняется тем, что Евросоюз состоит из отдельных государств. Единогласный выбор маловероятен, поэтому любое решение принимается большинством – иногда квалифицированным (необходима поддержка 14 из 27 государств, 258 из 345 голосов и 62% населения), иногда конституционным (число государств увеличивается до 15, кроме того, требуется поддержка 65% населения, число голосов роли не играет). Также может использоваться закон Пенроуза: для принятия решения требуется 62% голосов, притом, что количество голосов каждой страны пропорционально квадратному корню из ее населения.

Если считать, что страна поддерживает группу языков  $T$ , когда уровень ущемления для нее не превышает величины  $r$ , можно подсчитать минимальное количество языков, которое будет поддержано в соответствии с заданным правилом голосования: КвБ (квалифицированное большинство), КоБ (конституционное большинство) или ЗП (закон Пенроуза). Сведем данные, соответствующие модели для всего населения, для молодежи и с учетом матрицы расстояний между языками, в таблице 4.

Табл. 4. Минимальное число официальных языков в зависимости от уровня ущемления

$r$	Все население					Молодежь до 30 лет					С учетом близости языков				
	10%	20%	30%	40%	50%	10%	20%	30%	40%	50%	10%	20%	30%	40%	50%
КвБ	11	10	9	8	7	8	7	6	6	5	7	5	3	3	3
КоБ	11	10	9	6	4	7	5	4	4	3	7	4	3	2	2
ЗП	8	7	6	6	5	6	5	5	4	3	6	5	2	2	2

Видим, что в настоящее время при принятии решений квалифицированным большинством даже 6-язычный сценарий оказывается непроходным (ни при каком уровне  $r \leq 50\%$ !), если не учитывать близость языков. В то же время для молодежи (а следовательно, в будущем) он становится реальным, начиная с  $r = 30\%$ .

Также сильно облегчает проведение решений предположение о том, что если в числе официальных имеются «языки-соседи», то люди смогут их относительно легко изучить. Поскольку в данном случае уровень ущемления считается существенно меньшим, то проходным при  $r = 30\%$  может оказаться даже вариант всего с 3 языками.

#### СПИСОК ЛИТЕРАТУРЫ

1. **I. Dyen, J. Kruskal, P. Black** «An Indo-European Classification: a Lexicostatistical Experiment» // Transactions of the American Philosophical Society, 1992, №82 (5).
2. **S. Weber, V. Ginsburgh, I. Ortuno-Ortin** «Disenfranchisement in Linguistically Diverse Societies. The Case of the European Union» // Journal of European Economic Association, 2005, №3 (4).
3. **S. Weber, V. Ginsburgh** «Language Disenfranchisement in the European Union» // Journal of Common Market Studies, 2005, №43 (2).

#### ON THE PROBLEM OF CHOOSING THE SET OF OFFICIAL LANGUAGES OF THE EUROPEAN UNION

A.Yu. Filatov

*Institute of energy systems named after L.A. Melent'ev, Irkutsk*  
*e-mail: fial@irlan.ru*

**Abstract.** The paper considers one of the heterogeneity aspects – the problem of cross-country communications. We investigate some ways to the solving the problem of finding the optimal number and set of official languages of the European Union. It's shown that using different disenfranchisement indices leads to the essentially different order of inclusion languages into the official languages set.

**Key words:** game theory, the linguistic models, the Nash equilibrium, official languages of the European Union.