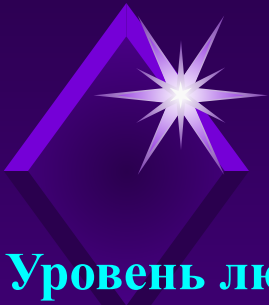


# Лингвистика глазами экономики и теории игр

**Филатов А.Ю.**

Институт систем энергетики им.Л.А.Мелентьева,  
Иркутский государственный университет

<http://math.isu.ru/filatov>,  
<http://polnolunie.baikal.ru/me>,  
[http://fial\\_.livejournal.com](http://fial_.livejournal.com),  
[alexander.filatov@gmail.com](mailto:alexander.filatov@gmail.com)



# Неоднородность

**Уровень людей:** индивидуальные ценности и интересы, доходы и потребности

**Уровень стран:** история, идеология, культура, доминирующая религия, титульная нация, экономика, **язык...**

## Альтруистическое поведение распространено в однородных группах

США: бедные – другие (перераспределение меньше);

Западная Европа: бедные – несчастные (перераспределение больше).

**Разнообразие – хорошо** (культурный и технический прогресс, экономич.рост).

**Поляризация – плохо** (разобщенность, дестабилизация, войны...).

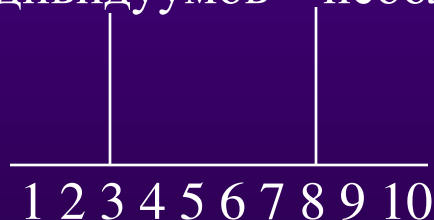
(2,2,2) и (3,0,3). Первая – более неоднородна, вторая – более поляризована.

**Поляризация общества наблюдается**, если имеют место следующие свойства:

1. Высокая степень однородности внутри каждой группы.
2. Высокая степень неоднородности между группами.
3. Число групп индивидуумов – небольшое, а численность групп – большая.



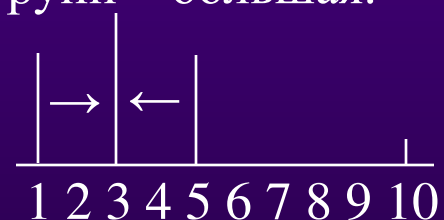
**разнообразие**



**поляризация**



**поляризация ↑**



**поляризация ↓**



# Общественный антагонизм

$N$  – население, разделенное на  $k$  групп по  $N_i$  человек. Доля  $s_i = \frac{N_i}{N}$ ,  $\sum_{i=1}^n s_i = 1$ .  
 $T = \{\tau_{ij}\} \in [0; 1]$  – матрица расстояний,  $\tau_{ii} = 0$ ,  $\tau_{ij} = \tau_{ji}$ ,  $\tau_{ij} \leq \tau_{ik} + \tau_{kj}$ .

## Варианты формул для расстояния:

$T^d = \{\tau_{ij}\}$ ,  $\tau_{ij} = 1, i \neq j$  – дихотомическое расстояние.

$T^c = \{\tau_{ij}\}$ ,  $\tau_{ij} = 0, i \neq c, j \neq c$  – важно расстояние между центром и периферией.

Возможный вариант – **расстояние между языками**.

**Отчуждение (alienation)** представителя  $i$ -группы к представителю  $j$ -группы, возрастает при  $\tau_{ij} \uparrow$ , часто просто равняется расстоянию.

**Вовлеченность (identification)** – увеличивается при росте размера группы, часто  $s_j^\alpha, \alpha > 0$ , иногда допускается  $\alpha = 0$ .

**Индивидуальный антагонизм** к представителю  $j$ -группы =  $s_i^\alpha \tau_{ij}$ .

**Индивидуальный антагонизм** ко всей  $j$ -группе =  $s_i^\alpha s_j \tau_{ij}$ .

**Антагонизм**  $i$ -группы к  $j$ -группе =  $s_i^{1+\alpha} s_j \tau_{ij}$ .

**Общественный антагонизм** (Esteban, Ray' 1994):  $A(\alpha, \tau) = \sum_{i=1}^n \sum_{j=1}^n s_i^{1+\alpha} s_j \tau_{ij}$ .



# Измерение неоднородности

**ELF (этнолингвистическая фрагментация; Shannon' 1949)** – индекс разнообразия; вероятность двух индивидов относиться к разным группам; 0 – однородное общество, 1 – максимально разобщенное

$$ELF = A(0, T^d) = 1 - \sum_{i=1}^k s_i^2. \text{ При росте числа групп } k \text{ ELF увеличивается.}$$

При фиксированном  $k$   $ELF \rightarrow \max$  при  $s_1 = s_2 = \dots = s_n$ .

**RQ (индекс поляризации; Reynal-Querol' 2002)**

$$RQ = A(1, T^d) = 1 - \sum_{i=1}^k s_i^2 (1 - s_i).$$

**GI (Greenberg' 1956)** – ожидаемое расстояние между двумя случайными представителями разных групп

$$GI = A(0, T) = \sum_{i=1}^k \sum_{j=1}^k s_i s_j \tau_{ij}.$$

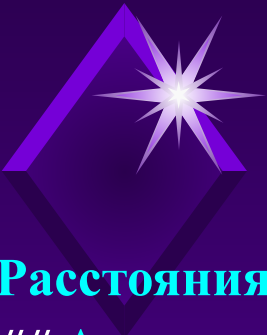
**ER (Esteban, Ray' 1994)**

$$ER = A(1, T) = \sum_{i=1}^k \sum_{j=1}^k s_i^2 s_j \tau_{ij}. \text{ Вариация: } ER' = A(\alpha, T) = \sum_{i=1}^k \sum_{j=1}^k s_i^{1+\alpha} s_j \tau_{ij}, \alpha \in [1; 1,6].$$

**PH (периферийная гетерогенность; Desmet, Ortuno-Ortin, Weber' 2009)**

$$PH = A(0, T^c) = 2 \sum_{i=1}^k s_i s_c \tau_{ic}. \text{ } C \text{ – максимально центральная группа.}$$

Отчуждение – только между центром и периферией.



# Расстояние между языками

**Расстояния необходимо принимать во внимание!**

**## Андорра:** каталонский (50%) + испанский (50%) – похожие романские языки

**Бельгия:** голландский (60%, герм. группа) + французский (40%, роман. группа)

Бельгия более неоднородна, чем Андорра; индексы ELF, RQ – наоборот!

**Becker' 1957** – важно не только свой / чужой, но и насколько чужой!

**Попутно решается проблема разграничения одинаковых и различных групп**

**## Мексика** – 291 язык, Венецианский  $\approx$  итальянский...

## Методы построения матрицы расстояний

**1. Fearon' 2003** – метод, основанный на лингвистических деревьях

$\tau_{ij} = 1 - (l/m)^\delta$ , где  $l$  – число общих ветвей,  $m$  – общее число ветвей.

**2. Dyen' 1992** – метод, основанный на доле общих корней

**95 индоевропейских языков, 200 основных понятий (списки Сводеша)**

$n_{ij}^0$  – число несовпадений корней,

$n_{ij}^1$  – число совпадений корней.

$\tau_{ij} = \frac{n_{ij}^0}{n_{ij}^0 + n_{ij}^1}$  – расстояние.

$\tau_{ij} = 0$  – абсолютно совпадающие языки,  $\tau_{ij} = 1$  – абсолютно различные языки.



# Эмпирические исследования

**Разнообразие:**  $r(ELF, GI) = 0,69$ , однако

1. Латинская Америка более диверсифицирована с учетом расстояния (есть индоевропейские языки + языки индейских племен).  
## Боливия (+54 места), Мексика (+69), Эквадор (+68).
2. Африка – более однородна (много похожих языков).  
## Мозамбик (-163), Замбия (-146), Ангола (-128), Того (-126).
3. Европа – по-разному.  
## Эстония (+71 – русские), Болгария (+65 – турки), Россия (+57),  
Андорра (-102 – испанцы и каталонцы).

**Поляризация vs разнообразие:**  $r(ER, GI) = 0,87$ , однако

## Папуа - Новая Гвинея – 829 мелких языковых групп,  
 $GI(4)$  (диверсифицирована)  $\rightarrow ER(186)$  (не поляризована).

**Периферия и центр:**  $r(PH, GI) = 0,93$ .

Индекс важен при недостатке данных (нужны расстояния только до центра).

**Резюме:**  $GI \sim ER \sim PH > ELF \sim RQ$ .

Неоднородность значима в моделях, учитывающих расстояние.

Рост неоднородности на  $\sigma \Rightarrow$  сокращение перераспределения на 10%.



# Евросоюз и задача выбора официальных языков

## не зона Евро

	Шенген	не Шенген
ЕС	Швеция	Болгария
не ЕС	Норвегия	Албания

## зона Евро

	Шенген	не Шенген
ЕС	Франция	Кипр
не ЕС	Сан-Марино	Черногория

### До 2004 г. – 15 государств

Австрия, Бельгия, Великобритания, Германия, Греция, Дания, Ирландия, Испания, Италия, Люксембург, Нидерланды, Португалия, Финляндия, Франция, Швеция.

### 2004 г. – 25 государств (+10)

Венгрия, Кипр, Латвия, Литва, Мальта, Польша, Словакия, Словения, Чехия, Эстония.

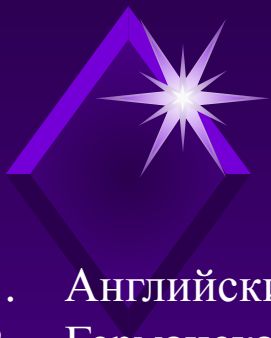
### 2007 г. – 27 государств (+2), 23 государственных языка

Болгария, Румыния.

### Дилемма:

**Малое число официальных языков** – ущемление.

**Большое число официальных языков** – высокие затраты  
(переводчики, задержки, ошибки и разночтения,...)



## 8 групп, представленных в ЕС

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

**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
Голландский	21,9	24,0	1,10
Румынский	21,0	22,2	1,06
Русский	4,2	22,4	5,33





# Расстояние между языками (Duen)

Табл. 2. Матрица расстояний между языками

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



# Индексы ущемления

$n(i)$  – родной язык  $i$ -человека,

$L(i)$  – множество языков, на которых он говорит,

$T$  – множество официальных языков Евросоюза.

$\delta_i^{(d,p)} = \begin{cases} 0, & n(i) \in T, \\ 1, & n(i) \notin T. \end{cases}$  – равно 0, если родной язык человека является официальным языком Евросоюза, иначе 1.

$\delta_i^{(d,s)} = \begin{cases} 0, & L(i) \cap T \neq \emptyset, \\ 1, & L(i) \cap T = \emptyset. \end{cases}$  – равно 0, если человек говорит на одном из официальных языков Евросоюза, иначе 1.

$\delta_i^{(\tau,p)} = \min_{t \in T} \tau(n(i), t)$  – равно минимальному расстоянию между родным языком и одним из официальных.

$\delta_i^{(\tau,s)} = \min_{l \in L(i), t \in T} \tau(l, t)$  – равно минимальному расстоянию между одним из языков, на котором говорит человек, и одним из официальных.

$\Delta^*(T)$  – один из индексов ущемления (среднее по всем жителям значение).

**Двухкритериальная задача:**  $\Delta^*(T) \rightarrow \min, |T| \rightarrow \min.$

# Наилучшие наборы офиц. языков

Табл. 3. Доля граждан ЕС, не говорящих на соотв. языках, %

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

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

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10a</b>	<b>10b</b>	<b>10c</b>
<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
<b>11</b>	<b>12</b>	<b>13</b>	<b>14a</b>	<b>14b</b>	<b>15</b>	<b>16a</b>	<b>16b</b>	<b>17</b>	<b>18a</b>	<b>18b</b>	<b>19</b>
<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



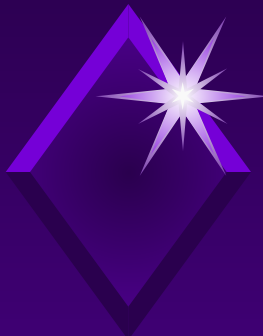
# Наилучшие наборы офиц. языков

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
EN	1 + FR	2 + GE	3 + IT	4 + SP	5 + PL	6 + RO	7 + HU	8 + PT	9 + CZ	10 + GR	10 + BG
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				
11a+ BG	12 + NL	13 + RU	13 + FI	13 + SL	13 + LT	13 + LV	13 + FI / SK / LT / LV				
2,3	1,8	1,5	1,5	1,5	1,5	1,5	0,7 $\Delta^{(\tau,s)}$				

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

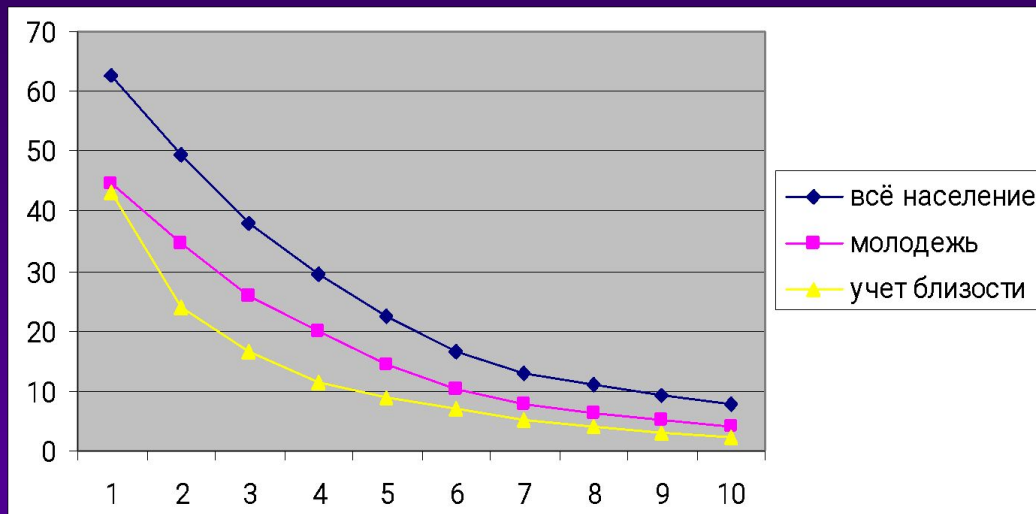
1	2	3	4	5	6	7	8a	8b	9
EN	1 + FR	2 + PL	3 + GE	4 + IT	5 + HU	6 + SP	7 + GR	7 + RO	8a + RO
43,1	24,0	16,6	11,4	9,0	6,9	5,2	4,0	4,0	2,9
10a	10b	10c	12	13a	13b	14			
9 + CZ	9 + FI	9 + BG	10a + FI/BG	12 + SW	12 + PT	13a+ PT			



# Оптимальное число офиц. языков

$\Delta^{(\tau, p)} : GE \rightarrow GE+IT \rightarrow GE+IT+PL.$

Зависимость индексов ущемления от числа официальных языков



**Квалифицирован. большинство:** 14 из 27 гос-в, 258 из 345 голос., 62% насел.

**Конституционное большинство:** 15 государств, 65% населения.

**Закон Пенроуза:** 62% голосов, пропорциональных корню из населения страны.

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

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



# Теоретико-игровые постановки лингвистических моделей

## **R. Selten, J.Pool, 1991:**

1. На каких языках говорят люди?
2. Каковы выгоды от изучения языка?
3. Каковы издержки изучения языка?

## **J. Church, I. King, 1993:**

2 языка, одинаковые способности, нет разницы родной/приобретенный

## **J. Gabszevicz, V. Ginsburgh, S. Weber, 2008:**

Различные способности к языкам

## **V. Ginsburgh, A. Savvateev, S. Weber, 2008:**

3 языка, в т.ч. вариант «эсперанто»

## **V. Ginsburgh, I. Ortuno-Ortin, S. Weber, 2007:**

Есть разница родной/приобретенный, полезность зависит от расстояния между языками, эмпирическое исследование



# Модель Черча-Кинга, 1993

$e_0$  – родной язык английский,  $\hat{e}$  – число выучившихся англичан,  
 $f_0$  – родной язык французский,  $\hat{f}$  – число выучившихся французов,  
 $e_0 + f_0 = N$ ,  $e_0 > f_0$ .

**Условия изучения иностранного языка:**  $E: u(N) - u(e_0 + \hat{f}) > c$ ,  
 $F: u(N) - u(f_0 + \hat{e}) > c$ .

## Равновесные решения

$$\begin{aligned} u(N) - u(e_0) < u(N) - u(f_0) < c &\Rightarrow (\hat{f} = 0, \hat{e} = 0) \\ u(N) - u(e_0) < c < u(N) - u(f_0) &\Rightarrow (\hat{f} = f_0, \hat{e} = 0) \\ c < u(N) - u(e_0) < u(N) - u(f_0) &\Rightarrow (\hat{f} = f_0, \hat{e} = 0), (\hat{f} = 0, \hat{e} = e_0) \end{aligned}$$

## Общественное благосостояние

$$W_{NL} = e_0 u(e_0) + f_0 u(f_0), \quad W_{EF} = Nu(N) - e_0 c,$$

$$W_B = Nu(N) - Nc, \quad W_{FE} = Nu(N) - f_0 c.$$

$$W_B < W_{EF} < W_{FE}, \quad W_{NL} < W_{FE} \text{ при } c < c^*, \quad c^* = (u(N) - u(e_0))e_0 / f_0 + (u(N) - u(f_0)).$$

# Эффективные решения

Табл. 5. Эффективные и равновесные решения

интервал 1	интервал 2	интервал 3	интервал 4
$c < u(N) - u(e_0)$	$c \in [u(N) - u(e_0); u(N) - u(f_0)]$	$c \in (u(N) - u(f_0); c^*)$	$c \geq c^*$
FE эффективно FE / EF равновесны	FE эффективно FE равновесно	FE эффективно NL равновесно	NL эффективно NL равновесно

Государство в состоянии регулировать долю изучающих язык:

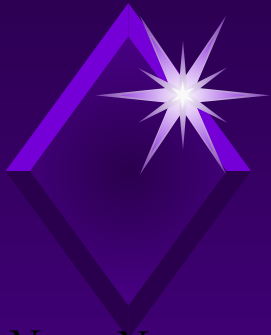
$$u''(x) < 0, \quad \rho \equiv -u''(x)x/u'(x) < 1 \quad \Rightarrow$$

$$c_1 = u(N) - u(f_0) + e_0 u'(N), \quad c_2 = u(N) - u(f_0) + e_0 u'(e_0).$$

Табл. 6. Эффективные и равновесные решения

	интервал 1	интервал 2	интервал 3	интервал 4	интервал 5
границы	$c < u(N) - u(e_0)$	$\left[ \begin{array}{l} u(N) - u(e_0); \\ u(N) - u(f_0) \end{array} \right]$	$(u(N) - u(f_0); c_1)$	$[c_1; c_2]$	$c > c_2$
эффективно	$\hat{f} = f_0$	$\hat{f} = f_0$	$\hat{f} = f_0$	$\hat{f} \in [0; f_0]$	NL
равновесно	$\hat{f} = f_0 / \hat{e} = e_0$	$\hat{f} = f_0$	NL	NL	NL





# Модель

## Габжевича-Гинзбурга-Вебера, 2008

$N_i$  и  $N_j$  – число жителей 2 стран,

$\theta$  – «неспособность к языкам», равномерно распределенная на  $[0;1]$ ,

Язык учит доля жителей страны, равная  $\alpha_i$ :  $\theta_i \in [0; \alpha_i] \Rightarrow \alpha_i = \theta_i$ .

**Полезность:**  $u(x) = x$ , **издержки:**  $c_i\theta_i = c_i\alpha_i$ .

$$\begin{aligned} (N_i + N_j) - (N_i + \alpha_j N_j) &= c_i \alpha_i, & \alpha_i &= (1 - \alpha_j) N_j / c_i, \\ & & \alpha_j &= (1 - \alpha_i) N_i / c_j. \end{aligned}$$

$b_i^j = N_j / c_i$  – выгоды от увеличения общения, нормированные на издержки.

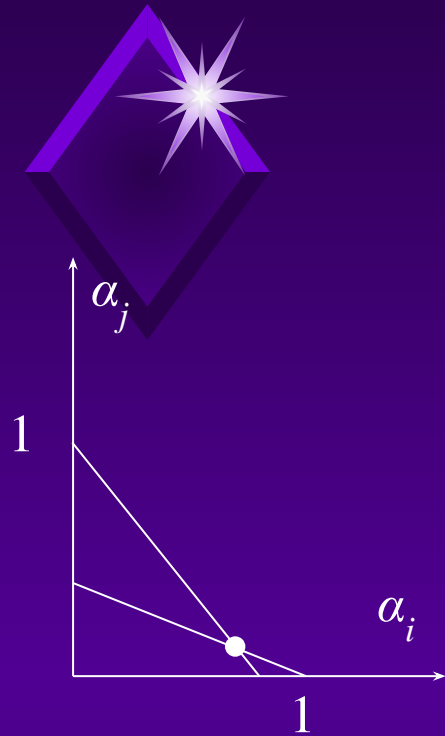
**Кривые реакции:**

$$\alpha_i = \min \left\{ (1 - \alpha_j) b_i^j; 1 \right\}, \quad \alpha_j = \min \left\{ (1 - \alpha_i) b_j^i; 1 \right\}.$$

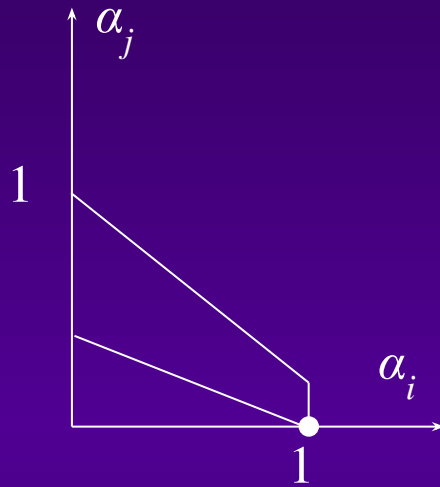
**Равновесное решение:**

$$\alpha_i^* = \min \left\{ \frac{b_i^j (1 - b_j^i)}{1 - b_i^j b_j^i}; 1 \right\}, \quad \alpha_j^* = \min \left\{ \frac{b_j^i (1 - b_i^j)}{1 - b_i^j b_j^i}; 1 \right\}.$$

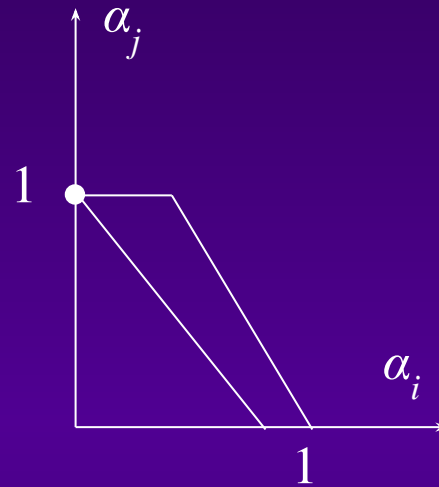
# Равновесные решения



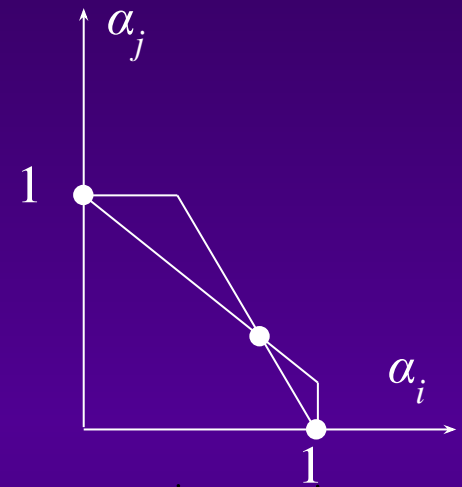
$b_i^j < 1, b_j^i < 1$   
устойчивое  
внутреннее  
равновесие



$b_i^j \geq 1, b_j^i < 1$   
угловое  
равновесие



$b_i^j < 1, b_j^i \geq 1$   
угловое  
равновесие

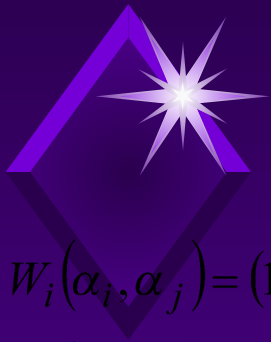


$b_i^j > 1, b_j^i > 1$   
2 угловых +  
+ неустойчивое  
внутр. равновесие

$$\alpha_i^* - \alpha_j^* = \frac{b_i^j - b_j^i}{1 - b_i^j b_j^i}$$

В стране, где выше выгоды, нормированные на издержки, выше и доля изучающих язык.

**Примеры:** Канада, Бельгия.



# Эффективные решения

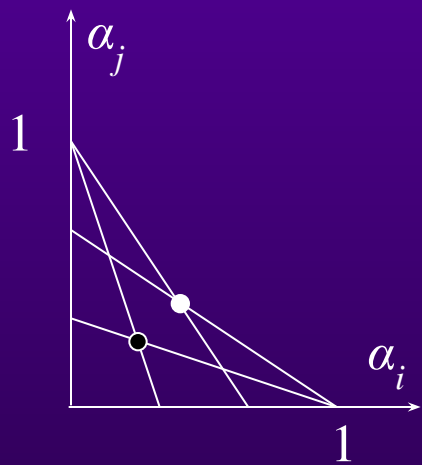
$$W_i(\alpha_i, \alpha_j) = (1 - \alpha_i)N_i(N_i + \alpha_j N_j) + \alpha_i N_i(N_i + N_j) - c_i \int_{\theta=0}^{\alpha_i N_i} \frac{\theta}{N_i} d\theta,$$

$$W_i(\alpha_i, \alpha_j) = N_i^2 + N_i N_j (\alpha_i - \alpha_i \alpha_j + \alpha_j) - \frac{1}{2} c_i \alpha_i^2 N_i,$$

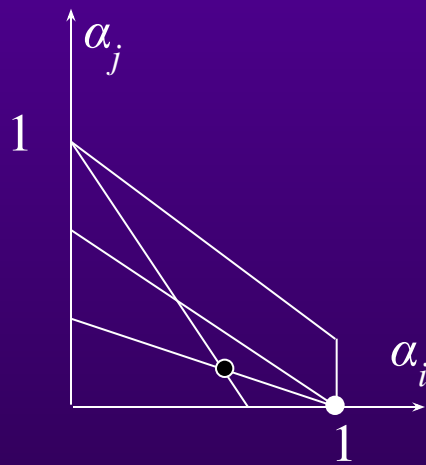
$$W(\alpha_i, \alpha_j) = W_i(\alpha_i, \alpha_j) + W_j(\alpha_i, \alpha_j) =$$

$$= N_i^2 + N_j^2 + 2N_i N_j (\alpha_i - \alpha_i \alpha_j + \alpha_j) - \frac{1}{2} c_i \alpha_i^2 N_i - \frac{1}{2} c_j \alpha_j^2 N_j.$$

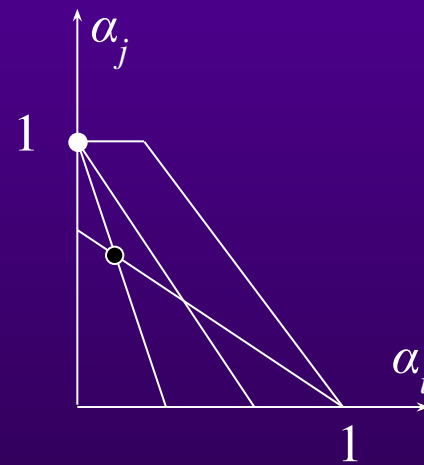
$$\alpha_i = \min \{ 2b_i^j (1 - \alpha_j); 1 \}, \quad \alpha_j = \min \{ 2b_j^i (1 - \alpha_i); 1 \}.$$



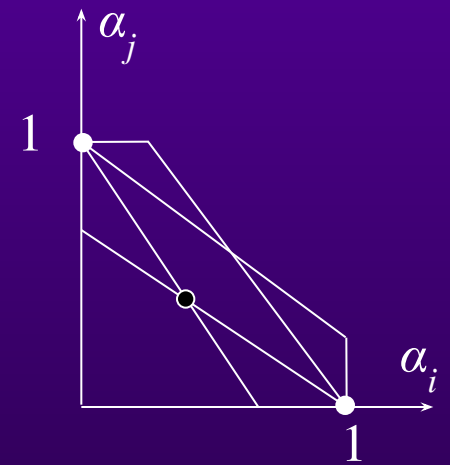
$$b_i^j < 0,5, \quad b_j^i < 0,5$$



$$b_i^j \geq 0,5, \quad b_j^i < 0,5$$



$$b_i^j < 0,5, \quad b_j^i \geq 0,5$$



$$b_i^j > 0,5, \quad b_j^i > 0,5$$



# Модель Гинзбурга, Савватеева, Вебера, 2008

Можно не ограничиваться 2 языками.

- 1)  $l_0$  и  $p_0$  – население 2 стран, в сумме составляющее  $N$  человек.
- 2) Сверхвысокие издержки изучения языков друг друга.
- 3) Существует третий язык  $R$  (эсперанто), издержки изучения которого равны  $c$ .
- 4)  $u(n)$  – функция полезности, зависящая от числа людей, с которыми возможен диалог.

## Равновесные решения

$\hat{l} = 0, \hat{p} = 0$  – равновесие, выполняющееся при любых  $c > 0$

$\hat{l} = l_0, \hat{p} = p_0$  – равновесие, выполняющееся при

$$\begin{cases} u(N) - c \geq u(l_0) \\ u(N) - c \geq u(p_0) \end{cases} \Leftrightarrow c \leq u(N) - \max\{u(l_0), u(p_0)\}$$

При  $u(n) = n$   $c \leq \min\{l_0, p_0\}$ .

$\begin{cases} u(l_0 + \hat{p}) - c = u(l_0) \\ u(p_0 + \hat{l}) - c = u(p_0) \end{cases}$  – неустойчивое внутреннее равновесие



# Эффективные решения

$$Nu(N) - cN \geq l_0 u(l_0) + p_0 u(p_0), \quad c \leq u(N) - \frac{l_0}{N} u(l_0) - \frac{p_0}{N} u(p_0).$$

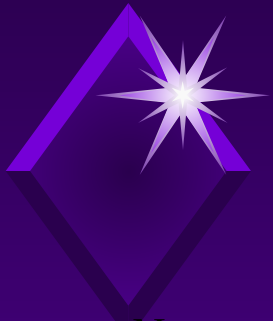
При  $u(n) = n$  равновесие  $(l_0, p_0)$  эффективно при  $c \leq 2 \frac{l_0 p_0}{N}$

Равновесие (0; 0)



## Возможные равновесия

	∅	F	R	F&R
∅	+	+	-	-
F	+	-	-	-
R	-	-	+	-
F&R	-	-	-	-



# Модель Гинзбурга, Ортуно-Ортина, Вебера, 2007

$N_i$  и  $N_j$  – число жителей 2 стран,

$N_{ij}$  – число жителей  $i$ -страны, выучивших иностранный,

**Полезность:**  $u(N_i, N_j, L_{ij})$  или  $u(N_i, N_{ji}, L_{ij})$ , **издержки:**  $c(L_{ij})$ .

**Функция спроса на языки:**

$D_i(N_i, N_j, L_{ij}) = \ln N_{ij} / N_i$  – доля жителей  $i$ -страны, изучающих  $j$ -язык.

**Утверждение 1:**  $D_i(N_i, N_j, L_{ij})$  возрастает по  $N_j$ .

**Утверждение 2:**  $D_i(N_i, N_j, L_{ij})$  убывает по  $N_i$ .

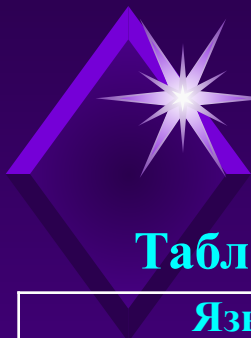
**Утверждение 3:**  $D_i(N_i, N_j, L_{ij})$  убывает по  $L_{ij}$ .



# Эмпирическое исследование

Табл. 7. Доля населения стран ЕС, говорящих на наиболее распространенных языках

Страна	Родн. язык млн.чел.	Доли знающих язык, %			
		Английский	Немецкий	Французский	Испанский
Австрия (GE)	100	46	100	11	1
Великобрит. (EN)	341	100	22	9	5
Германия (GE)	100	54	100	16	2
Греция (GR)	12	47	12	12	5
Дания (DA)	5	75	37	5	1
Ирландия (EN)	341	100	6	23	2
Испания (SP)	340	36	2	19	100
Италия (IT)	62	39	4	29	3
Нидерланды (NL)	20	70	59	19	1
Португалия (PT)	176	35	2	28	4
Финляндия (FI)	6	61	7	1	1
Франция (FR)	77	42	8	100	15
Швеция (SW)	9	79	31	7	4



# Матрица расстояний между языками

**Табл. 8.** Число говорящих на наиболее распространенных языках ЕС

Язык	Родной, ЕС, млн.чел.	Родной, млн.чел.	Говорят, млн.чел.
Английский	62,4	341	1800
Немецкий	85,3	100	126
Французский	60,7	77	169
Испанский	39,7	340	450

**Табл. 9.** Матрица расстояний между языками (\*1000)

	Английский	Немецкий	Французский	Испанский
Английский	0	422	764	760
Голландский	392	162	756	742
Греческий	838	812	843	833
Датский	407	293	759	750
Испанский	760	747	266	0
Итальянский	753	735	197	212
Немецкий	422	0	764	747
Португальский	760	753	291	126
Финский	1000	1000	1000	1000
Французский	764	756	0	266
Шведский	411	305	756	747



# Регрессионные модели

$$\ln \frac{\hat{N}_{iE}}{N_i} = 0,733^* - 0,153^* \ln N_i - 0,408^* \ln L_{ij},$$

(0,016)            (0,021)            (0,082)

$$\hat{R}^2 = 0,919.$$

$$\ln \frac{\hat{N}_{iG}}{N_i} = 0,586^* - 0,361^* \ln N_i - 1,362^* \ln L_{ij},$$

(0,077)            (0,072)            (0,214)

$$\hat{R}^2 = 0,910.$$

$$\ln \frac{\hat{N}_{iF}}{N_i} = 0,193^* + 0,355^* \ln N_i - 0,512 \ln L_{ij},$$

(0,121)            (0,138)            (0,416)

$$\hat{R}^2 = 0,599.$$

$$\ln \frac{\hat{N}_{iS}}{N_i} = 0,091 + 0,032 \ln N_i - 0,560 \ln L_{ij},$$

(0,109)            (0,168)            (0,385)

$$\hat{R}^2 = 0,232.$$

$$\ln \frac{\hat{N}_{ij}}{N_i} = 0,080 - 0,233^* d_G - 0,112 d_F - 0,514^* d_S -$$

(0,100)            (0,061)            (0,062)            (0,050)

$$- 0,058 \ln N_i + 0,625^* \ln N_j - 0,954^* \ln L_{ij},$$

(0,069)            (0,057)            (0,200)

$$\hat{R}^2 = 0,758.$$



# Регрессионные модели

$$\ln \frac{\hat{N}_{ij}}{N_i} = 0,070 - 0,340 * d_s - 0,055 \ln N_i +$$
$$+ 0,600 * \ln N_j - 0,789 * \ln L_{ij} + 0,249 \ln Tr_{ij}, \quad \hat{R}^2 = 0,712.$$

(0,096)      (0,044)      (0,070)      (0,067)      (0,205)      (0,134)

## Выводы по эмпирическому исследованию

1. Высокая доля прогнозируемой вариации (76%).
2. В больших странах меньше учат иностранные языки.
3. Люди чаще учат более распространенные языки.
4. Люди не склонны изучать языки, очень непохожие на их собственный.
5. Огромная разница между отличными результатами для английского и немецкого, посредственными для французского и плохими для испанского.
6. Торговые связи частично объясняют различия.
7. Особое место Испании из-за того, что большинство носителей языка в Латинской Америке, а не в ЕС, а также из-за длительной изоляции.



# Возможные направления применения подобных моделей

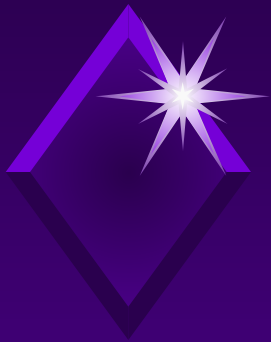
## Удобство, надежность, цена / совместимость

1. Операционные системы (**Windows** vs Linux).
2. Другие программные продукты (Word vs Tex).
3. Стандарты и форматы (**mp3** vs ogg...).
4. Социальные сети (Odnoklassniki.ru vs Vkontakte.ru...).
5. Блоги (Livejournal.com vs Liveinternet.ru...).
6. Валюты, в т.ч. в резервах.
7. Тарифы, в т.ч. сотовых операторов.

.....

### Важные вопросы:

1. Критическая доля пользователей, обеспечивающая переход к фактической монополии.
2. Минимальное количество пользователей, гарантирующее «невывымирание».



*Спасибо  
за внимание!*

<http://math.isu.ru/filatov>,  
<http://polnolunie.baikal.ru/me>,  
[http://fial\\_.livejournal.com](http://fial_.livejournal.com),  
[alexander.filatov@gmail.com](mailto:alexander.filatov@gmail.com)