

DIMENSION OF EMOTION SPACE: DIACHRONIC APPROACH

Solovyev V.D., Bochkarev V.V., Shevlyakova A.V., Bayrasheva V.R.

РАЗМЕРНОСТЬ ПРОСТРАНСТВА ЭМОЦИЙ: ДИАХРОНИЧЕСКИЙ ПОДХОД

Соловьев В.Д.¹, Бочкарев В.В.², Шевлякова А.В.³, Байрашева В.Р.⁴

Abstract.

Абстракт. Многочисленные исследования по структуре, в том числе, размерности пространства эмоций не привели к согласованным общепризнанным результатам. В данной работе предложен новый подход определения размерности пространства эмоций, основанный на многомерном шкалировании эмотивной лексики из корпуса Google Books Ngram с метрикой Кульбака-Лейблера. Впервые рассмотрен вопрос о изменении размерности пространства со временем.

Keywords: emotions, corpus approach, multidimensional scaling

Ключевые слова: эмоции, корпусный подход, многомерное шкалирование

ВВЕДЕНИЕ

Эмоции изучались с позиций различных подходов, включая психолингвистические исследования множества слов, обозначающих эмоции. Для пространства эмотивных слов естественно поставить вопрос о его структуре. В частности, вопрос о размерности этого пространства изучался во многих публикациях [1-4]. Одна из первых работ, в которой предложена четкая методология исследований размерности лингвистического пространства эмоций, была статья Рассела [1]. Он предложил несколько независимых техник. Одна из них состояла в анализе методом многомерного шкалирования матрицы степеней сходства (по мнению 36 участников эксперимента) отобранных 28 эмотивных прилагательных английского языка. Результатом явилось расположение эмотивных слов на плоскости в форме круга. Показано, что первые два измерения включают 45,8% общей вариативности, в то время как следующие 3 измерения – лишь 13,1%. Поэтому пространство эмоций уместно трактовать как двухмерное.

Сходные результаты получены в ряде последующих исследований. На рис.1. представлен круг эмоций из работы [2]. Определенно выделяется одно измерение – ось X – степень приятности/неприятности эмоции и второе измерение – ось Y – активный/пассивный характер эмоции. На рис. 1 на окружности жирным выделены базовые эмоции.

Первые две шкалы выглядят совершенно естественно и, как правило, повторяются и в других работах. В[1] предложена интерпретация и следующих трех компонент. Причем 3-я и 4-я являются монополярными, а пятая, как и первые две, биполярной. Третья – *sleery, tired* (сонливость, усталость), четвертая – *angry* (гнев), пятая компонента – *alarmed-bored* (встревоженный-скучающий).

¹ Казанский федеральный университет, Казань, Россия, maki.solovyev@mail.ru

² Казанский федеральный университет, Казань, Россия, vbochkarev@mail.ru

³ Казанский федеральный университет, Казань, Россия, vbochkarev@mail.ru

⁴ Казанский федеральный университет, Казань, Россия, vbayrasheva@gmail.com

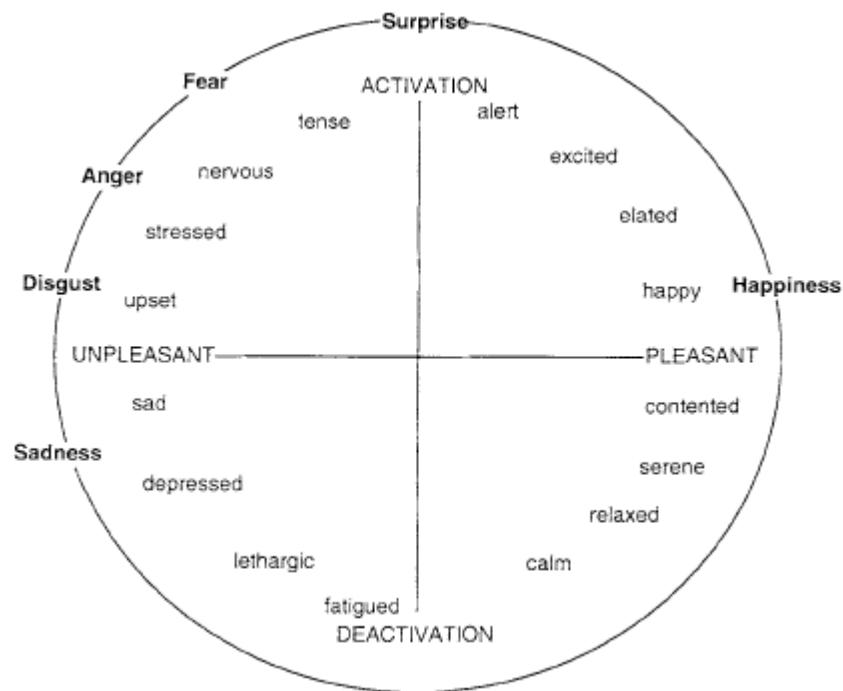


Рис. 1. 2-х мерная круговая модель структуры пространства эмоций [2].

Такая структура, подтвержденная во многих работах разными методами, представляется, в основном, верной, однако, возможно, является недостаточно точной. В недавней работе [8] выдвинута новая гипотеза о трехмерности пространства эмоций на основе психофизиологических соображений. Хотя данная гипотеза пока не получила серьезной экспериментальной поддержки, она возрождает вопрос о размерности пространства эмоций.

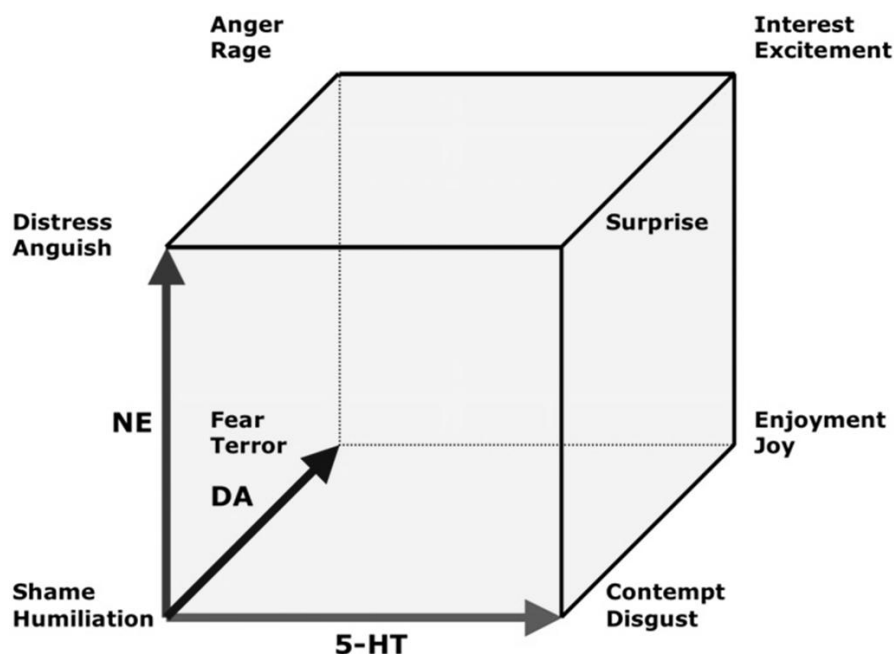


Рис.2. Трехмерная модель из работы [8].

В последние годы ранее полученные результаты были перепроверены на таких современных данных как твитты [5]. Большинство исследований описывают пространство эмоций в какой-то определенный момент времени. Диахронические аспекты изучались значительно меньше. В [6] установлено существование общей тенденции на уменьшение

употребления эмотивной лексики в основных европейских языках. В [7] изучалось изменение лексики, в том числе эмотивной, от эпохи к эпохе. В этих работах, однако, не ставился вопрос о размерности пространства эмоций.

ДАННЫЕ И МЕТОДЫ

В нашем исследовании мы привлекаем новые лингвистические данные – корпус Google Books Ngram (<https://books.google.com/ngrams>). Рассматриваются английский и русский языки. Мы используем набор из ____ слов из работы []. Для каждого слова строится вектор частот слов, располагающихся непосредственно слева от данного слова, и вектор частот слов, располагающихся непосредственно справа. Для каждых двух слов из нашего набора вычисляются левые и правые расстояния по Кульбаку-Лейбнеру между ними. Детали методологии изложены в [9]. К полученным матрицам расстояний далее применяются стандартные алгоритмы многомерного шкалирования.

Если в вышеупомянутых работах [1-4] рассматривались эмоции как таковые, то в нашем исследовании изучается пространство слов, обозначающих эмоции. Это определяет специфику исследования и позволяет выявить более тонкие, лингвистически кодируемые аспекты эмоций. В частности, способы их выражения через конструкции с различным местом расположения эмотивного слова в них. В данной работе мы ограничиваемся биграммами. X+Э будет обозначать конструкции, в которых эмотивное слово занимает вторую позицию, Э+X – где оно занимает первую позицию.

Кроме получения многомерного представления эмоций в современных текстах, прослеживается также динамика структуры пространства эмоций. Для этого все расчеты проводятся для всех 25-летних интервалов, начиная с 1800 г.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Приведем вклад первых 5 компонент в сравнении для английского и русского языков, а также для разного порядка слов. Все данные за 1975-2000 гг.

Таб. 1. Английский язык, X+Э

Номер компоненты	Вклад компоненты	Кумулятивный вклад
1	0,2829	0,2829
2	0,1775	0,4604
3	0,1539	0,6143
4	0,1064	0,7206
5	0,0635	0,7841

Таб. 2. Английский язык, Э+X

Номер компоненты	Вклад компоненты	Кумулятивный вклад
1	0,3800	0,3800
2	0,2935	0,6734
3	0,1557	0,8291
4	0,0458	0,8749
5	0,0442	0,9190

Таб. 3. Русский язык, X+Э

Номер компоненты	Вклад компоненты	Кумулятивный вклад
1	0,4403	0,4403
2	0,2079	0,6483

3	0,1798	0,8281
4	0,0825	0,9106
5	0,0207	0,9313

Таб. 4. Русский язык, Э+Х

Номер компоненты	Вклад компоненты	Кумулятивный вклад
1	0,5026	0,5026
2	0,2336	0,7362
3	0,0568	0,7930
4	0,0453	0,8382
5	0,0329	0,8712

Совершенно четких критериев выделения числа главных компонент, т.е. размерности пространства, не существует. В таблицах жирным шрифтом выделены возможные варианты выбора.

В конструкциях с эмотивным словом на первом месте первые несколько компонент выделяются более четко. Для английского языка уже первые две компоненты дают более 67% вклада, а первые три – даже более 82%, для русского языка первые две компоненты дают более 73% вклада. Как видно, в анализе, ориентированным на корпус, первые компоненты дают больший вклад, чем в вышеупомянутых психологических исследованиях.

В конструкциях с эмотивным словом на втором месте первые компоненты не выделяются четко. Очевидно, разница в позиции слова отражается на синтаксических связях слов в предложении.

При сравнении данных за разные периоды времени, мы получили следующие результаты.

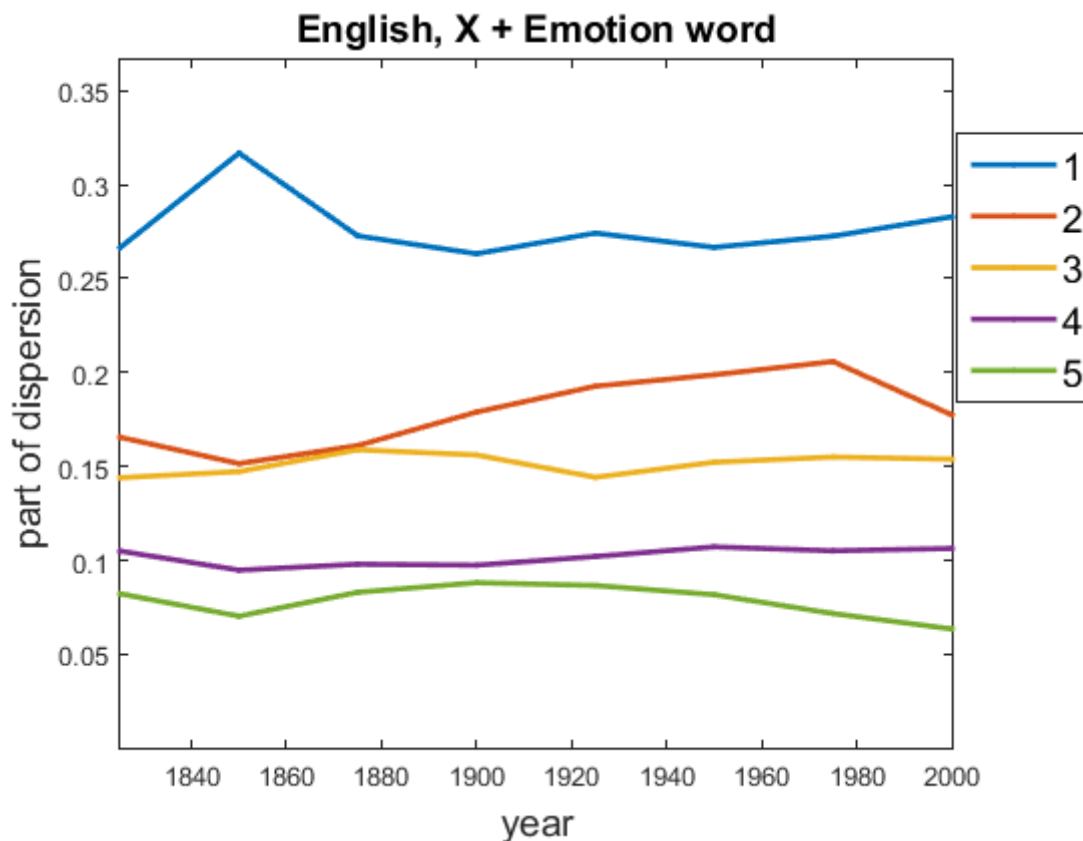


Рис. 3. Динамика вклада первых 5 компонент в английском, порядок слов Х+Э

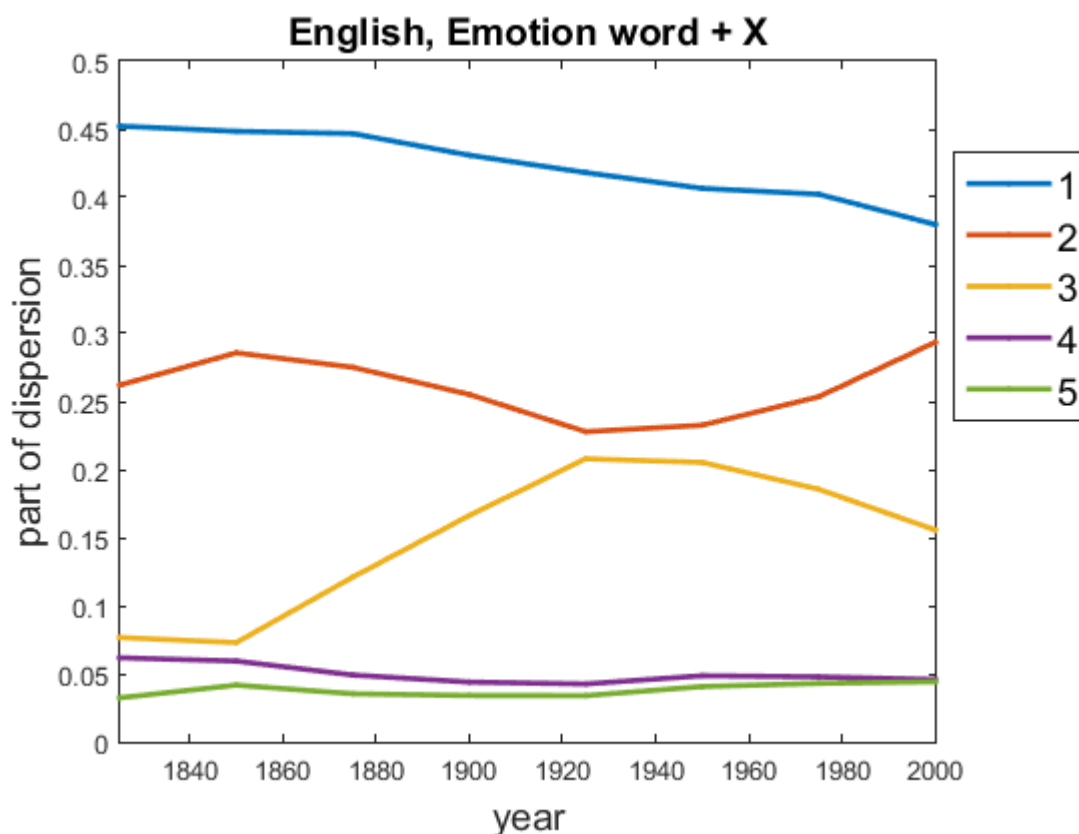


Рис. 4. Динамика вклада первых 5 компонент в английском, порядок слов Э+Х

Как видим, при второй позиции эмотивного слова вклад компонент с течением времени меняется слабо. При первой позиции эмотивного слова весьма определенно выделяются первые три компонента, вклад остальных невелик. Диахронические данные подтверждают, что для английского языка можно говорить о трехмерности пространства эмоций. Здесь наблюдаются любопытные диахронические закономерности. Во-первых, заметное падение вклада первой компоненты. Во-вторых, существенные колебания вклада второй и третьей компонент. Мы пока что не можем предложить объяснения этого эффекта. Падением вклада третьей компоненты за последние полвека можно интерпретировать, как тенденцию к понижению размерности пространства, сближение семантики эмотивных слов.

Это предположение коррелирует со следующими данными. На рис. 5 приведена динамика расстояний между словами с близкими значениями: *sadness* и *grief*. Как видим, в последние 70 лет расстояние уменьшается, т.е. контексты употребления этих слов сближаются. Метафорически, и очень предварительно говоря, пространство эмоций схлопывается.

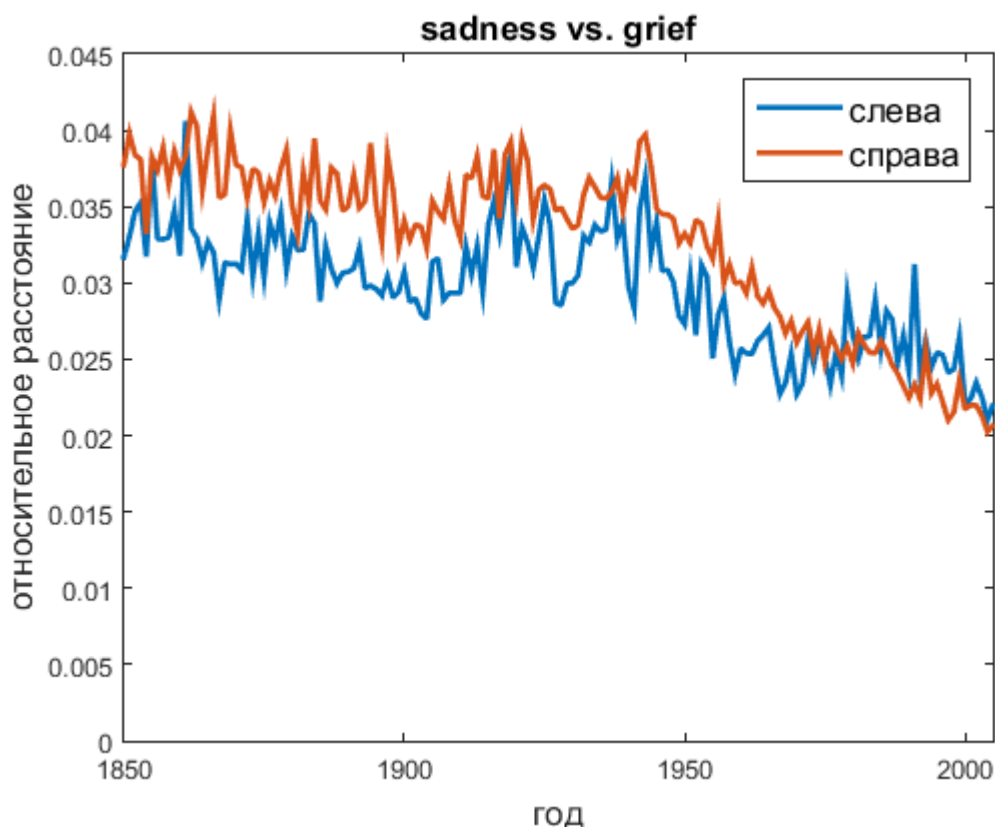


Рис. 5. Динамика расстояния между эмоциями sadness и grief

ЗАКЛЮЧЕНИЕ

Использование диахронического корпуса Google Books Ngram позволило по-новому взглянуть на проблему размерности пространства эмоций. Исследование проведено для английского и русского языков на биграмах вышеуказанного корпуса. Оказалось, что при применении корпусного подхода нужно учитывать позицию эмотивного слова в биграме. При второй позиции эмотивного слова в биграме размерность пространства не определяется четко, уменьшение вклада компонент происходит постепенно. Для первой позиции наиболее точная оценка размерности пространства - 2 для русского языка и 3 для английского языка.

Некоторые полученные данные можно интерпретировать так, что имеет место тенденция к уменьшению размерности пространства (для английского языка).

Работа выполнена при поддержке РФФИ, гранты № 15-29-01173, 15-06-07402 и Программы повышения конкурентоспособности Казанского федерального университета.

ЛИТЕРАТУРА

1. Russell J. A circumplex model of affect. *Journal of Personality and Social Psychology*. 1980. V. 39, N. 6, p. 1161-1178.
2. Feldman Barrett, L., & Russell. J. A. (1998). Independence and bipolarity in the structure of affect. *Journal of Personality and Social Psychology*, 74, 967-984.
3. Russell J., Barrett L. Core effect, prototypical emotional episodes, and other things called emotions: dissecting the elephant. *Journal of Personality and Social Psychology*. 1999. V. 76, N.5, p. 805-819.
4. Rubin D., Talarico J. A comparison of dimensional models of emotion: Evidence from emotions, prototypical events, autobiographical memories, and words, *Memory*. 2009. 17:8, p. 802-808.
5. Bann E., Bryson J. The Conceptualisation of Emotion Qualia: Semantic Clustering of Emotional Tweets. *Proceedings of the 13th Neural Computation and Psychology Workshop*, 2012.
6. Soloviev V., Bochkarev V., Shevlyakova A. Dynamics of emotive lexis use in European languages in the 19-20-th centuries. *The 1st International forum on cognitive modeling. Proceedings*. Southern Federal University Press. 2013, p. 222-225.
7. Popescu O., Strapparava C. Behind the Times: Detecting Epoch Changes using Large Corpora.

- International Joint Conference on Natural Language Processing. 2013. p. 347–355.
8. Lövheim H. A new three-dimensional model for emotions and monoamine neurotransmitters. *Medical Hypotheses*. 2012. 78, p. 341–348.
 9. Bochkarev V., Wichmann S. Solovyev V. Universals versus historical contingencies in lexical evolution/V. Solovyev, // *Journal of the Royal Society Interface*. - 2014. - V. 11.