




<https://doi.org/10.17323/jle.2024.22439>

Синхронические и диахронические предикторы рейтингов социальности слов

Бочкарев Владимир Владимирович , Шевлякова Анна Владимировна , Ачкеев Андрей Алексеевич 

Казанский федеральный университет, г. Казань, Россия

АННОТАЦИЯ

Введение: В недавних исследованиях была предложена и изучена новая психолингвистическая концепция — социальность слова. Рейтинг социальности отражает социальную значимость слова. Словари с рейтингами социальности составляются с использованием либо опросного, либо машинного метода. К сожалению, размер словарей с рейтингами социальности слов, созданных методом опроса, относительно невелик.

Цель: Составление большого словаря с рейтингами социальности английских слов с использованием машинной экстраполяции, а также перенос оценок рейтингов на другие языки и получение диахронических моделей рейтингов социальности.

Методология: Рейтинги социальности слов оценивались с помощью многослойных нейронных сетей прямого распространения. Для получения синхронических оценок на вход подавались предварительно обученные векторы fasttext. Для получения диахронических оценок использовалась статистика совместной встречаемости слов в большом диахроническом корпусе.

Результаты: Полученный коэффициент корреляции Спирмена между человеческими и машинными оценками социальности составил 0,869. Обученные модели позволили получить оценки социальности для двух миллионов английских слов, а также для большого количества слов на 43 других языках. Неожиданным результатом стало то, что линейная модель может обеспечить высокоточную оценку рейтингов социальности, которую практически невозможно существенно улучшить. По-видимому, это связано с тем, что в пространстве векторов, представляющих слова, выделяется направление, отвечающее за значения, и которое отражает степень социальной значимости, приписываемую словам. В статье также представлен диахронический нейросетевой предиктор оценок конкретности, использующий векторы совместной встречаемости слов в качестве входных данных. Показано, что с использованием годовых данных из большого диахронического корпуса Google Books Ngram можно получить точность, сопоставимую с точностью синхронических оценок.

Заключение: Созданный большой машинный словарь оценок социальности может быть использован в психолингвистических и культурологических исследованиях. Изменения в рейтингах социальности могут служить маркером изменения значения слова и применяться для выявления семантических изменений.

КЛЮЧЕВЫЕ СЛОВА

Социальность, психолингвистика, психолингвистические базы данных, предварительно обученные векторы слов, нейронные сети, семантические изменения

Для цитирования: Бочкарев, В.В., Шевлякова, А.В., & Ачкеев, А.А. (2024). Синхронические и диахронические предикторы рейтингов социальной значимости слов. *Journal of Language and Education*, 10(4), 45-57. <https://doi.org/10.17323/jle.2024.22439>

Контакты:

Бочкарев Владимир Владимирович
vbochkarev@mail.ru

Получена: 6 сентября 2024

Принята: 16 декабря 2024

Опубликована: 30 декабря 2024

ВВЕДЕНИЕ

Язык напрямую связан с восприятием человеком реальности и культурным контекстом. Поэтому были введены различные психолингвистические параметры слов, которые служат ключевыми характеристиками представления концептов и широко изучаются в современной науке. Среди упомянутых

параметров — конкретность слова, образность (imageability), тональность и т. д. Социальная значимость также является одной из ключевых характеристик представления концептов, поскольку социальность оказывает большое влияние на структуру концепта и познание. Социальность означает степень социальной значимости конкретного слова и описывает или ссылается на некоторые



социально значимые концепты, такие как социальная роль, социальное пространство, идеология и т. д. (Pexman et al., 2022). В последнее время были составлены словари психолингвистических параметров слов, включая социальность слов, которые могут быть использованы для решения различных практических задач.

Для изучения психолингвистических параметров слов используются опросные и машинные методы. Однако словари, составленные с использованием метода опроса, имеют относительно небольшой объем, поскольку для их составления требуется много времени и усилий. Создание больших текстовых корпусов и разработка методов обработки естественного языка позволили сформировать большие словари с психолингвистическими рейтингами слов путем машинной экстраполяции. В этом случае вычислительная модель обучается на небольшом количестве слов с имеющимися оценками человеком, а затем обученная модель генерирует машинные рейтинги для большого количества слов. Такой подход позволил получить большие машинные словари с рейтингами конкретности, аффективными рейтингами и др. (Mohammad et al., 2013; Koper & Schulte im Walde, 2016).

В последнее время ряд исследований был посвящен оценке рейтингов социальности слов с использованием метода опроса. В работе (Binder et al., 2016) был представлен первый английский словарь рейтингов социальности, включающий 535 слов. Позже в статье (Diveica et al., 2023) был представлен еще один словарь, который был составлен методом опроса и включал 8838 слов. Следует отметить, что инструкции, используемые в этой работе (Diveica et al., 2023), были гораздо более подробными, чем в первом словаре (Binder et al., 2016). Аналогичное исследование было проведено для китайского языка (Wang et al., 2023). Оно было выполнено в несколько этапов. Сначала был составлен словарь рейтингов социальности для 17 940 китайских слов с использованием метода опроса. Затем была обучена компьютерная модель, и с ее помощью рейтинги были экстраполированы на 900 тысяч китайских слов. Наконец, с помощью обученной модели и машинного перевода был получен машинный английский словарь, основанный на переносе рейтингов китайских слов на английский язык. Однако словарь, представленный в работе Diveica et al. (2023), имеет относительно небольшой объем, что может ограничивать его практическое применение. В связи с этим актуальной задачей становится создание компьютерной модели, которая позволила бы экстраполировать оценки социальности на максимально возможный набор слов.

Целью данной работы является создание большого английского словаря с рейтингами социальности слов на основе машинной экстраполяции. Кроме того, оценки рейтинга будут распространены на 43 других языка, что позволит получить диахронические модели рейтингов социальности. В работе используется модель, которая позволяет предсказывать рейтинги социальности для двух миллионов англий-

ских слов. В отличие от исследования Wang et al. (2023), в котором рейтинги были получены посредством переноса с китайского языка, в нашем исследовании модель обучалась с помощью рейтингов, полученных путем опросов носителей английского языка. Такой подход повышает точность оценивания рейтингов социальности для английских слов.

МЕТОДОЛОГИЯ

Исходные данные

Для обучения моделей использовались оценки рейтингов, полученные методом опросов в работе (Diveica et al., 2023), которая содержит рейтинги для 8388 английских слов. Значения рейтингов, приведенные в словаре, варьируются от 1 до 7. Высокие значения указывают на то, что слово имеет высокую социальную релевантность, а низкие — на ее отсутствие. Мода распределения значений рейтингов находится в середине шкалы. Таким образом, шкала оценок социальности по сути является биполярной. Для удобства мы преобразовали шкалу в диапазон от -1 до $+1$.

Используемые наборы векторов

Для оценки психолингвистических параметров слов используются векторные представления слов, разработанные в рамках дистрибутивной семантики. Общая идея дистрибутивной семантики заключается в том, что дистрибутивное сходство и смысловое сходство коррелируют друг с другом (Harris, 1970; Rubenstein & Goodenough, 1965; Firth, 1957). Поэтому значение слова может быть выявлено и оценено путем анализа его дистрибуции. Существуют различные алгоритмы получения дистрибутивного значения. В ранних работах в основном использовались представления, основанные на векторах совместной встречаемости (Weeds et al., 2004; Pantel, 2005; Bullinaria & Levy, 2007; Gulordava & Baroni, 2011). Bullinaria & Levy (2012) предложили использовать векторы, построенные на основе точечной взаимной информации (PMI). Одной из причин, препятствовавших эффективному применению ранних вложений слов, была высокая размерность векторов. Для снижения размерности применялись различные методы, например, с использованием SVD (Turney & Pantel, 2010; Bullinaria & Levy, 2012). В 2013 г. была предложена усовершенствованная техника вложений слов с использованием нейронных сетей (Mikolov et al., 2013; Wojanowski et al., 2017), которая открыла новые возможности в этой области исследований. Последние достижения в этой области включают использование контекстуализированных вложений слов (Peters et al., 2018; Devlin et al., 2019). Современные методы применения низкоразмерных векторных представлений слов подробно рассмотрены в работах Worth (2023), Pilehvar & Camacho-Collados (2020). В настоящее время в большинстве случаев используются методы, основанные на моделях векторных нейронных сетей. Однако и более простые представления, основанные на полных векторных представлениях слов,

находят применение, поскольку их использование имеет некоторые преимущества: легкая интерпретация результатов (Basile & McGillivray, 2018), возможность построения диахронических моделей (Bochkarev et al., 2022). В этой статье мы тестируем оба типа векторных представлений слов применительно к проблеме прогнозирования рейтингов социальности слов.

Во-первых, мы выбрали два набора предобученных векторов, которые были созданы на основе самых больших корпусов. Один из наборов включает векторы fasttext, обученные на корпусе CommonCrawl (Grave et al., 2018), включающем 650 миллиардов слов. В соответствии с рекомендациями Charbonnier & Wartena (2019) и результатами наших собственных экспериментов мы использовали векторы, обученные без применения информации сублексического уровня. Кроме того, мы использовали векторы Glove-840B, также обученные на корпусе CommonCrawl, включавшем 840 миллиардов слов на момент создания набора векторов (Pennington, 2014).

Во-вторых, для создания словарей для разных языков (кроме английского) мы использовали два многоязычных набора векторов. На странице проекта fasttext представлены вложения для 44 языков, которые были созданы по материалам Википедии (по состоянию на 2017 г.) и выровнены в едином векторном пространстве. Для выравнивания векторов был использован алгоритм, представленный в работе Juolin (2018). На странице проекта MuSE представлены выровненные вложения для 29 языков. Для создания этого многоязычного набора данных, как и в предыдущем случае, в качестве исходных были выбраны векторы fasttext, обученные на текстах Википедии, однако для выравнивания использовался другой алгоритм (Conneau et al, 2017). Все указанные выше наборы векторов относятся к классу контекстно-свободных моделей. Как уже упоминалось выше, более перспективными считаются контекстуализированные вложения слов. Однако следует отметить, что в существующем словаре, представленном в работе Diveisa et al. (2023), для каждой словоформы приводится только одно значение рейтинга социальности. Более того, для многозначных слов не указано, к какому из их значений относится рейтинг. В этом случае контекстуализированные вложения могут не иметь преимуществ по сравнению с контекстно-свободными моделями. Стоит также отметить, что все приведенные выше векторные наборы были получены в результате обучения на синхронических корпусах, поэтому они не могут быть использованы для получения диахронических оценок рейтингов социальности.

Помимо упомянутых выше низкоразмерных векторных представлений, мы также использовали явные векторы слов, построенные по методу CFW (совместная сочетаемость с наиболее частотными словами). Подробное описание метода можно найти в работах Xu & Kemp (2015) и Khristoforov et al. (2020). Согласно методу CFW, векторы были составлены из значений регуляризованной точечной

взаимной информации (в форме, предложенной в исследовании Bochkarev et al., 2021) для биграмм вида Wx и xW , где W — целевое слово, а x — одно из наиболее часто встречающихся слов. Данные о частотности слов и словосочетаний, необходимые для построения векторов, были извлечены из большого диахронического корпуса Google Books Ngram (Lin et al., 2012). Для обучения нейронной сети мы использовали усредненные частотные данные за период с 1900 по 2019 г. В данной работе, следуя Khristoforov et al. (2020), мы использовали список из 20 тысяч наиболее частотных слов. Таким образом, слово описывается вектором размерностью 40 000.

Обучение нейросетевых предикторов

Степень социальности слов оценивалась с помощью многослойных нейронных сетей прямого распространения. Для максимизации точности прогнозирования был протестирован ряд сетевых архитектур с различным количеством слоев и нейронов в каждом слое. Каждая из этих сетей обучалась с использованием нескольких алгоритмов (adadelta, adagrad, adam, SGD) с различными параметрами скорости обучения. Также проводились сравнительные тесты с применением регуляризации L1 и L2, а также с использованием метода «дропаута». На основании результатов тестирования была выбрана следующая архитектура предикторов нейронных сетей и параметров обучения для случая низкоразмерных векторных представлений:

- три плотных слоя из 3072 нейронов с функцией активации ReLU, выходной слой размерности 1 с линейной активацией;
- L2-регуляризация с коэффициентом $5 \cdot 10^{-4}$;
- метрика MSE для ранней остановки (улучшение не более $1 \cdot 10^{-6}$ в течение 100 эпох)

Аналогично были выбраны следующие параметры для предикторов, использующих полные векторные представления слов:

- шесть плотных слоев из 512 нейронов с функцией активации ReLU, выходной слой размерности 1 с линейной активацией;
- регуляризация с использованием «дропаута» между плотными слоями с коэффициентом 0,02;
- метрика MSE для ранней остановки (улучшение не более $1 \cdot 10^{-6}$ в течение 5 эпох).

В обоих случаях наилучшие результаты были получены с использованием функции потерь MSE и алгоритма оптимизации SGD.

Процедура кросс-проверки

Для повышения надежности результатов и контроля точности полученных оценок использовалась кросс-проверка. В соответствии с описанием в работе Бочкарева и др. (2024a)

список слов был разделен на шесть непересекающихся групп. В каждом случае четыре группы из шести использовались для обучения модели, а оставшиеся две группы служили в качестве тестового набора. поскольку существует 15 различных способов выбрать 4 группы из 6, было получено 15 независимо обученных моделей. При этом для каждого слова существует 5 моделей, для которых это слово было в тестовом наборе, а не в обучающем. Наличие нескольких моделей позволило дополнительно повысить точность путем усреднения оценок, а также определить стандартное отклонение полученной оценки.

Обучение линейных предикторов

Помимо нейронных сетей, мы также представим для сравнения результаты линейных моделей. В случае полных векторных представлений слов связь между отдельными компонентами вектора и содержательными характеристиками слова является, очевидно, нелинейной. Поэтому имеет смысл использовать линейные предикторы только для случаев, когда на вход подаются векторы малой размерности. Обучение линейного предиктора является по сути задачей линейной регрессии и осуществляется с использованием псевдообращения по норме L2. Так же, как и для нейросетевых предикторов, в этом случае набор моделей независимо обучается на 15 подмножествах выборки. Оценки, полученные с помощью независимо обученных моделей могут быть усреднены тем или иным способом.

Перенос оценок на другие языки

Наличие свободно доступных многоязычных наборов векторов, выровненных в едином векторном пространстве, позволяет легко переносить оценки с одного языка на другой. Существует словарь с рейтингами оценок социальности для английского языка, полученными путем опросов. Мы обучаем предиктор рейтингов социальности, используя векторы для английских слов из многоязычного векторного набора в качестве входных данных. Подставляя векторы слов для другого языка из того же набора, мы получаем словарь с оценками социальности для этого языка. Следует учиты-

вать, что ошибки в оценках рейтингов, связанные с несовершенством модели, будут суммироваться с ошибками в выравнивании векторов в двух языках. Поэтому для решения задачи переноса оценок в данной работе используются линейные предикторы, так как в этом случае легче предсказать значение ошибки выходного значения, если известны ошибки во входных данных.

РЕЗУЛЬТАТЫ

Мы обучили модели оценке социальности английских слов, используя четыре набора предобученных векторов и один набор полных векторных представлений слов. Для каждого из этих пяти наборов векторов были обучены нейросетевые предикторы. Кроме того, для четырех низкоразмерных наборов векторов были обучены линейные предикторы.

Как было упомянуто выше, оценки были получены для пяти моделей, в которых это слово было в тестовой выборке, и поэтому не было представлено нейронной сети на этапе обучения. Эти оценки были усреднены для каждого слова. Коэффициенты корреляции Пирсона и Спирмена между полученными таким образом усредненными оценками и человеческими оценками для различных представлений слов и архитектур предиктора приведены в Таблице 1. Усреднение по набору независимо обученных моделей позволяет повысить точность оценки. Например, для набора предобученных векторов fasttext-CommonCrawl среднее значение коэффициентов корреляции Пирсона и Спирмена между опросными рейтингами и их машинными оценками для 15 моделей составило 0,8531 и 0,8566 соответственно. Усреднение результатов по независимо обученным моделям позволило увеличить значения коэффициентов корреляции до значений 0,8655 и 0,8688 соответственно (Таблица 1).

Прежде всего, следует отметить, что точность линейных предикторов совсем немного уступает точности нейросетевых предикторов, использующих тот же набор предобученных векторов. При этом линейные предикторы имеют

Таблица 1
Коэффициенты корреляции Пирсона (r) и Спирмена (ρ) между усредненными оценками по независимо обученным моделям и рейтингам, полученными путем опроса

Набор векторов	Тип предиктора	r	ρ
fasttext-CommonCrawl	Нейросетевой	0.8655	0.8688
fasttext-CommonCrawl	Линейный	0.8411	0.8502
GloVe-840B	Нейросетевой	0.8541	0.8577
GloVe-840B	Линейный	0.8361	0.8418
fasttext-wiki	Нейросетевой	0.8390	0.8418
fasttext-wiki	Линейный	0.8179	0.8251
MuSE	Нейросетевой	0.8388	0.8414
MuSE	Линейный	0.8183	0.8255
Co-occurrence vectors (CFW)	Нейросетевой	0.8512	0.8540

число настраиваемых параметров, равное размерности входных векторов (в нашем случае — 300 параметров), тогда как нейросетевые предикторы, использующие те же входные данные, имеют 13,9 млн весовых коэффициентов (см. Раздел 2). Модель, использующая полные векторные представления слов, имеет даже 21,8 млн весовых коэффициентов. Таким образом, это незначительное увеличение точности нейросетевых предикторов достигается колоссальным усложнением модели и соответствующим увеличением времени обучения.

Следует также отметить, что 15 линейных моделей, независимо обученных на разных подмножествах слов, обладают высокой степенью согласованности друг с другом. Для обученного линейного предиктора градиент выхода модели постоянен во всем векторном пространстве. Таким образом, i -я модель может быть охарактеризована единичным вектором направления v_i , градиентом, нормированным на единичную длину. Например, для набора предварительно обученных векторов fasttext-CommonCrawl медианное значение косинусов углов между парами векторов направления независимо обученных моделей составило 0,9728. Мы можем синтезировать одну модель из 15 независимо обученных моделей. Для этого мы усредняем векторы направления отдельных моделей и, нормализуя полученный вектор на единичную длину, получаем вектор направления одной синтетической модели V :

$$V = \frac{\sum_i v_i}{\|\sum_i v_i\|}$$

Медианное значение проекций векторов направления 15 моделей v на направление V для набора предобученных векторов fasttext-CommonCrawl составляет 0,9842. Это доказывает высокую степень согласованности всех 15 моделей.

Интересно, что использование синтетической модели позволяет достичь более высокой точности по сравнению с простым усреднением оценок. Например, коэффициент корреляции Спирмена между человеческими оценками и средним машинным рейтингом для набора предобученных векторов fasttext-CommonCrawl равен 0,8354. Он был получен с помощью простого усреднения. Коэффициент корреляции Спирмена, полученный с помощью синтетической модели, равен 0,8502. Поэтому в Таблице 1 приведены значения коэффициентов корреляции для линейных предикторов, полученных с помощью синтетической модели.

Далее сравним точность предикторов, использующих различные наборы предобученных векторов. Вложения fasttext-wiki и MuSE были получены путем обучения с использованием корпуса текстов Wikipedia, который имеет гораздо меньший размер по сравнению с корпусом CommonCrawl, поэтому предикторы, использующие эти векторы, показывают меньшую точность. Следует отметить, что

предварительно обученные векторы Glove-840B, несмотря на больший размер обучающего корпуса, показали более низкий результат по сравнению с fasttext-CommonCrawl. Аналогичное наблюдение было описано в работе Wang et al. (2023). По-видимому, это обусловлено более высоким качеством обучения встраиваний, которое достигается при использовании алгоритма fasttext.

Диахронический предиктор рейтингов социальности слов

Как видно из Таблицы 1, использование полных векторных представлений слов, построенных с использованием метода CFW, позволяет получить практически такую же точность прогнозирования, как и использование векторов малой размерности, обученных на корпусах большого объема. Несмотря на несколько меньшую точность, метод CFW имеет большое преимущество — он позволяет легко получить диахроническую модель при наличии соответствующего корпуса (Bochkarev et al., 2022). Для этого требуется лишь построить явные векторы слов для целевых временных интервалов, используя данные о совместной встречаемости целевого слова в диахроническом корпусе. Затем полученные векторы подаются на вход предиктора, что позволяет получать диахронические оценки рейтинга социальности целевого слова.

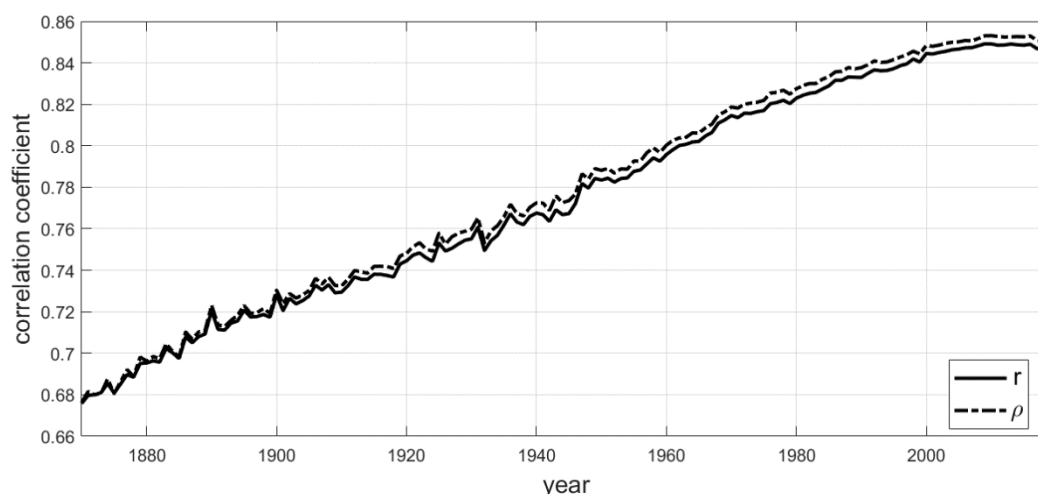
Корпус Ngram Google Books предоставляет ежегодные данные о частоте слов и словосочетаний. Используя данные корпуса GBN за каждый год с 1870 по 2019, мы построили векторы для каждого слова из словаря (Diveica et al., 2023) и рассчитали соответствующие оценки рейтинга социальности этих слов для каждого года. Коэффициенты корреляции Пирсона и Спирмена между рейтингами, полученными путем опроса, и машинными оценками, рассчитанными с использованием годовых данных, показаны на Рисунке 1.

Максимальное коэффициента корреляции Спирмена составляет 0,8531 (в 2010 г.), что всего на несколько десятых тысячных меньше значения коэффициента корреляции Спирмена приведенного в Таблице 1, полученного с использованием данных всего временного интервала с 1900 по 2019 г. В интервале с 2000 по 2019 г. коэффициент корреляции Спирмена не опускается ниже 0,8480. Годовой размер английского подкорпуса Google Books Ngram в рассматриваемые годы варьируется от 22,8 до 34,9 млрд слов. Как видно из Рисунка 1, а также исходя из представленных значений, такой объем данных уже позволяет получить оценки, не менее точные, чем полученные с использованием всех имеющихся данных.

В то же время для более ранних периодов, согласно Рисунку 1, коэффициенты корреляции между рейтингами, основанными на опросах, и соответствующими машинными оценками, рассчитанными на основе годовых данных, заметно ниже. Основной причиной этого является уменьшение годового размера корпуса в более ранние периоды.

Рисунок 1

Коэффициенты корреляции Пирсона (r) и Спирмена (ρ) между рейтингами, полученными путем опроса, и машинными оценками, рассчитанными с использованием ежегодных данных



Кроме того, эволюция языка может вызывать изменения в оценках социальности слов с течением времени. Поскольку использованные для обучения рейтинги были получены в результате опросов, проведенных в последние годы, это явление также может привести к падению коэффициента корреляции между рейтингами социальности на основе опросов и их машинными оценками для более ранних периодов.

Ниже мы проанализируем некоторые примеры слов, у которых рейтинг социальности изменяется со временем, и попытаемся выявить возможные закономерности изменения рейтинга.

Первый пример изменения рейтинга социальности представлен словом *apple* (яблоко) (Рисунок 2, А). До конца 70-х годов рейтинг был отрицательным и колебался около значения 0,7. На графике виден рост значений его рейтинга с конца 80-х годов. Это объясняется тем, что это слово, традиционно обозначавшее вид фруктов, приобрело новое значение — название транснациональной корпорации. Самый быстрый рост рейтингов социальности наблюдается в момент выхода модели компьютера Macintosh, когда персональные компьютеры Apple приобрели широкую популярность (Linzmayr, 2004). И сейчас Apple — всемирно известный бренд, ассоциирующийся с высоким социальным статусом. Таким образом, появление нового значения вызвало изменение рейтингов социальности. Второй всплеск на графике тоже не случаен — он совпадает с запуском семейств мобильных устройств компанией Apple.

Другим примером является слово *bush*, которое показывает схожую тенденцию, что и предыдущее слово (Рисунок 2, Б). Основное значение этого слова — «древовидное растение, разветвляющееся у самого основания». Поэтому рейтинг социальности этого слова изначально был отрицательным в течение длительного времени, оставаясь на уровне 0,6. Однако это слово также обозначает фамилию немецко-

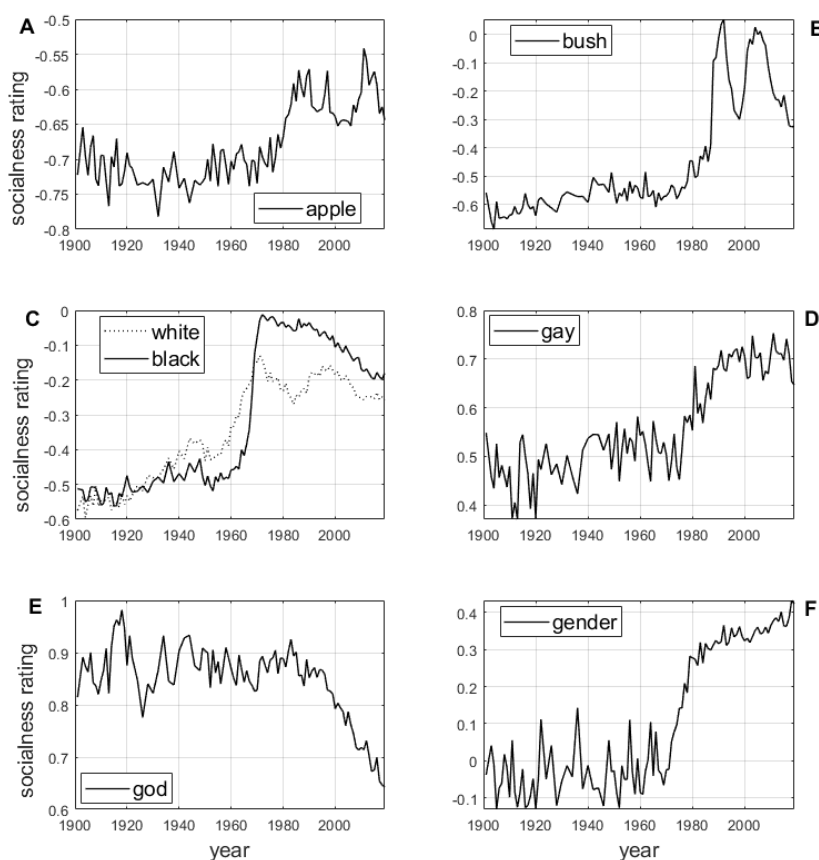
го происхождения, в частности, известной политической династии в США (Schweizer & Schweizer, 2004). Начиная со времен Второй мировой войны, можно наблюдать постепенный рост рейтинга социальности слова *bush*, связанный с ростом процента использования этого слова в качестве имени собственного. Рост ускорился в 70-е годы, когда карьера известного американского политика Джорджа Буша — старшего начала широко освещаться в прессе. Особенно значительные скачки рейтингов социальности наблюдаются в 1988 г. (когда Джордж Буш — старший победил на выборах президента США) и в 2000 г. (когда выборы выиграл его сын Джордж Буш — младший).

Следующие примеры слов, изменивших свои рейтинги социальности, — это *white* и *black* (Рисунок 2, В). Изначально они обозначали цвета. Однако из-за метонимического сдвига приобрели второе значение — обозначение социальных групп людей по цвету кожи. На графиках показано, как изменялись рейтинги социальности этих слов. Для XIX в. рейтинги социальности обоих слов колеблются от -0,6 до -0,55. Начиная с 1920 г., после Первой мировой войны, наблюдается тенденция к росту рейтингов, с резким скачком в 60-х годах, что связано с активизацией движения за гражданские права афроамериканцев. Интересно, что наибольшее изменение рейтингов произошло в 1968 г., когда был убит Мартин Лютер Кинг, видный лидер движения за гражданские права афроамериканцев. Это трагическое событие вызвало огромный резонанс в американском обществе.

Еще одним примером изменения значения, вызвавшего изменение рейтинга слова, является слово *gay* (Рисунок 2, Г), которое изменило и денотаты, и отношение к определенной части речи (POS). *Gay* (веселый) — это слово было с умеренно низкими рейтингами социальности (в диапазоне от 0,45 до 0,50). Однако с начала 80-х годов наблюдается быстрый рост рейтинга. В качестве существительного *gay* обозначает человека нетрадиционной ориентации. Несмо-

Рисунок 2

Примеры слов, для которых рейтинг социальности меняется с течением времени



тря на то, что это значение слова *gay* появилось давно, но до недавнего времени оно не было широко распространено. Увеличение частоты слова во втором значении в 1980-х гг. было связано с социальными процессами в США, что и спровоцировало рост рейтингов социальности.

Существуют также слова, обозначающие абстрактные понятия и социальные концепции, рейтинги социальности которых изменяются главным образом за счет изменения культурного контекста и восприятия. Рассмотрим слова *god* и *gender* (Рисунки 2, Д, Е). *God* всегда был словом с высоким рейтингом социальности. Однако его рейтинги также колеблются. Самые высокие пики наблюдались в 1914–1921 гг. (время Первой мировой войны и несколько лет после нее) и в 1940–1945 гг. (время Второй мировой войны). Максимальное значение рейтинга 0,982 было достигнуто в 1918 г., тогда как годовые значения рейтинга находились в основном в диапазоне 0,85–0,9. Особое внимание заслуживает текущая тенденция к снижению рейтинга социальности слова *god*. Эта тенденция наметилась в последние годы XX в. Этот вопрос требует дополнительных исследований, однако можно

предположить, что это связано с тенденцией к более личностному восприятию идеи Бога.

Еще один интересный пример — слово *gender*, которое в основном означает пол. Его рейтинги социальности были почти нейтральными до 1970-х гг. Однако, с начала 70-х гг. можно наблюдать стремительную тенденцию к росту. Согласно этимологическому словарю¹, никаких новых значений слова *gender* в XX в. не появилось, но в результате общественных дискуссий его восприятие меняется, это понятие переосмысливается как более социально значимое, что и отразилось на росте рейтингов социальности.

Перенос рейтингов социальности на другие языки

Используя обученные линейные предикторы для fast-text-wiki и предобученные векторные наборы MuSE для английского языка, мы получили машинные словари с рейтингами социальности для 43 и 28 других языков соответственно. Наибольшую трудность представляла проверка

¹ Online Etymology Dictionary. (n.d.). Gender. Последнее обращение — 15.07.2024, <https://www.etymonline.com/search?q=gender>

качества полученных рейтингов социальности для каждого из этих языков.

Во-первых, это можно было сделать путем выборочной ручной проверки полученных машинных оценок (полная ручная проверка практически невозможна из-за большого размера полученных словарей). Мы проверили русские слова, получившие максимально высокие и минимальные оценки. Проверка показала, что модель справилась с этой задачей очень хорошо. Слова с высокими рейтингами социальности оказались в верхней части списка (например, *общественность, партнерство, дружелюбие, вежливый, демократичность*) а слова с низкими рейтингами — в нижней части списка (например, *целлофан, брыжейка, нубук, стручок, песчаник*).

Во-вторых, при наличии предобученных векторов fast-text-wiki и MuSE для одного языка мы можем сравнивать рейтинги, полученные с использованием каждого из этих двух наборов, что (в некоторой степени) позволяет судить о качестве полученных рейтингов. В Таблице 2 показаны значения коэффициентов корреляции Пирсона и Спирмена между оценками социальности, полученными с использованием двух наборов выровненных, предобученных векторов.

Более высокие значения коэффициента корреляции свидетельствуют о большей степени сходства между машинными оценками, полученными с использованием двух разных наборов векторов, и, следовательно, о более высокой степени надежности результатов для данного языка.

Что касается китайского языка, для него доступны только векторы из многоязычного набора данных fasttext-wiki. Однако в работе Wang et al. (2023) представлены независимые

рейтинги социальности, полученные путем опроса, для 17 940 китайских слов. Коэффициенты корреляции Пирсона и Спирмена между рейтингами социальности по данным опросов из этой работы и нашими машинными оценками составляют 0,6010 и 0,6382 соответственно. Для 64 791 китайского слова доступны как машинные оценки, представленные в исследовании Wang et al. (2023), так и наши машинные оценки. Коэффициенты корреляции Пирсона и Спирмена между этими двумя наборами оценок равны 0,5828 и 0,5827 соответственно. Следует отметить, что среди 64,8 тыс. упомянутых слов значительную долю составляют редкие и малочастотные слова, для которых машинные оценки менее точны. Это обстоятельство повлияло на снижение уровня корреляции.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Метод опроса, используемый для создания словарей с психолингвистическими оценками, является достаточно трудоемким. Поэтому словари, полученные путем опросов, имеют относительно небольшой размер. Развитие технологий обработки естественного языка вызвало появление значительного количества работ, посвященных экстраполяции оценок, полученных методом опроса, на широкий диапазон слов. Таким образом, были созданы большие машинные словари для многих психолингвистических параметров, таких как словари аффективных рейтингов, конкретности и образности (Mohammad et al., 2013; Koper & Schulte im Walde, 2016; Charbonnier & Wartena, 2019).

Однако на сегодняшний день только одна работа посвящена недавно введенному рейтингу социальности, в ней делается попытка построить большие машинные словари для китайского и английского языков. При этом рейтинги для

Таблица 2
Коэффициенты корреляции Пирсона (r) и Спирмена (ρ) между оценками социальности слов, полученными с использованием выровненных, предварительно обученных векторных наборов Fasttext-Wiki и MuSE для того или иного языка

Язык	r	ρ	Язык	r	ρ
Болгарский	0,8106	0,7910	Индонезийский	0,8763	0,8638
Каталонский	0,9111	0,9005	Итальянский	0,9409	0,9316
Чешский	0,8730	0,8617	Македонский	0,8221	0,8067
Датский	0,8818	0,8728	Голландский	0,9300	0,9201
Немецкий	0,9062	0,8959	Норвежский	0,9206	0,9137
Греческий	0,8256	0,8032	Польский	0,8842	0,8745
Английский	0,9997	0,9996	Португальский	0,9436	0,9386
Испанский	0,9492	0,9427	Румынский	0,8889	0,8775
Эстонский	0,8922	0,8859	Русский	0,8649	0,8512
Финский	0,8540	0,8420	Словацкий	0,8240	0,8107
Французский	0,9259	0,9149	Словенский	0,8357	0,8216
Иврит	0,6948	0,6656	Шведский	0,9100	0,8922
Хорватский	0,8901	0,8833	Турецкий	0,8328	0,8162
Венгерский	0,8628	0,8528			

английского языка в данной работе были получены путем переноса рейтингов с китайского языка, несмотря на то, что социальная значимость одного и того же слова в разных языках может быть не одинакова.

В настоящей работе представлены следующие результаты. Во-первых, были обучены синхронические модели рейтингов социальности для английского языка, а также был составлен словарь с машинными рейтингами социальности для двух миллионов слов. Во-вторых, диахроническая модель рейтингов социальности также была обучена для английского языка, и были проанализированы изменения в восприятии слов как связанных, так и не связанных с социальной сферой. Наконец, с помощью выровненных наборов предобученных векторов полученные оценки рейтингов были перенесены на 43 языка.

Синхронические модели

Построенные модели позволяют получать оценки рейтингов социальности английских слов с достаточно высокой точностью. Максимальное значение коэффициента корреляции Спирмена между рейтингами, полученными методом опроса, и их оценками составило 0,8688. Этот результат сопоставим со значениями коэффициентов корреляции опросных рейтингов и машинных оценок, полученных при прогнозировании аффективных рейтингов, рейтингов конкретности и других психолингвистических характеристик английских слов (Buechel & Hahn, 2018; Charbonnier & Wartena, 2019; Bochkarev et al., 2021). Существует несколько английских словарей с оценками рейтингов тональности и конкретности слов, которые были разработаны независимо друг от друга различными исследовательскими группами. Сравнительный анализ, проведенный в работе Charbonnier & Wartena (2019), показал, что достигнутый уровень корреляции рейтингов по методу опроса и машинных оценок уже близок к уровню корреляции опросных оценок, представленных различными группами.

Провести аналогичные сравнения для рейтинга социальности проблематично, поскольку словарь, представленный в работе Diveica et al. (2023), по-прежнему является единственным большим английским словарем с рейтингами социальности.

Коэффициент корреляции Пирсона между рейтингами, полученными путем опросов и представленными в исследованиях Diveica et al. (2023) и Binder et al. (2016), составляет 0,76. Однако он рассчитан только для 258 слов, которые входят в состав обоих словарей. В работе Wang et al. (2023) отмечается, что коэффициент корреляции Пирсона между рейтингами, полученными при переводе китайских слов из словаря, предложенного авторами указанной работы и рейтингами словаря Diveica et al. (2023) составляет 0,724. Таким образом, в нашей работе (Таблица 1) уровень корреляции между рейтингами, полученными методом опросов и машинными оценками значительно превышает уровень

корреляции между опросными рейтингами, составленными разными исследователями. Мы также рассчитали коэффициенты корреляции Пирсона и Спирмена между полученными нами машинными оценками (нейронная сеть предиктор, векторы fasttext-CommonCrawl) и рейтингами 535 слов из словаря (Binder et al., 2016), которые составили 0,7545 и 0,7944 соответственно. Значения оказались не ниже коэффициентов корреляции между опросными рейтингами, приведенными в двух указанных выше словарях.

Кроме того, сходство полученного значения коэффициентов корреляции человеческих и машинных рейтингов социальности с коэффициентами корреляции человеческих и машинных рейтингов для других психолингвистических параметров позволяет предположить, что полученный уровень точности прогнозирования рейтингов социальности также близок к максимально достижимому.

Диахронические модели

Предварительно обученные векторы низкой размерности, находящиеся в открытом доступе, были получены путем обучения на синхронических текстовых корпусах. Таким образом, они не подходят для получения диахронических оценок рейтингов социальности слов. Напротив, модель, использующая полные векторные представления слов, позволяет легко получать диахронические оценки рейтингов социальности слов, используя любой диахронический корпус достаточного размера. Мы априори ожидаем, что восприятие социальности многих слов может изменяться со временем. Действительно, при рассмотрении конкретных примеров мы сразу же выявляем случаи таких изменений. Приведенные примеры показывают, что, во-первых, рейтинг социальности слова может претерпевать резкие изменения, когда слово приобретает новое значение, коннотацию, или вследствие изменения культурного контекста. Полная классификация случаев резких изменений оценок социальности слов требует отдельного большого исследования; в данной работе мы рассмотрели лишь несколько примеров, иллюстрирующих различные возможные направления дальнейшей работы. Тем не менее, рассмотренные примеры показывают, что изменение рейтинга социальности может быть маркером лексико-семантического изменения.

Значительное количество работ посвящено задаче обнаружения лексико-семантических изменений (Tang, 2018; Hengchen et al. 2021). В большинстве случаев в таких работах используется то или иное диахроническое векторное представление слов. Изменение вектора, представляющего слово, или изменение прямого контекста слова в векторном пространстве рассматривается как индикатор изменения значения. Возможен также альтернативный подход, впервые описанный в исследовании Ryzhova et al. (2021), когда рассматривается статистика употребления слов в тексте в той или иной грамматической форме (грамматические профили слова), и изменение такой статистики

служит маркером лексико-семантических изменений. В этой работе (Ryzhova, et al., 2021) рассматривались только грамматические признаки слов, однако для выявления лексико-семантических изменений могут быть использованы и такие признаки, как употребление слова в качестве имени собственного или нарицательного (Bochkarev et al., 2022), а также психолингвистические характеристики слов (Bochkarev et al., 2024a).

Изменения рейтинга социальности слов также могут служить дополнительными маркерами семантических трансформаций. Также в ряде рассмотренных случаев мы наблюдали, что значение слова сохранялось неизменным, однако его рейтинг социальности изменился под влиянием некоторых культурных факторов. Таким образом, диахроническая модель рейтингов социальности слов может быть полезна для культурологических исследований.

Перенос рейтингов на другие языки

Наличие векторных представлений слов, выровненных в едином векторном пространстве, позволило легко перенести рейтинги социальности с английского на 43 других языка. Основная проблема была связана с проверкой полученных машинных рейтингов для других языков. Выборочная проверка для русского языка показала, что слова с высокими положительными или отрицательными машинными рейтингами обычно оцениваются моделью адекватно. Полная ручная проверка рейтингов для всех языков чрезвычайно трудоемка и в настоящее время находится за пределами наших возможностей. Однако для 28 языков у нас есть независимые оценки рейтингов социальности, полученные с использованием двух наборов векторов — fasttext-wiki и MuSE. Сравнение рейтингов, полученных двумя независимыми методами, позволяет нам судить о качестве сформированных словарей. При интерпретации коэффициентов корреляции, представленных в Таблице 2, следует учитывать два фактора. Во-первых, первоначальные человеческие оценки были получены для английского языка, поэтому языки, более близкие к английскому, показывают лучшую корреляцию. Например, корреляция оценок с немецким и датским языками высокая, тогда как для вьетнамского языка она самая низкая. Очевидно, что социальная значимость слов воспринимается в этих языках одинаково. Во-вторых, объем текстов Википедии на европейских языках больше, а значит для этих языков был больше и объем обучающего корпуса, использованного для построения векторных представлений слов.

Очень важным результатом является неожиданно высокая эффективность линейных предикторов в прогнозировании рейтинга социальности. В предыдущем разделе было показано, что результаты линейных предикторов могут быть лишь немного улучшены более сложными нейросетевыми предикторами, обладающими на пять порядков большим числом параметров подгонки. В этом случае оценка рейтингов социальности резко отличается от того, что наблюдает-

ся для других психолингвистических параметров, таких как рейтинги тональности, конкретности и образности. Во всех упомянутых случаях, за исключением рейтинга социальности, линейные предикторы сильно уступают нейросетевым предикторам по точности.

Следует также отметить, что линейные предикторы, независимо обученные на разных подмножествах слов, оказались высоко согласованными друг с другом. Это доказывает, что в пространстве векторов, представляющих слова, есть выделенное направление, отвечающее за восприятие слов как связанных, так и не связанных с социальной сферой. Социальные связи и взаимодействия играют важнейшую роль в жизни человека, что закономерно находит отражение в языке. По-видимому, существенная роль социальных факторов улавливается существующими языковыми моделями, что приводит к появлению выделенного направления в векторном пространстве, отвечающем за степень восприятия слова как социально значимого. Таким образом, рейтинг социальности слова в векторном пространстве слов можно охарактеризовать очень просто: рейтинг социальности растет вдоль выделенного направления. Соответственно, оценки рейтинга в первом приближении могут быть получены как проекции векторов слов на это направление.

Ограничения настоящего исследования могут быть связаны, во-первых, с особенностями исходного словаря с рейтингами, полученными методом опроса, который использовался для обучения модели. Как было показано в работе Bochkarev et al. (2024b), различия в составе лексики аффективных словарей, созданных методом опроса, могут приводить к смещению полученных машинных оценок. Провести аналогичное исследование для рейтингов социальности слов пока не представляется возможным, поскольку, как уже упоминалось выше, на данный момент словарь (Diveica и др., 2023) является единственным большим словарем английского языка с рейтингами социальности. Во-вторых, очевидное ограничение связано с тем, что существующие модели не способны предоставлять оценки для различных значений многозначных слов. Прогресс в этом направлении может быть достигнут путем использования контекстно зависимых вложений слов.

ЗАКЛЮЧЕНИЕ

В данной статье была решена задача составления большого словаря с рейтингами социальности английских слов на основе предложенных компьютерных моделей. Точность разработанных моделей высока: наилучшее достигнутое значение коэффициента корреляции Спирмена между человеческими оценками и их машинными оценками составляет 0,8688. Использованные модели позволили нам экстраполировать рейтинги на очень широкий диапазон слов, в результате чего нам удалось получить машинные оценки для двух миллионов слов. Таким образом, полученный словарь рейтингов социальности слов значительно превосхо-

дит по объему ранее существовавшие аналоги. Кроме того, рейтинги для широкого диапазона слов из 43 других языков были получены с использованием свободно доступных вложений слов, выровненных в одном векторном пространстве. Также был построен диахронический предиктор рейтинга социальности с использованием полных векторных представлений слов.

Высокая эффективность линейных предикторов при прогнозировании рейтингов социальности оказалась неожиданной. Фактически достаточно найти проекцию вектора, представляющего слово, на некоторое выбранное направление в векторном пространстве, чтобы получить хорошую оценку социальности слова. Такая простая оценка может быть немного улучшена, однако это очень трудоемкий и длительный процесс. Мы предполагаем, что поскольку отношения в обществе играют фундаментальную роль в жизни человека, значительное влияние социальных факторов отражается в существующих языковых моделях и приводит к появлению выделенного направления в векторном пространстве, отвечающего за степень восприятия слова как социально значимого.

Также с использованием полных векторных представлений слов был построен диахронический предиктор рейтинга социальности. Показано, что с помощью статистики совместной встречаемости слов в большом диахроническом корпусе возможно выявлять изменения рейтингов социальности с течением времени.

Полученные результаты могут быть полезны для нескольких областей науки. Созданный словарь является хорошим материалом для психолингвистических и культурологических исследований. Более того, поскольку проанализи-

рованные примеры слов иллюстрируют, что изменение рейтинга социальности может быть маркером лексико-семантических изменений, диахроническая модель может быть использована для этимологических исследований.

Также отметим несколько направлений для дальнейшей работы. В первую очередь, это получение рейтингов социальности для полисемантических слов с использованием контекстно зависимых вложений слов. Вторым направлением является использование контекстно зависимых вложений для повышения эффективности переноса рейтингов на другие языки.

БЛАГОДАРНОСТИ

Исследование выполнено при финансовой поддержке Российского научного фонда, грант № 20-18-00206.

ВКЛАД АВТОРОВ

Владимир Бочкарев: концепция; методология; написание черновика; руководство работой.

Анна Шевлякова: концепция; написание черновика, рецензирование и редактирование.

Андрей Ачкеев: программное обеспечение, визуализация.

КОНФЛИКТ ИНТЕРЕСОВ

Авторы заявляют об отсутствии конфликта интересов.

ЛИТЕРАТУРА

- Basile, P., & McGillivray, B. (2018). Exploiting the web for semantic change detection. *Lecture Notes in Computer Science*, 11198, 194–208. https://doi.org/10.1007/978-3-030-01771-2_13
- Binder, J.R., Conant, L.L., Humphries, C.J., Fernandino, L., Simons, S.B., Aguilar, M., & Desai, R.H. (2016). Toward a brain-based componential semantic representation. *Cognitive Neuropsychology*, 33(3–4), 130–74. <https://doi.org/10.1080/02643294.2016.1147426>
- Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2017). Enriching Word Vectors with Subword Information. *Transactions of the Association for Computational Linguistics*, 5, 135–146. https://doi.org/10.1162/tacl_a_00051
- Bochkarev, V. V., Savinkov, A. V., & Shevlyakova, A. V. (2021). Estimation of imageability ratings of English words using neural networks. In I. Baturshin, A. Gelbukh, & G. Sidorov (Eds.), *Advances in Soft Computing. MICAI 2021. Lecture Notes in Computer Science* (vol. 13068, pp. 59–69). Springer. https://doi.org/10.1007/978-3-030-89820-5_5
- Bochkarev, V. V., Khristoforov, S. V., Shevlyakova, A. V., & Solovyev, V. D. (2022). Neural network algorithm for detection of new word meanings denoting named entities. *IEEE Access*, 10, 68499–68512. <https://doi.org/10.1109/ACCESS.2022.3186681>
- Bochkarev, V., Khristoforov, S., Shevlyakova, A., & Solovyev, V. (2024). Diachronic analysis of a word concreteness rating: Impact of semantic change. *Lobachevskii Journal of Mathematics*, 45, 961–971. <https://doi.org/10.1134/S1995080224600559>
- Bochkarev, V., Solovyev, V., Nestik, T., & Shevlyakova, A. (2024). Variations in average word valence of Russian books over a century and social change. *Journal of Mathematical Sciences*, 285, 14–27. <https://doi.org/10.1007/s10958-024-07419-z>
- Buechel, S., & Hahn, U. (2018). Word emotion induction for multiple languages as a deep multi-task learning problem. *Proceedings of NAACL-HLT 2018* (pp. 1907–1918). Association for Computational Linguistics. <https://doi.org/10.18653/v1/N18-1173>

- Bullinaria, J., & Levy, J. (2007). Extracting semantic representations from word co-occurrence statistics: A computational study. *Behavior Research Methods*, 39, 510–526. <https://doi.org/10.3758/BF03193020>
- Bullinaria, J. A., & Levy, J. P. (2012). Extracting semantic representations from word co-occurrence statistics: Stop-lists, stemming, and SVD. *Behavior Research Methods*, 44(3), 890–907. <https://doi.org/10.3758/s13428-011-0183-8>
- Charbonnier, J., & Wartena, C. (2019). Predicting word concreteness and imagery. *Proceedings of the 13th International Conference on Computational Semantics - Long Papers* (pp. 176–187). Association for Computational Linguistics. <https://doi.org/10.18653/v1/W19-0415>
- Conneau, A., Lample, G., Denoyer, L., Ranzato, M. A., & Jégou, H. (2017). *Word translation without parallel data*. arXiv preprint arXiv:1710.04087.
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (vol. 1: Long and Short Papers, pp. 4171–4186). Association for Computational Linguistics.
- Diveica, V., Pexman, P. M., & Binney, R. J. (2023). Quantifying social semantics: An inclusive definition of socialness and ratings for 8388 English words. *Behavior Research Methods*, 55, 461–473. <https://doi.org/10.3758/s13428-022-01810-x>
- Firth, J. (1957). A synopsis of linguistic theory, 1930–55. *Studies in linguistic analysis* (Special Volume of the Philological Society, pp. 1–31). Blackwell.
- Grave, E., Bojanowski, P., Gupta, P., Joulin, A., & Mikolov, T. (2018). Learning word vectors for 157 languages. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. ELRA.
- Gulordava, K., & Baroni, M. (2011). A distributional similarity approach to the detection of semantic change in the Google Books Ngram corpus. *Proceedings of the GEMS 2011 Workshop on Geometrical Models of Natural Language Semantics* (pp. 67–71). Association for Computational Linguistics.
- Harris, Z. (1970). *Papers in structural and transformational linguistics*. Reidel.
- Hengchen, S., Tahmasebi, N., Schlechtweg, D., & Dubossarsky, H. (2021). Challenges for computational lexical semantic change. In N. Tahmasebi, L. Borin, A. Jatowt, Y. Xu, & S. Hengchen (Eds.), *Computational approaches to semantic change* (pp. 341–372). Language Science Press.
- Joulin, A., Bojanowski, P., Mikolov, T., Jégou, H., & Grave, E. (2018). Loss in translation: Learning bilingual word mapping with a retrieval criterion. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing* (pp. 2979–2984). Association for Computational Linguistics.
- Khristoforov, S., Bockharyev, V., & Shevlyakova, A. (2020). Recognition of parts of speech using the vector of bigram frequencies. In W. van der Aalst et al. (Eds.), *Analysis of images, social networks and texts. AIST 2019, Communications in Computer and Information Science* (vol. 1086, pp. 132–142). Springer. https://doi.org/10.1007/978-3-030-39575-9_13
- Koper, M., & Schulte im Walde, S. (2016). Automatically generated affective norms of abstractness, arousal, imageability and valence for 350,000 German lemmas. *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)* (pp. 2595–2598). ELRA.
- Lin, Y., Michel, J.-B., Aiden, E. L., Orwant, J., Brockman, W., & Petrov, S. (2012). Syntactic annotations for the Google Books Ngram Corpus. In H. Li, C.-Y. Lin, M. Osborne, G. G. Lee, & J. C. Park (Eds.), *50th Annual Meeting of the Association for Computational Linguistics 2012: Proceedings of the Conference* (vol. 2, pp. 238–242). Association for Computational Linguistics.
- Linzmayr, O. (2004). *Apple confidential 2.0: The definitive history of the world's most colorful company* (2nd ed.). No Starch Press.
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems* (vol. 26, pp. 3111–3119). Curran Associates, Inc.
- Mohammad, S., Kiritchenko, S., & Zhu, X. (2013). NRC-Canada: Building the state-of-the-art in sentiment analysis of tweets. *Second Joint Conference on Lexical and Computational Semantics (SEM)* (vol. 2: Proceedings of the Seventh International Workshop on Semantic Evaluation, pp. 321–327). Association for Computational Linguistics.
- Pantel, P. (2005). Inducing ontological co-occurrence vectors. *Proceedings of the 43rd Conference of the Association for Computational Linguistics* (pp. 125–132). Association for Computational Linguistics. <https://doi.org/10.3115/1219840.1219856>
- Pennington, J., Socher, R., & Manning, C. (2014). GloVe: Global vectors for word representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1532–1543). Association for Computational Linguistics. <https://doi.org/10.3115/v1/D14-1162>
- Peters, M., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., & Zettlemoyer, L. (2018). Deep contextualized word representations. *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (vol. 1: Long Papers, pp. 2227–2237). Association for Computational Linguistics. <https://doi.org/10.18653/v1/N18-1202>

- Pexman, P. M., Diveica, V., & Binney, R. J. (2022). Social semantics: The organization and grounding of abstract concepts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 378(1870), 20210363. <https://doi.org/10.1098/rstb.2021.0363>
- Pilehvar, M. T., & Camacho-Collados, J. (2020). *Embeddings in natural language processing: Theory and advances in vector representations of meaning*. Morgan & Claypool Publishers.
- Rubenstein, H., & Goodenough, J. (1965). Contextual correlates of synonymy. *Communications of the ACM*, 8(10), 627–633. <https://doi.org/10.1145/365628.365657>
- Ryzhova, A., Ryzhova, D., & Sochenkov, I. (2021). Detection of semantic changes in Russian nouns with distributional models and grammatical features. *Komp'yuternaja Lingvistika i Intellektual'nye Tehnologii*, 20, 597–606.
- Schweizer, P., & Schweizer, R. (2004). *The Bushes: Portrait of a dynasty* (1st ed.). Doubleday.
- Tang, X. (2018). A state-of-the-art of semantic change computation. *Natural Language Engineering*, 24(5), 649–676. <https://doi.org/10.1017/S1351324918000220>
- Turney, P. D., & Pantel, P. (2010). From frequency to meaning: Vector space models of semantics. *Journal of Artificial Intelligence Research*, 37(1), 141–188. <https://doi.org/10.1613/jair.2934>
- Wang, S., Zhang, Y., Shi, W., et al. (2023). A large dataset of semantic ratings and its computational extension. *Scientific Data*, 10, 106. <https://doi.org/10.1038/s41597-023-01995-6>
- Weeds, J., Weir, D., & McCarthy, D. (2004). Characterising measures of lexical distributional similarity. *Proceedings of the 20th International Conference on Computational Linguistics* (pp. 1015–1021). COLING.
- Worth, P. (2023). Word embeddings and semantic spaces in natural language processing. *International Journal of Intelligence Science*, 13, 1–21. <https://doi.org/10.4236/ijis.2023.131001>
- Xu, Y., & Kemp, C. (2015). A computational evaluation of two laws of semantic change. *Proceedings of the 37th Annual Meeting of the Cognitive Science Society (CogSci 2015)*. Association for Computational Linguistics.