DOI 10.24249/2309-9917-2026-79-5-110-115

Н.В. Балабанова (Смоленск, Россия)

Кластерная организация семантического пространства социальной глагольной лексики английского языка с семантическим компонентом «Образование»

Аннотация: В данной статье рассматривается кластерная организация семантического пространства социальной глагольной лексики английского языка с семантическом компонентом «Образование». Для достижения цели исследования был использован метод сплошной выборки, компонентный анализ, кластерный анализ. Материалом для статьи послужили 65 глаголов современного английского языка, отобранных методом полной выборки. Версия данного кластерного анализа выполнена в компьютерной программе Statistica 5.5. Кластерный анализ учитывает два статистических параметра: «Кластерное расстояние» и «Способ группировки». По линии «кластерного расстояния» в настоящей статье используется корреляционный коэффициент «1-Pearson r». Для реализации компаративного анализа были введены несколько параметров, которые присутствовали в признаковом пространстве вне зависимости от конкретной организации признака: признаковая вариативность, плановая вариативность, ядерная вариативность уровней и вариативность оппозиций.

Ключевые слова: кластерная организация, семантическое пространство, социальная глагольная лексика, семантический компонент «Образование», выборка, кластерный анализ, компаративный анализ

N.V. Balabanova (Smolensk, Russia)

Cluster Organization of the Semantic Space of the Social Verbal Vocabulary of the English Language with the Semantic Component “Education”

Abstract: This article examines the cluster organization of the semantic space of the social verbal vocabulary of English language with the semantic component “Education”. To achieve the purpose of the study the continuous sampling method, component analysis, and cluster analysis were used. The material for the article includes 65 verbs of modern English, selected by the method of a complete sample. The version of this cluster analysis is performed in the computer program Statistica 5.5. Cluster analysis takes into account two statistical parameters: “Cluster distance” and “Grouping method”. Along the “cluster distance” line, the correlation coefficient “1-Pearson r” is used in this article. To implement the comparative analysis several parameters were introduced that were

present in the feature space regardless of the specific organization of the feature: feature variability, planned variability, nuclear variability of levels and variability of oppositions.

Key words: cluster organization, semantic space, social verbal vocabulary, semantic component “Education”, sampling, cluster analysis, comparative analysis

В нашей работе целью исследования является анализ кластерной организации семантического пространства социальной глагольной лексики английского языка с семантическим компонентом «Образование». Объектом нашего исследования является социальная глагольная лексика с семантическим компонентом «Образование». Материалом послужили 65 глаголов современного английского языка, отобранных методом полной выборки. Предметом исследования является семантическое пространство.

Семантическое пространство – это определенный набор диагностичных семантических признаков, организованных в определенную структуру за счет применения статистического метода кластерного анализа. Таким образом, семантическое пространство выделяет два аспекта: систему семантических признаков как структурных компонентов семантического пространства и кластерные отношения между этими признаками [1].

Семантические признаки выделяются при помощи метода компонентного анализа, основной целью которого является расчленение языкового значения на минимальные семантические компоненты (признаки) [2]. Применительно к нашей работе основной целью, достигаемой при помощи метода компонентного анализа, является выделение ограниченного списка (в количестве двенадцати одной единицы) репрезентативных семантических признаков глаголов социальной сферы «Образование», объединенных в семантические планы – обобщенные семантические категории в нашей трактовке, в рамках которых организуется как минимум два антонимичных семантических признака [3].

Для достижения цели исследования были использованы следующие методы:

1. метод сплошной выборки (осуществляется подбор примеров для анализа и иллюстрации теоретических положений: выписывание из оригинального (неадаптированного) текста подряд всех встречающихся в нем примеров анализируемого типа. В данном случае были выписаны все глаголы с лексическим компонентом «Образование»);
2. компонентный анализ (предполагает анализ значений языковых единиц через выделение основных компонентов значения, или сем. При изучении лексических значений часто используется анализ словарных дефиниций, т. е. толкований значений слов, которые приводятся в различных типах словарей. Эти толкования содержат в себе всю необходимую информацию о компонентном составе значения данного слова и о возможностях его использования);
3. кластерный анализ (кластерный анализ лексики – это метод, который позволяет определить содержательную близость лексических единиц при сопоставлении их синтагматических свойств (сочетаемости с другими элементами контекста). Суть метода в том, что множество контекстов употребления исследуемых лексем представляют как точки или вектора дистрибуций в N-мерном пространстве. Вычислив расстояния между точками или

сравнив вектора дистрибуций, можно получить количественную оценку тесноты семантических связей).

Для кластеризации используются результаты измерений, сохраняемые в матрице расстояний. Чем ближе синтагматические свойства лексем (а значит, чем ближе их значения), тем меньше расстояние между векторами их дистрибуций и тем больше вероятность их объединения в один кластер.

В своем анализе мы использовали 65 социальных глаголов с семантическим компонентом «Образование». В предыдущей работе мы проанализировали эти глаголы по 21 разноплановому признаку.

Полученные данные были сгруппированы в виде кластера, где каждый кластер представляет собой подмножество объектов, а расстояние между ними отображает степень их различия. Положительные и отрицательные связи между признаками были определены на основе корреляционного анализа. Если признаки находились в пределах положительного поля, то они относились к одному и тому же комплексу (кластеру). Если кластеры были соотнесены в рамках отрицательного поля, то они являлись противопоставленными и образовывали оппозицию.

По итогам анализа можно выделить 3 кластера:

1. субъекто не маркированный кластер. Признаки в этом кластере менее релевантны в этой системе. Они присутствуют, но являются субъекто не маркированными.
2. кластер субъект – ученик. Ядром этого кластера является именно этот семантический признак.
3. кластер субъект – учитель. Ядром кластера является этот признак и признак объект – ученик. Эти признаки взаимосвязаны.

Разделение на кластеры было произведено с учетом системно значимых признаков. Так как все глаголы имеют семантический компонент «Образование», то очевидно, что релевантным признаком будет являться субъект ученик и субъект учитель.

Вторым аспектом являются структурные отношения между семантическими признаками этих глаголов, которые подразумевают группировку всех исследуемых признаков в противопоставленные друг другу классы (кластеры). Группировка исследуемых признаков осуществляется при помощи метода кластерного анализа. Версия данного кластерного анализа выполнена в компьютерной программе Statistica 5.5. Кластерный анализ учитывает два статистических параметра: «Кластерное расстояние» и «Способ группировки».

«Кластерное расстояние» устанавливает бинарные расстояния между всеми элементами в рассматриваемой системе. По линии «кластерного расстояния» в настоящей статье используется корреляционный коэффициент «1-Pearson r».

Функция «правил группировки» определяет различные способы объединения элементов системы в кластеры. В нашей работе задействован вариант «невзвешенное попарное среднее», который группирует элементы (семантические признаки) в комплексы на основании средней суммы бинарных «кластерных расстояний» всех элементов комплексов [4].

В кластерной схеме на основании усредненных расстояний можно выделить три «поля»: «релевантное положительное», «нерелевантное положительное» и «отрицательное». Признаки, относящиеся к одному и тому же кластеру, группируются в рамках «нерелевантного положительного» поля и подразумевают положительную соотнесенность друг с другом. Различные кластеры противопоставляются

друг другу в пределах «отрицательного поля», и, следовательно, если признаки принадлежат разным кластерам, то они образуют оппозиции [5].

В структуре одного кластера могут выделяться «ядерные» и «периферийные» признаки. Первые группируются друг с другом в рамках «релевантного положительного» поля кластера, т. е. все признаки, относящиеся к одному и тому же «ядру», связаны друг с другом наиболее тесными (релевантными положительными) связями. В структуре кластера может быть как одно, так и несколько ядер. Периферийные признаки выявляют связь с другими признаками кластера в пределах «нерелевантного положительного» поля, т. е. данные признаки не образуют оппозиции с другими признаками этого же кластера, с одной стороны, но и не обнаруживают положительную релевантную связь ни с одним из признаков рассматриваемого кластера, с другой.

Изначально мы считаем релевантным показатель в 25. Соответственно, все, что ниже 75, – это ядро, если показатель выше, то это периферия, т. е. эти признаки относятся к этому кластеру, но находятся вне положительной связи.

Первоначально дадим общую характеристику кластерной организации семантических пространств исследуемых глаголов. Признаки мы делим на ядерные и периферийные.

Для реализации компаративного анализа нам потребовалось ввести несколько параметров, которые присутствовали в признаковом пространстве вне зависимости от конкретной организации признака:

1. Признаковая вариативность (количество признаков, входящих в состав одного кластера).
2. Плановая вариативность (количество планов, репрезентирующих один кластер).
3. Ядерная вариативность уровней (количество «ядер» в каждом кластере).
4. Вариативность оппозиций (степень противопоставленности кластеров относительно друг друга).

Исследование привело к следующим результатам.

Признаковая вариативность принимает следующие значения: 4 признака (отрицательная оценка, дивергентность, цель в ретроспективной ситуации, ретроспективность) в 1 кластере, 7 признаков (субординативность, слабая мотивация, цель в базовой ситуации, объект учитель, координативность, субъект ученик, адъект дидактический) во втором кластере, 9 признаков (субъект учитель, объект ученик, доминантность, процессуальность, сильная мотивация, проспективность, цель в проспективной ситуации, конвергентность, положительная оценка) в 3 кластере.

Плановая вариативность представлена следующим образом:

1 кластер (субъектно немаркированный)	**2 кластер (субъект – ученик)**	**3 кластер (субъект – учитель)**
аксиологический план	иерархический план	аксиологический план
интерактивный план	мотивационный план	интерактивный план
телеологический план	телеологический план	телеологический план
темпоральный план		темпоральный план
		иерархический план
		мотивационный план

Что касается ядерной вариативности, то в первом кластере все четыре признака (ретроспективность, цель в ретроспективной ситуации, дивергентность, отрицательная оценка) являются ядерными. Во втором кластере «субъект ученик» три ядерных признака (субъект ученик, субъект дидактический, координативность) и четыре периферийных признака (объект учитель, цель в базовой ситуации, слабая мотивация и субординативность). В третьем кластере восемь ядерных признаков (субъект учитель, объект ученик, доминантность, процессуальность, проспективность, цель в проспективной ситуации, конвергентность, положительная оценка) и один периферийный признак (сильная мотивация).

По параметру «вариативность оппозиций кластеров» мы установили степень противопоставленности кластеров относительно друг друга. Следует отметить, что рассматриваемый параметр учитывает следующие более частные характеристики рассматриваемой системы: количество признаков, которые участвуют в анализе, количество кластеров в системе (чем больше кластеров, тем больше оппозиций), а также степень равномерности распределения одноплановых признаков по кластерам. Отметим, что чем больше оппозиций в кластерном пространстве, тем более структурированным является признаковое пространство относительно вышеуказанных характеристик. В нашем исследовании мы учитываем эквиполентные оппозиции. Оппозиция называется эквиполентной, если один кластер обнаруживает наличие признака некоторого плана относительно прямого или переносного субъекта, а другой кластер выявляет наличие антонимичного признака того же плана относительно этого же субъекта.

Подведя итог вышесказанному, отметим, что проанализированные по вышеуказанным параметрам семантические пространства социальной глагольной лексики английского языка с семантическим компонентом «Образование» являются релевантными, что, по нашему мнению, находит отражение именно в системной организации учебного взаимодействия: противопоставленные кластеры репрезентируют типичные роли участников образовательного процесса (учитель – ученик), а выявленные оппозиции соответствуют базовым векторам коммуникативной и дидактической деятельности (проспективность / ретроспективность, доминантность / координативность, конвергентность / дивергентность). Полученные данные позволяют рассматривать кластерную модель как инструмент семантической типологизации лексики сферы «Образование». Кроме того, существует возможность применения предложенной кластерной модели в прикладных задачах: при составлении учебных словарей и тезаурусов, разработке лексических минимумов для изучающих английский язык в профессиональном контексте «Образование», а также при создании семантических разметок для корпусных исследований.

ЛИТЕРАТУРА

1. *Бордюкова А.В.* Структура разноуровневого признакового пространства отглагольных существительных в произведении Р. Брэдбери «Вино из одуванчиков» // Материалы международного молодежного научно- го форума «ЛОМОНОСОВ-2022» / Отв. ред. И.А. Алешковский, А.В. Андриянов, Е.А. Антипов, Е.И. Зимакова. М.: МАКС Пресс, 2022.
2. *Гулевич Е.В.* Роль подсознания в творческом процессе Р. Брэдбери // Вестник Мозырского государственного педагогического университета им. И.П. Шамякина. 2023. № 1(61). С. 113–118.

3. *Дедова Н.А*. Кластерный анализ семантических признаков образных метафор романа Дж. Оруэлла «1984» // Риторика. Лингвистика. 2020. № 15. С. 90–98.

4. *Сильницкий А.Г.* Кластерная организация семантического пространства полиситуативных юридических глаголов английского языка с субъектом «судья» // Известия Смоленского государственного университета. 2020. № 4(52). С. 101–116.

5. *Сильницкий А.Г.* Структура семантического пространства юридических глаголов английского языка с субъектом «Расследователь» // Известия Смоленского государственного университета. 2022. № 4(60). С. 91–105.

REFERENCES

1. Bordyukova A.V. The Structure of the Multi-Level Feature Space of Verbal Nouns in R. Bradbury's "Dandelion Wine". In: Proceedings of the International Youth Scientific Forum "LOMONOSOV-2022" / Ed. by I.A. Aleshkovsky, A.V. Andriyanov, E.A. Antipov, and E.I. Zimakova. Moscow. MAKS Press, 2022.

2. Gulevich E.V. The Role of the Subconscious in R. Bradbury's Creative Process. *Bulletin of the I.P. Shamyakin Mozyr State Pedagogical University*. 2023. No 1(61), pp. 113–118.

3. Dedova N.A. Cluster Analysis of the Semantic Features of Figurative Metaphors in George Orwell's Novel "1984". *Rhetoric. Linguistics*. 2020. No 15, pp. 90–98.

4. Silnitsky A.G. Cluster Organization of the Semantic Space of Polysituational Legal Verbs in the English Language with the Subject "Judge". *Izvestiya of Smolensk State University*. 2020. No 4(52), pp. 101–116.

5. Silnitsky A.G. Structure of the Semantic Space of Legal Verbs in the English Language with the Subject "Investigator". *Izvestiya of Smolensk State University.* 2022. No 4(60), pp. 91–105.

Сведения об авторе:

Наталья Владиславовна Балабанова,
студент
кафедра английского языка
Смоленский государственный университет

Natalia V. Balabanova,
Student
Department of English Language
Smolensk State University
andrewwin@yandex.ru