

ПРИМЕНЕНИЕ МЕТОДА КОРРЕЛЯЦИОННОГО АНАЛИЗА ПИРСОНА ДЛЯ ИССЛЕДОВАНИЯ РАЗНОУРОВНЕВЫХ ПРИЗНАКОВ АНГЛИЙСКИХ ИНСТРУМЕНТАЛЬНЫХ СУЩЕСТВИТЕЛЬНЫХ
И. Л. Петроченкова

THE USE OF PEARSON CORRELATION ANALYSIS FOR THE ANALYSIS OF HETEROGENEOUS CHARACTERISTICS OF ENGLISH INSTRUMENTAL NOUNS
I. L. Petrochenkova

Статья посвящена определению релевантных корреляций трех семантических подклассов инструментальных имен существительных – энергетического (EIN), информационного (IIN) и онтологического (OIN) с семантическими, морфематическими, словообразовательными и хронологическими признаками. На основе полученных данных демонстрируется ярко выраженная противопоставленность трех изучаемых семантических подклассов по отношению друг к другу, прослеживается история развития английской инструментальной лексики. Для выявления диагностических признаков подклассов инструментальных существительных в исследовании используется метод корреляционного анализа Пирсона в рамках компьютерной математической программы STATISTICA 6.0.

The paper is devoted to the determination of relevant correlations of three semantic subclasses of instrumental nouns – energetic, informational and ontological with their semantic, morphemic, word-formative and chronological characteristics. The obtained data show the distinct opposition between the three studied semantic subclasses, reveal the evolution of English instrumental nouns. To find out the diagnostic characteristics of instrumental nouns the author uses Pearson correlation analysis within the framework of the computer programme STATISTICA 6.0.

Ключевые слова: инструментальные существительные, энергетический, информационный, онтологический семантические подклассы, разноуровневые диагностические признаки, корреляционный анализ Пирсона, положительные, отрицательные корреляции.

Keywords: instrumental nouns, energetic, informational, ontological semantic subclasses, heterogeneous diagnostic characteristics, Pearson correlation analysis, positive, negative correlations.

Целью данного исследования является определение количественных характеристик отношений между тремя семантическими подклассами инструментальных имен существительных (энергетическим, информационным и онтологическим) и их разноуровневыми признаками (семантическими, словообразовательными, морфематическими, фонетическими и хронологическими) с помощью метода корреляционного анализа по критерию Пирсона.

Метод *корреляционного анализа Пирсона* [16, с. 84 – 91] является одним из наиболее эффективных способов определения системно-значимых отношений между языковыми признаками различных уровней и позволяет определить степень вероятности наличия/отсутствия одного признака при условии наличия/отсутствия другого, т. е. степень и характер (положительный или отрицательный) их взаимосвязи, независимо от их частотности [18; 19]. Как указывает Г. Г. Сильницкий, «коэффициенты корреляции по критерию Пирсона имеют ряд методологических преимуществ: наличие готовой формулы ошибки выборочности, различение позитивных и негативных связей, комонотонность с мерами связи; сравнительно с другими тетракорическими мерами этот коэффициент в наибольшей степени отображает среднестатистические отношения между сопоставляемыми признаками» [13, с. 25]. Следовательно, корреляционный анализ представляет собой одну из наиболее надежных статистических процедур для интерпретации соответствующих количественных данных.

Отличительной особенностью количественных методов является то, что сбор данных происходит еще до начала процедуры самого исследования. В исследуемом

явлении выделяются параметры, подлежащие измерению, которые называются переменными. Все полученные в ходе сбора данных количественные показатели оформляются в виде матрицы, которая затем исследуется с помощью одного или нескольких статистических методов. По результатам статистических исследований проводится интерпретационный анализ полученных данных. Поскольку любое исследование направлено на изучение содержания объекта и его внутренней структуры, то понятно в этой связи образное определение, данное Золтаном Дерьнеи [17]. Автор называет количественные методы «смыслом в цифрах», подразумевая, что качественные характеристики объекта имеют цифровое выражение не только в единичном ряду, но и во взаимосвязи.

На наш взгляд, основным преимуществом метода корреляционного анализа является его точность и однозначность результатов. Этот метод систематичен и легко контролируем, поскольку математический аппарат уже содержится в статистической программе. Кроме того, процесс подсчета достаточно экономичен по сравнению с подготовкой матрицы, он выполняется на компьютере с помощью пакета STATISTICA 6.0.

Описываемое исследование межуровневых соотношений признаков английских инструментальных существительных проводится в рамках коллективной темы школы количественной лингвистики под многолетним руководством профессора, доктора филологических наук Г. Г. Сильницкого [22], С. Л. Захаренкова [3], Е. В. Малец [6], А. В. Степанов [15] и др. Первоначально в количественном отношении был исследован английский глагол (С. Н. Андреев [1], О. Ю. Головинская [2], В. П. Луговской [5], Г. Г. Сильницкий

[11]). Аналогическим способом статистические методы были использованы в разноуровневом описании английского прилагательного и существительного (Л. А. Кузьмин [4], Н. В. Солодовникова [14]).

Эмпирической основой проведенного корреляционного анализа послужил корпус английских наименований инструментов [7, с. 340] (700 лексических единиц), взятых из словарных статей англоязычных словарей: The Concise Oxford Dictionary of Current English [23], Roget's Thesaurus of English Words and Phrases [20], Shorter Oxford English Dictionary on Historical Principles [21].

Каждому существительному было приписано наличие или отсутствие 161 «критериального» признака семантического, морфематического, словообразовательного и хронологического уровней, через соотношения с которыми определяется системный статус «базисных» признаков трех семантических подклассов (энергетического – EIN, информационного – IN и онтологического – OIN) [8] и всего семантического класса инструментальных существительных. Обработка и анализ данных осуществлялся при помощи программы «STATISTICA 6.0». При обработке искомой выборки мы ориентировались на опыт применения корреляционной статистики на языковом материале, зафиксированный в коллективной монографии представителей вышеупомянутой школы квантитативной лингвистики [13, с. 21 – 34]. Задаваясь уровнем статистической значимости $p = 0,05$ при указанном выше объеме выборки, мы считали значимыми коэффициенты корреляции Пирсона, для которых $r_{xy} \geq |0,06|$. При значениях, близких к 0, связь считается слабой, либо отсутствует вовсе. Таким образом, были вычислены коэффициенты корреляции выше-

упомянутых семантических подклассов инструментальных существительных с соответствующими разноуровневыми признаками.

Наиболее релевантными в рамках настоящего исследования являются следующие критериальные признаки:

Фонетические признаки: односложность/много-сложность существительного.

Морфематические признаки: морфемная членность/нечленность; наличие/отсутствие аффикса, префикса, суффикса в составе существительного.

Словообразовательные признаки: производность/не-производность существительного.

Семантические признаки: подтипы энергетических, информационных и онтологических значений [12, с. 219 – 224].

Хронологические признаки: принадлежность к староанглийскому, среднеанглийскому или новоанглийскому периоду [10].

При рассмотрении нижеприведённых примеров следует, однако, иметь в виду, что коэффициент корреляции представляет вероятностную, но не линейную меру связи двух признаков; поэтому положительный коэффициент не означает, что данные два признака совместно представлены во всех без исключения рассматриваемых случаях, равно как отрицательный коэффициент не означает их абсолютной несовместимости.

Результаты проведенного корреляционного анализа приводятся в нижеследующей таблице (статистически релевантные положительные корреляции выделены жирным шрифтом, значимые отрицательные корреляции – курсивом, нерелевантные коэффициенты обозначены прочерком) (таблица 1).

Таблица 1

Сопоставительный анализ корреляций разноуровневых признаков семантических подклассов инструментальных существительных

	<i>EIN</i>	<i>IN</i>	<i>OIN</i>
Подтипы энергетических значений			
Движение	.54	-.45	-.13
Форма	.34	-.28	-.08
Физические состояния	.20	-.18	-
Физиологические состояния	.16	-.15	-
Подтипы информационных значений			
Сенсорные	-.49	.58	-
Речь	-.16	.19	-
Семиотические	-.54	.64	-
Подтипы онтологических значений			
Наличие	-.11	-	.46
Количественные	-.09	-	.36
Невозможность	-	-	.21
Общее действие	-.19	-.09	.78
Формальные признаки			
Непроизводные нечленимые	.24	-.33	-
Непроизводные членимые	-.31	.37	-
Производные нечленимые	.11	-.11	-
Производные членимые	-	.09	-
От глагола	.15	-.13	-
От существительного	-.09	.09	-
Аффиксальные	.10	-.08	-

Префиксальные	-	-	-
Суффиксальные	.08	-	-
Свободный суффикс	.08	-	-
Связанный суффикс	.09	.47	-
Связанный префикс	-	-	.21
Свободная основа	-	-	-
Связанная основа	-.37	.45	-
-er/-or	.16	-.12	-
-ance, -ence, -ency	-	-	.16
-ion	-	-	.21
-ment	-	-	.29
Словосложение	-.17	.22	-
Фонетические признаки			
Односложность	.15	-.22	-
Многосложность	-.15	.22	-
Хронологические признаки			
Староанглийский период	.09	-.15	-
Среднеанглийский период	.14	-.20	-
Новоанглийский период	-.18	.26	-.08
16 век	.13	-.13	-
17 век	-	-	-
18 век	.09	-.08	-
19 век	-.11	.15	-.07
20 век	-.19	.27	-

Из приведённой таблицы видно, что большинство соотносимых признаков выявили статистически значимые корреляции.

В семантическом плане следует отметить, что все три семантических подкласса инструментальных существительных диагностируются положительными и отрицательными корреляциями. Так, подкласс EIN положительно сопряжен со всеми подтипами энергетических значений, кроме подтипа «местонахождение». У подкласса EIN наблюдается высокая степень корреляции с подтипом «движение» (.54): accelerator (педаль газа), bat (бита) и др., «изменение формы» (.34): auger (сверло), axe (топор) и т. д., «изменение физического состояния» (.20): boiler (кипятильник), condenser (холодильник) и т. д. и «физиологические процессы» (.16): aspirator (аспиратор), catheter (катетер) и др.

Подкласс IIN положительно коррелирует со следующими подтипами информационных значений: «семиотические» (.64): barometer (барометр), seismograph (сейсмограф) и др., «сенсорные» (.58): binoculars (бинокль), camera (камера) и др., «речь» (.19): loudspeaker (громкоговоритель), microphone (микрофон) и др. – и остается нейтральным к информационным подтипам значений: «эмоциональные», «волевые», «когнитивные», «коммуникативные».

Подкласс OIN позитивно связан со следующими подтипами онтологических значений: «общее действие» (.78): apparatus (аппарат), instrument (инструмент) и др., «наличие» (.46): mouse-trap (мышеловка), music stand (пюпитр для нот) и др., «количественные» (.36): copier (копировальный аппарат), duplicator (множительный аппарат) и др., «невозможность» (.21): brake (тормоз), manacle (наручники) и др.

На основании полученных корреляций можно сделать вывод о том, что в семантическом плане подклассы EIN, IIN и OIN положительно сопряжены с примерно одинаковым количеством семантических подтипов энергетических, информационных и онтологических значений соответственно.

В то же время мы можем говорить о ярко выраженной противопоставленности семантических подклассов EIN и IIN по отношению друг к другу в деривационно-морфематическом, фонетическом и словообразовательном планах [9, с. 36 – 38].

В деривационно-морфематическом плане подкласс EIN выявляет относительно сильные положительные корреляции с подтипами «непроизводные нечленимые» (.24): fork (вилка), knife (нож) и др. и «производные нечленимые» (.11): guillotine (гильотина), zimmer (ходунки) и др. Следует особо отметить, что при всей частотности конверсивного словообразования в английском языке в подгруппе инструментальных существительных случаи конверсивного образования существительных от глагола на нашем материале не представлены. Подкласс EIN формирует одну релевантную отрицательную корреляцию с подтипом «непроизводные членимые» (-.31), оставаясь нейтральным к подтипу «производные членимые».

Подкласс IIN положительно коррелирует с подтипами «непроизводные членимые» (.37): gastroscope (гастроскоп), semaphore (семафор) и др. и «производные членимые» (.09): amplifier (усилитель), calculator (калькулятор) и др., отрицательно коррелируя с подтипами «непроизводные нечленимые» (-.33) и «производные нечленимые» (-.11).

В фонетическом плане подкласс EIN характеризуется односложностью (.15): bat (ракетка), spade (лопата), spoon (ложка), whip (кнут, хлыст).

Напротив, в фонетическом плане подкласс IIN характеризуются многосложностью (.22): photocoagulator (светокоагулятор), radio-telephone (радиотелефон), spectroheliосore (спектрогелиоскоп) и др.

В словообразовательном плане EIN положительно соотносены с «аффиксацией» (.10), в частности, с «суффиксацией» (.08), в то время как IIN отрицательно коррелируют с аффиксальным способом словообразования в целом (-.08), но позитивно связаны со «словосложением» (.22) (Ср.: EIN – (-.17)). Ярко выражена положительная корреляция EIN с суффиксом -er (.16) (Ср.: IIN – (-.12)), в то время как OIN характеризуются статистически релевантными корреляциями с суффиксами -ment (.29), -ion (.21) и -ance, -ence, -ency (.16).

Таким образом, подкласс EIN характеризуется простотой формальной структуры:

- фонетической (односложность: .15);
- морфематической (морфемная нечленимость: .24).

Однако, для EIN характерно нестандартное сочетание морфемной нечленимости и отглагольной аффиксальной производности (.10), а также тяготение к

аффиксации (.08), в частности, к суффиксу -er / -or (.16).

IIN, в противоположность EIN, характеризуются сложной формальной структурой:

- фонетической (многосложность: .22);
- морфематической (морфемная членимость: .37), но в словообразовательном плане – аффиксальной непроезвучностью (-.08) и тяготением к словообразованию (.22). Кроме того, для IIN характерна производность не от глагольной, а от субстантивной производящей основы (.09) и сильная положительная корреляция со связанными суффиксами (.47) и связанными основами (.45).

Подкласс OIN одинаково нейтрален относительно фонетических и морфематических признаков. Однако в словообразовательном плане OIN представляют собой единственный класс, положительно соотносённый со связанными префиксами (.21) и суффиксами -ance, -ence, -ency (.16), -ion (.21), -ment (.29).

На основании вышеприведенных данных могут быть выделены следующие релевантные корреляционные признаки, определяющие системные отношения трех рассматриваемых семантических подклассов инструментальных существительных (таблица 2).

Таблица 2

Релевантные корреляционные признаки семантических подклассов EIN, IIN, OIN

Семантические подклассы	Формальные признаки	Формальная простота	Производность	Отглагольная деривация	Отсубстантивная деривация
EIN		+	-	+	-
IIN		-	+	-	+
OIN		0	0	0	0

Примечания: + – положительно релевантный признак, - – отрицательно релевантный признак, 0 – нерелевантный признак.

«Диагностическим» для каждого семантического подкласса инструментальных существительных будет считаться такой релевантный признак, по которому данный подкласс противопоставлен двум остальным.

Диагностические признаки EIN: формальная простота (односложность, нечленимость), непроезвучность, отглагольная деривация.

Диагностические признаки IIN: формальная сложность (многосложность, членимость), производность, отсубстантивная деривация.

Из вышесказанного следует, что наибольшей диагностической силой обладают семантические подклассы EIN и IIN.

В хронологическом плане [9, с. 39 – 40] контрастность двух основных семантических подклассов по отношению друг к другу также ярко выражена, хотя хронологические признаки более диагностически значимы для энергетических существительных и менее значимы – для информационных. Например, подкласс EIN отмечает наивысшими положительными коэффициентами *среднеанглийский* (.14): bodkin (шило), colander (дуршлаг), scissors (ножницы), spoon (ложка) и др. – и *староанглийский* периоды (.09): hammer (мо-

лоток), mattock (киркомотыга), rake (грабли) и др., отрицательно коррелируя с *новоанглийским периодом* (-.18) в целом, но положительно отмечая шестнадцатый (.13) и восемнадцатый (.09) века новоанглийского периода, в то время как подкласс IIN позитивно связан с *новоанглийским периодом* (.26), а именно с девятнадцатым столетием (.15) и особенно с двадцатым веком (.27): computer (компьютер), laparoscope (лапароскоп), scanner (сканер) и др., формируя отрицательные корреляции со староанглийским (-.15) и среднеанглийским периодами (-.20). Онтологические существительные определяются здесь только слабыми негативными зависимостями.

Итак, исходя из вышеприведенных данных, можно выделить следующие релевантные хронологические признаки семантических подклассов инструментальных существительных (таблица 3).

Релевантные хронологические признаки EIN, IIN, OIN

Хронологические признаки	Старо-английский период	Средне-английский период	Ново-английский период	XVI век	XVII век	XVIII век	XIX век	XX век
EIN	+	+	-	+	0	+	-	-
IIN	-	-	+	-	0	-	+	+
OIN	0	0	-	0	0	0	-	0

Диагностические хронологические признаки EIN: староанглийский, среднеанглийский периоды и XVI, XVIII века новоанглийского периода.

Диагностические хронологические признаки IIN: XIX – XX века новоанглийского периода.

Таким образом, корреляционный анализ, проведенный для сопоставления трех семантических подклассов инструментальных существительных с их разноуровневыми признаками позволил установить наличие статистически значимых зависимостей меж-

ду ними. Полученные корреляции позволили прояснить особую роль энергетических и информационных инструментальных существительных, которые отчетливо противопоставлены друг другу в семантическом, деривационно-морфематическом, словообразовательном и хронологическом планах и сильно диагностируются разноуровневыми и разноаспектными признаками. Онтологические существительные характеризуются минимальным количеством релевантных корреляций.

Литература

1. Андреев С. Н. Соотношение семантических, морфематических и синтаксических признаков глаголов речи современного английского языка // Проблемы соотношения словообразования с другими уровнями в германских языках. 1985. С. 22 – 47.
2. Головинская О. Ю. Семантические характеристики производных аффиксальных глаголов в современном английском языке // Проблемы соотношения словообразования с другими уровнями в германских языках. 1985. С. 102 – 111.
3. Захаренкова С. Л. Семантическая структура существительных с эмоциональным значением в английском языке // Известия Смоленского государственного университета. 2013. № 1(21). С. 142 – 152.
4. Кузьмин Л. А. Квантитативные критерии разграничения качественных и относительных прилагательных в современном английском языке // Актуальные проблемы романистики и германистики. Ч. II. Германистика. 1998. С. 3 – 11.
5. Луговской В. П. Словообразовательные характеристики простых глаголов английского языка: автореф. дис. ... канд. филол. наук: 10.02.04. Минск, 1986. 22 с.
6. Малец Е. В. Семантическая система информационных значений имен речи // Известия Смоленского государственного университета. 2013. № 3(23). С. 137 – 145.
7. Петроченкова И. Л. Применение методов многомерного анализа в исследовании инструментальных существительных английского языка // Вестник Брянского государственного университета. 2011. № 2.
8. Петроченкова И. Л. Семантическая классификация английских «инструментальных» существительных // Вестник Бурятского государственного университета. 2008. № 11. С. 137 – 141.
9. Петроченкова И. Л. Системный анализ разноуровневых признаков инструментальных существительных английского языка. Смоленск, 2012. 120 с.
10. Петроченкова И. Л. Хронологические характеристики инструментальных существительных в английском языке // Девятые Поливановские чтения. 2009. Ч. I. С. 49 – 53.
11. Сильницкий Г. Г. К проблеме соотношения семантических и формальных признаков английских глаголов // Семантика английского глагола в соотношении с признаками других уровней. 1988. С. 8 – 22.
12. Сильницкий Г. Г. Семантика. Грамматика. Квантитативная и типологическая лингвистика. Смоленск, 2006. Т. 2. 362 с.
13. Сильницкий Г. Г., Андреев С. Н., Кузьмин Л. А., Кусков М. И. Соотношение глагольных признаков различных уровней в английском языке. Минск, 1990. 182 с.
14. Солодовникова Н. В. Формальное и семантическое взаимодействие прилагательного с деадъективными знаменательными частями речи в современном английском языке: дис. ... канд. филол. наук: 10.02.04. Смоленск, 2007. 145 с.
15. Степанов А. В. Межуровневые корреляции характеристик «конъектных» существительных с пространственным значением в современном английском языке // Известия Смоленского государственного университета. 2013. № 1(21). С. 152 – 162.
16. Тулдава Ю. А. Проблемы квантитативных исследований текста и языка // Прикладная лингвистика и автоматический анализ текста. 1988.

17. Dornyei Z. Research Methods in Applied Linguistics. OUP, 2007. 230 p.
18. Pearson K. On the General Theory of Skew Correlation and Non-Linear Regression. L., Dulau and Co., 1905. – 68 p.
19. Pearson K. The Grammar of Science. 2 ed. L., Walter Scott, 1900. 655 p.
20. Roget P. M. Roget's Thesaurus of English Words and Phrases. Penguin Books, 2004. 824 p.
21. Shorter Oxford English Dictionary on Historical Principles. 5th ed. Oxford, 2002. 3750 p.
22. Silnitsky G. G. Correlation of Phonetic and Morphological Systems of Indo-European Languages // Journal of Quantitative Linguistics. 2003. Vol. 10. № 2. С. 129 – 141.
23. The Concise Oxford Dictionary of Current English / Ed. by Judy Pearsall. 10th ed. Oxford, 1999. 1666 p.

Информация об авторе:

Петроченкова Инна Леонидовна – кандидат филологических наук, доцент кафедры иностранных языков Смоленского филиала Московского государственного университета путей сообщения, petrocinna@yandex.ru.

Inna L. Petrochenkova – Candidate of Philology, Assistant Professor at the Department of Foreign Languages, Moscow State University of Railway Engineering (Smolensk Branch).

Статья поступила в редколлегию 01.10.2014 г.