DOI: https://doi.org/10.18454/RULB.2021.25.1.13

СЕМАНТИЧЕСКАЯ КЛАСТЕРИЗАЦИЯ И КОМБИНАТОРНОСТЬ АФФИКСОВ АНГЛИЙСКОГО ЯЗЫКА

Научная статья

Дмитриева Е.И. *
ORCID: 0000-0003-4421-8602,
Российский университет транспорта, Москва, Россия

* Корреспондирующий автор (eidmitrieva[at]mail.ru)

Аннотация

Статья посвящена применению метода кластерного анализа для исследования семантических особенностей аффиксальных комбинаций на материале современного английского языка. Использованы следующие методы: изучение и анализ научной литературы, сравнение, кластерный анализ, структурный анализ. В ходе работы выявлено 17 семантических кластеров английского языка. Далее в работе кластерный анализ был использован применительно к выявлению закономерностей возможных комбинаций аффиксов со значением качества Классифицированы поликластерные, смешанные и невключенные сочетания кластеров. Установлено, что наибольшей комбинаторностью обладают аффиксы с ярко выраженной семантикой качества.

Ключевые слова: аффикс, комбинаторность, кластерный анализ, семантика.

SEMANTIC CLUSTERING AND COMBINATIVITY OF AFFIXES IN THE ENGLISH LANGUAGE

Research article

Dmitrieva E.I.*
ORCID: 0000-0003-4421-8602,
Russian University of Transport, Moscow, Russia

* Corresponding author (eidmitrieva[at]mail.ru)

Abstract

The article discusses the application of the cluster analysis method to study the semantic features of affixal combinations based on the modern English language. The author of the study applies the following methods of research: the analysis of scientific literature, method of comparison, cluster analysis, and structural analysis. 17 semantic clusters of the English language were identified in the course of the study. Further, the author used the cluster analysis to identify patterns of possible combinations of qualitative affixes. Polycluster, mixed and unincorporated combinations of clusters were classified. It is established that affixes with a pronounced meaning of quality are the most combinative.

Keywords: affix, combinativity, cluster analysis, semantics.

Introduction

The purpose of this article is to study the semantic features of affixal combinations in the English language with the help of the cluster analysis method. Cluster analysis is the process of statistical data collection, which contains information about the selection of subjects for analysis and subsequent distribution into relatively homogeneous groups. A detailed description of the method of structural cluster analysis in word formation and a description of the semantic features of English affixes is featured in the author's earlier articles [3], [4].

Methods

The semantic clusters were formed based on the presence of at least one common semantic element in two or more affixes. The results are presented in the tables below.

Let us cluster the suffixes. Since the semantics of affixes of different parts of speech differ significantly, it is advisable to examine suffixes and prefixes separately in each case.

First, we perform the clustering of noun suffixes:

Table 1 – Noun suffixes

№	Suffix	Abstraction	Condition\Quality	Totality\Generality
1	ness	1	0	0
2	ity	1	0	0
4	ship	0	1	0
5	dom	0	0	1
6	hood	0	1	0
7	ation	0	0	0
8	ment	0	0	0
9	ery	0	1	1
10	acy	0	1	1
11	age	0	1	1
Cluster		A	B	C

As a result, 3 clusters were identified:

A is represented by the suffixes -ness, -ity with the meaning of abstraction
B is represented by the suffixes -ship, -hood, -ery, -acy, -age with the meaning of condition\quality
C is represented by the suffixes -dom, -ery, -acy, -age with the meaning of totality\generality.
Next, the clustering of the suffixes of adjectives is carried out:

Table 2 – Adjective suffixes

No.	Suffix	Presence of an attribute, quality	Propensity	Weakening of quality	Similarity	Completeness	Affiliation
1	ed	1	0	0	0	0	0
2	y	1	1	1	0	0	0
3	ish	1	0	1	0	0	1
6	ly	1	0	0	1	0	0
7	ful	0	0	0	0	1	0
8	some	1	1	0	0	1	0
10	like	0	0	0	1	0	0
12	ous	1	0	0	0	0	0
13	an	0	0	0	1	0	1
Cluster		D	F	G	H	I	J

Thus, at the second stage, 6 clusters were identified:
D is represented by the suffixes -ed, -y, -ish, -ly, -some, -ous with the meaning of the presence of an attribute or quality
F is represented by the suffixes -y, -some with the meaning of propensity
G is represented by the suffixes -y, -ish with the meaning of quality attenuation
H is represented by the suffixes -ly, -like, -an with the meaning of similarity
I is represented by the suffixes -ful, -some with the meaning of completeness
J is represented by the suffixes -ish, -an with the meaning of affiliation.
At the third stage, we conduct the clustering of verb suffixes:

Table 3 – Verb suffixes

No.	Suffix	Activity	Change	Transformation
1	ize	1	1	1
2	fy	1	1	1
3	ate	0	0	1
4	en	0	1	0
Cluster		K	L	M

At the third stage, 3 clusters were identified:
K is represented by the suffixes — ize, -fy with the meaning of activity
L-is represented by the suffixes -ize, -fy, — en with the meaning of change
M-is represented by the suffixes -ize, -fy, -ate with the meaning of transformation.
Next, the prefixes were clustered.
At the first stage, we perform the clustering of the adjective prefixes:

Table 4 – Adjective prefixes

Prefix	Negation
un	1
in	1
non	1
a	1
self	0
pre	0
post	0
Cluster	N

At this stage, only one cluster was found, N, that signifies negation and is represented by the prefixes un -, in -, non -, and a-.

At the second stage, the clustering of verb prefixes was carried out.

Table 5 – Verb prefixes

Prefix	Reverse activity	Deprivation, Disposal	Disadvantage, Absence	Direction of movement
un	1	0	0	0
de	1	1	0	0
dis	1	1	0	0
mis	0	0	1	0
under	0	0	1	1
over	0	0	0	0
up	0	0	0	1
re	0	0	0	0
be	0	1	0	1
Cluster	O	P	Q	R

Key results

At this stage, 4 clusters were identified:

O is represented by the prefixes un-, de-, dis- with the meaning of the reverse activity

P is represented by the prefixes de-, dis-, be- with the meaning of deprivation, disposal

Q is represented by the prefixes mis-, under- with the meaning of disadvantage

R is represented by the prefixes under-, over-, up-, and be- with the meaning of direction of movement.

17 semantic clusters have been identified in total:

1. A is represented by the suffixes -ness, -ity with the meaning of abstraction
2. B is represented by the suffixes -ship, -hood, -ery, -acy, -age with the meaning of condition\quality
3. C is represented by the suffixes -dom, -ery, -acy, -age with the meaning of totality\generality.
4. D is represented by the suffixes -ed, — y, — ish, -ly, -some, — ous with the meaning of the presence of an attribute or quality
5. F is represented by the suffixes – y, -some with the meaning of propensity
6. G is represented by the suffixes -y, -ish with the meaning of quality attenuation
7. H is represented by the suffixes -ly, -like, -an with the meaning of similarity
8. I is represented by the suffixes -ful, -some with the meaning of completeness
9. J is represented by the suffixes -ish, -an with the meaning of affiliation.
10. K is represented by the suffixes — ize, -fy with the meaning of activity
11. L is represented by the suffixes -ize, -fy, -en with the meaning of change
12. M is represented by the suffixes -ize, -fy, -ate with the meaning of transformation.
13. N is represented by the prefixes un-, in-, non-, a- with the meaning of negation
14. O is represented by the prefixes un -, de -, -dis with the meaning of the reverse activity
15. P is represented by the prefixes de-, dis-, be- with the meaning of deprivation, disposal
16. Q is represented by the prefixes mis-, under- with the meaning of disadvantage
17. R is represented by the prefixes under-, over-, up-, and be- with the meaning of direction.

In the course of the research, the authors found the cases of intersection of semantic clusters. These include cases where two or more affixes are included to two or more intersecting clusters. The intersections are presented in the diagrams below. The central element of the diagram shows the affixes at the intersection, the two side elements show the intersecting clusters:

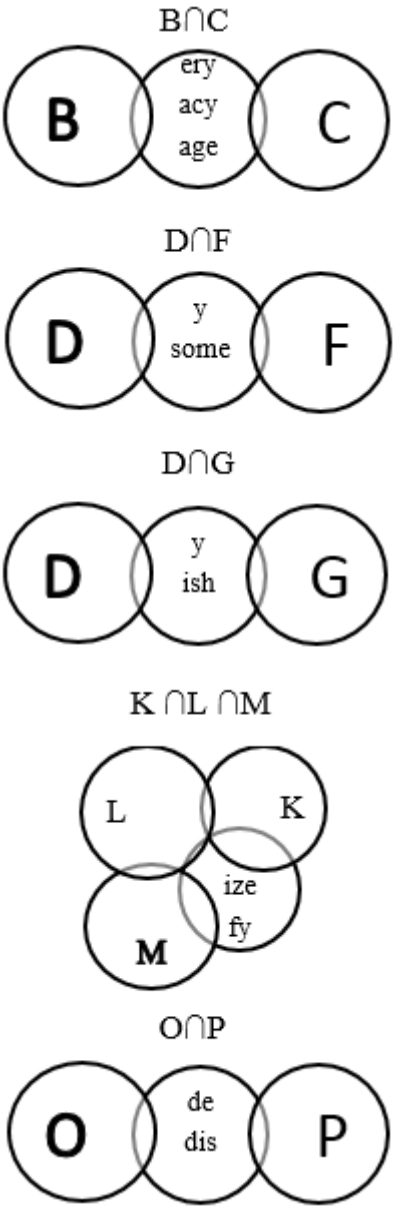

Figure 1 – cluster intersections

The analysis of the data above shows that 80% of intersections are observed among the semantic clusters of suffixes.

Later in the paper, the cluster analysis will be used to identify the patterns of possible combinations of affixes with the meaning of quality.

The affixal combinations obtained in the course of the analysis of practical language material, which include at least one qualitative affix, were analyzed in the context of the previously conducted semantic cluster analysis. For greater clarity, we first list the identified combinations: *able+ity, al+dom, al+ize+ation , ation+al, de _____ing, de_____ation , de_____ment , dis_____ed , en_____ed , en_____ing , fy+(c)+ation , ible+ity , il_____al, im_____al, in_____able, in_____able+ly , ing+ness , ir_____ant, ir_____able , ir_____al , ize+ation , ly+hood , ly+ness , ment+al, ment+ation, ment+ing , mis_____ation , mis_____ed , mis_____ing, non_____ate+ion, non_____fy+ed, non_____ing , ous+ly , some+ly, un _____ing, under_____ _ _ ed.*

The formulas of affixal combinations, taking into account the semantic cluster generalization, are presented below. If an affix is not included in the combination is not included in any of the selected clusters, it is marked as an unincorporated affix (NA): *NA+NA , NA+C, NA+M+NA , NA+NA , P _____NA, O_____NA, P_____NA , O_____D , NA_____D, NA_____NA, L+(c)+NA , M+A , NA_____NA , NA_____NA, N_____NA, N_____NA+H, NA+NA , NA _____NA , NA_____NA, NA_____NA , M+NA , D+B, H+A , NA+NA, NA+NA, NA+NA, Q_____NA, Q_____D, Q_____NA, N_____NA, N_____K+D, N_____NA, D+H, F+H, O _____NA, R_____D.*

The results obtained are presented in the diagram. Semantic clusters are marked with circles. The rectangles represent unincorporated affixes. Blue lines indicate connections between the clusters, red lines indicate connections between the clusters and unincorporated affixes, and yellow lines indicate connections between the unincorporated affixes.

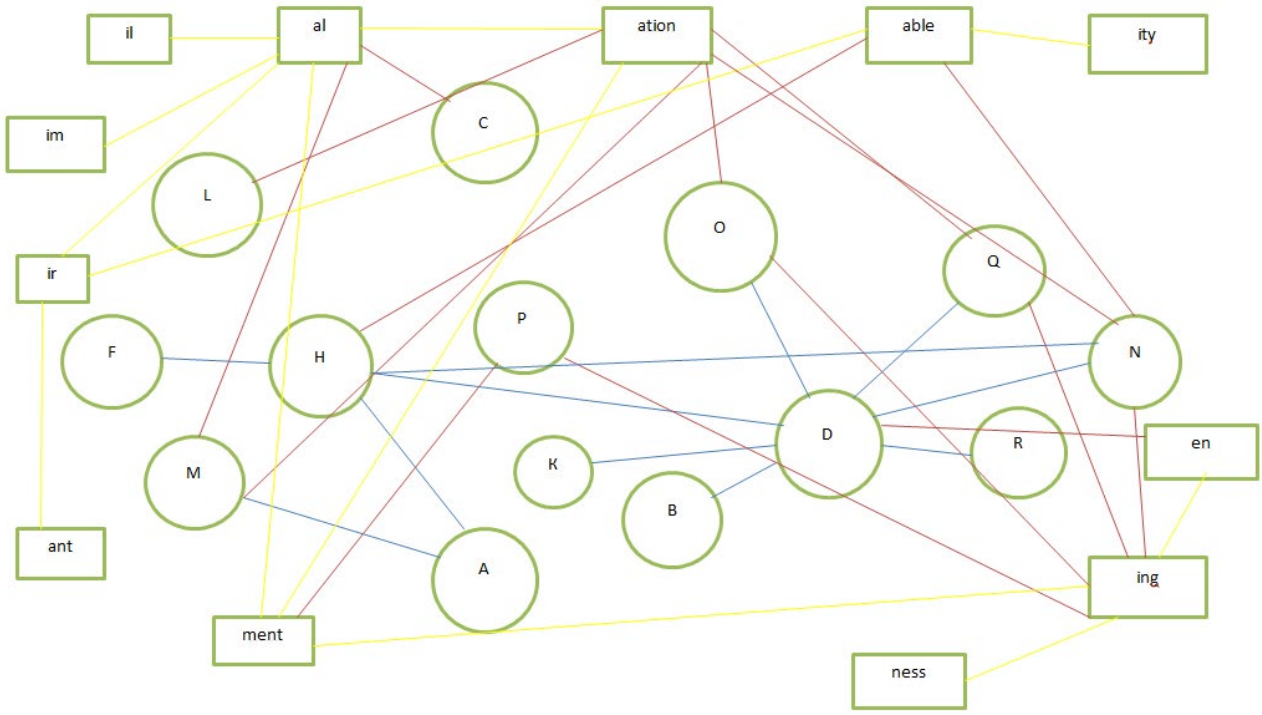

Figure 2 – compatibility of semantic clusters

When we talk about the application of cluster analysis of affix semantics to the combinative possibilities of qualitative affixes, we can distinguish 3 groups of combinations: polycluster, mixed, and unincorporated. In this case there are no monocluster combinations.

Polycluster combinations include:

1. O_____D
2. M+A
3. D+B
4. H+A
5. Q_____D
6. N_____K+D
7. D+H
8. F+H
9. R_____D

Mixed combinations include:

1. NA+C
2. NA+M+NA
3. P _____NA
4. O_____NA
5. P_____NA
6. NA_____D
7. L+(c)+NA
8. N_____NA
9. N_____NA+H
10. M+NA
11. Q_____NA
12. Q_____NA
13. N_____NA
14. N_____NA
15. O _____NA

Unincorporated combinations include:

1. able+ity
2. ation+al
3. en_____ing
4. il_____al
5. im_____al
6. ing+ness
7. ir _____ant

8. ir_____able
9. ir_____al
10. ment+al
11. ment+ation
12. ment+ing

Let us consider the first group in more detail since it is of the greatest interest from the point of view of establishing patterns in the compatibility of affixes. During the pairwise analysis, it is established which clusters and meanings are included in the combinations

1. O_____D
reverse activity+quality
2. M+A
transformation+abstraction
3. D+B
quality+condition
4. H+A
similarity+abstraction
5. Q_____D
disadvantage+quality
6. N_____K+D
negation+activity+quality
7. D+H
quality+similarity
8. F+H
propensity+similarity
9. R_____D
direction+quality

Conclusion

In the course of the research, 17 semantic clusters of the English language were identified in total with some of the clusters having the property of overlapping. Based on the analysis of the language material, it can be concluded that the affixes with a pronounced semantic attribute of quality are the most combinative (cluster D – 66% of semantic polycluster combinations).

Конфликт интересов

Не указан.

Conflict of Interest

None declared.

Список литературы / References

1. Архипов И.К. Семантика производного слова английского языка / Архипов И.К. – Москва: Просвещение, 1984
2. Гридина Т.А. Современный русский язык. Словообразование: теория, алгоритмы анализа, тренинг / Гридина Т.А., Коновалова Н.И. Изд. 3. 2009.
3. Дмитриева Е.И. Семантическая характеристика аффиксов английского языка / Дмитриева Е.И. // Общественные науки. 2017. № 1. С. 104-112.
4. Дмитриева Е.И. Кластерный анализ в исследовании аффиксального словообразования (на материале английского языка) / Дмитриева Е.И., Телегин Л.А. // Филологические науки. Вопросы теории и практики. 2020. Т. 13. № 5. С. 183-189.
5. Зятковская Р.Г. Суффиксальная система современного английского языка. Структурные аспекты слова и словосочетания / Зятковская Р.Г. – Калининград: КГУ, 1980
6. Каращук П. М. Словообразование английского языка / Каращук П. М. – М.: Высшая школа, 1977. 314 с.
7. Кубрякова Е.С. Теория номинации и словообразование / Кубрякова Е.С. – М.: Либроком, 2012.
8. Мешков О. Д. Словообразование современного английского языка / Мешков О. Д. – М.: Наука, 1975. 248 с.
9. Defays D. An efficient algorithm for a complete link method / Defays D. // The Computer Journal. 1977. Vol. 20. № 4. P. 364-366.
10. Rand W. M. Objective criteria for the evaluation of clustering methods / Rand W. M. // Journal of the American Statistical Association. 1971. Vol. 66. Issue 336. P. 846-850.

Список литературы на английском / References in English

1. Arkhipov I. K. Semantika proizvodnogo slova anglijjskogo jazyka [Semantics of a Derived Word in English] / I. K. Arkhipov. – Moscow: Prosveshchenie, 1984 [in Russian]
2. Gridina T. A. Sovremennyjj russkijj jazyk. Slovoobrazovanie: teorija, algoritmy analiza, trening [Modern Russian Language. Word Formation: Theory, Analysis Algorithms, Training] / T. A. Gridina, N. I. Konovalova. 3rd Edition. 2009 [in Russian]
3. Dmitrieva E. I. Semanticheskaja kharakteristika affiksov anglijjskogo jazyka Obshhestvennye nauki [Semantic Characteristics of English Affixes Social Sciences] / E. I. Dmitrieva. 2017. № 1, pp. 104-112 [in Russian]
4. Dmitrieva E. I. Klasternyjj analiz v issledovanii affiksal'nogo slovoobrazovanija (na materiale anglijjskogo jazyka) [Cluster Analysis in the Study of Affixal Word Formation (Based on the Material of the English Language)] / E. I. Dmitrieva, L. A. Telegin // [Philological Sciences. Theory & Practice]. 2020. Vol. 13. No. 5, pp. 183-189 [in Russian]

5. Zyatkovskaya R. G. Suffixalnaya sistema sovremennogo angliyskogo yazyka [The suffixal system of the modern English language] / R. G. Zyatkovskaya. – Kaliningrad: Kaliningrad State University, 1980 [in Russian]

6. Karashchuk P. M. Slovoobrazovanie anglijjskogo jazyka [Word Formation of the English Language] / P. M. Karashchuk. – Moscow: Vysshaya shkola, 1977. 314 p. [in Russian]

7. Kubryakova E. S. Teorija nominacii i slovoobrazovanie [Theory of Nomination and Word Formation] / E. S. Kubryakova. – Moscow: Librocom, 2012 [in Russian]

8. 8. Meshkov O. D. Slovoobrazovanie sovremennogo anglijjskogo jazyka [Word Formation of the Modern English Language] / O. D. Meshkov. – Moscow: Nauka, 1975. 248 p. [in Russian]

9. Defays D. An efficient algorithm for a complete link method / Defays D. // The Computer Journal. 1977. Vol. 20. № 4. P. 364-366.

10. Rand W. M. Objective criteria for the evaluation of clustering methods / Rand W. M. // Journal of the American Statistical Association. 1971. Vol. 66. Issue 336. P. 846-850.