

КЛАССИФИКАЦИЯ СТИХОТВОРНЫХ ТЕКСТОВ УИТЬЕРА

Статья посвящена классификации индивидуального стиля Дж. Уитъера путем сопоставления его текстов с произведениями авторов, сформировавшими основные тенденции в американской литературе XIX века. Исследование ведется на базе широкой парадигмы разноуровневых признаков при помощи дискриминантного анализа.

В результате определена степень сходства текстов Уитъера с критериальными классами текстов, выявлены языковые характеристики, позволяющие с высокой степенью точности дискриминировать стихотворные тексты различных авторов.

Важной задачей таксономических исследований является классификация объектов в системе заданных классов (сформированных априорно или на основании предыдущего опыта). В лингвистике решение данной таксономической задачи становится необходимой частью исследований по определению статуса языковых единиц, атрибуции текстов, сопоставлению идиостилей авторов и др.

Целью данной статьи является классификация идиостиля Дж. Уитьера — одного из наиболее известных американских поэтов середины и второй половины XIX века — путем сопоставления с идиостильями авторов, сформировавшими к тому времени основные тенденции в американской литературе. Материалом исследования служили выборки стихотворных текстов американских поэтов В. Брайента, Г. Лонгфелло, Р. Эмерсона и Э. По, заложивших основу американской стихотворной традиции, а также текстов Дж. Уитьера. Общий объем выборки составил более 4800 строк. Для анализа брались только лирические произведения всех указанных авторов, так как именно в лирике в наиболее полной мере проявляется своеобразие творческой манеры поэта.

Произведения больших жанров, такие как поэмы, драмы к анализу не привлекались. Также исключались из анализа переводные стихотворения и сонеты. В первом случае требовалось бы определить, насколько переводы вписываются в общую картину оригинального творчества. Сонеты не привлекались к анализу в связи с тем, что они обладают жесткими схемами структурного построения (рифмовки и строфической организации) и могут несколько деформировать пропорции строфического и рифменного материала.

В произведениях указанных авторов используются различные меры — метры и размеры. В данном исследовании анализируются только стихотворения, написанные ямбом (различными размерами). Это

объясняется, с одной стороны, тем, что в подавляющем большинстве случаев американскими романтиками использовался именно ямб. С другой стороны, такое метрическое сходство позволяет избежать ввода дополнительных параметров при сопоставлении текстов.

Выбор стихотворного текста в качестве материала исследования обусловлен тем, что он обладает особым способом структурной организации по сравнению с прозаическим текстом — повышенной упорядоченностью, что, с одной стороны, налагает на него определенные ограничения, а с другой — приводит к значительно более сильно выраженной соотнесенности составляющих его единиц. Стихотворный текст обладает внутренней мерой — метром. Он дробится на четко заданные, сопоставимые между собой единицы (строки, строфы). Соизмеримость (сопоставимость) стиховых отрезков соотносит образующие их слова не только по горизонтали, как в прозе, но и по вертикали¹. Налагаемые на стихотворный текст формальные ограничения, взаимодействуя с чисто языковыми закономерностями, ограничивают возможности автора в организации текста.

Сложная, многоаспектная природа объекта исследования обусловила необходимость привлечения статистических методов, позволяющих одновременно учесть большое количество варьирующих в разной степени признаков и эксплицировать не отмечаемые при непосредственном наблюдении тенденции.

Методы математической статистики широко применяются в ряде эмпирических наук (биологии, социологии, экономике и др.), позволяя получать ценные результаты. Имеется также успешный опыт применения математических процедур в лингвистике на языковом и текстовом материале².

Для классификации нами использовался метод дискриминантного анализа. Этот вид анализа является одной из про-

цедур многомерного анализа данных и позволяет, во-первых, ответить на вопрос о степени сходства (близости) нескольких групп объектов с выявлением признаков, дифференцирующих эти классы и, во-вторых, определить, к какой из изучаемых групп относится спорный (неясный) случай³.

Каждое исследуемое произведение было описано с помощью 43 фонетических, морфологических, синтаксических, рифменных и ритмических признаков. За основу были взяты характеристики, описанные в работах В. С. Баевского⁴, однако добавлен целый ряд языковых признаков.

На первом этапе исследования⁵ сопоставлялись тексты Брайента, Эмерсона, Лонгфелло и По, и на их базе определялись признаки, релевантные для разграничения идиостилей авторов. На втором этапе полученная признаковая модель использовалась для определения степени сходства идиостиля Уитьера и рассмотренных четырех авторов.

В результате анализа классов текстов авторов (критериальных классов) выявлена высокая дифференциальная значимость признаковой парадигмы. Из 43 привлекаемых к исследованию характеристик 80% оказались значимыми для их дифференциации:

- количество слогов в слове, замещающем первую и последнюю сильную позицию в строке;
- количество существительных, замещающих первую сильную позицию в строке;
- количество глаголов, замещающих первую и последнюю сильную позицию в строке;
- количество прилагательных, замещающих первую и последнюю сильную позицию в строке;
- количество наречий, замещающих первую и последнюю сильную позицию в строке;

- количество местоимений, замещающих первую и последнюю сильную позицию в строке;
- количество подлежащих, замещающих первую и последнюю сильную позицию в строке;
- количество сказуемых в первой сильной позиции в строке;
- количество дополнений, замещающих последнюю сильную позицию в строке;
- количество определений, замещающих последнюю сильную позицию в строке;
- количество обстоятельств, замещающих первую сильную позицию в строке;
- количество в первой сильной позиции в строке слов, не являющихся членами предложения;
- количество простых предложений в составе сложноподчиненного;
- количество простых предложений в составе сложносочиненного;
- количество стилистически маркированных строк;
- количество переносов;
- количество строк, разорванных синтаксической паузой;
- количество полных и неполных инверсий;
- количество точных рифм;
- количество мужских рифм;
- средняя длина строки в стопах, т. е. усредненный размер, которым написаны строки;
- количество видов размеров, задействованных в произведении;
- количество слогов в анакруссе;
- количество сверхсхемных ударений на анакруссе;
- количество пропусков первого икта;
- количество пропусков второго икта;
- количество строф;
- количество видов строф.

Таблица 1

Расстояние между центроидами классов

Классы текстов	Класс 1 (Брайент)	Класс 2 (Лонгфелло)	Класс 3 (Эмерсон)	Класс 4 (По)
Класс 1 (Брайент)	0,00	9,11	23,39	29,06
Класс 2 (Лонгфелло)	9,11	0,00	16,15	19,61
Класс 3 (Эмерсон)	23,39	16,15	0,00	22,90
Класс 4 (По)	29,06	19,61	22,90	0,00

В полученном признаковом пространстве определены расстояния (мера Махаланобиса) между центроидами четырех классов текстов (табл. 1). Под *центроидом* понимается точка в пространстве с координатами, являющимися средними значениями переменных всех объектов в данном классе.

Проверка на статистическую значимость результатов производится при помощи показателя уровня значимости⁶. Значения этого показателя приводятся в табл. 2.

Как видно из табл. 2, уровень значимости во всех случаях удовлетворяет наиболее жестким критериям ($p = 0,00001$).

Таблица 2

Уровень значимости

Классы текстов	Класс 1 (Брайент)	Класс 2 (Лонгфелло)	Класс 3 (Эмерсон)	Класс 4 (По)
Класс 1 (Брайент)	x	0,00	0,00	0,00
Класс 2 (Лонгфелло)	0,00	x	0,00	0,00
Класс 3 (Эмерсон)	0,00	0,00	x	0,00
Класс 4 (По)	0,00	0,00	0,00	x

Кроме того, диагностическую силу выделенной модели можно проверить, проведя разбиение текстов на группы, основываясь только на вошедших в модель признаках.

В табл. 3 приводятся данные о распределении произведений четырех классов в признаковом пространстве, образованном признаками модели. По горизонтали указаны классы, согласно естественной группировке (принадлежности текстов данному автору). По вертикали классы даны так, как они образуются в результате группировки текстов в пространстве признаков модели. Произведения рассматриваются как принадлежащие тому классу, расстояние до центроида которого минимально.

Сопоставление полученной группировки с исходной классификацией позволяет определить степень диагностической точности модели в целом. Эффективность модели тем выше, чем более доля правильно расклассифицированных случаев превышает вероятность случайного отнесения произведения в свой класс (такая вероятность в данном случае составляет 25%). Как видно из табл. 3, выводимая по признакам модели классификация в очень большой степени соответствует естественной группировке – модель классифицирует правильно более 95% всех произведений, т. е. в пространстве вошедших в модель признаков почти 100% текстов расположены ближе к центроиду своего клас-

са, чем к центроидам чужих классов. Превышение почти в четыре раза точности случайной классификации говорит как об удачном выборе признаков, сформировав-

ших признаков пространство, так и о высокой степени релевантности этих характеристик для различения индивидуальных стилей авторов.

Таблица 3

Распределение текстов четырех классов в пространстве признаков модели

Исходные классы текстов	Процент правильного распределения	Классы текстов согласно признаковой модели			
		Класс 1 (Брайент)	Класс 2 (Лонгфелло)	Класс 3 (Эмерсон)	Класс 4 (По)
Класс 1 (Брайент)	93,77	30	1	0	0
Класс 2 (Лонгфелло)	90,57	3	48	1	1
Класс 3 (Эмерсон)	97,06	0	1	33	0
Класс 4 (По)	100,00	0	0	0	23
Всего	95,04	33	50	34	24

Полученные данные позволяют также оценить степень диагностируемости отдельных классов. Наиболее диагностируемой является группа произведений По (100%). С очень большой точностью определяются и остальные классы текстов: Эмерсона (97,06%), Брайента (96,77%) и Лонгфелло (90,57%). Такая высокая точность классификации заслуживает внимания также и потому, что привлекаемые к анализу тексты характеризуют различные периоды творчества каждого автора, порой различающиеся между собой как по стилю, так и тематически.

Картина соотношения классов текстов может быть существенно дополнена с учетом порядка их разбиения дискриминантными функциями. Дискриминантные функции – это линейные функции вида

$$f_{km} = n + a_1 \times x_{1km} + a_2 \times x_{2km} + \dots + a_p \times x_{pkm},$$

где f_{km} – значение дискриминантной функции для m -го объекта в группе k ; n – постоянная; x_{ikm} – значение переменной x_i для m -го объекта в группе k ; a_i – коэффициент признака ($i = 1, p$).

Каждый признак является переменной этих функций и имеет собственный коэф-

фициент. Чем больше значение коэффициента признака по модулю, тем более релевантным является этот признак для дискриминации.

В результате анализа было получено три дискриминантные функции, осуществившие три разбиения четырех классов. Первая функция разграничивает классы Брайента и Лонгфелло, с одной стороны, и Эмерсона и По – с другой. Таким образом, основное противопоставление наблюдается между кластерами, включающими группы текстов Брайента и Лонгфелло (первый кластер) и Эмерсона и По (второй кластер). Вторая и третья функции дифференцируют классы текстов внутри, соответственно, первого и второго кластеров.

Как видно из табл. 1, Брайент максимально противопоставлен Эмерсону и По, а класс текстов Лонгфелло занимает центральное положение относительно остальных и удален от них приблизительно на одинаковое расстояние. Эмерсон и По оказываются достаточно удаленными друг от друга и объединяются в один кластер благодаря противопоставлению Брайенту. В целом, структура соотношения лирических текстов рассматриваемых поэтов имеет более сложный и диверсифицированный

характер, чем простое противопоставление «американских» традиций Эмерсона и По и «неамериканских» Брайента и Лонгфелло.

В ряде случаев были выявлены элементы корреляции между степенью удаленности текстов от центра и содержательной стороной текстов (Лонгфелло, Брайент, По). Так, в классе Брайента на периферии находятся произведения, в которых тема природы, которая является ведущей в его творчестве, не находится в центре внимания автора, играя второстепенную роль. У Лонгфелло более удаленными от ядра его класса являются произведения заключительного этапа его творчества, в которых поэт все более обращался к теме смерти. Напротив, в творчестве По на периферии находятся тексты, в которых центральным

является образ живой героини, судьба которой если и печальна, то определяется не нависшим над ней роком, а вполне земными обстоятельствами.

Следует отметить, что ни один из вошедших в модель 35 признаков по отдельности не может дискриминировать рассматриваемые классы произведений. Преимуществом используемого метода является то, что он позволяет установить комплексы дискриминантных признаков с уточнением вклада каждого из них в дифференциацию четырех классов.

В первой дискриминантной функции, дифференцирующей противопоставленные кластеры, более важную роль играют структурные характеристики: рифменные, ритмические, строфические (табл. 4).

Таблица 4

**Признаки с максимальным вкладом в дискриминантные функции
(наиболее релевантные для разграничения соответствующих классов текстов)**

Первая функция	Вторая функция	Третья функция
Количество подлежащих в первой сильной позиции; количество типов строф; количество мужских рифм; количество внесхемных ударений на анакрусе	Количество существительных в последней сильной позиции; количество глаголов в последней сильной позиции; количество прилагательных в последней сильной позиции; количество наречий в последней сильной позиции; количество местоимений в последней сильной позиции	Количество наречий в последней сильной позиции; количество местоимений в последней сильной позиции; количество сказуемых в первой сильной позиции; количество слов, не являющихся членом предложения, в первой сильной позиции; количество строк, разорванных синтаксической паузой

Разграничение внутри кластеров проводится дискриминантными функциями 2 (классы текстов Эмерсона и По) и 3 (классы текстов Брайента и Лонгфелло). В отличие от первой функции при дискриминации классов внутри кластеров более релевантными являются морфологические и синтаксические признаки.

Полученные данные о признаках позволяют провести сопоставительное исследование релевантности инициальных и финальных элементов строки по их роли в дискриминации текстов.

Если на морфологическом уровне наибольшую значимость имеет слово в последней сильной позиции, т. е. в конце стиха, то на синтаксическом и ритмическом уровнях наибольшей диагностической силой обладают признаки инициальной части стихотворной строки. При этом синтаксическая функция характеристик инициальной части стиха позволяет дифференцировать даже достаточно сходные между собой тексты Брайента и Лонгфелло.

На втором этапе исследования сформированная признаковая парадигма исполь-

зована для сопоставления группы текстов Уитьера и описанных классов текстов. Тексты Уитьера были охарактеризованы по тем же признакам, которыми ранее описывались тексты Брайента, Лонгфелло, Эмерсона и По, и были помещены в уже сформированное признаковое пространство наряду с указанными выше текстами критерийных классов.

В полученном признаковом пространстве были определены расстояния (мера Махаланобиса) между текстами Уитьера и центроидами остальных четырех классов текстов (табл. 5).

В результате оказалось, что тексты Уитьера оказались ближе к текстам кластера Брайент–Лонгфелло.

Как следует из данных табл. 5, 13 текстов наиболее близки к классу Брайента,

9 – к классу Лонгфелло и один текст оказался ближе всего к группе текстов Эмерсона.

Таким образом, языковые особенности текстов Уитьера сближают его с кластером Брайента–Лонгфелло. Такой вывод не является совершенно неожиданным, так как литературоведческие исследования⁷ отмечали сходство его произведений с лирикой этих поэтов. Однако столь значительное сходство результатов лингвистического анализа и литературоведческого исследования, основанного на совершенно иной парадигме признаков и иных методах обработки материала, довольно неожиданно свидетельствует о значительной корреляции содержательного и формального планов в стихотворном тексте.

Таблица 5

Расстояние (мера Махаланобиса) между текстами Уитьера и центроидами четырех классов текстов

Произведение	Расстояние до центроидов соответствующих классов			
	Класс 1 (Брайент)	Класс 2 (Лонгфелло)	Класс 3 (Эмерсон)	Класс 4 (По)
1. Clerical Oppressors	42,64	45,49	72,86	65,59
2. New Hampshire	82,11	76,44	80,21	87,52
3. The Rendition	49,28	50,25	45,06	69,00
4. The Emancipation Group	37,67	32,16	35,76	37,63
5. A Word for the Hour	33,61	43,31	65,22	48,84
6. The Watchers	26,04	33,20	49,15	59,91
7. Seed-time and Harvest	28,50	35,23	52,78	47,48
8. The Corn-Song	22,28	22,75	33,95	61,74
9. Poor Voter on Election	61,90	55,62	65,92	70,17
10. The Moral Warfare	34,80	33,25	39,82	44,28
11. At Port Royal	44,08	31,89	39,13	42,78
12. The Proclamation	30,15	35,88	53,05	51,48
13. The Hive at Gettysburg	34,58	22,84	41,07	39,13
14. What of the Day	53,84	58,75	78,69	63,25
15. Garrison	43,91	38,60	47,83	49,74
16. The Problem	42,78	50,15	64,12	46,33

Произведение	Расстояние до центроидов соответствующих классов			
	Класс 1 (Брайент)	Класс 2 (Лонгфелло)	Класс 3 (Эмерсон)	Класс 4 (По)
17. The Jubilee Singers	118,75	105,46	131,17	114,31
18. The Lost Statesman	39,80	60,80	85,84	92,16
19. Mithridates at Chios	35,87	39,46	52,26	63,15
20. The Battle of Autumn of 1862	26,19	22,73	30,58	42,18
21. Italy	18,51	20,09	30,70	37,13
22. Disarmament	27,89	30,75	47,92	51,17
23. Freedom in Brazil	13,45	22,38	33,52	44,35

Как указывалось выше, одно произведении Уитъера оказалось ближе к центроиду класса Эмерсона. Этим произведением является *The Rendition* с расстоянием 45,06 до центроида класса. Таким образом, оно противопоставлено остальному его творчеству, более сходному с кластером Брайента—Лонгфелло. Любопытно, что здесь наблюдается противопоставление остальным текстам Уитъера и на содержательном уровне. Будучи посвящено, как и многие другие его тексты, теме свободы, оно отличается от других лирических произведений пессимистическим настроением.

В других стихотворениях Уитъер опирается на библейские образы (*What of the Day*, *The Hive at Gettysburg* и др.), обращается к будущему, в котором восторжествует справедливость (*Freedom in Brazil*, *Italy*, *The Emancipation Group* и др.):

Yet, surely as He lives, the day
Of peace He promised shall be ours, <...>

Italy

And Samson's riddle is our own to-day,
Of sweetness from the strong,
Of union, peace, and freedom plucked away
From the rent jaws of wrong.

The Hive at Gettysburg

The Voice that spake at Nazareth speaks still,
No sound thereof hath died;

Alike thy hope and Heaven's eternal will
Shall yet be satisfied.

Freedom in Brazil

Stand in thy place and testify
To coming ages long,
That truth is stronger than a lie,
And righteousness than wrong.

The Emancipation Group

В тексте *The Rendition*, напротив, эти элементы содержания отсутствуют:

And Law, an unloosed maniac, strong,
Blood-drunken, through the blackness trod, <...>

«Mother of Freedom, wise and brave,
Rise awful in thy strength,» I said;
Ah me! I spake but to the dead;
I stood upon her grave!

Следующим шагом в исследовании стало выявление характеристик, дифференцирующих тексты Уитъера и тексты остальных авторов. Для этого тексты Уитъера были внесены в многомерное пространство как отдельный класс объектов, затем была рассчитана дискриминантная функция, отграничивающая группу текстов Уитъера от остальных четырех классов.

Степень диагностической точности полученной модели можно оценить на осно-

вании данных табл. 6, где приводятся данные о распределении произведений пяти

классов в признаковом пространстве, образованном признаками модели.

Таблица 6

Распределение текстов пяти классов в пространстве признаков модели

Исходные классы текстов	Процент правильного распределения	Классы текстов согласно признаковой модели				
		Класс 1 (Брайент)	Класс 2 (Лонгфелло)	Класс 3 (Эмерсон)	Класс 4 (По)	Класс 5 (Уитьер)
Класс 1 (Брайент)	90,32	28	1	0	0	2
Класс 2 (Лонгфелло)	90,57	2	48	1	1	1
Класс 3 (Эмерсон)	94,12	0	2	32	0	0
Класс 4 (По)	91,31	0	1	1	21	0
Класс 5 (Уитьер)	78,26	3	2	0	0	18
Всего	89,63	33	54	34	22	21

Привлекаемые к исследованию признаки разграничивают классы стихотворных текстов очень хорошо. Практически 90% текстов было отнесено к своим классам. Относительно более диффузной оказалась группа текстов Уитьера, где 5 из 23 текстов были отнесены к классам Брайента и Лонгфелло, но и в этом случае процент правильного распределения весьма высок. Полученные результаты свидетельствуют о том, что несмотря на сходство текстов Уитьера с текстами Брайента и Лонгфелло, его идиостиль обладает достаточно вы-

раженными чертами, отличающими его от других авторов.

О степени значимости характеристик при разграничении класса текстов Уитьера и классов остальных авторов можно судить, как указывалось выше, по коэффициентам переменных (признаков) в рассчитанной дискриминантной функции. В табл. 7 приводятся данные о стандартизированных коэффициентах признаков, релевантных для дискриминации текстов Уитьера и остальных авторов. Наибольший вклад вносят признаки с большим значением коэффициента по модулю.

Таблица 7

Признаки, дискриминирующие класс текстов Уитьера и классы остальных авторов

Признак	Коэффициент
Количество размеров	-0,11
Количество мужских рифм	-0,61
Количество точных рифм	0,28
Количество стихотворных переносов	0,05
Количество строф	0,12
Количество видов строф	0,17
Количество сверхсхемных ударений на анакрусе	-0,24
Количество в первой сильной позиции в строке слов, не являющихся членами предложения	0,09

Окончание таблицы 7

Признак	Коэффициент
Количество сказуемых в первой сильной позиции в строке	0,42
Количество пропусков второго икта	0,31
Количество обстоятельств, замещающих первую сильную позицию в строке	0,24
Количество строк, разорванных синтаксической паузой	-0,13
Количество дополнений, замещающих последнюю сильную позицию в строке	-0,30
Количество стилистически маркированных строк	-0,19
Средняя длина строки в стопах	-0,07
Количество существительных, замещающих последнюю сильную позицию в строке	1,00
Количество слогов в слове, замещающем последнюю сильную позицию в строке	-0,51
Количество прилагательных, замещающих последнюю сильную позицию в строке	0,62
Количество слогов в анакруссе	0,03
Количество простых предложений в составе сложносочиненного	-0,15
Количество дополнений, замещающих первую сильную позицию в строке	-0,04
Количество глаголов, замещающих первую сильную позицию в строке	-0,10
Количество слогов в слове, замещающем первую сильную позицию в строке	0,15
Количество существительных, замещающих первую сильную позицию в строке	-0,48
Количество неполных инверсий	0,48
Количество наречий, замещающих первую сильную позицию в строке	-0,06
Количество простых предложений в составе сложноподчиненного	0,30
Количество наречий, замещающих последнюю сильную позицию в строке	0,60
Количество глаголов, замещающих последнюю сильную позицию в строке	0,66
Количество определений, замещающих последнюю сильную позицию в строке	-0,13
Количество местоимений, замещающих последнюю сильную позицию в строке	0,25
Количество подлежащих, замещающих последнюю сильную позицию в строке	-0,22
Количество полных инверсий	0,20

Наиболее релевантными являются такие признаки, как количество сказуемых, замещающих первую сильную позицию в строке, количество существительных, глаголов, прилагательных и наречий, замещающих последнюю сильную позицию в строке, и др.

Наименьший вклад в дискриминацию дают признаки, характеризующие структурные свойства стихотворного текста: количество стихотворных переносов (enjambements), количество представленных в тексте стихотворных размеров, количество строф, количество видов строф, средняя длина

строки в стопах, количество слогов перед первой сильной позицией в строке и др.

Такое различие роли признаков может, на наш взгляд, объясняться тем, что структура стихотворного текста (особенности строфики и размера, способ рифмовки и др.) представлена эксплицитно, легко воспринимается и моделируется и, как следствие, во многом определяется направлением, к которому принадлежит автор, предписывается ему традицией. Следование традиции влечет за собой и следование закрепившимся структурным схемам, попытка отойти от традиции автоматически вызывает и отход от этих схем. Так, структурные признаки отличают кластер Эмерсона и По, искавших новые пути для создания собственно американской стихотворной традиции, от кластера Брайента и Лонгфелло, следовавших европейской традиции организации текста. Уитьер следует традиции Брайента и Лонгфелло, и структурные признаки слабо различают их тексты. Напротив, распределение по строке слов различных частей речи и синтаксических конструкций является в значительной степени имплицитным признаком, не регулируется традициями и отражает только особенности авторского стиля. Поэтому морфологические и синтаксические характеристики является более важными для дифференциации текстов различных авторов и позволяют разграничивать более сходные (близкие) классы текстов.

Рассматривая тексты Уитьера как отдельный класс, мы можем получить дополнительную информацию о структуре этого класса, о текстах, входящих в его ядро (т. е. более прототипических) и текстах, находящихся на периферии класса. В табл. 8 приводятся данные о степени удаленности текстов Уитьера от центроида его класса.

Как видно из табл. 8, наиболее прототипичными для данного класса являются

тексты *The Battle of Autumn of 1862, Italy* и *Freedom in Brazil*. Остальные тексты находятся значительно дальше от центроида — приблизительно в два раза. Наиболее удалены *New Hampshire, The Rendition, Poor Voter on Election Day, The Jubilee Singers, The Lost Statesman*.

Интересно, что в данном случае также наблюдается взаимосвязь между лингвистическими характеристиками стихотворного текста и его содержанием. Ядро класса Уитьера полностью сформировано текстами, посвященными борьбе за свободу, ведущуюся государствами, целыми народами. Тексты, находящиеся на периферии, относятся к аболиционистской лирике, т. е. к борьбе за отмену рабства (за исключением текста, *Poor Voter on Election Day*, воспевающего свободные демократические выборы). По нашему мнению, борьба за независимость представляет собой наиболее типичное выражение темы свободы, занимающей важное место в лирике Уитьера. Лингвистические особенности идиостиля Уитьера проявляются наиболее эксплицитно в текстах с этой тематикой. Борьба за отмену рабства и утверждение гражданских свобод является вариацией основной темы, отклонением от нее, что влечет и отклонения языковых характеристик текста от структуры прототипических произведений.

Таким образом, в результате проведенного исследования получена классификация текстов Уитьера, определены языковые характеристики, позволяющие с высокой степенью точности дискриминировать стихотворные тексты различных авторов. При этом выявлено значительное соответствие группировки, полученной на основе лингвистических признаков, с содержательными характеристиками, свидетельствующее о сильной взаимосвязи планов содержания и выражения в стихотворном тексте.

Таблица 8

Удаленность текстов Уитьера от центроида его класса (мера Махаланобиса)

Произведение	Расстояние	Произведение	Расстояние
1. Clerical Oppressors	23,06	13. The Hive at Gettysburg	18,45
2. New Hampshire	44,62	14. What of the Day	26,06
3. The Rendition	31,58	15. Garrison	23,33
4. The Emancipation Group	19,66	16. The Problem	29,76
5. A Word for the Hour	21,13	17. The Jubilee Singers	57,90
6. The Watchers	21,24	18. The Lost Statesman	32,37
7. Seed-time and Harvest	19,16	19. Mithridates at Chios	21,21
8. The Corn-Song	22,12	20. The Battle of Autumn of 1862	10,05
9. Poor Voter on Election	40,41	21. Italy	12,27
10. The Moral Warfare	22,70	22. Disarmament	21,34
11. At Port Royal	21,58	23. Freedom in Brazil	9,42
12. The Proclamation	20,18		

Перспективами исследования могут стать определение интегральных и дифференциальных характеристик тек-

стов различной жанровой принадлежности, стихотворных и прозаических текстов.

ПРИМЕЧАНИЯ

¹ Гаспаров М. Л. Современный русский стих: Метрика и ритмика. – М.: Наука, 1974.

² Пиотровский Р. Г. Анонимные произведения и их авторы (Предисловие). – Минск, 1988; Беляева Л. Н. Применение ЭВМ в лингвистических исследованиях и лингводидактике. – Л.: ЛГПИ, 1986; Тулдава Ю. А. Проблемы и методы количественно-системного исследования лексики. – Таллин: Валгус, 1987.

³ Клекка У. Р. Дискриминантный анализ // Факторный, дискриминантный и кластерный анализ. – М.: Финансы и статистика, 1989. – С. 78–138.

⁴ Баевский В. С. Лингвистические, математические, семиотические и компьютерные модели в истории и теории литературы. – М.: Языки славянской культуры, 2001.

⁵ Андреев В. С. Классификация стихотворных текстов методом дискриминантного анализа (на материале лирики американских поэтов-романтиков) // Вестник Минского государственного лингвистического университета. – Серия 1. Филология. – № 10. – Минск: МГЛУ, 2002. – С. 141–146.

⁶ Дженирич Р. И. Пошаговый дискриминантный анализ // Статистические методы для ЭВМ. – М.: Наука, 1986. – С. 94–112.

⁷ Засурский Я. Н., Коренева М. М., Стеценко Е. А. История литературы США. – Т. 2. – М.: Наследие, 1999.

V. Andreyev

CLASSIFICATION OF J. G. WHITTIER'S POETICAL TEXTS

The paper deals with the classification of the individual style of J. G. Whittier. His texts are compared with the works by poets who founded the main tendencies in the American literature of the 19th century. The research is carried out with the help of discriminant analysis on the basis of a wide range of characteristics belonging to the different levels of language.

The degree of similarity between texts by Whittier and other classes of texts is established, linguistic characteristics highly effective in differentiating verse texts of different authors are revealed.