

РАСПРЕДЕЛЕНИЕ СЛОГОВЫХ ТИПОВ В СТИХОТВОРНЫХ ТЕКСТАХ О.ХАКСЛИ

В настоящей работе исследуется распределение типов слогов в стихотворном творчестве Олдоса Хаксли на материале его стихотворений, написанных в начале XX века. Подсчет слогов разных типов позволяет составить базу данных, которая используется для сопоставления различных текстов. Для этого проводится аппроксимация распределения слогов с использованием экспонентной функции и функции Ципфа – Алексеева.

В ходе исследования выделяются последовательности, организованные по правилам неповторяемости своих элементов (мотивы), распределение которых позволяет установить закономерности на более высоком уровне. Результаты показали, что все случаи распределения различных слоговых типов в произведениях Хаксли хорошо аппроксимируются указанными функциями.

Ключевые слова: *Хаксли, экспонентная функция, типы слогов, мотивы.*

Статья посвящена исследованию распределения различных типов слогов в поэтическом творчестве Олдоса Леонарда Хаксли. Как было показано в ряде исследований, соотношение и распределение слоговых типов может служить основой классификации индивидуального стиля поэта. Более подробно об этом сказано в статье В.С. Андреева «Аппроксимация распределения типов слогов в стихотворном тексте» в этом сборнике. Наряду с индивидуальными особенностями авторов слоговой состав текстов подчиняется некоторым основным закономерностям [Andreev, 2018; Köhler, Altmann, 2014]. В настоящей статье ставится задача, используя методологию анализа слогового состава в указанных работах, проанализировать индивидуальные особенности стиля О. Хаксли в его поэтическом творчестве.

В качестве материала к анализу привлекается ряд поэтических произведений Хаксли начального периода творчества (сборник стихотворений «The burning wheel»).

Обозначив гласные в слоге как V, согласные – как C, получаем следующие типы слогов: V (are), CV (she), CCV (stir), CVC (walk), CCVC (black), CVCC (lift), CCCV (straw), VC (at), CCVCC (branch), VCC (and), CCCVC (spring), CVCCC (limps) и некоторые другие, крайне малочастотные типы.

Рассмотрим в качестве примера стихотворение Хаксли «Philoclea in the forest». Подсчет слоговых типов позволил получить следующие результаты.

Слоги, начинающиеся с гласной:

V – 25

VC – 46

VCC – 27

VCCC – 1

Среди слогов, начинающихся с согласной, максимальное количество зафиксировано для двух типов:

CV – 133

CVC – 132

Асимметричные слоги 4 и 5 имеют средние ранги частотности:

CCVC – 33

CVCC – 37

Еще один слог CCV, составляющий пару «обратному» по форме слогу VCC, имеет частоту 23.

Частоты остальных слогов достаточно малы: CCVCC встречается 5 раз, CVCCC зафиксирован 4 раза, слоги CCCVCC, CCCVC – 3 раза.

После этого все указанные слоги были расположены по убыванию их частот. Первый ранг получает слоговой тип CV, второй – CVC, третий VC и так далее. Этот ранговый ряд может быть проанализирован с различных точек зрения. Одним из возможных подходов является анализ распределения слоговых типов [Andreev, Mistecký, Altmann 2018]. Используя предложенные в работе формулы, можно попытаться достичь хороший уровень аппроксимации распределения.

Для этого используется экспонентная функция [Andreev, Lupea, Altmann 2017; Wimmer, Altmann 2005]:

$$f_x = a * e^{-bx},$$

где a и b – параметры.

В таблице 1 указываются ранжированные частоты слогов, которые были выявлены в тексте «Philoclea in the forest», а также ожидаемые частоты, исходя из функции.

Аппроксимация дает очень хороший показатель – коэффициент детерминации $R^2 = 0,9164$ (параметры – $a = 204,37$, $b = 0,360$).

Все другие стихи, которые были проанализированы, также показывают наличие регулярной тенденции в распределении слоговых типов. Во всех случаях коэффициент детерминации очень высок, что говорит о хорошей аппроксимации данной экспонентной функцией эмпирических распределений слогов.

Таблица 1

**Наблюдаемые и ожидаемые частоты слоговых типов
в стихотворении Хаксли «Philoclea in the forest»**

п/п	Тип слога	Наблюдаемые частоты	Ожидаемые частоты
1	CV	133	142,53
2	CVC	132	99,40
3	VC	46	69,32
4	CVCC	37	48,34
5	CCVC	33	33,71
6	VCC	27	23,51
7	V	25	16,40
8	CV	23	11,43
9	CCVCC	5	7,97
10	CVCCC	4	5,56
11	9CCCVCC	1	3,88
12	CCVC	1	2,70
13	VCCC	1	1,89

Из числа остальных текстов выделим следующие результаты.

QUOTIDIAN VISION.

Формула имеет параметры $a = 0,37,84$; $b = 0,35$, а коэффициент регрессии равен $R^2 = 0,931$.

VISION

$a = 49,53$; $b = 0,40$ $R^2 = 0,931$

THE MIRROR

$a = 41,15$; $b = 0,37$ $R^2 = 0,973$

VARIATIONS ON A THEME OF LAFORGUE

$a = 46,95$; $b = 0,44$ $R^2 = 0,962$

PHILOSOPHY

$a = 23,07$; $b = 0,46$ $R^2 = 0,983$

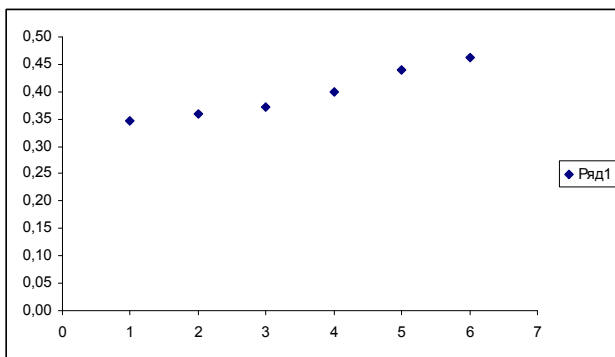


Рис. 1. График рассеяния стихотворений по параметру b

Прежде всего, отметим, что, как показывают полученные результаты, экспонентная функция хорошо аппроксимирует распределение слоговых типов в текстах, написанных более 100 лет тому назад. Во-вторых, следует подчеркнуть, что эта функция аппроксимирует распределения слогов в различных по тематике и объему стихотворениях.

Использование цепочек слогов для дальнейшего анализа может вызвать серьезные затруднения из-за большого количества вариантов их сочетаемости. Избежать этих трудностей можно, если использовать методику выделения цепочек на основе неповторяемости ее элементов. Такой подход был предложен Р. Келером [Köhler, 2008; Köhler, Naumann 2010; Altmann, Köhler 2015] и позволил получить очень хорошие результаты [Kellih et al. 2016].

В качестве примера возьмем первую строфу стихотворения Хаксли «Philoclea in the forest».

'TWas I that leaned to Amoret
 With: "What if the briars have tangled Time,
 Till, lost in the wood-ways, he quite forget
 How plaintive in cities at midnight sounds the chime
 Of bells slow-dying from discord to the hush whence
 they rose and met.

Этот отрывок можно представить в следующее записи, в которой отмечены типы слогов:

CCVC V CVC VCC CV V CV CVC
 CVC CVC VC CV CCV VC CVC CVC CCVC CVC
 CVC CVCC VC CV CVC CVC CV CCVC CV CVC
 CV CCVC CVC VC CV CVC VC CVC CVC CVCCC CV CVC

VC CVCC CCV2 CV VC CCVC CVC CVC CV CV CVC CVCC
CV CVC VCC CVC

Исходя из правила недопустимости повторения одинаковых слогов в первой строке выделяется: CCVC V CVC VCC CV. Далее следует слог V, который уже присутствует в первой последовательности. За слогом V идет слог CV, который мог бы быть включен во вторую последовательность. Однако здесь вступает в действие второе правило, которое запрещает в ней иметь два элемента, сходных с последовательностью предыдущей. Поэтому CV отходит к следующей последовательности. Используя эти правила, всю строфу можно разметить следующим образом. Наклонная черта обозначает конец каждой очередной последовательности.

CCVC V CVC VCC CV / V / CV CVC /
CVC / CVC VC CV CCV / VC / CVC / CVC CCVC / CVC /
CVC CVCC VC CV / CVC / CVC CV CCVC / CV / CVC
CV CCVC / CVC VC / CV CVC / VC CVC / CVC CVCC CV / CVC
VC CVCC CCV2 / CV / VC CCVC CVC / CVC CV / CV / CVC CVCC
CV / CVC VCC / CVC /

Всего в строфе выделяется 26 последовательностей, которые Р. Келлер назвал «мотивами» [Köhler, Naumann, 2016].

Как видно из записи, мотивы могут быть разной длины – по количеству входящих в каждый из них слогов. Так, первый мотив включает в себя пять слогов, второй короче – только один слог, третий два слога и так далее. Здесь следует отметить, что выделение мотивов происходит по всей строфе, а не в границах отдельных строк.

Таблица 2

**Аппроксимация типов мотивов в «PHILOCLEA IN THE FOREST»
при помощи формулы Цифа-Алексеева**

Длина	Частоты	
	наблюдаемые	теоретически ожидаемые
1	74	74,00
2	69	71,73
3	45	39,00
4	18	19,42
5	5	9,72
6	4	5,01
	$R^2 = 0,986017$ $a = 0,87498$ $b = -1,32714$	

Таким образом, представляется возможным подсчитать количество 1-членных (один слог), 2-членных (2 слога) мотивов и т.д. Максимальная длина мотива в стихотворениях Хаксли составляет 6.

В таблице 2 отражен количественный состав мотивов в рассматриваемом тексте. В ней также даются результаты аппроксимации распределения мотивов, взятых по их длине, с использованием функции Циффа–Алексеева. Более детально об этой функции и ее формула приводятся в статье В.С. Андреева в настоящем сборнике.

Результаты помещены в таблице 3.

Судя по коэффициенту детерминации, аппроксимация оказалась весьма успешной.

В таблице 3 приводятся результаты аппроксимации распределения количественных мотивов еще в пяти стихах.

Таблица 3

Аппроксимация типов мотивов в при помощи формулы Циффа-Алексеева

QUOTIDIAN VISION		VISION		THE MIRROR		VARIATIONS ON A THEME OF LAFORGUE		PHILOSOPHY	
Частота		Частота		Частота		Частота		Частота	
набл.	теор.	набл.	теор.	набл.	теор.	набл.	теор.	набл.	теор.
14	14,00	18	18,00	13	13,00	17	17,00	8	8,00
13	12,25	17	16,69	9	9,72	11	11,37	6	5,90
6	7,76	7	7,89	8	6,87	7	6,62	3	3,26
6	4,73	4	3,39	5	4,96	5	3,94	2	1,78
3	2,93	2	1,47	3	3,68	1	2,43		
R ² = 0,944 a = 0,398 b = -0,852		R ² = 0,993 a = 0,986 b = -1,581		R ² = 0,962 a = -0,144 b = -0,398		R ² = 0,977 a = -0,105 b = -0,686		R ² = 0,994 a = 0,209 b = -0,934	

Анализ показывает, что снова, независимо от размера и темы произведения, слоговой состав оказывается подверженным действию какой-то глубинной тенденции, упорядочен и по частоте использования индивидуальных слогов, и по их комбинациям.

Полученные результаты свидетельствуют о целесообразности дальнейших исследований за счет расширения материала. Кроме того, они могут составить базу для сопоставительного анализа распределение слоговых типов в разных языках.

Литература

- Altmann G., Köhler R. Forms and Degrees of Repetitions in Texts. Detection and Analysis. Berlin–Munich–Boston: de Gruyter Mouton, 2015.
- Andreev S. Distribution of Syllables in Russian Sonnets. *Glottometrics* 41, 2018. P. 13–23.
- Andreev S., Místecký M., Altmann G. Sonnets: Quantitative Inquiries. *Studies in Quantitative Linguistics*, 29. Lüdenscheid: Ram-Verlag, 2018.
- Andreev S., Lupea M., Altmann G. Belza chains of adnominals // *Glottometrics* 39. 2017. P. 72–86.
- Zornig P. et al. Studies in quantitative linguistics. Quantitative Inquiries into the Syllable Structure. Lüdenscheid: RAM-Verlag, 2019.
- Köhler R. Sequences of Linguistic Quantities. Report on a New Unit of Investigation // *Glottology* № 1 (1). 2008. P. 115–119.
- Köhler R., Altmann G. Problems in Quantitative Linguistics. Lüdenscheid: Ram-Verlag, 2014.
- Köhler R., Naumann S. A syntagmatic approach to automatic text classification. Statistical properties of F- and L-motifs as text characteristics // *Text and Language, structures, functions, interrelations, quantitative perspectives*. Wien: Praesens, 2010. P. 81–89.
- Köhler R., Naumann S. Syntactic text characterization using linguistic S-motifs. *Glottometrics* 34, 2016. P. 1–8.
- Wimmer G., Altmann G. Unified derivation of some linguistic laws // *Linguistic Motifs Sequences Quantitative Linguistics. An International Handbook*. Berlin: de Gruyter, 2005. P. 791–807.

N.A. Nikiforova

THE DISTRIBUTION OF SYLLABLE TYPES IN POEMS BY O. HUXSEY

Abstract. In the present research syllable distribution in poetry of Aldous Huxley in his poems, written at the beginning of the 20th century is studied. The distribution of different types of syllables are explored. Counting the types of syllables one obtains a data-base which can be used for fitting using the exponential and Zipf-Alekseev functions. In the present research strings of syllables based on non-repetition of their elements (motifs) are singled out whose distribution reflects the order on a higher level. The results revealed that the rank-frequency distribution of syllabic types in all poems under study is well fitted by both functions..

Keywords: *Huxley, the exponential function, syllable types, motifs.*