

В.В. ЕВДОКИМОВА
**АНАЛИЗ СПЕКТРА ГЛАСНЫХ НА ОСНОВЕ
НЕРАВНОМЕРНОЙ ПСИХОАКУСТИЧЕСКОЙ ШКАЛЫ ЭРБОВ
ДЛЯ ОПРЕДЕЛЕНИЯ СЛОВЕСНОГО УДАРЕНИЯ**

Евдокимова В.В. Анализ спектра гласных на основе неравномерной психоакустической шкалы эрбов для определения словесного ударения.

Аннотация. Рассматривается возможность определения в слове ударного слога путем изучения особенностей формантной картины. На основе анализа массива логатомов выявляются особенности формантной картины, причем сравниваются энергетические характеристики отдельных формант. Опираясь на допущение об оптимальности способа формирования диктором речевой посылки, спектр гласных разбивается по психоакустической шкале эрбов. Для удобства обработки материала полученный массив частичных дисперсий кодифицируется. Для выявления признака ударности полученный массив кодов разбивается по признаку гласной и по признаку ударности. Сравнение частичных массивов дает возможность подтвердить существование исследуемого признака ударности и выявить фонетические особенности этого явления, существующие в формантной картине. Сделан вывод о возможности применения признака ударности по особенностям формантной картины гласной как эффективно дополняющего определение ударности слога в слове по иным общепринятым признакам.

Ключевые слова: фонетика, акустический анализ речевого сигнала, формантная картина гласного, шкала эрбов, ударность гласного.

Evdokimova V.V. Vowel spectral analysis using the psychoacoustic ERB scale for the word stress detection.

Abstract. The paper presents a possibility to determine the stressed syllable in the word by examining features of the vowel formant structure. Based on the analysis of the array of logatoms the vowel formant structure is identified. The energy characteristics of individual formants of different vowels in one word are compared. Spectrum is split by the vowels psychoacoustic ERB scale. For ease of handling of the material resulting array of partial dispersions is codified. To determine the stressed vowel the resulting array codes are broken on the basis of the stressed/unstressed vowel characteristic and phoneme type. Partial arrays comparison enables to confirm the existence of the stress feature characteristic and identify phonetic features of this phenomenon, existing in the vowel formant structure. The vowel stress characteristic is defined by the vowel formant pattern and can effectively complement the determination of stressed syllables being added to conventional stressed vowel characteristics.

Keywords: phonetics, speech acoustic analysis, vowel formant structure, ERB scale, vowel stress.

1. Введение. В современных системах автоматической обработки и анализа речевого материала существует проблема определения границ слов в потоке речи. Одним из приемов решения этой проблемы является определение места словесного ударения, которое «объединяет звуки, образующие облик слова, а если слово состоит более чем из одного слога, то ударный слог связывает его в единое целое» [5].

Как отмечает Ю.С. Маслов, «словесное ударение заключается в том, что в слове с помощью тех или иных звуковых средств подчёркивается один вполне определённый слог» [7].

Известно, что информативность словесного ударения обеспечивается всей совокупностью речевых явлений [1, 3]. Ударный слог выделяется интенсивностью речевого сигнала, его длительностью, вариациями частоты основного тона и формантной картины, причем способ такого комплексного выделения ударения диктором всегда оптимизирован по критерию его восприятия слушателем.

Основным коррелятом словесного ударения в русском языке является длительность, что было доказано в диссертации Л.В. Златоустовой [6]. Ударный слог обычно является самым длинным в слове [1]. Этот признак достаточно устойчив и выделяется многими исследователями [4, 7, 9], но в ряде случаев длительность первого предупредительного слога может оказаться сравнимой с ним и даже превышать его длительность. Дополнительную неточность вносит и удлинение последнего слога в синтагме [2].

Другим общепринятым признаком ударения является повышенная интенсивность ударного слога, которая проявляется как в увеличении его акустической мощности, так и в повышении частоты основного тона из-за большей напряженности речевого тракта [3]. Однако признак интенсивности также не может считаться универсальным. Замеряемая акустическая мощность в значительной степени зависит от типа гласной в ударном слоге. Так, акустическая мощность ударного слога с закрытой гласной может оказаться меньше, чем у соседнего предупредительного либо первого заударного слога с открытой гласной. В свою очередь, частота основного тона существенно зависит от интонационного рисунка речи, который тоже вносит искажение в определение ударного слога.

Когда появилась возможность создавать системы автоматического синтеза речи, было выявлено, что акустически предупредительные, ударные и заударные аллофоны имеют разные спектральные картины [1, 3, 8], т. е. отмечено, что ударение формируется не только действием голосового источника, реализующего отмеченные выше признаки, но и деятельностью артикуляторного аппарата, трансформирующего формантную картину гласной [1]. Ряд исследователей обратил внимание на особенности спектра речевого сигнала: в ударных слогах повышена интенсивность низкочастотных первой и второй формант [11, 12], что соответствует участку высокого уровня слухового восприятия и способствует лучшему информационному контакту со слушателем.

Данная статья посвящена исследованию процесса формирования формантной картины гласных с целью оценки возможности использования особенностей трансформации этой картины в качестве признака ударности слога.

2. Оценка влияния вариативности формантной картины гласного на словесное ударение. Предметом исследования стал записанный одним диктором в условиях лаборатории экспериментальной фонетики Санкт-Петербургского государственного университета массив из пяти тысяч трехсложных логатомов [14].

Чтобы исключить влияние интонационных особенностей построения фразы, в данной работе было решено исследовать характеристики ударных и безударных гласных в изолированно произнесённых словах - логатомах. Каждый логатом представлял собой повторение трех слогов типа CVCVCV, где при одном и том же наборе в каждом слоге гласных и согласных ударение ставилось попеременно на одну из трех гласных в слове.

Логатомы составлялись таким образом, чтобы можно было анализировать гласные во всех акустически значимых фонетических контекстах [8]. Для этого массивы логатомов были записаны для гласных /a/, /i/, /u/, /e/, /o/, /ɨ/. Для каждой согласной записаны логатомы с различным местом ударного слога в логатоме. Общее количество аллофонов гласных фонем в материале составляет 5263 единицы. Общее количество слов в материале составило 1748 единиц. В дальнейшем они были сгруппированы как по гласным, так и по месту ударного слога в логатоме.

Далее производилась сегментация полученных файлов и их обработка с помощью стандартного программного обеспечения Wave Assistant. Спектральные характеристики гласных вычислялись при помощи программы «Фонетическая лаборатория v.3». Результатом обработки явились спектрограммы гласных, отражающие основные особенности их формантных картин [1]. Площади пиков на спектрограмме соответствовали энергии формант. Спектрограммы были построены в шкале эрбов [13], учитывающей психофизиологические особенности речи диктора. Выбор именно этой шкалы обусловлен тем, что она дает оптимальный с точки зрения дальнейшей обработки уровень градации спектрограммы на участке трансформации первой форманты (280...650 Гц), который соответствует седьмому-двенадцатому эрбам (284...643 Гц). Предварительно этот диапазон трансформации первой форманты был уточнен для данного диктора по полученным спектрограммам. Альтернативные шкалы барков и мелов также могли быть использованы, но они не имеют на этом участке естественной градации удобной для дальнейшей обработки.

При обработке спектральных плотностей гласных нормированные частичные дисперсии определялись для каждого эрба по формуле:

$$D_i = \frac{1}{D} \int_{f_{1i}}^{f_{2i}} S(f) df, \quad (1)$$

где i – номер эрба ($i = 1 \dots 30$), D – дисперсия сигнала, рассчитанная по спектрограмме, f_{1i} и f_{2i} – частотные границы i -го эрба, $S(f)$ – спектральная плотность (спектрограмма) гласной, полученная при обработке.

Таким образом, для каждой гласной была получена последовательность из тридцати частичных дисперсий, характеризующая распределение энергии по шкале эрбов. После этого проводилось сопоставление последовательностей, полученных для одного ударного и двух безударных слогов одного и того же логотома. Пример такого сопоставления приведен в таблице 1.

Таблица 1. Интенсивность спектра по эрбам

Вид логатома	Номер эрба									
	1	...	7	8	9	10	11	...	30	
бЕбебе	0,191	...	0,026	0,082	0,063	0,011	0,155	...	0,012	
бЕбебе	0,082	...	0,048	0,098	0,049	0,006	0,362	...	0,0075	
бЕбебе	0,312	...	0,160	0,195	0,055	0,041	0,189	...	0,0012	
бебЕбе	0,112	...	0,040	0,155	0,091	0,0119	0,184	...	0,017	
бебЕбе	0,038	...	0,039	0,077	0,042	0,009	0,302	...	0,001	
бебЕбе	0,225	...	0,135	0,121	0,038	0,033	0,127	...	0,001	
беббЕ	0,011	...	0,128	0,031	0,031	0,076	0,258	...	0,002	
беббЕ	0,056	...	0,134	0,154	0,041	0,065	0,272	...	0,001	
беббЕ	0,105	...	0,184	0,067	0,017	0,163	0,056	...	0,015	

Каждый опыт (каждый логатом) представлен тремя строками таблицы, в которых приведены последовательности распределения энергии, рассчитанные для первого, второго и третьего его слогов соответственно. По каждому эрбу проведено сравнение относительных дисперсий между тремя слогами данного логатома (в пределах столбца из трех строк логатома). Наиболее интенсивный слог выделен на каждом эрбе жирным шрифтом).

Таблица дает возможность сопоставить распределение интенсивности сигнала для различных слогов одного и того же логатома по каждому участку спектра, соответствующему одному эрбу. Появляется возможность перехода с ее помощью к вторичным показателям оценки относительной интенсивности в различных слогах

одного логатома. Для этого после сравнения по каждому эрбу относительных частичных дисперсий трех слогов трехстрочного столбца, выбирается слог с максимальной дисперсией и полученному таким путем событию присваивается код 1, 2 или 3 по номеру строки в этом трехзначном столбце. Таким образом, каждый опыт рассматривается как сложное событие, описываемое тридцатирядным троичным кодом [10]. Пример преобразования таблицы дискретизации приведен в таблице 2.

Таблица 2. Интенсивность спектра по эрбам

Вид логатома	Код дискретизации по эрбам
бЕбебе	3...33132...1
бебЕбе	3...31132...1
бебебЕ	3...32232...3

Собранный информационный массив велик и структурно избыточен для решения задачи оценки связи словесного ударения со структурой формантной картины. Исходя из поставленной задачи исследования и из известных особенностей физиологии работы речевого тракта, выберем структуру аппарата анализа.

Низкочастотный участок спектра речевого сигнала описывает работу голосового источника. Считаем, что структура первых шести эрбов (частный диапазон 30...284 Гц) в основном определяется вариациями частоты основного тона [1, 2]. Их исследование выходит за пределы данной работы, направленной на изучение трансформации формантной картины, поэтому рассмотрение начинаем с седьмого эрба, поставив задачу поиска особенностей формирования кодов – вторичных признаков трансформации формантной картины.

В соответствии с поставленной задачей сосредоточим внимание на участке спектра с седьмого до двенадцатого эрба (частотный диапазон 284...643 Гц). Проведенные пробные расчеты по большему числу эрбов показали, что увеличение их числа не только не помогает выявлению признака ударности, но наоборот нивелирует результаты. Причина этого в том, что начиная с диапазона второй форманты существенно сказываются различия формантных картин разных гласных.

Каждый из опытов рассматривается, таким образом, как событие, описываемое шестирядным кодом (см. таблицу 2) по числу эрбов на отобранном участке спектра (7-й...12-й эрбы).

Полученный код представляет собой формализованное отражение фонетических процессов, соответствующих определенному в конкретном опыте логатому и дает возможность оценки по вторичным признакам картины трансформации первой форманты в ударных слогах.

Разделим полученный по результатам обработки эксперимента общий числовой массив на три по признаку первого, второго и третьего ударных слогов, затем каждый из них – на шесть массивов по числу исследуемых гласных фонем русского языка /a/, /i/, /u/, /e/, /o/, /i/.

Сгруппируем массивы разных фонем по признаку ударного слога и сопоставим их. Оценим степень их перекрытия, т.е. наличие одинаковых кодов, что в фонетическом плане соответствует одинаковому функционированию артикуляторного аппарата. В таблицах 3-5 приведены оценки степени взаимоперекрывания этих массивов, а именно присутствия в массиве **у** кодов гласной /y/, кодов гласной /x/. Показатели приведены в процентной форме в соответствии с формулой:

$$y_{xy} = \frac{N_{xy}}{N_x} \times 100, \quad (2)$$

где y_{xy} - степень присутствия кодов массива **У** в массиве **X**, N_x - общее число опытов в массиве **X**, N_{xy} - число опытов массива **У**, имеющих коды, присутствующие в массиве **X**.

В таблицах 3-5 введено обозначение m_j , где m – одна из шести гласных, j – номер слога, на который падает ударение в логатоме. Как уже указано, эти таблицы отражают определенную по формуле (2) степень перекрытия пространства кодов по каждой гласной, т.е. случаев, когда формантная картина одной гласной неотличима от основного массива формантных картин другой гласной.

Таблица 3. Ударение выпадает на первый слог логатома

		Массив Y						
		%	e1	i1	a1	u1	o1	i1
Массив X	e1	X	19	16	21	15	10	
	i1	47	X	8	14	8	12	
	a1	36	14	X	13	37	5	
	u1	30	27	14	X	16	25	
	o1	23	11	26	11	X	10	
	i1	40	11	15	54	30	X	

Таблица 4. Ударение выпадает на второй слог логатома

		Массив Y						
		%	e2	i2	a2	u2	o2	i2
Массив X	e2	X	13	58	46	52	21	
	i2	28	X	14	11	10	6	
	a2	45	17	X	20	63	11	
	u2	52	3	43	X	53	22	
	o2	33	2	34	29	X	20	
	i2	34	8	24	41	40	X	

Оценим попарно по таблицам 3-5 степень взаимопроникновения исследуемых массивов кодов гласных. Качественный результат сравнения сведем в таблицу 6.

Количественно уровень взаимопроникновения максимален при втором ударном слоге, минимален при третьем.

Таблица 5. Ударение выпадает на третий слог логатома

		Массив Y						
		%	e3	i3	a3	u3	o3	i3
Массив X	e3	X	16	52	42	49	18	
	i3	24	X	13	12	9	7	
	a3	32	17	X	20	63	11	
	u3	48	5	42	X	51	23	
	o3	32	3	37	26	X	18	
	i3	36	9	22	41	42	X	

Таблица 6. Степень взаимопроникновения массивов кодов гласных

Ударный слог логатома	Массивы кодов гласных	
	Максимальное взаимопроникновение	Минимальное взаимопроникновение
Первый слог	/u/ и /i/, /o/ и /a/, /i/ и /e/	/o/ и /i/, /i/ и /a/, /i/ и /i/
Второй слог	/a/ и /e/, /a/ и /o/, /e/ и /u/	/o/ и /i/, /u/ и /i/, /i/ и /i/
Третий слог	/u/ и /i/, /o/ и /a/, /i/ и /u/	/o/ и /i/, /a/ и /i/

Результаты сравнения хорошо согласуются с известными положениями фонетики о близости артикуляции определенных пар гласных (а – е, е – і, о – u) [1], что подтверждает допустимость выбранного инструмента исследования.

Проведем сравнение массивов различных ударных слогов, так как именно перекрестная связь между этими массивами представляет собой погрешность при использовании алгоритма определения в логатоме ударного слога по кодам. Для этого объединим массивы гласных по признаку общего ударного слога (первый гласный в слове ударный, второй или третий) и сравним эти массивы. Допустимость такого обобщения обусловлена заложенным в постановку задачи ограничением частотного диапазона, которое исключает разброс кодов разных гласных от более высокочастотных формант.

По результатам сравнения можно отметить, что в каждом из трех массивов выделяются три основные группы кодов. Процентное отношение этих групп к общему числу кодов в данном массиве приведено в таблице 7. В каждом из них существует группа кодов, соответствующая исключительно этому массиву. Другой важной группой являются коды, многократно повторяющиеся во всех трех массивах. Фонетически они соответствуют наиболее привычным диктору режимам работы речевого тракта. Перечислим их: 311211, 311311, 311312, 311313, 311321, 313312, 313313, 321311, 321321, 331211, 331311, 331321, 332211, 332312, 332312, 332322.

Таблица 7. Перекрестное присутствие кодов

Вид кодов	Номер ударного слога		
	Первый слог (%)	Второй слог (%)	Третий слог (%)
Коды, неповторяющиеся в других массивах	43	33	53
Коды, однократно повторяющиеся в других массивах	7	8	6
Коды, многократно повторяющиеся во всех массивах	52	59	41

Кроме этих двух групп существует промежуточная группа, которая соответствует однократному появлению каком-то массиве кода, присутствующего также и в другом массиве.

3. Заключение. Проведенное исследование позволяет сделать следующие выводы. Использованный в работе метод кодирования формантных картин гласных в исследуемых логатомах дает возможность перевода набранного речевого материала в форму, удобную для проведения его анализа.

Поставленный эксперимент подтверждает, что ударному слогу в слове присуща общая для всех гласных характерная трансформация формантной картины. Примерно в 30-50% опытов в ударном слоге формируется комбинация кодов, которая позволяет выделить ударный слог в слове. В остальных случаях формируются коды, соответствующие привычным для диктора режимам работы речевого тракта, которые присутствуют как в ударных, так и в безударных слогах.

Анализ результатов подтверждает, что характерная трансформация формантной картины присутствует в ударном слоге, но этот признак недостаточно устойчив и может быть использован при определении ударения только как дополнительный к описанным выше.

Литература

1. *Бондарко Л.В.* Фонетика современного русского языка. СПб. 1998. 275 с.
2. *Вольская Н.Б., Степанова С.Б.* Предпаузальное удлинение в русском языке // Экспериментально-фонетический анализ речи: проблемы и методы. Фонетическое многообразие языков мира. СПб. 2004. № 5. С. 48-55
3. *Кодзасов С.В., Кривнова О.Ф.* Общая фонетика. М.: РГГУ, 2001, 592 с.
4. *Кривнова О.Ф.* Реализация словесного ударения в связанном тексте // Фонетические чтения в честь 100-летия со дня рождения Л.Р. Зиндера, Санкт-Петербург, СПбГУ, 2004. С. 150-154.
5. *Зиндер Л.Р.* Общая фонетика. СПб. 2007. 576 с.
6. *Златоустова Л.В.* Фонетическая природа русского словесного ударения: дисс. канд. филол. Наук. Л., 1953.
7. *Маслов Ю.С.* Введение в языкознание. М., 1998. 272 с.
8. *Скрелин П.А.* Фонетические аспекты речевых технологий. СПб, 1999. 76 с.
9. *Скрелин П.А.* Формальные методы анализа речи: проблемы интерпретации результатов // Материалы XXXVI международной филологической конференции, Санкт-Петербург, 2007. С. 3-14.
10. Цифровое кодирование URL: http://fmi.asf.ru/library/book/network2_2.html (дата обращения: 26.10.2013).
11. *Sluijter A.M.C., van Heuven V.J.* Spectral balance as an acoustic correlate of linguistic stress // *Transactions of the Acoustical Society of America*, 1996, vol. 100, pp. 2471-85. Труды SPIIRAS. 2014. Вып. 1(62). ISSN 2078-9181 (печ.), ISSN 2078-9580 (онлайн).

12. *Sluijter A.M.C., van Heuven V.J., Pacilly J.J.A.* Spectral balance as a cue in the perception of linguistic stress // *Journal of the Acoustical Society of America*, 1997. vol. 101, pp. 503-513.
13. *Smith J.O., Abel J.S.* Bark and ERB Bilinear Transforms // *IEEE Trans. on Speech and Audio Proc.*, Nov 1999. vol. 7, no. 6, pp. 697-708.
14. *Welge-Lüssen A., Hauser R., Erdmann J., Schwob Ch., Probst R.* Sprachaudiometrie mit Logatomen // *Laryngo-Rhino-Otologie*, 2008. vol. 76. №2. p. 57–64.

References

1. Bondarko L.V. *Fonetika sovremennogo russkogo jazyka* [The Phonetics of Contemporary Standard Russian]. SPb. 1998, 275 p. (In Russ.).
2. Volskaya N.B., Stepanova S.B. [Prepause extension in Russian]. *Eksperimentalno-foneticheski analiz: problem i metody – Experimental-phonetics apeech analysis: problems and methods*. Spb. 2004. no. 5. pp. 48-55 (in Russ.).
3. Kodzasov S.V., Krivnova O.F. *Obshhaja fonetika* [General Phonetics]. M., 2001. 592 p. (In Russ.).
4. Krivnova O.F. [Realization of word stress in continuous speech]. *Foneticheskije chtenija v chest' 100-letije so dnia rozhdenija L.R.Zindera*. [Phonetic readings in honour of 100-th anniversary of Zinder L.R.] Spb. SPbSU. 2004. pp. 150-154. (in Russ.).
5. Zinder L.R. *Obshhaja fonetika* [General phonetics]. SPb. 2007. 576 p. (in Russ.).
6. Zlatoustova L.V. *Foneticheskaja priroda russkogo slovesnogo udarenija* [Phonetic nature of Russian word stress] PhD thesis in Philological sciences. Leningrad. 1953. (in Russ.).
7. Maslov Y.S. *Vvedenie v jazykoznanie* [Introduction to linguistics]. M. 1998. 272 p. (in Russ.).
8. Skrelin P.A. *Foneticheskie aspekty rechevyh tehnologij* [Phonetic aspects of speech technologies]. SPb. 1999. 76 p. (in Russ.).
9. Skrelin P.A. [Formal methods of speech analysis: problems of the result interpretation]. *Materijaly XXXVI mezhdunarodnoj filologicheskoy konferencii* [XXXVI international philological conference]. SPb. 2007. pp. 3–14. (In Russ.).
10. Cifrovoe kodirovanie [Digital encoding]. Available at: http://fmi.asf.ru/library/book/network/2_2.html (accessed 26.10.2013). (in Russ.).
11. *Sluijter A.M.C., van Heuven V.J.* Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America*, 1996. vol. 100, pp. 2471–85.
12. *Sluijter A.M.C., van Heuven V.J., Pacilly J.J.A.* Spectral balance as a cue in the perception of linguistic stress. *Journal of the Acoustical Society of America*, 1997. vol. 101, pp.503-513.
13. *Smith J.O., Abel J.S.* Bark and ERB Bilinear Transforms. *IEEE Trans. on Speech and Audio Proc.*, Nov 1999. vol. 7, no. 6, pp. 697-708.
14. *Welge-Lüssen A., Hauser R., Erdmann J., Schwob Ch., Probst R.* Sprachaudiometrie mit Logatomen. *Laryngo-Rhino-Otologie*, 2008. vol. 76. no. 2, pp. 57–64.

Евдокимова Вера Вячеславовна — к-т филол. наук., ст.преподаватель кафедры фонетики и методики преподавания иностранных языков филологического факультета СПбГУ. Область научных интересов: Речевые технологии, фонетика, языкознание, фонетический анализ речевого сигнала. Число научных публикаций — 30. postmaster@phonetics.spbu.ru, www.phonetics.spbu.ru; СПбГУ, Университетская наб., д. 11, г. Санкт-Петербург, 199034, РФ; p.t. +7(812)328-9565.

Evdokimova Vera Viatcheslavovna — Ph.D., senior teacher, Department of Phonetics and Foreign Languages Teaching Methodology, Philological Faculty, SPbSU. Research interests: speech technology, phonetics, linguistics, phonetic analysis of speech. The number of publications — 30. postmaster@phonetics.spbu.ru, www.phonetics.spbu.ru; SPbSU, Universitetskaya emb., 11, St. Petersburg, 199034, Russia; office phone +7(812)328-9565.

РЕФЕРАТ

Евдокимова В.В. Анализ спектра гласных на основе неравномерной психоакустической шкалы эрбов для определения словесного ударения.

Рассматривается возможность определения в слове ударного слога путем анализа особенностей формантной картины. Приводится описание известных признаков ударности и работ авторов, обративших внимание на особенности формирования формант при ударении.

Для проведения исследования используется специально сформированный массив трехсложных логатомов, в которых каждый из трех слогов состоит из определенных гласной и согласной, а ударение поочередно делается на первый, второй и третий слог. Для каждого слога, как ударного, так и безударных строится спектр гласной, причем делается это в шкале эрбов, чтобы учесть присущее речи стремление диктора оптимизировать речевую посылку для лучшего ее восприятия слушателем. Полученный спектр разбивается по эрбам и определяются соответствующие каждому участку частичные дисперсии. Для удобства работы с материалом проводится сравнение частичных дисперсий по каждому эрбу между слогами логатома и последующее кодифицирование этих дисперсий по величине. Полученный массив кодов позволяет при соответствующей группировке по гласным и по ударности слога провести сравнение и выявить особенности, присущие ударному слогу.

Делается вывод о том, что признак ударности по особенностям формантной картины существует и может быть использован для дополнения действия общепринятых признаков.

SUMMARY

Evdokimova V.V. Vowel spectral analysis using the psychoacoustic ERB scale for the word stress detection.

The paper presents a possibility to determine the stressed syllable in the word by examining features of the vowel formant structure. Based on the analysis of the array of logatoms the vowel formant structure is identified. The energy characteristics of individual formants of different vowels in one word are compared. Spectrum is split by the vowels psychoacoustic ERB scale. For ease of handling of the material resulting array of partial dispersions is codified. To determine the stressed vowel the resulting array codes are broken on the basis of the stressed/unstressed vowel characteristic and phoneme type. Partial arrays comparison enables to confirm the existence of the stress feature characteristic and identify phonetic features of this phenomenon, existing in the vowel formant structure. The vowel stress characteristic is defined by the vowel formant pattern and can effectively complement the determination of stressed syllables being added to conventional stressed vowel characteristics.