
СИСТЕМЫ АНАЛИЗА РЕЧИ

УДК 57.087.1

М. В. ХИТРОВ, А. Ю. ВАСИЛЬЕВ

СТАТИСТИЧЕСКИЕ ЯЗЫКОВЫЕ ОСОБЕННОСТИ И ИХ ГЕНДЕРНЫЕ РАЗЛИЧИЯ НА ПРИМЕРЕ ЛИТОВСКОГО ЯЗЫКА

Выявлены речевые особенности, позволяющие решать задачи автоматической идентификации языка и идентификации диктора. Предложен метод, использующий статистические параметры, характеризующие мелодический контур фраз исследуемого языка.

Ключевые слова: речевые технологии, статистические аспекты языка, литовский язык.

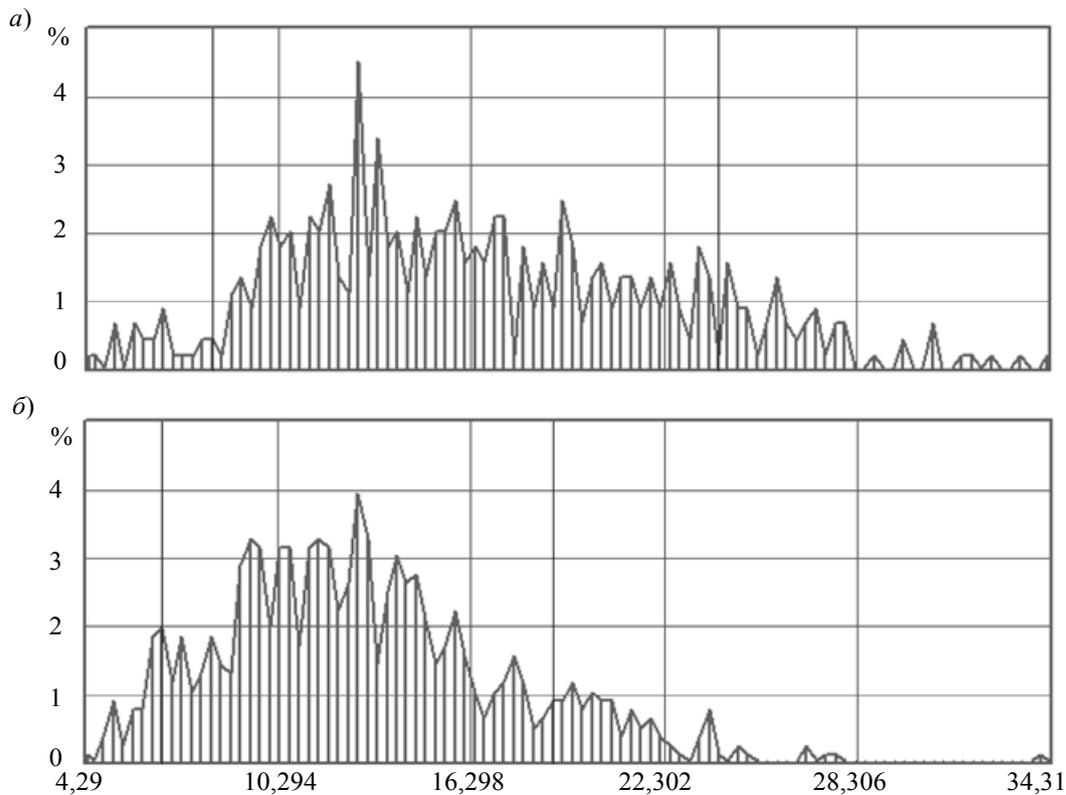
Введение. Определение речевых особенностей языка диктора — одна из актуальных задач современных речевых технологий. Одним из методов ее решения является сопоставление статистических характеристик основного тона говорящего и выявление типичных для выбранного языка интервалов статистических параметров различных речевых фрагментов [1—3]. Важно отметить, что границы данных интервалов и частота появления определенных речевых фрагментов существенно зависят от пола говорящего. В настоящей работе проанализированы статистические особенности языка и гендерные различия характеристик и речевых фрагментов на основе литовского языка.

База исследований. Исследование проводилось на 78 фонограммах литовских дикторов (на 55 фонограммах представлены 15 мужчин, на 23 — 7 женщин). Фонограммы содержали как литовскую, так и русскую речь. Обработка материала включала получение файлов основного тона, его минимальную коррекцию, а также подготовку таблиц, содержащих числовые значения параметров характерных интонационных единиц речи последовательностей фраз (длительность речевого абзаца 10—20 с). В качестве характерных интонационных единиц использовались синтагмы и их структурные элементы (ядерные слоги и ядерно-заядерные участки, шкалы, предшкалы [4—6]). Для анализа выбирались фрагменты речи, содержащие утвердительные высказывания, имеющие характер завершенности или незавершенности и одинаковую эмоциональную окраску.

В качестве параметров рассматриваемых речевых абзацев и локальных фрагментов выступали максимальная, минимальная и средняя частоты фрагмента (в герцах), частотный интервал фрагмента (в герцах и полутонах), скорость изменения тона (в полутонах в секунду), коэффициент изрезанности фрагмента [1, 2].

Анализ статистических характеристик. Значения каждого параметра для всех исследуемых речевых абзацев всех дикторов одного пола строятся в виде диаграмм. Специально разработанное программное обеспечение позволяет определять для полученного распределения основные статистические характеристики, включая математическое ожидание, среднеквадратичное отклонение, а в качестве конечного результата — интервал типичных значений с доверительной вероятностью 95 %.

Ниже приводятся два примера таких распределений параметра „интервал в полутонах“ для случаев литовских дикторов-женщин (рисунок, а) и мужчин (б). На рисунке отчетливо видны установленные границы доверительного интервала, соответствующие диапазону типичных значений характеристик для литовского языка.



Анализ распределений параметров показал, что наиболее близкий к нормальному характер, как у женщин, так и у мужчин, имеют изрезанность, максимальная частота и скорость изменения тона. Значения интервалов, минимальной и средней частот имеют распределение, промежуточное между нормальным и равномерным. Стоит отметить ряд аномальных пиков как вне, так и в области доверительного интервала, для женщин и средней частоты у мужчин, а также существенно выходящие за пределы доверительного интервала значения ряда статистических характеристик у мужчин (см. рисунок, б). Это может быть объяснено дефектом записанной фонограммы, процент таких отклонений входит в допустимую статистическую погрешность (табл. 1—4).

Таблица 1

**Интервалы типичных значений параметров
для литовского языка (мужчины)**

Параметр	Область значений (доверительный интервал 95 %)
Максимум, Гц	129,9—217,4
Минимум, Гц	63,9—103,6
Интервал-Гц	41,7—136,8
Интервал-Пг	6,7—18,8
Средняя частота, Гц	90,1—143,7
Скорость изменения тона	-9,1—1,3
Экссесс	-0,9—1,7
Скос	-0,15—1,68
Изрезанность	24,3—49,3

Таблица 2

**Интервалы типичных значений параметров
для тюркских языков (мужчины)**

Параметр	Область значений (доверительный интервал 95 %)
Максимум, Гц	133,6—218,8
Минимум, Гц	66,3—97,4
Интервал-Гц	49,8—131,6
Интервал-Пг	8,3—17,7
Средняя частота, Гц	95—141,9
Скорость изменения тона	-12,1—-0,07
Эксцесс	2,3—4,6
Скос	-0,17—1,47
Изрезанность	27,5—50,25

Таблица 3

**Интервалы типичных значений параметров
для литовского языка (женщины)**

Параметр	Область значений (доверительный интервал 95 %)
Максимум, Гц	225,8—395,7
Минимум, Гц	102,9—164,6
Интервал-Гц	88,4—254,5
Интервал-Пг	8,7—19,6
Средняя частота, Гц	156,3—242
Скорость изменения тона	-7,3—1,5
Эксцесс	-0,8—2,8
Скос	0,03—1,9
Изрезанность	26,6—46,4

Таблица 4

**Интервалы типичных значений параметров
для тюркских языков (женщины)**

Параметр	Область значений (доверительный интервал 95 %)
Максимум, Гц	255,1—383,4
Минимум, Гц	124,4—196,4
Интервал-Гц	96,3—221,7
Интервал-Пг	8,2—16,3
Средняя частота, Гц	177,7—251,8
Скорость изменения тона	-9,3—0,52
Эксцесс	2,5—5,3
Скос	0,09—1,6
Изрезанность	28,3—45,2

Определенные интервальные значения явно демонстрируют гендерные различия: речевые фрагменты женского голоса имеют тенденцию к более высоким частотам, чем у мужского голоса. Также для речи женщин характерен более широкий диапазон частот. Прочие статистические характеристики в целом совпадают для обоих полов, их особенности проявляются при сравнении с другими языками.

Для сравнительной оценки типичных значений параметров литовского языка использованы усредненные типичные значения параметров узбекского и азербайджанского языков. Интервалы типичных значений для обоих тюркских языков получены на материале речевых абзацев большой длительности, полученном от 59 носителей азербайджанского языка (13 женщин, 46 мужчин) и 84 носителей узбекского языка (49 женщин, 35 мужчин). Сравнение проводилось отдельно для мужчин и для женщин.

В результате установлено, что для мужчин-литовцев по сравнению с тюркоязычными дикторами минимальные значения частоты основного тона (ЧОТ) варьируют в более широком диапазоне, максимальные значения ЧОТ также более вариативны, но диапазон их варьирования лежит в низкочастотной области. Частотный диапазон (выраженный в герцах и в полутонах) у литовцев шире, чем у узбеков и азербайджанцев. Коэффициент изрезанности для литовцев ниже, чем для носителей узбекского и азербайджанского языков, что указывает на более ровную интонацию во фразах литовцев.

Для литовских женщин-дикторов максимальные значения ЧОТ варьируют в более широком диапазоне по сравнению с тюркоязычными женщинами, а минимальные лежат в низкочастотной области и варьируют в более узких пределах. Судя по значениям параметра „средняя частота“, голоса литовских женщин в целом ниже, чем голоса тюркоязычных. Скорость изменения тона у литовских женщин выше, чем у узбекских и азербайджанских; значение коэффициента изрезанности у них варьирует в более широком диапазоне.

Кроме сравнения статистических параметров литовского языка со значениями для конкретной языковой группы — тюркской — было проведено исследование общих усредненных параметров по всем имеющимся данным для различных языковых групп. Результаты показали следующие особенности. Максимальные значения характерных частот литовского языка ниже средних характеристик по всем языкам базы, особенно ярко это выражено у мужчин, поэтому в связи с тем, что значения минимальных частот в целом соответствуют усредненным данным, характерный частотный интервал у литовцев меньше (у мужчин — примерно на четверть). Средняя скорость изменения тона для литовского языка невысока, у мужчин отклонения от средних значений достигают 50 %. Изрезанность характерных речевых фрагментов свидетельствует о более плавных, в сравнении со средними значениями, изменениях частоты основного тона. В целом литовскому языку присуща плавность речи в сравнительно узком частотном диапазоне. Для речи дикторов-мужчин типичны существенные отклонения от усредненных характеристик по всей речевой базе, женщины демонстрируют большее соответствие. Поскольку база содержит характеристики различных языковых групп, трудно делать выводы о близости характеристик для литовского языка и родственных ему языков балтийской ветви.

Анализ речевых фрагментов. Анализ набора интонационных структур показал, что отличительной особенностью речи женщин является более частое использование восходящего мелодического контура.

Так, в приведенном материале (табл. 5) в речи мужчин на русском и литовском языках встретилось одинаковое количество реализаций синтагм с восходящим и нисходящим завершением, в то время как в речи женщин синтагмы с восходящим завершением реализуются значительно чаще, чем с нисходящим (363 и 189 реализаций соответственно).

Таблица 5

Тип	Подтип	Фрагменты речи		Дикторы	
		мужчины	женщины	мужчины	женщины
Синтагма	Нисходящее завершение	598	189	15	7
	Восходящее завершение	598	363	15	7
Ядерно-заядерный участок	Нисходящий	475	142	15	7
	Восходящий	415	195	15	7
	Нисходяще-восходящий	83	101	12	7
	Восходяще-нисходящий	58	18	13	6

Та же закономерность наблюдается в реализации ядерно-заядерных участков синтагм в речи литовских дикторов. Мужчины наиболее часто используют нисходящую интонацию ядерно-заядерного участка, восходящая интонация используется реже. Нисходяще-восходящее и восходяще-нисходящее оформление встречается значительно реже. Женщины чаще реализуют восходящий контур ядерно-заядерного участка, нисходящая интонация используется ими реже. В речи женщин значительно чаще, чем у мужчин, реализуется нисходяще-восходящая интонация ядерно-заядерного участка.

Заключение. По результатам проведенных исследований можно сделать следующие выводы. Литовский язык по своим речевым особенностям тяготеет к более низким частотам, что особенно выражено в речи мужчин. Для него также характерны достаточно узкий частотный диапазон и плавность изменения частоты основного тона. Женская и мужская речь в литовском языке различается не только очевидным тяготением мужских голосов к низким частотам, а женских — к высоким, но и тем, что статистические параметры мужского голоса сильнее отличаются от средних для различных языковых групп, а также тем, что женщины демонстрируют стремление к интонации незавершенности в утвердительных высказываниях. Сделанные выводы справедливы для утвердительных высказываний одинаковой эмоциональной окраски.

В настоящей работе впервые представлены возможности использования метода статистического анализа мелодического контура речи в различных задачах, включая задачи идентификации языка и диктора. Логичное продолжение исследований — сбор и анализ данных не только по отдельным языкам, но и по значимым языковым группам, поскольку это должно повысить точность результатов и послужить первым шагом в разработке инструментальных программных средств анализа особенностей языка. Точность анализа и идентификации параметров будет различаться для языка внутри своей языковой группы и языков разных групп. Другим направлением может служить анализ речевых фрагментов различной эмоциональной окраски.

Работа выполнена при государственной финансовой поддержке ведущих университетов Российской Федерации (субсидия 074-U01).

СПИСОК ЛИТЕРАТУРЫ

1. Коваль С. Л., Лабутин П. В., Пеховский Т. С., Процина Е. А., Смирнова Н. А., Таланов А. О. Методика идентификации дикторов по голосу и речи на основе комплексного анализа фонограмм // Тр. Междунар. конф. „Диалог 2007“. М., 2007. С. 256—262.
2. Смирнова Н. А. Идентификация дикторов на основе сравнения параметров реализации мелодических контуров высказываний // Тр. Междунар. конф. „Диалог 2007“. М., 2007. С. 502—507.
3. Хитров М. В. и др. Фоноскопическое исследование фонограмм речи: Исследование достоверности фонограмм. СПб: Изд-во „Юридический центр-Пресс“, 2011. Кн. I. 281 с.
4. O'Connor J., Arnold G. Intonation of colloquial English. London: Longman, 1973.
5. Брызгунова Е. А. Интонация // Русская грамматика / Н. Ю. Шведова (гл. ред.). М.: Наука, 1980. С. 96—122.
6. Светозарова Н. Д. Интонационная система русского языка. Л.: Изд-во ЛГУ, 1982.

Сведения об авторах

- Михаил Васильевич Хитров** — канд. техн. наук; ООО „ЦРТ“, Санкт-Петербург; генеральный директор; Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кафедра речевых информационных систем; заведующий кафедрой; E-mail: khitrov@speechpro.com
- Андрей Юрьевич Васильев** — ООО „ЦРТ“, Санкт-Петербург; программист; E-mail: vasilyev-a@speechpro.com

Рекомендована кафедрой
речевых информационных систем

Поступила в редакцию
22.10.13 г.