Компьютерная лингвистика и интеллектуальные технологии:
по материалам международной конференции «Диалог 2017»

Москва, 31 мая — 3 июня 2017

ФОНЕТИЧЕСКИЕ ХАРАКТЕРИСТИКИ ДЫХАТЕЛЬНЫХ ПАУЗ С РАЗНОЙ ТЕКСТОВОЙ ЛОКАЛИЗАЦИЕЙ[1]

Кривнова О. Ф. (okrivnova@mail.ru)

Московский государственный университет
им. М. В. Ломоносова, Москва, Россия

Ключевые слова: фонетика, устная речь, просодическое членение, интонационно-смысловая дыхательная пауза, текстовая локализация, инструментальный анализ

PHONETIC CHARACTERISTICS OF BREATHING PAUSES WITH DIFFERENT TEXT LOCALIZATION

Krivnova O. F. (okrivnova@mail.ru)

Lomonosov Moscow State University, Moscow, Russia

In spite of its fundamental role in the processes of speech production, speech breathing is studied experimentally quite insufficiently. The main task of this paper was to obtain phonetic-acoustic data on the breathing pauses (BP) with different text localization in oral Russian texts. In the introduction we formulate the problems being discussed with the main goal to analyze the correlation of BP's acoustics with the boundaries of principal text units—paragraphs, sentences, clauses, taking into account the interspeaker variability in reading of the same text. In our previous studies it was supposed that the distinctive function of BP as prosodic markers of macrosegmentation is realized by the stable differences in their general phonetic patterns and such acoustic parameters as duration, intensity and noise spectrum of the inhalation phase. The quantitative analysis of these differences comprises the main part of the paper. The second section

[1] Исследование проведено при поддержке гранта РФФИ 15-06-06103

describes the material and the methods of research. In the third section the analysis of the general acoustic-physiological pattern of BP with different text localization is given, which allows to set the regular objective differences between them. The fourth section contains the data of BP's duration in connection with their text localization. It is shown that BP's text localization is the main factor to control its duration. In the fifth section the similar information is presented for the average BP's intensity. The sixth section contains an analysis of the spectral characteristics of the inhalation noise in BP. It is shown that it is the most invariant feature for all BP types, almost independent of its text localization. The conclusion summarizes the study's results emphasizing the fact that the presence of BP in any oral text can serve as a sufficient indication of its structure, but interspeaker variability shows that BP's realization is not a necessary feature of the text boundary.

Key words: phonetics, spoken language, prosodic phrasing, breathing pause, text localization, interspeaker variability, instrumental analysis

1. Введение

Механизм речевого дыхания, несмотря на свою фундаментальную роль в процессах речепроизводства, в экспериментальном плане до сих пор изучен мало. Наиболее обстоятельные результаты в этой области были получены американским фонетистом П. Ладефогедом. Основные результаты обобщены в (Ladefoged 1967), где речевое дыхание рассматривается в разных аспектах — физиологическом, аэродинамическом, акустическом, и во взаимодействии с другими артикуляционными процессами. Нельзя не отметить, к сожалению, что физиологическая сторона речевого дыхания по-прежнему мало доступна для прямого анализа в естественных условиях. Однако в последнее время появились исследования, в которых используются компьютерные методы детектирования и анализа дыхательных пауз в устной речи, как при чтении, так и в спонтанном режиме. Так, в работе (Johansson et al. 2005) идентификация ДП основана на спектральных эталонах и алгоритме динамического программирования. Авторы отмечают, что алгоритм работает лучше на мужской речи по сравнению с женской, а в целом, результаты сильно зависят от степени зашумленности и интенсивности сигнала. В работе (Braunschweiler, Chen 2013) ДП идентифицируются с использованием скрытой Марковской модели с 3-мя состояниями, которая обучается на речевом корпусе и учитывает 3 акустических параметра: ЧОТ, среднюю амплитуду и спектр. Также показано, что добавление акустической модели ДП в систему синтеза речи на немецком языке значительно повышает ее естественность, особенно при озвучивании больших текстов, в частности аудиокниг. Оказалось также, что введение информации о локализации ДП в модель языка при распознавании речи, основанном на Марковских моделях, улучшает точность распознавания (Weilhammer, Schiel 1999).

На материале русского языка подобных исследований не проводилось. В связи с этим задача настоящей работы заключалась в том, чтобы получить более детальные акустико-фонетические сведения о ДП с разной текстовой локализацией для русской устной речи. Нас интересовала как общая

кустико-физиологическая картина ДП, так и их конкретные акустические характеристики: длительность, интенсивность и спектр шумовой части ДП (вдоха).

Вопрос о текстовой локализации ДП подробно обсуждался в (Кривнова 2016), поэтому ниже указаны кратко лишь наиболее важные результаты проведенного нами исследования, а основное внимание будет уделено анализу фонетических характеристик ДП учетом их текстовой локализации и ряда других факторов.

Анализ текстовой локализации ДП на материале прочтения одного и того же текста разными дикторами позволяет сделать следующие выводы:

1. Главным фактором, который влияет на организацию дыхания в репродуцированной речи, является стратегия интонационного паузирования диктора, для которой характерна тенденция к реализации темпоральных пауз после каждой клаузы в предложении. Однако эта достаточно регулярная тенденция взаимодействует с когнитивными характеристиками дикторов: некоторые дикторы в определенных синтаксических условиях «пропускают» конечные границы произносимых клауз, в то время как другие часто реализуют темпоральные паузы в определенных точках внутри произносимой клаузы.
2. В организации речевого дыхания находит отражение иерархическая структура текстовых единиц, основание которой образуют отдельные клаузы.
3. Текстовые фрагменты упорядочиваются по убыванию вероятности последующего вдоха следующим образом (в скобках дается частота реализации вдоха в среднем по 10 дикторам): абзац (100%) > самостоятельное предложение внутри абзаца (94%) > клауза внутри предложения (65%) > компонент внутри клаузы (34%).
4. Полученные результаты подтверждают мнение многих исследователей о центральной роли клаузы и текстовых границ более высокого ранга в процессах порождения, понимания и озвучивания русских текстов, и в том числе в организации речевого дыхания. Аналогичные результаты были получены и на материале других языков, причем не только для локализации ДП, но и для глубины вдоха и объема вдыхаемого воздуха в интервале ДП (Winkworth et al. 1994; 1995; Rochet-Capellan, Fuchs 2013). Для изучения физиологических параметров дыхания современные исследователи чаще всего используют миниатюрный респираторный плейтисмограф, регистрирующий сигналы дыхания на уровне грудной клетки и диафрагмы с одновременным использованием ротового микрофона. Данная технология подробно описывается в (Włodarczak et al. 2015). Авторы планируют применить ее для создания речевого корпуса разговорной речи с разметкой на речевые и дыхательные фрагменты и с разграничением разных фаз в ДП. Аналогичная техника уже использовалась на материале немецкого языка для исследования связи дыхательного ритма со сменой ролей в диалоге (Rochet-Capellan, Fuchs 2016). Полученные результаты довольно противоречивы: на протяжении диалога авторы не обнаружили отчетливой координации в дыхательной

активности участников, однако при этом выявили ряд признаков, которые показывают, что локально смена реплик связана с вдохами говорящего на стадии подготовки реплики. А в целом, как и в других, более «плавных» формах устного дискурса (чтении и спонтанном монологе), в диалоге имеется тенденция к координации речевого дыхания с важными дискурсивно-текстовыми событиями и их таймированием.

2. Материал и методика настоящего исследования

В качестве материала использовался корпус прочтений небольшого рассказа о посещении научного учреждения[2]. Текст был прочитан «с листа»[3] десятью дикторами, носителями русского языка с высшим образованием, но без специальной дикторской подготовки; средняя длительность озвученного текста 3–3,5 минуты. Материал записывался на компьютер (SR 22050 Гц, 16-bit, Mono) в условиях тихой комнаты с использованием высокочувствительного микрофона, что позволило в большинстве случаев выявить ДП в каждом прочтении текста, при некотором локальном усилении сигнала на темпоральной паузе.

Материал был отобран из большего массива, включавшего 30 прочтений текста разными дикторами (суммарный объем исходного речевого массива около 400 Мб). Для дальнейшего анализа было выбрано 10 наилучших прочтений, среди которых удачно оказалось 5 мужских и 5 женских — далее они обозначаются соответственно -m-i и f-i, где i меняется от 1 до 10 и обозначает место, которое занял диктор в десятке лучших текстовых прочтений.

Дыхательное заполнение пауз определялось на слух и визуально по осциллограммам с использованием анализатора Speech Analyzer SA SIL, версия 1.5-2002 и SA 3.1-2007. Акустические характеристики ДП и других речевых отрезков измерялись с помощью того же анализатора с учетом предварительной «ручной» сегментации речевого материала. Статистическая обработка инструментальных данных производилась средствами Excel MS-2000.

3. Общая фонетическая картина дыхательных пауз с разной текстовой локализацией

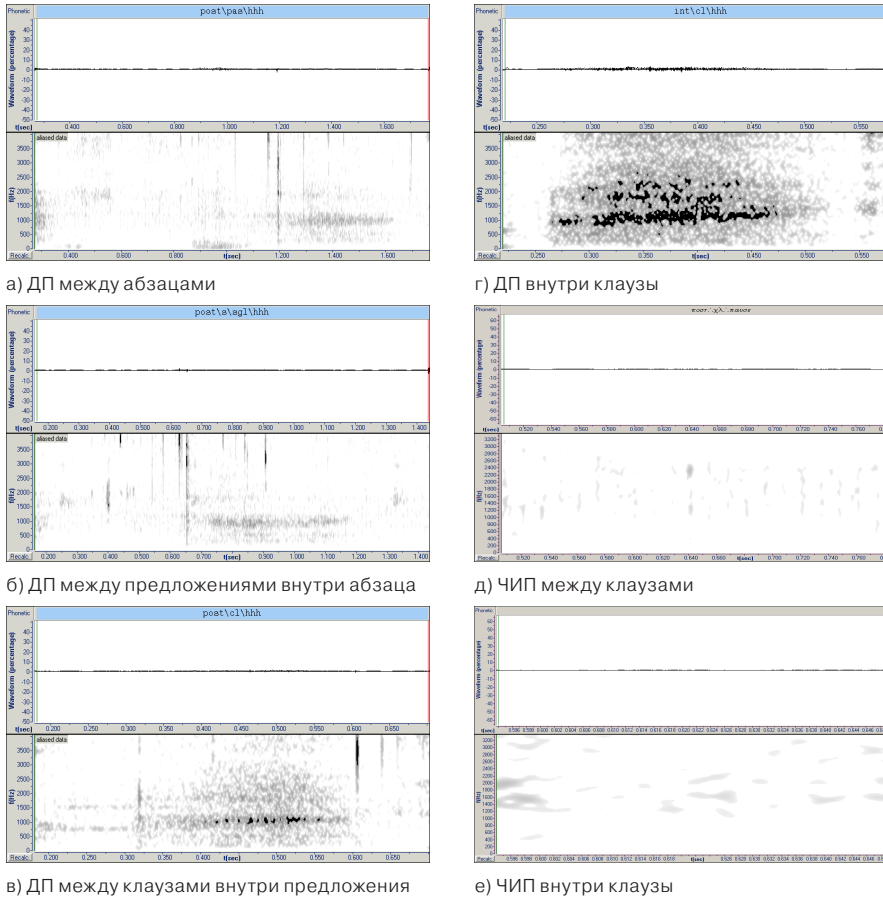
На рис. 1 представлены осциллограммы и спектрограммы ДП с разной текстовой локализацией в прочтении экспериментального текста диктором-женщиной f-2. Отметим, что по экспертному рейтингу это второе наилучшее прочтение в экспериментальном массиве. Аналогичные данные для наилучшего мужского прочтения того же текста приведены в работе (Кривнова 2010). Для

---

[2] Текст был взят из методической разработки по составлению текстовых массивов [(Штерн 1984), а в качестве основы для него использовался отрывок из книги С. Иванова «Схватка с роботом». М., 1977.

[3] Материал для эксперимента был любезно предоставлен московской компанией Stel Computer Systems, ведущей разработку систем автоматического распознавания речи для русского языка.

сравнения на рис. 1 даны также иллюстрации для «чистых» интонационных пауз (ЧИП) с тождественной текстовой локализацией.



а) ДП между абзацами

г) ДП внутри клаузы

б) ДП между предложениями внутри абзаца

д) ЧИП между клаузами

в) ДП между клаузами внутри предложения

е) ЧИП внутри клаузы

Рис. 1. Диктор f-2. Акустическая картина дыхательных пауз (ДП) — слева. ЧИП — чистая интонационная пауза (без элементов вдоха) — справа. В иллюстративных целях рисунки масштабированы по оси Y в соотношении 1x2, а затем приведены к одной высоте, а также приведены к одной длине по оси X. Паузы были выделены курсорами без захвата соседних участков речевого сигнала.

Даже простой визуально-слуховой анализ материала, аналогичного представленному на рис. 1, приводит к заключению, что ДП с разной текстовой локализацией имеют различное акустико-физиологическое наполнение, которое регулярно воспроизводится в прочтениях разных дикторов. А именно:

1. В ДП между абзацами отчетливо выделяются три фазы: сначала носовой вдох, которому может предшествовать краткий выдох, затем достаточно резкий ротовой шумный вдох, которому обычно предшествуют явления чмоканья, сглатывания и под. (это т. н. артикуляционный шум). На спектрограммах ДП они видны хорошо и реализуются подобно взрывам смычных согласных. За ротовым вдохом идет короткий относительно «немой» участок, на котором тоже может присутствовать артикуляционный шум. Спектрограммы обнаруживают заметные различия в интенсивности и спектре шума на носовой и ротовой фазах вдоха в ДП. В (Garman 1990), со ссылкой на исследования речевого дыхания (Butcher 1981), также говорится о трехфазной структуре ДП с выделением 3 участков pre-inhalation, inhalation, post-inhalation, с разными акустическими характеристиками. Отмечается, что временное соотношение фаз ДП различно для спонтанной речи и чтения, при этом фаза ротового вдоха практически постоянна и равна 400 мс, что, возможно, объясняется чисто физиологическими факторами.

   Предполагается, что ДП большей длительности оказываются таковыми за счет двух других фаз — предшествующей и следующей за вдохом. Именно эти фазы увеличиваются по длительности в спонтанной речи, что позволяет, по мнению автора, ввести в интегральную оценку плавности речи такой показатель, как *доля собственно вдоха в общей длительности ДП*. Однако, как показывает наш материал, этот показатель зависит не только от формы речи, но и от текстовой локализации ДП (см. ниже).
2. ДП между самостоятельными предложениями внутри абзаца характеризуются фонетической картиной, сходной с ДП между абзацами. Возможно, есть некоторые различия во временных характеристиках носовой и ротовой фаз ДП, что, в свою очередь, может быть связано с различиями в общей длительности ДП между и внутри абзаца.
3. В ДП между клаузами внутри предложения, как правило, отчетливо выражена только шумная фаза ротового вдоха, а явления чмоканья в ее начале менее заметны и встречаются реже, чем в ДП более высокого текстового уровня. Кроме того, шум на ротовой фазе вдоха имеет существенно бóльшую интенсивность.
4. ДП внутри клаузы демонстрируют дальнейшее нарастание явлений, отмеченных для ДП после клаузы внутри предложения. Так, темпоральная пауза практически полностью заполнена шумом ротового вдоха, инициальных явлений чмоканья не наблюдается, шум вдоха очень интенсивен и отчетливо фиксируется на спектре ДП, аналогично шумным фрикативным согласным.

Что касается гендерных различий, то при сходстве общей фонетической картины ДП определенные различия наблюдаются в выраженности и интенсивности шума вдоха, особенно в ротовой фазе. В среднем, дикторы-мужчины в своих текстовых прочтениях дышат более шумно, чем дикторы-женщины.

4. Длительность дыхательных пауз с разной текстовой локализацией

Мысль о том, что длительность интонационных пауз, включая ДП, отражает иерархическую структуру текста, высказывалась уже давно (Волконский 1913; Светозарова1982). Однако конкретных данных в литературных источниках немного.

Нас интересовала прежде всего средняя длительность ДП с различной текстовой локализацией в прочтениях одного текста разными дикторами, а также соотношение длительностей ДП и ЧИП с аналогичной текстовой локализацией. Междикторская и внутридикторская вариативность длительности текстовых пауз в настоящей работе не была предметом специального анализа, так как общее количество темпоральных пауз с разной текстовой локализацией в прочтении каждого отдельного диктора было сравнительно невелико (30–60). Некоторое представление о вариативности можно получить из показателей стандартного отклонения от средних значений паузальных длительностей (табл.1). Длительность пауз определялась как в абсолютных единицах (мсек), так и в условных, нормированных относительно темпа (средней длительности слога) в прочтении конкретного диктора. Подробные данные о длительности пауз разного типа представлены ниже в таблице 1 и в более наглядной форме на рис. 2 для нормированных длительностей ДП и ЧИП.

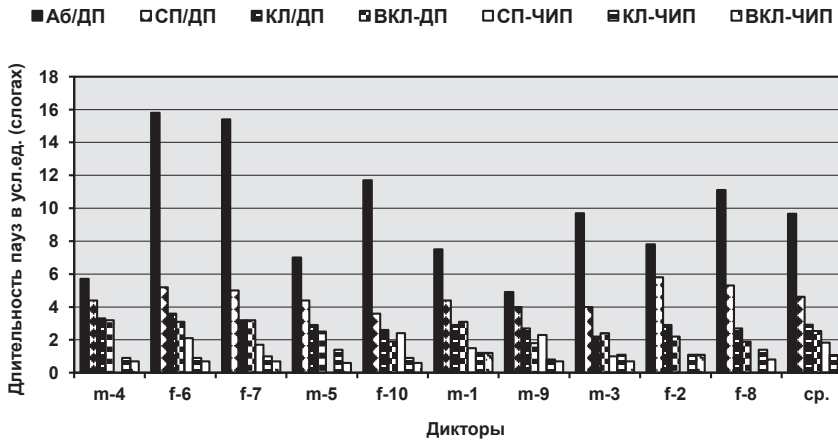


Рис. 2. Нормированная длительность ДП и ЧИП с разной текстовой локализацией в прочтениях разных дикторов, в среднем и по отдельным дикторам. Обозначения текстовых границ: Аб — между абзацами, СП — между предложениями внутри абзаца, КЛ — между клаузами внутри самостоятельного предложения, ВКЛ — внутри клаузы.

Таблица 1. Абсолютная (мс) и нормированная (в слогах) длительность ДП и ЧИП с различной текстовой локализацией. Данные упорядочены по темпу (средней длительности слога) в прочтении диктора (в сторону возрастания); по строкам m-5 — m-1 проходит примерная граница между быстрым и замедленным темпом произнесения относительно нейтрального.

Дикторы	Ср. длительность слога	Аб-ДП		СП-ДП		КЛ-ДП		ВКЛ-ДП		СП-ЧИП		КЛ-ЧИП		ВКЛ ЧИП	
		мс	сл	мс	сл	мс	сл	Мс	сл	мс	сл	мс	сл	мс	сл
m-4	0,141	0,796	5,7	0,624	4,4	0,462	3,3	0,455	3,2	–	–	0,126	0,9	0,099	0,7
		0,376	2,7	0,193	1,4	0,123	0,9	0,055	0,4	–	–	0,084	0,6	0,046	0,3
f-6	0,142	2,236	15,8	0,741	5,2	0,516	3,6	0,444	3,1	0,295	2,1	0,132	0,9	0,105	0,7
		0,475	3,3	0,193	1,4	0,102	0,7	0,048	0,3	–	–	0,114	0,8	0,079	0,6
f-7	0,146	2,247	15,4	0,727	5,0	0,446	3,2	0,467	3,2	0,248	1,7	0,145	1	0,106	0,7
		1,209	8,3	0,156	1,1	0,121	0,8	0,166	1,1	0,158	1,1	0,113	0,8	0,079	0,5
m-5	0,153	1,076	7,0	0,670	4,4	0,448	2,9	0,381	2,5	–	–	0,212	1,4	0,092	0,6
		0,210	1,4	0,165	1,1	0,107	0,7	0,033	0,2	–	–	0,143	0,9	0,070	0,5
f-10	0,165	1,941	11,7	0,592	3,6	0,434	2,6	0,309	1,9	0,398	2,4	0,148	0,9	0,104	0,6
		0,502	2,3	0,129	0,9	0,089	0,5	0,033	0,2	–	–	0,096	0,6	0,058	0,4
m-1	0,167	1,251	7,5	0,739	4,4	0,488	2,9	0,523	3,1	0,256	1,5	0,196	1,2	0,196	1,2
		0,249	1,5	0,253	1,5	0,206	1,2	0,265	1,6	0,110	0,7	0,098	0,6	0.097	0,6
m-9	0,174	0,858	4,9	0,695	4,0	0,463	2,7	0,321	1,8	0,398	2,3	0,132	0,8	0,117	0,7
		0,133	0,8	0,293	1,1	0,122	0,7	0,128	0,7	–	–	0,084	0,1	0,112	0,6
m-3	0,176	1,711	9,7	0,703	4,0	0,382	2,2	0,421	2,4	0,178	1,0	0,189	1,1	0,123	0,7
		0,430	2,4	0,168	1,0	0,081	0,5	0,166	0.9	0,069	0,4	0,095	0,5	0,070	0,4
f-2	0,177	1,381	7,8	1,025	5,8	0,511	2,9	0,397	2,2	–	–	0,187	1,1	0,187	1,1
		0,245	1,4	0,350	2,0	0,122	0,7	0,128	0,7	–	–	0,059	0,0	0,060	0,3
f-8	0,187	2,068	11,1	0,998	5,3	0,495	2,7	0,354	1,9	–	–	0,261	1,4	0,144	0,8
		0,133	0,2	0,520	2,8	0,159	0,9	0,091	1,5	–	–	0,165	0,9	0,082	0,4
Ср. длительность паузы		1,557	9,7	0,751	4,6	0,465	2,9	0,407	2,5	0,296	1,8	0,173	1,1	0,127	0,8
т. отклонение		0,557	3,8	0,147	0,7	0,040	0,4	0,068	0,6	0,088	0,5	0,044	0,2	0,037	0,2

Анализ полученных данных приводит к следующим выводам:

1. Дыхательные паузы имеют бóльшую длительность, чем чистые интонационные паузы независимо от текстовой категории последних. Это выполняется как в среднем, так и по отдельным дикторам. Однако для пауз одинаковой текстовой категории это соотношение выдерживается более отчетливо. Реализация вдоха увеличивает длительность соответствующей ЧИП условно на 2–2,5 слога (~ 300–400 мс).
2. Текстовые фрагменты упорядочиваются по убыванию длительности последующей ДП следующим образом: в скобках дается средняя нормированная длительность ДП: абзац(~10 слогов) > самостоятельное предложение внутри абзаца(~5 слогов) > клауза внутри предложения (~3 слога) > компонент внутри клаузы (~3 слога). Таким образом, предположение о зависимости длительности ДП от ее текстовой категории в целом подтверждается, хотя в этом отношении ДП между клаузами и внутри клауз

как самые короткие среди ДП, по длительности могут пересекаться, причем в различной степени у разных дикторов. То же самое наблюдается и для ЧИП. На рис. 3 приводится также формула, которая позволяет количественно оценивать данную зависимость. R2-показатель, который отражает степень соответствия наблюдаемых и расчетных данных по длительности ДП в зависимости от ее текстовой категории, близок к 1, что свидетельствует о хорошем согласовании данных.

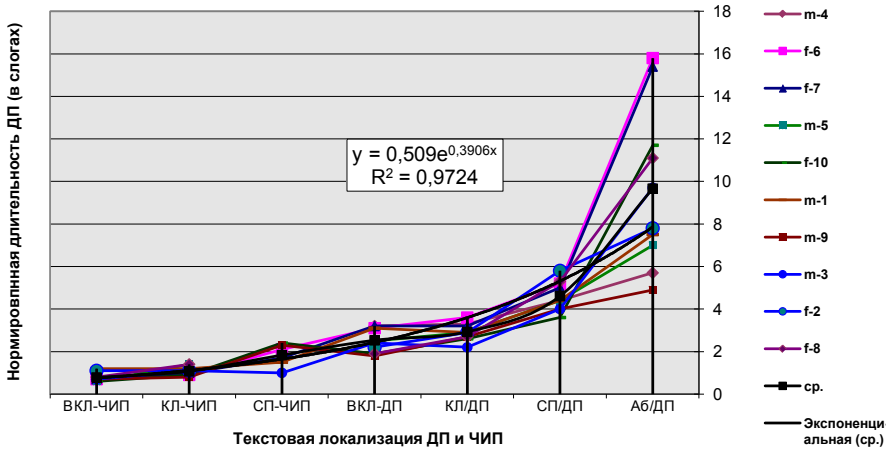


Рис. 3. Зависимость длительности ДП от ее текстовой локализации в среднем и по разным дикторам. Условные обозначения текстовых границ см. выше на рис. 2.

3. Дополнительный анализ показал, что текстовая локализация ДП является главным фактором, определяющим ее длительность. На нашем материале не обнаружено (рис. 4) значимых корреляций между длительностью ДП и таким фактором, как *длина* предшествующей и последующей дыхательных групп (ДГ), который часто указывается как контролирующий длительность ДП при порождении речи (Conrad et al.1983; Henderson et al. 1965; Mitchell et al. 1996).
4. Не обнаружено и корреляции с длительностью предшествующей дыхательной *паузы*, см. рис. 5, которую можно было бы ожидать, исходя из принципа экономии дыхательных усилий, на что указывается в работе (Златоустова 1968).

---

[4] В (Rochet-capellan, Fuchs 2013) отмечается также , что длительность ДП в спонтанной речи зависит не только от длины последующей ДГ, но и от синтаксического типа первой клаузы в ней.

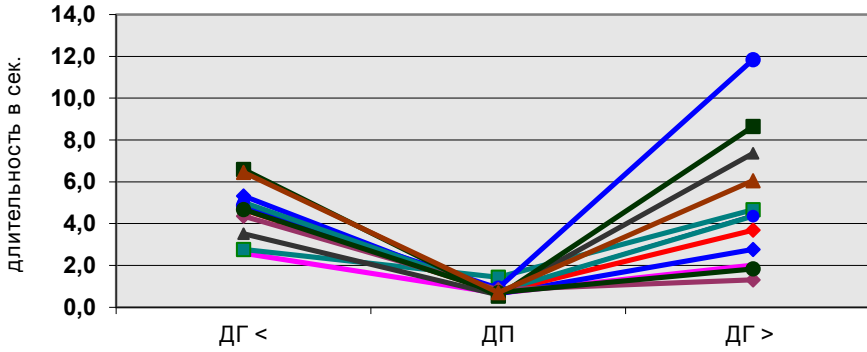


Рис. 4. Зависимость длительности ДП между самостоятельными предложениями внутри абзаца от длины соседних ДГ в сек. (диктор m-1). Условные обозначения: ДГ1 < дыхательная группа перед ДП на границе между предложениями, ДГ2 > — дыхательная группа после ДП на границе между предложениями.

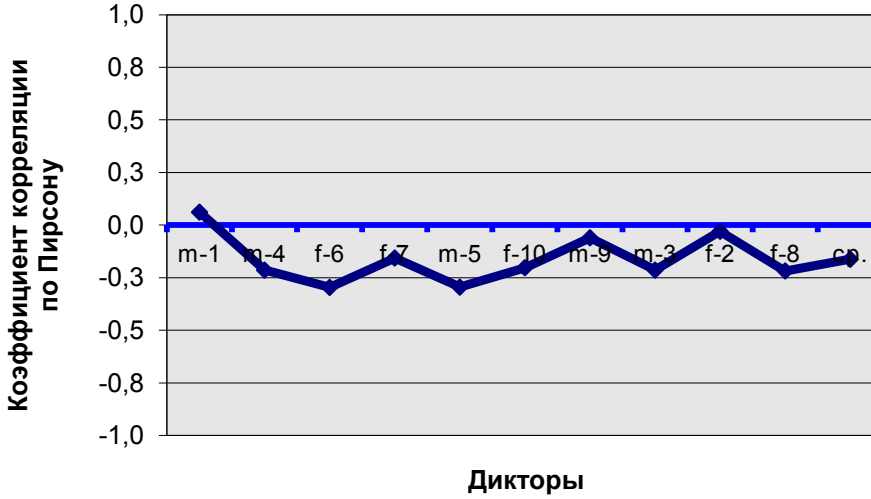


Рис. 5. Корреляция между длительностями соседних ДП (в среднем и по разным дикторам). Коэффициенты корреляции (по Пирсону) близки к 0, что не позволяет говорить о какой-либо корреляционной зависимости между длительностями соседних ДП.

5. Интенсивность на отрезке ДП с разной текстовой локализацией

Особенности общей картины ДП, о которых говорилось выше, позволяют предположить, что интенсивность акустического наполнения ДП также может отражать их текстовую локализацию. Как видно из рис. 6, это предположение подтверждается данными. Понижение ранга текстовой границы, совмещенной с вдохом, в направлении от абзацной ДП к внутриклаузальной, сопровождается увеличением *среднего* уровня интенсивности сигнала на интервале ДП, правда, весьма незначительным — до 10 дБ при усреднении по 10 дикторам. Более детальные данные показывают, что это происходит как за счет сокращения общей длительности ДП, так и за счет усиления шумовой части вдоха.

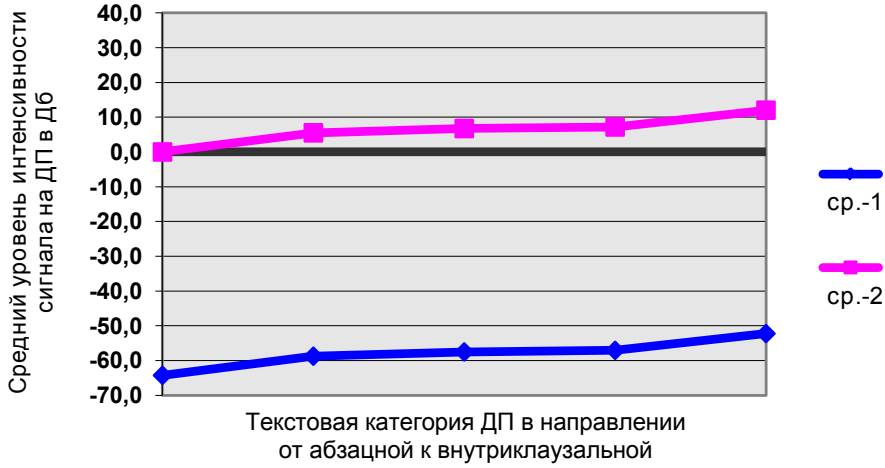


Рис. 6. Средний уровень интенсивности акустического наполнения ДП относительно максимального уровня записи (ср-1) и относительно уровня фона в начале записи (ср-2).

6. Спектральные характеристики шума вдоха на отрезке ДП с разной текстовой локализацией

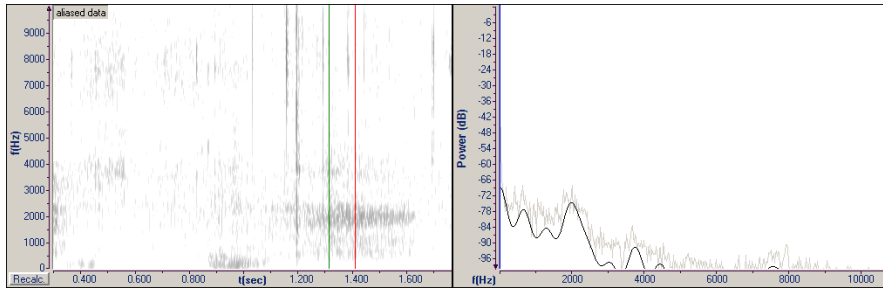
На спектрограммах рис. 1, видно, что ДП содержат участок шума приблизительно одинаковой длительности и разной интенсивности, спектр мощности которого находится в низкочастотной области с выраженным максимумом (формантой) на частоте 2000 Гц. В целом, такая спектральная картина характерна для фрикативных согласных заднего места образования, что естественно для шума речевого вдоха. Более точные спектральные данные можно получить из спектральных срезов на участке шума. Выборочные иллюстрации на рис. 7 подтверждают наблюдения об определенной инвариантности спектральных характеристик шума для ДП с разной текстовой локализацией. По спектрограммам видно также, что при уменьшении ранга текстовой локализации ДП от абзацной к внутриклаузальной увеличивается частотный диапазон шума за счет его верхней границы.

7. Заключение

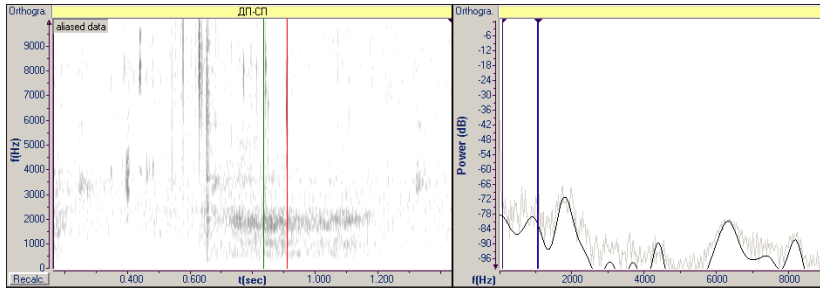
Устойчивые различия в общей фонетической картине и физических характерстиках ДП с разной текстовой локализацией создают возможность детектирования ранжированных границ между смысловыми отрезками текста как в естественном устном дискурсе, так и в задачах автоматической обработки речи, по крайней мере в режиме чтения. Отчетливая зависимость между акустическими и текстовыми показателями наблюдается для длительности: чем выше ранг текстовой границы, тем больше длительность ДП. Аналогичная зависимость, но менее яркая, характерна для среднего уровня сигнала на интервале ДП. Спектр шума в ДП («ротового» вдоха) — наиболее инвариантная характеристика для всех ДП, практически независимая от текстовой локализации. Этот акустический параметр может использоваться как достаточно надежный детектор ДП, что продемонстрировано на материале английского языка (Johansson et al. 2005). Корреляция ширины частотного диапазона шума ДП с разной текстовой локализацией требует отдельного исследования на более представительном речевом материале.

В итоге можно сказать, что, реализация вдоха в интонационной паузе является достаточным признаком наличия смысловой текстовой границы, а различия в длительности и интенсивности шума в ДП с разной текстовой локализацией несут информацию о смысловой и синтаксической структуре текста.

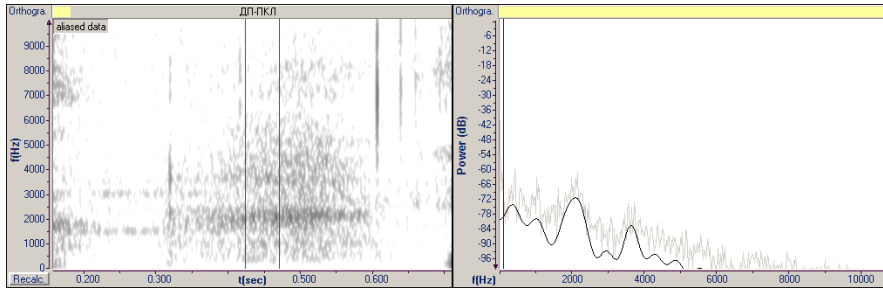
При этом нужно иметь в виду, что в общем случае *отсутствие* ДП *не* является абсолютно надежным признаком отсутствия текстовой границы. Так, даже в режиме чтения некоторые дикторы в определенных текстовых условиях не делают вдохов между самостоятельными предложениями внутри абзаца (Кривнова 2016).



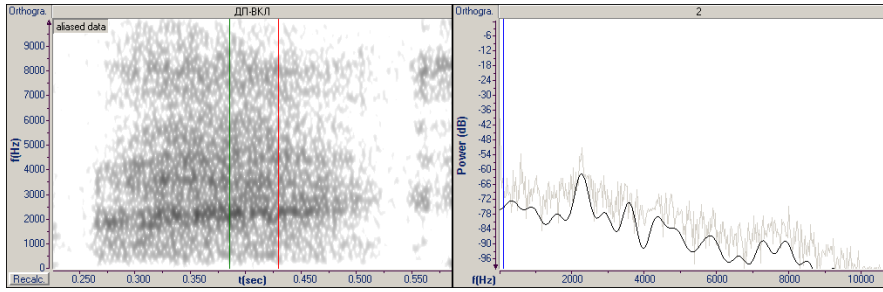
а) ДП-Абзац



б) ДП-СП



в) ДП-ПКЛ



г) ДП-ВКЛ

Рис. 7. Спектральные характеристики шума вдоха на ДП с разной текстовой локализацией (слева динамические спектрограммы, справа спектральные срезы, полученные на временных фреймах, выделенных курсорами).

References

1. *Garman M.* (1990), Psycholinguistics. Cambridge Univ. Press.
2. *Braunschweiler N., Chen L.* (2013), Automatic detection of inhalation breath pauses for improved pause modelling in HMM-TTS // 8th ISCA Speech Synthesis Workshop. Barcelona, Spain.
3. *Butcher A.* (1981), Aspects of the speech pause: Phonetic correlates and communicative functions // Arbeitsberichte. N15. Kill.
4. *Johansson A., Frid J., Horne M.* (2005), Using cepstral coefficients for inhalation pause detection in spontaneous speech. Department of Linguistics and Phonetics. Lund University, Lund, Sweden. VINNOVA Language Technology Program.
5. *Krivnova O. F., Chardin I. S.* (1999), Pausing for automatic speech synthesis [Pauzirovanije pri avtomaticheskom sinteze rechi], Theory and practice of speech research (ARSO-99). Proceedings of the conference [Teorija i praktika rechevyh issledovanij (ARSO-99). Materialy konferentsii]. M.
6. *Krivnova O. F.* (2010), General phonetic pattern of breathing pauses in a spoken Russian text. (reading mode) [Obshchaja foneticheskaja kartina dyhatel'nyh pauz v zvuchashchem russkom tekste (v rezhime chtenija)], Phonetics and grammar: present, past and future: to the 50th anniversary of the scientific activities of S. K. Pozharitskoj [Fonetika I grammatika: nastojashcheje, proshedsheje I budushcheje: k 50-letiju nauchnoj dejatel'nosti S. K. Pozharitsoj]. Ser. Issues of Russian linguistics. V. XIII. MGU, M., pp. 61–70.
7. *Krivnova O. F.* (2016), Prosodic phrasing in spoken text: localization of breathing pauses [Prosodicheskoje chlenenije v zvuchashchem tekste: lokalizatsija dyhatel'nyh pauz], Computer linguistics and intellectual technologies. Proceedings of the annual international conference "Dialogue [Komp'juternaja lingvistika I intellektual'nyje tehnikigii. Materialy jezhegodnoj mezhdunarodnoj konferentsii 'Dialog'] M., RGGU, v.14,t.1, pp. 340–354.
8. *Ladefoged P.* (1967), Three Areas of Experimental Phonetics, Oxford UP.
9. *Rochet-capellan A., Fuchs S.* (2013), The interplay of linguistic structure and breathing in German spontaneous speech // *Interspeech-2013*, 2014–2018.
10. *Rochet-Capellan A., Fuchs S.* (2016), Take a breath and take the turn: how breathing meets turns in spontaneous dialogue // Philos Trans R Soc Lond B Biol Sci. 2014 Dec 19; 369(1658).
11. *Shtern A. S.* (1984), Articulation tables. Methodical recomendations for developing listening skills and testing of auditory function [Artikuljatsionnyje tablitsy. Metodicheskaja razrabotka dlja razvitija navykov audirovanija I testirovanija sluhovoj funktsii], L.
12. *Svetozarova N. D.* (1982), Intonation system of the Russian language [Intonatsionnaja sistema russkogo jazyka] L.
13. *Volkonsky S.*(1913), Expressive Word [Vyrazitel'noje slovo]. St. Petersburg.
14. *Winkworth A. L., Davis P. J., Adams R. D., & Ellis, E.* (1994), Variability and Consistency in Speech Breathing During Reading: Lung Volumes, Speech Intensity, and Linguistic Factors // J. of Speech, Language, and Hearing Research, 37 (6), pp. 535–556.

15. *Winkworth A. L., Davis P. J., Adams R. D., Ellis E.* (1995), Breathing Patterns During Spontaneous Speech // J. of Speech, Language, and Hearing Research, 38(1), 124–144.
16. *Weilhammer K., Schiel F.* (1999), Investigation of language structure by means of language models incorporating breathing and articulatory noise // CPhS-14, pp. 1649–1652.
17. *Włodarczak M., Heldner M., Edlund J.* (2015). Breathing in Conversation: an Unwritten History// Proceedings of the 2nd European and the 5th Nordic Symposium on Multimodal Communication pp. 107–112.
18. *Zlatoustova L. V.* (1968), Some remarks on speech breathing [Nekotoryje zamechanija o rechevom dyhanii], Researches on speech information [Issledovanija po rechevoj informatsii], M.