

РЕАЛИЗАЦИЯ МЕЖЪЯЗЫКОВЫХ РАЗЛИЧИЙ ИНТОНАЦИИ ЗАВЕРШЁННОСТИ И НЕЗАВЕРШЁННОСТИ В СИНТЕЗАТОРЕ РУССКОЙ И ПОЛЬСКОЙ РЕЧИ ПО ТЕКСТУ

[1]

**Б. Лобанов, **Б. Пьорковска, **Я. Рафалко, *Л. Цирульник*

**Объединённый институт проблем информатики НАН Беларуси, Минск*

*** Институт информатики Белостокского университета, Белосток*

Описана методика анализа языковых особенностей интонации речи. Приводятся результаты, полученные для польскоязычных и русскоязычных дикторов для стиля речи «чтение текста». Обсуждаются особенности реализации интонационных контуров завершённости и незавершённости, показаны их межъязыковые различия.

Введение

Данное исследование проведено в рамках создания единой системы синтеза речи по тексту для славянских языков [1] с использованием методов компьютерного клонирования голоса и дикции личности [2]. Для создания речевого материала русскоязычным и польскоязычным дикторам предлагалось начитать тексты приблизительно одинакового научного содержания. Объём текстов на каждом из языков составлял около 300 слов. Ниже приведены фрагменты используемых текстов.

Фрагмент русскоязычного текста:

«Спектр результирующего речевого сигнала может быть представлен в виде произведения спектра источника звука на передаточные функции речевого тракта и каналов передачи звуковых сигналов, плюс – спектры различного рода сигналов акустических помех».

Фрагмент польскоязычного текста:

«Widmo sygnału mowy obliczamy przez pomnożenie aparatu wokalnego, źródła mowy, stylu mowy, mnożymy to też razy mikrofon, czyli jakie występują tam wypaczenia i dodajemy do tego wszystkiego lokalną akustykę».

Полученные в результате записи звуковые файлы в wav-формате подвергались дальнейшему анализу.

1. Методика проведения исследования

Исходными данными для исследований являются фонограмма речи и соответствующая ей стенограмма записей - исходный текст. Этапы исследования интонационных характеристик для создания набора мелодических портретов показаны на рис.1.

Прежде всего, речевой материал прослушивается, из него удаляются неинформативные участки: шумы, вдохи, «лишние» звуки (например, «э-э-э»). Затем опытный фонетист, пользуясь пакетом SoundForge, прослушивает фонограмму и разбивает её на фонетические синтагмы, где под синтагмой понимается самостоятельная в интонационном смысле часть фразы или вся фраза. Решение о наличии конца синтагмы принимается на основе ряда признаков, таких как присутствие дыхательной паузы, комплексная реализация одного из возможных интонационных типов синтагмы, наличие определённой динамической структуры (контура силы звуков) и определённой ритмической структуры (контура длительности звуков). При членении фонограммы на синтагмы во внимание принимается присутствие знаков препинания в соответствующем ей тексте, а также некоторых других формальных признаков текста. В фонограмме границы синтагм отмечаются регионами, в стенограмме – знаком «/».

Далее определяется тип синтагмы, который зависит от наличия или отсутствия знаков препинания и союзов после синтагмы. Используются следующие обозначения типов синтагм: (.) – перед точкой, (,) – перед запятой, (-) – перед дефисом, (:) – перед двоеточием, (()) – перед открывающей скобкой, ()) - перед закрывающей скобкой, (^) – перед союзом, () – после синтагмы отсутствует знак препинания или союз. Тип синтагмы отмечается в стенограмме.

Затем каждая синтагма разбивается на акцентные единицы. Под акцентной единицей (АЕ) понимается фонетическое слово или группа слов, объединённых одним сильным ударением. В фонограмме границы АЕ отмечаются маркерами, в стенограмме – знаком «/». В стенограмме, кроме того, для каждой АЕ указываются сильное (обозначается знаком /+) и слабое (обозначается знаком -/) ударения. Фрагменты обработанных таким образом стенограмм для русского и польского языков показаны в таблицах 1 и 2 соответственно.

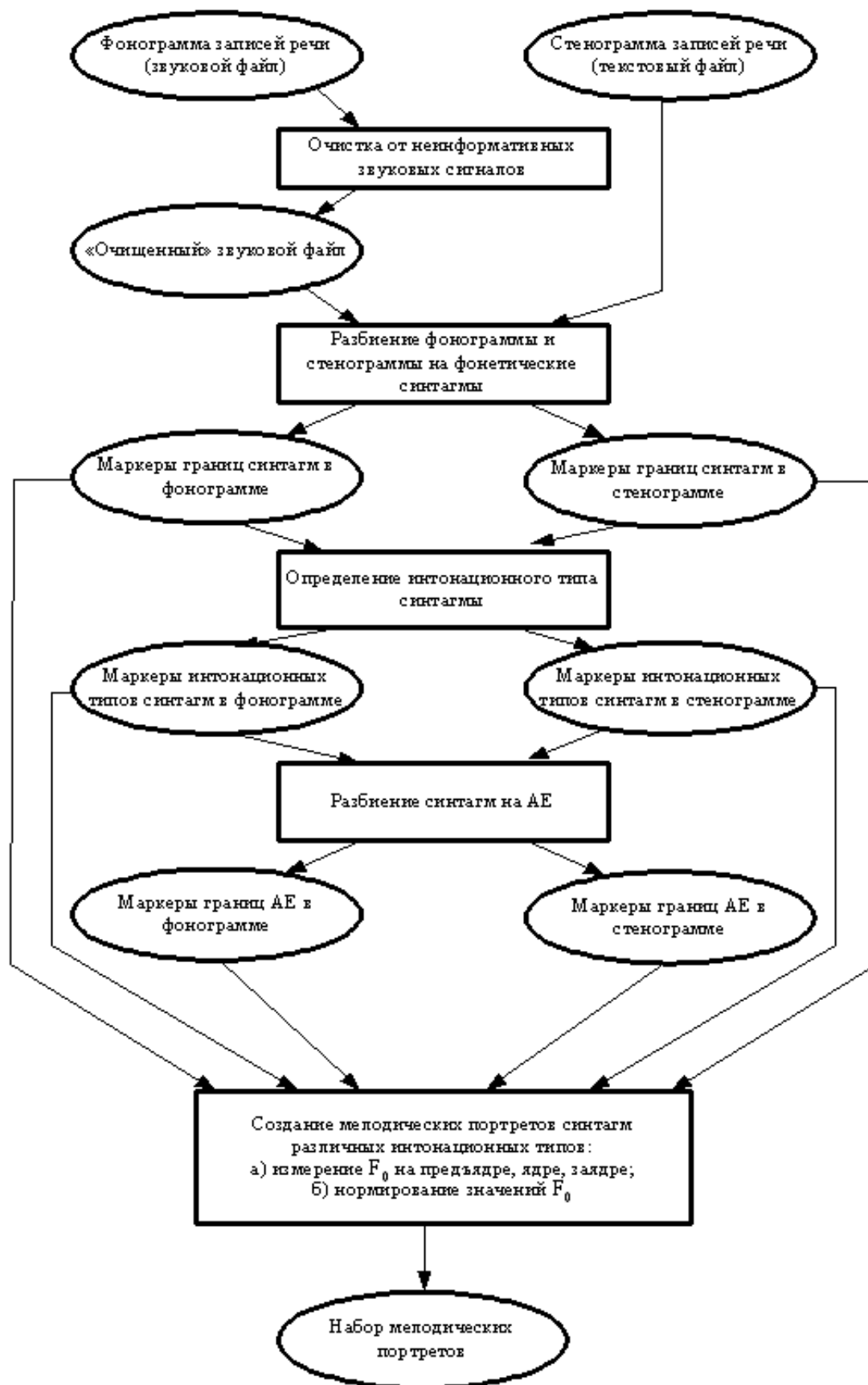


Рис. 1. Процедура создания набора мелодических портретов

№ п/п	Маркированный текст синтагмы	Кол. АЕ	Инт. тип
1.	/таки+м /о+бразом //	2	(.)
2.	/спе+ктр / речево+го /сигна+ла //	3	(.)
3.	/зави+сит/ не то+лько //	2	(.)
4.	/от его+/ смыслово+го/ /содержа+ния //	3	(.)
5.	/но+ и от/ ти+па/ /ре+чи //	3	(.)

6.	/индивиду+льных/ осо+бенностей/ произноше+ния/ ди+ктора //	4	(.)
7.	/индивиду+льных/ физиологи+ческих/ осо+бенностей/ ди+ктора//	4	(.)
8.	/характери+стик/ акусти+ческой/ среды+ //	3	(.)
9.	/в кото+рой/ /распростра+няется/ /речево+й/ /сигна+л//	4	(.)
10.	/и+ //	1	(.)
11.	/о+т/ характери+стик/ микрофо+на//	3	(.)
12.	/а та+кже//	1	(^)
13.	/и+ от/ характери+стик/ кана+ла/ свя+зи //	4	(.)

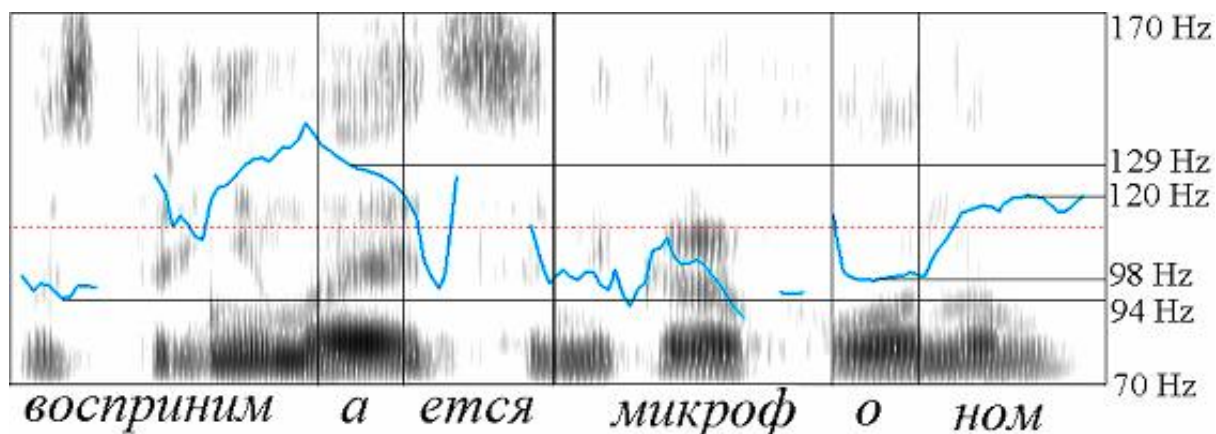
Таблица 1. Фрагмент деления текста на фонетические синтагмы и АЕ с указанием интонационного типа каждой синтагмы для русского языка

№ п/п	Маркированный текст синтагмы	Кол. АЕ	Инт. тип
1.	/oso+ba/ któ+ra/ mó+wi //	3	(.)
2.	/mo+że/ mó+wić //	2	(.)
3.	/z pe+wnym/ diale+ktem //	2	(.)
4.	/lub pod wpły+wem/ jaki+ś /emo+cji //	3	(.)
5.	/mo+że/ te+z/ wystą+pić/ loka+lna/ aku+styka //	5	(.)
6.	/na przy+kład //	1	(.)
7.	/rewerbera+cja //	1	(.)
8.	/czy+li/ po+głos//	2	(.)

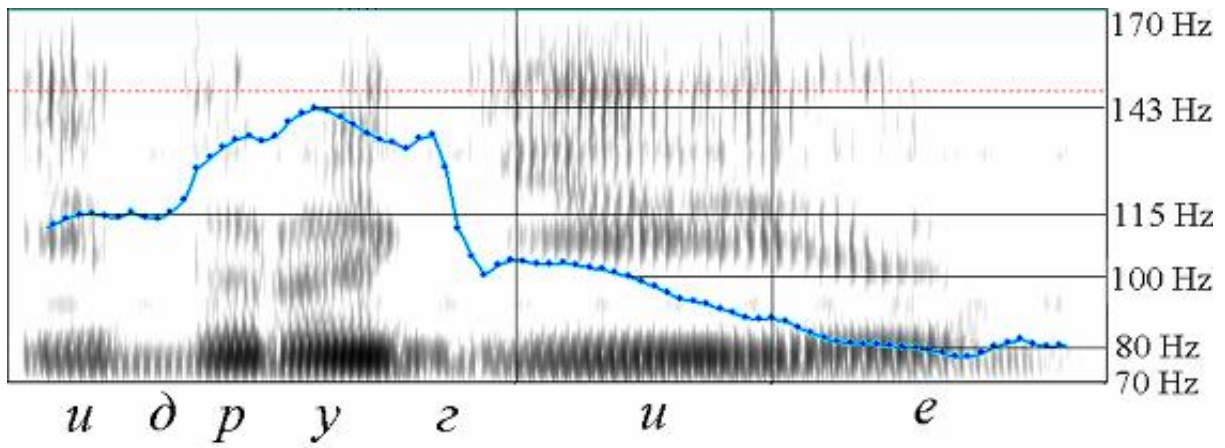
Таблица 2. Фрагмент деления текста на фонетические синтагмы и АЕ с указанием интонационного типа каждой синтагмы для польского языка

Следующий этап обработки – получение интонационных контуров синтагм – значений F_0 на вокализованных участках речи. Для этого используется программный пакет PRAAT. Каждая АЕ синтагмы предварительно делится на ядро – ударный гласный, предъядро – участок АЕ перед ударным гласным, и заядро - участок АЕ после ударного гласного. Примеры таких интонационных контуров с указанием значений на предьядре, ядре, заядре для синтагм русского языка: «воспринимается микрофоном» (интонационный тип синтагмы: /^/, 2 АЕ), «и другие» (интонационный тип синтагмы: /./) приведены на рисунке 2.

Примеры интонационных контуров для синтагм польского языка: «mikrofonu» (тип синтагмы: /^/, 1 АЕ) и «czyli roglós» (тип синтагмы: /./, 2 АЕ) приведены на рисунке 3.

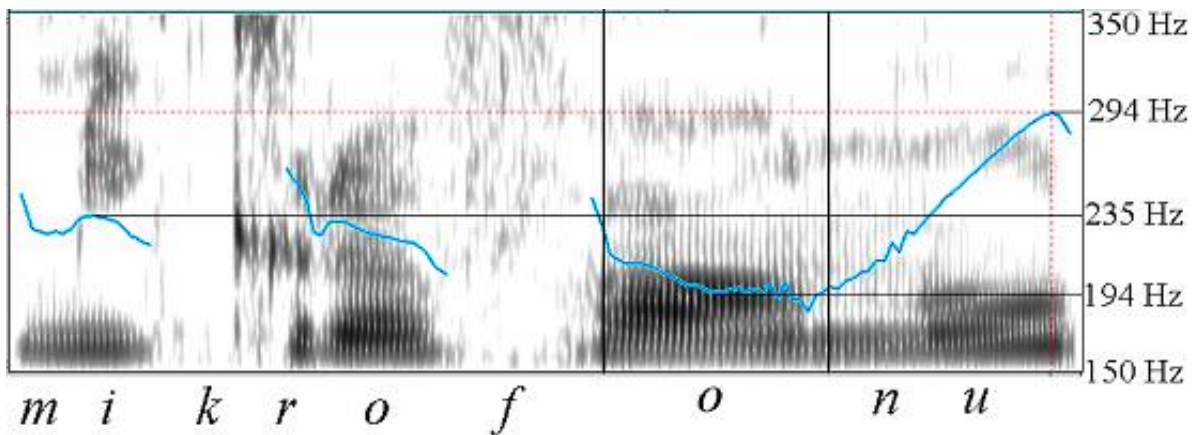


а)

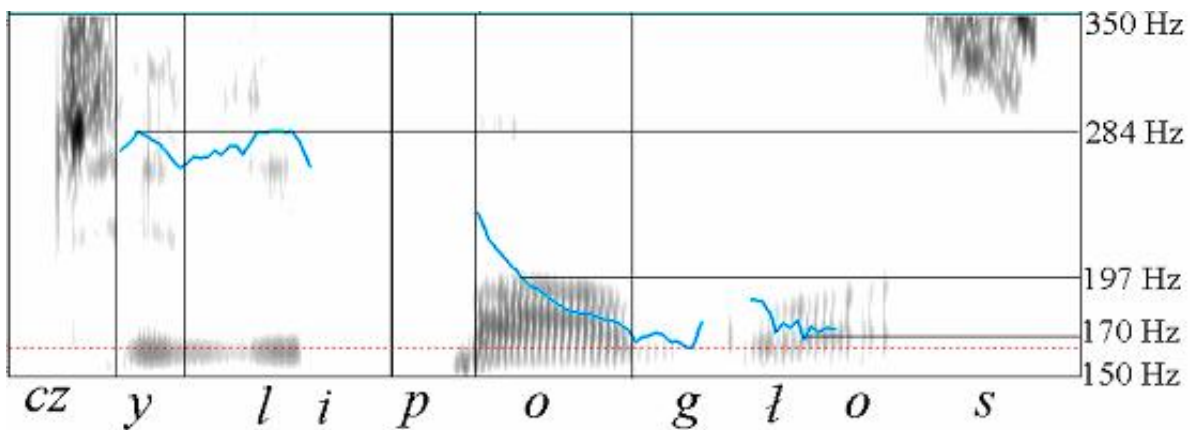


б)

Рис. 2. Примеры интонационных контуров для синтагм русского языка: а) «воспринимается микрофоном», б) «и другие»



а)



б)

Рис. 3. Примеры интонационных контуров для синтагм польского языка: а) «mikrofonu», б) «czyli pogłos»

2. Методика создания мелодических портретов АЕ

Методика представления мелодических портретов АЕ для синтеза речи по тексту была предложена ещё в 80-х годах [3, 4] и успешно используется до настоящего времени. Процесс создания мелодического портрета синтагмы показан на рис.4 на примере фразы “О чём написать Миле?”, произнесённой женским голосом. Данная фраза является синтагмой вопросительного типа первого подтипа, состоящей из 3 АЕ (обозначается Q1–3).

Для речевого сигнала (рис. 4а) вычисляются значения F_0 на всех вокализованных участках (рис.4б) и осуществляется их нормирование. Для этого определяются минимальное и максимальное значения F_0 на всём речевом участке исследуемой фонограммы. При анализе речи минимальное значение F_0 оказалось равным 180Hz,

а максимальное - 500Hz (см. рис.4б). Затем осуществляется разметка границ АЕ и границ областей предъядра, ядра, заядра для каждой АЕ (рис.4в), экстраполяция значений на невокализованных участках и нормирование каждой из областей по длительности (рис.4г).

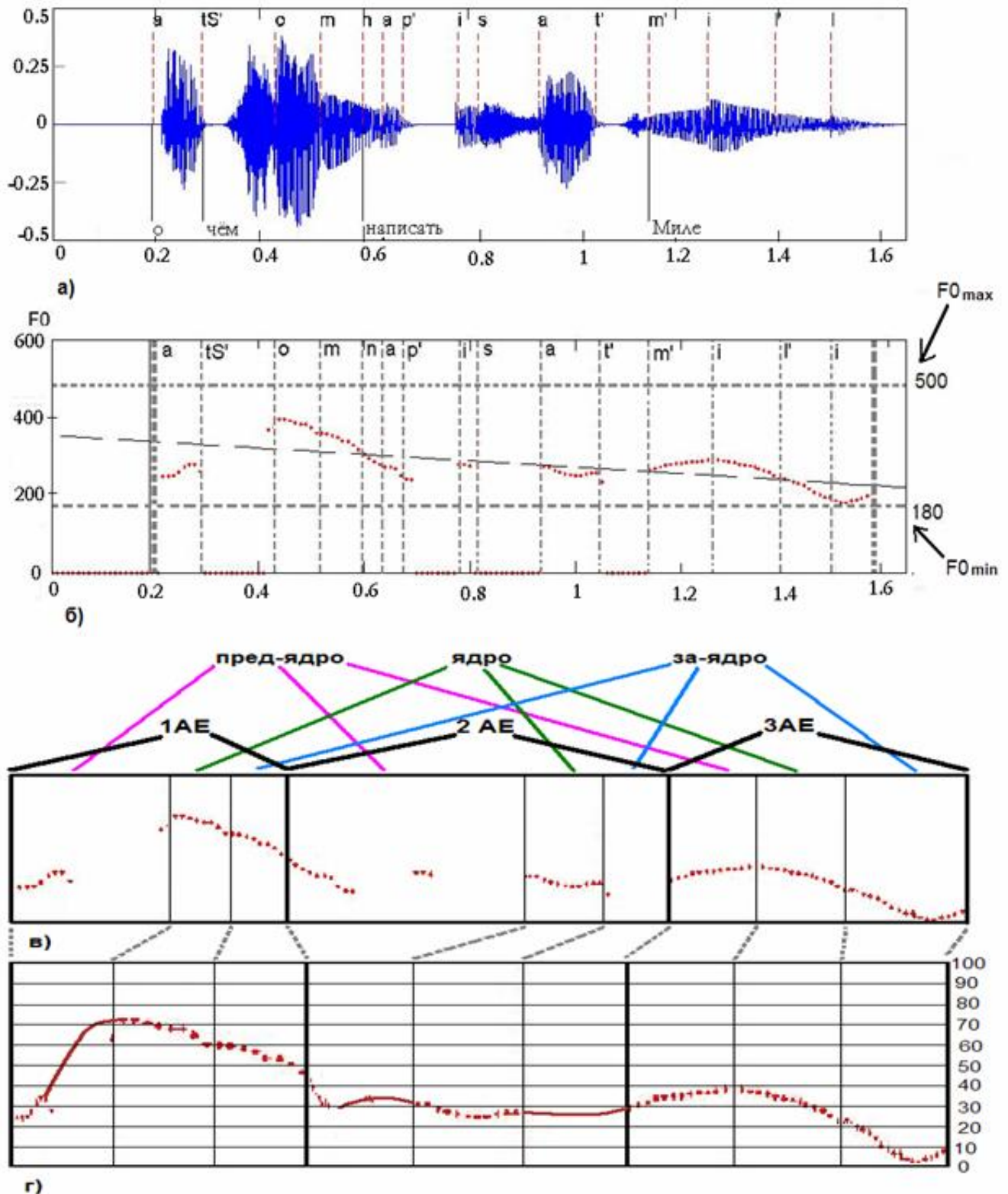


Рис. 4. Процесс создания мелодического портрета синтагмы типа Q1-3: а) речевой сигнал, б) значения F_0 , в) границы АЕ и участков предъядра, ядра, заядра, г) экстраполяция и нормирование контура F_0

3. Результаты исследования

Специфика исследуемого интонационного стиля - чтение научного текста- такова, что главное внимание

уделено наиболее «массовым» явлениям - интонации незавершённости и завершённости. Другие интонационные типы, такие как вопрос или восклицание, не рассматриваются. В соответствии с описанной в пп. 1, 2 методикой были построены мелодические портреты различных подтипов интонаций незавершённости и завершённости для русской и польской речи. Как хорошо известно, наибольшую информационную нагрузку несёт интонационный контур конечной АЕ синтагмы, в котором наиболее ярко проявляются особенности того или иного интонационного типа. На рис. 5 отражены наиболее типичные (наиболее частотные) закономерности реализации мелодических контуров конечной АЕ для интонаций незавершённости и завершённости в русской и польской речи. Как видно из рисунков, мелодические контуры как интонации незавершённости, так и завершённости весьма существенно отличаются в русской и польской речи. Интонация незавершённости, характеризующаяся в общем случае восходящим тоном, реализуется в русском языке, в основном, на ядерном участке АЕ, в то время как в польском – на заядерном участке. Аналогичная картина наблюдается и для случая интонации завершённости. Интонация завершённости, характеризующаяся в общем случае нисходящим тоном, реализуется в русском языке, в основном, на ядерном участке АЕ, в то время как в польском – на заядерном участке. Такое явление может быть легко интерпретировано, исходя из того факта, что в польском языке практически всегда присутствует заядерный участок слова (как правило, ударение падает на предпоследний слог), в то время как в русском языке заядерный участок слова очень часто может вообще отсутствовать (свободная позиция ударения).

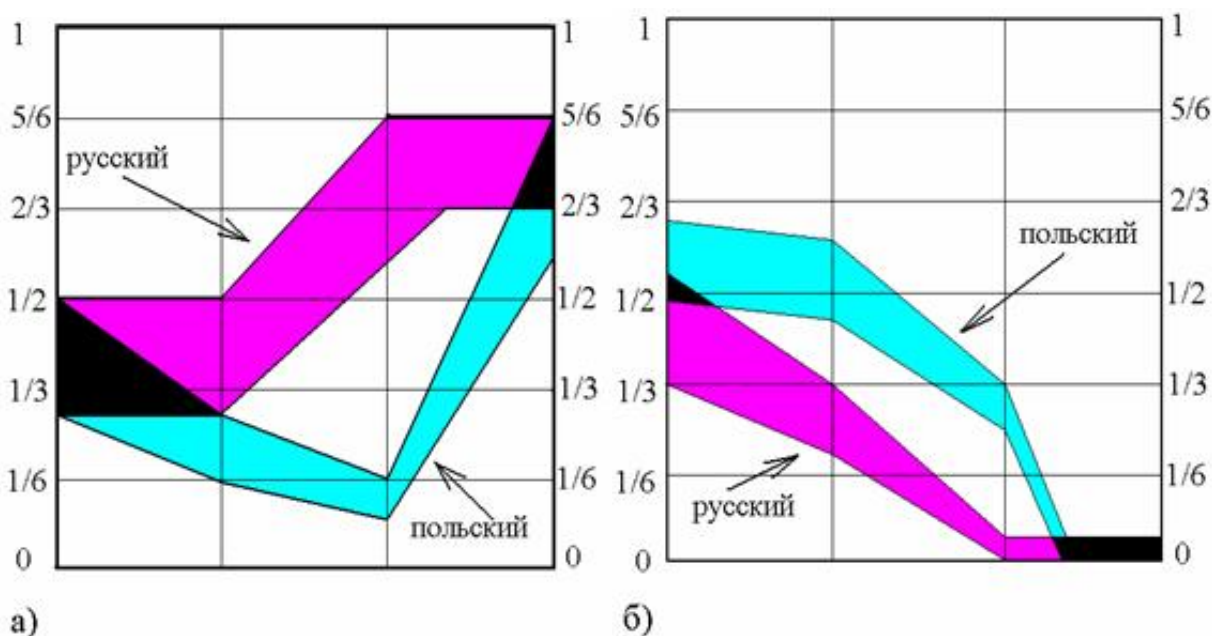
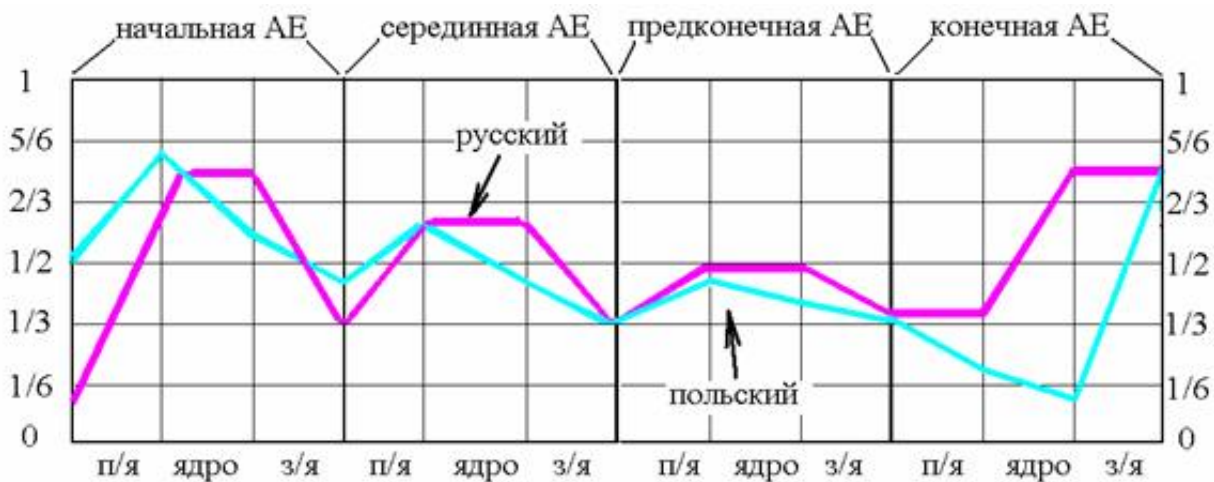
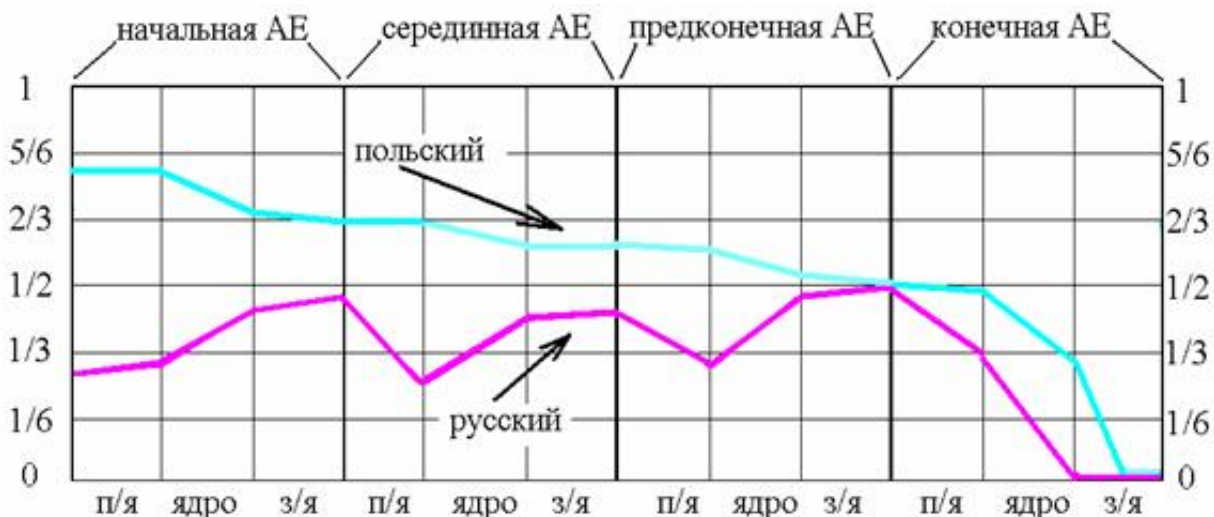


Рис. 5. Закономерности реализации мелодических контуров конечной АЕ: а) для интонации незавершённости, б) для интонации завершённости.

Закономерности реализации мелодических контуров для других АЕ (начальной, срединной и предконечной) в 4-хакцентных синтагмах для интонаций незавершённости и завершённости в русской и польской речи представлены на рис. 6. Из представленных графиков видно, что мелодические контуры русской и польской речи различаются не только на конечной АЕ синтагмы, но и, хотя и менее существенно, на начальной, срединной и предконечной АЕ.



а)



б)

Рис 6. Закономерности реализации мелодических контуров для 4-х акцентных синтагм: а) для интонации незавершённости, б) для интонации завершённости.

Заключение

Проведенное исследование позволило создать основной набор мелодических портретов интонации завершённости и незавершённости для синтеза русской и польской речи по тексту в стиле «чтение». За рамками данной работы остались вопросы реализации интонационных подтипов и правила их выбора, исходя из анализа структуры текста. Дальнейшее расширение исследований связано с охватом как можно большего набора интонационных конструкций, изучением особенностей их реализации, связанных с персональными дикторскими различиями, стилем речи и её эмоциональным наполнением.

Доклад иллюстрируется фонограммами русской и польской речи, синтезированной в соответствии с разработанными интонационными правилами.

Литература

- 1) Hoffmann R., Shpilewsky E., Lobanov B., Ronzhin A. Development of multi-voice and multi-language Text-to-Speech (TTS) and Speech-to-Text (STT) conversion system (languages: Belorussian, Polish, Russian) // Proc.of the Ninth International Conference "Speech and Computer" SPECOM'2004, Saint-Petersburg, 2004, p. 657-661.
- 2) Lobanov B.M., Tsiurulnik L.I. Phonetic-Acoustical Problems of Personal Voice Cloning by TTS // Proc.of the Ninth International Conference "Speech and Computer" SPECOM'2004, Saint-Petersburg, 2004, p. 17-21.
- 3) Lobanov B. The Phonemophon Text-to-Speech System // Proc. of the XI International Congress of Phonetic Sciences, Tallin, 1987, p. 61-64.
- 4) Boguslavsky I., Karnevskaia E., Lobanov B. Generation of Intonation and Accentuation of Synthetic Speech on the Basis of Morpho-Syntactic Knowledge // Proc.of International Workshop "Integration of Language and Speech", Moscow, 1995, p. 11-28.

[\[1\]](#) Данное исследование было проведено благодаря поддержке европейского фонда INTAS в рамках проекта «Разработка многоголосовой и многоязыковой системы синтеза и распознавания речи (языки: белорусский, польский, русский)» в соответствии с грантом INTAS- 04-77-7404