

Keywords in retrieval of prosodic data

Yanko T. E.

Institute of Linguistics,
Russian Academy of Sciences /
1 bld. 1 Bolshoy Kislovsky
Lane, 125009 Moscow,
tanya_yanko@list.ru

Abstract

Data retrieval can be based on the keywords found in the orthographic records of sound speech. This paper is aimed at analyzing the communicative meanings, compositions of meanings expressed by prosody, and the structure of segmental material aligned with prosodic variations. The meanings and their compositions expressed by prosody can have segmental counterparts, e.g., lexemes, or some other linguistic units. These counterparts, in turn, are used as keywords for retrieval. The results of searching are used as a material for analyzing the syntactic structure of the communicative components of sentences, such as themes and rhemes of statements, and the components of some specific illocutions (dreams, recollections, explanations), and combinations of illocutions with discourse incompleteness. The data source for this analysis is the Russian National Corpus (Multimodal sub-corpus Murko), Prosodically Annotated Corpus of Spoken Russian (SpokenCorpora.ru), and video-hosting Youtube. As the instrument, the computer system for sound speech analyzing Praat is used. The paper is illustrated throughout with pitch contours of sound records.

Keywords: sound speech; illocutions; data retrieval; keywords

DOI: 10.28995/2075-7182-2022-21-577-585

Метод поиска просодических данных по ключевым словам

Янко Т. Е.

Институт языкознания РАН
/ 125009, Москва, Большой
Кисловский пер. 1 стр. 1
tanya_yanko@list.ru

Аннотация

К поиску информации в массиве звучащих данных может быть применен метод ключевых слов. В работе анализируются коммуникативные значения, их композиции, которые выражаются просодией, и структура сегментного материала, несущего эти значения. Значения и их композиции, выражающиеся просодически, могут иметь свои сегментные корреляты, лексемы и другие единицы языка. Эти корреляты и используются в качестве ключевых слов поиска. Результаты поиска служат материалом для анализа синтаксической структуры коммуникативных компонентов предложений, таких, как темы и ремы сообщений, компоненты особых типов иллокуций (мечтаний, воспоминаний, обоснований), а также композиции иллокуций с дискурсивной незавершенностью. В качестве источника звучащего материала использован Мультимодальный подкорпус Национального Корпуса русского языка НКРЯ, корпус «Рассказы о сновидениях и другие корпуса звучащей речи» (SpokenCorpora.ru) и видео-хостинг Youtube. Инструментальный анализ данных проведен с помощью компьютерной системы анализа звучащей речи Praat. Работа иллюстрирована графиками изменений частоты звуковых данных.

Ключевые слова: звучащая речь; поиск информации; иллокуции; ключевые слова

1 Введение

В настоящее время исследователями просодии активно создаются просодически аннотированные звучащие корпуса. Для русского материала это корпус «Русский мультимедийный курс» [7] и «Рассказы о свидениях и другие корпуса звучащей речи» [9]. Аннотирование материала предусматривает ручную просодическую и иную разметку орфографической записи текста. Единицам текста приписываются частотные (в терминах частоты основного тона), паузальные, долготные и другие характеристики. Это создает возможность для поиска по движениям тона (подъемам, падениям и их комбинациям), длине пауз и другим параметрам.

Между тем при анализе просодии часто возникает необходимость поиска не по просодическим показателям, а по тем значениям, которые просодия выражает. Эти значения могут выражаться лексически, т.е. через сегментные единицы языка. Соответственно, мы предлагаем заменить поиск просодических маркеров значений на поиск по сегментным единицам языка: ключевым словам (или, дескрипторам). Ключевое слово — это традиционный термин поиска информации. Через ключевые слова и логические связи формулируется поисковый запрос, множество ключевых слов — это т.н. поисковый образ текста или массива текстов. В качестве поискового образа текста может быть использован и сам текст. Примеры ключевых слов в смысле подхода к поиску, предлагаемого здесь, мы приведем ниже, потому что коммуникативные значения представляют собой весьма абстрактные смыслы.

В соответствии с советом анонимного рецензента этой работы, мы должны отмежеваться от другого способа поиска информации в звучащем массиве речи, тоже называемого поиском по ключевым словам. При этом способом ключевыми словами объявляются собственно носители акцентных пиков на акцентной кривой [18] или слова, обозначающие эмоции [13], ибо слова, выражающие эмоции, по [13], могут иметь просодическое выражение. Таким способом через анализ просодической кривой автоматически индексируются транскрипты звучащих текстов. В дальнейшем это обеспечивает информационный поиск в массиве транскриптов. При нашем же подходе речь идет только о словах, непосредственно выражающих коммуникативные значения или косвенно на них указывающих. Коммуникативные значения формируют предложение как речевой акт и отвечают за встраивание речевого акта в связный нарратив или диалог, а применение метода ключевых слов происходит не в процессе индексирования текста, а на этапе формирования запроса.

Как могут быть использованы результаты поиска по ключевым словам, связанным с прагматическими значениями? Возьмем результаты поиска. Перед нами просодический материал с заданными заранее прагматическими параметрами речевых актов. Это фрагмент рабочего корпуса для проведения научного анализа: анализа синтаксиса речевых актов, семантики и способов наложения акцентной кривой на сегментный материал. Так, например, результаты поиска могут не только указать на словоформу — носитель определенного движения тона, но и представить достаточный материал для определения общих принципов выбора акцентоносителя коммуникативно значимой просодической структуры.

Таким образом, в этой работе предлагается рабочий метод оптимизации формулирования информационного запроса. Подбор ключевых слов предлагается производить методом интроспекции и с использованием традиционных для информационного поиска тезаурусных связей. При поиске используются ключевые слова, коррелирующие с коммуникативными значениями, а также их синонимы, родовидовые значения и ассоциаты. В дальнейшем результаты поиска подвергаются лингвистическому анализу, поэтому в работе, кроме обоснования методики поиска, говорится и о конкретных эпизодах использования его результатов.

Приведем пример. Пусть рассматривается просодический контур предложений, которые выражают воспоминания говорящего. Воспоминание — это особый иллокутивный тип предложения. Говорящий погружается в свой внутренний мир и воссоздает картины прошлого: они возникают перед его мысленным взором. Для выражения этого значения просодия предусматривает специальное средство. В русском языке это конструкция типа ИК-6 (по Е.А. Брызгуновой [1]), совершающаяся с существенным растяжением ударного слога словоформы-акцентоносителя.

При этом поиск планируется проводить не по ИК-6 (включая ее тональные варианты), а по ключевому слову *вспоминаю*, его синонимам и ассоциатам. Рассмотрим тональный рисунок ИК-6 с растяжкой:

1) *Мы заехали на поля-янку, на березовую, погуля-яли там* [9].

Говорящий вспоминает, какой ему снился сон. Акцентоноситель здесь и ниже выделен жирным шрифтом. Ключевых слов воспоминания в этом примере и в его контексте нет. Соответствующее значение выражается только просодией. Обратимся к другому примеру. Герой-вертолетчик, чудом оставшийся в живых после Чернобыльской аварии, вспоминает, как он тушил пожар в ядерном реакторе:

2) *И вот я завис над этим самым реа-актором, на высоте двести ме-етров, пятьсот рентген зашкаливает* [14].

Последний пример получен уже в результате поиска по ключевому слову «воспоминания», которое значилось в названии записи. В результате поиска в массиве данных, полученных по ключевым словам *воспоминания, мемуары, вспоминаю, помню*, возникают данные для синтаксического анализа сегментного материала с целью определения принципов выбора словоформы-акцентоносителя. В примерах 1) и 2) это словоформы *полянку, погуляли, реактором* и *метров*. Как мы увидим ниже, выбор этих словоформ в качестве носителя акцента ИК-6 (с растяжкой) не случаен, он согласуется с определенными принципами, которые были выделены ранее для предложений другого иллокутивного типа и с другим типом акцента [15-16].

При поиске по ключевым словам мы в целом придерживаемся традиционной структуры значений, отмечаемых в словарной статье информационно-поискового тезауруса [12]. Семантические отношения между заглавным ключевым словом и единицами, находящимися с ним в синонимических, родовидовых, ассоциативных (т.е. имеющих референцию к той же ситуации, что и заглавное ключевое слово, ср. *покупка* vs. *продажа*) отношениях и в отношении *часть* vs. *целое* нацелены, как известно, на увеличение полноты поиска информации. В качестве ключевых слов мы используем не только однословные именные группы и именные группы с зависимыми, как это делается традиционно при информационном поиске, но и языковые единицы других типов вплоть до короткого предложения. Соответственно, ассоциативную связь по сравнению с традиционным тезаурусным подходом мы понимаем расширительно, как имеющую ту же референцию, что и исходное ключевое слово.

В разделе 2 рассматривается вопрос *Что случилось?* как ключевое слово, его аналоги, просодии ответа на него или обоснования, почему такой вопрос задается, а также выбор словоформы-акцентоносителя в ответе. В разделе 3 мы возвращаемся к воспоминаниям и показываем, что воспоминания и мечты формируют акцентоноситель по принципам, аналогичным тем, которые дает ответ на вопрос *Что случилось?*. В разделе 4 рассматриваются ключевые слова, называющие жанры пресс-конференции и интервью. Это жанры, демонстрирующие композицию дискурсивной незавершенности с недостаточно исследованной ранее на этот счет иллокуцией *да-нет*-вопроса.

2 Вопрос *Что случилось?* как ключевое слово

Вопрос *Что случилось?* и его аналоги (*В чем дело?*; *Что происходит?*; *Что такое?*; *Ты что?*; *Ты чего?*; *Ты чё?*; *А что?*; *А чё?*; *А ты чё?*) имеют в качестве ответа или в качестве объяснения, почему задается этот вопрос, иллокуцию обоснования. Рассмотрим пример: *Дядя Вася Денисюк // Говорит: //– Послушай, друг, // Ты скажи мне, // что случилось: // На лице твоём испуг! // – Дядя Вася Денисюк, – // Отвечает старый друг, – // Ты скажи мне, // что случилось: // У тебя же дым вокруг!* (Л. Каминский). Дядя Вася задает вопрос *что случилось?*, потому что видит испуг на лице своего друга, а его друг задает вопрос *что случилось?*, потому что видит дым в квартире у дяди Васи. Ответы и объяснения, почему вопрос задается, имеют две существенные особенности: 1) они выражают иллокуцию обоснования; 2) они имеют нерасчлененную комму-

никативную структуру, потому что состоит из одной ремы. Методологическая ценность анализа нерасчлененных предложений состоит в том, что они дают материал для определения акцентоносителя в коммуникативной составляющей, имеющей в качестве синтаксического сегментного материала предложение в целом (S). Обратимся к просодии обоснования и рассмотрим пример 3).

3) — *Что случилось?* — *У Максим Ивановича **мама** умерла* [8].

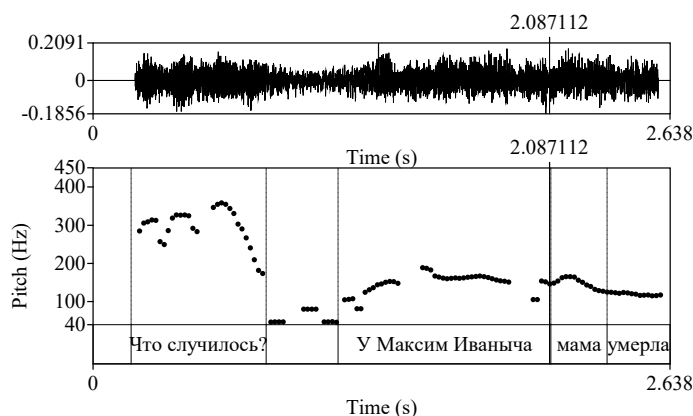


Рисунок 1: График изменения частоты (нижняя панель) в примере 1).

Онтологическая ситуация, в которой фигурирует пример 3), состоит в том, что первый говорящий слышит телефонный звонок и задает вопрос *Что случилось?* второму говорящему, ответившему на звонок. Второй говорящий сообщает о смерти матери общего коллеги. Рисунок 1 говорит о том, что акцентоноситель в ответной реплике примера 3) — словоформа *мама*. Она несет акцент ИК-2 (по Е. А. Брызгуновой [1]). Всего в ответной реплике четыре фонетических слова: *у Максим, Ивановича, мама и умерла*. Возникает вопрос, какому принципу подчиняется выбор акцентоносителя в предложениях обоснования, таких, как ответная реплика в примере 3).

Рассмотрим информационную (с точки зрения членения на данное vs. новое) и семантико-синтаксическую структуру предложения. Все элементы события, о котором говорится в ответе, — Максим Иванович, его мама и факт ее смерти — соотносятся с новой информацией. С семантико-синтаксической же точки зрения в предложении выделяется именная группа *у Максим Ивановича* как внешний посессор (EP), или второй участник ситуации (A2), именная группа *мама* — как субъект (A1) и финитный глагол — как предикат (P).

Для определения иерархии компонентов предложения как претендентов на роль акцентоносителя продолжим рассмотрение ответов на вопрос типа *Что случилось?*.

4) — *А че ты в тулупе-то парисься?* — ***Наручники** мешают* [8].

5) — *В чем дело?! Что?! Кто приказал?* — ***Дезинфекция** же в области. Деревянный жук все **деревья** поел* [8].

6) — *Да ты че?! — На Арбате **танки** стоят* [8].

7) ... *ты чего реवेशь?* — ***Пела** божественно* [8].

8) — *Что с вельветом?* — *Вельвет **берут*** [8].

Все результаты поиска здесь не приводятся, приведенные примеры позволяют проиллюстрировать следующие приоритеты, наблюдаемые при выборе акцентоносителя.

(1) Первый участник ситуации A1 (подлежащее) имеет приоритет перед финитным глаголом P, что иллюстрирует пример 4), ср. *Наручники* (A1) *мешают* (P).

(2) Второй участник ситуации A2 имеет приоритет перед первым участником A1 и перед финитным глаголом P, о чем говорит пример 5), ср. *Древесный жук* (A1) *все деревья* (A2) *поел* (P).

(3) Основные участники ситуации имеют приоритет перед внешним посессором (пример 3) *У Максим Ивановича мама умерла*) и локативным обстоятельством L (примеры 5) и 6), ср. *Дезинфекция же в области* (L) и *На Арбате* (L) *танки* (A1) *стоят* (P).

(4) Финитный глагол P имеет приоритет перед обстоятельством образа действия (M), ср. пример 7): *Пела* (P) *божественно* (M).

(5) Если синтаксически приоритетная группа соотносится с известным O (старым, данным, предупомянутым), она отдает свое право на роль акцентоносителя группе, следующей за ней в иерархии, т.е. такой, у которой при выборе акцентоносителя меньший вес. Так, в примере 8) акцентоноситель — глагол в личной форме при том, что прямое дополнение A2, в соответствии с п. (2) имеет больший вес, чем глагол. Дополнение A2 *вельвет* теряет свое право на роль акцентоносителя, т.к. в ответе оно заимствовано из вопроса: *Вельвет* (O) *берут* (P). Здесь мы исходим из предположения, что в ответе на вопрос типа *что случилось?* могут быть местоименные, предупомянутые и другие компоненты, соответствующие известной информации и что при отсутствии просодических показателей темы это не влияет на нерасчлененность предложения. Иначе говоря, в составе ремы в определенных условиях могут быть известные компоненты.

Таким образом, можно наблюдать следующие приоритеты иерархии: $O < M < P < L < A1 < A2$. Наивысший приоритет в этой (пока неполной) иерархии у второго участника ситуации, ср. прямое дополнение в примере 5). Эти приоритеты совпадают с теми, которые были получены методами интроспекции и эксперимента [15-16], однако корпусный анализ говорит о неполноте результатов, достигнутых ранее, и необходимости проведения дальнейшего анализа. Некоторые из этих приоритетов были выделены и на английском материале [2], [3: 208], [6].

Добавим также, что, если приоритетная группа имеет внутреннюю структуру (так, имя может иметь атрибутивные зависимые, оно может быть встроенным в сочиненную или числовую группу, оно может представлять собой имя и отчество, имя и фамилию, имя и прозвище, титул и имя), акцентоноситель в такой группе выбирается в соответствии с локальными принципами, по сумме представленных в работах [4: 146]; [10: 203–206]; [11]; [5: 202]. Так, в примере *На базаре шесть гусей продал* [8]) в приоритетном актанте *шесть гусей*, который имеет структуру числовой группы, по локальному правилу выбирается акцентоноситель *гусей*, а в примере *...я его на собственные деньги купил* [8] в приоритетной именной группе *собственные деньги* — словоформа *деньги*. Локальные правила применяются к именным группам рекурсивно вплоть до достижения терминального узла.

Цель этого раздела состояла в том, чтобы продемонстрировать возможность разработки звучащего массива исследования путем использования ключевого слова *что случилось* и его аналогов.

3 Воспоминания и мечты

Для анализа предложений, отражающих погружение говорящего в его внутренний мир, были использованы ключевые слова *мечтаю*, *вспоминаю*, *помню*, *воспоминания*. Вернемся к примерам 1) и 2), рассмотренным во Введении. В примере 1) имеются два отдельных акцентоносителя *полянку* (A2) и *погуляли* (P), а в примере 2) акцентоносители — это словоформы *реактором* (A2) и *метров*. Эти акцентоносители соответствуют иерархии приоритетов, полученной в разделе 2. Акцентоноситель *метров* соответствует локальному правилу выбора акцентоносителя в

числовой группе *двести метров*. Рассмотрим еще несколько примеров. Пример 9) служит результатом поиска в Мультимедийном подкорпусе НКРЯ по ключевому слову запроса *помню*.

9) *Я даже помню: доро-о-жка (A1) была-а (P) [8].*

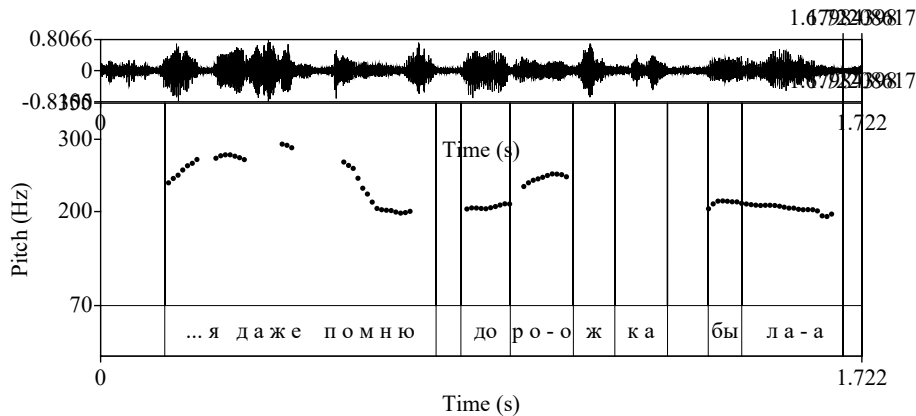


Рисунок 2: График изменения частоты в примере 9)

Тонотрама (нижняя панель) говорит о том, что на акцентоносителе словоформе *дорожка* фиксируется пологий подъем типа ИК-6, который сопровождается растянутой артикуляцией ударного слога. Заударный слог и последующая словоформа характеризуются плавным понижением тона. Перед нами просодическая структура, маркирующая воспоминания. Ср. также примеры 10)-11), отражающие мечты говорящего.

10) *Знаешь, я мечтаю, чтобы ты приехал на неде-ельку (A2), мы б с тобой погуля-яли (P), сходили бы в кино-о (A2) [8].*

11) *Она (O) уже седа-ая (P), наверное, теперь... За-амужем (P) давно (M). А все-таки мечтаю, хоть голос ее услышать [8].*

Примеры 9)-11) также подтверждают гипотезу о том, что принципы выбора акцентоносителя, полученные на примере ответов на вопрос типа *что случилось?*, применимы и к предложениям погружения говорящего в его внутренний мир.

Итак, дан краткий анализ просодии и принципа выбора акцентоносителя в предложениях с иллюкативной силой мечтаний и воспоминаний. Показано, что подбор звучащего материала для анализа можно осуществить через ключевые слова *помню, вспоминаю, мечтаю*. Отметим также, что просодический контур, условно названный нами просодией воспоминания и мечты, имеет более широкое толкование: соответствующей просодией оформляются любые речевые акты, связанные с имитацией мыслительной деятельности, которая понимается говорящим как протекающая во времени. Говорящий вспоминает, мечтает, размышляет, подбирает ускользящее слово, ищет подходящие примеры для того, чтобы слушающий лучше его понял, говорящий «шевелит мозгами».

4 Дискурсивная незавершенность и да-нет-вопрос

Дискурсивная незавершенность — одно из коммуникативных значений, имеющее определенные просодические показатели. Незавершенность, говорит о том, что текущий речевой акт позиционируется говорящим как неконечный в ряду речевых актов. Говорящий дает понять слушающему, что он еще не готов передать слово и собирается продолжить речь. Пример 12) демонстрирует незавершенность в речевом акте сообщения. Просодическими показателями незавершенности служит подъем частоты на акцентоносителях незавершенности *неофициальное, маленькая, называется и обозначили*.

12) Орлинка — это название **неофициальное**, потому что бухта **маленькая**, официально она никак не **называется**, ее так **обозначили**, чтоб хоть как-то [8].

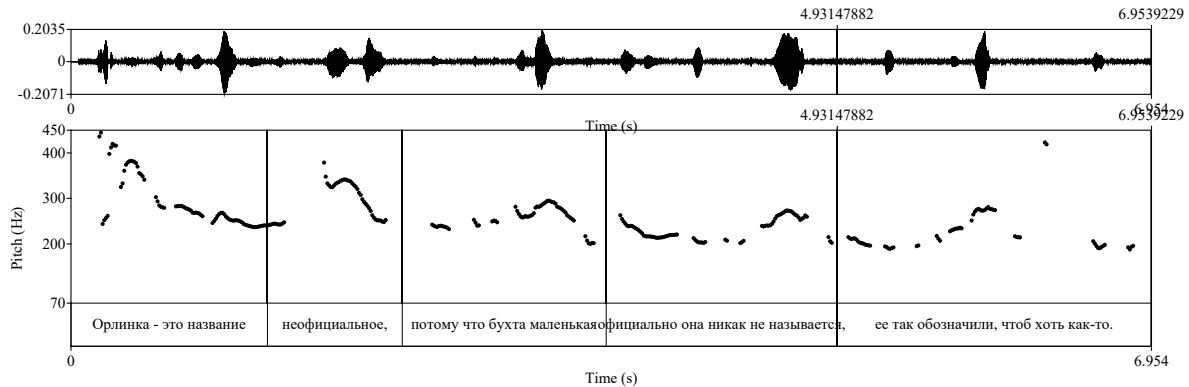


Рисунок 3: График изменения частоты в примере 12).

В работе [17] показано, что незавершенность совместима со всеми основными типами речевых актов: императивом, вокативом и вопросом с вопросительным словом. Все эти типы речевых актов способны присоединять показатель незавершенности. Единственным типом речевого акта, который оставался не исследованным окончательно, был *да-нет*-вопрос. Очевидно, что сочетаемость *да-нет*-вопроса с незавершенностью затруднена. Причиной этого представляется то, что показатель *да-нет*-вопроса, акцент типа ИК-3, это также и наиболее частый показатель незавершенности. *Да-нет*-вопрос, в котором иллюкуция вопроса передается частицей *ли*, в качестве *да-нет*-вопроса из рассмотрения исключается, потому что в вопросе с *ли* иллюкутивная сила выражается не просодически, а сегментно: частицей *ли*. Соответственно, вопросы с *ли* легко комбинируются с незавершенностью.

В настоящее время исследование корпусного материала говорит о том, что сочетаемость *да-нет*-вопроса с незавершенностью возможна. Анализ дискурсивной ситуации, в которой представлено образование вопроса в режиме незавершенного дискурса, свидетельствует о том, что за вопросом, за которым в наиболее естественном случае должен следовать ответ, иногда может следовать и другой вопрос или обоснование того, почему этот вопрос задается. Будет ли при переходе от одного вопроса к вопросу, следующему за ним, или повествовательному предложению выражаться значение незавершенности? Это не бессмысленный вопрос, потому что незавершенность, вообще говоря, может оставаться и не выраженной просодически.

В каких же ситуациях может быть задано более одного вопроса подряд? Мы предполагаем, что наиболее благоприятным дискурсивным контекстом для серии вопросов служит контекст пресс-конференции. Когда журналист получает слово, он, как правило, делает попытку задать максимальное количество вопросов, приличествующее ситуации, потому что второй раз слова ему уже не дадут. Заметим также, что вопрос, имеющий продолжение, которое нуждается в указании на незавершенность, должен содержать достаточно объемный сегментный материал, чтобы акцентам, выражающим отдельно вопрос и незавершенность, было где разместиться.

Кроме пресс-конференции более двух вопросов подряд может встретиться и в жанре интервью, когда интервьюеру естественно подготовить ряд вопросов заранее. Кроме того, может быть рассмотрен материал раздела «вопросы» на научных конференциях и лекциях. Все эти жанры послужили источником ключевых слов для поиска *да-нет*-вопросов, имеющих указание на то, что текущий вопрос — не последнее предложение в речи. Наиболее плодотворный результат дал анализ материалов пресс-конференций:

13) **Можете** рассказать вообще, есть ли такие, кто хочет приобрести пятьдесят процентов в капитале **Открытие**? Или, может быть, **всю** группу целиком? И какие шансы у IPO **Открытие**? [youtube]

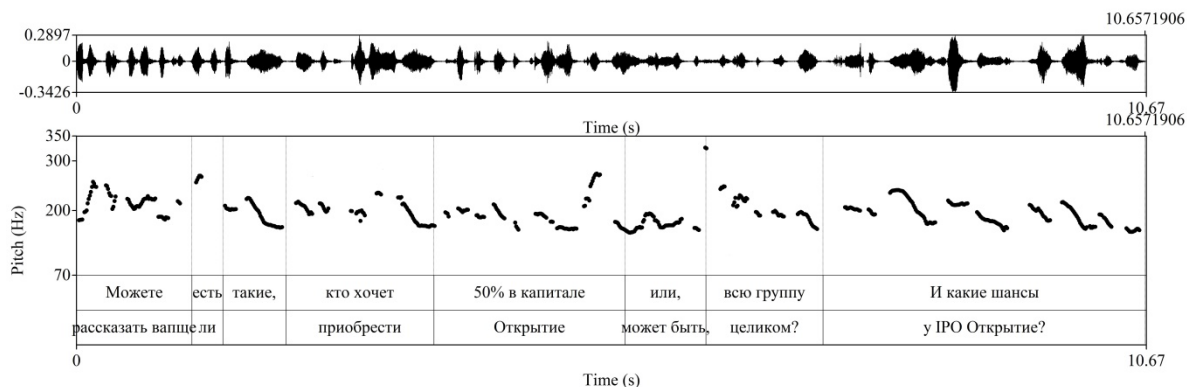


Рисунок 4: График изменения частоты в примере 13).

В примере 13) журналистка задает три вопроса подряд. Первый вопрос содержит подъем, маркирующий иллокуцию *да-нет*-вопроса на акцентоносителе вопроса словоформе *можете*, и подъем, маркирующий незавершенность на акцентоносителе незавершенности словоформе *Открытие* (первая встречаемость этой словоформы). Второй вопрос содержит только акцентоноситель вопроса, несущий подъем. Это словоформа *всю*. Последний вопрос с вопросительным словом *какие* завершает серию вопросов. Второй и третий вопрос указаний на незавершенность не содержат. Пример 13) весьма многословный, что, как известно, затрудняет анализ, но именно те контексты, которые содержат обширный сегментный материал, и составляют материал для анализа совместимости *да-нет*-вопросов и незавершенности.

Аналогично в примере 14) имеется подъем вопроса на словоформе *правильно* и подъем незавершенности на словоформе *получалось*:

14) *Правильно* я понимаю, что дружественными путями не *получилось*? И тогда решили действовать *агрессивно*? [youtube].

Всего в нашей коллекции около двух десятков примеров, аналогичных 13) и 14). Большинство примеров получено путем поиска по ключевому слову *пресс-конференция*. Несколько примеров найдено в жанре интервью с использованием соответствующего ключевого слова. Примеров достаточно для того, чтобы показать, что *да-нет*-вопрос совместим с просодией незавершенности.

5 Заключение

Показано, что поиск в массиве звучащих данных может быть проведен с использованием метода ключевых слов. Материалом для поиска служат транскрипты звучащих текстов, заголовки в орфографической записи и рубрики массива, в котором хранится текст. В качестве иллюстрации метода были использованы ключевые слова, обозначающие иллокутивную силу искомого речевого акта, и слова, связанные с обозначением искомой иллокутивной силы синонимическими, родовидовыми отношениями и отношениями, которые могут указывать на соответствующую искомой иллокутивной силе коммуникативную ситуацию. Последний тип отношений мы трактуем как ассоциативную связь в терминологии информационно-поискового тезауруса.

Кроме демонстрации метода поиска звучащего материала по ключевым словам в работе дается анализ проблем, для решения которых результаты поиска послужили рабочим материалом. Так, для анализа иллокутивной силы обоснования было использовано ключевое слово *что случилось*, его синонимы и ассоциаты. Результаты поиска позволили проанализировать принципы выбора словоформы-акцентоносителя в предложениях нерасчлененной коммуникативной структуры, которые как раз и представлены иллокуцией обоснования. Кроме того, анализ ответов на вопрос типа *что случилось?* позволяет выявить синтаксические приоритеты, которые действуют при выборе акцентоносителя в русском языке. Это типологически значимая языковая шкала, которая применима не только к русскому языку.

Далее. Для анализа выбора акцентоносителя в иллокуциях, соответствующих погружению говорящего в его внутренний мир, использованы ключевые слова *помню, вспоминаю, воспоминания, мечтаю*. Результаты поиска позволяют проанализировать просодическую структуру предложений, отражающих мыслительную деятельность, которая понимается как продолжающаяся во времени, и способы наложения просодической кривой на сегментный материал, в частности, выбор словоформы-акцентоносителя.

И наконец, поиск по ключевым словам *пресс-конференция* и *интервью* дает материал для анализа *да-нет*-вопроса в контексте дискурсивной незавершенности, потому что жанры, обозначенные этими ключевыми словами, способны представить вопрос в контексте других вопросов, следующих за ним, или объяснений того, почему этот вопрос задается, т.е. в контексте незавершенного дискурса.

Теоретическую значимость предложенного метода мы видим в разработке толкований прагматических значений, имеющих просодическое выражение, и в сопоставлении полученным толкованиям их лексических коррелятов, способных служить ключевыми словами запроса. Создание толкований фактически представляет собой разработку понятийного аппарата и соответствующего метаязыка для описания коммуникативно-просодического уровня языка.

References

- [1] Bryzgunova E. A. Intonation [Intonatsija] // Russian Grammar [Russkaja grammatika]. Vol. 1, Nauka, Moscow, 1982. P. 103-118.
- [2] Enkvist N. E. Marked focus: functions and constraints // Greenbaum, S. et al. (Eds.) Studies in English Linguistics for Randolph Quirk. London: 1979. P. 134–152.
- [3] Halliday M. A. K. Notes on transitivity and theme in English. Part 2 // Journal of Linguistics. 1967. Vol. 3. P. 199–244.
- [4] Kovtunova I. I. Modern Russian. Word order and theme-rheme division of a sentence [Sovremennyj russkij jazyk. Porjadok slov i aktual'noe chlenenije predlozhenija]. Prosveshchenije, Moscow, 1976. P. 103-118.
- [5] Kodzasov S. V. Phrase accentuation laws [Zakony frazovoi aktsentuatsii] // Prosodicheskie stroi russkoi rechi. 1996.
- [6] Schwarzschild R. GIVENness, AvoidF, and other constraints on the placement of accent // Natural Language Semantics. 1999. 7. 141–177.
- [7] Russian multichannel discourse [Russkij mul'tikanal'nyj diskurs] (Multidiscourse 2020): <<http://multidiscourse.ru>>
- [8] Russian National Corpus [Nacional'nyj korpus russkogo jazyka]: <www.ruscorpora.ru>.
- [9] Prosodically Annotated Corpus of Spoken Russian “Night dream stories” [Rasskazy o snovidenijah i drugie korpusa zvuchashhej rechi]: <<http://spokencorpora.ru>>.
- [10] Russian Grammar [Russkaja grammatika]. Vol. 2, Nauka, Moscow, 1982.
- [11] Svetozarova N. D. Accent-rhythmical innovations in Russian spontaneous speech [Aktsentno-ritmicheskie innovatsii v russkoi spontannoj rechi]. Problemy fonetiki, I. 1993.
- [12] STATE STANDARD GOST 7.25-2001 SIBID. Thesaurus for information retrieval. The principles of development, structure, and presentation. The Forth of November 2001. [GOST 7.25-2001 SIBID. Tezaurus informacionno-poiskovyj odnojazychnyj. Pravila razrabotki, struktura, sostav i forma predstavlenija ot 04 sentjabrja 2001 - docs.cntd.ru]: <<http://docs.cntd.ru>>.
- [13] Toivanen J., Seppänen T. Prosody-based search features in information retrieval. Proceedings of FONETIK, Stockholm, Finland. 2002. Pp. 105-108.
- [14] Video-hosting youtube: <<https://www.youtube.com>>.
- [15] Yanko T. E. The communicative structure with a non-inherent theme [Kommunikativnaia struktura s neingerentnoi temoi]. Nauchno-tehnicheskaja informatsija, 1991. 2 (7).
- [16] Yanko T. E. Accent placement principles in Russian // Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference “Dialogue” (2011). [Komp'juternaja lingvistika i intellektual'nye tehnologii]. Issue 10. [Vyp. 10]. Moscow: RGGU Publ. 2011. P. 288-301.
- [17] Yanko T. Speech acts within the structure of discourse: the incompleteness cues in sound corpora [Rechevyje akty v strukture svjaznogo diskursa: pokazateli nezavershennosti po dannym korpusov zvuchashhej rechi] // Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference “Dialogue” (2018). [Komp'juternaja lingvistika i intellektual'nye tehnologii]. Issue 17 (24). [Vyp. 17 (24)]. Moscow: RGGU Publ. 2018. P. 791-802.
- [18] Lezhenin Iu., Zhuikov A., Bogach N., Boitsova E., Pyskin E. PitchKeywordExtractor: Prosody-based Automatic Keyword Extraction for Speech Content // Proceedings of the Federated Conference on Computer Science and Information Systems, pp. 265–269.