

# МОДЕЛИРОВАНИЕ ОСНОВНЫХ ТИПОВ ИНТОНАЦИЙ ПРИМЕНИТЕЛЬНО К СИНТЕЗУ ТАТАРСКОЙ РЕЧИ<sup>1</sup>

## MODELING OF THE BASIC INTONATION TYPES FOR TATAR SPEECH SYNTHESIS

*Р.Р. Зарипова*

*Казанский государственный университет*

*Т.И. Ибрагимов*

*Казанский государственный университет, филологический факультет*

[Tavzich.Ibragimov@ksu.ru](mailto:Tavzich.Ibragimov@ksu.ru)

*Ф.И. Салимов*

*Казанский государственный университет, факультет ВМК*

[Farid.Salimov@ksu.ru](mailto:Farid.Salimov@ksu.ru)

**Ключевые слова:** татарский язык, ритмомелодическая структура, интонация, ритмическая группа, ритмика, мелодика, актантная структура, актант,

Предложены методика описания интонации предложений и общий алгоритм интонационного оформления синтезированных фраз основанные на фонетических и грамматических особенностях языка. Представлены результаты исследований ритмомелодического строя татарского языка на примере отдельных фраз.

Интонация - обязательный компонент звучащей речи, который делает ее более осмысленной, более понятной, живой и выразительной. В связи с этим возникает необходимость разработки системы просодического оформления речи при ее синтезе.

Исследования, результаты которых излагаются в данной статье, направлены на разработку ритмомелодических моделей татарских предложений, различающихся как по синтаксическому строю, так и по коммуникативной направленности. Цель работы не ограничивается решением лишь практических задач. Она имеет непосредственное отношение к рассмотрению таких теоретических проблем, как зависимость интонационных моделей от грамматики языка, а также характерную для данного языка систему противопоставлений высказываний, различающихся по интонации.

Наличие связи между ритмикой и интонацией вызывает интерес к сопоставительному изучению интонаций русского и татарского языков. Как известно, для русского языка характерно ярко выраженное словесное ударение и ритмический строй, основанный на чередовании ударных и безударных слогов. Что касается татарского языка, то данные, полученные при синтезе татарской речи, показали отсутствие в ней словесного ударения как такового. В частности при синтезе односоставных предложений типа *Салиха бакчадан арып кайтты* более качественное звучание достигалось тогда, когда слова, входящие в группу сказуемого, озвучивались как одно слово. Относительно приведенного примера это означало замену всех начальных дифонов (кроме дифона, входящего в слог *ба*) и конечных (кроме, входящего в сочетание *ты*) на срединные.

---

<sup>1</sup> Работа выполнена при финансовой поддержке гранта РФФИ №03-07-90285

Отсутствие словесного ударения свидетельствует о различных принципах, лежащих в основе ритмики фраз татарского и русского языков. Отметим, что имеющиеся в академических грамматиках сведения по интонации татарского языка зачастую являются поверхностными и не имеют отношения к разговорному стилю речи [1].

Уже на начальных этапах разработки синтезатора в рамках проблемы улучшения качества синтезированной речи нами были проведены определенные исследования ритмики татарской фразы. Согласно полученным в этот период результатам, основной ритмической единицей в татарском языке является ритмическая группа [2]. Ритмика татарской фразы представляет определенную последовательность РГ. Отметим основные свойства РГ. По определению, ритмическая группа - это фонетическое и смысловое единство. [3; 4]. В разговорной речи минимальные ритмические группы (фонетические слова) объединяются в синтагмы и фразы. При этом группа подлежащего и группа сказуемого сохраняют свою самостоятельность. Объединение минимальных РГ, входящих в группу подлежащего и группу сказуемого, происходит раздельно. Величина РГ зависит от темпа речи, смыслового содержания и длины предложения.

Существенной, на наш взгляд, является, связь между РГ и актантной структурой предложения, а именно, во фразе границы РГ совпадают с границами составляющих ее актантной структуры. Эта связь позволяет целенаправленно проводить исследования по систематизации огромного количества интонаций, использовать те семантические отношения между составляющими фразы, которые могут отразиться и на супrasegmentной структуре предложения.

Данная работа посвящена решению общих задач – систематизации большого количества ритмомелодических конструкций фраз и ритмических групп; построению наиболее общих ритмомелодических моделей языка относительно семи базовых интонационных контуров [5] и разработке системы правил, способной озвучивать синтезированную фразу с данной интонацией. Объектом исследования явились одноактантные и многоактантные простые предложения. Хотя полученные результаты прошли апробацию на предмет состоятельности при синтезе татарской речи, часть выводов носит, тем не менее, предварительный характер.

Приведем краткое описание алгоритма интонационного оформления синтезированной фразы, который лег в основу методики исследования и моделирования интонационных контуров различных типов предложений татарского языка. Работа алгоритма начинается с установления коммуникативного типа фразы. На втором этапе фраза проверяется на наличие слов и словосочетаний с дополнительным смысловым

значением (вопросительные слова, усилительные частицы, отрицательные местоимения и другие). Далее на базе актантной структуры входного предложения формируются ритмические группы озвучиваемой фразы.<sup>2</sup>

В формировании РГ наиболее существенным с точки зрения интонации является выделение групп подлежащего и сказуемого. В случае, когда актанты, входящие в группу подлежащего или сказуемого выражены отдельными словами, т.е. не представляют распространенные обстоятельства и дополнения, они могут быть объединены в ритмическую группу подлежащего или в ритмическую группу сказуемого соответственно.

Приведем некоторые детали процедуры объединения слов предложения в ритмические группы. Были приняты следующие обозначения: Пр – предложение;  $\oplus$  - объединение частей предложения; Аг – агент, выраженный одним словом, т.е. без конкретизации; САг – агент, выраженный словосочетанием; Гл – глагол-сказуемое, выраженное одним словом; СГл – глагол-сказуемое, выраженное словосочетанием; Об – объект, выраженный одним словом; СОб – объект, выраженный словосочетанием; Об → - объект, удаленный от глагола-сказуемого; « ' » - знак логического ударения; Вр - актант время.

Образование РГ во фразе зависит от количества актантов, от их способа выражения (с конкретизацией или без конкретизации) и от длины предложения. В одноактантной фразе, произнесенной в разговорной речи, ритмические группы подлежащего и сказуемого могут образоваться в следующих случаях:

1а) Пр = САг  $\oplus$  Гл;

1б) Пр = Аг  $\oplus$  СГл;

1в) Пр = САг  $\oplus$  СГл;

1г) Пр = Аг'  $\oplus$  Гл.

В многоактантной фразе каждый актант, входящий в группу сказуемого, может образовать отдельную РГ, если он выражен словосочетанием или расположен на некотором расстоянии от глагола сказуемого. К примеру, слово *Китаны* во фразе *Китаны о'ста'лга' куйдым*<sup>3</sup> входит в РГ

<sup>2</sup> Согласно блочной схеме синтеза татарской речи, работы по интонационному оформлению фразы производятся вслед за семантико-синтаксическим анализом входного предложения. Основная функция семантико-синтаксического анализа состоит в выявлении актантного состава и установлении границ актантов.

<sup>3</sup> В примерах, где встречаются татарские буквы будет использоваться следующая нотация: а', о', у', ж', н', h (со знаком апострофа в правом верхнем углу). Символ а' произносится как [a] в слове [cat], символ о' произносится как [ir] в слове [bird], символ у' произносится как [w] в слове [down],

сказуемого. Во другой фразе *Укытучы бирга'н китаны о'ста'лга' куйдым* в силу выраженности определительным словосочетанием это слово составляет группу объекта, а во фразе *Китаны о'ста'л тартмасына салып куйдым* оно представляет отдельную ритмическую группу в силу того, что располагается на значительном расстоянии от основного глагола.

В данной работе мы ограничились анализом ритмического строя группы глагола-сказуемого, включающей только актанта Объект. При этом в двухактантной фразе могут формироваться ритмические группы подлежащего, объекта и сказуемого в следующих случаях:

2а) Пр = (Аг или САг) ⊕ СОб ⊕ Сгл;

2б) Пр = (Аг или САг) ⊕ Об ⊕ СГл;

2в) Пр = (Аг или САг) ⊕ СОб ⊕ Гл;

Из перечисленных актантных структур более подробно нами исследованы ритмомелодические особенности конструкций 1г) и 2б). Кроме того, на предложениях типа *Кайчан тары ча'чтегез?* (*Тары кайчан ча'чтегез?*) изучено ритмомелодическое построение частного вопроса. Учитывая активную употребительность в языке предложений с указанием времени действия, анализу были подвергнуты предложения типа *Кича' машина то'зу'че сабакташын Актанышка озатты.*

Методика проведения эксперимента и представления полученных данных состояла в следующем. Длительности, частоты основного тона (ЧОТ) и интенсивности гласных измерялись на фразах, произнесенных мужским голосом. Фразы были озвучены тем же диктором, голосом которого произносятся синтезированные фразы. Записи фраз, выделение границ звуков и измерения производились с помощью программы WinPitchPro. ЧОТ и интенсивность гласных фиксировались на стабильных участках звуков, поскольку собственно синтез речи основан на дифонах и переходные участки сочетающихся звуков входят в исходные единицы синтезатора. Данные по каждой РГ были усреднены и представлены графически, а затем для всей фразы в целом.<sup>4</sup>

### Обсуждение полученных результатов

Рисунки 1 – 4 получены по данным анализа предложений со структурой 1г) (предложения типа *Сандугач кунган*) и 2б) (предложения типа *Шакира катык эча'*). На всех приведенных диаграммах по

символ ж' произносится как [sio] в слове [adhesion] или как [зж] в слове [дожжек], символ н' произносится как [ng] в слове [sing], символ h произносится как [h] в слове [hat].

<sup>4</sup> В разговорном стиле речи изменения интенсивностей оказались незначительными и поэтому не подвергались обработке

горизонталь отложены номера слогаобразующих гласных. Рисунки 1 – 2 отображают характер изменения длительностей гласных, 3 и 4 – особенности изменения ЧОТ.

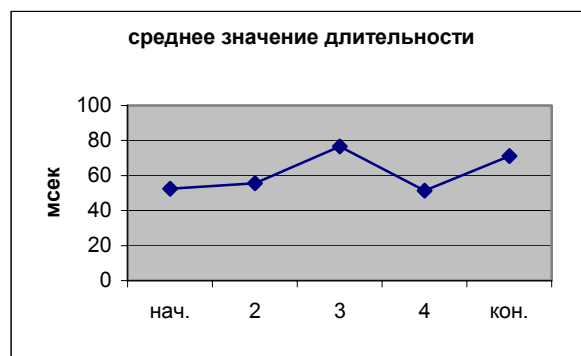


Рис.1



Рис.2

Как видно (рис. 1 – 2), в предложениях данных структур выделяются две РГ - РГ подлежащего и РГ сказуемого. Заметное выделение границы РГ подлежащего в структуре 1г) скорее объясняется падением на данный слог логического ударения, поскольку нередко в произношении двусоставных предложений в стиле разговорной речи наблюдается объединение подлежащего и сказуемого в одну РГ.

Изменения в ЧОТ гласных (рис. 3 - 4) не так значительны, как в их длительностях.

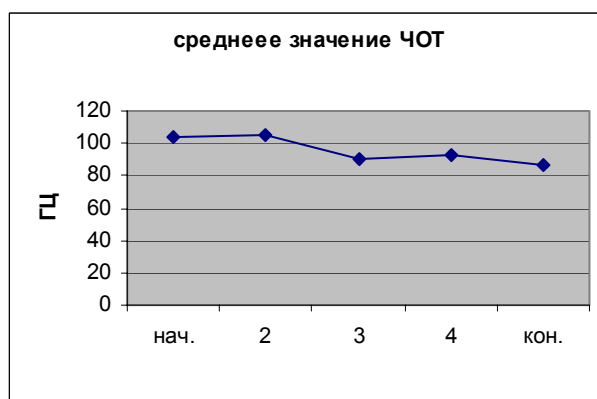


Рис.3



Рис.4

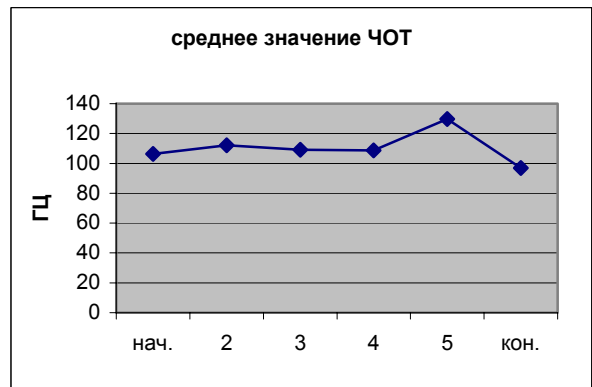


Рис.7

Однако нетрудно заметить сходство данных графиков. Прежде всего, в обоих случаях тон падающий, что характерно для интонации завершения. На падающий тон накладываются некоторый отпечаток интонации РГ – в итоге более заметным падение тона оказывается на границе ритмической группы.

Графики 5 – 8 построены на базе данных анализа тех же предложений, но с интонацией общего вопроса (получаемого добавлением частицы *ма', ма' к* сказуемому).

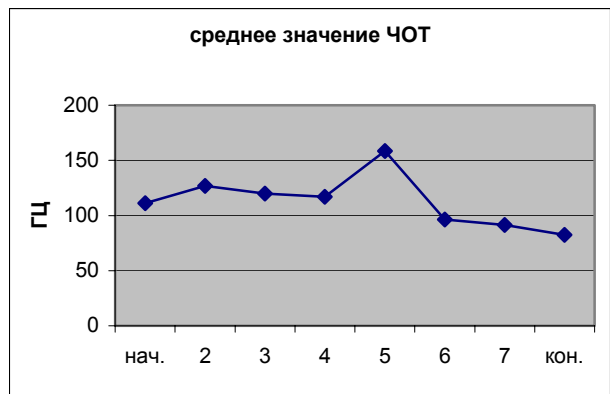


Рис.8

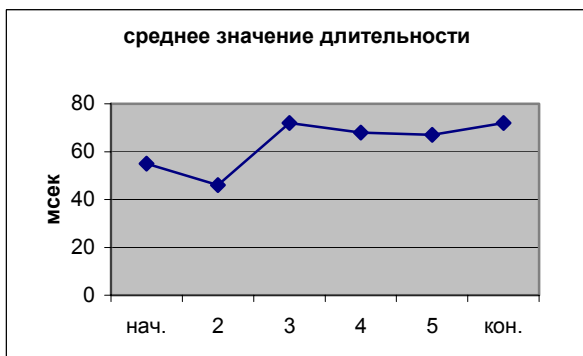


Рис.5

Как видно, расхождения в распределениях длительностей (рис. 5 - 6) и в изменениях ЧОТ (рис. 7 - 8), представленных на этих графиках, являются существенными. Эти расхождения могут быть обусловлены, в первую очередь, позицией логического ударения. В предложениях типа *Сандугач кунганмы?* ударным оказывается предпоследний слог группы сказуемого (в нашем примере *ган*), а в простых распространенных предложениях ударным является объект, и ударение падает на последний слог обозначения этого объекта (в нашем примере на слог *тык* в слове *катык*).

Данные, представленные на рис. 9 и 10, позволяют проследить за формированием ритмики и интонации вопросительного предложения с вопросительным словом.

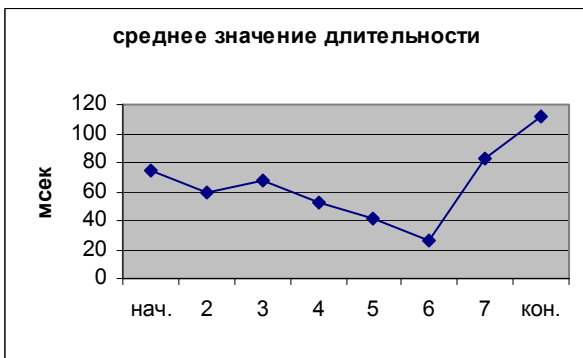


Рис.6



Рис.9



Рис.10

Кривые построены по результатам анализа предложений типа *Тары кайчан ча'чтегез?* Конфигурация кривых показывает, что характерной для данного типа вопросительной интонации является высокая тональность вопросительных слов. Отметим, что наличие в языке словесного ударения также должно было бы обнаружить себя по этому признаку. Но, как видно из представленных рисунков, заметных изменений высоты основного тона в пределах слова не происходит. Отличительной чертой кривых, приведенных на рисунках 9 и 10, является равномерное падение тона после вопросительного слова. Крутизна падения при этом по сравнению с интонацией завешенности является более значительной.

Ритмическая структура данного типа фраз, по-видимому, представляет одну ритмическую группу. Для кривой рис. 9 не характерно присущее конечному слогу РГ увеличение длительности. Пробный синтез фразы, при котором она (фраза) была оформлена в согласии с приведенными на графиках длительностями и ЧОТ, дал более качественное звучание по сравнению с озвучиванием путем деления ее на две ритмические группы.

Рисунки 11 и 12 получены на основе данных



Рис.11

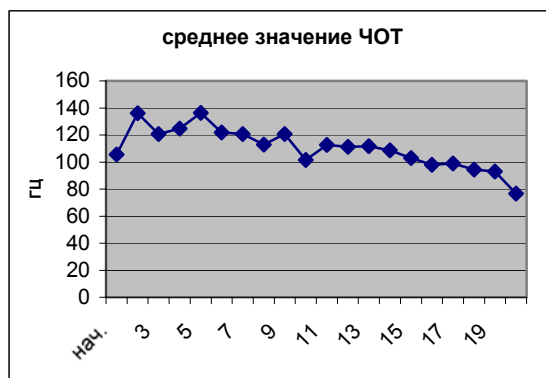


Рис.12

анализа трехактантных предложений со структурой).ПР = Вр ⊕ САг ⊕ Об ⊕ СГл. (*Кича' физика укытучысы /уктучысы/ кече кызын карточкага то'шерде*

В представленной структуре актант *Время* находится на значительном расстоянии от РГ глагола-сказуемого. Актант при этом, как правило выделяется в отдельную РГ, что и нашло отражение на графике (рис.11). Форма кривой этого графика позволяет констатировать, что в многоактантных структурах наиболее четко выделяются границы между актантами *Время* и *Агент*, равно как и между группой подлежащего и сказуемого. В группе сказуемого нет четких границ между входящими в нее актантами. Создается впечатление, что чем «многословнее» предложение, тем больше в нем объединений ритмических групп.

Основной тон достигает максимума на конечном слоге обозначения актанта *Время*, затем постепенно падает и к концу фразы достигает своего минимума (рис.12). Разность между максимальным и минимальным значениями основного тона в среднем составляет более 50 Гц, что несколько больше падения ЧОТ, наблюдаемого при произнесении более коротких фраз.

Моделирование ритмомелодической структуры языка возможно лишь при условии привлечения более обширного по объему языкового материала. Данная работа представляет начальный этап исследований интонационного строя татарского языка. Она направлена на решение частной чисто практической задачи – приблизить звучание синтезированной речи к естественной.

***Список литературы:***

1. Татарская грамматика // Казань: Татарское книжное издательство, 1993. том.1. 539с.
2. Ибрагимов ТИ, Хусаинов РР Синтезатор татарской речи: формирование ритмических групп в речевом потоке // Труды казанской школы по компьютерной и когнитивной лингвистике TEL-2000. Казань: Сэлэт, 2000.
3. Щерба ЛВ Фонетика французского языка. // М.: Иностранная литература, 1957.
4. Виллер МА О зависимости французской ритмической группы от стиля речи. Звуковой строй языка. // М.: Наука, 1979,. 268 с.
5. Брызгунова ЕА Практическая фонетика и интонация русского языка. // М., 1963. 307 с.