

**БАЗА ДАННЫХ «ИНТОНАЦИЯ РУССКОГО ДИАЛОГА»:  
РЕПЛИКИ-СООБЩЕНИЯ<sup>1</sup>**  
**«INTONATION OF RUSSIAN DIALOGUE» DATABASE:  
DECLARATIVE UTTERANCES**

*Кодзасов С. В. (sankod@philol.msu.ru), Архипов А. В. (arxipov@philol.msu.ru),  
Бонч-Осмоловская А. А. (abonch@gmail.com), Захаров Л. М. (leon@philol.msu.ru),  
Кривнова О. Ф. (okri@philol.msu.ru)*  
*Московский государственный университет*

В докладе подводятся итоги завершающего этапа работы по созданию базы данных «Интонация русского диалога». Единица базы содержит звуковой файл реплики-сообщения, графический файл с интонограммой и многопараметрическое описание просодии реплики. В докладе предлагается детальная классификация реплик-сообщений.

### *1. Введение*

При создании информационной системы «Интонация русского диалога» применялись методы когнитивно-дискурсивного анализа функций интонации и современные средства ее фонетического анализа на основе современных компьютерных программ. БД реализована в среде Microsoft Access и позволяет осуществлять поиск и сортировку данных как по семантическим, так и по фонетическим основаниям, а также текущим образом прослушивать объекты базы данных и наблюдать их акустические характеристики.

Результатом работы являются объединенные общей оболочкой три самостоятельных базы данных, содержащих массивы вопросительных, побудительных и повествовательных реплик, снабженных звуковыми файлами, просодическими дескрипциями, семантическими классификаторами, а также иллюстрациями в виде графиков звуковой волны и частоты основного тона.

#### *1.1. Отбор и запись материала*

Сбор текстового материала осуществлялся на основе анализа записей естественного дискурса в средствах массовой информации, а также записей естественного бытового диалога. Далее этот материал сопоставлялся с существующими в литературе описаниями функций и средств, характерных для диалогических реплик соответствующих типов (вопросов, побуждений, сообщений). На этой основе составлялись массивы образцов для записи. При этом авторы стремились максимально полно покрыть все разнообразие соответствующих реплик, наблюдаемое в русской речи. Оказалось, что массивы порядка 300–400 реплик (для каждого из трёх компонентов) достаточно полно покрывают пространство дискурсивных возможностей русской речи.

Запись материала производилась на высококачественной аппаратуре в студии звукозаписи филологического факультета МГУ. На первых двух этапах работы запись производилась на магнитную ленту, а затем оцифровывалась. На третьем этапе работы был использован цифровой рекордер, записывающий непосредственно файлы формата *wav*. В качестве дикторов выступили участники проекта, все — носители литературной нормы, жители Москвы, с высшим образованием.

#### *1.2. Просодическая разметка*

В ходе интонационного анализа каждому файлу базы сопоставлена просодическая дескрипция — результат просодической разметки. При этом использовалась система комбинаторного описания интонации, разработанная С. В. Кодзасовым ранее [Кодзасов 1996, 1999].

Каждая реплика представлена в БД как в стандартной орфографической форме, так и в расширенной орфографической записи, включающей элементы просодической разметки: обозначения границ просодических блоков и локальных акцентов. Все просодические маркеры далее получают расшифровку в отдельной группе полей.

<sup>1</sup> Работа выполнена при поддержке гранта РГНФ № 04-04-12027в.

Помимо просодической разметки, в базу данных включены графики звуковой волны (осциллограммы) и частоты основного тона (интонограммы).

## 2. Третий компонент БД: реплики-сообщения

Третий компонент базы данных разрабатывался по модели, апробированной на предыдущих этапах работы [Кодзасов и др. 2005; 2006; Кобозева 2005]. Вначале была разработана комплексная классификация реплик-сообщений, которая суммирует результаты фундаментальных работ последних десятилетий, а также результаты анализа массива высказываний, собранных по материалам дискурса в средствах массовой информации. Классификация опирается как на коммуникативную и лексико-синтаксическую структуру предложений, так и на их иллокутивно-модальные характеристики.

Произведена цифровая запись массива реплик с семантикой сообщений, покрывающего все выделенные классы. Всего в базу данных введено 277 файлов, содержащих 313 диалогических реплик. Произведен интонационный анализ всех файлов базы; просодические дескрипции опирались на систему описания интонации, используемую на двух первых этапах выполнения проекта.

Полученные результаты введены в базу данных в формате Microsoft Access. Типовое описание реплики содержит следующие зоны:

а) стандартная орфографическая запись реплики; б) орфографическая запись высказывания с просодической разметкой, позволяющей соотносить акценты и интегральные просодии с компонентами предложения; в) интонационная расшифровка акцентов; г) интонационная расшифровка блоков; д) тип реплики по классификации (максимальная глубина классификации — 5 уровней). Форма снабжена кнопками, которые дают возможность прослушать запись, а также просмотреть график звуковой волны и интонограмму.

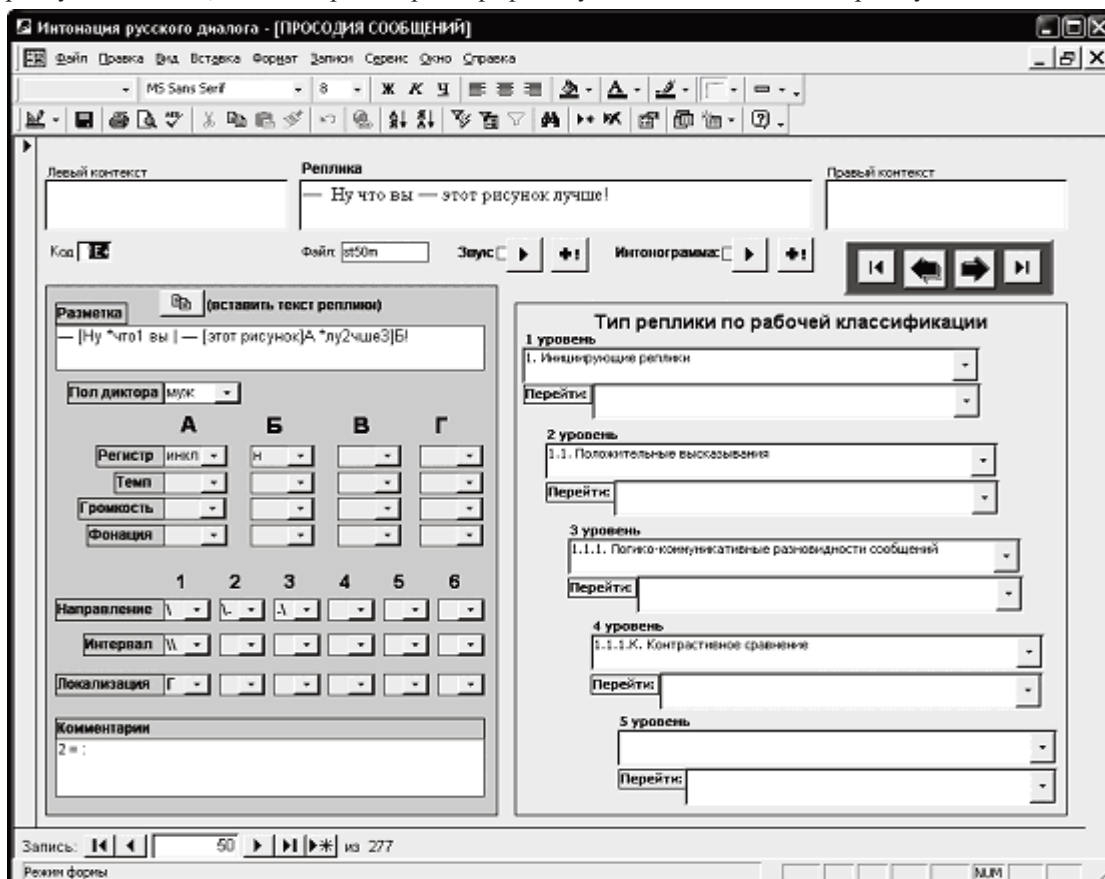


Рис. 1. Вид рабочего окна MS Access для реплики «Ну что вы — этот рисунок лучше!»

## 3. Классификация сообщений и реакций на них

Если при подготовке материалов для двух первых компонентов БД мы в значительной степени опирались на существующие классификации вопросов и побуждений, то детальной классификации сообщений, которой

можно было бы воспользоваться, до сих пор не было. При её создании отчасти были использованы логико-коммуникативные оппозиции, которые предлагались лингвистами независимо от иллокутивных характеристик высказываний ([Арутюнова 1998], академические грамматики русского языка). Однако основным источником при подготовке материала послужили тексты диалогов в средствах массовой информации (радио и телевидение).

Поскольку полученная классификация может послужить далее основой для более детальных исследований в области типологии высказываний, мы приводим её целиком с минимальным количеством иллюстраций.

Следует предупредить читателей, что настоящая классификация преследует цель не столько исчисления всех возможных комбинаций логико-коммуникативных параметров, сколько выявления наиболее значимых типов сообщений, которые встретились в эмпирическом материале, и выделения самих параметров варьирования.

## 1. ИНИЦИИРУЮЩИЕ РЕПЛИКИ

### 1.1. Положительные высказывания

#### 1.1.1. Логико-коммуникативные разновидности сообщений

##### А. Сравнение сообщений с диктальной и модальной ремой

(1) — *По вечерам он плавает. У нас тут бассейн недалеко.*

(2) — *Да не утонет он, не бойся! Он уже давно плавает, притом неплохо.*

(3) — *Позвони Ване. Он вчера приехал.*

(4) — *А-а, значит он все-таки приехал.*

##### Б. Сравнение фактового и событийного представление информации

(5) — *Это упражнение опасно — можно сломать руку.*

(6) — *Можешь себе представить — Иван Иванович сломал руку.*

##### В. Сравнение предложений немаркированной и противопоставительной идентификации

(7) — *Это Володя мне рассказал.*

(8) — *Как раз Володя мне это и рассказал.*

(9) — *Эту классификацию предложил еще Трубецкой.*

(10) — *Эту классификацию предложил не кто иной как Трубецкой.*

##### Г. Бытийные предложения

(11) — *На третьем этаже есть буфет. Там можно перекусить.*

(12) — *Перекусить можно на третьем этаже. Там есть буфет.*

(13) — *Существует феномен коллективного разума.*

##### Д. Предложения подобия и тождества — Ты знаешь, она очень похожа на мать.

(14) — *Смотри-ка! Точь-в-точь как ваш дом.*

(15) — *Северо-восточный регион Испании — это Страна Басков.*

##### Е. Предложения номинации

(16) — *Это животное называется «бегемот».*

(17) — *Завхоза зовут Иван Петрович.*

##### Ж. Состояния онтологической среды

(18) — *Начинает темнеть. Пора уходить.*

(19) — *Начало темнеть и нам пришлось уйти.*

З. Констатация и вывод

(20) — *У него уже поражены легкие. Операция бесполезна.*

(21) — *Вы окружены. Сопротивление бессмысленно.*

И. Нейтральное сравнение

(22) — *Ну, этот рисунок лучше.*

(23) — *Лучше второй вариант выбрать.*

(24) — *Ваня приехал во вторник, а Петя в среду.*

К. Контрастивное сравнение

(25) — *Ну что вы — этот рисунок лучше!*

(26) — *Ваня-то сдал все во время, а Петя — с опозданием!*

1.1.2. Маркированные оценочные характеристики в сравнении с нейтральными

1.1.2.1. Эпистемический статус сообщений

А. Нейтральные сообщения, напоминания

(27) — *Ты не забыл — завтра день рождения Юлика.*

(28) — *Напомним, что Александр уже был вчера в нашей студии.*

Б. Фокализованные сообщения

(29) — *Какая-то она странная последнее время.*

(30) — *Уж кто-кто, а она-то всегда поможет.*

В. Сообщения об ожидаемом или привычном

(31) — *Что-то интересное было? — Кое-что было. Но в общем ничего особенного.*

Г. Повтор чужого сообщения

(32) — *Я слышал — ты новую машину купил.*

1.1.2.2. Значимостный статус сообщений

А. Нейтральные сообщения, констатации

(33) — *Днем опять дождь. Придется тебе взять плащ.*

Б. Мобилизация внимания и снятие значимости сообщения

(34) — *Дождь обещали. Не забудь пальто взять!*

(35) — *Да пальто мне не понадобится. Я на машине.*

1.1.2.3. Аксиологический статус сообщений

А. Дистанцированные сообщения о положительном и отрицательном

(36) — *Уборщица у них работает на совесть.*

(37) — *Им не удалось открыть новое отделение.*

Б. Сообщения о положительном и акты одобрения

(38) — *Хорошо, что родители были дома.*

(39) а. — *Как у тебя чисто! б. — Это дочка помогает.*

В. Сообщения об отрицательном и акты осуждения

(40) — *Беда с тобой. Ты все делаешь не так как надо.*

(41) — *Надо было раньше об экзаменах думать.*

Г. Отрицательный комментарий

(42) — *Вряд ли он успеет это сделать.*

(43) — *Я ему не доверяю — слишком хорошо его знаю.*

1.1.3. Высказывания возможности и долженствования

(44) — *Ты можешь придти завтра вечером.*

(45) — *Исправлять ситуацию нужно на уровне всего высшего образования.*

1.2. Отрицательные высказывания

1.2.1. Логико-коммуникативные разновидности отрицательных сообщений

1.2.1.1. Отрицательные сообщения о факте

(46) — *Учти — по четвергам они не работают.*

(47) — *Ваня-таки не приехал.*

(48) — *Ты знаешь — экзамена завтра не будет.*

(49) — *Смотри-ка — здесь уже давно никого не было. Никаких свежих следов.*

1.2.1.2. Отрицательная идентификации — *Это не Володя.*

(50) — *Как раз четверг мне не подходит.*

1.2.1.3. Бытийные отрицания

(51) — *Пообедать негде. Здесь поблизости ничего такого нет.*

(52) — *Совсем нет времени. Мы ничего не успеем.*

1.2.2. Маркированные оценки

1.2.2.1. Эпистемические разновидности отрицательных сообщений

А. Нейтральные сообщения

(53) — *Он еще не приехал.*

(54) — *Ответ неверный.*

Б. Сообщения о контрастивном или странном

(55) — *Что-то ее нет до сих пор.*

(56) — *Вроде все делал правильно, а ответ неверный.*

(57) — *Дом большой, но неудобный.*

В. Сообщения об ожидаемом или привычном

(58) — *А Сидоров как всегда не пришел.*

Г. Повтор чужого отрицательного сообщения

(59) — *Ваня говорит — Лена не придет.*

1.2.2.2. Значимостный статус отрицательных сообщений

А. Нейтральные сообщения, констатации

(60) — *Учти, что занятий завтра не будет — день науки.*

Б. Сообщения о важном

(61) — *Иванов, к большому нашему сожалению, не пришел. А он как раз был очень нужен.*

В. Сообщения о маловажном

(62) а.— *Как съездил? б. — Да ничего интересного.*

1.2.2.3. Аксиологический статус отрицательных сообщений

А. Сообщения о положительном и акты одобрения

(63) — *Ты ничуть не постарел. Здорово выглядишь.*

Б. Сообщение об отрицательном

(64) — *Он вообще не соблюдает никакие правила игры.*

В. Прескриптивное отрицание

(65) — *С этим подонком лучше дела не иметь.*

Г. Акты осуждения

(66) — *Ты за весь день не сделал ничего полезного.*

(67) — *В этой истории никто, кроме тебя, не виноват.*

1.2.3. Отрицательная эпистемическая оценка

(68) — *Не думаю, чтобы в этом озере была рыба.*

(69) — *Не уверен, что он придет.*

## 2. РЕАКТИВНЫЕ РЕПЛИКИ

2.1. Реакции на сообщения разных типов

2.1.1. Подтверждения и согласия

(70) а. — *Говорят, Иванов женился. б. — Так и есть. Уже полгода как.*

2.1.2. Констатации- интерпретации

(71) а. — *У жены теперь своя машина. б. — Да ты, брат, разбогател.*

2.1.3. Дополнения

(72) — *Он, кстати, был первым, кто признал это открытие.*

2.1.4. Уточнения

(73) а. — *Ему уже, наверное, пятьдесят исполнилось. б. — Ну, пятьдесят ему уже давно. Скоро шестьдесят будет.*

2.1.5. Опровержения, возражения, несогласия

(74) а. — *Они проиграли, потому что играли в открытый футбол. б. — Но ведь соперники тоже в открытый футбол играли.*

2.1.6. Сравнения

(75) — *Я-то всегда вовремя прихожу. А вот другие...*

2.1.7. Мнение

(76) — *Я считаю, что у него ничего не получится.*

(77) а. — *А где Иванов? б. — Не знаю. Может, в столовую пошел.*

2.1.8. Уступка, извинение

(78) — *Честно говоря, я с вами согласен. Хотя и не во всем.*

2.1.9. Дистанцирование

(79) — *Ничего не подделаешь. Дети есть дети.*

2.2. Реакции на да/нет вопросы

2.2.1. Подтверждение

(80) а. — Как дела у дочки? Дети есть? б. — Да, уже двое.

2.2.2. Отрицание

(81) а. — Как дела у сына — диссертацию защитил? б. — Пока нет. Все собирается.

2.3. Реакции на диктальные вопросы

2.3.1. Прямой ответ

(82) а. — Как прошла поездка? б. — Прекрасно. Погода была отличная.

2.3.2. Замена пресуппозиции вопроса

(83) а. — Ты диссертацию когда защитил? б. — А я ее не защищал.

(84) а. — Ты кого больше любишь — Толстого или Достоевского? б. — Ни того, ни другого.

2.3.3. Уход от ответа

(85) а. — Сколько ты теперь зарабатываешь? б. — На хлеб с маслом хватает.

2.3.4. Отказ от ответа

(86) а. — А жена у тебя прежняя? б. — Все-то ты хочешь знать.

3. НЕФАКТИВНЫЕ СООБЩЕНИЯ

3.1. Предположения, мнения

(87) — Возможно, поезда пустят в обход.

(88) — Это может привести к обвалу рубля.

3.2. Утверждения, констатации, аргументы, выводы

(89) — Я уверен, что он с этим не справится.

(90) — Это сегодня должно решиться.

3.3. Возражение, упрек, автоупрек, вынужденное согласие

(91) — Мы ему ничего не должны доказывать. Он все равно не поверит.

(92) — Я в это не верю. Не такой он человек, чтобы это сделать.

(93) — Надо было позвонить, прежде чем приходить.

(94) — Зря пришли. Нам бы остаться дома.

3.4. Пояснения

(95) — Но кончил он в низком чине: был разжалован за поражение на южном фронте.

(96) — На этом мы сегодня кончаем. Мне необходимо уйти раньше.

4. Специфические виды сообщений

4.1. Объявления

4.1.1. Объявления положительные

(97) — Поезд номер 23 из Одессы прибыл на третий путь.

(98) — Пассажиры, прошедшие регистрацию, приглашаются на посадку.

4.1.2. Объявления отрицательные

(99) — Лица, не прошедшие медосмотр, к экзамену допущены не будут.

4.2. Напоминания, предупреждения

(100) — Напоминаю, что всем нужно сделать прививки.

(101) — *Предупреждаю, что без справки от врача в бассейн вас не запищут.*

#### 4.3. Сравнения

(102) — *Они были левее даже коммунистов.*

(103) — *Эта фирма гораздо надежнее.*

#### 4.4. Перформативы

(104) — *Объявляю собрание открытым.*

(105) — *Вы обвиняетесь в государственной измене.*

### 5. СЛОЖНЫЕ СООБЩЕНИЯ

(106) а. — *Ты знаешь, у нас вчера защита была. Девочка из провинции защищалась. Первый оппонент дал отрицательный отзыв. Работа, видимо, слабая. И провалили — большинство было против. б. — Ну, это руководитель виноват. Не нужно было ее допускать.*

### 6. ВВЕДЕНИЕ СООБЩЕНИЙ РАЗНЫХ ТИПОВ С ПОМОЩЬЮ НАЧАЛЬНЫХ ДИСКУРСИВНЫХ ОБОРОТОВ

#### 6.1. Интродукция неожиданного

(107) — *Ты знаешь — мне только что звонил начальник. Посылает в командировку.*

#### 6.2. Интродукция противоречащего предыдущей информации

(108) — *В действительности, дело обстояло совершенно иначе.*

#### 6.3. Возражение или отказ с интродукцией

(109) — *Прости, пожалуйста, но я не согласен.*

#### 6.4. Аргументы против

(110) — *Прежде всего, это слишком дорого.*

#### 6.5. Аргументы за

(111) — *Кроме того, это даст тебе полезные связи.*

#### 6.6. Причинно-следственные связи

(112) — *Его вчера видели в парке. Следовательно, он куда-то не уезжал.*

#### 6.7. Обобщения и выводы

(113) — *В целом, это должно принести облегчение.*

#### 6.8. Мнения

(114) — *Я считаю, что ничего из этого не получится.*

### 7. ЭКСПЛИЦИТНОЕ ВВЕДЕНИЕ ОЦЕНОК РАЗНЫХ ТИПОВ С ПОМОЩЬЮ НАЧАЛЬНЫХ ДИСКУРСИВНЫХ ОБОРОТОВ

#### 7.1. Эпистемические оценки

(115) — *Несомненно, это картина Айвазовского.*

(116) — *На самом деле, это картина его ученика.*

(117) — *Разумеется, он опоздал.*

#### 7.2. Попутная или малозначимая информация

(118) — *Кстати, ему удалось починить часы.*

(119) — *Причем тогда речь шла о старом варианте.*

#### 7.3. Значимая информация

(120) — *Но самое важное — ему удалось починить компьютер.*



7.4. Аксиологические и эмотивные оценки

(121) — *К счастью, поезд опоздал ненадолго.*

(122) — *К сожалению, встречу пришлось отложить.*

8. БЕЗЛИЧНЫЕ КОНСТРУКЦИИ

8.1. Желательность

(123) — *Неплохо бы перекусить.*

8.2. Побудительность

(124) — *Просили поторопиться.*

(125) — *Можно ехать.*

8.3. Констатация состояния среды

(126) — *Здесь не пройти.*

(127) — *Времени мало.*

8.4. Сенсорные состояния

(128) — *Холодно сегодня.*

(129) — *Что-то меня знобит.*

8.5. Сенсорная идентификация

(130) — *Гудок слышно.*

(131) — *Откуда-то дым.*

(132) — *Кто-то идет.*

(133) — *Рыба играет.*

8.6. Констатация прагматической ситуации

(134) — *Нам не успеть.*

(135) — *Пять часов уже.*

**Список литературы**

1. Арутюнова Н.Д. Язык и мир человека. М., 1998.
2. Грамматика русского языка. Т. 1. М., 1953.
3. Грамматика русского языка в 2х тт. Т. 2. М., 1960.
4. Кобозева И.М. Опыт разработки признаковой базы для характеристики лексико-синтаксических, семантических и прагматических свойств вопросительных реплик // Труды международной конференции Диалог'2005. М.: 2005. С. 238-244.
5. Кодзасов С.В. Комбинаторная модель фразовой просодии // Просодический строй русской речи. М., 1996.
6. Кодзасов С.В. Уровни, единицы и процессы в интонации // Вопросы фонетики. Вып. 3. М., 1999.
7. Кодзасов С.В., Бонч-Осмоловская А.А., Захаров Л.М., Кобозева И.М., Кривнова О.Ф. База данных «Интонация русского диалога»: вопросительные реплики // Труды международной конференции Диалог'2005. М.: 2005. С. 245-249.
8. Кодзасов С.В., Архипов А.В., Бонч-Осмоловская А.А., Кобозева И.М., Кривнова О.Ф. База данных «интонация русского диалога»: побудительные реплики // Труды международной конференции Диалог'2006. М.: 2005. С. 236-242.