

## К ВОПРОСУ О СТРУКТУРЕ МУЛЬТИКАНАЛЬНОГО ДИСКУРСА<sup>1</sup>

### Введение

Лингвистический анализ дискурса, являясь частью более широкой междисциплинарной научной дисциплины, сложившейся в 1970-х гг. (см., однако, [Якубинский 1923]), изучает как процесс языковой коммуникации, так и образующиеся в ее результате тексты-дискурсы. В последние десятилетия использование термина «дискурс» в обоих значениях – и как процесса, и как объекта, то есть текста – уже закрепилось в научном обиходе. Как показывает анализ в [Демьянков 2005], история употребления слов *текст* и *дискурс* в обиходном и научном русском языке идет из XVIII в. и оказывается весьма причудливой. Слово *текст* начинает употребляться в русском языке со времен Ломоносова, слово *дискурс* – несколько позже, на границе XVIII и XIX вв., при этом в те времена *текст* употреблялся в контексте издательского дела, а *дискурс* – в качестве синонима устного монолога: *Слушай же мой дискурс* (И.И. Лажечников. Последний Новик, цитируется по [Демьянков 2005: 43]). В XIX в. *текст* начинает превалировать в художественной литературе, *дискурс* в этот период активно встречается только у Н.С. Лескова: *Начались уже дискурсы в дамском вкусе* (Н.С. Лесков. Заячий ремиз, [Демьянков 2005: 44]); прилагательное же *дискурсивный* встречается в это время только в научных текстах в значении ‘рассудочный’. Наконец, в начале XX в. *дискурс* полностью исчезает из языка художественной литературы, его заменяют *речь*, *слова*, *разговор*, и, пройдя период некоторого забвения, к концу XX в. *дискурс* постепенно приобретает исключительно терминологическое употребление [Демьянков 2005: 50]. В настоящей работе будет описана одна из разновидностей дискурса, привлекающая внимание исследователей в самые последние годы – мультимодальный дискурс.

### 1. Постановка проблемы

Данное исследование выполнено в рамках мультимодальной (мультимодальной, бимодальной)<sup>2</sup> лингвистики – молодого и одновременно одного из старейших направлений

---

<sup>1</sup> Работа выполнена при финансовой поддержке гранта РФФИ №14-18-03819 «Язык как он есть: русский мультимодальный дискурс».

<sup>2</sup> В настоящее время более распространен термин «мультимодальный», однако корректнее говорить именно о «мультимодальном» или «бимодальном» дискурсе, так как пока в основном изучаются только две модальности – вокальная (слуховая) и кинетическая (зрительная), а остальные модальности, напр., обоняние

исследований языка в широком смысле слова. Мультиканальная лингвистика изучает все реальное многообразие «живой» коммуникации между людьми: слова, интонацию, жестикуляцию, направление взгляда, мимику, см. рис. 1. Интерес к изучению мультиканальности возник еще в древности, однако современная лингвистика, берущее начало в XX в., пошла другим путем и исследовала исключительно письменные тексты и выражения, то есть вербальный канал [Linell 1982]. В конце XX в. ситуация начала меняться в сторону изучения устных дискурсов, то есть к вербальному каналу прибавился просодический, см., в частности, коллективную монографию под редакцией Уоллеса Чейфа «Рассказы о грушах» [Chafe ed. 1980]. Наконец, в XXI в. на наших глазах формируется новый мультиканальный подход, принимающий во внимание все каналы общения между людьми (см. в том числе [Крейдлин 2002; Бутовская 2004; Кибрик 2010; Kress 2010; Knight 2011; Adolphs, Carter 2013; Müller et al. eds. 2013; Гришина 2017; Кибрик 2018a]).

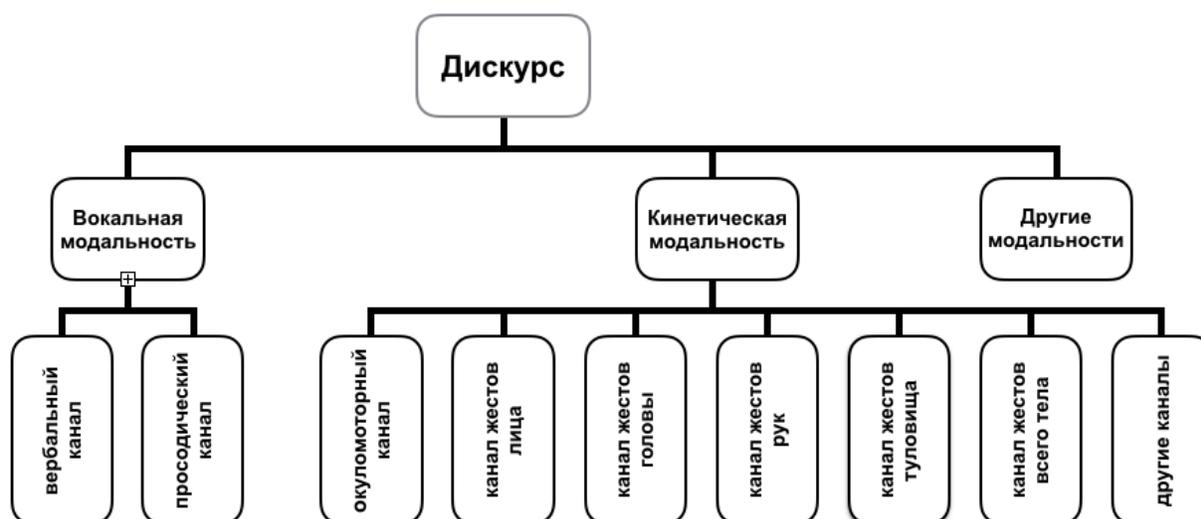


Рисунок 1. Модель мультиканального дискурса

Одним из ключевых вопросов в области мультиканальных исследований является вопрос о выделении для каждого канала базовых единиц, при помощи которых происходит сегментация информации. Базовая единица вербального канала – *клауза* (предикация), которая описывает единичное событие или состояние. Базовая единица просодического канала – *элементарная дискурсивная единица* (ЭДЕ)<sup>3</sup>, которая выделяется на основании

или осознание, остаются за пределами рассмотрения. См., однако, недавнюю работу [Mondada 2016], в которой наравне с этими двумя модальностями изучается также и осознание.

<sup>3</sup> Понятие ЭДЕ было введено в работах, описывающих вокальный компонент коммуникации (в частности, см. [Кибрик, Подлесская ред. 2009]). Описывая мультиканальный дискурс, корректнее было бы говорить об элементарных вокальных или просодических (если их выделение основывается на чисто

преимущественно просодических критериев и прототипически соответствует одной клаузе [Кибрик, Подлесская ред. 2009; Кибрик 2015]. В частности, ЭДЕ имеет единый контур частоты основного тона; основной акцентный центр, обычно расположенный на реме; характерный громкостный паттерн – затихание к концу; характерный темповый паттерн – ускорение в начале, замедление к концу; характерный паттерн паузации – пауза сопровождает планирование ЭДЕ и отсутствует внутри ЭДЕ. Базовая единица канала жестов рук (мануальных жестов) – *мануальный жест* (gesture; подробнее см. [Литвиненко и др. 2017]; в работах Адама Кендона и его последователей принят термин *gesture phrase* или *G-phrase* [Kendon 1980, 2004]). Жесты, с одной стороны, делятся на фазы (gesture phase, G-phase по [Kendon 1980]), с другой – объединяются в жестовые цепочки (gesture chains; gesture units, G-units по [Kendon 1980]). Базовой единицей канала жестов головы (цефалических жестов) является *цефалический жест*, цефалические жесты объединяются в жестовые цепочки [Сухова 2017]. Наконец, базовая единица окуломоторного канала (канала движений глаз) – *взгляд* (gaze); взгляды делятся на фиксации, саккады и другие более мелкие движения глаз [Федорова 2018].

Таким образом, изучая каналы коммуникации по отдельности, мы опираемся на соответствующие базовые единицы. Но как устроена структура мультиканального дискурса в целом, когда мы рассматриваем взаимодействие каналов? В этом вопросе возможно, по крайней мере, два принципиально разных подхода. С одной стороны, это вокальноцентрический подход, при котором мы считаем вокальную модальность приоритетной, а базовой единицей всего мультиканального дискурса признаем ЭДЕ; другой вариант – мы исходим из равноправия всех каналов, каждый из которых «имеет право» называться базовой мультиканальной единицей, и описываем структуру мультиканального дискурса как соположение базовых единиц всех каналов.

Этот вопрос также пришел к нам из глубокой древности. Так, Марк Фабий Квинтилиан, римский ритор I в. н.э., автор самого первого полного учебника ораторского искусства, считал вокальную модальность приоритетнее кинетической: «А поелику Действие (лат. *actio*, прим. авторов), как я сказал, разделяется на две части, на голос (лат. *vox*) и телодвижение (лат. *motum*, позднее *gestus*), из которых первый действует на слух, а другое на зрение, чрез которые два чувства все страсти проникают в душу: то прежде будем говорить о голосе, к коему приспособляется и телодвижение» [Квинтилиан 1834: XI]. В

---

просодических критериях) единицах. Однако термин ЭДЕ уже успел закрепиться в современной литературе, поэтому мы продолжим его использование с данной оговоркой.

дальнейшем изложении мы проанализируем взаимодействие вокальной модальности с тремя основными не-вокальными, кинетическими каналами – мануальным, цефалическим и окуломоторным – и предложим свой взгляд на данную проблему. Но сначала кратко опишем материал настоящего исследования.

## 2. Ресурс «Рассказы и разговоры о грушах»

Ресурс «Рассказы и разговоры о грушах» создается в ходе реализации проекта «Язык как он есть: русский мультимодальный дискурс», который осуществляется с 2014 г. в Институте языкознания РАН (вебсайт проекта: [multidiscourse.ru](http://multidiscourse.ru), более подробное описание см. в [Кибрик 2018б]). Ресурс включает 40 записей суммарной длительностью около 15 часов; в записях приняли участие 160 человек в возрасте от 18 до 36 лет, включая 110 женщин и 50 мужчин.

Для проведения исследования была разработана новая оригинальная методика. В каждой записи принимали участие четыре человека с заранее распределенными ролями. Три участника – Рассказчик, Комментатор и Пересказчик – участвовали в основной части записи, последний – Слушатель – присоединялся в конце. Сначала Рассказчик и Комментатор смотрели «Фильм о грушах» У. Чейфа [Chafe ed. 1980] и старались его запомнить. Затем к ним присоединялся Пересказчик. Задача Рассказчика состояла в том, чтобы рассказать сюжет фильма Пересказчику; это был этап *рассказа* в режиме монолога. На следующем интерактивном этапе *разговора* Комментатор дополнял рассказ Рассказчика, а Пересказчик уточнял детали. Наконец, на этапе *пересказа* Пересказчик в режиме монолога пересказывал сюжет фильма Слушателю. После этого Слушатель записывал услышанный пересказ. Таким образом, задача каждого участника состояла в том, чтобы максимально понятно донести до других полученную информацию, минимизировав эффект «испорченного телефона».

При сборе материала было использовано современное высокоточное оборудование. Три промышленные видеокамеры JAI с частотой 100 к/с записывали с фронтальной позиции каждого из трех основных участников; кроме того, камера GoPro фиксировала общий план. Для регистрации движений глаз были использованы две пары очков-айтрекеров Tobii Glasses с частотой 50 Hz; один из двух айтрекеров был надет на Рассказчика, второй – на Пересказчика.

Аннотирование ресурса осуществляется по всем основным каналам, независимо друг от друга. В частности, вокальная аннотация в программе PRAAT ([fon.hum.uva.nl/praat](http://fon.hum.uva.nl/praat))

состоит в членении речевого потока на значимые фрагменты (ЭДЕ, слова, (не)заполненные паузы), а также в приписывании определенных свойств ЭДЕ и отдельным их частям. Для аннотации мануальных движений в программе ELAN ([tla.mpi.nl/tools/tla-tools/elan](http://tla.mpi.nl/tools/tla-tools/elan)) была разработана новая методика [Литвиненко и др. 2017], включающая деление жестового потока на периоды неподвижности, в том числе позы, и движения, среди которых выделяются жесты, смены поз и адапторы – физиологически мотивированные движения (*adaptors*, [Ekman, Friesen 1969]). Аннотация движений головы также включает деление на жесты, смены поз и адапторы, которые противопоставлены периодам покоя [Сухова 2017]. В ходе окуломоторного аннотирования был произведен экспорт данных на видеосцену и с помощью программы Tobii Analyzer извлечены данные о временной развертке фиксации длительностью выше 100 мс, на которые потом в ручном режиме была наложена аннотационная схема [Федорова 2017].

### **3. Координация вокальной модальности с мануальным, цефалическим и окуломоторным каналами**

Рассмотрим попарное взаимодействие вокальной модальности с мануальным, цефалическим и окуломоторным каналами. Нас будет интересовать вопрос, являются ли они взаимонезависимыми или между ними существует определенная координация.

Несмотря на то, что в реальной коммуникации важны все каналы, большинство современных мультиканальных исследований до сих пор посвящены *мануальной* жестикуляции в ее взаимодействии с вокальным компонентом [Kendon 2004, McNeill 2005]. Важный теоретический вопрос в этой области – насколько жесты и речь связаны между собой в когнитивной системе человека. Автор наиболее влиятельной гипотезы Дэвид Макнилл утверждает, что жесты и речь одновременно активируются в одном общем источнике, и, следовательно, должны быть синхронизированы [McNeill 1992]. Макнилл предположил, что обычно время начала жеста опережает время начала речи [McNeill 1992], что было затем подтверждено в корпусных исследованиях на материале английского [Loehr 2012], французского [Ferré 2010] и польского [Karpíński et al. 2009] языков. Одно из возможных объяснений этого феномена состоит в том, что общий когнитивный источник, находящийся на досемантическом уровне, одновременно запускает активацию как абстрактных семантических репрезентаций, так и более конкретных моторных. Однако время, которое тратится на поиск моторных репрезентаций, обычно оказывается меньше, чем время, необходимое для поиска семантических репрезентаций. Отметим, что

исследование, проведенное на материале нашего корпуса, не подтвердило упомянутую закономерность [Федорова и др. 2016].

Тем не менее, если мы принимаем гипотезу Макнилла в целом, а также полагаем, что речь является обязательной составляющей коммуникации, то это означает, что мануальные жесты обычно используются только *говорящим*<sup>4</sup> собеседником. Данные нашего корпуса подтверждают это предположение – мануальная жестикуляция без вокальной поддержки встречается крайне редко, и только в том случае, когда коммуникант, которому на данном этапе запрещено говорить, использует пантомиму.

Иначе обстоит дело с мануальными адапторами и сменами поз – небазовыми элементами мануального канала. По данным нашего корпуса, мануальные адапторы и смены поз могут сопровождать как речь говорящего собеседника, так и речь слушающего, причем в самых разных пропорциях – от почти полного отсутствия (2% от общей длительности) до почти постоянного присутствия (98%).

*Цефалический* канал в целом в настоящее время изучен значительно хуже мануального, однако вопрос о взаимодействии движений головы с речью исследуется уже около полувека, причем работы 1960–1970 гг. до сих пор остаются актуальными и высокоцитируемыми. В статье [Dittmann, Llewellyn 1968] авторы проанализировали цефалическое поведение *слушающего* собеседника. Они пришли к выводу, что слушающий начинает кивать и говорить не в любой случайный момент, а на границах интонационных единиц в речи говорящего; авторы выделили два паттерна такого поведения – (1) перед тем, как задать вопрос или что-то прокомментировать и (2) в ответ на прямой вопрос говорящего; на долю двух этих паттернов приходится около 70% всего цефалического поведения слушающего; остальные 30% представляют собой маркеры обратной связи, как они были названы чуть позднее Виктором Ингве (backchannels, [Yngve 1970]). В работе [Dittmann, Llewellyn 1968] было также установлено, что задержка между кивком и началом вербализации составляет около 175 мс. Авторы предположили, что таким образом слушающий посылает говорящему сигнал о начале своего высказывания, не желая перебивать собеседника.

В 1972 г. Адам Кендон исследовал цефалическое поведение *говорящего* собеседника [Kendon 1972]. Проведя независимую сегментацию вокального и цефалического компонентов, он пришел к выводу, что выделенные единицы имеют тенденцию к

---

<sup>4</sup> Здесь и далее мы будем использовать термины «говорящий» и «слушающий» в узком смысле: не как передающий / принимающий информацию, а буквально – как относящийся к вокальной модальности.

синхронизации. Кендон показал, в частности, что к концу очередной просодической единицы голова говорящего собеседника часто возвращается в то место, в котором она была в начале предыдущей просодической единицы, тем самым обеспечивая кинетическую связность дискурса.

Данные нашего корпуса подтверждают выявленные закономерности; в частности, говорящий часто «передает ход» собеседнику поворотом головы, а слушающий кивает в ответ на слова собеседника и/или перед началом своего высказывания.

Наконец, ключевая работа в области *окуломоторного* канала коммуникации также была написана около полувека назад. Проанализировав пятиминутное общение семи пар испытуемых, Кендон (1967) выявил следующие закономерности: (1) испытуемый чаще смотрит на собеседника, когда слушает его, чем когда сам говорит; (2) взгляды на собеседнике длиннее, когда испытуемый молчит, чем когда говорит; (3) когда испытуемый молчит, его взгляды на собеседнике длиннее, чем взгляды на окружение; (4) когда испытуемый говорит, его взгляды на собеседнике короче, чем взгляды на окружение; (5) наблюдаются сильные индивидуальные различия. Так, испытуемые-слушатели фиксировали взгляд на собеседнике от 32% до 81% всего времени, а испытуемые-говорящие смотрели на собеседника от 20% до 68% всего времени [Kendon 1967].

Проанализировав данные нашего корпуса, мы выявили аналогичные закономерности: (1) *говорящий* собеседник смотрит на адресата своего сообщения от 21% до 72% всего времени; остальное время он смотрит на окружение; (2) *слушающий* собеседник смотрит на говорящего 80%–90% всего времени, причем на этапе *рассказа* Пересказчик смотрит на говорящего Рассказчика 95%–100%, а на этапе *пересказа* Рассказчик смотрит на говорящего Пересказчика чуть реже, 81%–99%; (3) на этапе *разговора* оба *слушающих* собеседника в 95% случаев смотрят на говорящего собеседника (подробнее см. [Федорова в печати]).

Таким образом, мы показали, что для каждого из трех основных не-вокальных каналов существует координация с вокальным компонентом<sup>5</sup>. Прототипический *говорящий* собеседник, время от времени смотря на слушающего, обычно жестикулирует при помощи рук, а при помощи жестов головы структурирует собственную речь и регулирует процесс смены реплик. Прототипический *слушающий* собеседник почти постоянно смотрит на

---

<sup>5</sup> Необходимо отметить, однако, что если мы рассмотрим взаимодействие не вокального, а, например, мануального компонента с окуломоторным, мы также получим определенную координацию: «когда человек жестикулирует при помощи рук, он смотрит на собеседника только время от времени, а когда не жестикулирует – то почти постоянно». Однако очевидно, что эта координация опосредована вокальной модальностью и связана с процессом порождения / понимания звучащей речи.

говорящего, не жестикулирует при помощи рук, а при помощи жестов головы обеспечивает обратную связь и сигнализирует говорящему о начале собственной речи.

### **Заключение. Стержневая роль вокальной модальности**

Как уже было отмечено выше, традиционная лингвистика XX в. придавала исключительное значение вербальному каналу, а все остальные компоненты в лучшем случае обозначались несколько уничижительным термином «паралингвистика». В практической психологии была представлена другая крайность – часто цитировались цифры из работы [Mehrabian 1971], согласно которой язык тела (=кинетическая модальность) передает 55% информации, просодия – 38%, а вербальный канал – лишь 7%. Эти цифры, однако, были получены в очень специфическом исследовании – при анализе вклада каналов в формирование установки слушающего по отношению к эмоциональным сообщениям.

Около 10 лет назад один из авторов этой статьи разработал экспериментальный подход, позволяющий количественно оценить вклад каждого из трех коммуникативных компонентов – вербального, просодического и кинетического – в общий процесс понимания дискурса. В ходе эксперимента испытуемый получал задание просмотреть / прослушать запись коммуникации и затем ответить на вопросы к содержанию этой записи. Всего в эксперименте было восемь экспериментальных условий: изолированный вербальный компонент (каждое слово было произнесено отдельно, а затем они были склеены в нужном порядке), изолированный просодический компонент (на исходный звук был наложен фильтр, создающий эффект «разговора за стеной»), изолированный кинетический компонент (только видеоряд), пары компонентов в трех возможных комбинациях, все три компонента и контрольная группа для проверки случайного угадывания правильных ответов. В результате оказалось, что относительный вклад вербального, просодического и кинетического компонентов составил 39%, 27% и 35%, соответственно. Таким образом, все три компонента играют существенную роль в понимании естественного дискурса. Оказалось, однако, что в условии «кинетический плюс просодический» наблюдается заметный провал, то есть испытуемые испытывали трудности с интеграцией информации из кинетического и просодического компонентов в отсутствие вербального подкрепления. Это говорит о том, что вербальный компонент играет роль якоря, к которому прикрепляется информация, поступающая по другим каналам [Kibrik, Molchanova 2013].

Исследование, описанное в [Kibrik, Molchanova 2013], было выполнено при помощи *экспериментальной* методологии при изучении *понимания* дискурса. Тем не менее,

полученные выводы о необходимости учета всех коммуникативных каналов и связующей, опорной, стержневой роли вербального компонента перекликаются с выводами настоящей работы, проведенной на материале анализа *порождения* и *понимания* речи в *естественной* коммуникации. Отметим, однако, что в естественной коммуникации лицом к лицу вербальный и просодический каналы образуют неразрывное целое – слова и просодия не существуют друг без друга, так что мы можем разделить их только в результате непростых экспериментальных манипуляций. Поэтому, обобщая результаты обоих исследований, мы говорим о связующей роли вокальной модальности в целом. Нельзя безоговорочно утверждать, что вокальная модальность всегда является ведущей в процессе передачи информации между коммуникантами. В отдельные моменты коммуникации самым важным может оказаться отдельный мануальный жест говорящего собеседника<sup>6</sup>, или его взгляд, или отрицательное покачивание головой слушающего собеседника. Однако в целом именно вокальный компонент является стержневым. Этот вывод хорошо согласуется с тем, что в обычной бытовой коммуникации кинетические каналы являются факультативными – например, мы без труда понимаем сообщение собеседника, переданное по телефону, хотя отсутствие кинетического компонента и делает его беднее.

Таким образом, мы приходим к *вокальноцентрическому* подходу, при котором приоритетной считается вокальная модальность, а базовой единицей всего мультиканального дискурса является ЭДЕ, выделенная преимущественно по просодическим основаниям.

## Литература

*Бутовская М.Л.* Язык тела: природа и культура (эволюционные и кросскультурные основы невербальной коммуникации человека). М.: Научный мир, 2004.

*Гришина Е.А.* Русская жестикуляция с лингвистической точки зрения. Корпусные исследования. М.: Языки славянской культуры, 2017.

*Демьянков В.З.* Текст и дискурс как термины и как слова быденного языка // В.Н. Топоров (ред.) Язык. Личность. Текст: Сборник к 70–летию Т.М. Николаевой. М.: Языки славянских культур, 2005. С. 34–55.

*Квинтилиан М.Ф.* Двенадцать книг риторических наставлений / Пер. с лат. А. Никольского. 1834. Ч. 2. 522 с.

---

<sup>6</sup> Ср. цитату из [Корогаева 2015]: «в фольклоре ведущая роль отводится не смысловому значению слова, а рукодвижению. Слова могут иметь неопределенное значение, быть заумными, но жесты, сопровождающие их, имеют конкретное назначение, производятся с определенной целью. Не важно, какие слова человек говорит, а важен сам жест как способ воздействия.»

*Кибрик А.А.* Мультимодальная лингвистика // Ю.И. Александров, В.Д. Соловьев (ред.) Когнитивные исследования. Вып. IV. М.: ИП РАН. С. 134–152.

*Кибрик А.А.* Когнитивный анализ дискурса: локальная структура // А.А. Кибрик, А.Д. Кошелев, А.В. Кравченко, Ю.В. Мазурова, О.В. Федорова (ред.) Язык и мысль: современная когнитивная лингвистика. Москва, ЯСК: 2015. С. 595–634.

*Кибрик А.А.* Русский мультимедийный дискурс. Часть I. Постановка проблемы // Психологический журнал, 2018а. Т. 39(1). С. 70–80.

*Кибрик А.А.* Русский мультимедийный дискурс. Часть II. Разработка корпуса и направления исследований // Психологический журнал, 2018б. Т. 39(2). С. 78–89.

*Кибрик А.А., Подлеская В. И. (ред.).* Рассказы о сновидениях: корпусное исследование устного русского дискурса. М.: ЯСК, 2009.

*Коротаева Е.В.* Рукодвижения в фольклоре и авангарде // В.В. Фещенко (ред.) Живое слово: логос – голос – движение – жест: Сборник статей и материалов. М.: Новое литературное обозрение, 2015. С. 303–312.

*Крейдлин Г.Е.* Невербальная семиотика. М.: НЛЮ, 2002.

*Литвиненко А.О., Николаева Ю.В., Кибрик А.А.* Аннотирование русских мануальных жестов: теоретические и практические вопросы // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог 2017», М., 2017. С. 271–286.

*Сухова Н.В.* Значения жестовых единиц: к вопросу об аннотировании жестов головы // А. Д. Кривоносов (ред.) Российская психология-4: тренды и драйверы. Сборник трудов в честь профессора Л. В. Минаевой. СПб.: Изд-во СПбГЭУ, 2017. С. 82–87.

*Федорова О.В., Кибрик А.А., Коротаев Н.А., Литвиненко А.О., Николаева Ю.В.* Временная координация между жестовыми и речевыми единицами в мультимодальной коммуникации // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции Диалог (Москва, 1–4 июня 2016 г.). Вып. 15 (22). М.: Изд-во РГГУ, 2016. С. 159–170.

*Федорова О.В.* Распределение зрительного внимания собеседников в естественной коммуникации: 50 лет спустя // Е.В. Печенкова, М.В. Фаликман (ред.) Когнитивная наука в Москве: новые исследования. Материалы конференции 15 июня 2017 г. М.: БукиВеди, ИППиП, 2017. С. 370–375.

*Федорова О.В.* О коммуникативной функции взгляда // Труды Института русского языка им. В.В. Виноградова РАН, в печати.

*Якубинский Л.П.* О диалогической речи // Русская речь, 1, 1923.

*Adolphs S., Carter R.* Spoken corpus linguistics: From monomodal to multimodal. N.-Y.: Routledge, 2013.

*Chafe W. (ed.).* The pear stories: Cognitive, cultural, and linguistic aspects of narrative production. Norwood, 1980.

*Dittmann A., Llewellyn L.* Relationship between vocalizations and head nods as listener responses // Journal of Personality and Social Psychology, 1968. Vol. 9. P. 79–84.

*Ekman F., Friesen W.V.* The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding // Semiotica, 1969. Vol. 1(1). P. 49–98.

*Ferré G.* Timing Relationships between Speech and Co-Verbal Gestures in Spontaneous French // Language Resources and Evaluation, Workshop on Multimodal Corpora, May 2010, Malta. P. 86–91.

*Karpiński M., Jarmolowicz-Nowikow E., Malisz Z.* Aspects of gestural and prosodic structure of multimodal utterances in Polish task-oriented dialogues // *Speech and Language Technology*, 2011. Vol. 11. P. 113–122.

*Kendon A.* Some functions of gaze direction in social interaction // *Acta Psychologica*, 1967. Vol. 26. P. 22–63.

*Kendon A.* Some relationships between body motion and speech // A.W. Siegman, B. Pope (eds.) *Studies in dyadic communication*. Elmsford: Pergamon, 1972. P. 177–210.

*Kendon A.* Gesticulation and speech: Two aspects of the process of utterance // M.R. Key (ed.) *The relationship of verbal and nonverbal communication*, 1980. P. 207–227.

*Kendon A.* *Gesture. Visible action as utterance*. Cambridge, 2004.

*Kibrik A.A., Molchanova N.B.* Channels of multimodal communication: Relative contributions to discourse understanding // M. Knauff, M. Pauen, N. Sebanz, I. Wachsmuth (yds.) *Proceedings of the 35th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society, 2013. P. 2704–2709.

*Knight D.* *Multimodality and active listenership: A corpus approach*. London: Bloomsbury, 2011.

*Kress G.* *Multimodality: A social semiotic approach to communication*, Routledge Falmer, London, 2010.

*Linell P.* *The written language bias in linguistics*. Linköping, Sweden: University of Linköping, 1982.

*Loehr D.* Temporal, structural, and pragmatic synchrony between intonation and gesture // *Laboratory Phonology*, 2012. Vol. 3(1). P. 71–89.

*McNeill D.* *Hand and mind: What gestures reveal about thought*. Chicago, 1992.

*McNeill D.* *Gesture and thought*. Chicago, 2005.

*Mondada L.* Challenges of multimodality: Language and the body in social interaction // *Journal of sociolinguistics*, 2016. Vol. 20. P. 336–366.

*Müller C., Fricke E., Cienki A., McNeill D. (eds.)*. *Body — Language — Communication*, Mouton de Gruyter, Berlin, 2013.

*Yngve V.* On getting a word in edgewise // *Chicago Linguistic Society*, 1970. P. 567–578.