

О.В. Федорова, А.А. Кибрик (Россия)
Институт языкознания РАН,
Московский государственный университет имени М.В. Ломоносова
olga.fedorova@msu.ru, aakibrik@gmail.com

ОБЩЕЕ, ИНДИВИДУАЛЬНОЕ И КОНТЕКСТ В МУЛЬТИКАНАЛЬНОЙ КОММУНИКАЦИИ

Вопрос об общих закономерностях и индивидуальных различиях принадлежит к числу ключевых вопросов мультимедийных исследований. Изучая этот вопрос на материале нового ресурса «Рассказы и разговоры о грушах», мы вводим понятия индивидуального «портрета» и общей / индивидуальной закономерности. Предложенная методология важна для скрупулезного анализа полученных данных.

Ключевые слова: мультимедийная лингвистика, индивидуальные различия, портрет говорящего, жестикация, движение взора

Мультимедийная (бимодальная, мультимодальная) *лингвистика* – молодое и в то же время одно из старейших направлений исследований языка в широком смысле слова, которое изучает все реальное многообразие «живой» коммуникации между людьми: слова, интонацию, жестикацию, направление взора, мимику. Интерес к изучению мультимедийности возник еще в древности. Так, Квинтилиан, римский ритор I в. н.э., автор самого полного учебника ораторского искусства, дошедшего до нас из античности, писал: «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].

Слова Квинтилиана звучат очень современно. Как и в те времена, мы выделяем две основные модальности – *вокальную* (слуховую) и *кинетическую* (зрительную), и считаем первую более базовой. Вокальная модальность далее подразделяется на два канала – вербальный и просодический; кинетическая модальность (язык тела) имеет сложную структуру, которая включает как динамический компонент – жесты лица, головы, рук и туловища, так и статический – выражение лица и позы; подробнее см. [Кибрик 2018а,б].

Одним из первых мультимедийных исследований на современном этапе стала работа классика изучения жестов Адама Кендона о распределении зрительного внимания собеседников в ходе естественной коммуникации [Kendon 1967]. Проанализировав общение семи пар испытуемых (И.), Кендон заключил, что:

1. И. чаще смотрит на собеседника, когда слушает его, чем когда сам говорит.
2. Взгляд на собеседника длиннее, когда И. молчит, чем когда говорит.
3. Когда И. молчит, его взгляд на собеседника длиннее, чем взгляд на окружение.

4. Когда И. говорит, его взгляд на собеседнике короче, чем взгляд на окружение.
5. Наблюдаются сильные индивидуальные различия. Так, И.-слушатели фиксируют взгляд на собеседнике от 32% до 81% всего времени, а И.-говорящие смотрели на собеседника от 20% до 68% всего времени [Kendon 1967].

Как мы видим, Кендон (1967) обсуждает два полярных вопроса – общие закономерности, с одной стороны, и индивидуальные различия, с другой. Имея работу Кендона (1967) в качестве исходного ориентира, в настоящей статье мы рассмотрим оба этих ключевых вопроса на материале создаваемого нами ресурса *«Рассказы и разговоры о грушах»*, проект по сбору которого осуществляется в Институте языкознания РАН (сайт multidiscourse.ru). Ресурс включает 40 записей суммарной длительностью около 15 часов. При сборе материала было использовано современное оборудование, в том числе промышленные видеокамеры с частотой 100 к/с и две пары очков-айтрекеров Tobii Glasses II с частотой 50 Гц.

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

Аннотирование ресурса осуществляется по всем основным каналам, включая вербальный, просодический [Кибрик, Подлесская ред. 2009], кинетический и окуломоторный. Вокальная аннотация в программе PRAAT (fon.hum.uva.nl/praat) состоит в членении речевого потока на значимые фрагменты (элементарные дискурсивные единицы – ЭДЕ, слова, (не)заполненные паузы), а также в приписывании свойств ЭДЕ и отдельным их частям. Для аннотации мануальных движений (движений рук) в программе ELAN (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 (self-adaptors, [Ekman, Friesen 1969]). 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].

К настоящему моменту мы полностью аннотировали и выложили на сайт три записи № 04, 22 и 23 суммарной длительностью около 1 часа, которые составляют

эталонный подкорпус. *Эталонный подкорпус* – это своего рода экспериментальная площадка, на которой мы тестируем различного рода гипотезы, чтобы потом верифицировать их на более обширном материале. Прежде чем переходить непосредственно к теме данной работы, надо убедиться, что три записи подкорпуса *сопоставимы* между собой. Как можно видеть из Табл. 1, несмотря на то, что мы не устанавливали для испытуемых временных ограничений, все три записи похожи друг на друга по длительностям отдельных этапов – на *рассказ* приходится около 20% от времени всей записи, на *разговор* – примерно 50%, и на *пересказ* – 30%.

№	общая длительность записи	рассказ	разговор	пересказ
04	24:36.240	05:22.640 (21.9%)	12:37.920 (51.3%)	06:35.680 (26.8%)
22	18:04.960	03:37.960 (20.1%)	08:48.280 (48.7%)	05:38.700 (31.2%)
23	16:26.520	03:52.400 (23.5%)	07:41.240 (46.8%)	04:52.880 (29.7%)

Таблица 1. Дескриптивная статистика эталонного подкорпуса по этапам

В данной работе мы сначала остановимся на индивидуальных различиях и предложим «портретный» подход к мультиканальности. Мы покажем, что создание таких портретов является важным условием скрупулезного аннотирования и последующего анализа дискурса. Затем мы опишем разного рода закономерности, которые можно заметить при анализе данных индивидуальных портретов. В силу небольшого размера статьи мы ограничимся рассмотрением двух каналов в рамках кинетической модальности – мануального и окуломоторного.

Портрет – это совокупность характеристик мультиканального поведения человека. Мы выделяем просодический [Кибрик 2009], окуломоторный и жестикуляционный портреты. В частности, *окуломоторный портрет* создается после процедуры полуавтоматической аннотации и используется на этапе анализа материала. Различия между людьми оказываются весьма существенными, поэтому мы не можем корректно сравнивать эти данные без усреднения для последующего дисперсионного анализа. Эту функцию и выполняют окуломоторные портреты. В Табл. 2 представлен окуломоторный портрет испытуемого R04.

Жестикуляционный портрет выполняет двойную функцию – он необходим как на этапе аннотирования жестов, так и на этапе их анализа. Данные, необходимые для аннотирования, включают: (не)склонность к состоянию покоя, (не)склонность к использованию адапторов, типичную скорость и амплитуду жестов, преобладание двуручных или одноручных жестов, распределение по рукости. Данные, необходимые для анализа, включают: суммарное количество мануальных жестов, их длительность, среднюю, минимальную, максимальную длительности, а также распределение по 25, 50 и 75% квантилям.

Total duration	1170.667
Total count	2249
Mean duration & std	0.52 & 0.66
Min, Quantile 25, 50, 75, Max	0.06, 0.16, 0.26, 0.58, 10.5
R: count & ratio	1033 & 0.46
R: duration & ratio	848.55 & 0.72
R: mean & std	0.82 & 0.85
R: Min, Quantile 25, 50, 75, Max	0.06, 0.24, 0.5, 1.16, 10.5
C: count & ratio	129 & 0.057
C: duration & ratio	43.8 & 0.036
C: mean & std	0.33 & 0.28
C: Min, Quantile 25, 50, 75, Max	0.06, 0.16, 0.24, 0.4, 1.32
L: count & ratio	16 & 0.007
L: duration & ratio	3.28 & 0.003
L: mean duration & std	0.2 & 0.12
L: Min, Quantile 25, 50, 75, Max	0.06, 0.12, 0.19, 0.26, 0.44
Surroundings: count & ratio	1071 & 0.48
Surroundings: duration & ratio	275.02 & 0.23
Surroundings: mean & std	0.26 & 0.21
Surroundings: Min, Quantile 25, 50, 75, Max	0.06, 0.13, 0.2, 0.3, 2.2

Таблица 2. Окуломоторный портрет испытуемого R04 (разработчик И.Ю. Жердев)

Перейдем теперь к описанию *закономерностей*, обнаруженных на материале собранных портретов. Табл. 3 представляет *мануальные* данные, где для каждого участника различаются жесты (Ge) и адапторы (Ad), условно включающие смены поз; аннотация записей произведена А.О. Литвиненко и Ю.В. Николаевой:

№	рассказ						разговор						пересказ					
	N		C		R		N		C		R		N		C		R	
	Ge	Ad	Ge	Ad	Ge	Ad	Ge	Ad	Ge	Ad	Ge	Ad	Ge	Ad	Ge	Ad	Ge	Ad
04	81	8	0	18	0	98	14	61	13	41	26	59	0	12	0	28	52	4
22	57	4	0	2	0	16	23	28	31	37	2	40	0	4	0	28	33	54
23	57	13	0	2	0	4	30	43	32	18	9	8	0	54	0	2	65	15

Таблица 3. Распределение длительностей движений рук, в % от общей длительности

Вводя две оппозиции «общее vs. индивидуальное» и «контекстно-зависимое vs. контекстно-независимое», мы получаем следующую классификацию (далее «мало» значит $\leq 1/4$, «немного» – $\leq 1/3$, «много» – $\geq 3/4$ и «постоянно» – $\geq 95\%$):

I. Общие контекстно-независимые. Это самые общие закономерности, которые верны для всех испытуемых на всех этапах записи независимо от их роли:

- И. использует как жесты, так и адапторы;

II. Общие контекстно-зависимые. Эти закономерности зависят от роли И. и/или характерны для всех И. на некотором этапе:

- На этапах *рассказа* и *пересказа* жесты есть только у говорящего – N или R, соответственно;

- С использует немного жестов;

III. Индивидуальные контекстно-независимые. Эти закономерности характерны для конкретных И. на всех этапах записи:

- R22 использует больше адапторов, чем жестов;
- R23 использует мало адапторов;

IV. Индивидуальные контекстно-зависимые. Такие закономерности характерны для конкретных И. на некотором этапе записи:

- R04 на этапе *рассказа* постоянно использует адапторы.

Наконец, рассмотрим **окуломоторные закономерности**, см. Табл. 4 и 5:

Этап	Направление взора	R	C	Окружение (включая L)
Рассказ	№04	61	0	39
	№22	21	0	79
	№23	72	0	28
Разговор	№04	68	5	27
	№22	34	56	10
	№23	21	57	22
Пересказ	№04	91	0	9
	№22	99	0	1
	№23	82	0	18

Таблица 4. Распределение зрительного внимания N по этапам, данные анализа суммарной длительности фиксации больше 100 мс в программе Tobii Analyzer, в %

Этап	Направление взора	N	C	Окружение / L на этапе пересказа
Рассказ	№04	98	1	1
	№22	100	0	0
	№23	95	1	4
Разговор	№04	47	40	13
	№22	41	59	0
	№23	45	48	7
Пересказ	№04	0	0	47 / 53
	№22	2	0	38 / 60
	№23	4	1	61 / 32

Таблица 5. Распределение зрительного внимания R по этапам, данные анализа суммарной длительности фиксации больше 100 мс в программе Tobii Analyzer, в %

Мы выделяем следующие закономерности:

I. Общие контекстно-независимые.

- И. чаще смотрит на собеседника, когда слушает, чем когда говорит сам;

II. Общие контекстно-зависимые.

- С притягивает взгляды собеседников только на этапе *разговора*;
- На этапе *пересказа* внимание R распределено между L и окружением;

- На этапе *рассказа* R постоянно смотрит на N;
- III. Индивидуальные контекстно-независимые.
- N04 не смотрит на C04;
- IV. Индивидуальные контекстно-зависимые.
- На этапе *рассказа* N22 мало смотрит на R22.

Как представляется, создание индивидуальных портретов и выявление на их основе общих и индивидуальных закономерностей поведения испытуемых для каждого канала является необходимым условием для проведения более детального и скрупулезного исследования обнаруженных явлений как по отдельности, так и во взаимодействии с явлениями других каналов.

Работа выполнена при финансовой поддержке РФФ, грант №14-18-03819 «Язык как он есть: русский мультимодальный дискурс».

Литература

- Кибрик А. А.* Просодические портреты говорящих // Кибрик А. А., Подлесская В. И. (ред.). Рассказы о сновидениях: корпусное исследование устного русского дискурса. М.: ЯСК, 2009. С. 478–487.
- Кибрик А. А.* Русский мультимодальный дискурс. Часть I. Постановка проблемы // Психологический журнал. 2018а. Т. 39 (1). С. 70–80.
- Кибрик А. А.* Русский мультимодальный дискурс. Часть II. Разработка корпуса и направления исследований // Психологический журнал, 2018б. Т. 39 (2) С. 78–89.
- Кибрик А. А., Подлесская В. И.* (ред.). Рассказы о сновидениях: корпусное исследование устного русского дискурса. М.: ЯСК, 2009.
- 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.
- Марка Фабия Квинтилиана* Двенадцать книг риторических наставлений / Пер. с лат. А. Никольского. 1834. Ч. 2. 522 с.
- 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, ИПИП, 2017. С. 370–375.
- Ekman F., Friesen W. V.* The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding // *Semiotica*. 1969. Vol. 1 (1). P. 49–98.
- Kendon A.* Some functions of gaze direction in social interaction // *Acta Psychologica*. 1967. Vol. 26. P. 22–63.

O.V. Fedorova, A.A. Kibrik (Russia)
Institute of Linguistics RAS,
Lomonosov Moscow State University
olga.fedorova@msu.ru, aakibrik@gmail.com

**GENERAL, INDIVIDUAL, and CONTEXT
IN MULTICHANEL COMMUNICATION**

Among the key issues in multichannel studies is the question of general trends and individual variation. We address these issues with the help of a new multichannel resource “Russian Pear Chats and Stories”. We introduce the notions of a speaker’s individual portrait and a general / individual regularity. The proposed methodology is important for fine-grained analysis of multichannel data.

Keywords: multichannel linguistics, individual differences, speaker's portrait, gesticulation, eye gaze