

Компьютерная лингвистика и интеллектуальные технологии:
по материалам международной конференции «Диалог 2016»

Москва, 1–4 июня 2016

ВЕРБАЛЬНАЯ РАБОЧАЯ ПАМЯТЬ И ЛЕКСИКО-ГРАММАТИЧЕСКИЕ СИГНАЛЫ РЕЧЕВЫХ ЗАТРУДНЕНИЙ: ДАННЫЕ РУССКОГО МУЛЬТИМОДАЛЬНОГО КОРПУСА¹

Потанина Ю. Д. (yulia.potanina.msu@gmail.com)

МГУ имени М. В. Ломоносова, Москва, Россия

Подлеская В. И. (vi_podlesskaya@il-rggu.ru)

РГГУ, РАНХиГС, Москва, Россия

Федорова О. В. (olga.fedorova@msu.ru)

МГУ имени М. В. Ломоносова, Институт языкознания РАН,
РАНХиГС, Москва, Россия

Хорошо известно, что объем вербальной рабочей памяти коррелирует с индивидуальными различиями испытуемых при понимании речи (Daneman, Carpenter 1980). В тоже время связь объема рабочей памяти и речепорождения пока изучена относительно слабо. В настоящей работе предпринята попытка восполнить этот пробел и проверить гипотезу о корреляции между объемом рабочей памяти и количеством лексико-грамматических маркеров затруднений при порождении спонтанного пересказа. Девятнадцать русскоязычных испытуемых приняли участие в двух тестах: тесте «Speaking span», в котором им был измерен объем рабочей памяти, и тесте на порождение речи, основанном на пересказе «Фильма о грушах» (Chafe 1980). Собранные пересказы были расшифрованы и вручную аннотированы с точки зрения лексико-грамматических сигналов речевых затруднений. Оказалось, что количество сигналов речевых затруднений в пересказах колеблется от 0,77 до 8,58 на 100 слов, что соответствует средним значениям, характерным для спонтанной речи. Проведенный статистический анализ показал, что объем рабочей памяти действительно коррелирует с беглостью речи, измеренной в числе лексико-грамматических маркеров речевых затруднений.

Ключевые слова: вербальная рабочая память, порождение речи, беглость речи, лексико-грамматические сигналы речевых затруднений

¹ Работа выполнена при финансовой поддержке РФФ (проект № 14-18-03819).

VERBAL WORKING MEMORY AND SPEECH PRODUCTION DIFFICULTIES: DATA FROM RUSSIAN MULTIMODAL CORPUS

Potanina Y. D. (yulia.potanina.msu@gmail.com)

Lomonosov Moscow State University, Moscow, Russia

Podlesskaya V. I. (vi_podlesskaya@il-rggu.ru)

RSUH, RANEPa, Moscow, Russia

Fedorova O. V. (olga.fedorova@msu.ru)

Lomonosov Moscow State University, Institute of Linguistics
RAS, RANEPa, Moscow, Russia

It's a well-known fact that working memory capacity correlates with individual differences in comprehending speech (Daneman, Carpenter 1980). At the same time, the relationship between working memory capacity and speech production remains relatively unexplored. In this paper, we attempt to partially fill the gap and check the hypothesis about correlation between working memory capacity and number of lexical and grammatical markers of difficulties in production of spontaneous narratives. 19 Russian participants took part in two tests: the "Speaking Span" test by which we have measured their working memory capacity and the speech production test based on retelling the Pear Film (Chafe 1980). The Speaking Span test was designed in (Daneman and Green 1986) for English-speaking individuals: during the test increasingly longer sets of words are presented to participants; at the end of each set, they are supposed to use each word to generate a separate sentence (the word should be in the same grammatical form as it has been presented). Speaking span is measured as the maximum number of semantically and grammatically correct sentences produced in the experiment. This test was adapted to Russian: words in the set were balanced by syntactic categories, frequencies of individual lexemes and frequencies of grammatical forms. Collected narratives have been transcribed and manually annotated for lexical and grammatical markers of production difficulties. The documented number of lexical and grammatical markers of speech production difficulties varied between 0.77 and 8.58 per 100 words, which matches average rates reported previously in the literature. The study demonstrates the statistically significant correlation between working memory capacity measured by the "Speaking Span" test and verbal fluency measured in number of lexical and grammatical markers of production difficulties.

Key words: verbal working memory, speech production, verbal fluency, lexical and grammatical markers of production difficulties

1. Исследования рабочей памяти в когнитивной психологии и психолингвистике

Начало современного этапа изучения памяти в области *когнитивной психологии* связано с именем Г. Эббингауза, который в конце XIX в. разработал первые экспериментальные методики ее измерения. Примерно в то же время У. Джеймс предложил разделить память на первичную и вторичную (James 1890). В начале второй половины XX в. в работах Дж. Миллера (Miller 1956) и Р. Аткинсона & Р. Шиффрина (Atkinson, Shiffrin 1968) был сформулирован многокомпонентный подход к памяти. Согласно наиболее известной трехкомпонентной модели Аткинсона и Шиффрина (1968), сначала некоторая входящая информация попадает в сенсорные регистры (так называемый иконический (в зрительной модальности) и эхоический (в слуховой модальности) виды памяти), затем на 10–25 с. переводится в кратковременное хранилище, после чего попадает в долговременную память. Термин «рабочая память» (РП) был впервые использован в (Miller et al. 1960). Употребление термина «рабочая память» вместо «кратковременная память» подчеркивает функциональную значимость этой системы. Согласно современным представлениям, РП состоит из четырех модулей: (1) центрального исполнителя, (2) фонологической петли, (3) визуально-пространственной матрицы и (4) эпизодического буфера (Baddeley 2000). Центральный исполнитель является ядром системы, отвечающим за координацию работы всех ее подсистем, а три других модуля выполняют вспомогательные функции. Вербальная информация, поступающая из первичного сенсорного хранилища, попадает в фонологическую петлю, которая в свою очередь состоит из пассивного фонологического хранилища и подсистемы, обеспечивающей субвокальное повторение, которое препятствует угасанию следа речевого сигнала в памяти; без поддержки системы субвокального повторения информация в фонологическом хранилище угасает примерно через 1,5 с. Зрительная информация, поступающая из первичного сенсорного хранилища, попадает в визуально-пространственную матрицу, также состоящую из двух подсистем — зрительной и пространственной. Многомодальный эпизодический буфер используется для синтеза информации из фонологической петли и визуально-пространственной матрицы, а также для связи с долговременной памятью.

В области *психолингвистики* термин «рабочая память» используется с начала 1980-ых гг. Сначала тест на определение объема РП был разработан в области *понимания речи* (Daneman, Carpenter 1980). Данеман и Карпентер предположили, что в процессе интерпретации текста в РП происходят процессы, связанные как с пассивным хранением поступающей информации, так и с ее активной обработкой. Разработанный ими тест “Reading span” учитывал оба этих процесса: испытуемые читали отдельные предложения (обработка) и одновременно удерживали в памяти последние слова ранее прочитанных предложений (хранение). По словам Данеман, «теория кратковременной памяти была заменена теорией РП, а методика измерения кратковременной памяти — методикой измерения РП» (Daneman 1994: 443, *перевод наш*). Используя этот тест, авторы показали, что деление людей на «хороших читателей»,

которые умеют эффективно распределять ресурсы РП между хранением и обработкой поступающей информации, и «плохих читателей», которые делают это хуже, имеет под собой экспериментальные свидетельства. Чуть позже аналогичный тест был создан и для изучения процессов *порождения речи* (Daneman, Green 1986; Daneman 1991). Тест «Speaking span» состоял в следующем. Для эксперимента было отобрано 100 слов, которые были распределены в блоки по 2, 3, 4, 5 и 6 слов, в эксперименте было использовано по пять блоков каждого типа. Каждое слово появлялось на экране на 1 с. Испытуемый получал инструкцию читать слова, появляющиеся на экране, и, увидев пустой экран, придумывать с каждым прочитанным словом по одному предложению, причем целевое слово в этом предложении должно было стоять в той же грамматической форме. Например, прочитав слова *shelter*, *muscles* и *dangers*, англоязычный испытуемый произносил: *Trees provide poor shelter during a thunderstorm; Mr. Universe has very big muscles; There are dangers associated with every occupation*. Объем РП приравнивался к количеству слов, с которыми испытуемый смог придумать предложения.

Оба теста были адаптированы к русскому материалу, об адаптации теста «Reading span» см. работу (Федорова 2003), об адаптации «Speaking span» — работу (Федорова, Потанина 2014).

2. Рабочая память и лексико-грамматические сигналы речевых затруднений

Еще в работе (Daneman 1991) было показано, что объем РП, измеренный с помощью теста «Speaking span» коррелирует с беглостью речи. В работе (Федорова, Потанина 2014) эта корреляция была проверена на русском материале. Исследование, проведенное с 32 испытуемыми, состояло из пяти тестов: 1) «Speaking span»; 2) порождение речи; 3) чтение вслух; 4) тест с провоцированием испытуемых к оговоркам; 5) чтение скороговорок. В результате были получены значимые корреляции между объемом РП и (1) количеством слов в тесте на порождение речи ($\text{cor} = 0,522$, $p\text{-value} < 0,05$); (2) временем чтения вслух ($\text{cor} = -0,704$, $p\text{-value} < 0,01$); (3) количеством оговорок в тесте на оговорки ($\text{cor} = -0,706$, $p\text{-value} < 0,01$); (4) временем чтения скороговорок ($\text{cor} = -0,500$ и $-0,471$ для первой и второй попыток прочтения, соответственно, $p\text{-value} < 0,05$). Однако, в отличие от результатов Данеман, значимой корреляции между объемом РП и количеством ошибок при чтении вслух (как для художественного текста, так и для скороговорок) обнаружить не удалось. Цель данного исследования состоит в изучении взаимосвязи объема РП и маркеров речевых затруднений при порождении устных пересказов. Мы предполагаем, что объем РП будет обратно коррелировать с количеством маркеров речевых затруднений, т. е. при относительно большом объеме РП испытуемого мы получим относительно небольшое количество маркеров речевых затруднений в его речи, и, наоборот, небольшой объем РП будет сопровождаться большим количеством речевых затруднений.

2.1. Сбор подкорпуса «Рассказы о грушах»

Исследование по определению взаимозависимости между объемом РП и маркерами речевых затруднений было выполнено на материале 19 пересказов известного «Фильма о грушах» У. Чейфа (Chafe 1980). Корпус, отобранный для проведения данного исследования, состоял из двух частей. Первую часть (9 пересказов) составили записи, собранные по традиционной процедуре: каждый испытуемый смотрел незнакомый ему шестиминутный фильм, а затем пересказывал его содержание другому человеку, который этот фильм не видел. Вторая часть пересказов (10 записей) была взята из подкорпуса «Рассказы и разговоры о грушах», являющегося частью проекта РФ «Язык как он есть: русский мультимодальный дискурс», включающего 24 записи общей длительностью 10 часов и объемом более 110 000 слов. В каждой записи принимали участие 4 человека с заранее распределенными ролями Рассказчика, Комментатора, Пересказчика и Слушателя; для настоящего исследования из этих записей были отобраны части, в которых Рассказчик рассказывал о содержании фильма Пересказчику в присутствии Комментатора.

Все отобранные пересказы представляют собой монологическую речь продолжительностью от 2 мин. 50 с. до 7 мин. 37 с., объем каждого пересказа лежит в интервале от 279 до 967 слов. Суммарная длительность записей составила 84 мин., суммарный объем — 10 700 слов. С каждым испытуемым был проведен тест по определению объема его РП по описанной выше процедуре; объем РП подсчитывался в процентах и находится в интервале от 49% до 84%.

2.2. Аннотация лексико-грамматических маркеров речевых затруднений

Во всех собранных пересказах была вручную произведена разметка лексико-грамматических сигналов речевых затруднений. Долексические сигналы, такие как заполненные паузы и фонологически не мотивированные удлинения звуков, нами не рассматривались, возможная их связь с объемом РП должна составить в перспективе предмет отдельного исследования. Аннотирование маркеров речевых затруднений производилось на основе классификации, предложенной в работах (Кибрик, Подлеская (ред.) 2009; Подлеская 2013; Подлеская 2014) с некоторыми уточнениями. Нами учитывались следующие явления²:

- а) *Микрокоррекции* — самоисправления говорящего в пределах ЭДЕ, напр., замена слова в (1):

² Индекс примера содержит ссылку к ярлыку пересказа в составе выборки. Примеры из текстов выборки даны в дискурсивной транскрипции, разработанной в (Кибрик, Подлеская (ред.) 2009) и принятой также в коллекции корпусов устной речи (SpokenCorpora 2013). В соответствии с базовыми принципами этой транскрипции тексты были разбиты на элементарные дискурсивные единицы (ЭДЕ) — минимальные кванты дискурса, обладающие грамматической и коммуникативно-просодической целостностью; строка транскрипта соответствует одной ЭДЕ.

(1) #16

сначала показывается || ээ ээ описывается /пейзаж,

б) **Макрокоррекции** — самоисправления говорящего, затрагивающие более одной ЭДЕ. Напр., в (2) строка 2 отменяется, так как она является преждевременной попыткой, удачно реализованной лишь позднее, в строке 4:

(2) #5

ээ он на неё /→засматривается,

/∧и-и ээ значит колесо натывается ==

у него во-п= || сначала сдувает его /шляпу светл=|| светлую,

а потом он колесом натывается на-а /камень,

в) **Гибридная коррекция** — строка, планировавшаяся изначально как единая ЭДЕ, обрывается и достраивается в отдельной ЭДЕ, при этом материал исходной ЭДЕ не отбраковывается полностью. Так, в (3) в строке 2 происходит коррекция сказуемого, но эта строка разделяет со строкой 1 подлежащее и содержит анафорическую отсылку к прямому дополнению *груши*:

(3) #18

/мужчина собирает \груши с де= ==

обрывает их с /деревя,

складывает в большие /→корзины,,,

г) **Маркеры препаративной подстановки** — слова, обычно местоименного происхождения, которые выполняют функцию временного замещения искомого выражения при затрудненном поиске. Типичным для дескриптивных фрагментов дискурса, напр., оказывается использование в этой функции местоименного прилагательного *такой*, проецирующего грамматическую форму отложенного атрибута, ср. женский род, родительный падеж в (4), типичным является и дублирование акцента — на маркере и на отложенном атрибуте:

(4) #16

мальчик \такой || \европейской /внешности,

д) **Маркеры нечеткой номинации**, в частности, так называемые маркеры-аппроксиматоры (см. (Подлесская 2013)), которые сопровождают «пробные» попытки вербализовать некоторый смысл, сигнализируя о том, что предпринятая попытка нуждается в обобщении или уточнении, но адекватного языкового выражения в арсенале говорящего в данный момент не нашлось. Так, в (5) после маркера препаративной подстановки *такая* в строке 1, следуют два маркера нечеткой номинации — *знаешь* во второй строке и *что ли* — в третьей:

(5) #16

ээ перед передним /колесом у него есть ээ такая вот ==
/знаешь;
/→седушка что ли...

е) *Маркеры эмоциональной реакции на речевую проблему* — формирующие автономную реплику междометия, сигнализирующие об эмоциональном состоянии говорящего в связи с обнаружением сбоя. Это может быть, напр., удивление при обнаружении собственной ошибки, как ой в (6), строка 4, или досада от невозможности выбрать адекватную номинацию, как уфф в (7), строка 2. В (6) маркер эмоциональной реакции комбинируется с макрокоррекцией (обрыв строки 2), в (7) — с микрокоррекцией в строке 1 и маркером нечеткой номинации, конструкцией «X не X» в строке 3:

(6) #16

на первом /плане,
ээ яблоневое ==
ой \нет!,
\грушёвое \дерево.

(7) #16

у одного б-была-а || ну такое /приспособление,
уфф!,
как теннис не /теннис,
вот теннисная ... ээ /ракетка небольшая,
и на ней на верёвочке вот теннисный \мячик.

ж) *Оффлайн коррекции, или редактирование* — отложенные эксплицитные самоисправления, не нарушающие ни грамматическую, ни просодическую когерентность текста, постфактум направляющие слушающему инструкцию о том, какого рода ошибка содержалась в предшествующем фрагменте, ср. (8), где редактирование происходит в строках 2–3:

(8) #18

мимо проходит какая-то /→коза-а,,,
ээ точнее проходит мимо \↑мужчина,
с \козой,

Редактирование нередко сочетается с другими сигналами речевых затруднений. Так, в (9) «внахлест» реализуются макрокоррекция и редактирование — в строке 3 эксплицитно редактируется строка 1 и подтверждается правильность смены номинации, произошедшей в строке 2 (правильно *девочка*, а не *девушка*), при этом строка 2 обрывается (макрокоррекция) и реализуется уже как строка 4, после редактирования:

(9) #18

видит как навстречу /-ему едет \девушка на \↑велосипеде,,
ээ ну девочке где-то-о ==
\девочка даже скорее,
девочке лет ну /-тринадцать-\четырнадцать,

з) *Сплит* — разрыв текущей ЭДЕ в связи необходимостью ввести некоторый фрагмент дискурса безотлагательно. Так, в (10) разорвана клауза *он аккуратнo её протирает*, с тем, чтобы внутрь её, не дожидаясь завершения, вставить адвербиальное уточнение *бережно*:

(10) #9

он мм аккуратнo её —
\бережно,
— /п-протирает,
и кладёт в корзину.

Нередко внутри сплита происходит оффлайн коррекция, см. (11), где разрывается клауза *потом проезжает мальчик на велосипеде* в связи с безотлагательной необходимостью заменить *проезжает* на *приезжает*:

(11) #18

/потом проезжает —
\приезжает точнее,
— /мальчик на \велосипеде.

2.3. Анализ и обсуждение результатов

В результате проведенной аннотации мы получили 529 лексико-грамматических сигналов речевых затруднений, следующим образом распределенных по 19 пересказам (см. табл. 1).

Как видно из табл. 1, количество лексико-грамматических сигналов речевых затруднений в пересказах «Фильма о грушах» колеблется в пределах от 0,77 до 8,58 на 100 слов (среднее 4,57) и от 0,71 до 14,46 в минуту (среднее 6,02), что в целом соответствует средним значениям, характерным для спонтанной речи. Согласно данным, представленным в работе (Подлесская 2014), частотность самоисправлений обычно фиксируется в интервале от 1 до 7 самоисправлений в минуту; в диалогах она обычно выше, чем в монологической речи, а в бытовой речи — выше, чем в официальной. Так, для китайских диалогов приводятся сведения о 5,4 случаях самоисправлений в минуту (Tseng 2006); для английских устных пересказов — 1,9–3,7 на 100 слов (Fraundorf, Watson 2008); для венгерских диалогов — 3,8 в минуту (Németh 2012); для японских монологов — 1,2 на 100 слов (Maruyama, Sano 2006); для спонтанных диалогов в иорданском варианте арабского языка — 1,6 на 100 слов (Al-Harashsheh 2015).

Таблица 1. Основные количественные характеристики аннотированных пересказов

#	Объем РП	Кол-во затруднений	Длительность (в секундах)	Объем (в словах)	Кол-во затруднений (на 100 слов)	Кол-во затруднений (в минуту)
1	84	34	322	709	4,80	6,33
2	82	51	385	956	5,33	7,94
3	82	34	336	733	4,64	6,07
4	76	7	177	361	1,94	2,37
5	75	14	174	372	3,76	4,83
6	74	2	170	279	0,77	0,71
7	71	10	345	584	1,71	1,74
8	70	11	249	351	3,13	2,65
9	70	28	216	481	5,82	7,78
10	64	42	221	745	5,64	11,41
11	64	12	214	419	2,86	3,36
12	63	13	205	347	3,75	3,80
13	62	17	183	294	5,78	5,57
14	59	72	422	967	7,45	10,24
15	58	18	224	333	5,41	4,83
16	56	24	221	449	5,36	6,52
17	52	40	457	963	4,16	5,26
18	49	60	249	699	8,58	14,46
19	48	40	285	664	6,02	8,42

Проведя статистический анализ, мы нашли значимую корреляцию между объемом РП и отношением числа сигналов речевых затруднений к объему пересказа (в словах): $\text{cor} = -0,483^*$, $p\text{-value} < 0,05$, см. рис. 1.

Таким образом, мы подтвердили гипотезу о том, что вербальный объем РП является фактором, коррелирующим с количеством лексико-грамматических сигналов речевых затруднений, т. е. с плавностью речепорождения: чем больше у испытуемого объем РП, тем меньше подобных маркеров обнаруживается в его спонтанной монологической речи.

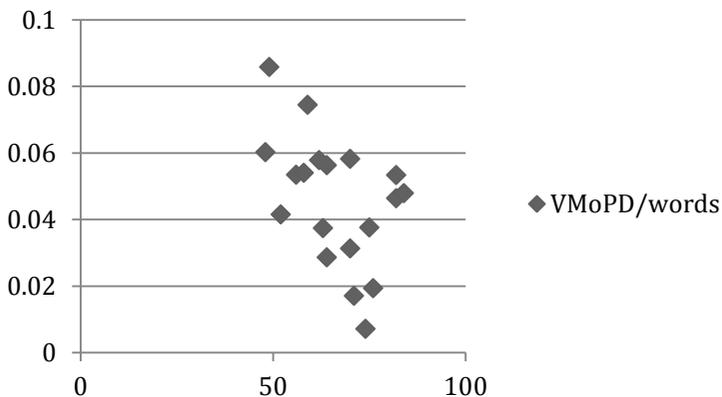


Рис. 1. Точечная диаграмма для объема РП (ось X) и отношения числа сигналов речевых затруднений к объему рассказа (ось Y)

2.4. Перспективы дальнейших исследований

В настоящей работе мы показали, что объем РП коррелирует с индивидуальными различиями при речепорождении. Более конкретно, впервые на материале русского языка была продемонстрирована значимая корреляция между вербальным объемом РП и числом лексико-грамматических сигналов речевых затруднений при порождении пересказа. Данное исследование является пилотным в области анализа взаимосвязи объема РП и маркеров вербальных затруднений, оно проведено на сравнительно небольшом объеме материала и оставляет множество вопросов для дальнейшего изучения. Опишем некоторые из них более подробно.

Во-первых, нуждается в проверке гипотеза о наличии корреляции между объемом РП, полученным в тесте «Speaking span», и количеством речевых затруднений, которые возникали у испытуемых в ходе прохождения самого этого теста. Процедура прохождения испытуемыми теста «Speaking span» была записана на диктофон, так что расшифровка и анализ лексико-грамматических сигналов речевых затруднений даст ответ на вопрос о наличии корреляций как между объемом РП и количеством речевых затруднений при прохождении теста, так и между речевой продукцией испытуемых в ходе прохождения теста и в ходе пересказа.

Во-вторых, в настоящей работе была рассмотрена взаимосвязь между объемом РП и лексико-грамматическими сигналами речевых затруднений. Как было упомянуто выше, другой класс речевых затруднений включает различные **долексические маркеры**, в первую очередь заполненные паузы и фонологически не мотивированные удлинения звуков. Подобные маркеры более частотны в спонтанной речи, но их связь с объемом РП не так очевидна, так как интуитивно они менее когнитивно трудозатратны. Данная гипотеза также может быть проверена на рассмотренном материале 19 пересказов.

В-третьих, представляет интерес вопрос о взаимозависимости объема РП и количества сигналов вербальных затруднений в ситуации дефицита времени, т. е. в таких тестах, в которых испытуемым нужно уместить как можно больше информации в ограниченный по длительности пересказ. Эти и многие другие вопросы требуют как дополнительных исследований вербальной РП, так и более детального анализа лексико-грамматических сигналов речевых затруднений.

Наконец, обнаружение корреляции между объемом РП и вербальными маркерами речевых затруднений может означать, что (а) низкий объем рабочей памяти увеличивает затруднения при порождении, (б) затруднения снижают объем РП, (в) другой фактор или факторы влияют на рабочую память и речевые затруднения³. Чтобы определить, каким образом объем РП связан с речевыми затруднениями, необходимо более детальное исследование механизма синтеза речи и роли рабочей памяти в этом процессе. Кроме того, необходимо более тщательное изучение различных классов речевых затруднений и анализ вклада каждого класса в корреляцию с объемом РП. На наших данных не удалось продемонстрировать значимых корреляций между объемом РП и определенными типами затруднений. Мы предполагаем, что для данного исследования понадобится больший объем корпуса рассказов, в котором будут представлены все известные типы затруднений.

Литература

1. *Al-Harahsheh A. M. A.* (2015), A Conversation Analysis of self-initiated repair structures in Jordanian Spoken Arabic. *Discourse Studies* 17(4), 397–414
2. *Atkinson R. C., Shiffrin R. M.* (1968), Human memory: A proposed system and its control processes, K. W. Spence, J. T. Spence (eds.) *The psychology of learning and motivation: Advances in research and theory*, New York.
3. *Baddeley A. D.* (2000), The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences* 4(11), 417–23.
4. *Chafe W.* (ed.) (1980), *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production*, Norwood: Ablex.
5. *Daneman M.* (1994), Working memory and language, *Language and Speech*, 37.
6. *Daneman M.* (1991), Working memory as a predictor of verbal fluency. *Journal of Psycholinguistic Research*, 20(6), 445–464.
7. *Daneman M., Carpenter P. A.* (1980), Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior*, 19(4), 450–466.
8. *Daneman M., Green I.* (1986), Individual differences in comprehending and producing words in context. *Journal of memory and language*, 25(1), 1–18.

³ Мы выражаем искреннюю признательность анонимному рецензенту, справедливо указавшему на то, что сама по себе корреляция между объемом РП и частотой речевых сбоев не объясняет механизма связи между этими феноменами и нуждается в интерпретации в рамках той или иной модели речепорождения. Эта работа нам еще предстоит.

9. *Fedorova O. V.* (2003), Survey of the State of the Art in Verbal Span Test [Test po opredeleniju ob"ëma operativnoj pamjati: Istorija i sovremennoe sostojanie], Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialog 2003" [Komp'uternaia Lingvistika I Intellektual'nye Tehnologii: Trudy Mezhdunarodnoj Konferentsii "Dialog 2003"], Moscow.
10. *Fedorova O. V., Potanina Ju. D.* (2014), Working memory and Russian language: from comprehension to production [Rabochaya pamyat' I russkiyazyk: ot recheponimaniya k recheporozhdeniyu], Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialogue 2014" [Komp'yuternaya Lingvistika i Intellektual'nye Tekhnologii: Trudy Mezhdunarodnoy Konferentsii "Dialog 2014"], Bekasovo.
11. *Fraundorf S. H., Watson D. G.* (2008), Dimensions of variation in disfluency production in discourse, in J.
12. *James W.* (1890). The principles of psychology (Vol. 1). New York: Holt.
13. *Kibrik A. A., Podlesskaya V. I.* [Eds.] (2009), Night Dream Stories: A corpus study of spoken Russian discourse [Rasskazy o snovidenijax: korpusnoe issledovanie usnogo russkogo diskursa]. Moscow: Jazyki Slavjanskix Kul'tur.
14. *Maruyama T. and Sano Sh.* (2006), Classification and Annotation of Self-Repairs in Japanese Spontaneous Monologues, in LPSS — Linguistic Patterns in Spontaneous Speech, Taipei, 283–298.
15. *Miller G. A.* (1956), The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, 63(2), 81.
16. *Miller G. A., Galanter E., Pribram K. H.* (1960), Plans and the structure of behavior, New York.
17. *Németh Z.* (2012), Recycling and replacement self-repairs in spontaneous Hungarian conversations, in Proceedings of the First Central European Conference in Linguistics for postgraduate Students, 211–224.
18. *Podlesskaya V. I.* (2014), They shot him dead, oh, no, they knifed him dead with a saber: self-repairs in oral stories [To est', ne ubili, a zarezali sablej: samoispravlenija govorjashhego v ustnyh rasskazah]. Computational Linguistics and Intellectual Technologies: papers from the Annual International Conference "Dialog". V. 13 (20). Moscow: RSUH, 2014, 526–540.
19. *Podlesskaya V. I.* (2013), Vague reference in Russian: evidence from spoken corpora [Nechetkaja nominacija v russkoj razgovornoj rechi: opyt korpusnogo issledovanija]. Computational Linguistics and Intellectual Technologies: papers from the Annual International Conference "Dialog". V. 12 (19). Moscow: RSUH, 2013, 561–573.
20. *Spoken corpora* (2013), Night dream stories and other spoken corpora [Rasskazy o snovidenijah i drugie korpusa zvuchashhej rechi]. Prosodically Annotated Corpus of Spoken Russian. Pilot version. Online: <http://spoken corpora.ru>
21. *Tseng S.-C.* (2006), Repairs in Mandarin conversation, *Journal of Chinese Linguistics* 34(1), 80–120.