

**КОГНИТИВНАЯ СЛОЖНОСТЬ ХУДОЖЕСТВЕННЫХ ТЕКСТОВ  
ДЛЯ ДЕТЕЙ: КВАНТИТАТИВНЫЕ МЕТОДЫ ОЦЕНКИ<sup>1</sup>**

**Валуева Екатерина Александровна**

кандидат психологических наук  
научный сотрудник, Институт психологии РАН  
ведущий научный сотрудник,  
Московский городской психолого-педагогический университет  
*ekval@list.ru*

**Данилевская Нина Михайловна**

филолог, независимый исследователь  
*ninadan@mail.ru*

**Лаптева Екатерина Михайловна**

кандидат психологических наук  
ведущий научный сотрудник,  
Федеральный институт развития образования  
*ek.lapteva@gmail.com*

**Ушаков Дмитрий Викторович**

член-корреспондент РАН, доктор психологических наук, профессор  
заведующий лабораторией, Институт психологии РАН  
*dv.usakov@gmail.com*

В статье излагаются результаты исследования языковых маркеров простоты/сложности художественных текстов на материале сравнения детской и подростковой художественной литературы, издававшейся с 1900 по 2016 гг. Для проведения анализа были подобраны тексты произведений, наиболее популярных в соответствующий временной период. Произведения были подобраны по категориям «русская литература», «европейская литература», «американско-канадская литература». Все тексты были разбиты на 4 возрастные категории: «дошкольный возраст», «младший школьный возраст», «средний школьный возраст», «старший школьный возраст». В анализ вошли 588 книг на русском языке и 118 текстов на английском языке. В каждом тексте подсчитывались значения ряда количественных показателей, измеряющих объем текста, а также ряд показателей, характеризующих структуру и содержание текстов. Результаты статистического анализа показали, что существуют компоненты художественных текстов, которые можно расценивать как маркеры усложнения когнитивной организации. Выявленные параметры могут составить основу инструментария для дальнейших сравнительных операций с текстами, связанных с исследованием речевого развития человека в процессе чтения и сопряженных с этим эффектов.

**Ключевые слова:** детская художественная литература, XX в., когнитивная сложность.

---

<sup>1</sup> Исследование поддержано грантом президента РФ № МК-7584.2015.6

### Введение

Понимание текстов, в том числе – художественных, предполагает разворачивание когнитивной деятельности, которая тем сложнее, чем более насыщенным и трудным для понимания является текст [Dehaene 2007]. Каждый образованный человек, имевший в жизни дело с большим количеством текстов, по опыту знает, что понимать некоторые из них намного труднее, чем другие: например, текст из «Мурзилки» понять заведомо проще, чем фрагмент «Поминок по Финнегану» Дж. Джойса или «Бытия и времени» М. Хайдеггера. Однако вопрос о причинах этой сложности и ее формальных признаках остается весьма непростым.

Очевидно, что источником сложности могут быть как выражаемые мысли, так и форма выражения. Авторы, как правило, стремятся выразить свои мысли по возможности доступно, но это не всегда удается в равной мере.

Согласно влиятельной и достаточно хорошо эмпирически фундированной теории Ф. Джонсон-Лэрда, процесс понимания начинается и заканчивается невербальной репрезентацией, или умственной моделью. Однако для передачи другому человеку эта репрезентация должна быть преобразована говорящим в вербальную, или пропозициональную, форму. Слушающий должен произвести обратное преобразование пропозициональной репрезентации в умственную модель. Поскольку коммуниканты обладают индивидуальными особенностями, их вербальный интеллект и языковая компетентность неодинаковы. На стороне говорящего это приводит к тому, что умственная модель отражается в тексте с большей или меньшей степенью адекватности, а сам текст обладает большей или меньшей ясностью, то есть требует от слушателя или читателя различных усилий для восстановления умственной модели. На стороне слушателя получается, что один и тот же текст приводит к созданию умственных моделей, с различной степенью адекватности отражающих смыслы текста.

Согласно модели Т.Н. Ушаковой, понимание текста предполагает соответствие между когнитивными системами говорящего и слушающего (или пишущего и читающего) на нескольких уровнях: на уровнях вербальных сетей, т.е. ассоциативных связей между элементами, грамматических структур и контекста [Ушакова 1998; Ушакова, Белова, Валуева 2010]. Тем самым в когнитивную проблематику вводится более широкий контекст понимания – к реконструкции понимающим умственной модели, лежащей в основе текста, добавляются ассоциативные коннотации и контекстуальный смысл. Действительно, если трудность понимания, например, научного текста может быть в большой мере сведена к сложности умственной модели, то в отношении поэтического текста должно приниматься во внимание ассоциативное богатство [Ушакова 1989, 1990].

К анализу причин сложности текстов, определяемых умственными моделями, намечены продуктивные подходы. Начиная с Ж. Пиаже, психология онтогенетического развития сумела показать последовательность появления у ребенка различных возможностей структурирования репрезентаций и сопутствующих им речевых форм. Так, понимание причинно-следственных отношений у ребенка, согласно Ж. Пиаже, возникает примерно в 7-8 лет. С этого же возраста формируется адекватное употребление причинных союзов. Как показал Ж. Пиаже, дифференцированное

употребление союзов «потому что» и «поэтому» происходит вместе с формированием понимания причинности.

Аналогично в пиажеанской психолингвистике (Э. Синклер де Зварт) установлено, что сравнительные формы (больше-меньше) начинают употребляться ребенком с того момента, как у него формируется возможность репрезентации асимметричных транзитивных отношений ( $A < B < C \dots$ ) и решения основанных на таких репрезентациях задач, например, сериации. Таким образом, сложность структурирования умственной модели оказывается важным фактором трудности понимания текста [Ушаков 1997, 2009].

Для разработки обсуждаемой проблематики существенным оказывается вопрос о языковых маркерах когнитивной сложности текста. С одной стороны, этот вопрос позволяет пролить свет на источники сложности текстов, связанные со структурированием репрезентаций. Из сказанного выше следует, что наличие в тексте определенных маркеров, например, слов «потому что» или «больше», указывает на такие характеристики репрезентаций, которые связаны с определенной сложностью когнитивных процессов, необходимых для их обработки. Соответственно исследование таких маркеров в связи со сложностью текстов позволяет проверять гипотезы относительно когнитивных процессов, происходящих при понимании текстов, а также создаваемой ими когнитивной нагрузки и их онтогенетического развития.

С другой стороны, это исследование имеет важный методический аспект. Выявленные параметры могут составить основу инструментария для дальнейших сравнительных операций с текстами, связанных с исследованием речевого развития человека в процессе чтения и сопряженных с этим эффектов. Например, можно проследить глобальные тенденции упрощения или усложнения текстов художественной литературы на протяжении отрезков времени либо сравнить соответствующие средние показатели по разным странам. Это могло бы выявить интересные корреляции и закономерности.

Понятно, что сложность текста растет по мере увеличения возрастной категории читателей, которой адресован текст. Сравнительный психолингвистический анализ текстов, предназначенных для разных возрастных категорий, может выявить конкретные параметры простоты/сложности их восприятия. В связи со сказанным нами было проведено исследование языковых маркеров простоты/сложности художественных текстов на материале сравнения детской и подростковой художественной литературы, издававшейся с 1900 по 2016 гг.

### Материалы

Для выполнения поставленных задач мы подобрали детские и подростковые художественные произведения, издававшиеся в период с 1900 по 2016 гг. Произведения были подобраны по категориям «русская литература», «европейская литература», «американско-канадская литература». Все проанализированные зарубежные произведения издавались также на русском языке. Дополнительно к категории «американско-канадская литература» были найдены оригиналы текстов на английском языке для сопоставления результатов анализа и оценки адекватности использования переводных источников. В каждой из этих трех категорий исходя

из авторских и издательских рекомендаций тексты делились на возрастные диапазоны: «дошкольный возраст», «младший школьный возраст», «средний школьный возраст», «старший школьный возраст». Для проведения анализа были подобраны тексты произведений, наиболее популярных в соответствующий временной период. Главный критерий, по которому отобраны книги, – их популярность на родине, выражаемая либо в количестве переизданий, либо в завоевании престижных премий (Медаль Джона Ньюбери и Медаль Кальдекотта в США; Премия имени Ханса Кристиана Андерсена, Международная литературная премия Астрид Линдгрэн; Сталинская премия, Государственная Премия СССР и др.). При выборе русских произведений учитывались также отзывы известных деятелей культуры (например, Марина Цветаева о Лидии Чарской) и современные читательские интернет-рейтинги.

Всего в анализ вошли 588 книг на русском языке и 118 текстов на английском языке. В таблице 1 представлено общее количество произведений в соответствии с категориями.

Таблица 1

Количество анализируемых произведений по периодам издания, возрастам и регионам на русском/английском языках

Возраст	Регион	Период издания				Всего
		1900-1930	1931-1969	1970-2000	2001-2016	
Дошкольники	Америка, Канада	0/0	3/3	3/3	1/0	7
	Европа, Англия	3	4	2	9	18
	Россия	4	9	6	3	22
Младшая школа	Америка, Канада	8/8	15/12	9/6	25/0	57
	Европа, Англия	20	22	20	14	76
	Россия	10	8	10	16	44
Средняя школа	Америка, Канада	19/19	16/15	14/13	19/0	68
	Европа, Англия	13	13	16	18	60
	Россия	16	17	16	14	63
Старшая школа	Америка, Канада	13/13	14/14	15/15	16/0	58
	Европа, Англия	19	12	14	16	61
	Россия	14	15	13	12	54
<b>Всего</b>		<b>139</b>	<b>148</b>	<b>138</b>	<b>163</b>	<b>588</b>

Нам необходимо было выявить те характеристики текста, которые имеют возрастную динамику, т.е. связаны с увеличением умственного возраста читателя. Мы выделили широкий спектр показателей. Анализ этих показателей позволил нам отобрать те из них, которые значимо изменяются при переходе от одной возрастной группы к другой.

Для каждой книги были рассчитаны 2 типа показателей: абсолютные и относительные. К абсолютным показателям относятся количественные показатели, которые характеризуют объем текстов: 1) общее количество знаков в тексте; 2) общее количество слов в тексте; 3) количество предложений в тексте; 4) количество запятых в тексте. К относительным показателям мы отнесли, во-первых, формальные характеристики, такие как: 5) средняя длина слова, 6) средняя длина предложения (медиана количества слов в предложении), 7) количество запятых в тексте (относительно общего количества слов в тексте); во-вторых, содержательные характеристики, которые, по нашему мнению, связаны с усложнением когнитивной организации целевой аудитории. В последнюю группу были отнесены морфологические и синтаксические компоненты.

### *Морфологические компоненты:*

8) количество вводных слов и словосочетаний (относительно общего количества слов в тексте). Примерами вводных слов являются такие слова, как «однако», «в общем», «видимо» и т.д. (полный список анализируемых вводных слов см. в Приложении). Вводные слова имеют множество особенностей и функций, усложняющих структуру высказывания. В первую очередь, они связаны не с содержанием предложения, а с миром говорящего («*Виджу*, ты боишься») или адресата высказывания («*Видишь ли*, ты боишься»). Они могут выражать отношение автора к высказыванию («*К счастью*, дождь кончился») или неуверенность в достоверности содержащейся в нем информации («*Вероятно*, дождь кончился»). Вводные слова могут также упорядочивать и иерархизировать большой поток информации («*во-первых*», «*во-вторых*», «*прежде всего*», «*главное*»), организовывать сложные (причинно-следственные, уступительные, противительные и др.) связи между высказываниями или их частями («*Ты, стало быть*, боишься»; «*Дождь, правда*, не кончился»; «*К вечеру, впрочем*, должно проясниться»). Кроме того, вводные слова и словосочетания придают смысловые оттенки суждениям, зачастую повышая неоднозначность и неопределенность высказываемых мыслей (ср., например: «Он профессионал своего дела» и «Он, так сказать, профессионал своего дела»: вторая фраза содержит имплицитный иронический оттенок, уловить который способен только человек, находящийся на достаточно высокой ступени речевого развития);

9) количество компаративов, то есть прилагательных и наречий в сравнительной степени (относительно общего количества слов в тексте). Опираясь на частотные словари русского языка [Частотный словарь 1977; Ляшевская, Шаров 2009], мы взяли наиболее распространенные прилагательные и наречия и поставили их в степень сравнения («дальше», «ближе», «веселее», «проще» и т.д. (см. Приложение)). Слова, выражающие степень сравнения, вносят дополнительное логическое звено в характеристику объектов действительности, требуют осознания непростых взаимоотношений между ними, поэтому могут служить маркерами

усложнения когнитивной организации. Кроме того, компаративы прилагательных и наречий имеют идентичную форму; адекватная интерпретация этой омонимии требует от читателя дополнительной умственной нагрузки и работы с контекстом (например, фраза «Она лучше» может соответствовать как высказыванию «Она более хорошая», так и к эллиптическому высказыванию «Она (что-то делает) более хорошо»; фраза «Это хуже» может относиться как к характеристике существительного среднего рода – «Это платье хуже», так и к оценке действия или ситуации – «Так хуже»);

10) количество простых производных предлогов (относительно общего количества слов в тексте). Согласно нашей гипотезе, одним из проявлений сложности текста должна была стать его большая детализация и визуализированность по сравнению с элементарными субъектно-предикатными текстами, адресованными дошкольникам и детям младшего школьного возраста (ср.: «Мама мыла раму»). Грамматически за эти построения в значительной степени отвечают предлоги. Во избежание амбивалентных ситуаций (типа «в течение дня» / «в течение реки»), не поддающихся статистическому анализу в крупных объемах, мы ограничились простыми производными предлогами, такими как: «в», «за», «на», «под», «перед» и т.д. (см. Приложение);

11) количество междометий и 13) ономотопеической лексики [Тимошук 1994] (относительно общего количества слов в тексте). Логично предположить, что такие лексические средства, как междометие («ах», «ох», «ну») и ономотопея («мычать», «квакать», «журчать» и т.д. (см. в Приложении)), передающие информацию не денотативно, а непосредственно через звукоподражание, должны преобладать в более «детских» текстах (еще в «Риторике для Геренния» (I в.н.э.), иногда приписываемой Цицерону [Rhetoricon ad Herennium libri IV 1890-1907], говорилось о примитивном характере «этого вида украшений»). Как известно, вообще в сравнении со взрослыми детские книги в большей степени апеллируют к эмоциям читателей, нежели к рассудку [Якобсон 1998]. В частности, представляется очевидным, что с точки зрения познания мира и развития личности ситуация, описанная фразой: «Ой, доска кончается...» для ребенка 3-х лет будет намного более эффективной, чем нейтральная констатация: «Бычок испугался края доски». Таким образом, эмотивная и ономотопеическая лексика в нашей концепции, является механизмом, «упрощающим», а не «усложняющим» текст;

12) количество модальных частиц (относительно общего количества слов в тексте). Модальные частицы вносят в предложение эмоционально-оценочные, вопросительные, уступительные, персуазивные (со значением достоверности-недостоверности) и другие семантические оттенки. Но, в отличие от междометий с пересекающимися функциями, делают это более «по-взрослому» – реализуя стилистическое смещение от разговорного языка к художественному. Давно замечено, что уменьшение возрастной категории книги коррелирует с уменьшением зазора между повседневно-разговорным языком, составляющим коммуникативную практику ребенка, и языком текста [Бибик 2013]. На наш взгляд, на пути «овзроствления» текста частицы являются промежуточным звеном между акустической денотацией (междометие, звукоподражание) и дескриптивным способом передачи соответствующих смыслов. Например, цепочка денотативного усложнения в ситуации

удивления могла бы выглядеть следующим образом: «Ого!» (междометие) – «Подумать только!» (использование частицы) – «Ты меня удивляешь!» (местоименно-глагольная дескрипция).

Однако, сохраняя стилизацию устно-разговорной стихии, частицы могут отражать сложную взаимосвязь контекстов. Например, частица «все-таки» подразумевает, что в предшествующем текстовом пространстве содержались условия, частично или полностью противоречащие высказыванию (Ср.: в предложении: «Мы все-таки пойдем гулять» заложена информация о каких-то негативных сопутствующих обстоятельствах: плохой погоде, альтернативных планах, запрете и т.п.).

Таким образом, в целом, согласно нашим предсказаниям, частица должна быть маркером когнитивного усложнения.

*Синтаксические компоненты:*

14) количество сложноподчиненных предложений (относительно общего количества предложений в тексте). Эти синтаксические конструкции характерны для предложений, описывающих сложную картину взаимосвязей в окружающем мире. Зачастую эти взаимосвязи трудно различимы, и для того, чтобы воспринять их согласно авторскому замыслу, требуется соответствующий жизненный опыт. Например: «Он пришел пораньше, дабы...» (цель) и «Он пришел пораньше, ибо...» (причина); «Я грущу, если ты веселишься» (условие) и «Я грущу, хотя ты веселишься» (уступка). Количество сложноподчиненных предложений оценивалось по количеству в тексте подчинительных союзов, таких как «где», «дабы», «если», «который» и т.д. (см. «Приложение»).

Во избежание «шума» в данных мы постарались минимизировать количество омонимичных в русском языке синтагм, наподобие «лишь», «лишь только», способных выступать как в качестве подчинительного союза в придаточном предложении времени («Лишь только вечер затеплится синий..., отвори потихоньку калитку»), так и в качестве выделительно-ограничительной частицы («Лишь только с вами я хотел бы жизнь проведать...»).

15) количество вопросительных и 16) восклицательных предложений (относительно общего количества предложений в тексте). На наш взгляд, концентрация восклицательных предложений (исходя из предшествующих рассуждений) должна падать по мере «взросления» текста. Концентрация же вопросительных предложений, судя по всему, должна наоборот расти, придавая повествованию полемическую остроту. Причем это может происходить как в рамках диалоговой/полилоговой организации текста, так и в рамках вопросно-ответной формы авторских рассуждений (в качестве примера можно привести особую концентрацию вопросительных предложений в романе И.С. Тургенева «Отцы и дети», сложная социальная и экзистенциальная проблематика которого практически целиком вычитывается из разговоров персонажей и риторических вопросов автора: «Неужели их молитвы, их слезы бесплодны? Неужели любовь, святая, преданная любовь не всемогуща?» и т.п. [Мартыанова 1981; Пустовойт 1981]).

### **Обработка текстов**

Вся обработка производилась в автоматическом режиме с использованием программы R. Перед проведением обработки все тексты были переведены в формат txt. Все слова книги переводились в нижний регистр.

Предложением считалась часть текста, которая заканчивается точкой, вопросительным или восклицательным знаком. В предложениях подсчитывалось количество слов как отдельных объектов, состоящих из одной или более букв алфавита. Количество слов в тексте является суммой количества слов в предложениях. Средняя длина слова рассчитывалась по формуле: *(количество знаков в книге – количество знаков пунктуации)/количество пробелов*.

Количество слов, относящихся к морфологическим показателям (частицы, союзы, компаративы и др.) производилось суммированием всех случаев совпадения слов книги со словами из этой подгруппы. Количество подчинительных союзов было скорректировано на количество вопросительных слов, которые совпадают с ними по форме. Для этого из текста книги выделялись все вопросительные предложения и по ним подсчитывалось общее количество слов, совпадающих в подгруппах «вопросительные слова» и «подчинительные союзы». Полученное число было вычтено из показателя, рассчитанного для всего текста по подгруппе «подчинительные союзы».

### Результаты

В первую очередь были сопоставлены русские переводы и оригинальные англоязычные тексты для американско-канадской литературы. Корреляции всех показателей<sup>2</sup> оказались высокими и значимыми – коэффициенты корреляции варьировали от 0.49 до 0.96 (см. таблицу 2). Наиболее низкая корреляция (0.49) была получена для показателя «относительное количество запятых», что, по всей видимости, связано с различиями в правилах пунктуации в русском и английском языках.

Таблица 2

Корреляции показателей между английскими оригиналами и русскими переводами (все корреляции значимы на уровне  $p < 0.001$ )

Количество знаков	0.96
Количество слов	0.96
Количество предложений	0.96
Количество запятых	0.92
Средняя длина слова	0.70
Средняя длина предложения	0.86
Относительное количество запятых	0.49

На втором этапе анализа для каждого выделенного показателя была оценена его возрастная динамика. В таблице 3 приведены средние значения соответствующих показателей для каждой возрастной группы, а также результаты однофакторного дисперсионного анализа. Для показателя «Ономатопеическая лексика», имеющего сильные выбросы, предварительно была произведена процедура упорядоченной ранговой трансформации [ Wobbrock, Findlater, Gergle, & Higgins 2011].

Так как четыре количественных показателя (общее количество знаков в тексте, количество слов в тексте, количество предложений в тексте, количество

<sup>2</sup> Для англоязычных текстов не рассчитывались показатели, связанные с содержательным анализом слов, т.к. привлечение соответствующих специалистов на данном этапе исследования не представлялось возможным.



запятых в тексте) очень высоко коррелировали между собой (коэффициенты корреляции от 0.93 до 0.99), они были объединены в общую шкалу «объем текста». Значения шкалы «объем текста» рассчитывались как среднее z-оценок входящих в нее показателей.

Из таблицы 3 следует, что все показатели, кроме количества вводных слов и компаративов, демонстрируют значимую возрастную динамику. Количество междометий, звукоподражательных глаголов и восклицательных предложений – падает. Значения остальных показателей растут по мере увеличения возраста читателей.

На рисунках 1-3 изображена возрастная динамика стандартизованных показателей (стандартизация показателей была проведена с целью унификации единиц измерения для более наглядного представления результатов).

Несмотря на значимость фактора возраста в целом, апостериорные сравнения (Games-Howell post-hoc test) показывают более детальную динамику показателей с возрастом. Так, например, для фактора «Объем текста» и входящих в него показателей, а также для показателей «Относительное количество запятых» и «Количество восклицательных предложений» различия значимы между всеми возрастными группами (все  $p < 0.001$ ). Для показателя «Средняя длина предложения» – различия значимы между дошкольниками и всеми остальными возрастными группами (все  $p < 0.01$ ). Для показателя «Непроизводные предлоги» – между всеми возрастными группами (все  $p < 0.01$ ), кроме пар дошкольники – младшая школа и средняя школа – старшая школа. Показатель «Междометия» значимо ниже в старшей школе, чем во всех остальных возрастных группах ( $p$  от 0.001 до 0.065), в средней школе по сравнению с младшей ( $p = 0.09$ ). По показателю «Ономатопеическая лексика» различия значимы между старшими школьниками и всеми остальными возрастными группами (все  $p < 0.01$ ). Для показателя «Частицы» различия не значимы лишь между младшей и средней школой (все остальные различия значимы на уровне  $p < 0.01$ ). Для показателя «Количество сложноподчиненных предложений» – между дошкольниками и всеми остальными возрастными группами (все  $p < 0.001$ ) и между младшими и старшими школьниками ( $p = 0.03$ ). По показателю «Количество вопросительных предложений» на уровне тенденции различаются средняя школа и старшая школа ( $p = 0.06$ ); значимые различия наблюдаются между дошкольниками и старшими школьниками ( $p = 0.02$ ).

Анализ отдельно по регионам («Америка, Канада», «Европа, Англия», «Россия») показал, что основные паттерны результатов остаются теми же независимо от принадлежности к региональной группе. Можно лишь отметить, что в некоторых группах выявленные общие тенденции не достигают принятого уровня значимости. Так, например, в группах «Европа, Англия» и «Россия» незначимыми оказываются рост количества вопросительных предложений и средней длины предложения. Также в русской литературе незначимой оказалась возрастная динамика междометий.

Таблица 3

Возрастная динамика показателей

	Показатель	Возрастная группа				F	p
		Дошколь-ники	Млад-шая школа	Сред-няя школа	Стар-шая школа		

1	Кол-во знаков	90791	209415	361214	552822	<b>45.49</b>	<b>0.000</b>
2	Кол-во слов	15119	34416	58679	89106	<b>44.60</b>	<b>0.000</b>
3	Кол-во предложений	1601	3624	5938	8614	<b>38.22</b>	<b>0.000</b>
4	Кол-во запятых	1549	3605	6505	10149	<b>44.95</b>	<b>0.000</b>
	Объем текста	-0.73	-0.39	0.03	0.56	<b>44.34</b>	<b>0.000</b>
5	Средняя длина слова	4.71	4.85	4.93	5.01	<b>28.72</b>	<b>0.000</b>
6	Средняя длина предложения	7.10	8.04	8.40	8.50	<b>5.64</b>	<b>0.001</b>
7	Относительное кол-во запятых	0.094	0.104	0.109	0.114	<b>28.46</b>	<b>0.000</b>
8	Вводные слова	0.0046	0.0049	0.0048	0.0049	0.34	0.797
9	Компаративы	0.0042	0.0043	0.0042	0.0042	0.59	0.619
10	Непроизводные предлоги	0.083	0.088	0.091	0.092	<b>15.85</b>	<b>0.000</b>
11	Междометия	0.0026	0.0025	0.0021	0.0018	<b>8.30</b>	<b>0.000</b>
12	Ономатопеическая лексика <sup>3</sup>	0.00027	0.00017	0.00013	0.00009	<b>9.60</b>	<b>0.000</b>
13	Частицы	0.0026	0.0038	0.0039	0.0045	<b>15.04</b>	<b>0.000</b>
14	Сложноподчиненные предложения	0.221	0.307	0.317	0.345	<b>13.44</b>	<b>0.000</b>
15	Вопросительные предложения	0.0794	0.0920	0.0906	0.0986	<b>5.02</b>	<b>0.002</b>
16	Восклицательные предложения	0.173	0.133	0.098	0.075	<b>37.05</b>	<b>0.000</b>

<sup>3</sup> Для показателя «Ономатопеическая лексика» приведены значения медиан, т.к. в этом случае перед анализом проводилась ранговая трансформация.

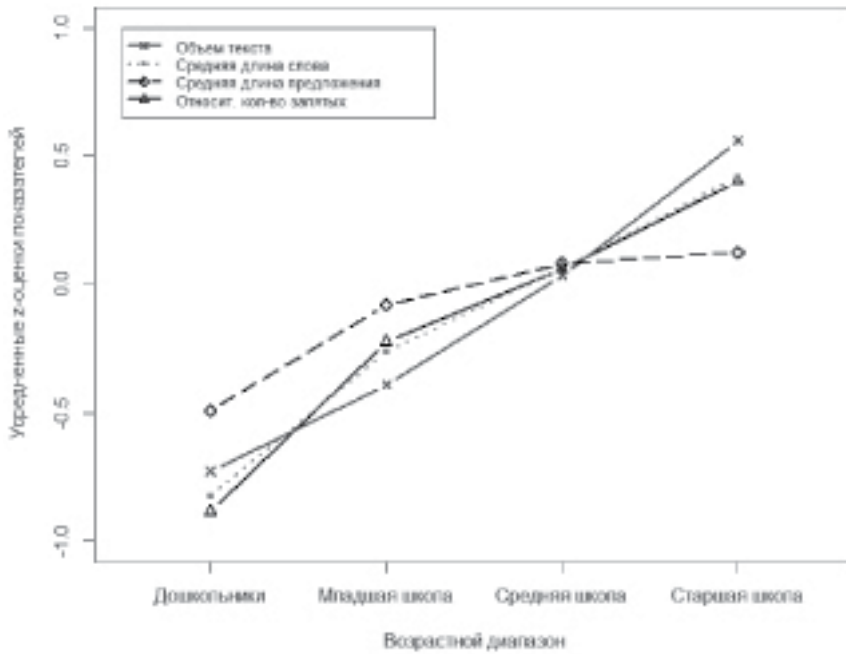


Рисунок 1. Возрастная динамика абсолютных и относительных формальных показателей

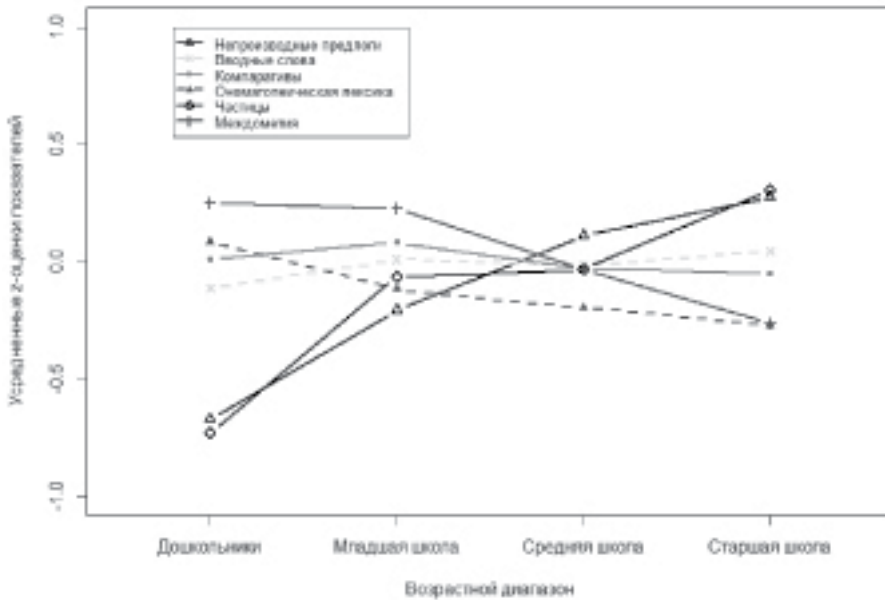


Рисунок 2. Возрастная динамика морфологических показателей

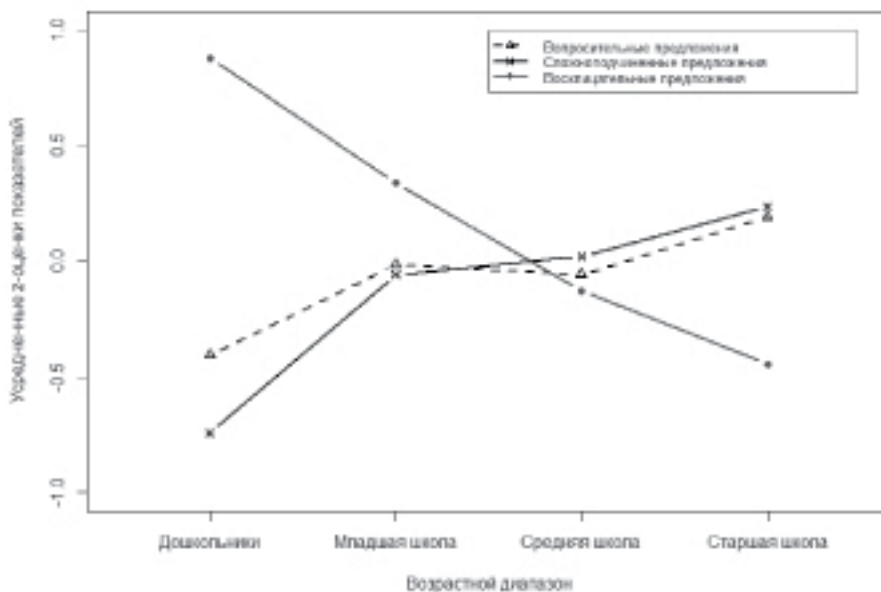


Рисунок 3. Возрастная динамика синтаксических показателей

### Обсуждение результатов

Результаты статистического анализа показали, что существуют компоненты художественных текстов, которые можно расценивать как маркеры усложнения когнитивной организации. Во многом подтвердив ожидания, эти результаты, однако, поставили перед нами и ряд вопросов.

#### Абсолютные показатели

Совокупный рост всех абсолютных показателей по мере «взросления» текста демонстрирует очевидное: чем старше читатель, тем больший объем информации он способен удержать в голове; тогда как младшая читательская аудитория предпочитает получать информацию в раздробленном виде (сборник сказок/рассказов, а не, допустим, равные по общему объему повесть/роман).

#### Относительные показатели

##### Относительные формальные показатели

Усложненная структура (длина) предложения и, соответственно, концентрация запятых являются очевидными показателями «взрослости» текста. Более любопытно осмыслить результат, касающийся средней длины слова. Действительно (подобно цепочке: «сад» – «школа» – «колледж» – «университет»), по мере взросления читателя, лексемы, пополняющие его словарный запас, удлинняются и/или обрастают «шубой» морфем (ср.: «хитр-ый» – 2 морфемы – и «бес-хитр-ост-н-ость» – 6 морфем, включая нулевую флексию).

##### Относительные содержательные показатели

##### Морфология

Обращает на себя внимание отсутствие возрастной динамики вводных слов и компаративов. Наперекор нашим ожиданиям оказалось, что восприятие этих речевых элементов, не представляет дополнительной сложности для человека, находящегося на «читающей» стадии развития. Возможно, подобно таблице сложения до десяти, раз выученной и работающей автоматически, с равной скоростью, на протяжении всего дальнейшего периода интеллектуального развития, усвоение моделей «лучше-хуже», «добрее-злее» с определенного этапа уже приобретает совершенный автоматизм и не прогрессирует. С восприятием вводных слов, очевидно, работает аналогичный принцип: исследование показало, что расширение по крайней мере пассивного словарного запаса ребенка на эту языковую единицу происходит в «дочитательском» возрасте. Свойство же вводных слов в определенных контекстах приобретать коннотативные значения с точки зрения когнитивного усложнения оказалось несущественным.

Довольно интересным явился также результат статистического анализа показателя «Частицы». Медленное нарастание их количества показывает, возможно, то, что в процессе «взросления» текста действуют два взаимокомпенсирующих механизма: с одной стороны, уменьшение «детской» эмоционально-разговорной парадигмы (типа «ведь не я же!»), с другой стороны, нарастание «взрослой» сложной семантической парадигмы (включающей уточнительные, персуазивные, причинно-следственные, интенциональные и другие смыслообразующие возможности таких частиц, как «даже», «вряд ли», «ведь», «разве что», «все-таки»). В целом кривая роста частиц свидетельствует о «перевесе» этого широкого спектра модальностей.

Абсолютно предсказуемым оказалось поведение ономастической лексики, «населяющей» в основном дошкольно-младшешкольный литературный мир «хрюкающих свинок» и «крякающих уток». Впоследствии этот пласт языка практически сходит на нет, в более серьезных текстах в основном бытуя не в животнораздражательном, а в шумовом амплуа (Ср. тютчевское: «...весенний первый гром... *грохочет* в небе голубом»).

«Падение» междометий тоже хорошо согласуется с теоретическими предсказаниями: эмоциональная сфера человеческой жизни во взрослых текстах определенно теряет акустическую денотацию. Однако удивительным оказался тот факт, что этого не происходит в русской литературе. По всей вероятности, традиции, заложенные классикой («*Ба!* Знакомые всё лица!»; «*Татьяна ах!* А он реветь...»), на русской почве двадцатый век не изжил.

### *Синтаксис*

Несильно выраженная динамика вопросительных конструкций, предположительно, связана с их амбивалентной функцией в художественном тексте. Если в дошкольной и младшешкольной литературе вопросительные предложения участвуют в весьма частотной организации простой разговорной структуры («Кто там?» – «Мышка норушка»), то к старшей школе на первый план выходят и резко прогрессируют другие структуры: вопросно-ответная форма изложения авторской мысли и риторические вопросы (вот, например, фрагмент из повести В.Пелевина «Омон Ра»: «...меня по-прежнему удивляло – как это я вижу? И кто этот я, который видит? И что это вообще значит – видеть? Вижу ли я что-то внешнее или

просто гляжу сам на себя? И что такое – вне меня и внутри меня?»). При этом сохраняется и диалогово-полилоговая структура текста, но, в отличие от дошкольно-младшешкольных текстов, она нагружена глубоким идейно-философским и драматическим пафосом. Например, такой диалог:

– А ты думаешь, что с этой вот ногой меня еще признают годным? – спрашивает он с горечью.

– Ты ее не спеша подлечишь. Ведь сустав цел. Может, все еще уладится.

(Э.М. Ремарк «На западном фронте без перемен»).

Таким образом, здесь (подобно феномену частиц) мы имеем основание фиксировать два взаимокомпенсирующих процесса, с превалированием «взрослой» тенденции (что интересно, не выраженной в русской литературе, отражающей, таким образом, повествовательно-рассужденческие авторские предпочтения).

В полном согласии с предсказаниями повели себя восклицательные предложения, явив совершенно отчетливый падающий тренд и тем самым показав сопутствующий «взрослению» текста спад экспрессивности.

Симметричную закономерность продемонстрировали сложноподчиненные предложения, относительное преобладание которых во «взрослых» текстах интуитивно не подлежало сомнению.

Осмывая механизмы усложнения когнитивной организации текста, мы пришли к необходимости альтернативной (дополнительной) систематики значимых показателей. Помимо деления на *абсолютные* и *относительные* (включающие в себя *формальные* и *содержательные* – разделенные в свою очередь на *морфологические* и *синтаксические*), мы разделили все показатели исходя из разных принципов отношения с текстом, разных способов его усложнения/упрощения. Мы заметили, что одна часть показателей относится к структуре текста – так сказать, к его внешнему аспекту, а другая часть – к смыслу, то есть к внутреннему аспекту текста. Таким образом, мы разделили все показатели на *структурные* (это совокупный объем текста, относительное количество запятых, относительная длина предложения, относительная длина слова и количество сложноподчиненных предложений) и *смысловые* (это количество междометий, ономастопей, частиц, предлогов, а также вопросительных и восклицательных предложений). Эта упрощенная систематика поможет избежать терминологической путаницы в дальнейших исследованиях.

### **Выводы**

Подводя итоги сказанному, мы можем наметить гипотетические области применения методик оценки простоты/сложности текстов, а также перспективы дальнейших статистических исследований, связанных с когнитивными возможностями художественной литературы.

Легко заметить, что, в связи с понижением конкурентоспособности литературы среди прочих, менее трудозатратных, активно развивающихся средств познания мира, наиболее популярные современные тексты становятся лаконичнее, остросюжетнее, эмоциональнее [Шапошников 2010]. Они содержат меньше описаний и рассуждений (наглядным примером чему может служить, в частности, сверхпопулярность книжных серий наподобие «Иллюстрированной адаптированной классики», выпускаемой издательством «Астрель»; это произведения

Ж. Верна, М. Твена, Р.Л. Стивенсона, Ч. Диккенса, из которых изъяты «скучные» длинноты, угрожающие свести на нет нетерпеливый интерес сегодняшнего читателя). Было бы интересно и важно, во-первых, оценить совокупные тенденции к усложнению/упрощению современной художественной литературы; во-вторых, выявить актуальные для сегодняшних вызовов специфические маркеры простоты/сложности текстов. Например, интересно было выработать алгоритм подсчета глаголов и деепричастий (концентрация которых, на наш взгляд, обратно пропорциональна «сложности» текста), а также прилагательных и наречий (связанных с описательным, то есть более «сложным» компонентом текста).

И, безусловно, наиболее интересным дальнейшим направлением представляется разработка на основе статистических подсчетов и сопутствующих психолингвистических исследований общей методики оценки когнитивной эффективности текста. Это позволит более точно определять возраст целевой аудитории и специфику влияний, которые может оказать произведение на интеллектуальное и личностное развитие человека.

### Литература

*Библик С.П.* Разговорный стиль в контексте национальной культуры // Вестник Ленинградского государственного университета им. А.С. Пушкина, 2013. №1, том 1, С. 164-171.

*Ляшевская О.Н., Шаров С.А.* Частотный словарь современного русского языка. М.: Азбуковник, 2009.

*Мартьянова И.А.* Синтаксические способы выражения и роль обобщений в романах Тургенева // Функционирование синтаксических категорий в тексте. Л., 1981. С. 132-141.

*Пустовойт П.* Диспут поколений (Роман «Отцы и дети» И. С. Тургенева) // Вершины. М., 1981. С. 193-210.

*Тимошук М.* Ономапоеические глаголы в русском языке // Rozprawy Uniwersytetu Warszawskiego. T. 425. С. 10. Wydawnictwa Uniwersytetu Warszawskiego, 1994. 150 с.

*Ушаков Д.В.* Одаренность, творчество, интуиция // Основные современные концепции творчества и одаренности. Москва, 1997. С. 78-89.

*Ушаков Д.В.* Когнитивная система и развитие // Когнитивные исследования. Проблема развития / Под ред. Д.В. Ушакова. М., ИПРАН, 2009. С. 6-12.

*Ушакова Т.Н.* Текст как объект психологического анализа // Психологический журнал, 1989. Т. 10. № 1. С. 107-116.

*Ушакова Т.Н.* Природная организация речевого процесса и восприятие поэтического произведения (на примере лирических стихов А. Ахматовой) // Современные модели психологии речи и психолингвистики. М., ИПАН СССР, 1990. С. 104-112.

*Ушакова Т.Н.* Природные основания речезыковой способности ребенка (анализ раннего речевого развития) // Языковое сознание: формирование и функционирование. Москва, Институт языкознания РАН, 1998. С. 7-23.

*Ушакова Т.Н., Белова С.С., Валеева Е.А.* Лингвопсихологическое исследование вербальной семантики // Психологический журнал, 2010. Т. 31. № 6. С. 83-97.

- Частотный словарь / Под ред. Л.Н. Засориной. М.: «Русский язык», 1977.
- Шапошников В.* Тенденции современной литературы в языковом аспекте // Сборник статей по итогам конференции «XXI век. Итоги литературного десятилетия: Язык – культура – общество / Под ред. А.Ю. Большаковой и А.А. Дырдина. Ульяновск: УлГТУ, 2011. С.406.
- Якобсон П.М.* Психология чувств и мотивации. Воронеж: МОДЭК; Москва: Институт практической психологии, 1998.
- Dehaene S.* Les Neurones de la lecture. Paris: Éditions Odile Jacob, 2007.
- Rhetoricon ad Herennium libri IV* // Энциклопедический словарь Брокгауза и Ефрона: в 86 т. (Т. 26, с. 622). СПб., 1890 – 1907.
- Wobbrock J.O., Findlater L., Gergle D., Higgins J.J.* (2011). The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. CHI '11 Conference on Human Factors in Computing Systems. P. 143–146. <http://doi.org/10.1145/1978942.1978963>

### **Приложение**

Использованные для анализа слова

#### **Вводные слова**

безусловно, в двух словах, в общем, в самом деле, в частности, вернее сказать, вероятно, видимо, во-вторых, возможно, вообще, вообще сказать, во-первых, впрочем, главное, грубо говоря, действительно, допустим, естественно, знать, значит, итак, казалось бы, конечно, короче, кроме того, кстати, лучше сказать, между прочим, можно сказать, можно сказать, мягко говоря, на самом деле, наверное, наверняка, надо сказать, наконец, наоборот, например, напротив, однако, одним словом, очевидно, по сути говоря, по-видимому, подлинно, пожалуй, по-моему, предположительно, прежде всего, проще говоря, прямо сказать, разумеется, с другой стороны, с одной стороны, сказать правду, скорее, скорее всего, следовательно, словом, собственно говоря, стало быть, так сказать, таким образом, тем не менее, точнее сказать, честно говоря, честно сказать.

#### **Компаративы**

белее, ближе, более, больше, быстрее, важнее, веселее, выше, главнее, глубже, горячее, дальше, длиннее, добрее, дольше, дороже, живее, западнее, известнее, интереснее, короче, красивее, крупнее, левее, легче, лучше, любимее, менее, меньше, моложе, необходимее, ниже, новее, нужнее, огромнее, опаснее, позже, полнее, правее, проще, прямее, раньше, роднее, светлее, свободнее, серьезнее, сильнее, скорее, слабее, сложнее, современнее, способнее, старше, страшнее, счастливее, темнее, теплее, тоньше, точнее, тяжелее, увереннее, холоднее, хуже, целее, чаще, человечнее, чернее, чище, шире, ярче.

#### **Подчинительные союзы**

будто, где, дабы, едва, ежели, если, зачем, ибо, как, как бы, какой, когда, коли, который, кто, куда, лишь, насколько, несмотря на то что, откуда, пока, потому что, почему, поэтому, пускай, пусть, с тем чтобы, сколько, словно, так как, хотя, что, чтобы.



**Частицы**

же, неужели, разве, вряд ли, даже, ведь, уж, все-таки.

**Междометия**

ах, ох, ух, уф, фу, ага, ну, ой, ай, эй, ха, ха-ха-ха, о-о-о, а-а-а, тс, гм, хм.

**Подчинительные союзы**

будто, где, дабы, едва, ежели, если, зачем, ибо, как, как бы, какой, когда, коли, который, кто, куда, насколько, несмотря на то что, откуда, пока, потому что, почему, поэтому, с тем чтобы, сколько, словно, так как, хотя, что, чтобы.

**Ономатопейческая лексика**

гавкать, гавкал, гавкала, гавкали, гавкаю, гавкает, загавкал, загавкала, загавкали, загавкаю, загавкает, мяукать, мяукал, мяукала, мяукали, мяукаю, мяукает, замюкал, замюкала, замюкали, замюкаю, замюкает, мычать, мычал, мычала, мычали, мычу, мычит, замычал, замычала, замычали, замычу, замычит, квакать, квакала, квакали, квакаю, квакает, квакала, квакали, квакаю, квакает, крякать, крякал, крякала, крякали, крякаю, крякает, закрякал, закрякала, закрякали, закрякаю, закрякает, кукарекать, кукарекал, кукарекали, кукарекаю, кукарекает, закукарекал, закукарекали, закукарекаю, закукарекает, рычать, рычал, рычала, рычали, рычу, рычит, зарычал, зарычала, зарычали, зарычу, зарычит, хрюкать, хрюкал, хрюкала, хрюкали, хрюкаю, хрюкает, захрюкал, захрюкала, захрюкали, захрюкаю, захрюкает, жужжать, жужжал, жужжала, жужжали, жужжу, жужжит, жужжал, жужжала, жужжали, жужжу, жужжит, бахнула, бахнуло, бахнули, бахнул, бахну, бахнет, бухнуться, бухнулся, бухнулась, бухнулись, бухаюсь, бухается, бухнется, грохнуть, грохнул, грохнула, грохнули, грохаю, грохнет, грохает, грохотать, грохотал, грохотала, грохотали, грохочу, грохочет, загрохотал, загрохотала, загрохотали, загрохочу, загрохочет, хохотать, хохотал, хохотала, хохотали, хохочу, хохочет, захохотал, захохотала, захохотали, захохочу, захохочет, ахать, ахал, ахала, ахали, ахаю, ахает, заахал, заахала, заахали, заахаю, заахает, охать, охал, охала, охали, охаю, охает, заохал, заохала, заохали, заохаю, заохает, чихать, чихал, чихала, чихали, чихаю, чихает, чихнуть, чихнул, чихнула, чихнули, ахаем, бахнем, бухаемся, гавкаем, грохаем, грохочем, жужжим, заахаем, загавкаем, загрохочем, жужжим, заквакаем, закрякаем, закукарекаем, замюкаем, замычим, заохаем, зарычим, захохочем, захрюкаем, квакаем, крякаем, кукарекаем, мяукаем, мычим, охаем, рычим, хохочем, хрюкаем, чихаем, ахаеть, бахнешь, бухаеться, гавкаешь, грохаешь, грохочешь, жужжишь, заахаеть, загавкаешь, загрохочешь, жужжишь, заквакаешь, закрякаешь, закукарекаешь, замаукаешь, замычишь, заохаешь, зарычишь, захохочешь, захрюкаешь, квакаешь, крякаешь, кукарекаешь, маукаешь, мычишь, охаешь, рычишь, хохочешь, хрюкаешь, чихаеть, ахаете, бахнете, бухаетесь, гавкаете, грохаете, грохочете, жужжите, заахаете, загавкаете, загрохочете, жужжите, заквакаете, закрякаете, закукарекаете, замаукаете, замычите, заохаете, зарычите, захохочете, захрюкаете, квакаете, крякаете, кукарекаете, маукаете, мычите, охаете, рычите, хохочете, хрюкаете, чихаете, ахают, бахнут, бухаются, гавкают, грохают, грохочут, жужжат, заахают, загавкают, загрохочут, жужжат, заквакают, закрякают, закукарекают, замаукают, замычат, заохают, зарычат, захохочут, захрюкают, квакают, крякают, кукарекают, маукают, мычат, охают, рычат, хохочут, хрюкают, чихают, тарыхтеть, тарыхтел, тарыхтела, тарыхтели, тарыхтит, затарыхтел, затарыхтела, затарыхтели, затарыхтит.

## COGNITIVE COMPLEXITY OF CHILDREN FICTION: QUANTITATIVE METHODS OF EVALUATION

**Ekaterina A. Valueva**

Ph.D. Researcher,  
Institute of psychology of Russian Academy of Sciences  
Leading researcher,  
Moscow State University of Psychology & Education  
29 Sretenka str., Moscow, 127051, Russian Federation  
*ekval@list.ru*

**Nina M. Danilevskaya**

Philologist, Independent researcher  
*ninadan@mail.ru*

**Ekaterina M. Lapteva**

Ph.D. Leading Researcher  
Federal Institute of Education Development  
9/1 Chernyakhovskogo Str., Moscow, 125319, Russian Federation  
*ek.lapteva@gmail.com*

**Dmitry V. Ushakov**

Corresponding member of Russian Academy of Sciences, D.Sc., Professor  
Head of the laboratory,  
Institute of psychology of Russian Academy of Sciences  
*dv.usakov@gmail.com*

The current study addresses the language markers of the text complexity in children and adolescent literature published from 1900 to 2016. The literature for the analysis consists of the most popular books of the appropriate period. The design of the study has three independent variables: time, book origin (Russian literature, European literature and American and Canadian literature) and the target age (preschool, primary school, middle school and high school). In total 588 texts in Russian and 118 texts in English have been analyzed. We have calculated various volume measures, structure and content measures for each text. The statistical analysis has revealed that some text characteristics have the age-related dynamics and therefore they may be accounted as the markers of text cognitive complexity. These parameters may be used as the basic tool for the further studies of the human speech development in reading that involve text comparison.

**Keywords:** children fiction, the XX<sup>th</sup> century, cognitive complexity.

### References

*Bibik S.P.* Razgovornyj stil' v kontekste natsional'noj kul'tury [Conversational style in the context of national culture] // Vestnik Leningradskogo gosudarstvennogo universiteta im. A.S. Pushkina, 2013. №1, tom 1, P. 164-171.

Chastotnyj slovar' [Frequency dictionary] / Ed. L.N. Zazorina. M.: «Russkij yazyk», 1977.

*Dehaene S.* Les Neurones de la lecture. Paris: Éditions Odile Jacob, 2007.

*Lyashevskaya O.N., Sharov S.A.* Chastotnyj slovar' sovremennogo russkogo yazyka [Frequency dictionary of the contemporary Russian language]. M.: Azbukovnik, 2009.

*Mart'yanova I.A.* Sintaksicheskie sposoby vyrazheniya i rol' obobshhenij v romanakh Turgenjeva [Syntactic styles and a role of generalizations in I.S. Turgenev's novels] // Funkcionirovanie sintaksicheskikh kategorij v tekste [Syntactic categories functioning in literary works]. L., 1981. P. 132-141.

*Pustovojt P.* Disput pokolenij (Roman «Ottsy i deti» I.S. Turgenjeva) [Generation dispute (I.S. Turgenev's novel "Fathers and Sons")] // Vershiny. M., 1981. P. 193-210.

Rhetoricon ad Herennium libri IV // Entsiklopedicheskij slovar' Brokgauza i Efrona [Brokgauz's and Efron's encyclopaedic dictionary]: v 86 t. (Vol. 26, p. 622). SPb., 1890 – 1907.

*Shaposhnikov V.* Tendentsii sovremennoj literatury v yazykovom aspekte [Contemporary literature tendencies from the linguistic point of view] // Sbornik statej po itogam konferentsii «XXI vek. Itogi literaturnogo desyatiletija: Yazyk – kul'tura – obshhestvo» [Papers from the conference "XXI century. Summing up the literary decade: Language – culture – society"] / Eds. A.Yu. Bol'shakova i A.A. Dyrdin. Ul'yanovsk: UIGTU, 2011. P. 406.

*Timoshuk M.* Onomatopeicheskie glagoly v russkom yazyke [Onomatopoeic verbs in Russian] // Rozprawy Uniwersytetu Warszawskiego. Vol. 425. P. 10. Wydawnictwa Uniwersytetu Warszawskiego, 1994. 150 pp. Ushakov D.V. Kognitivnaya sistema i razvitie [Cognitive system and development] // Kognitivnye issledovaniya. Problema razvitiya [Cognitive research. The problem of development] / Ed. D.V. Ushakov. M., IPRAN, 2009. P. 6-12.

*Ushakov D.V.* Odarennost', tvorcestvo, intuitsiya [Giftedness, creativity, intuition] // Osnovnye sovremennye kontseptsii tvorcestva i odarennosti [Mainstream theories of creativity and giftedness]. Moskva, 1997. P. 78-89.

*Ushakova T.N.* Prirodnaya organizatsiya rechevogo protsessa i vospriyatие poeticheskogo proizvedeniya (na primere liricheskikh stikhov A. Akhmatovoj) [Natural organization of vocal process and poetry perception (by the example of A. Akhmatova's lyrics)] // Sovremennye modeli psikhologii rechi i psikholingvistiki [Contemporary models in psycholinguistics]. M., IPAN SSSR, 1990. P. 104-112.

*Ushakova T.N.* Prirodnye osnovaniya recheyazykovoj sposobnosti rebenka (analiz rannego rechevogo razvitiya) [Natural basis of child's speech and language abilities (analysis of early speech development)] // Yazykovoe soznanie: formirovanie i funkcionirovanie [Language consciousness: formation and functioning]. Moskva, Institut yazykoznanija RAN, 1998. P. 7-23.

*Ushakova T.N.* Tekst kak ob'ekt psikhologicheskogo analiza [Text as an object of psychological studies] // Psikhologicheskij Zhurnal [Psycholinguistic Journal], 1989. T. 10. № 1. P. 107-116.

*Ushakova T.N., Belova S.S., Valueva E.A.* Lingvo-psychological study of verbal semantics // Psikhologicheskij Zhurnal [Psycholinguistic Journal], 2010. T. 31. № 6. P. 83-97.

*Wobbrock, J. O., Findlater, L., Gergle, D., & Higgins, J. J. (2011). The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. CHI '11 Conference on Human Factors in Computing Systems, 143–146. <http://doi.org/10.1145/1978942.1978963>*

*Yakobson P.M. Psikhologiya chuvstv i motivatsii [Psychology of feelings and motivation]. Voronezh: MODEK; Moskva: Institut prakticheskoy psikhologii, 1998.*

Ψ λ