

УДК 81'33

ВЗАИМОСВЯЗЬ ПАРАМЕТРОВ РУССКОГО ПИСЬМЕННОГО ТЕКСТА И ПРОФИЛЯ ФУНКЦИОНАЛЬНОЙ СЕНСОМОТОРНОЙ АСИММЕТРИИ ЕГО АВТОРА¹

РЫЖКОВА Екатерина Сергеевна,

аспирант, научный сотрудник Лаборатории корпусной социолингвистики и автороведческих исследований

ЛИТВИНОВА Татьяна Александровна,

кандидат филологических наук, руководитель Лаборатории корпусной социолингвистики и автороведческих исследований,

Воронежский государственный педагогический университет

АННОТАЦИЯ. Предлагается принципиально новый подход к изучению типологических особенностей письменной речи лиц с разным типом профиля функциональной сенсомоторной асимметрии его автора на материале электронного корпуса русских письменных текстов *RusNeuroPsych*, содержащего, кроме собственно текстов, метаразметку в виде информации об их авторах (пол, возраст, образование, результаты психологических тестов и нейропсихологического обследования). Дается краткая характеристика корпуса *RusNeuroPsych*. Описывается процесс обработки собранного авторами материала с помощью средств автоматического анализа языка и методов математической статистики. Обсуждаются результаты корреляционного анализа, направленного на выявление взаимосвязей между параметрами текста и типом профиля функциональной сенсомоторной асимметрии его автора.

КЛЮЧЕВЫЕ СЛОВА: письменный текст, корпус текстов, формально-грамматические параметры текста, естественная письменная речь, профиль функциональной сенсомоторной асимметрии, нейронауки, математическая статистика, корреляции.

RYZHKOVA E.S.,

Postgraduate Student, Research Assistant of the Corpus Sociolinguistics and Authorship Profiling Researches Lab

LITVINOVA T.A.,

Cand. Philol. Sci., Head of the Corpus Sociolinguistics and Authorship Profiling Research Laboratory Voronezh State Pedagogical University

THE INTERRELATIONSHIP BETWEEN RUSSIAN WRITTEN TEXT PARAMETERS AND FUNCTIONAL SENSORIMOTOR ASYMMETRY PROFILE OF ITS AUTHOR

ABSTRACT. The article proposes a fundamentally new approach to studying typological features of written speech of individuals with different types of functional sensorimotor asymmetry profile. The research is based on Russian written texts electronic corpus *RusNeuroPsych* containing, in addition to the actual texts, the meta-marking as information about the authors (gender, age, education, results of neuropsychological survey and psychological testing). A brief description of *RusNeuroPsych* corpus is provided. The study considers the details of data collection and procession using automatic language analysis and the methods of mathematical statistics. The paper discusses the results of the correlation analysis, aimed at identifying the relationship between the text parameters and the type of its author's functional sensorimotor asymmetry profile.

KEY WORDS: written text, corpus of texts, formal-grammar parameters of a text, natural written speech, functional sensorimotor asymmetry profile, neurosciences, mathematical statistics, correlations.

Введение. Речь человека позволяет узнать о его некоторых индивидуально-личностных особенностях [1]. Проблема отражения характеристик личности в ее речевой продукции на протяжении длительного времени привлекает внимание лингвистов, литературоведов, психологов, криминалистов, социологов, математиков, а в последнее время и специалистов по информационным технологиям [2, с. 5].

Весьма заметный всплеск интереса к этой проблеме в последние несколько лет связан со стремительным развитием различных форм интернет-коммуникации и широкой распространенностью ситуации, при которой единственным источником информации о человеке становится текст [там же].

Уровень развития современной науки позволяет исследователям шире рассматривать проблему отражения в создаваемых человеком речевых произведениях его различных индивидуально-личностных особенностей, не ограничиваясь с

¹ Исследование выполнено при поддержке гранта РФФИ «Языковые параметры письменного текста и нейропсихологические характеристики его автора: корпусное исследование», № 16-36-00036.

изучением представленности в речи лишь пола, возраста, уровня образования, психологических характеристик личности, а учитывать особенности мозговой организации, в частности профиля латеральной организации функций, который, по мнению ученых, «может проявляться на уровне особенностей построения и оформления текста» [3, с. 89–97]. В работах современных исследователей данное понятие (одна из важнейших нейропсихологических характеристик человека) получает различные номинации: *индивидуальный латеральный профиль* [4], *индивидуальный профиль латеральности* [5], *латеральный фенотип человека* [6] и др. В нашем исследовании мы используем термин «*профиль функциональной сенсомоторной асимметрии*» (далее – ПФСА), предложенный Т.А. Доброхотовой и Н.Н. Брагиной и определяемый как сочетание сенсорных и моторных асимметрий [7].

Исходя из вышесказанного, мы считаем необходимым проводить исследования русских письменных речевых произведений с учетом достижений современных нейронаук – нейролингвистики нормы [8] и нейропсихологии индивидуальных различий, посвященных обследованию здоровых лиц [9].

Целью нашего исследования является выявление типологических параметров письменных русскоязычных текстов, создаваемых условно здоровыми людьми с разными профилями функциональной сенсомоторной асимметрии, с применением программных средств обработки языка и методов математической статистики, а также опорой на данные нейронаук (нейролингвистики нормы и нейропсихологии индивидуальных различий).

Материал. Для решения поставленной задачи в течение 2015 года нами проводился психолингвистический эксперимент, в ходе которого был собран электронный корпус связанных русских письменных текстов *RusNeuroPsych* (специально написанных респондентами по заданию исследователя), структурированный определенным образом и содержащий, помимо собственно речевых произведений, метаразметку в виде информации об их авторах (пол, возраст, родной язык, уровень образования, результаты психологического тестирования и нейропсихологического обследования) [10]. Названный корпус представляет собой подкорпус корпуса текстов *RusPersonality*, создаваемого с 2012 года научными сотрудниками Регионального центра русского языка при ВГПУ под руководством Т.А. Литвиновой [11; 12; 13].

Общее количество речевых произведений формируемого нами электронного корпуса *RusNeuroPsych* – 644 текста от 455 авторов. Все собранные тексты являются образцами естественной письменной русской речи [14]. На данном этапе исследования нами было проанализировано 242 текста 121 респондента (каждый респондент писал два текста – письмо другу и описание картины) в возрасте от 24 до 45 лет, среди них 17 лиц мужского пола и 104 – женского. Средняя длина текстов корпуса составляет 130 слов.

Методы

Лингвистическая разметка корпуса

Для решения задач настоящего исследования нами была осуществлена лингвистическая разметка всех текстов корпуса *RusNeuroPsych* с помощью морфологического анализатора *rumorphy2* и онлайн-сервиса *istio.com*. На данном этапе всего нами получены числовые значения 22 формально-грамматических параметров собранных текстов:

1. Водность, %, т.е. доля в тексте неполнозначных слов, вычисляемая сервисом *istio.com*.
 2. TTR (число разных слов в 100 первых словах текста).
 3. TTR ядра (число разных слов в 100 первых в ядре текста).
 4. ПП2 (процент покрытия словаря текста пятью самыми частотными словами (со стоп-словами)).
 5. ПП3 (доля повторов стоп-слов среди повторов 5 самых частотных слов).
 6. Доля служебных слов + местоимений в тексте.
 7. Доля служебных слов.
 8. Доля союзов.
 9. Доля предлогов.
 10. Доля прилагательных.
 11. Соотношение (предлоги + союзы + частицы + глаголы-связки) / (сущ. + прилагательные + местоимения + глаголы).
 12. P-density (показатель, выражающий степень преобладания утверждений или вопросов над простой отсылкой к предметам или событиям, т.е. доля в тексте слов тех частей речи, которые выражают пропозиции: *глаголов, прилагательных, наречий, предлогов и союзов*) [15, с. 540–542].
 13. Формальность текста (F-measure), вычисляемая как $F = (\text{существительные} + \text{прилагательные} + \text{предлоги} - \text{местоимения} - \text{глаголы} - \text{причастия} - \text{наречия} - \text{союзы} - \text{междометия}) + 100) / 2$ [16, с. 227].
 14. Доля прилагательных и местоименных прилагательных.
 15. Доля предлогов и местоименных прилагательных.
 16. Доля слов, чья длина превышает шесть букв.
 17. Доля строевых слов (местоимения, предлоги, модальные глаголы, союзы, числительные и междометия).
 18. Доля слов, обозначающих мыслительные процессы и отношения (глаголы: *думать, считать, полагать*, а также подчинительные союзы типа *потому что, так как* и т.д.).
 19. Отношение числа всех знаков препинания в тексте к числу слов в тексте.
 20. Отношение числа точек в тексте к числу слов в тексте.
 21. Отношение числа запятых в тексте к числу слов в тексте.
 22. Доля в тексте слов, входящих в 100 самых частотных слов русского языка [17].
- Выбор данных параметров обусловлен тем, что они, во-первых, присущи каждому тексту, во-вторых, слабо зависят от темы текста и не подвержены сознательному искажению, и, в-третьих, как показали наши исследования (см. обзор в [13]), коррелируют с различными личностными характеристиками их авторов.

Методики обследования респондентов

Существование множество различных способов определения профиля функциональной сенсомоторной асимметрии [подробнее об этом см.: 7; 9; 18 и др.]. В нашем исследовании мы использовали те пробы определения моторного и сенсорного профилей (определение ведущей руки, ноги, глаза, уха), выполнение которых несложно, не требует об • Гуманитарные науки

нимает много времени [4; 19; 20]. методика обследования респондентов подробно описана нами [10].

В нашей работе мы рассчитывали индекс латеральной организации функций головного мозга респондентов (показатели моторного и сенсорного профилей) как разницу между количеством «правых»,

«левых» и «смешанных» ответов информантов, разделенных на число тестов. Интегральный индекс профиля функциональной сенсомоторной асимметрии также подсчитывался как разница между общим количеством «правых», «левых» и «смешанных» ответов испытуемых во всех тестах, разделенных на число тестов. Для осуществления названных расчетов мы использовали следующую формулу:

$$\frac{\text{Правые} - \text{левые} - \text{смешанные}}{\text{всего проб}}$$

К примеру, для определения ведущей руки респондент выполнял всего 7 проб, 5 из них он выполнил правой рукой, одну – левой, в одной из проб не было зафиксировано преобладание правой/левой руки. Соответственно, показатель «ведущая рука» у данного респондента составит $(5-1-1)/(7)=0,4$.

Математическая обработка данных

Программное обеспечение SPSS Statistics 24.0 было применено нами для вычисления коэффициента корреляции (the Pearson coefficient) между параметрами письменных текстов и индексами латеральной организации функций головного мозга их авторов.

Для большей объективности мы проводили анализ одного и того же языкового материала в двух сериях эксперимента. Так, в первой серии эксперимента оба текста одного автора (письмо другу и описание картины) мы рассматривали как один текст («суммарный корпус»), а во второй – два текста были проанализированы нами отдельно («индивидуальный корпус»). В процессе обработки собранного языкового материала мы учитывали только те формально-грамматические параметры текстов,

которые показали корреляцию с характеристиками моторного и сенсорного профилей их авторов в двух сериях эксперимента.

Отмеченные выше обстоятельства обуславливают значительную трудоемкость данного этапа работ в рамках нашего исследования.

Результаты. Наибольшее количество корреляций для уровня значимости $p < 0.05$ было найдено между текстовыми параметрами и моторными асимметриями (8), ведущей рукой (8), ПФСА_{интегральный} (7) (коэффициент корреляции от 0.27 до 0.41). Гораздо меньше значимых корреляций обнаружено между показателями сенсорной асимметрии, исключение составляет параметр «ведущий глаз» (5).

Так, положительная корреляция была обнаружена между показателем «ведущая рука», «моторные асимметрии» и ПФСА_{интегральный} и индексом лексического разнообразия TTR (число разных слов в 100 первых словах текста), т.е. чем больше правосторонних признаков у человека, тем выше индекс лексического разнообразия его текстов. Отрицательная корреляция выявлена между названными параметрами и долей служебных слов + местоимений; долей служебных слов; долей слов, обозначающих мыслительные процессы и отношения; долей знаков препинания; долей 100 самых частотных русских слов, т.е. чем больше правосторонних признаков в ПФСА респондента, тем ниже значения этих показателей.

В приведенной ниже таблице представлены данные о корреляциях между параметрами письменных текстов и характеристиками профиля функциональной сенсомоторной асимметрии их авторов.

Таблица – Корреляции между характеристиками профиля функциональной сенсомоторной асимметрии респондентов и параметрами их текстов

ПФСА	Параметры текстов	Знак корреляции
ПФСА _{интегральный}	TTR (число разных слов в 100 первых словах текста)	+
	доля служебных слов + местоимений	-
	водность	-
	доля союзов	-
	доля служебных слов	-
	доля слов, обозначающих мыслительные процессы и отношения	-
	доля частотных слов	-
Моторные асимметрии	TTR (число разных слов в 100 первых словах текста)	+
	TTR ядра (число разных слов в 100 первых словах ядра текста)	+
	доля служебных слов + местоимений	-
	водность	-
	формальность	+
	служебные слова	-
	доля слов, обозначающих мыслительные процессы и отношения	-
частотные слова	-	
Ведущая рука	TTR (число разных слов в 100 первых словах текста)	+
	TTR ядра (число разных слов в 100 первых словах ядра) • Гуманитарные науки	+
	доля служебных слов + местоимений	-
	доля служебных слов	-
	доля слов, обозначающих мыслительные процессы и отношения	-
	• Филологические науки	

ПФСА	Параметры текстов	Знак корреляции
Ведущая нога	отношения	
	частотные слова	–
	формальность	+
Ведущее ухо	частотные слова	–
	доля служебных слов	–
	доля союзов	–
	водность	–
	доля слов, обозначающих мыслительные процессы и отношения	–
	частотные слова	–

Заключение. Полученные нами в ходе исследования значимые корреляции между параметрами текстов корпуса *RusNeuroPsych* и характеристиками типа ПФСА их авторов подтверждают целесообразность использования предлагаемого нами подхода к изучению типологических особенностей русской письменной речи условно здоровых лиц с различными типами ПФСА функций головного мозга с применением программных средств обработки языка, методов математической статистики и опорой на данные современных нейронаук (нейролингвистики нормы и нейропсихологии индивидуальных различий). Для дальнейшего исследования проблемы взаимосвязи характеристик текста и нейропсихологических особенностей его автора необходимо расширение корпуса связанных письменных текстов на русском языке *RusNeuroPsych* и его лингвистическая разметка.

Далее мы надеемся провести корреляционно-регрессионный анализ для построения математиче-

ских моделей, позволяющих на основе формально-грамматических параметров текстов определить тип ПФСА их авторов. Кроме того, нами также планируется исследовать взаимосвязи между типом профиля функциональной сенсомоторной асимметрии и другими характеристиками авторов собранных нами связанных письменных русских текстов (возраст, пол, данные психологического тестирования).

Думается, что разрабатываемый нами подход, состоящий в выявлении типологических особенностей письменной речи лиц с разным типом профиля функциональной сенсомоторной асимметрии, найдет свое применение в исследованиях теоретического и прикладного характера в различных областях знаний: нейролингвистике нормы, судебном автороведении, корпусной лингвистике и т.д.

СПИСОК ЛИТЕРАТУРЫ:

1. Фомина, Н.А. Свойства личности и особенности речевой деятельности [Текст] / Н.А. Фомина. – Рязань : Узорочье, 2002. – 412 с.
2. Литвинова, Т.А. Идентификация и диагностирование личности автора письменного текста : монография [Текст] / Т.А. Литвинова, О.А. Литвинова. – Воронеж : Воронежский государственный педагогический университет, 2015а. – С. 5.
3. Шубин, А.В. Асимметрия мозга и особенности вербальной креативности [Текст] / А.В. Шубин, Е.И. Серпионова // Вопросы психологии. – 2007. – № 4. – С. 89–97.
4. Сиротюк, А.Л. Нейропсихологическое и психофизиологическое сопровождение обучения [Текст] / А.Л. Сиротюк. – М. : ТЦ Сфера, 2003. – 288 с.
5. Москвин, В.А. Межполушарные асимметрии и индивидуальные различия человека [Текст] / В.А. Москвин, Н.В. Москвина. – М. : Смысл, 2011.
6. Леутин, В.П. Психофизиологические механизмы адаптации и функциональная асимметрия мозга [Текст] / В.П. Леутин, Е.И. Николаева. – Новосибирск : Наука. Сибирское отделение, 1988.
7. Брагина, Н.Н. Функциональные асимметрии человека [Текст] / Н.Н. Брагина, Т.А. Доброхотова – 2-е изд., перераб. и доп. – М. : Медицина, 1988. – 240 с.
8. Ахутина, Т.В. Нейролингвистика нормы [Текст] / Т.В. Ахутина // I Международная конференция памяти А.Р. Лурия : сб. докладов. – М., 1998. – С. 289–298.
9. Нейропсихология индивидуальных различий: учеб. пособие для студ. учреждений высш. проф. образования. – М. : Издательский центр «Академия», 2011.
10. Литвинова Т. А. Электронный корпус письменных текстов RusNeuroPsych: состав, структура и возможности использования [Текст] / Т.А. Литвинова, Е.С. Рыжкова // *Studia Humanitatis*. – 2016. – № 3.
11. Литвинова, Т.А. Индивидуально-личностные характеристики автора и количественные параметры его текста: корпусное исследование [Текст] / Т.А. Литвинова, Е.В. Диброва, П.В. Середин, О.А. Литвинова // Вопросы психологии. – 2015б. – № 4. – С. 98–108.
12. Литвинова, Т.А. Корпусные исследования письменной речи в решении задач судебного автороведения [Текст] / Т.А. Литвинова, Е.В. Диброва, О.А. Литвинова, Е.С. Рыжкова // *Филологические науки. Вопросы теории и практики Тамбов : Грамота, 2015в. – № 8. – Ч. 1. – С. 107–113.*
13. Litvinova, T., Litvinova O., Zagorovskaya O., Seredin P., Sboev A., Romanchenko O. "Ruspersonality": a Russian Corpus for Authorship Profiling and Deception Detection // *Proceedings of Proceedings of International FRUCT conference on Intelligence, Social Media and Web (FRUCT 2016)*. pp. 1–7.
14. Лебедева, Н.Б. Естественная письменная русская речь как объект лингвистического исследования [Текст] / Н.Б. Лебедева // *Вестник ВГПУ*. – 2001. – № 1. – С. 4–10.
15. Brown, C., Snodgrass, T., Kemner S., Herman R. and Covington M. (2008). Automatic measurement of propositional idea density fr. • *Гуманитарные науки* • *Филологические науки*
16. Nini A. Authorship profiling in a forensic context. PhD thesis, 2014. (<http://library.college-police.uk/docs/theses/Nini-Authorship-profiling-2014.pdf>).
17. Ляшевская, О.Н. Частотный словарь современного русского языка (на материалах Национального корпуса русского языка) [Текст] / О.Н. Ляшевская, С.А. Шаров. – М. : Азбуковник, 2009.

18. Николаева, Е.И. Сравнение разных способов оценки профиля функциональной сенсомоторной асимметрии у дошкольников [Текст] / Е.И. Николаева, Е.Ю. Борисенкова // Асимметрия. – 2008. – № 1. – С. 32–39.
19. Леутин, В.П. Функциональная асимметрия мозга. Мифы и действительность [Текст] / В.П. Леутин, Е.И. Николаева – СПб. : Речь, 2005. – 368 с.
20. Хомская, Е.Д. Нейропсихология индивидуальных различий [Текст] / Е.Д. Хомская, И.В. Ефимова, Е.В. Будыка, Е.В. Ениколопова. – М. : Рос. пед. агентство, 1997. – 282 с.