

ПРОФИЛИРОВАНИЕ АВТОРА ПИСЬМЕННОГО ТЕКСТА

Т.А. Литвинова

Аннотация. В настоящее время считается доказанным, что в тексте отражается личность его автора. Установлено, что одним из эффективных путей выявления личностных особенностей является анализ дейктических единиц, особенно личных местоимений, предлогов и союзов. Предложено применить методики установления особенностей личности автора текста путем анализа дейктических единиц к русскому языку и рассмотреть возможность анализа дейксиса как средства выявления характеристик автора письменного текста.

Ключевые слова: текст; автор; дейксис; языковая личность; профилирование.

В настоящее время является общепризнанным, что текст как продукт индивидуальной речевой деятельности может служить средством диагностики индивидуально-типических особенностей свойств личности, поскольку за каждым текстом стоит своеобразная языковая личность [1].

Теория языковой личности Ю.Н. Караулова [2] является одной из основополагающих для теоретического осмысления и обоснования классификационных исследований в криминалистической экспертизе в целом, будь то судебное автороведение, фоноскопия или психология [3]. Она может применяться при диагностировании личности по тексту в маркетинговых исследованиях, а также в других прикладных областях.

В отечественной и зарубежной науке в течение последних 50 лет активно формируется направление, связанное с интеграцией собственно лингвистических и психолингвистических методов идентификации личности по устной и письменной речи.

Общий подход заключается в том, чтобы на основе исследования формальных (собственно языковых – фонетических, грамматических, стилистических и пр. – и текстовых; в зарубежной терминологии *style-based* – Т.Л.) и неформальных (содержательных, смысловых, мотивационных и др.; в зарубежной терминологии – *content-based* – Т.Л.) характеристик речевого произведения определить отдельные черты породившего его человека либо, в идеале, его целостный личностно-психологический портрет [4].

В зарубежной лингвистике вошел в употребление термин *author profiling* [5] – профилирование, т.е. установление характеристик автора по созданному им речевому произведению.

Мы рассмотрим формальные (*style-based*) характеристики текста, релевантные для профилирования автора письменного текста (ПАПТ). Исследователи, рассматривающие возможности применения формальных характеристик для ПАПТ, основываются на предположении, что слова, которые люди употребляют, несут, кроме своего основного значения, психологическую информацию независимо от контекста [5].

Кроме того, было установлено, что особенности выбора лексико-грамматических разрядов языковых единиц для конкретного человека являются достаточно стабильными показателями и сохраняются независимо от контекста, причем это касается как базовых лексико-грамматических категорий, так и более связанных с психологией разрядов языковых единиц, например оценочных слов [5].

Разрабатывают данную проблему и отечественные ученые. Так, еще в 1961 г. Л.Е. Ароцкер писал: «В языковом поведении каждого человека отражается динамический стереотип, т.е. человек склонен употреблять одни синонимы, а не другие» [6. С. 142].

Первоначально этой проблемой занимались ученые-криминалисты в рамках разработки методик судебно-автороведческой экспертизы. С.М. Вул в 1980 г. в статье «Предмет и объекты судебно-автороведческой экспертизы» писал: «Предмет судебно-автороведческой экспертизы – фактические данные о личности автора документа, устанавливаемые экспертом на основе специальных познаний», причем эти сведения (возраст, национальность, профессия, черты характера и др.) «устанавливаются с помощью специальных познаний путем исследования структурных особенностей текста документа» [7. С. 110].

К настоящему времени накоплено немало исследований, свидетельствующих о выраженности разнообразных индивидуальных свойств субъекта говорения (письма) в продуцируемом им тексте. Однако значительно прогресса в этой сфере до сих пор достигнуто не было [8. С. 205].

В настоящей статье мы рассмотрим основные работы отечественных и зарубежных исследователей, связанные с проблемой отражения в письменной речи различных особенностей человека, т.е. проблем профилирования автора письменного текста, и наметим вопросы, которые должны быть решены в дальнейшем.

Отражение особенностей человека в письменной речи

Традиционно в психологии и лингвистике человек рассматривается как единство физического и психологического [9].

В зависимости от области применения (судебная экспертиза, медицина, маркетинг) набор извлекаемых характеристик человека в автоматическом профилировании может меняться. Но одними из наиболее

значимых для идентификации личности автора характеристик являются его гендерный признак и возраст, а также психологические особенности.

Определение физических особенностей автора текста

Пол. Самым разработанным в отечественном лингвистике параметром ПАПТ является пол. Работы, направленные на определение пола автора анонимного текста, начали проводиться в 70-е гг. XX в. в Харьковском НИИСЭ и связаны с именами таких исследователей, как С.М. Вул, Е.И. Горошко и др. [3].

В одном из последних исследований по данной проблеме установлено, что мужской текст короче женского, в нем больше ошибок, меньше отрицательных частиц и приставок, меньше терминов *семантической безысключительности* («вечно», «всегда», «никогда», «никто»), большее разнообразие словаря и меньше речевых клише [10. С. 101].

Качественный анализ мужских текстов показал, что в них намного меньше описаний и упоминаний своих эмоций.

Женский текст, напротив, напоминает больше речь в состоянии эмоционального напряжения: в нем будет много отрицательных частиц, терминов *семантической безысключительности*, речевых клише и штампов, часто употребляемых неправильно, кроме того, в текстах женщин часто присутствуют и упоминания собственных разнообразных эмоций [Там же. С. 102].

Связь пола и языка говорящего широко изучается и западными учеными. Основополагающей является работа Р. Лакофф «Язык и место женщины», которая считала, что из-за ограниченной власти в обществе женщины используют менее агрессивный язык, что проявляется в употреблении формул вежливости, интенсификаторов, слов, выражающих неопределенность и т.д.

Обзор современных исследований на эту тему приведены в [11]. М. Мель и Дж. Пеннебакер [12] на основе сочинений студентов в жанре «поток сознания» выяснили, что в целом мужчины, по сравнению с женщинами, используют намного больше длинных слов, артиклей, предлогов, слов с негативной оценочностью, меньше местоимений «я», «мой» и больше заместителей.

Женщины употребляют больше слов-заполнителей, слов с семантикой долженствования и положительной оценочностью, хотя в целом количество оценочных слов одинаково. Женскую речь также отличает высокий уровень местоимений первого лица единственного числа. Примечательно, что те же самые данные были подтверждены на материале научных текстов: женщины используют больше местоимений, а мужчины – больше

определителей существительных (с предлогом *of*). Женщины употребляют больше маркеров «вовлеченности», тогда как мужчины – больше маркеров «информированности».

Возраст. Исследований, посвященных выявлению возраста автора анонимного текста, насколько нам известно, российскими исследователями не проводилось. Зарубежными учеными этот параметр также недостаточно изучен.

Одно из немногих исследований на эту тему [13] посвящено связи возраста и языковых характеристик. Было установлено, что по мере увеличения возраста люди используют больше слов с положительной оценкой, меньше – с негативной, меньше личных местоимений первого лица, больше глаголов в будущем времени и меньше – в прошедшем. Возрастает также показатель «когнитивной сложности»: увеличивается процент слов, указывающих на причинно-следственные отношения (из-за, по причине, вследствие, потому что и др.), длинных слов (больше 6 букв) и др.

Эти данные опровергают устойчивое мнение о том, что с возрастом люди становятся более обращенными в прошлое и страдают от снижения интеллектуальных способностей. Надо сказать, что последние исследования в психологии также опровергают это мнение [13].

Определение психических особенностей автора текста

В отечественной лингвистике в последнее время стали активно разрабатываться вопросы, связанные с установлением психологических особенностей автора текста. Так, В.Ф. Енгальчев пришел к выводу, что «по тексту можно с достаточной степенью определенности описать психологический портрет его автора» [8. С. 213].

Ученые установили связь некоторых формальных критериев и определенных личностных характеристик (например, выявлена положительная корреляционная зависимость между категорией «отношение глаголов к прилагательным и причастиям») и фактором «прямолинейность – дипломатичность») в текстах сочинений.

На основе анализа личностных смыслов в художественной литературе была создана типология текстов, в основу которой положена эмоционально-смысловая доминанта, соотносимая со структурами того или иного типа акцентуации личности [14]. Данная теория находит практическое применение в проблеме идентификации личности по речи [15].

Очевидно, что выявленные группы, объединенные эмоционально-смысловой доминантой, формирующей тип текста, являются вербальной манифестацией сознания личности. Есть основания говорить о том, что если слова тех или иных типов текстов в количественном отношении преобладают над словами других типов и частотность их употребления

в анализируемом тексте превышает их частотность в языке, то это свидетельствует об определенном типе личности автора текста [15].

Была установлена связь между типами личности и способами построения текстов [16]. Н.А. Фомина [1] в своей докторской диссертации на соискание степени доктора психологических наук «Индивидуально-типические особенности свойств личности и их проявления в речи» создает классификацию типов организации свойств личности на основе характеристик общительности, настойчивости и инициативности и приходит к выводу, что данные особенности обнаруживают общие и специфические проявления в лингвистических характеристиках текста на языковом, речевом и содержательно-смысловом уровнях, а также в различных психологических компонентах речевых действий: операционально-динамическом, мотивационном, когнитивном, эмоциональном, регуляторном и результативном.

На Западе проблема отражения личности в языке начала активно изучаться еще в 40-х гг. XX в. В 1942 г. Ф. Сэнфорд [17] предположил, что речевое поведение является надежным источником сведений о личностных свойствах человека. С тех пор накоплено много эмпирических данных, однако связь между языковыми характеристиками и личностными свойствами только начинает получать свое теоретическое обоснование.

В последние годы в американской лингвистике и психологии все больше исследований доказывают, что анализ использования местоимений, предлогов, союзов и других дейктических едениц позволяет получить достаточно информации об авторе текста.

Так, находясь в депрессии или в сложной ситуации, люди используют больше местоимений (особенного первого лица единственного числа), меньше артиклей и больше модальных глаголов в настоящем времени. Оказываясь перед лицом национальных бедствий, как, например, 11 сентября 2001 г., локальных катастроф или узнавая о смерти известного человека, люди чаще употребляют местоимения первого лица множественного числа и реже – единственного [18].

Одно из последних исследований выявило, что служебные слова служат надежным коррелятом личностных черт, таких как невротизм, экстравертность, открытость для нового, самоуважение и стремление к социальному доминированию [19].

Распознавание лжи в письменном тексте

Распознаванию лжи в письменном тексте на основе анализа языковых характеристик посвящено множество работ ученых [20]. В целом исследователи рассматривают три языковые категории: местоимения, слова с эмоциональной оценкой и слова, указывающие на последовать и логику изложения мыслей (*cognitive complexity words*).

Установлено, что ложные высказывания содержат меньше местоимений первого лица, поскольку люди пытаются дистанцироваться от своих ложных высказываний и отказываются принимать ответственность за собственное поведение [21–23]. Также ложные высказывания содержат больше слов негативной оценочности, чем правдивые, поскольку ложь связана с беспокойством и, в некоторых случаях, чувством вины.

Показатели когнитивной сложности ассоциируются с правдивыми высказываниями. Одна из групп таких показателей – так называемые союзы и предлоги с семантикой исключения (но, кроме, исключая, без). Такие слова требуют от говорящего различать то, что находится в категории, от того, что находится вне ее. При построении лживого высказывания намного сложнее придумывать то, что было, чем то, чего не было [24].

Дейксис как личностный предиктор

Нам представляется продуктивным подход, предложенный, в частности, Дж. Пеннебакером и подробно описанный в его недавно вышедшей книге [5].

Результаты, полученные автором, впечатляют: основываясь на анализе служебных слов, он определяет с точностью, намного превышающей 50% (т.е. случайную величину), пол, возраст, склонность к депрессии, психологические особенности и прогнозирует с высокой вероятностью, основываясь на данных лонгитюдных исследований, развитие отношений у пар на начальном этапе, вводя понятие LSM (*Language Style Matching*) – чем больше совпадений имеют кривые, отражающие формальные характеристики стиля (в частности, процент употребления тех или иных категорий служебных слов, местоимений, наречий и др.), тем лучше прогноз в развитии отношений мужчины и женщины.

Как замечает Дж. Пеннебакер, так называемые функциональные слова (к ним он относит местоимения, артикли, предлоги, союзы и модальные глаголы) особенно ценны для ПАПТ по следующим причинам. В английском языке менее 200 наиболее часто употребляемых функциональных слов, и они покрывают более 50% всех слов, которые используют говорящие на данном языке. Что особенно важно, исследование афазий показывает, что такие слова хранятся в отличных от полнозначных слов зонах мозга и обрабатываются по-другому.

Например, повреждение зоны Брока часто сопровождается корректным использованием существительных, прилагательных, глаголов, но отсутствием в речи служебных слов и местоимений. Повреждение зоны Вернике заставляет пациентов употреблять в большом количестве служебные слова и местоимения, но при этом процент использования знаменательных слов крайне мал [25]. Служебные слова соединяют знаменательные

слова в высказывании. Но это не единственная их функция. Это отсылочные слова, которые имеют колоссальное социологическое и психологическое значение [5].

Автор особо выделяет категорию личных местоимений. Являясь маркерами личностной идентичности (я vs мы), они указывают на локус контроля и несут ценную информацию о личности [5].

Как утверждает сам Дж. Пеннебакер, интересно было бы сравнить результаты, полученные для английского языка, с данными других языков. Кроме того, он рассматривает разные языковые явления (личные, указательные местоимения, временные формы глаголов) вне единой лингвистической концепции.

На наш взгляд, в основе такой концепции могла бы лежать категория дейксиса, который, как известно, делится на три вида: 1) личный – соотнесение участников обозначаемой ситуации с участниками акта речи, выражаемый личными местоимениями и формами лица глагола; ср. *я читаю, ты читаешь, мы читаем, он читает* и т.д.; 2) пространственный – характеристика предмета по степени его пространственной близости к говорящему, выражаемый указательными местоимениями *тот, этот*, частицами *вот, вон*, местоименными наречиями *здесь, там, сюда, туда, отсюда, оттуда*; 3) временной – соотнесение времени обозначаемой ситуации с моментом речи, выражаемый формами глагольного времени (ср. *читал, читаю, буду читать*) и обстоятельственными выражениями типа *сейчас, тогда, завтра, на прошлой неделе, будущей зимой* и т.п. [26. С. 543].

На наш взгляд, рассмотрение дейксиса как средства выявления личностных особенностей автора письменного текста имеет большие перспективы как в научном, так и в практическом плане.

Таким образом, в лингвистике и психологии считается на сегодняшний день неоспоримым то, в речи и языке отражается личность человека, что делает возможным установление характеристик (профилирование) автора по созданному им тексту. Определенные результаты в этом направлении уже получены. Особенно перспективным в настоящее время считается анализ дейктических единиц в целях установления характеристик автора текста: такие слова не зависят от контекста, являются самыми частотными, и методики, основанные на их анализе, могут быть универсальными.

Однако большинство работ, посвященных проблеме выявления личностных характеристик автора текста, выполнено на материале английского языка.

Нам представляется весьма перспективным рассмотреть на материале русского языка возможность выявления личностных характеристик автора текста путем анализа служебных слов и – шире – путем анализа ролевого и пространственно-временного дейксиса.

Литература

1. **Фомина Н.А.** Свойства личности и особенности речевой деятельности. Рязань : Узоречье, 2002. 362 с.
2. **Караулов Ю.Н.** Русский язык и языковая личность. 7-е изд. М. : ЛКИ, 2010. 264 с.
3. **Горошко Е.И.** Судебно-автороведческая классификационная экспертиза // Теорія та практика судової експертизи і криміналістики : сб. наук.-практ. матеріалів конф. Харків : Право, 2003. Вип. 3. С. 105–112.
4. **Галяшина Е.И.** Основы судебного речеведения. М. : СТЭНСИ, 2003. 236 с.
5. **Pennebaker J.W.** Secret life of pronouns: what our words say about us. N.Y. : Blumberry Press, 2011.
6. **Ароцкер Л.Е., Войнов В.К.** Об использовании лингвистической статистики для установления автора анонимного текста // Криминалистика и судебная экспертиза. 1961. Вип. 3. С. 141–151.
7. **Вул С.М.** Предмет и объекты судебно-автороведческой экспертизы // Криминалистика и судебная экспертиза. 1980. Вип. 20. С. 110–113.
8. **Енгальцев В.Ф., Белянин В.П., Константинова Е.С., Ощепкова Е.С.** Отражение личностных особенностей автора в письменном тексте // Труды регионального конкурса научных проектов в области гуманитарных наук. Калуга : Эйдос, 2000. Вип. 1. С. 205–217.
9. **Литвинова Т.А.** Лексемы, характеризующие человека по отношению к труду, в воронежских говорах // Вестник Челябинского государственного университета. Филология. Искусствоведение. 2011. Вип. 52. С. 84–89.
10. **Ощепкова Е.С.** Идентификация пола автора по письменному тексту (Лексико-грамматический аспект) : дис. ... канд. филол. наук. М., 2003. 192 с.
11. **Mulac A., Bradac J.J., Gibbons P.** Empirical support for the gender-as-culture hypothesis: an intercultural analysis of male / female language differences // Hum. Commun. Res. 2001. № 27. P. 121–152.
12. **Mehl M.R., Pennebaker J.W.** The social dynamics of a cultural upheaval: Social interactions surrounding September 11, 2001 // Psychological Science. 2002. № 14. P. 579–585.
13. **Pennebaker J.W., Stone L.D.** Words of Wisdom: Language Use Across the Lifespan // Journal of Personality and Social Psychology. 2002. № 85. P. 291–301.
14. **Белянин В.П.** Авторский смысл эмоционально отмеченного текста и проблема реконструкции личности автора // Теория и практика судебной экспертизы. 2008. № 4 (12). С. 181–184.
15. **Белянин В.П.** Психолингвистический анализ текста в целях идентификации личности // Актуальное состояние и перспективы развития судебной психологии в Российской Федерации : материалы Всерос. науч.-практ. конф. Калуга : КГУ им. К.Э. Циолковского, 2010. С. 67–73.
16. **Кулишова Н.Д.** Языковая личность в аспекте психолингвистических характеристик : автореф. дис... канд. филол. наук. Краснодар : Кубан. гос. ун-т, 2001. 24 с.
17. **Sanford F.H.** Speech and personality // Psychol. Bull. 1942. № 39. P. 811–45.
18. **Gortner E.M., Pennebaker J.W.** The anatomy of a disaster: Media coverage and community-wide health effects of the Texas A&M bonfire tragedy // Journal of Social and Clinical Psychology. 2002. № 22. P. 580–603.
19. **Argamon S., Dawhle S., Koppel M., Pennebaker J.** Lexical Predictors of Personality Type. In Proceedings of Classification Society of North America, St. Louis MI, June 2005.
20. **Vrij A.** Detecting Lies and Deceit: The Psychology of Lying and the Implications for Professional Practice. Chichester, UK : Wiley, 2000.
21. **Feldman Barrett L., Williams N.L., Fong G.T.** Defensive verbal behavior assessment // Personal. Soc. Psychol. Bull. 2002. № 28. P. 776–788.

22. *Schutz A., Baumeister R.F.* The language of defense: linguistic patterns in narratives of transgression // J. Lang. Soc. Psychol. 1999. № 18. P. 269–286.
23. *Shapiro D.* Psychotherapy of Neurotic Character. N.Y. : Basic Books, 1989.
24. *Newman M.L., Pennebaker W.J., Berry D.S., Richards J.M.* Lying words: predicting deception from linguistic styles // Personal. Soc. Psychol. Bull. 2002.
25. *Miller G.* The Science of Words. N.Y. : Sci. Am. Library, 1995.
26. *Росси́йский* гуманитарный энциклопедический словарь : в 3 т. М. : ВЛАДОС: Филол. фак-т С.-Петерб. гос. ун-та, 2002. Т. 1. 688 с.

WRITTEN TEXT AUTHOR PROFILING

Litvinova T.A.

Summary. At present, there is no doubt that a text reflects its authors personality. It is found out that one of the most efficient ways to reveal authors personality in the text is the analysis of function words, in particular, personal pronouns, prepositions, and conjunctions. We propose to apply function words related techniques for author personality profiling to Russian, as well as to consider the possibility of analysis of a deixis as a personal predictor.

Key words: text; author; function words; language personality; author profiling.