

*Исследование* влияния пола и психологических характеристик автора на количественные параметры его текста с использованием программы *Linguistic Inquiry and Word Count* / Т. А. Литвинова, О. А. Литвинова, Е. С. Рыжкова, Е. Д. Бирюкова, П. В. Середин, О. В. Загоровская // Научный диалог. — 2015. — № 12 (48). — С. 101—109.



УДК 81'33+81'322.5

## **Исследование влияния пола и психологических характеристик автора на количественные параметры его текста с использованием программы *Linguistic Inquiry and Word Count*\***

© **Литвинова Татьяна Александровна (2015)**, кандидат филологических наук, научный сотрудник Регионального центра русского языка, Воронежский государственный педагогический университет (Воронеж, Россия), [centr\\_rus\\_yaz@mail.ru](mailto:centr_rus_yaz@mail.ru).

© **Литвинова Ольга Александровна (2015)**, ассистент кафедры английского языка, Воронежский государственный педагогический университет (Воронеж, Россия), [olga\\_litvinova\\_teacher@mail.ru](mailto:olga_litvinova_teacher@mail.ru).

© **Рыжкова Екатерина Сергеевна (2015)**, аспирант кафедры русского языка, современной русской и зарубежной литературы, Воронежский государственный педагогический университет (Воронеж, Россия), [ryzhkova.katerina@yandex.ru](mailto:ryzhkova.katerina@yandex.ru).

© **Бирюкова Елизавета Дмитриевна (2015)**, аспирант кафедры русского языка, современной русской и зарубежной литературы, Воронежский государственный педагогический университет (Воронеж, Россия), [olzagor@yandex.ru](mailto:olzagor@yandex.ru).

© **Середин Павел Владимирович (2015)**, доктор физико-математических наук, старший научный сотрудник, Воронежский государственный университет (Воронеж, Россия), [paul@phys.vsu.ru](mailto:paul@phys.vsu.ru).

© **Загоровская Ольга Владимировна (2015)**, доктор филологических наук, профессор, заведующий кафедрой русского языка, современной русской и зарубежной литературы, Воронежский государственный педагогический университет (Воронеж, Россия), [olzagor@yandex.ru](mailto:olzagor@yandex.ru).

Рассматриваются вопросы диагностирования психологических и иных характеристик личности автора письменного текста. Авторы берут за основу тезис о том, что в тексте на разных его уровнях отражаются характеристики его автора (пол, возраст, психологические особенности и т. д.). Приводятся результаты пилотного исследования текстов на русском языке с использованием программы *Linguistic Inquiry and Word Count* (LIWC). С ее помощью выявляются взаимосвязи между характеристиками личности автора и количественными параметрами текста (доля слов некоторых частей речи, лексико-тематических групп, частота знаков препинания и т. д.). Актуальность и новизна исследования обусловлены тем, что на русскоязычных текстах программа ранее не применялась. Исследование выполнено на материале специального корпуса текстов *Personality*, снабженного метаразмечкой, содержащей информацию об их авторах. Авторы считают необходимым использо-

\* Исследование выполнено при поддержке гранта РГНФ 15-34-01221 «Детекция лжи в письменном тексте: корпусное исследование».

вать для выявления языковых коррелятов тех или иных устойчивых психологических характеристик текста только стабильные параметры текста, то есть сохраняющие устойчивость в разных текстах одного и того же автора. Отмечается, что в целом результаты исследования согласуются с данными, полученными с использованием этой программы на материале английского языка.

Ключевые слова: письменный текст; математические методы в лингвистике; моделирование личности по тексту; автороведение; компьютерная лингвистика; Linguistic Inquiry and Word Count; LIWC.

## 1. Введение

В современной науке является общепринятым положение о том, что текст несет информацию о личности его автора, однако нет единого мнения по поводу того, какие именно уровни текста наиболее информативны в этом отношении.

В настоящее время в мировой науке наблюдается всплеск интереса к исследованиям проблемы атрибуции (установления авторства) и диагностирования личностных характеристик личности автора письменного текста, что во многом обусловлено активным развитием интернет-коммуникации и практической необходимостью получения информации об авторах интернет-текстов в криминалистике, маркетинге и пр. Этой тематике посвящены тысячи научных публикаций, проводятся конкурсы на самые точные методики атрибуции текста и диагностирования характеристик личности его автора [PAN 2015]. В большинстве зарубежных исследований (выполненных преимущественно на материале английского языка) проблемы атрибуции текста и диагностирования личности по ее речевым произведениям изучаются на большом корпусном материале — образцах естественной письменной речи (в том числе текстах интернет-коммуникации), которые анализируются с использованием различных программных средств (морфологических и синтаксических парсеров, программ для подсчета лексического разнообразия текста и т. д.) (см., например: [Gender..., 2003; Automatically profiling..., 2009; Author profiling..., 2007; Rangel et al., 2013 и мн. др.; см. краткий обзор: Литвинова, 2013а]). В дальнейшем, анализируя эти данные, ученые с применением методов математической статистики ищут корреляции между параметрами текста и разнообразными характеристиками его автора и строят математические модели, в которых входными параметрами служат параметры текста, а выходными — характеристики личности. Подобный подход к решению задачи диагностирования пола и психологических особенностей автора текста на русском языке был применен нами в предыдущих работах [Литвинова, 2013б; Литвинова и др., 2014; Диагностирование..., 2015]. На материале специального созданного корпуса текстов *Personality*, содержащего метаразметку в виде информации об их авторах (пол, возраст, баллы по психологическим тестам и т. д.) [Корпусные исследования..., 2015], нами был выявлен ряд корреляций между формально-грамматическими параметрами текста и характеристиками личности его автора, и на основе этих корреляций построены математические модели, показавшие достаточно высокую точность определения той или иной характеристики автора.

Во многих зарубежных работах, посвященных моделированию личности по письменной речевой продукции, для анализа текстов используется компьютерная программа *Linguistic Inquiry and Word Count* [The Development...] (обзор исследований, выполненных с использованием данной программы, см. в работе [Tausczik et al., 2010]).

Программа *Linguistic Inquiry and Word Count* подсчитывает в тексте доли (что позволяет избежать зависимости от объема текста) слов некоторых частей речи (местоимений, наречий, предлогов, союзов), слов длиннее 6 букв, частотности знаков препинания, доли в тексте слов тех или иных лексико-семантических категорий (слов, обозначающих положительные и отрицательные эмоции, восприятие, когнитивные процессы и т. д.). Разработаны версии программы для английского, датского, китайского, арабского, французского, немецкого, итальянского, португальского, сербского, испанского, турецкого, русского языков. Для анализа текстов на русском языке программа впервые была использована нами – для выявления стабильных характеристик идиостиля [Литвинова, 2015], установления статистически значимых различий между «ложными» и «правдивыми» текстами [Литвинова и др., 2015]. Для выявления корреляций между параметрами текста на русском языке и устойчивыми психологическими характеристиками личности его автора данная программа ранее не использовалась, чем и определяется научная новизна настоящего исследования.

## 2. Материал и методы

В качестве материала исследования использовался созданный нами корпус текстов *Personality* [Корпусные исследования..., 2015]. Корпус содержит как образцы естественной письменной речи, так и данные об их авторах — пол, возраст, образование, результаты психологического тестирования и т.д.

В настоящей работе нами использовались два подкорпуса указанного корпуса текстов: первый (назовем его K1) содержит 150 текстов 75 респондентов одного возраста и результаты тестирования авторов по пятифакторному личностному опроснику, второй (K2) — 958 текстов от 479 респондентов одного возраста и результаты тестирования респондентов по Фрайбургскому многофакторному личностному опроснику FPI.

Для анализа текстов нами была использована версия программы *Linguistic Inquiry and Word Count* (LIWC) 2007 года со встроенным словарем для русского языка, который был получен в результате перевода словаря, составленного для английского языка.

В качестве параметров текста для дальнейших расчетов нами были выбраны только те вычисленные LIWC параметры текста, которые показали стабильность в наших предыдущих исследованиях [Литвинова, 2015]. Насколько нам известно, такой подход к решению проблемы моделирования личности автора текста используется впервые. Мы считаем необходимым использовать для выявления языковых коррелятов тех или иных устойчивых психологиче-

ских характеристик текста только стабильные параметры текста, то есть сохраняющие устойчивость в разных текстах одного и того же автора.

В указанной работе [Литвинова, 2015], посвященной определению степени стабильности тех или иных параметров в текстах одного автора, с помощью программы *LWSC* нами было проанализировано 1188 текстов от 594 респондентов (по два текста от каждого автора) из корпуса *Personality*. Для определения устойчивых параметров в разных текстах одного и того же автора мы рассчитали среднее абсолютное отклонение значений всех параметров от их средней величины для конкретного автора. Далее мы определили коэффициент вариации каждого параметра для всех авторов корпуса текстов, что позволило нам оценить степень разбросанности значений параметров текста и понять, насколько она велика относительно их среднего значения. Статистический анализ показал, что рассчитанный коэффициент вариации для выбранных параметров текста лежит в достаточно широких пределах. Принимая во внимание, что значение коэффициента вариации менее 33 % свидетельствует об однородности совокупности данных, можно заключить, что в нашем случае это означает устойчивость параметра в текстах одного автора.

В группе устойчивых параметров нами были выделены две подгруппы: 1) подгруппа параметров, имеющих низкий коэффициент вариации (до 17 %) (*линейно-стабильные параметры*); 2) подгруппа параметров с более высоким коэффициентом вариации (от 17 до 33 %) (*рекуррентно-стабильные*).

К *подгруппе линейно-стабильных* в проанализированных нами текстах относятся следующие параметры: 1) доля слов длиннее 6 букв; 2) доля строевых слов; 3) доля предлогов; 4) доля слов группы «Мыслительные процессы»; 5) доля слов группы «Пространство»; 6) доля знаков препинания.

К *подгруппе рекуррентно-стабильных параметров* относятся следующие параметры: 1) доля местоимений; 2) доля личных местоимений; 3) доля союзов; 4) доля слов группы «Эмоции»; 5) доля слов группы «Положительные эмоции»; 6) доля слов группы «Время».

Именно эти параметры и использовались нами для расчета корреляций с характеристиками личности авторов (в виде баллов по Пятифакторному личностному опроснику и Фрайбургскому многофакторному личностному опроснику FPI).

### 3. Результаты и обсуждение

Для первого исследовательского корпуса (K1) посредством корреляционного анализа с использованием критерия Пирсона при тесноте связи  $p = 0,05$  были установлены корреляции для пола автора с такими параметрами, как доля в тексте строевых слов (служебных слов и местоимений) ( $r = -0,258$ ,  $p = 0,0286$ ), доля личных местоимений ( $r = -0,251$ ,  $p = 0,0333$ ), союзов ( $r = -0,357$ ,  $p = 0,00208$ ), доля слов, обозначающих мыслительные процессы ( $r = -0,368$ ,  $p = 0,00145$ ). У женщин все эти показатели в среднем выше, чем

у мужчин, что согласуется с результатами, полученными на материале английского языка [Nini, 2014; Tausczik et al., 2010].

Из психологических характеристик, измеряемых при помощи Пятифакторного личностного опросника (теста, наиболее часто применяемого в зарубежных работах по моделированию личности на основе анализа текста), были найдены корреляции для характеристик «**Экстраверсия**» (доля слов длиннее 6 букв,  $r = -0,230$ ,  $p = 0,0469$ ; более высокие значения этого параметра характерны для интровертов), «**Доброжелательность**» (доля слов длиннее 6 букв,  $r = -0,222$ ,  $p = 0,05$ ; более высокие значения этого показателя характерны для людей с низкими показателями по этой шкале, то есть для людей, стремящихся к независимости, обособленности, малообщительных), «**Добросовестность**» (доля слов, обозначающих эмоции,  $r = 0,233$ ,  $p = 0,459$ ; более высокие значения этого показателя характерны для людей с низкими показателями по этой шкале, то есть для людей, характеризующихся рациональным, практичным подходом к жизни).

Для второго подкорпуса K2 были получены следующие корреляции. Для **пола** были выявлены корреляции с долей местоимений в целом ( $r = -0,258$ ,  $p = 0,00000001$ ) и личных местоимений в частности ( $r = -0,246$ ,  $p = 0,00000005$ ) (для женщин характерны более высокие значения этих показателей).

Из всех характеристик, которые измеряются при помощи Фрайбургского многофакторного личностного опросника, были найдены корреляции параметров текстов только с уровнем **невротичности** (доля местоимений,  $r = 0,131$ ,  $p = 0,0039$ ; доля личных местоимений,  $r = 0,138$ ,  $p = 0,002$ ; более высокие показатели характерны для лиц с более высокими значениями по этой шкале), **депрессивности** (доля местоимений,  $r = 0,127$ ,  $p = 0,005$ ; доля союзов,  $r = 0,141$ ,  $p = 0,00194$ ; более высокие показатели характерны для лиц с более высокими значениями по этой шкале), **эмоциональной лабильности** (доля местоимений,  $r = 0,157$ ,  $p = 0,0005$ ; доля личных местоимений,  $r = 0,137$ ;  $p = 0,0026$ , доля союзов,  $r = 0,147$ ,  $p = 0,001$ ; для текстов эмоционально лабильных авторов характерны более высокие значения этих параметров; доля слов тематической группы «Пространство»,  $r = -0,149$ ;  $p = 0,00120$ ; для эмоционально лабильных авторов характерны низкие показатели по этому параметру).

Полученные результаты во многом согласуются с данными, полученными на материале других языков (преимущественно английского) [Nini, 2014; Tausczik et al., 2010]. Заметим, однако, что сам по себе факт корреляционной зависимости не даёт основания утверждать, что одна из переменных предшествует или является причиной изменений, или то, что переменные вообще причинно связаны между собой, а не наблюдается действие третьего фактора [Елисеева и др., 2002, с. 229]. Так, корреляции параметра «доля местоимений» с такими характеристиками автора, как пол, уровень депрессивности, невротичности, эмоциональной лабильности, могут иметь в основе разные причины. Связь уровня личных местоимений и пола может быть опосредованной тем фактом,

что среди женщин в среднем выше уровень депрессивности, невротичности, эмоциональной лабильности, что в свою очередь может быть связано с особенностями гормонального статуса [Nini, 2015, p. 34]. Заметим, что поиск причин устойчивых корреляций между параметрами личности и текста — отдельная сложная научная проблема (см. подробнее: [Диагностирование..., 2015]).

#### 4. Выводы

Таким образом, впервые для текстов на русском языке с использованием программы LIWC, широко применяющейся в зарубежных исследованиях проблемы моделирования личности на основе анализа текста, на обширном корпусном материале был выявлен ряд корреляций между полом, психологическими характеристиками автора и параметрами его текста. Установлено, что в целом эти корреляции согласуются с результатами, полученными с использованием этой программы на материале английского языка. Однако выявленных корреляций было недостаточно для построения регрессионных моделей и диагностирования личностных характеристик автора на основе количественных параметров текста, как это было сделано в наших предыдущих работах [Литвинова, 2013б; Литвинова и др., 2014; Диагностирование..., 2015].

На наш взгляд, относительно небольшое число выявленных корреляций с использованием программы LIWC объясняется следующим:

1. Впервые в исследованиях проблемы моделирования личности по тексту в качестве параметров отбирались только продемонстрировавшие устойчивость в текстах одного автора.
2. Поскольку русскоязычный словарь программы LIWC был создан путем перевода английского словаря, требуется его независимая оценка и валидизация с привлечением русскоязычных респондентов.

#### Литература

1. *Елисеева И. И.* Общая теория статистики / И. И. Елисеева, М. М. Юзбашев ; под ред. И. И. Елисеевой. — 4-е изд., перераб. и доп. — Москва : Финансы и Статистика, 2002. — 480 с.
2. *Диагностирование* склонности автора письменного текста к аутоагрессивному поведению / Т. А. Литвинова, П. В. Середин, О. А. Литвинова, О. В. Загоровская, М. Е. Сердюк // Вестник Воронежского государственного университета. Серия: Лингвистика и межкультурная коммуникация. — 2015. — № 3. — С. 98—104.
3. *Корпусные* исследования письменной речи в решении задач судебного автороведения / Т. А. Литвинова, Е. В. Диброва, О. А. Литвинова, Е. С. Рыжкова // Филологические науки. Вопросы теории и практики. — 2015. — № 8. — Ч. 1. — С. 107—113.
4. *Литвинова Т. А.* Исследование лингвистических характеристик текстов, содержащих намеренно искаженную информацию, с помощью программы Linguistic Inquiry and Word Count / Т. А. Литвинова, О. А. Литвинова // Вестник МГОУ. Серия: Лингвистика. — 2015. — № 4. — С. 71—77.
5. *Литвинова Т. А.* К проблеме стабильности характеристик идиостиля / Т. А. Литвинова // Известия ЮФУ. Филология. — 2015. — № 3. — С. 98—106.

6. Литвинова Т. А. Профилирование автора письменного текста идиостиля / Т. А. Литвинова // Язык и культура. — 2013а. — № 3 (23). — С. 64—72.
7. Литвинова Т. А. Формально-грамматические корреляты личностных особенностей автора письменного текста / Т. А. Литвинова // Филологические науки. Вопросы теории и практики. — 2013б. — № 12 (30). — Ч. 1. — С. 132—135.
8. Литвинова Т. А. Частоты встречаемости последовательностей частей речи в тексте и психофизиологические характеристики его автора: корпусное исследование / Т. А. Литвинова, О. А. Литвинова, П. В. Середин // Вестник Иркутского государственного лингвистического университета. — 2014. — № 2. — С. 8—12.
9. Author profiling for English emails / D. Estival, T. Gaustad, S. B. Pham, W. Radford, B. Hutchinson // Proceedings of the 10<sup>th</sup> Conference of the Pacific Association for Computational Linguistics. — 2007. — Pp. 263—272.
10. Automatically profiling the author of an anonymous text / S. Argamon, M. Koppel, J. W. Pennebaker, J. Schler // Commun. ACM. — 2009. — Vol. 52(2). — Pp. 119—123.
11. Gender, genre, and writing style in formal written texts / S. Argamon, M. Koppel, J. Fine, A. Shimoni // Text. — 2003. — Vol. 23(3). — Pp. 321—346.
12. Nini A. Authorship profiling in a forensic context : PhD thesis / A. Nini. — Birmingham, United Kingdom: Aston University, 2014. — 250 p.
13. PAN 2015 [Electronic resource]. Access mode : <http://www.uni-weimar.de/medien/webis/events/pan-15/pan15-web/author-profiling.html>.
14. Rangel F. Use of Language and Author Profiling : Identification of Gender and Age [Electronic resource] / F. Rangel, P. Rosso // Natural Language Processing and Cognitive Science: Proceedings of NLPCS 2013 10th International Workshop on Natural Language Processing and Cognitive Science (Marseille, France, October, 2013). Access mode : <http://www.kicorangel.com/wp-content/uploads/2013/10/NLPCS2013-proceedings.pdf#page=179>.
15. Tausczik Y. R. The Psychological Meaning of Words : LIWC and Computerized Text Analysis Method / Y. R. Tausczik, J. W. Pennebaker // Journal of Language and Social Psychology. — 2010. — Vol. 29. — Issue 1. — Pp. 24—54.
16. The Development and Psychometric Properties of LIWC2007 [Electronic resource] / J. W. Pennebaker, C. K. Chung, M. Ireland, A. Gonzales, R. J. Booth. — Access mode : [http://homepage.psy.utexas.edu/homepage/faculty/pennebaker/reprints/liwc2007\\_languagemanual.pdf](http://homepage.psy.utexas.edu/homepage/faculty/pennebaker/reprints/liwc2007_languagemanual.pdf).

---

### Studying Influence of Author's Gender and Psychological Characteristics on Quantitative Parameters of Text Using "Linguistic Inquiry and Word Count" Program

© Litvinova Tatyana Aleksandrovna (2015), PhD in Philology, research scientist, Regional Centre of Russian Language, Voronezh State Pedagogical University (Voronezh, Russia), [centr\\_rus\\_yaz@mail.ru](mailto:centr_rus_yaz@mail.ru).

© Litvinova Olga Aleksandrovna (2015), assistant lecturer, Department of English Language, Voronezh State Pedagogical University (Voronezh, Russia), [olga\\_litvinova\\_teacher@mail.ru](mailto:olga_litvinova_teacher@mail.ru).

© Ryzhkova Yekaterina Sergeevna (2015), post-graduate student, Department of Russian Language, Modern Russian and Foreign Literature, Voronezh State Pedagogical University (Voronezh, Russia), [ryzhkova.katerina@yandex.ru](mailto:ryzhkova.katerina@yandex.ru).

© Biryukova Yelizaveta Dmitriyevna (2015), post-graduate student, Department of Russian Language, Modern Russian and Foreign Literature, Voronezh State Pedagogical University (Voronezh, Russia), [olzagor@yandex.ru](mailto:olzagor@yandex.ru).

- © **Seredin Pavel Vladimirovich (2015)**, Doctor of Physics and Mathematics, senior research scientist, Voronezh State Pedagogical University (Voronezh, Russia), paul@phys.vsu.ru.  
 © **Zagorovskaya Olga Vladimirovna (2015)**, Doctor of Philology, professor, Head of Department, Department of Russian Language, Modern Russian and Foreign Literature, Voronezh State Pedagogical University (Voronezh, Russia), olzagor@yandex.ru.

The problems of diagnosis of psychological and other personality characteristics of the author of the written text are studied. The authors take as a basis the idea that the text at its various levels reflects the characteristics of the author (gender, age, psychological characteristics, etc.). The results of the pilot study of Russian texts using *Linguistic Inquiry and Word Count* (LIWC) program. It reveals the links between the personality characteristics of the author and quantitative parameters of the text (number of the words of some of the parts of speech, of some lexical-thematic groups, the frequency of punctuation marks, etc.). The relevance and novelty of the research lies in the fact that the program has never been used with the Russian-language texts. The study is based on a special corpus of texts *Personality*, equipped with meta-marking that contains information about the authors. The authors consider it necessary to use only stable parameters of the text, that is, preserving stability in the different texts of the same author, to identify linguistic correlates of those or other stable psychological characteristics of the text. It is noted that, in general, the results of the study are consistent with the data obtained using this program on the English language.

Key words: written text; mathematic methods in linguistics; modeling personality in the text; studying of the author; computational linguistics; Linguistic Inquiry and Word Count; LIWC.

## References

- Argamon, S., Koppel, M., Fine, J., Shimoni, A. 2003. Gender, genre, and writing style in formal written texts. *Text*, 23(3): 321—346.
- Argamon, S., Koppel, M., Pennebaker, J. W., Schler, J. 2009. Automatically profiling the author of an anonymous text. *Commun. ACM*, 52(2): 119—123.
- Eliseeva, I. I., Yuzbashev, M. M. 2002. *Obshchaya teoriya statistiki*. Moskva: Finansy i Statistika. 480. (In Russ.).
- Estival, D., Gaustad, T., Pham, S. B., Radford, W., Hutchinson, B. 2007. Author profiling for English emails. In: *Proceedings of the 10<sup>th</sup> Conference of the Pacific Association for Computational Linguistics*. 263—272.
- Litvinova, T. A. 2013a. Profilirovaniye avtora pis'mennogo teksta idiostilya. *Yazyk i kultura*, 3 (23): 64—72. (In Russ.).
- Litvinova, T. A. 2013b. Formalno-grammaticheskiye korrelyaty lichnostnykh osobennostey avtora pis'mennogo teksta. *Filologicheskkiye nauki. Voprosy teorii i praktiki*, 12 (30)/1: 132—135. (In Russ.).
- Litvinova, T. A. 2015. K probleme stabilnosti kharakteristik idiostilya. *Izvestiya YuFU. Filologiya*, 3: 98—106. (In Russ.).
- Litvinova, T. A., Dibrova, E. V., Litvinova, O. A., Ryzhkova, E. S. 2015v. Korpusnyye issledovaniya pis'mennoy rechi v reshenii zadach sudebnogo avtorovedeniya. *Filologicheskkiye nauki. Voprosy teorii i praktiki*, 8(1): 107—113. (In Russ.).
- Litvinova, T. A., Litvinova, O. A. 2015b. Issledovaniye lingvisticheskikh kharakteristik tekstov, soderzhashchikh namerenno iskazhennuyu informatsiyu, s pomoshchyu programmy Linguistic Inquiry and Word Count. *Vestnik MGOU. Seriya: Lingvistika*, 4: 71—77. (In Russ.).
- Litvinova, T. A., Litvinova, O. A., Seredin, P. V. 2014. Chastoty vstrechaemosti posledovatel'nostey chastey rechi v tekste i psikhofiziologicheskkiye kharakteristiki yego av-



- tora: korpusnoye issledovaniye. *Vestnik Irkutskogo gosudarstvennogo lingvističeskogo universiteta*, 2: 8—12. (In Russ.).
- Litvinova, T. A., Seredin, P. V., Litvinova, O. A., Zagorovskaya, O. V., Serdyuk, M. E. 2015a. Diagnostirovaniye sklonnosti avtora pis'mennogo teksta k autoagressivnomu povedeniyu. *Vestnik Voronezhskogo gosudarstvennogo universiteta. Seriya: Lingvistika i mezhkulturnaya kommunikatsiya*, 3: 98—104. (In Russ.).
- Nini A. *Authorship profiling in a forensic context*: PhD thesis. — Birmingham, United Kingdom: Aston University, 2014. — 250 p.
- PAN 2015. Available at: <http://www.uni-weimar.de/medien/webis/events/pan-15/pan15-web/author-profiling.html>.
- Pennebaker, J. W., Chung, C. K., Ireland, M., Gonzales, A., Booth, R. J. *The Development and Psychometric Properties of LIWC2007* / The University of Texas at Austin and The University of Auckland, New Zealand. Available at: [http://homepage.psy.utexas.edu/homepage/faculty/pennebaker/reprints/liwc2007\\_languagemanual.pdf](http://homepage.psy.utexas.edu/homepage/faculty/pennebaker/reprints/liwc2007_languagemanual.pdf).
- Rangel, F., Rosso, P. Use of Language and Author Profiling : Identification of Gender and Age. In: *Natural Language Processing and Cognitive Science: Proceedings of NLPCS 2013 10th International Workshop on Natural Language Processing and Cognitive Science (Marseille, France, October, 2013)*. Available at: <http://www.kicorangel.com/wp-content/uploads/2013/10/NLPCS2013-proceedings.pdf#page=179>.
- Tausczik, Y. R., Pennebaker, J. W. 2010. The Psychological Meaning of Words: LIWC and Computerized Text Analysis Method. *Journal of Language and Social Psychology*, 29(1): 24—54.