Министерство образования и науки Российской Федерации

Федеральное государственное бюджетное образовательное учреждение

высшего образования

**«Московский государственный лингвистический университет»**

**(ФГБОУ ВО МГЛУ)**

Аннотация

к выпускной квалификационной работе

**Карповой Екатерины Константиновны**

Факультет английского языка

Направление подготовки - 45.03.02 - Лингвистика

Направленность (профиль) - Теоретическая и прикладная лингвистика

Квалификация - бакалавр

Группа 5-8-31

На тему:

**Установление личности автора в Интернете по лингвистическим признакам**

**Ключевые слова:** идентификация личности, экспертный анализ письменного текста, идентификационные признаки, «портрет» автора, автороведческая экспертиза.

Данная работа включает теоретический обзор литературы и практическую часть по проблеме идентификации личности автора письменного текста по лингвистическим признакам.

В настоящее время проблема идентификации личности автора в Интернете, а именно в социальных сетях очень актуальна, особенно в области криминалистики. Из-за того, что сегодня практически каждый человек пользуется интернетом, оставляет там письменные сообщения, данная дисциплина продолжает активно развиваться, и чем больше людей пользуются социальными сетями, тем более широкое применение она находит в жизни людей. Именно поэтому у специалистов данной области постоянно появляются новые и новые вопросы связанные с возможностями идентификации по лингвистическим признакам.

Автороведческая экспертиза, которая считается сложным и трудоемким процессом, становится все более популярной, так как все чаще в социальных сетях нарушаются авторские права, появляются призывы к экстремистской деятельности или происходит унижение человеческого достоинства, что всё является статьями УК РФ. Большое внимание уделяется не только проблеме идентификации личности автора, но и анализу самого лингвистического материала. Разговорная и письменная речь является продуктом сложной полифункциональной человеческой деятельности, системы, которая включает в себя все компоненты психофизической, функционально-двигательной и ментально-когнитивно-интеллектуальной подсистем относительно научных аспектов биологии, психологии, физиологии, биомеханики, антропофоники, семиотики, когнитологии, когитологии и социопсихологии [Potapova R. 2015][[1]](#footnote-2). Объектом исследования служит письменная речь, проявляющиеся в ней языковые навыки, которые подразделяются на общие (лексико-фразеологические, синтаксические, стилистические, орфографические, пунктуационные) и частные (устойчивые языковые нарушения речи, использование определенных языковых средств и прочее.

Автороведческая экспертиза решает идентификационные задачи: является ли определенное лицо автором текста определенного документа, соответствующего фрагмента документа, нескольких документов; классификационные и диагностические задачи: является ли язык исследуемого документа родным для его автора; носителем какого наречия (говора) определенного языка является автор конкретного документа; каков образовательный уровень автора документа; обладает ли автор исследуемого документа навыками научного (публицистического, делового) стиля письменной речи; составлен ли текст конкретного документа с намеренным искажением письменной речи и др.[Бабаева Э.У.,2000]

Известно, что уровень лингвистических знаний у каждого человека разный. Он зависит от того, насколько начитан человек, от его образования, профессии, окружающего общества и так далее. Так или иначе всё это влияет на письменную речь.

Данная область знаний нашла своё применение в криминалистике еще в конце XX века. И по сей день эксперты пытаются решить проблемы и задачи, связанные с этим разделом прикладной лингвистики/криминалистики. К числу активно разрабатываемых особенно относят интернет-коммуникации, так как в чатах существует новая форма языкового взаимодействия, формально являющаяся письменной, но, по сути, это обычная устная речь.

Главная **цель** работы – наметить возможные признаки, достаточные для идентификации автора письменных сообщений в интернете, что непосредственно может быть связано с реконструкцией «портрета»[Потапова 2006][[2]](#footnote-3) автора данного текста.

Достижение цели требует выполнения многих **задач**, в том числе:

1) анализ литературы по соответствующей теме;

2) формирование базы данных текстов для проведения эксперимента;

3) анализ текста по лексическим признакам;

4) анализ тональности текста;

5) анализ текста по стилям речи;

6) анализ темпоральности текста;

7) интертекстуальность текста;

8) проведение эксперимента, с целью определения испытуемыми личности автора текста;

9) анализ полученных результатов;

10) перевод полученных результатов в относительные данные (в %);

**Методы исследования**

В данной работе использовалось 2 метода исследования:  
1. Анализ на материале текстов Интернета

2. Эксперимент с привлечением испытуемых

**Анализ на материале текстов Интернета**

В качестве материала для проведения эксперимента были использованы 250 текстов 25 разных авторов с сайта [**www.vk.com**](http://www.vk.com/). Данный сайт является социальной сетью «Вконтакте», где люди могут обмениваться текстовыми сообщениями и размещать свой контент. Все материалы взяты с публичных страниц и находятся в открытом доступе. Каждый автор текста свободно излагал мысли в сфере интересующих их различных аспектов жизни, искусства и т.д. Исследование, проведенное в рамках данной работы включает:

1. Подбор текстов (n=25)(В социальной сети Вконтакте было подобрано 250 текстов 25 разных авторов)
2. Разбитие текстов на тематические группы:
3. Книги - 1 текст, 1 автор
4. Дневник – 56 текстов, 19 авторов
5. Семья – 7 текстов, 6 авторов
6. Друзья – 2 текста, 1 автор
7. Здоровье - 13 текстов, 5 авторов
8. Отпуск – 5 текстов, 3 автора
9. Воспоминания – 21 текст, 12 авторов
10. Мнения – 82 текста, 20 авторов
11. Политика – 4 текста, 2 автора
12. Наука – 2 текста, 2 автора
13. Путешествия – 36 текстов, 9 авторов
14. Работа – 21 текст, 10 авторов
15. Разбор текста по частям речи, помимо этого части речи были поделены на 2 группы – знаменательные и служебные, далее полученные результаты были переведены в относительные данные (в %).
16. Анализ текста по лексическим признакам (по *С.М. Вулу*)[[3]](#footnote-4).
17. Анализ тональности текста, все слова были поделены на 3 группы: негативные, позитивные и нейтральные, результаты опять же были переведены в относительные данные (в %).
18. Стилистический анализ текста
19. Установление темпоральности текста, все глаголы в тексте были поделены на 3 временные группы: настоящее, прошедшее и будущее время, и, исходя из того, в какой группе больше глаголов, была выявлена темпоральность.
20. Установление интертекстуальности текста – а именно наличие или отсутствие в тексте 1) цитации, 2) пересказа, 3) фоновых ссылок на идеи, высказанных ранее.
21. Установление наличия/отсутствия инвективов в тексте.

Проанализировав **результаты**, полученные в ходе **обработки текстов**, удалось установить:

1. в целом, во всех текстах одного автора есть повторяющиеся элементы, в основном *лексические*, благодаря которым его можно идентифицировать
2. однако в большинстве текстов превалирует нейтральная тональность, разговорно-обиходный стиль и знаменательные части речи (в особенности существительные, местоимения, глаголы), что усложняет задачу идентификации личности, так как данные характеристики достаточно распространены.

**Эксперимент с привлечением испытуемых**

Также был проведен эксперимент, в котором участвовало 13 человек: мужчины и женщины, в возрасте от 21 до 44 лет. Испытуемым было предъявлено 92 текста 20ти авторов, на темы: дневник и путешествия. Для начала они должны были заполнить анкету о себе (см. Табл.11). Анкета[[4]](#footnote-5) включала в себя такие вопросы как:

1. имя,
2. возраст,
3. социальный статус,
4. род занятий/профессия,
5. родной язык,
6. владение иностранными языками,
7. билингвизм/трилингвизм и
8. хобби

Затем, прочитав тексты, испытуемые должны были приблизительно указать в анкете о авторе текста(см.Табл.12, Табл.13):

1. пол,
2. возраст,
3. социальный статус,
4. род занятий/профессию,
5. владение языками,
6. образование,
7. наличие билингвизма/трилингвизма,
8. хобби
9. происхождение автора текста, а также указать
10. психологический тип по Юнгу
11. темперамент

Испытуемыми были предоставлены результаты в таблицах, в которых они давали характеристику авторам текстов, а именно:

1. пол,
2. возраст,
3. социальный статус,
4. род занятий/профессию,
5. владение языками,
6. образование,
7. наличие билингвизма/трилингвизма,
8. хобби
9. происхождение автора текста, а также указать
10. психологический тип (по Юнгу)
11. темперамент

Если испытуемый не мог точно определить какой-либо из параметров, он пропускал его.

Проанализировав **результаты экспериментов**, можно сделать следующие выводы:

1. При помощи лексических признаков можно установить пол, социальный статус, происхождение и темперамент в 100% случаях.
2. Сложнее всего определить род занятий/профессию (в 75% случаев), хобби (в 85% случаев), психологически й тип (в 90% случаев)

**Заключение**

В данной работе было проведено исследование, целью которого было узнать, можно ли наметить возможные признаки для идентификации автора письменных сообщений в социальных сетях. В будущем предполагается разработка методики проведения эксперимента, увеличение количества лингвистических признаков, использование большего количества текстов различных категорий. Что же касается текущего исследования, то проведенный эксперимент является пилотным и предполагает дальнейшие более глубокие исследования разрабатываемой проблемы для получения более объективных результатов.

Студент КАРПОВА Е.К.

Согласовано:

Научный руководитель ПОТАПОВА Р. К.

д-р филол. наук, проф., зав. кафедрой

1. Potapova R. FROM DEPRIVATION TO AGGRESSION: VERBAL AND NON-VERBAL SOCIAL NETWORK COMMUNICATION - Global Science and Innovation November 18-19th, 2015. Vol. I, 2015. – c. 129-137. [↑](#footnote-ref-2)
2. *Потапова, Р. К., Потапов В.В.* Язык, речь, личность / Р. К. Потапова, В. В. Потапов. – М.: Языки славянской культуры, 2006. – 491 с. [↑](#footnote-ref-3)
3. По книге *Потапова, Р. К., Потапов В.В.* Язык, речь, личность / Р. К. Потапова, В. В. Потапов. – М.: Языки славянской культуры, 2006. – 491 с. [↑](#footnote-ref-4)
4. Из вопросника профессора, доктора наук, академика Потаповой Р.К. [↑](#footnote-ref-5)