

О ГЕНДЕРНО-МАРКИРОВАННЫХ ЕДИНИЦАХ В РУССКОЯЗЫЧНЫХ ТЕКСТАХ КАК СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИКАХ СЛУЖЕБНЫХ ЧАСТЕЙ РЕЧИ

Н. Д. Голев, Д. А. Реттих

ABOUT GENDER-MARKED UNITS IN RUSSIAN TEXTS AS STATISTIC CHARACTERISTICS OF LINK PARTS OF SPEECH

N. D. Golev, D. A. Rettikh

Статья посвящена проблеме наличия в русском языке гендерно-маркированных лексем. Ставится и обсуждается вопрос о степени такой маркированности у разных лексем, разрабатывается методика измерения лексических единиц в данном аспекте. Статья выполнена на материале служебные части речи – частиц и союзов. Источник материала – авторские профили, принадлежащие авторам-мужчинам и авторам-женщинам, на одном дискуссионно-информативном портале русского Интернета. Для решения поставленных задач авторами применяется лексико-квантитативный метод, связанный с определением частотности избранных лексем в текстах разных авторских профилей с их дальнейшим ранжированием в данном параметре для определения степени маркированности.

The article is devoted to the problem of gender-marked lexemes in the Russian Language. The question about the degree of marked lexemes is introduced and discussed, the methods of lexical units measuring being worked out. The article is done on the materials of link parts of speech – conjunctions and particles. The source of the materials is author profiles belonging to men and women on the discussion-informative portal of Russian Internet. The research is done by means of lexical-quantitative method which is connected with the definition of chosen lexemes frequency in the texts of different author profiles, further ranking to define the degree of marked lexemes being carried out.

Ключевые слова: гендерная лингвистика, идентификационная лингвистика, частицы русского языка, лексико-квантитативный метод, мужская и женская речь, гендерно-маркированные тексты.

Keywords: gender linguistics, identification linguistics, Russian particles, lexical-quantitative method, male and female speech, gender-marked texts.

Статья включена в парадигму сибирской лингвоперсонологии, она продолжает линию разработки лингвоперсонологической гипотезы языка, согласно которой между оппозициями типов языковой личности и значимостями, формируемыми оппозициями в системе языка, существуют устойчивые коррелятивные отношения [2, с. 512 – 522]. В данной статье осуществляется попытка выстроить гендерную оппозицию языковых личностей и представить ее как оппозицию гендерно-маркированных лексем. Попытка противопоставления лексем по гендерному признаку осуществлена нами в статье «Общий род» и гендерная семантика русских имен существительных: бигендерность или агендерность?» [3, с. 14 – 28], в которой гендерная сема понималась в системно-языковом плане как узловая «способность» персоной лексемы обозначать лиц мужского и женского пола, а зоолексемы – самцов и самок. В данной статье поиск гендерной семы переведен в речевую плоскость, ее проявления авторы стремятся обнаружить не в объекте обозначения, а в субъекте речи, точнее, в образе автора в тексте, который так или иначе проявляет себя в нем, в том числе и в плане его гендерной принадлежности. Текст, таким образом, выступает как персонотекст. Гендерные персонотексты исследовались нами ранее в работе «К основаниям гендерной типологии читателей (на материале рефлексий о художественном тексте младших подростков)» [6, с. 119 – 125].

Основной способ обнаружения гендерных проявлений автора в тексте в настоящей статье – квантитативный, связанный с определением частотности избранных лексем в текстах разных авторов (относи-

тельно которых известна их гендерная принадлежность) с их дальнейшим ранжированием в данном параметре для определения степени маркированности. Квантитативное измерение системных оппозиций в языке также имеет свою линию в исследованиях авторов [1; 4; 5].

Изучение языка в аспекте гендерных различий имеет в российской лингвистике достаточно давнюю традицию, начало которым положили социолингвистические исследования. Гендерная лингвистика изучает языковые явления, связанные с различием носителей языка по полу. Среди факторов, определяющих дифференциацию языковых единиц, используемых мужчинами и женщинами, могут быть социальные, религиозные и др. Чаще всего различия проявляются в речи, на лексическом уровне, но могут также проявляться на орфоэпическом, фонологическом уровнях [16, с. 42]. В большинстве исследований гендерных проявлений в речи под мужским и женским полом подразумевается «...не биологический пол, а скорее, социальная роль (быть мужчиной/женщиной и совершать в связи с этим соответствующие данной культуре действия, в том числе и речевые)» [11, с. 46 – 47].

Существует множество работ, связанных с исследованием закономерностей в мужской и женской речи, среди них: Т. В. Гомон, Е. И. Горошко, Е. А. Земская, А. В. Кирилина, Е. С. Ощепкова [8 – 11; 15]. В результате исследований учёные приходят к выводу, что чётких дифференцирующих признаков, способных идентифицировать мужскую и женскую речь, не существует, однако, можно говорить о степени проявленности тех или иных единиц, указывающих на принадлеж-

ность автора к мужскому или женскому полу. Среди этих исследований мы выделяем работу А. В. Кириловой, в которой говорится об определенных особенностях речевого стиля мужчин и женщин, который проявляется на двух уровнях, которые с известной долей условности мы назвали *симптомами первого и второго порядка*. К симптомам первого порядка относятся признаки, обнаруживаемые более четко. Они могут быть замечены непосредственно в период речевого общения: перебивания, длительность речевого периода, категоричность высказывания и связанные с ней предпочтения в выборе типа речевого акта, управление тематикой диалога и т. д. К симптомам второго порядка относятся особенности речи, для выявления которых требуется специальная статистическая процедура: частотность употребления определенных частей речи, частиц, синтаксических конструкций. Данное исследование имеет своим предметом, по терминологии А. В. Кириловой, симптомы второго порядка.

Целью настоящего исследования является выявление в тексте симптомов второго порядка на основе служебных слов русского языка (прежде всего частиц), которые несут в себе признак гендерной принадлежности языковой личности автора данного текста. Основным способом обнаружения названного признака является количественный подсчет интересующих нас лексем в текстах, написанных мужчинами и женщинами, и дальнейшим их сопоставлением, направленным на выявление лексем, которые можно было бы квалифицировать как преимущественно мужские или женские. Маркированные по гендерному признаку тексты мы называем персонотекстами. Обращение к формально количественным методам обусловлено тенденциями развития лингвоперсонологии вообще (см. опыты квантитативных исследований в [13 – 14]) и гендерной лингвистики в частности. А. Г. Фомин писал в связи с последней: «Введение статистических принципов обработки данных в лингвогендерные исследования не является новым для мировой науки, но современная российская лингвистика находится на таком уровне обобщения лингвистических данных, когда обнаруживается необходимость обязательной, точной обработки данных для получения репрезентативных выводов» [17, с. 5].

Исходной установкой исследования послужило предположение о том, что в русском языке существуют слова, которые в плане частотности имеют гендерную маркированность. Следовательно, в текстах, написанных женщиной, будут чаще встречаться одни слова, а в текстах, написанных мужчиной, другие слова. На основе данной гипотезы за основу исследования был взят лексико-квантитативный метод изучения (см. о нём: [13; 14; 17]).

Материалом послужили 9 текстов (*средней размер текста – 10 тыс. слов.*) пользователей портала “Newsland”, при этом известно, что 3 из них принадлежат мужчинам (м1, м2, м3), 3 – женщинам (ж1, ж2, ж3) и 3 текста на начальном этапе идентифицированы не были.

Newsland – информационно-дискуссионный портал, в котором пользователи могут оставлять комментарии к новостным статьям. У каждого зарегистриро-

ванного пользователя сайта имеется профиль. За основу материала исследования были отобраны профили пользователей данного портала, и из данных профилей извлечены авторские комментарии. Комментарии были извлечены подряд, без деления на темы и т. п.

Для исследования были выбраны наиболее частотные частицы и союз русского языка [18]: *же, ведь, бы, вот, только, ещё, если, ну, не*.

На начальном этапе работы было подсчитано количество слов, принадлежавших данному пользователю, и по формуле:

$$\frac{\text{Количество частиц в тексте} \times 100}{\text{общее количество слов}} =$$

$$= \text{частотность частиц, выраженная в \%}$$

найдена частотность употребления заданных частиц в соотношении с общим числом слов.

Ниже представлены результаты подсчётов, которые показывают процентное соотношение исследуемого слова ко всему тексту (все данные указаны в %):

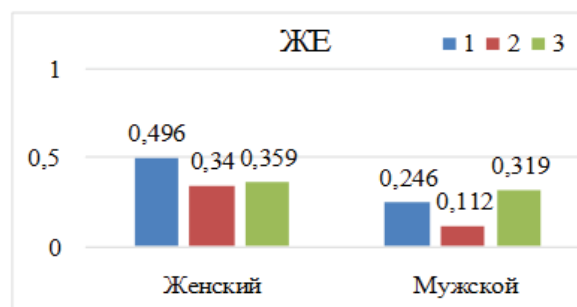


Рис. 1. Частотность употребления частицы **ЖЕ**

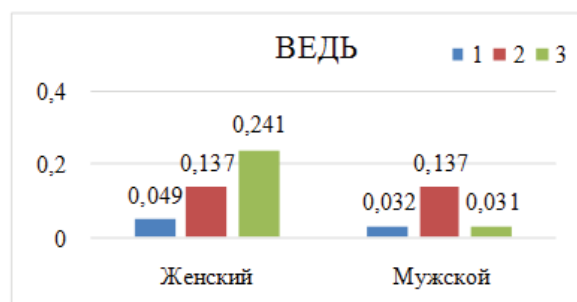


Рис. 2. Частотность употребления частицы **ВЕДЬ**

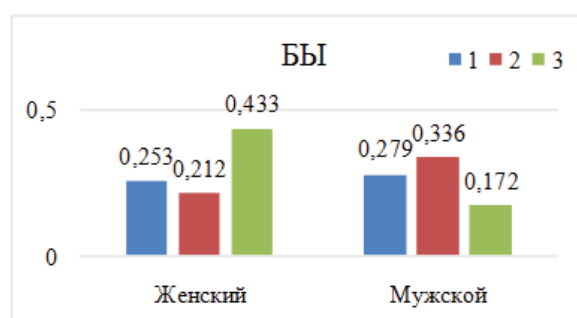


Рис. 3. Частотность употребления частицы **БЫ**

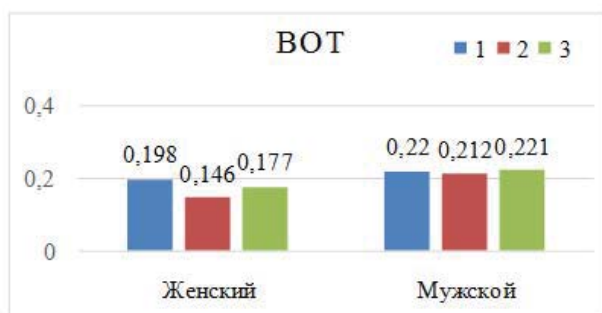


Рис. 4. Частотность употребления частицы **ВОТ**

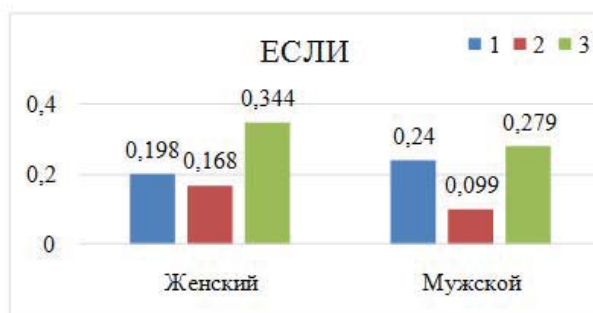


Рис. 7. Частотность употребления союза **ЕСЛИ**



Рис. 5. Частотность употребления частицы **ТОЛЬКО**

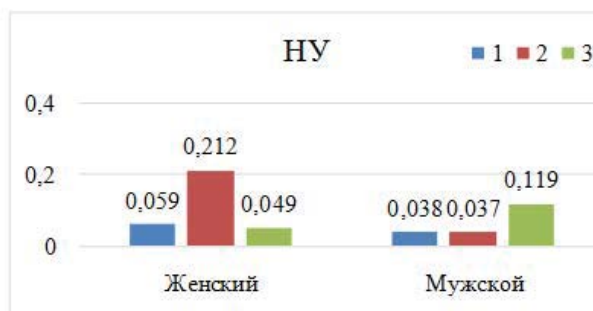


Рис. 8. Частотность употребления частицы **НУ**

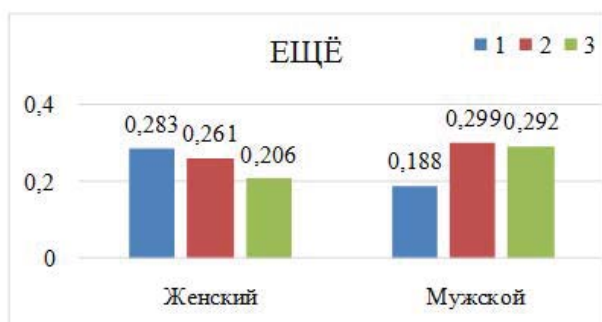


Рис. 6. Частотность употребления частицы **ЕЩЁ**

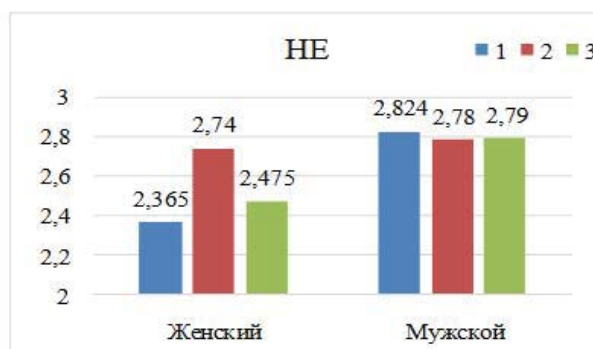


Рис. 9. Частотность употребления частицы **НЕ**

Результаты позволяют провести линию гендера у трёх частиц: *же*, *вот*, *не*.

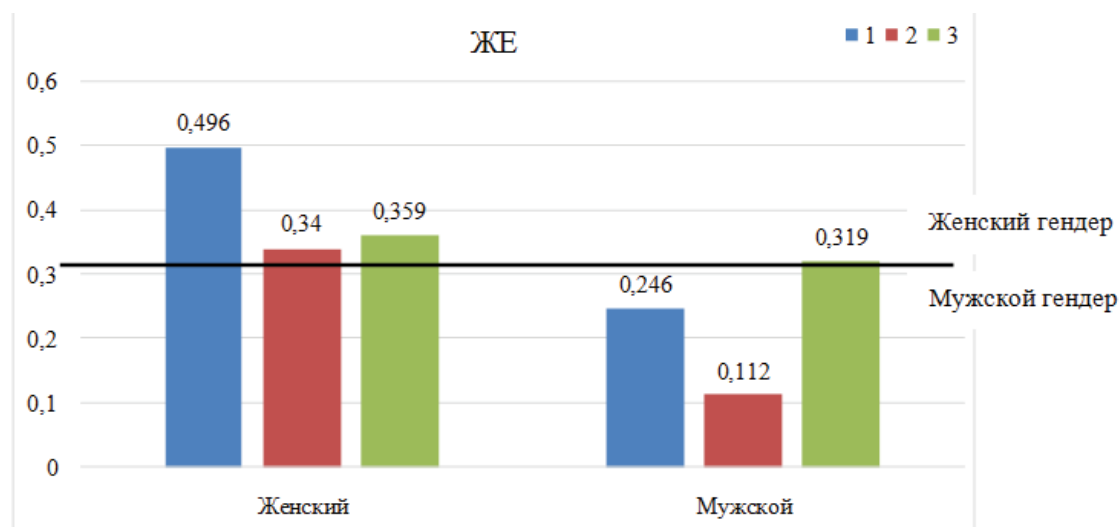


Рис. 10. Сопоставление частотностей употребления частицы **ЖЕ** в женских и мужских текстах

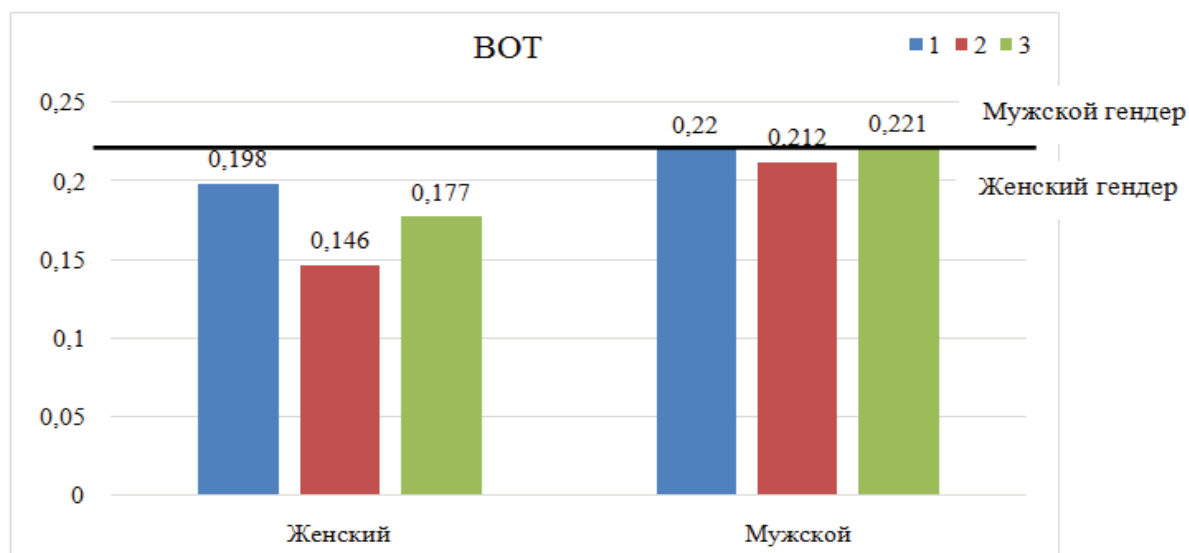


Рис. 11. Сопоставление частотностей употребления частицы ВОТ в женских и мужских текстах

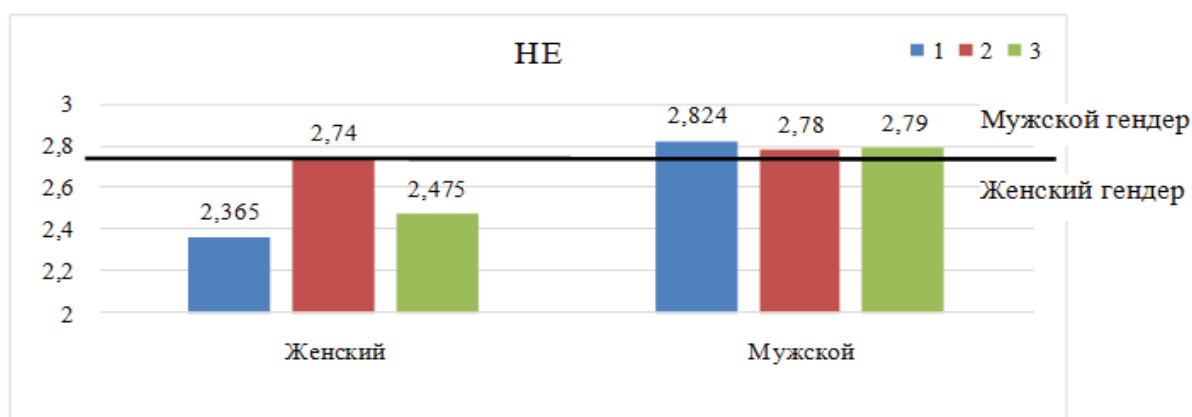


Рис. 12. Сопоставление частотностей употребления частицы НЕ в женских и мужских текстах

Данная линия говорит о возможной гендерной маркированности данных частиц русского языка. Следующим этапом исследовательской работы является апробация полученных результатов. Проверка проводилась на основе трёх текстов других авторов, взятых из того же источника. Для частоты проверки изначально пол авторов известен не был.

Принцип проверки: в проверочных текстах (п1, п2, п3) высчитывалась частотность гендерно-маркированных частиц по той же формуле, что в исходных текстах. Далее результаты соотносились с предыдущими данными. Реальный пол авторов был проверен по глаголам прошедшего времени: все тексты были написаны мужчинами.

Результаты проверки: рис. 13 – 15.

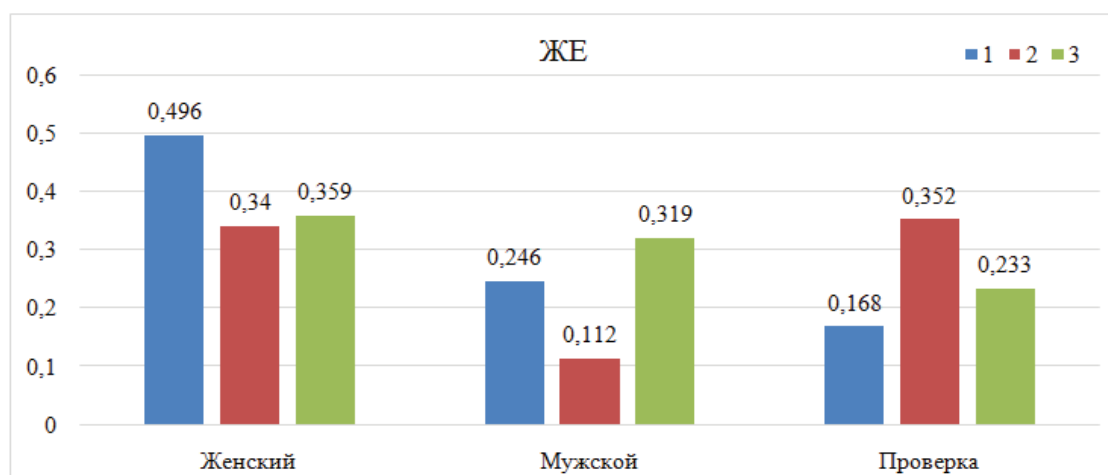


Рис. 13. Сопоставление частотностей проверочных текстов с имеющимися данными по частице ЖЕ

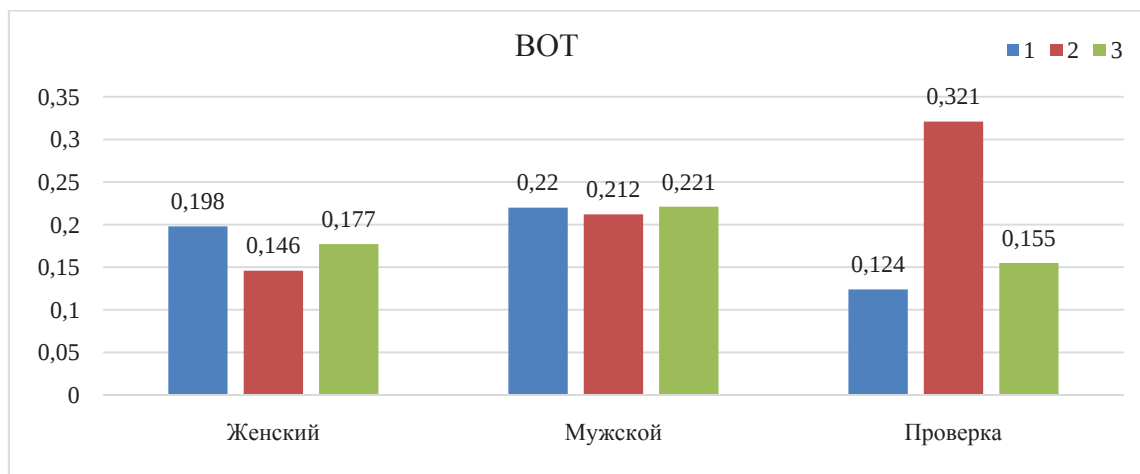


Рис. 14. Сопоставление частотностей проверочных текстов с имеющимися данными по частице BOT

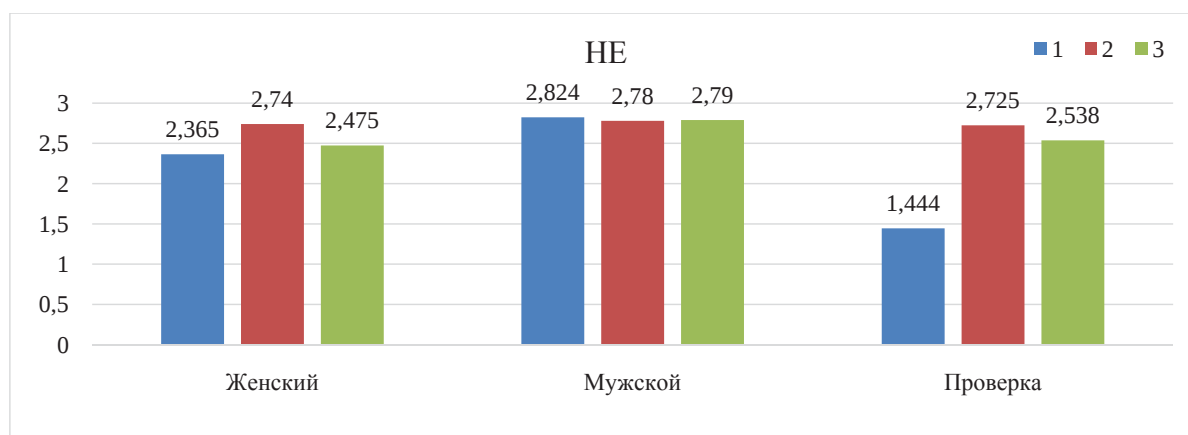


Рис. 15. Сопоставление частотностей проверочных текстов с имеющимися данными по частице HE

Результаты анализа позволяют сделать следующие обобщения: p1 = женский (по частице *не*), мужской (по частице *вот*), мужской (по частице *же*), следовательно p1 = 1/3 женский, 2/3 = мужской; p2 = женский + мужской + женский, следовательно p2 = 2/3 женский, 1/3 мужской; p3 = женский + женский + мужской, следовательно p3 = 2/3 женский, 1/3 мужской. Что говорит нам о том, что текст № 1 принадлежит мужчине, № 2 и № 3 – женщине.

Данные результаты получены на ограниченном материале (и в плане массива «просчитанных» текстов и в плане набора «просчитанных» лексем и языковых личностей – авторов текстов), поэтому экстраполировать эти результаты в плоскость системной лингвистики и говорить об устойчивости гендерной семы для той или иной лексики преждевременно. Также из того, что полученные результаты не соответствуют реальному полу автора, следует: из приведённых частиц нет такой, которая смогла бы чётко определить гендер автора. Есть лишь некоторые проявления различий, носящих слабо выраженный градуальный характер, так, частица *же* проявляет достаточно сильную позицию в гендерно-идентификационном аспекте. Это во многом объясняется специфическими функционально-семантическими характеристиками служебных слов, употребление которых обусловлено не столько семантикой (тем более номи-

нативной), сколько синтагматикой. По терминологии представителей Пражской лингвистической школы, они являются номинативно слабыми, но синтагматическими сильными (в большей мере последнее касается предлогов и союзов). Продолжая линию обоснования слабости гендерной семы, заметим также, что среди проанализированных частиц и союза: *же, ведь, бы, вот, только, ещё, если, ну, не* – нет такой, с помощью которой можно чётко идентифицировать пол автора. В это плане значимы результаты, полученные в исследовании Т. Б. Крючковой [12], в поле зрения которой попадает частица *не*. Согласно данному исследованию «в речи женщин также наблюдалась тенденция к более частому употреблению частицы "не"». Нетрудно убедиться, что в нашем исследовании получены данные, позволяющие интерпретировать частицу «не» как гендерно (слабо) маркированную. Нарушение закономерности можно объяснить разными дискурсами текстов, взятых для исследований. Все сказанное приводит к выводу о необходимости продолжения исследований в рамках сформулированной гипотезы на другом (дискурсном, жанровом, персонном) материале.

Литература

1. Голев Н. Д. О возможностях Интернета в количественном описании лексической семантики // Славянская филология: исследовательский и методический аспекты: материалы I Международной научной конференции (23 – 24 марта 2006 г.) / под ред. Н. Б. Лебедевой. Вып. 1. Кемерово: Меркурий, 2006. С. 41 – 49.
2. Голев Н. Д. О корреляционных связях между системой типов языковых личностей и системой категорий русского языка // VI Международная научная конференция «Русский язык в языковом и культурном пространстве Европы и Мира: Человек. Сознание. Коммуникация. Интернет». 22 – 25 мая 2014 года. Лёвен, Бельгия. Department of Slavonic and Eastern European. Studies of the University of Leuven. Instytut Rusycystyki Uniwersytetu Warszawskiego). С. 512 – 522.
3. Голев Н. Д. «Общий род» и гендерная семантика русских имен существительных: бигендерность или агендерность? // Вестник Томского государственного университета. (Серия: Филология). 2013. № 6(26). С. 14 – 28.
4. Голев Н. Д. Поисковые системы Интернета как лингвистический источник (на примере решения некоторых теоретических и прикладных вопросов русского словообразования) // Новые явления в славянском словообразовании: система и функционирование: Доклады XI Международной научной конференции Комиссии по славянскому словообразованию при Международном комитете славистов (Москва, МГУ им. М. В. Ломоносова, филологический факультет, 24 – 26 марта 2009 г.) / под ред. Е. В. Петрухиной. М.: Изд-во Моск. ун-та, 2012. С. 197 – 212.
5. Голев Н. Д. Понятие «квантитативная мощность языковых единиц» и его исследовательские возможности (по статистическим данным поисковых систем Интернета) // Метаязык науки: мат. Международной научной конференции / гл. ред. С. В. Лесников. Сыктывкар: СыГГУ, 2012. С. 245 – 254.
6. Голев Н. Д., Максимова С. А. К основаниям гендерной типологии читателей (на материале рефлексий о художественном тексте младших подростков) // Вестник Кемеровского государственного университета. 2008. № 4. С. 119 – 125.
7. Головина Т. А. Лингвоперсонологическое функционирование частей речи: статистический аспект: на материале художественных текстов: дис. ... канд. филол. наук. Барнаул, 2008.
8. Гомон Т. В. Судебно-автороведческая экспертиза текстов документов, составленных с намеренным искажением письменной речи: автореф. дис. ... канд. юрид. наук. М., 1992.
9. Горошко Е. И. Особенности мужского и женского вербального поведения: (психолингвистический анализ): автореф. дис. ... канд. филол. наук. М., 1996.
10. Земская Е. А., Китайгородская М. А., Розанова Н. Н. Особенности мужской и женской речи // Русский язык в его функционировании / под ред. Е. А. Земской, Д. Н. Шмелева. М.: Наука, 1993.
11. Кирилина А. В. Гендер: лингвистические аспекты: монография. М., 1999.
12. Крючкова Т. Б. Некоторые экспериментальные исследования особенностей использования русского языка мужчиной и женщиной // Проблемы психолингвистики. М., 1975. С. 186 – 200.
13. Напреенко Г. В. Идентификация текста по его авторской принадлежности на лексическом уровне (формально-количественная модель) // Вестник Томского государственного университета. 2014. № 379. С. 17 – 23.
14. Напреенко Г. В. Интернет-дневники и проблема идентификации личности // Языковая личность: моделирование, типология, портретирование. Сибирская лингвоперсонология. М.: ЛЕНАНД, 2014. С. 544 – 559.
15. Ощепкова Е. С. Идентификация пола автора по письменному тексту (лексико-грамматический аспект): автореф. ... дис. канд. филол. наук. М., 2003.
16. Словарь социолингвистических терминов / отв. ред. В. Ю. Михальченко. М., 2006.
17. Фомин А. Г. Психолингвистическая концепция гендерной языковой личности: дис. ... д-ра филол. наук. Барнаул, 2004.
18. Частотный словарь русского языка / под ред. Л. Н. Засориной. М., 1977.

Информация об авторах:

Голев Николай Данилович – доктор филологических наук, профессор кафедры русского языка КемГУ, ngolevd@mail.ru.

Nikolay D. Golev – Doctor of Philology, Professor at the Russian Language Department of Kemerovo State University.

Реттих Дмитрий Александрович – студент по направлению «Отечественная филология» КемГУ, hcitteramid@gmail.com.

Dmitriy A. Rettikh – student of Kemerovo State University.

(Научный руководитель – Н. Д. Голев). (Research advisor – N. D. Golev).

Статья поступила в редколлегию 01.06.2015 г.