

Бурцева Е. А.

Московский государственный университет имени М. В. Ломоносова

СТИЛЬ ЙОХАННЕСА БРЮХМАНА И ЛЕКСИКО-СТАТИСТИЧЕСКИЙ АНАЛИЗ СБОРНИКА ПРОПОВЕДЕЙ ИЗ МОНАСТЫРЯ СВЯТОЙ АГНЕС В МААСЕЙКЕ

Статья посвящена стилометрическому исследованию проповедей Йоханнеса Брюхмана, собранных в рукописи Gent UB hs. 1301 и использовавшихся монахинями монастыря Святой Агнес в Маасейке. В качестве метода исследования выступает анализ наиболее частотных служебных слов в текстах. Анализ десяти наиболее частотных слов приводит к выводу, что в девяти из десяти проповедей демонстрируются закономерности в употреблении: союзов *ende* ('и' — наиболее частотная лемма), *die* и *dat*, которые могут означать 'который, которая, которое, тот, та, то, те'. Вероятность погрешности (смещение одного из элементов далее 3-го места в ранге или *ende* не на первом месте) увеличивается обратно пропорционально объему текста проповеди. Не удалось исключить специфические индивидуальные стилистические черты в проповедях Брюхмана.

Ключевые слова: стилометрия, стилистика, рукопись, проповедь, жанровое своеобразие, средненидерландский язык.

Evgeniia A. Burtseva

Lomonosov Moscow State University

JOHANNES BRUGMAN'S STYLE AND LEXICOSTATISTICAL ANALYSIS OF A COLLECTION OF SERMONS FROM SAINT AGNES CONVENT IN MAASEIK

This article is devoted to a stylometric study of Johannes Brugman's sermons collected in the manuscript Gent UB hs. 1301, created and used by the nuns of the convent of Saint Agnes in Maaseik. The study was an attempt to refute the claim that the 10 most closely related texts of Brugman's oeuvre would show common patterns that can be detected by lexico-statistical methods. The method of research is the analysis of the most frequent function words in the texts. The analysis of the ten most frequent words leads to the conclusion that nine out of ten sermons demonstrate patterns in the use of conjunctions: *ende* 'and' is the most frequent lemma, followed by *die* and *dat*, which can mean 'which, this, these, that, those'. The probability of a deviation (shifting one of the elements further than third place in the ranking or *ende* not in the first place) increases in inverse proportion to the volume of the sermon text. The discovery of patterns in the use of service words suggests that the study has failed to refute the assumption that individual stylistic features of Brugman's sermons can be detected.

Keywords: stylometry, stylistics, manuscript, sermon, genre specificity, Middle Dutch language.

Изучение средневековых рукописных текстов неразрывно связано с проблемой авторства. Если автор известен заранее, возникает вопрос, как много авторского текста и авторских черт осталось после многократного переписывания, в процессе которого в рукопись могли быть внесены значительные изменения. Концепция авторства в том виде, в каком она представляется современному человеку после эпохи романтизма, оказывается в значительной степени анахроничной для Средних веков. Процесс воспроизведения авторского текста в Средние века был неизбежно связан с внесением изменений: от орфографической правки до полного переписывания [Kestemont, Moens, Deploige, 2014, p. 200]. Более того, некоторые жанры словесности предполагали вмешательство в авторский текст в большей степени, чем другие. Проповеди можно отнести к жанру, для которого характерно существенное вмешательство переписчика в авторский текст.

Описывая способы передачи и сохранения произнесенных проповедей на письме, Г. С. Зилеман отталкивается от книги известной английской исследовательницы Берил Смаллей «Изучение Библии в Средние века» [Smalley, 1952, p. 200], в которой она сообщает о специфическом жанре записанных проповедей — *reportatio* (от лат. *reporto* — ‘передавать, доносить’). Такие проповеди представляли собой конспект, часто содержали слова *inquit, dicit* (‘сказал’) и несли утилитарную информативную функцию, обычно использовались в образовании и схоластике. В противоположность Зилеман ввел термин *recompositio* (от лат. *recompono* — ‘воссоздавать’), описывающий особый тип проповедей, которые в дальнейшем должны были служить текстом для благочестивого чтения, такие проповеди подлежали последующей обработке и редактированию [Zieleman, 1978, p. 17].

Проповеди *recompositio* подходили для чтения во время трапез и досуга, в XV в. они массово воспроизводятся в так называемых среднидерландских *devotieboeken* — книгах для благочестивого чтения.

Проблема устойчивости авторского текста при копировании, а также устойчивости индивидуального авторского стиля оказывается далеко не праздной в том случае, если мы сталкиваемся с авторским стилем, который высоко ценился современниками и оставил глубокий след в культуре. В Нидерландах XV в. обладателем такого стиля был Йоханнес, или Ян, Брюхман (умер в 1473 г.) [Mertens, 2008, p. 253].

Своеобразие Брюхмана как автора делает вопрос о выделении характеристик его индивидуального языкового стиля крайне интересным. Он может быть разрешен разными методами, например при помощи создания и анализа корпуса текстов среднидерландских проповедей XV в., куда вошли бы и проповеди Брюхмана. Приступая к созданию массивного корпуса из сотен тысяч слов, а также к его анализу, важно обосновать потенциальную возможность положительного результата. Представляется уместным обратиться к идеям Карла Поппера, считавшего фальсификацию, то есть опровержение или попытку опровержения теории, намного более перспективным подходом, чем верификацию, то есть попытку доказать утверждение. По мнению Поппера, именно в выдвижении предположений и их дальнейшей фальсификации заключается истинная теория познания [Поппер, 2004, с. 260].

В данной статье сделана попытка опровергнуть предположение о том, что индивидуальный стиль Йоханнеса Брюхмана может быть выделен и описан. Используемые тексты были собраны вместе одним писцом (монахиней монастыря Святой Агнес в Маасейке) и представляются наиболее близкими друг другу с точки зрения их происхождения. Значительные стилистические расхождения, в частности используемые для авторской атрибуции, например частотность использования служебных слов и их положение с точки зрения ранга, если бы они были найдены в этих текстах, указали бы на принципиальную невозможность выделения общих авторских черт, составляющих индивидуальный стиль. Так как в данном исследовании в качестве основного материала для выделения индивидуальных авторских черт был выбран лексико-статистический метод, считающийся наиболее объективным при определении авторства [Литвинова, Громова, 2020, с. 78], и так как ключевую роль в квантитативных лексико-статистических исследованиях играют служебные части речи [Binongo, Smith, 1999, p. 450], значительные расхождения в употреблении служебных частей речи указали бы на невозможность выделения единого языкового стиля автора.

Понимание стиля в современной лингвистике двояко: стиль может восприниматься как набор индивидуальных черт, характерных для того или иного автора, либо как функциональный стиль, характерный для определенной сферы общения [Степанов, 1990, с. 440]. Так как в данном случае интерес представляет именно авторский стиль, речь идет об индивидуальном стиле — наборе индивидуальных лингвистических особенностей или характеристик, которые объединяют разные тексты, принадлежащие одному автору [Но, 2011, p. 6]. Индивидуальные лингвистические особенности, составляющие авторский стиль, в письменном тексте могут быть выделены на лексическом и грамматическом уровнях [Doherty, 2015, p. 1943].

В качестве объекта для первого этапа исследования были выбраны 10 проповедей Йоханнеса Брюхмана, находящихся в рукописи из коллекции Университета Гента (шифр рукописи — Gent UB hs. 1301). Помимо фотокопий рукописи на сайте библиотеки Университета Гента, проповеди также можно найти в издании проповедей Брюхмана 1948 г. [Dijk van, 1948, p. 1]. Рукопись датирована библиотекой 1476 г. и таким образом оказывается практически прижизненной (Брюхман умер в глубокой старости в 1473 г.). В то же время примерные годы активности Брюхмана как странствующего проповедника — 1450–1464 гг. [Pollmann, 1978, p. 95], что увеличивает временную дистанцию между произнесением проповедей и созданием рукописи в монастыре Святой Агнес в Маасейке. Как было сказано ранее, обнаруженное методами лексико-статистического анализа сходство указывает на потенциальную возможность выделения авторского стиля Брюхмана, но не указывает на авторский стиль сам по себе и не доказывает выделение специфических авторских черт: языковой стиль памятника может оказаться индивидуальным стилем монахини, а не самого Брюхмана.

Над вопросом о способах и стратегиях применения лексико-статистических и иных стилометрических методов для определения стилистического сходства средневековых религиозных текстов работает группа из Университета Антверпена, в частности Мике Кестемонт и Вальтер Далемнс. С начала 2010-х годов они изучают средневековые тексты на средненидерландском языке и латыни, применяя методы стилометрии.

Прежде чем привести в пример результаты, полученные исследователями, необходимо объяснить, в чем заключаются принципы стилометрии. Введенный революционером-народником Н. А. Морозовым [Мартыненко, 2015, с. 16] термин содержит соединительную гласную *e*. В работе Мартыненко он определен следующим образом:

«Стилометрия — прикладная филологическая дисциплина, занимающаяся измерением стилиевых характеристик с целью систематизации и упорядочения (типологии, атрибуции, датировки, диагностики, реконструкции и т. д.) текстов. Важную роль в стилометрии играют лексико-статистические методы» [Мартыненко, 1988, с. 10].

Исследователи стилометрии исходят из того, что каждый автор склонен использовать те или иные, характерные для него специфические лингвистические средства, которые М. Кестемонт и другие называют «стиломами» [Kestemont, Moens, Deploige, 2014, p. 205]. Индивидуальный набор таких «стилом» служит отличительной чертой и может быть выделен статистическими методами. Г. Я. Мартыненко подчеркивает важность именно лексических методов в статистических подсчетах.

Группа исследователей вместе с М. Кестемонтом занята изучением частотности лексики и ее распределения по текстам. Выбор лексики как основного предмета исследования, вероятнее всего, продиктован техническими возможностями: исследования выполнены при помощи программ, которые подсчитывают наиболее частотные лексемы, а затем строят графики их распределения по корпусу текстов. После небольшой технической подготовки, лемматизации, простейшие программы легко подсчитывают наиболее частотные лексемы в тексте. Автоматизация исследования синтаксиса средненидерландского языка на данный момент едва ли представляется возможной, хотя и существуют программы, способные строить синтаксические деревья для современного английского языка.

Составляя списки наиболее частотных, но при этом не связанных с темой текста слов, ученые из Университета Антверпена пытаются выяснить, насколько может быть устойчив авторский стиль в тексте рукописи. Одна из наиболее интересных работ посвящена произведениям Хильдегарды Бингенской, написанным ею с двумя разными секретарями. Совместная работа над текстом двух разных авторов приводит к появлению третьего авторского стиля, не являющегося ни авторским стилем Хильдегарды, ни авторским стилем ее секретаря, но в то же время полученный новый стиль обладает чертами авторских стилей и того и другого. Их индивидуальные предпочтения в лексике могут быть обнаруже-

ны при помощи стилометрического анализа даже после редактуры [Kestemont, Moens, Deploige, 2014, p. 220].

В другой работе, посвященной стилометрическому анализу рифмованной поэзии, исследователи из Антверпена объясняют значимость служебных слов, приводя следующие аргументы. Во-первых, служебные слова обычно являются наиболее частотными в практически любом корпусе текстов; во-вторых, они не связаны с темой текста напрямую, что позволяет сравнивать тексты разных авторов и тексты, посвященные разным темам. В-третьих, по мнению ученых, такие слова не подконтрольны воле автора, автор с меньшей долей вероятности будет осознанно выбирать союз, чем прилагательное, занимающее место важного эпитета [Kestemont, Daelemans, Sandra, 2012, p. 61].

Чтобы проанализировать проповеди Брюхмана в рукописи Gent UB hs. 1301 лексико-статистическими методами, была составлена таблица наиболее частотных слов в каждой из проповедей, полученных при помощи автоматического компьютерного анализа. Для анализа использован бесплатный веб-интерфейс стемминга текста GSGen.ru, оценивающий частоту употребления слов. После ввода текста веб-интерфейс формирует таблицу, содержащую слова в порядке убывания числа вхождений и число вхождений для каждого элемента.

Сопоставление наиболее частотных слов показало, что в большинстве проповедей обнаруживаются заметные сходства в частотности союзов. В восьми из десяти случаев наиболее частотным словом оказывается союз *ende* ('и'), за которым следуют *die* ('тот, который') или *dat* ('что, то, которое') в разной очередности. Сложно определить, какую именно функцию и в какой пропорции несут слова *die* и *dat*, которые могут быть как указательными местоимениями, так и союзами. Таблица также демонстрирует, что устойчивость закономерности, поднимающей *ende*, *die* и *dat* на самые высокие позиции в рейтинге частотности слов, коррелирует с объемом текста: чем больше объем текста, тем выше шанс, что самым частотным словом окажется *ende*, а за ним будут следовать *die* или *dat*. С большой вероятностью такое распределение связано с тем, что чем больше в тексте слов, тем меньше роль случайного отклонения как статистической погрешности. В целом распределение союзов в списках наиболее частотных слов указывает на общие закономерности во всех текстах, кроме текста № 3, что, согласно лексико-статистическим методам, используемым в квантитативной стилистике, косвенно указывает на стилистическое единство большинства проповедей.

Одна из проповедей — проповедь № 3 «Две вещи свойственны для благочестивой девы» — заметно отличается от остальных как распределением наиболее частотных слов, так и общей длиной — это самый короткий текст в сборнике. Проповедь оказывается не законченным текстом, предназначенным для чтения во время досуга, а набором списков. Автор перечисляет три черты благочестивой девы, три признака благочестивого человека, каждый из пунктов состоит из одного слова, словосочетания или предложения и вводится словами «во-первых», «во-вторых» и т. д.

Организация текста в виде конспекта сближает проповедь № 3 с жанром, определенным Б. Смаллей как *reportatio*, что, в свою очередь, указывает на жанровые различия между данной проповедью и остальными.

Несмотря на заметные, но объяснимые различия в употреблении наиболее частотных служебных слов, можно сделать вывод о том, что определение авторского стиля Брюхмана потенциально возможно, так как большинство наиболее близких друг другу текстов не демонстрирует радикальных различий по результатам лексико-статистического анализа. Заметные различия могут быть объяснены статистическими погрешностями и жанровым разнообразием памятника. Теоретическая возможность выделения индивидуального, авторского стиля проповедей Йоханнеса Брюхмана не была опровергнута.

Литература

- Литвинова Т. А., Громова А. В. Компьютерные технологии в судебной автороведческой экспертизе: проблемы и перспективы использования // Вестник Волгоградского государственного университета. Серия: Языкознание. 2020. Т. 19, № 1. С. 77–88.
- Мартыненко Г. Я. Основы стилиметрии. Л.: Изд-во ЛГУ, 1988.
- Мартыненко Г. Я. Стилиметрия: возникновение и становление в контексте междисциплинарного взаимодействия. Ч. 2. Первая половина XX века: расширение междисциплинарных контактов стилиметрии // Структурная и прикладная лингвистика. 2015. Вып. 11. С. 9–28.
- Поппер К. Р. Предположения и опровержения: рост научного знания / пер. с англ. М.: АСТ, 2004.
- Степанов Ю. С. Стилистика. Лингвистический энциклопедический словарь. М.: Советская энциклопедия, 1990. С. 440–442.
- Binongo J., Smith W. The application of principal components analysis to stylometry // Literary and Linguistic Computing. 1999. 14 (4). P. 446–466.
- Johannes Brugman, *Verspreide sermoenen* / ed. by Dijk A. van. Antwerpen: De Nederlandse Boekhandel, 1948.
- Doherty M. *Syntax and Stylistics* // *Syntax — theory and analysis: An international handbook*. Berlin: De Gruyter Mouton, 2015. P. 1942–1962.
- Ho Y. *Corpus Stylistics in Principles and Practice: A Stylistic Exploration of John Fowles' The Magus*. London: Continuum, 2011.
- Kestemont M., Daelemans W., Sandra D. Robust Rhymes? The Stability of Authorial Style in Medieval Narratives // *Journal of Quantitative Linguistics*. 2012. 19 (1). P. 54–76.
- Kestemont M., Moens S., Deploige J. Collaborative authorship in the twelfth century: A stylometric study of Hildegard of Bingen and Guibert of Gembloux // *Digital Scholarship in the Humanities*. 2014. 30 (2). P. 199–224.
- Mertens Th. The Sermons of Johannes Brugman OFM († 1473): Preservation and Form // *Constructing the Medieval Sermon*. Turnhout: Brepols Publishers, 2008. P. 253–274.
- Pollmann R. J. *Jan Brugman en Observanten-beweging. 750 jaar minderbroeders in Nederland*. Dl. 1. Utrecht: Werkgroep K, 1978. P. 91–102.
- Smalley B. *The Study of the Bible in the Middle Ages*. Oxford: Blackwell, 1952.
- Zieleman G. C. Overleveringsvormen van middeleeuwse preken in de landstaal // *Nederlands Archief voor Kerkgeschiedenis*. 1978. 59 (1). P. 11–20.

References

- Binongo J., Smith W. The application of principal components analysis to stylometry. *Literary and Linguistic Computing*, 1999, 14 (4), pp. 446–466.
- Dijk A. van (ed.). *Johannes Brugman, Verspreide sermoenen*. Antwerpen: De Nederlandse Boekhandel, 1948.
- Doherty M. *Syntax and Stylistics. Syntax — theory and analysis: An international handbook*. Berlin: De Gruyter Mouton, 2015, pp. 1942–1962.
- Ho Y. *Corpus Stylistics in Principles and Practice: A Stylistic Exploration of John Fowles' The Magus*. London: Continuum, 2011.
- Kestemont M., Daelemans W., Sandra D. Robust Rhymes? The Stability of Authorial Style in Medieval Narratives. *Journal of Quantitative Linguistics*, 2012, 19 (1), pp. 54–76.
- Kestemont M., Moens S., Deploige J. Collaborative authorship in the twelfth century: A stylometric study of Hildegard of Bingen and Guibert of Gembloux. *Digital Scholarship in the Humanities*, 2014, 30 (2), pp. 199–224.
- Litvinova T. A., Gromova A. V. Computer Technologies in Forensic Authorhoveistic Examination: Problems and Prospects of Use. *Vestnik Volgogradskogo gosudarstvennogo universiteta. Serii: Iazykoznanie*, 2020, vol. 19, no. 1, pp. 77–88. (In Russian)
- Martynenko G. Ya. *The Basics of Stylometry*. Leningrad: LGU Publ., 1988. (In Russian)
- Martynenko G. Ya. Stylometry: Emergence and Formation in the Context of Interdisciplinary Interaction. Part 2: The First Half of the 20th Century: Expanding Interdisciplinary Contacts of Stylometry. *Strukturnaya i prikladnaya lingvistika*, 2015, 11, pp. 9–28. (In Russian)
- Mertens Th. The Sermons of Johannes Brugman OFM († 1473): Preservation and Form. *Constructing the Medieval Sermon*. Turnhout: Brepols Publishers, 2008, S. 253–274.
- Pollmann R. J. *Jan Brugman en Observanten-beweging. 750 jaar minderbroeders in Nederland. Dl. 1*. Utrecht: Werkgroep K, 1978. S. 91–102.
- Popper K. R. *Conjectures and Refutations: The Growth of Scientific Knowledge*. Moscow: AST Publ., 2004. (In Russian)
- Smalley B. *The Study of the Bible in the Middle Ages*. Oxford: Blackwell, 1952.
- Stepanov Yu. S. Stylistics. *Lingvisticheskii entsiklopedicheskii slovar'*. Moscow: Sovetskaia entsiklopediia Publ., 1990, pp. 440–442. (In Russian)
- Zieleman G. C. Overleveringsvormen van middeleeuwse preken in de landstaal. *Nederlands Archief voor Kerkgeschiedenis*, 1978, 59 (1), S. 11–20.

Сведения об авторе:

Бурцева Евгения Андреевна

Преподаватель

Московский государственный университет имени М. В. Ломоносова

Российская Федерация, 119991, Москва, Ленинские горы, 1

Evgeniia A. Burtseva

Teacher

Lomonosov Moscow State University

1, Leninskie gory, Moscow, 119991, Russian Federation

E-mail: jenybur@gmail.com