

DOI: 10.31648/pw.12140

WOJCIECH KAJTOCH

ORCID: <https://orcid.org/0000-0003-3000-2384>

Cracow

КОЛИЧЕСТВЕННЫЙ ЛЕКСИЧЕСКИЙ КОНТЕНТ-АНАЛИЗ В ОТНОШЕНИИ К ДРУГИМ МЕТОДАМ ИССЛЕДОВАНИЯ ЯЗЫКА ПРЕССЫ¹

Lexical quantitative content analysis as a method of studying the language of print media

ABSTRACT: The article looks at different analyses of media content. The quantitative content analysis is a necessary first step in the research of any media flow. This type of analysis provides a technique for objectively pointing out themes and thematic lines in a particular medium. Lexical quantitative content analysis is defined as both analyzing and measuring words. It has two distinct paradigms: the leading words method and the keyword method. The text provides definitions for the terms “textual image of the word” and “linguistic image of the world”. The methods explore are presented in a way that accommodates their application.

KEYWORDS: lexical quantitative content analysis, quantitative lexical textual image of the world, content analysis, media studies

Количественный контент-анализ (далее: ККА) всегда представляет собой обязательный первый этап исследования потока медийных сообщений. Сказанное выше подразумевает исследования любых СМИ: номеров газет, телевизионных программ или периодических изданий редактируемых интернет-сайтов.

¹ Данная статья продолжает мои предыдущие публикации (Kajtoch 2023a; 2023b) и имеет своей целью ознакомление русскоязычных читателей из Восточной Европы с оригинальной концепцией лексико-количественного контент-анализа. В силу ограниченного объема статьи я не имел возможности представить и проанализировать хорошо известные мне публикации восточноевропейских исследователей в этой области (например, таких авторов, как А. Д. Шмелев, Ж. Е. Фомичева, Л. Я. Аверьянов, Е. Я. Таршис, А. В. Семенова, М. В. Корсунская, В. В. Осин, В. А. Бубнов и др.).

1. Краткая история изучения классической формы ККА

Согласно Б. Р. Берельсону (Berelson 1970, 7 и сл.) целью контент-анализ является такое описание явного содержания текста, которое обладает следующими свойствами: 1) принимает во внимание только синтаксические и семантические элементы текста; 2) имеет объективный характер; 3) характеризуется систематичностью; 4) обязательно основывается на количественных данных. Методология такого анализа стала известна в Польше благодаря работам В. Писарека (Pisarek 1983a; 1983b). Общеизвестное определение выглядит следующим образом:

Контент-анализ представляет собой совокупность различных техник систематического исследования потоков или сборов текстов, используемых в целях познания элементов и условий процесса коммуникации посредством познания содержания текстов, заключающегося в по мере возможности объективном (на практике: интерсубъективно согласованном) выделении и идентификации самым однозначным образом конкретизированных, формальных или содержательных элементов, а также на как можно более точной (на практике: количественной) оценке дистрибуции появления этих элементов и выводах, полученных, в основном, в результате сопоставления (Pisarek 1983a, 45).

Писарек выделил две методики проведения количественного контент-анализа. Первая позволяет упорядоченно представить тематический хаос газетных текстов, с указанием тематических предпочтений – то есть представляет собой анализ газеты в целом. Другая облегчает нахождение в газете определенной темы таким образом, чтобы потом можно было представить ее наиболее полно. Обе методики оформились в публикациях, появившихся на рубеже 50-х и 60-х гг. прошлого века. Первыми из них были статьи А. Кłosковской (Kłoskowska 1969) и И. Тетелевской (Tetelowska 1960).

2. Почему ККА необходим?

Количественный контент-анализ позволяет объективно представить важные тематические мотивы в газете или другом СМИ, а также показать, какие из них экспонированы в большей степени (находятся на переднем плане), а какие – менее заметны. Количество и экспозиция – это объективные показатели важности темы. Если исследователь определяет важность темы сам, то он руководствуется, как правило, не столько объективными показателями, сколько собственным вкусом и взглядами. По-другому просто не получится. Кроме того, невозможно без предварительного анализа составить список всех тем, затрагиваемых в конкретном СМИ. Эту возможность может предоставить

исследователю ККА. Сущность исследования, проводимого с его помощью, можно сравнить с дефрагментацией твердого диска в компьютере.

Дело в том, чтобы информация, разбросанная по страницам газеты (см. таб. 1) и подаваемая в линейном порядке (в колонках) или хронологически (на радио и в телевидении), а даже при случайной выборке (в интернете), была представлена исследователю в тематической классификации, демонстрирующей, например, идейное, жанровое или любое другое содержание.

Таблица 1. Схематическое представление порядка расположения тем на страницах газеты.

| 1 страница | 2 страница | 3 страница | 4 страница | 5 страница | 6 страница |
|------------|------------|------------|------------|------------|------------|
| Z | N | M | Прочие | Прочие | M |
| Y | X | Прочие | Z | Y | Прочие |
| Y | Прочие | Прочие | Z | N | X |
| X | M | Прочие | 0 | N | Прочие |
| X | Прочие | X | X | 0 | Z |
| Z | M | N | Прочие | Прочие | Прочие |
| Z | 0 | Прочие | N | M | 0 |
| Z | Прочие | M | X | Прочие | Прочие |

Символы X, Y, Z, M, N, 0 кодируют информацию, иллюстрирующую главные темы, обсуждаемые в газете; Прочие – эта группа остальных тем, не интересующих исследователя в данный момент (рассматриваемых обобщенно). Порядок следования символов в направлении сверху вниз показывает, условно говоря, степень экспонирования на странице текста, в котором обсуждается та или иная тема.

Таким образом, целью проведения анализа является упорядочение содержания согласно принципу, позволяющему без помех найти, понять и интерпретировать соответствующую информацию, а также оценить степень ее важности. Результаты анализа представлены схематически в Таблице 2.

Таблица 2. Схематическое представление результата ККА – степень важности выбранных для исследования тем газеты.

| Тема Y | Тема X | Тема Z | Тема N | Тема M | Тема 0 |
|--------|--------|--------|--------|--------|--------|
| Y | X | Z | N | M | 0 |
| Y | X | Z | N | M | 0 |
| Y | X | Z | N | M | 0 |
| Y | X | Z | N | | |
| Y | X | Z | | | |
| Y | X | Z | | | |
| Y | | | | | |

Количество букв | цифр, приписанных данной теме, информирует о том, какой процент пространства газетных колонок занимает представление каждой темы. Чем выше процентное содержание, тем важнее тема. В результате формируется рейтинг: на первом месте – тема Y, на втором месте – тема X и Z, на третьем – N, на четвертом – M и 0.

Информация в ней упорядочена тематически, причем тематическая структура сначала определяется исследователем, а потом верифицируется в ходе анализа.

Естественно, дефрагментация содержания газеты, телевизионной или радиопрограммы намного глубже, чем то, что я показал: в практике она доходит до уровня абзаца или – в устной речи – до уровня фразы (отдельные тексты фельетонов, репортажей и т.д. бывают политематическими, только отдельно взятый абзац всегда остается монотематическим по определению). Техника измерения бывает в реальности более сложной, поскольку только при условии, что единицей анализа является текст, в котором принимается во внимание одна доминирующая тема, могли быть получены результаты, представленные в Таблицах 1 и 2.

3. Как верифицировать самый простой количественный контент-анализ?

Анализ будет проведен успешно, если исследователь правильно, т.е. в соответствии с интересующей его темой, сможет подобрать 1) ключ категоризации; 2) единицы анализа; 3) единицы измерения. Ниже я определю эти термины и дам к ним комментарии.

1. Ключ категоризации:

Ключ категоризации представляет собой схему классифицирования содержания, которая служит основным инструментом контент-анализа СМИ. Правильно подобранный ключ должен соотноситься со всеми категориями (однородными множествами элементов) содержания СМИ, которые достаточным образом описывают проблематику исследуемого материала. Отдельные категории – с точки зрения методологии – должны конкретно и логично разделяться (чтобы исследователи не колебались при кодировании материала) и получить чёткие определения. От правильности ключа категоризации зависит качество последующих этапов проводимого контент-анализа (Kołodziej 2006, 95).

В качестве иллюстрации представлю ключ Клосковской (Таб. 1), при помощи которого автор определила черты «социалистической» модели семьи, представленной на страницах журнала „Przyjaciółka” в 1950-1951 гг., а затем сопоставила эту модель с другой, более традиционной моделью, популяризированной в том же журнале в 1956-1957 гг. За единицу анализа была принята отдельная публикация, а в качестве единицы измерения – появление мотива, связанного с различными условиями семейной жизни.

Таблица 3. Ключ Антонины Кłosковской (1969, 437):
главные темы деятельности и ценностей в рассказах и в рубрике
Radości i smutki в период 1950-1951 гг. в журнале „Przyjaciółka”

| Темы и ценности | Рассказы | <i>Radości i smutki</i> | Итого | % |
|---|----------|-------------------------|-------|-------|
| Работа и семейная жизнь | 17 | 18 | 35 | 33,3 |
| Общественно-политические проблемы и семейная жизнь | 12 | 19 | 31 | 29,5 |
| Семья и брак как высшие ценности | 8 | 9 | 17 | 16,2 |
| Общественные факторы, отрицательно воздействующие на семейную жизнь | 10 | – | 10 | 9,5 |
| Семейный бюджет | – | 8 | 8 | 7,6 |
| Семья и алкоголизм | – | 3 | 3 | 2,9 |
| Неклассифицированные | 1 | – | 1 | 1,0 |
| Всего | 48 | 57 | 105 | 100,0 |

Второй образец анализа, проведенного в 1956-57 гг., принимает во внимание 140 произведений. Семейная тематика затрагивалась в 81% рассказов (для сравнения: первый анализ 1950-1951 гг. выявил эту тематику в 60% текстов), притом рубрика *Radości i smutki* включала 72,7 процента произведений (в первом исследовании эта тематика составила 68,6%). В первом исследовании в 33 из 48 текстов объектом повествования были семьи из «рабочей» среды, а в другом исследовании только 18 из 81 семей главного героя происходила из рабочих, тогда как остальные представляли свободные профессии и интеллигенцию или их социальное происхождение вообще не упоминалось.

В 36 рассказах из 40 главной проблемой были любовь и семейная жизнь (в первом случае – только в 8 из 48). Во втором исследовании практически не говорилось о передовиках труда и не упоминалось обсуждение в семейном кругу политических проблем (Kłoskowska 1969, 435 и сл.). Выводы, вытекающие из исследований Кłosковской (контролируемых цензурой), следующие: после окончания пропагандистского наступления и институционально навязанного сверху соцреализма, сопровождавших приход к власти в Польше сталинистов², „Przyjaciółka” стала публиковать такие материалы, которые были интересны женщинам практически с самого начала существования женской прессы: посвященные любви, семейной жизни, моде, воспитанию

² Крайне просоветские и просталинские деятели ПНР пришли к власти в Польше в результате фальсифицированных выборов в Законодательный Сейм (Собрание) в 1947 году и последующего постановления пленума ПРП «О правом и националистическом уклоне в руководстве партии» (от 3 сентября 1948 г.). Период правления этих деятелей (1949-1955), опиравшихся на террор служб безопасности и пропаганду, позже был назван «периодом ошибок и извращений».

детей и другим личным темам, что в свою очередь предоставляет возможность для дальнейших политологических интерпретаций.

2. Единицы анализа представляют собой те элементы содержания, которые при проведении исследования будут проанализированы и классифицированы согласно категориям принятого ключа. Выборка единиц производится в зависимости от изучаемых СМИ: 1) в случае СМИ, оперирующего письменной речью, это могут быть тексты или отдельные их абзацы; 2) фильмы или телевизионные программы исследуются путем разбиения на сцены и кадры; 3) в принципе, в случае каждого СМИ, в котором появляется натуральный язык, могут быть выделены такие элементы, как имена собственные, предложения, суждения, высказывания целиком (напр., выдержанные в одном и том же жанре); 4) при исследовании СМИ с точки зрения визуализации, например, исследуя фоторепортажи или визуальные элементы содержания газеты или интернета, в качестве единицы анализа мы должны учесть фотографию, рисунок, инфографику и т.п.; 5) каждое СМИ может быть исследовано посредством выбора в качестве единицы анализа фигуры (образа) или реализации тематического мотива, поскольку, как известно, каждое медиа служит в конце концов для передачи информации о том, что с кем произошло, где и когда.

3. В качестве единиц измерения, в зависимости от того, что в случае анализа конкретной единицы мы хотим измерить и что имеет смысл в случае конкретного СМИ, могут использоваться: 1) принимая во внимание площадь: квадратные сантиметры, условные точки на площади, знаки и пробелы, процент площади; 2) принимая во внимание время (напр., эфирное): минуты и секунды; 3) с точки зрения интенсивности: т.н. баллы экспозиционности, присваиваемые в зависимости от концентрации внимания получателя; 4) в качестве единиц измерения могут использоваться также текстовые слова, определяемые как последовательности знаков от пробела до пробела, предложения и всё, что может служить для измерения медийных текстов.

4. Что представляет собой лексический количественный контент-анализ (ЛККА)?

Лексический количественный контент-анализ представляет собой такой количественный анализ содержания (ККА), в котором как единицей анализа, так и единицей измерения является слово (Flont | Kajtoch | Pielużek 2015, 129 и сл.). К типам ЛККА, разрабатываемым в настоящее время в Польше, я отношу метод ключевых слов Писарека (2002), метод коллективных

символов Михаэля Фляйшера (см. Kołodziej 2011, 265 и сл.) (а в принципе только специальное использование результатов исследований, проведенных при помощи этих методов), аксиологическое измерение (Kajtoch 2008 | 1, 127 и сл.), метод слов-ключей (Jeżyk 2005, 100 и сл.; Wyka 1969, 153 и сл.), попытку продолжения исследований *General Inquirer* (Planeta 2018), некоторые типы количественного стилистического анализа (Flont | Kajtoch | Pielużek 2015, 133 in). Поскольку я описал эти методы в другой работе (Kajtoch 2018, 641 и сл.), а кроме того они были исчерпывающим образом охарактеризованы учеными, которые занимались подобными исследованиями, я остановлюсь здесь на кратком схематическом представлении сути этих исследований (Таб. 4). В таблице я не рассматриваю метода определения текстовой лексической количественной картины мира – этому вопросу я посвящаю отдельный фрагмент статьи (см. также Таб. 5). Легко заметить, что первой чертой, которая позволяет заметить разницу между перечисленными течениями ЛККА, является тот или иной выбор подсчитываемых слов (Pisarek 2016). Второй отличительной чертой следует считать цель исследования.

Таблица 4. Техники и исследовательские цели направлений лексического количественного, контент-анализа (ЛККА), применённого к анализу текстов или их корпусов

| Направление ЛККА | Техника проведения исследования | Цели | Форма полученного результата |
|---|---|--|---|
| 1 | 2 | 3 | 4 |
| Аксиологическое измерение | Поиск в данном корпусе тексте на предмет наличия в нем т.н. лексики со значением оценки (наименования ценностей, наименования оценивающих черт и существительных, называющих оцениваемые десигнаты) | Определение степени «насыщения» корпуса текстов текста определенными ценностями и антиценностями, реконструкция присутствующих в них систем оценки | Набор диаграмм, представляющих рейтинг присутствия данных типов оценки и интерпретация результатов, указывающая на наличие систем |
| Определение наличия коллективных символов | Обработка с помощью анкетирования состава слов, выявляющих наличие в мышлении определённой группы людей так называемых коллективных символов, то есть понятий, представляющих образ мировосприятия, характерного для данной культуры субкультуры, а затем исследование наличия этих слов в текстах. | Определение принадлежности данных текстов к актуальным для данной культуры субкультуры | Чаще всего сравнение рейтинговых или алфавитных списков слов |
| Определение наличия „знаменных слов” (слов, которые используют на знаменах) | Обработка с помощью анкетирования. Разрабатывается перечень лексем, репрезентирующих совокупность мифов и кондемандов*, нормативно предписанных для адептов данной идеологии, с последующим анализом их представленности в текстах. | Определение степени насыщенности данного текста корпуса (напр., всех номеров газеты за год) данной идеологией | Рейтинговые списки с указанием частотности |

Продовження таблиці 4

| 1 | 2 | 3 | 4 |
|--|---|---|--|
| Обнаружение наличия слов-ключей (<i>mots-clés</i>) | Выявление в тексте лексем с гипертрофированной частотностью, определение их «значимости» и приобретаемых в данном тексте значений, а также интуитивное обнаружение формулы, объясняющей тот или иной их набор. | Представление черт мировоззрения, зашифрованного в текстах | Рейтинговый список слов-ключей с указанием степени их значимости, сопровождаемый для читателя выводом в форме эссе, с многочисленными примерами из исследованного текста |
| Количественные исследования стилистики | Составление рейтингов частотности лексики данного текста в целом (повторяющихся полнзначных и неполнзначных слов), обнаружение в списках лексем, вводящих данные стилистические категории с последующим обобщением их частотности (в рамках тех же самых категорий) | Представление – главным образом с целью сравнения – степени реализации в тексте корпусе данной стилистической категории типа: поэтичности, директивности, научности, канцеляризма, разговорности, вульгарности и т.п. | Набор «стилистических полей», т.е. групп слов, реализующих данные стилистические категории (с указанием общей частотности), и диаграмма, иллюстрирующая степень важности этих категорий в тексте |
| Польский <i>General Inquirer</i> | Ранжирование корпуса, выбор существительных, занимающих самый высокий рейтинг и обслуживающих данную тему; определение взаимных семантических и количественных корреляций между ними. | Определение в данном корпусе тексте «зашифрованных» смыслов, не сводимых к содержанию фавулы, которые можно отчитать главным образом в процессе нелинейного восприятия. | Образ действительности, определяемой значениями лексем, представленный с какой-то конкретной точки зрения и под каким-то углом зрения в форме «облака слов» |

* Миранды – это репрезентируемые различными лексемами ценности, установки и явления, которые в рамках определённой идеологии интерпретируются как положительные и заслуживающие подражания. Они формируют семантическое ядро идеологической нормы и функционируют как маркеры лояльности к данной системе взглядов. Кондеманнды – это антиценности, репрезентируемые различными лексемами, то есть понятия, явления, установки или действия, подлежащие осуждению, критике или отрицанию. Они выполняют функцию негативной идеологической маркировки и закрепляют границы допустимого поведения и мышления для адептов данной идеологической системы.

Самым сложным и самым важным из исследовательских методов в рассматриваемом направлении является определение лексической количественной текстовой картины мира, представленной в наборе или в корпусе текстов. ЛКТКМ (лексическая количественная текстовая картина мира) представляет собой одну из разновидностей текстовой картины мира:

Текстовая картина мира (выступающая на уровне *parole*) представляет собой своеобразную, представленную в конкретном тексте или в наборе текстов

реализацию языковой картины мира (выступающей на уровне *langue*), то есть является множеством закономерностей, вытекающих из факта предпочтения в данном тексте или наборе текстов определенных морфологических, словообразовательных, синтаксических конструкций, а прежде всего – определенной лексики. Текстовая картина мира указывает на господствующий в данном тексте (наборе текстов) взгляд на существование и функционирование отдельных элементов мира, их взаимосвязей и взаимных пропорций, т.е. на такое понимание организации мира, существующих в нем иерархий и ценностей, которое является предпочтительным для отправителя данного текста и приемлемым для пользователей этого же текста [...] Каждая группа текстов, обладая индивидуально подобранной лексикой и так же индивидуально ее используя, строит свой текстовый мир отлично от других (Kajtoch 2008 | I, 15).

Определение ЛКТКМ можно проводить несколькими способами.

Таблица 5. Разные типы лексических количественных текстовых картин мира

| Тип ЛКТКМ | Способ проведения исследований | Цель исследований | Форма результата |
|---|--|---|--|
| 1 | 2 | 3 | 4 |
| ЛКТКМ в полном варианте – «карта» | Определение частотности, или всех повторяющихся полнозначных лексем, или – что представляется более верным – только существительных (включая – или нет – имена собственные) и прилагательных*, выступающих в тексте корпусе текстов и составление рейтинга, в котором будут учитываться только самые высокие позиции; разработка тематической классификации, включающей всю лексику из выборки | Определение в данном корпусе тексте «зашифрованных» смыслов, не сводимых к смыслам фавулы, которые можно обнаружить в нелинейном процессе восприятия. | Цельная картина действительности, представленная в форме «ментальной карты действительности» (более или менее подробной), состоящей из определенного набора понятийных полей (может быть – тематических) и указывающей на интенсивность обслуживания этих полей, выражающаяся в частотности текстовых слов, семантически приписанных этим полям. |
| ЛКТКМ в версии частичной – ближайшая картина данного тематического поля «карты» | Выбор в рамках общепринятой классификации одной из тематических или понятийных категорий, разделение ее на подкатегории и определение частотности относящейся к ней лексики | См. выше – только определяемые смыслы, связанные с одной темой | Картина одного из фрагментов действительности – подход, полезный в особенности в сопоставительных исследованиях типа «текстовый образ человека там-то и там-то» |

Продовження таблиці 5

| 1 | 2 | 3 | 4 |
|---|--|--------------------------|---|
| ЛКТКМ в версии концен- трированной «месседж» | Составление частотного списка, включающего 25-30 самых частотных существительных из данного текста корпуса, дополнение списка необходимыми глагола- ми и сочинение с ними истории. | Так же, как в рубрике 1. | Связный текст «анно- тации», передающий основ- ное содержание текста корпуса |

* Следует подсчитывать их вместе и помещать в одном списке. Прилагательных в корпусе обычно намного меньше, чем существительных, они редко занимают высокие места в рейтингах, иногда, однако, могут незначительно модифицировать результаты по сравнению с полученными только на основании оценки частотности существительных. Лучше всего, определяя полную и частичную версию ЛКТКМ, считать существительные и глаголы вместе.

5. «Карты» и месседжи

«Карты» создаются после предварительного тематического упорядочения лексики, представленной в исследуемом корпусе тестов, и последующей ее категоризации, причем в соответствующих разделах, категориях и подкатегориях размещаются данные о частотности лексем, обслуживающих семантическое поле, связанное с данным элементом классификации. Поскольку довольно сложно составить полную тематическую классификацию, в качестве ее основы можно использовать созданную ранее (Batko-Tokarz 2019). На практике таким образом исследуются более мелкие корпуса с классификацией только существительных или существительных и прилагательных. В этих случаях можно получить, например, такие результаты, которые получил я, применяя описанную выше процедуру к корпусу польской патриотической песни (см. Таб. 6).

Очередным этапом может быть, к примеру, составление классификации лексем, обслуживающих данное поле, и подсчет количества текстовых лексем, обслуживающих поля данных подкатегорий. Например, категорию «Сознание человека» можно разбить на подкатегории «Память», «Опыт», «Разум», «Эрудиция» и др. Таким образом мы получим то, что в Таблице 6 я определил как ЛКТКМ в версии частичной – приближенный образ данного тематического поля «карты».

Месседжи, однако, можно обнаружить в текстовых корпусах при получении рейтинговых списков существительных и последующем составлении текстов с существительными, занимающими самые высокие позиции

Таблица 6. Рейтинг тем в рамках ЛКТКМ на примере корпуса патриотических песен (Kajtoch 2021, 348)

| Рейтинг темы | Темы в корпусе | % в ЛКТКМ (100% = 5998 текстовых лексем) |
|--------------|------------------------------|--|
| 1. | Общество, общественная жизнь | 42,53 |
| 2. | Человеческое тело | 16,95 |
| 3. | Сознание человека | 10,00 |
| 4. | Время и его измерение | 5,41 |
| 5. | Произведения культуры | 5,20 |
| 6. | Цивилизация | 4,65 |
| 7. | Неживая природа | 4,31 |
| 8. | Основные категории быта | 4,11 |
| 9. | Чувства и эмоции | 2,81 |
| 10. | Пространство и его измерения | 2,45 |
| 11. | Живая природа | 2,11 |
| 12. | Личность и ее составляющие | 1,11 |
| 13. | Комические явления | 0,76 |
| 14. | Движение | 0,20 |
| 15. | Прочие | 0,15 |

(первые 25-30 мест). Чтобы «месседж» содержал какую-то информацию, по возможности как можно более объективную, существительные следует – по минимуму – дополнить словами других частей речи, подобранными в произвольном порядке, в результате чего сообщение будет помещено в какую-то рамку. Оформление таких рамок не должно нарушать информации, содержащейся в существительных. Известно, что именно существительные передают главную и достаточно однозначную информацию о мире, тогда как прилагательные говорят о свойствах, т.е. о ценностях (хотя и не всегда), а глаголы о действиях, движениях, процессах и т.п., но не точно; местоимения позволяют распределить элементы в пространстве и т.д. – но это всего лишь дополнение к основной информации, самой важной, повторяемой чаще всего, конкретной информации о мире, передаваемой наиболее частотными существительными.

«Месседжи» можно также составлять для отдельных, в особенности для больших текстов. В Таблице 7 я представляю иллюстрации к описанной выше методике, используя рейтинговый список нарицательных существительных, составленный несколько лет назад на основании 2 корпусов по 30 000 текстовых лексем для публикации «Перспективы количественного лексического контент-анализа» (Flont | Kajtoch | Pielużek 2015, 136-137).

Таблица 7. Существительные, занимающие первые 25 мест в 2 корпусах;
указано число встречающихся словоформ

| P | Польские фашистские зины | Портал <i>Pudelek</i> |
|----|--|---|
| 1 | 170 – <i>борьба</i> | 117 – <i>год</i> |
| 2 | 162 – <i>год</i> | 65 – <i>фильм</i> |
| 3 | 160 – <i>человек люди</i> | 59 – <i>актриса, ребенок</i> |
| 4 | 91 – <i>раса</i> | 52 – <i>время, жизнь</i> |
| 5 | 76 – <i>народ (нация), жизнь</i> | 45 – <i>сцена, связь</i> |
| 6 | 70 – <i>скин скинхед</i> | 40 – <i>актер, медиа</i> |
| 7 | 67 – <i>группа, движение (организация)</i> | 39 – <i>роль, фото</i> |
| 8 | 65 – <i>член (группы)</i> | 37 – <i>день, муж, личность</i> |
| 9 | 59 – <i>ансамбль</i> | 35 – <i>интервью</i> |
| 10 | 55 – <i>мир</i> | 34 – <i>таблоид</i> |
| 11 | 54 – <i>национал-социализм</i> | 31 – <i>женщина, место, певица</i> |
| 12 | 53 – <i>час (промежуток времени)</i> | 30 – <i>дом, программа, дело</i> |
| 13 | 46 – <i>война</i> | 29 – <i>план</i> |
| 14 | 44 – <i>страна</i> | 28 – <i>концерт, танец, неделя, выступление</i> |
| 15 | 43 – <i>организация</i> | 27 – <i>карьера</i> |
| 16 | 42 – <i>концерт</i> | 26 – <i>человек люди, месяц, журналист, выпуск (программы), слово, жена</i> |
| 17 | 40 – <i>народ (коллектив)</i> | 25 – <i>информатор, разговор</i> |
| 18 | 37 – <i>история</i> | 24 – <i>любовница</i> |
| 19 | 36 – <i>способ</i> | 23 – <i>шоу, брак</i> |
| 20 | 34 – <i>расизм, дело</i> | 22 – <i>пара (два человека), секс, причина, шоу-бизнес</i> |
| 21 | 31 – <i>журнал, работа, право</i> | 21 – <i>деньги, семья, любовные отношения</i> |
| 22 | 30 – <i>цивилизация</i> | 20 – <i>информация (сообщение), мир</i> |
| 23 | 28 – <i>деятельность, место, деньги</i> | 19 – <i>отец, продюсер, зритель, ситуация</i> |
| 24 | 26 – <i>музыка, религия</i> | 18 – <i>мероприятие, диск, работа, страница, друг, тысяча</i> |
| 25 | 25 – <i>часть, идея, задание</i> | 17 – <i>фан, злотый (валюта)</i> |

Источник: Flont | Kajtoch | Pielużek 2015, 136-137.

Рейтинговые списки позволяют сконструировать следующие «меседжи». В случае фашистских зинов:

Nasza **walka** to **ruch ludzi** ceniących swą **rasę** i **naród**. **Członkowie** naszej **grupy**, **skinheadzi** chcą, aby w **świecie** zapanował **narodowy socjalizm** i **rasizm**. Tej **sprawie** poświęcamy cały nasz **czas**, każdy **rok** naszego **życia**. **Miejscem** działalności jest nasz **kraj** i cały **świat**. Szanujemy nasz **lud**, jego **religię** i **historię**. Mamy **prawo**, by prowadząc **wojnę** o naszą **cywilizację** i **idee**, działać w różny **sposób**. Nasza **organizacja** urządza **koncerty**, na których śpiewają nasze **zespoły**. Ich **rzecz** to tworzyć **muzykę** i **pisma**. Nie chcą **pieniędzy**. To ich **część** naszej **pracy**.

Наша **борьба** – это движение людей, которые ценят свою **расу** и свой **народ**. **Члены** нашей **группы**, **скинхеды** хотят, чтобы в **мире** воцарился **национал-социализм** и **расизм**. Этому делу мы посвящаем все своё **время**, каждый год нашей **жизни**. **Местом деятельности** является наша **страна** и весь **мир**. Мы ценим наш **народ**, его **религию** и **историю**. Мы имеем **право** разными **способами** вести **войну** за нашу **цивилизацию** и **идеи**. Наша **организация** проводит **концерты**, на которых выступают наши **ансамбли**. Их **задание** – создавать **музыку** и **журналы**. Им не нужны **деньги**. Это **часть** нашей **работы**.

В случае портала для сплетников *Pudelek*:

Czas biegnie. Jak co **roku producenci** nakręcili dużo **filmów**, w których **aktorzy** i **aktorki** grają swoje **role**. Co **miesiąc** ukazują się **plyty**, na **scenach** podczas **show** widzimy **tańce par**, a podczas **koncertów** śpiewają **piosenkarki**. Patrzcie **fani**, jak biegnie **życie** tych **ludzi**, jak robią **karierę** w **świecie show-biznesu**, co mówią o nich **tabloidy** i inne **media**, ile **złotych** zarabiają, ile mają **pieniędzy**. Co **tydzień** można obejrzeć nowe, serialowe **odcinki**, w każdym **dniu** posłuchać **rozmów**, na naszych **stronach** przeczytać **wywiady**, zobaczyć **zdjęcia**. **Powód**, by nas czytać? – To proste. Nasi **informatorzy** i **dziennikarze** zdobyli dla was – **widzów** – **tysiące informacji** nie tylko o tym, jak wygląda **praca** znanych **osób**, jakie mają **plany występów**, na jakich **imprezach** się pokażą. Piszemy, jakie mają **domy**, **rodziny**, **przyjaciół**, także o ich ważnych sprawach, jak np. **ślub**. W tym **miejscu** znajdziesz też **słowo** o ich **romansach**, o tym, jak uprawiają **seks**. Dowiesz się jakimi są **mężami** dla swych **żon** (i jakie są ich inne **kobiety** – **kochanki**), **ojcami** dla swych **dzieci**. Pomyśl, jak w ich **sytuacji** byś się zachował.

Время бежит. Как каждый **год продюсеры** сняли много **фильмов**, в которых **актеры** и **актрисы** играют свои **роли**. Каждый **месяц** выходят **диски**, на **сценах** во время **шоу** мы видим **танцы пар**, а во время **концертов** поют **певицы**. Смотрите, **фаны**, как протекает **жизнь** этих **людей**, как они делают **карьеру** в **мире шоу-бизнеса**, что пишут о них **таблоиды** и другие **медиа**, сколько **злотых** они зарабатывают, сколько у них **денег**. Каждую **неделю** можно увидеть новые **выпуски** сериалов, каждый **день** прослушать **разговоры**, на наших **страницах** прочитать **интервью**, увидеть **фото**. Какая нужна **причина**, чтобы нас читать? – Это просто. Наши **информаторы** и **журналисты** добыли для вас – **зрителей** – **тысячи сообщений** не только о том, как выглядит **работа** известных **личностей**, какие у них **планы выступлений**, на каких **мероприятиях** их можно будет увидеть. Мы пишем о том, какие у них **дома**, **семьи**, **друзья**, а также о важных делах, таких как **брак**. В этом **месте** ты найдешь также несколько **слов** об их **любовных отношениях**, о том, как они занимаются **сексом**. Ты узнаешь о том, какими **мужьями** они являются для своих **жен** (и какие у них есть другие **женщины** – **любовницы**), какие они **отцы** для своих **детей**. Подумай о том, как бы ты себя повел в их **ситуации**.

Понятно, что при написании таких текстов требуется также предварительная информация о данном СМИ и о характере текстов, входящих в состав корпуса. Известно, что польская фашизирующая молодежь представляет собой очень маленькие группки, членам которых приходится прежде всего убеждать самих себя в том, что их существование имеет смысл, а порталы типа *Pudelek* наряду с публикацией сплетен занимаются саморекламой. Без этого знания нам бы не удалось придать текстам жанровый контекст, необходимый для каждого сообщения.

Кроме того, следовало бы поставить вопрос о том, возможны ли иные варианты подобного рода методики исследования ЛКТКМ. Я полагаю, что такие варианты возможны, однако различия между ними будут незначительные. В каждом из них будут заключаться те же «месседжи». В первом случае о том, что фашизм представляет собой общественно значимую идеологию и следует ее распространять, претворять в жизнь, а в другом – о том, что жизнь «звезд» привлекательна, следует его наблюдать, а самым лучшим инструментом для этого является портал *Pudelek*.

6. С какой целью проводятся описанные выше исследования

Как и любой вид контент-анализа, составление карт и «месседжей» для лексической количественной картины мира на первый взгляд подтверждает очевидные истины. Мы видим результаты и соглашаемся: «Ну, да, все правильно». Проблема, однако, состоит в том, что для подтверждения чего-либо следует сначала понять и принять факт существования чего-либо, иметь знания об этом – тогда как ЛКТКМ исследует ту сферу, которой отправитель и получатель не замечают, потому что никто не в состоянии точно оценить частотность употребления лексики в больших текстовых корпусах. А ведь мы существуем в кругу этих текстов. Больше того, кроме крайне ограниченного объема информации, доступной непосредственно нашим органам чувств, свои знания о мире мы получаем прежде всего из средств массовой информации, а говоря точнее, из текстовой информации, которую они нам предоставляют, значительная часть которой имеет языковой характер. Картинки, иллюстрации – как мне кажется – в основном навязывают эмоции и нуждаются в уточнении своего содержания, а это возможно только при посредстве естественного (или искусственного) языка, в форме текстов.

Текстовые картины мира имеют принудительный характер. Они оставляют свой отпечаток на нашем образе мышления, формируют его – подобно тому как натуральный язык формирует культуру символов. В итоге мы в состоянии принять во внимание в мире то только, что позволяет нам заметить культура, а наш натуральный язык – точно разместить в системе известных

нам понятий и назвать (и соответственно передать другим информацию об этом). По крайней мере, так выглядит радикальное понимание гипотезы Сэпира-Уорфа, сторонником которой я являюсь. Из факта существования во всех типах текстов (устных и любым способом записанных), а выражаясь точнее, в больших их собраниях (в том числе медийных) скрытых структур, которые можно обнаружить только посредством ЛККА, можно сделать вывод о том, что на этом дело не кончается, что наши познавательные возможности ограничены не только грамматикой и лексиконом известного нам языка, но также и самой сферой окружающих нас текстов. При этом сфера текстов, во-первых, сообщает нам скрытые истины, над которыми мы не властны, а во-вторых, способна нарушить – такую неуверенную и не до конца точную – коммуникацию с действительностью, обеспечиваемую нам языком и являющуюся объектом исследований прагмалингвистики, когнитивистики, а прежде всего – науки о языковых картинах мира.

7. Теория двух шлемов

Предположение о том, что в процессе человеческого познания существенную роль играет не только язык, но и его текстовые реализации, влияющие на наше мышление, в том числе посредством элементов, не имеющих исключительно языкового значения, таких как частотность употребления лексем, которая указывает на их значимость, экспозиционность и интенсивность исследования в данном тексте | корпусе текстов тематического | понятийного поля – заставило меня предпринять попытку введения в языковедческий (и медиаведческий) дискурс новой категории, которая принимала бы во внимание эти элементы. Я называл эту категорию текстовой картиной мира, и полагаю, что среди всех ее вариантов сильнее всего и результативнее всего – поскольку абсолютно не явно – на нас влияют лексические количественные текстовые картины мира.

Если учесть существование текстовых картин мира, как нам следует оценивать баланс приобретений и потерь в познании мира с помощью языка и его творений – текстов? Человек или создает тексты, или воспринимает – и это для него единственный возможный (и всегда опосредованный) контакт с системой языка, а следовательно, и с той частью интересубъективной символической культуры, с которой система языка связана неразрывно, являясь ее важнейшим звеном. Производя или воспринимая тексты, человек использует слова, функционирующие и на уровне текста, и на уровне системы. Таким образом, производя тексты, человек изучает систему, интериоризирует ее, а в результате расширяет свои знания в области культуры символов, свойственной его группе – группе пользователей языка – и получает определенные возможности участия в обогащении этой системы. В случае,

если использованный им неологизм будет создан в соответствии с правилами семиозиса, т.е. будет всем понятен, то, подхваченный другими носителями языка, закрепится в системе. В результате можно было бы признать, что между сознанием пользователя, уровнями текстовой картины мира (ТКМ) и языковой картины мира (ЯКМ), intersубъективными и интериоризированными, существуют взаимодействия (пусть только опосредованные в случае контакта интериоризированной в сознании ЯКМ с intersубъективной ЯКМ), причем ситуация выглядит вполне оптимистически. Человек порождает тексты, создавая сферу ТКМ, сфера ТКМ влияет на intersубъективную систему ЯКМ, которая под этим влиянием надлежащим образом упорядочивается. Под воздействием воспринимаемых одного за другим текстов сознание конкретного человека воспринимает принципы и лексикон intersубъективной ЯКМ, интериоризирует культуру и язык группы и одновременно участвует в создании культуры сообщества. Все это происходит по правилам особой обратной связи. Данный индивид (т.е. носитель данного естественного языка и связанной с ним культуры) самостоятельно, благодаря присутствию в его сознании интериоризированной формы ЯКМ создает тексты, сотрудничая с другими носителями языка – членами группы, которые также порождают и воспринимают тексты, формирует последующие реализации ТКМ, регистрируемые и обобщаемые в сфере intersубъективной ЯКМ и подчиненные действию присутствующих там механизмов. В конце эта несколько модифицированная форма ЯКМ, вновь поддаваемая интериоризации в сознаниях носителей языка, становится базой для очередного круга изменений: язык развивается и совершенствуется.

Единственное сомнение в этом буколическом пейзаже возникает тогда, когда мы задумываемся о соответствии ЯКМ (т.е. подсистемы символической культуры, связанной с языком) реальной действительности, то есть – с языковедческой точки зрения – той сфере действительности, в которой функционируют десигнаты. Это очень старая философская проблема, которая решается по-разному. Вспомним хотя бы концепцию солипсизма. Тем не менее мы с чистой совестью можем оставить проблему соответствия между языком и миром нерешенной, поскольку – во-первых – она нерешаема, а – во-вторых – как сложна бы она ни была, это не мешает человеческому языку функционировать как положено, если наша цивилизация все еще существует.

Следует, однако, помнить, что возможна и ситуация критического несоответствия языка | культуры и действительности. Она например имела место, как мне кажется, во время испанской конкисты (колонизации Америки). Мощные империи ацтеков и инков после нападения групп конкистадоров, насчитывающих несколько сот или самое большее несколько тысяч солдат, распались главным образом из-за того, что индейцы явно неадекватно

оценили действительность, приняв грабителей за богов, что находило обоснование в местных мифологических системах (т.е. и в ЯКМ, присущей их языку).

Человек черпает свои знания о мире не только из слов, он пользуется также визуальными представлениями и разными имеющимися у него органами чувств, существует также интуиция, а значит общие сведения, передаваемые языковой системой, в достаточной степени верифицируемы. Ситуация меняется, когда в игру входят средства массовой информации, большие корпуса текстов, большие группы людей и коммуникация становится на самом деле массовой. В принципе, я не вижу других, равно эффективных исследований смыслов, привносимых медийными бурями в человеческое сознание, которые могли бы дать в достаточной степени объективные результаты, кроме исследований текстовых картин мира, и именно лексических количественных текстовых картин мира. Существует крайне актуальная потребность проведения таких исследований, если мы на самом деле хотим знать, как СМИ распоряжаются в наших головах, какие системы мышления они чрезвычайно успешно нам предлагают. Здесь, естественно, мы услышим возражения исследователей визуальной информации, которая так же играет значительную роль в СМИ, я лично, однако, считаю, что передача информации при помощи языка не только по-прежнему доминирует, но и может быть подвергнута в достаточной степени точному изучению.

В настоящее время человек, стремящийся познавать окружающий мир, может рассчитывать на собственные чувства и опыт только в случае знакомства с самым близким к нему пространством и ее истории в перспективе, которую может охватить своей памятью. В других случаях в его распоряжении оказываются средства массовой информации в широком значении этого слова: не только газеты, радио, телевидение, интернет, но также школьные учебники, комиксы, популярные романы и фильмы. Вообще человек со всех сторон обстреливается мегабитами информации, которые несут ТКМ отдельных текстов, с которыми он имеет дело. Инструмент действует в одну сторону, потому что если столько времени и энергии приходится посвящать получению информации – то когда ее отправлять? Энтузиасты теории о развитии интерактивного характера масс-медиа не замечают этой простой зависимости.

Если человек отправляет информацию по остаточному принципу, он перестает участвовать не только в формировании ТКМ, но также перестает формировать свой язык и культуру. Кроме того, он постепенно перестает ее понимать, поскольку не упражняясь в понимании, утрачивает способность обобщения, формулировки выводов, а прежде всего обогащения лексического запаса собственного идиолекта. Ко всему прочему, поскольку информационное содержание атакующих человека текстов довольно примитивно

и противоречиво (скажем, утверждение о том, что каждый рекламируемый продукт является самым лучшим, просто абсурдно, ведь самым лучшим может быть только один продукт данного вида; прививки не могут одновременно защищать от болезней и их вызывать), а практика инфотейнмента и создания симулякров приводит к тому, что СМИ работают не на получателя, а на самих себя, окружающая человека информация, и прежде всего лексика начинает уничтожать человеческое знание, напоминают стальной шлем с забралом без отверстий для глаз. Получатель все реже понимает то, что читает, вместо этого он впитывает слоганы, лозунги, крылатые фразы, которые способствуют отчуждению его от инструментов, служащих познанию и вращению в культуру (ТКМ представляет собой не только реализацию, но и строительный материал для ЯКМ). Это первый шлем, который я вижу.

Итак, человек со шлемом на голове оказывается еще и под куполом. Этот купол называется интересубъективной языковой картиной мира и по существу почти идентичен культуре. Он и является вторым шлемом, отделяющим человека от действительности. Случается, что ЯКМ заставляет видеть мир искривленным (что может закончиться катастрофой), чаще однако превосходство культуры над человеком приводило к добру, культура оказывалась мощным инструментом приспособления *homo sapiens* к действительности. Однако культура не существует без своей самой важной составляющей, языка, а язык владеет собственной картиной мира (ЯКМ), которая остается в тесной связи с текстовыми картинами мира (ТКМ), стабилизирующими ее.

Из сказанного выше следует вывод о том, что нельзя безнаказанно провозглашать абсурд, распространять пропаганду, создавать постправду, заниматься только постмодернистической наукой, разрушать систему ценностей и т.д., потому что если беспорядки на уровне *parole* проникнут на уровень *langue*, наша культура перестанет соответствовать действительности. Результаты будут плачевны.

Библиография

- БАТКО-ТОКАРЗ, В. (2019), *Tematyczny podział słownictwa współczesnego języka polskiego: teoria, praktyka, leksykografia*. Kraków.
- BERELSON, B. (1970), *Analiza treści*, [w:] b.a. (red.), *Teoria i metodologia badań nad prasą za granicą (przekłady)*, 3. *Analiza zawartości*. Kraków, 5-62.
- FLONT, M. | КАЙТОХ, В. | ПІЕЛУЖЕК, М. (2015), *Perspektywy leksykalnej ilościowej analizy zawartości*, [w:] Kiklewicz, A. | Uchwanowa-Szmygowa, I. (red.), *Dyskurs: aspekty lingwistyczne, semiotyczne i komunikacyjne*. Olsztyn, 129-145.
- ЖЕЖЫК, А. (2005), *Najmłodsza polska poezja w świetle analizy zawartości*, [w:] *Zeszyty Prasoznawcze* 3-4, 100-116.
- КАЙТОХ, В. (2008), *Языковые образы świata и человека в прасе мłодzieżowej i alternatywnej*, т. 1-2. Kraków.

- KAJTOCH, W. (2018), O niektórych rodzajach badań prasy drukowanej, [w:] Jęczeń, J. | Guzdek, P. (red.), *Medioznawstwo personalistyczne*, t. 2: Wybrane zagadnienia z kultury mediów. Lublin, 625-668.
- KAJTOCH, W. (2021), Uwagi o współczesnej polskiej piosence patriotycznej, [w:] Kawka, M. | Prażuch, W. | Planeta, P. (red.), *Medialne i transmedialne procesy narracyjne w identyfikacji postaw patriotycznych Polaków*. Kraków, 333-356.
- KAJTOCH, W. (2023a), Leksykalna ilościowa analiza zawartości jako technika badania języka prasy, [w:] *Rocznik Medioznawczy* 2, 8-36.
- KAJTOCH, W. (2023b), Существует ли текстовая картина мира, [Sushchestvuyet li tekstovaya kartina mira.] [in:] *Przegląd Wschodnioeuropejski XIV* (1), 333-349.
- KŁOSKOWSKA, A. (1969), *Z historii i socjologii kultury*. Warszawa.
- KOŁODZIEJ, J. (2006), *Klucz kategoryzacyjny*, [w:] Pisarek, W. (red.), *Słownik terminologii medialnej*. Kraków, 95.
- KOŁODZIEJ, J. (2011), *Wartości polityczne. Rozpoznanie, rozumienie, komunikowanie*. Kraków.
- LICHAŃSKI, J. | KAJTOCH, W. | TROCHA, B. (2015), *Literatura i kultura popularna. Metody: propozycje i dyskusje*. Wrocław.
- PISAREK, W. (1983a), *Analiza zawartości prasy*. Kraków.
- PISAREK, W. (1983b), *Analiza zawartości prasy: drogi rozwoju*, [w:] *Zeszyty Prasoznawcze* 2, 5-16.
- PISAREK, W. (2002), *Polskie słowa sztandarowe i ich publiczność*. Kraków.
- PISAREK, W. (2016), *Słowa ważne i ważniejsze*, [w:] *Przegląd Humanistyczny* 3, 11-20.
- PLANETA, P. (2018), *Komputerowa analiza tekstu w dyskursach medialnych*, [w:] Szymańska, A. | Lisowska-Magdziarz, M. | Hess, A. (red.), *Metody badań medioznawczych i ich zastosowanie*. Kraków, 67-90.
- TETELOWSKA, I. (1960), *Analiza i ocena treści dzienników*, [w:] *Zeszyty Prasoznawcze* 1, 7-22.
- WYKA, K. (1969), *O potrzebie historii literatury. Szkice polonistyczne z lat 1944-1967*. Warszawa.

