

УДК 316

**О.С. Нагорный, А.Т. Мухетдинова**  
*Национальный исследовательский университет  
«Высшая школа экономики», г. Санкт-Петербург, Россия*

**ИССЛЕДОВАНИЕ ДИСКУРСА О БИОПЕДАГОГИКЕ  
ПРИ ПОМОЩИ ТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ  
И СИНТАКСИЧЕСКОГО АНАЛИЗА ТЕКСТОВ**

В данной работе мы хотели изучить, как дискурс о биопедагогике проявляет себя в Интернете, какие лингвистические средства для этого используются и какие темы затрагиваются.

Биопедагогика – термин тесно связанный с понятием био-власти, введённым в употребление Мишелем Фуко [1]. Фуко считал, что в современном обществе широко распространены практики контроля над телом отдельного человека и популяцией в целом, а человеческое тело является, таким образом, объектом различных манипуляций. Этот контроль проявляется, прежде всего, в дискурсе о теле и здоровье, который задаётся государством и другими институтами, прививающими людям определённые привычки и образ мыслей. В последние годы одним из основных каналов трансляции этого дискурса стали специализированные сайты в сети Интернет [2].

Эмпирической основой данного исследования послужили 1714 статей с сайта Лайфхакер (<https://lifehacker.ru/topics/health/>), посвящённые тематике здорового образа жизни. Этот сайт был выбран по причине того, что он является одним из самых популярных коллективных блогов с шестимиллионной аудиторией [3] и имеет раздел с чётко выраженной тематикой – советами о том,

как вести здоровый образ жизни, и, следовательно, является местом бытования этого дискурса.

Для определения того, на какие темы распадается дискурс о здоровом образе жизни, был использован метод вероятностного тематического моделирования. Мы использовали алгоритм LDA, конечным продуктом которого являются матрица вероятностей принадлежности слов к темам и матрица вероятностей принадлежности текстов к темам.

Результаты ручного кодирования тремя кодировщиками тематического решения на 15 тем показали, что тематика здорового образа жизни распадается на следующие направления: 1) уверенность в себе и своих силах, мотивация окружающих; 2) гигиена и защита организма; 3) эффективные упражнения, советы для новичков и профессионалов; 4) спортивные тренировки (плавание, бег, разогрев мышц); 5) кулинарные рецепты, правильное питание; 6) правильное питание, диета, вредная пища, вредные привычки; 7) здоровье, забота о себе; 8) здоровье глаз дома и в офисе, сидячий образ жизни; 9) исследования о пользе спортивного образа жизни; 10) сон и пробуждение, привычки, техники сна; 11) активный образ жизни, путешествия, спорт, полезные советы; 12) упражнения на все случаи жизни; 13) спортивное питание – набор массы и диеты; 14) спортивное приложения и фитнес-гаджеты; 15) курсы, обучающие здоровому образу жизни.

Более глубокое исследование дискурса о здоровом образе жизни включало изучение того, какие именно слова использовались при описании человека, тела и здоровья. Для этого все тексты были сегментированы на отдельные предложения, а затем, при помощи нейронной сети Syntaxnet [5], созданной исследователями из компании Google и обученной на размеченном Национальном корпусе русского языка SynTagRus из коллекции Universal Dependencies ([http://universaldependencies.org/#ru\\_syntagrus](http://universaldependencies.org/#ru_syntagrus)), был произведён синтаксический разбор каждого из 70149 получившихся предложений. Таким образом, мы установили синтаксические связи между словами внутри предложений.

Анализ получившихся структур показал, что существительное «тело» чаще всего описывается прилагательными «человеческий», «красивый», «здоровый» и «стройный». При этом над

ним предлагается совершать следующие действия: «расслабиться», «сохранять», «вернуть». Анализируя определения, описывающие понятие «человек», мы заметили, что среди них часто фигурируют слова «многие», «другие», «любой», что может свидетельствовать о применении такого риторического приёма как отсылка к большинству. Анализ списка существительных, которые описывались прилагательным «здоровый» («питание», «пища», «тело», «сон», «спортсмен»), позволяет сделать вывод, что выделенные с помощью тематического моделирования темы действительно присутствуют в тексте.

Такой подход – тематическое моделирование на первом, разведывательном этапе, и синтаксический анализ для более глубокого проникновения в суть явления – может быть использован для анализа различных типов дискурса.

## **Литература**

1. Фуко М. Рождение биополитики. Курс лекций, прочитанных в Коллеж де Франс в 1978–1979 учебном году / пер. с фр. А.В. Дьяков. СПб.: Наука, 2010.
2. Wright J. Biopower, Biopedagogies and the obesity epidemic // Faculty of Education – Papers (Archive). 2009. P. 1–14.
3. Как блог Лайфхакер удерживает 6-миллионную аудиторию с помощью советов. URL: <https://habrahabr.ru/company/surfingbird/blog/298242/>
4. Blei David M., Ng Andrew Y., Jordan Michael I. Latent Dirichlet Allocation // J. Mach. Learn. Res. 2003. March. Vol. 3. P. 993–1022.
5. Andor D. et al. Globally Normalized Transition-Based Neural Networks // arXiv:1603.06042 [cs]. 2016.