

Библиографический список

1. *Левкиевская Е.Е.* Восточнославянский мифологический текст: семантика, диалектология, прагматика: дис. ... д-ра филол. наук. М., 2007. 634 с.
2. *Васнева О.И.* Типы и способы номинации демонологической лексики (на материале пермских говоров) // Филологические науки. Вопросы теории и практики. 2013. № 4 (22), ч. 1. С. 37–39.
3. *Новичкова Т.А.* Русский демонологический словарь. СПб.: Петербургский писатель, 1995. 640 с.
4. *Власова М.Н.* Новая АБЕВЕГА русских суеверий. СПб.: Северо-Запад, 1995. 383 с.

ABOUT THE PRINCIPLES OF PRESENTATION THE MYTHOLOGICAL CHARACTERS MULTIWORD NOMINATIONS IN THE ELECTRONIC DICTIONARY

Rusinova I.I., Granova M.A.

Perm State University, 15, Bukireva st., Perm, Russia, irusinova@mail.ru

The authors substantiate the relevance of the electronic “Dictionary of the Perm Krai mythological stories” creation. The greatest lexicographical difficulty is the multiword nominations of the spirits-“owners” of the various loci (brownie, yard spirit, bath spirit, goblin, etc.). In the article a thematic classification of these units is provided and the question of the principles of these units presentation in the dictionary is raised. Analysis of the research papers and the existing dictionaries of the demonological vocabulary allow drawing a conclusion about the absence of clear principles of theoretical consideration of the mythological characters multiword nominations and their dictionary descriptions in modern science. In our opinion, such units should be included in the dictionary. Criteria for their selection are, firstly, a regular occurrence in many texts, and secondly, reflection of the character’s relevant features in their semantics.

Key words: mythological stories of the Perm Krai, demonological vocabulary, electronic dictionary.

УДК 005

ОНТОЛОГИЯ МАРКЕРОВ АКАДЕМИЧЕСКОГО СТИЛЯ¹

Филипсон С.К., Ланин В.В.

Национальный исследовательский университет «Высшая школа экономики», Россия, Пермь, ул. Студенческая, 38, vlanin@hse.ru

В докладе описывается подход к разработке онтологии, предназначенной для систематизации стилевых маркеров академического английского языка. Онтология содержит описание общелингвистических понятий, на основе которых

¹ Статья подготовлена в ходе проведения исследования № 17-05-0020 в рамках Программы «Научный фонд Национального исследовательского университета «Высшая школа экономики» (НИУ ВШЭ)» в 2017 г. и в рамках государственной поддержки ведущих университетов Российской Федерации “5-100”.

описываются конкретные стилевые конструкции. Представление онтологии в человекочитаемом и машиночитаемом форматах позволяет использовать ее как в процессе изучения предметной области, так и для автоматической обработки языкового материала.

Ключевые слова: маркер стиля, научная публикация, онтология.

Публикация результатов исследований в научных изданиях является одним из существенных показателей результативности ученых и исследовательских коллективов. Английский язык значительно расширяет аудиторию читателей, но для авторов, не являющихся носителями языка, подготовка таких материалов довольно часто вызывает сложность, т. к. текст должен удовлетворять нормам «академического английского». На данный момент экспертами разработан значительный объем методических материалов по академическому письму, проводятся специализированные учебные курсы. Анализ литературы по данной тематике показал, что приведенные рекомендации не систематизированы, а в некоторых случаях даже имеют явные внутренние противоречия. Следует также учесть, что многие из изданий имеют свои стилистические «традиции», которые необходимо учитывать при подготовке материалов. В настоящее время ведутся исследования текстов с применением компьютерных технологий, благодаря которым доступна обработка объемных корпусов текстов. Корпусные исследования дают эмпирический материал, на основании которого могут быть построены эталонные языковые модели и исследованы языковые закономерности, описаны языковые явления, свойственные той или иной области языка, т. е. выявлены так называемые «маркеры стиля».

На протяжении нескольких лет сотрудниками и студентами департамента иностранных языков и кафедры информационных технологий в бизнесе Пермского филиала «Высшей школы экономики» проводятся исследования стиля научной речи на английском языке с помощью специализированного программного обеспечения [1]. К настоящему времени выделенные маркеры стиля представлены в виде несистематизированного набора. Появилась острая необходимость систематизации маркеров стиля, притом используемый способ должен предусматривать возможности расширения и адаптации, так как язык является динамической и постоянно развивающейся системой. В качестве способа классификации набора маркеров было решено создать онтологический ресурс [2, 3], описывающий основные понятия лингвистики и выявленные маркеры стиля. Онтология не только позволит решить широкий спектр задач, но и сможет стать отличным инструментом для исследователя, изучающего академический английский, и основой автоматизированного поиска маркеров в тексте.

Библиографический список

1. *Strinyuk S. A., Shuchalova Y., Lanin V.* Academic Papers Evaluation Software, in: Application of Information and Communication Technologies (AICT), 2015 9th International Conference on, 14–16 Oct. 2015. Rostov-on-Don: IEEE, 2015. P. 506–510.
2. *Загорулько Ю.А.* Построение порталов научных знаний на основе онтологий // Вычислительные технологии. 2007. Т. 12. Спецвып. 2. С. 169–177.
3. *Kent D. Bimson, Richard D. Hull* Unnatural Language Processing: Characterizing the Challenges in Translating Natural Language Semantics into Ontology Semantics // Semantic Web 2015. P. 119–135.

ACADEMIC STYLE MARKER ONTOLOGY

Philipson .K., Lanin V.V.

National Research University Higher School of Economics, 38,
Studencheskaya st., Perm, Russia, vlanin@hse.ru

The report describes an approach to an ontology development. The ontology is intended to organize academic English style markers. Ontology contains two main parts: the first part is general linguistic concepts description and the second part is specific style constructions. The ontology representation in human readable and machine readable formats allows to use it as a data domain study recourse and as a basis for automatic language processing.

Key words: style marker, scientific paper, ontology.

УДК 81'33

МАШИННОЕ ОБУЧЕНИЕ В СОЦИОЛИНГВИСТИЧЕСКИХ ЗАДАЧАХ

Худякова Е.С.

Пермский государственный национальный исследовательский
университет, Россия, Пермь, ул. Букирева, 15,
khudiakova.es@gmail.com

Проблемой исследования является оценка эффективности методов машинного обучения при решении социолингвистических задач. Конкретно на основе метода машинного обучения с учителем (в учебной выборке тексты сведены до лексем, каждой автоматически присвоены частеречные характеристики) проведена попытка определения этничности автора текста. Результаты эксперимента показали, что только 1 группа текстов («русские») устойчиво определяются машиной. Тексты «татар» определяются, но с низкой вероятностью, тексты «коми-пермяков» отнесены в результате переобучения то к группе «русских», то «татар». Эффективность атрибуции этничности автора текста объяснима малым количеством параметров обучения, а также лингвистическим эффектом дисперсии языковых средств, обусловленным внутренними свойствами самого языка.

Ключевые слова: машинное обучение, социолингвистика, этничность, нейронная сеть, алгоритм обучения с учителем.