

УДК 004.89:004.93

DOI 10.24412/2413-7383-4-11

А. В. Ниценко, В. Ю. Шелепов

ФГБНУ «Институт проблем искусственного интеллекта»,
283048, г. Донецк, ул. Артема, 118-б

О РОДИТЕЛЬНОМ ПАДЕЖЕ СУЩЕСТВИТЕЛЬНОГО ПОСЛЕ ГЛАГОЛА В СЛУЧАЕ ЕГО ОМОНИМИИ С ВИНИТЕЛЬНЫМ ПАДЕЖОМ

A. V. Nicenko, V. Ju. Shelepov

Federal Budgetary State Scientific Institution «Institute of Artificial Intelligence Problems»,
Donetsk, Artema str., 118-b.

ABOUT THE GENITIVE AFTER A VERB IN THE CASE OF ITS HOMONYMY WITH THE ACCUSATIVE

В статье предложен способ автоматического снятия омонимии родительного и винительного падежа существительного, идущего вслед за глаголом. Результат определяется самим этим глаголом (глагольный критерий). Способ реализован в экспериментальном программном обеспечении для снятия омонимии.

Ключевые слова: обработка естественного языка,
родительный и винительный падеж вслед за глаголом, снятие омонимии.

The article suggests a method for automatic disambiguation of genitive and accusative for a noun following the verb. The result is determined by this verb itself (the verb criterion). The method is implemented in experimental software for homonymy resolution.

Keywords: natural language processing, genitive and accusative cases after the verb, homonymy resolution.

Введение

Наиболее типично употребление существительного в родительном падеже, когда оно стоит после другого существительного, которое поясняет. Пример: *Мы слушали доклад президента*. Разумеется, указанному родительному падежу может предшествовать согласованное с ним прилагательное или причастие. Пример: *Мы слушали доклад нового президента*. Типичными являются также цепочки родительных падежей, следующих за поясняемым существительным. Пример: *Мы слушали доклад президента Соединенных Штатов Америки*.

Родительный падеж и его омонимы после существительного – это отдельная тема. Настоящая работа посвящена случаю, когда родительный падеж существительного следует за глаголом и имеет в качестве омонима винительный падеж этого существительного. **Цель** – разработка методов и программы автоматического снятия упомянутой омонимии. Окно этой программы представлено на рисунке 1. О ее работе сказано ниже. Снятие омонимии является одной из основных проблем при автоматическом анализе текста и активно исследуется в настоящее время (см., например, [1-13]).

Результаты работы с текстами

Нами извлечены более двенадцати тысяч соответствующих предложений, содержащиеся в Национальном корпусе русского языка со снятой омонимией [14]. Их использование дополнено результатами собственной работы с русскоязычным материалом. Вот список глаголов, после которых обнаруживается только родительный падеж.

алкать, беречься, бояться, вдохнуть, вкладываться, вколоть, вломить, вскипятить, выдавиться, выйти, выпасть, выпивать, глотнуть, гнущаться, держаться, дичиться, добавиться, добиваться, добиться, дождаться, дожидаться, доискиваться, докричаться, долить, помогать, допускать, допустить, допытываться, достаться, достигать, достигнуть, достичь, дотронуться, жаждать, желать, желаться, заботиться, ждать, заслуживать, застесняться, захотеться, зачерпнуть, избегнуть, избежать, исполниться, испробовать, испросить, испугаться, истечь, казаться, капнуть, касаться, коснуться, курнуть, лакнуть, лишать, лишаться, лишить, лишиться, ловить, набежать, набирать, набираться, набиться, наболтать, набрать, набрести, набулькать, набухать, набухаться, навезти, наvertеть, навидаться, наворовать, наглотаться, наглядеться, надавать, наддавать, надергать, надоить, надраться, надуться, наедаться, наестся, наехать, нажарить, нажеваться, нажевывать, нажраться, накашивать, накидаться, накипятить, наколдовать, накопиться, накосить, накупить, накуриться, накушаться, налакаться, нализаться, наломать, намесить, наместить, наметать, намешать, намолотить, намять, нанести, наносить, нанюхаться, напасать, напечь, напиваться, напиться, напороть, напринимать, напугаться, напускать, напустить, нарасти, нарубить, наскребать, наскрести, наслушаться, насмотреться, насобирать, насолить, насосаться, настрелять, насушить, насыпать, натаскать, натаять, натворить, натерпеться, натрескаться, нахватать, нахвататься, нахлебаться, нацедить, начитаться, недодать, недоедать, недооценивать, недооценить, недополучить, недопонимать, недоставать, недостать, недосчитаться, недосчитываться, нюхнуть, обкушаться, объесться, ожидать, оказаться, оказываться, опасаться, оробеть, ослушаться, остаться, остерегаться, переест, плеснуть, поддавать, поддать, подлить, подождать, поесть, пожелать, покушать, понабравшись, попить, пригубить, прихлебнуть, пролить, сгрудиться, сойти, сохнуть, становиться, стать, толпиться, умереть, умирать, хлебнуть, хлопнуть, чаять, черпануть, черпнуть. (1)

Далее приведем список глаголов, после которых обнаруживается как родительный, так и винительный падеж, но родительный имеет приоритет (встречается чаще, чем винительный): бежать, брать, вкусить, возжаждать, возжелать, выждать, выжидать, выпить, выпросить, давать, дать, доливать, доставлять, достать, жажнуть, заварить, завезти, заготовить, задавать, задать, заказать, закупить, залить, заменить, запросить, заслужить, захватить, захотеть, избегать, искать, испить, испрашивать, кланяться, купить, кушать, молить, набить, наболтать, набраться, набросать, навалить, наварить, нагнать, наговорить, нагреть, наддать, наделать, нажать, накапать, накапливать, накапливаться, накидать, накладывать, накопать, накопить, накопиться, накапливаться, накрошить, наливать, налить, нарвать, настелить, насыпать, натаскивать, натащить, натечь, натолкать, начистить, оставить, отварить, отведать, отведывать, откусить, отломить, отпивать, отпить, отсыпать, отхлебнуть, отчерпнуть, перелить, переложить, плюнуть, подать, подбавить, подбавлять, подвалить, поджарить, поджидать, подкинуть, подкладывать, подливать, подложить, поднести, подносить, подпустить, подрезать, подсыпать, пожалеть, пожаловать, поискать, пойти, покупать, положить, понабрать, понавезти, понавешать, понагнать, понапихать, поносить, понюхать, понять, пообещать, попортить, попробовать, попросить, попытать, пососать, поставить, постелить, постлать, посулить, потребовать, потянуть, поубавить, похлебать, почерпнуть, предлагать, предложить, прибавить, привезти, придать, приказать, прикупить, притащить, прихватить, прождать, произвести, просить, пытаться, развести, раздобыть, разлить, разрешать, сварить, сдать, сделать, слопать, снести, согреть, содержать, спрашивать, спросить, съесть, требовать, убавить, увидеть, удостаивать, удостоить, ухватить, хватануть, хватать, хватить, хлебать, хотеть. (2)

В этом случае наша программа будет по умолчанию выдавать родительный падеж. В случае необходимости пользователь сможет заменить его на винительный путем двойного щелчка в окне программы по строчке с нужным омонимом.

Наконец приведем список глаголов, после которых обнаруживается как родительный, так и винительный падеж, но приоритет имеет винительный падеж (встречается чаще, чем родительный):

вкладывать, вставить, втянуть, выдавать, выдать, выделить, выжать, вылить, вынести, выставлять, добавить, добыть, затратить, затребовать, испечь, испытать, качнуть, лизнуть, навернуть, называть, наловить, наставить, настрогать, настругать, обращать, оказывать, опережать, определять, оставлять, осудить, осушить, отвести, отвлекать, открыть, отметить, отнять, оторвать, отправить, оценить, показать, покинуть, получать, получить, построить, править, превышать, преградить, предать, предвидеть, пресечь, придавать, прикрывать, прикрыть, приметить, принести, приносить, производить, проявить, радовать, разделить, свалить, скрывать, сорвать, составить, сыскать, тащить, уделять, удержать, уменьшать, уничтожать, утишить, утратить, чувствовать, являть, взяться, подойти, взять, воздать, выдать, доставать, жалеть, ждать, зацепить, кинуть, наворотить, наложить, обойти, отпустить, отрезать, отрезать, перебрать, вложить, доставить, загубить, замешать, замутить, занести, запасать, навести, наводить, навязать, нагонять, наметать, недоговорить, обдумать, оплатить, перенести, поделать, подогреть, понаделать, послать, потратить, привозить, приготовить, прислать, присыпать, присыпать, прихлебнуть, причинить, продать, сбавить, скушать, спроворить, схватить, таить, уделить, цапнуть. (3)

Отметим также случаи, когда существительное в родительном падеже после глагола имеет кроме омонима в винительном падеже также омоним в именительном

падеже. Это возможно после большинства непереходных глаголов. В этом случае омонимия снимается на основании согласованности глагола.

Примеры:

Его походка исполнилась упругости(род) и даже некоторой прыти. Заводу исполнилось полвека(им).

Она испугалась старухи(род). Испугались старухи(им).

Об инверсиях

Будем называть изменение естественного порядка слов инверсией. Такие изменения характерны в первую очередь для поэтических текстов. Представляется невозможным использовать глагольный критерий для снятия омонимии родительного падежа при наличии инверсии существительных. За счет инверсии родительный падеж можно поставить вслед за любым глаголом. Пример:

Ночами он видел облик подруги – Ночами он видел подруги(род) облик.

Однако такое снятие возможно при наличии инверсии с числительным. Пример:

Мы видим два персонажа – Мы видим персонажа(род) два.

Это можно программно реализовать не только для числительных, но и для их комбинаций. Пример:

Мы поработали часа(род) два-три.

Здесь в скобках указаны падежи, выбираемые нашей программой.

Омонимия снимается также в случае инверсии с предлогом. Пример:

Биатлоном я занялся шутки(род) ради.

Здесь применяется алгоритм, основанный на следующем правиле:

Если в предложении есть предлог родительного падежа и за ним нет существительного в родительном падеже, то следует двигаться к началу предложения, пока мы не найдем омоним родительного падежа. Наша программа выберет для этого омонима родительный падеж.

О предложениях с отрицанием

В предложениях с отрицанием родительный падеж выбирается в следующих случаях [15]:

- 1) со словом *нет* в функции предикатива (*У него нет сомнения*)
- 2) при сочетании отрицания с глаголом *иметь* (*не имеет права, не имеет значения*).
- 3) при наличии перед глаголом с отрицанием слов: *никто, ничто, никакой, ни один* (*никто не взял ответственности на себя, ни одной не прочитал статьи*).
- 4) при глаголах восприятия, мысли: *видеть, увидеть, слышать, услышать, осязать, ощущать, ощутить, нюхать, понюхать, воспринимать, воспринять, осмысливать, осмыслить, чувствовать, почувствовать* (*не видел дороги, не чувствовал боли, не ощущал тревоги*).

Винительный падеж при наличии отрицания выбирается в следующих случаях:

- 1) при наличии после существительного придаточного предложения со словом *который* (*он не прочитал книги, которые я ему дал*)
- 2) если существительное является одушевленным (*медики не обследовали младенцев*)
- 3) Если отрицание входит в состав частицы *едва не, чуть не, чуть-чуть не* (*она чуть не уронила книги*)

- 4) если существительное относится к инфинитиву, отделенному от глагола с отрицанием другим инфинитивом (*он не хочет начинать писать статьи*)
- 5) если отрицание выражено словами *некому, некого, негде, некуда* (*некому показать записки*)

О работе программы

Окно нашей программы снятия омонимии выглядит следующим образом (рисунок 1).

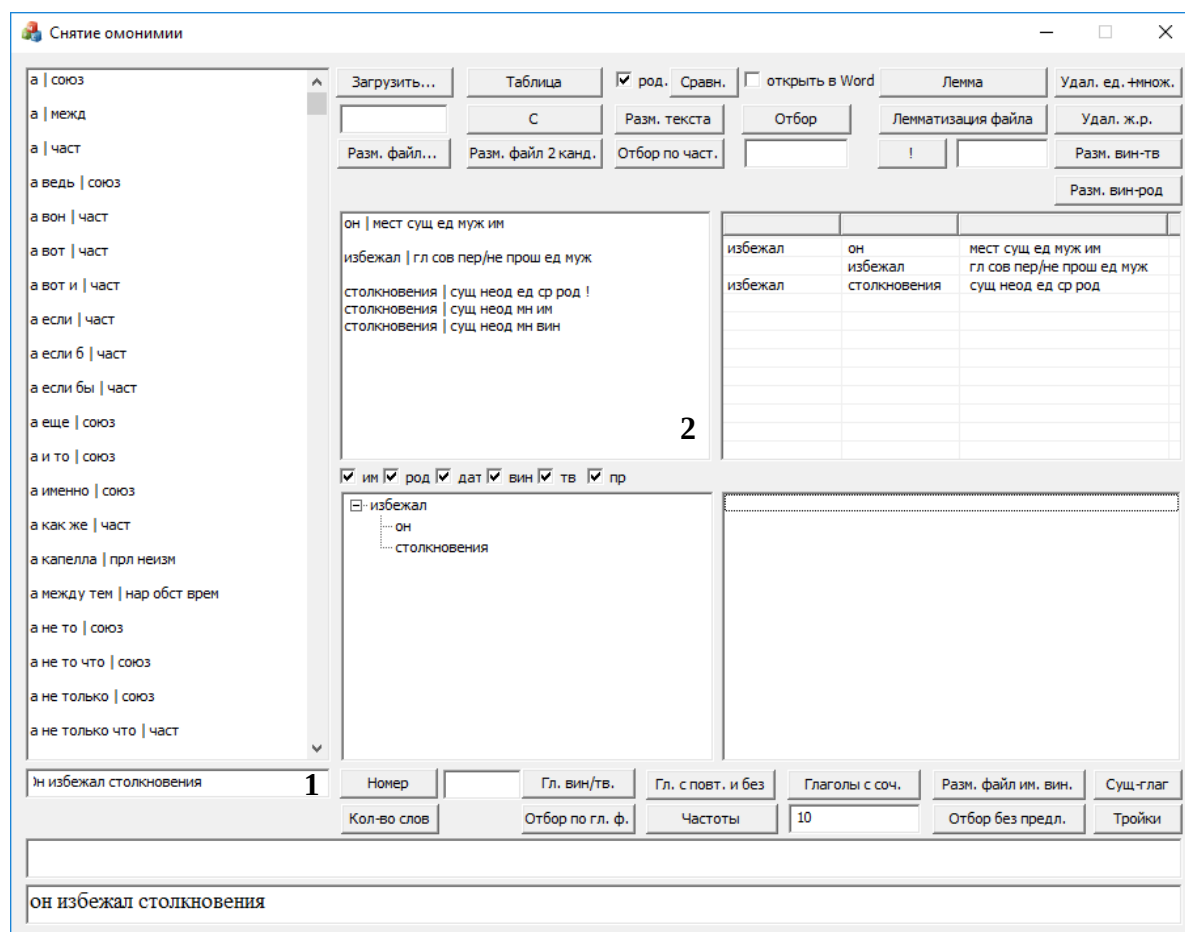


Рисунок 1 – Результат снятия омонимии для предложения «он избежал столкновения»

Предложение для обработки вводится с клавиатуры в поле 1. При нажатии кнопки «таблица» или клавиши «Enter» в поле 2 программа (в реальном времени) выводит в виде столбца последовательность групп соответствующих омонимов, снабжая выбранный омоним восклицательным знаком. Работа с перечисленными выше глаголами (1)-(3) осуществляется с помощью текстового файла *Род после глагола с последующим омон род-вин.txt*. Если омониму, с которым нужно определиться, предшествует форма какого-либо из глаголов (1), то программа выбирает для омонима родительный падеж. То же происходит и для глагола из (2). Но в этом случае пользователю, возможно, придется самостоятельно изменить выбор падежа на винительный путем двойного щелчка мышкой по нужной строчке в группе омонимов. Если глагол не попадает в (1) или (2), то для нашего омонима автоматически выбирается винительный падеж. Это статистически наиболее частый случай.

Дополнительные замечания

а) Отметим, что приведенные результаты распространяются на случай, когда между глаголом и существительным присутствует прилагательное или причастие, согласованное с существительным. Пример:

Он избежал лишней работы(род).

б) Замечание о словах с частицей пол (полу)

В предложении «Он выпил полстакана молока» слово полстакана стоит в винительном падеже, так же как в предложении «Он выпил стакан молока».

в) После слов МАЛО, МНОГО, МЕНЬШЕ, БОЛЬШЕ, МЕНЬШЕ ВСЕГО, БОЛЬШЕ ВСЕГО, МЕНЕЕ ВСЕГО, БОЛЕЕ ВСЕГО, ЧУТОЧКУ употребляется родительный падеж. Пример:

Он много уделяет внимания(род) занятиям с детьми.

Пример результата работы нашей программы отражен в рисунке 1.

Заключение

В статье приведено описание способа автоматического снятия омонимии родительного и винительного падежа существительных, следующих за глаголами. Разработанные правила были реализованы с использованием языка программирования C++ в экспериментальном программном обеспечении для снятия омонимии. Полученные результаты могут быть использованы для повышения точности морфологической разметки при автоматической обработке текстов на русском языке.

Список литературы

1. Боярский, К. К. Использование словарной информации при анализе текста / К. К. Боярский, Е. А. Каневский, С. К. Стафеев // Научно-технический вестник информационных технологий, механики и оптики. – 2012. – № 3(79). – С. 87-91.
2. Зарипова, Д. А. Подходы к автоматическому разрешению многозначности на основе неравномерности распределения значений слов в корпусе / Д. А. Зарипова, Н. В. Лукашевич // Вестник Московского университета. Серия 9: Филология. – 2023. – № 6. – С. 40-51.
3. Сорокин, А. Б. Автоматическая морфологическая разметка текста / А. Б. Сорокин, Л. М. Железняк // Научно-технический вестник Поволжья. – 2020. – № 2. – С. 44-46.
4. Бручес, Е. П. Гибридный подход к снятию частеречной омонимии на примере русского языка / Е. П. Бручес // Материалы 54-й Международной научной студенческой конференции МНСК-2016: Прикладная лингвистика, Новосибирск, 16–20 апреля 2016 года. – Новосибирск: Новосибирский национальный исследовательский государственный университет, 2016. – С. 13-14.
5. Литвинов, М. И. Метод повышения качества снятия омонимии с использованием статистики совместного употребления слов / М. И. Литвинов // Качество. Инновации. Образование. – 2010. – № 11(66). – С. 53-58.
6. Черепанова, Ю. Ю. Разрешение неоднозначности слов естественного языка / Ю. Ю. Черепанова // Восточно-Европейский журнал передовых технологий. – 2010. – Т. 5, № 2(47). – С. 50-52.
7. Метод разрешения частеречной омонимии на основе применения корпуса синтаксической сочетаемости слов в русском языке / Э. С. Клышинский, Н. А. Кочеткова, М. И. Литвинов, В. Ю. Максимов // Научно-техническая информация. Серия 2: Информационные процессы и системы. – 2011. – № 1. – С. 31-35.
8. Гашков, А. В. Математическое моделирование омонимии имени существительного в тексте и снятие омонимии приближенными методами (на материале русского языка) / А. В. Гашков // Вестник Санкт-Петербургского университета. Серия 9. Филология. Востоковедение. Журналистика. – 2012. – № 1. – С. 96-100.
9. Хорошилов А. А. Контекстное разрешение омонимии на основе центроидно-контекстной модели / А.А. Хорошилов, Ю.В.Никитин, А.В. Кан, Я.Д.Козловская, Е.А.Евдокимова // Труды ИСП РАН. – 2022. – т. 34, вып. 5. – С. 171-182.
10. Каневский, Е. А. К вопросу о снятии омонимии при компьютерном анализе текстов / Е. А. Каневский, К. К. Боярский // Проблемы преобразования и регулирования региональных социально-

- экономических систем / Под редакцией С.В. Кузнецова. Вып. 45. – Санкт-Петербург: Санкт-Петербургский государственный университет аэрокосмического приборостроения, 2019. – С. 37-41.
11. Афиногенов, Т. Ю. Решение проблемы лексической многозначности русскоязычного текста при помощи тезауруса / Т. Ю. Афиногенов, А. Ю. Бородащенко, Д. Л. Жусов // Информационные системы и технологии. – 2020. – № 6(122). – С. 70-78.
 12. Ниценко, А. В. Об использовании семантической информации для снятия омонимии именительного и винительного падежа (как элемента создания онтологии) [Текст] / А.В. Ниценко, В.Ю. Шелепов // Проблемы искусственного интеллекта. – 2024. – № 4 (35). – С. 16–24. – ISSN 2413-7383. – DOI 10.24412/2413-7383-2024-4-16-24
 13. Ниценко, А. В. О снятии омонимии «предикатив-предложная группа» для некоторых русских словосочетаний [Текст] / А. В. Ниценко, В. Ю. Шелепов, С. А. Большакова // Проблемы искусственного интеллекта. – 2023 г. – № 2 (29). – С. 49–57. – ISSN 2413-7383.
 14. Национальный корпус русского языка [Электронный ресурс]. – Режим доступа: URL: <http://www.ruscorpora.ru/> (дата обращения: 19.03.2025).
 15. Розенталь Д. Э. Справочник по русскому языку. Управление в русском языке. – М.: ОНИКС 21 век: Мир и Образование, 2003. – 304 с.

References

1. Boyarsky, K. K. Using dictionary information in text analysis / K. K. Boyarsky, E. A. Kanevsky, S. K. Stafeyev // Scientific and technical bulletin of information technologies, mechanics and optics. - 2012. - No. 3 (79). - P. 87-91.
2. Zaripova, D. A. Approaches to automatic resolution of polysemy based on the uneven distribution of word meanings in a corpus / D. A. Zaripova, N. V. Lukashevich // Bulletin of Moscow University. Series 9: Philology. - 2023. - No. 6. - P. 40-51.
3. Sorokin, A. B. Automatic morphological tagging of text / A. B. Sorokin, L. M. Zheleznyak // Scientific and technical bulletin of the Volga region. – 2020. – No. 2. – P. 44-46.
4. Bruches, E. P. Hybrid approach to removing part-of-speech homonymy using the example of the Russian language / E. P. Bruches // Proceedings of the 54th International Scientific Student Conference ISSC-2016: Applied Linguistics, Novosibirsk, April 16–20, 2016. – Novosibirsk: Novosibirsk National Research State University, 2016. – P. 13-14.
5. Litvinov, M. I. Method for improving the quality of homonymy removal using statistics of joint use of words / M. I. Litvinov // Quality. Innovations. Education. – 2010. – No. 11(66). – P. 53-58.
6. Cherepanova, Yu. Yu. Disambiguation of words in natural language / Yu. Yu. Cherepanova // East European Journal of Advanced Technologies. - 2010. - Vol. 5, No. 2 (47). - P. 50-52.
7. Method of resolving part-of-speech homonymy based on the use of the corpus of syntactic combinability of words in the Russian language / E. S. Klyshinsky, N. A. Kochetkova, M. I. Litvinov, V. Yu. Maksimov // Scientific and technical information. Series 2: Information processes and systems. - 2011. - No. 1. - P. 31-35.
8. Gashkov, A. V. Mathematical modeling of noun homonymy in the text and disambiguation by approximate methods (based on the Russian language) / A. V. Gashkov // Bulletin of St. Petersburg University. Series 9. Philology. Oriental studies. Journalism. - 2012. - No. 1. - P. 96-100.
9. Khoroshilov A. A. Contextual resolution of homonymy based on the centroid-context model / A. A. Khoroshilov, Yu. V. Nikitin, A. V. Kan, Ya. D. Kozlovskaya, E. A. Evdokimova // Proceedings of the ISP RAS. - 2022. - Vol. 34, issue. 5. - P. 171-182.
10. Kanevsky, E. A. On the issue of removing homonymy in computer analysis of texts / E. A. Kanevsky, K. K. Boyarsky // Problems of transformation and regulation of regional socio-economic systems / Edited by S. V. Kuznetsov. Issue. 45. – Saint Petersburg: Saint Petersburg State University of Aerospace Instrumentation, 2019. – P. 37-41.
11. Afinogenov, T. Yu. Solving the Problem of Lexical Polysemy of Russian-Language Text Using a Thesaurus / T. Yu. Afinogenov, A. Yu. Borodashchenko, D. L. Zhusev // Information Systems and Technologies. – 2020. – No. 6(122). – P. 70-78.
12. Nitsenko, A. V. On the Use of Semantic Information to Remove Homonyms between the Nominative and Accusative Cases (as an Element of Ontology Creation) [Text] / A.V. Nitsenko, V.Yu. Shelepov // Problems of Artificial Intelligence. – 2024. – No. 4 (35). – P. 16–24. – ISSN 2413-7383. – DOI 10.24412/2413-7383-2024-4-16-24
13. Nitsenko, A. V. On the removal of the “predicative-prepositional group” homonymy for some Russian phrases [Text] / A. V. Nitsenko, V. Yu. Shelepov, S. A. Bolshakova // Problems of Artificial Intelligence. – 2023. – No. 2 (29). – P. 49–57. – ISSN 2413-7383.
14. National Corpus of the Russian Language [Electronic resource]. – Access mode: URL: <http://www.ruscorpora.ru/> (date of access: 19.03.2025).
15. Rosenthal D. E. Handbook of the Russian language. Management in the Russian language. – М.: ONIX 21st century: Peace and Education, 2003. – 304 p.

РЕЗЮМЕ

А. В. Ниценко, В. Ю. Шелепов

О родительном падеже существительного после глагола в случае его омонимии с винительным падежом

Одной из ключевых задач в области обработки естественного языка является автоматическое снятие неоднозначности слов в текстах. Она заключается в выборе того значения многозначного слова, в котором оно употреблено в конкретном контексте. Неоднозначность, свойственная естественному языку, является серьёзным препятствием для компьютерного анализа текстов.

В статье приведено описание способа автоматического снятия омонимии родительного и винительного падежа существительных в сочетаниях с глаголами на основе контекстных правил.

Предложен способ автоматического снятия омонимии родительного и винительного падежа существительного, идущего вслед за глаголом. Результат определяется самим этим глаголом (глагольный критерий). Способ реализован в экспериментальном программном обеспечении для снятия омонимии.

Разработанные правила были реализованы с использованием языка программирования C++ в экспериментальном программном обеспечении для снятия омонимии. Полученные результаты могут быть использованы для повышения точности морфологической разметки при автоматической обработке текстов на русском языке.

RESUME

A.V. Nicenko, V. Ju. Shelepov

About the genitive after a verb in the case of its homonymy with the accusative

One of the key tasks in natural language processing is automatic disambiguation of words in texts. It consists of selecting the meaning of a polysemantic word in which it is used in a specific context. The ambiguity inherent in natural language is a serious obstacle to computer analysis of texts.

The article describes a method for automatic disambiguation of the genitive and accusative cases of nouns in combinations with verbs based on contextual rules.

A method is proposed for automatic disambiguation of the genitive and accusative cases of a noun following a verb. The result is determined by the verb itself (verb criterion). The method is implemented in experimental software for disambiguation.

The developed rules were implemented using the C++ programming language in experimental software for disambiguation. The results can be used to improve the accuracy of morphological markup in automatic processing of texts in Russian.

Ниценко Артём Владимирович – к.т.н., Федеральное государственное бюджетное научное учреждение «Институт проблем искусственного интеллекта», г. Донецк. *Область научных интересов:* Искусственный интеллект. Обработка естественного языка. Компьютерная лингвистика. Эл. почта: nav_box@mail.ru, адрес: 283048, г. Донецк, ул. Артема, д. 118 б, телефон: +7949 4315356.

Шелепов Владислав Юрьевич – д.ф.-м.н., профессор, Федеральное государственное бюджетное научное учреждение «Институт проблем искусственного интеллекта», г. Донецк. *Область научных интересов:* Искусственный интеллект. Обработка естественного языка. Компьютерная лингвистика. Эл. почта: vladislav.shelepov2012@yandex.ru, адрес: 283048, г. Донецк, ул. Артема, д. 118 б, телефон: +7949 4219188.

Статья поступила в редакцию 11.01.2025.