

Topośław—narzędzie służące tworzeniu elektronicznych słowników fleksyjnych jednostek wieloczłonowych

Piotr Sikora

IPI PAN

22 lutego 2016

Cel

Wszystkie formy jednostki wielocłonowej wraz ze znacznikiem morfosyntaktycznym. Np. *sztuczne ognie*:

sztuczne ognie ◀ sztuczne ognie:subst:pl:nom:m3

sztucznych ogni ◀ sztuczne ognie:subst:pl:gen:m3

sztucznych ogniów ◀ sztuczne ognie:subst:pl:gen:m3

sztucznym ogniom ◀ sztuczne ognie:subst:pl:dat:m3

sztuczne ognie ◀ sztuczne ognie:subst:pl:acc:m3

sztucznymi ogniami ◀ sztuczne ognie:subst:pl:inst:m3

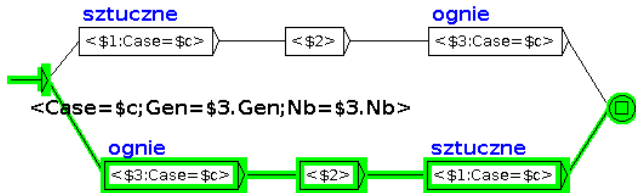
sztucznych ogniach ◀ sztuczne ognie:subst:pl:loc:m3

sztuczne ognie ◀ sztuczne ognie:subst:pl:voc:m3

ognie sztuczne ◀ sztuczne ognie:subst:pl:nom:m3

itd.

Opis odmiany jednostek wieloczłonowych



- Graficzna reprezentacja paradygmatów odmiany w formie grafów.
- *Multiflex* Agaty Savary — formalizm, narzędzie odmiany jednostek wieloczłonowych.
- *Morfeusz SGJP* — analiza morfologiczna, generacja form pojedynczych fleksemów.

Wkrótce Werbosław

- Hasła na poziomie leksemów.
- Grupy grafów odmiany dla leksemów.
- Automatyczna generacja fleksemów.
- Morfeusz 2

Wymagania techniczne

- Java w wersji 8 lub nowszej.
- Systemy operacyjne Windows lub Linux.
- Wolnodostępny na licencji GPL v 3.

Bibliografia

- SAVARY, A., RABIEGA-WIŚNIEWSKA, J., WOLIŃSKI, M. (2009): Inflection of Polish Multi-Word Proper Names with Morfeusz and Multiflex, in MARCINIAK, M., MYKOWIECKA, A. (eds.) *Aspects of Natural Language Processing*, Lecture Notes in Computer Science 5070, Springer Verlag, pp. 111–141.
- MARCINIAK, M., RABIEGA-WIŚNIEWSKA, J., SAVARY, A., WOLIŃSKI, M., HELIASZ, C. (2009): Constructing an Electronic Dictionary of Polish Urban Proper Names, in *Recent Advances in Intelligent Information Systems (Proceedings of the Balto-Slavonic Natural Language Processing Workshop, Kraków)*, Academic Publishing House EXIT, Warsaw, pp. 743–749.
- GRALIŃSKI, F., SAVARY, A., CZEREPOWICKA, M., MAKOWIECKI, F. (2010): Computational Lexicography of Multi-Word Units: How Efficient Can It Be?, in *Proceedings of Multiword Expressions: from Theory to Applications (MWE 2010)*, Workshop at COLING 2010, Beijing, China, August 28.
- CZEREPOWICKA, M., KOSEK, I. (2011): Problemy opisu związków frazeologicznych w formalizmie „Multifleks” (na przykładzie rodzaju wyrażen frazeologicznych), in Kopcińska, D., Bańko, M. (eds.) *Żózne formy, różne treści*, pp. 117–126, Warszawa 2011.
- CZEREPOWICKA, M. (2011): „Toposław” jako narzędzie znakowania jednostek wieloczlonowych, in Matusiak-Kempa, I., Przybyszewski, S. (eds.) *Nowe zjawiska w języku, tekście, komunikacji. Kontekst a komunikacja*, Olsztyn, pp. 28–35.
- CZEREPOWICKA, M. (2014): Jednostki obce w słowniku języka polskiego na przykładzie „Słownika elektronicznego jednostek frazeologicznych” (SEJF), in *LingVaria (IX)*, vol. 1 (17), pp. 59-68 [doi: 10.12797/LV.09.2014.17.04].
- CZEREPOWICKA, M. (2014): SEJF - Słownik elektroniczny jednostek frazeologicznych, in *Język Polski (XCIV)*, v. 2, pp. 116-129.

Bibliografia c.d.

- Małgorzata Marciniak, Agata Savary, Piotr Sikora, and Marcin Woliński. Toposław – a lexicographic framework for multi-word units. In Zygmunt Vetulani, editor, Human Language Technology. Challenges for Computer Science and Linguistics: 4th Language and Technology Conference, LTC 2009, Poznań, Poland, November 6–8, 2009, Revised Selected Papers, volume 6562 of Lecture Notes in Artificial Intelligence, pages 139–150. Springer-Verlag, Berlin, 2011.
- Piotr Sikora. Narzędzia do tworzenia elektronicznych słowników fleksyjnych jednostek wielocłonowych, Uniwersytet Warszawski, Warszawa 2011.
- Piotr Sikora and Marcin Woliński. Toposław — a dictionary creation tool. In Mieczysław A. Kłopotek, Adam Przepiórkowski, Sławomir T. Wierzchoń, and Krzysztof Trojanowski, editors, Recent Advances in Intelligent Information Systems, pages 743–749. Akademicka Oficyna Wydawnicza EXIT, Warsaw, 2009.
- Marcin Woliński, Agata Savary, Piotr Sikora, and Małgorzata Marciniak. Usability improvements in the lexicographic framework Toposław. In Zygmunt Vetulani, editor, Proceedings of the 4th Language and Technology Conference, pages 321–325, Poznań, Poland, 2009.
- Małgorzata Marciniak, Joanna Rabięga-Wisniewska, Agata Savary, Marcin Wolinski, Celina Heliasz. Constructing an Electronic Dictionary of Polish Urban Proper Names, In Recent Advances in Intelligent Information Systems, pages 233–246, Exit, Warszawa 2009.
- Cvetana Krstev, Ranka Stankovic, Ivan Obradovic, Dusko Vitas, Milos Utvic. Automatic Construction of a Morphological Dictionary of Multi-Word Units. Advances in Natural Language Processing, 7th International Conference on NLP, IceTAL 2010, Reykjavik, Iceland, August 16-18, 2010; 01/2010.
- Monika Czerepowicka, Agata Savary. Kodowanie czasowników w Toposławie. Raport techniczny, instrukcja leksykografa.