

Terminological research in Ukraine

Natalia Grabar	Nataliia Shyshkina, Halyna Zorko	Thierry Hamon
UMR8163 STL	V.M. Glushkov Institute of Cybernetics	LIMSI-CNRS, Orsay
CNRS, U Lille 3	National Academy of Sciences of Ukraine	U Paris 13
Villeneuve d'Ascq	Kyiv, Ukraine	Sorbonne Paris Cité
France	{nata_shy, zorko_gv}@voliacable.com	France
natalia.grabar@univ-lille3.fr		hamon@limsi.fr

Abstract

Our purpose is to present research done in Ukraine on terminology-related questions. We address three aspects: theoretical questions studied, automatic terminology-related tools and existing terminological resources.

1 Terminological research in Ukraine

Terminological research in Ukraine has long tradition:

- the journal *Лексикографічний бюлетень* (*Lexicographical bulletin*) has been founded back in 1951;
- the activity of *Інститут української наукової мови ВУАН* (*Institute of Ukrainian Scientific Language of Academy of Sciences*) and *Технічний комітет стандартизації науково-технічної термінології* (*Technical Committee for Scientific and Technical Terminology Standardization*) consists in normalization and standardization of Ukrainian language and terminology;
- a dedicated conference *Українська термінологія і сучасність* (*Ukrainian Terminology and Contemporaneity*) has been held for the 10th year in 2015.

On the whole, the research in terminology is similar to the research work done in other countries, although the geographical and political situation in Ukraine is specific: poor development during and after the soviet period, overwhelming neighborhood of the Russian society, first years of the opening of Ukrainian researchers to

the internationally recognized journals and conferences. Accordingly, the main body of the scientific work is currently published in Ukrainian. Another fact is the domination of theoretical work, although computational methods are emerging. We present various realizations in the terminology area across three lines: theoretical questions, automatic terminology-related tools and existing terminological resources and their utilization.

1.1 Theoretical questions

Terminology research in Ukraine is an active area, although the main terminological work shows mainly theoretical and linguistic orientation, such as:

- history of national lexicology, computational lexicography and terminology (Самойлова, 2004; Ivashchenko, 2013);
- history of national lexicography and terminography dictionaries starting from XVI century and up to the modern times (Балог, 2004);
- computer problems related to the standardization of terminology (Рожанківський and Кузан, 2000);
- semantic processes in the modern Ukrainian terminology (Тищенко, 2004).

1.2 Tools

There are very few automatic tools for the terminology work. We can cite Part-of-Speech taggers: UGtag tagger (Kotsyba et al., 2009) which does not perform syntactic disambiguation and the TNT model for Ukrainian (Babych, 1997).

1.3 Terminological resources

Several specialized areas have been addressed in order to propose their terminological description:

- physics (Кочерга and Мейнарович, 2010);
- law (Тименко, 2004);
- computer sciences (Коссак, 2000; Dmytruk, 2009);
- religion (Пуряєва, 2004);
- literature (Shatalina, 2005);
- Crimean Tatar language (Alieva, 2005; Memetova, 2007).

These are mainly descriptive studies which aim is to confine the area and to identify the main notions and terms.

In accordance with the recent international orientations, new research directions of the area aim at:

- building the electronic corpora (Kelikh et al., 2009; Демська, 2011),
- using electronic corpora for the building of terminologies and dictionaries (Монахова, 2009; Бугаков, 2006; Глибовец and Решетнев, 2014),
- transforming traditional dictionaries in electronic format (Левченко and Кульчицький, 2013).

Up to now, very few works are oriented on the use of terminologies. The currently addressed tasks are related to the software localization (Shyshkina et al., 2010), and to the indexing of speech therapy terminology (Лалаєва et al., 2004).

The majority of these resources are not freely available, although some of them can be queried online: the national corpus of the Ukrainian language¹ (Дарчук, 2010) and online dictionaries².

2 Conclusion

The terminological research is an active area in Ukraine. It is dominated by theoretical work and is mainly published in Ukrainian. Nevertheless, as

the topics researched are similar to those studied in other countries, we believe that the Ukrainian research community will be soon well positioned and recognized at the international level.

Acknowledgments

This work is funded by the LIMSI-CNRS AI project *Outiller l'Ukrainien*.

References

- VN Alieva. 2005. Onomatopoeic words in the Crimean Tatar language. *Uchenye zapiski*, 18(57):8--11.
- B Babych. 1997. Representation and interpretation of ambiguous deep syntactic structures. *Ukrainian Linguistics*, 21:89--100. in Ukrainian.
- V Dmytruk. 2009. Typological features of word-formation in computing, the internet and programming in the first decade of the XXI century. In *УДК*, pages 1--11.
- VL Ivashchenko. 2013. Historiography of terminology: metalanguage and structural units. In *UDC*, pages 1--22.
- E Kelikh, S Buk, P Grzybek, and A Rovenchak. 2009. Project description: designing and constructing a typologically balanced ukrainian text database. In *Методи аналізу тексту*, pages 125--132.
- N Kotsyba, A Mykulyak, and IV Shevchenko. 2009. Utag: morphological analyzer and tagger for the ukrainian language. In *Proceedings of the international conference Practical Applications in Language and Computers (PALC 2009)*.
- ES Memetova. 2007. Lexicophraseological expressive means of the Crimean Tatar language. *Uchenye zapiski*, 18(57):37--39.
- OF Shatalina. 2005. Literature terminology of the Old Ukrainian literature of the 18th century. *Uchenye zapiski*, 18(57):5--7.
- N Shyshkina, G Zorko, and L Lesko. 2010. Terminology work and software localization in Ukraine. In *Problems of Cybernetics and Informatics*, pages 17--20.
- В Балог. 2004. Параметрична система тлумачних словників української мови та її використання в комп'ютерній лексикографії. *Лексикографічний бюлетень*, 10:12--18.
- ОВ Бугаков. 2006. Создание семантического словаря предложных конструкций на основе украинского национального лингвистического корпуса. Technical report, Украинский языково-информационный фонд НАН Украины, Киев, Украина.

¹<http://www.mova.info/corpus.aspx?l1=209>

²<http://lcorp.ulif.org.ua/dictua/>

- АН Глибовец and ИВ Решетнев. 2014. Метод итеративного построения терминологии в коллекциях научных текстов на украинском языке. *Кибернетика и системный анализ*, 50(6):53--62.
- Н. Дарчук. 2010. Дослідницький корпус української мови: основні засади і перспективи. *ВІСНИК Київського національного університету імені Тараса Шевченка*, 21:45--49.
- О Демська. 2011. *Текстовий корпус: ідея іншої форми*. ВПЦ НаУКМА, Київ, Україна.
- О Коссак. 2000. Українська комп'ютерна термінологія. In *Сучасні проблеми в комп'ютерних науках*, pages 39--42.
- О Кочерга and Є Мейнарович. 2010. *Англійсько-Українсько-Англійський словник наукової мови. Фізика та споріднені науки*. Нова книга, Вінниця, Україна.
- Р Лалаєва, Ю Суrowанець, and В Тищенко. 2004. Індексція польсько-, російсько- та українськомовної логопедичної термінології. *Лексикографічний бюлетень*, 10:29--36.
- ОП Левченко and ІМ Кульчицький. 2013. Технологія перетворення п'ятимовного словника порівнянь в електронну форму. In *Інформаційні системи та мережі*, pages 129--138.
- ТВ Монахова. 2009. Застосування прийомів корпусної лінгвістики в лексикографії. *Наукові праці*, 98(85):55--60.
- Н Пуряєва. 2004. Дещо про мову богослужіння взагалі та про словник мови богослужіння зокрема. *Лексикографічний бюлетень*, 10:36--42.
- Р Рожанківський and М Кузан. 2000. Комп'ютерні проблеми стандартизації термінології. In *Сучасні проблеми в комп'ютерних науках (CCU'2000)*, pages 42--44.
- І Самойлова. 2004. З історії відділу лексикології та комп'ютерної лексикографії. *Лексикографічний бюлетень*, 10:7--12.
- Л Тищенко. 2004. Лексико-тематичні групи української юридичної термінології початку хх століття. *Лексикографічний бюлетень*, 10:65--70.
- О Тищенко. 2004. Семантичні процеси в українській термінології початку ххі століття (психолінгвістичний аспект). *Лексикографічний бюлетень*, 10:50--55.

