The Writer Text Corpus in Native Language Teaching
(Homonymy in the Corpus of the Long Prose Fiction by Ivan Franko)

Solomija BUK
Department for General Linguistics
Ivan Franko National University of Lviv
Lviv, Ukraine
solomija@gmail.com

This article presents a possible way of using a text corpus (in particular Ivan Franko text corpus). A corpus is a powerful instrument in linguistic studies and can be used in both theoretical and practical research. Learner, national and writer's text corpora can also be used in foreign and native language teaching. The application of a text corpus in university lectures is demonstrated with the example of homonymy.

The presented analysis is part of a larger project involving the creation of a complex corpus and liguostatistical study of the long prose fiction by Ivan Franko, a famous Ukrainian writer (1856-1916). The corpus has a structural, morphological, anaphora and semantic mark-up.

In the 21st century the important aim of the university is the preparation of highly qualified linguists who can freely operate on a text corpus. The familiarity with the main concepts of corpus linguistics is an integral part of linguistic competence of present-day students. That is why integrating corpus exercises into the study programs promotes the rethinking of academic status of linguistic disciplines as well as their social function.

There is good practice in the world linguistics to include text corpora in the study process (Braun *et al.*, 2006; Campoy *et al.*, 2009; O'Keeffe, 2007; Sinclair, 2004; McCarthy, 2008). Students are often invited to take part in the creation and tagging of text corpora (Český národní korpus; Korpus języka polskiego IPI PAN; Slovenský národný korpus, etc.).

University teachers should explain, for example, the homonymy not only from the theoretical point of view (the way that dominates in the education systems of many countries) but also practically. Due to its rich morphology, Ukrainian, as all Slavic languages, has a lot of word forms that can have different grammatical interpretations. Being part of automatic morphosyntactic analysis of the languages, homonymy disambiguation is an important stage of text corpus construction. That is why it is useful to draw students' attention to the applied value of homonym resolution in a text. A text corpus (the Ivan Franko text corpus in particular) helps to demonstrate the behavior of language items in the natural text, specially their prevalence, frequency dependent on the text type (direct, author's speech) and other. In this way the Ivan Franko text corpus is used in the training, research and independent homework of students.

Practically all the types of homonymy were found in the Franko corpus. In the exercises based on this corpus students are asked to distinguish the homonyms (e.g., "*а*" conjunction/particle/interjection; "*мати*" noun/verb; "*та*" pronoun/conjunction/particle; "*що*" pronoun/conjunction/particle; "*так*" adverb/conjunction/particle etc.), to calculate their

frequencies, and then compare the results with the Franko corpus with disambiguated homonymy.

These exercises help students understand deeper the homonymy as a phenomenon, the value of its resolution for automatic processing of text, for morphological tagging, etc. They develop the computational and corpus linguistics skills as well as language intuition of students. The application of a text corpus will also provide the specification for lexicographical description of homonyms.

References

Braun, S., Kohn, K. & Mukherjee, J. (eds). (2006). *Corpus Technology and Language Pedagogy*: *New Resources, New Tools, New Methods.* Frankfurt am Main: Peter Lang.

Campoy, M.C., Bellés-Fortuno, B. & Gea-Valor, M. (eds). (2009). *Corpus-Based Approaches to English Language Teaching*. London/New York: Continuum.

Český národní korpus. Available from: <http://ucnk.ff.cuni.cz>

Korpus języka polskiego IPI PAN. Available from: <http://korpus.pl>

McCarthy M. (2008). 'Accessing and interpreting corpus information in the teacher education context.' *Language Teaching* 41: 563–574.

O'Keeffe, A., McCarthy, M. & Carter, R. (2007) *From corpus to classroom: language use and language teaching*. Cambridge: Cambridge University Press.

Sinclair, J. McH. (ed.). (2004). *How to Use Corpora in Language Teaching*. Amsterdam: John Benjamins.

Slovenský národný korpus. Available from: <http://korpus.juls.savba.sk>