PROJECT MUSE®

Frekvenční slovník mluvené češtiny (review)

Charles E. Townsend

Journal of Slavic Linguistics, Volume 16, Number 1, Winter - Spring
2008, pp. 177-179 (Review)

Published by Slavica Publishers

DOI: <https://doi.org/10.1353/jsl.0.0014>

➔ *For additional information about this article*

<https://muse.jhu.edu/article/257809>

František Čermák a kolektiv. *Frekvenční slovník mluvené češtiny*. Univerzita Karlova v Praze – Nakladatelství Karolinum. 2007. pp. 511.

Reviewed by Charles E. Townsend

The *Frequency Dictionary of Spoken Czech (FDSC)* reviewed here is based on the Prague Spoken Corpus (Pražský mluvený korpus), assembled by František Čermák and his collective, based on some 300 recordings of native speakers in the Prague area and cataloguing about 15,000 utterances amounting to some 780,000 words. About two thirds of the recordings were “formal” (elicited by questionnaire) and one third “informal” (conversations). The recordings were made from 1988 to 1995, hence during a politically transitional period. The 278 participants were divided binarily according to four criteria: sex, age, education, and type of conversation. This breakdown is not reflected in the dictionary itself but is available from an accompanying CD or from the project’s website (<http://ucnk.ff.cuni.cz>).

The entries are actual Spoken Czech (SC) words, e.g., *starej* rather than Literary Czech (LC) *starý* ‘old’, *von* instead of *on* ‘he’. Most important lexical items are also listed, e.g., *barák* ‘(apartment) house’, useful for the foreigner because the word evokes military housing. Many words, of course, belong to both registers and hence their listing coincides with LC. Entries are alphabetized and easy to find. The forms of individual head words are listed under them, and all forms are marked for frequency. Since SC and LC forms are spelled with the same letters and SC has no special sounds, words and forms are transparent. Each headword is followed by a number designating part of speech, followed by sets of numbers indicating categories appropriate for nouns, adjectives, and so forth. The explanatory front matter is clearly written but requires fairly intense study; this is helped by examples of entries (27–28) and a “decoder” (29–35). Still, the uninitiated user of *FDSC* need not master all the instructions to find the copious basic information in it. The dictionary proper comprises 436 clearly formatted pages. The end matter includes useful “frequency surveys”: 100 most frequent words, followed by a breakdown into the 100 com-

monest auto- and synsemantic words; then nouns, adjectives, verbs, and adverbs.

An immediate objection to any corpus-dependent list is that it is dependent on what happens to occur. However, this SC corpus is massive and well-processed enough to yield what the authors term a “reliable and representative basis for further research and description of the language spoken by the majority of Czechs and is for them their own native tongue, in contrast to Literary Czech, which they encounter only when they enter school at six and which is a second language to them.” And this leads us to something more important than the mere contents of the dictionary: *FDSC* is a published dictionary of SPOKEN Czech, not normative, of course, but with a potential for that status. As such it marches directly into the fabled controversy about Czech diglossia: whether and, if so, which SC elements should be admitted to LC.

Frequency dictionaries have been of limited interest to those outside their precincts, and are probably mostly useful for language teachers looking to include frequently used words in their pedagogical materials. *FDSC* gives more information than most, but its real significance is elsewhere. It rests firmly on *obecná čeština* (‘generally spoken Czech’), in marked contrast to the previous frequency dictionary (Jelínek, Bečka, and Těšitelová 1961), which accepted colloquial elements “only from the speeches of (communist icons) like Zápotocký, Gottwald, and Nejedlý” (7). But Čermák and his colleagues have gone much further. Ignoring the forbidding scrutiny of the Czech Language Institute and the Moravians (who are often considered to speak “the most correct Czech”) they unashamedly base *FDSC* on the speech of Prague. As they openly state: “The language of the capital, due to the huge migration and confluence of people from all corners of the land, is not and cannot be just one of many local languages” (7).

This reviewer hopes fervently that *FDSC* can be a forerunner to proper description of SC, ultimately with grammars and “real” dictionaries. Usage, after all, should determine norms, and Čermák has given us a start on documented usage. A proscriptive language like LC freely ignores usage, so why should the “norms” of Spoken Czech be less legitimate? There will be huge arguments over choices, obviously. Townsend 1990 was a hesitant effort to help foreign learners but, unlike *FDSC*, was not based on massive recorded material. Now real Czechs need to do the job. Jiří Hronek started things off in 1972 with

his ground-breaking *Obecná čeština*. Not much has happened in the 35 years since then. Maybe things can get moving again.

References

- Hronek, Jiří. (1972) *Obecná čeština*. Prague: Univerzita Karlova.
- Jelínek, Jaroslav, Josef Václav Bečka, and Marie Těšitelová. (1961) *Frekvence slov, slovních druhů a tvarů*. Prague: Ústav pro jazyk český.
- Townsend, Charles E. (1990) *A description of Spoken Prague Czech*. Columbus, OH: Slavica.

Princeton University
247 East Pyne
Princeton, NJ 08544
townsend@princeton.edu

Received: December 2007