

About TAMBiC

The Tampere Bilingual Corpus of Finnish and English

[Size and content](#)

[Format](#)

[Availability](#)

[Source texts](#)

This is a project which has been in progress for a number of years. The person in charge is [Robert Cooper](#) from the English Department from [School of Modern Languages and Translation Studies](#) at [University of Tampere](#) . The corpus currently consists of 1.8 million words, the eventual target being 2 million words.

[Back To Top](#)

Size and content

2 million words: 1 million words of fiction and 1 million words of non-fiction.

Fiction sub-corpus: long extracts (15,000 words/50 pages) from 16 English novels (plus their Finnish translations) and similar extracts 16 Finnish novels (plus their English translations).

Non-fiction sub-corpus: long extracts (and their translations) from the following: books on history, art, music, science, nature and linguistics; tourist guides, cookery books and instruction manuals; and a large number of newspaper articles.

[Back To Top](#)

Format

The corpus has the following fixed format:

The Sons of Freedom are a breakaway sect of the Doukhobor faith, an obscure group of Russian Christians whose beliefs include pacifism and total equality.

(PRESS 1:1:10)

Vapauden veljet on Doukhobor uskosta irtautunut lahkko, hämäräperäinen ryhmä venäläisiä kristittyjä joiden uskomuksiin kuuluu rauhanaate ja täydellinen tasa arvo.

The group immigrated to Western Canada in the late 1800s to escape brutal persecution.

(PRESS 1:1:11)

Ryhmä muutti Länsi Kanadaan 1800 luvun lopulla paetakseen julmaa vainoa.

Searches give complete sentences, each accompanied by a full translation into Finnish (or English) – as above.

[Back To Top](#)

Availability

The corpus is available to all teachers, researchers and students within the School of Modern Languages and Translation Studies at the University of Tampere. However, access can also be granted on request to anyone else interested in obtaining data from the corpus.

[Back To Top](#)

Source texts

[See the source texts](#)

[Back To Top](#)

To know more information about how this project was done please visit [Tambic Wiki](#)

For information on the background to the corpus see Cooper W. R. (1998) “The Tampere Bilingual Corpus of Finnish and English: Development and Applications.” In Cooper, W. R. (ed.) *Compare or Contrast? Current Issues in Cross-Language Research*. Tampere: University of Tampere.

