

The Bilingual Concordancer TransSearch

Guy Lapalme Philippe Langlais Fabrizio Gotti

RALI

Université de Montréal

H3C 3J7, Montréal, Québec, Canada

<http://rali.iro.umontreal.ca>

Overview

- **TRANSSEARCH** is a web-based translation search engine, developed by **RALI** and commercialized by **Terminotix** (<http://www.terminotix.com>).
- The system replies to a user query with a set of sentence pairs whose source sentence contains the query.
- The query is highlighted and the corresponding target expression is also identified.

Screenshot

The screenshot shows the TransSearch interface with the user 'lapalme'. The search expression is 'in keeping with'. The results show 208 translations of 'in keeping with' in 901 occurrences. The interface includes a navigation bar with 'REQUÊTES', 'MON COMPTE', 'PRÉFÉRENCES', 'AIDE', and 'QUITTER'. The search results are displayed in a table with columns for the source and target expressions, and a count of occurrences.

A few facts

- Available since 1996 through a Web interface by Université de Montréal [2]
- Currently used by more than 3000 professional translators
- Treats over xxxxx queries a week
- A Spanish-English version is also available
- French-English Corpora indexed:

Corpus	docs	M pairs	M words
Hansards	2 854	10.5	339.6
Senate	1 025	1.2	47.6
Canadian Courts	13 068	3.2	142.6
Total	16 947	14.9	528.2

Features

- ▶ Querying discontinuous expressions:

The screenshot shows the TransSearch interface with the user 'felipe'. The search expression is 'make .. hair stand on end'. The results show 2 translations of 'make .. hair stand on end' in 4 occurrences. The interface includes a navigation bar with 'REQUÊTES', 'MON COMPTE', 'PRÉFÉRENCES', 'AIDE', and 'QUITTER'. The search results are displayed in a table with columns for the source and target expressions, and a count of occurrences.

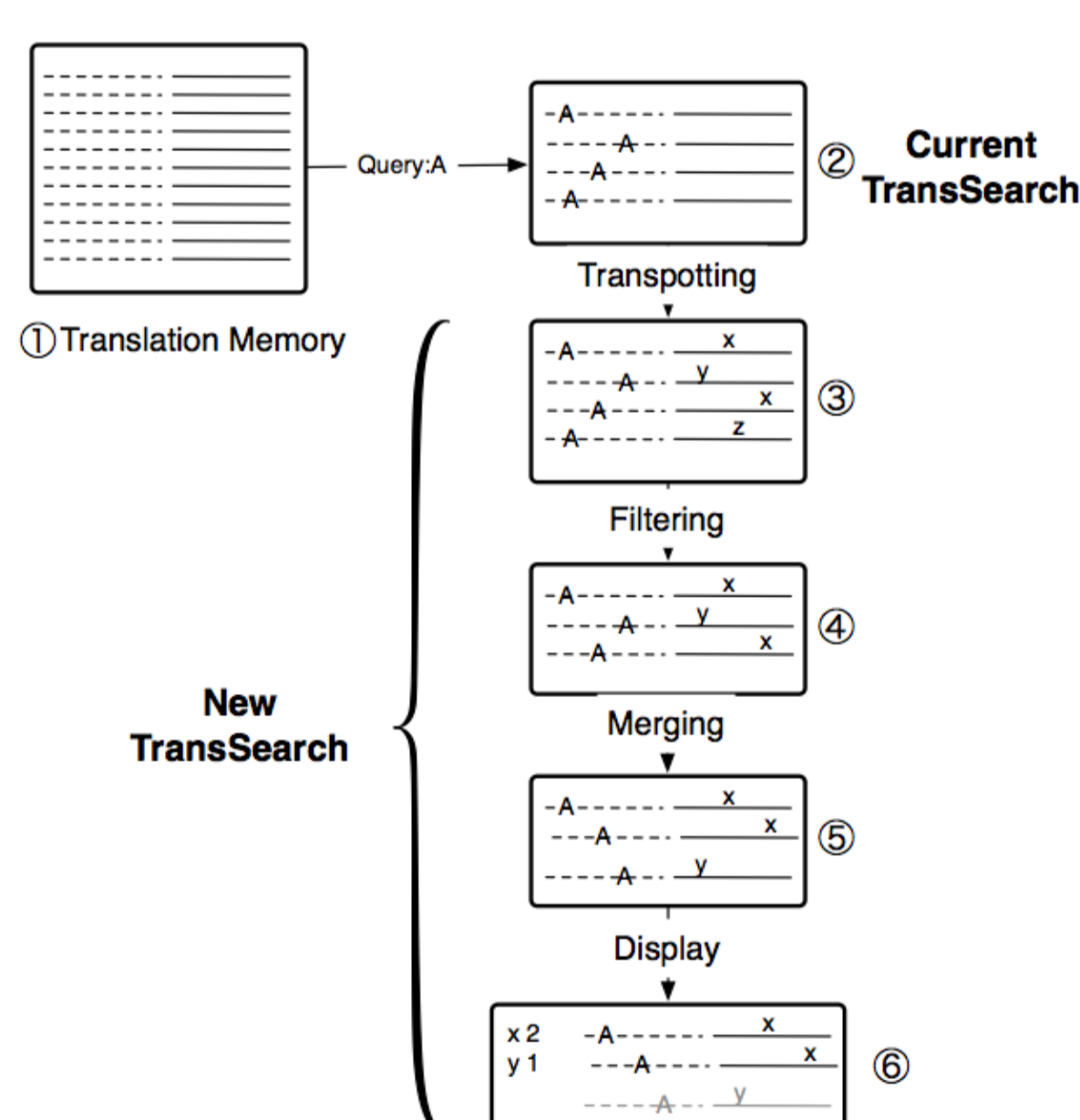
- ▶ Morphological expansion:

The screenshot shows the TransSearch interface with the user 'felipe'. The search expression is 'take+ no for an answer'. The results show 13 translations of 'take+ no for an answer' in 16 occurrences. The interface includes a navigation bar with 'REQUÊTES', 'MON COMPTE', 'PRÉFÉRENCES', 'AIDE', and 'QUITTER'. The search results are displayed in a table with columns for the source and target expressions, and a count of occurrences.

- ▶ Bilingual Query:

The screenshot shows the TransSearch interface with the user 'felipe'. The search expression is 'take+ no for an answer'. The results show 13 translations of 'take+ no for an answer' in 16 occurrences. The interface includes a navigation bar with 'REQUÊTES', 'MON COMPTE', 'PRÉFÉRENCES', 'AIDE', and 'QUITTER'. The search results are displayed in a table with columns for the source and target expressions, and a count of occurrences.

Technology



- Developed with a mix of PHP
- Database indexed with Lucene
- Word Alignment computed online (See comparison of word alignment techniques in [1])

References

- [1] J. Bourdaillet, S. Huet, P. Langlais, and G. Lapalme. "TransSearch: from a bilingual concordancer to a translation finder". In: *Mach. Transl.* 24(3-4) (2010), pp. 241–271.
- [2] E. Macklovitch, M. Simard, and P. Langlais. "TransSearch: A Free Translation Memory on the World Wide Web". In: *2nd LREC*. Athens, Greece, 2000, 1201–1208.

Acknowledgments

- **Julien Bourdaillet** and **Stéphane Huet** for their research contributions as postdoctoral fellows at **RALI**.
- **Gilles Gamas**, **Micheline Cloutier** and **Jean-François Richard** from **Terminotix** who supported this research via their contribution to a Collaborative Research and Development (CRD) grant from the Natural Science and Engineering Research Council (NSERC) of Canada.
- **Jacques Steinlin** and **Elliott Macklovitch**.