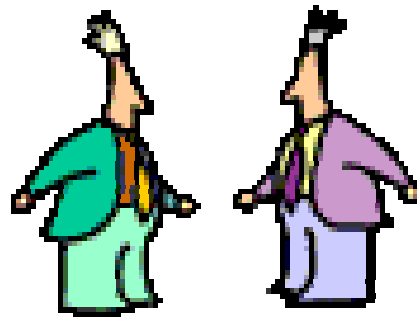# Machine Translation and Translation Memory:
A Message of Reconciliation

Overview

- *Machine Translation (MT) and Translation Memory (TM) have wrongly been viewed as separate entities*

- *MT and TM can benefit from (re-)merging their technologies*

- *MT can benefit from the large collections of data that have been assembled over the years of TM use*

"Disdain on the side of the professional translators for the hilarious and stupid MT mistakes gave birth to a new variant of MT called translation memory (TM). TM started off as a lower-level feature of commercial MT systems (...). But the success of TM came with dedicated products such as IBM TM/2 and Trados. The marketing message was tuned in to what the professional translation industry wanted to hear: 'Forget about MT; it doesn't work well. Instead, use our TM product because it leaves you in full control of the process.'

"The message worked well. Within a period of 10 to 15 years, TM products have found their way to the workstations of more than 50,000 translators in the world. But the message had also caused a 'cognitive disorder' in the translation industry, namely that TM is good and MT is evil, foregoing the fact that TM is just a new variant of MT (...). The damage is done, however, and it will take years to convince the community of business translators that post-editing fuzzy matches from TM databases is, in fact, not different from post-editing fuzzy matches from any other MT system."

Jaap van der Meer in MultiLingual 71 (2005)
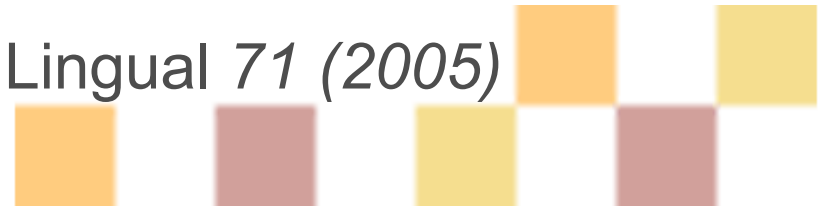
What is the translation memory aspect of TM tools?

- *Simple but amazingly effective system for perfect and fuzzy matching through previously translated data*

  - *Some tools provide more complex parameters ("EBMT")*
  - *Project management*
  - *Terminology management*
  - *File protection and conversion*
  - *...*

If it is so effective, and the parameters are so simple . . .

- *… why not create an environment that allows for this processing of data:*
  - *Look for perfect match*
  - *Look for high-percentage fuzzy match*
  - *If found, attempt to "repair" fuzzy match*
  - *Generate MT match*
- *All of these processes are based on the same database(s) and being fed into the same database*
- *If terminology management becomes part of this equation, results could be phenomenal*
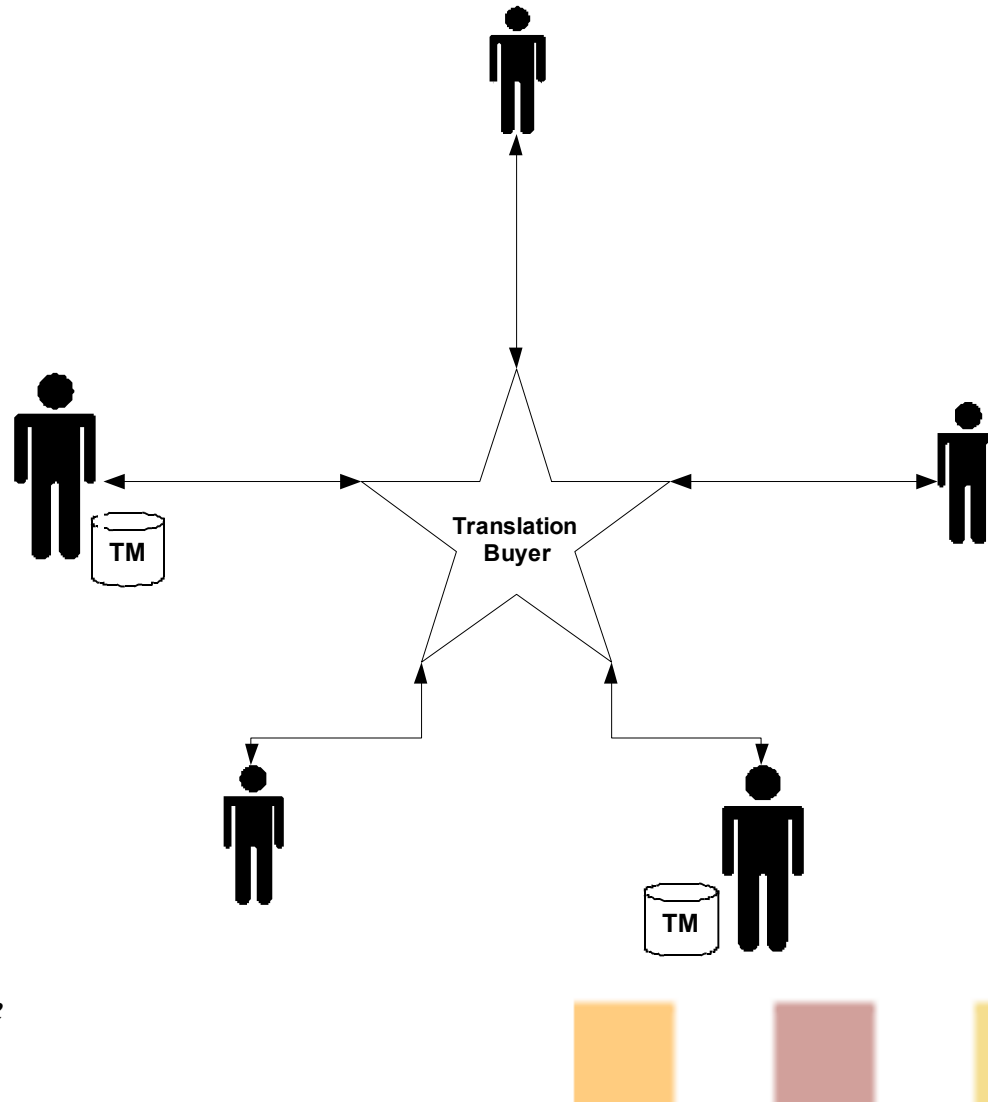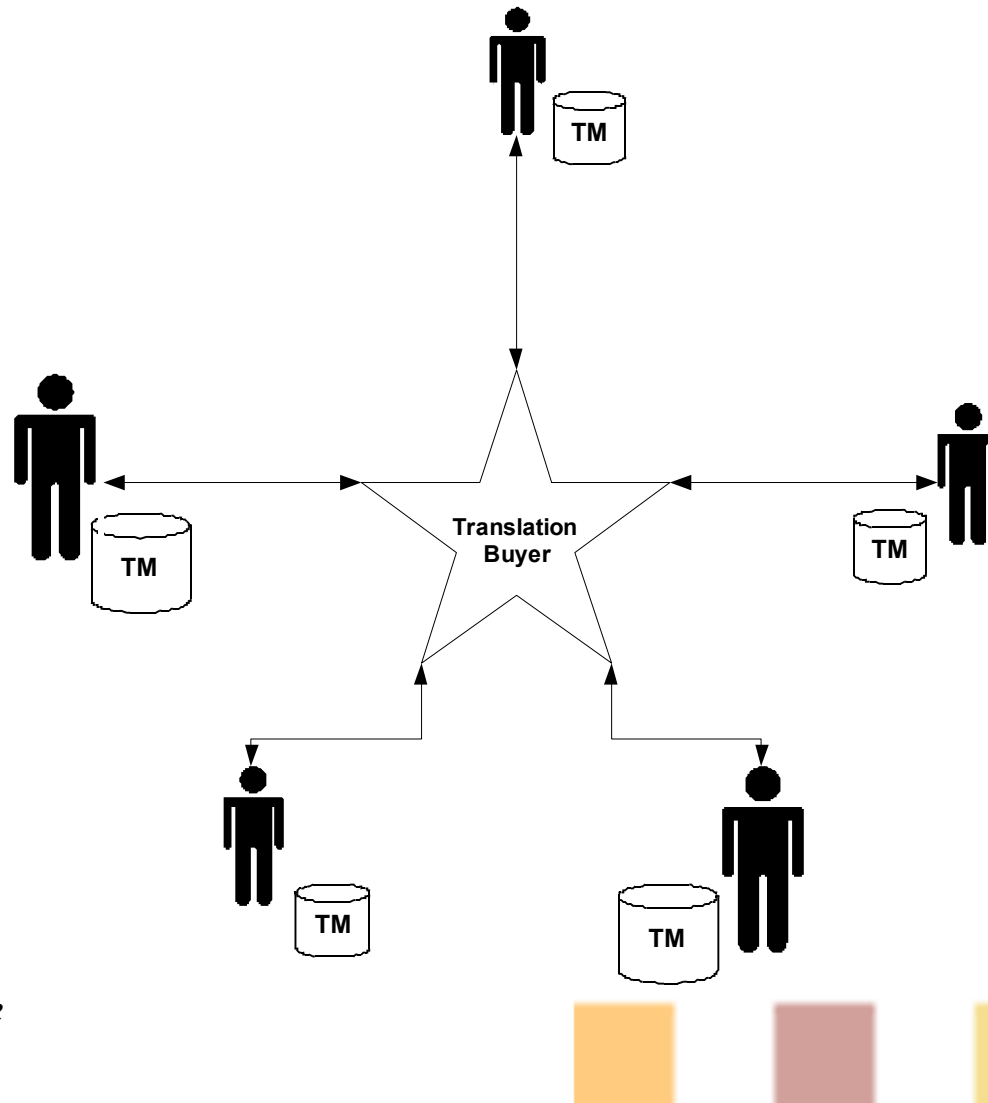
Roadmap (?)

- *Let's not look at a parallel use of separate technologies, but let's integrate them into one tool.*

- *Within the past three or so years, five completely new TM systems have entered the market, some of which are quite open to new ideas . . .*
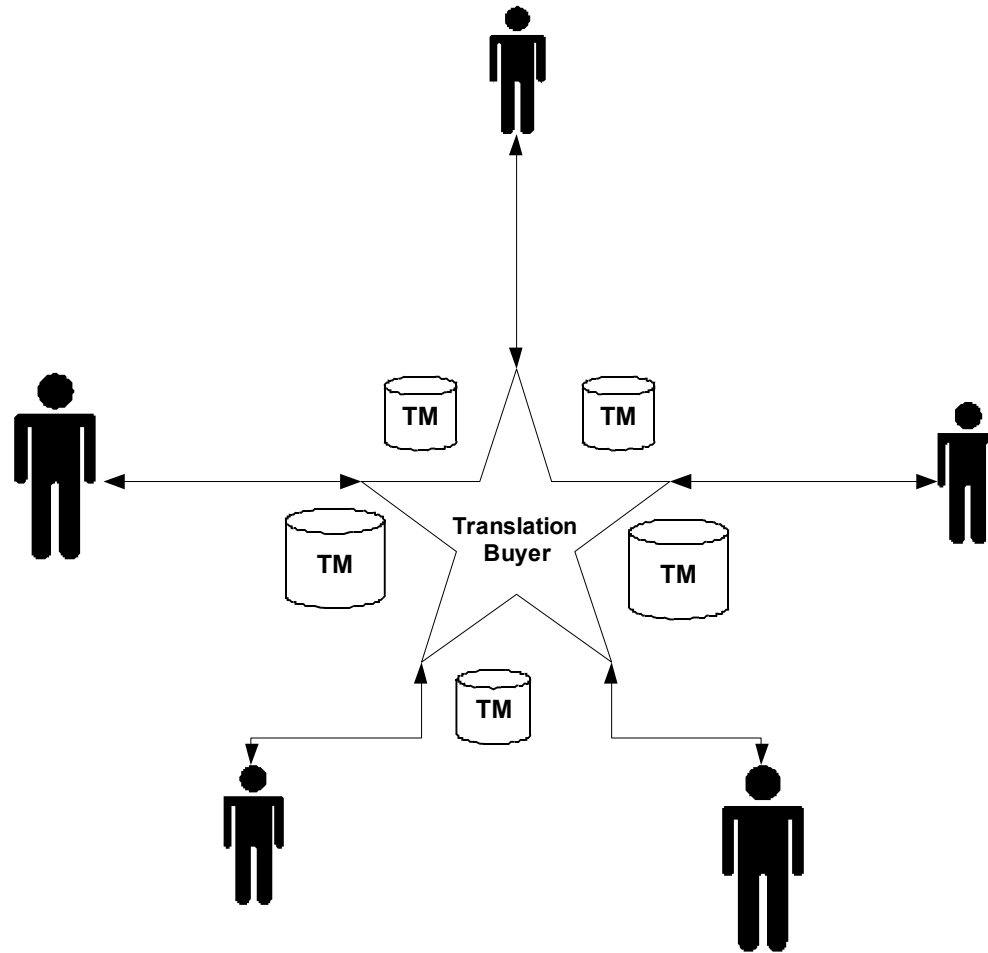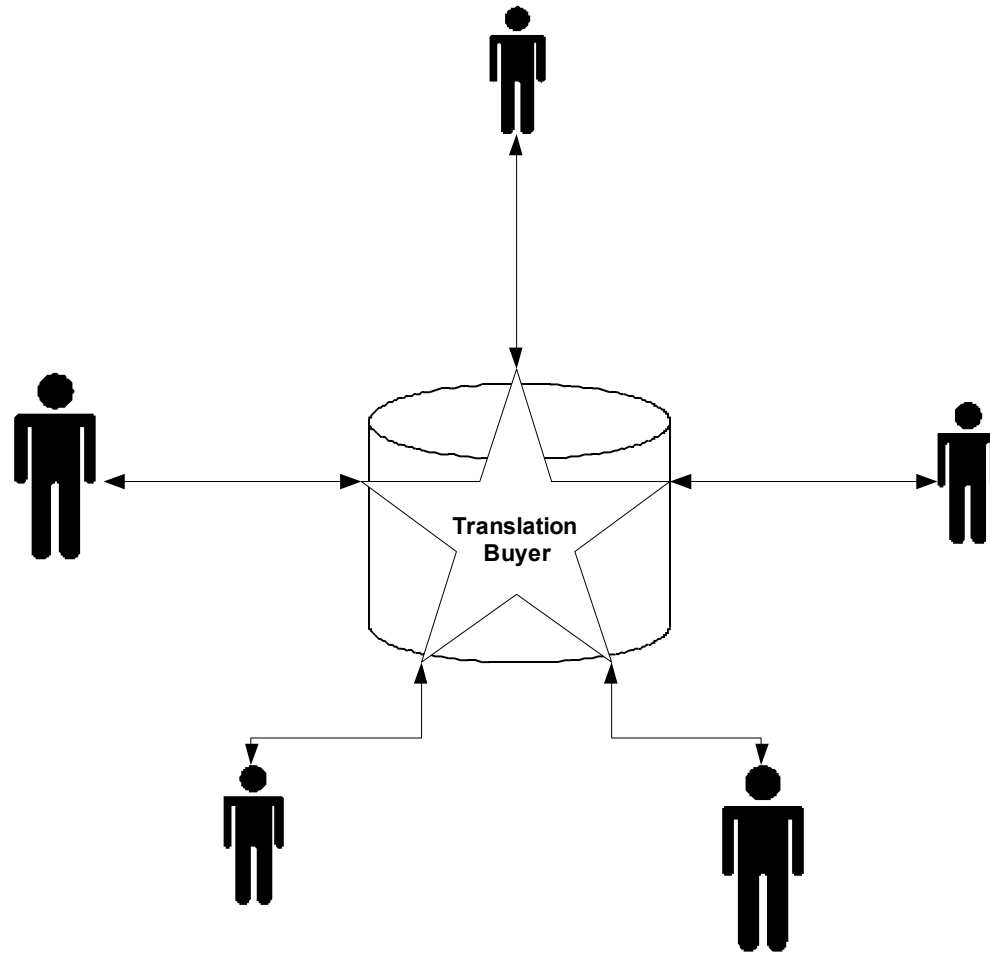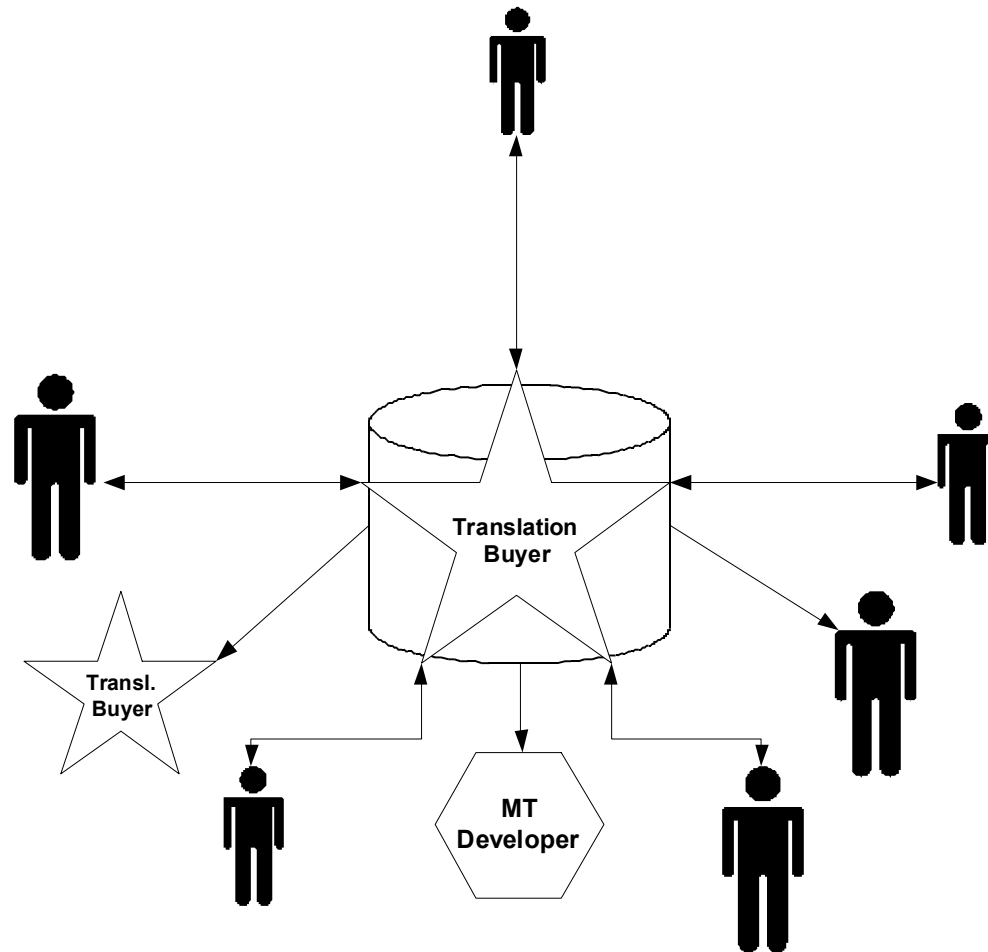
History of TM

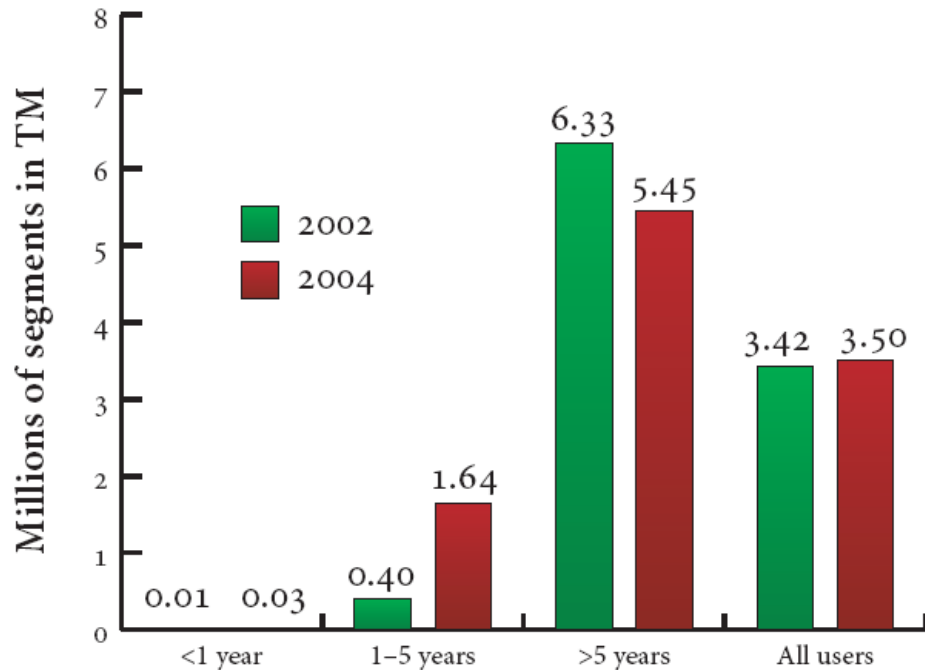# History of TM

History of TM

History of TM

Translation
Buyer

History of TM

Adding up the numbers

- *According to the LISA Translation Memory Survey 2004, TM users of five years or more have more than 5.45 million translation units (TUs).*

- *Looking only at Fortune 500 companies, this would come out to be 2.725 billion verified and industry-specific TUs.*

With statistically-based machine translation being on the forefront of development in MT, this is obviously of great significance.

Ways to get to the data

- *Industry-specific online repositories that can be accessed and written to on a subscription basis*

- *Licensing of translation memory data on a per-match basis*

- *Selling/licensing of translation memory data on the basis of number of translation units/words*

THE TRANSLATION MEMORY BROKERS

www.tmmarketplace.com