

Using free online MT in multilingual websites

Harold Somers

Federico Gaspari

*School of Languages, Linguistics and Cultures
University of Manchester*

Timetable

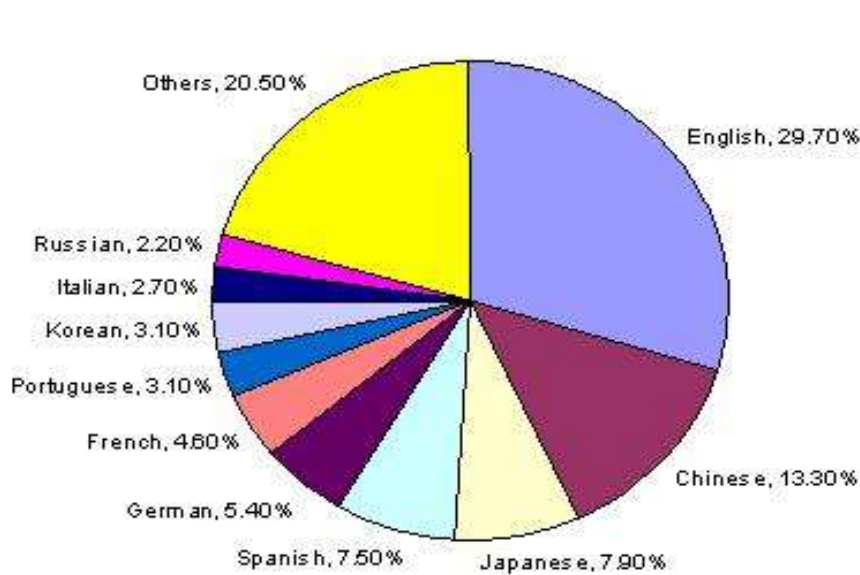
- 9.00 **Introduction**
- 9.30 **Good practice**
Common current practice, good practice issues
- 10.00 **Exercise I**
Looking at some (real) examples, how good are they?
- 10.30 **COFFEE BREAK**
- 10.45 **Key steps**
Practical issues, HTML, MT-friendliness
- 11.15 **Recommendations for good practice**
Check-list of desiderata for MT-friendly web page design
- 11.45 **Exercise II**
Evaluation of redesigned webpages
- 12.15 **Wrap-up**
Opportunities and challenges
- 12.30 **LUNCH**

Introduction

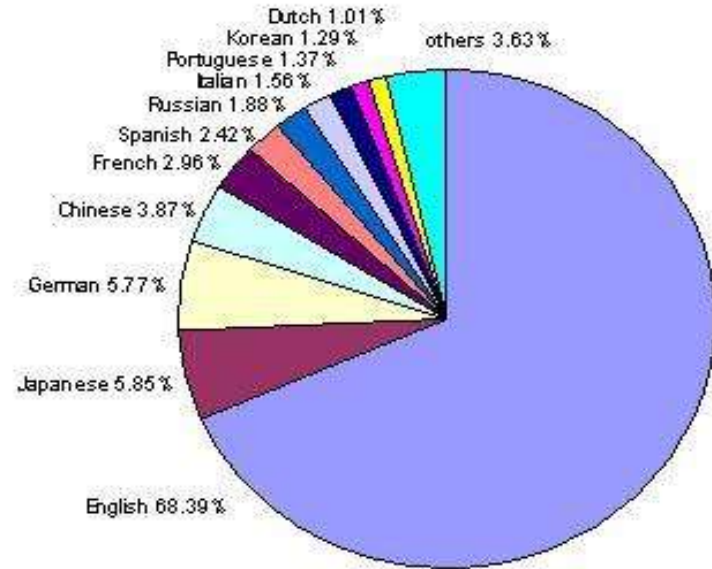
- Timetable for the morning
- Aims and goals of this tutorial
- Brief history of free online MT
- What is available (services, language pairs, underlying engines)
- Assimilation vs dissemination
- Integration

On-line MT is a good solution for short-lived web-pages with volatile, perishable content which make human translation not viable

English and the web



**Languages used online
(based on a total of 1.085 bn
users) – 75% of internet
users don't know any
English**



**Percentage of webpages
in various languages
(from a total of 313.2 m
web pages)**

Aims and goals

- Looking at use of free online MT for on-the-fly “localisation” of web pages
 - i.e. click to translate
- Not concerned with general webpage localisation, whether using MT or not
 - though distinction might be an issue
- Translation quality of online MT systems is not an issue (here)
- Especially interested in web-page designer’s point of view

Brief history of online MT

- Babelfish is 10 years old ...
- ... but online MT is nearly 20 years old
- 1988: Systran offered MT service to Minitel users in France (via file transfer)
- 1992: CompuServe used MT to translate messages posted to certain forums
- 1996: CompuServe offered document translation service

Brief history of online MT

- December 9, 1997: Babel Fish launched
 - AltaVista search engine
 - Systran MT system
 - First free system
- Babel Fish soon had competitors
 - More than 10 companies involved by 2000
 - Currently 30+



What is available?

- Services
 - Text box translation of typed-in (or pasted) text
 - Web-page translation by submitting URL
 - For example ...
- Evidence of misuse
 - Back and forth translation (as a game, or for evaluation)
 - Looking up individual words
 - Nonsense or x-rated material

Sites



Il traduttore inglese-italiano e italiano-inglese direttamente **online**.

ΑΥΤΟΜΑΤΗ ΜΕΤΑΦΡΑΣΗ



Language pairs

- Eng-Spa has always been most popular ...
- ... followed by combinations of Eng, Fre, Ger, Spa
- “English to X” pairs popular locally
- BabelFish currently offers 36 pairs involving 13 languages

Select from and to language

Chinese-simp to English
Chinese-trad to English
English to Chinese-simp
English to Chinese-trad
English to Dutch
English to French
English to German
English to Greek
English to Italian
English to Japanese
English to Korean
English to Portuguese
English to Russian
English to Spanish
Dutch to English
Dutch to French
French to English
French to German
French to Greek
French to Italian
French to Portuguese
French to Dutch
French to Spanish
German to English
German to French
Greek to English
Greek to French
Italian to English
Italian to French
Japanese to English
Korean to English
Portuguese to English
Portuguese to French
Russian to English
Spanish to English
Spanish to French

Underlying engines


- Several sites use Systran (Babelfish, Lycos, Music, Systranet, Voila)
- Many companies have their own sites
- Often not clear which engine is used
 - Multilingual sites may use several
 - Translation quality often similar from site to site, implies same engine (esp. e.g. Chinese)

Assimilation vs dissemination

- Assimilation
 - user is reader
 - many SLs (poss. unidentified), one TL
 - rough translation OK
- Dissemination
 - user is author
 - one SL, many TLs
 - good translation needed
 - “localization”

This is what
online MT is
especially good
for

In order to achieve
this, web-page
designer needs
help



Integration

- From both perspectives (assimilation/dissemination) the translated website should be smoothly integrated
 - Easy to generate on the fly
 - Should behave like the source site in terms of navigation