

# **Outline of the Evolution of Machine Translation**

**John Hutchins**

**Presentation on Panel at Aslib T&C conference,  
November 2008**

# The beginnings: 1950s and 1960s

- Patents: Artsrouni and Troyanskii (1933)
- Weaver's memorandum 1949
- First conference, June 1952
- First demonstration, January 1954
- Characteristics of first decades
  - Military and intelligence funding
  - Russian/ English concentration
  - Direct translation systems
  - Interlinguas
  - Formal syntax
- Bar-Hillel's survey, 1960
  - non-feasibility of FAHQT

# The 1970s

- ALPAC (1966)
  - “no immediate or predictable prospect of useful machine translation”
- European Commission acquires Systran (1976)
- Météo starts (1977)
- EC conference “Overcoming the language barrier”, Luxembourg 1977
- First Aslib “Translating and the Computer” conference 1978

# The early 1980s

- **Widening perspectives**
  - Japanese companies and universities
  - linguistics-based, transfer-based MT (Ariane, Mu, Eurotra)
  - interlingual MT, knowledge-based MT (CMU, Rosetta, DLT), artificial intelligence
- **Office automation**
  - translation aids, terminology management, word processing

# The later 1980s

- **Human-aided MT production**
- **Controlled language input**
- **Large-scale operations**
  - Systran, Logos, METAL, etc.
- **Software localisation**
- **Evaluation: economic, quality, acceptability**
- **Software for PCs (wider use by translators)**
  - database access, terminology software, desktop publishing, telecommunications

# The early 1990s

- **Cheap PC software**
- **Translator's workstation**
  - translation memories (1993)
- **Corpus-based methods**
  - statistical (1988), example-based (1990)
- **Spoken language translation (since 1986)**
- **New ancillary technologies for MT**
  - OCR, voice input, voice output

# The later 1990s and early 2000s

- **MT on the Internet**
  - email, Web pages, free services, online MT (1997)
- **MT as component of:**
  - information retrieval, summarization, information extraction (intelligence)
- **Evaluation metrics**
  - automated comparison with human translations

# Current usages of MT

- **Systems for dissemination (publication)**
  - traditional use by corporations, agencies, localisation
  - rough drafts for authors
- **Systems for assimilation**
  - 'unedited' MT, intelligence/analysis, online MT
- **Systems for interchange**
  - electronic mail, correspondence, Web pages, tourism
- **Language coverage**
  - good for English, French, German, Spanish, Japanese, Chinese, Korean, Arabic,
  - Poor for: African, Indian, S.E.Asian, E.European, UK minorities



# Future system types

- Hybrid systems (rules and statistical data)
- MT integral component of WP/Internet software (with speech I/O); and fully integrated with translation memories
- multilingual access to information, databases
- authoring software and MT (integrated)
- Summarisation and MT
- MT of television captions/subtitles, MT for the deaf
- MT on PDAs, mobiles; photocopier-translator; camera-translator
- Special-purpose online MT systems (e.g. medical)
- Speech translation; translation telephone, etc.

# Resources

- **Historical:**
  - Articles on my website ([www.hutchinsweb.me.uk](http://www.hutchinsweb.me.uk))
- **Associations:**
  - European Association for Machine Translation ([www.eamt.org](http://www.eamt.org))
  - Association for Machine Translation in the Americas ([www.amtaweb.org](http://www.amtaweb.org))
- **Current research and some older materials:**
  - *Machine Translation Archive* ([www.mt-archive.info](http://www.mt-archive.info))
  - All major MT conferences since 1985; eventually all Aslib conferences