Research Methods and System Designs in Machine Translation

a ten-year review
1984-1994

John Hutchins

University of East Anglia
Norwich, England
Outline

Five eras of MT
Situation in 1984
From 1984 to 1989
  rule-based transfer systems
  rule-based interlingua systems
From 1989 to the present
  corpus-based methods
  rule-based methods
  lexicalist tendency
  constraint-based formalisms
Lexicon acquisition
Generation and database MT
Controlled language MT
Domain- and user-specific MT
Speech translation
"Third generation" systems
Wider perspectives
Situation in 1984

Systran
   at USAF (1970)
   at European Commission (1977)
Meteo (1976)
Fujitsu ATLAS

GETA (Grenoble)
SUSY (Saarbrücken)
METAL (Siemens)
Mu (Kyoto)

Eurotra (1982)
From 1984 to 1989

Rule-based transfer
   Ariane (GETA)
   SUSY (Saarbrücken)
   Eurotra (CEC)

Rule-based interlingua
   DLT (Utrecht)
   Rosetta (Philips)
   PIVOT (NEC)
   KBMT89 (Carnegie Mellon Univ.)
From 1989 to the present

Corpus-based methods
  statistics-based: Candide (IBM)
  example-based: ATR, SHALT
connectionist methods

Rule-based methods
  Eurolang
  CAT2, PaTrans
  LMT
  KANT (Carnegie Mellon Univ.)
  CATALYST (CMU and Caterpillar)
  ULTRA (Nex Mexico State Univ.)
  UNITRAN
  Pangloss
Speech Translation

British Telecom

ATR (Osaka)

C-STAR (JANUS)
   Carnegie Mellon Univ.
   ATR
   Karlsruhe Univ.

Verbmobil
Further perspectives

global research
Western Europe
North America
Japan
China
Taiwan
Korea
India
Thailand
Malaysia
Indonesia
Eastern Europe
Russia

MT for translation services/agencies
MT for occasional non-professional translation
MT for information gathering
MT for monlingual users
MT for spoken language
MT on networks
   Minitel, PC-VAN, Niftyserve, CompuServe
   Systran, Logos

MT for translators
   Translator workstations