Latest Developments in Machine Translation Technology

John Hutchins

University of East Anglia, Norwich, England
Introduction

- not a comprehensive survey of all MT research in the past few years
- examples of projects only indicative of trends
- highlight new developments/approaches/methods
- no discussion of projects and systems well established in the late 1980s (e.g. Systran, Logos, EDR, ATR, Mu/JICST, METAL)

Contents

MT in the 1980s
Rule-based MT (transfer, interlingua, formalisms)
Corpus-based MT (statistical, example-based)
Text corpora and alignment
Generation
Controlled, domain-specific, user-specific MT
Workstations
Eras of MT
Towards "third generation" systems
MT during the 1980s

at first MT Summit conference (1987):

- dominant frameworks
  
  indirect translation
  (transfer, interlingua)
  linguistics-based

- knowledge-based MT still innovative

- multinational-multilingual projects

- some well-established systems
  (Systran, Meteo, Weidner/Bravice, PAHO systems, Logos)

- many new systems/projects from
  Japanese companies
  (NEC, Toshiba, Fujitsu, Oki, Ricoh, Sharp, etc.)
Rule-based MT

Transfer-based (typical "second generation")

- three stages: analysis, transfer, synthesis
- abstract semantico-syntactic interfaces
- multiple levels/strata:
  - morphology, syntax, semantics
- syntax-oriented, tree-transduction approach
- batch processing, post-edited
- little pragmatic/discourse information

projects: [Ariane, Eurotra], Eurolang (SITE, METAL), LMT 'Logic programming MT' (IBM centres)

Interlingua-based

- two stages: analysis, synthesis
- abstract language-neutral representation
- multistratal: morphology, syntax, semantics
- semantics-oriented ('understanding')
- [knowledge bases]

projects: [DLT, Rosetta], PIVOT (NEC), Carnegie-Mellon University (KBMT89, KANT, CATALYST) ULTRA (New Mexico State University) MCC (Microelectronics and Computer Technology Corporation) UNITRAN
Formalisms

Non-transformational

Constraint-based formalisms

Unification formalisms

Lexical-Functional Grammar
Logic programming (Q-systems, Prolog)
Definite Clause Grammar
Slot grammar (LMT)
Generalized Phrase Structure Grammar
Head-driven Phrase Structure Grammar
Categorial Grammar

Principles-based MT (Government-Binding Theory)

projects: ITS (Geneva), UNITRAN

Lexicalist approaches
[transfer rules = simple bilingual lexical equivalences]

projects: CRITTER, ACQUILEX,
'Shake-and-bake' (Sharp)

Reversibility
projects: [Rosetta], ISSCO, CRITTER, UNITRAN
Tree transduction model (Eurotra)

Input text in \( L \)

Analysis

Grammar (1) \( \rightarrow \) Representation (1)

\(<---\) Mapping rules \( T (1/2) \)

Grammar (2) \( \rightarrow \) Representation (2)

\(<---\) Mapping rules \( T (2/3) \)

Transfer

Grammar (n) \( \rightarrow \) Representation (n)

\(<---\) Mapping rules \( T (n/n') \)

Grammar (n') \( \rightarrow \) Representation (n')

\(<---\) Mapping rules \( T (n'/n'-1) \)

Synthesis

Grammar (2') \( \rightarrow \) Representation (2')

\(<---\) Mapping rules \( T (2'/1') \)

Grammar (1') \( \rightarrow \) Representation (1')

Output Text in \( L' \)
Constraint-based formalism (LFG)

John likes Mary  \iff  Marie plait à Jean

\textit{like}, V:
\begin{align*}
(\uparrow\text{PRED}) &= \text{like} <\text{SUBJ}, \text{OBJ}> \\
(\tau^{\uparrow}\text{PRED FN}) &= \text{plaire} <\text{SUBJ}, \text{OBJ}> \\
(\tau^{\uparrow}\text{AOBJ OBJ}) &= \tau(\text{SUBJ}) \\
(\tau^{\uparrow}\text{SUBJ}) &= \tau(\uparrow\text{OBJ})
\end{align*}

\textit{john}, N: \quad \textit{mary}, N:
\begin{align*}
(\uparrow\text{PRED}) &= \text{john} & (\uparrow\text{PRED}) &= \text{mary} \\
(\tau^{\uparrow}\text{PRED FN}) &= \text{jean} & (\tau^{\uparrow}\text{PRED FN}) &= \text{marie}
\end{align*}

TL f-structure:
\begin{align*}
&\begin{bmatrix}
\text{PRED} & \text{plaire} \\
\text{SUBJ} & [\text{PRED marie}] \\
\text{AOBJ} & [\text{OBJ} [\text{PRED Jean}]]
\end{bmatrix}
\end{align*}

Student is likely to work  \iff  Il est probable que l'étudiant travaillera

\textit{likely}, A: \quad \textit{probable}, A:
\begin{align*}
(\uparrow\text{PRED}) &= \text{likely} <\text{XCOMP}> \text{SUBJ} & (\uparrow\text{PRED}) &= \text{probable} <\text{COMP}>\text{SUBJ} \\
(\uparrow\text{SUBJ}) &= (\uparrow\text{XCOMP} \text{SUBJ}) & (\uparrow\text{SUBJ FORM}) &= \text{il} \\
(\tau^{\uparrow}\text{PRED FN}) &= \text{probable} & (\tau^{\uparrow}\text{COMP}) &= \tau(\uparrow\text{XCOMP}) \\
(\tau^{\uparrow}\text{COMP}) &= \tau(\uparrow\text{COMP})
\end{align*}

TL f-structure:
\begin{align*}
&\begin{bmatrix}
\text{PRED} & \text{probable} \\
\text{SUBJ} & [\text{FORM} \text{il}] \\
\text{COMP} & [\text{COMPL que}] \\
\end{bmatrix}
&\begin{bmatrix}
\text{PRED} & \text{travailler} \\
\text{SUBJ} & [...] \\
\end{bmatrix}
\end{align*}
Knowledge-based MT: lexical entry (CMU)

(find
  (make-frame
   +find-v1
   (CAT (value v))
   (STUFF
    (DEFN "to discover by chance, to come across")
    (EXAMPLES "found X in the bedroom"
      "found X sleeping upstairs")
    "found that X was sleeping on the porch"
    (MORPH
     (IRREG (*v+past* found) (*v+past-part* found))
     (SYN-STRUC
      (*OR* ((root $var0)
          (subj (root $var1)(cat N))
          (obj (root $var2)(cat N))
          ((root $var0)
           (subj (root $var1)(cat N))
           (xcomp(root $var2)(cat N)(form pres-part)))
          ((root $var0)
           (subj (root $var1)(cat N))
           (comp (root $var2)(cat V)(form fin))))))
    (SEM
     (LEX-MAP
      (%involuntary-perceptual-event
       (experiencer (value ^$var1))
       (theme (value ^$var2))))))
UNITRAN system design

Example parameters:

Syntactic divergences: parameters

Constituent order  I have seen him – Ich habe ihn gesehen
Null subject  I saw the book – (Yo) Vi el libro

Lexical–semantic divergences: parameters

structural  X entered the house – X trat ins Haus hinein
thematic  I like Mary – Mary gefällt mir
categorial  I am hungry – Ich habe Hunger
demotional  I like eating – Ich esse gern
Lexicalist approach ('Shake-and-bake')

monolingual lexical entry (English):

\[
\begin{align*}
\text{ORTH} & : \text{like} \\
\text{SEM} & : \begin{cases} 
\text{role (E1, experiencer, X1)}, \\
\text{role (E1, stimulus, Y1)}
\end{cases} \\
\text{ARG0} & : E1 \\
\text{ARG1} & : X1 \\
\text{ARG2} & : Y1
\end{align*}
\]

monolingual lexical entry (Spanish):

\[
\begin{align*}
\text{ORTH} & : \text{gust-} \\
\text{SEM} & : \begin{cases} 
\text{role (E2, stimulus, X2)}, \\
\text{role (E2, experiencer, Y2)}
\end{cases} \\
\text{ARG0} & : E2 \\
\text{ARG1} & : X2 \\
\text{ARG2} & : Y2
\end{align*}
\]

bilingual lexical entry for \textit{like-gustar}:

\[
\begin{align*}
\text{SPANISH} [13] & : \text{SEM} \begin{cases} 
\text{ARG0} & : E \\
\text{ARG1} & : X \\
\text{ARG2} & : Y
\end{cases} \\
\text{ENGLISH} [12] & : \text{SEM} \begin{cases} 
\text{ARG0} & : E \\
\text{ARG1} & : Y \\
\text{ARG2} & : X
\end{cases}
\end{align*}
\]

Reversibility (CRITTER)

\text{eat} <-> \text{manger}
\text{miss (1: X, 2: Y)} <-> \text{manquer (1: Y', 2: X')}
\text{walk (inv-1: across (2: X))} <-> \text{traverser (2: X, inv-1: $manner (2: à_pied)$)}
General-purpose NLP systems

ELU
(Environnement Linguistique d'Unification) (ISSCO)

KIELIKONE (Finland)

Core Language Engine (SRI, Cambridge)

PLNLP
(Programming Language for Natural Language Processing) (IBM):
[SHALT-2, C-SHALT, KSHALT, PORTUGA]
## Summary of trends in rule-based approaches

<table>
<thead>
<tr>
<th>mid-1980s</th>
<th>mid-1990s</th>
</tr>
</thead>
<tbody>
<tr>
<td>syntax-orientation</td>
<td>lexicalist orientation</td>
</tr>
<tr>
<td>complex transfer rules</td>
<td>simple lexical transfer</td>
</tr>
<tr>
<td>stratificational/multi-level</td>
<td>single</td>
</tr>
<tr>
<td>representation/monostratal</td>
<td></td>
</tr>
<tr>
<td>tree transduction [filters]</td>
<td>constraints/unification</td>
</tr>
<tr>
<td>analysis/transfer</td>
<td>generation</td>
</tr>
<tr>
<td>understanding/disambiguation</td>
<td>style/quality output</td>
</tr>
<tr>
<td>uni-directional</td>
<td>reversible</td>
</tr>
<tr>
<td>linguistic information</td>
<td>lexical/conceptual</td>
</tr>
<tr>
<td>databanks</td>
<td></td>
</tr>
<tr>
<td>lexicon compilation</td>
<td>lexicon acquisition</td>
</tr>
</tbody>
</table>
Corpus-based MT

a) the direct use of information derived from corpora for the analysis, transfer and generation of translations

b) the indirect use of corpora
   - as sources of information for deriving or compiling lexical, grammatical and knowledge databases,
   - as sources of statistical information about source and target languages.
Statistics-based MT

Alignment: (IBM Candide, using corpus of the Canadian Hansard)

The proposal will not now be implemented

Les propositions ne seront pas mises en application maintenant

______________________________

English: not

<table>
<thead>
<tr>
<th>French</th>
<th>Probability</th>
<th>Fertility</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>pas</td>
<td>.469</td>
<td>2</td>
<td>.758</td>
</tr>
<tr>
<td>ne</td>
<td>.460</td>
<td>0</td>
<td>.133</td>
</tr>
<tr>
<td>non</td>
<td>.024</td>
<td>1</td>
<td>.106</td>
</tr>
<tr>
<td>pas du tout</td>
<td>.003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>faux</td>
<td>.003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>plus</td>
<td>.002</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

etc.

______________________________

Later modifications: syntactic transformations

E.g.

English questions:

Has the store any eggs? -> The store has any eggs QINV

English 'adverbs':

John does not like turnips -> John likes do_not_M1 turnips

French negation:

Je ne sais pas -> Je sais ne_pas

French pronouns:

Je vous le donnerai -> Je donnerai le_DPRO vous_IPRO
Example-based MT

DLT: Bilingual Knowledge Bank

<table>
<thead>
<tr>
<th>English</th>
<th>French</th>
</tr>
</thead>
<tbody>
<tr>
<td>the main fields</td>
<td>les principaux domaines</td>
</tr>
<tr>
<td>the following fields</td>
<td>les domaines suivantes</td>
</tr>
<tr>
<td>these two fields</td>
<td>ces deux domaines</td>
</tr>
<tr>
<td>the specialized fields</td>
<td>les domaines spécialisés</td>
</tr>
<tr>
<td>the para-medical fields</td>
<td>activités paramédicales</td>
</tr>
<tr>
<td>the magnetic fields</td>
<td>les champs magnétiques</td>
</tr>
<tr>
<td>the coal fields</td>
<td>les bassins-houilliers</td>
</tr>
<tr>
<td>the corn fields</td>
<td>les champs de blé</td>
</tr>
<tr>
<td>le livre de mon père</td>
<td>my father's book</td>
</tr>
<tr>
<td>un verre d'eau</td>
<td>glass of water</td>
</tr>
<tr>
<td>il est certain de réussir</td>
<td>certain to succeed</td>
</tr>
<tr>
<td>il est capable de résister</td>
<td>capable of resisting</td>
</tr>
<tr>
<td>il vient de Paris</td>
<td>he comes from Paris</td>
</tr>
<tr>
<td>le train de Paris</td>
<td>the train to/from Paris</td>
</tr>
<tr>
<td>il partit de nuit</td>
<td>he left at night</td>
</tr>
<tr>
<td>il partit de bonne heure</td>
<td>he left in good time</td>
</tr>
<tr>
<td>je suis âgé de trente ans</td>
<td>I am thirty years old</td>
</tr>
</tbody>
</table>

English ...have a/an effect on...   ->   French

<table>
<thead>
<tr>
<th>English</th>
<th>French</th>
</tr>
</thead>
<tbody>
<tr>
<td>have a direct effect on</td>
<td>ont une influence directe à</td>
</tr>
<tr>
<td>have a direct effect on</td>
<td>intéressent directement</td>
</tr>
<tr>
<td>have a direct effect on</td>
<td>ont eu une répercussion directe sur</td>
</tr>
<tr>
<td>has had a marked effect on</td>
<td>a largement influencé</td>
</tr>
<tr>
<td>had a positive effect on</td>
<td>s’est avérée positive dans</td>
</tr>
<tr>
<td>had a highly negative effect on [X]</td>
<td>[X] en auraient été gravement affectés</td>
</tr>
<tr>
<td>will have a decisive effect on</td>
<td>influencera de façon déterminente</td>
</tr>
<tr>
<td>would have a detrimental effect on</td>
<td>aurait de fâcheuses répercussions sur</td>
</tr>
</tbody>
</table>
Example-based method with probabilistic scores

SL input (Japanese):

US ga ... wo fusegu tame ni buhin ni kanzei wo kakeru

kakeru

ga wo ni ni

US kanzei buhin tame

mod

fusegu

similarities (with probabilities):

5.988 (meishi) ni zeikin no kakeru impose tax on (noun)
3.077 (meishi) no saibin ni kakeru take (noun) to court
2.717 (meishi) no made ni kakeru hang (noun) in window
2.554 (meishi) no sutoobu ni kakeru put (noun) on stove

5.988 (meishi) ni zeikin no kakeru

 impose

 tax on

pred

meishi

noun

'for + ing'

subj

mod

TL output (English):

US customs on in order + inf blockade
US imposes tax on parts in order to blockade...
Connectionist approaches

Spreading activation: Modification Deciding Network (Matsushita)

example of positive (co-operative) links:

1) single phrase modifying a number of verb phrases simultaneously:
   *Kare wa hon o kai, sore o yonda*
   *He book buy it read*
   *'He bought a book and read it'*
   *kare modifies both kai and yonda*

examples of negative (exclusive) links:

2) one word cannot have two different meanings in same sentence:
   *Watashi wa junen no eigo o motte iru*
   *I freshmen English take charge of*
   *'I teach freshmen English'*

   *motte means either 'take charge of' or 'have'*
   *here it can only be 'take charge of'*

3) modifications should not intersect:
   *Watashi wa kare ga kinoo sakkyokushita kyoku o kiita*
   *I he yesterday compose song listened*
   *'I listened to the song he composed yesterday'*

   *the modification of kiita by kare and of sakkyokushita*
   *by watashi would violate the non-intersection condition.*

examples of control rules:

4) obligatory case is given precedence to optional case

5) modifications close to verbs are given precedence
   to modifications more distant

6) modifications with modifier closely related semantically
   to verb are preferred
Corpora as sources of information

text corpora:
  LDC (Linguistic Data Consortium)
  EAGLES (Expert Advisory Group on Language Engineering Standards)

sublanguage or domain-specific systems

lexical/knowledge databases:
  UNITRAN
  Carnegie-Mellon systems
  ULTRA, Pangloss
  LMT
  Electronic Dictionary Research

Bilingual text databases (aligned):
  - for example-based systems
  - for translator's workstations
    AT&T
    ACQUILEX
    CWARC (Canadian Workplace Automation Research Center)

Direct use of statistical information
  PIVOT (NEC)
  ArchTran
Generation

- stylistic improvement
- discourse features
- dialogue translation
- illocutionary acts

element-based methods

multilingual generation (e.g. RAREAS)

dialogue-based MT
  (translation by monolinguals not knowing target language)

UMIST
Brussels (Babel-2)
Grenoble (LIDIA)
Kuala Lumpur (Malaysia)
Controlled, domain- and user-specific MT

Controlled input MT systems
Xerox (Systran)
Smart Corporation
Perkins Engines (Weidner)
CATALYST (Carnegie-Mellon for Caterpillar)

Domain-specific and sublanguage MT
Meteo, CRITTER, ELU
Pangloss, CATALYST
ATR speech translation
VERBMOBIL

User-specific and custom-built MT systems
Winger
Le Routier
Volmac Lingware Services
TRADEX
CSK (ARGO system)
HESS (Hangul-English Support System)
Translator Work Stations

- multilingual word processing
- optical character recognition, electronic receipt and transmission of texts
- terminology management software
- automated access to dictionaries, terminology databanks and other information sources (on-line, remote access, CD-ROM, local network, etc.)
- concordance software
- storage of and access to existing translations (for later (partial) reuse or revision)
- access to example translations [aligned bilingual text corpora]
- access to automatic translation facilities (individual words; phrases; sentence by sentence; full text)
- 'pre-translation' facilities

examples:

Canadian Workplace Automation Research Center
Carnegie-Mellon Centre for Machine Translation
TWB (Translator's Workbench): [METAL]
IBM workstation TranslationManager/2: [LMT]
English-Malay translation workstation: [JEMAH]
Momentary loads can occur repeatedly during the duty cycle but are of short duration, not exceeding 1 min. at any occurrence. When several momentary loads occur within the same 1 min. period and a discrete sequence cannot be established, the load shall be assumed to be the sum of all momentary loads occurring within that minute.

charges_momentanées peuvent occur de façon répétée pendant le cycle opératoire mais sont de courte durée, not exceeding 1 min at any occurrence. Lorsque several charges_momentanées occur within the same 1 min period and a discrete sequence ne peut pas être établi, la charge doit être assumed to être la somme de all charges_momentanées occurring within that minute.
The global view

Commercial systems since 1988:
- Globalink
- PC-Translator
- Tovna
- DP/Translator
- Toltran
- Translate (Finalsoft Corporation)
- XLT (Socatra)
- AppTek
- Hypertrans
- Lexitrans
- Language Assistant (Microtac)
- RMT/EJ (Ricoh)
- DuetQt (Sharp)
- STAR (Catena), LogoVista E to J
  (Language Engineering Corporation)
- EZ JapaneseWriter
- Meltran (Mitsubishi)

Worldwide activity:
mainly:
  United States, Canada
  Western Europe
  Japan
also:
  Malaysia (e.g. JEMAH)
  Thailand
  China (mainland and Taiwan)
  India
  Korea (HESS, KSHALT, MATES, etc.)
  Eastern Europe (AMPAR, ETAP-2, PARS)
The five eras of MT history

[1] 1947-1954:
  - Weaver’s memorandum (July 1949)
  - Bar-Hillel at MIT (1951-52)
  - word for word translation (Booth, Richens, Reifler)
  - statistical methods (Kaplan, RAND)
  - first MT conference (MIT, 1952)

  - IBM-Georgetown demonstration (Jan 1954)
  - direct translation ("first generation")
  - syntactic analysis (MIT, ATN)
  - conceptual interlingua (CLRU)
  - Georgetown systems installed (1963, 1964)

[3] 1966-76:
  - ALPAC report (1966)
  - indirect approach (‘interlingua’ CETA, LRC)
  - rule-based formalisms
  - sublanguage (Meteo 1976)
  - Systran at USAF (1970) and at CEC (1976)

  - transfer "second generation" (Ariane, Eurotra, Mu)
  - multinational, multilingual projects (Eurotra, CICC)
  - linguistics-based interlingua (DLT, Rosetta)
  - knowledge-based interlingua (CMU)
  - operational systems (Systran, Meteo, PAHO, CSK)
  - controlled input (Xerox, Smart)
  - commercial systems (Logos, Weidner, ATLAS, PENSEE, PIVOT, HICATS, ATRANSAC, DUET, METAL, Tovna, etc.)
  - PC systems (PC-Translator, Globalink)

[5] 1989-
Developments since late 1980s

International Association for Machine Translation (1991)
US government hearings (1990), JTEC report (1992)

statistics-based IBM Candide (1989)
example-based MT (1989)
constraint-based/lexicalist tendencies in rule-based MT
systems for monolingual use (Babel-R, LIDIA, UMIST)
spoken language (ATR 1986, VERMOBIL 1992)
lexicon acquisition (two conferences 1993)

end of European "second generation" projects (Ariane, Eurotra, DLT, Rosetta) and of Eastern European projects after 1989/90 political changes
start of new multilingual European project (Eurolang)
operational implementation of knowledge-based MT (Pangloss, CATALYST)
integration of MT/MAT in documentation systems (e.g. Xerox DocuTran, Krupp Industrietechnik)
user-specific custom-built systems (e.g. TRADEX, Volmac, Winger, etc.)
generation from non-textual sources
evaluation: methodology, benchmarks, etc.
<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1945</td>
<td>Weaver letter</td>
</tr>
<tr>
<td>1946</td>
<td>Weaver memo</td>
</tr>
<tr>
<td>1950</td>
<td>IBM-GU demo</td>
</tr>
<tr>
<td>1955</td>
<td>Mark II (IBM)</td>
</tr>
<tr>
<td>1960</td>
<td>Systran, GETA, LRC, TAUM</td>
</tr>
<tr>
<td>1965</td>
<td>ALPAC</td>
</tr>
<tr>
<td>1970</td>
<td>GEOMET, LRC, TAUM</td>
</tr>
<tr>
<td>1975</td>
<td>CEC Systran, Stanford, KBMT</td>
</tr>
<tr>
<td>1980</td>
<td>Weidner, DLT, Rosetta, CMU</td>
</tr>
<tr>
<td>1985</td>
<td>Fujitsu ATLAS, EBMT, Unification</td>
</tr>
<tr>
<td>1990</td>
<td>Trados, Eurolang</td>
</tr>
</tbody>
</table>
"Third generation" systems?

- Rule base less abstract than that of the 'indirect' models
- Syntactic analysis restricted to surface constituency and dependency relations (roles and cases); single monostratal representation (unification/constraint-based analysis)
- Semantic analysis limited mainly to identification of sentence and clause roles (agents, patients, etc.)
- Broad-brush disambiguation by simple semantic features (human, animal, etc.)
- Lexical information derived primarily from standard dictionary sources ('crude' syntactic categories and semantic features)
- Lexical/structural transfer rules (constraint-based) operating on monostratal ('shallow') representations

- Example translations (aligned bilingual corpora)
- Statistical data about lexical collocations and vocabulary frequencies (monolingual)
- Probabilities of lexical transfer (bilingual)
- Domain-specific knowledge bases (both linguistic and subject knowledge)
- Feedback ('learning') for grammar/lexicon improvement (connectionist ?)
- Greater emphasis on discourse and text stylistic aspects
- Integration into documentation processing and publishing systems (as in translator's workstations)
Possible precursors of the "third generation"

SHALT (IBM Japan), with five types of knowledge sources:
- grammar rules
- concept definitions
- mapping rules
- conceptual paraphrasing rules
- corpus of example sentences

ArchTran (Taiwan)
- language model (inductively derived)
- corpus as data source
- statistical translation scores (lexical, syntactic, semantic)
- probabilistic transfer
- feedback to adjust parameters for specific users

Transfer-Driven MT (ATR, Japan)
- thesaurus coding for distance measurement
- string level transfer:
  \[ \text{sochira} \rightarrow \text{this (desu), you (okura), it (miru)} \]
  \[ \text{e.g. sochira ni tsutaeru} \rightarrow \text{you convey} \]
  \[ [\text{okura (send)} \leftrightarrow \text{tsutaeru (convey)}] \]
- pattern level transfer:
  \[ X \text{ onegaishimasu} \rightarrow \text{please.../may I...} \]
- grammar level transfer:
  [grammatical category sequence]
- example database:
  \[ \text{e.g. Japanese } N_1 \text{ no } N_2 \]

fee for the conference not: fee of the conference,
conference in Tokyo not: conference of Tokyo
week's holiday not: holiday of a week
hotel reservation not: reservation of hotel
MT and users

Multilingual/multinational corporations
Translation agencies/services
  - post-edited, pre-edited, controlled input
  - single source text and multiple target languages
  - partial retranslation of revised texts

occasional non-professional translator.
  - bilingual non-translators (PC-based systems)
  - monolingual users knowing only the source language (UMIST, Babel-R, LIDIA)
  - monolingual users knowing only the target language (mainframe/batch systems, PC-based systems)

research implications:
  adoption of realistic and realizable aims
  recognition of usefulness of the less than perfect definition of user environments
  no attempt to 'mimic' the human translator
  user/developer collaboration
  user/researcher collaboration
  evaluation of systems, benchmarks, standards of performance