

Exploiting the World Wide Web for MT Teaching

Lorna Balkan,
Department of Language and Linguistics,
University of Essex

Looked at:

A) What is currently available:

- reference material
- software tools
- course material

B) Ways of improving:

- access to resources for MT teaching
- content of resources for MT teaching

WWW increasingly first port of call for seeking information

Advantages:

- fast, easy access to reference material that may otherwise be difficult to obtain
- distance no object – makes access available to a wider range of students
- access usually free

Disadvantages:

- resource discovery sometimes difficult
- quality of material not always guaranteed
- skill required in writing/using online material

MT teaching for whom?

- Computational linguists
 - Translating/interpreting students/practitioners
 - Interested public
 - Other computing personnel, e.g. I.R. researchers/developers
- Number and type of person interested in learning about MT set to increase as MT becomes more accessible
- Different users have different needs

Methods of Resource Discovery

- search engines (e.g. Altavista)
- directories (e.g. Yahoo)
- subject gateways (e.g. Sosig)
- organisational homepages (inc. academic departments)
- home pages of known experts
- project home pages
- publishers' home pages

Locating Resources:

- EAMT** and sister organisations
- ACL** (Association of Computational Linguistics)
- ELSNET** (European Network for Language and Speech)
- HLT** (Human Language technology) gateway
- Natural Language Translation Group**, British Computer Society
- Human Language Page**
- Linguist List**

Reference material

- books:
 - Cole et al.: Survey of the State of the Art in HLT
 - Arnold et al.: MT: An Introductory Guide
- Scholarly articles
- Popular articles
- Conference Proceedings
- Phd Theses

Software Tools

- MAT (Machine Aided Translation) tools include:
 - dictionaries
 - terminologies
 - thesauri
 - concordancers (e.g. Tsrali)
- MT
 - commercial
 - research
- Other NLP tools (e.g. parsers, etc.)

Other teaching materials

•most difficult category to research

•Some initiatives:

—JEWELS project: has searchable database on H.E. curricula for language and speech, including NLP

—information available on producing web-based materials

e.g. Lee et al.: Online teaching: Tools and Projects

Improving Resource Creation, Management and Exploitation

Problem common to all academic disciplines

Some initiatives:

•**ELRA**: European Language Resources Association

•**ACL/EACL 2001 Workshop** on Sharing Tools and Resources

•**DNER** (Distributed National Electronic Resource), part of JISC (Joint Information Systems Committee) in UK.

DNER

2 components:

• the underlying information environment: how to find, access and disseminate quality information resources for Higher and Further Education

• a collection of information resources of particular value to the research and education community

Resource discovery served by

• **RDN Training suite**:

a suite of teach yourself tutorials delivered over the web for students and researchers who want to improve their internet skills and learn what the internet offers in their subject discipline

• the notion of subject gateways

select/describe/develop resource collection

An MT Gateway

should:

- allow browsing and searching of material
- search should be on the basis of different criteria
(e.g. type/level of material)
- multilingual resources and interface
- Content?
 - Users need to be consulted
 - more “classical” papers
 - up to date conference proceedings