Machine Translation and Language-related Teaching

One-Soon Her
何萬順
政大語言所

Abbreviations

- **MT**: *M*achine *T*ranslation
- **LRT**: *L*anguage-*R*elated *T*eaching

Organization of the Talk

- Motivation and Goal
- Language-related Teaching
- Machine Translation Software Packages
- Potential Applications

Motivation and Goal

- MT software involves solid language research
- Its potential in LRT application is thus worth exploring

Why me?

- Experienced in MT R&D
- Had training in TESOL & linguistics
- Used MT in teaching English, translation and linguistics

# Language-related Teaching (LRT)

- Scope
- Aims

LRT: Scope

- Teaching Language
  e.g. English

- Teaching Translation and Interpretation
  e.g. English-Chinese

- Teaching Linguistics
  e.g. English linguistics

LRT: Aims (1)

- To improve students' knowledge in and/or about the target language

- in: 知其然　about: 知其所以然

LRT Aims (2)

Knowledge *in* the language

*Lang.

*Interp.

*Trans.

*Ling.

Knowledge *about* the language
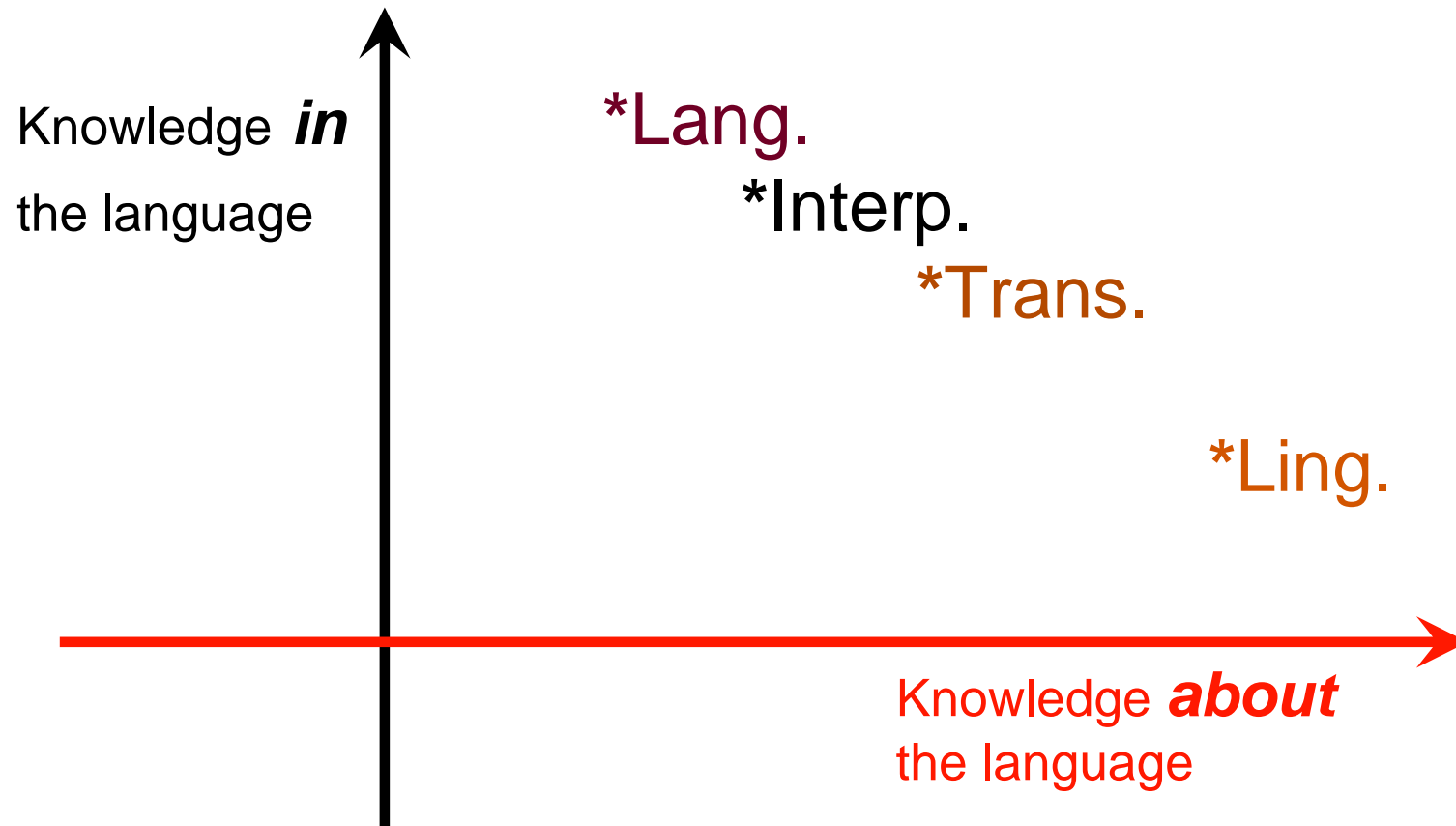
MT Software Packages

- Scope
- Aims
- Approaches to MT development

MT Software: Scope

- Automatic translation
- Related tools for interactive translation
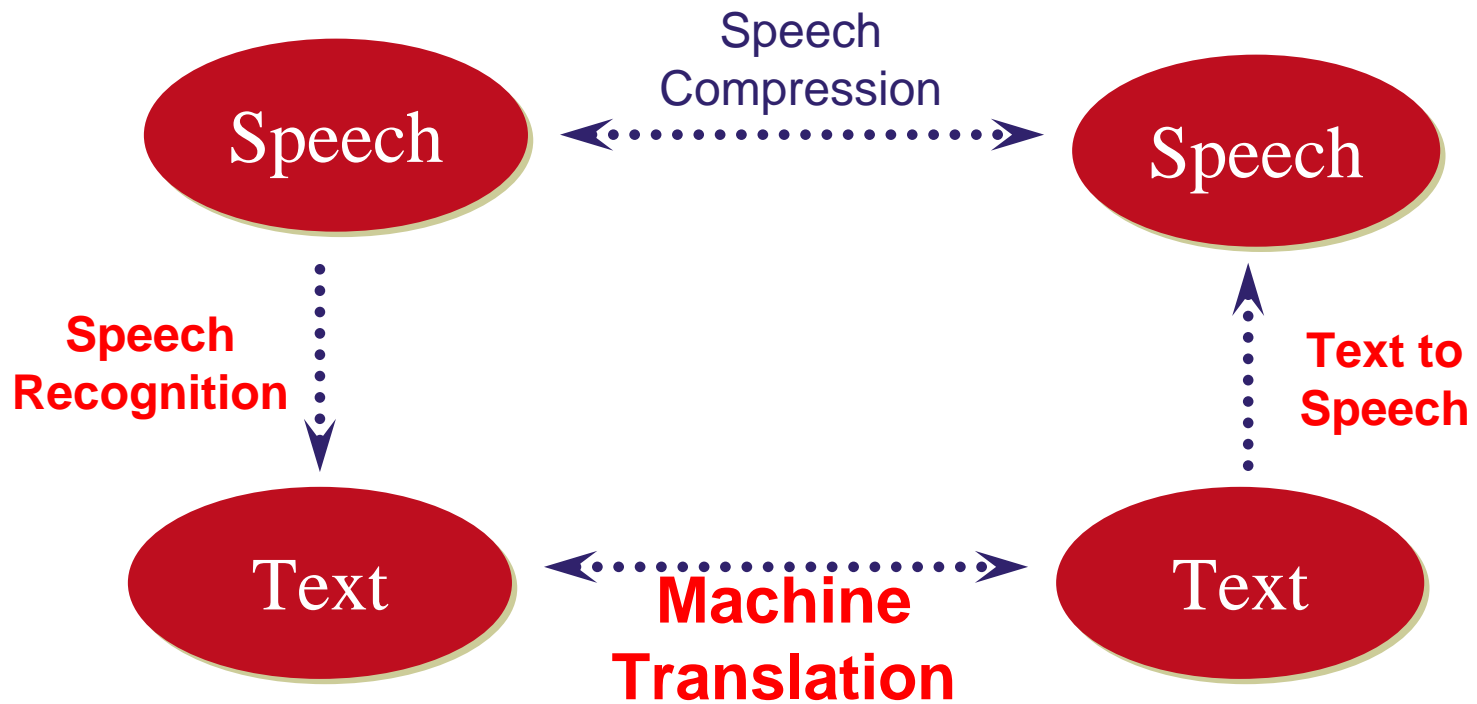- Translator's workstation

MT Software: Aims (1)

- Immediate focus:

  to translate a text fully automatically or, in most cases, *interactively with the aid of tools* provided

MT Software: Aims (2)

Long-term goal: to enable a complete cycle of language processing

Approaches to MT Development

- Linguistics-based
  - Based on lexicons and linguistic analyses
- Statistics-based
  - Based on bilingual corpuses and statistics
- Hybrid Approach
  - Uses both approaches for best results

Potential Application of MT in LRT

- In ways that may directly help improve students' knowledge *in and about* the target language

- In ways that may increase students' *interest* in learning

MT Features and potential applications (1)

- Dictionary and thesaurus on-demand:
  - *reading, writing, vocabulary*

- Special domain lexicons:
  - *ESP (English for Special Purposes), translation, and interpretation*

- Text-to-Speech utility:
  - *listening comprehension, pronunciation, speaking, interpretation*

MT Features and potential applications (2)

- Automatic translation mechanisms:
  - all LRT subjects

- Bilingual corpuses:
  - all LRT subjects

Caution!

- MT software is NOT designed specifically with LRT in mind

- MT quality varies drastically, even within the same system

- Careful planning and close supervision are thus necessary

Examples of MT Packages

- ***Dr.eye***

- ***Tranwhiz 譯經***

- ***Power Translator Pro***
  - ***former consultant (Lernout & Houspie, L&H)***

- ***TranSphere***
  - ***original developer, consultant (AppTek)***

Dr.eye

- Features worth exploring:
  - Dictionary lookup, thesaurus, grammar hints:
    can assist in reading and writing
  - Limited TTS (text-to-speech):
    can assist in the pronunciation of words
  - Limited MT: can be used as a game
  - **Low cost and high market share:**
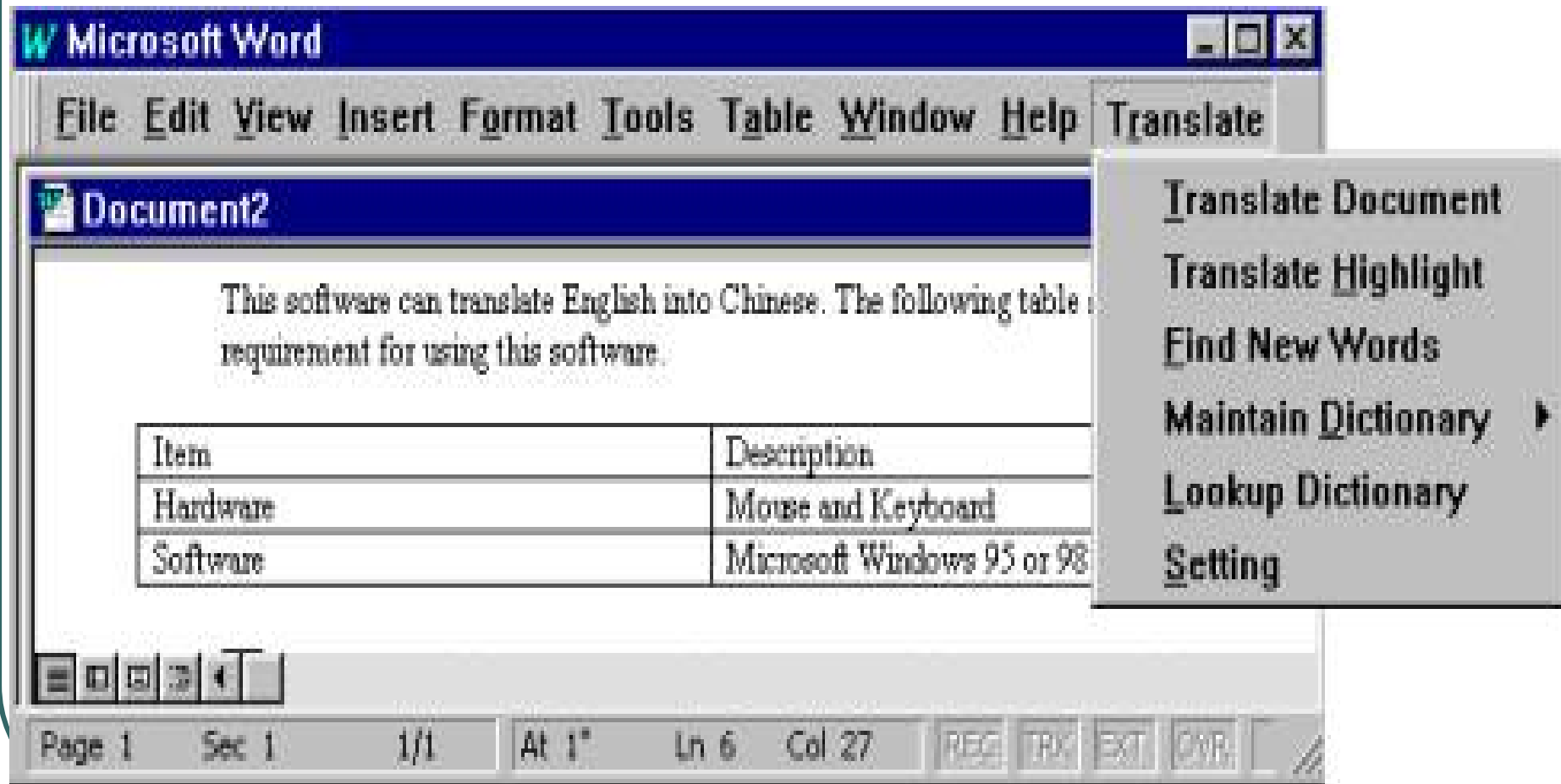   easier to promote

Dr.eye

即時寫作　多語輸入　　即時語音　　網頁翻譯　文件通　　多語複製

即時翻譯　　即時辭典　生字筆記　　全文翻譯　本地通　　多語瀏覽　加值服務

Transwhiz 譯經

- Features worth exploring:
  - (Limited) Translation Memory (TM):
    can build up your own corpus and use it for LRT purposes
  - Limited MT:
    can assist in teaching translation and linguistics
  - Text-to-Speech: ??
 - Special domain dictionaries: ESP, T&I

Transwhiz 譯經

Power Translator Pro

- Features worth exploring:
  - Voice command
  - **High-end TTS (text-to-speech):**
    can help improve listening and pronunciation
  - **Multiple European Languages**:
    English, French , Italian, Spanish,
   Portuguese, German
  - Dictionary lookup

TranSphere (AppTek) (1)

- Features worth exploring:
    - **Sophisticated TM technology**

        large tagged corpuses can be used in all areas all LRT purposes

    - **Sound linguistics-based MT**

    - Multiple diversified language pairs:

        English, Arabic, Turkish, Persian,

        Chinese, Korean, Japanese

    - Dictionary lookup and maintenance

TranSphere (AppTek) (2)
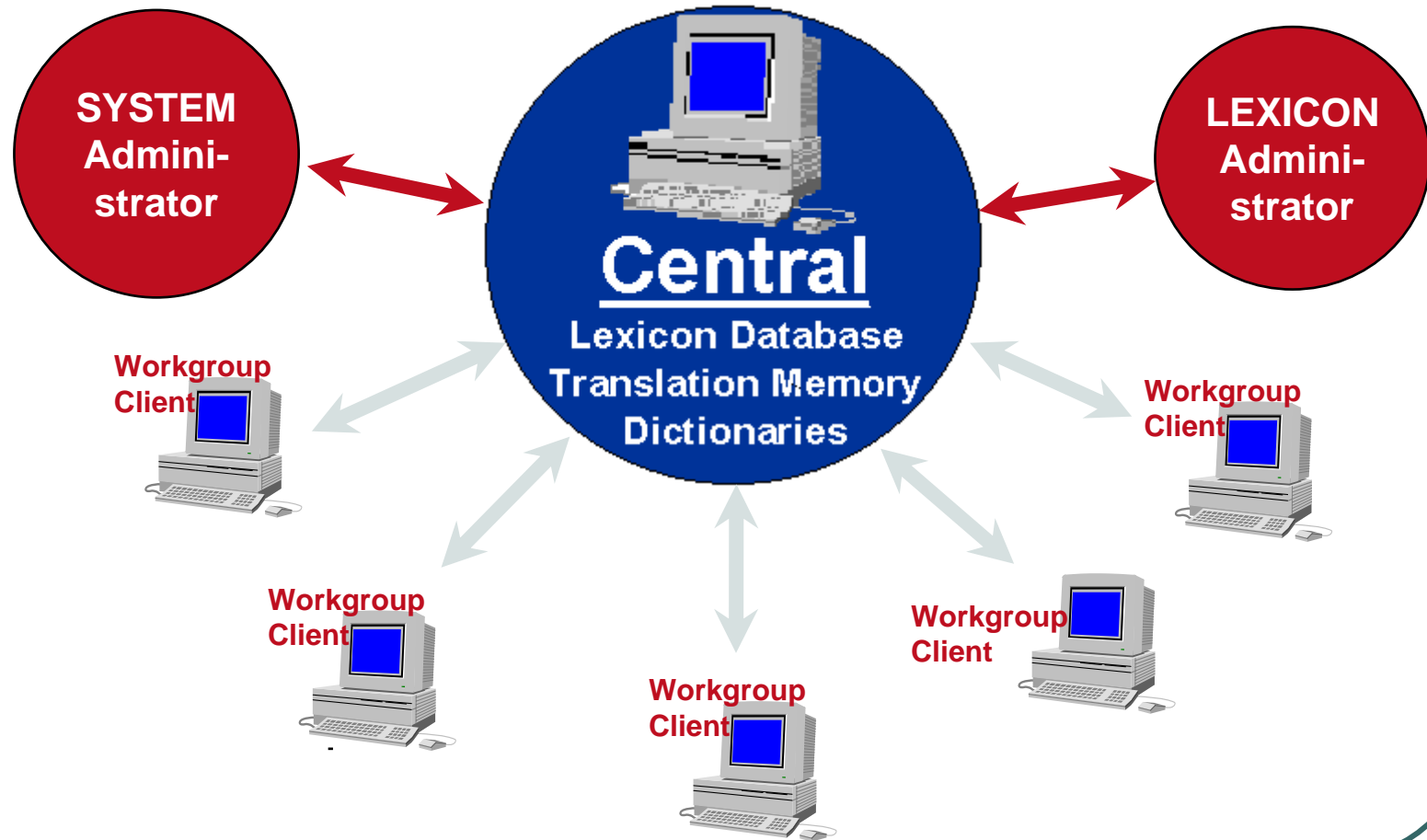
TranSphere (AppTek) (3)
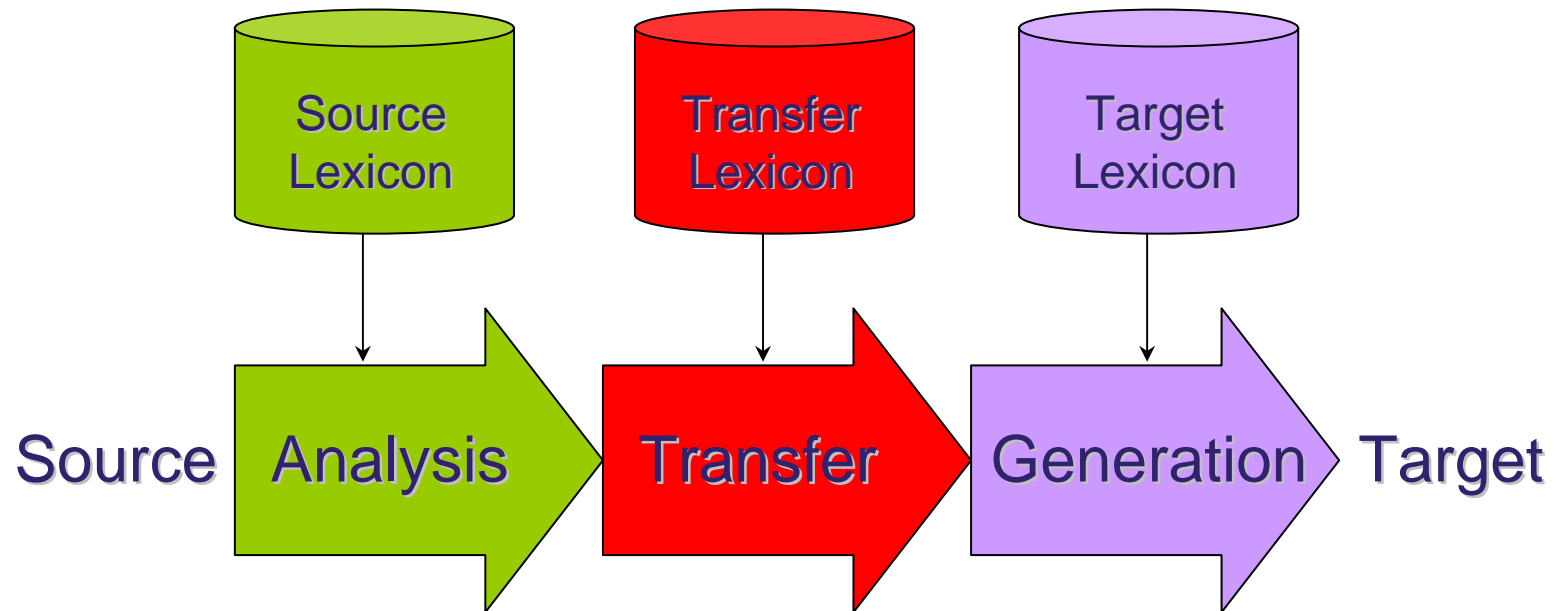
Transfer/Interlingua based on *Lexical-Functional Grammar*

(1/2)

Input

Ten thousand companies.

Analysis

(tree discarded)

```
[ FORM company
  QFR [ FORM thousand
        MULTIPLIER [ FORM ten ]
  NUM PL
]
```

Transfer

```
[ FORM 公司
  QFR [ FORM 萬
        MULTIPLIER [ FORM 一 ]
  NUM PL
]
```

Generation

```
[ FORM 公司
  QFR [ FORM 萬
        MULTIPLIER [ FORM 一 ]
  CLASS 家
  NUM PL
]
```

Output

一萬家公司。

Conclusion

- MT is worth exploring for LRT
- Advantages of MT in LRT application
    - The bonus factor
    - The high-tech image factor
    - The on-line utility factor
    - The fun factor

The Bonus Factor

- No extra cost: the MT package may already be in your computer

- Lot of useful things: MT, dictionaries, special-domain lexicons, bi-lingual corpuses, etc

The High-Tech Image Factor

- MT is on the cutting edge of technology
- High-tech gauges are attractive to many students and teachers

The On-line Utility Factor

- MT utilities are on demand and thus time-saving
- MT capacity and utilities can usually be updated on-line

The Fun Factor

- People love to beat the machine

- MT output can be quite entertaining
  e.g. I left my heart in San Francisco.

　　　我把我的心臟遺留了在舊金山裡。 *Dr.eye*

Machine Translation and Language-related Teaching

One-Soon Her
何萬順
政大語言所