

# 통제언어(Controlled Language)와 그 가능성에 대하여

2005년 5월 28일  
한국전자통신연구원(ETRI)  
언어처리연구팀  
홍 문 표

Dr. Munpyo Hong

## 목차

- Controlled Language
- AECMA Simplified English
- Controlled Language의 응용분야
- Controlled Korean

Dr. Munpyo Hong



## 특허 문장의 예 (1)

상기 뉴시제비에 들어가 걸리어 올라오면 상기 스톱(9)의 회전을 제공하여주는 제동부재(3a)를 해제위치에 두면은 상기 뉴시제비의 뉴시바늘(79)에 걸린 상기 봉어의 무게에 의해 상기 뉴시줄(73)에 힘의 장력 생겨 상기 스톱(9)을 역회전시켜 상기 뉴시줄(73)이 풀어지고 상기 스톱(9)이 역회전 할 때에 상기 스톱(9)의 측면에 돌출시킨 축(23)에 설치한 제1기어(33)도 역회전하고 상기 제1기어(33)의 수직상향방향에 설치한 제2기어(31)도 역회전하고 상기 제2기어(31)와 중일 축에 설치한 제3기어(29)도 역회전하고 상기 제3기어(29)의 수직하향방향에 설치한 제4기어(27)도 역회전하고 상기 제4기어(27)와 중일 축에 설치한 태엽(25)도 역회전하면서 상기 태엽(25)이 감아지며 상기 스톱(9)에서 뉴시줄(73)의 길이가 상기 빌 대(63)의 길이 만큼 줄어들면은 상기 스톱(9)의 회전을 제공하여주는 제동부재(3a)를 제동위치에 두고 상기 뉴시바늘(79)에 걸린 상기 봉어를 빼어 내주고 상기 뉴시바늘(79)에 띠끼를 끼우고 상기 스톱(9)의 회전을 제공하여주는 제동부재(3a)를 해제위치에 두면은 상기 감겨진 태엽(25)이 원래의 상태로 복귀하려는 물리는 힘에 의해 상기 태엽(25)과 중일 축에 설치한 제4기어(27)가 회전하고 상기 제4기어(27)의 수직상향방향에 설치한 제3기어(29)도 회전하고 상기 제3기어(29)와 중일 축에 설치한 제2기어(31)도 회전하고 상기 제2기어(31)의 수직하향방향의 상기 스톱(9)측면에 돌출시킨 축(23)에 설치한 제1기어(33)도 회전하고 상기 제1기어(33)가 회전하면서 상기 스톱(9)도 회전하면서 뉴시줄(73)을 감아준다 이때에 상기 초리대(67)의 뉴시줄가이드(69)와 중일(77)사이의 뉴시줄(73)의 길이를 조정하여주는 스톱퍼(71)가 상기 초리대(67)의 뉴시줄가이드(69)에 걸리면 상기 스톱(9)의 회전이 멈추고 상기 스톱(9)의 회전이 멈추면 상기 스톱(9)의 회전을 제공하여주는 제동부재(3a)를 제동 위치에 두고 상기 뉴시제비를 수조사이나 얼음구멍에 드리우면은 된다.

Dr. Myung-Hong



## 특허 문장의 예 (2)

- ▣ 또한, 하나의 memory chip으로 여러권의 정형책(2)을 사용할수 있는 책선택버튼(10)에 의한 책선택 기능을 채택함으로써, 언어 학습에 있어서 편리성과 효율성을 갖춘 음성재생전자책을 제공하기 위한 것이다.
  - ◆ 주어와 술어가 불명확함: “제공하기 위한 것이다”의 주어?
- ▣ 구성부품의 수가 줄어 소요비용이 저감되고, 래치의 내구성과 작동신뢰성이 향상되며
  - ◆ 저감되고 -> 절감되고
- ▣ 증폭기의 출력단으로 흐르는 전류의 소실과 싱킹을 용이하게 하여 큰 전류로 외부 회로를 구동할 수 있는 장점이 있다.

Dr. Myung-Hong



## Controlled Language (1)

- A Controlled Language is a subset of a natural language which has a restricted lexicon and controlled grammatical structures
- The objective of a Controlled Language is to improve readability, accessibility and translatability of documentation
- This is achieved through increased terminological consistency and standardization, generally simplified sentence structure, and standardized document format and layout.

Dr Munpyo Hong



## Controlled Language (2)

- particularly effective in commercial or industrial applications such as the authoring of user manuals or maintenance manuals, where large quantities of complex documents are generated and updated on a regular basis, and where terminology is domain specific
- The best-known example is European Association of Aerospace Industries (AECMA) Simplified English (SE)

Dr Munpyo Hong



## Controlled Language (3)

### ■ CL의 예

- Agilent Technologies English (Smartny, 2002)
- Attempto Controlled English (ACE)
- Boeing Technical English (Wojcik, Holmback, and Hoard, 1998)
- Caterpillar Technical English (Kamprath, Adolphson, Mitamura, and Nyberg, 1998)
- Diebold Controlled English (Moore, 2000)

Dr. Mungyo Hong



## Controlled Language (4)

- Ericsson English (Ericsson, 2000)
- General Motors CASL (Means and Godden, 1996)
- Global English (Means, Chapman and Liu, 2000)
- Kodak English (Kodak, 2000)
- Nortel Standard English (Smartny, 2002)
- Océ Technologies English (Smartny, 2002)
- Perkins PACE (Douglas and Hurst, 1996)
- Xerox MCE (Xerox, 2001)

Dr. Mungyo Hong



## Controlled Language (5)

- **Controlled Chinese (Zhang and Shiwen, 1998)**
- **Controlled Siemens Documentary German (Schachtl, 1998)**
- **GIFAS FR for French (Barthe, 1998)**
- **ScaniaSwedish (Almqvist and Hein, 1996)**

Dr Munpyo Hong



## AECMA Simplified English (1)

### ■ Background

- ◆ 항공산업 분야에서 사용되는 영어의 다양성, 복잡성
- ◆ 1970년대 후반 유럽의 European Association of Aerospace Industries (AECMA)와 미국의 Aerospace Industries Association (AIA)는 민간 항공기 제작 분야의 maintenance 문서들의 가독성 (readability)에 대해 연구를 시작
- ◆ Simplified English는 영어를 모국어로 하지 않는 독자 뿐만 아니라 영어가 모국어인 독자를 위해서 고안됨

Dr Munpyo Hong



## AECMA Simplified English (2)

### ■ History

- ◆ First SE specification was released in 1986
- ◆ Underwent several minor changes
- ◆ Issue 1 was released in 1995
- ◆ ASD Simplified Technical English (Issue 3) released in Jan. 2005
- ◆ [www.simplifiedenglish-aecma.org](http://www.simplifiedenglish-aecma.org)

Dr. Munpyo Hong



## AECMA Simplified English (3)

### ■ What is SE?

- ◆ A controlled general vocabulary and a set of writing rules
- ◆ The vocabulary has sufficient words to express any technical sentence.
- ◆ 1,565 words together with a set of 57 rules for using that vocabulary
- ◆ Besides the basic words, the writer can also use those words which he/she decides belong to one of two categories: either Technical Names or Manufacturing Processes

Dr. Munpyo Hong



## AECMA Simplified English (4)

### • Vocabulary of AECMA SE

- A list of 1,565 General Words from the SE Guide. This includes 196 verbs in four forms: the infinitive, the third person singular, the past simple and the past participle
- A list of Technical Names specific to the company
- A list of Manufacturing Processes for the company
  - The latter lists can be long. One version of Boeing Simplified English Checker used 7,000 terms (Wojcik, 1998)

Dr Mungyo Hong



## AECMA Simplified English (5)

### ▣ 사전 구성

Keyword (POS)	Assigned Meaning/Use	Approved Example	Not Acceptable
ACCURATE (adj)	"Exact"	THE ADJUSTMENT MUST BE ACCURATE	
act (v)	<u>NOTE:</u> Do not use this verb Use a "specific action" verb.	THE HYDRAULIC FLUID OPENS THE VALVE.	The hydraulic fluid acts on the inlet valve.
active (adj)	ON	WHEN THE SYSTEM IS ON, YOU WILL GET A REPORT.	When the system is active, you will get a report.

Dr Mungyo Hong



## AECMA Simplified English (6)

### ■ Writing Rules (총 57개)

#### ◆ 9개의 section으로 구성

- Section 1 - Words (13개)
- Section 2 - Noun Phrases (3개)
- Section 3 - Verbs (7개)
- Section 4 - Sentences (4개)
- Section 5 - Procedures (5개)
- Section 6 - Descriptive Writing (8개)
- Section 7 - Warnings and Cautions (5개)
- Section 8 - Punctuation and Word Counts (9개)
- Section 9 - Writing Practices (3개)

Dr. Manryo Hong



## AECMA Simplified English (7)

### ■ WORDS - Part of Speech

**RULE: 1.2** Use approved words from the Dictionary only as the part of speech given.

Each approved word in the Dictionary has a part of speech. Do not use it as another part of speech for which it is not approved. For example, if a word is given only as a noun, do not use it as a verb.

**Example:** "Test" is approved as a noun but not as a verb.

**WRITE:** *Do the leak test for the system.*

or

*Do a test for leaks in the system.*

**NOT:** *Test the system for leaks.*

**Example:** "Close" is a verb (and not an adverb).

**WRITE:** *Do not go near the landing gear if ...*

**NOT:** *Do not go close to the landing gear if ...*

Dr. Manryo Hong





## AECMA Simplified English (8)

### ▣ WORDS

**RULE: 1.8 Do not use different Technical Names for the same thing.**

If you decide to call something e.g. “servo control unit”, call it that in all places (including on illustrations). Do not call it “actuator” for example, somewhere else.

Dr Munpyo Hong



## AECMA Simplified English (9)

### ▣ NP

**RULE: 2.1 Do not make noun clusters of more than three nouns.**

**Example:**

**WRITE:** The bolt that attaches the uplock to the nose landing gear is ....

**NOT:** The nose landing gear uplock attachment bolt is ....

Dr Munpyo Hong



## AECMA Simplified English (10)

### ▣ VERB - How to Describe an Action

**RULE: 3.7** If there is an approved verb to describe an action, use the verb (not a noun or other part of speech).

If there are different ways to write something, use the construction in which an approved verb shows the action. This is because the primary function of a verb is to express action.

Example:

**WRITE:** The ohmmeter shows  $450 \pm 22$  ohms.

**NOT:** The meter gives an indication of  $450 \pm 22$  ohms.

**WRITE:** Make sure that you hear the 600 Hz tone at the same time.

**NOT:** Ensure that the 600 Hz is audible when the GPW GND ACT switch is depressed.

Dr. Manpyo Hong



## AECMA Simplified English (11)

### ▣ Sentence: Tabular Layout

**RULE: 4.3** Use a tabular layout (vertical layout) for complex texts.

Example:

**WRITE:** The controls on the main panel, from the top to the bottom, are:

- An OFF/ON main switch
- A START push button
- A STOP/O.S. TEST push button.

**NOT:** From top to bottom, the controls on the main panel consist of an OFF/ON main switch, a START push button, and a STOP/O.S. TEST push button.

Dr. Manpyo Hong



## AECMA Simplified English (12)

### ▣ Procedures

**RULE: 5.1** Keep procedural sentences as short as possible (20 words maximum)

The maximum length of a procedural sentence is 20 words. (Refer to Section 8 for the rules about word count.)

### Example:

**WRITE:** (1) Remove the electrical power from the system.  
(Seven words)

(2) Make sure the refueling panel switches go back to their correct position.  
(Twelve words)

**NOT:** After you have removed the electrical power from the system, make sure that the refueling panel switches go back to their normal position. (Twenty-three words)

Dr. Manpyo Hwang



## AECMA Simplified English (13)

### ▣ Descriptive Writing

**RULE: 6.1** Keep sentences in descriptive writing as short as possible (25 words maximum).

In descriptive writing, a sentence length of 25 words is permitted. This is because descriptive text is more complex than procedural text.

Dr. Manpyo Hwang



## AECMA Simplified English (14)

### ■ Warnings

**RULE: 7.1** Start a warning or a caution with a simple and clear command.

Warnings and cautions must start with a simple and clear command. This command tells the technicians what they must do to avoid the danger and/or damage. Do not hide this command as a general comment with other background information. Start with the command, and then give the supporting information.

#### Example:

**WRITE:WARNING:DO NOT GET THE ENGINE OIL ON YOUR SKIN. THE OIL IS POISONOUS. IT CAN GO THROUGH YOUR SKIN AND INTO YOUR BODY.**

**NOT:WARNING:THE SYNTHETIC LUBRICATING OIL USED IN THIS ENGINE CONTAINS ADDITIVES WHICH, IF ALLOWED TO COME INTO CONTACT WITH THE SKIN FOR PROLONGED PERIODS, CAN BE TOXIC THROUGH ABSORPTION.**

Dr. Munyoo Hong



## AECMA Simplified English (15)

### ■ Punctuation

**RULE: 8.1** Use colons (:) and dashes ( - ) to make tabular layouts.

#### Example:

When you set the equipment to ON:

- Adjust the BRT switch to MAX
- Adjust the CONT switch to MIN.

Dr. Munyoo Hong



## AECMA Simplified English (16)

### ▣ Writing Practices

**RULE: 9.2** When you combine words to make a phrase, make sure that the words continue to obey the meanings given to them in the Dictionary (Part 2).

*Example:*

*Instead of:* When the fire is put out, close the valve.

*WRITE:* When the fire is extinguished, close the valve.

(PUT is approved in SE, and OUT is approved also. However, "put out" in the example does not obey the approved meanings of either word in SE.)

Dr Mungyo Haug



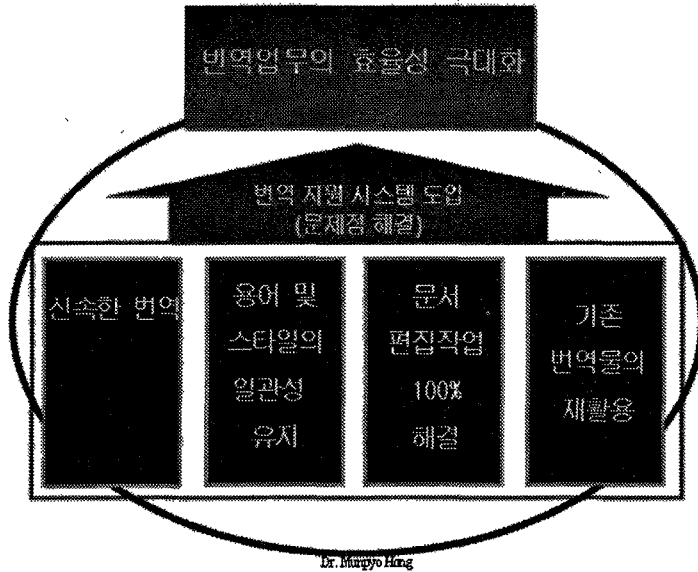
## 응용분야 (기계번역)

- EU 연방법원의 자체 번역센터 (Übersetzungsbüro am Europäischen Gerichtshof)
  - 230명의 전문 번역가 보유
  - 매년 33만 페이지의 법률문서 번역
  - 신속한 번역을 위하여 MT 시스템 사용 (시간당 2천 페이지 번역)  
→ 번역 결과는 번역가의 수작업을 통해 개선됨
  - 정확하고 일관성 있는 번역을 위해 TM 시스템 사용

Dr Mungyo Haug



## 비용절감을 위한 기업측의 요구사항



## 적용 사례: Océ Technologies

- ▣ Products: Copiers, plotters, printers and office solutions for office and engineering systems
- ▣ Employees: 3250 head office , 23000 worldwide, R&D 1000
- ▣ Head office: Venlo, The Netherlands operating companies in 30 countries, active in 80 countries
- ▣ Annual sales :Euro 3.2 billion

Dr. Munpyo Hong



## Translation background info

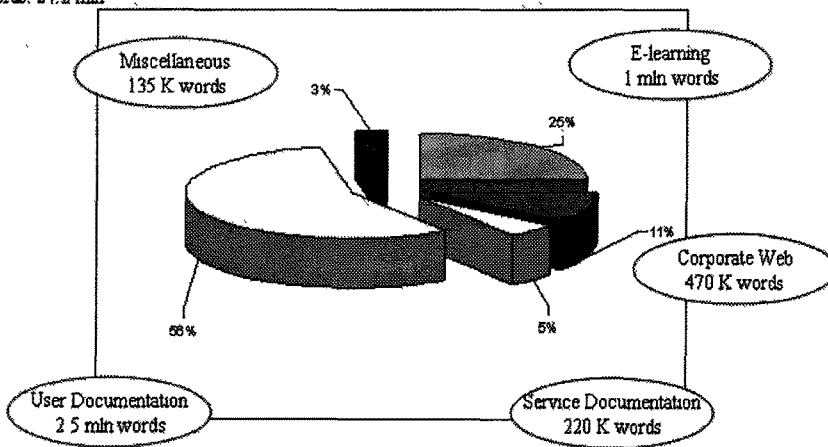
<b>Source language</b>	English
<b>Document formats</b>	Frame, Word, HTML, Lotus Notes, XML
<b>Document types</b>	User documentation 17 languages Service documentation 5 languages Software, Web content, E-learning .....
<b>Translation volume</b>	2.5 mln source -> 23 mln target words (2002)
<b>Savings</b>	approx. 55%
<b>Tooling</b>	TM (Trados TWB), MT (Logos), ...

Dr. Munpyo Hong



## Documentation Sources

Period: 15 months  
 Source words: 4.4 mln  
 Target words: 24.2 mln



Dr. Munpyo Hong



## Translation Services

**Goal**      make all kinds of documentation translatable,  
produce large volume of high quality translations  
in a short time at low cost

**Methods**     highly automated workflow  
 standardisation (terminology, file formats etc)  
 automated translation tools

Automation in-house, actual translation out-sourced

Dr. Mungyo Hong



## Processes introduced 1995-2003

**Localisation Formula:** MT - TM - CE - T&S - XML

**MT:** Machine Translation

**TM:** Translation Memory

**CE:** Controlled English

**T&S:** Internal course; teaches writing principles

**XML based Content Management System**

**Reasons:**

Cost reduction

Shorter release cycles

Quality improvement

Increasing translation volume

Dr. Mungyo Hong





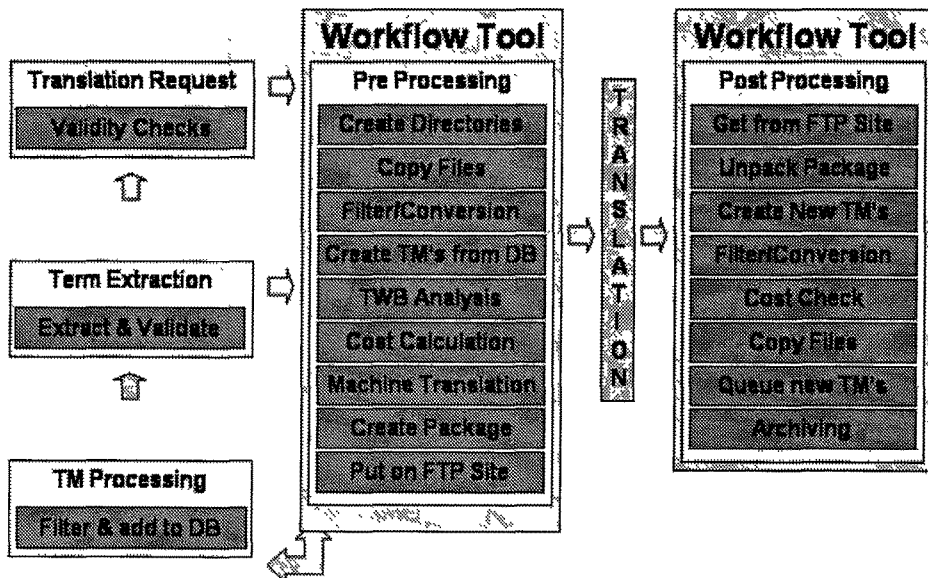
## Controlled English

### Benefits

- consistent use of terminology
  - grammatical English
  - improves efficiency XML CMS
    - structured generic elements –reuse
  - improves performance MT system
  - reduction of superfluous information
- 
- initially time consuming
  - "black" authoring
  - terminology control

Dr Munpyo Hong

## Translation Service Workflow



Dr Munpyo Hong



## Controlled Language/Controlling Language

- ▣ CL을 사용하면 읽기 쉽고, 번역하기도 쉬운 문서가 될 것은 확실하다. 그러나 어떤 Technical Writer가 AECMA SE 같은 세세한 제약들을 다 숙지하고, 모든 제약에 위배되지 않는 문장을 쓸 수 있을까?
- ▣ CL이 Technical Writer의 생산성과 의욕을 떨어뜨리지 않을까?

### ▣ 현재의 추세:

Controlled Language → Controlling Language

“쓰고 싶은대로 써라... 일단 문서 작성이 끝난 후 무엇을 잘못 썼는지 알려주마... 그대신 CL Checker가 제안하는대로 수정하라”

Dr. Manpyo Hong



## CL Checker

- ▣ KANT CLC (CMU, USA)
- ▣ CLAT (IAI, Germany)
- ▣ XEROX CL Checker (Xerox, France)
- ▣ ...
- ▣ 주요기능
  - ◆ Unapproved Term Detection
  - ◆ Spell, Syntax Error Detection
  - ◆ Style Error Detection
  - ◆ Term Inconsistency Checking
  - ◆ ...
  - ◆ 주로 CAT (Computer Aided Translation)-Tool에 embedding된 형태로 사용됨

Dr. Manpyo Hong



## Controlled Korean

▣ 한국어를 위한 CL 개발의 시기가 왔음

▣ CL에 대한 정부, 기업의 인식 변화

▣ 예:

◆ 특허문서 자동번역 (다음 슬라이드)

- 번역 엔진 성능 향상만으로는 한계가 있음
- 특허문서 일본 자체의 문제가 많음: 심사관, 사무관 인식 변화

◆ 글로벌 기업

- 중국 내에 번역만을 위해 연구소를 별도로 두고 있는 실정
- 유면번역 및 기계번역을 위해 깔끔하고 이해가 쉬운 문서작성이 절실히 요구됨

Dr. Munpyo Hong



## 한영 특허 자동번역 - 배경 (1)

▣ 자국 기술보호를 위한 특허정보의 해외보급 강화 시급

◆ 각국 특허청에서 모든 심사처리 정보를 인터넷 공개 추진 (일본, 미국, 유럽 등)

◆ 해외에서의 한국 기술에 대한 관심으로 인해 특허정보에 대한 영문 보급 요구 급증

◆ 매년 94,000 여건의 신규 특허 문서 발생 -> 수동 번역 불가능

◆ 일본 특허청 - 인터넷 일정한 자동번역 서비스 제공 시작함

Dr. Munpyo Hong



## 한영 특어 자동번역 - 배경 (2)

### ▣ 특허청에서 한영 자동번역 기술 사전 조사 후 ETRI에 요청(2003.11)

◆ 특허청 : 국내 최고 수준의 한영 자동 번역 시스템들에 대한 사전 조사 수행

○ 현재 국내 기술 수준(번역을 50%)으로는 자동번역 서비스가 불가능하다고 파악함

◆ ETRI : 샘플 문서 simulation 통해 실용화 가능성을 보여줌

Dr. Manryo Hwang



## 한영 특어 자동번역 - 배경 (3)

### ▣ 자국 기술보호를 위한 특허정보의 해외보급 강화 시급

◆ 각국 특허청에서 모든 심사처리 정보를 인터넷 공개 추진 (일본, 미국, 유럽 등)

◆ 해외에서의 한국 기술에 대한 관심으로 인해 특허정보에 대한 영문 보급 요구 급증

◆ 매년 94,000 여건의 신규 특허 문서 발생 -> 수동 번역 불가능

◆ 일본 특허청 - 인터넷 밀명안 자동번역 서비스 제공 시작함

Dr. Manryo Hwang



## 연구목표

### ■ 올해 최종 연구 목표 (2005)

- ◆ 한영 특어문서 자동번역 기술 : 특어형 상용 서비스 실시

### ■ 2004년도 연구성과

- ◆ 일영 특어문서 자동번역 서비스 수준 (번역률 : 78%)
- ◆ 한영 특어문서 자동번역 특어형 내부 시범 서비스 (전기전자 분야)

### ■ 2006년도 연구목표

- ◆ 특어문서 작성을 위한 한국어 CL 개발
- ◆ 한국어 CL 적용을 위한 CL Checker 개발

Dr. Mungyo Hong



## Controlled Korean 개발을 위한 제안

### ■ 기본 어휘 사전 구성

- ◆ 최소의 규모로 의미적 모호성이 없고, 어떠한 내용도 충분히 표현할 수 있는 단어리스트 확보가 필요
  - 인지언어학적 관점
  - 전산언어학적 관점
- ◆ 각 분야별 전문용어 리스트의 구축 및 정비가 필요

### ■ 문장 규칙

- ◆ 문장 유형별로 문장 이해에 가장 적합한 평균 단어수 파악
- ◆ 이해가 어려운 문장 유형 파악 (인지언어학적 관점)
- ◆ 분석 및 번역이 어려운 문장 유형 파악 (전산언어학적 관점)

### ■ 언어학자, 전산학자, 심리학자, Technical Writer 등의 공동작업이 필요함

Dr. Mungyo Hong