



# 기계번역 시스템

박 철 제<sup>†</sup> 김 태 완<sup>††</sup>

## ◆ 목 차 ◆

- |                |                |
|----------------|----------------|
| 1. 서 론         | 4. 해외의 연구개발 동향 |
| 2. 영한 기계번역 시스템 | 5. 결 론         |
| 3. 일한 기계번역 시스템 |                |

## 1. 서 론

최근의 고도 정보화 사회를 맞이하여 국제간의 경쟁이 심화되고 전략적인 제휴가 빈번해 짐에 따라 실제 업무에 이용 가능한 기계번역 시스템의 필요성이 급속도로 증대하고 있다. 1952년, 기계번역에 관한 최초의 국제회의가 미국의 MIT에서 개최된 것을 계기로 1954년의 러시아어-영어 번역 시스템의 데모, 1956년의 프랑스어-러시아어의 데모가 행하여져 기계번역 시스템에 관한 연구개발이 활발히 이루어 졌다. 그러나 기계번역처리에는 현실 세계의 지식이 필요하며 그러한 지식을 정형화하는 것이 불가능하므로 양질의 자동번역 시스템 개발이 불가능하다는 Bar-Hillel의 지적을 계기로 많은 토론이 행하여졌다. 그 이후 1966년 ALPAC(Automatic Language Processing Advisory Committee) 보고서의 영향으로 인하여 기계번역 시스템에 관한 연구 개발은 동면기를 맞이하였으나 일본, 유럽, 캐나다에서의 꾸준한 연구개발로 인하여 오늘날 많은 상용 기계번역 시스템이 발표되고 있다[4,9,12,13]. 언어학적 연구에 있어서는 Chomsky의 「변형생성문법이론」이

혁명적인 현대언어 이론으로서 인식되어져 1970년대 중반까지 영어 구조를 중심으로 광범위한 언어현상의 해명에 이용되었으며, 1968년 Fillmore가 제창한 동사의 의미구조를 기초로 한 「격문법」은 일본어의 연구에 이용되어 오늘날 일본에서의 기계번역 시스템 대부분이 격문법을 채용하고 있다[2].

현재 국내에서 가장 많이 연구되고 또한 보급되어 있는 기계번역 시스템은 주로 일본어와 영어를 대상으로 하고 있다. 컴퓨터가 다루기 어려운 인간의 언어를 그 처리 대상으로 하고 있다는 점에서 기계번역은 상당히 어려운 연구 분야 중 하나이다. 그러나, 일본어를 대상으로 하는 일한 기계번역 시스템은 양 언어의 문법적 성질과 어순의 유사성으로 인해 현재의 기술 수준으로도 일정한 번역 품질을 확보할 수 있다. 일본어와 한국어는 언어 계통상 알타이어에 속하며 문법체계가 많은 부분에서 유사하다. 특히, 어절 단위에서는 양 언어의 어순이 대부분 일치하기 때문에 어절을 기본 단위로 하여 양 언어를 일대일로 대응시켜 번역 시스템을 구축하여도 상당한 수준의 번역 결과를 얻을 수 있으므로, 일한 기계번역 시스템에서는 문법적 유사성을 이용한 직접 번역 방식을 채용하고 있다. 그러나 직접 번역 방식을 이용한 기계번역 시스템을 구축할 때 직면하는

† 정회원 : 동의대학교 경영정보학과 교수  
 †† 정회원 : 한국전자통신연구원 선임연구원

문제점인 어휘의 중의성(Lexical ambiguity)과 부분적인 어순의 조정 문제에 의하여 양 언어를 일대일로 대응시킬 수 없는 경우가 발생하며, 일한 기계번역 시스템을 실용화하기 위해서는 특히 어휘의 중의성 문제 해결이 중요한 과제로 되어 있다.

한편, 영어를 그 처리 대상으로 하고 있는 영한 기계번역 시스템은 아직 이러한 수준에 도달하지 못하고 있다. 한국어-일본어와는 달리 한국어-영어는 양 언어의 문법적 성질과 문장 구조, 어순 등이 전혀 다르기 때문에 일한 기계번역에서와 같이 비교적 단순한 처리 방법으로는 영한 번역을 달성할 수 없기 때문이다.

본 논문에서는 영한, 일한 기계번역 시스템을 중심으로 연구 현황과 기본 특성을 살펴보고 시스템 구축시의 문제점 및 앞으로의 발전 방향에 대해서 고찰한다. 또한 해외의 연구 개발 동향에 대해서 개략적으로 살펴보기로 한다.

## 2. 영한 기계번역 시스템

### 2.1 영한 기계번역의 문제점

이 장에서는 영한 기계번역 기술 개발에 있어서 해결해야 할 여러 가지 문제점에 대해 설명한다[10,11].

#### 2.1.1 영어 문장 분석에 있어서의 문제

##### 1) 어휘 중의성(Lexical ambiguity)

하나의 단어가 2가지 이상으로 분석이 가능한 것을 나타낸다. 여기에는 품사 중의성(Category ambiguity), 동형의의어(Homograph), 다의어(Poly semy), 대역어 중의성(Transfer ambiguity) 등이 있다.

- 품사 중의성(Category ambiguity)

하나의 단어가 여러 가지 품사를 갖고 있는 경우이다. 분석 단계에서 올바른 품사를 선택하지 못하면 전혀 엉뚱한 번역문이 만들어진다.

예) I can do it : 나는 그것을 할 수 있다.(can

조동사)

나는 짱통 그것 한다.(can 명사)

- 동형의의어(Homograph)

하나의 단어가 전혀 다른 두 가지 의미를 가지고 있는 경우이다. 역시 올바른 의미로 분석하지 못하면 전혀 다른 의미의 번역이 되어 버린다.

예) Bank of Ohio : 오하이오 은행, 오하이오 강둑

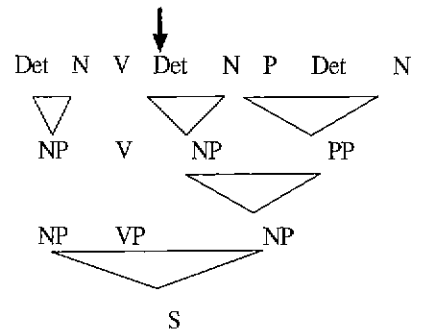
하나의 단어가 의미는 유사하나 전혀 다른 상황에서 쓰여질 수 있는 경우이다.

예) mouth of river : 강의 입 (?), 하구

##### 2) 구문 구조의 중의성(Structural ambiguity)

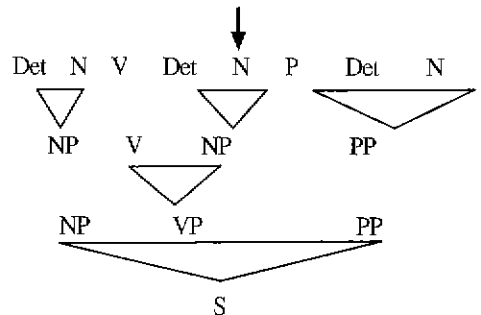
영어 문장을 구문 분석(Parsing) 하였을 때 여러 가지 분석 결과(Parsing Tree)가 만들어지는 경우이다. 만들어 낸 분석 결과에 따라 번역문의 의미에 큰 차이가 생길 수 있다.

예) The man saw the girl with the telescope.



그 사람은 망원경을 가지고 있는 소녀를 보았다.

The man saw the girl with the telescope.



↓

그 사람은 그 소녀를 망원경으로 보았다

### 2.1.2 영한 변환에 있어서의 문제(Transfer problem)

영한 기계번역은 2.1.1에서 설명한 영어 분석에 있어서의 문제점 외에 한국어로의 변환 시에 발생하는 다음과 같은 문제점을 해결할 필요가 있다.

#### 1) 양 언어간 어휘 개념의 차이(Lexical differences)

영어에서 하나의 의미로 결정되었어도 한국어에서는 여러 가지 대역어로 번역될 수 있는 경우를 말한다. 예를 들면 영어 단어 home은 한국어에서 문서의 성격에 따라 가정으로도 집으로도 번역될 수 있으며, driver라는 단어는 한국어에서는 운전자, 마부, 기관사 등으로 대역어가 달라질 수 있다.

#### 2) 양 언어의 구조적 차이(Structural differences)

간단한 단문(Simple sentence) 구조에서도 영어는 주어 + 술어 + 목적어의 어순이지만 한국어는 주어 + 목적어 + 술어의 형태를 취한다. 이러한 구조의 차이는 여러 개의 절을 포함하는 중문, 복문의 경우 더욱 커지며 양 언어간의 번역을 어렵게 하는 주요 요인이다.

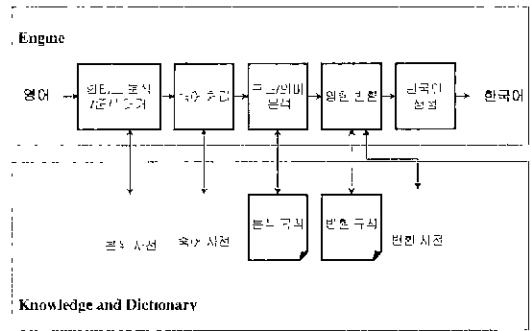
## 2.2 영한 기계번역 시스템의 개요

이 장에서는 영한 기계번역 시스템의 일반적인 구성과 번역 수행 절차에 대해서 설명한다.

영한 기계번역 시스템의 일반적인 시스템 구성은 (그림 1)과 같다.

이하 (1)의 예문을 대상으로 영한 번역이 수행되는 절차의 개요를 설명하도록 한다. 2.1에서 설명한 여러 가지 문제점을 해결하기 위한 방법들이 시스템에 따라 다양하게 도입되고 있으나, 이러한 방법들에 대한 자세한 설명은 본고의 범위를 넘어서는 것이므로 언급하지 않는 것으로 한다.

입력문장: I abstained him from going home. (1)



(그림 1) 영한 기계번역 시스템의 구성도

### 2.2.1 형태소 분석/품사 태깅

영어 문장에는 예문의 abstained, going과 같이 동사는 물론 명사, 형용사 등이 활용된 형태로 나타난다. 반면, 분석 사전에는 정보의 중복 기술, 불필요한 사전 관리를 배제하기 위하여 단어의 원형만을 등록하여 두는 것이 일반적이다. 따라서, 올바른 사전 정보를 찾기 위해서는 입력문장에 나타나는 활용형을 원형으로 복원하여 주어야 한다. 또한, 각 단어의 품사 정보 등 번역에 필요한 각종 정보를 추출하는 것이 필요하다. 영어 형태소 분석은 바로 이러한 역할을 담당하는 것이다. 최근에는 품사 태거를 이용하는 경우도 있다. 품사 태거란 영어 단어의 앞뒤에 오는 문맥 정보를 보고 해당 단어의 품사를 결정하는 것이다. 다음은 예문의 형태소 분석 결과를 보인 것이다.

```
((pron ((ROOT I) (NLEX I) (LEX I) (CAT pron)))
(verb ((ROOT abstain) (NLEX abstained) (LEX abstained) (CAT verb)))
(pron ((ROOT him) (NLEX him) (LEX him) (CAT pron)))
(preposition ((ROOT from) (NLEX from) (LEX from) (CAT prep)))
(verb ((ROOT go) (NLEX going) (LEX going) (CAT verb)))
(verb ((ROOT go) (NLEX going) (LEX going) (CAT verb)))
(noun ((ROOT home) (NLEX home) (LEX home) (CAT noun)))
```

### 2.2.2 속어 처리

영어에는 예문의 abstained ~ from ~ing와 같이 속어가 자주 사용된다. 영어를 한국어로 유연하게 번역하기 위해서는 여러 개의 단어로 구성된 속

어를 찾아 내어 하나의 단어처럼 취급하는 등, 적절한 처리가 필요하다. 다음은 예문에 대한 속어 처리 결과를 보인 것이다.

```
((pron ((ROOT I) (NLEX I) (LEX I) (CAT pron)))
(verb ((CU T) (CU_KEY restrain) (ROOT abstain-#1-from)
(NLEX abstained) (LEX abstained-him-from) (CAT verb)
(TRV_K_LEX1 기를 삼가도록 #1에게 권하)))
(verb ((CU T) (CU_KEY go) (ROOT go-home) (NLEX going)
(LEX going-home) (CAT verb) (TRV_K_LEX1 귀가하))))
```

### 2.2.3 구문 분석

영어를 한국어로 올바르게 번역하기 위해서는 영어 문장을 구성하는 각 구성 성분(단어 또는 구) 사이의 수식 관계를 분석하여야 한다(2.1.2 구문 구조 중의성 참조). 또한 각 구성 성분의 문법적 역할(주격, 목적격 등), 능동태/수동태, 긍정문/부정문/의문문 등 번역에 필요한 정보를 결정하여 주어야 한다. 이러한 기능을 수행하는 것이 구문 분석이다. 아래는 예문에 대한 구문 분석 결과이다.

```
(S (NP (PRON_ (PRON ((ROOT I)(NLEX I)(LEX I)(CAT PRON)(NUMBER S)(CASE SUBJ)(ROOT_LEX I))))
(VP (VERB_ (VERB ((ROOT ABSTAIN-#1-FROM)(NLEX ABSTAINED)(LEX ABSTAINED-HIM-FROM)(CAT VERB)(FORM PAST EN)(V_TYPE T4)(SEM_MK DURATIVE)(#1 #1_PARSE) (TRV_K_LEX1 기를 삼가도록 #1에게 권하)(ROOT_LEX abstain-#1-from))))
(INGP (VERB_ (VERB ((ROOT GO-HOME)(NLEX GOING)(LEX GOING-HOME)(CAT VERB)(FORM ING)(V_TYPE IO)(SEM_MK DURATIVE)(TRV_K_LEX1 귀가하)(ROOT_LEX go-home))))))
```

### 2.2.4 영한 변환

영어 구문 분석 결과로부터 한국어 구문 구조를 만들어 낸다. 변환 사전으로부터 영어 단어의 대역어 및 한국어 번역에 필요한 한국어 품사 정보, 격 정보, 조사, 어미 등을 생성하고, 나아가서 한국어 문법에 맞는 구문 구조 및 어순 등을 결정한다. 2.2.2에서 설명한 영한 변환시의 여러 중

의성을 해결하기 위한 기법들이 이 단계에서 도입된다. 아래는 예문에 대한 영한 변환 결과이다.

```
(PRED ((VOICE ACTIVE) (TENSE PAST) (MOOD DECL)
(TRV_K_GCODE1 EV00016) (TRV_K_LEX1 기를 삼가도록 그
에게 권하) (TRV_EK_DEEP_CASE1 AGT)(EOMI ㄴ다 )
(COMN (( (S_CASE SUBJ) (TRN_K_POS PRON)
(TRN_K_LEX1 나) (TRN_K_GCODE1 NN00001)(JOSA 는)))
(PREN ((VOICE ACTIVE) (TRV_K_GCODE1 NV00016) (TRV_K_LEX1 귀가하) (S_CASE DOBJ)
(TRV_E_POS V) (TRV_E_VTYPE L9) (TRV_EK_DEEP_CASE1 AGT))))
```

### 2.2.5 한국어 생성

최종적으로 한국어의 활용, 띄어쓰기 등 한국어 맞춤법에 맞는 자연스러운 한국어 문장을 만들어 내는 단계이다.

출력 문장: 나는 귀가하기를 삼가도록 그에게 권하였다.

## 2.3 영한 기계번역 시스템 기술 동향

90년대 중반부터 여러 업체들이 번역 소프트웨어들을 출시하고 있으며 영한 번역 소프트웨어와 일한 번역 소프트웨어들이 주종을 이루고 있다. 영한 번역 소프트웨어로는 언어공학연구소의 트레니와 IBM의 앙코르, 정소프트의 워드체인지, 리틀컴퓨터의 번역 마당 등의 제품이 있다. 그러나, 일한 번역 소프트웨어들이 사용자들에게 어느 정도 인정 받을 수 있는 번역 품질을 내놓고 있는 것에 비해 영한 번역 소프트웨어들은 실생활에 적용하기에는 아직 만족스럽지 못한 성능을 보이고 있는 것이 현실이다.

다음은 현재 시판되고 있는 일반 문서 영한 번역 시스템과 웹 영한 번역 시스템의 현황을 살펴본다.

### 2.3.1 일반 문서 영한 번역 시스템 현황

- 앙코르

한국 IBM과 서울대학교 자연어처리 연구실이

공동 개발한 영한 번역 소프트웨어이다. 속어 및 관용어를 포함한 모두 20만 단어의 영한사전을 가지고 있으며, 고교 1학년 영어 교과서를 소화해 낼 수 있는 수준이며 국내 최초로 화일 단위는 물론 문장 및 절 단위 번역이 가능하다.

● 트래니

언어공학연구소에서 개발한 영한 번역 소프트웨어이며 한 시간에 10만 단어를 처리하는 빠른 번역 속도를 자랑하며 20단어 이상의 긴 문장도 번역이 가능하다는 점이 특징이다. 20만 단어와 속어가 내장된 영어 기본사전과 20개 분야별 전문 용어 사전을 가지고 있다.

● 워드체인지

정소프트에서 개발한 영한번역 소프트웨어이다. 10만 단어 영한사전, 8만 단어 속어 관용어사전과 전문가 사전을 가지고 있으며 30초 이내에 1백 개 문장을 처리할 수 있으며 윈도즈 상에서 네트워크와 연결시켜 온라인상에서 캡처 받은 내용을 즉시 번역해 활용할 수 있다.

2.3.2 웹 영한 자동 번역 시스템 현황

월드 와이드 웹(WWW)의 보급에 따라 영어 정보에 접할 기회가 증가하고 있다. 현재는 이러한 영어 장벽의 불편을 해소하기 위한 방법으로 일반 사전이나 전자 사전을 이용하는 비율이 높지만 이제 인터넷 영한 번역 시스템의 출사로 인하여 사용자는 영어의 장벽에 구애 받지 않고 웹상에서 정보를 수집하는 것을 기대하게 되었다.

현재 출시된 인터넷용 영한 번역 시스템으로는 앙꼬트(IBM), 워드체인지(정소프트), 사이버트랜스(언어공학연구소), 세종대왕(미래 소프트웨어) 등이 있으며, 인터넷 번역 기능과 브라우저 기능을 통합한 영한 번역 시스템으로 성운시스템의 세계로96이 있다. 그러나, 웹에서 사용되는 영어는 일반 저작과 달리 그 표현 체계가 자유롭고 Image 정보, Hyper-link 정보 등 문서 내에 다양한

HTML tag가 나타나며, 저작자의 강조, 인용, 생략 등의 의도에 따라 비문법적인 문장이 다수 출현하는 등 여러 가지 문제점을 내포하고 있어 일반 문서용 영한 기계번역 시스템의 번역 품질보다 더욱 낮은 품질을 보이고 있다.

3. 일한 기계번역 시스템

3.1 일한 기계번역 시스템의 문제점

일한 기계번역 시스템에서도 어휘의 중의성 문제는 크게 품사 중의성, 동형이의어 및 다의어의 3가지로 나누어 진다. 이 중에서 품사의 중의성 문제는 형태소 분석의 단계에서 휴리스틱(Heuristic)을 이용하여 해결되어지나 동형이의어 및 다의어의 경우는 구문, 의미정보를 가지고서 해결하지 않으면 안된다. 일한 기계번역 시스템 구축에 있어서 어휘의 중의성 문제는 고품질의 번역 결과를 얻기 위한 중요한 과제가 되고 있다.

예) 彼は金/父母を失う.

예문의 失う의 경우 한국어로 번역될 때 잃다, 여의다 두 가지로 나누어 진다. 여의다의 의미로서 사용될 때는 金 앞에 사람의 의미가 포함된 체언이 올 경우이다. 이와 같이 같은 단어가 다른 의미로 쓰여지는 것을 중의성이라고 하며, 직접번역 방식에 의한 일한 기계번역 시스템에서 잘못된 번역 결과의 대부분은 다의어의 경우를 제외하고는 용언과 조사의 처리 부분에서 발생한다.

용언의 중의성 문제를 해결하기 위한 방법으로서 문헌[5]에서는 격문법을 이용한 단문의 의미해석 수법을 제안하고 있다. 동사와 명사의 공기관계를 의미소 대신으로 공기관계테이블에 의하여 표현하고 있다. 조사의 중의성 문제를 해결하기 위해서 ATLAS-I에서는 조사 앞에 나타나는 단어의 연결관계를 이용하고 있다. 또한 문헌 [7]에서는 각각의 조사 앞에 접속하는 단어의 의미와 문법정보를 고려한 단어번역 테이블을 이용하고

있다.

또한, 한국어와 일본어의 어절간에는 어순이 일치하지만 술부내의 문법형태소에서 어순의 차이점을 보여 주고, 일본어와 한국어 형태소 간에 일대일 대응이 되지 않는 경우가 많다. 이러한 점으로 인하여 직접 번역 방식의 시스템에서는 자연스러운 한국어 술부를 생성하기 어렵다. 이러한 문제를 해결하기 위하여 문헌[14]에서는 한국어의 술부에 나타나는 양상류 의미 자질들의 관계를 표현한 자료 구조인 MFOLT(Modality Feature Ordering and Lexicalizing Table)와 한국어 접속 정보를 이용하는 생성 방법을 제안하고 있다.

### 3.2 일한 기계번역 시스템 연구현황

일본어와 한국어의 유사성을 이용하여 대학 및 연구소를 비롯한 기업에서도 활발한 연구가 이루어져 최근에는 성능이 우수한 상용화 시스템이 발표되고 있다. 또한 각 제품에 대한 성능의 비교 분석도 활발하게 이루어져 학회에 보고되고 있으며 그 중 대표적인 제품으로서 J-SEOUL/JK, 오경박사, 한글가나, 명품 등이 알려져 있다. 대학 및 기업에서의 상용화 현황에 대해서도 여러 보고 문건이 발표되었으며 자세한 사항은 문헌[1,8]에 잘 나타나 있다.

일한 기계번역 시스템의 상용화 시대를 맞이하여 기계번역 시스템 도입을 원하는 사용자의 입장에서 시스템을 평가하기 위해서는 여러 가지 부분을 고려하지 않으면 안된다. 번역 결과물의 평가 방법을 비롯하여 번역 속도, 번역을 위한 비용(컴퓨터 사용료, 번역 소프트웨어 및 사전의 가격 등), 번역 시스템 운용에 필요한 인원수, 후편집을 위한 인건비의 예측, 사전 데이터의 관리 및 경신을 위한 인원, 사전의 단어 증가를 위해 필요한 인건비 등등[6] 수 많은 요소들을 고려해서 사용자에게 편리한 시스템을 제공 할 수 있어야 하겠다.

## 4. 해외의 연구개발 동향

### 4.1 미국의 연구개발 동향

미국에서는 1986년 카네기 멜론 대학(CMU)에 기계번역센터(CMT)가 설치되어 기계번역 및 관련 기술 연구가 본격적으로 재개되었다. 남가주의 정보과학연구소(ISI), 뉴멕시코 주립대학의 계산연구소(CRL)와 공동으로 분석, 생성, 사전의 개발을 분담하는 공동연구 「범언어(PAN-GLOSS)」 프로젝트가 1991년 여름부터 3년간 약 50억원 규모의 예산으로 국방성에서 자금을 지원하였다. 미국에서의 자연언어처리 연구는 정보검색, 인공지능과의 연계테마로 설정되어 기계번역도 지식처리를 가미한 의미분석 프로세스를 시스템에 구축하지 않으면 평가받지 못하는 경향이 있다. 1991년에는 음성번역·대화 시스템도 발표되었다. CMU에서는 회의용 영·독·일 대화 번역 시스템(JANUS)이 개발되었으며 ATT 벨 연구소, MIT에서도 대화형 데이터 베이스 검색 시스템(ATT:VEST, MIT:VOYAGER/TINA)을 개발하였다. 또한, 국방성 고등연구계획국(ARPA)에서는 항공권 구입용의 음성대화시스템의 개발을 장려하기 위하여 기업에 연구비를 지원하고 있다. 상용의 번역 시스템으로서는 「SYSTRAN」, 「METAL」 「LOGOS」가 있으며 전미보전기구(PAHO)에서 개발한 스페인어-영어, 영어-스페인어 번역 시스템(SPANAM/ENGSPAN)이 각각 1980년, 1985년부터 가동되고 있다.

### 4.2 일본의 연구개발 동향

일본에서는 교토(京都) 대학의 나가오(長尾) 연구실을 중심으로 Mu 프로젝트(1982 ~ 86)가 과학논문 초록 일-영 번역 시스템의 개발을 추진하였다. 통산성 전자기술종합연구소, 정보전산센터, 과학기술청의 일본과학기술정보센터가 참가하여 기업에서 대학에 연구원을 파견하는 국가적 사업으

로 추진되었으며 여기에서 쌓여진 기술이 기계번역 시스템 상용화의 기반이 되었다. 1993년에는 문부성 중점영역 연구 「음성·언어·개념의 통합적 처리에 의한 대화 이해 및 생성에 관한 연구」가 시작되어 주요 전기·전자·통신 및 소프트웨어 회사들이 참가하여 번역 시스템을 개발하였으며 ATR, NEC에서 음성 대화 시스템을 발표하였다. 기계번역을 중심으로 한 자연언어처리는 교토(京都) 대학, 오사카(大阪)대학, 동경공업대학, 규슈(九州)대학, 전자기술종합연구소 등이 기초 연구를 수행하고 있다. 또한, 정보처리 진흥사업 협회에서는 1981년부터 국어학자를 중심으로 「계산기용 일본어사전 IPAL」을 구축하였다.

정부와 민간 기업 공동 프로젝트로서는 거대한 전자사전을 구축하여 국제적으로 주목을 받고 있는 일본전자화사전연구소(EDR), 대화체 언어의 완전 자동번역을 목표로 하는 국제전기통신기초기술연구소(ATR)의 자동번역전화연구소가 1986년 설립되었다. 자동번역전화연구소에서는 회의사무용 대화번역 시스템(SL-TRANS(일·영)/ASURA(일·영·독))을 개발하였으며 1993년 음성번역통신연구소로 개명하였다. 일본에서의 각 프로젝트는 기업이 연구원을 파견하여 운영자금을 부담하고 있으며 70% 정도를 통산·우정성의 기초기술연구 촉진센터가 자금을 지원하고 있다.

## 5. 결 론

기계번역 시스템에 관한 연구가 시작 된지 50여년이 지나 이제는 상용화 시스템이 우리들 생활 속으로 침투되고 있으며 기술적 한계도 명확하게 인식되어 지고 있다. 본 논문에서는 영한, 일한 기계번역 시스템을 중심으로 연구 개발 동향과 현 시스템이 가지고 있는 문제점을 살펴 보았다.

일한 기계번역 시스템은 수준 높은 제품들이

지속적으로 발표되고 있으며 향후 번역 품질의 향상 뿐만 아니라 사용자의 편리성을 고려한 여러 기능들이 갖추어진 시스템이 출시될 것으로 기대된다. 영한 기계번역 시스템의 번역 품질은 실생활에 적용하기에는 아직 미흡한 수준이다. 그러나, 상용화된 시스템의 출현은 시장으로부터 기술 개선 요구를 Feed-back 받아 기술 향상이 이루어진다고 하는 기술진화의 한 과정으로서 높게 평가 받아야 할 것이다. 또한, 영한 기계번역 시스템의 경우 완전 자동 번역기의 역할을 수행할 수 있는 기술 수준을 확보하기까지는 번역 생산성 향상 도구로서의 개발 전략이 필요하리라 생각되며, 아울러 현재의 번역 품질 수준에서도 실용성을 확보할 수 있는 응용분야의 탐색과 그에 따른 특성화된 시스템의 개발이 필요하다.

## 참고문헌

- [1] 김태석, “일한 기계번역 시스템의 연구 및 개발”, 한국정보과학회지, 제 15권, 제 10호, pp.9-15, 1997.
- [2] 成田 一 外 2人, こうすれば使える機械翻譯, パベル書プレス, 1994.
- [3] 朴 哲濟 外 3人, 日-韓機械翻譯における連語パターンを用いた變換手法, 日本情報處理學會論文誌, Vol. 38, No. 4, pp.707-718, 1997.
- [4] 野村 浩郷 外 1人, 機械翻譯, bit別冊, 日本共立出版, 1988.
- [5] 李 震東 外 2人, 語と語の關係を用いた意味解析による日韓單文機械翻譯システム, 日本電子情報通信學會論文誌, Vol. J72-D-II, No. 10, pp.1689-1695, 1989.
- [6] 長尾 眞, 機械翻譯はどこまで可能か, 岩波書店, 1989.
- [7] 金 泰錫 外 1人, 日韓機械翻譯における意味接続關係を用いた韓國語の生成方法, 日本情報處

理學會論文誌, Vol. 33, No. 12, pp.1578-1588, 1992.

[8] 崔 杞鮮, 金 泰完, 日韓機械翻譯システムの現状および分析, 日本言語處理學會 第 2回年次大會論文集, pp.433-443, 1997.

[9] Bonnie Jean Dorr, Machine Translation : A view from the Lexicon, The MIT Press, 1992.

[10] Doug Arnold, Lorna Balkan et al, Machine Translation : An Introductory Guide,



박 철 제

1986년 연세대학교 수학과(학사)  
1991년 일본 와세다대학교 대학원  
정보과학과(석사)  
1998년 일본 와세다대학교 대학원  
정보과학과(박사)  
1986년-1987년 포항종합제철  
시스템개발부

1987년-1989년 Japan Knowledge Industry  
1995년-1997년 일본 와세다대학교 미디어 네트워크  
센터  
1996년-1998년 현대정보기술연구소 책임연구원  
1998년-현재 동의대학교 경영정보학과 교수  
관심분야 : 기계번역, 정보검색, 한국어 정보처리,  
인터넷, CALS/EC



김 태 완

1983년 한양 대학교 전자공학과  
(학사)  
1985년 한양 대학교 대학원 전지  
공학과(석사)  
1985년-현재 ETRI 자연어처리  
연구부 선임연구원

관심분야 : 기계번역, 한국어 정보 처리, 다국어 정보  
검색

## SQMS '98

### 제2회 소프트웨어 품질관리 심포지움 개최 발표논문 및 사례발표 모집

- ◎ 일 시 : 1998년 11월 11일(수) ~ 12일(목)
- ◎ 장 소 : 한국과학기술회관(강남역 부근)
- ◎ 내 용 :
  - 11일(수) : 등록, 개회식, 초청강연, 튜토리얼, 패널토의
  - 12일(목) : 품질전문가 튜토리얼, 논문발표 및 사례발표
- ◎ 논문투고 요령 : 학회 논문투고 양식에 준하여 6페이지 이내
- ◎ 논문마감 : 1998년 10월 20일(화)까지
- ◎ 문의전화 : (02)593-2894 팩스 (02)593-2896