

佛語—韓國語 翻譯：人間과 機械의 接近*

尹 愛 善
(부 산 대 학 교)

목 차

| | |
|-----------------------|-----------------------|
| 제 1 장 서 론 | 제 3 장 불한 기계 번역에 대한 접근 |
| 제 2 장 인간의 번역과 기계 번역 | 제 1 절 형태소 분석 |
| 제 1 절 번역 과정 | 제 2 절 단일화에 의한 문장 분석 |
| 제 2 절 기호 전환 방식과 그 문제점 | 제 3 절 불어—한국어 변환 |
| 제 3 절 기계 번역의 종류 | 제 4 장 결 론 |

제 1 장 서 론

본 연구에서는 전통적으로 행하여 왔던 인간의 번역(human translation)과 현재 논의가 활발히 진행되고 있는 기계 번역(machine translation)의 제 문제점을 비교, 검토하여 기존 기계 번역 시스템보다 향상된 불어—한국어 기계 시스템 구현을 위한 방향을 제시하고자 한다.

현재 번역에 대한 연구는 크게 두 가지로 나뉜다. 하나는 인간 번역사를 양성하면서, 실제 번역 경험을 바탕으로 연구가 진행되는 번역학 분야다. 세계 대전과 국제 교류 증대에 따라 외교, 경제, 학술 분야 등에 회담, 회의, 세미나 등이 급격하게 늘어나 많은 수의 인간 번역사를 요구했고 이에 부응하여 전문적인 양성기관이 생겨나면서부터, 번역을 더 이상 언어에 대해 특별한 재주를 타고난 사람만이 가진 기술이 아니라, 가르치고 습득할 수 있는 분야로 간주하려는 경향이다. 기계에 비해 고도의 지능을 가진 인간이, 피번역어와 번역어에 대한 전문 지식 등을 통해 번역의 목적인 “피번역문의 실제 의미를 번역문으로 재표현하는 것”¹⁾에 도달하는 것이다. 번역에 대한 또 다른 추세로는, 인간번역에 비해 질은 떨어지지만 경제성과 빠른 속도로 다량의 번역을 가능하게 해주는 기계 번역에 대한 연구를 들 수 있다. 20세기 중반과 후반에 들어오면서 특히 기술과학 분야의 빠른 발전은, 우리로 하여금 단시간 내에 많은 정보 습득을 요구하고 있다. 그런데 20세기 중반의 기대했던 대로 영어나 러시아어가 과학분야의 링가 프랑카(lingua franca)가 되는 경향과는 반대로 이 분야의 중요한 논문이

* 이 논문은 학술진흥재단의 “1989 자유공모 과제 연구비 지원”에 의하여 연구된 것임.

1) SELESKOVITCH & LEADERER (1986:9~19).

자국어로 쓰여지는 추세가 날로 확산되고 있다. 따라서 인간의 번역처럼 고도로 정밀하지 못하나, 대략 정보의 내용의 유효성을 판단할 수 있게 해주는 수준의 값싼 번역이 다량으로 필요하게 되었다. 더우기 computer의 hardware와 software의 끊임없는 발전, 언어학 이론의 도입에 따른 형태, 통사, 의미 분석 등 lingware 개발, 또 인공 지능(artificial intelligence)의 적용으로 기계 번역의 난점이었던 중의성(ambiguity)의 상당 부분을 해결함으로써, 기계 번역은 50년대에 시작되어, 60~70년대 암흑기를 거쳐 80년대부터 다시 각광을 받기 시작했다.

한 언어를 다른 언어로 번역하는 경우 실제로 여러 어려움에 직면하게 된다. 자연언어는 그 특성상 형태적, 통사적, 의미적 또는 화용적으로 중의성(ambiguity)을 띠고 있어 의미 해석 과정에서 번역에는 문맥(context)이 결정적인 역할을 한다. 또 피번역어(source language)에서 쉽게 표현되는 개념이 다른 언어에서는 대응되는 형태소나 표현을 찾기 어려운 경우가 많다. 그리고 분야에 따라 같은 단어가 다른 의미를 나타내는 용어로 사용되므로 번역 대상 영역에 대한 전문 지식(domain specific knowledge)이 필요하다.

인간과 기계 번역이 동시에 가지는 이러한 어려움 외에도 기계 번역에는 수많은 문제가 발생한다. 표현(expression)에 직결된 문제를 포함해서 입력 문장의 정확한 해석 및 중의성 해결에 필요한 지식과 문맥 처리 기법이 아직 확립되어 있지 않을 뿐더러 자연언어 처리에 대한 이론이나 알고리즘이 제시되지 못하였다.

번역 정보의 증대, 빠른 속도, 경제력 등의 장점으로 그 수요 증가가 확실시되는 기계 번역은 인간의 번역에 비해 그 질의 차이는 엄청나지만, 현재 기계 번역이 안고 있는 문제점은 결국 인간 번역의 모델에서 그 해결의 실마리를 찾아, 더욱 개발되어야 할 것이다. 본 연구의 목적은 경험과 이론을 바탕으로 한 번역학과 언어학에서 연구되는 인간 번역 과정을 살펴봄으로써, 특히 불어-한국어 기계 번역이 추후해야 할 방향을 제시해 보고자 하는 데 있다.

제 2 장 인간의 번역과 기계 번역

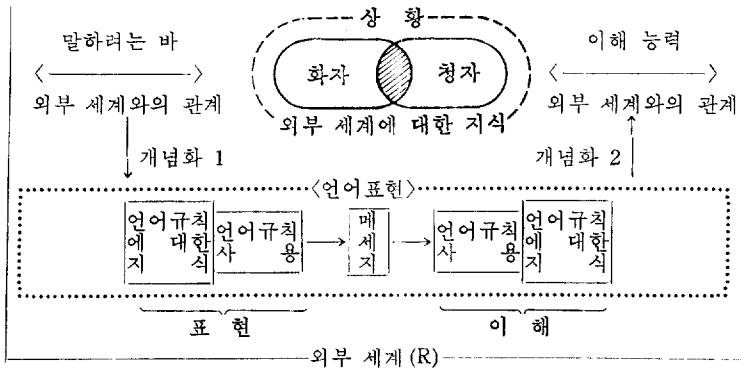
제 1 절 번역 과정

우선 두 언어를 재료로 하는 번역 과정은 한 언어를 사용하는데 나타나는 이해(comprehension)와 표현(expression) 과정에 비견된다. 화자는 자신이 말하려는 바를 “언어 기호”로써 표현하고 청자는 전달된 언어 기호를 통해 화자의 말하려는 바를 “이해”한다. 번역은 이러한 의사 소통(communication) 과정을 두개의 언어를 사용하여 이중으로 수행하는 것이며, 번역사는 첫번째 과정에서는 청자 역할을, 두번째 과정에서는 화자의 역할을 수행한다. 즉 그림 1에서 번역사는 청자 1과 화자 2의 역할을 순서대로 행하고, 그 목적은 실제 의미 A, B, C가 동등하도록(equivalent) 하는 것이다.



[그림 1 번역 수행 과정]

그러면 실제로 번역이 어떤 작업이어야 하는 가를 알기 위해, 언어 사용과정을 자세히 살펴볼 필요가 있다.



[그림 2 communication의 수행 과정²⁾]

복잡해 보이는 이 그림은 communication에서 기능을 수행하는 여러 함수를 보여 준다. 즉, 화자와 청자는 “외부 세계”에 대한 “지식”을 갖고 있으며, 그것과 “관계”를 맺고, 발화 당시 “상황”의 영향을 받는다. 다음 화자가 “말하려는 바”는 언어로 표현되기 이전에 “개념화 1” 되어, 이것이 화자가 그 언어에 대해 알고 있는 음, 형태, 통사, 의미 규칙 등 “언어 규칙”을 “사용”하여 “메세지”로 구성된다. 이것이 메세지 “표현” 과정이다. 청자는 이 메세지를 받아들여 화자와 같은 언어규칙을 통하여 이를 분석하여 “개념화 2”를 시키고, 이런 작업을 통해 화자가 말하려는 바를 “이해”하게 된다. 이것이 메세지 “이해” 과정이다.

언어 표현 부분을 좀 더 구체적으로 살펴 보면 크게 세 단계로 나눈다.³⁾ 첫째, 우리가 지각하는 사건은 개별 언어로 표현되기 이전에, 보편적 (universal) 성질을 띤 인간의 사고를 바탕으로 개념적 차원 (concept)에서 그 사건의 구성이 파악된다. 둘째, 개별 언어 단계에 들어가 이 사건을 표현할 어휘를 선택하고, 화자가 그 사건을 바라보는 관점에 따라 선택된 문장 구조를 형성한다. 이 단계는 개별 언어가 지닌 언어 규칙의 지배를 받는 랑그 (langue) 차원이다. 셋째, 담화 (discours) 단계로서 사건을 표현하는 언어 규칙이 실제로 사용되는 단계다.

2) POTTIER, (1987:9~12).

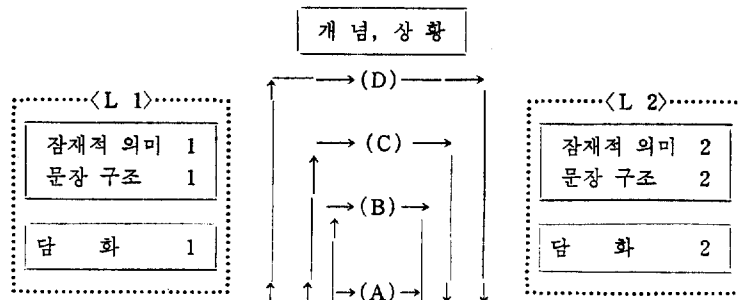
3) Ibid., p. 105.

상황(situation, context)이 개입하기 때문에 명시적(explicit)으로 표현되지 않을 수도 있을 것이고, 중복(redundant)되어 나타날 수도 있다.

| 보편성 | 1. 개념적 단계 | 사건의 구성 파악 | 상황, 사고 |
|-----|-----------|----------------------------|-------------------------|
| 개별성 | 2. 랑그 단계 | 어휘 선택 관점 선택 문장 구조 선택 | 어휘부 (잠재적 의미부) 통사부 |
| | 3. 담화 단계 | 언어 규칙의 실제 사용 | 상황 |

[표 1 언어 사용 단계]

언어 사용 과정이 다차원적이므로 어느 부분까지 고려하느냐에 따라 상이한 번역 방법이 존재한다.



[그림 3 언어 사용 단계와 번역 단계]

(A)는 이른바 단어 대 단어 번역 방식(word-to-word translation)으로 기계 번역의 직접 번역(direct) 방식이다. L1의 담화에 나타난 언어 기호를 L2의 언어 기호로 바꾸는 단순한 방법이다. (B)는 L1의 통사구조에 대응되는 통사구조를 L2에서 찾아 어휘를 대치하는 방법으로 기계 번역에서 가장 널리 사용되는 구조 변환(structural transfer) 방식이다. (C)는 흔히 (B)에 추가되어 발달된 의미 전환(semantic transfer) 방식으로 이중어 사전(bilingual dictionary)을 통해 L1과 L2의 어휘의 잠재력 의미 간에 동등성을 결정하고, 문장구조, 단어, 동음이의어 등으로 인한 중의성(ambiguity) 문제를 해결하는 방식이다. (A), (B), (C)는 D. Selekovitch가 기호 전환(transcodage)이라고 부르는 방식으로 인간 번역에서 하기 쉬운 오류를 나타낸다. 즉, 인간 번역은 (D) 과정을 따라야 한다고 주장한다. 즉, L1의 언어 표현에 의해 제약을 받지 않고 상황과 개념화를 통해 담화 1의 실제 의미를 파악하고 다시 이 실제 의미를 L2로 표현해야 한다는 것이다. 이 방법이 초기 기계 번역 CETA program⁴⁾이 추구하다 실패한 interlingua(또는 pivot) 방식이다.

4) CETA(Centre d'Etudes sur la Traduction Automatique)는 Grenoble 대학에서 1960년에서 1970년까지 러시아-불어 기계번역을 위해 B. Vauquois 의해 추진된 project로서 자연 언어와 독립된 의미 표현을 통해 번역하고자 하는 interlingua 방식을 채택했다. 이것의 실패로 1970년부터 GETA(Groupe d'Etudes sur la Traduction Automatique)은 변환 방식(transfer approach)으로 추진되고 있다.

제 2 절 기호 전환 방식과 그 문제점

그러면 인간 번역에서도 발견되고 현대 기계 번역이 주로 사용하는 기호 전환방식(trans-codage)과 그 문제점을 알아보자.

(A)나 (B)방식은 특히 피번역어와 번역어가 같은 어족이거나 근친 어족이어서 형태소 결합방식이나 문장 구조가 유사한 언어 간의 번역에서 자주 나타난다.⁵⁾

| 피번역어 | 번역어 | 피번역어의 실제 의미 |
|---|---|--|
| (1) (a) J'aime la vie (b) surchargé (c) Where are you? | Ich liebe das Leben überlasten Où êtes-vous? | |
| (2) (a) actually (b) conventional weapons (c) birth control | actuellement armes conventionnels contrôle des naissances | en fait armes classiques limitation des naissances |
| (3) (a) Il n'y a rien de tel que... | There' s no such thing as... | There's nothing like... |
| (4) Out with it! | Dehors avec cela! | Vas-y! Accouche! |
| (5) I met him on the train | Je l'ai rencontré sur le train | Je l'ai rencontré dans le train. |

영어, 독어, 불어의 경우로 (1)은 단어 대 단어 대응인 직역이 실제 의미를 그대로 전달할 수 있는 경우로 그 가능성은 극히 제한된다. (2)는 인간 번역에서 흔히 있는 실수다. 어원이나 형태(signifiant)의 유사성 때문에 나타내려는 의미를 제대로 전달하지 못하는 것인데, 이것은 인간의 기억력 소관이므로 이를 어휘부에 수록할 수 있는 기계 번역에서는 이러한 오류가 많이 줄어들 수 있다.

(3), (4), (5)는 대응되는 문장 구조를 가진 직역이 실제 의미와 차이를 보이는 것인데, (3)에서는 피번역어의 직역과 실제 의미가 전혀 다른 경우이고, (4)에서 직역은 아무런 뜻도 전달하지 못하게 되며, (5)에서는 영어와 불어에서 공간에 대한 관점의 차이로 인해 *on*의 직역인 *sur*가 문장 구조상 불가능한 경우다.

- (6) (a) Capable married mothers should have career opportunities.
 (b) capable: capable; compétent;
 to marry: se marier; épouser; faire un mariage; convoler; s'allier; s'apparenter etc.
 mother: mère
 career: (course précipitée); carrière
 opportunity: occasion.
 (c) Des mères mariées capables devraient avoir des occasions de carrière.
 (d) -Il faut que les femmes qui ont des enfants puissent, elles aussi, exercer un métier.

5) VINAY, (1982:729~757)에서 예 (1)-(5) 인용.

- Les mères de famille doivent avoir la possibilité d'exercer un métier.
 -Il faut donner aux femmes ayant des enfants la possibilité de travailler.

예문(6-a)는 영국의 Sunday Times 지가 독자들의 취향을 파악하고 그들의 미래 요구에 부합하기 위해 조사한 설문지 문항이다.⁶⁾ 이를 불어로 옮기기 위해 이중어 사전을 찾아보면 (6-b)와 같이 등가치의 잠재적 의미를 담고 있는 어휘 항목이 제시된다. 이를 이용하여 불어로 옮겨보면 (6-c)와 같이 언어 기호만이 전환된 결과가 나타날 수 있다. (6-c)는 문법적으로 흠잡을 데 없으나 구성 어휘들이 서로 연결되어서는 잘 쓰이지 않기 때문에 매우 어색한 불어가 된다. (6-c)에서는 *mères mariées*가 자연스럽게 못한 표현일 뿐 아니라, *mères capables*은 불어에서 다른 능력보다는 “어머니로서 아내로서 살림을 잘하는 가정 주부”를 뜻하기 때문에, (6-a)가 전달하고자 하는 실제 의미인 “아이가 있는 기혼 여성이 능력이 있을 경우라면 사회적 활동을 해야 하는지”에 대한 의견 조사와는 다른 의미를 나타내게 된다. (6-a)에서 전달하고자 하는 바는 (6-d)처럼 재표현되어야 할 것이다.

문체(style)와 같은 어려움을 제기하는 문학 text나, 다른 전문 지식을 요하는 text 문제를 제쳐 놓고도, 일상 생활에서 자주 접하는 text를 번역하는 것도 앞서 그림 (1)에서 나타난 바와 같이, 피번역어와 번역어에 대한 정확한 언어적 지식 외에, 그 언어 문화, 번역되는 text의 주제, 외부세계, 상황등에 대한 지식이 모두 요구되며 이것이 피번역 text의 실제 의미를 찾는 데 모두 참여한다. 그림 (3)의 (D)방식으로 제시되는 “이상적인” 번역 방법을 익히기 위해 D. Seleskovitch는 다음 네 단계의 훈련 방법을 제시한다.

- (a) text를 몇 부분으로 분할하여 한 부분의 실제 의미(sens)를 파악
- (b) (a)의 실제 의미를 번역어(대부분 전문 번역사의 모국어)로 표현
- (c) text를 다시 보고 번역된 (b)의 세부 사항을 확인
- (d) 피번역 text와 번역된 text를 비교

물론 도식적이거나 수리적인 번역 모델 제시를 기대하지는 못하지만, Seleskovitch의 이상적인 번역 방식은 성공도의 많은 부분을 인간이 논리와 사고라는 인지 과학의 대상에 할당한다.

제 3 절 기계 번역의 종류

컴퓨터가 개발된 초기부터 컴퓨터의 부호(symbol) 처리 능력을 이용하여 인간의 언어, 즉 자연언어를 다른 언어로 번역하려는 시도가 꾸준히 이루어졌다. 그러나 인간의 언어 능력은 고도의 지적 능력을 바탕으로 이루어지는 만큼, 초기의 기계 번역은 지적 문제 해결에 대한 경험의 부족과 컴퓨터 분야의 전반적 기술 수준이 낮았으므로 실패했다. 그러나 70년대를 지나면서 컴퓨터 기술의 급격한 발달과 지적 문제에 대한 해결 경험의 축적은 기계 번역에 새로운 가능성을 보여주기 시작했다. 또한 국제 교류의 증대에 따라서 번역의 필요성은 증대되

6) SELESKOVITCH & LEDERER(1986:104~115)에서 예 (6) 인용.

였으나 번역을 감당할 인력이 부족함으로써 새로운 돌파구가 요구되었다.

이에 따라서 기계 번역에 대한 연구가 급속히 증대되었다. 특히 영·불 번역을 목적으로 몬트리올대학에서 개발된 TAUM의 성공은 기계번역 연구에 큰 영향을 주었다. TAUM은 90~95%의 번역 성공율을 보이고 있으며, 번역 오류도 대부분 입력 잘못된 것으로 알려져 있다. 이 시스템은 1977년부터 실용화되어 캐나다 기상청에서 사용되고 있다.

우리는 여기서 TAUM이 성공할 수 있었던 이유를 살펴볼 필요가 있다. 캐나다는 영어를 사용하는 지역과 불어를 사용하는 지역으로 나뉘어지며, 따라서 캐나다 기상청은 영어로 작성된 일기 예보들 불어로 번역할 필요성이 있었다. 그런데 캐나다 기상청의 일기 예보관들은 비교적 일정한 문장 구조와 단어를 이용하여 영어문장을 작성하는 것이 일반화되어 있었다. 이와 같은 상황에서 영어문장을 불어문장으로 번역하는 작업은 매우 따분하고 비창조적이며 단순한 작업이었다. 따라서 번역사는 6개월을 넘기지 못하고 이직하는 것이 일반적이었다.

위에서 제시된 캐나다 기상청의 상황은 인간에게 있어서는 최악이지만 기계의 입장에서 보면 상황은 완전히 바뀌게 된다. 첫째, 기상예보는 처리 영역이 매우 제한되어 있으므로, 처리 영역 지식을 이용하기 쉽다. 둘째, 단어와 문장구조가 매우 제한되어 있다. 셋째, 번역의 오류가 크게 문제되지 않는다. 네째, 인간에게는 단조로와 비교적 부적합한 작업이다. 이와 같은 상황에 따라 TAUM은 크게 성공할 수 있었다.

그런데 여기서 재미있는 것은 TAUM이 문장의 이해를 전제로 번역을 수행하는 것이 아니라라는 점이다. 즉 TAUM은 영어 통사분석 결과를 불어 통사구조로 변환하는 구조 변환 방식(transfer approach)을 이용했다는 것이다. 이 결과는 고도의 지적 처리능력이 없이도 분야에 따라서는 효율적인 번역시스템을 구성할 수 있다는 것을 보여준다.

번역에 기계가 이용되는 방법은 두 가지가 있다. 첫번째는 인간의 번역을 기계가 도와주는 방법(Machine Aided Human Translation)이 있다. 이와 같은 시스템으로는 철자교정기(spell checker), 전자사전(electronic dictionary), 문체교정시스템(style analyzer)등을 들 수 있다.

두번째는 기계에 의한 번역이다. 이 기계에 의한 번역도 인간의 도움에 의한 기계번역(Human Aided Machine Translation)과 완전 자동번역(Fully Automatic Machine Translation)으로 나눌 수 있다. 인간이 기계의 번역과정에 개입하는 방법은 전처리(pre-editing), 후처리(post-editing) 및 번역과정에 직접 개입하는 방법이 있다. 전처리는 피번역어를 모국어로 하는 사람이 번역될 문장의 중의성이나 문장구조가 어려운 것 등을 제거해주는 작업이다. 이에 대해 후처리는 번역어를 모국어로 하는 사람이 번역된 문장을 다시 다듬는 작업이다. 번역과정에 인간이 직접 개입하는 방법은 기계가 문장의 중의성이나 미등록어를 인간에게 문의하면서 번역을 진행하는 방법이다. 완전 자동번역은 기계번역의 궁극적 목적이지만 현재의 기술로는 처리 영역이 매우 제한되고, 문장구조가 일정한 문장만이 가능하다.

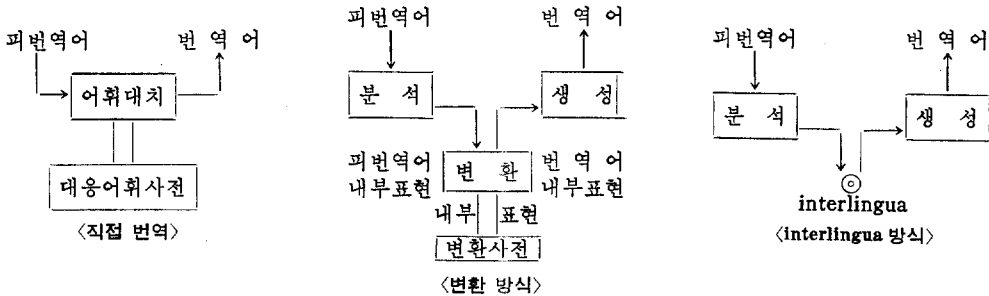
그러면 기계번역방법을 살펴보자. 기계번역은 번역기법에 따라 크게 직접 번역(direct translation), 변환 방식(transfer approach) 그리고 interlingua 세 가지로 나눌 수 있다.

직접 번역은 어휘 분석 정도의 간단한 처리를 한 후에 피번역어의 어휘를 번역어의 어휘로 대치시키는 방법이다. 이 방법은 시스템이 간단하고 처리속도가 빠른 장점이 있으나, 근친 어족간의 번역만 가능하며 번역의 질이 매우 떨어진다. 따라서 기계 번역의 연구가 시작된 초기에는 부분적으로 시도되었으나 현재는 거의 시도되지 않고 있다.

변환 방식은 피번역어를 분석하여 피번역어 내부구조를 만들고, 이 내부구조를 번역어의 내부구조로 변환(transfer)한다. 이 번역어의 내부구조로부터 번역어를 생성함으로써 번역을 완료한다. 이 변환 방식은 피번역어와 번역어의 내부구조를 상이하게 설정하는 면에서 interlingua 방식과 차이가 있다.

interlingua 방식은 피번역어를 분석하여 interlingua 로 만들고, 이 interlingua 로부터 번역어를 생성하는 방법을 이용한다. 따라서 별도의 변환 루틴이 없다.

그러므로 interlingua 방식은 개별언어와 무관한 interlingua 가 필요하다. 일반적으로 interlingua 는 문장이 가진 개념(concept) 혹은 의미(meaning)를 표현한다.



[그림 4 기계번역의 방법]

D. Seleskovitch 에 의한 번역의 분류 방식에 의하면, 직접번역은 그림 (3)의 (A)에 해당하며, 변환 방식은 (B)와 (C)에 해당한다. 그리고 직접 번역과 변환 방식은 기호전환 방식의 일종으로 파악될 수 있다. interlingua 방식은 (D)에 해당하지만, 기계번역에서는 문장이 가진 실제적 의미를 파악한다는 것이 기존의 기술로는 불가능하므로 인간과 비교될 수는 없다.

제 3 장 불-한 기계 번역에 대한 접근

의미에 기반한 기계번역은 상식(commom sense)과 처리 영역에 대한 지식(domain specific knowledge)을 요구한다. 그러나 D. Seleskovitch 가 제시한 모형에서 인간의 번역에 요구되는, 그림 (3)의 (D) 유형의 번역이 가능한 기계가 가까운 시일내에 개발될 것으로 보여지지는 않는다. 그 이유는 기존 인공지능 기술로는 상황을 비롯한 미묘한 언어적 문제를 처리할 기술 수준에 도달하지 못하고 있으며, 또한 CETA 의 경험은 interlingua 방식이 언어 자체가

가진 구조를 파괴할 수 있다는 것을 보여주었기 때문이다.

따라서 처리영역을 제한함으로써, 상황 혹은 문맥으로 해결해야 할 문제를 처리 영역에 대한 지식으로 대체하여 해결할 수 밖에 없다. 또한 피번역어와 번역어 간의 비교분석을 통해 문장이 가진 특성을 이중어 사전을 바탕으로 처리함으로써, 문장이 가진 개념구조의 분석이 어려움을 회피해야 할 것이다.

이와 같은 측면에서 불·한 번역을 실용화하기 위해서는 지식베이스에 근간하여, 의미의 중요성을 강조한 변환방식을 이용할 수 밖에 없다고 생각된다. 그리고 처리 영역도 문체가 일정하며, 용어의 중의성이 적고, 문장구조가 단순한 형태를 취하는 과학 및 기술 분야의 논문과 사용자 지침서(manual)가 대상이 될 수 밖에 없다.

우리가 구현하려는 불어-한국어 번역 시스템은 변환 방식을 기초로 하되 가능한 많은 의미 정보와 불어-한국어의 비교분석을 통한 의미가 추가된 이중어 사전을 이용함으로써, 기계 번역의 가장 큰 어려움인 중의성(ambiguity)을 해결하고자 한다. 이러한 원칙에 입각하여, 본 불어 분석 시스템도 하위 범주화(subcategorization)와 논리적 표현을 이용하고, 술어와 일치할 수 없는 격을 분리 표현하여 분석과정에서 제거하는 방법을 사용한다. 또 의미 자질과 통사 자질의 처리를 절차적(procedure) 방법인 복사를 대신하여, 비절차적(non-proceeding) 방법으로 수행할 수 있도록 단일화에 의한 문장 분석을 시도한다. 본 불어 분석 시스템은 단일화 기반 문법을 위해 Prolog 언어를 확장한 GULP(Graphic Unification Logic Programming)라는 언어로 구현되었다. 단일화 기반 문법은 단일화(unification)라는 연산을 통해 자질값(feature value)에 제약을 가함으로써 자연언어 문장 분석을 시도하는 문법 이론을 지칭한다.

제 1 절 형태소 분석

본 시스템은 Prolog 언어를 이용하여 형태소 분석을 행하였으며, 형태소 분석 결과로부터 자질구조를 생성하여 형태소 분석기에 결과를 넘겨주게 했다.

사전은 각 어휘의 어간과 이에 수반되는 정보를 가지고 있다. 동사에 관한 표현을 중심으로 사전 구성을 살펴보자.

dict(어간) → 통사 정보(자질구조, 하위범주화)
의미 정보(mood, tense, 술어)
불규칙 종류

어간은 동사의 변화과정에서 불변하는 부분을 지칭한다. 그리고 각 어간은 그에 대응되는 통사 정보와 의미 정보를 가지며, 또한 불규칙 유형도 보여준다.⁷⁾

규칙 동사의 경우 어간을 수록하는 사전에 다음과 같은 자질 정보가 수록된다. 1군 변화

7) 불규칙 동사와 대명동사 등의 자질 정보(feature information) 수록에 대한 자세한 예는 윤애선(1991)을 참조하시오.

에 속하는 규칙 동사 “parler(말하다)”를 살펴보자.

```
(7) dict-verb(parl, Feature): -Feature=sem: (lex:parler.. arg1: hum: pos.. arg2: hum: pos),
(
  Feature= syn: subcat: [nont, dat, del, [inst, del, agentive, coi,
                        dat]]
;
  Feature=syn: subcat: [nont, [inf, de], dat, [inst, del, agentive, coi,
                        dat]]
),
Feature=syn:(cat: verb.. type:1).
```

어미 활용 중 불변하는 어간 *parl*의 통사적 정보(syntactic informations)는 다음과 같다.

첫째, 문법 범주(grammatical category)가 동사(verb)임이 표시된다.

둘째, 어미 활용의 유형(type)의 종류가 표기되어 그에 따라 형태소 분석기에서 해당되는 부분과 연결된다.

셋째, 하위범주(subcategory)로 []안의 첫번째 항에는 동사의 종류가, 다음 두 항에는 이 동사가 취하는 보어의 격이 표시된다. *parler*의 경우, 직접 목적보어(direct object 또는 accusative)를 취하지 않는 비타동사(nont)이며,⁸⁾ 보어의 종류에 따라 두 가지의 다른 하위범주 항을 설정한다.

하나는 *Pierre parle à Marie de Jean*(Pierre는 Jean에 대해 Marie에게 말한다)에서처럼 말을 건네는 상대(dative)와 언급되는 대상(delocutive)을 취하는 경우이고, 다른 하나는 *Pierre parle à Marie de venir*(Pierre는 Marie에게 오라고 말한다)에서와 같이 delocutive 대신 《전치사 de+infinitive》를 취하는 경우다. infinitive를 보어로 취할 때 붙어에서는 전치사 *de* 또는 *à*를 선행하거나 전치사를 수반하지 않은 경우가 있다. 이를 명시하기 위하여 [inf, de], [inf, à], [inf, o]를 분리하여 사용하였다.

보어가 절인 경우, 보어절 내의 범(mood)이 제시된다. *vouloir*인 경우에 sbar(sub)로, *dire*의 경우에 sbar(ind)가 보어항에 기재된다.

subcat의 네번째 항에는 해당되는 동사와 함께 나타날 수 없는 상황보어를 명시하여 상황보어의 의미 해석에 제한을 가함으로써 잘못된 문장분석과 수행시간을 매우 효과적으로 줄일 수 있다. *parler*의 경우 도구(instrument), 언급되는 대상(delocutive), 동작주보어(agentive), 여격(dative)과 다른 간접 목적 보어격(coi)⁹⁾이 상황 보어로 나타날 수 없다.

8) 동사의 종류를 속사(attributive)의 유무와 대격(accusative)의 유무에 따라 나누었다. 대격을 취하는 동사를 타동사(transitive), 대격 대신 속사를 취하는 동사(copula), 이외의 동사를 비타동사(non-transitive)라고 칭하고, 이 모든 동사는 전치사구(prepositional phrase)를 취할 수 있되, 이 전치사구가 동사의 보어(complement)인지 상황보어(circumstance)인지는 사전에 수록된 subcat항과 통사 규칙부의 보어 통제규칙에 의해 나타낼 수 있도록 했다.

9) 동사가 취하는 보어 전치사구에 격을 분류하는 경우 도구격(instrument), 근원격(source), 장소격(locative)등 처럼 의미를 부여하는 것이 항상 가능하지는 않다. 이 경우 간접 목적격(Complément d'Objet Indirect)이라고 칭하였다.

네째, 불어 동사는 복합형 시제일 때 시간 조동사로 *avoir* 또는 *être* 를 취한다. 이를 구분하기 위하여 수가 적은 *être* 를 취하는 동사에 temp-type: être 로 표시하고, 경제성을 위해 *avoir* 를 취하는 동사에는 이 항목을 수록하지 않는다. *parler* 는 후자에 해당한다.

parl 의 의미정보는 다음과 같다.

첫째, 이 어간의 어휘(lexicon)가 *parler* 임이 표시되는데, 이 항목은 규칙동사보다 불규칙동사에서 다양한 어간이 한 어휘에 해당됨을 나타내는 데 중요한 항목이다.

논항(arguments)의 의미 제한(semantic constraints)이 표시된다. 주어인 argument 1 은 사람(human: positive)이고 dative 인 argument 2 도 같은 의미 제한을 가짐이 표시된다.

이상과 같은 *parl* 의 자질 정보는 형태소 분석기에 넘겨지고 해당되는 어미 활용 유형 항에서 어미에 관한 다음과 같은 정보를 취하게 된다.

```
(8) inflexion-check-verb0 (Inf, Feature):-
    Feature=syn:type:X,
    X==1,!,
    name(Suf, Inf),
    (
    Suf== 'er',!,
    Feature=syn:form:root
    ;
    Suf== 'e',!,
    (
    Feature=sem:(mood:ind...tense:pres),
    Feature=syn:(per:first...num:sg)
    ;
    Feature=sem:(mood:ind...tense:pres),
    Feature=syn:(per:third...num:sg)
    ;
    Feature=sem:(mood:sub...tense:pres),
    Feature=syn:(per:first...num:sg)
    ;
    Feature=sem:(mood:sub...tense:pres),
    Feature=syn:(per:third.. num:sg)
    )
    ;
    ⋮
    compare-word (Inf, "é", R),
    inflexion-check-normal (R, Feature),
    Feature=syn:form:past part
    ).
```

규칙동사는 형태소 분석기에서 활용형의 경우 단순형시제의¹⁰⁾ 각 어미가 나타내는 법(mood),

10) 불어의 복합형 시제는 <시간 조동사+과거분사>로 구성되므로 규칙을 통해 분석될 수 있다. 따라서 사전과 형태소 분석에는 단순형 시제만이 수록된다.

시제 (tense), 인칭 (person), 수 (number)가 표현된다. 위의 예는 1군동사 활용형 어미 *-e*가 4가지 정보를 전달해 줌을 보여준다. (1) 직설법 현재이고 주어가 일인칭 단수임과 (indicative, present, first, singular), (2) 직설법 현재 삼인칭 단수, (3) 접속법 현재 일인칭 단수, (4) 접속법 현재 삼인칭 단수이다.

비활용형의 경우인 동사원형 (root)과 과거분사 (past participle)는 형태 (form)로 명시된다. 과거분사 항목에는 추가 정보가 필요하다. 불어에서 과거 분사는 주어나 직접 목적 보어의 성, 수에 일치하는 경우가 있으므로, compare-word (Inf, "é", R), inflexion-check-normal (R, Feature)는 기본 과거분사 형태 *parlé*를 남성 단수형으로 하여, 여성 단수형, 남성 복수형, 여성 복수형을 분석할 수 있다.

제 2 절 단일화에 의한 문장 분석

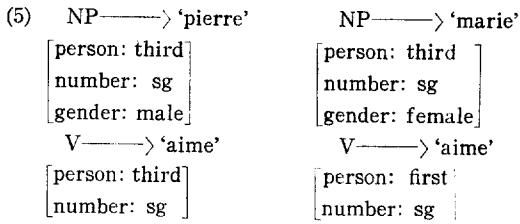
가. 단일화 기법

초기의 Chomsky 문법은 범주 자질 (category feature)을 제외한 모든 자질을 무시했다. 그러나 최근에 와서는 파스나무 (parse tree) 구조의 각 노드 (node)에 자질다발 (feature bundles)을 부가함으로써 자질 간의 다양한 조응문제를 해결하고 있다. Chomsky의 최근 이론은 문법적 조응을 한 노드에서 다른 노드로 자질을 복사하는 변형에 의해 해결한다. 그러나 자질의 복사는 절차적 수행 (procedural processing)을 요구하며, 이 절차적 수행은 전산처리에 부적합하다.¹¹⁾

단일화 기반 문법은 이 문제를 비절차적으로 해결한다.

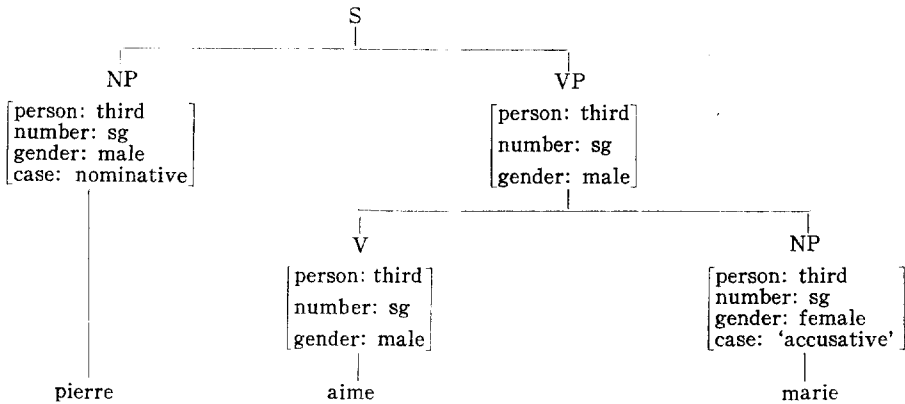
- (1) S → NP VP
- | | |
|--|------------------------------------|
| -person:P number:N gender:G -case: 'nominative' | [person:P number:N gender:G] |
|--|------------------------------------|
- (2) VP → V NP
- | | | |
|------------------------------------|------------------------------------|----------------------|
| [person:P number:N gender:G] | [person:P number:N gender:G] | [case: 'accusative'] |
|------------------------------------|------------------------------------|----------------------|
- (3) VP → V
- | | |
|------------------------------------|------------------------------------|
| [person:P number:N gender:G] | [person:P number:N gender:G] |
|------------------------------------|------------------------------------|
- (4) VP → VP PP
- | | |
|------------------------------------|------------------------------------|
| [person:P number:N gender:G] | [person:P number:N gender:G] |
|------------------------------------|------------------------------------|

11) COVINGTON (1989:8).



(그림 5 예제 문법)

〈그림 5〉는 본 시스템의 구조를 보이기 위해 설정한 간단한 예제 불어 문법을 보여준다. 이 문법을 이용하여 *Pierre aime Marie* 를 분석하면 〈그림 6〉와 같은 결과가 나온다. (5)는 위의 예문을 형태소 분석한 결과를 개략적으로 보여준다. 설명의 편의를 위해 하위 범주화와 의미 정보에 대한 내용을 제외했다.



(그림 6)

위의 결과에서 'S'에 직접 지배되는 'NP'의 'case'는 'nominative'이다. 이 'nominative'는 'S'가 주어의 격을 'nominative'로 요구함을 나타낸다. 그런데, 사전에서 *pierre*가 가진 정보에는 'case'에 관한 정보가 없다. 따라서, 이 정보는 〈그림 5〉의 규칙 (1)에 의해 단일화 과정을 거쳐 구체화(instantiate)된 것이다. 반대로 'VP'의 자질값은 'V'로부터 부여받아 구체화되었다. 이 결과는 상호 반대되는 방향에서 자질구조가 구체화되는 것을 보여주며, 이런 의미에서 단일화는 비절차적(non-procedural) 수행과정이다.

*aime*는 1인칭인 경우와 3인칭인 두 경우 모두에 사용될 수 있다. 그런데, 1인칭인 경우에는 'S'의 직접 지배를 받는 'NP'의 인칭이 3인칭이므로 단일화 연산이 실패하게 된다. 따라서, *Pierre aime Marie*는 한 가지 파스 나무구조(parse tree)를 가지게 된다.

나. 의미의 표현

문장의 의미는 술어와 술어의 논항 및 시제, 양상 등을 포함하는 확정된 논리적 표현을 이

용하며, 파서는 이 의미 표현만을 출력한다. 본 분석기의 구현 목적은 불·한 기계번역이므로, 불어의 통사구조는 출력에서 제외했다. 그 이유는 통사구조보다는 의미 표현이 더 중립적인 표현이기 때문이다.

```
sem: lex: aller
      tense: passé
      mood: ind
      arg1: lex: pierre
            hum: pos
      arg2: lex: chez
            obj: lex: paul
            hum: pos
```

위의 그림은 *Pierre est allé chez Paul* 을 본 시스템으로 분석한 결과를 보여준다. *aller* 동사는 두 개의 논항을 요구하며, 두번째 논항은 전치사구가 된다. 그리고 그 전치사구 내의 목적어는 *Paul* 이다.

단일화를 이용한 불어 분석기의 효율성을 높이기 위해서 앞서 언급한 하위 범주화(subcategorisation)나 의미 표현 이외의 몇가지 기법이 사용된다. 하향식 분석방법(top-down parsing)에서 흔히 일어나는 좌반복(left-recursion)을 방지하기 위해 tail recursion 이 사용되고 동시에 예측을 도입하고 cut(!)을 이용한다.¹²⁾

제 3 절 불어-한국어 변환

불어와 한국어는 어족이 다른 언어이므로 직접 번역이 불가능하다. 한편, interlingua 방식은 방법적인 면의 우수함에도 불구하고, 기존 기술로는 실용화가 불가능하다. 따라서 불·한 기계 번역을 위해서는 변환방식이 가장 현실적인 방법이다.

그런데 변환방식은 피번역어의 분석 정도에 따라 다양한 형태를 취한다. 단순히 통사구조만을 분석할 수도 있으며, 어느 정도 의미까지를 분석할 수도 있다. 단순한 통사구조의 분석만으로는 불·한 번역의 질을 높일 수 없으므로, 자질 구조를 근간으로 하는 의미 분석을 바탕으로 번역함으로써 기존 전산 기술과 번역의 질 간에 타협점을 찾을 수 있다.

또한 번역을 위한 이중어 사전의 의미 표현을 불어와 한국어의 단어를 쌍으로 하면서, 여기에 추가의 제약이 가해진 구조를 취함으로써 단어 선택과 변환을 쉽게 하고, 두 언어 간의 비교 결과가 번역에 직접 이용될 수 있는 구조를 제시하겠다.

불어 “mémoire”는 “논문”과 “기억력”이라는 두 가지의 다른 번역이 가능하다. 그런데 “논문”은 “남성”의 자질을 가지며, “기억력”은 “이성”의 자질을 가진다. 또한 “기억력”은 복수

12) 본 분석기에 적용된 이러한 기법은 윤애선(1991)을 참조하시오.

로는 사용될 수 없다. 이를 바탕으로 번역을 위한 사전을 구성하면 다음과 같다.

| | | | | | | | | | | | | | | | | | | | | | | |
|----------------|--|---------------|---|-----------|---|----------------|---|-----------|---|------------|---|-------------|-----|---|----------------|---|----------|---|--------------|---|-----------|---|
| (1) | <table style="border: none;"> <tr><td style="border: none;">-lex: mémoire</td><td style="border: none;">-</td></tr> <tr><td style="border: none;">hlex: 기억력</td><td style="border: none;">-</td></tr> <tr><td style="border: none;">gender: female</td><td style="border: none;">-</td></tr> <tr><td style="border: none;">number: X</td><td style="border: none;">-</td></tr> <tr><td style="border: none;">-cat: noun</td><td style="border: none;">-</td></tr> </table> | -lex: mémoire | - | hlex: 기억력 | - | gender: female | - | number: X | - | -cat: noun | - | <-X/=plural | (2) | <table style="border: none;"> <tr><td style="border: none;">-lex: mémoire-</td><td style="border: none;">-</td></tr> <tr><td style="border: none;">hlex: 논문</td><td style="border: none;">-</td></tr> <tr><td style="border: none;">gender: male</td><td style="border: none;">-</td></tr> <tr><td style="border: none;">cat: noun</td><td style="border: none;">-</td></tr> </table> | -lex: mémoire- | - | hlex: 논문 | - | gender: male | - | cat: noun | - |
| -lex: mémoire | - | | | | | | | | | | | | | | | | | | | | | |
| hlex: 기억력 | - | | | | | | | | | | | | | | | | | | | | | |
| gender: female | - | | | | | | | | | | | | | | | | | | | | | |
| number: X | - | | | | | | | | | | | | | | | | | | | | | |
| -cat: noun | - | | | | | | | | | | | | | | | | | | | | | |
| -lex: mémoire- | - | | | | | | | | | | | | | | | | | | | | | |
| hlex: 논문 | - | | | | | | | | | | | | | | | | | | | | | |
| gender: male | - | | | | | | | | | | | | | | | | | | | | | |
| cat: noun | - | | | | | | | | | | | | | | | | | | | | | |

본 연구에서는 모든 불어 자질에 대응되는 한국어 자질을, 자질 명칭에 “h”를 추가함으로써 표현했다. 즉 (1)의 “lex” 자질값 “mémoire”에 대응되는 한국어 자질값은 “hlex”의 “기억력”이 된다. 그리고 “mémoire”가 “기억력”으로 번역되기 위해서는 “mémoire”가 “female”이고 “noun”이지만, 수는 “plural”이 될 수 없음을 보여준다. “mémoire”가 “논문”으로 번역되기 위한 제약이 (2)에 있다. 위의 사전을 효과적으로 이용하면, 불어-한국어 간의 양방향 기계 번역을 위한 사전의 구성이 가능하다.

- (a) Pierre a écrit un mémoire.
- (b) Il y a des trous dans sa mémoire.
- (c) Pierre a écrit des mémoires.
- (d) Le mémoire de Pierre était bon.
- (e) La mémoire de Pierre était bonne.

위의 예문에서 (a)의 ‘mémoire’의 관사가 남성형이므로 “논문”으로 번역되어야 하며, (b)는 여성형이므로 “기억력”으로 번역되어야 한다. (c)에서는 “성”에 대한 정보는 없으나, “수”가 “복수”이므로 “논문”으로만 번역된다. (d)에서는 “bon”이 남성형이므로, 단일화 과정에서 “mémoire”가 남성형이 되며, 따라서 “논문”이 된다. 그러나 (e)에서는 여성형이므로 “기억력”이 된다. 이 결과는 단일화 연산과 이를 기반으로 한 이중어 사전이 기계번역 과정에서 어휘적 중의성을 해결하는 방법을 보여준다.

위의 사전을 직접 GULP 로 구현하면, 불어-한국어 양방향 번역이 가능하다. 그러나 불어-한국어 번역을 목적으로 하는 본 시스템에서는 처리 속도와 프로그램의 편의를 위해 다음과 같이 변화시켜 표현했다.

```

bi-dict(Feature, Hmor):-
    Feature=lex: memoire.. number: X.. cat: noun
    .. gender:G,
    (G= female,
    X/=plural,
    Hmor="기억력"
    ;
    G= male,
    Hmor="논문"
    ).
  
```

다음은 ‘belle-mère’에 대한 이중어 사전의 내용이다.

- | | |
|--|--|
| (3) $\left\{ \begin{array}{l} \text{lex: belle-mère} \\ \text{cat: noun} \\ \text{hlex: 장모} \end{array} \right.$ | (4) $\left\{ \begin{array}{l} \text{lex: belle-mère} \\ \text{cat: noun} \\ \text{hlex: 시어머니} \end{array} \right.$ |
|--|--|

(3)과 (4)는 “belle-mère”가 “장모”와 “시어머니” 둘 다로 번역될 수 있음을 보여준다. 그런데 한국어에는 이 둘 다를 동시에 표현하는 단어가 없다. 따라서 본 연구에서는 둘 다를 선택하도록 했다.

(f) Je suis allée chez ma belle-mère

(f)의 경우에 “나(Je)”는 “allée”에 의해 여성임이 밝혀지므로 “belle-mère”는 “시어머니”로 번역되어야 한다. 그러나 이와 같은 처리는 대용어(anaphora)의 처리가 이루어져야 가능하다. 그러나 본 연구에서 구현한 시스템은 대용어 문제를 해결하지 못하므로 다음과 같이 번역된 결과가 나온다.

(f') 나는 나의 $\left[\begin{array}{l} \text{장모} \\ \text{시어머니} \end{array} \right]$ 의 집에 갔다.

이와 같은 문제는 처리가 불가능한 것은 아니지만, 처리의 어려움과 필요한 지식의 양 등을 고려하여 현재 처리하지 않고 있다.

본 연구에서는 통사구조를 선정하지 않고, 자질 구조에 의한 의미표현을 바탕으로 번역을 수행한다. 처리 알고리즘은 다음과 같다.

1. 번역될 자질구조의 중심어(head)를 찾아서, 자질구조를 바탕으로 대응되는 한국어 어휘를 대치한다.
2. 중심어에 의해 하위범주화되는 요소를 찾아서 한국어 어휘를 대치한다.
3. 부가어가 있으면, 부가어에 대응되는 한국어 어휘를 대치한다.
4. 어휘에 기능어를 삽입하고 어순을 고려하여 한국어 문장을 생성한다.

위의 알고리즘이 재귀적(recursive)으로 적용되어, 대응되는 한국어 문장이 출력된다. 이 방법은 단일화 기반의 분석을 바탕으로 한 변화 방식을 적용한다. 이는 기존의 통사 변환 방식과는 차이가 있다.

제 4 장 결 론

본 고에서는 인간의 번역 수행 과정을 검토하여 기존의 기계 번역 시스템을 보완해 좀 더 효율적인 불어-한국어 기계 번역 시스템 구현에 방법을 제시하려 하였다.

인간의 번역에서나 기계 번역에서 모두 “피번역어의 의미를 정확히 해석하고 그 의미를 번역어로 재표현”하는 것이 가장 중요한 목표다. 그러나 기존 기계 번역 시스템이 사용하는 직접 번역 방식과 구조 변환 방식에서는 의미를 처리하기가 어렵다. 따라서 본 고에서는 인간 번역과정을 모델로 삼아 입력된 문장의 의미를 정확히 파악하기 위한 방법을 모색하였다.

불어-한국어의 특성과 현단계의 기계 번역 기술 수준에서 취할 수 있는 가장 현실적인 방법은 자질 구조(feature structure)를 근간으로 하여 의미 분석(semantic analysis)을 가능하게 하는 의미 변환 방식(semantic transfer)이다. 사전(dictionary)에 상세한 의미, 통사 정보를 수록하고 하위범주(subcategory)와 문장 허용 제약을 명기함으로써 문장 분석에서 나타날 수 있는 중의성(ambiguity)을 가능한한 미리 해결하고자 하였다. 이러한 자질 연산이 용이하고 비절차적으로 수행되는 단일화 기반 분법(Unification Based Grammar)을 시스템 구현에 사용하였다. 또한 번역을 위한 이중어 사전의 의미 표현을 불어-한국어 단어를 쌍으로 하면서, 여기에 추가의 제약이 가해진 구조를 취함으로써 단어 선택과 변환을 쉽게 하고, 두 언어 간의 비교 결과가 번역에 직접 이용될 수 있는 구조를 제시하였다.

본 시스템은 시제품으로, 현재 약 200 단어 수준에서 간단한 불어 문장을 한국어로 번역할 수 있다. 또한 불어 문장 분석에서 한국어 문장 생성과 사전이 GULP(Graphic Unification Language Programming)로 프로그램되어 있다. 물론 실용화를 위해서는 시스템 확장 뿐 아니라 어휘 의미 자질의 분석이론 제시등 여러 문제가 계속 연구되어야 한다.

참 고 문 헌

- 강승식, 심광섭, 장병탁, 권혁철, 우치수, 김영택(1987), “KSHALT: 영-한 기계 번역 시스템”, 춘계 인공지능 학술 발표회 논문집, 정보과학회 인공지능 연구회, pp.113~132, 1987.
- 과학기술처(1987), 기계 이해 시스템에 관한 연구, 1987년 6월.
- 과학기술처(1988), 한영 변환기 및 영어 생성기 모형 개발, 특정 연구 개발 사업 '88년도 중간보고서.
- 김재훈, 이진선, 정희성(1988), “한·영 양국간의 위상구조 변환을 위한 transfer 모델”, '88 가을 학술 발표 논문집, 한국 정보 과학회, 제15권, 2호, pp.233~236, 1988.
- 김재훈, 김철호, 최기선(1987), 김길창, “PIVOT 방식의 기계 번역에서 한국어 격구조 설정과 중간 언어로부터 조사 생성”, '87 가을 학술 발표 논문지, 제14권, 2호, 한국 정보 과학회.
- 윤애선(1991), “단일화 기반 문법을 이용한 불어 분석기”, 언어 15집 수록 예정
- 조혁규, 장명길, 권혁철(1989), “KPSG에 기반한 한국어 해석기의 구현”, '89 봄 학술 발표 논문집, pp.229~232.
- 한국 과학 재단, 자연언어 처리의 기초연구('86, '87 중간 보고서).
- COVINGTON, M. A. (1989), *GULP 2.0: An Extension of Prolog for Unification-Based Grammar*, AI-1989-01, University of Georgia.
- GROSS, M. (1968), *Grammaire transformationnelle du français*, Larousse: Paris, 1968.
- ISABELLE, P. & BOURBEAU, L. (1985), “TAUM-AVIATION: Its Technical Features and Some Experimental Results”, colling 85 vol II, no. 1, pp.18~27.
- LE GOFFIC, P. & COMBE-McBRIDE, N. (1975), *Les constructions fondamentales du français*, Hachette/Larousse: Paris, 1975.

- MEL'CUK, I. & alii (1984, 1988), *Dictionnaire explicatif et combinatoire du français contemporain I (1984), II (1988)*, Les Presses de l'Université de Montréal: Montréal.
- MOUNIN, G. (1976), *Les problèmes théoriques de la traduction*, Paris: Gallimard, 1976.
- POTTIER, B. (1987), *Théorie et analyse en linguistique*, Paris: Hachette, 1987.
- SELESKOVITCH, D. & LEDERER, M. (1986), *Intépréter pour traduire*, Didier Erudition: Paris, 1986.
- SELLS, P. (1985), *Lectures on Contemporary Syntactic Theories*, CSLI, 1985.
- SLOCUM, J. (1985), "Machine Translation: its History, Current Status, and Future Prospects", *COMPUTATIONAL LINGUISTICS*, Jan-Mar., 1985, pp. 1~17.
- SEO J.Y. & SIMMON, F. (1989), "Syntactic Graphs: A Representation for the Union of All Ambiguous Parse Trees" in *Computational Linguistics*, vol. 15, no. 2, 1989.
- VAUQUOIS, B. & BOITET, C. (1986), "Automated Translation at Grenoble University", *Colling* 86, vol. II, no. 1, pp. 28~39, 1986.
- VINAY, J.P. (1982), "Traduction humaine", in *LE LANGAGE* sous dir. d'A. MARTINET, Pléiade: bruce (Belgique), 1982, pp. 729~757.

《Résumé》

Traduction humaine et traduction automatique: du français en coréen

Le présent travail vise à donner quelques suggestions pour développer un système plus efficace de la traduction automatique (TA) français-coréen. Les méthodes de la traduction humaine (TH) sont examinées afin de résoudre, au moins, une part des problèmes sémantiques que des systèmes déjà établies de la TA n'ont pas osé surmonter.

Notre système adopte une méthode de "Transfer Sémantique" qui nous permet une analyse sémantique des lexiques. Le dictionnaire comprend des informations syntaxico-sémantiques d'un lexique, sa soucatégorie, et ses contraintes phaséologiques. Ce dernier élément peut réduire des ambiguïtés possibles de l'interprétation. La "Grammaire à la Base de l'Unification (Unification Based Grammar)" facilite la composition de ces traits lexicaux, en particulier avec son procédé "Non-Procédural". Le dictionnaire bilingue présente une paire des mots français-coréen ayant un sens approximatif, où sont marquées également d'autres contraintes syntaxico-sémantiques pour la traduction.

Notre système est un proto-type décrit en GULP (Graphic Unification Language Programming) qui peut traduire des phrases simples du français en coréen et dont le dictionnaire contient à peu près 200 mots. L'étude sur l'analyse sémantique doit être, donc, poursuivie afin de l'application pratique de ce système.