

일-한 기계 번역에 있어서 한국어 술부의 생성과 평가

김 정인[†], 문 경희[†], 이 종혁[‡], 이 근배[‡]

포항공과대학교 정보통신연구소[†], 포항공과대학교 전자계산학과[‡]

Generation of Korean Predicates for Japanese-Korean Machine Translation System and its Evaluation

Jung-In Kim[†], Kyong-Hi Moon[†], Jong-Hyeok Lee[‡], Geunbae Lee[‡]

POSTECH Information Research Laboratories[†], Dept. of Computer Science & Engineering, POSTECH[‡]

요약

일-한 기계 번역을 연구하는 많은 연구자들은 양국어의 문절-어절 단위의 어순 일치와 같은 구조적 유사성을 최대한 이용하기 위해 직접 번역 방식을 채택하고 있다. 그러나, 일본어와 한국어 술부간에는 대응하는 품사의 불일치 및 국부적인 어순의 불일치 등이 어려운 문제로 남아 있다. 본 논문에서는 이들 술부 표현의 불일치를 해결하기 위해 이미 제안하였던 “양상 테이블을 기반으로 한 한국어 술부의 생성 방법”에 대해 좀더 체계적인 평가를 하고자 한다. 이 방법은 술부만을 대상으로 하는 추상적이고 의미 기호적인 양상 자질(modality feature)을 테이블화(양상 테이블)하여, 양국어의 술부 표현의 피벗(pivot)으로 이용함으로써 술부 양상 표현의 효과적인 번역을 가능하게 하였다. 일본어 499 문을 대상으로 실제 술부의 번역처리를 시행해 본 결과, 약 97.7%가 자연스럽게 번역됨을 확인하였다. 특히, 술부의 생성 부분은 일본어에 의존하지 않는 양상 테이블을 도입함으로써 일-한뿐만 아니라 다른 언어로부터의 한국어 술부 생성에도 적용시킬 수 있을 것이다.

1. 서론

직접 번역 방식은 보통의 변환 번역 방식에서 행하고 있는 구문 해석이나 의미 해석 부분을 생략하거나 간소화할 수 있어, 유사성을 가진 언어들 사이의 번역에는 자주 채택되어지고 있으며, 현재, 알려진 거의 모든 일-한, 한-일 번역 시스템은 직접 번역 방식을 채택하고 있다. 최근, 구문 해석 분야나 의미 해석 분야의 언어 처리 기술이 발달하고 컴퓨터의 하드웨어 성능이 향상된 시점에서 직접 번역 방식을 사용하는 것은 번역에 필요한 막대한 정보의 손실을 가져오므로, 변환 번역 방식을 사용할 것을 권하고 있으나, 적어도 일-한 기계번역에 있어서는 그렇게 많은 정보가 요구되지 않는다. [2]에 의하면 실제로 아무런 정보 처리 없이 양언어의 단어를 1:1로 대응시킨 경우에도 약 65% 정도의 번역이 이루어졌다고 보고하고 있으며, 약간의 추가 정보를 사용하면 상당한 품질의 역어 생성이 가능할 것이라고 예상하고 있다.

이러한 유사성을 이용한 일-한 직접 번역 시스템이 최근에 와서 상용화되기 시작했다. 최초의 상용 시스템은 1983년, 후지쯔의 ATLAS-I를 기본으로 일본의 후지쯔와 한국의 시스템공학연구소가 공동으로 개발한 ATLAS-J/K이며, 그 후 많은 상용 시스템들이 속속 등장하기 시작했다. 일-한 기계 번역 부분의 대표적인 상용 시스템으로는 일본 고텐사의 “J-SEOUL” 및 히다찌 정보네트워크가 개발한 “HICOM / MT” 등을 들 수 있으며, 국내에서는 최근 유니소프트가 개발한 “오경박사”와 창신컴퓨터의 “한글가나” 등이 있다. 그러나 대다수 이들 시스템은 번역어의 품질이 만족할 만한 상태까지 이르지 못했으며, 형태소 해석, 다의성, 대역어 선정, 품사 판정, 미등록 단어 처리 등의 부분에서 아직 많은 문제점들을 가지고 있다[13].

일-한 직접 번역 방식에는 크게 나누어 (1)동형 이의어 처리, (2)술부의 양상 및 활용 처리가 문제로 남아 있다. 첫 번째 문제를 해결하기 위한 선행

연구로서는 “격형식을 개입시킨 말과 말의 관계를 이용한 한-일 단문 기계 번역 시스템[8]”, 비교적 많은 다의성을 가지는 동사들만을 대상으로 한 “일-한 기계 번역에 있어서 동사의 다역성 처리[9]”, 전후에 접속되는 단어들과의 접속 관계를 연어 패턴으로 표현한 “연어 패턴에 의한 일-한 기계 번역 시스템의 구축과 그 평가[6]” 등이 발표되었으며, 계속하여 개선을 위한 연구가 진행되고 있다. 또한 두 번째 문제의 선행 연구로는 한국어 슬부의 양상 표현 및 활용 형태를 전후 단어와의 의미 접속 관계에 따라 테이블 형식으로 미리 준비한 후 번역 처리를 행하는 “일-한 기계 번역에서 의미 접속 관계를 이용한 한국어의 생성 방법 [10]” 과 부정문을 대상으로, 부정 표현을 긍정 어간으로 변화시켜 처리하는 “일-한 기계 번역에서 부정문의 처리[11]”, 번역 테이블 방식을 더욱 개선한 “일-한 기계 번역에서 확장 번역 테이블을 이용한 한국어의 생성 처리[12]” 등이 있다. 그러나, 이 방법들은 양국어 간 활용어 대응에서 규칙을 찾기가 어렵다고 판단하여, 복잡한 활용 규칙을 이용하지 않고 표층 표현을 1:1로 대응시키는 선언적 처리를 택하고 있다. 이 경우 표층 표현이 1:n으로 대응할 때는 1:1의 관계를 만들기 위한 전후 단어와의 접속 규칙 및 형태가 다른 n 개의 모든 대역어를 사전에 준비해야 한다. 또한, 양상 표현과 활용 형태를 구분하지 않고 한꺼번에 처리함으로써 이형태의 대역어 수가 상대적으로 늘어나는 문제가 있다. 이것은 실용화를 염두에 둘 경우 방대한 사전 용량이 문제가 되며, 사전의 정보 입력 작업에도 막대한 시간과 인력이 소요된다.

이에 비하여 본 논문에서는 양상 표현과 활용 형태를 분리하여 처리하는 것이 특징이다. 양상 처리는 의미 기호적 양상 자질을 모아둔 양상 테이블을 이용하며, 활용 처리는 한국어 슬부의 불규칙 활용 및 음운 변화 등을 규칙화 시켜 한국어 슬부의 생성에 이용한다. 이로 인해 대역어를 사전에 전부 준비해 둘 필요가 없으며, 사전 입력 작업이 간소화 되고, 용량도 적게 차지하게 될 것이다.

2. 양상류 테이블을 이용한 한국어 슬부의 생성 방법

일본어의 슬부와 한국어의 슬부는 그 표현 방법에 상이점이 있다. 양국어의 슬부는 1:1의 대응을 기본으로 하는 직접 번역 방식에서 상당한 문제가 되며, 슬부의 번역이 잘못될 경우 치명적인 번역 실패가 되어버리기 쉽다. 따라서 한국어의 슬부를 정확하고 자연스럽게 표현해주는 방법은 일-한 기계 번역에 있어서 상당히 큰 비중을 차지한다.

2.1 일본어와 한국어의 슬부 표현의 차이점

일본어의 슬부는 용언의 어간 뒤에 조동사나 보조용언, 그리고 격조사를 제외한 조사들을 사용하여 표현하며, 대부분의 조동사, 보조용언 및 조사들은 정해진 순서에 따라 나열되어야 한다. 즉, 일본어의 양상류는 일정한 표현의 순서를 가지고 있다.

일본어 슬부의 구성

- (1)체언 + [조동사*] + [일부 조사]
- (2)용언 + [조동사*] + [{조사(테)} + 보조용언] + [조동사*] + [일부 조사]

그러나, 한국어는 일본어와 달라서 조동사가 존재하지 않으며 어간과 보조용언, 서술격 조사 “이다”, 선어말어미, 어말 어미 등으로 구성한다.

한국어 슬부의 구성

- (1)어간¹ + [선어말 어미*] + 어말 어미* + [보조사]
- (2)보조용언 + [선어말 어미*] + 어말어미* + [보조사]

한국어의 경우 (1)은 단독으로 슬부를 구성하지만, (2)는 (1)에 붙어서만 사용이 가능하다. 또한, (2)는 복수로 붙을 수 있으며, 그 경우 (2)의 선어말 어미는

¹ 어간 : {체언 + 서술격 조사 | 용언}

가장 뒤에 붙은 보조 용언에만 나타날 수 있다. 일본어를 기준으로 양국어의 대응 관계를 살펴보면,

조동사 - 선어말 어미, 어말 어미, 보조 용언
 일부 조사 - 어말 어미, 보조사

로 나타낼 수 있다. 양국어의 술부 표현은 1:1 대응이 어려우며, 일본어 조동사의 반복 표기에 의한 양상 표현 순서도 한국어와 일치하지 않는다. 뿐만 아니라, 한국어의 보조 용언, 선어말 어미, 어말 어미 등은 전 접하는 단어의 속성에 따라 이형태를 띄고 있으며, 여기에 11 종류의 불규칙 음운 변화, 음운 축약 등이 적용된다. 일-한 기계 번역에 있어서 자연스러운 한국어 술부의 표현을 위해서는 이러한 부분들을 세심하게 처리할 필요가 있다. 표 1은 일본어의 조동사를 기준으로 한 양상류와 표현 순서를 나타낸다[1].

표 1 일본어 술부의 양상류

순서	일본어의 양상류	예
1	사역	せる, させる
2	피동/가능/자발/존경	れる, られる
3	정중	ます
4	희망	たい
5	부정, 상태, 속성	ない(ぬ/ん), そうだ, らしい
6	과거/완료/단정 (1)	た(だ)
7	정중	です
8	의지/추정/권유	う, よう, まい
9	전문	そうだ
10	과거/완료/단정 (2)	た(だ)

여기에 비하여 한국어의 양상류는 그 종류가 더욱 다양하다. 또한, 술부 양상류의 표현 순서가 일치하지 않음으로 인한 술부의 어순 조정이 필요하며, 특히 한국어에서는 전후 단어와의 음운 표현을 위한 이형태 처리도 필요하다. 다음의 예문은 일본어와 한국어의 술부의 순서에 차이가 있음을 나타낸다.

예문(1) 行 き ます
 | | |
 가 [∅] | |

가 [∅] 비니다
 음운 축약 처리 → 갑니다

예문(2) 行 き まし た
 | | | |
 가 [∅] 았 습니 다
 음운 축약 처리 → 갔습니다

예문(1)은 어미 “끼”의 번역 생략 외에는 1:1의 대응에 의해 번역이 가능한 경우를 보여주고 있다. 그러나, 예문(2)을 보면, 일본어는 (정중+과거/단정)의 순서인데 비하여, 한국어는 (과거+정중+단정)의 순서이어야 하므로, 번역어의 나열 순서가 바뀌어져 있는 것을 볼 수 있다. 조금 더 복잡한 예문을 보자.

예문(3) 行 き ませ ん
 | | | |
 가 [∅] 지 않 습니 다

예문(4) 行 き ませ ん でし た
 | | | | | |
 가 [∅] [∅] 지 않 았 습니 다

예문(3)과 예문(4)에서는 양국어의 술부 표현이 1:1로 대응되고 있지 않음을 보여주고 있다. 예문(3)은 “ませ:정중”과 “ん:부정”의 순서가 바뀌어 역어가 생성되는 것을 보여준다. 또한 “ん”으로 문이 끝나는 경우는 단정의 의미도 포함되어 있어서, “ん”으로부터 “지 않”과 “다”를 모두 생성시켜야 한다. 예문(4)는 정중 표현이 2번 나타난 경우이나, 한국어에서는 정중 표현을 한 번으로 표현하므로, “ませ”의 번역어는 생성되지 않게 처리해야 하며, 또, “でし”와 “た”의 번역어 순서를 바꾸어 나열하고, “た”로부터 단정의 의미까지 추출하여 문의 종결 부분에 “다”를 생성시켜야 한다. 이와 같이, 일본어와 한국어는 술부 표현 부분에서 양상 표현의 순서 불일치, 1:n 대응, 대역어

생략 등 많은 차이점을 보이고 있으며, 이런 차이점들은 직접 번역 방식을 이용한 일-한 기계 번역에서 고품질의 번역문 생성에 큰 장애물이 되고 있다.

2.2 양상 테이블 (MFOLT : Modality Feature Ordering and Lexicalizing Table) 의 도입

술부 표현의 불일치를 해결하기 위하여, 우리는 술부의 양상 정보를 언어에 의존하지 않는 의미 기호적 양상 자질 (modality feature)로 표현하여 양국어의 술부 표현의 피복으로 이용하는 방법을 제안했다[4]. 이 의미 기호적 양상 자질을 이용하여, 한국어 술부의 상대적인 순서를 결정한 다음, 이형태 처리 및 불규칙 변화에 대한 처리, 음운 축약 처리 등을 행하여 자연스러운 한국어 술부를 생성한다. 양상 자질의 구성은 한국어 보조용언들에다가, 번역에 필요한 선어말 어미와 어말 어미, 그리고 특별히 일-한에만 적용되는 기타 자질들도 다수 포함시켜, 보다 부드러운 술부의 표현이 가능하도록 구성했다.

본 논문에서는, 경험에 따라 보조 용언, 선어말 어미, 어말 어미를 통폐합하여, 한국어를 생성시키는데 필요한 어미류를 정규화하였다. 같은 의미를 가지는 다른 형태의 어미류를 하나의 어미로 표현하는 것이 번역 작업의 단순화를 가져오며, 실지로 통폐합된 어미를 이용하면, 현재의 단계에서 거의 모든 한국어 술부를 표현하는데 모자람이 없다고 판단된다. 이것은 경험에 의한 것이며 일-한 번역에서 성과를 최대 높이기 위하여 준비한 것이다. 표 2에 각 자질들을 표시한다.

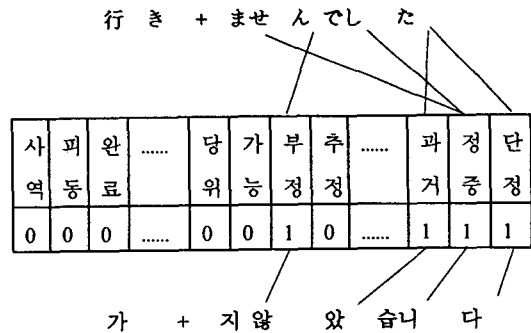
양상 자질들을 테이블 형식으로 기술한 것을 MFOLT (Modality Feature Ordering and Lexicalizing Table) 라고 하며, 테이블 속의 활성화된 자질의 상태에 따라 한국어의 접속표[3]를 참고하며, 술부의 자연스러운 표현을 유도하여 한국어의 술부를 생성한다. MFOLT의 자질들은 통계적인 경험에 의한 것이며, 따라서 자질의 구성도 통계치의 변화와 함께 필요에 따라 변경될 수 있다.

2.3 양상 테이블을 이용한 한국어 술부의 생성
자연스러운 한국어의 술부를 생성하는데는 몇 가지 단계를 필요로 한다. 즉, MFOLT의 각 자질의 활성화, 부정어 처리, 피동 표현 처리, 보조 용언 및 어미의 이형태 처리, 음운 축약 및 불규칙 음운 처리 등이며, 여기서는 각 단계별 특징에 대해서 설명한다.

2.3.1 의미 자질의 활성화

입력된 일본어의 술부에서 의미 자질을 추출하여, MFOLT의 각 양상 자질들을 활성화하는 과정이다. 각 자질들은 자유롭게 활성화시킬 수 있는데 비하여, 어말어미의 자질들은 연결형, 종결형, 그리고 전성형으로 나뉘어지는데, 각 형태 속에서는 1개의 자질만을 활성화할 수 있다. 그림 1은 예문 4 “行きませんでした”의 활성화된 상태를 나타낸다.

그림 1 일본어의 술부“行きませんでした”의 활성화 처리



2.3.2 부정어 처리

일본어와 한국어는 부정 표현의 상이점이 존재한다. [11]에 의하면, 일본어의 부정 표현에는 형용사 혹은 조동사로서 “ない”, 그리고 조동사인 “ん(ぬ)”와 “まい” 등을 사용한다. 그리고 접두사로서 “非,不,未,無”를 붙혀 한 단어만을 부정하는 한어적 부정 표현과 부정의 추량 혹은 의지를 나타내는 조동사 “まい”는 그 대역어가 한국어에도 존재하므로 일-한에 있어서는 1:1 대응 처리로서 번역이 가능하다. 결국, 일본

어의 부정어 중에서 부정어 처리를 필요로 하는 부정 표현은 “ない” 와 “ん(ぬ)” 이 들어가는 부정문이다. 그에 비하여 한국어의 부정 표현은 “못,안”이라는 부정소를 용언의 앞에 배치하는 방법과 부정 보조 용언 “지 않다” 혹은 “없다” 를 사용하여 부정문을 나타낼 수 있다. 부정소 “못,안” 은 사용상의 많은 제약이 따르며, 부정 보조 용언 “지 않다” 와 “없다”로도 모든 표현이 가능하므로, 일본어의 부정 표현 “ない” 와 “ん(ぬ)” 은 한국어의 부정 표현 “지 않다” 와 “없다” 로 번역하여도 큰 무리가 없다. 부정에는 다음의 3가지로 나누어서 생각할 수 있다[11].

예문 5) 太郎は背が小さくない

타로우는 키가 작지 않다

타로우는 키가 작다는 사실이 단순히 부정되어 있다. 이런 부정을 단순 부정이라 한다.

예문 6) 太郎は家に歸らない

타로우는 집에 가지 않는다

이문은 2가지의 의미로 해석할 수 있다. 한가지는 예문 5와 마찬가지로 단순 부정을 나타낸다. 다른 한가지는, “타로우”는 “집에 가고 싶지 않기 때문에 가지 않는다”는 의미를 나타낸다. 이 경우를 의도 부정이라 한다.

예문 7) 太郎は家に歸れない

타로우는 집에 돌아갈 수 없다

이문은 주어인 “타로우”가 무엇인가의 원인에 의해 “갈 수 없다”는 것을 나타낸다. 이 경우를 능력 부정이라 한다.

단순 부정과 의도 부정은 한국어의 부정 보조 용언 “지 않다”로 번역 가능하며, 능력 부정은 부정 보조 용언 “없다”로 번역 가능하다. 이와 같은 처리는 MFOLT 의 각 자질의 활성화를 조정하는 것에 의해 처리 가능하다. 또한, 일본어의 “知る, 分かる, 解かる” 가 부정의 의미로 쓰이는 경우는 “알지 않다” 대신 한국어의 대립 긍정 용언 “모른다” 로, 존재를 나타내는

동사 “有る, 在る, 居る” 는 “있지 않다” 대신 “없다” 로 번역하는 것이 생성된 한국어의 표현이 자연스러우며, 이것을 위해 부정 표현이 부정의 의미를 포함하는 긍정 표현으로 바뀌어지는 특수 처리를 행할 필요가 있다.

그림 2 일본어의 부정어 “知らない” 의 번역 처리

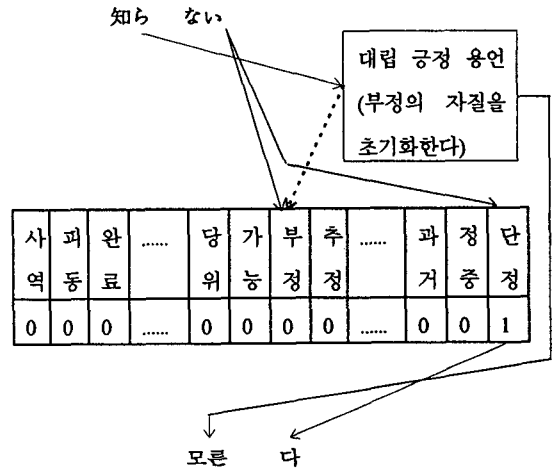


그림 2는 조동사 “ない”가 부정 의미의 자질을 활성화시키지만, 술부의 어간 “知ら”가 “모른다”로 바뀌면서 부정의 의미를 포함하는 대립어가 되므로 부정 의미의 자질을 다시 초기화한다.

2.3.3 피동 표현 처리

한국어의 피동 표현은 “아어 지다”의 일반적인 표현 방법과 동시에, “이,히,리,기” 의 보조 어간을 사용하여 표현한다. 이러한 처리는 용언의 적절한 분류와 MFOLT 의 자질들을 이용함으로써 해결 가능하지만, 하다 동사의 일부는 특별한 처리를 필요로 한다. 즉, 피동으로 바뀔 때, 그 형태가 보통의 형태와는 다르게 나타난다.

보호하다. > 보호하여 지다. (X)

보호 받다. (O)

사랑하다. > 사랑하여 지다. (X)

사랑 받다. (O)

이러한 동사들은 MFOLT 를 이용해서 술부의 순

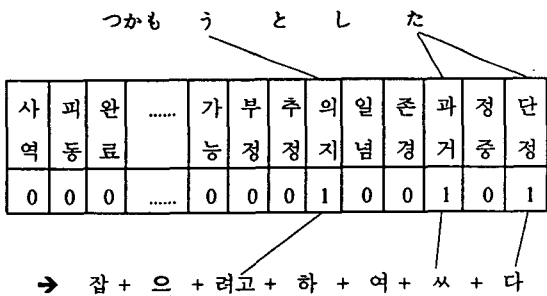
서가 조정된 후, 피동 표현을 위한 특수 처리를 해 줌으로써 해결하고 있다.

2.3.4 보조 용언, 선어말 어미, 어말 어미의 이형태 처리

양상류의 순서 테이블을 사용함으로써 양국어 간의 국부적인 어순의 불일치는 해결할 수 있다. 그러나, 술부의 표현은 한 개 이상의 자질들로 구성되며, 순서가 맞게 나열되더라도 각 표층 표현들 간의 음운 법칙에 따라 형태가 다른 표현이 되어야 한다.

각 자질들은 전접하는 단어의 종자음의 유무에 따라 보조어간 “으”가 접속될 수 있으며, 또한 모음조화에 따라 “아,야”의 모음으로 끝나는 표층 표현에는 “아”를, “어,여,으,이”의 모음으로 끝나는 표층 표현에는 “어”를 삽입한다. 그림 3은 어간의 종자음의 영향으로 “으”를 삽입하는 과정을 보여준다.

그림 3 “つかもうとした”의 번역 처리



2.3.5 음운 축약 및 불규칙 음운의 처리

MFOLT에 의해 생성된 한국어의 술부는 음운 법칙에 의해 나열된 상태이나, 음운의 축약화 및 불규칙 음운들의 처리를 더해 줌으로써 보다 부드러운 표현의 술부를 생성한다.

보이 + 었다 -음운 축약 처리-> 보였다
 굶 + 어서 -불규칙 처리-> 그어서

3. 번역 실험 및 분석

3.1 번역 실험

양상류 테이블을 이용한 한국어의 술부 생성 방법이 유효한가를 검증하기 위하여, 일본어 499 문장을 대상으로 술부의 번역 처리를 행하였다. 실험에 사용된 문장들은 1994년 10월 1일에서 1995년 2월 31일 사이에 발행된 아사히 신문에서 정치, 경제, 사회, 국제, 칼럼면 등을 이용했다. 표 2는 MFOLT를 심볼 별로 분석한 결과이다.

예문에 나타난 양상 자질 중에서 가장 많은 빈도를 보인 것은 연결(아/어/여): 899, 종결(다/나다/는다): 701, 관형사 전성(은/는/ㄴ): 600의 순이며, 각각 96%, 93%, 99%의 번역률을 보였다. 전체적으로는 3,693개의 양상 자질 중 85개가 번역 실패되어 97.7%의 높은 번역 성공률을 보였으며, 이는 MFOLT를 이용한 술부의 생성 처리가 효과적임을 보인다.

3.2 술부의 생성 실패 분석

MFOLT를 이용한 술부의 생성은 상당한 성공률을 보였으나, 다음 예문들은 생성된 번역문이 자연스럽지 못했다. 대역문 중에서 첫 번째 번역문(I)은 본문에서 제안한 생성 방법에 의해 생성된 것이며, 두 번째 번역문(II)은 올바른 결과를 나타낸다.

1) pCAN(가능)의 생성 실패

延期は避けられないと判断した

- (I) 연기는 피하여지지 않다고 판단하였다
 (II) 연기는 피할 수 없다고 판단하였다

2) sSTATE(상태)의 생성 실패

舊ソ連領土内の捕虜收容所に抑留されていた日本国籍保有者に對し,

- (I) ...포로수용소에 억류되고 있은 일본 국적...
 (II) ...포로수용소에 억류되어 있었던 일본 국적...

3) tCOND(가정)의 생성 실패

路線の表明であると判断,

- (I) 노선의 표명이면 판단,

(II) 노선의 표명이라고 판단,

표 2 MFOLT의 빈도 및 매칭 결과

양상 자질 기호	의미	예	빈도수	실패수	번역률
cAUS	사동	게 하 (이/히/리/기/우/구/추)	6	0	100
pASS	피동	게 되(이/히/리/기/ 아/어/여) 지	107	6	94.39
sFIN	완료	{아/어/여} 버리	0	0	
sWISH	희망	고 싶	14	2	85.7
sPROG	진행	고 있	188	0	100
sSTATE	상태	{아/어/여} 있	184	10	94.57
sBECM	진행	{아/어/여} 고	8	0	100
sMUST	당위	{아/어/여}야 하	7	0	100
sCAN	가능	{르/을} 수 있	11	0	100
sNEG	부정	지 않	84	0	100
sPLAN	추정	{르/을} 것이	0	0	
sONLY	일념	{르/을} 뿐이	0	0	
sHONO	존경	(으)시	0	0	
sPAST	과거	왔/였/였	435	2	99.54
sPLT	정중	{비/습)니	5	0	100
tSYN	동시동작	자마자	0	0	
tINYONG	전문	ㄴ다고/는다고	1	0	100
tBUT	역접	지만/다만	54	5	90.74
tBTW	역접	는데/은데	0	0	
tEVEN	양보	더라도	14	1	92.86
tSOIH	추량	도록	12	0	100
tREA1	원인/결과	므로/으므로	11	0	100
tREA2	연결	아/어/여	899	34	96.22
tCOND	가정	면/으면	39	5	87.18
tAND1	연결	고	130	9	93.08
tAND2	선택	느니	0	0	
tOR/tORI	선택	든지/거나	5	0	100
tOBJ	목적	러/으러	2	0	100
tING	동시동작	면서/으면서	3	0	100
tINTR	의문	까니/인가	1	0	100
tYOUKA	추량	{르/을)가	1	0	100
tEXCL	감탄	(는)구나	6	0	100
tIMPE	명령	{아/어/여)라	0	0	
tIMPE	명령	{거/너)라	0	0	
tDECL	종결	다/ㄴ 다/는다	701	7	93.08
tNO_1	명사전성	미/음	6	0	100
tNOUN	명사전성	기	19	0	100
tPREN	관형사전성	은/는/ㄴ	600	3	99.5
tADVE	부사전성	게	37	1	97.30
sSEEM	추량	니/은/는 것 같	0	0	
sMONO	지정	니은 것	0	0	
sUNAV	필연	리/을 수 밖에 없	0	0	
sUNCAN	불가능	리/을 수 없	2	0	100
sDISALW	금지	면/으면 안되	1	0	100
sMAY	추측	리/을지도 모르	2	0	100
sSUP	미래추측	겠	0	0	
sWANT	(3 자)희망	고 싶어하	0	0	
sLAY	결과/보존	아/어/여 두	0	0	

tSAME	동시성	口/음과 동시에	0	0	
tINTHERE	한정동작	口/음에 있어서	1	0	100
tHOPE	의지	려고으려고	3	0	100
tMORE	심화/접진	리/을 수록	0	0	
tORNOT	의심	리/을지 어떨지	2	0	100
tPAST	과거	니/은	76	0	100
tAFTR	경과	니/은 후	0	0	
tQUEST	의문	니/는가	0	0	
tWHETR	추측	리/을지	5	0	100
tUNTIL	한도	기까지	2	0	100
tACCO	의거	口/음에 따라	1	0	100
tSEEM	추량	리/을 것이다	8	0	100
tLETs	청유	읍시다	0	0	
tSUP	유도	지요	0	0	
합 계			3693	85	97.69

4) tAND1(연결)의 생성 실패

國連安保理決議違反であり,

- (I) 국연 안보리 결의 위반이어,
- (II) 국연 안보리 결의 위반이며,

5) tPREN(관형사 전성)의 생성 실패

自衛隊の活動をめぐる本質的な...

- (I) 자위대의 활동을 둘러싸는 본질적인
- (II) 자위대의 활동을 둘러싼 본질적인

여기에 소개한 번역 실패는 술부의 어색한 표현들을 유형별로 나열한 예문 들로써, pCAN, sSTATE의 경우와 같이 생성 처리 부분에서의 실패 만으로는 보기 어려운 문장들이 같이 포함되어 있다. 이런 종류의 생성 실패는 해석 부분이나 변환 부분에서 생기는 오류가 영향을 미친 경우이며, 해석 부분이나 변환 부분의 개선에 의해서, 생성 부분의 성공률은 자연스럽게 향상될 것이다. 그리고, 생성에 실패한 경우도 대체로 원문의 본질은 전달되는 정도의 가벼운 실패가 많이 포함되어 있어서, 실용 시스템에 바로 적용해도 별 문제는 없을 것으로 예상된다.

4. 결론

일본어의 술부와 한국어의 술부는 대응하는 품사

의 불일치, 국부적인 어순의 불일치 등의 차이를 가지고 있다. 일-한 번역에 있어서, 양국어의 술부 표현의

불일치를 해결하기 위하여, 우리는 술부만을 대상으로 하는 의미적 양상 자질(modality feature)을 테이블화(양상 테이블)하여, 양국어의 술부 표현의 피벗(Pivot)으로 이용하는 방법을 제안하였다. 즉 일본어의 술부를 양상류로 분해한 다음, 양상 테이블(Modality Feature Ordering and Lexicalizing Table)에 각각의 양상 자질을 활성화시킨다. 이 양상 테이블을 이용하여, 한국어 술부의 상대적인 순서를 결정한 다음, 한국어의 접속 정보와 각 자질의 상태를 참고하여 한국어 술부를 생성한다. 또한, 생성 규칙에 기반한 활용 어미의 이형태 처리 및 음운 축약 처리, 불규칙 변화에 대한 처리 등을 행하여 자연스러운 한국어의 술부를 표현하도록 했다. 일본어 499 문을 대상으로 실제로 술부를 번역 처리해 본 결과, 약 97.7%가 자연스럽게 번역이 되는 것을 확인하였다. 특히, 술부의 양상 표현의 생성 부분은 일본어에 의존하지 않는 양상 테이블의 도입을 제안함으로써 일-한 뿐만 아니라 다른 언어로부터의 술부 생성에도 적용시킬 수 있을 것이다.

참고문헌

[1] 田近 洵一: くわしい國文法, 文英堂(日本), 1987
 [2] Tae-Suk Kim, Syouji Ura: Study on "JAPANESE-TO-

- KOREAN" Machine Translation by Similarity Between the Two Languages, *Technical Report*, No 91001, *Dept. of Administration Eng., KEIO Univ, JAPAN*, 1991 (in Japanese)
- [3] 이은철, 이종혁: 계층적 기호 접속 정보를 이용한 한국어 형태소 분석기의 구현, *한글 및 한국어 정보처리 논문집*, pp 95-104, 1992
- [4] EunJa Kim, Jong-Hyeok Lee, Geunbae Lee: A Table-Driven Modality Generation In COBALT J/K, *PRICAI'94*, pp 759-763, Beijing, China, 1994
- [5] EunJa Kim, Jong-Hyeok Lee: A Collocational-Based Transfer Model for Japanese-to-Korean Machine Translation, *NLPRS'93*, pp.223-231, Fukuoka, Japan, 1993
- [6] Chul-Jae Park, Jyong-Hee Moon, Jong-Geun Kwak, Jong-Hyeok Lee, Geunbae Lee: Collocation-Based MT system from Japanese to Korean and Its Evaluation, *NLP Proceeding of Information Processing Society of Japan*, Vol.109-2, pp.9-26, 1995 (in Japanese)
- [7] 황찬호 외: *한일어 대조분석*, 명지출판사, 1993
- [8] Heedong Lee, M Nakajima, T Agui: A Korean-Japanese Simple Sentence Machine Translation Using Relation Words through Case Forms, *Transactions of the Institute of Electronics, Information and Communication Engineers*, Vol.J-73-D-II, No.3, pp.418-426, 1990 (in Japanese)
- [9] Jung-In Kim, Seiichi Okoma: An Appropriate Selection Method for Multi-translatable Verbs on Japanese-Korean Machine Translation, *Proceeding of Information Processing Society of Japan*, Vol.3, pp.97-98, 1992 (in Japanese)
- [10] Tae-Suk Kim, Syouji Ura: Generation of Korean Based on Connection Forms of the Semantics in Japanese-Korean Machine Translation, *Transactions of Information Processing Society of Japan*, Vol.33, No.12, pp.1578-1588, 1992 (in Japanese)
- [11] Tae-Suk Kim, Syouji Ura: Processing of Negative Sentence in Japanese-Korean Machine Translation, *Transactions of Information Processing Society of Japan*, Vol.34, No.5, pp. 892-904, 1993 (in Japanese)
- [12] Jung-In Kim, Seiichi Okoma: A Method of Generating Korean Language By Using Extended Translation Tables on Japanese-Korean Machine Translation, *Transactions of Information Processing Society of Japan*, Vol.37, No.9, 1996 (in Japanese)
- [13] 崔紀鮮, 金泰完: 日韓機械翻譯システムの現状および分析, *言語處理學會 第2回年次大會, 發表論文集*, pp.433-443, 1996 (in Japanese)