

한국과 일본의 기계번역 예에 관한 비교연구

-한국어의 일본어 역을 중심으로-

辛碩基*

(e-mail : dct@konyang.ac.kr)

目 次

1. 서론
 2. 본론
 - 2-1. 기계번역 일반
 - 2-2. 번역 프로그램에 의한 신문기사 번역
 - 2-2-1. 명사·대명사·수사
 - 2-2-2. 동사·형용사·부사
 - 2-2-3. 문형·관용구 표현
 - 2-2-4. 조사
 - 2-2-5. 문체
 3. 결론
-

1. 서론

최근의 IT 산업의 눈부신 발전에 따라서 컴퓨터 하드웨어와 소프트웨어의 발전은 가히 놀랍다고 할 수 있다. 이에 따라 예전보다 더 많은 양의 데이터를 저장할 수 있게 되었고 이를 활용하는 많은 소프트웨어의 개발이 이루어지고 있다.

또한 컴퓨터의 활용 면을 살펴보면, 개인용 컴퓨터의 보급과 성능의 눈부신

* 건양대학교 일본언어문화학과 일본어학

발전에 의해, 예전에 주로 이루어졌던 문서작성이나 계산과 같은 단순한 작업에서 인터넷의 발달에 따른 웹 서핑¹⁾ 쪽으로 흐르고 있다. 이에 따라 다른 나라의 웹 페이지에의 접속이 많아지게 되는데, 이 때 발생하는 것이 자국어가 아닌 다른 나라의 언어로 된 홈페이지의 번역문제이다.

이러한 언어장벽의 문제를 해결하기 위해서 각국의 주요 인터넷 포털 사이트²⁾는 웹 페이지의 번역 서비스를 제공하고 있으며 사용자의 설정여하에 따라 외국의 웹 페이지를 자동으로 모국어로 번역해서 사용자에게 제공하는 수준에 이르고 있다.

이렇듯 일반적으로 쉽게 접할 수 있는 인터넷 포털 사이트의 번역 서비스 외에도, 현재 번역을 대행해주는 번역 사이트도 다수 존재하고 있으며, 사람이 아닌 컴퓨터에 의한 번역을 목적으로, 여러 번역 전문 소프트웨어의 개발도 활발하게 이루어지고 있다.

이러한 전문 번역 프로그램들은 현재, 주로 이공계나 의학 분야의 외국어 논문 번역 등에 사용되고 있으며 특정한 분야의 전용 전자사전을 갖추고 있어 번역 성공률은 일반적으로 상당히 높다고 한다.

본고의 목적은 이렇듯 현재 인터넷 사용자에게 의해 빈번하게 이용되고 있는 번역 프로그램 중에서 일반인이 쉽게 접할 수 있는 인터넷 포털 사이트에서 제공하고 있는 번역 프로그램의 번역 결과를 재료로 하여, 자동번역 결과의 품질과 문제점에 대해서 살펴보고자 하는 데에 있다. 이러한 연구가 시급히 이루어져야 하는 이유는 한국과 일본의 일반인들이 정확하지 못한 번역 결과로 인하여 잘못된 정보를 얻거나, 나아가 무심결에 갖게 되는 서로에 대한 오해와 불신을 불식시킬 수 있다는 점에 있다. 이후 詳述하겠지만 너무나도 잘못된 번역 결과를 일반인들은 무의식적으로 받아들이고 사용하고 있는 현상을 시급히 개선하기 위한다는 점에 연구의 의의를 찾을 수 있을 것이다.

방법론적으로는 분석대상을, 우선, 한국과 일본의 인터넷 포털 사이트의 번역 프로그램에 의해 제공되고 있는 번역문으로 하고자 한다. 이들 양국에서 제공되는 번역문의 비교분석을 통해, 양국의 번역문의 문제점을 파악하고, 나아가 양국의 자동번역 수준을 검토하고, 또한 한국과 일본의 포털사이트의 번역문의 품질 비교를 통해 앞으로 보다 나은 자동번역 서비스를 위한 방안을 모색하고자 한다.

비교분석을 위한 번역 예는 한국과 일본의 인터넷 포털 사이트에서 제공하는

1) web surfing : 흥밋거리나 관심사를 찾아 인터넷에 개설된 여러 사이트에 이리저리 접속하는 일.

2) 포털 사이트(portal site) 또는 웹 포털(web portal)은 월드 와이드 웹에서 사용자들이 인터넷에 접속할 때 거쳐 가도록 만들어진 사이트를 말한다. 포털(Portal)이라는 단어는 본래 '정문' 또는 '입구'를 뜻한다.

번역 프로그램에 의한 번역문을 대상으로 하는데, 이 중 우선 한국어를 일본어로 번역한 文(이하 「韓日 번역문」으로 부름)만을 분석대상으로 하고자 한다.

2. 본론

2장에서는 한국과 일본의 포털사이트의 번역 프로그램에 의한 韓日 번역문의 비교고찰을 하고자 한다.

현재의 한국과 일본에는 많은 인터넷 포털 사이트가 존재하고 있는데 구글(Google)을 제외한 거의 모든 큰 규모의 유명 포털 사이트에서는 한국어와 일본어의 상호 번역 서비스를 제공하고 있다. 이에 필자는 한국과 일본에 존재하는 거의 모든 포털 사이트에서 제공하는 자동 번역 프로그램을 이용하여 한국어의 일본어 번역문 결과를 추출해 내었다.

이들 포털 사이트의 번역문은 조사 초기에는 각 포털 사이트마다 그 번역 결과가 다르게 나타날 것이 예상되었으나 거의 모든 포털 사이트가 완전히 동일한 번역결과를 내었다. 즉 한국의 모든 포털 사이트에 탑재되어 있는 번역 프로그램의 번역 결과는 완전히 동일하다는 것이다. 이러한 사실에서 모든 포털 사이트가 같은 번역 프로그램을 사용하고 있다는 것을 예상할 수 있다.

한편 일본의 포털 사이트의 번역 결과도 대조를 위하여 조사해 보았는데, 의외로 OCN(<http://www.ocn.jp>)이라는 포털 사이트를 제외한 나머지 모든 일본의 포털 사이트는 한국의 포털 사이트의 번역문과 완전히 동일한 번역문을 만들어 내었다. 일본에서는 한국의 번역 프로그램과는 다른 것을 사용하여 자동 번역 서비스를 제공하고 있을 것으로 예상하였으나, 조사 결과 한국 포털 사이트와 완전하게 동일한 결과물을 내었다. 결론적으로 OCN을 제외한 한국과 일본의 포털 사이트에서 제공하는 자동번역은 같은 공히 같은 번역 프로그램을 사용하고 있다고 할 수 있을 것이다.³⁾

한편 상기에서 언급한 일본의 OCN(<http://www.ocn.jp>)이라는 포털 사이트에서는 전혀 다른 번역 결과를 보였다. 이 사이트에서 사용하고 있는 프로그램은 일본의 高電社에서 제공하는 번역프로그램을 사용하고 있었는데, 따라서 이 사이트의 번역결과는 동일한 결과를 보인 한국과 일본의 포털사이트에서 사용

3) 각 포털 사이트에 문의한 결과 구체적인 프로그램명은 공개할 수 없다는 답변을 받아서 실제 사용하고 있는 번역 프로그램은 밝힐 수가 없었다. 다만 필자의 조사 결과 한국의 「ez Trans XP(창신소프트)」라는 번역 프로그램의 번역결과가 이들 포털 사이트에서 제공하는 자동번역의 결과와 완전하게 일치하는 것으로 봐서 이 프로그램을 한국과 일본의 각 포털 사이트에서 공동으로 사용하고 있다고 생각할 수 있다.

하는 번역프로그램(ez Trans XP)와는 전혀 다른 결과를 내었다.

따라서 결론적으로 거의 모든 한국과 일본의 포털 사이트는 하나의 번역 프로그램(ez Trans XP)을 사용하고 있다는 결론을 내릴 수 있으며 인터넷 포털 사이트는 프로그램 개발회사와 라이선스 계약을 맺고 프로그램을 공유하고 있다고 예상할 수 있다.⁴⁾

2-1. 기계번역 일반

컴퓨터 프로그램을 사용하여 번역을 할 때 일반적으로 다음의 단계별로 이루어진다. 우선 원문(Source Language)를 形態素 별로 나누는 형태소 해석 단계와 구문의 구조를 해석하는 구문 해석 단계 그리고 원문이 어떤 의미를 갖고 있는지를 파악하기 위한 의미 해석 단계로 나뉜다.⁵⁾

우선 가장 기본이 되는 형태소 해석 단계를 살펴보면, 한국어의 경우 일본어와는 달리 띄어쓰기를 하기 때문에 일본어에 비해 형태소 해석이 용이하다고 할 수 있다. 스페이스와 스페이스 사이에 존재하는 말을 「語」 즉 단어로 인식하도록 하면 되기 때문이다. 일본어는 이에 반해 띄어쓰기가 없기 때문에 仮名만을 입력하여 변환하는 과정에서 종종 형태소 해석이 애매해지는 경우가 있다.

- (1) a. ここではきものをぬいでください。
- b. ここで履物を脱いでください。
- c. ここでは着物をぬいでください。

이러한 난점은 仮名만을 입력하는 일본어문 작성에서는 빈번하게 나타나지만 한자 혹은 가타카나가 섞인 일본어 원문의 형태소해석은 히라가나 이외의 글자에 의해 어느 정도 그 어려움이 해소될 수 있다.

상기에서 한국어는 띄어쓰기에 의해 단어의 인식이 일본어만큼의 어려운 점은 없을 것으로 예상된다고 하였으나 다음과 같은 잘못된 형태소 해석이 발생하는 것으로 나타났다.

- 4) 기계번역 연구는 전문적으로 이공계에서 자연언어처리를 전공하는 전공자에 의해 연구되고 있으며 국가차원에서는 국어 정보화사업인 '21세기 세종계획'이 2007년에 완료되어, 2억 어절의 말뭉치가 구축됐으며 60만개에 달하는 어휘를 갖춘 전자사전도 개발됐다. 다만 이러한 사전에 탑재한 번역 프로그램은 아직 개발되고 있지 않다. 따라서 본고는 현재 일반인이 쉽게 접할 수 있는 포털사이트의 번역 예를 대상으로 비교 고찰하고자 한다.
- 5) 형태소 해석, 구문 해석, 의미 해석의 단계를 거쳐 문맥해석이나 언어학 분야의 語用論에 해당하는 지식 해석 단계가 있으나 의미해석 단계조차 언어학 분야에서도 많은 난점이 존재하고 있다. 의미 해석 단계 이후의 단계에 대한 검토는 현실적으로 거의 불가능하기에 본고에서는 구문해석 단계까지만 검토하고자 한다. 이에 대한 언급은 草薙(1988) pp.75~81 참조.

- (2) a. ...가끔 칭찬을 받는다는 응답이 67.2%에 그쳐 큰 차이를 보였는데...
 ...たまたま褒め言葉を受けるという回答が 67.2%に止めて大きい差を見せたが...
 b. 엄마와 딸의 모습을 담은 사진전을...
 ママと娘の姿を盛った写真展を...
 c. ...큰 영향을 끼치지만 대화와 이해 부족으로...
 大きい影響をかけるが対話と理解不足で...
 d. ...12월19일선거야말로 이 나라가 바르게 가느냐를 결정하는...
 ...12月19日選挙こそこの国が正しく細いかを決める...
 e. ...연설회에서 직접 사과하는 등 사실관계를 어느 정도 인정하면서도...
 ...合同演説会で直接謝るなど事実関係をどの位認めながらも...

상기의 오류들은 동음이의어의 처리가 문제인 것으로 생각된다. 즉 「받다=受ける」, 「담다=盛る」, 「끼치다=かける」, 「어느 정도=どの位」와 같은 1대1 대응식의 변환에 그치고 있다는 사실을 알 수 있다. 이러한 단순한 1대1 대응식의 변환에서 벗어나기 위해서는 소위 말하는 단어의 「連語 환경」을 프로그램이 번역과정에서 사용하는 사전에 기술할 필요가 있을 것으로 생각된다. 이를 위해서는 보다 정밀한 전자사전의 기술이 필요한데 이에 대한 기술이 아직은 초보적인 단계에 머물고 있는 것을 알 수 있다. 다만 일본의 高電社の 번역 프로그램에서는 상기의 초보적인 오류가 거의 나타나지 않는 것을 알 수 있었다. 이에 대한 구체적인 검토는 형태소 해석 부분에서 다루도록 하겠다.

다음의 구문 해석 단계는 단어가 구성이 되어 문을 형성할 때의 문 구조 해석 단계인데, 이 단계에서 나타나는 문제로는 수식관계, 격조사, 副助詞, 생략 표현 등이 문제시되는데⁶⁾, 한일 번역문 비교에서는 格助詞와 보조동사 등의 해석에서 문제가 발생하는 것을 알 수 있었다.

- (3) a. 청소년에게 외모에 대한 올바른 가치관을 심어주기 위한 ...
 青少年に外貌に対する正しい価値観を植えてくれるための
 b. 그 다음해 그는 집집마다 전화를 가설해 줬다.
 その翌年彼は家ごとに電話を仮設してくれた.
 c. 90년 빚이 160여 만 위안이던 식당을 인수해 알짜기업으로 만들었다.
 90年借金が160余万慰安だった食堂を引き受けて選り抜き企業で作った.
 d. 1년이 지나자 손님이 하루 최대 3000명 수준으로 늘었다.
 1年が経つとお客さんが一日最大 3000人水準で増えた.
 e. 폭력적인 콘텐츠가 범람해 시청률 지상주의로 치닫고 있다고...
 暴力的なコンテンツが氾濫して視聴率至上主義で走り上げている”と...

즉 한국어를 일본어로 옮길 때의 「주다」의 해석은 인칭과 깊은 관련이 있는

6) 草薙(1985) pp.38~74 참조.

데 이러한 인칭의 해석을 하기 위해서는 주는 사람과 받는 사람의 인칭관계의 해석이나 심지어 생략되어 있는 주어의 복원까지도 가능해야 올바르게 해석이 이루어질 것으로 예상되나 현재의 기술수준으로는 거의 불가능에 가깝다고 할 수 있다. 따라서 이러한 「授受動詞」의 해석은 한일 양국의 어떤 프로그램도 올바르게 해석하지 못하고 있는 것을 알 수 있었다. 그리고 助詞의 해석은 명사와 술어와의 「連語環境」의 정확한 기술로 어느 정도 해결할 수 있을 것으로 생각되기에 조사의 해석 단계까지만을 구문해석에 포함하고자 한다.

이어서 마지막으로 의미해석단계인데, 의미해석에서는 입력문이 어떠한 의미를 갖고 있는 지를 해석하는 단계이다. 여기서 말하는 의미에는 단어가 갖고 있는 고유의 어휘적인 의미와 문법어가 갖는 문법적인 의미를 말한다. 일반적으로 보다 정확한 의미해석을 위해 단어나 문법어의 상세한 하위분류를 실시하는데 이러한 하위분류는 해당 어휘를 둘러싼 다른 주변 단어와의 관계를 기술하는 것이 일반적이다. 이러한 다른 단어와의 관계기술이라는 것을 구문해석이라고 생각할 수 있어서 종종 의미해석과 구문해석의 경계는 애매해지게 되고 이러한 경계를 규정하는 것 자체가 무의미하다고 생각할 수 있을 것이다.

따라서 본고에서는 구문해석과 의미해석을 하나의 단계로 보고자 하는데 상기에서도 언급했듯이 의미해석을 포함한 구문해석은 현 단계에서는 거의 초보적인 수준에 머물고 있는 실정이어서 본고에서는 형태소 해석 단계와 助詞의 해석을 포함한 구문해석의 일부분에 머물고자 한다.⁷⁾

2-2. 번역 프로그램에 의한 신문기사 번역

이하 현재 많은 사람들이 사용하고 있는 한국과 일본의 인터넷 포털사이트의 번역 프로그램에 의한 한일번역문을 비교분석하고자 한다.

일반적으로 번역프로그램에는 고유의 사전을 갖고 있고 또한 사용자가 임의로 자신만의 사전을 구축할 수 있도록 사전 섹터가 따로 존재하는데 본고의 번역 例 작성 시에는 소프트웨어의 고유사전만을 사용하였다. 번역대상문은 신문기사(정치, 경제, 사회, 문화, 사설)를 사용하였다⁸⁾.

上記에서 한국어는 띄어쓰기에 의해 단어의 인식이 일본어만큼의 어려운 점은 없을 것으로 언급하였으나 실제 번역 결과, 형태소 해석 단계의 수많은 오

7) 형태소 해석, 구문 해석, 의미 해석의 단계를 거쳐 문맥해석이나 언어학 분야의 語用論에 해당하는 지식 해석 단계가 있으나 의미해석 단계조차 언어학 분야에서도 많은 난점이 존재하고 있다. 의미 해석 단계 이후의 단계에 대한 검토는 현실적으로 거의 불가능하기에 본고에서는 구문해석 단계까지만 검토하고자 한다. 이에 대한 언급은 草薙(1988) pp.75~81 참조.

8) 신문기사는 2007년 10월1일~15일 동아일보 인터넷판 기사를 사용하였다.

류들이 나타났다.

각각의 오류들을 품사별, 문형별, 기타(관용구, 문체) 등으로 분류하였다. 몇 가지 예를 들기로 한다. 번역 예를 제시할 때의 형식 (A/B) 는 A가 한국과 일본의 대부분의 포털 사이트가 사용하고 있는 ez Trans XP(이하 한국의 번역프로그램으로 부름)의 결과이고 B는 일본의 OCN 사이트가 사용하고 있는 高電社の 번역 프로그램(이하 일본의 번역 프로그램으로 부름)의 결과이다.

2-2-1. 명사·대명사·수사

명사, 대명사, 수사에서는 무수한 오류들이 나타났는데 이 중 대부분의 오류들은 한국의 번역 프로그램에서 나타났다.

- (4) a. 허상을 과장하는 식 (式/ような) 의 기회주의적 (機會主義的/日和見主義的) 행태는….
- b. 어느 정도 (どの位/ある程度) 불안감을 느껴 왔다고 할 수 있다.
- c. 그의 2007년 신작 20여 점 (20余点/20点余り) 을 선보인다.
- d. 정 교수 (本堂に教授/チョン教授) 는 “AK2를 활성화시키는….
- e. 실물자산 선호 등 중장기 변수 (変数/変化の要素) 도 작용하고 있다.
- f. 내성을 키우려는 몸부림 (身震い/苦闘) 도 보인다.
- g. 검사받기 위해 (検事受けるために/検査を受けるために) 병원을….
- h. 군부독재 시대처럼 낙하산 인사 (落下傘人事/天下り人事) 가 없다….
- I. 또 기술주 (技術株/ハイテク株) 중심의 나스닥 종합지수도….
- j. 중국이 조속히 위안화 (慰安化/元貨) 절상을 단행해야 한다고 ….
- k. 한 누리꾼 (一ヌリクン/あるインターネットユーザ) 은 ‘명명명’.
- l. …맞다고 보지만 거품성장 (泡成長/バブル成長) 이나 경기 (競技/景気) 후퇴 등….
- m. 환권 정 후보는 “당과 (糖菓/党と) 선거대책위원회가 따로 같….
- n. 처음부터 경선 (経線/選挙戦) 에 참여하지 않고 막판 (どん詰まり/終盤) 에 끼어들어….
- o. 주변 사람들에게 자신 (自分/自身) 은 “경선은….

예문 (4)의 o.를 제외한 나머지 오류들은 한국의 번역 프로그램에서 나타난 반면 일본의 번역 프로그램에서는 문맥에 맞는 단어의 선택이 잘 이루어지고 있는 것을 알 수 있다. 특히 한국의 번역 프로그램은 가령 「~식, 기회주의적, (20)여점, 변수」 등을 비롯한 많은 한자로 구성된 단어를 그대로 옮기는 경향이 있는 반면, 일본의 번역프로그램은 「~ような (~식), 日和見主義的 (기회주의적), 20点余り(20여점), 変化の要素 (변수) 」 등과 같이 문맥에 맞게 잘 번역되고 있는 것을 알 수 있었다. 이것은 몇 개의 예에 불과 하지만 거의 모든 명사, 대명사의 번역에 있어서 이러한 경향이 있는 것으로 파악되었다. 또한

「기술주(4I)」나 「누리꾼(4k)」, 「거품성장(4I)」과 같은 최근에 생긴 단어에 대한 번역의 결과도 일본 쪽의 번역이 최근의 신조어를 잘 반영하고 있는 것으로 나타났다. 이것은 즉 사전의 갱신에 관한 문제인데 일본 쪽 사전의 업데이트가 잘 이루어지고 있는 반면 한국 쪽은 거의 이루어지고 있지 않은 것으로 나타났다.

일본 번역 프로그램 쪽의 오류로 지적될 수 있는 것은 「자신」을 「自身」으로 번역한 결과뿐이었다. 이는 사전의 간단한 수정으로 오류를 나타나지 않게 할 수 있을 것으로 생각된다.

2-2-2. 동사·형용사·부사

- (5) a. 헌법이 유린당하고 (踏み躪られて／蹂りんされて) 법치가 사라져다….
 b. 최재천 대변인은 “가치가 곧 (すなわち／まもなく) 정책이다.
 c. 능력이 탁월하다는 (卓越だという／卓越するという) 평가와….
 d. 위원회 위원장을 지낸 (勤めた／過ごした) 서울대 장호완 교수는….
 e. 신 재생에너지 개발 등에 드라이브를 걸어야 한다 (かけなければならない／歩かなければならない) ….

동사, 형용사 등과 같은 술어를 중심으로 하는 예문을 살펴보았는데, 이들 예문 중에서 5a~e는 한국의 번역 프로그램이 바른 결과를 낸 예이다⁹⁾. 이 예들 중에서 「탁월하다」와 같은 형용사를 일본어 번역 프로그램이 한어동사로 잘못 인식한 것은 사전을 구성할 때의 단순한 오류로 판단되기에 간단한 수정으로 처리할 수 있을 것으로 생각된다. 다음으로 한국어의 「곧」의 번역으로 「이제 곧」의 「まもなく」와 서로 동일하다는 의미의 「곧」이 있는데 이는 전후 문맥의 판단이나 「곧」을 전후한 連語환경을 검색해야하는 과정이 필요한데 이는 문법적, 담화적 이론의 적용이 요구되기에 현 시점에서는 불가능할 것으로 판단된다. 따라서 한국어 프로그램이 바른 결과를 낸 것은 우연한 결과로 생각된다. 이를 확인하기 위해 다음의 문을 번역해 보았다.

- 6) a. 기차가 앞의 역을 지나 곧 도착할 것이다.
 b. 汽車が前駅を通過してすなわち到着するでしょう。
 c. 汽車が前の駅をすぎて、まもなく到着するだろう。

예문 6) a.의 한국어문을 번역한 결과, 예상대로 한국 번역 프로그램은 6) b.의

9) 「유린당하고」의 번역예로 일본 번역 프로그램에서는 「蹂躪されて」를 들고 있는데 이 번역이 틀린 것은 아니나 한국 번역프로그램의 「踏み躪られて」쪽이 보다 자연스럽다는 느낌을 주기에 한국 번역프로그램의 번역예를 바른 것으로 간주하였다.

결과를 보였고, 일본 번역 프로그램은 6) c.의 결과를 보였다. 따라서 한국어 「곧」에 대한 번역 결과는 그 의미의 여하에 따라 번역 결과가 다르거나 틀리거나 할 수 있을 것으로 판단된다. 따라서 이 한국어 「곧」에 한해서는 한국이나 일본의 번역 프로그램 모두 1대1 대응식의 사전구성을 하고 있다는 사실을 지적할 수 있다.

다음으로 예문 5) d.의 「(위원장을) 지내다」의 번역 예를 검토해 보기로 하겠다. 예문 5) d.의 번역 예는 한국 번역 프로그램이 바르게 결과를 내었는데 반해 일본 번역프로그램은 「(시간을) 보내다」의 「過ごす」로 잘못된 결과를 내고 있는데 이 「지내다」라는 단어에 한해서 한국 번역 프로그램은 문맥에 맞게 「勤める」와 「過ごす」를 구분하여 그 결과를 내었다.

- 7) a. 나는 중학교 때 늘 축구를 하며 지냈다.
 b. 私は中学校の時常にサッカーをしながら過ごした.
 c. 私は中学校の時, 常に班長を勤めた.

그러나 일본 번역 프로그램은 문맥의 의미와는 상관없이 「過ごす」만을 제시하였다.

- 8) a. 나는 중학교 때 늘 축구를 하며 지냈다.
 b. 私は中学校の時いつもサッカーをしながら過ごした。
 c. 私は中学校の時いつも班長を過ごした。

따라서 한국 번역 프로그램의 경우 「지내다」에 관해서는 문맥에 맞게 번역결과를 제시하는 데 반해 일본 번역 프로그램의 경우는 단순한 1대1 즉 「지내다」를 문맥에 상관없이 단순히 「過ごす」와 대응시키고 있는 것을 알 수 있다.

다음으로 한국어의 「걸다」의 번역 예인데, 일본 번역 프로그램이 「(드라이브를) 걸었다」의 「걸다」를 「あるく」로 오역한 경우이다. 다음과 같은 한국어의 「걸다」의 여러 예문을 번역한 결과 일본 번역 프로그램의 경우 「(소송을) 걸다」와 「(최면을) 걸다」의 경우, 「かける」대신에 「あるく」를 사용하여 번역하였다.

- 9) a. 그는 부당 해고라고 회사에 소송을 걸었다.
 b. 彼は不当でと会社に訴訟を歩いた。
 10) a. 그는 관객들에게 최면을 걸어 모두 잠들게 했다.
 b. 彼は観客らに催眠を歩いて、皆寝つくようにした。
 11) a. 彼は不当解雇だと会社に訴訟をかけた。
 b. 彼は観客たちに催眠をかけて皆眠るようにした。

예문 11)의 한국 번역 프로그램의 경우는 모두 「かける」를 사용하여 바르게 옮긴 것을 알 수 있다. 일본 번역 프로그램이 「걸다」의 다른 예문은 바르게 옮기면서 예문 9)와 10)을 「歩く」를 사용하여 번역하는 이유는 현재로서는 알 수가 없다. 이와 같은 여러 가지 뜻을 갖고 있는 기본 동사의 경우, 동사를 중심으로 하는 명사구의 환경을 미리 사전에 기술해 놓는 것으로 해결할 수 있을 것이기에 이 부분에 대한 수정이 필요할 것 같다. 이러한 수정을 위해서는 이들 동사 이외의 다른 동사에 대한 전반적인 번역의 수준여부를 살펴보아야 할 것이다.

이상과 같은 몇몇 동사나 부사 등에서 일본 번역 프로그램의 오류가 나타나기는 했으나 이를 제외한 거의 모든 예문에서는 일본의 번역 프로그램이 올바른 결과를 내었다. 즉 상기에서 논한 것처럼, 한국 번역 프로그램은 술어를 둘러싼 連語환경을 무시한 단순한 1대1 대응방식에 의해 번역결과를 내는 오역의 과정을 거치고 있는 것이다. 몇 가지의 예를 들어보겠다.

- (12) a. 그 다음으로 교육과 가족 행복의 상실을 꼽았다 (指折った／挙げた) .
 b. 내용이 없다는 지적이 잇갈리고 있다(行き違っている／交錯している) .
 c. 원자재 가격 상승세도 이어지고 있다 (つながっている／続いている) .
 d. 영토전쟁에서 밀리지 않기 (滞らないために／押されないために) 위해….
 e. 안정적인 성장에 초점이 맞춰질 것으로 보인다 (合わせられるように見える／合わせられるものと見られる) ….
 f. 검사받기 위해 병원을 찾으면 (尋ねれば／訪れれば) “CT를 찍어….
 g. 침입 가능성이라는 단기 요인도 있지만 (いるが／あるが) ….

다만 위의 예문 중에서도 가령 「(‘영토전쟁’에서) 밀린다」의 경우, 일본 번역 프로그램이 바른 번역 결과를 제시하였으나(押される), 「사고가 나서 차가 많이 밀리고 있다」와 같은 문의 번역결과도 똑같이 「押される」로 번역하고 있다. 따라서 예문 12)d의 경우도 1대1 대응식의 번역과정을 거치고 있다. 따라서 일본 번역 프로그램 역시 단어에 따라서는 1대1 대응식의 번역 과정이 존재하고 있다는 것을 알 수 있으며 이에 대한 연구와 辭典記述의 보완이 뒤따라야 할 것으로 생각된다.

이상의 동사, 형용사 같은 술어 중에서 한자어가 들어가는 동사나 형용사의 경우, 한국의 번역 프로그램은 한자를 글자 그대로 즉 한국어식 한자를 그대로 일본어로 옮기는 경향이 보인다. 이러한 경향은 명사에서도 확인되었는데 이러한 예는 여러 품사에 걸쳐 나타나므로 이에 대한 보완작업은 반드시 이루어져야 할 것으로 판단된다.

- (13) a. 유리한 방향으로 만들기 위해 벌이는 (する／行う) 경쟁은 치열한 (熾烈な／厳しい) ‘영토전쟁’을 방불케 한다.
 b. 안정적인 (安定的な／安定した) 성장에 초점이 맞춰질 것으로...
 c. 대선후보의 구호는 숫자와 연관된 (連関された／関連になった) 것이...
 d. 국가주석의 ‘과학 발전관’이 들어간 (入って行った／入った) ...
 e. CT는 심장처럼 빨리 움직이는 장면도 담을 (盛る／入れる) 수 있다.
 f. 높은 (高い／偉い) 사람, 잘사는 사람은 대통령과 대통령후보들...

예문 (13)a.~c.의 경우, 한국어식 한자를 그대로 일본어로 번역하는 경향을 확인할 수 있으며 또한 예문 (13) d. 의 경우는 「들어가다」를 글자 그대로 「入っていく」로 옮기는 것을 확인할 수 있다. (13)e. f.는 상기의 1대1 대응 방식에 의한 오류라고 할 수 있다.

2-2-3. 문형, 관용구 표현

여기에서는 표현문형이나 관용구와 관련된 예문을 살펴보고자 한다.

- (14) a. 선명하게 보일 수 있도록 (見えるように／見えられるように) ...
 b. 시장에서 자취를 감추고 말았다 (影をひそめてしまった／痕跡をなくしてしまった) .
 c. 즉 우파·보수층의 표는 분산되기 마련이고 (分散するものと決まってい／分散するようになっていて) ...
 d. AK2 유전자를 손상시켰더니 (損傷させたら／損傷させたところ) ...

상기의 예문 (14)a.~c.는 표현문형이나 관용구의 번역 결과 중에서 한국 번역 프로그램이 바른 결과를 낸 예들이다. (14)a.의 「보일 수 있도록」의 경우, 일본 번역 프로그램은, 「~할 수 있다」라는 한국어문을 「可能」으로 해석하여 「見える」를 「見えられる」형태로 번역하는 오류를 보이고 있다.

관용구로 생각되는 예문 중에서 (14) b. c.의 경우, 「자취를 감추다」, 「~하기 마련이다」와 같은 관용구의 경우, 한국 번역 프로그램 쪽이 바른 결과를 낸 것으로 생각된다. 이러한 문형이나 관용구는 거의가 정해져 있는 경우가 대부분이기에 사전의 보완으로 충분히 해결될 수 있을 것으로 생각된다. (14) d. 의 경우는, 양 쪽 모두 바른 결과를 보이고 있는 예이다.

이러한 예문들은 모두 한국 번역 프로그램이 바른 결과를 낸 예문들이지만 이들 예문을 제외한 下記の 예문들은 일본 번역 프로그램 쪽이 바른 번역 예를 나타내고 있다.

- (15) a. 이씨는 요즘 세상의 욕이란 욕은 다 먹고 있다 (悪口と言う悪口はすべて食べている／悪口という悪口はみな言われている) .

- b. 전혀 사리에 맞지 않는다 (全然私利に当たらない／理にかなわない) .
- c. 언론에 재갈을 물려 (言論にはみをくわえて／言論にクツワをはめて) 국민의….
- d. 한 누리꾼은 ‘명명명시리즈’라면서 (シリーズと／シリーズとしながら)
- e. 조영제에 과민반응을 보일 수 있으므로 (見せることができるので／見せることがあるので) ….

상기의 예문 중에서 특이한 점은 (15)e. 의 경우인데, 일본번역 프로그램에서 上記의 (14)a. 의 「보일 수 있도록」과 같은 예문은 「見えられるように」와 같이 오역의 결과를 내었지만 (15)e.의 문은 「見せることができるので」라는 바른 번역 결과를 나타내고 있다. 반대로 한국 번역 프로그램은 (15)e.의 예문을 가능으로 처리하여 번역하고 있다. 즉 「～할 수 있다」를 모두 가능으로 처리하여 「～見せることができるので」라는 오역문을 내고 있다. 즉 한국번역 프로그램은 「～할 수 있다」의 문을 가능의 의미로 1대1 대응시키고 있는 반면 일본 번역 프로그램은 문의 의미에 따라 결과를 달리 내고 있다는 사실을 알 수 있다.

또한 이 표현문형이나 관용구 번역 부분에서도 한국어문을 있는 그대로 일본어로 옮기려는 경향이 한국의 번역 프로그램에는 많이 보인다.

- (16) a. 이회창씨가 그것을 모를 리 없다 (分らないわけない／分らないはずがない).
- b. 이탈도 결정적 요인이 될 수 있다 (なることができる／なり得る) .
- c. 전혀 사리에 맞지 않는다 (全然私利に当たらない／理にかなわない) .
- d. 증식해 암 같은 질병이 생긴다 (疾病ができる／病気ができる) .
- e. 정부 기업 할 것 없이 (するまでもなく／関係なく) 세계 경제….
- f. 전환중이라곤 하지만 (転換の中とはしかし／転換の中とはいえ) ….

2-2-4. 조사

다음으로 助詞에 대해 살펴보면, 조사의 번역오류는 그리 많지 않은 것으로 나타났다. 한국어와 일본어가 助詞를 사용하고 있고 그 쓰임새가 비슷하기 때문일 것이다. 가장 빈번하게 나타나는 것은 양국의 프로그램 공히 「で」와 「から」의 오류이다.

- (17) a. 이번 대통령 선거에서 (で／から) ‘가치’로 승부하자”며….
- b. 표준전쟁에서 패하자 시장에서 (で／で) 자취를 감추고 말았다.
- c. 대학 안팎의 부당한 공격에서 (で／で) 교수를 보호하기 위한 ….
- d. 에너지 다소비형 산업에서 (で／で) 고부가가치 산업으로 전환 ….
- e. 남북 평화 등을 자신이 추구하는 ‘5대 가치’로 (で／と) 규정했다.

이러한 「で」와 「から」의 오류는 초급일본어 학습자에게서도 빈번하게 나타나는 오류인데, 동작이 일어나는 장소와 출발점이 한국어에서는 「～에

서」로 표현되기에 모국어의 간섭으로 생각할 수 있다.

그러나 번역 프로그램에서는 초급 학습자가 범하는 모국어의 간섭이란 것이 없기 때문에 조사를 둘러싼 구문환경을 사전에 정확히 기술하고 이를 프로그램이 번역할 때 참조할 수 있도록 해야 할 것이다. 가령 예문 (17)d.의 경우, 「에너지 다소비형 산업에서 고부가가치 산업으로 전환 중이라곤 하지만…」와 같이 「~에서」와 쌍을 이루어 나타나는 「~으로」를 프로그램에서 참조할 수 있으면 「~で」로 오역하는 오류를 피할 수 있을 것이다. 또한 (17)b.와 같은 예문에서는

「표준전쟁에서 패하자 시장에서 자취를 감추고 말았다.」와 같이 동사술어구(자취를 감추다)와의 참조를 통해서 오류를 피할 수 있을 것이다.

이러한 조사나 동사술어와의 참조를 통해서 조사의 오류를 충분히 해결할 수 있을 것으로 생각되기에 한국과 일본의 번역프로그램은 辭典의 보다 정확한 기술이 필요할 것으로 생각된다.

2-2-5. 문제

문제의 오류는 사전의 간단한 정비로 인하여 해결할 수 있는 문제들인데 다음 예와 같이 한국의 프로그램은 구어체 표현을 우선시하고 있다.

- (18) a. 잘 들여다보면 그의 그림 속 오리 같다는 (みたいだという／のようだという) 생각이 든다.
 b. 이런 (こんな／このような) 경우 항암제를 써도 간암세포가….
 c. 교육공약이 공교육을 어떻게 (どんなに／どのように) 정상화….

3. 결론

지금까지 한국과 일본의 포털 사이트가 제공하고 있는 번역 프로그램에 의한 자동번역의 예를 비교 검토하였는데 그 번역 결과에 많은 오역이 있다는 것을 확인할 수 있었다. 또한 한국과 일본의 많은 포털 사이트들은 각각 서로 다른 프로그램을 사용하고 있을 것으로 예상했으나 한 곳을 제외하고 같은 프로그램을 사용하고 있었다.

한국과 일본의 거의 모든 포털 사이트가 사용하는 번역 프로그램에 의한 번역은 여러 품사에 걸쳐 많은 오역을 보인 반면 高電社가 제공하는 번역 프로그램을 사용하는 OCN의 번역 결과는 이에 비해 우수한 결과를 보였으나 양쪽 프로그램 모두 각각 사용하는 사전의 정비가 필요한 것으로 판단된다. 이 두

프로그램 모두 사전에 기술되어 있는 連語情報에 따라 번역을 실시하고 있는데 그 상세함의 정도에 따라 번역 결과는 매우 다르게 나타났다.

여러 포털 사이트에서 사용하고 있는 번역 프로그램의 경우는 한국어 단어를 문맥의 구문환경과 상관없이 1대1 대응식으로 일본어로 옮기는 경향을 확인할 수 있었다. 따라서 단어의 의미와 함께 문맥환경 즉 단어를 둘러싼 連語環境의 정보까지도 사전에 충실하게 포함시켜서 명사고유의 어휘정보에 따른 올바른 번역이 이루어지도록 해야 할 것이다. 이에 비해 高電社 제공의 프로그램을 사용하고 있는 OCN의 경우에도 오류가 나타나기는 했으나 그 수가 적었고 번역 결과도 우수하였다. 이러한 차이는 결국 프로그램이 사용하는 사전에 의해 좌우되는 것으로 생각되는데, 어휘의 連語情報를 보다 충실히 기술하는 것이 급선무인 것으로 판단된다. 또한 나날이 새로 생겨나는 어휘의 업데이트 또한 충실히 하여 인터넷이라는 매체의 속도에 맞추어갈 필요가 있을 것이다.

품사적인 면에서 구체적으로 살펴보았는데, 다른 품사에 비해 명사의 경우, OCN의 번역 성공률이 뛰어난 것을 확인할 수 있었다. 명사를 사전에 탑재할 때 해당 명사의 의미뿐만 아니라 명사를 둘러싼 기본적인 구문환경에 대한 보다 섬세한 기술이 필요할 것으로 판단된다. 이러한 상세한 사전에의 기술은 다른 품사에게도 동일하게 요구되는 사항이다.

마지막으로 관용구의 경우는, 술어를 포함한 관용구를 하나의 단위로 생각하여 사전에 정확하게 기술하는 것이 필요하다. 소위 말하는 「기능동사」의 보다 면밀한 사전에의 기술이 요구된다고 할 수 있다.

몇 십 년 전에 비해 현재는 컴퓨터 하드웨어의 눈부신 발전이 있기에 보다 방대한 양의 데이터를 사전에 기술할 수 있게 되었다. 따라서 이러한 하드웨어를 십분 활용할 수 있는 소프트웨어적인 사전의 세밀한 기술 및 이에 대한 연구가 시급히 이루어져야 할 것으로 생각한다. 이러한 연구와 개발이 이루어짐으로 인하여 일반인이 쉽게 그리고 보다 정확한 자동번역 결과물을 얻어 이를 충분히 활용할 수 있게 될 것이고 오역으로 인한 불필요한 오해를 줄여 서로에 대한 보다 정확한 정보제공과 공유가 가능할 것이다.

【参考文献】

草薙裕, 『LISPによる自然言語処理』, 工学図書株式会社, 1988.

_____, 『自然言語とコンピュータ言語』, 講談社, 1985.

要 旨

本稿は日本と韓国のポータル・サイトで提供している外国語翻訳サービスの翻訳結果を比較分析することを目的としている。このような比較分析を通して、普通に使われているサイト上の翻訳サービスの質を検討すると共にさらによりよい翻訳のための提案を試みた。

比較のための検討資料としては新聞記事の翻訳結果を利用している。本稿ではまず韓国語の日本語訳のみを検討した。翻訳結果物は韓国と日本のポータルサイトが提供している翻訳サービスを利用して出した。最初の予測は、韓国と日本のポータルサイトはそれぞれ異なる翻訳結果を出すだろうからそれぞれの結果が比較できるということであった。しかし、その結果物を比較したら日本のOCN（高電社提供のプログラム使用）というサイトの結果を除いて、すべての日本と韓国のポータルサイトの翻訳結果は同じであった。従って、結果的にはいくつかの翻訳結果が得られず二つの翻訳結果を比較することになった。

比較は機械翻訳における形態素解析のレベルにとどまっている。授受動詞の場合、主語の人称と述語が関係しているのでこれらの動詞の翻訳は構文解析の処で行われるがポータルサイトで提供している翻訳サービスではこのレベルまでには達していないのが確認された。ほとんどの翻訳は、プログラムが使っている辞書に記述されている単語を構文や意味に関係なく一対一に対応する単語を拾ってくるという単純なプロセスによって出されている。

分析は、韓国語を日本語に訳した翻訳文を品詞別に分けて比較をした。まず、名詞などの体言は、圧倒的にOCNの方が正確であった。一見名詞などの体言は語形変化がないので翻訳されやすいだろうという予測とは逆になかなか厄介な問題点が多かった。さらに新造語などの更新の部分でもOCNの方がよい結果を見せた。そして、大部分のサイトが使っている韓国の翻訳プログラムは韓国の漢語名詞などをそのまま日本語に訳すという誤りがよく見受けられた。次の動詞をはじめとする用言の方では、名詞ほどではないが用言の翻訳でもOCNの翻訳の質が高かった。一部誤った結果を出してはいるが、大体文脈にあっている結果を出している。これに対して韓国の翻訳プログラムはやはり韓国の漢語動詞などのような漢字で出来ている述語を日本語にそのまま訳すという傾向がまた見られる。この述語の場合は連語環境の正確な記述が両方とも求められる。最後に慣用語句などの部分でも、韓国の翻訳プログラムの質は劣っている。OCNにもいくつかの誤った結果はあったが、連語関係の正確な記述がより多くなされているところがうかがえる。

これまでの分析を通して、大部分のポータルサイトが提供している翻訳サービスの質は極めて低いということが分かった。これを改善するためには、翻訳プログラムの辞書の整備とアップデートが必要であろう。より正確な形態素解析のために連語関係の記述を辞書に載せなければならないのである。

キーワード：ポータルサイト、自動翻訳文対照、形態素解析、プログラム辞書
連語環境、辞書記述

투 고 : 2008. 8. 31
1차 심사 : 2008. 9. 12
2차 심사 : 2008. 9. 27