

## 한국어 스타일 생성 패턴에 의한 영한 번역 품질 개선

한국전자통신연구원, 음성/언어정보연구센터, 언어처리연구팀  
최승권 · 홍문표 · 박상규

### Enhancement of English-to-Korean Translation Quality by Korean Style Generation Patterns

Sung-Kwon Choi, Munpyo Hong, Sangkyu Park

*ETRI, NLP Research Team, Speech/Language Technology Research Center, Daejeon, Korea*

#### 요 약

본 논문에서는 영한 자동번역 시스템에 한국어 스타일 생성 패턴을 적용함으로써 영한 번역 품질을 향상하고자 하는 것이 목표이다. 이러한 목표는 기존의 원문에 대한 번역문의 정보 전달 정확성을 측정하는 1차원적인 번역률 평가 방법에서 벗어나 번역문의 정보 정확성뿐만 아니라 자연스러움도 평가할 수 있는 2차원적인 번역률 평가방법으로써 정확성과 스타일을 동시에 평가하는 방법을 제안한다. 2차원적인 번역률 평가 방법에 따라 스타일 생성 패턴이 적용되기 전과 적용된 후의 평가 결과는 100 문장의 샘플문을 대상으로 하였을 때, 스타일 생성 패턴에 의해서만 0.5%의 번역률이 향상되는 것을 관찰하였다. 본 논문에서의 스타일 생성 패턴은 단순히 언어간 스타일 차이만 적용한 것이며 향후에는 신문, 일기예보, 기술 매뉴얼과 같은 특정 그룹을 위한 스타일 생성 패턴을 적용할 계획이다.

#### 서 론

영한 자동번역이 국내에서 연구 개발된 지도 벌써 10여 년이 지나고 있다. 그 동안 영한 자동번역기의 응용은 시대의 변천에 따라 텍스트 영한 자동번역기로부터 웹 영한 자동번역기, 방송자막 자동번역기, 모바일 영한 자동번역기 등으로 그 응용이 확대되고 있다.

그러나 영한 자동번역기에 대한 일반인들의 의식은 지난 10년 전에 비하면 개선되었는지 몰라도 아직도 영한 자동번역기는 영어에 대한 언어장벽을 해소해 줄 수 있는 편리한 소프트웨어로 인정받고 있지 못한 실정이다.

이러한 이유는 영한 자동번역기의 번역률이 일반인들이 만족하기에는 아직도 낮기 때문일 것이다. 그러나 여러 평가 결과에 따르면 영한 자동번역기의 번역률은 평균

60~65%에 달하는 것으로 발표되고 있다.<sup>1,2)</sup>

이 정도의 번역율이라면 일반인들도 어느 정도 번역기의 품질을 인정할 수 있지 않을까 하지만 위의 번역율은 대부분 영어 원문에 대한 한국어 번역문의 충실도 항목들인 어휘, 관용어, 논항, 용법등과 같은 언어학적 문법 항목들이 평가 항목 수에서 많은 부분을 차지하다 보니 한국어를 모국어로 하는 독자들이 영한 자동번역기의 번역문을 읽었을 때 번역의 자연스러움을 반영하는 한국어 스타일에 대한 반영이 충분히 고려되지 않음으로써 한국어 번역문을 보는 독자들이 번역문을 이해하는 데 어려움을 갖는다.

본 논문에서는 이러한 영어 원문 구조가 한국어 번역문 구조에 반영됨으로써 한국어 번역문의 이해도가 낮아지는 것을 방지하기 위해 한국어 스타일 생성 패턴에 의한 영한 번역 품질을 개선하고자 하며 또한 영한 번역 품질 평가에서도 원문의 정보를 번역문에 얼마나 정확히 전달하였는가를 평가하는 기존의 정확성 평가 방법 이외에 스타일 측면에서의 평가 방법을 제시함으로써 번역문의 자연스러움도 평가의 대상으로 확대하고자 한다.

E-mail : choisk@etri.re.kr  
E-mail : hmp63108@etri.re.kr  
E-mail : parksk@etri.re.kr

## 스타일 번역

### 1. 정 의

스타일이란 라틴어 ‘stilus’에서 유래한 것으로 본래는 ‘쓰는 기구’를 의미한 말이었다. 일찍이 고대 라틴어에서 ‘글을 쓰는 법’에 응용되었고 그 뒤 글이나 말로 ‘자신의 특징을 표현하는 방법’ 또는 이 때 나타나는 사람의 특징으로 그 의미가 일반화되었다.<sup>3)</sup>

그러므로 스타일과 관련한 스타일 번역의 정의는 원문 저자의 의도를 유지하면서 가장 자연스러운 목표 언어 문장을 번역하는 것을 말한다. 영어를 한국어로 번역한 스타일 번역의 예를 들면 다음과 같다 :

- 1) 가. Poverty drove her to despair.  
 나. 가난은 그녀를 절망시켰습니다(영어 스타일).  
 다. 가난 때문에 그녀는 절망했습니다(한국어 스타일).
- 2) 가. His failure in the examination was due to lack of time.  
 나. 그 시험에 그의 실패는 시간의 부족 때문이었습니다(영어 스타일).  
 다. 그가 그 시험에 실패한 것은 시간의 부족 때문이었습니다(한국어 스타일).

위의 예에서 1) 가)와 2) 가)의 번역을 1) 나)와 2) 나)와 같이 하는 것은 영어 구조를 그대로 한국어로 옮겨 놓은 것으로써 한국어를 이해하는 데 어려움이 있으나 1) 다)와 2) 다)는 한국어 구조가 반영된 자연스러운 한국어 번역문들이다.

### 2. 스타일 번역의 종류

스타일 번역은 스타일을 형성하는 주체에 따라 크게 2개의 유형으로 나눌 수 있다.

하나는 유형적인 스타일 번역이며 다른 하나는 개성적 스타일 번역이다. 이들의 비교는 다음과 같다 :

이러한 스타일 번역을 결정짓는 요소는 단어의 선택(예 : 한글체 v 한자체), 어미의 선택, 문장의 길이(예 : 간결체 v 만연체), 서술의 방법(예 : 건조체 v 화려체, 강건

Table 1. 스타일 번역의 유형

유형적 스타일 번역	개성적 스타일 번역
- 특정 사회의 특성 반영	- 특정 개인의 특성 반영
- 공통적 표현을 대상으로 함	- 독자적 표현을 대상으로 함
- 언어학적 입장	- 문학적 입장
- 기술적 연구	- 발생적 연구

체 v 우유체) 등으로 형성된다.

본 논문에서는 개성적 스타일 번역은 향후 기회가 있으면 다루기로 하고 유형적 스타일 번역 중에서 언어학적 입장에서 언어적 스타일만을 다루고자 한다.

## 관련 연구

번역에서 스타일과 관련한 언급은 동종언어보다도 이종 언어간에 스타일 처리가 중요하며 그것의 위치를 생성 모듈에서 설명한 것,<sup>4)</sup> 스타일 번역을 위한 스타일 분석 규칙을 최초로 소개한 것,<sup>5)</sup> 한국어와 관련해 영한 자동번역에서 스타일 변환 규칙에 의해 영한 스타일의 중요성을 언급한 것<sup>6)</sup>이 있다.

본 논문이 기존의 스타일 처리 연구들과 차이가 나는 것은 우선 패턴 기반 자동번역에서의 스타일 처리의 적용이며, 두번째는 영어 워드넷의 명사 시소로스<sup>7)</sup>를 토대로 한 영한 번역용 의미코드체계<sup>1)</sup>로 스타일 생성 패턴을 기술하였다는 것이다.

## 한국어 스타일 생성 패턴

한국어 스타일 생성 패턴은 ETRI에서 개발중에 있는 웹 영한 자동번역 시스템인 Tellus-EK에 직접 활용되었기에 본 논문에서는 Tellus-EK 시스템을 토대로 설명하고자 한다.

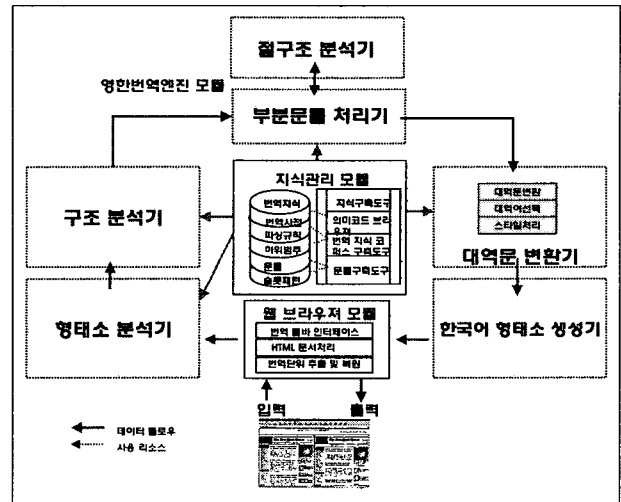


Fig. 1. Tellus-EK에서의 스타일 생성 패턴의 위치.

1 영한 웹 자동번역 시스템인 Tellus-EK에서 번역패턴에 활용되는 의미코드체계는 워드넷 1.7.1 버전의 명사 시소로스를 기반으로 하고 있으나, 번역패턴의 커버리지(Coverage)를 높이기 위해 워드넷의 모든 Synset을 활용하지는 않고 최상위로부터 420개의 의미코드를 활용하고 있다. 이 의미코드 체계는 앞으로 계속 세분화될 계획이다.

### 1. 스타일 생성 패턴의 시스템 위치

한국어 스타일 생성 패턴의 시스템에서의 위치는 Fig. 1과 같다 :

Tellus-EK 시스템을 요약하여 설명하면 다음과 같다 : 영어 형태소 분석(형태소 분석기)과 구조분석(구조 분석기)이 이루어진 후, 부분문틀 처리기에서는 분석 결과가 번역 패턴 DB에 존재하면 대역문 변환기로 넘기고, 번역 패턴이 존재하지 않으면 절구조 분석기에서는 문장을 절 단위로 분절 분석하여 절단위 부분번역패턴을 번역 패턴 DB에서 검색할 때까지 분석한다. 대역문 변환기에서는 영어 패턴과 한국어 대역패턴에 따라 번역을 실시하고 그 이후 한국어 형태소 생성(한국어 형태소 생성기)이 이루어진다.<sup>8)</sup>

한국어 스타일 생성 패턴은 대역문 변환기에 위치하며 대역어 선택 처리가 된 한국어 슬롯구조를 입력으로 받아서 스타일 처리된 한국어 슬롯구조를 출력해주는 기능을 한다. 한국어 스타일 생성 패턴이 대역어 선택 세부블록 다음에 오는 위치의 장점은 기존의 번역 패턴의 수에 영향을 미치지 않는다는 것이다. 그러나 단점은 번역 패턴에서의 자의적인 한국어 대역 어순에 의존적일 수 있다는 것이다.

### 2. 스타일 생성 패턴의 포맷

스타일 생성 패턴 DB의 BNF 표현 및 예는 다음과 같다.

```

<스타일 생성 패턴> : = <대역어 선택후 한국어 슬롯구조> -> <스타일 적용 후 한국어 슬롯 구조>
<대역어 선택후 한국어 슬롯구조> : = { <슬롯명> ! ( <슬롯명> : <슬롯자질> ) * }
<스타일 적용후 한국어 슬롯구조> : = { <슬롯명> ! ( <슬롯명> : <슬롯자질> ) * }
<슬롯명> : = NP ; VERB ; PP ; CONJ
<슬롯자질> : = nil ; <슬롯속성> <연산자> [ <슬롯값 > ] ( , _AND < 슬롯자질> ) *
<슬롯속성> : = kcase ; kend ; kroot ; sem ; voice
<슬롯값> : = <의미> * <타입> <어휘>
<연산자> : = = ; * ; ** ; =
    
```

#### 3) 가. Poverty drove her to despair.

나. 가난은 그녀를 절망시켰습니다. -> 가난 때문에 그녀는 절망했습니다.

다. {NP1 : [kcase == [topic] \_AND sem \*\* [condition]] NP2 : [kcase == [obj] \_AND sem \*\* [person]] VERB1! : [voice == [caus]] }  
 -> {NP1 : [kcase : = [때문에]] NP2 : [kcase : = [topic]] VERB1! [voice : = [activ]] }

스타일 생성 패턴은 화살표를 경계로 좌측부에 스타일 처리가 되지 않은 구문 슬롯과 슬롯의 제약조건이 명시되며 우측부에 스타일 처리된 구문 슬롯과 해당 슬롯의 자질 값이 명시된다. 한편, 좌측부에 명시하지 않은 슬롯이지만 보다 매끄러운 대역어 생성을 위해 우측부에 새로운 슬롯

을 삽입할 수도 있다.

### 3. 스타일 생성 패턴의 종류

한국어 스타일 생성 패턴은 현재 영한 웹 자동번역 시스템인 Tellus-EK에서 계속해서 수정, 보완되고 있다. 현재는 언어 스타일 패턴과 관련해 5개의 그룹으로 구성되어 있지만 향후 세분화될 가능성이 있음을 명기한다.

#### 1) 수동태의 능동태화

국어에는 수동을 표현하는 문장이 잘 발달되어 있지 않으며, 수동태의 모양도 통일되어 있지 않다. 그와 반면에 영어에는 수동태가 잘 발달되어 있다. 영한 자동번역에서 영어의 수동태를 능동태로 바꿔줌으로써 번역이 자연스러울 때가 많다. 특히 복잡한 문장일수록 수동태를 사용하기가 어려워지므로 능동태로 바꿔주는 것이 좋다.

[스타일 생성 패턴 : passive-to-active]

##### • 조 건

주어가 무생물 동사화 명사구이고, 주동사가 수동태이면

##### • 실 행

주어를 부사구로, 주동사를 능동태로 스타일 변환

##### • 패 턴

```

{NP1 : [kcase == [topic] _AND sem ** [act]] NP2 : [kcase == [obj]] VERB1! : [etype == [d1] _AND voice == [passive]]}
-> {NP1 : [kcase : = VERB1.KFLEX2] NP2 VERB1 : [voice : = [active]]}
    
```

##### • 예

Preservation of Antarctica's natural resources must be given the highest priority

스타일 적용 전 : 남극 대륙의 천연 자원의 보존은 가장 높은 우선권을 주어져야 합니다.

스타일 적용 후 : 남극 대륙의 천연 자원을 보존하는 것에 가장 높은 우선권을 주어야 합니다.

#### 2) 양상술어를 양태부사로

영어의 화법조동사인 must나 확신과 관련된 양상표현을 한국어로 번역한 경우 "...임에 틀림이 없다.... 임이 분명하다"와 같이 주어를 길게 번역하는 경우를 한국어 번역 문장에서 흔히 볼 수 있다. 틀린 번역은 아니지만 한국어에서는 조금 어색한 번역이 아닐 수 없다.

[스타일 생성 패턴 : modaux-to-modaladverb]

• 조 건

주어가 Heavily Left-branched 절이고, 주동사가 형용사인 양상술어이면,

• 실 행

주어를 주문장으로, 양상술어를 양상부사로 스타일 변환

• 패 턴

{NP1 : [kcase==[subj]] NP2 : [kcase==VERB1.KFLEX2] VERB1 : [kend==[ㄴ것은]] VERB2! : [kpos ==[adj]]}

-> ADV : [kroot : =VERB2.kroot, kend : =[계]]

NP1 NP2 VERB1 : [kend : =[decl]]}

• 예

It is pitiful that he failed in the entrance examination.

스타일 적용 전 : 그가 그 입학 시험에 실패한 것은 불쌍합니다.

스타일 적용 후 : 불쌍하게 그는 그 입학 시험에 실패했습니다.

3) 보조동사 '돕다'를 가능성 양상보조사로

[스타일 생성 패턴 : help-to-can]

• 조 건

주어가 무생물이고, 목적어가 생물이고, 주동사가 '돕다'이면

• 실 행

주어를 원인 부사구로, 목적어를 주어로, 주동사를 보조사로 스타일 변환

• 패 턴

{NP1 : [kcase==[obj]] VERB1 : [kend==[는\_것은]] NP2 : [kcase==[subj]] NP3 : [kcase==[obj]] VERB2 : [kend==[도록]] VERB3! : [kroot==[돕다]]}

-> {NP1 VERB1 : [kend : =[ㄴ으로써]] NP2 : [kcase : =[topic]] NP3 VERB2 : [kend : =[decl], modal : =[ㄴ수있]]}

• 예

Learning a foreign language will help you find a solution.

스타일 적용 전 : 외국어를 배우는 것은 당신이 해법을 발견하도록 도울 것입니다.

스타일 적용 후 : 외국어를 배우으로써 당신은 해법을 발견할 수 있을 것입니다.

4) 사역동사를 자동사로

<sup>4)</sup>에서도 언급한 바와 같이 이중언어인 영어와 한국어 사이의 번역에서 가장 흔하게 접할 수 있는 스타일 번역의 예는 영어의 무생물 주어가 한국어에서는 원인 부사구로 번역되어야 하는 것이다. 이러한 스타일 생성은 기존의 <sup>9)</sup>에서 언급한 7가지의 번역의 차이<sup>2)</sup>에서는 찾아볼 수 없는 번역차이의 유형이며 명사의 의미가 기반이 되지 않으면 설명할 수 없는 현상이다.

[스타일 생성 패턴 : causative-to-intransitive]

• 조 건

주어가 무생물 주어이고, 주동사가 사역동사이고, 목적어가 생물이면,

• 실 행

주어를 원인 부사구로, 주동사를 자동사로, 목적어를 주어로 스타일 변환

• 패 턴

{NP1 : [kcase == [topic] \_AND sem \*\* [condition]]

NP2 : [kcase==[obj] \_AND sem \*\* [person]] VERB1! : [voice==[caus]] }

-> {NP1 : [kcase : =[때문에]] NP2 : [kcase : =[topic]] VERB1! [voice : =[activ]] }

• 예

Poverty drove her to despair.

스타일 적용 전 : 가난은 그녀를 절망시켰습니다.

스타일 적용 후 : 가난 때문에 그녀는 절망했습니다.

5) 동사화 명사구를 절로

영어에서는 논항을 하위범주로 가지는 동사화 명사가 명사구로 사용하는 것이 일반적이거나 한국어에서는 동사화 명사구는 절로 풀어서 번역하는 것이 이해하는 데 더욱 일반적이다. 이러한 동사화 명사구를 절로 풀어서 사용하는 한국어의 언어적 스타일 생성 패턴은 다음과 같다:

[스타일 생성 패턴 : verbalnp-to-clause]

• 조 건

명사구가 동사화 명사구이면

• 실 행

동사화 명사구를 절로 스타일 변환

• 패 턴

2 <sup>9)</sup>는 두 언어간의 번역 차이(Translation Divergence)를 7가지로 나누어 설명하였다 : thematic divergence, promotional divergence, demotional divergence, structural divergence, conflational divergence, categorial divergence, lexical divergence.

{ [DET1 NOUN1 PRON1 NOUN2 : [sem\*\*[act]]]  
 -> PRON1 : [kcase : =[subj]] DET1 NOUN1 :  
 [kcase : =NOUN2.KFLEX2] VERB2 : [kroot : =  
 NOUN2.kvroot, kend : =[ㄴ것]] }

• 예

His failure in the examination was due to lack of time.  
 스타일 적용 전 : 그 시험에 그의 실패는 시간의 부족  
 때문이었습니다.

스타일 적용 후 : 그가 그 시험에 실패한 것은 시간의 부  
 족 때문이었습니다.

## 실 험

### 1. 스타일 생성 패턴 크기

한국어 스타일 생성 패턴은 현재 규칙이 아닌 패턴의 형  
 태로 기술되고 있으며 적용될 때마다 카운팅이 되고 있어  
 향후 사용 빈도가 매우 낮은 스타일 생성 패턴은 삭제할  
 예정이다.

스타일 생성 패턴은 최근에 그 모델이 설정되었기 때문  
 에 모든 가능한 언어현상에 대해 수동구축의 어려움으로  
 기존에 구축되어 있는 번역 패턴 중에 3회 이상 적용되고  
 있는 1,162개의 번역패턴을 가지고 반자동으로 스타일 생  
 성 패턴을 구축한 후 이를 사용빈도에 따라 재정리하고 있  
 다. 따라서 스타일 생성 패턴은 계속 검증되고 있으며 계  
 속해서 확장될 계획이다. 현재까지 구축된 스타일 생성 패  
 턴의 크기는 Table 2와 같다 :

### 2. 번역률 평가 방법

번역률 평가는 정확성(Accuracy) 평가와 스타일(Style)  
 평가 두가지 측면에서 이루어졌다. 정확성 평가는 원문의  
 내용이 얼마나 충실하게 대역문에 전달되었는가를 평가하  
 는 방법으로써 원문과 번역문을 비교하여 원문의 내용이  
 번역문에 올바르게 표현되었는가를 평가하는 방법이며, 스타  
 일 평가는 원문은 보지 않고 번역문만을 읽고 평가하는  
 방법으로써 한국어 모국어 화자의 입장에서 생성된 한국어

Table 2. 스타일 생성 패턴의 크기

스타일 그룹명	설 명	크기
Passive-to-active	수동태의 능동태화	420
Modaux-to-modaladverb	양상술어를 양태부사로	116
Help-to-can	보조동사 '돕다'를 가능성 양상보조사로	116
Causative-To-Intransitive	사역동사를 자동사로	116
Verbalnp-To-Clause	동사화 명사구를 절로	206
합 계		974

문장의 자연스러움을 평가하는 방법이다. 정확성 평가와  
 스타일 평가를 위한 평가 기준은 각각 Table 3과 같다 :

### 3. 번역률 실험

영한 번역 시스템의 번역률 평가를 위하여 100개의 샘  
 플문을 문장 길이별로 주로 CNN웹페이지, 토익 지문, 온  
 라인 쇼핑 관련 페이지 등으로 부터 무작위로 추출한 후  
 번역 결과를 스코어링하는 방법을 채택하였다.

전체 100문장은 20단어 이하의 문장 40개, 21~30단어  
 급의 문장 40개, 31단어 이상급의 문장 20개로 구성되어  
 있다. 100문장의 총 단어 수는 2,357개로서 문장 당 평균  
 약 23.6단어로 구성되었다.

문장 길이별 평균 단어수는 Table 5와 같다.

자동 번역결과는 2명의 언어학자에 의해 Blind Test 형  
 식으로 점수가 매겨지었다. 각 문장별로 정확성 평가와 스  
 타일 평가 기준에 따라 스코어링 되었다. 각 문장은 8점  
 만점으로서 각 문장의 점수를 합계하여 800으로 나눈 후  
 100을 곱한 값이 번역율로 결정되었다.

$$\text{번역률(\%)} = (\text{점수합산}/800) * 100$$

Table 3. 정확성 평가 기준

점수	설 명
4점	원문의 의미를 완전히 전달함
3점	사소한 1-2단어 정도를 제외하고 원문의 의미를 거의 전달함
2점	구 단위, 일부 절 단위 정도로 부분적으로 원문의 의미를 전달함
1점	단어 레벨 정도로만 원문의 의미를 전달함
0점	원문의 의미를 전혀 전달하지 못함

Table 4. 스타일 평가 기준

점수	설 명
4점	1번 읽고도 완전히 내용이 이해됨.
3점	내용은 대체적으로 쉽게 이해가 되나 어색한 어순 및 표현, 어색한 대역어가 4점짜리 보다 많음
2점	여러 번 읽고 한참을 생각해야 내용 파악이 어느 정도 가능함
1점	대부분 한국어 문장으로 되어 있으나 내용 파악이 거의 안 되는 경우
0점	대부분 원문 그대로 나와 내용 파악이 전혀 안됨.

Table 5. 평가 문장 통계

	총단어수	문장당 평균 단어수
20단어 이하 (40문장)	661	16.53
21-30단어 (40문장)	995	24.88
31단어 이상 (20문장)	701	35.05
전체 (100문장)	2,357	23.60
20단어 이하 (40문장)	661	16.53

Table 6. 점수별 번역률

	적용 전		적용 후	
	정확성	스타일	정확성	스타일
4점	10	1	10	1
3점	42	10	42	11
2점	41	50	41	50
1점	6	32	6	31
0점	1	7	1	7

Table 7. 단어수별 번역률

	적용 전		적용 후	
	정확성	스타일	정확성	스타일
20단어이하	68.75%	43.13%	68.75%	43.75%
21-30단어	61.88%	40.00%	61.88%	40.00%
31단어이상	56.25%	41.25%	56.25%	42.50%
전체	63.50%	41.50%	63.50%	42.00%

Table 8. 스타일 생성 패턴 적용 전·후 비교표

	적용 전	적용 후
총 점	420점	422점
전 체	52.50%	52.75%

각 문장의 평가 결과는 스타일 생성 패턴이 적용되기 전의 정확성과 스타일 평가 점수와 스타일 생성 패턴을 적용한 후의 정확성 및 스타일 평가점수를 대비하여 만들었다. 평가결과는 Table 6과 같이 측정되었다. 점수별 분포를 보면 다음과 같다:

단어수별 번역률은 Table 7과 같이 평가 되었다.

평가 결과 스타일 생성 패턴 적용 전과 스타일 생성 패턴 적용 후의 총점과 번역률은 각각 Table 8과 같이 측정되었다:

Table 6에 따르면 샘플문 100문장을 대상으로 스타일 생성 패턴이 적용되기 전보다 적용한 후에 정확성 점수 3점짜리 1문장과 1점짜리 1문장이 개선된 것을 발견할 수 있으며, 가중치에 따라 Table 7에서 스타일 번역률이 0.5% 향상되었고 이에 따라 Table 8에서는 전체적으로 0.25%의 정확성과 스타일 번역률이 향상되었다는 것을 관찰할 수 있었다.

## 결 론

본 논문에서는 영한 자동번역 시스템에 한국어 스타일

생성 패턴을 적용함으로써 영한 번역 품질을 향상하고자 하는 것이 목표였다.

이러한 목표는 기존의 원문에 대한 번역문의 정보 전달 정확성을 측정하는 1차원적인 번역률 평가 방법에서 벗어나 번역문의 정보 정확성 뿐만 아니라 자연스러움도 평가할 수 있는 2차원적인 번역률 평가방법으로써 정확성과 스타일을 동시에 평가하는 방법을 제안하였다.

2차원적인 번역률 평가 방법에 따라 스타일 생성 패턴이 적용되기 전과 적용된 후의 평가 결과는 100문장의 샘플문을 대상으로 하였을 때, 스타일 생성 패턴에 의해서만 0.5%의 번역률이 향상되는 것을 관찰하였다.

0.5%의 스타일 번역률 향상은 현재는 아주 미미한 번역률 향상으로 간주될 수 있으나, 스타일 생성 패턴의 구축량이 매우 소규모의 초기 구축량을 고려할 때 그 커버리지(Coverage)는 구축량이 증가함에 따라 점증적으로 높아질 것이며,<sup>10)</sup> 스타일 번역률도 한국어 스타일 생성 패턴의 증가에 따라 점증적으로 향상될 것으로 예측된다.

향후 더욱 향상된 스타일 번역률을 위한 계획은 본 논문에 소개된 단순한 언어간 스타일 생성 패턴의 적용에서 한 걸음 더 나아가 신문, 일기예보, 기술 매뉴얼과 같은 특정 그룹의 스타일 패턴을 적용하는 것이며, 또한 스타일 번역률의 향상을 위해 한국어 스타일 생성 패턴의 대응량 반자동 구축 방법의 개발일 것이다.

## REFERENCES

- 1) 이민행, 지광신, 정소우(1998) : "기계번역 시스템 측정 장치 연구". 언어와 정보, Volume 2, Number 2
- 2) 시정곤, 김원경, 고창수(2000) : "영-한 기계번역 성능 평가 방안 연구". 언어와 정보, Volume 4, Number 2
- 3) [http : //www.chunndong.or.kr/middroom/kslh/kor/style.htm](http://www.chunndong.or.kr/middroom/kslh/kor/style.htm)
- 4) John Hutchins W, Somers HL(1992) : *An Introduction to Machine Translation*, pp140-143
- 5) Chrysanne DiMarco, Graeme Hirst(1988) : *Stylistic Grammars in Language Translation*. COLING88, Budapest, pp148-153
- 6) 양승현(1997) : 영한 기계번역을 위한 언어 스타일의 변환 박사학위논문
- 7) Christian Fellbaum(1998) : *WordNet An Electronic Lexical Database*. MIT Press
- 8) Hong MP, Lee KY, Roh YH, Choi SK, Park SK(2003) : *Sentence Pattern-based MT revisited*. ICCPOL2003
- 9) Bonnie Dorr J(1993) : *Machine Translation : A View from the Lexicon*. The MIT Press, Cambridge
- 10) Choi SK, Kim TW, Yuh SH, Jung HM, Sim CM, Park SK(1999) : *English-to-Korean Web Translator : FromTo/Web-EK*, MT-Summit 1999