

번역 : 대응과 평가

장석진

서울대학교

Chang, Suk-Jin. 1998. Translation: Mapping and Evaluation. *Language and Information* 2.1, 1-41. Evaluation of multilingual translation fundamentally involves measurements of meaning equivalences between the formally mapped discourses/texts of SL (source language) and TL (target language) both represented by a metalanguage called IL (interlingua). Unlike a usual uni-directional MT (machine translation) model (e.g.: SL → analysis → transfer → generation → TL), a bi-directional (by 'negotiation') model (i.e.: SL → IL/S ↔ IL ↔ IL/T ← TL) is proposed here for the purpose of evaluating multilingual, not merely bilingual, translation. The IL, as conceived of in this study, is an English-based predicate logic represented in the framework of MRS (minimal recursion semantics), an MT-oriented off-shoot of HPSG (Head-driven Phrase Structure Grammar). In addition, a list of semantic and pragmatic checkpoints are set up, some being optional depending on the kind and use of the translation, so as to have the evaluation of translation fine-grained by computing matching or mismatching of such checkpoints. (Seoul National University)

¹이 논문은 1996년도 한국학술진흥재단의 외국석학과의 공동연구과제(연구책임자: 이기용) 연구에 의하여 연구되었으며, 한국정보학회에서 발표(조선대학교; 1997. 11.15)한 것을 수정 보완하였음.

1. 들어가기

번역은 원어(source language, SL)의 의미와 번역 대상어(target language, TL)의 의미와의 등치 여부에 따라 평가된다.² 이 글은 최소귀환의미론(minimal recursion semantics, MRS)에 입각한 SL-TL의 중간 언어(interlingua, IL) 표상을 의미 항목별로 대비하여 번역을 객관적으로 평가하는 방안을 제시하는 시도이다. 의미는 기본적으로 진리조건(truth condition)적 의미를 말하며, 부수적으로 담화에서의 화행(speech act), 전제(presupposition), 주제(topic), 초점(focus) 등의 화용적 의미를 포함한다. 번역은 종류와 용도에 따라 설정되는 의미 항목들이 조절되며 이들 항목의 상대적 비중에 따라 배점이 달라진다. 여기서 구상하는 메타 언어인 IL은 영어 어휘에 기반을 둔 술어 논리(predicate logic)이다.

이 IL은 단일 방향의 기계 번역(MT) 과정(예: SL → 분석(analysis) → 전이(transfer) → 생성(generation) → TL)과는 달리, SL과 주어진 번역 TL을 각각 IL로 표상하여 ‘협상(negotiation)’을 통하여 양방향으로 오가며 공동 IL을 모색하는 과정(즉, SL → IL/S ↔ IL ↔ IL/T ← TL)이다. 이 번역 모형의 바탕에는 통사·의미·화용의 세 부분의 균형 있는 기술을 시도하는 ‘통합문법’(장석진 1994)이 있고, 담화 지향적인 이 문법의 의미·화용 부문에 MT 평가를 위해 MRS를 도입하여 SL-TL의 IL 표상을 의미 항목별로 대비하여 번역의 정/오-역을 검색한다. 여기서 말하는 MT의 평가란 흔히 말하는 ‘MT 시스템의 평가’(Arnold의 1993, Minnis 1994, 이민행외 1997)가 아니라, 효율적이고 경제적인 MT 시스템과 이런 MT 시스템을 이용하여 출력된 번역이 있음을

²번역은 전달하는 정보가 없어지기도 하고 생기기도 하는 절차이며 따라서 번역은 의미 보존적이 아니라는 견해도 있으나(Kay외 1994:85), 의미 보존성을 번역평가의 기본 수칙으로 삼는다.

³통합문법은 정보기반적(information-based)이고 제약기반적(constraint-based) 문법인 HPSG(Pollard and Sag 1994, Sag and Wasow 1997)를 수용하고 담화 분석을 시도하는 문법이론이다.

전제로 하여 번역된 결과를 객관적으로 점수화하여 평가함을 말한다.

이 글은 다음 차례—2절에서 담화/텍스트의 구성 성분간의 형식적 대응 관계, 3절에서 IL 표상과 MRS의 소개, 4절에서 번역 평가의 기준, 5절에서 번역의 해례—로 기술된다.

2. 번역의 형식적 대응

번역의 대상은 구어체 담화(discourse, D)나 문어체 텍스트(text)이며 한 담화/텍스트는 화단—담화의 단락(discourse segment, D S)—으로 구성된다.⁵ DS는 하나 또는 그 이상의 문(S)으로 구성되고, S는 하위 구성 성분들—절(clause, C), 구(phrase, P), 단어(word, W), 형태소(morpheme, M)의 총위로 나뉜다. 번역에 있어서의 이들 총위간의 형식적 대응 관계는 대충 (1)과 같고,⁶ (2)는 그 예시이다.

(1) 형식적 대응

[M(형태소), W(단어), P(구), C(절), S(문), DS(화단), ↔ (대응)]

SL ↔ TL

- a. i. MM
- ii. MW
- iii. MP
- iv. MC/S
- b. i. WW
- ii. WM
- iii. WP
- iv. WC/S

⁵ 구어와 문어를 구분하지 않고 전자에 쓰인 ‘담화(discourse)’를 양자를 포괄하는 말로 쓰기로 한다.

⁶ 절(C)과 문(S)의 대응은 정밀하게 구분하지 않고 C/S 로 묶어서 제시한다.

- c. i. PP
- ii. PM
- iii. PW
- iv. PC/S
- d. i. C/SC/S
- ii. C/SM
- iii. C/SW
- iv. C/SP
- v. C/SDS
- e. i. DS DS
- ii. DSC/S

(2) 형식적 대응의 예

[E(nglish), G(erman), J(apanese), K(orean)]

	SL	↔	TL
a. i.	M E: <i>post-(war)</i> E: <i>-ly (honestly)</i>		M K: (전)후 K: -계 (정직하게)
ii.	M K: -겠- (하-겠-다)		W E: <i>will (will do)</i>
iii.	M K: -자- (가-자)		P E: <i>let's (let's go)</i>
iv.	M K: -더- (하-더-라)		C E: <i>I recall</i>
b. i.	W E: <i>dog</i> J: <i>inu</i> E: <i>liberty/freedom</i> J: <i>ziyuu</i>		W K: 개 G: <i>Hund</i> K: 자유 G: <i>Freiheit</i>
ii.	W E: <i>will</i>		M K: -겠-
iii.	W E: <i>shoot</i> K: 약속하다 K: 광복절		P K: 총(을) 쏘다 E: <i>give one's word</i> E: (The Korean) <i>Independence Day</i>
iv.	W E: <i>hungry</i>		C K: 배(가) 고프다
c. i.	P E: <i>from 7 to 11</i>		K: 7시부터 11시까지

		J: 7 zikara 11 zimadeii.
ii. P E:	<i>Let's</i>	M K: -자-
iii. P K:	총(을) 쏘다	W E: shoot
	E: <i>kick the bucket</i>	K: 죽다
iv. P E:	<i>become hungry</i>	C K: 배가 고파지다
d. i. C E:	<i>what I need</i>	C K: 내가 필요한 것
	E: <i>Harry met Sally.</i>	K: 해리가 샬리를 만났다.
ii. C E:	<i>I recall . . .</i>	M K: -더-
iii. C K:	태어난 곳	W E: hometown
		G: <i>Heimat</i>
iv. C K:	내가 다음 멈춘 곳	P K: my next stop
v. S E:	<i>Though he was tired,</i>	DS K: 존은 힘이 들었다.
	<i>John stayed up all night.</i>	그러나 밤새 깨어있었다.
e. i. DS E:	<i>He came home. And he</i>	DS K: 그이는 집으로 돌아왔
	<i>lived there until he died.</i>	어요. 그리고 거기서
		죽을 때까지 살았어요.
ii. DS E:	<i>He came home. And he</i>	DS K: 그이는 집으로 돌아와
	<i>lived there until he died.</i>	서 거기서 죽을 때까지
		살았어요.

위에서 i-형은 SL-TL 간의 동일 층위끼리 대응하는 형이며 이 형이 번역에서 선호된다. 특히 단어(W) 층위에서 1-대-1 대응이 가장 선호되고 동일 계통/어족의 언어들 사이(예: 불어-이탈리아어-스페인어)나 동일 유형의 언어들 사이(예: 한국어-일어)에서는 비교적 높은 1-대-1 의 대응 비율을 나타내 효율성 있는 MT의 정렬(alignment) 시스템의 개발이 기대된다(최기선 1996, 이주호의 1998). 그러나 동일 계통, 동일 유형이 아닌 언어들 사이(예: 한국어-영어)에서는 단어 층위에서도 1-대-1 대응이란, 전문 학술 분야의 용어를 제외하고는, 극히 드물고 다-대-다 대응이 보통이다.⁴ ii-iv

⁴한국어(K)와 일어(J)는 영어(E)와 다음과 같은 어순의 차로 달리 유형화된다.

형은 SL의 문법 범주가 TL의 다른 범주로 대응하는 형이며 P 층위에서 관용어구, C/S 층위에서는 능동태, 피동태의 대응이 주목된다. DS 층위에서는 담화의 연결(S+S+...) 관계의 대응 여부가 주의를 요한다(3.4절 참조).

어휘 층위에서의 적절한 역어 선정의 문제도 크지만, 동일한 의미 정보가 SL과 TL에서 각각 다른 형식적 층위로 나타나는 대응의 구조적 다기화(divergent) 현상을 체계적으로 유형화하는 문제가 IL 방식을 추구하는 MT가 봉착하는 난제이다. 개별 언어들의 용언(동사, 형용사)의 하위범주화에 의한 상세한 문형(sentence pattern) 분류가 되어 있음을 전제해야 한다. 이를테면, IL을 지향하는 MT에서 Dorr(1994)은 다음 7 가지의 다기화 유형을 제시하고 있다.⁷

(3) 구조적 다기화 유형(Dorr 1994:598)

[1] Thematic(테마적): 테마어가 달리 배열⁸

	SL	↔	TL
K/J	E		
S-O-V	S-V-O		K/J에서는 동사(V)가 머리(head)로서 문말에, E에서는 주어(S) 다음에,
NP-P	P-NP		K/J에서는 후치사(P)가 NP 뒤에, E에서는 전치사 (P)가 NP 앞에,
V-v	v-V		K/J에서는 조동사(v)가 본동사(V) 뒤에, E에서는 조동사가 본동사 앞에,
RelCL-N	N-RelCL		K/J에서는 관형절(RelCl)이 N 앞에, E에서는 관계절(ReLCL)이 N 뒤에.

⁷영어 문형 분류로는 Longman 사전(Longman Dictionary of Contemporary English), Hornby사전(Oxford Advanced Dictionary of Current English), 한국어의 동사 유형 분류로는 홍재성의 1997 ([한국어 동사 구문 사전]이 유용하다. HPSG식 한국어 용언 하위범주화에 대해서는 신효필(1990), 장석진(1993) 참조.

⁸이 7 가지의 대응 유형은 영어, 독일, 스페인어 등 서구어에서 추출한 것으로 한국어, 일어와 같이 유형이 다른 언어에서는 유형이 더 다기화한다. 여기서는 Dorr(1994)의 분류명과 SL인 영어와 대응하는 독일, 스페인어의 예를 제시하고, 덧붙여 한국어, 일어를 추가하여 유형의 동일 여부를 밝혔다.

⁹테마어인 Mary가 E에서는 목적어, S에서는 주어로. K, E에서는 E와 동일하나 목적격 조사 '를' 외에 주격 조사 '가'를 취할 수도 있다.

J: Watakusi-wa onaka-ga suitea.

[7] Lexical(어휘적): 다른 유형에 부수적으로 일어나는 유형

E: John broke into the room. ↔ S: Juan forzo la entrada al cuarto.
room. 'John forced (the) entry to the room.'

K: 존이 방으로 쳐들어왔다.

J: John-ga heya-ni rannyusita.

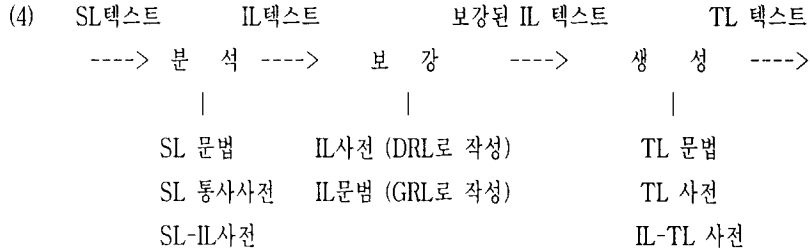
Dorr(1994)의 번역 다기화 유형 분류는 Jackendoff(1983, 1990)의 어휘개념구조(lexical conceptual structure, LCS)를 근간으로 삼았고, 이 어휘-의미적 IL 표상과 이를 결합 확장한 중간언어 GLR (generalized linking routine)과 CSL(canonical syntactic realization)을 표층의 형태로 사상한다. 이 사상에는 7 가지 유형이 GLR과 CSL의 예외 항목으로 기재되어 이들 유형의 대응 문제에 대한 해결을 시도한다. 이 시도는 어휘적 개념구조에 입각한 방법으로서 IL 표상에 이 유형 정보가 SL 사전의 해당 어휘에 기재됨으로써 TL과의 형식적 대응을 유형화하는 데에 유용하다.

3. IL 표상과 MRS

3.1. IL의 모형

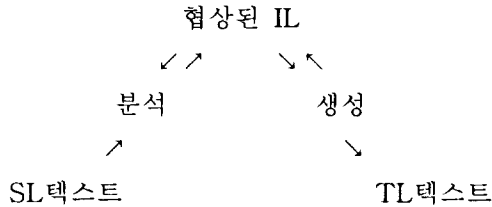
중간언어(IL)를 개재하는 번역"은 SL을 분석하여 IL을 구축하고 TL을 생성하는 모형(즉, SL->IL->TL)이 일반적이다. 이를테면, Nierenburg외(1987:91)는 다음과 같은 모형을 제시하고 있다.

"IL 방식의 이점중의 하나는 동일한 IL 표상이 수 많은 언어의 MT뿐 아니라 여러 종류의 자연 언어처리의 응용될 수 있기 때문에 MT 시스템 구성에 드는 비용과 노력이 절감된다는 점이다(Kay외 1994:93).



위 모형에서 분석(analysis)과 생성(synthesis/generation)에 보강(augmentation)된 IL은 사전표상언어(Dictionary Representation Language, DRL)와 문법표상언어(Grammar Representation Language, GRL)를 갖추고 있다. 이런 IL 모형에서 분석부의 IL과 생성부의 IL의 언어별 차이를 조정하는 것이 Kay의 협상(negotiation)식 MT 이론이다.

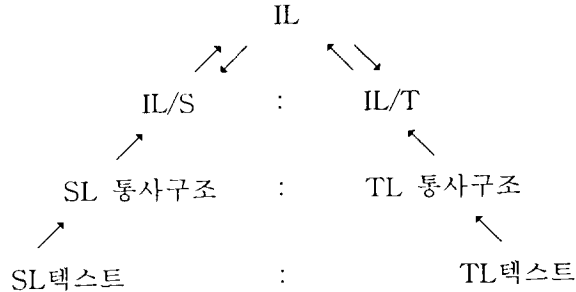
(5) Kay의 협상 모형



이 모형에서는 IL로 표상된 SL-TL의 번역이 Dorr식 다기화 유형이나 오역의 문제 등을 SL의 분석과 TL의 생성을 오가며 조정 타협하고, 필요하면 선행하는 담화/텍스트 기록을 검색하여 IL 표상에서 생기는 의미의 모호성과 중의성 문제를 해결한다.

이 글은 기존의 번역된 담화를 놓고 그 번역의 정/오-역을 검증하고 평가하는 입장이므로 SL과 TL이 동일한 통사이론과 의미론의 틀에서 이루어지고 의미구조는 MRS이론을 따른 IL로 각각 표상되고, 협상되고 두 IL 표상—IL/S와 IL/T—의 대응을 대비하는 번역 평가 모형으로서 다음과 같이 도식할 수 있다.

(6) 번역 평가를 위한 IL 모형



이 모형은 SL에서 TL로 단일 방향(즉, 분석 → 생성)으로 가는 Nierenberg나 Kay의 모형과는 달리 이미 번역된 TL을 SL과 대조하는 입장이어서 분석과 생성의 구분이 없어진다. 관련 언어들의 텍스트는 의미가 중간언어로 번역된 표상 IL/S와 IL/T는 협상을 통해 조정되고(쌍방향(\rightleftharpoons))으로 표시), IL/S와 IL/T의 동치 여부가 대비된다(IL/S : IL/T). 이를 위해, 개별 언어의 사전에는 Longman 사전식 용언의 문형 분류 정보, WordNet나 이에 바탕을 둔 CoreLex식 명사의 의미 분류 정보,¹⁾ 내지는 Dorr식 구조적 다기화 유형의 정보가 적절히 등재되어 있어야 하고, 이런 개별 언어의 사전에 바탕을 둔 번역 평가를 위한 IL 사전이 편찬되어야 한다."

¹⁾NLP에서 널리 알려진 WordNet는 어의(word sense)의 유의(synonymy) 집합(synset)에 입각해서 11 가지의 최상위 유형(top type)과 이를 포함한 39 가지의 기본 유형으로 구축된 데이터베이스이며, 이를 확장하여 단어의 다의성(polysemy)을 미명시(underspecify)한 유형으로 포착하려는 CoreLex(Core Lexical Engine) (Buitelaar 1998) 등이 있다.

²⁾사전의 어휘 항목은 자질체계에 바탕을 둔 AVM(attribute-value matrix)식으로 기재한다. (SUBCAT를 ARG-ST(argument structure)로 대체, SYNSEM은 SYN과 SEM으로 분리; SEM은 화용적 의미를 포함하는 SEMPR(semantic-pragmatic)라 부르고, 새로운 속성 SIT(uation)(상황)을 추가한다. 용언의 유형(type) 정보는 'dtv (ditransitive verb, 이중타동사)'는 용언의 유형(type) 계층에 따라 상위 정보를 승계한다. 세분된 동사 분류 정보 (<give>: D1(Longman사전); <주다>: V10(장석진(1993)) 번역 평가 항에 기재된다.

3.2. 중간언어의 특성과 구현

여기서 표방하는 IL은 자연 언어인 영어 어휘와 그 어의(word sense)를 바탕으로 하는 논리어이며 이를 구축함에 있어서 다음과 같은 사항이 요망된다(Kay의 1994, Copestake의(1996) 참조).

- (7) IL의 일반 특성
 - a. 추론(inference) 가능한 의미론이어야 한다.
 - b. 술어-논항(predicate-argument) 관계를 명시할 수 있어야 한다.
 - c. 명제내용 외에 화용적 정보—신구(new-given), 주제(topic), 초점(focus) 등—를 전달할 수 있어야 한다.
 - d. 개념 해체(decomposition)는 절제해야 한다.
 - e. 양화의 작용권을 미명시(underspecify) 할 수 있어야 한다.

영어 어휘를 근간으로 하여 IL—즉, 추상적인 술어논리(predicate logic)언어—을 구축할 때 술어나 논항은 실제 영어 어휘이므로 IL의 보편성을 위해 영어 고유의 어원적, 민속적, 문화적 의의를 되도록 배제한다. 다음에 IL 구축에 있어서 생기는 몇 가지 문제를 예

(i) E: <give, [dtv	K: <주다, [dtv
ARG-S <[i, [j, [k]>	ARG-S <[i, [j, [k>
SEMPRA [INDEX s	SEMPRA [INDEX s
RESTR <[RELN give	RESTR <[RELN give
SIT s	SIT s
GIVER i	GIVER i
GIVEE j	GIVEE j
GIFT k]	GIFT k]>]]>

위 예시에서 <E: give/K: 주다>는 ARG-ST, SEMPRA의 값이 각각 같다. 이 SEMPRA는 다음 MRS식 IL 표상에서 mrs-str 유형으로 재정립되어 손잡이(handle), 지표(index) 등 자질을 가진다. MRS에서 의미 논항 GIVER, GIVEE, GIFT는 대신 ARG1, ARG2, ARG3으로 표기하고, 차례로 주어(SBJ), 간접목적어(OBJ2), 직접목적어(OBJ)임을 상징하는 방식을 쓴다.

시한다. 어휘의 개념(concept) 표상에서 개념의 해체는 최소화하며, Jackendoff-Dorr식 원초적 의미 요소로서 행위(DO), 사동(CAUSE), 기동/변성(BECOME), 존재/상태(BE)를 도입하지만 생성의미론(generative semantics)에서 시도하던 방식(McCawley 1971) (예 : 영어 *kill*을 'cause(x, become(not(alive x)))'로 해체)을 지양하여 다음과 같이 영어 단어를 (대문자로) 써서—즉, ALIVE, DEAD, DIE, KILL—IL 표상으로 한다.

(8) 어휘와 개념 해체/합성

개념 구성	E	K	J	IL
CAUSE(y, _) ⇒	kill	죽이다	korosu ⇒	KILL(y,x)
/				
BECOME(_) ⇒	die	죽다	sinu ⇒	DIE(x)
/				
NOT(_) ⇒	dead (살아 있지 않다)		(ikiteinai) ⇒	DEAD(x)
/				
ALIVE(x) ⇒	alive(살아 있다)		(ikiteiru) ⇒	ALIVE(x)

개별 언어 고유의 관용적 표현은 비관용적이고 단순한 개념으로 표상한다. Dorr의 대응 유형(유형-6)에서 예시한 한국어의 '배(가) 고프다'(또는 '목(이) 마르다')는 주어-술어로 구성된 단문 구조이지만(일어도 마찬가지로 : 'hara-ga heru') 영어의 'be hungry', 독일어의 'haben Hunger'는 술부만으로 나타나고 술부 구성에 차이가 있다. 이럴 경우에 IL 표상은 '배고픈' 상태(state)라는 일항 술어 'HUNGRY'로 한다.

(9) K: x-가 배(가) 고프다.

J: x-ga hara-ga heru.

E: x is hungry.

G. x haben Hunger.

IL: HUNGRY(x)

관용어구의 경우는 비관용적인 정상 표현을 IL 표상으로 한다. 경동사(light verb) 구문은 동작성 명사를 본동사로 하는 단순 구문으로 표상한다.

(10) 관용 표현

관용	비관용	IL
a. E: kick the bucket	die	DIE(x)
b. E: go to bed	sleep	SLEEP(x)
c. K: 깡통(을) 차다	가난해지다	BECOME-POOR(x)

(11) 경동사 구문

경동사	일반동사	IL
a. take a look	look	LOOK(x)
b. have a drink	drink	DRINK(x)
c. make an attempt	attempt	TRY(x, y) ¹⁴
d. give y a kick	kick y	KICK(x, y)
e. 공부를 하다	공부하다	STUDY(x)
f. 장가를 가다 ¹⁵	장가가다	MARRY(x) ^ MAN(x)

능동-피동이 대응하는 표현에서는 능동 표현을 기본으로 하되, 보통 주어에 수반하는 주제성(topicality)은 담화구조의 주제-초점(topic-focus) 정보의 추가로 변별한다. 또한 표층구조의 동일성 유지를 번역 평가 기준에 넣되 부자연스러운 표현보다는 번역의 자연성이 선호된다.¹⁶

¹⁴<E: attempt/try>에서 try가 보다 기본적 어휘라는 관점에서 TRY를 IL 동사로 사용.

¹⁵ '장가(를)가다,' '시집(을)가다'는 관용구 처리를 하지만 '가다'는 경동사로 본다. '시집에 가다'에서 '가다'는 본동사 구실을 한다.

¹⁶ 'John hit Bill.'에서 'hit' ('때렸다')의 피동 'was hit'에 대응하는 한국어의 어휘적 피동이 없으므로—*매려졌다, ?매림을 당했다/받았다—대신 피동의 의미를 가진 '맞았다'가 자연스러운 번역이 된다. 능동문에서는 '때렸다'고 피동문에서는 '(매를)맞았다'. '맞다'는 목적어인 '매를'이 임의로 내장(incorporate)된 경우이며 본래의 주어가 필수적인 앞에서 본 '배(가)고프다' 구문과 대조된다. 'be hit'와 '맞다'는 Dorr의 어휘

(12) 피동-능동 구문과 IL

- | E | | K |
|--|---|------------------|
| i. John was hit by Bill. | → | 존이 빌에게 (매를) 맞았다. |
| \ | | / |
| IL: HIT(s1, Bill, John) ∧ PAST(s1) ∧ TOPIC(s1, John) | | |
| ii. Bill hit John. | → | 빌이 존을 때렸다. |
| \ | | / |
| IL: HIT(s1, Bill, John) ∧ PAST(s1) ∧ TOPIC(s1, Bill) | | |

SL-TL 간의 어휘의 공백은, IL 설정에서 영어 어휘를 기준으로 하되 필요에 따라—특히 영어 어휘가 없거나, 미분화(예 : bank: 1. 은행, 2. 둑)된 경우—대상어의 어휘를 도입하고 지표를 붙여 (예 : BANK1, BANK2,...) 매꾼다. 한국어 ‘보다’와 이에 대응하는 일본어 ‘miru’는 다음 예에서 보듯이 영어에서는 보는 대상에 따라 적어도 ‘see’, ‘look’, ‘watch’로 대응된다.” 이 경우는 어휘적으로 분화된 영어를 본 따서 IL은 SEE, LOOK, WATCH 등으로 한다.

- | (13) | K | J | E | IL |
|------|--------|------------|----------------|----------------|
| | 보다 | miru | see/look/watch | SEE/LOOK/WATCH |
| a. | 영화를보다 | eiga-omiru | see the movie | SEE(x,y) |
| b. | 미아를 보다 | Mia-o miru | look at Mia | LOOK-AT(x,y) |
| c. | TV를 보다 | TV-o miru | watch TV | WATCH(x,y) |

문장의 의미구조는 약식으로는 선형 논리형태로, 정식으로는 아래 (19)에서와 같이 속성-값 행렬(AVM) 도식으로 나타낸다.

적 다기화의 경우다.

”한자는 영어처럼 어휘적으로 분화된다 : 見/視/觀/...

3.3. 최소귀환의미론(MRS)

MRS는 통합 연산(unification)에 바탕을 두고 HPSG의 문법이론을 수용한 의미론으로서 다음과 같은 특성을 지니고 있다 (Copestake외 1996).

(14) MRS의 특성

- a. 자질(feature)과 유형(type)에 기반을 둔 HPSG 문법 이론의 틀을 따른다.
- b. '손잡이(handle)와 사건(event)—여기서는 상황(situation)—을 자질 구조에 도입하여 평탄(flat)한 논리 구조를 가진다.
- c. 양화의 작용권이나 전치사구 수식 구조의 중의성을 미명시(under-specify) 할 수 있다.
- d. 관계(relation)의 유형(type)에 의해 어휘적 중의성을 해소한다.

다음 (15)의 예문은 전치사구(PP)의 작용권에 따라 (16)의 a, b 와 같이 내포 구조를 달리 표상할 수 있으나 이런 피상적 중의문을 (16c)와 같이 상황(s)를 도입하여 평탄하게 나타낸다.

(15) On Monday Kim ran in Green Park.

(16) a. on(in(run(Kim), Green-Park), Monday)

b. in(on(run(Kim), Monday), Green-Park)

c. on(s1, Monday) \wedge run(s1, Kim) \wedge in(s1, Green-Park)

(17)과 같은 양화 표현은 양화의 작용권을—(18a)의 n, m 처럼—명시 안 할 수도 있고, 전칭과 존재 양화의 영역의 차를—(18b,c) 처럼—손잡이 4, 1을 써서 각각 달리 명시할 수도 있다.

(17) Every dog chases some cat.

(18) a. top p

1: every(x,3,n), 3: dog(x), 7: cat(y), 5: some(y,7,m),

- 4: chase(s1,x,y) - 작용권 미명시
- b. top 5
 1: every(x,3,4), 3: dog(x), 7: cat(y). 5: some(y,7,1),
 4: chase(s1,x,y) - 광역: some
- c. top 1
 1: every(x,3,5), 3: dog(x), 7: cat(y). 5: some(y,7,4),
 4: chase(s1,x,y) - 광역: every

MRS는 정식으로는 자질구조를 갖는 AVM 도식으로 나타낸다. HPSG 문법 틀에서 MRS는 mrs-struct의 유형으로 설정되어 HANDEL, INDEX, LISZT의 자질로 구성되고 각각 손잡이(handle: h1, h2,...), 지사체(referent: s1, s2; x1, x2; t1, t2;...), 리스트(list: <...>)를 값으로 취한다.⁸ 여기서는 표기의 간결을 위해 각각 HAN, IND, LIS로 약기한다.

(19) MRS: 유형과 자질 구성

[mrs-struct
 HAN: handle
 IND: ref
 LIS: <list>]

리스트의 값은 평탄한 관계(relation, REL)들이고 이 관계들은 다 손잡이(HAN) 자질을 갖는다.⁹ 양화사 every-관계는 세 가지 자질—BV(bound variable, 결속 변항), RESTR (restriction, 제약)과 SCOPE(작용권)—로 구성된다. 위 (17)에서 some과 every의 작용권의 차이—(18b)와 (18c)—는 (20b'), (20c')와 같이 첫 최상위의 손잡

⁸HANDEL, LISZT는 작곡가의 이름이지만 'handle', 'list' 대신 멍을 부린 것이다. LISZT의 값은 리스트(list)이지만 실은 집합(set) 구실을 한다.

⁹관계는 유형(type)의 계층 체계에 입각한 유형으로 기술하는 것이 선호되지만 (Carpenter 1994의 ALE, Pollard & Sag 1994의 9장), 여기서는 자질 REL과 그 값의 쌍으로 표기한다: 즉, 'every_rel', 'chase_rel' 대신 [REL: every], [REL: chase].

이의 차 h5(some-관계)와 h1(every-관계)과 그 안에서의 작용권의 손잡이의 차로 파악된다.

- (20) a. Every dog chases some cat. (17a)
- | | |
|-------------------------|--------------------------|
| b. some이 광역인 경우(=(18b)) | c. every가 광역인 경우(=(18c)) |
| b'. [HAN: h5 | c'. [HAN: h1 |
| IND: s1 | IND: s1 |
| LIS: <[REL: every | LIS: <[REL: every |
| HAN: h1 | HAN: h1 |
| BV: x2 | BV: x2 |
| RESTR: h3 | RESTR: h3 |
| SCOPE: h4], | SCOPE: h5], |
| [REL: dog | [REL: dog |
| HAN: h3 | HAN: h3 |
| INST: x2], | INST: x2], |
| [REL: some | [REL: some |
| HAN: h5 | HAN: h5 |
| BV: x6 | BV: x6 |
| RESTR: h7 | RESTR: h7 |
| SCOPE: h1], | SCOPE: h4], |
| [REL: cat | [REL: cat |
| HAN: h7 | HAN: h7 |
| INST: x6], | INST: x6], |
| [REL: chase | [REL: chase |
| HAN: h4 | HAN: h4 |
| SIT: s1 | SIT: e1 |
| ARG1: x2 | ARG1: x2 |
| ARG2: x6]>] | ARG2: x6]>] |

이제 앞(9)에서 예시한 ‘배고프다’의 중간언어 표상을 MRS식으로 기술하면 (22)와 같다. 여기서 개체의 변항 x 는 이름(철수, Taroo,

John 등)으로 지시될 때 명명(naming)-관계로 의미 기술에 추가된다.

(21) (=9))

K: x-가 배가 고프다.

J: x-ga hara-ga heru.

E: x is hungry.

G. x hat Hunger.

IL: HUNGRY(s, x)

(22) MRS 표기: (x가 John인 경우)

IL/K,J,E,G

a. [mrs_struct

HAN: h1

IND: s1

LIS: <[REL: hungry

HAN: h1

SIT: s1

ARG: x1],

[REL: naming

HAN: h1

NAMED: xi

NAME: John]>]

b. 약식

HUNGRY(s1, x1) \wedge NAMED(x1, John)

시간 관계는 화맥지표(contextual indices, C-INDS)의 하나인 발화시(utterance time, UT), 사건시(event time, ET), 참조시(reference time, RT) (Reichenbach 1947)와의 선행, 동시, 후행 등 시간관계로 나타낸다(장석진 1994 참조). 또한, 1,2 인칭은 C-INDS의 화자(speaker, SP), 청자(hearer, HR)의 값, 의미지표(IND)의 값

의 공지시로 표상한다. 다음 (23)은 과거 시제와 1인칭의 예로 (24a)는 MRS 도식의 IL 표상, (24b)는 약식 표기이다.*

(23) a. SL/E: I went happily to school.

b. TL/K: 나는 즐겁게 학교에 갔다.

(24) a. MRS 도식*

[HAN: h1

IND: s1 (ET: t2, RT: t2)

LIS: <[REL: go

*AVM 도식에서는 상황(SIT)을 자질명으로 사용하고 사건(e1,...)과 상태(s1,...)를 구별한다. Jackendoff(1983, 1990), Dorr(1994)의 어휘개념구조(lexical conceptual structure, LCS)는 개념을 원초적 의미 요소로 해체하여 예문 (23)은 다음과 같은 LCS로 분석된다.

i. John went happily to school.

[Event GOLoc

([Thing JOHN],

[Path TOLoc

([Position ATLoc ([Thing JOHN], [Location SCHOOL])])])

[Manner HAPPILY]

Sowa(1984, 1991)의 개념구조론에 따르면 이 문은 다음과 같이 도시된다.

ii. (PAST) → [SITUATION: [PERSON: John] ← (AGNT) ← [GO] → (DIRECT) → [SCHOOL]

즉, 과거(PAST)에 이런 상황(SITUATION)—GO라는 행위, 행위자(AGNT)는 John, 방향(DIRECT)은 SCHOOL, 양태(MANR—manner)는 HAPPY. 위에서 [] 안에 있는 SITUATION, PERSON, GO, SCHOOL, HAPPY는 개념이고 () 안에 있는 PAST, AGNT, DIRECT, MANR는 개념 관계라고 한다. Sowa의 개념구조 이론을 한국어 분석에 적용한 연구로는 충남대학교의 박사학위논문 5편을 바탕으로 하여 펴낸 서상욱(1996)이 있다.

*MRS 도식에서 시간 관계는 상황 변수(s1,...) 뒤 괄호 안에 'ET: t1, RT: t2' 등으로, 또 개체변수 (예: x1, x2)는 그 뒤 괄호 안에 인칭(1,2,3), 수(sg, pl), 성(m, f, n)을 약식으로 각각 표기한다. 이 방식은 TDL(Type Description Language for HPSG; Krieger외 1994 참조) 틀에서 현재 진행중인 Verbmobil 프로젝트인 영-독 번역에서 쓰이고 있다.

HAN: h1
 SIT: s1
 ARG: x1
 PATHTO: x2],
 [REL: happy
 HAN: h1
 MNN: m1],
 [REL: temp_precede
 HAN: h1
 ET: t2
 UT: t1],
 [REL: school
 HAN: h1
 INST: x2]>
 C-INDS: [SP: x1
 UT: t1]>]

b. 선형 IL: $GO(s1, x1, TO-x2) \wedge HAPPY(m1) \wedge$
 $T-PRECEDE(et2, ut1) \wedge SPEAKER(x1)$
 $\wedge SCHOOL(x2)$

3.4. 담화의 연결과 IL

번역에서는 문 단위의 명제, 지시체, 시제 등의 번역 못지 않게 문과 문의 연결 관계(discourse linking relation, DLR)의 정확한 번역이 요구된다. 이 DLR은 그 수가 한정적이고 (29)에 열거한 것들이 대표적인 DLR이다.

(25) 담화 연결 관계(DLR)

관계	IL
[1] 연결	AND(s1, s2)

[2]	계기	AND-THEN(s1, s2)
[3]	이접(포괄적)	OR(s1, s2)
[4]	이접(배타적)	ORex(s1, s2)
[5]	동시	WHILE(s1, s2)
[6]	선행시	BEFORE(s1, s2) (또는, T(EMP)-PRECEDE(s1, s2))
[7]	후행시	AFTER(s1, s2) (또는, T(EMP)-FOLLOW(s1, s2))
[8]	대조	BUT(s1, s2)
[9]	양보	THOUGH(s1, s2)
[10]	원인	BECAUSE(s1, s2)
[11]	이유	WHY(s1, s2)
[12]	결과	SO(s1, s2)
[13]	조건	IF(s1, s2)
[14]	비교(동등)	AS-AS(s1, s2)
[15]	비교(더)	MORE(s1, s2)
[16]	비교(덜)	LESS(s1, s2)

다음은 위 [1]의 and-연결 관계의 MRS 식 AVM이다.

(26) a. 선형 IL: AND(s1, s2) - 위 (29[1])

b.² HAN: hi

IND: s0 (ET: t1, RT: t2)

LIS: <[REL: and

HANS: c1 (h2, h3)

INDS: c2 (s1, s2)]>

²동접(coordination)의 경우 HAN과 IND는 복수(HAN(DEL)S, IND(ICE)S)로 하여 그 값은 임의의 c-변항—c1, c2,...을 매기고 그 구성은 각각 다중치 h1, h2,...와 s1, s2, ...로 표기한다.

4. 번역 평가

번역의 평가는 의미의 등치 여부의 정도를 따지는 것이며 서두에서 밝혔듯이 의미에는 명제 내용의 진리조건적 의미(truth-conditional meaning)와 이밖에 담화의 참여자를 포함한 화맥, 배경에 관한 화용적 의미(pragmatic meaning)가 있다. 다음 항목 중에서 [1]-[4]는 명제내용의 의미, [5]-[10]은 화용적 의미에 속한다. 후자는 수의적 항목으로 번역의 장르, 목적, 용도 등에 따라 달라진다. 경어체계가 확립되어 있는 한국어/일어에 있어서는 담화의 화계가 형태적으로 실현되는 문계(speech level, SL)의 층위, 존대(HON) 여부 등 항목이 필요하다.²⁾ 여기서 세분된 평가 측정 항목을 바탕으로 하여 번역의 분야, 목적, 용도 등에 따라 필수와 수의 항목으로 분리하고 조절하는 신축성 있는 번역 항목을 설정한다. 이밖에 SL의 문이 TL에 번역되지 않고 빠진(missing) 경우는 완전 실점이다. 또한, 주어진 TL은 그 자체가 문법적인 정문임을 전제한다.

(27) 평가 항목과 지침

[X]는 필수 항목, []는 임의 항목

항목	감점 (100-)
[X] 1 SITUation (명제 상황)	(50)
1-1 RELation (관계)	(30)
1-2 REF-INDices (지시체)	(20)
1-3 POLarity (긍정/부정의 극성)	(50)
[X] 2 MODification (수식)	(20)
2-1 V-MOD (용언 수식)	(10)
2-2 N-MOD (체언 수식)	(10)
2-3 S-MOD (문 수식)	(10)
[X] 3 TAM (Tense, Aspect, Mood, 시·상·태)	(20)

²⁾한국어의 문계(SL), 문형(ST)의 분류와 명칭에 대해서는 장석진(1994) 참고.

	3-1	TNS (Tense)	(10)
	3-2	ASP (Aspect)	(10)
	3-3	MD (Mood)	(10)
[X]	4	DLR (Discourse Linking Relation, 담화연결관계)	(20)
[]	5	QS (Quantification Scope, 양화영역)	(10)
[]	6	SA (Speech Act, 화행)	(10)
[]	7	DF (Discourse Function, 담화기능)	(20)
	7-1	TOPic (주제)	(10)
	7-2	FOCus(초점)	(10)
[]	8	BKGR (Background, 배경 화맥)	(10)
[]	9	HON (orification, 존대)	(20)
[]	10	STYle (문체)	(10)

채점 지침:

- 번역의 최소 평가 단위는 화단(DS)이다. 화단을 구성하는 문(S) 번역에 기본점으로—만점을 가정하여—100점을 부여한다.
- 위 평가 항목에 따라 오역에 대해서 이 100 만점에서 감점한다.
- 소항목의 합계(예: 1-1, 1-2, 1-3)는 상위 항목(예: 1)의 배당 점수(예: 50점)를 초과하지 않는다.(예: 1에서 소항목이 다 오역인 경우에도 50점 이상 감점하지 않는다.)
- 문의 평균치가 DS의 평점이고, DS의 평균치가 담화의 평점을 얻는다.

이렇게 해서 산정된 담화/텍스트의 평점은 다음과 같은 등급으로 종합 평가할 수 있다.

(28) 번역의 종합 평가 (예시).

수(Excellent)	95-100 점
우(Very Good)	85-94
미(Good)	70-84
양(Fair)	69-40

불가(Bad)

39-0

번역 평가를 위한 참고 자료로 다음 (29)과 같은 형태통사적 정보—문(S)의 단어 수, 화단(DS) 수, 문절(S/CL) 유형, 문형(ST)의 유형, 용언의 하위범주화 등—를 대비한다. IL 표상을 위해 SL의 관용어구는 이에 대응하는 비관용 어구를 제공하여 이에 대한 TL의 역어의 선정과 IL 표상을 돕는다.

(29) SL-TL의 구조적 대응

[ST = sentence type, SL = sentence level, CT = clause type, SPL = simple, CPX = complex, CPD = compound, CORD = co-ordinate, SP = sentence pattern, IDM = idiom, dcl = declarative, int = interrogative, imp = imperative, exl = exclamatory, frm = formal, pol = polite, blt = blunt, fml = familiar, pln = plain, int = intimate]

SL/E	TL/K
[1] D의 수: n ((D1, D2,...))	n ((D1,D2,...))
[2] DS의 수: n ((DS1, DS2,...))	n ((DS1, DS2,...))
[3] S의 수: n (S1, S2,...)	n ((S1, S2,...))
[4] W의 수: n (W1, W2, ...)	n ([W1, W2, ...])
[5] ST: {dcl, int, imp, exl}	{dcl, int, imp, prp, exl}
[6] SL: {...}	{frm, pol, blt, fml, pln, int}
[7] CT: {SPL, CPX, CPD, CORD}	{SPL, CPX, CPD, CORD}
[8] SP: (예) <i>give</i> : D1	주다: V10
[9] IDM: (예) <i>go to bed</i> => sleep	깡통을 차다 => 가난해지다

²⁴한국어의 문계(SL)에 대응하는 영어의 문계는 없지만. 문체(STYle)면에서 한국어나 영어는 담화/텍스트의 구어체(spoken), 문어체(written), 격식체(formal), 비격식체(informal) 등 유형으로 분류할 수 있어서 이 항목을 추가한다.

(i)	E	K
	STY: {spk, wrt, frm, infm,...}	STY: {spk, wrt, frm, infm,...}

(IL: SLEEP(x)) (IL: BECOME-POOR(x))
kick the bucket → *die* (IL: DIE(x))

번역문과 원문의 대조 평가를 쉽게 하기 위해 (27)의 평가 항목의 차례를 따라, 3.3.절에 도입한 MRS식 표상을 병렬로 수정 표기하여, 다음 절에 예제를 제시한다. 평가 항목이 담화 화용적 의미까지 포함하므로 통사(SYN)와 대응하는 의미·화용(SEMPRA)의 평가 항목은 다음과 같다.

(30) SEMPRA (예시)

- a. HAN: handle (h1, h2 ...)
 IND: ref (x1, t1,...)
 x1: person-ref
 t1: time-ref
 LIS: list (< ... >)
- b. LIS: <[REL: relation
 SIT: situation-ref (s1, s2,...)
 ARG1:
- c. 화맥 정보(CONTEXT)
 C-INDS: (SP, HR, UT)
 SP: person-referent
 HR: person-referent
 UT: time-referent
- d. 담화기능(DF)
 DF: (FOC, TOP)
 FOC: {set of referents}
 TOP: <list of referents>

5. 해례

위에 제시한 평가 기준을 적용하여 세 가지 유형의 번역을 검토한다. 예제-1은 단일한 문에서 주제(TOP)가 명시된 영문의 번역, 예제-2는 부자(父子) 사이의 대화를 전제로 하는 영어의 번역으로 한국어의 문계가 관련되고, 예제-3은 국제 무역 협정문의 번역이다. 위의 평가 기준에 따라 간소화한 MRS식 표기로 SL, TL을 각각 IL 표상하고, IL 표상간의 일치(match), 불일치(mismatch)를 적요(comment) 난에 기재한다.

첫째, (31)의 영어 번역 문제에 네 명의 학생(a,b,c,d)이 각각 다른 번역을 했다고 가정하자. 출제자의 의도는 문두에 위치한 주제(John)의 파악 여부에 있다고 하자. 번역 평가의 틀에서 검색하면 다음과 같다.

- (31) SL/E.S1 John, Mary adores. (대조: Mary adores John.)
 TL/K.S1 a. 존이 매리를 숭배해요.
 b. 존은 매리가 아침해요.
 c. 존을 매리가 숭배해요.
 d. 존은 매리가 숭배해요.

(31a)는 주어-목적어의 구문 파악이 잘못된 오역이고, (31b)는 단어 ‘adore’의 오역이다.⁵⁾ (31c)와 (31d)는 구문이나 단어의 번역에서 둘 다 잘못이 없는 정역이다. 그러나 (31c) 보다 (31d)가 좋은 번역이다. 원문에 나타난 주제(topic) 개념이 주제 표시의 ‘는’을 통하여 제대로 나타나 있다. 출제자의 의도가 바로 주제 파악 여부에 있으므로 (31d)가 완벽한 100점 짜리 번역이고 (31c)는 주제 파악을 못한 데서 평가 기준 [7-1]에서 10점 감점된다. 주제를 고려하지 않는 입장에서는 (31c)도 감점 없는 만점을 받을 것이다.⁶⁾ (31d)의 IL 표상

⁵⁾이 두 (31a), (31b)의 오역은 평가항목 [1] 상황에 해당하며, (31a)는 [1-1] REL에서 30점, (31b)는 [1-2] 논항의 문법관계의 오역으로 40점이 각각 감점된다.

은 다음과 같다.

(32) MRS식 IL 표상 (31d)

SL/E.S1: John, Mary adores.	TL/K.S1d: 존은 매리가 숭배해요.
HAN: hi	HAN: hi
IND: s1	IND: s1
LIS: <[REL: adore	LIS: <[REL: adore
HAN: hi	HAN: hi
SIT: s1	SIT: s1
ARG1: x1	ARG1: x1
ARG2: x2],	ARG2: x2],
[REL: t_overlap	[REL: t_overlap
HAN: h1	HAN: h1
UT: t1	UT: t1
ET: t2],	ET: t2],
[REL: naming	[REL: naming
HAN: h1	HAN: h1
NAMED: x1	NAMED: x1
NAME: Mary],	NAME: Mary],
[REL: naming	[REL: naming
HAN: h1	HAN: h1
NAMED: x2	NAMED: x2
NAME: John],	NAME: John],
[REL: topic	[REL: topic
HAN: h1	HAN: h1

*아래 (i)을 뒤섞기(scrambling) 한 것이 (31c)이므로 정상어순(S-O-V)의 (i) 보다는 (31c)가 더 낫다고 할 수 있겠으나 전치된 '존을'이 초점일 수 있어 양자간의 번역의 우열을 논하기가 쉽지 않다.

- i. 매리가 존을 숭배한다.

ARG: x2]>

ARG: x2]>

이제 위 MRS식 IL을 바탕으로 한 (31d)의 번역 평가를 통사부 (SYN)와 의미부(SEMPRA)로 나누어 제시한다.

(33) 번역 평가

SL/ES1: John, Mary adores.

TL/K.S1d: 존은 매리가 숭배해요.

(D#: 1, DS#: 1, S#: 1)

(D#: 1, DS#: 1, S#: 1)

I. SYN(통사부)

Comment

W#:3

WP#:²⁷

- match

ST:dcl

ST: dcl

- match

SL:[]

SL: pol

- N/A(해당 무)

CT:spl :

CT:spl

- match

II. SEMPRA(의미부)

1 SIT: s1

1 SIT: s1

- match

1-1 REL: adore

1-1 REL: adore

- match

ARG1: x1

ARG1: x1

- match

ARG2: x2

ARG2: x2

- match

1-2 REF-IND

1-2 REF-IND

- match

x1: Mary

x1: Mary

- match

x2: John

x2: John

- match

1-3 POL: +

1-3 POL: +

- match

3-1 TNS: prs

3-1 TNS: prs

- match

7-1 TOP: x2

7-1 TOP: x2

- match

감점: 0

다음 예제-2는 영어 대화문을 네 가지로 번역한 것이다. 영어에서는 무의미한 화제/문제가 번역에서는 문제가 된다. MRS 표상에

²⁷한국어는 영어의 W에 어구(WP)로 대비(즉, <N+조사>를 한 WP로 계산).

담화의 참여자를 HPSG식으로 화맥지표(contextual-indices, C-INDS)인 화자(SP), 청자(HR), 발화시(UT) 정보를 추가한다.

(34) 예제-2

SL/E: (Father to his 15-year-old son) TL/K: (아버지가 15살 난 아들에게)

S1 I give you my word.	S1a 내가 약속한다.
S2 So go to bed.	S2a 가 자라.
	S1b 내가 네게 내 말을 준다.
	S2b 그러니 침대로 가라.
	S1c 제가 약속합니다.
	S2c 그러니 주무십시오.
	S1d 내가 약속한다.
	S2d 그렇지만 가 자라.

(35) IL 표상 (MRS식)

SL/E:	TL/K:
S1:	S1a:
C-INDS: (SP:x1, HR:x2, UT:t1)	C-INDS: (SP:x1, HR:x2, UT:t1)
HAN: h1	HAN: h1
IND: s1 (ET: t2, RT: t3)	IND: s1 (ET: t2, RT: t3)
LIS: <[REL: promise	LIS: <[REL: promise
HAN: h1	HAN: h1
SIT: s1	SIT: s1
ARG: x1	ARG: x1
SOA-ARG: s2],	SOA-ARG: s2],
[REL: t_overlap	[REL: t_overlap
UT: t1	UT: t1
ET: t2],	ET: t2],
[REL: father	[REL: father
HAN: h3	HAN: h3

ARG1: x1	ARG1: x1
ARG2: x2]>	ARG2: x2],
	[REL: not_honor
	HAN: h3
	ARG1: x1
	ARG2: x2]>
S2:	S2a:
C-INDS: (SP:x1, HR:x2, UT:t3)	C-INDS: (SP:x1, HR:x2,, UT:t3)
HAN: h4	HAN: h4
IND: s3 (ET: t4, RT: t4)	IND: s3 (ET: t4, RT: t4)
LIS: <[REL: sleep	LIS: <[REL: sleep
HAN: h4	HAN: h4
SIT: s3	SIT: s3
ACT: x2],	ACT: x2],
[REL: t_precede	[REL: t_precede
HAN: h4	HAN: h4
UT: t3	UT: t3
ET: t4],	ET: t4],
[REL: direct	[REL: direct
HAN: h5	HAN: h5
SIT: s4	SIT: s4
ARG1: x1	ARG1: x1
ARG2: x2	ARG2: x2
SOA-ARG: s3]>	SOA-ARG: s3],
	[REL: not_honor
	HAN: h3
	ARG1: x1
	ARG2: x2]>
DLR(담화연결):	DLR(담화연결):
HAN: h7	HAN: h7

IND: s4	IND: s4
LIS: <[REL so	LIS: <[REL so
HAN: h1	HAN: h1
SIT: s4	SIT: s4
ARG1: s1	ARG1: s1
ARG2: s2]>	ARG2: s2]>

위 IL 표상에서 한국어 번역에는 화맥에 ‘화자가 청자를 ‘honor’하지 않는다’는 정보가 들어 있고 이 정보가 문계(SL)가 평어체(plain level)라는 통사적 자질에 대응하여 ‘약속한다’, ‘가자라’는 형태와 호응하고 (34) TL/K의 S1c, S2,c와는 호응하지 못한다. 즉, (36a)가 정형, (36b)가 비정형으로 각각 판정된다. 따라서 번역 S1c, S2c는 [8]의 HON의 오역이 된다.

(36) a. [SL: plain]	b. # [SL: formal]
[REL: not_honor	[REL: not_honor
SP: x1	SP: x1
HR: x2]	HR: x2]

(34b)의 번역은 ‘give my word’와 ‘go to bed’의 관용적 의미를 파악 못한 오역이다. (34d)에서는 담화연결관계가 ‘그렇지만’으로 잘못 연결되었다. 이들에 대한 IL 표상은 생략하고 번역 평가로 들어간다.

(37) 번역 평가 (예제-2)

SL/E: (Father to his 15-year-old son)	TL/K: (아버지가 15살 난 아들에게)
S1 I give you my word.	a. S1 내가 약속한다.
S2 So go to bed.	S2 가 자라.
	b. S1 내가 네게 내 말을 준다.

S2 그러니 침대로 가라.

c. S1 재가 약속합니다.

S2 그러니 주무십시오.

d. S1 내가 약속한다.

S2 그렇지만 가자라.

I. SYN(통사부)

(D#1, DS#1, S#2)	(D#: 1, DS#: 1, S#: 2)	
S1:	S1a:	Comment
W#	WP#: 2	
ST: dcl	ST: dcl	- match
SL: []	SL: frm	- N/A
CT: spl	CT: spl	- match
IDM: <give x's word> => <x promise>		
S2:	S2a:	
W#: 4	WP#: 2	
ST: imp	ST: imp	- match
SL: []	SL: frm	- N/A
CT: spl	CT: spl	- match
IDM: <go to bed> => <sleep>		

II. SEMPRA(의미부)

S1:	S1a:	
1 SIT: s1	1 SIT: s1	- match
1-1 [REL: promise	1-1 [REL: promise	- match
ARG1: x1	ARG1: x1	- match
ARG2: s2]	ARG2: s2]	- match
1-2 REF-IND	1-2 REF-IND	- match
x1: father	x1: father	- match
x2: son	x2: son	- match
1-3 POL: +	1-3 POL: +	- match
3-1 TNS: prs	3-1 TNS: prs	- match

4	DLR:	4	DLR:	- match
	[REL: so		[REL: so	
	ARG1: s1		ARG1: s1	
	ARG2: s3]		ARG2: s3]	
6	SA:	6	SA:	- match
	x1: speaker		x1: speaker	
	x2: hearer		x2: hearer	
	t1: utterance time		t1: utterance time	
7	HON:	7	HON:	- match
	[REL: not_honor		[REL: not_honor	
	ARG1: x1		ARG1: x1	
	ARG2: x2]		ARG2: x2]	
			↓	
			[SL: plain]	- match
S2:		S2.a:		
1	SIT: s3	1	SIT: s3	- match
1-1	[REL: sleep	1-1	[REL: sleep	- match
	ARG: x2		ARG: x2]	- match
1-2	REF-IND	1-2	REF-IND	
	x1: father		x1: father	- match
	x2: son		x2: son	- match
1-3	POL: +	1-3	POL: +	- match
3-1	TNS: prs	3-1	TNS: prs	- match
4	DLR:	4	DLR:	- match
	[REL: so		[REL: so	
	ARG1: s1		ARG1: s1	
	ARG2: s3]		ARG2: s3]	
6	SA:	6	SA:	
	x1: speaker		x1: speaker	- match
	x2: hearer		x2: hearer	- match

t1: utterance time	t1: utterance time	- match
7 HON:	7 HON:	- match
[REL: not_honor	[REL: not_honor	
ARG1: x1	ARG1: x1	
ARG2: x2]	ARG2: x2]	
	↓	
	[SL: plain]	

(37b)의 경우는 관용어구가 문자대로 번역되어 그 IL 표상이 관용의 의미와 부합하지 않으므로 오역이 된다. 즉,

(38) SL/E:S1:TL/K:b.S1:

I give you my word	내가 네게 내 말을 준다.	
S1:	S1b:	
1 SIT: s1	1 SIT: s1	
1-1 [REL: promise	1-1 [REL: give	- mismatch
ARG1: x1	ARG1: x1	
ARG2: s2]	ARG2: x2	
	ARG3: x10]	
1-2 REF-IND	1-2 REF-IND	
x1: father	x1: father	
x2: son	x2: son	
	x10: x1'sword	- mismatch
1-3 POL: +	1-3 POL: +	

감점: 40

(39) SL/E:S2:

So go to bed.

S2:

1 SIT: s3

1-1 [REL: sleep

TL/K:b.S2:

그러니 침대로 가라.

S2a:

1 SIT: s9

1-1 [REL: go

- mismatch

ARG: x2]	ARG: x2	
	PATHTO: x9]	- mismatch
		감점: 40

S1, S2에서 각각 40점 감점이므로 번역 b의 평균은 40점이다.

번역-c는 대우 표현의 오역, 번역-d는 담화연결관계의 오역이다. c-문의 번역을 원문과 대비하면 다음과 같다. 적요 칸에는 오역 (mismatch)만 표기한다.

(40) SL/E:S2:	TL/K.c.S1:	
I give you my word.	제가 약속합니다.	
7 HON:	7 HON:	
[REL: not_honor	[REL: honor	- mismatch
ARG1: x1	ARG1: x1	
ARG2: x2]	ARG2: x2]	
	↓	
	[SL: formal]	
		감점: 20

(41) SL/E:S2:	TL/K:c.S2:	
So go to bed.	그러니 주무십시오	
7 HON:	7 HON:	
[REL: not_honor	[REL: honor	- mismatch
ARG1: x1	ARG1: x1	
ARG2: x2]	ARG2: x2]	
	↓	
	[SL: formal]	

번역-c는 S1, S2에서 각각 20점씩 감점, 따라서 평균 80점으로 평가된다. d-문은 연결관계가 'so-관계를 '그렇지만'-관계로 오역한 것으로 S2에서 20점 감점, 따라서 평균 90점으로 평가된다.

(42)	SL/E.S2:	TL/K.c.S2:	
	So go to bed.	그렇지만 가자라.	
	4 DLR:	4 DLR:	- mismatch
	[REL: so	[REL: but	
	ARG1: s1	ARG1: s1	
	ARG2: s3]	ARG2: s3]	

예제-3은 우루과이라운드(UR) 무역 협정문의 첫 부분의 번역이다. 먼저 원문과 국문 번역을 병기한다.

(43) 예-3(UR 최종협정문, p.8)

I. 원문(SL/E)

MTO Agreement Establishment

Article II Scope of the MTO

1. The MTO shall provide the common institutional framework for the conduct of trade relations among its Members in matters related to the agreements and associated legal instruments included in the Annexes to this Agreement.

II. SYN(통사부)

(D#: 1, DS#: 1, S#: 1)

S1:

W#: 34

ST: dcl

SL: []

CT: cpd

III. SEMPra(의미부)

SL/E.S1

I. 번역문(TL/K)

다자간 무역기구 설립 협정

제 II 조 다자간 무역기구의 범위

1. 다자간 무역기구는 본협정의 부속서에 포함된 사항에 있어서 회원국들 간의 무역관계가 이루어지는 데 있어서 필요한 공통의 제도적인 틀을 제공한다.

(D#: 1, DS#: 1, S#: 1)

S1:

WP#: 18

ST: dcl

SL: pln

CT: cpd

TL/K.S1

Comment

- match

- match

HAN h1	HAN: h1	
1 SIT: s1	1 SIT: s1	
1-1 [REL: provide	1-1 [REL: provide	(:제공한다)
HAN: h1	HAN: h1	
ARG1: x1	ARG1: x1	
ARG2: x2],	ARG2: x2],	
1-3 POL: +	1-3 POL: +	
2-1 [REL: for	2-1 [REL: needed_for	(:필요한)
HAN: h2	HAN: h2	
ARG1: h1	ARG1: x2	
ARG2: x3],	ARG2: x3],	
2-2 [REL: among	2-2 [REL: among	(:간의)
HAN: h3	HAN: h3	
ARG1: x3	ARG1: x3	
ARG2: x4],	ARG2: x4],	
2-1 [REL: in	2-1 [REL: in	(:있어서)
HAN: h4	HAN: h4	
ARG1: h1	ARG1: h2	
ARG2: x5],	ARG2: x5],	
2-2 [REL: related_to	2-2 [REL: done	- mismatch
HAN: h5	HAN: h7	
ARG1: x5	ARG: x3],	
ARG2: x6+x7],		
2-2 [REL: included_in	2-2 [REL: included_in	- mismatch
HAN: h6	HAN: h6	
ARG1: x6+x7	ARG1: x5	
ARG2: x8]	ARG2: x8],	
1-2 REF-IND	1-2 REF-IND	
x1: MTO	x1: MTO	(:다자간 무역기구)
x2: common_	x2: common_	

institutional_ framework	institutional_ framework (:공통의 제도적인 틀)
x3: conduct_of_ trade_relations	x3': trade_relations (:무역관계)
x4 x1's members	x4: x1's members (:회원국들)
x5: matters	x5: matters (:사항)
x6: agreement	x6: ∅ - missing
x7: associated_legal_ instrument	x7: ∅ - missing
x8: annex_to_this_ agreement	x8: annex_to_his agreement (:본 협정의 부속서)
3-1 TNS: prs	3-1 TNS: prs
4 DLR: none	4 DLR: none

예제-3은 통사적 대응에서 SL/E의 34 단어가 TL/K의 18어절로 대응되고, 각각 수식구문과 내포문을 포함한 복합문이며, 의미 표상에서도 다양한 관계로 구성되어 있어 협정문의 특색을 잘 들어내고 있다. 오역이 두 군데, 미번역이 두 군데 검색되었다. 참고로 보드나은 번역을 다음에 제시한다.

(44) 개역

다자간 무역기구는 본 협정의 부속서에 포함된 합의 및 이와 관련된 법적 기제에 관한 사항에 있어서 회원국들 간의 무역 관계를 이루는 데에 필요한 공통의 제도적 틀을 제공한다.

6. 마무리

번역 평가를 위해 의미의 등치와 그 정도를 측정하는 구체적인 검색 항목을 제시하여, 각 의미 항목을 MRS에 바탕을 둔 IL의 표

상으로 나타내고, 문-화단-담화 층위로 번역의 정역/오역을 검색하여 종합 평가하는 방안을 제시하였다. 이 바탕에는 개별 언어의 IL 사전을 전제하였다. 그러나 이런 IL 사전이 없는 시점에서 볼 때, 여기서 제시한 방향의 언어별 IL—IL/E, IL/K, IL/J 등—사전에 관한 연구와 사전 편찬이 번역 평가에 선행하는 과제로 남는다.

참고 문헌

- 상공자원부. 1993. UR 최종협정문 (국영문).
- 서상옥. 1996. (편). 영어개념구조연구. 서울: 형설출판사.
- 신효필. 1990. *HPSG를 기초로 한 한국어 동사의 하위범주화*. 서울대학교 석사학위논문 (언어학연구 7, 서울대학교)
- 이민행, 정소우, 지광신. 1997. “기계번역 시스템의 평가방안 연구. 최종보고서,” 시스템 공학연구소.
- 이주호, 이재성, 김정구, 최기선. 1998. “대응 번역 문장의 자동정렬.” MIRAC 발표회(1998.5.2).
- 장석진. 1993. 정보기반 한국어문법. 서울: 언어와 정보.
- 장석진. 1994. 통합문법론. 서울: 서울대학교 출판부.
- 최기선. 1996. “한국어/영어 ‘정렬시스템,” 국어정보베이스, 187-213 (한국과학기술원 제2차년도 최종보고서), 과학기술처.
- Arnold, Doug, Louisa Sadler and R. Lee Humprheys. 1993. “Evaluation: An Assessment,” *Machine Translation* 8, 1-24.
- Buitelaar, Paul. 1998. CoreLex: Systematic Polysemy and Under-specification. Brandeis University dissertation.
- Copestake. Ann. 1995. “Semantic Transfer for Verbmobil”. Draft.
- Copestake, Ann, Dan Flickinger, Robert Malouf, Susanne Riehemann, Ivan A. Sag. 1996. “Translation using Minimal Recursion Semantics,” *Proceedings of the Sixth*

International Conference on Theoretical and Methodological Issues in Machine Translation. Leuven..

- Dorr, Bonnie J. 1993. "The Use of Lexical Semantics in Interlingual Machine Translation," *Machine Translation* 7, 135-93.
- Dorr, Bonnie J. 1994. "Machine Translation Divergences: A Formal Description and Proposed Solution," *ACL* 20, 597-633.
- Dorr, Bonnie J., J. Lee, Dekang Lin, and Sungki Suh 1995. "Efficient Parsing for Korean and English: A Parameterized Message Passing Approach.," *ACL* 21.2, 255-63.
- Jackendoff, Ray S. 1983. *Semantics and Cognition*. Cambridge, MA: MIT Press.
- Jackendoff, Ray S. 1990. *Semantic Structures*. Cambridge, MA: MIT Press.
- Krieger, Hans-Ulich. Ulich Schäfer 1994. TDL—A Type Description Language for HPSG: Part1 1-2. Deutsches Forschungszentrum für Künstliche Intelligenz GmbH manuscript.
- Kay, Martin., Jean Mark Gawron, and Peter Norvig 1994. *Verbmobil: A Translation System for Face-to-Face Dialog*. Stanford: CSLI, Stanford University.
- McCawley, James. D. 1971. "Prelexical Syntax," *Monograph Series in Language and Linguistics* 24, 19-33.
- Minnis, Stephen. 1994. "A Simple and Practical Method for Evaluating Machine Translation Quality," *Machine Translation* 9, 133-49.
- Nierenburg, S. ed. 1987. *Machine Translation*. Cambridge: Cambridge University Press.
- Pollard, Carl. and Ivan A. Sag. 1994. *Head-Driven Phrase*

- Structure Grammar*. Stanford: CSLI, Stanford University.
- Reichenbach, Hans. 1947. *Elements of Symbolic Logic*. New York:Free Press.
- Sag, Ivan. A. and Thomas Wasow 1997. *Syntactic Theory: A Formal Introduction*. Stanford University manuscript.
- Sowa, John. F. 1984. *Conceptual Structures: Information Processing in Mind and Machine*. Reading, MA: Addison-Wesley.
- Sowa, John. F. ed. 1991. *Principles of Semantic Networks*. San Mateo, CA: Morgan Kaufmann Publishers.

서울시 관악구 신림동 산56의 1
서울대학교 인문대학 언어학과
151-742
sjchang@plaza.snu.ac.kr
FAX +82-2-3487-6669

접수일자: 1998. 5. 25.

게재결정: 1998. 6. 29.