

# DaMaN: 좌우접속정보를 이용한 한국어 생성기 (A Korean Generator using Left-Right Connectivity Information)

장원, 여상화, 정한민, 김태완, 황도삼, 박동인  
(Won Chang, Sang Hwa Yuh, Han Min Jung, Taewan KIM, DoSam Hwang, Dong-In Park)

국어공학센터 / 시스템공학연구소  
(Center for Korean Language Engineering / System Engineering Research Institute)

## 요약

기계번역은 대상 언어를 해석하고 변환하여 목적언어의 대역어를 선정한 후, 목적언어를 생성하는 과정을 거친다. 이때, 대상언어의 분석 단위에 따라 대역어의 생성 단위 또는 깊이가 다르다. 그러므로, 특정한 시스템을 위한 생성기는 그 시스템의 해석 또는 변환단계에서 추출되는 대역어에 의존하게 되어 시스템 호환성을 상실한다. 따라서, 중복된 생성기의 개발을 피하기 위하여 번역시스템 특성에 국한되지 않고 독립적으로 이용될 수 있는 한국어 형태소생성기 개발이 필요하다.

본 논문에서는 한국어 해석에 사용되는 한국어형태소 좌우인접정보를 이용하여 한국어 형태소를 생성하는 시스템인 DaMaN을 소개한다. 세분류된 형태소의 활용과 접속, 조사의 변동, 띄어쓰기를 고려한 형태소 좌우접속정보를 임의 조합 가능한 복합형태(합성어)에도 적용할 수 있도록 확장하였다. 따라서, 대상언어의 분석단위에 제한받지 않으므로 시스템 호환성이 있다.

## 1. 서론

기계번역은 대상언어의 분석깊이에 따라 대역어의 생성 단위 또는 깊이가 다르다. 예를 들어, 일본어의 ‘字が書ける’를 ‘字/が/書/け/る’, ‘字/が/書け/る’, ‘字/が書け/る’ 등으로 분석할 수 있다. 이와 같이, 기계번역시스템마다 최소성분의 형태소를 분석하거나 관용어구 또는 속어를 분석하는 등 각기 분석 단위가 다르고 처리 기법이 다르므로 특정한 번역시스템의 생성기는 그 시스템의 해석 또는 변환처리에 종속되어 다른 시스템과의 호환성을 상실한다. 지금까지의 실용화된 기계번역시스템의 생성기로서는 ATLAS-I J/K[1], MATES E/K[2] 등에서 볼 수 있다. 이 시스템들은 선어말어미 또는 보조동사, 활용접미사 등의 활용어 처리에 있어 형태소를 세분류하지 않았으며, 특정한 번역시스템의 생성 부분으로 개발되어 다른 시스템과의 호환성이 약하다. 그러므로, 한국어를 목적언어로 하는 다른 번역시스템을 개발한다면 기존의 생성기를 수정하거나 다시 개발해야 한다. 이러한 중복연구를 피하기 위하여 DaMaN은 활용어의 어간, 선어말어미, 보조동사 등과 같은 최소 단위 형태소를 분류하고, 이들의 상호 접속 관계에 의한 좌우접속정보를 도출한 후, 이 좌우접속정보를 임의 조합가능한 복합형태소(합성어)에까지 확장하였다. 이로써 형태소와 형태소, 형태소와 합성어, 합성어와 합성어의 조합을 자유롭게 하여 대상언어의 해석 단계 또는 변환 단계에 의존하지 않는 한국어 형태소 생성이 가능하다. 이 논문에서는 DaMaN에서 고려한 한국어 특성, 설계, 구성, 실험결과를 기술한다.

## 2. DaMaN에서 고려한 한국어 특성

아래에서는 한국어 형태소생성을 위하여, DaMaN에서 고려한 띄어쓰기, 활용어의 분류, 조사의 처리에 대하여 기술한다.

## 2.1 띄어쓰기

모든 형태소는 띄어쓰는 것을 원칙으로 하고 한글 띄어쓰기에 맞게 붙여써야 할 형태소를 [표1]과 같이 한정한다.

[표1] 붙여쓰는 형태소 분류

품사	비고
조사	좌측의 명사, 조사, 부사, 어미에 붙여 쓴다. (예: 연결어미 - ~지는, 종결어미 - ~다그러, ~다마는, 명사형 어미 - ~기는)
선어말어미	좌측의 활용언의 어간 또는 다른 선어말어미에 붙여 쓴다. (예: 했었다)
어말어미	좌측의 활용언의 어간 또는 선어말어미에 붙여 쓴다. (예: 한다)
접미사	좌측에 오는 체언에 붙여 쓴다. (예: 불안스럽다)
기타	일부 부사와 부사를 붙여 쓴다. (예: 똑같이)

## 2.2 활용어의 분류

활용어들을 [표2]와 같이 어간류 (활용시에 동사, 형용사, 존재사, 서술격조사, 조동사, 활용접미사 등의 변하지 않는 부분)와 선어말어미, 어말어미로 분류한다.

[표2] 활용어의 분류

구	분	비	고
어간 류	동사, 형용사	좌측은 반드시 띄어쓰기되며 우측에는 선어말 또는 어말어미가 붙여쓰여진다. 어간의 규칙/불규칙활용, 모음의 양성/음성, 종성유무에 따라 활용 형태를 분류한다.	활용규칙에 따른 처리를 거쳐
	존재사, 서술격 조사, 조동사, 활용 접미사	동사와 형용사의 어간은 반드시 앞 단어와 띄어쓰기하지만, 이 부류는 앞 형태소에 붙여 쓸수 있는 형태이므로 별도로 분류한다. 그러나, 우접속형태는 동/형용사 어간의 활용범주에 포함된다. 존재사는 체언에 붙지 않는 형용사의 쓰임새가 있고, 조동사와 활용접미사는 좌접속 형태에 차이가 있다.	
어미 류	선어말 어미	어간류에 접속할 때 어간의 활용형태에 따라 변형되며, 어말어미와 접속할 때 다시 변형되기도 한다. 그러나 어간과 접속 또는 다른 선어말어미의 접속시에도 변형의 틀은 어간과 어미의 접속형태를 벗어나지 않았으므로 좌우접속정보테이블 안에서 선어말어미의 활용형태도 처리할 수 있다.	
	어말어미	어간이나 선어말어미에 좌접속한다. 어말어미는 보조사나 문장부호 등에 우접속하게 된다.	

## 2.3 조사의 변형

조사는 앞에 오는 형태소의 종성유무와 ‘ㄹ’종성의 여부에 따라 ‘이/으’ 매개모음이 개입하거나 ‘을/를’, ‘이/가’처럼 형태를 달리 하는 조사가 있다. 이런 종류의 조사를 분간조사[3]로 분류하고 이외의 조사를 공통조사[3]로 분류한다. DaMaN은 분간조사의 경우 ‘이’, ‘으’ 매개모음이 포함된 조사형태를 입력하는 것으로 하여 무종성 형태소나 ‘ㄹ’종성형태소 뒤에 올 때 ‘이/으’를 삭제한다. 또, ‘이/가’, ‘을/를’, ‘와/과’등의 분간조사류는 둘중에 어느 것이 생성기에 넘어 오더라도 앞 형태소에 맞추어 바른 형태를 생성한다.

[표 3] 조사의 분류

분 1	구분 2	구분 3	분간조사	공통조사
격조사	주격	개체	[이,가]	
		여격 공동격	[와,과]	에,에게,한테,께 하고
	부사격	변성격 인용격	(으)로 (이)라(고)	
접속조사	단어접속		[와,과],[이]고, (이)며,(이)랑	하고, 에(다)
	문절접속			마는

### 3. DaMaN의 설계

여기서는 DaMaN의 시스템 구성, 데이터 입력 형태, 생성처리에 참조하는 생성기호, 좌우접속정보테이블, 띄어쓰기테이블, 출력형태에 대하여 기술한다.

#### 3.1 입력형태

입력은 한 문장분의 형태소와 생성기호를 쌍으로 하는 리스트로 구성되어 [그림1]과 같은 형태이다.

("철수" NN00001)("을" JJ0102)("찾" VV00003)("려고" EE00300)("선생님" NN00001)("깨서" JJ00202) ("오" VV00001)("서" BV00421)("았" BV00103)("겠" BV00004)("니" EE00000)("?" NN00001))
[그림 1] 입력문장의 형태

형태소는 단일형태소 또는 복합형태소로 구성되고 생성처리에 필요한 정보로 생성기호를 갖고 있다.

- ◎ 단일형태소 - 한국어를 해석할 때 최소의 의미를 가지는 단위로 구별되는 형태소다.  
(예: 먹, -있, -겠, -습니다)
- ◎ 복합형태소(합성어) - 대상언어를 분석함에 있어 관용구나 속어처럼 더 이상 분석할 필요가 없는 경우, 한국어 대역어는 단일형태소의 복합형태다.  
(예: '~하지 않으면 안되', '~가 되', '어서오십시오' 등)
- 편의상 이하에서는 단일형태소와 복합형태소를 하나로 묶어 '형태소'라 한다.

#### 3.2 생성기호

##### 3.2.1 생성기호의 구성

생성기호는 형태소의 성분을 구별하는 좌우품사기호, 다른 형태소와 좌우접속시 참조되는 좌우접속정보로 구성한다.

영문자1자리 좌품사기호	영문자1자리 우품사기호	숫자 2자리 좌접속정보	숫자 2자리 우접속정보
-----------------	-----------------	-----------------	-----------------

[그림2] : 생성기호의 구성

생성기호의 각 구성요소는 [표4]와 같은 의미로 구분한다.

[표4] 생성기호의 구성요소

구 분	의 미
좌품사기호	대역어의 좌측형태소의 품사
우품사기호	대역어의 우측형태소의 품사
좌접속정보	대역어의 좌측에 오는 형태소들과의 접속정보
우접속정보	대역어의 우측에 오는 형태소들과의 접속정보

### 3.2.2 생성기호 분류

여기서는 위의 생성기호구성 요소들을 분류에 대하여 기술한다.

#### 가. 좌/우 품사기호의 분류

좌우품사기호는 형태소간의 접속처리에 있어서 띄어쓰기와 활용처리를 구분하기 위하여 분류한다. 좌우품사기호 식별에 의해 활용어와 조사 이외의 품사는 띄어쓰기만 고려하면 된다. 분류기준은 활용어의 처리, 조사의 변형, 띄어쓰기를 충족하는 범위 내에서 최적화한다. 또한, 같은 부류의 활용어처리를 일원화하기 위하여 다른 품사라도 처리형태가 같으면 좌우품사기호 중 해당하는 쪽에 동일한 품사기호를 준다. (예: [표5]에서 형용사와 서술격조사의 우품사기호)

[표 5] 생성기호의 좌우품사기호

구 분	좌품사기호	우품사기호
동사, 1'간음화	V(erb)	V(erb)
형용사	A(djjective)	A(djjective)
어말어미	E(nding)	E(nding)
선어말어미	B(efore Ending)	V
활용접미사	I	V
서술격조사, 존재사	I	A
조사	J	J
기타	N(ominal)	N(ominal)

예를 들어, '가 있'이라는 대역어의 좌측부분의 '가'는 조사이므로 좌측품사는 조사의 좌품사기호(J)가 되며, 우측형태소 '있'은 존재사 '있다'의 어간이므로 우측품사는 존재사의 우품사기호(A)가 된다.

#### 나. 좌우접속정보의 분류

좌우품사기호에 의해 띄어쓰기, 활용어처리, 조사의 변형 등의 처리 종류를 판단할 수 있고 활용어의 좌우접속정보는 '[표 6] 좌우접속정보테이블'로부터 처리방법의 수치값을 참조하는데 이용한다.

##### 1) 좌접속정보의 분류

좌접속정보의 분류는 [표 7]과 같다. 앞에서 예를 들었던 '가 있'이라는 대역어의 좌접속정보를 보자. 조사의 좌측접속정보는 분간조사와 공통조사로 분류되어 있다. 조사 '가'는 앞단어의 종성유무에 따라 '이/가'로 변화하는 분간조사이므로 분간조사의 좌접속정보(01)를 취한다.

[표6] : 좌우접속정보 테이블

어간 분류	어미분류	다, 고 겠, 었	아/어+ 았, 었, 4)	느, 는	(으)면서, (으)면-	(으)리지, (으)니들	(느)니다	(스) 비니다	(으)ㅁ
00									
무종1	ㅏ, ㅑ, ㅣ, ㅓ, ㅕ 모음 :가다, 차다		E-아 12	-1			-1		
01	ㅓ :보다, 오다		E+아 ㅓ 16	-1			-1		
무종2									
02	ㅑ, ㅓ, ㅣ, ㅕ, ㅓ :쉬다, 기다		E+아어 13	-1			-1		
무종3									
03	ㅏ, ㅑ :잡다, 좋다, 밝다								
유종1									
04	ㅏ, ㅑ 이외		E+아어 13	-1			1+느 -22	1+스 23	1+으 21
유종2	:먹다, 놀다, 넓다								
09	ㅓ :끓다, 돌다		S-, E+아와 33	-1			1+느 -22	1+스 23	1+으 21
10	ㅓ 이외		S-, E+아워 34						
14	:출다, 길다, 아를닫다 ㅏ, ㅑ 이외 :부르다, 이르다 (EARLY, INFORS) :이르다 (ARRIVE), 푸르다		S-, E+아려 39	-1			-1		
15									
러			E+아려 15	-1			-1		
16	하다		E+아여 14	-1			-1		
여	:깨끗하다, 착하다								
20	푸다		S+T ㅓ, E-아 43	-1					
T									
21	피다		S+ ㅓ, E-아 44	-1					
1간음화	(계시다, 파래지다)								
22	있다, 없다		E+아어 13				-1	1+스 23	1+으 21
존제사									
23	이다		E+아여 14	-1			-1		
서술조사									

[표 7] 좌접속정보 분류

분류 및 품사기호	분류 방법	예	좌접속정보
동사(V)	어간	먹	00
형용사(A)	어간	아름답	00
선어말어미(B)/ 어말어미(E)	변화없이 앞에 붙이기만 하면 되는 형태	다,게,지,고 선어말어미: 았(대과거),겠	00
	앞에 오는 어간이나 선어말어미의 끝 음절 중성모음이 ㅏ,ㅓ인가 아닌가에 따라 어미의 첫음절도 양성이나 음성으로 조화시켜 주어야 하는 것 ('아/어'로 시작되는 어미)	아서,어서 아야,어야 아도,어도 선어말어미: 았(았)	01
	명사형어미 '(으)ㄴ'	(으)ㄴ	07
서술격조사(I)		이	04
존제사(I)		있,없	04
조사(J)	분간조사	이/가,을/를,와/과,로/으로 ...	01
기타 (N)			00

### 2) 우접속정보의 분류

우접속정보의 분류는 [표 8]과 같다. 앞에서 예를 들었던 '가 있'이라는 대역어의 우접속정보를 보자. 존재사 '있'은 우품사기호가 'A'이고 우접속정보는 '23'이다. 따라서, 대역어 '~가 있'의 생성기호는 'JA0123'이다.

[표 8] 우접속정보 분류

구분	활용형	어간 끝음절 중성	동/형용사의 예	선어말어미의 예	우접속정보
동사(V)	무중성 규칙어간	ㅏ, ㅓ, ㅕ, ㅗ, ㅛ	가, 차		00
		ㅓ	보, 오		01
형용사(A)		ㅓ, ㅕ, ㅗ, ㅛ, ㅕ, ㅗ, ㅛ	쉬, 기		02
선어말어미(V)	유중성 규칙어간	ㅏ, ㅓ	잡, 속	았	03
		ㅏ, ㅓ 이외	먹, 읽, 뛸	았, 겠, 았, 았, 았, 았	04
활용접미사 (v)		되(03), 하(16)	당하(16), 받(03)	시키(21)	
어말어미(E)					00
존제사 (A)			있, 없		22
서술격조사(A)			-이(다)		23
기타(N)					01

## 3.3 좌우접속정보 테이블

### 3.3.1 활용어의 처리패턴

활용어들은 어간을 기준으로 선어말어미, 어말어미 순으로 붙여쓰면서 활용규칙과 음운 현상을 맞춘다. 어간은 일반적인 분류형태로 정하고 선어말어미와 어말어미 형태를 양성(아/어의 경우 '아'를), 매개모음이 생략된 형태(예: (으)면서)를 기본형으로 한 접속형태를 분류하였다. 그 결과 아래와 같이 가.~다.의 형태 또는 이들의 복합적인 형태로 모두 27가지의 형태로 분류되었다. 이 27가지 패턴이 모든 활용관련어들의 접속처리를 가능하게 하였다.

- 가. 앞에 오는 형태소 S(어간 또는 선어말어미)에 변형을 주는 접속형태
  - S[-J] -> 앞에 온 형태소에서 종성이 탈락되는 형태
  - S[JJ'] -> 앞에 온 형태소에서 종성 J를 J'로 교체하는 형태
  - S[-M] -> 앞에 온 형태소에서 모음이 탈락되는 형태
  - S[MM'] -> 앞에 온 형태소에서 모음M을 M'로 교체하는 형태
- 나. 두 형태소간에 매개모음 삽입 형태 I
  - I[+M] -> 두 형태소 사이에 매개모음을 삽입하는 형태
- 다. 뒤에 오는 형태소 E(선어말 또는 어말어미)에 변형을 주는 접속형태
  - E[-M] -> 뒤에 오는 형태소의 모음 탈락 형태
  - E[MM'] -> 뒤에 온 형태소에서 모음M을 M'로 교체하는 형태

### 3.3.2 좌우접속정보테이블의 구성

좌우접속정보테이블은 활용어들의 접속처리를 위한 정보테이블이다. 이 테이블은 [표 7]과 [표 8]의 좌우접속정보를 행과 열로 하고 있으며, 27가지로 분류된 처리패턴을 수치값으로 가지고 있다. [표 6] 좌우접속정보테이블'을 설명하면 다음과 같다.

제1행의 2자리의 숫자는 어간류의 우접속정보에 해당하고 무종/유종/ㄷ불규칙/ 등은 규칙/불규칙 활용을 나타낸다.

제2행의 'ㄱ, ㄴ, ㄷ, ㄱ'와 같이 나열된 모음은 어간의 마지막 중성을 나타내며, 여기에 해당하는 어간의 부류를 예로 나열하고 있다.

제3행부터 10행까지는 어미의 좌접속정보(00 ~ 07)로서 8가지로 분류된다.

제4행 1열을 보면 'E-아'는 앞에서 설명한 처리패턴을 나타내고 '12'는 그 패턴에 대한 참조값이다. 테이블 안에서 참조값에 마이너스 부호가 첨가된 것은 형용사 어간에는 접속되지 않는다는 의미이다.

### 3.3.3 좌우접속정보 테이블을 이용한 예

예를 들어, '불안'이라는 명사와 '스럽다'라는 접미사를 조합하여 '불안스러우셨었겠습니까그려.' 라는 어절을 생성한다고 가정하면

- 불안(NN0001) + '스럽(IA0409)
  - 불안스럽(IA0409) + '으시(BV0321)' -> S-ㅂ, I+우
  - 불안스러우시(BV0321) + '았(BV0103)' -> S-ㅣ ㄷ, E-아
  - 불안스러우셨(BV0103) + '었(BV0004)' -> 무변환
  - 불안스러우셨었(BV0004) + '겠(BV0004)' -> 무변환
  - 불안스러우셨었겠(BV0004) + '바니다(EE0600)' -> I+스
  - 불안스러우셨었겠습니까(EE0600) + '그려(JJ0202)' => 불안스러우셨었겠습니까그려
- 와 같이 생성하게 된다.

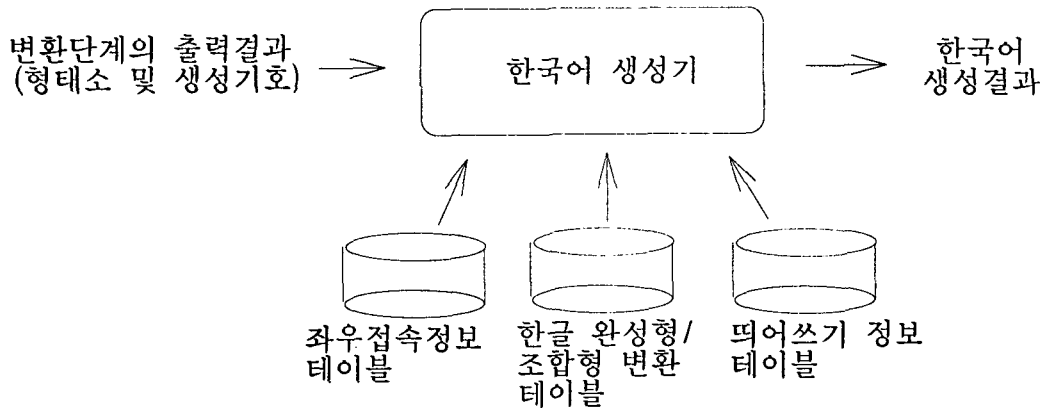
## 3.4 띄어쓰기 정보 테이블

활용어 이외에 붙여써야 하는 형태소들의 처리를 위하여 8 x 8 매트릭스인 '띄어쓰기 정보테이블'을 구성하여 처리하였다. 행과 열은 활용관계어 이외의 붙여쓰기할 대상 형태소의 좌우접속정보로 구성되고 그 값은 0과 1로 붙여쓰기 여부를 나타낸다.

## 4. DaMaN의 구현

본 형태소 생성기는 SUN 호환 국산 W/S SDT-400에서 C언어로 구현하였고 한글 코드는 KS C-5601 2byte완성형을 사용하였으며,음소의 접속을 위해 2byte 조합형으로 변환하여 처리하는 방식을 채택하고 있다. 생성시 2byte 조합형을 다시 완성형으로 변환한다. 다음에 시스템 구성도와 알고리즘을 소개한다.

### 4.1 시스템 구성도



[그림3] : 한국어 생성기의 구성도

### 4.2 생성 처리 알고리즘

- 1) 한 문장분의 형태소와 생성기호 list 를 입력하여 분류 저장
- 2) 한 문장분의 list 를 좌에서 우방향으로 앞 쪽 형태소의 우품사기호 및 우접속정보와 뒷 쪽 형태소의 좌품사기호 및 좌 접속정보를 추출
- 3) 만약 두 대역어의 생성기호가 활용어 기호이면  
두 형태소의 좌우접속정보를 index 로하여 좌우접속정보테이블로부터 처리구분 값을 찾아 이 값을 parameter로 하여 접속처리 루틴을 호출한다.
- 4) 아니고 뒤쪽 형태소의 좌품사기호가 조사이면 조사처리 루틴을 호출한다.
- 5) 아니고 뒤쪽 형태소의 좌품사기호가 활용접미사이면 앞 형태소에 붙여쓴다.
- 6) 아니면 모두 띄어쓰기로 출력한다. 그리고 뒷쪽 형태소의 우품사기호와 우접속정보를 분리하여 다음 형태소와 접속을 위한 좌측 형태소의 생성기호로 저장한다.
- 7) 문말이 아니면 pointer 를 우로 옮기고 2)항부터 반복한다.
- 8) 문말이면 생성결과를 출력한다.



## 5. 실험 결과

[그림 4]는 실험하기 위한 대역어와 생성기호 쌍이 나열되어 있고 생성결과가 출력되어 있다. 띄어쓰기와 활용어의 처리, 조사의 변동에 중점을 두어 개발한 결과 만족할 만한 실험 결과를 얻었다. 실험은 모든 규칙/불규칙 용언의 어간과 선어말어미, 어말어미, 조동사, 보조형용사, 활용접미사의 생성을 고려한 총 800 여 문장의 실험결과에서 어떠한 형태의 조합이라도 띄어쓰기와 조사의 변형 그리고 활용규칙에 따른 선어말/어말어미의 접속이 정확히 출력되었다.

1>	((("나" NN0001)("만" JJ00202)("이" JJ00102)("알" VV00007)("고" EE00100)("있" VV00022)("다" EE00000))
	생성결과 : 나만이 알고 있다
2>	((("하" VV00016)("시" BV00421)("았" BV00103)("르"지" EE00400)("모르" VV00013)("겠" BV00004)("느냐" EE00600)("." NN00001))
	생성결과 : 하겠을지 모르겠습니다 .
3>	((("m" NN00001)("는" JJ00102)("또는" AD00505)("o" NN00001)("은" JJ00102)("1" NN00001)("와" JJ00102)("2" NN00001)("과" JJ00102)("3" NN00001)("과" JJ00102)("4" NN00001)("을" JJ00102))
	생성결과 : m 은 또는 o 는 1 과 2 와 3 과 4 툴
4>	((("김" NN00707)("씨" NN00808)("과" JJ00102)("김" NN00707)("가" NN00808)("는" JJ00102)("들어"가" VV00000)("기" EE00000)("는" JJ00102)("커"녕" JJ00202))
	생성결과 : 김씨와 김가는 들어가는커녕
5>	((("더" AD00505)("빨리" AD00505)("똑" AD00505)("같이" AD00505)("달리" VV00021)("아라" EE00100))
	생성결과 : 더 빨리 똑같이 달려라
6>	((("문" NN00001)("밖" NN00001)("에서" JJ00202))
	생성결과 : 문밖에서
7>	((("일" NN00001)("-" NN00001)("한" NN00001)("기계"번역" NN00001)("시스템" NN00001)("을" JJ00102)("개발" NN00001)("중" NN00001)("이" IA00423)("느냐" EE00600)("." NN00001))
	생성결과 : 일 - 한 기계번역 시스템을 개발 중입니다 .
8>	((("철수" NN00001)("을" JJ00102)("찾" VV00003)("려고" EE00300)("선생"님" NN00001)("께서" JJ00202)("오" VV00001)("시" BV00421)("았" BV00103)("겠" BV00004)("니" EE00000)("?" NN00001))
	생성결과 : 철수를 찾으려고 선생님께서 오셨겠니 ?
9>	(((" " NN00001)("그" NN00001)("여자" NN00001)("에게" JJ00202)("아름"답" AA00010)("다" EE00000)("." NN00001)(" " NN00001)("이라고" JJ00102)("하" VV00016)("시" BV00421)("았" BV00103)("겠" BV00004)("느냐" EE00600)("그"려" JJ00202)("." NN00001))
	생성결과 : ` 그 여자에게 아름답다 .`라고 하셨습니다그려 .
10>	((("받지" 않" VV00003)("아"야"만" 하" EV00116)("느냐" EE00600))
	생성결과 : 받지 않아야만 합니다
11>	((("회"전" NN00001)("시"키" IV00421)("다" EE00000))
	생성결과 : 회전시키다
12>	((("먹" VV00004)("시" BV00421)("았" BV00103)("었" BV00004)("겠" BV00004)("다" EE00009))
	생성결과 : 먹으셨겠다

[그림 4] 실험결과

## 6. 결론

DaMaN 에서 생성의 주안점으로 하였던 점들 이외에도 해결해야 할 것들이 있다. 일부 붙여쓰기하는 부사와 부사, 숫자와 알파벳에 접속하는 조사의 변형, 맞춤법상 붙여쓰기하는 일부 명사와 명사의 결합을 부분적으로 수행하여 만족할 수 있었으나 개선의 여지가 아직 남아 있다. 이외에도 사이시옷 삽입 문제('수' + '자' = '수자' -> '숫자'), 특수기호 다음의 조사처리(예: 교황(~설명문-)가 ~한다), 부사어미 '이', '히'의 생성 또는 음운 축약(딸 + 님=따님)과 같은 음운 변동, 같은 뜻의 어휘라도 문맥이나 문체에 어울리는 더 나은 한국어다운 표현 영역을 고려해야만 독립적인 한국어 생성기가 될 것이다.

## 참고문헌

- [1] 시스템공학연구소, 일한기계번역시스템개발 연구(3), 1989
- [2] 시스템공학연구소, 기계번역을 위한 언어모델링 및 골격시스템 개발(3), 1992
- [3] 하희주, 고교국문법 교실, 성문각, 1989
- [4]. 한광록, '한일 기계번역시스템에 관한 연구', 박사학위논문, 인하대학교, 1989
- [5] 김상국, '영한기계번역을 위한 형태소생성시스템의 설계 및 구현', 석사학위논문, 한양대학교, 1991

본 생성기의 활용어의 처리는 한광록[4]씨의 박사학위 논문중 한국어생성 부분과 이를 실용시스템에 적용시킨 김상국[5]씨의 알고리즘을 참조하였음을 밝혀둔다.