

# 의사 형태소 단위의 연속 음성 인식<sup>1)</sup>

이경님, 정민화

서강대학교 전자계산학과

## Pseudo-Morpheme-Based Continuous Speech Recognition

Kyongnim Lee and Minhwa Chung

Department of Computer Science, Sogang University

E-mail : {knlee, mchung}@nlpeng.sogang.ac.kr

### 요 약

본 논문에서는 언어학적 단위인 형태소의 특성을 유지하면서 음성인식 과정에 적합한 분리 기준의 새로운 디코딩 단위인 의사 형태소(Pseudo-Morpheme)를 정의하였다. 이러한 필요성을 확인하기 위해 새로이 정의된 37개의 품사 태그를 갖는 의사 형태소를 표제어 단위로 삼아 발음사전 생성과 형태소 해석에 조점을 두고 한국어 연속음성 인식 시스템을 구성하였다. 각 음성신호 구간에 해당되는 의사 형태소가 인식되면 언어모형을 사용하여 구성된 의사 형태소 단위의 상위 5개 문장을 기반으로 시작 시점과 끝 시점, 그리고 확률값을 가진 의사 형태소 격자를 생성하고, 음성사전으로부터 태그 정보를 격자에 추가하였다. Tree-trellis 탐색 알고리즘 기반에 의사 형태소 접속정보를 사용하여 음성언어 형태소 해석을 수행하였다. 본 논문에서 제안한 의사 형태소를 문장의 디코딩 단위로 사용하였을 경우, 사전의 크기면에서 어절 기반의 사전 entry 수를 현저히 줄일 수 있었으며, 문장 인식률면에서 문자기반 형태소 단위보다 약 20% 이상의 인식률 향상을 얻을 수 있었다. 뿐만 아니라 형태소 해석을 수행하기 위해 별도의 분석과정 없이 입력값으로 사용되며, 전반적으로 문장을 구성하는 디코딩 수를 안정화 시킬 수 있었다. 이 결과값은 상위레벨 언어처리를 위한 입력값으로 사용될 뿐만 아니라, 언어 정보를 이용한 후처리 과정을 거치지 않은 인식률 향상을 꾀할 수 있다.

### 1. 서 론

한국어는 교착어로서 한 어절은 실질 형태소와 조사나 어미와 같은 형식 형태소로 이루어진다. 따라서 문장을 단어의 연결로 보는 것이 아니라 형태소의 연결로 형식 형태소의 수를 어느 정도 고정함으로써 사전 크기를 안정화 시키고 문장 재인식부에서는 언어적인 정보를 사용하여 음성 인식률을 높이는 방법을 제안하고자 한다.

음성언어 처리의 경우, 입력단위나 인식단위의 경계가 모호하고 신호처리의 결과값에 이미 오류를 포함하고 있기 때문에 문자기반의 자연어 처리 방법을 그대로 음성인식 시스템에 적용하기가 그리 쉽지 않다. 일례로 형태소 단위의 음성인식 시스템에서는 다음과 같은 문제점을 볼 수 있다. 일반적인 언어학에서 말하는 형태소들은 단음소나 단음절로 이루어져 있다. 예를 들면, '나', '리', '이' 등과 같은 단음소와 대부분 의존명사, 접미사의 경우는 단음절로서 하나의 형태소이다. 이와 같은 형태소는 음성인식에서 매우 짧은 시간 동안에 발생되기 때문에 이를 인식하기에는 많은 어려움이 있다. 따라서 음성인식에서는 이를 보완하기 위한 방법이 강구되어야 한다. 이러한 문제점을 피하기 위한 한가지 방법으로, 주로 형태소를 분리하지 않고 어절을 그대로 사용하는 경우를 들 수 있겠다. 발성의 지속 시간의 제법 길기 때문에 인식률을 어느 정도 이상으로 높일 수 있으나, 인식 대상 단어수가 늘어나고 활용에 많은 제

<sup>1)</sup> 본 연구는 1998년도 LG 전자의 연구비 지원에 의하여 이루어졌음

## 의사 형태소 단위의 연속 음성 인식 시스템

약을 받는다는 단점 때문에 역시 잠정적인 해결책이 될 뿐이다. 결국 인식률을 높이기 위해서는 적절한 길이의 발성 시간과 적절한 수의 단어를 가질 수 있는 단위가 새로이 정의되어야 한다. 이를 위해서 가급적으로 언어학적 단위인 형태소를 유지하면서 음성인식에 그다지 부리가 가지 않는 범위 내에서 적절한 분리 기준에 따른 새로운 디코딩 단위가 필요하게 된다[10][13].

본 논문에서는 언어학적 단위인 형태소의 특성을 유지하면서 음성인식 과정에 적합한 분리 기준의 새로운 디코딩 단위인 의사 형태소를 정의하고, 37개의 품사 태그를 갖는 의사 형태소를 표제어 단위로 삼아 발음사전 생성과 형태소 해석에 초점을 두고 한국어 연속 음성 인식 실험을 수행하였다.

## 2. 본 론

### 2.1 의사 형태소 (Pseudo-Morpheme)

디코딩 단위는 사전의 표제어 단위이며 분장의 구성 단위가 된다. 연속 음성 인식 시스템에서 인식률을 높이기 위해서는 적절한 길이의 발성시간과 적절한 수의 사전 표제어를 가질 수 있는 디코딩 단위가 정의되어야 한다. 이 장에서 설명하는 의사 형태소는 발음사전의 표제어, 더 나아가서는 음성언어 사건의 표제어 단위가 되기 때문에 선정 기준이 매우 중요하다. 본 논문에서 제안하는 의사 형태소는 기존 문자기반의 형태소 단위 [9][12]와 기본 골격은 같으나 다음과 같은 경우에 의해 구분되어진다. 첫째, 합성어의 경우로 복합어나 파생어의 경계점에서의 'ㄹ'첨가 현상의 발생과 한자어의 결합에서만 발생하는 두음법칙, 사이 'ㅅ'의 비첨가 현상, 역행적 유음화, 그리고 한자어의 선택적 경음화 현상 등을 위한 역주적 규칙은 그 노력에 비해 손실이 크기 때문에 지금까지 이는 발음 생성시 예외사전 처리를 해왔다. 이와 같은 것은 의사 형태소로 두고, 이와 같은 특징을 형태소 정보에 싣고 이를 이용하여 그 특징에 맞춘 발음열을 생성하고, 또한 분석시 의사 형태소로 구성된 이 정보를 이용하여 효율적인 분석도 시행 가능하다. 둘째, 빈번히 생략이 가능한 조사의 경우로 다음과 같이 나누어 살펴 볼 수 있다. 격조사는 여러 개의 결합된 형태 자체를 하나의 단위로 본다. 종결 보조사는 앞의 형태소와 결합한 형태로 인식단위를 선정한다. 셋째, 시술어의 경우 어미의 결합은 필수적이다. 음성인식에서 용언구의 인식은 전체 성능 평가에 거대한 영향을 미치므로 가능한 한 그 음운변화를 그대로 사전에

반영하는 것이 시스템의 성능에 도움이 된다. 따라서 여기서는 선어말어미는 연결어미와 선성어미, 종결어미와 결합된 형태로 사전에 등록한다.

이에 따라 의사 형태소 태그는 '한국어 표준 형태 품사 태그 분류 표준안'[14]에 기반하여 기술하였으며 표준안과 더불어 KIS에 사용된 품사태그[9]를 참조하여 음성언어 처리시 고려되는 새로운 기준에 적합한 37개의 품사 태그로 정의하였다.

[표 1] 의사 형태소 태그 Set

대분류	태그	명칭	결합규칙
체언 (n)	ncpa	동작성 보통명사	
	ncps	상태성 보통명사	
	ncn	비서술성 보통명사	
	nct	시간성 보통명사	← nnc+nbu
	nq	고유명사	
	nbu	단위성 의존 명사	
	nbu	비단위성 의존명사	
	npp	인칭대명사	
	npd	지시대명사	
	anc	양수사	← nnc+nbu
nno	서수사		
nx	보조체언	= ctm+nbu	
용언 (p)	pvd	지시동사	
	pvg	일반동사	
	pa	형용사	
	px	보조용언	← 양상
수식언 (m)	mm	관형사	← mmd, mma
	maj	접속부사	
	mag	일반부사	← +xsa
독립언	ii	감탄사	
관계언 (i)	jes	주격조사	
	jeo	부격조사	
	jec	보격조사	
	jcm	관형격조사	
	jev	호격조사	
	jea	부사격조사	
	jei	접속격조사	
	jet	공동격조사	
	jeq	인용격조사	
jp	서술격조사		
어미 (e)	ecc	대동작 연결어미	← ep + ecc
	ecs	중속적 연결어미	← ep + ecs
	etn	명사형어미	← ep + etn
	etm	관형사형어미	← ep + etm
	ef	종결어미	← ep + ef
접사 (x)	xsn	명사파생접미사	
	xsv	동사파생접미사	

음성인식의 특성상 구체적으로 고려되는 사항을 나누어서 살펴보면 다음과 같다.

한국어에서는 양상을 나타내는 보조용언과 어미 등의

낱절접사가 대단히 발달되어 있어 다양한 형태와 분포, 의미를 보이고 있다. 또한 문장에서 대부분의 문법 기능을 담당하고 있으므로 문장을 이해하는데 매우 중요한 역할을 한다. 일반적인 한국어 술어의 구조를 나타내면 다음과 같다[1].

술어 = 어간 + 보조용언 + 선어말어미 + 어말어미  
 보조용언 = (피동 | 사동 | 상 | 범성 | 부정)  
 선어말어미 = (존칭 | ε) + 시제 + 겸칭  
 어말어미 = 서법

서술어의 경우 어미의 활용에 따라 인식성능의 저하를 가져오기 때문에 그 기준 설정이 매우 중요하다. 대부분 음성인식 수행시 용언구 전체를 하나의 표제어로 삼아 인식을 수행하게 되는데 자유로운 문형을 표현하기 어려울 뿐만 아니라 화자의 자연스러운 문장 구성을 어렵게 만드는 요인이 된다. 이를 고려하여 한국어 술어 구조를 의사 형태소 단위로 재정리하면 다음과 같다.

술어 = 어간 + 보조용언 + 어말어미  
 보조용언 = 양상정보  
 어말어미 = 선어말어미 + 종결어미 + 종결보조사

용언구를 하나의 어절로 묶어서 사전에 기재하는 것 보다는 다양한 어미의 활용을 위해 어간과 어미를 분류하여 인식과정에서 결합하고, 다양한 문형의 표현을 위하여 양상정보를 포함하고 있는 보조 용언구를 하나의 인식단위로 선정하였다. 이렇게 어간과 보조용언, 그리고 어말어미의 결합으로 활용이 자유로운 의사 형태소 단위의 문장을 구성할 수 있게 된다.

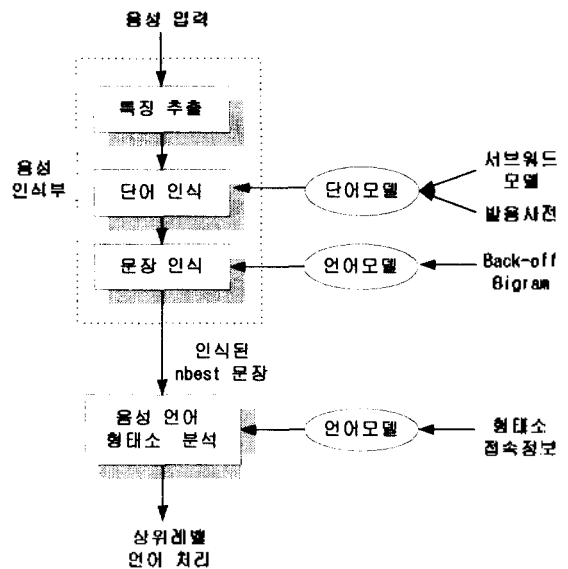
또한 기존의 한국어 형태소 품사태그 집합으로는 표기하기 힘든 경우들이 대화체 문장에서 자주 발생하게 되는데, 대표적으로 준말 표현의 경우 음성언어 분석을 위한 태그가 필요하다. 대화체에서 자주 발생하는 '뭐지', '뭐라고', '해줘' 등은 '무엇이지', '무어라고', '무엇이라고', '하여주어'의 준말로써 원형으로부터 음성 인식을 수행하기 힘들 뿐만 아니라 그러한 노력을 통하여 반복된 되돌림 작업을 수행할 필요가 없으므로 처음부터 구어체 기반의 사전을 구성하여 준말을 그대로 인식단위로 사용하는 것이 경제적이다. 발음상 정확한 줄임말 표현이 되어 정보를 서로 공유될 수 있는 것들이 있는가 하면 '마지막 게 뭐지에서' '게'의 경우 '것이'의 준말로 실질 형태소와 형식 형태소가 발음상의 경계가 없이 나타나게 되고 또한 그 음운변화상의 규칙을 찾을 수가 없게 되어 '명사+조사'의 구성으로 품사를 선정한다는 것

이 매우 모호해진다. 따라서 새로이 준말에 대한 새로운 품사의 선정 및 도입이 필요하며 자주 사용되는 어휘에 대한 준말사전이 필요하게 된다. 이에 따른 규칙이 설정되면 대화체 문장의 분석이 용이해질 것이다.

## 2.2 의사 형태소 단위의 한국어 연속음성인식

본 논문에서 구현하고자 하는 의사 형태소 단위의 한국어 연속음성인식 시스템 구성도는 [그림 1]과 같다.

대용량 어휘 인식을 구현하기 위해 고려되어야 할 분야 중의 하나는 이에 적합한 인식 단위에 대한 연구이다[4]. 한국어의 경우 어형의 변화와 첨가가 심하고 음운변화 현상이 다양하기 때문에 많은 음소 문맥이 고려된다. 조음현상에 민감하고 공유성을 통해 혼련성의 단점을 극복할 수 있는 트라이폰과 같은 문맥종속단위가 현재로서는 가장 우수한 인식단위로 알려져 있다. 트라이폰을 인식단위로 삼는 경우 데이터 부족현상이 발생하게 되는데 이를 해결하기 위해 조음 방식에 따른 분류에 따라 자음군들의 노드를 공유하도록 신장한 Tied-stat 트라이폰을 인식단위로 삼았다.



[그림 1] 전체 시스템 구성도

인식기는 입력 발화에 대해 발음 사전을 참조하여 확률적으로 높은 HMM 모델 순서와 매칭이 되는 의사 형태소를 인식하게 된다. 의사 형태소 단위로 인식이 이루어지면 언어 모델을 이용하여 문장을 구성해 나간다. 이 때 언어 모델을 사용하는 것은 특정 단어와 단어의 결합 가능성이 정의된 문법에 의해 제한함으로써 탐색

## 의사 형태소 단위의 연속 음성 인식 시스템

해야 할 단어의 수를 줄일 수 있을 뿐만 아니라 좀 더 정확한 인식이 이루어지도록 하기 위해서 이다. 문장 인식 단계를 거쳐 인식된 문장을 기반으로 형태소 해석을 수행하게 된다.

### 2.3 음운변화 현상을 반영한 발음사전의 구성

한국어 음운변화 현상을 반영하기 위한 처리 방법에는 대표적으로 두 가지 방법을 들 수 있다. 하나는 인식된 음소열 간의 결합을 보고 음운변화 역주적 규칙 [7]을 사용하여 예상되는 표제어를 사전에서 찾아내는 규칙적 처리 방법과 표제어에 대해 가능한 모든 발음열을 사전에 기재하여 사전에 정의된 표제어 단위로 인식하는 사전적 정의에 의한 처리를 들 수 있다. 본 논문에서는 다중 발음열 발음사전을 사용한 사전적 처리 방식으로 한국어 음운변화 현상을 처리하였다.

Short pause와 silence 등 포함하여 45 개로 구성된 유사 음소 단위(PLU, phone-like unit)로 발음사전을 표현하였으며, 이중 모음의 경우 반모음과 단모음의 결합으로 보았다. 음성언어의 경우 두개의 형태소가 하나의 어절을 이룰 때 음운이 생략되거나 축약되는 현상이 빈번히 발생하므로 형태소의 경계가 모호하게 된다. 이때 축약되는 음절의 경우 음운적으로 볼 때 이중모음이 대부분이므로 반모음 이후를 중심으로 형태소의 경계를 삼기 위해 반모음과 단모음의 결합형태로 본다.

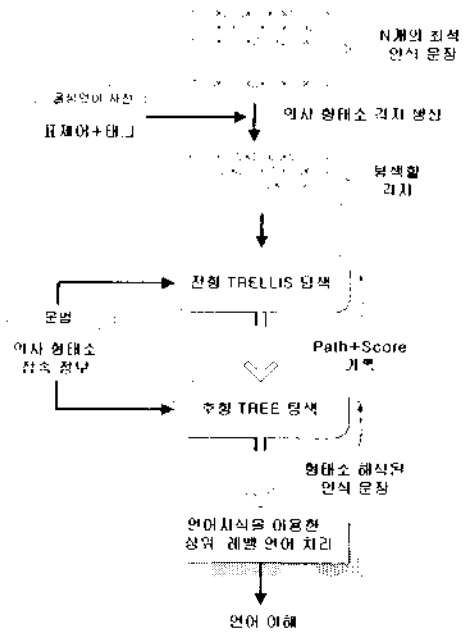
표제어 내부에 대해서는 18 가지 음운변화 현상[3]을 모델링 하였으며 표제어 간 - 의사 형태소 간에 발생하는 음운변화 현상은 크게 체언과 조사, 어간과 어미 또는 보조 용언, 서술성 보통명사와 동사파생접미사 간에 생기는 음운변화 현상을 모델링 하여 발음사전을 구성하였다. 또한 용언의 불규칙 현상[8]의 경우, 발음에 기반한 어간과 어미의 이형태를 사전에 등록하는 형식으로 발음사전을 모델링 하였다.

### 2.4 음성언어 형태소 해석

[그림 2]는 Tree-Trellis 알고리즘[1]에 기반하여 인식된 N 개의 최적 문장을 탐색하며 음성 언어 형태소 해석을 수행하는 전체적인 구성도이다. 여기서는 인식된 의사 형태소 격자를 기반으로 인식을 수행하게 되며[5], 각 격자에는 시작 시점과 끝 시점, 발견 가능성(Probability)가 포함되어 있다.

기본적으로 두가지 탐색 단계를 거친다. 첫번째 단계는 Modified Viterbi 알고리즘의 변형인 Trellis search 알고리즘으로 전향·시간 동기적 탐색을 수행하며, 길과

로 N 개의 최적 격자들을 생성하고 그에 대한 부분 경로(Partial Path)와 Score 를 저장한다. 이때 음성언어 사전에서 표제어의 품사를 찾아낸 뒤 서로 인접한 의사 형태소 쌍이 결합 가능한지를 의사 형태소 접속정보를 참조하며 수행한다. 두번째 단계는 A\* 알고리즘 또는 stack 알고리즘의 변형인 tree search 알고리즘으로 후향·시간 비동기적 탐색이다. 전방향 탐색에서 구해진 부분 경로값과 현재까지의 역방향 부분 경로값의 합을 사용하여 전 경로(Full Path)에 대한 정확한 확률값을 알 수 있게 되므로 불필요한 경로의 확장이 줄게 된다. 이를 기반으로 격자들의 결합을 문장 결과로 내놓는다. 이때 격자가 가지고 있는 시간 정보와 의사 형태소 태그를 함께 출력한다.



[그림 2] 한국어 음성언어 형태소 해석

## 3. 실험 및 평가

인속 HMM 을 기반으로 한 화자 독립 시스템으로 Entropic 사의 HTK(Hidden Markov Toolkit)[6]를 사용하여 전반적인 인식 실험을 수행하였다. 본 실험에서는 16 KHz로 샘플링된 음성신호에 대해 25msec의 해밍 윈도우를 사용하여 프레임 분석을 한 후, 프레임 단위로 13 차 mel-frequency cepstral coefficients, 13 차 delta-cepstral coefficients, 13 차 delta-delta cepstral coefficients로 구성된 특성벡터를 학습 및 인식을 위한 벡터로 사용하였다.

단어 인식 과정에서는 발음사전과 서브워드 모델을 사용한다. 발음사전은 본문에서 설명한 바와 같이 구성하였으며 서브워드 모델은 각 유사 음소에 대해 학습된 HMM이다. 이 HMM은 유사 음소와 같은 모노폰의 형태에서 확장된 음소 문맥 정보를 포함하는 트라이폰으로 학습되어졌다. 단어인식이 이루어지면 언어모델을 이용하여 문장 인식을 수행한다. 여기서는 back-off bigram 모델을 사용하여 실험을 수행하였다.

실험 A는 호텔예약 영역 29 종류의 문장에 대해 5명의 화자가 5번씩 발화한 725문장과 추가로 8 종류의 문장에 대해 앞에 화자 중 동일 1명과 다른 2명의 화자가 5번씩 발화한 120문장에 대한 실험으로 총 845문장에 대한 실험 결과이며, 실험 B는 위의 학습 참여자 7명의 학습에 참여하지 않은 169개의 발화문장에 대한 인식실험 결과이다.

문장을 구성하는 디코딩 단위의 선정 기준은 다음과 같다. 어절은 띄어쓰기 단위이며, 단어(i)은 음성인식 실험에서 일반적으로 사용되는 단위로 용인구를 제외한 어절에 대해 내용어와 기능어를 분리한 단위이며, 단어(ii)는 단어(i)과 같은 구성으로 단어의 조합으로 이루어지는 모든 단위를 나눈 경우이며, 형태소는 문자기반의 형태소로 5가지의 경우에 대해 실험을 수행하였다.

[표 2] 실험 A의 문장 인식 결과

디코딩 단위		어절	단어 (i)	단어 (ii)	의사 형태소	형태소
단어	Correctness	98.22	96.87	95.81	96.55	92.49
	Accuracy	96.93	95.08	92.68	94.52	88.94
문장	Correctness	91.83	83.67	71.95	79.88	53.14

[표 3] 실험 B의 문장인식 결과

디코딩 단위		어절	단어 (i)	단어 (ii)	의사 형태소	형태소
단어	Correctness	98.02	96.15	94.51	94.82	90.90
	Accuracy	96.53	93.29	90.36	92.03	86.84
문장	Correctness	91.02	79.88	66.27	69.82	49.11

실험 결과를 보면 전반적으로 디코딩 단위에 따른 단어 인식률의 차이는 크게 나타나지 않았다. 반면 문장 인식률 면에서 보면 형태소를 디코딩 단위로 삼았을 경우에 인식률이 현저하게 떨어지는 것을 볼 수 있다. 이는 짧은 음성신호 상의 문제 뿐 아니라, 하나의 문장을 구성하게 되는 디코딩 단위의 개수에 따라 영향을 받은 것으로 구성수가 n배로 증가되는 경우 탐색 영역은 2<sup>n</sup> 배로 증가하게 된다. 따라서 계산량이 많이짐에 따라 속도가 떨어지고 인식률의 정확도가 떨어지게 된다. 참고로 한 문장을 구성하는 평균 디코딩 단위수는 다음과

같다. 어절은 3.62, 단어(i)은 6.05, 의사 형태소는 6.19, 형태소는 7.81이다. 반면 의사 형태소의 경우에는 인식률면에서 단어(i)과는 약간의 차이가 나타나지만 단어(ii)와 비교할 경우에는 거의 차이가 없음을 볼 수 있다.

음성 인식기에서 고려되는 것은 인식을 뿐만 아니라 실시간 처리를 위해 그 인식속도도 중요한 고려 대상이 된다. 인식기에서의 인식속도는 일반적으로 사전의 크기에 비례하게 된다. 사전의 표제어 단위에 따른 사전 entry 수이다.

[표 4] 사전의 표제어 단위에 따른 사전 크기

실험영역	어절	단어(i)	의사 형태소	형태소
호텔예약	104	99	115	108
LG 음성 DB	1192	811	765	762

호텔예약 영역의 경우 문장의 수가 적기 때문에 사전의 크기를 비교하기에 적당하지 않으므로 대용량 음성 인식시 발생하는 문제를 살펴보기 위해 LG 음성 DB 중 1000문장을 선정하여 분석하였다. 어절 단위의 경우 사전 entry 수가 가장 크며 형태소의 경우 가장 작다. 1000문장에 대한 크기 비교이지만 대규모 음성 인식 실험에서는 기능어의 수가 한정되어 있으므로, 사전에는 고유명사나 용언의 어간과 같은 새로운 미등록어가 첨가되므로 어절과 형태소 사이의 사전크기가 크게 차이를 추측할 수 있다.

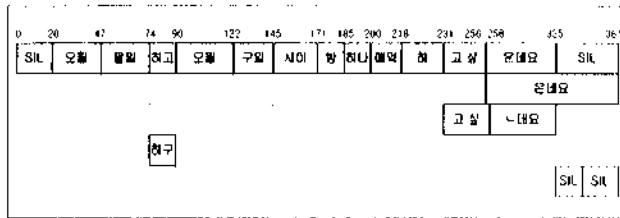
다음 [표 5]는 의사 형태소 단위로 인식 실험을 수행하였을 경우 N-best 문장인식 결과이다. 문장 순위별 아무리 늘려도 인식률의 증가에는 한계가 있음을 볼 수 있다. 이는 기본적으로 인식이 어려운 사항들이 포함되어 있기 때문이다. 따라서 형태소 해석시 상위 5문장을 기반으로 실험을 수행하였다.

[표 5] 실험 A의 N-best 문장인식 결과

문장순위	1	2	3	4	5
hit 개수	675	60	26	6	6
누적수	675	735	761	767	773
문장 인식률	79.88	86.98	90.06	90.77	91.48

문장순위	6	7	8	9	10
hit 개수	1	7	2	1	0
누적수	774	781	783	784	784
문장 인식률	91.60	92.43	92.66	92.78	92.78

다음은 실험에 사용된 "오월 팔일하고 오월 구일 사이 망 하나 예약하고 싶는데요."라는 입력문장에 대한 음성언어 형태소 해석 과정이다.



[그림 3] 형태소 해석의 입력값인 의사 형태소 격자

[그림 3]은 인식된 상위 5 번째까지의 문장 결과를 기반으로 채구성된 중간 탐색 격자이다. 앞장에서 기술된 trec-trellis 알고리즘에 기반하여 위의 예제 문장에 대해 다음과 같은 형태소 해석 결과를 얻었다.

오월/nct + 팔일/nct + 하고/jcj + 오월/nct + 구일/nct + 시이/nbn + 방/mcn + 하나/nnc + 예약/ncpa + 하/xsv + 고 싶/px + 은대요/ef

#### 4. 결론 및 향후 연구과제

본 논문에서는 형태소 해석을 수반하는 한국어 연속 음성 인식 시스템을 구현하였다. 음성언어 형태소 해석을 위해 문장을 구성하는 새로운 디코딩 단위가 필요함에 따라 의사 형태소(Pseudo-Morpheme)를 정의하였고, 이를 기반으로 인식실험을 수행하였다. 디코딩 단위가 어절이나 단어의 경우에는 의사 형태소나 형태소 보다 인식률 면에서는 좋으나, 형태소 결과를 얻기 위해서는 별도의 후처리 과정을 거쳐서 형태소 분석을 수행해야만 한다. 그러나 의사형태소나 형태소로 인식하는 경우에는 별도의 형태소 분석 과정 없이 형태소 단위로 결과를 얻을 수 있는 이점이 있다. 디코딩 단위가 분자기반의 형태소인 경우에는 발성시간이 짧은 기능어로 인해 인식률이 현저히 낮았으나 이러한 문제점을 보완한 의사 형태소의 경우 약 20% 이상의 인식률 향상을 얻었다. 상위 5 개의 인식 분장에 대해 재탐색시 한국어 대화체 문형의 특성을 고려한 접속정보를 사용하여 가장 올바른 형태소 해석 결과를 탐색하였다.

앞으로 안정화된 베이스 라인 시스템 구성과 더불어 격자 탐색 경로에 대한 좀 더 분별력 있는 score 를 구하기 위해서는 통계적 정보와 상위 언어 모델이 요구된다. 더 나아가서는 이러한 정보를 이용하여 오인식된 문장을 수정하는 기능을 추가할 수 있다. 또한 현재 실험에 사용된 접속 정보는 영역 의존적으로 구성되었지만, 앞으로 보다 많은 대화체 문형에 대한 연구에 기반하여 영역 독립적인 형태소 접속정보를 규칙화 할 필요가 있다.

#### 참고문헌

[1] F. K. Soong and E.-F. Huang, "A Tree-Trellis Based Search for Finding the N-best Sentence Hypothesis in Continuous Speech Recognition", *Proceeding of the DARPA Speech and Natural Language Workshop*, 1990, pp.12-19.

[2] Egidio Giachin, Chin-Hwa Lee, Lawrence R. Rabiner, Aaron E. Rosenberg, Roberto Pieraccini, "On the use of interword context-dependent units for word juncture modeling," *Proceedings of the DARPA Speech and Natural Language Workshop*, Vol.6, pp.197-213, 1992.

[3] Jehun Jeon, Sunhee Wee, and Minhwa Chung, "Generating Pronunciation Dictionary by analyzing Phonological Variations Frequently Found in Spoken Korean", *Proceedings of International Conference on Speech Processing*, 1997.

[4] Kai-Fu Lee, *Automatic Speech Recognition : The Development of the SPHINX System*, Kluwer Academic Publisher, Boston, 1989.

[5] Pierre Dupont and Ronald Rosenfeld, *Lattice based language models*, Ph.D. Thesis, CMU-CS-97-173, 1997.

[6] Steve Young, Julian Odell, Dave Ollason, Valtcho Valtcher, Phil Woodland, *The HTK Book (for HTK Version 2.1)*, Entropic Cambridge Research Laboratory, 1997.

[7] Youngsung Jang, Kyongnim Lee, and Minhwa Chung, "Morphological Analysis of Spoken Korean Using Phonemic Change Tracing Rules," *Proceedings of International Conference on Speech Processing*, 1997.

[8] 강승식, 김영택, "한국어 형태소 분석기에서 불규칙 용언의 분석 모형," *한국정보과학회 논문지*, Vol.19, No.2, pp.151-164, 1992.

[9] 김재훈, 서정연, "자연언어 처리를 위한 한국어 품사 태그," 한국과학기술원, 인공지능연구센터, CAIR-TR-94-55, 1994.

[10] 김재훈, "음성인식과 기계번역의 통신규정 : 의사형태소 정의," 한국전자통신연구원, TM97-1420, 1998.

[11] 안동연, *Corpus 를 기반으로 하는 한국어 슬어의 양상 생성*, 박사학위논문, 한국과학기술원, 1995.

[12] 이상호, 김재훈, 조경미, 서정연, "한국어 품사 모호성 해소를 위한 통계적 모델," *제 11 회 음성 통신 및 신호처리 워크샵 논문집*, pp.71-74, 1994.

[13] 장두성, 종영국, 김우성, 구병완, "형태소 분석을 사용하는 한국어 음성인식 시스템," *제 13 회 음성통신 및 신호처리 워크샵 논문집*, 1995.

[14] 최기선, 남영준, 김진규, 한영균, 박석분, 김관수, 이준택, 김덕봉, 김재훈, 최경진, "한국어 정보베이스를 위한 형태-품사 태그 표준에 관한 연구," *한국언어과학회 논문지*, Vol.7, No.4, pp.43-61, 1996.