

한국어 자소 음가 분류에 관한 연구

유승덕, 김학진, 김순협
광운대학교 컴퓨터공학과

A Study on the Korean Grapheme Phonetic Value Classification

Seung-Duk Yu, Hack-Jin Kim, Soon-Hyop Kim
Department of Computer Engineering, Kwangwoon University

요 약

본 논문에서는 한국어 대용량 음성인식 시스템의 기초가 되는 자소(grapheme)가 지니는 음가를 분류하였다. 한국어 자소를 음성-음운학적으로 조음 위치와 방법에 따라 분류하여, 그 음가 분석에 관한 연구와 함께 한국어 음성인식에서 앞으로 많이 논의될 청음음성학(auditory phonetics)에 대하여 연구하였다.

한국어는 발음상의 구조와 특성에 따라 음소 분리가 가능하여 초성, 중성, 종성 자소로 나눌 수 있다. 본 논문에서 초성은 자음음소 18개, 중성은 모음 음소(단모음, 이중모음) 17개, 그리고 'ㅅ' 추가 8중성체계의 자음음소로 하였다. 청음음성학적 PLU(Phoneme Like Unit)의 구분 근거는 우리가 맞춤법 표기에서 주로 많이 틀리는 자소(특히, 모음)는 그 음가가 유사한 것으로 판단을 하였으며, 그 유사음소를 기반으로 작성한 PLU는 자음에 'ㅅ' 중성을 추가하였고, 모음에 (ㄱ, ㅎ)를 하나로, (ㅂ, ㅌ)를 하나로, 그리고 모음(ㄴ, ㄷ, ㄹ)을 하나의 자소로 분류하였다. 혀의 위치와 조음 방법과 위치에 따라 분류한 자음과 모음의 자소를 HTK를 이용하여 HMM(Hidden Markov Model)의 자소 Clustering하여 그것의 음가를 찾는 결정트리를 검색하여 고립어인식과 핵심어 검출 시스템에 적용 실험한 결과 시스템의 성능이 향상되었다.

I. 서 론

근래에 들어 정보통신의 수단과 방법이 급속도로 발전하면서, 음성인식에 있어서도 대용량 음성인식의 필요성이 대두되고 있으며, 많은 연구소와 학교에서 대용량 음성인식에 관한 연구가 진행되고 있다.

대용량 음성인식을 하기 위해서는 인식 시스템은 언어모델과 어휘모델, 그리고 음향 모델의 지식 정보를 사용하고 있다. 이 세 가지 요소 중 어휘모델과 음향 모델의 근간이 되는 것은 자소의 음가구분에 따라 많은 인식률의 차이를 보이고 있다. 이러한 점을 고려하여 대용량 음성인식분야(연속 음성 인식, 연결 음성 인식, Diction 등)에서 기초가 되는 한국어의 발성에 따른 조음의 위치와 방법에 따른 자소 음가의 체계적인 분류방법이 필요하다.

본 논문에서는 조음의 위치와 방법에 따른 자소 음가에 따른 분류방법에 관한 연구한 결과를 제시하고자 하며, 음소-레벨의 음성인식에서 음소에 따른 질의어 집합을 구성하기 위한 PLU를 45개로 분류하였으며, 분류 근거는 한국어 맞춤법 표기에서 주로 많이 틀리는 (ㄴ, ㄷ, ㄹ)를 하나로, 그리고 (ㄱ, ㅎ)를 하나로, 그리고 (ㅂ, ㅌ)를 하나의 음소로 분류하였는데, 이것은 청음음성학(auditory phonetics)적인 분류를 강조한 것이다[4].

한국어 표준발성법에서 7중성으로 분류하고 있는 것을 청음음성학적으로 접근하여 본 실험에서는 중성 음가의 'ㅅ'의 유효성을 검토하였으며, 일차 실험으로 사이시옷 현상이 나는 것을 'ㄷ'이 아닌 'ㅅ'중성으로 구분하여 'ㅅ'추가 8중성 체계로 하여 실험을 하였으며, 사이시옷은 'ㅅ' 중성이 우수한 결과를 얻어, 'ㅅ'추가 8중성 체계 분류 방법을 제안 하고자 한다.

II. 한국어 자소의 분류방법

한국어는 한국어의 자소는 자음 19개와 모음 21개로 구성되어 있다. 이 자소를 혀의 위치와 입 모양에 따라 모음을 구분하고, 조음의 위치와 방법에 따라 자음을 구분한다[14].

국제음성표기법(IPA)에 의하면 자음의 분류는 조음위치에 따라 양순음(bilabial),순치음(labiodental),치음(dental),치경음(alveolar),경구개치경음(postalveolar),권설음(retroflex),경구개음(palatal),연구개음(velar),구개수음(uvular),인두음(pharyngeal),성문음(glottal)으로 분류하며, 조음방법에 따라 파열음(plosive), 비음(nasal),전동음(trill),설탄음(tap/flap),마찰음(fricative),설측마찰음(lateral fricative),접근음(approximant),설측 접근음(lateral approximant)으로 분류하고 있다. 그러나, 순치음(labiodental), 권설음(retroflex), 구개수음(uvular), 인두음(pharyngeal)은 한국어에 없는 자음이다[4][12].

본 논문에서는 IPA와 TIMIT DB와 영문음성학, 국문음성학 문헌을 참조하여 한국어 자소를 다음과 같이 분류하였다. 우선 자음의 분류는 조음방법과 위치에 따라 구분 하였으며, 모음은 혀의 위치와 입술의 모양으로 구분하였다[2][3][4][6][7][8][9][11].

2.1. 자음의 자소 음가 분류

자음은 폐에서 올라온 기류가 심도로부터 방해를 받아서 발생하는 음으로, 우리가 각 자음의 조음 과정을 이해하기 위해서 기류가 구강의 중앙부에서 어떠한 방해(조음방법)를, 성도의 어느 부분(조음위치)에서 받게 되는가와, 나머지 한 가지 중요한 것은 그 자음의 조음이 어떠한 발성유형을 갖는가인데, 이것은 IPA에서도 자음의 조음에 중요한 세 변수들에 따른 분류하고 있다. 유성음의 서로 다른 유형이 존재하듯이 무성음도 서로 다른 발성 유형이 존재한다. 언어에 따라서는 자음의 조음에서 이러한 발성 유형의 차이가 소리를 변별하는 데 아주 중요한 역할을 한다.

같은 조음방법과 조음 위치를 가지고 조음되는 자음이라도 조음시 성대의 상태가 어떠한가에 따라서 서로 다른 질의자음이 만들어질 수 있기 때문이다. 서로 다른 발성 유형의 모음은 변별력이 떨어지나, 그 반면 자음의 발성유형의 차이는 아주 많은 언어에서 체계적으로 소리를 변별하게 하는 중요한 특성으로 사용되고 있다[4].

2.1.1 조음에 따른 분류

한국어의 자음은 조음방법에 따라 파열음, 마찰음, 파찰음, 비음, 유음으로 분류한다. 파열음의 조음방법은 두 조음기관으로 구강의 한 부분을 막는 것과 막음을 지속하는 것, 구강을 개방하는 단계에 의해 발생된다. 파열음은 조음위치에 따라 양순음, 치조음, 그리고 연구개음으로 나뉘고, 긴장(tense) 및 기식성(aspiration) 유/무에 따라 파열음으로 나뉜다. (ㅂ, ㄷ, ㄱ)는 조음기관의 긴장을 없이 약간의 기식성을 동반하는 연음이고, (ㅍ, ㅌ, ㅋ)는 조음기관의 긴장과 강한 기식성을 수반하는 유기경음이고, (ㅃ, ㄸ, ㄱ)는 조음기관의 긴장을 있고, 기식성은 거의 없는 무기경음이다. 우리말의 파찰음은 모두 경구개에서 조음된다. 마찰음은 파열음과 같이 막음 및 지속단계, 개방단계와 개방 후 마찰단계로 발생된다. 비음은 구강을 막고 연구개를 내려 폐로부터 나오는 기류를 비강으로 내보내어 발생되는데, 조음위치에 따라 양순음, 치음, 연구개음으로 나뉜다[11].

표 1. 한국어 조음 방법과 위치에 따른 자음분류

위치 방법	양순음	치경음	경구개음	연구개음	성문음
파열음	ㅂ, ㅍ, ㅃ	ㄷ, ㅌ, ㄸ		ㄱ, ㅋ, ㆁ	
마찰음		ㅅ, ㅆ			ㅎ
파찰음			ㅈ, ㅊ, ㅉ		
비음	ㅁ	ㄴ		ㅇ	
설측음		ㄹ			

표 1에서와 같이 자음은 크게 조음의 위치와 조음 방법에 따라 분류하며, 그 외에도 세분화하여 분류할 수 있다. 표 2는 여러 참고 문헌과 실험에 의하여 분류한 자음 음가 분류 표이다[4][11][13][14].

표 2. 한국어 자음 음소 리스트 구분

음가구분	자음 음소 리스트
연음	ㄱ, ㄷ, ㅂ, ㅈ
비음	ㄴ, ㅁ, ㅇ
설측음	ㄹ
격음	ㅍ, ㅌ, ㅃ, ㅆ, ㅉ
평음	ㄱ, ㄷ, ㅂ
경음	ㅅ, ㅈ, ㅋ, ㅌ, ㅍ
변이음	ㄱ, ㄴ, ㄷ
마찰음	ㅅ, ㅆ, ㅎ
파찰음	ㅈ, ㅊ, ㅉ
유음	ㄹ
양순음	ㅂ, ㅃ, ㅍ, ㅁ
치경음	ㄴ, ㄷ, ㄸ, ㄹ, ㅅ, ㅆ, ㅌ
치경경구개음	ㅈ, ㅊ, ㅉ
치경마찰음	ㅅ, ㅆ
치경파열음	ㄷ, ㄸ, ㅌ
연구개음	ㄱ, ㅋ, ㆁ, ㅇ
성문음	ㅎ
파열음	ㅂ, ㅃ, ㅍ, ㄷ, ㅌ, ㄸ, ㄱ, ㅋ, ㆁ
무성자음	ㅅ, ㅈ, ㅋ, ㅌ, ㅍ, ㅎ
유성자음	ㄱ, ㄴ, ㄷ, ㄹ, ㅁ, ㅂ, ㅇ, ㅈ

표 2에서는 자음 자소분석을 정리한 것의 일부를 보여주고 있다. 이 자음의 분석을 보다 많은 실험과 학자들의 자문이 필요하다고 판단된다.

2.2. 모음의 자소 음가 분류

모음은 성대의 진동으로 작은 공기가 성도에서 공명을 일으켜 만들어지는데, 이때 성도가 어떠한 모양을 하고 있는가에 따라서 소리의 질이 결정된다. 자음의 조음에 비하여 모음의 조음은 단순한 편이다.

모음의 조음에서 가장 중요한 변수들은 혀의 고저와 전후 위치, 그리고 입술의 모양 이 세 가지가 모음의 조음에 있어서 가장 중요한 변수로 작용한다.

혀의 고저에 의하여 모음이 결정되는 데, 입을 적게 벌리고 내는 소리는 폐모음(close), 혹은 고모음(high), 입을 크게 벌리고 내는 소리를 개모음(open), 저모음(low), 고모음과 저모음의 사이에서 나는 소리를 중모음(mid)이라고 한다.

혀의 전후 위치에 따라, 앞쪽에 있으면 전설모음(front), 뒤쪽에 있으면 후설모음(back), 그 중간에 있는 중설모음(central)이라고 한다.

마지막으로 모음의 조음에서 중요한 변수가 되는 것은 입술의 모양이다. 조음 시 입술이 둥그렇게 모아져 앞으로 튀어나오는 모음을 원순모음(rounded)이라고 하고, 그러지 않은 모음을 평순모음(unrounded)이라고 한다[4][6][7].

표 3. 혀의 높이 및 위치에 따른 단모음의 분류

위치 높이	전설		중설		후설	
	평순	원순	평순	원순	평순	원순
고	ㅣ	ㅍ	ㅡ			ㅗ
중	ㅑ, ㅓ	ㅕ	ㅑ			ㅛ
저			ㅓ			

표 3과 같이 모음은 혀의 높이와 위치에 따라 분류한다[4][6][7][14]. 다음의 표 4는 여러 참고 문헌과 실험에 의하여 분류한 모음 음가 분류 표이다.

표 4. 한국어 모음 음소 리스트 구분

음가구분	모음 음소 리스트
단모음	ㅏ, ㅑ, ㅓ, ㅕ, ㅗ, ㅛ, ㅜ, ㅠ, ㅡ, ㅝ, ㅞ, ㅟ
이중모음	ㅑ, ㅓ, ㅕ, ㅗ, ㅛ, ㅜ, ㅠ, ㅝ, ㅞ, ㅟ, ㅑ, ㅓ, ㅕ, ㅗ, ㅛ, ㅜ, ㅠ, ㅝ, ㅞ, ㅟ
저모음	ㅏ
중모음	ㅑ, ㅓ, ㅕ, ㅗ, ㅛ, ㅜ, ㅠ
고모음	ㅝ, ㅞ, ㅟ
전설모음	ㅏ, ㅑ, ㅓ, ㅕ, ㅗ, ㅛ, ㅜ, ㅠ
후설모음	ㅝ, ㅞ, ㅟ
중설모음	ㅑ, ㅓ, ㅕ
전설-평순	ㅏ, ㅑ, ㅓ, ㅕ
후설-평순	ㅝ, ㅞ, ㅟ
후설-원순	ㅝ, ㅞ, ㅟ
원순	ㅑ, ㅓ, ㅕ, ㅗ, ㅛ, ㅜ, ㅠ
평순	ㅏ, ㅑ, ㅓ, ㅕ, ㅗ, ㅛ, ㅜ, ㅠ, ㅝ, ㅞ, ㅟ
활음	ㅑ, ㅓ, ㅕ
A모음	ㅏ, ㅑ
E모음	ㅑ, ㅓ, ㅕ
I모음	ㅑ, ㅓ
O모음	ㅓ, ㅕ
U모음	ㅕ, ㅗ, ㅛ, ㅜ, ㅠ

표 4에서 단모음으로 구분하는 (ㅑ)는 청음음성학적 지식으로 이중모음(diphthong)인 유사음(ㅑ, ㅓ)와 같은 그룹으로 묶었고, 모음의 (ㅓ)도 청음 정보를 가지고 (ㅑ) 묶었으며, (ㅓ)와 (ㅕ)도 같은 지식에 근거하여 묶었다.

III. 실험 및 결과 해석

3.1 음성 데이터 구축 및 실험 환경

표 5. 입력 데이터 설정

설정내용	설정 값
Sampling Rate	8000 Hz
Channel	Mono (1)
양자화 비트수	16 bit PCM (8bit alaw to linear)
전화녹음	A-law, Analog Line
환경	조용한 사무실
음성보드	D/120JCT-LS (Dialogic)

표 5와 같은 조건으로 증편저래 문장 72개와 거래소 상장회사 명을 남성화자 20명 여성화자 10명에게 전화로 녹음 받았다. 녹음 형식은 A-law로 하였으며, G.711의 규약에 의거하여 alaw to linear PCM을 16 bit 변환 음성 DB를 가지고 HTK에서 실험하였다.

HTK에서 적용한 특징 파라미터는 MFCC를 12차로 하였으며, 차의 에너지와 각각의 특징 파라미터에 대하여 델타(delta) Cepstrum을 적용하였다. HMM State는 5로 하였고, Tri-Phone 모델을 사용하였다. 대상 어휘 단어와 문장은 그림 1과 같은 KW음운 변동기를 이용하여 자소열로 전사된다.

KW음운 변동기는 음절의 끝소리, 자음동화, 구개음화, 모음동화, 모음조화, 축약과 탈락, 된소리현상, 사잇소리현상, 그리고 예외 사전으로 구성하였다.

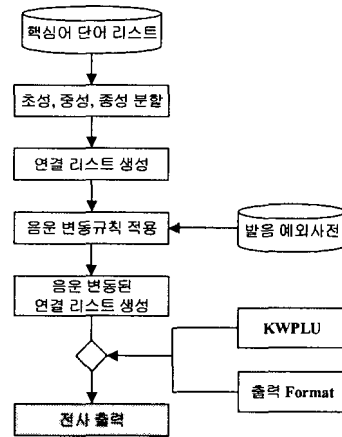


그림 1. KW음운 변동기의 구성

KWPLU는 KW음운변동기에 의한 발음법칙과 예외법칙에 의해 자소열로 구분한 음소를 가지고 적용하였고, 이 음운변동기에는 저자가 제안한 사이시옷현상을 'ㅅ'추가 8중성으로 구분하여 작성하였다.

표 5. KWPLU 집합의 구성

자음	PLU	설명	모음	PLU	설명
ㄱ	gs	초성ㄱ	ㅏ	aa	
	ge	중성ㄱ	ㅑ	av	
ㄴ	ns	초성ㄴ	ㅓ	oo	
	ne	중성ㄴ	ㅕ	uu	
ㄷ	ds	초성ㄷ	ㅡ	ww	
	de	중성ㄷ	ㅣ	ii	
ㄹ	ls	초성ㄹ	ㅞ, ㅟ	ee	청음음성학적
	le	중성ㄹ	ㅑ, ㅓ, ㅕ	we	청음음성학적
ㅁ	ms	초성ㅁ	ㅑ	ui	
	me	중성ㅁ	ㅓ	ya	
ㅂ	bs	초성ㅂ	ㅕ	yv	
	be	중성ㅂ	ㅗ	yo	
ㅅ	sb	초성ㅅ	ㅛ	yu	
	se	중성ㅅ	ㅜ, ㅠ	ye	청음음성학적
ㅇ	ne	중성ㅇ	ㅗ	wa	
ㅈ	js	초성ㅈ	ㅑ	wv	
ㅊ	cs	초성ㅊ	ㅓ	wi	
ㅋ	ks	초성ㅋ			
ㅌ	ts	초성ㅌ	단목음	sp	short term
ㅍ	ps	초성ㅍ	장목음	sil	long term
ㅎ	hs	초성ㅎ			
ㄱ	gg	초성ㄱ			
ㄷ	dd	초성ㄷ			
ㅂ	bb	초성ㅂ			
ㅅ	ss	초성ㅅ			
ㅈ	jj	초성ㅈ			

표 5는 자음과 모음음소의 KWPLU의 구분이며, 장-목음(long-term silence)을 [sil]로 그리고 단-목음(short-term silence)을 [sp]로 구성하였다.

표 2와 표 4를 이용하여 음소 집합을 다음 그림 2와 같이 구성하였다[1][2][3][10].

QS "R_Rounded"	{*+ㅓ, **ㅓ, **ㅓ, **ㅓ}
QS "R_Unrounded"	{*+ㅣ, **ㅣ, **ㅣ, **ㅣ, **ㅣ, **ㅣ, **ㅣ, **ㅣ, **ㅣ, **ㅣ}
QS "R_Avowel"	{*+ㅏ, **ㅏ}
QS "R_Evowel"	{*+ㅓ, **ㅓ, **ㅓ}
QS "R_Ivowel"	{*+ㅣ, **ㅣ}
QS "R_Ovowel"	{*+ㅓ, **ㅓ}
QS "R_Uvowel"	{*+ㅓ, **ㅓ, **ㅓ, **ㅓ, **ㅓ, **ㅓ}

그림 2. 음소 질의어 집합

위와 같은 질의어 집합(음소는 다음 표 5와 같은 KW PLU에 의거하여 영문 표기로 되나, 이해를 돕기 위해 한글 자소로 표기함)을 구성하고, Tied Block과 결정 트리를 거쳐 다음과 같은 실험결과를 얻었다.

질의어 집합의 질의 순서는 각 자소가 지니는 음가의 우선 순위(분류의 큰 범위)로부터 구분한다.

3.2 청음 음성학 지식 기반 실험 결과

표 6. Old KWPLU에 'ㅅ'추가 8중성 인식결과

인식률		7중성 인식률	8중성 인식률
		(/ㅅ/ 중성)	(/ㅅ/, /ㅅ/중성)
남자 A	핵심어	89.46 %	89.75 %
	단어	89.33 %	88.63 %
남자 B	핵심어	82.37 %	83.74 %
	단어	87.52 %	87.92 %
남자 C	핵심어	83.69 %	84.86 %
	단어	92.73 %	93.52 %
여자 D	핵심어	86.82 %	87.48 %
	단어	87.84 %	88.30 %
여자 E	핵심어	81.53 %	82.87 %
	단어	90.34 %	92.70 %

표 6은 종전 KWPLU에서 7중성과 'ㅅ'추가 8중성에 의한 실험 결과로, 인식률이 떨어지는 경우도 있으나, 전체적으로는 약 1% 이상의 성능 향상을 보였다.

표 7. (ㅅ, ㅅ), (ㅅ, ㅅ), (ㅅ, ㅅ, ㅅ) 구분 인식결과

인식률		Old KWPLU	New KWPLU
		(/ㅅ/ 중성)	(/ㅅ/, /ㅅ/중성)
남자A	핵심어	89.46 %	91.22 %
	단어	89.33 %	88.37 %
남자B	핵심어	82.37 %	92.50 %
	단어	87.52 %	89.73 %
남자C	핵심어	83.69 %	85.42 %
	단어	92.73 %	94.38 %
여자D	핵심어	86.82 %	89.70 %
	단어	87.84 %	89.86 %
여자E	핵심어	81.53 %	81.38 %
	단어	90.34 %	92.93 %

New KWPLU실험은 'ㅅ'추가 8중성으로 (ㅅ, ㅅ), (ㅅ, ㅅ), (ㅅ, ㅅ, ㅅ)를 유사음소로 묶어서 실험을 한 결과이다. 표 7에서의 결과와 같이 청음 음성학적 요소를 고려했을 경우 이전보다 약 2% 정도의 성능 향상을 얻었다. 그리고 본 실험이 Off-Line상에서 이루어져서 정확한 실시간 인식시간을 얻지는 못하였으나, PLU의 수를 이전의 51개보다 6개를 줄였으므로 인식 결과 시간 또한 줄일 수 있는 효과도 있다.

IV. 결 론

본 논문에서는 대용량 음성인식분야에서 기초가 되는 한국어의 자소 음가에 따른 분류방법에 관한 연구를 하였으며, 종전의 방법과 비교하면, 본 연구에서 분류한 방법으로 실험을 하였을 경우 약 2%의 성능 향상을 얻었다. 그리고, 본 논문의 실험은 Off-Line 환경의 실험으로 이루어 졌지만, 핵심어 검출 시스템이나 연속음성인식에 적용시 제일 큰 문제가 되는 계산시간 문제를 PLU-Set의 감소로 인하여 인식시간 감소에도 도움을 줄 수 있다.

앞으로의 과제는 청음음성학적으로 음성을 분류하기 위한 보다 체계적인 연구방법의 모색과, On-Line 상에서의 실험과 보다 많은 검증실험으로 'ㅅ'추가 8중성의 유효성 유/무를 검증하는 것과, 현재 45개로 분류한 KWPLU의 구분을 보다 많은 음운학자와 음성학자들과의 논의와 청음 음운학적인 실험을 통하여 분류하는 과제가 남아 있다고 하겠다.

V. 참 고 문 헌

- [1] 김동화, "연속음성인식을 위한 향상된 결정 트리 기반 상태공유기법 연구", 박사학위논문, 부산대학교, 1999
- [2] 전남대, "가변어휘 음성 인식기 구현 및 탐색시간 단축 알고리즘 비교", 석사학위논문, 전남대학교, 2001
- [3] 신영옥, "가변어휘 핵심어 검출 시스템의 구현 및 성능개선", 석사학위논문, 부산대학교, 2001
- [4] 신지영, "말소리의 이해; 음성학·음운론연구의 기초를 위하여", 한국문화사, 2000
- [5] 존 R. 테일러 지음; 조명원, 나익주 옮김, "인지언어학이란 무엇인가?", 한국문화사, 1999
- [6] 한영목 "언어화사회", 한신문화사, 1997
- [7] 남기심, 이정민, 이홍배, "언어학 개론", 탐출판사, 1984
- [8] 이영호, "국어음성학", 태학사, 1996
- [9] 허웅, "국어음운론", 샘문화사, 1965
- [10] Steve Young, Dan Kershaw, Julian Odell, Dave Ollason, Valtcher, Phil Woodland, The HTK Book (for HTK Version 2.2), Entropic Cambridge Research Laboratory, 1999
- [11] 마이크로소프트 12월호, "제5회 음성합성", 1996
- [12] <http://www.arts.gla.ac.uk/IPA/ipa.html>
- [13] 안승신, 한국방송대학교 영어영문학과 <http://spnode10.knou.ac.kr/phonetics/>
- [14] 문교부고시 제88-2호(1988), "국어 어문 규정집", 문화 체육부, 1996