

한국인을 위한 중국어 발음 교정 시스템

김효숙*, 김선주*, 강효원*, 김무중*, 하진영**

* (주) 언어과학 언어공학연구소

** 강원대학교 컴퓨터공학과

Chinese Pronunciation Correction System for Korean learners

HyoSookKim*, SunJuKim*, HyoWonKang*, MuJungKim* and Jin-YoungHa**

*Center for Linguistic Engineering, Eoneo Inc.

**Department of Computer Engineering, Kangwon National Univ.

Email: {hyosook, sunju, kanghyowon, donaldos}@eoneo.co.kr, jyha@kangwon.ac.kr

Abstract

This study is about constructing L2 pronunciation correction system for L1 speakers using speech technology. Chinese pronunciation system consists of initials, finals and tones. Initials/finals are in segmental level and tones are in suprasegmental level. So different method could be used assessing Korean users' Chinese. The recognition rate of initials is 81.9% and that of finals is 68.7% in the standard acoustic model. Differ from native speech recognition, nonnative speech recognition could be promoted by additional modeling using L2 speakers' speech. As a first step for the those task we analysed nonnative speech and then set a strategy for modeling Korean speakers'.

I. 서론

본 연구는 음성기술을 응용하여 L2의 발음체계를 학습하는 시스템을 구성하는 방법에 관한 것이다[1]. 중국어의 발음체계는 성모 및 운모와 성조의 세 가지로 이루어져 있다[2]. 성모 및 운모와 성조는 각각 분절음

층위와 초분절음 층위로 구분되는 현상이므로 서로 다른 방식으로 분석될 수 있다[3]. 실제로 기존의 여러 중국어 음성인식에서 다음과 같이 성모 및 운모와 성조의 인식이 독립적으로 수행되고 있음을 볼 수 있다 [4].

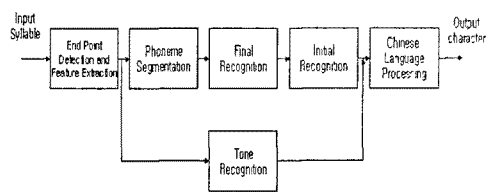


그림 1. 중국어 음절 인식 기본 구조 (cf.[4])

한국인을 위한 중국어 발음교정 시스템을 구성하는 주요 기술은 다음과 같다. 첫째는 음소(성모 및 운모) 모델에 기반을 둔 고립단어인식이며, 둘째는 화자별 음역의 차이를 고려한 성조 패턴 비교이다. 그리고 위의 두 가지 요소의 설계와 구현 및 성능 평가의 단간이 되는 중국어 발음학습 지식베이스로 이루어지 있다.

본고에서는 분절음 층위에서의 교정 기능 구현을 위한 중국어 음소인식 방법론 및 이의 성능 향상을 위한 방법론에 대해 논의하도록 한다.

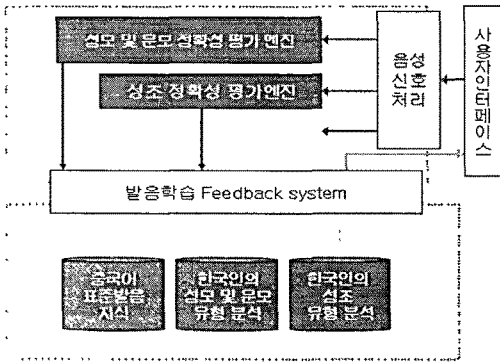


그림 2 한국인을 위한 중국어 발음교정 시스템

II. 표준 중국어 음성DB

언어학습, 특히 발음학습에 있어서는 표준 모델로 삼을 대상을 확정하는 것이 중요하다. 본고에서 중국어 표준 모델을 생성하기 위해 채택한 것은 중국 대륙의 표준어로 정의된 普通话이다[5].

중국어 표준 음성 DB는 SiTEC의 Chinese01, Chinese02를 이용하였으며 DB 각각의 내용은 다음과 같다.

Chinese02는 중국어 음성인식 시스템 개발을 위한 음향 모델 생성을 목적으로 구축된 음성 데이터베이스이다. 발성 목록은 무제한 어휘 인식 시스템에 적합하며, 단어와 문장으로 구성되어 있다. 이 DB는 본 연구에서 중국어 표준 음향모델을 훈련시키는 데 이용되었다.

표 1. Chinese02 DB 남/녀 화자수

성별	남성	여성	합계
화자수	150명	150명	300명

Chinese01 DB는 수출 지원을 위한 외국어 음성 DB로 북경어를 사용하는 현지 중국인 성인 남녀를 대상으로 PC환경에서 병음, 단어, 사연숫자, 날짜 관련 단어, 문장으로 구성된 총 4,752 토큰을 발성목록으로 구성하여 수집한 것이다. 이 DB는 본 연구에서 중국어 표준 음향모델의 성능을 테스트하는 데 이용되었다.

표 2 Chinese01 DB 남/녀 화자수

성별	남성	여성	합계
화자수	120명	180명	300명

III. 표준 성모 및 운모 모델링

중국어 표준 음향모델은 성모 21개 및 운모 38개에¹⁾ 대해 각각 3 states left-to-right continuous HMM을 사용하였다. 38개의 운모는 다시 5개의 성조(4가지 성조 및 경성 포함)로 구분하여 모델링하였다. Pre-emphasis와 10ms마다 25ms 크기의 hamming window를 적용하여 음성데이터로부터 특징벡터를 추출하였다. 특징벡터는 12차 MFCC와 log-energy 그리고 각각의 delta 및 delta-delta를 구한 39차 특징벡터를 사용하였다. 훈련 방식은 음소별 전사 정보를 필요로 하는 boot-strap 방식을 사용하였으며, mixture는 12로 하였다.

IV. 한국 화자의 성모 및 운모 분석

Native speech recognition과 달리 nonnative의 모국어 간섭에 의한 현상을 반영한 음향모델을 구성하기 위해 한국인 화자의 중국어 발화를 수집하였다. 발화 목록은 Chinese01을 참고하였으며 중국어를 배운 지 6개월 이상이 되는 남녀 20대 화자들로부터 수집하였다.

표 3 한국인 화자의 중국어 DB 남/녀 화자수

성별	남성	여성	합계
화자수	100명	100명	200명

중국어는 21개의 성모와 36개의 운모로 구성되어 있다. 운모는 5개(1-4성 및 경성)의 성조로 구별된다. 본고에서는 특히 한국인의 중국어 발음 가운데에서 성모에서 발견되는 특징적인 오류에 한해서 설명하기로 한다. 중국어의 성모와 운모는 다음과 같이 구성되어 있다.

표 4. 중국어 성모 목록(21개)

조음위치	성모
쌍순음	b p m
순치음	f
설첨중음	d t n l
설면후음	g k h
설면전음	j q x
설첨전음	z c s
설첨후음	zh ch sh r

1) 표 5의 36개에 환경을 고려하여 2개의 운모를 추가하였다.

표 5 중국어 운모 목록(36 개)

		i	u	ü	
단운모	a	ia	ua		결합운모
	o		uo		
	e	ie		üe	
복운모	ai		uai		
	ei		uei		
	ao	iao			
비운모	ou	iou			
	an	ian	uan	uan	
	en	in	uen	un	
	ang	iang	uang		
	eng	ing	ueng		
ong	iong				
특수 운모	er				

한국어의 자음 18개(음가가 없는 초성 'ㅇ'를 제외하고)와 조음 위치 및 방법 면에서 차이가 많이 나는 것은 /f, z, c, zh, ch, sh, r/의 7개이다[6]. 이 가운데 특히 한국인 화자의 발음에서 많은 오류가 발생하는 것은 설첨전음/z, c/와 설첨후음(권설음)인 /zh, ch, sh/이다. 한국인에게서 흔히 발견되는 오류의 양상은 권설음을 제대로 발음하지 못하는 것이었다.

V. 발음교정 시스템의 성능 향상

III에 의해 완성된 중국어 표준 음향 모델을 다음과 같은 언어 네트워크를 이용하여 성능을 테스트하였다.

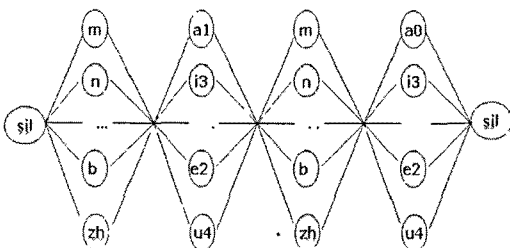


그림 3 'māma(어머니)'에 대한 언어네트워크

이러한 언어 네트워크를 이용한 인식률은 다음과 같다.

표 6 중국어 성모/운모 및 단어 인식률

	성모	운모	단어 ²⁾
인식률	81.9%	68.7%	96%

2) 1200 단어에서 하나를 고르는 task에서의 인식률이다.

그 가운데에서 /c ch/에 대한 confusion matrix는 다음과 같다.

표 7 /c ch/의 confusion matrix(단위 갯수)

	c	ch	q	s	sh	x	z	zh
c	530	31	10	12	0	1	5	0
ch	113	544	16	4	9	5	1	6

위의 confusion matrix에서 확인할 수 있는 것처럼 성모 가운데 c와 ch는 표준 음향모델에서도 가장 구분이 어려운 것 중에 하나이다. 각각의 스펙트로그램을 살펴보면 다음과 같다. 파찰음 구간에서 주파수 대역에 따른 에너지 차이가 관찰된다.

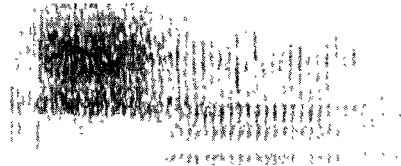


그림 4 설첨후음(권설음) 'ch'의 스펙트로그램

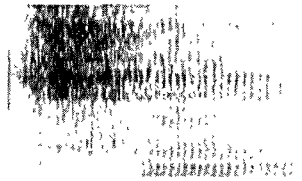


그림 5 설첨전음 'c'의 스펙트로그램

기존의 연구와 본고에서의 한국인의 중국어 DB분석 결과에 따르면 중국어 발음교정 시스템의 성능은 성모에서는 특히 중국어의 /j q x z c s zh ch sh/에 대한 에러 검출의 정확성에 달려 있다고 할 수 있다.

한국인의 중국어 발화에서 발생하는 에러는 대치, 삭제, 삽입 가운데에서 대부분 대치 범주에 해당한다. 어두 자음군이나 어말 자음군을 가지는 등 음절구조가 복잡한 영어를 발화할 때 흔히 가지는 삽입과 삭제 오류는 중국어에서 잘 나타나지 않았다.

발음교정 시스템의 성능 향상을 위해서는 첫째 언어 네트워크의 구성이 발음교정의 목적에 타당해야 할 것이다. 즉, 특정한 성모의 탐색공간은 그 성모와 유사하게 발음되어 한국인들이 흔히 혼동하는 예로 한정되어야 할 것이다. 둘째는 한국인 화자의 일반화된 오류 패턴을 탐색 공간에 포함시켜야 한다. 이를 위해서는 발음교정의 검출된 에러 유형에 적합한 학습 정보를

추출해야 한다.

VI. 결론

본고는 고립단어 인식을 이용한 중국어 발음교정 시스템 구성 중 분절음 층위에 속하는 성모 및 운모에 관한 내용이다. 중국어 화자의 DB를 토대로 이루어진 표준 음향모델의 성능은 성모에서 81.9%, 운모에서 68.7%로 나타났다. 발음교정 시스템에서 LI의 간섭에 의한 에러를 정확하게 검출하기 위한 방법론의 하나는 비모국어 화자의 발화를 분석하고 그 결과를 음향모델에 반영하는 것이다. 본고에서는 한국인 화자의 음성 DB를 수집하여 한국인 화자에게서 발견되는 일관된 오류 유형을 분석하였다.

주후 과제는 이상의 분석 결과를 반영한 한국인의 중국어 음향모델을 추가하여 발음교정 시스템의 성능을 향상시키는 것이다.

감사의글

본 연구는 정동부의 정보통신 산업경쟁력 강화사업의 지원으로 수행되었습니다.

(과제번호: A1200-0400-0013)

참고문헌

- [1] Hincks, R., "Speech technologies for pronunciation feedback and evaluation", *ReCALL* 15(1):3-20, 2003.
- [2] 맹주역, "중국어 발음교육에서 제기되는 문제점에 관하여", *중국연구*, 제 23권,33-44,1993.
- [3] 고미숙, 한중 분절음소와 초분절음소와의 관계 연구, *中國言語研究(第12輯)*, 235-256, 2001.
- [4] Stephen W.K.Fu, C.H.Lee, Orville L.Clubb, " A Survey on Chinese Speech Recognition", 1996.
- [5] 맹주역, "北京語, 台灣國語및 中國語의 標準", *中國言語研究, (第2輯)* 제 2집, 1993.
- [6] 전광진, "중국어 자음의 한글 표기법에 대한 음성학적 대비 분석", *중국문학연구* 제19집, 260-275, 1999