

# 음성인식기를 이용한 한국인의 외국어 발화오류 자동 검출\*

권철홍(대전대), 강효원(대전대), 이상필(대전대)

## <차 례>

- |                                  |                                        |
|----------------------------------|----------------------------------------|
| 1. 서론                            | 3.2. 유사 발음 음소별 HMM 모델<br>및 발음 네트워크의 구성 |
| 2. 언어별 음소인식기                     | 4. 실험 및 결과                             |
| 2.1. 언어별 모노폰 셋의 구성               | 4.1. 언어별 음소인식기 실험 결과                   |
| 2.2. 언어별 음성 DB 구축                | 4.2. 발음 오류 자동 분류 음성인식기<br>실험 결과        |
| 2.3. 언어별 HMM 모델 및 발음<br>네트워크의 구성 | 5. 결론                                  |
| 3. 발음 오류 자동 분류 음성인식기             |                                        |
| 3.1. 유사 발음 음소 셋의 구성              |                                        |

## <Abstract>

### **Automatic Detection of Mispronunciation Using Phoneme Recognition For Foreign Language Instruction**

**Chul-Hong Kwon, Hyo-Won Kang, Sang-Pil Lee**

An automatic pronunciation correction system provides learners with correction guidelines for each mispronunciation. In this paper we propose an HMM based speech recognizer which automatically classifies pronunciation errors when Korean speak Japanese. For this purpose we also develop phoneme recognizers for Korean and Japanese. Experimental results show that the machine scores of the proposed recognizer correlate with expert ratings well.

\* Keywords: Language instruction, Automatic detection of mispronunciation, Phoneme recognition

## 1. 서 론

발음 자동 교정 시스템은 한국인의 외국어 발화시 발음에 따른 오류 양상을 음성인식기로 자동으로 검출하여 오류 유형별로 발음 교정 지침을 제시하는 시스템을 말한다. 발음 오류 유형을 자동으로 분류하기 위한 음성인식기는 오류 유형별 발화 모델들을 미리 수집하여 HMM 모델을 통해 음소인식을 수행한다. 발음 교정 교수법은 언어·음성학 전문 지식을 이용하여 발음 오류에 정확히 해당하는 교정 사항을 제공하여 학습자가 효과적으로 발음 교정을 수행하도록 교정 지침을 제공한다.

기존의 음성인식 엔진을 이용한 발음 학습 시스템은 원어민의 표준발음 모델을 통계적으로 추출하여 미리 제작해 놓은 후, 학습자의 발화 음성이 입력되었을 경우 발화했으리라 추측되는 가장 유사한 음소로 인식하는 것을 목표로 한다. 즉, 학습자의 발음과 원어민의 발음을 특징 신호 추출을 통한 유사성의 판별을 통해 발음의 정확도만을 점수로 제시해 주는 시스템이다. 본 연구에서 제안한 발음 오류 자동 분류 음성인식기는 오류 유형별 발화 음성 데이터를 수집하여 HMM 모델을 통해 발화오류 모델을 만들어 놓고, 학습자의 발화 음성에 대하여 오류 유형별로 자동으로 분류하므로, 그 오류에 정확히 해당하는 교정 지침의 제공을 가능하게 한다.

본 논문에서는 한국인이 일본어 발화시 음소별로 발음 오류를 자동으로 분류하는 HMM 음성인식기를 제안한다. 또한 이를 위하여 각 언어별 음소인식기가 필요하므로 각 언어별 모노폰 셋을 선정하고 언어별 음소인식기를 구현하였다. 본 논문의 구성은, 2장에서 언어별 음소인식기의 구현 방법에 대하여 기술하고, 3장에서 제안한 발음 오류 자동 분류 음성인식기를 설명하고, 4장에서 실험 방법 및 결과를 논하고, 5장에서 결론 및 향후 연구 과제에 대하여 기술한다.

## 2. 언어별 음소인식기

본 장에서는 한국어 및 일본어 음소인식기의 구현 방법에 대하여 설명한다. 언어별 모노폰 셋의 구성, 음성 DB 구축 방법, 음소별 HMM 모델 생성 및 발음 네트워크의 구성 방법은 다음과 같다.

### 2.1. 언어별 모노폰 셋의 구성

#### 2.1.1. 한국어 모노폰 셋

한국어 모노폰 셋은 자음은 변이음을 고려하여 29개를, 모음은 음성학적 차이를 보이는 모음만을 고려하여 17개를 선정하여[1], 총 46개의 모노폰으로 구성하였다. <표 1>에 나타난 한국어 자음 셋에서 첫 번째 열에 자음을, 두 번째 열에 그 기호를 보여주고, 세 번째 열에 그 자음이 속한 단어를 예로 들었다. <표 2>의 한국어 모음 셋에서 첫 번째 열에 모음을, 두 번째 열에 그 기호를 보여준다.

<표 1> 한국어 자음 셋 (29개)

자음	기호	예제	자음	기호	예제	자음	기호	예제	자음	기호	예제	자음	기호	예제
ㄱ	kc	가시	ㅂ	pc	바지	ㅋ	khc	칼날	ㅌ	chc	차레	ㅅ	sc	사장
	gc	대구		bc	나방	ㄱ	kkc	까치	ㅈ	zc	자다		sic	쉽다
	kqc	독자		pqc	입다	ㅌ	thc	타자		zhc	이제	ㅆ	ssc	쓰다
ㄷ	tc	도시	ㅇ	ngc	강사	ㅌ	ttc	따님	ㅈ	zzc	짜다	ㅎ	hc	하지
	dc	사다	ㅁ	mc	감사	ㅍ	phc	파도	ㄹ	lc	사슬		hic	동해
	tqc	받기	ㄴ	nc	간소	ㅃ	ppc	뽀다		rc	다리			

<표 2> 한국어 모음 셋 (17개)

모음	기호	모음	기호	모음	기호	모음	기호	모음	기호
ㅏ	axc	--	euc	ㅑ	jac	ㅓ	jec	ㅕ	wic
ㅗ	eoc	ㅣ	ixc	ㅛ	jeoc	ㅜ	wac	ㅟ	euic
ㅜ	oxc	ㅞ, ㅟ	aec	ㅠ	joc	ㅡ	woc		
ㅠ	uxc			ㅢ	juc	ㅤ, ㅥ, ㅦ	wec		

주) 일본어와 한국어 음소표기의 구분을 위해 한국어 음소표기는 'c'로 끝난다.

### 2.1.2. 일본어 모노폰 셋

일본어 모노폰 셋은 일본음향학회에서 선정한 모노폰 셋을 참조하여[2], 38개의 모노폰을 선정하였다. <표 3> 일본어 모노폰 셋에서 첫 번째 열에 음소가 속하는 음절의 예를 들고, 두 번째 열에 음소의 기호를 보여준다.

&lt;표 3&gt; 일본어 모노폰 셋 (38개)

자음	기호	자음	기호	자음	기호	자음	기호	자음	기호	자음	기호	모음	기호	모음	기호
ぴ	pi	べ	be	さ	sa	も	mo	は	ha	ず	zu	あ	a	あー	ax
ぴよ	pyo	びゅ	byu	し	shi	みゅ	myu	ひよ	hyo	じゅ	ju	い	i	いー	ix
て	te	だ	da	つ	tsu	な	na	ら	ra	반모음	기호	う	u	うー	ux
か	ka	が	ga	ちよ	cho	によ	nyo	りよ	ryo	わ	wa	え	e	えー	ex
きゃ	kya	ぎゃ	gya			h	N		q	や	ya	お	o	おー	ox

주) ax ~ ox 는 장음 모음을, q는 이중 자음을 의미한다.

## 2.2. 언어별 음성 DB 구축

음성 DB는 연구실 수준의 조용한 방에서 1명씩 동시에 3곳에서 녹음하였다. 음성 데이터를 PC에서 수집하였고, 사운드카드는 Soundblaster Audigy를, 마이크는 SHURE 565SD를 사용하였다. 한국어 음성 DB는 SiTEC에서[3] 작성한 한국어 PBW (Phonetically Balanced Words) 452개 단어를 대전대학교 대학생 70명을 대상으로 수집하였다. 일본어 음성 DB는 ATR(일본 자동통역 연구소)에서 작성한 PBW 216개 단어를 고려대학교에서 한국어 연수 중인 일본인 70명을 대상으로 녹음하였다. 각 언어별 음성 DB에서 50명의 음성 DB를 HMM 모델 생성을 위한 훈련용으로, 20명분을 인식기의 성능을 테스트하기 위한 음성 DB로 사용하였다.

## 2.3. 언어별 HMM 모델 및 발음 네트워크의 구성

한국어 및 일본어 음소인식기를 다음과 같은 방법으로 구현하여 성능을 평가하였다. 음성신호의 분석은 매 10msec 마다 25msec의 Hamming 창함수를 사용하여 MFCC 39차를 추출하였고, HMM의 구조는 3 state left-to-right continuous HMM을 사용하여 언어별 음소별 음향모델을 생성하였다. 일본어 음소인식기에서 사용한 발음 네트워크는 <표 4>와 같다. <표 4>에서 < >는 한 번 또는 둘 이상의 반복을 의미한다.

&lt;표 4&gt; 일본어 음소인식기의 발음 네트워크

```

$phones = a | i | u | e | o | ax | ix | ux | ex | ox | w | y | N | q |
          p | py | t | k | ky | b | by | d | g | gy | s | sh | ts | ch |
          m | my | n | ny | h | hy | r | ry | z | j ;
( SENT-START <$phones> SENT-END )

```

### 3. 발음 오류 자동 분류 음성인식기

본 논문에서 제안한 음성인식기는, 한국인의 일본어 발화시 발음의 교정을 목표로 하고 있으므로 그 오류에 정확히 해당하는 교정지침을 제공하기 위하여, 오류 유형별 발화 모델들을 HMM 모델의 훈련 과정을 통해 미리 구축하고, 학습자의 발화에 따른 오류 양상을 자동으로 분류하는 기능을 수행한다. 본 논문에서는 오류발음의 유형별 분류를 위한 최적화된 HMM 모델을 계산적으로 도출하여 HMM 인식기가 최상의 성능을 발휘하도록 하였다.

#### 3.1. 유사 발음 음소 셋의 구성

발화오류 유형의 분류는 음소별로 수행하므로 일본어 음소와 유사하게 발음하는 음소들의 집합에 대한 정의가 필요하며, 본 연구에서 수집한 음성 DB를 일본어 음성학 전문가가 청취하여 한국인의 일본어 발화시 나타나는 음소별 발음 오류 유형을 분석하여 유사 발음 음소 셋을 구성하였다. <표 5> 유사 발음 음소 셋은, 자음인 경우 일본어 자음 20개와 유사 발음 한국어 자음 23개로 구성되고, 모음인 경우 일본어 모음 10개와 유사 발음 한국어 모음 6개로 구성하였다.

#### 3.2. 유사 발음 음소별 HMM 모델 및 발음 네트워크의 구성

이와 같은 유사 발음 음소 셋을 이용하여 <표 6> 및 <그림 1>과 같은 발음 네트워크를 구성하여 학습자의 오류 발음을 검출하였다. <그림 1>에서, 일본어 [k]를 일본어 [k, ky, g, gy]로 발화하거나 한국어 [khc/ㅋ/, kkc/ㄱ/, kc/무성 ㄱ/, gc/유성 ㄱ/]로 발화하는 오류가 발생함을 의미한다. 일본어 및 유사 발음 한국어 모노폰 모델은 2.2.절에서 언급한 일본어 및 한국어 음성 DB로 생성하였다. 그리고 발음 오류 자동 분류 음성인식기를 테스트하기 위한 음성 DB는, 일본어 전공자 120명

과 비전공자 120명 등 한국인 대학생 240명이 일본어 PBW 216개 단어를 발성한 것을 수집하였다.

<표 5> 유사 발음 음소 셋

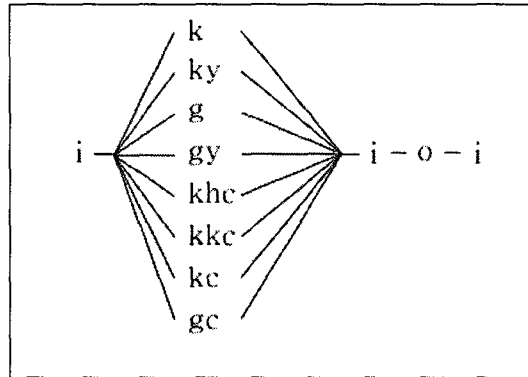
일본어 음소	일본어 유사 발음 음소	발음 오류 한국어 유사 음소	일본어 음소	일본어 유사 발음 음소	발음 오류 한국어 유사 음소
p	py, b, by	phc, ppc, pc, bc	a	ax	axc
py	p, b, by	phc, ppc, pc, bc	i	ix	ixc
b	p, py, by	phc, ppc, pc, bc	u	ux	uxc, euc
by	p, py, b	phc, ppc, pc, bc	e	ex	aex
t	d, ts, ch	thc, ttc, tc, dc	o	ox	oxc
d	t, ts, ch	thc, ttc, tc, dc	ax	a	axc
k	ky, g, gy	khc, kkc, kc, gc	ix	i	ixc
ky	k, g, gy	khc, kkc, kc, gc	ux	u	uxc, euc
g	k, ky, gy	khc, kkc, kc, gc	ex	e	aec
gy	k, ky, g	khc, kkc, kc, gc	ox	o	oxc
s	sh, z, j	sc, ssc, sic, zc, zhc	h	hy	hc, hic
sh	s, z, j	sc, ssc, sic, zc, zhc	hy	h	hc, hic
z	j, s, sh	zc, zhc, sc	r	ry	rc, lc
j	z, s, sh	zc, zhc, sc	ry	r	rc, lc
ts	ch, z, j, t	chc, zc, zhc, thc	ch	ts, z, j, t	chc, zc, zhc, thc

<표 6> 유사 발음 음소 셋을 이용한 발음 네트워크

```

$phone1 = i ;
$phone2 = k | ky | g | gy | khc | kkc | kc | gc ;
$phone3 = i ;
$phone4 = o ;
$phone5 = i ;
( SENT-START $phone1 $phone2 $phone3 $phone4 $phone5 SENT-END )

```



<그림 1> 유사 발음 음소 셋을 이용한 발음 네트워크  
(일본어 단어 「いきおい」 [i kio i]인 경우)

## 4. 실험 및 결과

### 4.1. 언어별 음소인식기 실험 결과

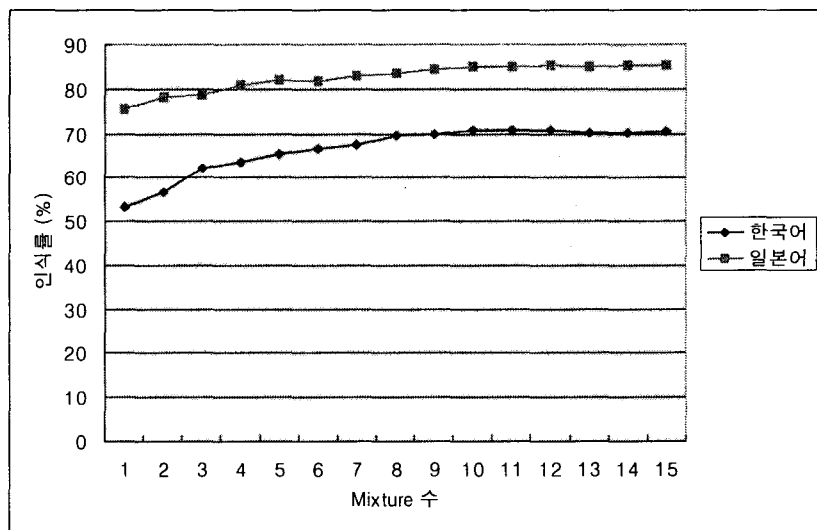
2. 언어별 음소인식기에서 설명한 방법으로 한국어 및 일본어 음소인식기를 구현하여 성능을 평가하였다. Mixture 수가 1인 경우 한국어 및 일본어 음소인식기의 인식률은 각각 52.2%, 75.4%로 나타났다. 실험 결과를 살펴보면, 일본어 음소인식기의 성능이 한국어인 경우보다 23.2% 더 좋은 인식률을 보여 주었다. 이와 같은 실험 결과는 일본어보다 한국어 음소끼리의 음성 자질이 유사하다는 사실을 반영한 것으로 생각된다.

<표 7>을 보면, 일본어 음소인식기는 자음 및 모음 모두 비슷한 인식률을 보여 주었으나, 한국어인 경우 모음의 인식률이 자음인 경우보다 상당히 저조함을 알 수 있다. 이와 같은 결과는, 본 연구에서 선정한 일본어 음소 셋에서 자음 수는 26개로 한국어 자음 수 29개와 비슷하나, 일본어 모음 수는 장음을 포함하여 12개이고 한국어 모음 수는 17개로 차이를 보여, 이러한 모음 셋의 차이가 인식 성능 차이의 원인으로 볼 수 있다. 또한 한국어 모음의 인식 결과를 분석해 보면, '아'가 '야'나 '와'로, '어'가 '여'나 '워'로, '이'가 '위'로, '애, 에'가 '웨'로 오인식되는 비율이 높았는데, 이것은 단모음이 이중모음으로 오인식되는 경우들이다. 그런데 일본어 모음은 단모음, 이중모음의 구분이 없어 한국어보다 인식률이 높다고 생각되며, 단지 장·단음의 구분이 존재하여 단음이 장음으로 오인식되는 경우가 일부 나타났다.

&lt;표 7&gt; 언어별 음소인식기의 자음 및 모음 인식률 (Mixture 수=1)

	자 음	모 음	평 균
한국어 음소인식기	57.4%	50.2%	52.2%
일본어 음소인식기	75.8%	75.0%	75.4%

Mixture 수에 따른 언어별 인식 성능의 변화를 실험하였다. <그림 2>를 살펴보면, Mixture 수가 증가함에 따라 인식 성능의 큰 향상을 볼 수 있는데, Mixture 수가 10인 경우부터 인식률이 포화되고 있음을 알 수 있다. 일본어인 경우 Mixture 15인 경우 인식률이 84.4%로 Mixture 1인 경우보다 9.0% 향상되었고, 특히 한국어인 경우는 71.3%로 19.1%의 큰 성능 개선이 이루어졌다. 따라서 두 언어의 인식 성능 차이는, Mixture 1인 경우 약 23.2%에서 Mixture 10인 경우 약 13.1%로 차이가 좁혀짐을 알 수 있다. 그러나 여전히 두 언어의 인식 성능은 큰 차이를 보이고 있다.



&lt;그림 2&gt; Mixture 수에 따른 인식률의 변화

본 논문에서 구현한 언어별 음소인식기는 발음 오류 자동 분류 음성인식기를 위한 사전 작업으로서, 발음 오류 분류를 음소별로 수행하므로 언어별 음소인식기의 인식 단위를 트라이폰이 아닌 음소로 선정하였다. 또한 대용량 어휘 음성인식기에서 사용하는 언어 모델과 같은 추가적인 정보 없이 음소인식을 수행하였다. 이와 같은 조건 하에서 본 논문에서 구현된 언어별 음소기의 인식 성능은 비교적 우수하다고 볼 수 있다.



## 4.2. 발음 오류 자동 분류 음성인식기 실험 결과

<표 5>의 유사 발음 음소 셋을 이용하고 <그림 1>과 같은 발음 네트워크를 구성하여 발음 오류 자동 분류 음성인식기를 구현하였다. 음소별 HMM 모델은, 4.1. 절에서 분석한 음소인식기의 언어별 성능 차이를 고려하여 일본어는 Mixture 1(인식률 75.4%)로, 한국어는 Mixture 15(71.3%)로 훈련하였다. 같은 Mixture 수를 사용한 경우에 한국어보다 일본어 음소인식기의 인식률이 높으므로, <그림 1>과 같은 발음 네트워크에서 일본어 음소로 바이어스 되어 인식되는 경향을 배제하기 위함이다.

발음 오류를 음성인식기로 분류한 결과가 <표 8>에 보인다. <표 8>의 각 행에서, 유사 발음 음소 셋 중 인식률이 크게 나온 3개의 음소에 대해서만 인식 결과가 보인다.

(1) 일본어 (k, t, p)인 경우, 비전공 대학생은 (kkc/ㄱ/, ttc/ㄷ/, ppc/ㅍ/) 등 한국어 경음으로 발음하는 경우가(평균 40%), (khc/ㅋ/, thc/ㅌ/, phc/ㅍ/) 등 한국어 격음으로 발음하는 경우(평균 27%)보다 많았다. 전공 대학생은 비전공 대학생보다 격음보다 경음으로 발음하는 경향이 더 컸다(평균 22% 대 41%). 이와 같은 실험 결과는 일본어 (k, t, p) 발음은 한국어의 경음과 유사한 발음이라는 사실과 일치한다 [4].

(2) 일본어 (ky, gy, by)인 경우, 전공 및 비전공 대학생 모두 인식결과가 (ky, gy, by)로 나타나는 비율이 제일 컸는데, 이는 /y/라는 음가가 존재하기에 다른 한국어 발음과 구분되었기 때문이라고 생각된다. 그리고 전공 대학生の 실험 결과 (ky, gy, by)로 인식되는 비율(40%, 47%, 74%)이 비전공 대학生の 비율(28%, 40%, 62%)보다 더 큰 사실로부터 전공 대학生の 발음이 보다 일본어에 가깝다는 것을 알 수 있다.

(3) 일본어 (g, d, b)인 경우, 비전공 대학생은 한국어 유성 자음 (gc/ㄱ/, dc/ㄷ/, bc/ㅂ/)으로 발음하는 경향을 보여 주었고, 전공 대학생은 일본어 (g, d, b)와 한국어 유성 자음 (gc, dc, bc)으로 발음하는 비율이(평균 32% 대 33%) 비슷한 분포를 보여 주었다.

(4) 일본어 모음 (a, i, e, o)인 경우, 비전공 대학생은 한국어 모음 (axc/ㅏ/, ixc/ㅣ/, aec/ㅓ/, ㅕ/, oxc/ㅗ/)으로 발음하는 경향을 전공 대학생은 일본어 모음 발음과 한국어 모음 발음이 비슷한 분포를(평균 43% 대 44%) 보여 주었다. 일본어 모음 (u)인 경우, 비전공 대학생은 한국어 모음 (uxc/ㅜ/, 43%), (euc/ㅡ/, 30%) 및 일본어 모음 (u, 24%) 순으로, 전공 대학생은 일본어 모음 (u, 39%) 및 한국어 모음 (euc, 31%), (uxc, 25%) 순으로 발음하는 경향을 보여 주었다. 이와 같은 실험 결과는 일본어 모음 (u)가 한국어 모음 (ㅜ)와 (ㅡ)의 중간 발음이라는 사실과 일치한다[4].

&lt;표 8&gt; 발음 오류 자동 분류 음소인식기 실험 결과 (%)

	비전공 대학생			전공 대학생		
k	k(7)	kkc(39)	khc(34)	k(16)	kkc(39)	khc(23)
ky	ky(28)	kc(29)	khc(19)	ky(40)	kc(21)	khc(20)
g	g(19)	gc(46)	kkc(13)	g(37)	gc(32)	kc(8)
gy	gy(40)	gc(20)	ky(15)	gy(47)	ky(16)	gc(14)
t	t(14)	ttc(46)	thc(19)	t(19)	ttc(47)	thc(18)
d	d(16)	dc(37)	ttc(19)	d(26)	dc(29)	t(16)
p	p(10)	ppc(34)	phc(29)	p(22)	ppc(36)	phc(26)
py	py(13)	phc(73)	ppc(5)	py(27)	phc(54)	ppc(9)
b	b(19)	bc(51)	pc(10)	b(32)	bc(39)	p(8)
by	by(62)	bc(15)	py(5)	by(74)	py(8)	bc(6)
ts	ts(18)	chc(47)	zc(22)	ts(39)	chc(22)	ch(15)
ch	ch(18)	chc(53)	zc(15)	ch(31)	chc(40)	zc(11)
h	h(26)	hc(37)	hic(30)	h(33)	hic(32)	hc(30)
hy	hy(43)	hc(48)	hic(6)	hy(61)	hc(33)	hic(4)
s	s(13)	sc(44)	ssc(36)	s(24)	sc(43)	zc(22)
sh	sh(23)	sc(59)	zc(8)	sh(38)	sc(49)	zc(6)
z	z(19)	zhc(47)	zc(21)	z(35)	zhc(27)	zc(15)
j	j(14)	zhc(42)	zc(33)	j(25)	zhc(32)	zc(26)
r	r(25)	rc(71)	lc(2)	r(42)	rc(53)	ry(2)
ry	ry(62)	rc(29)	r(7)	ry(79)	rc(14)	r(5)
a	a(27)	axc(67)	ax(4)	a(46)	axc(47)	ax(6)
i	i(27)	ixc(65)	ix(6)	i(43)	ixc(43)	ix(12)
u	u(24)	uxc(43)	euc(30)	u(39)	euc(31)	uxc(25)
e	e(30)	aec(63)	ex(5)	e(43)	aec(41)	ex(9)
o	o(25)	oxc(65)	ox(8)	o(40)	oxc(46)	ox(12)
ax	ax(24)	axc(64)	a(11)	ax(38)	axc(42)	a(19)
ix	ix(25)	ixc(64)	i(10)	ix(55)	ixc(39)	i(4)
ux	ux(29)	uxc(47)	euc(17)	ux(55)	euc(21)	uxc(15)
ex	ex(27)	aec(59)	e(12)	ex(44)	aec(41)	e(13)
ox	ox(32)	oxc(63)	o(3)	ox(59)	oxc(36)	o(3)

(5) 일본어 장음 모음 (ax, ix, ux, ex, ox)은 비전공 대학생인 경우 한국어 모음 (axc, ixc, uxc, exc, oxc)으로 발음하는 경향이 가장 컸고(평균 60%), 그 다음으로 일본어 장음 모음으로 발음하는 비율이(평균 28%) 일본어 단음 모음으로 발음하

는 비율(평균 12%)보다 더 커서, 비전공 대학생도 일본어 모음의 장·단음을 구분한다는 사실을 알 수 있다. 전공 대학생은 일본어 장음 모음으로 발음하는 비율이(평균 50%) 한국어 모음으로 발음하는 비율(평균 36%)보다 더 큰 사실을 발견할 수 있어 전공 대학생들이 보다 일본어에 더 가까운 발음을 한다는 사실을 알 수 있다.

이와 같은 실험 결과에서, (1)에서 '실험 결과는 일본어 (k, t, p) 발음은 한국어의 경음과 유사한 발음이라는 사실과 일치'와, (4)에서 '실험 결과는 일본어 모음 (u)가 한국어 모음 (ㅓ)와 (ㅡ)의 중간 발음이라는 사실과 일치'라는 결론, 그리고 (1), (2), (4), (5)에서 보듯이 비전공 대학생의 발음에 비해 전공 대학생의 발음을 보다 일본어 음소에 가깝게 인식하는 결론으로부터, 본 논문에서 제안한 음성인식기의 발음 오류 분류가 제대로 수행되었음을 알 수 있다.

일본어 음성학 전문가가 일본어 /k/ 발화를 청취하여 비전공 대학생의 발음 양상을 분석한 결과와[5] 본 논문에서 제안한 음성인식기의 자동 분류 결과를 비교·분석하였다. <표 9>에 음성학 전문가의 청취 판단 결과와 <표 8>의 음성인식기를 이용한 자동 분류 결과에서 일본어 /k/에 대하여 어두, 어중으로 재분류한 결과를 보여 준다. 어두 /k/에 대하여 청각적 분석은 한국어 /ㄱ/로 66%, /ㅋ/로 32%, /ㄱ/로 2% 비율로 분류하였고, 일본어 /k/로 발화했다고 분류된 경우는 없었다. 자동 분류 결과는 한국어 /ㄱ/로 41%, /ㅋ/로 18%, /ㄱ/로 21%, 일본어 /k/로 20% 분류하였다. 두 분석 결과는 어두 /k/에 대하여 한국어 /ㄱ/로 분류하는 비율이 가장 크다는 결과가 일치한다. 어중 /k/에 대하여 청각적 분석은 한국어 /ㄱ/로 62%, /ㅋ/로 26%, /ㄱ/로 12% 분류하였고, 일본어 /k/로 발화했다고 분류된 경우는 없었다. 자동 분류 결과는 한국어 /ㄱ/로 48%, /ㅋ/로 33%, /ㄱ/로 9%, 일본어 /k/로 10% 분류하였다. 두 분석 결과는 어중 /k/에 대하여 일본어 /k/로 분류한 결과를 제외하고 한국어 /ㄱ, ㅋ, ㄱ/로 분류하는 순서가 일치함을 알 수 있다.

<표 9> 비전공 대학생의 일본어 /k/ 발화의 청각적 분석 및 자동 분류 결과 (%)

	음성 환경	/k/	ㄱ	ㅋ	ㄱ
청각적 분석	어두	0	66	32	2
	어중	0	12	26	62
음성인식기의 자동 분류	어두	20	41	18	21
	어중	10	9	33	48

## 5. 결 론

본 논문에서는 외국어 학습자의 발음 품질을 평가하여 표준 및 오류 발음을 자동으로 분류하는 음성인식기를 제안하였다. 또한 이를 위한 사전 작업으로 한국어 및 일본어 음소인식기를 구현하여 언어별 음소인식기의 성능을 비교, 분석하였다. 본 논문에서 구현된 언어별 음소기의 인식 성능은 비교적 우수하다고 볼 수 있다. 또한 발음 오류 자동 분류 음성인식기의 실험 결과와 음성학 전문가의 청취 판단과의 비교로부터 음성인식기에 의한 자동 분류가 제대로 수행되었음을 알 수 있었다.

향후 연구에서는 언어별 음소인식기의 성능을 향상시키는 연구와, 자동 분류의 정확도를 높여 자동 분류 결과와 음성학 전문가의 청취 판단을 좀 더 일치시키는 연구를 수행할 계획이다. 이를 위하여 오류 발음의 자동 검출을 위한 발음 네트워크의 구성에 관한 연구와 학습자의 발음에 대한 자동 스코어링(scoring)에 관한 연구를 진행할 계획이다.

## 참 고 문 헌

- [1] 신지영, “우리말 소리의 이해”, *대한음성학회 창립25주년기념학술대회 논문집*, pp.15-23, 2002.
- [2] T. Kawahara et al., “Sharable software repository for Japanese large vocabulary continuous speech recognition”, *Proc. ICSLP '98*, pp.3257-3260, 1998.
- [3] SiTEC, 음성정보기술산업지원센터, <http://www.sitec.or.kr>
- [4] 민광준, *일본어 음성학 입문*, 건국대학교출판부, 2002.
- [5] 이재강, 권철홍, “청각인상과 음성파형 간의 관계 규명을 위한 일본어 /k/의 기초 연구”, *대한음성학회 봄 학술대회 논문집*, pp.52-55, 2003.

접수일자: 11월 6일

게재결정: 12월 12일

▶ 권철홍(Chul-Hong Kwon)

주소: 300-716 대전광역시 동구 용운동 96-3 대전대학교

소속: 대전대학교 정보통신공학과 BMW 연구실

전화: 042) 280-2555

E-mail: chkwon@dju.ac.kr

▶ 강효원(Hyo-Won Kang)

주소: 300-716 대전광역시 동구 용운동 96-3 대전대학교

소속: 대전대학교 정보통신공학과 BMW 연구실

전화: 042) 280-2567

E-mail: kanghyowon@hotmail.com

▶ 이상필(Sang-Pil Lee)

주소: 300-716 대전광역시 동구 용운동 96-3 대전대학교

소속: 대전대학교 정보통신공학과 BMW 연구실

전화: 042) 280-2567

E-mail: nishot@hanmail.net