



Journal of Knowledge Information Technology and Systems

ISSN 1975-7700

<http://www.kkits.or.kr>

A Comparative Study of the Speech Signal Parameters for the Consonants of Pyongyang and Seoul Dialects - Focused on the affricates “ㅈ/ㅉ/ㅊ”

Kwang-Bock You¹, Shinae So², Kang-Hee Lee³

¹*School of Electronic Engineering, Soongsil University*

²*Department of Korean Language and Literature, Soongsil University*

³*Global School of Media, Soongsil University*

ABSTRACT

In this paper, from the point of view of speech signal processing as an engineering application, the comparative study of the Pyongyang and Seoul dialects is performed. In special, the affricates "ㅈ, ㅉ, ㅊ", which would have different phonetic values between Pyongyang and Seoul dialects, are focused and compared. For these consonants, the speech parameters such as the spectrogram, pitch, and Formant frequencies are extracted (measured) and the differences in their phonetic values of these two regions have been compared. It is confirmed that for these consonants, the Pyongyang dialects have higher energy cohesion than the Seoul dialects, and the distribution of Formant frequency was well distinguished. With the vowel study of Pyongyang dialect, which was carried out with phonological or experimental phonological methods, this paper presents the study of consonants in Pyongyang dialect by using the speech signal parameters. In this paper, a method is proposed for verifying linguistic investigations and experimental phonetic results by using signal processing.

© 2019 KKITS All rights reserved

KEYWORDS : Affricates consonants, Speech signal processing, Pyongyang dialect, Spectrogram, formant frequencies, Pitch

ARTICLE INFO: Received 9 July 2019, Revised 3 August 2019, Accepted 9 August 2019.

*Corresponding author is with the Global School of Media, Soongsil University, 369 Sangdo-ro, Dongjak-gu

Seoul, 06978, KOREA.

E-mail address: kanghee.lee@ssu.ac.kr

1. 서론

음성신호에 대한 음운론적 연구 (phonological research)이든 언어학적 연구이든지 대부분 경우는 모음 (많은 경우 유성음) 혹은 모음의 체계에 관한 것이다. 음성신호처리라는 공학적 연구의 대부분도 모음 (유성음)인 것이 사실이다.

자음 (많은 경우 무성음)은 조음방법 (manner of articulation), 조음위치 (place of articulation), 그리고 position of vocal cords 등에 따라서 발음 혹은 그 소리 값을 설명할 수 있다. 그러나 모음은 혀 (tongue), 입술 (lips), 그리고 턱 (jaw) 등의 3차원적 움직임에 더불어 vocal tract이 갖는 특성 등이 표현되는 것이라 할 수 있다. 이런 많은 변수들이 포함되는 모음의 조음 방식은 자음에 비해 상당히 이해하기 어려운 것이다. 모음은 언어와 언어 사이의 차이 그리고 같은 언어에서도 지역 간의 차이가 매우 크다고 한다. 언어의 진화에서도 자음은 상대적으로 안정적인 변화가 있지만 이에 비해 모음은 그 변화가 급격하고 매우 크다. 이런 이유들로 인해 모음에 관한 연구들이 상대적으로 활발한 것으로 보인다. 이런 경향들은 유성음 신호는 준주기적 특성을 갖고 있어 신호처리의 많은 수학적 개념들을 적용하기가 용이하기에 음성신호의 연구에서도 그러하다. 유성음에 일반적으로 적용되고 있는 음성신호 파라미터들에는 신호의 에너지, 피치, 그리고 포먼트 주파수 등이 있다. 그러나 무성음에 관한 연구는 유성음 연구에 비해 상대적으로 그 방법론 그리고 추출 파라미터들이 많지 않다고 할 수 있다. 만일 어떤 특정 자음들이 그 특성 혹은 소리 값이 지역에 따라서 상당한 차이를 나타낸다면 이에 대한 연구가 있어야 하는 것이 정상적인 것이다.

본 논문은 신호처리의 수학적 개념에 기초하여 한국어의 자음에 관한 비교 연구를 수행하였다. 언

어학적 연구들도 기본적으로 검토를 하였다. 표준어 (서울 지역어)와 평양을 중심으로 한 서북지역 방언의 자음에서 그 소리 값에서 많은 차이를 보이는 파찰음을 중심으로 비교하였다.

표준어를 비롯한 남한 지역어의 경우에는 파찰음 /스, 츠, 짜/가 경구개음으로 실현되는 데 반해, 평안도, 함북 육진 지역, 황해도 일부 지역 등에서는 파찰음 /스, 츠, 짜/가 치음 내지 치조음으로 조음이 되고 있다. 즉, 파찰음의 조음 위치에 있어서 표준어와 평양 지역어가 큰 차이를 보이는 것이다. 대부분의 남한 지역어에서 폐쇄음 /ㄷ, ㅌ, ㅌ/나 비음 /ㄴ/가 치조음으로 조음되는 반면, 북한 지역어에서는 이들이 치음으로 조음되는 것이다. 이러한 음성적 차이들은 구개음화 및 ‘어두 ㄴ 탈락’ 과 같은 음운 현상의 차이로 나타나고 있다 [1-3]. 이 파찰음에 관련하여 2010년대에 새터민들의 발음으로 확인한 연구가 있다[4]. 이 연구의 결과를 <표 1>로 정리하였다.

표 1. 치경음 실현율

Table 1. Table for Alveolar pronunciation rate

	치경음 실현율(%)			
	전체	동북	서북	육진
자다	17.9	17.4	33.3	10
짜다	28.2	17.4	33.3	50
차다	5.1	4.3	16.7	0
지다	12.8	8.7	16.7	20
찌다	17.9	8.7	33.3	30
치다	17.9	21.7	0.0	20

이 두 지역의 모음들에서도 몇 가지 차이가 있다. 모음 “ㄴ”의 음가가 다르다. 서울 지역어에서는 [ㄴ] 발음인데, 평양 방언은 [ɲ]의 발음으로 입

술이 더 등글어지고 좁아진 형태에서의 발음이다. 이외에도 “ㅡ”의 음가가 다른 것으로 조사되었는데, 학자들에 따르면 이 발음은 “ㅓ”에 흡수되었다고 한다. 따라서 평양 지역어의 모음체계가 현재 변화하고 있는 것이다. 이외에도 그 음조가 서로 다르지만 이 차이에 대한 연구들이 충분히 진행되고 있지 않은 것이다[1,4-10].

기존에 남북한 방언들의 모음에 관한 연구는 비교적 활발히 이루어져 온 반면, 이들의 자음에 관한 음성적 연구는 거의 이루어지지 못했다는 점에서, 남북한의 개별 방언의 자음 음성에 대한 본 논문의 중요성은 상당히 크다고 할 수 있다.

본 논문의 구성은 다음과 같다. 2장에서는 본 논문에서 추출하고 비교한 음성신호의 파라미터들에 관해 신호처리적인 분석에 대해 기술하였고, 3장에서는 시뮬레이션의 결과들을 분석 비교 하였다. 마지막으로 4장에서는 본 논문의 결과들과 그에 대한 검토를 하였다.

2. 음성신호의 파라미터와 신호의 분석

본 논문은 평양지역 방언에서 /ㅅ, ㅆ/, /ㅈ, ㅊ, ㅊ/, /ㄷ, ㅌ, ㅌ/ 등의 자음 발음이 서울 지역어와는 다르다는 점에 착안하여 연구를 진행하였다. 음운론적 관점에서 자음의 연구는 조음 방법과 조음 위치에 따른 소리 값에 관심을 갖고 있다. 이런 음운론적 연구와 함께 실험음향적 연구가 더불어 수행되고 있다. 이런 실험음향적 연구에 대표적으로 사용되고 있는 프로그램으로 Computerized Speech Lab (CSL)과 Praat이 있다[11]. 이런 프로그램들은 음성신호처리 분야에서 일반적으로 사용하고 있는 알고리즘들을 채용하고 있다. 예를 들면, 포먼트 주파수들을 측정하기 위해서 Linear Prediction Coding (LPC) 알고리즘을 사용하였고, 알고리즘의 유연성을 위해서 LPC 방정식의 차수를 선택할 수

있도록 하고 있다. 그러나 음성신호처리 관점에서 보면 알고리즘은 꼭 하나일 필요는 없고 심지어 본 논문에서 시도한 것 같이 새로운 방법으로 측정할 수도 있다. 그리고 이 방정식의 차수는 10차 이상일 필요는 없다는 것이 일반적인 정설이다[12]. 실험음향적인 분석이 기존의 알고리즘만으로 혹은 제공된 tools들로 수행된다는 것은 신호처리의 공학적 접근에서 보면 제한적인 방법으로 측정된 데이터이기에 그 한계가 분명히 존재한다. 그러므로 이런 프로그램을 사용하기 위해서는 음성신호의 특성들을 표현하는 파라미터들의 속성들을 정확히 이해하고 실행해야 한다[12-15].

본 장에서는 이러한 한계들에 대해 신호처리의 기본 개념에 근거한 수학적인 모델링들을 분석한다. 이러한 분석을 통해서 음성신호에 적절한 파라미터들을 추출하였다. 본 논문은 이 파라미터들을 남북한 자음 음성 비교에 활용하였다.

2.1 자음 생성 모델의 분석

시간에 따라서 변하는 음성신호는 선형 시변 시스템으로 모델링하는 것이 타당하다. 그러나 충분히 짧은 프레임으로 신호를 처리하면 그 수학적 표현이 매우 간결해지는 선형 시불변 시스템으로 나타낼 수 있다[14].

그림 1은 무성음 생성 모델을 보인다. 출력 음성 신호, $s[n]$,은 아래의 식으로 나타낸다.

$$s[n] = G \cdot u[n] * h[n] \quad (1)$$

여기서 G 는 무성음 구간 프레임의 이득 조정 (Gain control) 값이고, $u[n]$ 은 랜덤 노이즈 여기신호이고, $h[n]$ 은 무성음의 성도 임펄스 응답 (vocal tract impulse response)이다. *는 신호들의 컨볼루션 (convolution) 계산을 나타낸다. <그림 1>의 무성

음 생성 모델에서 합성필터의 입력이 랜덤 신호이기에 무성음은 autocorrelation 함수로 표현할 수 있다.

$$\phi_{ss}[n] = G^2 \phi_{vv}[n] * \phi_{rr}[n] \quad (2)$$

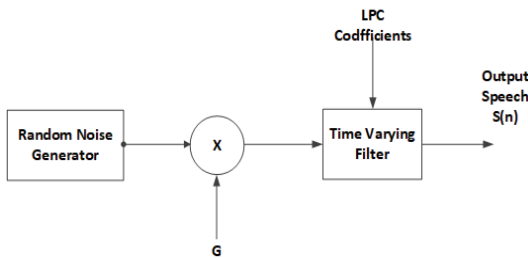


그림 1. 자음 (무성음) 생성 Model
Figure 1. Unvoiced production Model

$\phi_{vv}[n]$ 과 $\phi_{rr}[n]$ 은 각각 성도와 입술에서의 autocorrelation 함수들이다. 그래서 autocorrelation 함수의 푸리에 변환인 power spectrum으로 무성음을 다음과 같이 표현할 수 있다[14,15].

$$\Phi_{ss}(e^{j\omega}) = G^2 |V(e^{j\omega})|^2 |R(e^{j\omega})|^2 \quad (3)$$

식 3에서 알 수 있는 것 같이 무성음의 특성을 나타내는 파라미터로 power spectrum이 중요하다.

본 논문에서는 MatLab의 내장함수인 spectrogram function을 사용하였다. 이 함수는 음성신호가 시간에 따라서 변할 때 이에 대응하는 주파수 성분이 어떻게 분포하는지를 보여준다[15].

2.2 포먼트 주파수의 분석

음성신호의 주기적인 특성을 피치라고 하며 대부분의 경우 유성음에서 나타난다. 피치를 정확하

게 측정하는 것은 매우 어려운 일이다. 이는 음성신호가 갖는 준주기적인 특성과 시간에 따라 변하는 통계적 특성에 기인한다.

피치를 추정하는 방법들에는 시간영역에서는 Autocorrelation Function (ACF) 알고리즘과 Average Magnitude Difference Function (AMDF) 알고리즘이 있고, 주파수 영역에서는 Cepstrum Function (CF) 알고리즘이 있다. 본 논문에서는 ACF 방법을 사용하여 피치를 추정하였다. 이에 대한 정의식은 아래와 같다[12,14].

$$R[m] = \frac{1}{N} \sum_{n=0}^{N-1-m} s[n]s[n+m] \quad (4)$$

이 식에서 지연 (lag 혹은 delay)인 m 의 범위는 $0 \leq m \leq M_0$ 이다. 이 함수는 $m = 0$ 에서 최댓값을 갖는 우함수이다.

본 논문에서는 분석하려는 자음에 따르는 모음의 피치를 측정하였고 그것의 변화를 보았다.

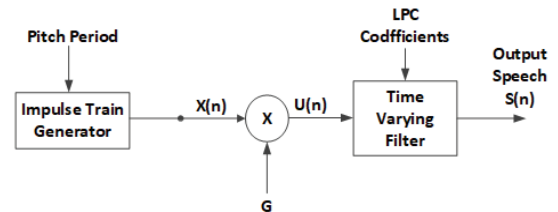


그림 2. 모음 (유성음) 생성 Model
Figure 2. Voiced Production Model

<그림 2>는 유성음에 대한 Linear Prediction Coding 분석 모델을 보여준다. 이 모델에서 디지털 필터는 음성신호 생성에서 성도 특성을 나타내는 포먼트 주파수들을 추출할 수 있고, 이 필터의 전달함수는 All-pole rational function으로 아래의 식으로 표현한다.

$$H(z) = \frac{1}{1 - \sum_{k=1}^p a_k z^{-k}} \quad (5)$$

식 5에서 예측계수 a_k 들을 구하는 것은 포먼트 주파수를 추출한다는 것과 같은 의미가 된다.

식 5에서의 예측계수 a_k 들은 다음과 같이 정의되는 $e(n) = s[\widehat{n}] - s[n]$ 오차를 최소화하는 값들이고, 이 계수들로 구성되는 다항식의 근들은 서로 켈레 복소수 관계이다. 본 논문에서는 이 계수들을 ACF를 기초한 Levinson-Durbin 알고리즘으로 구하였다[12][14][15]. 본 논문에서는 분석하는 자음에 따르는 모음의 포먼트 주파수를 측정하였다. 이 모음의 포먼트 주파수들은 서울 표준어와 평양 지역에서 어떤 차이를 보이는지 비교하였다.

2.3 유무성음 탐지 알고리즘

음성신호 처리의 연구 분야에서 고전적인 문제 중에 하나가 유무성음의 구간을 구별하는 문제이다. 이 문제의 어려움은 음성신호가 non-stationary 이고 준주기적 (quasi-periodic)인 것에 기인하고 또 glottal excitation과 성도 사이의 상호작용에 근거한다. 그래도 음성신호에서 유성음과 무성음은 에너지에서 큰 차이를 보인다. 이와 함께 신호가 (+) 값에서 (-) 값으로 또 그 반대로 변하는 수를 Zero Crossing Rate (ZCR)이라 한다. ZCR은 유성음 구간이 랜덤 신호 같은 무성음 구간에 비해 적다. 아래의 식 6으로 ZCR을 표현 할 수 있다[14].

$$ZCR[m] = \frac{1}{2} \sum_{n=m-N+1}^m |sgn(s[n]) - sgn(s[n-1])| \quad (6)$$

여기서 m 은 프레임의 번호이고 $sgn(\cdot)$ 함수는 인수의 부호에 따라 +1 혹은 -1의 값을 돌려준다.

일반적으로 식 7과 같이 신호의 에너지를 구한다. 실제적으로는 프레임 별로 에너지를 구하는 것이므로 windowing을 하여 계산하게 된다[14,15].

$$E = \sum_{m=-\infty}^{\infty} s^2[m] \quad (7)$$

본 논문은 위의 두 식 6과 7을 사용하여 유무성음을 구분하는 알고리즘을 구현하였다. 각 프레임의 ZCR 값으로 3개의 영역으로 나눈다. 먼저, ZCR 값이 최댓값의 50%보다 작으면 유성음 프레임으로 하고 산술값 “-1” 을 부여한다. 두 번째는 50 ~ 67%이면 천이 구간으로 처리하고 그 값을 “0” 으로 한다. 마지막으로 67% 이상이면 그 값을 “+1” 로 하고 무성음 프레임으로 처리한다. 각 프레임의 에너지 값들을 비교하여 프레임에 대한 판단을 수행한다[12]. 본 알고리즘은 경구개음에 좋은 성능을 보여준다. 향후 본 알고리즘은 음성신호의 다른 파라미터들을 포함하는 새로운 알고리즘으로 그 성능을 높이는 일을 해야 한다.

3. 시뮬레이션 결과

본 논문에서는 2010년에 수집된 평양 지역어 노년층 화자들의 음성 자료를 같은 시기에 조사된 서울 지역어 노년층 화자들의 음성 자료와 비교하였다. 본 데이터는 언어 이주 상황에서 수집된 것이기에 현 평양 지역의 화자들보다는 더 보수적인 발화 상태를 유지하는 것으로 볼 수 있다. 사실상 북한 지역어 음성 자료를 수집하는 데에는 현실적인 제약이 따른다[1-3].

<표 2>에서 시뮬레이션에 사용한 음성 데이터를 지역과 조음방법으로 분류하였다. 각 음성 데이터는 16 kHz로 샘플링 하였다. 분석 프레임의 길이는 충분히 짧아서 본 논문에서 분석한 식들을 사용하

기에 문제가 없는 480 샘플 (30ms)으로 하였다.

표 2 비교 음성 Data (ㅈ/ㅉ/ㅊ) 분류표
Table 2. Classification table for comparative speech data (ㅈ/ㅉ/ㅊ)

지역	분류	자음	모음				
			아	에	이	어/오	애
서울	경구개음	ㅈ	잠을 자다		짐승 말	저 달	재다
		ㅉ	소금 은 짜다		살이 짜다		쩌다
		ㅊ	차다	인체 간			채다
평양	치조음	ㅈ	잠을 자다		짐승 말	저 달	재다
		ㅉ	소금 은 짜다		살이 짜다		쩌다
		ㅊ	차다	인체 간			채다

본 시뮬레이션은 랜덤 신호를 처리하는 것이기에 각 음성신호를 표준화 (normalization)하였다. 이는 시뮬레이션의 속도를 높일 수 있고 또 locally optimal 되는 것을 방지할 수 있다. 신호들의 평균 (Mean)과 분산 (Variance)을 사용하여 표준화를 하였기에 음성신호가 갖고 있는 그 특성들이 그대로 나타날 수 있도록 하였다. 이와 더불어 식 8로 나타낸 High-Pass Filter (HPF)로 불필요한 신호의 저주파 성분들을 제거하였다[2,3].

$$H(z) = \frac{0.946 - 1.892z^{-1} + 0.946z^{-2}}{1 - 1.889033z^{-1} + 0.8948743z^{-2}} \quad (8)$$

각 데이터의 그림들에서 “P”는 평양 지역 방언을, “S”는 서울 표준어를 나타낸다. 예시로 “P-kimkyljaeda.wav”는 평양 방언 “재다”를 화자 “kimky1”가 1차로 발화한 데이터임을 나타낸다[1].

<그림 3>은 자음 “ㅈ” (단어 “재다” 에서)에 대한 스펙트럼을 보였다. 이 스펙트럼은 식 3으로 계산된다. 첫 번째 행은 단어 “재다”의 파형이고, 두개의 막대사이에 있는 부분이 “재”이고, 두 번째 행에서 “ㅈ (진한 회색) + ㅉ (검은 색)”으로 무성음과 유성음으로 분리하였다. 세 번째 행에서는 이에 대한 스펙트럼을 보였다.

왼쪽 스펙트럼 (평양 지역어)의 앞부분 약 0.125 (sec)와 이후를 비교해보면 확실하게 “ㅈ+ㅉ”로 구분되는 것을 알 수 있다. 이 평양 지역어의 “ㅈ”은 주파수가 4 ~ 8 kHz 사이에 많은 에너지 모여 있다. 반면에 서울 표준어의 “ㅈ”은 0.075 (sec)에서 주파수가 3 ~ 6 kHz 사이에 있음을 알 수 있다. 더불어 서울 표준어의 발화 속도가 0.14 (sec) 정도 빠르다고 할 수 있다.

<그림 4>는 단어 “재다”에서 식 6을 적용하여 Zero Crossing Rate (ZCR)을 구하고 이를 유무성음 구간을 탐지하는 것을 보였다. 먼저, 첫 번째 행에서 ZCR의 변화를 프레임별로 보였고, 다음 행에서 천이 구간을 제거하였다. 이 결과를 이용해서 유무성음을 구분 한 결과를 세 번째 행에서 보였다. 마지막 행에서는 식 4로 ACF를 구해서 피치를 측정하고 이의 프레임별 변화를 보였다.

<표 3, 4, 5>에서는 본 논문에서 시뮬레이션을 수행한 파찰음 /ㅈ, ㅉ, ㅊ/의 모음에 대한 포먼트 주파수들의 값을 구하고 서울 표준어와 평양 방언과의 그 차이를 보였다. 포먼트 주파수는 식 5에서의 전달함수 H(z)의 poles (극점)가 된다.

같은 방식으로 <그림 5>와 <그림 6>은 단어 “쩌다”의 분석을 보이고 있으며, <그림 7>과 <그림 8>은 단어 “채다”에 대해서 분석하였다.

4. 결론

평양 지역어의 자음 /ㅈ/이 서울 표준어보다 에

너지 응집도가 좋다. 새터민의 발음으로 북한 지역들의 치경음 실현율을 조사한 <표 1>의 결과와 같이 유사하게 /ㅈ/에서 더 다름이 확인된다. 파찰음 /ㅈ, ㅉ, ㅊ/에서 서울 표준어와 평양 지역어는 표 3, 4, 5에서 보이는 것 같이 포먼트 주파수들의 차이에서 다른 특성을 보이고 있다. 그러나 피치의 변화는 개인차가 많은 것으로 보여서 이 부분은 더 많은 데이터로 조사해 볼 필요가 있다.

포먼트 주파수의 f1과 f2의 분포로 보면 /ㅈ, ㅉ/이 /ㅈ/ 보다 상대적으로 서울 표준어와 평양 방언을 잘 구분이 되고 있다. 선형적 함수로 이 분포를 분류하는 것이 가능해서 support vector machine과 같은 간단한 인공 지능 알고리즘의 적용이 용이할 것으로 보인다. 이런 결과는 <표 3, 4, 5>에서 확인 할 수 있다. 그러나 /ㅈ/은 실험 결과 /ㅈ, ㅉ/과 다르게 그 분포의 구분이 선형적이지 못하다. 이 부분은 향후 더 많은 데이터로 확인이 필요하다. 식 5로 표현한 합성 필터의 극점들은 서로 결

레 복소수라는 성질을 이용하여 이 필터를 2차의 다항식으로 표현되는 필터들로 분리하였다. 분리된 각 필터에서 이에 대응하는 포먼트 주파수를 구하여 <표 3, 4, 5>를 완성하였다.

본 논문에서 사용한 유무성음 영역 탐지 알고리즘은 ZCR과 에너지의 간단한 구성이기에 많은 천이 영역이 있을 것으로 예상하였으나 성공적으로 유무성음을 분리하였다. 그래서 본 알고리즘은 파찰음 /ㅈ, ㅉ, ㅊ/의 유무성음 영역 탐지에 잘 적용됨을 알 수 있다. 이외의 다른 한국어의 자음들에게도 적용이 가능한지는 많은 데이터와 실험을 수행해볼 필요가 있다.

본 논문은 한국어 방언 연구에서 비교적 소홀했던 자음의 연구에 새로운 방법을 제시하였다. 또 언어학적인 조사와 실험 음성학적 결과들에 신호처리의 공학적 방법을 더하여 그 결과들을 응용할 수 있도록 음성처리의 새로운 연구 방법을 제시하였다.

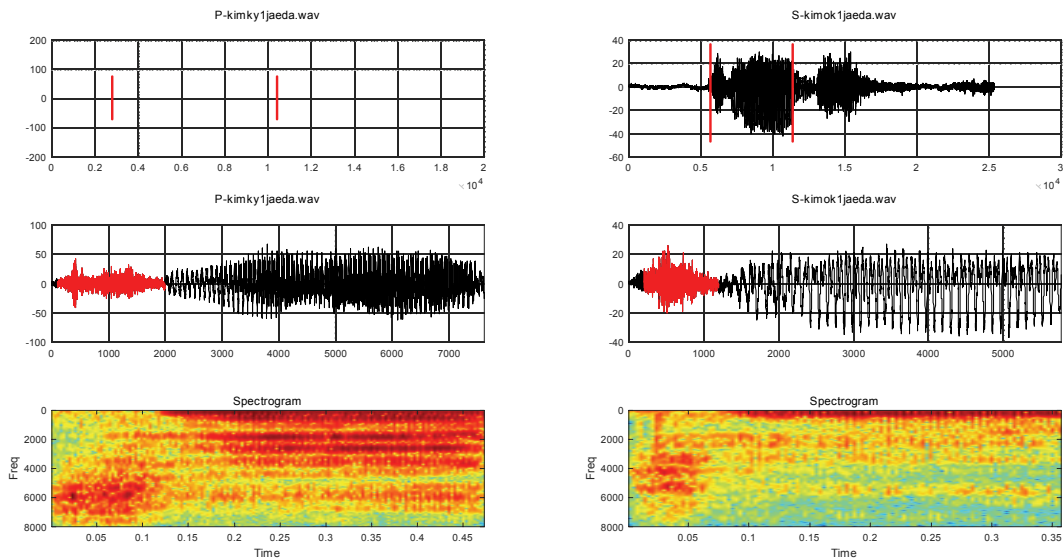


그림 3. 자음 “ㅈ” (제다)에 대한 스펙트럼의 비교
Figure 3. Comparison spectrum of the consonant “ㅈ” (제다)

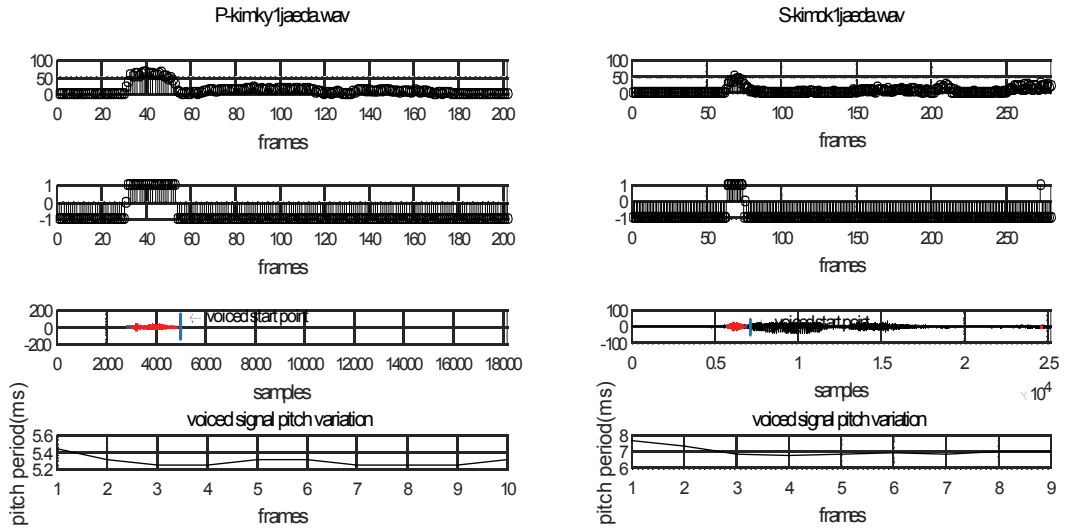


그림 4. “재” (재다)에서 유무성음 탐지 비교
 Figure 4. Comparison UV/V Detection in the word “재” (재다)

표 3. 자음 “ㄷ”에 관련하여 포먼트 주파수의 비교
 Table 3. Comparison table of the Formant frequencies regarding to the consonant “ㄷ”

Data	재다			잠을 자다			짐승 말		
	서울	평양	Difference	서울	평양	Difference	서울	평양	Difference
F1	281			218	578	360	218	203	15
F2	2000							1344	
F3	3234	2422	812		3094		2969	3703	734
F4	5469	3500	1969	4625	5203	578		5422	
F5		5875			6359				

표 4. 자음 “ㅈ”에 관련하여 포먼트 주파수의 비교
 Table 3. Comparison table of the Formant frequencies regarding to the consonant “ㅈ”

Data	살이 찌다			소금은 짜다			짜다		
	서울	평양	Difference	서울	평양	Difference	서울	평양	Difference
F1	125	218	93	187	625	438	265	281	16
F2	1922	1766	156				2016	2234	218
F3	3828	3766	62		3234				
F4	5477	5406	71	5688			5922	4515	1407
F5					6266			6531	

표 5. 자음 “ㄷ”에 관련하여 포먼트 주파수의 비교

Table 5. Comparison table of the Formant frequencies regarding to the consonant “ㄷ”

Data	차다			채다			인체 간		
	서울	평양	Difference	서울	평양	Difference	서울	평양	Difference
F1	344	1047	703	299	344	45	328	187	141
F2		1859		2094				2266	
F3	3640			2828	3234	406	3109	3500	391
F4		4766			5000		5219	5578	359
F5	6688	6422	266	6656	6922	266		6844	

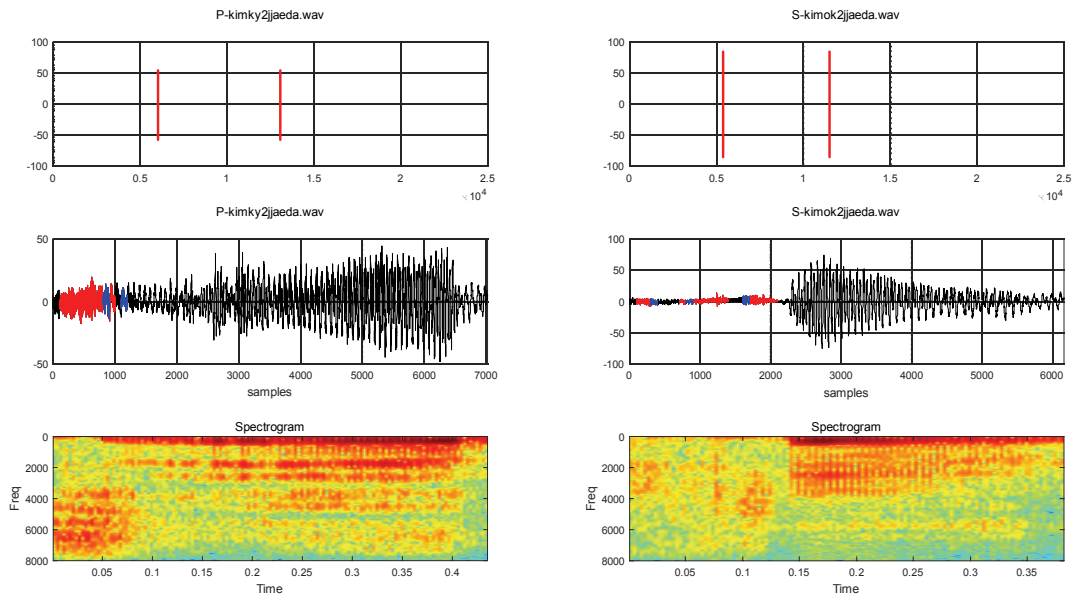


그림 5. 자음 “ㄷ” (ㄷ)에 대한 스펙트럼의 비교

Figure 5. Comparison spectrum of the consonant “ㄷ” (ㄷ)

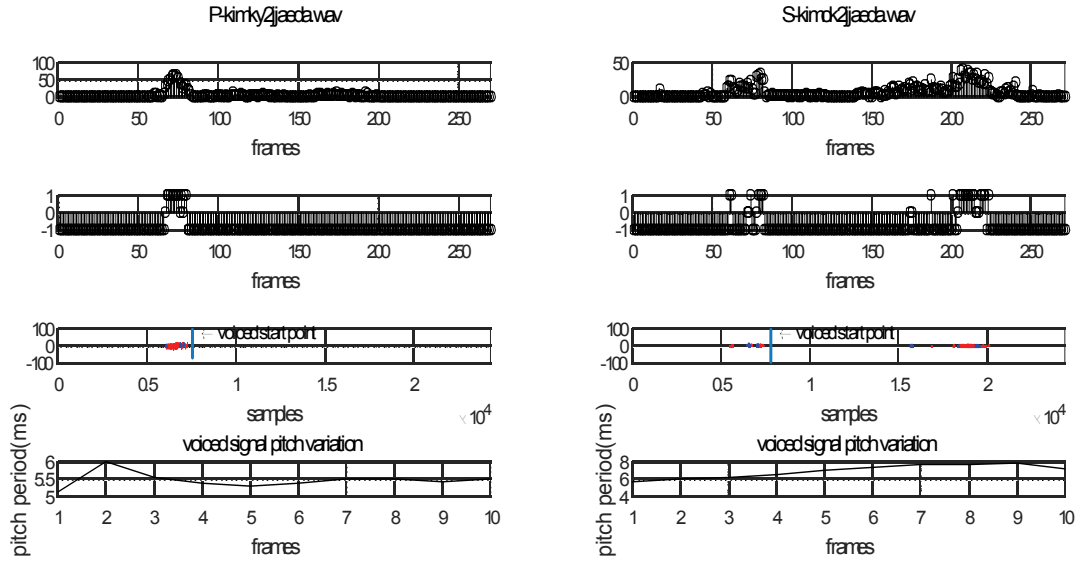


그림 6. “채” (채다)에서 유무성을 탐지 비교
 Figure 6. Comparison UV/V Detection in the word “채” (채다)

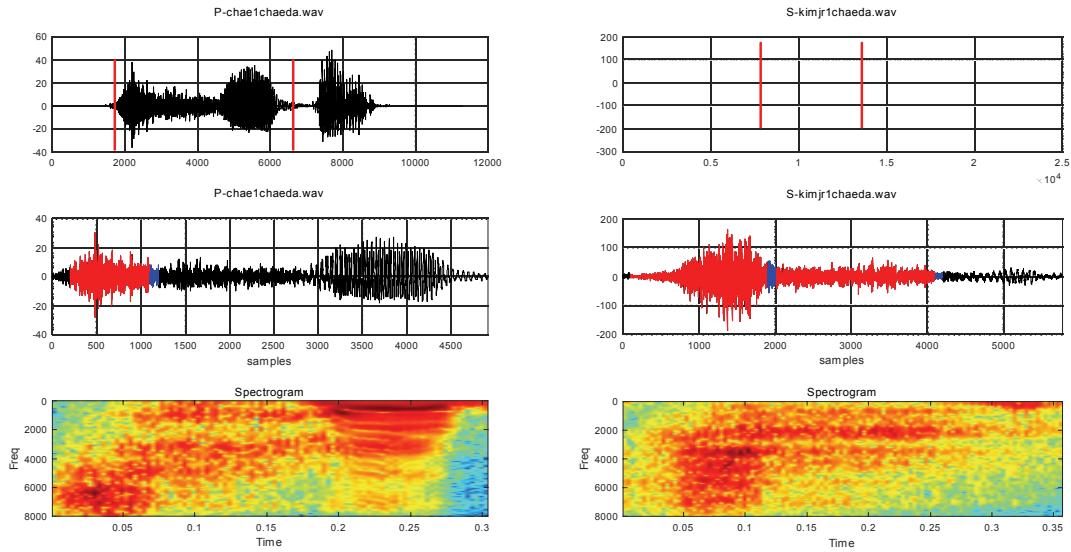


그림 7. 자음 “츠” (채다)에 대한 스펙트럼의 비교
 Figure 7. Comparison spectrum of the consonant “츠” (채다)

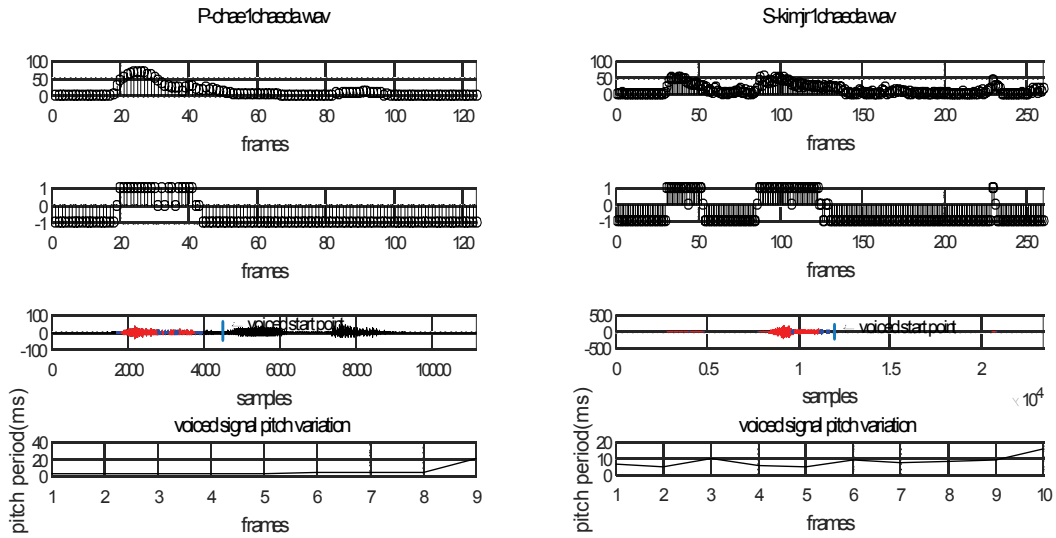


그림 8. “ㄷ” (체다)에서 유무성음 탐지 비교
 Figure 8. Comparison UV/V Detection in the word “ㄷ” (체다)

References

[1] S-A. So, *Experimental phonetic study on the Pyeong'an dialect: focusing on the vowels of the Pyeong'yang subdialect*, Journal of Korean Linguistics, The Society of Korean Linguistics, 58, pp. 231-254, 2010.

[2] K-H. Lee, S-A. So, and K-B. You, Annual report for the project - A study on extracting speech parameters for intonational and consonantal differences between the standard dialect (Seoul dialect) and the other dialects of Korean, 2018.

[3] S-A. So, K-H. Lee, K-B. You, and H-Y. Lim, *A comparative study of the speech signal parameters for the consonants of Pyongyang and Seoul dialects - focused on “ㄸ/ㅃ”*, Journal of Multimedia Services Convergent with Art, Humanities, and Sociology, Vol. 8, pp. 927-937, 2018.

[4] H-Y. Lee, *Report the result of pronunciation education for the north Korean refugee*, National Institute of Korean Language, Seoul, 2009.

[5] <https://ko.wikipedia.org/wiki>, Apr. 2019.

[6] <https://namu.wiki/w/>, Apr. 2019.

[7] H-S. Jung, S-K. Yang, and B-S. Kang, *Standard language education textbook for the north Korean refugee*, National Institute of Korean Language, Seoul, 2011.

[8] S-K. Kahng, *The hypercorrection of vowel /u/→/i/ in north Korean dialects*, Speech Science, The Korean Association of Speech Sciences, Vol. 6, No. 1, pp. 43-54, 1999.

[9] Y. B. Kim, *About the consonants of Pyongan dialect*, The Korean Language and Literature, 76, pp. 112-120, 1977.

[10] S-K. Kahng, *The merger of back vowels in north Korean, Speech Sciences*, The Korean Association of Speech Sciences, Vol. 5, No. 2, pp. 41-55, 1999.

[11] B. G. Yang, *Theory and practice of speech analysis using Pratt*, Mansu Publishing Co., Seoul, 2010.

[12] K-B. You, K-H. Lee, S-A. So, H-Y. Lim, and J. Park, *A study of peak finding algorithms for the autocorrelation function of speech signal*, Journal of The Korea Society of Computer and Information, Vol. 21, No. 12, pp. 131-137, 2016.

[13] P. C. Pandey, and M. S. Shah, *Estimation of place of articulation during stop closures of vowel-consonant-vowel utterances*, IEEE Trans. on Audio, Speech, and Language Processing, Vol. 19, No. 3, 2011.

[14] L. R. Rabiner, and R. W. Schafer, *Theory and applications of digital speech processing*, 1st edition, Prentice-Hall, Upper Saddle River, NJ. 2011.

[15] I. V. McLoughlin, *Speech and audio processing*, Cambridge University Press, New York, 2016.

평양 지역어와 서울 표준어의 음가가 다른 파찰음 /스, ㅍ, ㅊ/을 비교 연구하였다. 서울 지역어에서는 경구개음으로 서북 지역어에서는 치경음으로 실현되는 이 자음들에 대해 스펙트럼, 피치, 포먼트 주파수 같은 음성 신호의 파라미터들을 추출 (측정)하여 이 두 지역의 이 자음들에 대한 음가의 차이를 신호처리적인 방법들로 비교 연구하였다. /스, ㅍ, ㅊ/의 자음에서 평양 지역이 서울 지역 보다 에너지의 응집도가 높았고, 포먼트 주파수의 분포가 잘 구분이 되는 것으로 확인 되었다. 본 논문은 대부분의 경우 음운론적 혹은 실험음성학적 방법들로 수행했던 평양 방언의 모음 혹은 그 체계 연구에 더하여 음성신호처리의 파라미터들로서 자음 연구를 수행하는 연구 방법을 제시 하였다. 본 논문에서는 언어학적 조사와 실험 음성학적 결과들을 신호처리의 공학적 방법으로 검증하는 새로운 연구 방법을 제시 하였다.

감사의 글

이 논문은 2016년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2016S1A5A2A03927522)



Kwang-Bock You received the Ph.D., in ECE from Stevens Institute of Technology, Hoboken, NJ., USA, in 1998. Dr. You joined the School of

Electronic Engineering, Soongsil University as an Associate Professor in 2010. He is interested in speech signal processing, sampling theorem, and wireless communications.

E-mail address: kwangbockyou@ssu.ac.kr



Shin-Ae So received the B.S., M.S. and Ph.D. degrees in Korean Language and Literature from Sogang University, Korea, in 1999, 2002 and 2006, respectively.

평양 방언과 서울 표준어의 자음에 대한 음성신호 파라미터들의 비교연구 - 파찰음 “스/ㅍ/ㅊ”을 중심으로

유광복¹, 소신애², 이강희³

¹ 숭실대학교 IT대학 전자정보공학부 부교수

² 숭실대학교 인문대학 국어국문학과 부교수

³ 숭실대학교 IT대학 글로벌미디어학부 부교수

요 약

본 논문은 공학적 응용인 음성신호처리의 관점에서

Dr. So is currently an Assistant Professor in the Department of Korean Language and Literature, Soongsil University. She is interested in Korean phonology, Korean dialectology, and linguistic variation and change.

E-mail address: hetaira@ssu.ac.kr



Kang-Hee Lee received the B.S., M.S. and Ph.D. degrees in EECS from KAIST, Korea, in 1999, 2001 and 2006, respectively. Dr Lee is currently an Associate Professor in the Department of Digital Media, Soongsil University. He is interested in robotic intelligence & emotion and media robotics.

E-mail address: kanghee.lee@ssu.ac.kr