

구개인두부전증 환자와 모의 음성의 모음과 자음 분석

성미영¹ · 김희진² · 권택균³ · 성명훈³ · 김우일^{1*}

Analysis on Vowel and Consonant Sounds of Patent's Speech with Velopharyngeal Insufficiency (VPI) and Simulated Speech

Mee Young Sung¹ · Heejin Kim² · Tack-Kyun Kwon³ · Myung-Whun Sung³ · Wooil Kim^{1*}

^{1*}School of Computer Science & Engineering, Incheon National University, Incheon 406-772, Korea

²Otorhinolaryngology, Kyungpook National University School of Medicine, Daegu 700-842, Korea

³Otorhinolaryngology, Seoul National University College of Medicine, Seoul 110-744, Korea

요 약

본 논문에서는 구개인두부전증(VPI) 환자 발음과 정상인의 모의 발음에 대한 듣기 평가와 음향 분석을 실시한다. 본 연구를 위해 음성 데이터 수집을 위해 50개의 단어, 모음 및 단음절로 이루어진 발음 목록을 설정한다. 듣기 평가 실험의 편의를 위해 웹 기반의 듣기 평가 시스템을 구축한다. 듣기 평가 결과는 실제 VPI 환자의 발음에 대한 오인식 경향과 모의 발음의 오인식 경향이 유사함을 나타낸다. 이러한 유사성은 모음의 포먼트 위치와 자음의 스펙트럼의 비교를 통해서도 확인할 수 있다. 실험 결과는 본 연구에서 사용한 정상인의 VPI 모의 발화 기법이 실제 환자의 음성을 비교적 효과적으로 모의하는 것을 반영하는 결과이다. 향후 VPI 환자의 음성 인식 과정에서 정상인의 모의 발화 음성 데이터를 음향 모델의 적용 기법과 같은 분야에 유용하게 사용할 수 있을 것으로 기대한다.

ABSTRACT

This paper focuses on listening test and acoustic analysis of patients' speech with velopharyngeal insufficiency (VPI) and normal speakers' simulation speech. In this research, a set consisting of 50-words, vowels and single syllables is determined for speech database construction. A web-based listening evaluation system is developed for a convenient/automated evaluation procedure. The analysis results show the trend of incorrect recognition for VPI speech and the one for simulation speech are similar. Such similarity is also confirmed by comparing the formant locations of vowel and spectrum of consonant sounds. These results show that the simulation method for VPI speech is effective at generating the speech signals similar to actual VPI patient's speech. It is expected that the simulation speech data can be effectively employed for our future work such as acoustic model adaptation.

키워드 : 구개인두부전(VPI), 모의 발화, 듣기 평가, 음향 분석, 포먼트

Key word : VPI (Velopharyngeal Insufficiency), Simulation speech, Listening test, Acoustic analysis, Formant

접수일자 : 2014. 03. 06 심사완료일자 : 2014. 03. 18 게재확정일자 : 2014. 03. 31

* **Corresponding Author** Wooil Kim(wikim@incheon.ac.kr, Tel:+82-32-835-8459)

School of Computer Science and Engineering, Incheon National University, Incheon 406-772, Korea

Open Access <http://dx.doi.org/10.6109/jkiice.2014.18.7.1740>

print ISSN: 2234-4772 online ISSN: 2288-4165

©This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Copyright © The Korea Institute of Information and Communication Engineering.

I. 서론

구순구개열은 얼굴의 선천성 기형 중 빈도가 높은 장애의 하나로, 선천적으로 입술(구순) 또는 입천장(구개)이 갈라져서 구강과 비강이 연결된 상태를 말하며, 우리나라의 경우 약 700명의 신생아 중 1명 꼴로 발생하는 유병률이 높은 질환이다[1, 2]. 정상인이 발성할 때 연구개가 비강과 구강을 차단시켜 비음을 막는 것에 반해, 구순구개열 환자는 경구개 또는 연구개가 갈라져 있거나 연구개가 선천적으로 짧기 때문에 성도로부터 나온 공기 흐름이 비강과 구강 동시에 공명하게 됨으로써 발성 및 조음 장애가 발생한다. 이러한 증상을 구개인두부전증(Velopharyngeal Insufficiency, VPI)이라고 한다.

본 연구는 VPI 환자의 음성 복원 및 인식 기술을 개발하기 위한 기초적인 과정으로서, VPI 환자의 발음과 이를 실험적으로 유사하게 발생시킨 정상인의 모의 발음을 분석한다. 이전 연구에서는 VPI 환자의 음성 처리 연구를 위해 공동 음성 데이터베이스 구축을 위한 발음 목록 설계 및 수집 환경 조성 등을 실시하였고, 수집된 일부 음성에 대해 비음도를 측정하였다[2]. 본 연구는 그 후속 작업으로서, VPI 환자의 음성 인식 연구를 위한 음성 데이터베이스 수집을 목적으로 발음 목록을 재정의하고, 수집된 데이터에 대해 듣기 평가와 음향학적 분석을 실시하였다. 실험적으로 유발시킨 정상인의 모의 발음과 동일인의 정상 발음 비교를 통해 음향학적으로 어떤 변화가 일어나는지 관찰한다. 또한, 모의 발음과 실제 환자의 발음의 유사성을 관찰함으로써, 대용량의 수집이 어려운 실제 VPI 환자 음성 데이터에 대한 연구 및 실험 과정에 모의 발음 음성 데이터를 효과적으로 사용할 수 있는지 여부를 타진하고자 한다.

II. 음성 데이터베이스 수집

본 절에서는 본 연구에서 사용한 음성 데이터베이스의 설계 및 수집 과정을 설명한다.

2.1. 단어 목록 선정

본 연구에서는 한국어 PBW (Phoneme Balanced Words) 452 단어 목록[3]에서 VPI 환자의 발음 오류의 양상이 주로 나타나는 50개의 단어를 언어치료가 선정하여 발음 목록으로 사용하였다. 향후 VPI 환자 음성 인식 실험 및 연구를 위해 음성 인식 분야에서 단어 인식 시스템에 공통적으로 사용되고 있는 PBW452 목록을 사용하였다. 표 1은 본 연구에서 발음 목록으로 사용한 50개의 단어 리스트를 나타낸다.

본 연구에서는 VPI 환자 음성의 음향 분석을 위해 단어 목록 뿐 아니라 한국어 음소 발음을 함께 수집하였다. 발음 목록으로는 단모음 7개 (/ㅏ/, /ㅑ/, /ㅓ/, /ㅕ/, /ㅗ/, /ㅛ/, /ㅜ/, /ㅠ/, /ㅡ/, /ㅣ/, /ㅞ/, /ㅟ/)를 선정했고, 자음 발음 평가를 위해서는 화자(Speaker)의 발음과 듣기 평가 과정을 용이하게 하기 위해 초성 자음과 모음 /ㅏ/가 결합된 19개의 단음절 (/가/, /나/, /다/, /라/, /마/, /바/, /사/, /아/, /자/, /차/, /카/, /타/, /파/, /하/, /까/, /따/, /빠/, /싸/, /짜/) 목록을 선정했다).

표 1. PBW452 목록으로부터 선정된 50개의 단어 목록
Table. 1 50-word list selected from PBW452 list

컴퓨터, 계획하고, 뒷받침해, 쾌적한, 그쪽으로, 과부가, 분석됐다, 감수성이, 변증법적, 부딪혀, 툭툭히, 빼앗기고, 깨끗이, 세탁비누, 수필집, 바쁘게, 보컬선거, 뜻밖에, 두번째, 깨닫게, 표적수사, 펼치고, 코방귀를, 국제청, 집계됐다, 뒤쪽에, 뱃속에, 재검토, 끌어안고, 불빛이, 솜씨가, 붙잡고, 종소리가, 수법을, 교통사고, 아카데미, 스포츠, 생태계, 손쉽게, 백화점, 수수께끼, 첫번째, 총체적, 편집국장, 플라스틱, 뿔났다, 깨닫고, 재조정, 덮었다, 불법복제

2.2. 수집 대상

음성 녹음에 협조가 잘 되고, 발화 목록에 따른 발음이 적절히 되도록 하기 위해 만 10세 이상의 VPI 환자와 정상 발음을 가진 성인을 대상으로 하였다. 모집과정에서 정상 모의 환자와 VPI 환자의 녹음 의지를 확인한 후 피험자 동의를 받았다. 대부분의 발화자들은 구개구순열 수술 후 언어 치료를 위해 외래에 정기적으로 내원하는 환자로 구성되었다.

1) /아/는 자음+/아/의 형식이 아니지만, 듣기평가 비교에서 /아/를 사용하기 위해 단음절 목록에 포함시켰다.

2.3. 수집 방법

녹음은 주변 환경 소음을 최대한 피하기 위해 외래 진료 후 언어치료실 한 곳을 지정하여 시행하였다. 녹음 과정에는 음성 언어치료사, 의공학과 연구원, 이비인후과 의사가 참여하였고, 언어치료사의 주도하에 발화자가 긴장하지 않도록 최대한 편한 분위기를 조성하였다. 입으로부터 40cm 떨어진 위치에 고감도 마이크를 설치하여 녹음을 실시하였다. 마이크로부터 나오는 음성신호를 다목적 USB 녹음장치인 U46XL (SuESI Audiotechnik GmbH, Leonberg, Germany)을 이용하여 주파수 44.1 kHz, 양자화 비트수 16 bit로 디지털화하고 Cubase LE5 소프트웨어를 이용하여 녹음 파일을 취득하였다.

정상인으로부터 실험적으로 VPI 환자의 발음을 유발하기 위해 1mm 내경을 가지는 고무관 (Nelaton Catheter)을 사용하였다 [2]. 그림 1과 같이 카테터를 양측 비강을 통해 넣고, 긴장도가 없는 상태에서 위치 (1)을 지혈겸자 (Hemostatic Clamp)로 표시해 놓고, 통증을 유발하지 않는 선에서 최대의 긴장도가 생성되는 위치 (2)를 표시하였다. 벨라톤 카테터가 (2)의 위치에 있을 때를 VPI 모의 환자 중증 (Severe) 상태로 하였고, (1)가 (2)의 중간위치에 있을 때 VPI 발음이 녹음된 것일 정도 (Mild) 상태로 정의하였다.

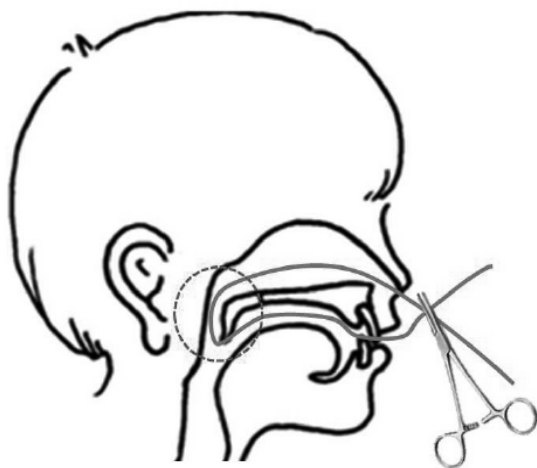


그림 1. 벨라톤 카테터를 사용하여 구개인두부전증을 모의한 그림 [2]
Fig. 1 Illustration of simulated velopharyngeal Insufficiency with a nelaton tube [2]

본 논문에서는 수집된 데이터베이스 중 VPI 환자 2명과 정상인 2명의 데이터를 사용하였다. VPI 환자는 남녀 각 1명으로 구성되고, 정상인은 30대의 남성 2명으로 구성되었다. VPI 환자는 단어 목록 및 음소 세트를 3회 반복 발음하였고, 정상인은 정상 발음과 모의 발음에 대해 단어 목록은 3회, 음소 세트를 5회 발음하였다.

III. 듣기 평가 시스템 및 결과 분석

본 절에서는 VPI 환자의 발음과 정상인의 모의 발음의 음소별 명료도 (Intelligibility)를 측정하기 위해 실시한 듣기 평가 실험의 결과를 기술한다. 이에 앞서 듣기 평가 실험을 위해 구축한 웹 기반의 듣기 평가 시스템에 대해서 설명한다.

3.1. 웹기반 평가 시스템 설계 및 구축

본 연구에서는 다수의 듣기 평가 시험자를 확보하고 실험의 편의를 위해 인터넷으로 접속하여 듣기 평가 실험에 참여할 수 있는 웹 기반의 듣기 평가 시스템을 구축하였다. 평가 시험자들은 인터넷을 통해 듣기 평가 시스템에 접속하여 부여받은 아이디로 로그인을 하여 평가에 참여하게 된다. 그림 2와 같이 해당 평가 시험자에 부여된 평가 목록에서 각 파일을 클릭하여 해당 음성 (모음 또는 단음절)을 들을 수 있게 하고, 평가 시험자가 인식하여 판단한 음성의 내용을 선택하고 저장함으로써 평가를 실시한다. 음성 파일은 16 kHz로 다운 샘플링되어 Windows *.wav 형식으로 저장되었다. 각 음성 파일 당 듣기 횟수는 최대 5회로 제한하였으며, 각 평가 시험자는 집, 사무실, 연구실 등과 같이 조용한 환경에서 헤드폰 또는 이어폰을 사용하여 듣기 평가에 참여하는 것을 원칙으로 했다.

주관적 듣기 평가 실험에는 총 15명의 정상 듣기 능력을 가진 20-30대의 남녀로 구성된 시험자가 참여하였다. 한명의 피시험자 (즉, VPI 환자 또는 정상인 발화자) 음성의 각 음소 종류에 대해 20회의 듣기 평가가 실시될 수 있도록 평가 시험자의 듣기 평가 목록을 설계하였다. 그 결과 각 시험자는 총 8개의 발음 세트 (모음 7개, 단음절 19개)에 대해 듣기 평가에 참여하였다. 시험자에게는 발화자에 대한 어떠한 정보도 제공되지 않고, 각 평가 목록의 음소 및 단음절 순서는 무작위로 나열

되도록 설계하여 평가 시험자가 예측에 의한 판단이 불가능하도록 하였다.



그림 2. 웹 기반 듣기 평가 시스템의 샘플 페이지.
 Fig. 2 A sample page of the web-based listening evaluation system developed for this research.

3.2. 듣기 평가 결과 및 분석

표 2는 앞 절에서 설명한 웹 기반 듣기 평가 시스템을 이용하여 실시한 정상인 화자의 정상 발화 및 모의 발화 음성에 대한 듣기 평가 결과이다. 정상인의 경우 정상 발화에 대해서는 모음과 단음절 모두 80-90% 정도의 인식정확도를 나타냈다. 모의 발음의 경우 모음에 대해서는 70% 정도의 정확도를 나타냈지만, 단음절의 경우 정상인 1은 28.95%, 정상인 2는 59.21%의 인식정확도를 보였다. 특히 정상인 1의 모의 발음의 자음 성분을 포함한 단음절에 대해 오인식 정도가 심한 것을 알 수 있다. 이러한 결과는 모의 발화의 경우 모음에 비해 자음의 음성 변형이 심한 것을 반증한다.

표 3부터 표 6까지는 정상인 화자의 모의 발음에 대한 듣기 평가 결과를 오인식표 (Confusion Matrix)로 나타낸 것이다. 본 논문에서의 오인식표는 각 행의 발음에 대해 각 열의 어떤 발음으로 몇 회 인식되었는지를 나타낸다. 표 3과 4는 정상인 화자 1과 2의 모음 발음에 대한 오인식표를 나타낸다. 총 20회 듣기 평가 중 10회 이상 정확하게 인식된 결과를 불투명한 칸으로 표시하였다. 모의 발화 모음의 경우 /우/와 /이/ 발음에 대해 인식 정확도가 매우 낮은 것을 알 수 있다. 발음 /우/는 주로 /오/ 발음으로 오인식 되었으며, 발음 /이/는 주로 /으/ 발음으로 오인식된 것을 확인할 수 있다.

표 2. 정상인의 정상 발화 및 모의 발화에 대한 듣기 평가 결과 (인식 정확도, %)

Table. 2 Result of listening test for normal speakers' normal speech and simulation speech (recognition accuracy, %)

		모음	단음절
정상인 1	정상발화	87.86	80.79
	모의발화	70.71	28.95
정상인 2	정상발화	92.86	90.00
	모의발화	68.57	59.21

표 3. 정상인 1의 모의 발화 모음에 대한 오인식표

Table. 3 Confusion matrix for normal speaker 1's simulation vowel speech

	아	어	오	우	으	이	애	인식률(%)
아	18	2	0	0	0	0	0	90
어	0	18	2	0	0	0	0	90
오	0	0	20	0	0	0	0	100
우	0	1	11	8	0	0	0	40
으	0	0	1	3	16	0	0	80
이	0	0	0	0	19	1	0	5
애	0	2	0	0	0	0	18	90

표 4. 정상인 2의 모의 발화 모음에 대한 오인식표

Table. 4 Confusion matrix for normal speaker 2's simulation vowel speech

	아	어	오	우	으	이	애	인식률(%)
아	20	0	0	0	0	0	0	100
어	0	20	0	0	0	0	0	100
오	0	3	17	0	0	0	0	85
우	0	0	14	4	2	0	0	20
으	0	0	0	1	19	0	0	95
이	0	0	0	0	20	0	0	0
애	0	1	0	0	4	0	16	80

표 5와 6은 정상인의 모의 발화의 단음절에 대한 듣기 평가를 실시한 결과로서, 자음 성분에 관한 인식 명료도를 나타낸다. 자음의 경우에는 총 20회 중 8회 이상 인식된 결과를 불투명한 칸으로 나타내었다. 정상인 1과 2 모두 발음 /나/를 /아/로 오인식 하는 경우가 많았고, 발음 /다/와 /자/의 경우 인식정확도가 극히 낮았다. 또한 많은 자음 성분이 발음 /하/로 오인식되는 경우가 많았다.

표 7은 VPI 환자의 음성에 대한 듣기 평가 결과이다. 정상인 화자에 비해 전체적으로 인식 정확도가 낮았다. 특히 환자 1의 경우 모음과 자음 모두 인식 정확도가 특히 낮은 것을 알 수 있다. 정상인 화자와 마찬가지로

VPI 환자의 음성에 대한 듣기 평가 결과를 표 8에서 11까지의 오인식표로 나타내었다. 표 8과 9는 환자 1과 2의 모음 발음에 대한 오인식표이다. 환자 1의 발음이 상대적으로 오인식한 정도가 많은 것을 확인할 수 있다. 환자 1의 경우 /어/, /우/, /오/의 발음이 모두 /오/로 오인식되는 경우가 많은 것을 확인할 수 있다. 환자 1과 2의 모음 발음 듣기 평가에서 공통적으로 나타나는 결과는, (1) /우/의 발음이 오인식 정도가 상대적으로 높았고, (2) /이/의 발음이 대부분 /오/로 오인식되었다. 이와 같은 결과는 앞에서 정상인의 모의 발화 모음의 오인식표에서 살펴본 결과와 유사한 양상이다. 이와 같은 유사성은 본 연구에서 사용한 모의 발화 방법이 VPI 환자의 모음의 장애 발음을 비교적 유사하게 모의하는 것을 입증하는 결과이다.

표 10과 11은 환자 1, 2의 자음 발음에 대한 듣기 평가 결과의 오인식표이다. 두 환자의 평가에서 공통으로 나타나는 사항은 다음과 같다. (1) 발음 /다/에 대한 인식정확도가 극히 낮다. (2) /카/, /타/의 발음이 모두 /파/로 오인식되는 정도가 높다. (3) 경음 /까/, /따/, /빠/, /싸/, /짜/ 모두 인식 정확도가 매우 낮다. (4) 많은 자음이 /파/로 오인식되는 경우가 많다. 경음을 발음하기 위해서는 연구개의 작용이 동반되는 후두 긴장을 요구하는데, VPI 환자의 경우 연구개의 동작이 불완전해짐에 따라 정확한 경음을 발음하기 어려워지는 것으로 해석할 수 있다.

표 5. 정상인 1의 모의 발화 자음에 대한 오인식표.

Table. 5 Confusion matrix for normal speaker 1's simulation consonant speech.

	ㄱ	ㄴ	ㄷ	ㄹ	ㅁ	ㅂ	ㅅ	ㅇ	ㅈ	ㅊ	ㅋ	ㆁ	ㅅ	ㅇ	ㅈ	ㅊ	ㅋ	ㆁ	ㅅ	ㅇ
ㄱ	2					14	2					2								10
ㄴ		1	1	8		1	8					1								5
ㄷ			1	5	1	2	1	2			1	7								25
ㄹ				18			1	1												90
ㅁ				2	16	1	1													80
ㅂ				1	19															95
ㅅ	1		4	1			2	2			1	4	5							0
ㅇ					1	17						2								85
ㅈ		3	4	2	2	3	1	3			2									15
ㅊ	1		1				3				1	4	10							0
ㅋ			1		2	1			1	1	4	10								5
ㆁ			1			6			1	1	3	8								5
ㅅ					8					1	10	1								50
ㅇ					7	1			1	4	7									35
ㅈ	1			1	1	8	9													0
ㅊ				8			9		1	1								1		5
ㅋ					6	1					4	1						8		40
ㆁ		1	1	3		4	6		1	2	1									0
ㅅ			5	4			8		1	1									1	5

표 6. 정상인 2의 모의 발화 자음에 대한 오인식표

Table. 6 Confusion matrix for normal speaker 2's simulation consonant speech

	ㄱ	ㄴ	ㄷ	ㄹ	ㅁ	ㅂ	ㅅ	ㅇ	ㅈ	ㅊ	ㅋ	ㆁ	ㅅ	ㅇ	ㅈ	ㅊ	ㅋ	ㆁ	ㅅ	ㅇ
ㄱ	15							2			2			1						75
ㄴ		6		4				10												30
ㄷ		5	4					1			3	1		5	1					20
ㄹ				10			1							9						50
ㅁ					20															100
ㅂ	1					18		1												90
ㅅ			1	1			13							5						65
ㅇ								18						2						90
ㅈ	3			1				4		4	1	4		3						20
ㅊ				1			1	1	11	2	3								1	55
ㅋ										18		1	1							90
ㆁ			2				5			2	9	1	1							45
ㅅ			1			4					1	14								70
ㅇ						2		2					16							80
ㅈ	4			1	1			1		1	1			9			2			45
ㅊ			4	2					2		1			1	8				2	40
ㅋ		1			2	3		3				2				9				45
ㆁ							8									12				60
ㅅ									6								3			55

표 7. VPI 환자의 발화에 대한 듣기 평가 결과 (인식 정확도, %)

Table. 7 Result of listening test for VPI patients' speech (recognition accuracy, %)

	모음	자음
환자 1	37.86	26.32
환자 2	63.57	44.74

표 8. 환자 1의 발화 모음에 대한 오인식표.

Table. 8 Confusion matrix for patient 1's vowel speech.

	아	어	오	우	으	이	애	인식률(%)
아	18	2	0	0	0	0	0	90
어	0	3	16	0	0	1	0	15
오	0	0	20	0	0	0	0	100
우	0	0	17	3	0	0	0	15
으	0	0	11	8	1	0	0	5
이	0	0	0	0	20	0	0	0
애	0	8	1	0	3	0	8	40

표 9. 환자 2의 발화 모음에 대한 오인식표

Table. 9 Confusion matrix for patient 2's vowel speech

	아	어	오	우	으	이	애	인식률(%)
아	20	0	0	0	0	0	0	100
어	0	15	5	0	0	0	0	75
오	0	0	11	9	0	0	0	55
우	0	2	1	9	8	0	0	45
으	0	1	0	3	16	0	0	80
이	0	0	0	0	19	1	0	5
애	0	0	0	0	3	0	17	85

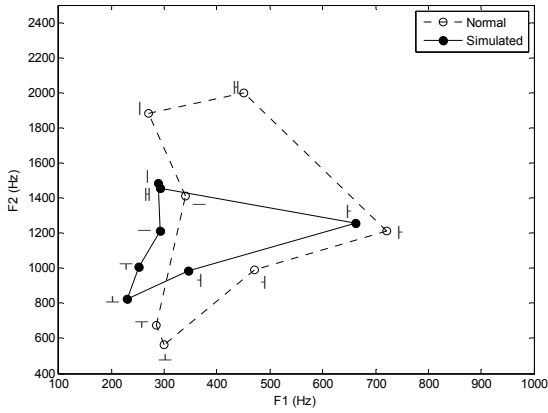


그림 4. 정상인 2의 정상 및 모의 발화 중 모음의 F1과 F2 위치 비교

Fig. 4 F1 & F2 location comparison for vowels of normal speaker 2's normal & simulation speech

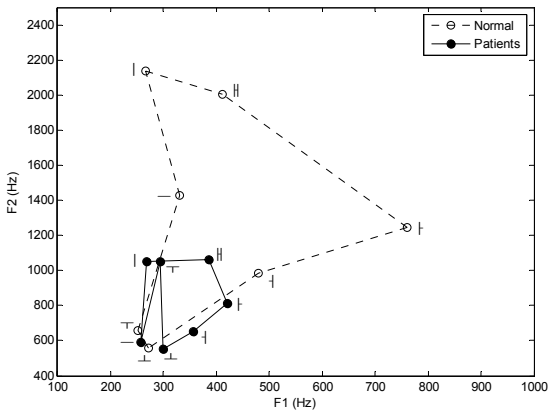


그림 5. VPI 환자 1과 정상 발음의 모음의 F1과 F2 위치 비교

Fig. 5 F1 & F2 location comparison for vowels of VPI patient 1 and normal speakers' normal speech

그림 5는 VPI 환자 1의 모음 음성의 포먼트 위치를 나타낸 것이다. 정상 발음과의 비교를 위해 그림 3과 4에서 살펴본 정상인 1과 2의 정상 발화 모음의 F1, F2 위치의 평균을 점선으로 나타내었다. 정상인 발음과 비교하여 /오/와 /우/ 발음을 제외하고 모두 제 2 포먼트 주파수가 대폭적으로 하락하는 것을 관찰할 수 있다. 특히, /어/, /우/, /으/ 발음 모두 정상인의 발음 /오/의 포먼트 위치와 유사하게 나타나는 것을 알 수 있으며, /이/와 /애/ 발음은 제 2 포먼트 위치가 급격하게 낮아지는 것을 알 수 있다[4].

한국어 단모음 /이/, /애/와 같은 전설 모음에서는 제

2 포먼트의 위치가 높은 주파수에서 형성되는데[5], VPI 환자의 경우에는 연구개의 차단 동작이 불완전해지면서 후설 모음화되는 것으로 해석할 수 있다. 이와 같은 결과는 표 8에서 확인한 오인식 경향을 설명해주는 결과이다. 제 2 포먼트 주파수가 급격하게 낮아지는 현상은 그림 3과 4와 같이 정상인 1과 2의 경우에서도 유사하게 관찰할 수 있다. 이러한 양상은 듣기 평가의 결과와 유사한 결과로서 본 실험에서 사용한 VPI 모의 발화 방법이 실제 환자와 유사한 경향의 모음 발음을 발생시키는 것을 음향학적 측정값 (즉, 포먼트의 위치)으로 확인한 것이다. 환자 2는 여성이고 정상인 화자는 모두 남성이므로 본 실험에서 포먼트 위치 비교는 생략하였다.

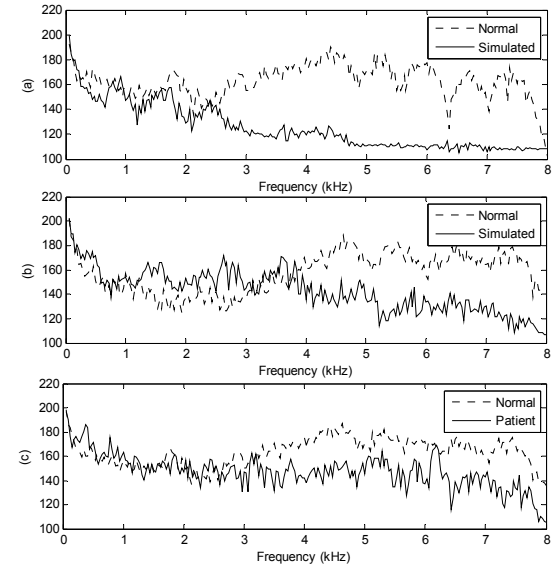


그림 6. 자음 /ㅈ/의 스펙트럼 비교; (a) 정상인 1의 정상 및 모의 발음, (b) 정상인 2의 정상 및 모의 발음, (c) VPI 환자 및 정상 발음

Fig. 6 Spectrum comparison for consonant /ㅈ/ (a) Normal speaker 1, (b) normal speaker 2, and (c) VPI patient

그림 6은 VPI 환자 및 모의 발화의 듣기평가에서 공통적으로 오인식이 많이 나타난 단모음 /자/에서 자음 /ㅈ/ 부분의 평균 스펙트럼을 (a) 정상인 1, (b) 정상인 2, (c) 환자 1에 대해 나타낸 것이다. 그림 (a)와 (b)에서 점선은 정상 발음, 실선은 모의 발음을 나타내며, 그림 (c)에서 점선은 정상인 1과 2의 평균, 실선은 환자 1의 발

음이다. 모의 발음과 실제 VPI 환자의 자음 /ㅈ/ 발음의 경우 정상 발음에 비해 3kHz 이상의 고주파수 성분의 크기가 작아진 것을 알 수 있다. 이러한 대폭적인 주파수 성분의 변화가 자음 /ㅈ/의 음가 성질을 변형함으로써 높은 오인식률을 가져온 것으로 해석할 수 있다.

V. 결 론

본 논문에서는 VPI 환자의 발음과 정상인의 모의 발음에 대한 듣기 평가와 음향 분석을 실시하였다. 음성 데이터 수집을 위해 PBW452 목록에서 50개의 단어를 선택하고, 모음과 단음절로 이루어진 발음 목록을 설정하였다. 다수의 평가자를 확보하고 듣기 평가의 편의를 위해 웹 기반의 듣기 평가 시스템을 구축하였다. 듣기 평가 결과를 오인식표를 통해 분석하였으며, 모음 성분에 대한 포먼트 위치 비교와 자음 성분에 대한 스펙트럼 비교를 통해 음향학적 분석을 실시하였다. 듣기 평가 결과는 실제 VPI 환자의 발음에 대한 오인식 경향과 모의 발음의 오인식 경향이 유사함을 나타내었다. 이러한 유사성은 포먼트의 위치와 스펙트럼의 변화를 통해서도 확인할 수 있는 결과이다. 이러한 결과는 본 연구에서 사용한 정상인의 VPI 모의 발화 기법이 실제 VPI 환자의 음성을 비교적 효과적으로 모의하는 것을 반영하는 결과이며, 향후 VPI 환자의 음성 인식 연구에서 정상인의 모의 발화 음성 데이터를 음향 모델의 적용 기법과 같은 분야에 유용하게 사용할 수 있을 것

로 기대한다.

감사의 글

이 논문은 미래창조과학부 공공복지안전연구사업 (No. 2013-2244) 지원에 의하여 연구되었음.

REFERENCES

- [1] S. G. Fletcher, "Theory and instrumentation for quantitative measurement of nasality," *Cleft Palate Journal*, vol. 7, pp. 601 - 609, 1970.
- [2] J.-E. Lee, et al., "Research on Construction of the Korean Speech Corpus in Patient with Velopharyngeal Insufficiency," *Korean Journal of Otorhinolaryngol - Head & Neck Surgery*, vol.55, no.8, pp.498-507, 2012 .
- [3] B.-W. Kim, et al., "A Study on the Design and the Construction of a Korean Speech DB for Common Use," *Journal of the Acoustic Society of Korea*, vol.16, no.4, pp.35-41, 1997.
- [4] J. Yun, et al., "Nasometric and Acoustic Analysis in Experimentally Induced Velopharyngeal Insufficiency in Human," *Journal of the Korean Society of Logopedics and Phoniatrics*, vol18, no.2, pp.210-216, 1997.
- [5] J. Shin, *Voice of Korean Language*, Jisk and Gyoyang, 2011.



성미영(Mee Young Sung)

1990년 프랑스 INSA de Lyon 컴퓨터공학 박사
 1990년 ~ 1993년 한국전자통신연구소 선임연구원
 1993년 ~ 현재 인천대학교 컴퓨터공학부 교수
 2001년 ~ 2002년 미국 카네기 멜론 대학교 교환교수
 2008년 ~ 2009년 미국 UC 버클리 대학교 교환교수
 ※관심분야: 멀티미디어, 가상현실, 햅틱스, 음성인식



김희진(Heejin Kim)

2007년 서울대학교 의과대학 학사
 2011년 ~ 2013년 서울대학교 의과대학 이비인후과학 석사
 2014년 ~ 서울대학교 의과대학 이비인후과학 박사
 2013년 ~ 현재 경북대학교 의과대학 이비인후과 임상조교수
 ※관심분야: 두경부 종양, 음성수술, 음성질환



권택균(Tack-Kyun Kwon)

2006년 서울대학교 의과대학 의학박사
2003년 ~ 2004년 미국 피츠버그 의과대학 Voice Clinic Fellow
2012년 ~ 2013년 미국 샌디에고 대학 Clinical Research 석사과정
2012년 ~ 현재 서울대학교 의과대학 이비인후과학 부교수
※관심분야: 음성수술, 음성질환진단, 음성분석, 임상시험 및 연구



성명훈(Myung-Whun Sung)

1991년 서울대학교 의과대학 의학박사
1990년 ~ 1999년 서울대학교 의과대학 이비인후과 조교수
1993년 ~ 1995년 미국 피츠버그 의과대학 Research Fellow
1999년 ~ 2004년 서울대학교 의과대학 이비인후과 부교수
2004년 ~ 현재 서울대학교 의과대학 이비인후과 교수
※관심분야: 두경부 종양, 음성수술, 음성질환, 임상시험 및 연구



김우일(Wooil Kim)

2003년 고려대학교 전자공학과 공학박사
2004년 ~ 2005년 미국 카네기 멜론 대학교 박사후 연구원
2005년 ~ 2012년 미국 텍사스 주립대 (University of Texas at Dallas) 연구원 및 연구교수
2012년 ~ 현재 인천대학교 컴퓨터공학부 조교수
※관심분야: 신호처리, 패턴인식, 음성인식, 휴먼 컴퓨터 인터페이스