(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0031888
(43) 공개일자 2021년03월23일

(51) 국제특허분류(Int. Cl.)
*G10L 15/25* (2013.01) *G10L 15/02* (2006.01)
*G10L 15/04* (2006.01) *G10L 21/06* (2006.01)
(52) CPC특허분류
*G10L 15/25* (2013.01)
*G10L 15/02* (2013.01)
(21) 출원번호 10-2021-0033765(분할)
(22) 출원일자 2021년03월16일
심사청구일자 2021년03월16일
(62) 원출원 특허 10-2018-0116479
원출원일자 2018년09월28일
심사청구일자 2018년09월28일
(30) 우선권주장
1020170126470 2017년09월28일 대한민국(KR)
1020170126769 2017년09월29일 대한민국(KR)

(71) 출원인
**인하대학교 산학협력단**
인천광역시 미추홀구 인하로 100(용현동, 인하대학교)
(72) 발명자
**이우기**
경기도 군포시 산본로432번길 25, 1222동 1302호(산본동, 한양목련아파트)
**심봉섭**
인천광역시 연수구 컨벤시아대로42번길 96, 601동 603호(송도동, 더샵 엑스포 6단지)
*(뒷면에 계속)*
(74) 대리인
**네이트특허법인**

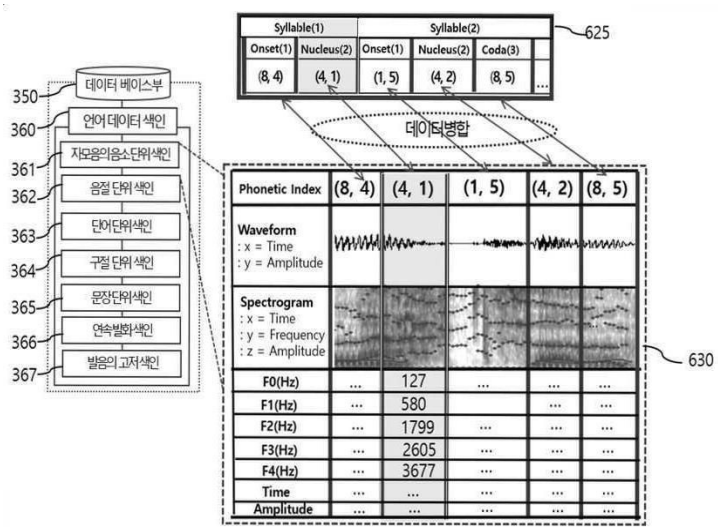
전체 청구항 수 : 총 19 항

(54) 발명의 명칭 **조음기관의 물리 특성과 음성 및 문자 간 매칭을 통한 발화 의도 측정 및 발화 구현 시스템**

## (57) 요 약

본 발명은, 화자의 두경부의 일면에 인접하여 조음기관의 물리특성을 측정하는 센서부; 상기 센서부의 위치와 상기 조음기관의 물리특성을 기반으로 화자의 발화 특징을 파악하는 데이터해석부; 상기 센서부의 위치와 상기 발화특징을 언어데이터로 변환하는 데이터변환부; 상기 데이터변환부(300)의 상기 언어데이터를 음성정보로 생성하는 텍스트정보부(600); 상기 텍스트정보부와 연동되어, 언어데이터 기반의 상기 음성정보를 외부로 표현하는 데이터표현부; 상기 센서부는, 구강설에 대응되는 구강설 센서를 포함하는 발화 구현 시스템 을 제공한다.

## 대 표 도



(52) CPC특허분류
*G10L 15/04* (2013.01)
*G10L 21/06* (2013.01)
(72) 발명자
**권헌도**
경상북도 안동시 강남1길 98-23, 305동 506호 (정하동, 현진에버빌3차)
**김덕환**
서울특별시 양천구 목동로 230, 103동 1004호 (목동, 아이파크아파트)
**신진호**
인천광역시 남구 능해길 32, 1동 402호 (용현동, 동아아파트)

# 명 세 서

## 청구범위

### 청구항 1

화자의 두경부의 일면에 인접하여 조음기관의 물리특성을 측정하는 센서부;

상기 센서부의 위치와 상기 조음기관의 물리특성을 기반으로 화자의 발화 특징을 파악하는 데이터해석부;

상기 센서부의 위치와 상기 발화특징을 언어데이터로 변환하는 데이터변환부;

상기 데이터변환부의 상기 언어데이터를 음성정보로 생성하는 텍스트 정보부;

상기 텍스트정보부와 연동되어, 언어데이터 기반의 상기 음성정보를 외부로 표현하는 데이터표현부;

상기 센서부는, 구강설에 대응되는 구강설 센서를 포함하며,

상기 데이터해석부는,

상기 센서부에서 측정되는 상기 구강설과 다른 조음기관과의 물리특성을 통해 상기 화자가 발화하는 자모음, 어휘 단위 강세 (Lexical Stress), 문장 단위 강세(Tonic stress) 중 적어도 하나의 발화 특징을 파악하고,

상기 센서부에 의해 측정되는 상기 조음기관의 물리특성에 의한 발화 특징을 파악함에 있어서, 2진수 내지 실수를 포함하는 수치로 구성된 표준 발화 특징 행렬을 기반으로 상기 화자의 발음과 강세의 유사근접도, 발화 의도 중 적어도 하나의 발화 특징을 측정하며,

상기 센서부의 위치, 상기 화자의 발화 특징, 상기 화자의 음성에 대응하는 적어도 하나의 언어 데이터 색인을 포함하는 데이터베이스부와 연동되며,

상기 데이터표현부는,

상기 데이터베이스부의 언어 데이터 색인과 연동되어, 상기 화자의 발화 특징을 자모음의 음소(Phoneme)단위, 적어도 하나의 단어단위, 적어도 하나의 구절단위(Citation Forms), 적어도 하나의 문장단위, 연속발화단위(Consecutive Speech) 중 적어도 하나의 발화 표현을 나타내는 발화 구현 시스템.

### 청구항 2

제 1 항에 있어서, 상기 구강설 센서는,

상기 구강설의 일측면에 고착되거나, 상기 구강설의 표면을 감싸거나, 상기 구강설 내부에 삽입되고, 발화에 따른 상기 구강설의 x축, y축, z축 방향 기반의 시간에 따른 벡터량의 변화량을 파악하여, 상기 구강설의 저고도, 전후설성, 굴곡도, 신전도, 회전도, 긴장도, 수축도, 이완도, 진동도 중 적어도 하나의 물리특성을 파악하는 발화 구현 시스템.

### 청구항 3

제 1 항에 있어서, 상기 구강설 센서는,

상기 구강설의 일측면에 고착되거나, 상기 구강설의 표면을 감싸거나, 상기 구강설 내부에 삽입되고, 발화에 따른 상기 구강설의 x축, y축, z축 방향 기반의 단위 시간 당 회전하는 각도의 변화량을 파악하여, 상기 구강설을 포함한 상기 조음기관의 물리 특성을 파악하는 발화 구현 시스템.

### 청구항 4

제 1 항에 있어서, 상기 구강설 센서는,

상기 구강설의 일측면에 고착되거나, 상기 구강설의 표면을 감싸고, 발화에 따른 상기 구강설의 수축 및 이완으로 발생하는 물리력에 따라 결정구조의 변화에 기인하는 편극에 대응되는 전기신호가 발생하는 압전소자를 통해 상기 구강설의 굽힘도를 파악하여, 상기 구강설의 저고도, 전후설성, 굴곡도, 신전도, 회전도, 긴장도, 수축도, 이완도, 진동도 중 적어도 하나의 물리특성을 파악하는 발화 구현 시스템.

**청구항 5**

제 1 항에 있어서, 상기 센서부는,

상기 구강설이 상기 두경부 내외의 다른 조음기관과의 상호작용에 기인하는 접근 및 접촉에 대응되는 마찰전기(Tribo Electric Generator)에 따른 파열도, 마찰도, 공명도, 접근도 중 적어도 하나의 물리특성을 파악하는 마찰대전소자를 포함하는 발화 구현 시스템.

**청구항 6**

제 1 항에 있어서, 상기 데이터해석부는,

상기 센서부에 의해 측정되는 상기 조음기관의 물리특성을 발화 특징을 파악함에 있어서, 상기 조음기관의 물리특성을 각 자모음 단위의 패턴으로 인식하는 단계, 상기 자모음 단위의 패턴의 특징을 추출하고, 추출된 상기 자모음 단위의 패턴의 특징을 유사도에 따라 분류하는 단계, 분류된 상기 자모음 단위의 패턴의 특징을 재조합하는 단계, 상기 조음기관의 물리특성을 상기 발화특징으로 해석하는 단계에 따라 상기 발화 특징을 파악하는 발화 구현 시스템.

**청구항 7**

제 8 항에 있어서, 상기 데이터해석부는,

상기 센서부에 의해서 측정되는 상기 조음기관의 물리특성에 의해, 자모음의 동화(Assimilation), 이화(Dissimilation), 탈락(Elision), 첨가(Attachment), 강세(Stress)와, 약화(Reduction)로 야기되는 기식음화(Asperation), 음절성자음화(Syllabic cosonant), 탄설음화(Flapping), 경음화(Tensification), 순음화(Labilalization), 연구개음화(Velarization), 치음화(Dentalizatiom), 구개음화 (Palatalization), 비음화(Nasalization), 강세변화(Stress Shift), 장음화(Lengthening) 중 적어도 하나의 이차조음현상인 발화 변이를 측정하는 발화 구현 시스템.

**청구항 8**

제 1 항에 있어서,

상기 구강설 센서는, 센서 작동을 위한 회로부, 상기 회로부를 감싸는 캡슐부, 상기 구강설 일면에 부착되는 접착부를 포함하는 발화 구현 시스템.

**청구항 9**

제 1 항에 있어서,

상기 구강설 센서는, 박막 회로를 가진 필름 형태로서 상기 구강설에 인접하여 작동하는 발화 구현 시스템.

**청구항 10**

제 1 항에 있어서,

상기 센서부는, 두경부 근육의 신경신호 측정의 기준 전위를 생성하는 적어도 하나의 레퍼런스 센서와, 상기 두경부 근육의 신경신호를 측정하는 적어도 하나의 양극센서 및 적어도 하나의 음극센서로 구성된 안면부 센서를 포함하는 발화 구현 시스템.

**청구항 11**

제 10 항에 있어서,

상기 데이터해석부는, 상기 안면부 센서에 기반하여 상기 센서부의 위치를 획득함에 있어서, 상기 레퍼런스 센서를 기준으로 하여 상기 적어도 하나의 양극센서 및 상기 적어도 하나의 음극센서의 전위차를 파악하여 상기 안면부 센서의 위치를 파악하는 발화 구현 시스템.

**청구항 12**

제 10 항에 있어서,

상기 데이터해석부는, 상기 안면부 센서에 기반하여 상기 화자의 발화 특징을 획득함에 있어서, 상기 레퍼런스 센서를 기준으로 하여 상기 적어도 하나의 양극센서 및 상기 적어도 하나의 음극센서의 전위차를 파악하여 상기 화자의 두경부에서 발생하는 상기 조음기관의 물리 특성에 의한 발화 특징을 파악하는 발화 구현 시스템.

**청구항 13**

제 1 항에 있어서,

상기 센서부는, 상기 화자의 두경부 중 성대에 인접하여 성대의 근전도 내지 떨림을 파악하여, 상기 화자의 발화 시작, 발화 정지, 발화 종료 중 적어도 하나의 발화 내역 정보를 파악하는 성대 센서를 포함하는 발화 구현 시스템.

**청구항 14**

제 1 항에 있어서,

상기 센서부는, 치아의 일면에 인접하여 상기 구강설 및 아랫 입술의 접촉에 기인하여 발생하는 전기적 용량 변화에 따른 신호발생 위치를 파악하는 치아센서를 포함하는 발화 구현 시스템.

**청구항 15**

제 1 항에 있어서,

상기 데이터해석부는, 상기 화자의 두경부 일면에 인접한 음성 취득 센서를 통해 발화에 따른 상기 화자의 음성을 취득하는 발화 구현 시스템.

**청구항 16**

제 1 항에 있어서,

상기 센서부는 화자의 발화 의도에 따라 발생하는 두경부의 상하좌우의 기울임, 흉곽부의 들썩임, 두경부와 흉곽부 간의 핏대 및 근육의 긴장, 상지부의 떨림, 상지부의 제스쳐, 하지부의 떨림, 하지부의 제스쳐 중 적어도 하나의 비언어적 표현을 파악하기 위해 상기 화자의 두경부를 촬상하는 촬상센서를 더욱 포함하는 것을 특징으로 하는 발화 구현 시스템.

**청구항 17**

제 1 항에 있어서,

상기 센서부의 구강설 센서, 안면 센서, 음성취득 센서, 성대 센서, 치아 센서, 촬상센서 중 적어도 하나에 전원을 공급하는 전원부를 더 포함하는 발화 구현 시스템.

**청구항 18**

제 1 항에 있어서,

상기 데이터표현부에 의해 나타나는 발화 표현은, 문자 기호, 그림, 특수기호, 숫자 중 적어도 하나로 시각화되거나, 소리 형태로 청각화되어, 상기 화자와 청자에게 제공되는 발화 구현 시스템.

**청구항 19**

제 1 항에 있어서,

상기 데이터표현부에 의해 나타나는 발화 표현은, 진동, 스누징, 태핑, 압박, 이완 중 적어도 하나의 촉각적 방법으로 상기 화자와 청자에게 제공되는 발화 구현 시스템.

## *발명의 설명*

### *기 술 분 야*

[0001] 본 발명은 조음 센서와 촬상 센서를 통해 구강설을 포함한 두경부의 조음기관의 물리 특성을 인지하여 두경부 전반의 발화에 따른 변화를 측정하고 이를 통해 발화 의도를 파악하여, 시각, 청각, 촉각화를 통해 발화 의도를 화자 본인 내지 외부로 제공하는 시스템 및 그 방법에 관한 것이다.

### *배 경 기 술*

[0003] 조음기관에서 생성되는 문자가 언어학적 정보전달인 의사소통을 위한 경우에는 발화 혹은 언어음으로 불리며 비언어학적인 경우에는 발성으로 불린다. 문자의 생성에 관여하는 인체의 주요한 기관은 신경계통과 호흡기계통이다.

[0004] 신경계통은 중추신경계와 말초신경계가 관여하는데 중추신경 중 뇌간에는 언어의 생성에 필요한 두개골 혹은 뇌신경 세포핵이 위치하며 소뇌는 동작에 대한 근육의 제어를 정밀하게 조율하는 기능이 있으며, 대뇌의 반구는 언어기능에 지배적인 역할을 한다. 언어음 생성을 위해 관여하는 두개골 신경에는 턱의 움직임을 관여하는 제 5 뇌신경, 입술운동에 관여하는 제 7 뇌신경, 인두 및 후두의 운동에 관여하는 제 10 뇌신경, 인두의 운동에 관여하는 제 11 뇌신경, 그리고 혀의 운동에 관여하는 제 12 신경 등이 있다. 말초신경 중에는 특히 미주신경에서 분지되는 상후두신경과 반회후두신경이 후두운동에 직접 관여하게 된다.

[0005] 또한 언어음은 하부 호흡기계, 후두와 성도가 상호 밀접하게 작용하여 생성된다. 성대는 문자의 근원으로, 폐로부터 송출되는 호기의 흐름이 성대를 진동시키고 발성 시 호기조절은 소리 에너지를 적절히 능률적으로 공급한다. 성대가 적당히 긴장하여 폐쇄되면 호기에 의해 성대가 진동하고 성문을 일정한 주기로 개폐시켜 성문을 통과하는 호기류를 단속하는데 이 호기의 단속류가 문자의 음원이다.

[0006] 사람이 의사소통을 목적으로 말을 사용하기 위해서는 여러 가지 생리적인 과정을 거쳐야 한다. 조음과정은 발성된 소리가 공명과정을 거쳐 증폭 및 보완된 후, 말소리의 단위인 음소를 형성해 가는 과정을 의미한다.

[0007] 조음기관으로는 혀가 가장 중요하게 생각하지만, 실제로 음소를 만드는 데는 혀뿐 아니라 구강 및 안면의 여러 가지 구조들이 관여한다. 이러한 조음기관에는 혀, 입술, 여린입천장(연구개, soft palate), 턱 등과 같이 움직일 수 있는 구조와 치아나 굳은입천장(경구개, hard palate)과 같이 움직일 수 없는 구조들이 포함된다. 이러한

조음기관들이 공기의 흐름을 막거나 제약하여 자음과 모음을 형성하게 되는 것이다.

[0008] 첫 번째 조음기관으로서의 혀는 그 부위들이 뚜렷한 경계선을 나타내지 않기 때문에 구별하는 것이 쉽지는 않으나 기능적인 측면에서 혀의 외부구조를 구별하는 것은 정상적인 조음뿐 아니라 병리적인 조음을 이해하는데 도움이 된다. 혀는 앞에서부터 혀끝(apex, tip), 혀날(blade), 혀등(dorsum), 혀몸통(body), 그리고 혀뿌리(root)로 나눌 수 있다. 혀끝을 우리가 혀를 뾰족하게 내밀거나 음절의 첫소리로 오는 /ㄹ/(예: "라라라")를 조음할 때 사용되는 부위이고, 혀날은 잇몸소리(치조음 alveolar sounds)와 같은 입의 앞쪽에서 만드는 음소들을 조음할 때 주로 사용되며, 혀등은 여린입천장소리(연구개음 velar sounds)와 같은 뒷소리 음소들을 조음할 때 주로 사용되는 혀의 부분이다.

[0009] 두 번째 조음기관으로서의 입술은 입의 입구를 이루는 부분으로 두경부 표정이나 조음에 중요한 기능을 한다. 특히 여러 가지 모음들은 혀의 움직임뿐만 아니라 입술의 모양에 의하여 음소가 구별되며, 두입술자음(양순자음 bilabial sound)들은 입술이 닫혀져야만 발음될 수 있다. 입술의 모양은 주변의 근육들에 의하여 변형된다. 예를 들어, 입술 주변을 둘러싸고 있는 입둘레근(구륜근 orbicularis oris muscle)은 입술을 다물거나 오므라들게 하여 두입술자음이나 /우/와 같은 원순모음들 발음하는 데 중요한 역할을 하며, 윗입술올림근(quadratus labii superior muscle)과 아랫입술내림근(quadrates labii inferior muscle)은 입술을 열게 한다. 또한, 입꼬리당김근(소근 risorius muscle)은 입술의 모서리를 잡아당겨 미소를 짓거나 입술을 수축시켜서 발음해야 하는 /이/와 같은 소리를 낼 때 중요한 역할을 한다.

[0010] 세 번째 조음기관으로서의 턱과 치아 중, 턱은 움직이지 않는 위턱(상악 maxilla)과 상하 및 좌우 운동을 하는 아래턱(하악 mandible)으로 구분된다. 이들 턱은 얼굴 뼈 중에서 가장 튼튼하고 큰 뼈로서 4쌍의 근육들에 의해서 움직인다. 아래턱의 움직임은 입안의 크기를 변화시키기 때문에 씹기뿐 아니라 모음산출에 있어서도 중요하다.

[0011] 네 번째 조음기관으로서의 잇몸 및 굳은입천장 중, 잇몸은 /ㄷ/나 /ㅅ/계열의 말소리들이 조음되는 부위이며, 굳은 입천장은 잇몸 뒤의 단단하고 다소 편편한 부분으로 /ㅈ/계열의 소리들이 조음되는 부위이다.

[0012] 마지막 조음기관으로서의 여린입천장은 움직이는 조음기관으로 분류되는데, 이는 여린입천장의 근육들이 수축함으로써 연인두폐쇄를 이루고 그에 따라 입소리들(oral sounds)을 조음하기 때문이다.

[0013] <조음과정>

[0014] 소리들 중에는 성대를 거친 기류가 성도를 통과하는 동안 구강에서, 더 정확히 말하면 구강 통로의 중앙부에서 어떠한 방해(장애)를 받으면서 생성되는 것과, 이와는 달리 아무런 방해를 받지 않고 생성되는 것이 있다. 보통 전자를 자음(consonant) 후자를 모음(vowel)이라고 한다.

[0016] 1) 자음의 조음

[0017] 자음은 발성되는 방법과 위치에 따라 살펴보아야 하는데 국제문자기호표상에서 각 칸은 조음위치를, 각 줄은 조음방법을 각각 나타내고 있다.

[0018] 우선 조음방법에 따라 분류해 본다면, 기류가 중앙부에서 어떤 종류의 방해를 받아서 조음되는가에 따라서 다막음 소리와 덜막음 소리로 크게 나누어 볼 수 있다. 다막음 소리는 구강에서 기류를 완전히 막았다가 터트리면서 내는 소리이고, 덜막음 소리는 성도의 한 부분을 좁혀서 그 좁아진 통로로 기류를 통과시켜 내는 소리이다.

[0019] 다막음 소리는 다시 비강의 공명을 동반하고 나는 소리와 동반하지 않고 나는 소리로 나눌 수 있다. 성도의 일부를 완전히 막음과 동시에 연구개를 내려 비강 통로를 열고 비강의 공명을 동반하면서 내는 비강 다막음 소리(비강 폐쇄음, nasal stop)들이 전자에 속하며, 연구개를 올려 인두벽에 대고 비강 통로를 차단하여, 기류가 비강으로 통하는 것을 막은 상태로 내는 구강 다막음 소리(구강 폐쇄음, oral stop)들이 후자에 속한다. 구강 다막음 소리는 폐쇄의 길이와 방법에 따라서 폐쇄음(막음소리, stop) 혹은 파열음(터짐소리, plosive), 전동음(떨소리, trill), 탄설음(혹을 설탄음, flap/tap)으로 생각해 볼 수 있다.

[0020] 그리고 덜막음 소리는 마찰음(갈이소리, fricative)과 접근음(approximant)으로 나누는데, 기류의 통로가 혀의 측면에 만들어지는 경우 이를 통틀어 설측음(lateral)이라고 한다.

[0021] 또한 다막음과 덜막음의 조음방법을 복합적으로 사용하는 파찰음(터짐갈이, affricate)이 있으며, 마지막으로 알파벳으로는 "r"이나 "l"로 표현되나 국어의 경우 /ㄹ/로 표현되는 유음(liquid)과 국어에는 없지만 조음기관

을 진동시켜서 소리를 말하는 전동음이 있다.

[0022] 조음위치에 따라 분류해보면, 양순음(bilabial)이란, 두 입술이 그 조음에 관계하는 소리를 지칭하는 것으로, 한국어의 /ㅂ, ㅃ ,ㅍ, ㅁ/등이 이에 속한다. 현대 한국어(표준어)에 존재하는 양순음들은 모두 두 입술을 막아서 내는 소리들이지만, 두 입술의 간격을 좁혀서 그 사이로 기류를 마찰시켜 낼 수도 있으며(양순 마찰음) 두 입술을 떨어서 낼 수도 있다(양순 전동음).

[0023] 순치음(labiodentals)이란 아랫입술과 윗니가 조음에 관계하는 소리를 지칭하는 것으로 한국어에는 존재하지 않는다. 한국어에는 순치음이 없지만, 영어에 있는[f, v]가 바로 이 순치음(순치 마찰음)에 속한다.

[0024] 치음(dental)은 기류의 협착이나 폐쇄가 윗니의 뒷부분에서 일어나는 소리를 말하는데, 이 사이에서 마찰이 이루어지기도 해서 치간음(interdental)이라고도 한다.

[0025] 치경음(alveolar)은 윗잇몸 부근에서 기류의 협착이나 폐쇄가 일어나면서 나는 소리로 한국어의 /ㄷ, ㄸ, ㅌ, ㄴ, ㅆ, ㅅ/등이 이에 속한다. 한국어의 /ㅅ, ㅆ/는 치경 부분에서 기류의 협착이 이루어져 나는 소리로 영어의 /s, z/와 기류의 협착이 이루어지는 장소가 거의 비슷하다.

[0026] 경구개치경음(palatoalveolar)은 후치경음(postalveolar)이라고도 불리는데, 혀끝이나 혓날이 후치경부에 닿아서 나는 소리로 국어에는 존재하지 않지만, 영어나 불어에는 존재한다.

[0027] 치경경구개음(alveolopalatal)은 전경구개음(prepalatal)이라고도 불리는데, 이 소리가 경구개의 앞쪽 즉 치경과 가까운 쪽에서 조음되기 때문이다. 국어의 세 파찰음 /ㅈ, ㅊ, ㅉ/가 이에 속한다.

[0028] 권설음(retroflex)은 혀끝이나 혀의 위 표면이 입천장에 닿거나 접근하여서 조음되는 여타의 설음들과는 달리 혀의 아래 표면이 입천장에 닿거나 접근하여서 조음된다는 점에서 뚜렷한 차이가 있다.

[0029] 경구개음(palatal)은 혓몸이 경구개부에 닿거나 접근하여 조음되는 소리를 말한다.

[0030] 연구개음(velar)은 혓몸이 연구개부에 닿거나 접근하여 조음되는 소리를 말한다. 국어의 폐쇄음/ㄱ, ㅋ, ㄲ/와 비음 /ㅇ/이 이에 속한다.

[0031] 구개수음(uvular)은 혓몸이 연구개의 끝부분인 구개수에 닿거나 접근하여 조음되는 소리를 말한다.

[0032] 인두음(pharyngeal)은 그 조음이 인두강에서 이루어지는 음을 지칭한다.

[0033] 마지막으로 성문음(glottal)은 성대가 조음기관으로 사용되어 조음되는 소리를 지칭하며 우리말에는 음소로서 성문 무성 마찰음 /ㅎ/만이 존재한다.

[0034] 2) 모음의 조음

[0035] 모음의 조음은 혀의 고저와 전후 위치, 그리고 입술의 모양 등 세가지가 가장 중요한 변수로 작용한다.

[0036] 첫 번째 변수로, 혀의 고저에 의하여 모음의 개구도, 즉 입을 벌린 정도가 결정되는데, 입을 적게 벌리고 내는 소리를 폐모음(close vowel), 혹은 고모음(high vowel)이라고 하며, 입을 크게 버리고 내는 소리를 개모음(open vowel), 혹은 저모음(low vowel)이라고 한다. 그리고 고모음과 저모음의 사이에서 나는 소리를 중모음(mid vewel)이라고 하는데, 이중모음은 다시 입을 벌린 정도가 더 작은 중고모음(close-mid vowel), 혹은 반폐모음(half-close vewel)과 입을 벌린 정도가 더 큰 중저모음(open-mid vewel), 혹은 반개모음(half-open vewel)으로 세분할 수 있다.

[0037] 두 번째 변수인 혀의 전후 위치란 사실 혀의 어느 부분이 가장 좁혀졌는가, 다시 말해서 혀의 어느 부분이 입천장과 가장 가까운가를 기준으로 앞뒤를 따지는 것이다. 그 좁아진 부분이 혀의 앞쪽에 있는 모음을 전설모음(front vowel), 뒤쪽에 있는 모음을 후설모음(back vowel)이라고 하며, 그 중간쯤에 있는 모음을 중설모음(central vowel)이라고 한다.

[0038] 마지막으로 모음의 조음에서 중요한 변수가 되는 것은 입술의 모양이다. 조음 시 입술이 동그랗게 모아져 앞으로 튀어나오는 모음을 원순모음(rounded vowel)이라고 하고, 그렇지 않은 모음을 평순모음(unrounded vowel)이라고 한다.

[0040] 발화 장애란 음도, 강도, 음질, 유동성이 성별, 연령, 체구, 사회적 환경, 지리적 위치에 적합하지 않은 것을 이야기 한다. 이는 선천적으로 혹은 후천적으로 만들어 질 수 있으며, 수술을 통해 후두의 일부분인 성대를 늘

이거나 줄여 어느 정도 치료하는 것이 가능하다. 하지만 완벽한 치료는 되지 않으며, 그 효과 또한 정확하다고 할 수 없다.

[0041] 이러한 후두의 기능으로는 삼킴, 기침, 폐색, 호흡, 발성 등의 기능을 가지고 있으며, 이를 위한 다양한 평가 방식(ex. 발화 내역 검사, 발화패턴, 음향학적 검사, 공기역학적 검사...)이 있다. 이러한 평가를 통해 발화 장애의 여부를 어느 정도 판단할 수 있다.

[0042] *44발화 장애의 유형도 다양하며 크게 기능적 발화장애와 기질적 발화장애로 나뉘게 된다. 이러한 유형의 대부분은 후두의 일부분인 성대에 이상이 생기는 경우가 많으며, 이러한 성대가 외부의 환경적 요인으로 인해 부어오름, 찢어짐, 이상 물질의 발생 등에 의해 장애가 오는 경우가 많다.

[0043] 이러한 성대의 기능을 대신하기 위해 인위적으로 진동을 발생시킬 수 있는 진동발생기를 이용할 수 있다. 진동발생기의 방법은 스피커의 원리를 사용할 수 있는데 스피커의 구조를 보면, 자석과 코일로 이루어져 있으며, 이러한 코일에 전류를 흘려주는 상태에서 전류의 방향을 반대로 하면 자석의 극이 반대로 바뀌게 된다. 따라서 자석과 코일의 전류의 방향에 따라 인력과 척력이 작용하게 되고, 이는 코일의 왕복운동을 발생시킨다. 이러한 코일의 왕복운동이 공기를 진동하여 진동을 발생시킨다.

[0044] 다른 방식으로 압전 현상을 이용한 방식이 있는데 압전 결정 유닛이 저주파 신호 전압을 받아서 일그러짐을 발생하고, 그에 의해서 진동판이 진동하여 음향을 발행하도록 만들 수 있다. 따라서 이러한 원리들을 이용한 진동발생기를 이용하여 성대의 기능을 수행하도록 할 수 있다.

[0045] 하지만 이러한 방법의 경우 외부의 위치하여 단순히 성대를 진동시켜 주는 기능에 불과하기 때문에 나타나는 음이 매우 부정확할 뿐 아니라 화자의 말하기 의도를 파악하는 것이 쉽지 않다. 또한 진동 발생기를 가지고 성대에 위치하여 항상 소지해야 되며 말할 때는 한 손을 이용하기 때문에 일상생활에 어려움을 준다. 전술한 발화장애와 이러한 발화 이상에 대해서는 후두나 성대의 일부를 수술하는 등의 치료적 방법을 모색할 수 있으나, 이러한 수술 방법이나 치료가 불가능한 경우가 있어서 완전한 해결책이 되지 못하고 있다.

[0047] 특히 관련 업계에 있어서는 유럽 및 홍콩을 구심점으로 WinEPG, Articulate Instruments Ltd 등의 회사에서 사용 중인 University of Reading, 일본의 Fujimura, Tatsumi가 1973년에 개발하여 Rion 이라는 회사 이름으로 널리 상용화 시킨 The Rion EPG, Flecher이 출원하여 UCLA Phonetics Lab이 연구목적으로 개발하여 사용하는 Kay Palatometer, Schmidt가 개발하여 Complete Speech(Logomertix) 등이 있다.

[0048] 그러나 종래의 기술들은 수동적 조음기관을 기반으로 발화하는 것에 한계가 있으며, 능동적 조음기관 자체인 구강설을 이용하거나, 구강설과 다른 조음기관과의 연계성에 의한 실제 조음 방식에 따른 발화를 구현하는 데 명확한 한계가 있었다.

[0049] 기존에 상태 변화나 움직임을 파악하기 위한 다양한 센서가 개발되어 있으며, 센서를 바탕으로 압력, 온도, 거리, 마찰 등의 변화를 파악하는 것이 가능하다.

[0050] 더불어, 발화 및 표정 동기화(Lip Sync)는, 대상 내지 객체의 아이덴티티를 결정하는 가장 중요한 요소인 말하는 음성 및 조음을 포함하는 발화, 표정 등을 캐릭터, 로봇, 다양한 전자제품, 자율주행 운송수단 등에 복제 적용하여 개인의 아이덴티티를 결정하고 확장하는 핵심 수단이다. 특히, 고품질의 Lip Sync 에니메이션을 만들기 위해서 전문 에니메이션 팀이 작업하므로, 높은 비용 및 많은 시간을 필요로 하고 대량의 작업에 어려움이 존재한다. 종래의 일반적인 기술은 단순한 입술 모양 라이브러리를 사용하여 저급한 에니메이션을 생성하는 수준에 국한되었다. Pixar나 Disney와 같은 해외의 에니메이션 콘텐츠 제작사들은 Lip Sync를 통한 실감나는 캐릭터 에니메이션을 생성하는데 많은 비용과 시간을 투입하는 실정이다.

## *발명의 내용*

### *해결하려는 과제*

[0052] 본 발명은 전술한 문제점을 해결하기 위하여 제안된 것으로, 본 발명의 목적은, 사용자의 발화 의도에 따른 사용자의 조음 방식을 구강설을 포함한 두경부의 센서를 통해 파악하고, 이를 청각, 시각, 촉각의 형태로 나타내어 양호한 품질의 음성 형성, 즉 발성이 표출될 수 있는 발화 보완용 기기 및 그 방법을 제공하는 것이다.

[0053] 본 발명의 다른 목적은, 발화에 있어서 정상적인 기능을 수행하지 못하고 교정이나 치료가 불가능한 경우에 양질의 적절한 발화를 구현하는 것이다.

[0054] 본 발명의 또 다른 목적은, 발화를 위한 조음 의도에 따라 사용자가 원하는 정도의 정확한 발화가 외부로 표출될 수 있는 내/외부에 위치한 발화 보완용 기기 및 그 제어 방법을 제공하는 것이다.

## *과제의 해결 수단*

[0056] 상기 목적을 달성하기 위하여, 본 발명은, 화자의 두경부의 일면에 인접하여 조음기관의 물리특성을 측정하는 센서부;

[0057] 상기 센서부의 위치와 상기 조음기관의 물리특성을 기반으로 화자의 발화 특징을 파악하는 데이터해석부;

[0058] 2상기 센서부의 위치와 상기 발화특징을 언어데이터로 변환하는 데이터변환부;

[0059] 상기 데이터변환부(300)의 상기 언어데이터를 음성정보로 생성하는 텍스트정보부;

[0060] 상기 텍스트정보부와 연동되어, 언어데이터 기반의 상기 음성정보를 외부로 표현하는 데이터표현부;

[0061] 상기 센서부는, 구강설에 대응되는 구강설 센서를 포함하는 발화 구현 시스템을 제공한다.

[0063] 그리고, 상기 구강설 센서는, 상기 구강설의 일측면에 고착되거나, 상기 구강설의 표면을 감싸거나, 상기 구강설 내부에 삽입되고, 발화에 따른 상기 구강설의 x축, y축, z축 방향 기반의 시간에 따른 벡터량의 변화량을 파악하여, 상기 구강설의 저고도, 전후설성, 굴곡도, 신전도, 회전도, 긴장도, 수축도, 이완도, 진동도 중 적어도 하나의 물리특성을 파악할 수 있다.

[0064] 또한, 상기 구강설 센서는, 상기 구강설의 일측면에 고착되거나, 상기 구강설의 표면을 감싸거나, 상기 구강설 내부에 삽입되고, 발화에 따른 상기 구강설의 x축, y축, z축 방향 기반의 단위 시간 당 회전하는 각도의 변화량을 파악하여, 상기 구강설을 포함한 상기 조음기관의 물리 특성을 파악할 수 있다.

[0065] 그리고, 상기 구강설 센서는, 상기 구강설의 일측면에 고착되거나, 상기 구강설의 표면을 감싸고, 발화에 따른 상기 구강설의 수축 및 이완으로 발생하는 물리력에 따라 결정 구조의 변화에 기인하는 편극에 대응되는 전기신호가 발생하는 압전소자를 통해 상기 구강설의 굽힘도를 파악하여, 상기 구강설의 저고도, 전후설성, 굴곡도, 신전도, 회전도, 긴장도, 수축도, 이완도, 진동도 중 적어도 하나의 물리특성을 파악할 수 있다.

[0066] 또한, 상기 센서부는, 상기 구강설이 상기 두경부 내외의 다른 조음기관과의 상호작용에 기인하는 접근 및 접촉에 대응되는 마찰전기(Tribo Electric Generator)에 따른 파열도, 마찰도, 공명도, 접근도 중 적어도 하나의 물리특성을 파악하는 마찰대전소자를 포함할 수 있다.

[0067] 그리고, 상기 데이터해석부는, 상기 센서부에서 측정되는 상기 구강설과 다른 조음기관과의 물리특성을 통해 상기 화자가 발화하는 자모음, 어휘 단위 강세 (Lexical Stress), 문장 단위 강세(Tonic stress) 중 적어도 하나의 발화 특징을 파악할 수 있다.

[0068] 또한, 상기 데이터해석부는, 상기 센서부에 의해 측정되는 상기 조음기관의 물리특성에 의한 발화 특징을 파악함에 있어서, 2진수 내지 실수를 포함하는 수치로 구성된 표준 발화 특징 행렬을 기반으로 상기 화자의 발음과 강세의, 유사근접도, 발화 의도 중 적어도 하나의 발화 특징을 측정할 수 있다.

[0069] 그리고, 상기 데이터해석부는, 상기 센서부에 의해 측정되는 상기 조음기관의 물리특성을 발화 특징을 파악함에 있어서, 상기 조음기관의 물리특성을 각 자모음 단위의 패턴으로 인식하는 단계, 상기 자모음 단위의 패턴의 특징을 추출하고, 추출된 상기 자모음 단위의 패턴의 특징을 유사도에 따라 분류하는 단계, 분류된 상기 자모음 단위의 패턴의 특징을 재조합하는 단계, 상기 조음기관의 물리특성을 상기 발화 특징으로 해석하는 단계에 따라 상기 발화 특징을 파악할 수 있다.

[0070] 또한, 상기 데이터해석부는, 상기 센서부에 의해서 측정되는 상기 조음기관의 물리특성에 의해, 자모음의 동화(Assimilation), 이화(Dissimilation), 탈락(Elision), 첨가(Attachment), 강세(Stress)와, 약화(Reduction)로 야기되는 기식음화 (Asperation), 음절성자음화(Syllabic cosonant), 탄설음화(Flapping), 경음화(Tensification), 순음화(Labilalization), 연구개음화(Velarization), 치음화(Dentalizatiom), 구개음화

(Palatalization), 비음화(Nasalization), 강세변화(Stress Shift), 장음화(Lengthening) 중 적어도 하나의 이차조음현상인 발화 변이를 측정할 수 있다.

[0071] 그리고, 상기 구강설 센서는, 센서 작동을 위한 회로부, 상기 회로부를 감싸는 캡슐부, 상기 구강설 일면에 부착되는 접착부를 포함할 수 있다.

[0072] 또한, 상기 구강설 센서는, 박막 회로를 가진 필름 형태로서 상기 구강설에 인접하여 작동할 수 있다.

[0073] 그리고, 상기 센서부는, 두경부 근육의 신경신호 측정의 기준 전위를 생성하는 적어도 하나의 레퍼런스 센서와, 상기 두경부 근육의 신경신호를 측정하는 적어도 하나의 양극센서 및 적어도 하나의 음극센서로 구성된 안면부 센서를 포함할 수 있다.

[0074] 또한, 상기 데이터해석부는, 상기 안면부 센서에 기반하여 상기 센서부의 위치를 획득함에 있어서, 상기 레퍼런스 센서를 기준으로 하여 상기 적어도 하나의 양극센서 및 상기 적어도 하나의 음극센서의 전위차를 파악하여 상기 안면부 센서의 위치를 파악할 수 있다.

[0075] 그리고, 상기 데이터해석부는, 상기 안면부 센서에 기반하여 상기 화자의 발화 특징을 획득함에 있어서, 상기 레퍼런스 센서를 기준으로 하여 상기 적어도 하나의 양극센서 및 상기 적어도 하나의 음극센서의 전위차를 파악하여 상기 화자의 두경부에서 발생하는 상기 조음기관의 물리 특성에 의한 발화 특징을 파악할 수 있다.

[0076] 또한, 상기 센서부는, 상기 화자의 두경부 중 성대에 인접하여 성대의 근전도 내지 떨림을 파악하여, 상기 화자의 발화 시작, 발화 정지, 발화 종료 중 적어도 하나의 발화 내역 정보를 파악하는 성대 센서를 포함할 수 있다.

[0077] 그리고, 상기 센서부는, 치아의 일면에 인접하여 상기 구강설 및 아랫 입술의 접촉에 기인하여 발생하는 전기적 용량 변화에 따른 신호발생 위치를 파악하는 치아센서를 포함할 수 있다.

[0078] 또한, 상기 데이터해석부는, 상기 화자의 두경부 일면에 인접한 음성 취득 센서를 통해 발화에 따른 상기 화자의 음성을 취득할 수 있다.

[0079] 그리고, 상기 센서부는, 상기 화자의 두경부 조음기관의 위치 변화 정보, 상기 화자의 두경부 표정 변화 정보, 상기 화자의 발화 의도에 따라 움직이는 두경부, 흉곽부, 상지부, 하지부의 비언어적 표현 중 적어도 하나를 파악하기 위해 상기 화자의 두경부를 촬상하는 촬상센서를 포함할 수 있다.

[0080] 또한, 상기 발화 구현 시스템은, 상기 센서부의 구강설 센서, 안면 센서, 음성취득 센서, 성대 센서, 치아 센서, 촬상센서 중 적어도 하나에 전원을 공급하는 전원부를 더 포함할 수 있다.

[0081] 그리고, 상기 발화 구현 시스템은, 상기 데이터해석부 및 상기 데이터베이스부가 외부에 위치하여 작동할 경우, 연동되어 통신할 수 있는 유선 또는 무선의 통신부를 더 포함할 수 있다.

[0082] 또한, 상기 데이터해석부는, 상기 센서부의 위치, 상기 화자의 발화 특징, 상기 화자의 음성에 대응하는 적어도 하나의 언어 데이터 색인을 포함하는 데이터베이스부와 연동될 수 있다.

[0083] 그리고, 상기 데이터베이스부는, 발화의 진행 시간, 발화에 따른 주파수, 발화의 진폭, 발화에 따른 두경부 근육의 근전도, 발화에 따른 두경부 근육의 위치 변화, 구강설의 굽힘 및 회전에 따른 위치 변화 중 적어도 하나의 정보를 기반으로, 자모음의 음소단위 색인, 음절단위 색인, 단어단위 색인, 구절단위 색인, 문장단위 색인, 연속발화단위 색인, 발음의 고저 색인 중 적어도 하나의 언어 데이터 색인을 구성할 수 있다.

[0084] 또한, 상기 데이터표현부는, 상기 데이터베이스부의 언어 데이터 색인과 연동되어, 상기 화자의 발화 특징을 자모음의 음소(Phoneme)단위, 적어도 하나의 단어단위, 적어도 하나의 구절단위(Citation Forms), 적어도 하나의 문장단위, 연속발화단위(Consecutive Speech) 중 적어도 하나의 발화 표현을 나타낼 수 있다.

[0085] 그리고, 상기 데이터표현부에 의해 나타나는 발화 표현은, 문자 기호, 그림, 특수기호, 숫자 중 적어도 하나로 시각화되거나, 소리 형태로 청각화되어, 상기 화자와 청자에게 제공될 수 있다.

[0086] 또한, 상기 데이터표현부에 의해 나타나는 발화 표현은, 진동, 스누징, 태핑, 압박, 이완 중 적어도 하나의 촉각적 방법으로 상기 화자와 청자에게 제공될 수 있다.

[0087]

## *발명의 효과*

[0088] 본 발명의 두경부 조음기관 물리 특성 기반의 발화 구현 시스템은, 화자의 구강설을 중심으로 한 두경부 조음기관 활용에 발화 의도를 파악하고 청각, 시각, 촉각의 형태로 나타내어 양호한 품질의 발화 형성, 즉 발성이 표출될 수 있는 효과를 갖는다. 더불어, 화자 스스로 제 1차 발화 내용과 적어도 제 2차 발화 내용간 비교를 통한 피드백을 통해 스스로 발화 개선을 도모할 수 있다.

[0089] 본 발명에서는 말하기 의도를 파악하기 위해 구강설을 포함한 두경부 내외의 조음기관을 이용하게 되며, 이러한 움직임의 예를 들면, 구강설의 독립적 물리 특성이나 수동적 조음기관과 입술, 성문, 성대, 인두, 후두개 중 하나이상으로 구성된 능동적 조음기관 중 하나이상의 조음기관과의 상호 작용에 의해 생기는 폐쇄도, 파열도, 마찰도, 공명도. 접근도 중 하나이상의 특성을 파악해야 하며, 이러한 특성을 파악하기 위하여 방위각, 앙각, 회전각, 압력, 마찰, 거리, 온도, 소리 등을 파악할 수 있는 다양한 센서를 이용하게 된다.

[0090] 기존에 제안된 인공 성대의 경우 외부에서 진동을 통해 소리를 내는 정도로, 한 손의 움직임이 부자연스럽고 발화의 질이 매우 낮다는 단점이 있었고, 인공 구개의 경우에는 수동적 조음기관인 경구개에 의존한다는 단점이 있었다.

[0091] 더불어, 음성학적으로는 인공구개를 활용하여 화자의 발화를 측정하고자 하는 조음음성학(Articulatory Phonetics)이 지금껏 주류로서 인정 되어 왔으나, 발화 측정에 있어서 특정 자모음의 조음에 따른 발화의 이산적인 발화 유무만 파악할 수 있었다. 하지만, 이러한 조음음성학의 주장은 인간의 발화가 이산적인 특징을 가지는 것이 아니라. 각 음소(Phoneme), 특히 모음에 있어서, 각 모음들이 분절되어 존재할 수도 없고 분절되어 발음될 수도 없는 연속적인 체계임을 주장하는 음향음성학(Acoustic Phonetics)에 의해 학계의 의문을 불러일으키고 있다. 자세히 말하자면, 인간의 발화는 조음을 하여 "발화를 한다." 또는 "발화 하지 못했다."와 같이 이산적으로 나누어 질 수 있는 것이 아니라, 유사정도에 따른 비례적, 비율적, 단계적 특성을 지닌다는 것이다.

[0092] 그렇기에 음향음성학은 화자의 발화에 따른 언어음 자체의 물리적 속성을 수치화(Scaling)하여, 발화의 유사도 또는 근접도를 파악함으로서, 종래의 조음음성학이 구현할 수 없었던 발음의 비례적, 비율적, 단계적 유사정도에 따른 발화 측정에 대해 가능성을 열어두고 있다.

[0093] 이러한 종래 관련 기술동향과 관련 학문적 배경을 참고하였을 때, 본 발명은 조음음성학의 기반을 두고서, 음향음성학이 추구하고자 하는 조음의 수치화(Scaling)에 따른 보다 정확한 발화 의도를 파악하고 구현할 수 있는 매우 획기적인 장점을 가지고 있다고 할 수 있다.

[0094] 자세히 말하자면, 본 발명에서는 화자의 조음기관 작용에 의해 발생하는 조음도를 수치화(Scaling)하여 발화 의도를 청각, 시각, 촉각의 형태로 직관적으로 제시하기 때문에 의사소통의 질 및 생활 편의도가 매우 탁월해질 것으로 기대된다는 것이다.

[0095] 더불어, 화자의 발화에 따른 발화 의도를 문자로서 표현할 경우, Speech to Text로 응용되어, Silent Speech(침묵 대화)가 가능해진다. 이를 통해, 청각 장애인과 의사소통을 할 시에, 화자는 발화를 하고 청자인 청각 장애인은 이를 시각적 자료로 인지하기에 소통상의 어려움이 없어진다. 더불어, 의사전달에 있어서 소음에 영향을 받는 대중 교통, 공공 시설, 군사 시설 및 작전, 수중 활동 등에 활용 될 수 있다.

[0096] 더불어, 발화에 따라 변화하는 화자의 두경부 조음기관의 외상을 촬상함으로서, 발화와 발화에 따른 조음기관의 외적 변화의 연관성을 파악해, 언어학적 방면과 보완 대체 의사소통 방면, 휴머노이드의 안면 구현 방면으로 활용될 수 있다.

[0097] 특히, 애니메이션 및 영화 제작 업계에서는 현재까지 애니메이션 캐릭터를 포함한 영상 객체의 발화와 표정의 일치를 달성하는 데 어려움을 겪고 있다. 가장 문제가 되는 부분은 바로 조음 기관의 작동, 발화의 영역이다. 인간의 복잡한 조음기관의 물리적 특성을 제대로 반영하지 못해, 월트 디즈니 및 픽사 등의 거대 기업들조차도 캐릭터가 입만 뻥긋거리는 정도의 개발 수준에 그치며, 대사와 발화 및 표정 간의 낮은 일치도를 보여준다. 이러한 문제를 해결하기 위해 영상 제작팀에게 높은 비용을 지불하고서 전신에 성우 혹은 모사 배우에게 특징점을 부착한다. 하지만, 이러한 방법은 영상 객체의 발화나 표정에 대한 근본적인 부분을 해결하지 못하고, 신체 전반에 걸치는 거시적인 움직임을 표현하는 것을 주 목적으로 하는 한계가 있다. 그러나, 본 발명은 실제 인간 화자의 조음기관 물리특성을 측정하여 이를 영상 객체의 두경부에 맵핑함으로서 영상 객체의 발화나 표정을 실제 인간 화자와 비슷하게 구현할 수 있도록 한다.

[0098] 특히, 본 발명은 화자 조음정보를 로봇 객체의 두경부의 움직임을 구현하는 액추에이터에 전달하여 매칭시킴으로서, 인간 화자와 유사한 로봇의 조음, 발화, 표정을 포함하는 두경부 움직임을 재현하는 것으로, 이는 일본의

모리 마사히로(Mori Masahiro)가 주장한 휴머노이드 로봇이 인간에게 유발하는 만성적인 인지적 부조화인 "불쾌한 골짜기(Uncanny Valley)"를 극복할 수 있는 효과가 있다. 더불어, 휴머노이드 및 그 외 일반 로봇의 인간 친화적인 조음 구현이 가능해짐에 따라, 로봇 및 안드로이드의 인간 역할 대체가 가능해지고, 더 나아가 인간-로봇의 대화가 달성됨으로서 고령화에 따른 노인 인구 증대로 대두되는 노인 사회의 고립 현상 및 우울증과 같은 정신/심리적 질환을 예방할 수 있는 효과가 있다.

## 도면의 간단한 설명

[0100] 도 1은 본 발명의 제1실시예에 따른 발화 구현 시스템의 센서부를 도시한 도면.

도 2는 본 발명의 제1실시예에 따른 발화 구현 시스템의 센서부의 위치를 도시한 도면.

도 3은 본 발명의 제1실시예에 따른 발화 구현 시스템을 도시한 도면.

도 4는 본 발명의 제1실시예에 따른 발화 구현 시스템에서 활용되는 구강설의 위치적 명칭을 도시한 도면.

도 5는 본 발명의 제1실시예에 따른 발화 구현 시스템에서 활용되는 모음 발화를 위한 구강설의 작용을 도시한 도면.

도 6 내지 도 10은 각각 본 발명의 제1실시예에 따른 발화 구현 시스템의 다양한 구강설 센서를 도시한 도면.

도 11 및 도 12는 각각 본 발명의 제1실시예에 따른 발화 구현 시스템의 구강설 센서의 부착상태를 도시한 단면도 및 사시도.

도 13은 본 발명의 제1실시예에 따른 발화 구현 시스템의 구강설 센서의 회로부를 도시한 도면.

도 14는 본 발명의 제1실시예에 따른 발화 구현 시스템의 구강설 센서의 다양한 활용상태를 도시한 도면.

도 15는 본 발명의 제2실시예에 따른 발화 구현 시스템을 도시한 도면.

도 16은 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터해석부가 발화 특징을 파악하는 원리를 도시한 도면.

도 17은 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터해석부가 측정된 조음기관의 물리 특성을 발화 특징으로 파악하는 원리를 도시한 도면.

도 18은 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터해석부가 활용하는 모음에 관한 표준 발화 특징 행렬을 도시한 도면.

도 19는 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터해석부가 활용하는 자음에 관한 표준 발화 특징 행렬을 도시한 도면.

도 20은 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터해석부가 조음기관의 물리 특성을 발화 특징으로 파악하기 위하여 활용하는 알고리즘 프로세스를 도시한 도면.

도 21은 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터해석부가 조음기관의 물리 특성을 발화 특징으로 파악하기 위하여 활용하는 알고리즘 프로세스를 상세히 도시한 도면.

도 22는 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터해석부가 조음기관의 물리 특성을 발화 특징으로 파악하기 위하여 활용하는 알고리즘 프로세스의 원리를 상세히 도시한 도면.

도 23은 본 발명의 제2실시예에 따른 발화 구현 시스템의 구강설 센서가 발화된 특정 모음을 발화 특징으로 파악하는 알고리즘 프로세스를 도시한 도면.

도 24는 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터해석부가 Alveolar Stop을 활용하는 경우를 도시한 도면.

도 25는 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터해석부가 Bilabial Stop을 활용하는 경우를 도시한 도면.

도 26은 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터해석부가 Voiced Bilabial Stop을 활용한 실험

결과를 도시한 도면.

도 27 및 도 28은 각각 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터해석부가 Voiced Labiodental Fricative를 활용하는 경우를 도시한 도면.

도 29는 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터해석부 및 데이터베이스의 연동을 도시한 도면.

도 30은 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터해석부가 특정 단어로 파악하는 경우를 도시한 도면.

도 31은 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터베이스부를 도시한 도면.

도 32는 본 발명의 제3실시예에 따른 발화 구현 시스템을 도시한 도면.

도 33 및 도 34는 각각 본 발명의 제3실시예에 따른 발화 구현 시스템의 데이터베이스부의 실제 형태를 도시한 도면.

도 35는 본 발명의 제4실시예에 따른 발화 구현 시스템을 도시한 도면.

도 36은 본 발명의 제4실시예에 따른 발화 구현 시스템의 센서부, 데이터해석부, 데이터표현부 및 데이터베이스부의 연동을 도시한 도면.

도 37 내지 도 41은 각각 본 발명의 제4실시예에 따른 발화 구현 시스템의 데이터표현부가 언어데이터를 표현하는 수단을 도시한 도면.

도 42는 본 발명의 제4실시예에 따른 발화 구현 시스템의 데이터표현부가 언어데이터를 시각적 및 청각적으로 표현하는 경우를 도시한 도면.

도 43은 본 발명의 제4실시예에 따른 발화 구현 시스템의 데이터표현부가 언어데이터를 시각적으로 표현하는 경우를 도시한 도면.

도 44는 본 발명의 제4실시예에 따른 발화 구현 시스템의 데이터표현부가 언어데이터를 시각적으로 표현하는 경우를 도시한 도면.

도 45는 본 발명의 제4실시예에 따른 발화 구현 시스템의 데이터표현부가 언어데이터를 연속 발화 단위로 표현하는 경우를 도시한 도면.

도 46은 본 발명의 제4실시예에 따른 발화 구현 시스템이 활용하는 Confusion Matrix를 도시한 도면.

도 47은 본 발명의 제4실시예에 따른 발화 구현 시스템이 활용하는 Confusion Matrix를 백분율로 도시한 도면.

도 48은 본 발명의 제4실시예에 따른 발화 구현 시스템이 화면을 통해 화자로 하여금 언어 교정 및 지도를 돕는 경우를 도시한 도면.

도 49는 본 발명의 제4실시예에 따른 발화 구현 시스템이 두경부 조음기관의 외상을 촬상하고 파악하는 경우를 도시한 도면.

도 50은 본 발명의 제4실시예에 따른 발화 구현 시스템이 표준 발화 특징 행렬을 통해 상호 정보들을 결합시키는 경우를 도시한 도면.

도 51은 본 발명의 촬상 센서가 촬상하는 화자의 비언어적 표현을 도시한 도면.

도 52는 본 발명의 텍스정보부가 음성정보를 생성하는 경우를 도시한 도면.

도 53은 본 발명의 텍스트정보부가 활용하는 음소 단위 색인의 Frequency Table일부 예시를 도시한 도면

도 54는 본 발명의 텍스트정보부가 활용하는 음소 단위 색인의 Frequency Table일부 예시를 도시한 또 다른 도면

## *발명을 실시하기 위한 구체적인 내용*

[0101] 이하, 첨부된 도면을 참조하면서 본 발명의 일실시예에 따른 발화 개선 가이드 및 피드백 을 위한 두경부 조음기관의 물리 특성을 이용한 발화 구현 시스템에 대해 상세히 설명하기로 한다.

[0102] 본 발명의 하기의 실시예는 본 발명을 구체화하기 위한 것일 뿐 본 발명의 권리범위를 제한하거나 이를 한정하

는 것이 아님은 물론이다. 본 발명의 상세한 설명 및 실시예로부터 본 발명이 속하는 기술 분야의 전문가가 용이하게 유추할 수 있는 것은 본 발명의 권리범위에 속하는 것으로 해석된다.

[0103] 본 발명을 실시하기 위한 내용을 도 1부터 도 54을 기반으로 상세히 설명하고자 한다.

[0104] 도 1은 본 발명의 제1실시예에 따른 발화 구현 시스템의 센서부를 도시한 도면이고, 도 2는 본 발명의 제1실시예에 따른발화 개선 가이딩 및 피드백 시스템의 센서부의 위치를 도시한 도면이고, 도 3은 본 발명의 제1실시예에 따른 발화 구현 시스템을 도시한 도면이다.

[0105] 도 1, 도 2 및 도 3에 도시한 바와 같이, 본 발명의 제1실시예에 따른 발화 구현 시스템에서, 센서부(100)는 두경부에 위치하는 구강설 센서(110), 안면 센서(120), 음성 취득 센서(130), 성대 센서(140), 치아센서(150)로 구성된다.

[0106] 더욱 자세히는 두경부에 위치하는 구강설 센서(110), 안면 센서(120), 음성 취득 센서(130), 성대 센서(140), 치아센서(150)는, 각 센서들이 위치하는 센서부의 위치(210), 화자(10)의 발화에 따른 발화 특징(220), 화자의 음성(230), 발화 내역 정보(240), 발화 변이(250)에 대한 데이터를 제공한다.

[0107] 데이터해석부(200)는 이러한 데이터들을 취득하고, 데이터변환부(300)는 이러한 데이터를 언어데이터(310)로 처리한다.

[0108] 도 4는 본 발명의 제1실시예에 따른 발화 구현 시스템에서 활용되는 구강설의 위치적 명칭을 도시한 도면이고, 도 5는 본 발명의 제1실시예에 따른 발화 구현 시스템에서 활용되는 모음 발화를 위한 구강설의 작용을 도시한 도면이다.

[0109] 도 4 및 도 5에 도시한 바와 같이, 구강설 센서(110)의 경우, 구강설(12)의 일측면에 고착되거나, 그 표면을 감싸거나, 그 내부에 삽입되며, 저고도, 전후설성, 굴곡도, 신전도, 회전도, 긴장도, 수축도, 이완도, 진동도 중 하나이상의 구강설 자체의 독립 물리 특성을 파악한다.

[0110] 도 6 내지 도 10은 각각 본 발명의 제1실시예에 따른 발화 구현 시스템의 다양한 구강설 센서를 도시한 도면이다.

[0111] 도 6 및 도 7에 도시한 바와 같이, 구강설(12) 자체의 독립 물리 특성을 파악함에 있어서, 구강설 센서(110)는 x축, y축, z축 방향의 가속도 및 단위 시간 당 회전하는 각도의 변화량(각속도) 중 적어도 하나를 파악함으로써, 구강설(12)을 포함한 다른 조음기관의 물리 특성에 의한 발화 특징(220)을 파악한다.

[0112] 도 8에 도시한 바와 같이, 구강설 센서(110)는 발화에 따른 구강설(12)의 수축 내지 이완으로 발생하는 물리력에 따라 결정 구조(111)의 변화에 의해 편극이 발생하여 전기신호가 발생하는 압전소자(112)를 통해 구강설(12)의 굽힘도를 파악함으로써, 구강설(12)을 포함한 조음기관의 물리 특성에 의한 발화 특징(220)을 파악할 수 있다.

[0113] 도 9에 도시한 바와 같이, 구강설 센서(110)는 구강설(12)이 두경부 내외의 다른 조음기관과의 상호작용에 의해 생기는 접근 및 접촉에 의해 발생하는 마찰전기(Tribo Electric Generator)에 따른 연계 물리 특성을 파악하기 위해 마찰대전소자(113)를 사용하여 화자의 발화 특징(220)을 파악한다.

[0114] 도 10에 도시한 바와 같이, 통합된 구강설 센서(110)는 x축, y축, z축 방향의 가속도 및 각속도, 압전에 의한 전기신호, 접촉에 의한 마찰전기를 이용하여 구강설(12)을 포함하는 조음기관의 물리 특성에 의한 발화 특징(220)을 파악한다.

[0115] 도 11 및 도 12는 각각 본 발명의 제1실시예에 따른 발화 구현 시스템의 구강설 센서의 부착상태를 도시한 단면도 및 사시도이다.

[0116] 도 11 및 도 12에 도시한 바와 같이, 구강설 센서(110)는 복합 박막 회로로 구성되어 단일한 필름 형태로 구현될 수 있다. 이때, 구강설 센서(110)는 센서부(100)를 작동하기 위한 회로부(114), 회로부(114)를 감싸는 캡슐부(115), 구강설 센서(110)를 구강설(12)의 일면에 고착시키는 접착부(116)로 구성된다.

[0117] 도 6 내지 도 9에 도시한 바와 같이, 구강설 센서(110)는 각 센서의 특징에 따라 두경부 내외의 다른 조음기관과의 인접 내지 응접에 의해 생기는 파열도, 마찰도, 공명도, 접근도 중 하나 이상의 물리 특성을 파악할 수 있다.

[0118] 도 13은 본 발명의 제1실시예에 따른 발화 구현 시스템의 구강설 센서의 회로부를 도시한 도면이다.

[0119] 도 13에 도시한 바와 같이, 구강설 센서(110)의 회로부(114)는 통신칩, 센싱회로, MCU로 구성된다.

[0120] 도 14는 본 발명의 제1실시예에 따른 발화 구현 시스템의 구강설 센서의 다양한 활용상태를 도시한 도면이다.

[0121] 도 14에 도시한 바와 같이, 구강설 센서(110)는 화자의 다양한 자모음의 발화에 따른 구강설(12)의 상태를 파악하여, 자모음 발화에 따른 발화 특징(220)을 파악할 수 있다.

[0122] 예를 들어, 구강설 센서(110)는 Bilabial Sound (양순음), Alveolar Sound (치경음), Palatal Sound (구개음)에 따른 발화 특징(220)을 파악할 수 있다.

[0123] 도 15는 본 발명의 제2실시예에 따른 발화 구현 시스템을 도시한 도면이다.

[0124] 도 15에 도시한 바와 같이, 본 발명의 제2실시예에 따른 발화 구현 시스템에서, 구강설 센서(110), 안면 센서(120), 음성취득센서(130), 성대센서(140), 치아센서(150)로 이루어진 두경부 조음기관 인근의 센서부(100)는, 두경부 조음기관에서 센서가 위치한 센서부의 위치(210), 발화에 따른 발화특징(220), 발화에 따른 화자의 음성(230), 발화의 시작, 발화 정지, 발화 종료를 포함하는 발화 내역 정보(240)을 파악한다.

[0125] 이때, 발화 특징(220)은, 인간이 발화할 때 발생하는 폐쇄파열음화, 마찰음화, 파찰음화도, 비음화, 유음화, 활음화, 치찰음화, 유무성음화, 성문음화 중 하나이상의 기본적인 물리적 발화 특징을 의미한다. 또한, 화자의 음성(230)은, 발화 특징으로 인해 함께 수반되는 청각적인 발화 특징이다. 그리고, 발화 내역 정보(240)는, 성대센서(140)를 통한 것으로, 성대의 근전도 내지 떨림으로 그 정보를 파악한다.

[0126] 데이터해석부(200)는, 구강설 센서(110), 안면 센서(120), 음성취득센서(130), 성대센서(140), 치아센서(150)로 이루어진 두경부 조음기관 인근의 센서부(100)가 측정한 화자의 조음기관 물리특성에서 화자의 성별, 인종, 나이, 모국어에 따라 발생하는 발화 변이(250)를 파악한다.

[0127] 이때, 발화 변이(250)는 자모음의 동화(Assimilation), 이화(Dissimilation), 탈락(Elision), 첨가(Attachment), 강세(Stress), 약화(Reduction)로 야기되는 기식음화 (Asperation), 음절성자음화(Syllabic cosonant), 탄설음화(Flapping), 경음화(Tensification), 순음화(Labilalization), 연구개음화(Velarization), 치음화(Dentalizatiom), 구개음화 (Palatalization), 비음화(Nasalization), 강세변화(Stress Shift), 장음화(Lengthening) 중 하나이상의 이차조음현상을 포함한다.

[0128] 데이터변환부(300)는, 두경부 조음기관 센서들(110, 120, 130, 140, 150)에 의해 측정된 센서부의 위치(210), 발화에 따른 발화특징(220), 발화에 따른 화자의 음성(230), 발화 내역 정보(240), 발화 변이(250)를 언어데이터(310)로 인지하여 처리한다.

[0129] 이때, 데이터변환부(300)가 언어데이터(310)를 인지하여 처리함에 있어서, 데이터해석부(200)는 데이터베이스부(350)와 연동된다.

[0130] 데이터베이스부(350)는, 자모음의 음소단위(361), 색인 음절 단위 색인(362), 단어단위 색인(363), 구절단위 색인(364), 문장단위 색인(365), 연속 발화 색인(366), 발음의 고저 색인(367)을 포함하는 언어 데이터 색인(360)을 가지고 있다. 이러한 언어 데이터 색인(360)을 통해, 데이터해석부(200)는 센서부(100)에서 취득된 다양한 발화 관련 정보들을 언어데이터로 처리할 수 있게 된다.

[0131] 도 16은 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터해석부가 발화 특징을 파악하는 원리를 도시한 도면이고, 도 17은 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터해석부가 측정된 조음기관의 물리 특성을 발화 특징으로 파악하는 원리를 도시한 도면이고, 도 18은 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터해석부가 활용하는 모음에 관한 표준 발화 특징 행렬을 도시한 도면이고, 도 19는 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터해석부가 활용하는 자음에 관한 표준 발화 특징 행렬을 도시한 도면이다.

[0132] 도 16, 도 17, 도 18 및 도 19에 도시한 바와 같이, 데이터해석부(200)는, 구강설 센서(110)를 포함한 센서부(100)로부터 측정된 조음기관의 물리 특성을 먼저 획득한다. 구강설 센서(110)로 인해 조음기관 물리 특성이 획득된 경우, 구강설 센서(110)는 조음기관 물리 특성을 센싱하면서 센싱된 물리 특성의 행렬값을 만든다.

[0133] 이후, 데이터해석부(200)는, 자모음의 표준 발화 특징 행렬(205)에서 이러한 물리 특성의 행렬값에 대응하는 자모음의 발화특징(220)을 파악한다. 이때 자모음의 표준 발화 특징 행렬(205)는 그 내부의 값들이 자모음 발화 기호, 2진수 내지 실수 중 하나이상으로 존재할 수 있다.

[0134] 도 20은 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터해석부가 조음기관의 물리 특성을 발화 특징

으로 파악하기 위하여 활용하는 알고리즘 프로세스를 도시한 도면이다.

[0135] 도 20에 도시한 바와 같이, 데이터해석부(200)가 활용하는 알고리즘 프로세스는, 센서부(100)에 의해 측정된 조음기관의 물리 특성을 파악함에 있어서, 조음기관의 물리 특성을 취득하는 단계, 취득된 조음기관의 물리 특성이 가지고 있는 각 자모음 단위의 패턴을 파악하는 단계, 각 자모음 패턴으로부터 고유한 특징을 추출하는 단계, 추출된 특징들을 분류하는 단계, 분류된 패턴의 특징들을 재조합하는 단계로 이루어지고, 이를 통해 최종적으로 특정 발화 특징으로 파악한다.

[0136] 도 21은 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터해석부가 조음기관의 물리 특성을 발화 특징으로 파악하기 위하여 활용하는 알고리즘 프로세스를 상세히 도시한 도면이고, 도 22는 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터해석부가 조음기관의 물리 특성을 발화 특징으로 파악하기 위하여 활용하는 알고리즘 프로세스의 원리를 상세히 도시한 도면이고, 도 23은 본 발명의 제2실시예에 따른 발화 구현 시스템의 구강설 센서가 발화된 특정 모음을 발화 특징으로 파악하는 알고리즘 프로세스를 도시한 도면이다.

[0137] 도 21, 도 22 및 도 23에 도시한 바와 같이, 데이터해석부(200)가 진행하는 발화 특징 파악 알고리즘에 있어서, 각 자모음의 단위의 패턴을 파악하는 단계는, 조음기관 물리특성을 파악한 센서부(100)가 구강설(12)일 경우에 x, y, z축을 기반으로 그 자모음 단위의 패턴을 파악한다.

[0138] 이때, 알고리즘은 K-nearset Neihbor(KNN), Artificial Neural Network(ANN), Convolutional Neural Network(CNN), Recurrent Neural Network(RNN), Restricted Boltzmann Machine(RBM), Hidden Markov Model(HMM) 중 하나이상의 알고리즘에 기반할 수 있다.

[0139] 예를 들어,도 22, 도 23에서, 구강설 센서(110)가 벡터량의 변화량 내지 각도 변화량을 파악하는 센서로 구동될 경우, 화자의 발화를 측정함으로써 벡터량의 변화량, 각도 변화량을 파악하고, 이를 통해 고설성(Tongue Height)과 전설성(Tongue Frontness)을 가지는 모음 [i]으로 인지한다.

[0140] 또한, 구강설 센서(110)가 압전신호 내지 마찰전기신호의 원리로 구동되는 센서일 경우, 압전에 따른 전기 신호 변화와 구강설 센서(110)와 구강 내외부의 조음기관과 근접 내지 마찰하여 발생하는 마찰전기신호를 파악하여 고설성과 전설성을 가지는 모음 [i]으로 인지한다.

[0141] 모음 [u]의 경우에도 같은 원리들을 기반으로, 고설성(Tongue Height: High)과 후설성(Backness)를 측정하여 해당 모음으로 파악하게 된다. [ ]의 경우에도 같은 원리들을 기반으로, 저설성(Tongue Height: Low)r과 전설성(Tongue Frontness)를 측정하여 해당 모음으로 파악한다.

[0142] 도 23에서, 구강설 센서(110)는 화자의 발화에 따라 발생한 [i], [u], [ ]과 같은 모음을 발화 특징(220)으로 측정한다. 이러한 모음의 발화 특징(220)은 데이터베이스부(350)의 자모음의 음소 단위 색인(361)에 대응한다.

[0143] 도 24는 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터해석부가 Alveolar Stop을 활용하는 경우를 도시한 도면이다.

[0144] 도 24에 도시한 바와 같이, 구강설 센서(110)는 화자에 의해 발화된 특정 자음을 발화 특징(220)으로 측정한다. 이러한 자음의 발화 특징(220)은 데이터베이스부(350)의 자모음의 음소 단위 색인(361)에 대응되며, 이를 데이터해석부(200)가 언어데이터(310)인 Alveolar Stop으로 파악한다.

[0145] 도 25는 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터해석부가 Bilabial Stop을 활용하는 경우를 도시한 도면이다.

[0146] 도 25에 도시한 바와 같이, 구강설 센서(110)와 안면 센서(120)는 화자에 의해 발화된 특정 자음을 발화 특징(220)으로 측정한다. 이러한 자음의 발화 특징(220)은 데이터베이스부(350)의 자모음의 음소 단위 색인(361)에 대응되며, 이를 데이터해석부(200)가 언어데이터(310)인 Bilabial Stop으로 파악한다.

[0147] 도 26은 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터해석부가 Voiced Bilabial Stop을 활용한 실험결과를 도시한 도면이다.

[0148] 도 26에 도시한 바와 같이, 구강설 센서(110)와 안면 센서(120)는 화자에 의해 발화된 특정 자음을 발화 특징(220)으로 측정한다. 이러한 자음의 발화 특징(220)은 데이터베이스부(350)의 자모음의 음소 단위 색인(361)에 대응되며, 이를 데이터해석부(200)가 언어데이터(310)인 Voiced Bilabial Stop인 /버/와 Voiceless Bilabial Stop인 /퍼/로 파악하였다.

[0149] 도 27 및 도 28은 각각 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터해석부가 Labiodental Fricatives를 활용하는 경우를 도시한 도면이다.

[0150] 도 27 및 도 28에 도시한 바와 같이, 구강설 센서(110), 안면 센서(120), 음성취득센서(130). 성대센서(140), 치아센서(150)은 화자에 의해 발화된 특정 자음을 발화 특징(220)으로 측정한다. 이러한 자음의 발화 특징(220)은 데이터베이스부(350)의 자모음의 음소 단위 색인(361)에 대응되며, 이를 데이터해석부(200)가, 도 27의 경우 언어데이터(310)인 Voiceless Labiodental Fricative, 도 28의 경우에는 Voiced Labiodental Fricative로 파악한다.

[0151] 도 29는 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터해석부 및 데이터베이스의 연동을 도시한 도면이다.

[0152] 도 29에 도시한 바와 같이, 촬상 센서(160)는 화자가 구강설 센서(110), 안면 센서(120), 음성취득센서(130). 성대센서(140), 치아센서(150) 중 하나이상을 사용하는 상황에서 발화할 시에 발생하는 두경부의 조음기관 위치 변화 정보(161), 두경부 표정 변화 정보(162), 비언어적 표현 정보(163)를 음성데이터(310)로 인지하여 처리한다.

[0153] 특히, 두경부의 일면에 위치한 안면 센서는 레퍼런스 센서(121)를 기준으로 양극 센서(122)와 음극 센서(123)의 전위차를 가지고 그 자체 위치를 제공하며, 이는 촬상 센서(160)가 촬상함으로써 파악되는 물리적인 두경부의 조음기관 위치 변화 정보(161), 두경부 표정 변화 정보(162), 비언어적 표현 정보(163)와 함께 언어데이터(310)로 데이터변환부(300)에 전달된다.

[0154] 도 30은 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터해석부가 특정 단어로 파악하는 경우를 도시한 도면이다.

[0155] 도 30에 도시한 바와 같이, 구강설 센서(110), 안면 센서(120), 음성취득센서(130), 성대센서(140), 치아센서(150)가 화자에 의해 발화된 특정 자음과 모음을 측정하고, 데이터해석부(200)는 자음과 모음을 발화 특징(220)으로 파악한다. 이러한 각 자모음의 발화 특징(220)인 [b], [i], [f]는 데이터베이스부(350)의 자모음의 음소 단위 색인(361)에 각각 대응되며, 데이터해석부가 이를 /beef/ 내지 [bif]라는 단어로 파악한다.

[0156] 도 31은 본 발명의 제2실시예에 따른 발화 구현 시스템의 데이터베이스부를 도시한 도면이다.

[0157] 도 31에 도시한 바와 같이, 데이터베이스부(350)의 언어데이터 색인(360)은 자모음의 음소단위 색인(361), 음절 단위 색인(362), 단어 단위 색인(363), 구절 단위 색인(364), 문장 단위 색인(365), 연속 발화 색인(366), 발음의 고저 색인(367)으로 구성된다.

[0158] 도 32는 본 발명의 제3실시예에 따른 발화 구현 시스템을 도시한 도면이다.

[0159] 도 32에 도시한 바와 같이, 데이터해석부(200)와 데이터표현부(도 34의 500) 중 하나 이상이 외부에 위치하여 작동할 경우, 연동되어 통신할 수 있는 통신부(400)를 포함된다. 통신부(400)는, 유선 및 무선으로 구현되며, 무선의 경우 블루투스, 와이파이, 3G, 4G, NFC 등 다양한 방법이 사용될 수 있다.

[0160] 도 33 및 도 34는 각각 본 발명의 제3실시예에 따른 발화 구현 시스템의 데이터베이스부의 실제 형태를 도시한 도면이다.

[0161] 도 33 및 도 34에 도시한 바와 같이, 데이터해석부(200)와 연동되는 데이터베이스부(350)는 언어데이터 색인을 가지고서 실제 발화에 따른 발화 특징(220), 화자의 음성(230), 발화 내역 정보(240), 발화 변이(250)을 언어데이터(310)으로 파악한다.

[0162] 도 33은, 도 23의 High Front tense Vowel과 High Back tense Vowel, 도 24의 Alveolar Sounds, 도 27의 Voiceless labiodental fricative를 포함하는 다양한 자모음 발화 특징을 센서부(100)가 측정하고 데이터해석부(200)가 반영한 데이터베이스부(350)의 실제 데이터이다.

[0163] 도 34는, 도 23의 High Front lax Vowel, 도 24의 Alveolar Sounds, 도 25의 Bilabial Stop Sounds를 포함하는 다양한 자모음 발화 특징을 센서부(100)가 측정하고 데이터해석부(200)가 반영한 데이터베이스부(350)의 실제 데이터이다.

[0164] 도 35는 본 발명의 제4실시예에 따른 발화 구현 시스템을 도시한 도면이고, 도 36은 본 발명의 제4실시예에 따른 발화 구현 시스템의 센서부, 데이터해석부, 데이터표현부와 텍스트정보부 및 데이터베이스부의 연동을 도시

한 도면이다.

[0165] 도 35에 도시한 바와 같이, 본 발명의 제4실시예에 따른 발화 구현 시스템은 유기적으로 연동되어 작동하는 센서부(100), 데이터해석부(200), 데이터변환부(300), 데이터베이스부(350), 데이터표현부(500) 및 텍스트정보부(600)를 포함한다. 이에 관한 자세한 설명은 도 52, 도 53, 도 54에서 후설할 것이다.

[0166] 도 36에 도시한 바와 같이, 센서부(100)가 실제 조음기관에 위치하여 화자의 발화에 따른 조음기관 물리특성을 측정하고 이를 데이터해석부(200)로 전달하고 데이터해석부(200)는 이를 언어데이터로 해석한다. 해석된 언어데이터는 데이터표현부(500)로 전달되며, 텍스트정보부(600)의 화자데이터취득(S710)부터 데이터생성단계(S760)에 이르는 처리 단계(S700)에 따라 음성정보(640)은 데이터표현부(500)으로 전달된다. 그 언어데이터(310) 및 음성정보(640)의 해석 과정과 표현 과정에서 데이터베이스부(350)가 연동되어 작동함을 알 수 있다.

[0167] 도 37 내지 도 41은 각각 본 발명의 제4실시예에 따른 발화 구현 시스템의 데이터표현부가 언어데이터를 표현하는 수단을 도시한 도면이다.

[0168] 도 37 내지 도 41에 도시한 바와 같이, 센서부(100)에 의해 획득된 화자의 두경부 조음기관의 물리특성은 데이터해석부(200)를 통해 센서부의 위치(210), 발화 특징(220), 화자의 음성(230), 발화 내역 정보(240), 발화 변이(250)로 파악된다.

[0169] 촬상센서(160)는 화자의 두경부 조음기관의 외관상 변화를 촬상하고, 데이터해석부(200)는 이를 통해 화자의 두경부 조음기관의 위치 변화 정보(161), 두경부 표정 변화 정보(162)를 파악한다.

[0170] 이후, 이러한 정보들은 데이터해석부(200)에서 언어데이터(310)로 변환되며, 데이터표현부(500)에서 외부로 표현된다.

[0171] 이때, 도 37은 언어데이터(310)를 데이터표현부(500)가 상기 텍스트정보부와의 연동에 따라 음성정보(640)의 형태로 청각적으로 표현하는 것을 나타낸 것이고, 도 38은 데이터표현부(500)가 언어데이터(310)를 시각적으로 표현함에 있어서, 데이터해석부(200)가 측정한 화자의 조음기관의 물리특성을 데이터베이스부(350)의 언어데이터 색인(360)과 비교하여, 실제 표준 발음의 광역표기 (broad description)와 함께 강세의, 유사근접도, 발화 의도 중 하나이상을 측정한 수치를 함께 제공하는 것을 나타낸 것이다.

[0172] 도 39는 데이터표현부(500)가 언어데이터(310)를 시각과 청각적으로 표현함에 있어서, 데이터해석부(200)가 측정한 화자의 조음기관의 물리특성을 데이터베이스부(350)의 언어데이터 색인(360)과 비교하여, 실제 표준 발음의 미세표기(narrow description)와 함께 강세의, 유사근접도, 발화 의도 중 하나이상을 측정한 수치를 함께 제공하는 것을 나타낸 것이다.

[0173] 도 40은 데이터표현부(500)가 언어데이터(310)를 시각적으로 표현함에 있어서, 데이터해석부(200)가 측정한 화자의 조음기관의 물리특성을 데이터베이스부(350)의 언어데이터 색인(360)과 비교하여, 실제 표준 발음의 미세표기(narrow description)와 함께 강세의, 유사근접도, 발화 의도 중 하나 이상을 측정한 수치, 그리고 해당 언어데이터(310)가 단어로서 단어 단위 색인(363)에 대응할 경우, 그에 해당하는 이미지를 함께 제공하는 것을 나타낸 것이다.

[0174] 도 41은 데이터표현부(500)가 언어데이터(310)를 시각과 청각적으로 표현함에 있어서, 데이터해석부(200)가 측정한 화자의 조음기관의 물리특성을 데이터베이스부(350)의 언어데이터 색인(360)과 비교하여, 실제 표준 발음의 미세표기(narrow description)와 함께 강세의, 유사근접도, 발화 의도 중 하나이상을 측정한 수치를 제공하고 화자에 의한 언어데이터(310)를 교정하고 강화할 수 있도록 해당 발음을 발화할 수 있는 발화 교정 이미지를 함께 제공하는 것을 나타낸 것이다.

[0175] 도 42는 본 발명의 제4실시예에 따른 발화 구현 시스템의 데이터표현부가 언어데이터를 시각적 및 청각적으로 표현하는 경우를 도시한 도면이다.

[0176] 도 42에 도시한 바와 같이, 데이터표현부(500)가 언어데이터(310)를 문자로 시각화하고 상기 텍스트정보부(600)에 따라 음성정보(640)로 청각화하여 제공함에 있어서, 데이터해석부(200)가 측정한 화자의 조음기관의 물리특성을 데이터베이스부(350)의 자모음 음소 단위 색인(361), 음절단위 색인(362), 단어단위 색인(363), 구절 단위 색인(364), 문장 단위 색인(365) 중 하나이상의 언어데이터 색인(360)과 비교한다.

[0177] 이러한 언어 데이터(310)를 데이터표현부(500)가 화자의 언어데이터(310)에 관련된 실제 표준 발음의 미세표기(narrow description)와 함께 강세의, 유사근접도, 발화 의도 중 하나이상을 측정한 문자와 소리로 제공하여 화

자가 언어데이터(310)를 교정하고 강화할 수 있도록 돕는다.

[0178] 도 43은 본 발명의 제4실시예에 따른 발화 구현 시스템의 데이터표현부가 언어데이터를 시각적으로 표현하는 경우를 도시한 도면이다.

[0179] 도 43에 도시한 바와 같이, 데이터표현부(500)가 언어데이터(310)를 문자, 그림, 사진, 영상 중 하나이상으로 시각화하여 제공한다.

[0180] 이때, 데이터해석부(200)는 측정한 화자의 조음기관의 물리특성을 데이터베이스부(350)의 자모음 음소 단위 색인(361), 음절단위 색인(362), 단어단위 색인(363), 구절 단위 색인(364), 문장 단위 색인(365) 중 하나 이상의 언어데이터 색인(360)과 비교한다.

[0181] 더불어, 문자로 시각화 될 경우, 실제 표준 발음의 미세표기(narrow description)와 광역표기 (broad description)를 모두 제공한다. 이를 통해 언어 데이터(310)를 데이터표현부(500)가 화자의 언어데이터(310)에 관련된 실제 표준 발음의 미세표기(narrow description) 및 광역표기 (broad description)와 함께 강세의, 유사근접도, 발화 의도 중 하나이상을 측정한 문자와 소리로 제공하여 화자가 언어데이터(310)를 교정하고 강화할 수 있도록 돕는다.

[0182] 도 44는 본 발명의 제4실시예에 따른 발화 구현 시스템의 데이터표현부가 언어데이터를 시각적으로 표현하는 경우를 도시한 도면이다.

[0183] 도 44에 도시한 바와 같이, 데이터표현부(500)가 언어데이터(310)를 문자로 시각화하여 제공함에 있어서, 데이터해석부(200)가 측정한 화자의 조음기관의 물리특성을 데이터베이스부(350)의 자모음 음소 단위 색인(361), 음절단위 색인(362), 단어단위 색인(363), 구절 단위 색인(364), 문장 단위 색인(365), 연속발화색인(366) 중 하나 이상의 언어데이터 색인(360)과 비교한다. 이러한 언어 데이터(310)를 데이터표현부(500)가 화자의 언어데이터(310)에 관련된 실제 표준 발음의 미세표기(narrow description) 및 광역표기 (broad description)와 함께 강세의, 유사근접도, 발화 의도 중 하나이상을 측정한 연속 발화 단위의 문자와 소리로 제공하여 화자가 언어데이터(310)를 교정하고 강화할 수 있도록 돕는다.

[0184] 도 45는 본 발명의 제4실시예에 따른 발화 구현 시스템의 데이터표현부가 언어데이터를 연속 발화 단위로 표현하는 경우를 도시한 도면이다.

[0185] 도 45에 도시한 바와 같이, 데이터표현부(500)가 언어데이터(310)를 문자로 시각화하고 소리로 청각화하여 제공함에 있어서, 데이터해석부(200)가 측정한 화자의 조음기관의 물리특성을 데이터베이스부(350)의 자모음 음소 단위 색인(361), 음절단위 색인(362), 단어단위 색인(363), 구절 단위 색인(364), 문장 단위 색인(365), 연속발화색인(366), 발음의 고저 색인(367) 중 하나 이상의 언어데이터 색인(360)과 비교한다. 이러한 언어 데이터(310)를 데이터표현부(500)가 화자의 언어데이터(310)에 관련된 실제 표준 발음의 미세표기(narrow description) 및 광역표기 (broad description)와 함께 강세의, 유사근접도, 발화 의도 중 하나이상을 측정한 문자와 소리로 제공하여 화자가 언어데이터(310)를 교정하고 강화할 수 있도록 돕는다.

[0186] 도 46은 본 발명의 제4실시예에 따른 발화 구현 시스템이 활용하는 Confusion Matrix를 도시한 도면이고, 도 47은 본 발명의 제4실시예에 따른 발화 구현 시스템이 활용하는 Confusion Matrix를 백분율로 도시한 도면이다.

[0187] 도 46 및 도 47에 도시한 바와 같이, 데이터해석부(200)는 언어데이터(310)를 파악함에 있어서, Time Domain의 Variance, 주파수 영역의 Cepstral Coefficient, Linear Predict Coding Coefficient를 사용하는 하나 이상의 특징 추출 알고리즘을 대표하여 사용한다.

[0188] 데이터의 분산 정도를 나타내는 Variance는 다음 [수학 식1]에 따라 계산된다. 여기서, n은 모집단의 네트워크, $\bar{x}$는 수집된 조음기관 물리 특성 인 데이터의 모집단의 평균, xi는 수집된 조음기관 물리 특성인 데이 터들을 나타낸다.

[0189] [수학식1]

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{N}(x_i - \bar{x})^2$$

[0190]

[0191] Cepstral Coefficient는 주파수의 세기를 정형화하기 위해 다음 [수학식 2]로 계산된다. 여기서, F-1는 역 푸리에급수 변환인 Inverse Fourrier Transform을 나타내고, X(f)는 신호에 대한 주파수의 스펙트럼을 나타낸다. 본 예시에서는 Cepstral의 Cofficent는 n=0일 때의 값을 활용하였다.

[0192] [수학식2]

$$C_n = F^{-1}\log|X(f)|$$

[0193]

[0194] Linear Predict Coding Coefficient는 주파수의 선형적 특성을 나타내는 것으로 다음[수학식 3]에 따라 계산된다. 여기서, n은 표본의 개수를 나타내며, ai는 Linear Predict Coding Coefficient 계수이다. Cepstral의 계수는 n=1일때의 값을 사용하였다.

[0195] [수학식 3]

$$C_n = -a_n - \frac{1}{n}\sum_{i=1}^{n-1}(n-i)a_i c_{n-i}$$

[0196]

[0197] 더불어, 조음기관 물리 특성인 데이터를 유사도에 따라 묶고 예측데이터를 생성하여 각 데이터를 분류하는 ANN을 활용하였다. 이를 통해, 화자가 최초 발화에 대해 표준 발화에 대비하여 본인 발화 내용의, 근접유사도, 발화 의도를 파악할 수 있게 된다. 이를 바탕으로 화자는 자신에 발화 내용에 대한 피드백을 얻고 지속적으로 발화 교정을 위한 재발화를 실시한다. 이러한 반복적 조음기관 물리특성 데이터 입력 방식을 통해 많은 조음기관 물리 특성 데이터가 모이고 ANN의 정확도가 증가한다.

[0198] 여기서, 입력 데이터인 조음기관 물리 특성을 10개의 자음으로 선정하였고, 추출 과정에서 5개의 조음위치인 Bilabial, Alveolar, Palatal, Velar, Glottal로 분류되었다. 이를 위해, 상기 5개의 조음위치에 해당하는 10개의 자음을 순서대로 100번씩, 총 1000번 무작위로 50번씩 총 500번 발음을 하였다.

[0199] 이에 따라, 도 45에 도시한 바와 같이, 자음 분류를 위한 Confusion Matrix가 형성되었다. 이를 기반으로 각 조음위치마다 발화되는 자음의 개수가 상이하다는 것을 고려하여, 도 46에 도시한 바와 같이, 백분율로 나타내었다.

[0200] 이를 통해, 화자는 표준 발화와 대비하여 발음의 및 근접유사도가 낮은 Palatal과 관련하여, 자음을 제대로 발화하지 못함을 알 수 있다. 또한 도 46에 도시한 바와 같이, Palatal과 관련된 자음을 발화하고자 하였으나 Alveolar와 관련된 자음으로 잘못 발화한 경우는 17%이다. 이는 화자가 Palatal과 관련된 자음과 Alveolar와 관련된 자음 간의 차이를 명확히 인지하지 못함을 의미한다.

[0201] 도 48은 본 발명의 제4실시예에 따른 발화 구현 시스템이 화면을 통해 화자로 하여금 언어 교정 및 지도를 돕는 경우를 도시한 도면이다.

[0202] 도 48에 도시한 바와 같이, 영어를 모국어로 하지 않는 한국인 화자는 [kiŋ]을 의도하고서, 발화하였고 센서부(100)는 상기 발화에 따른 조음기관 물리 특성을 파악한다.

[0203] 그러나, 화자의 경우, 한국어에 존재하지 않는 Velar Nasal Sound인 [ŋ]에 대해 조음과 발화 방법에 대해 미숙하였다.

[0204] 이에 데이터해석부(200)는 화자가 제대로 발화하지 못한 [ŋ]를 표준 발화 특징 행렬(205)과의 비교를 통해 파악한다. 이후, 데이터표현부(300)은 화자의 발화, 유사도를 제공하였고, 결과는 46%에 그쳤다. 이후, 데이터표현부(300)는 화면 등을 통해, 화자로 하여금 [kiŋ]을 정확히 발음할 수 있도록 돕는다.

[0205] 이때, 데이터표현부(300)는 화자가 어느 조음 기관을 어떻게 조작해야 하는지 직관적으로 보여주기 위해 Speech Guidance(Image)를 제공한다. 데이터표현부(300)가 제시하는 Speech Guidance(Image)는 상기 [ŋ]을 발화하기 위한 조음기관에 부착되거나 인접한 센서부를 기반으로 발화 교정 및 안내를 실시한다. 예를 들어, 상기 [kiŋ]의 경우, [k]은 혀의 뒷부분(Tongue Body, Root)을 Velar(연구개) 방향으로 승강시켜 붙이고 유격시키면서

파열음을 내고, 성대의 떨림 없이 무성음으로, 입을 통해 /크/로 발화해야 한다.

[0206] 이때, 혀의 뒷부분이 Velar(연구개) 방향으로 승강시키고 붙였다 유격되는 파열음을 내는 것은 구강설 센서(110)가 측정하게 된다. [i]의 경우에는, 전설 고설 긴장 모음(High Front Tense Vowel)임으로, 이 역시 구강설 센서(110)가 혀의 고설성(Hight)과 전설성(Frontness)을 파악한다. 더불어, [i]를 발화할 때, 입술의 양 끝이 양 볼로 각 각 당겨지는 현상이 발생한다. 이를 상기 안면 센서(120)가 파악하게 된다. [ŋ]의 경우에는, 혀의 뒷부분(Tongue Body, Tongue Root)를 Velar(연구개) 방향으로 승강시키고 코를 울려 발화해야 한다. 그렇기에 역시 구강설 센서(110)가 혀의 고저설성 및 전후설성을 파악한다.

[0207] 더불어, 상기 발음은 비음이기에 코와 그 주변의 근육이 떨리게 된다. 이러한 현상은 상기 안면 센서(120)가 코 주변에 부착됨으로써 파악될 수 있다.

[0208] 도 49는 본 발명의 제4실시예에 따른 발화 구현 시스템이 두경부 조음기관의 외상을 촬상하고 파악하는 경우를 도시한 도면이다.

[0209] 도 49에 도시한 바와 같이, 촬상 센서(160)는 발화에 따른 화자의 두경부 조음기관의 외관상 변화를 촬상하고, 데이터해석부(200)는 이를 통해 화자의 두경부 조음기관의 위치 변화 정보(161), 두경부 표정 변화 정보(162)를 파악한다. 이때 센서부(100)의 구강설 센서(110), 안면센서(120), 음성 취득 센서(130), 성대 센서(140), 치아 센서(150)을 통해 파악된 화자의 발화 특징(210)도 함께 데이터해석부(200)가 고려하게 된다.

[0210] 도 50은 본 발명의 제4실시예에 따른 발화 구현 시스템이 표준 발화 특징 행렬을 통해 상호 정보들을 결합시키는 경우를 도시한 도면이다.

[0211] 도 50에 도시한 바와 같이, 센서부(100)의 구강설 센서(110), 안면센서(120), 음성 취득 센서(130), 성대 센서(140). 치아센서(150)는 화자의 발화 특징(210)을 파악하고, 촬상 센서(160)는 화자의 두경부 조음기관의 위치 변화 정보(161), 두경부 표정 변화 정보(162)를 파악한다. 이를 통해, 데이터해석부(200)가 표준 발화 특징 행렬(205)을 기반으로 두경부 조음기관의 위치 변화 정보(161), 두경부 표정 변화 정보(162)에 대응하는 발화 특징을 결합시킨다.

[0212] 도 51에 도시한 바와 같이, 상기 촬상 센서(160)는 화자의 두경부 조음기관의 변화 정보(161), 두경부 표정 변화 정보(162)뿐 만 아니라, 화자가 발화하면서 표현하는 비언어적 표현(153)을 촬상한다. 즉, 비언어적 표현(153)는 화자의 발화 의도에 따라 발생하는 두경부의 상하좌우의 기울임, 흉곽부의 들썩임, 두경부와 흉곽부 간의 핏대 및 근육의 긴장, 상지부의 떨림, 상지부의 제스쳐, 하지부의 떨림, 하지부의 제스쳐 중 하나이상의 표현을 포함한다.

[0213] 도 52에 도시한 바와 같이, 상기 텍스트정보부(600)는 언어데이터를 음성정보로 출력함에 있어서, 화자데이터처리단계(S700)를 거치는 것을 특징으로 한다. 상기 단계는 다음과 같다.

[0214] (a) 단계에서는 상기 데이터변환부에서 적어도 하나의 언어데이터를 획득하는 화자데이터를 취득한다. (S710) 그리고 (b) 단계에서는 상기 데이터베이스부의 언어데이터색인의 단어 단위 색인(363)을 기반으로 상기 언어데이터에서 적어도 하나의 기능어(611) 내지 공백은 0 내지 NULL로 라벨링하여 제외하고, 적어도 하나의 화자내용어(610)를 상기 데이터해석부(200)의 상기 표준 발화 특징행렬(205)로 라벨링하여 화자내용어 테이블(615)을 생성한다. (S720)

[0215] 상기 기능어(Function word)(611)는 앞서 취득된 화자의 언어데이터(310) 중 /the/ 또는 /to/, /of/과 같은 단어들이다. 상기 기능어(611)는 문장 내에서 의미를 생성하지 아니하고, 문장 내의 독립적으로 쓰일 수 없는 조사와 어미를 의미한다. 반면에 내용어(Content Word)는 앞서 취득된 화자의 언어데이터(310) 중 명사, 동사, 형용사, 부사, 관형사를 포함하는, 문장 내에서 독립적으로 쓰일 수 있는 단어의 최소 단위이다. 예를 들어, 도 34에서 측정된 "The postman brings the letter."에서, 기능어는 /the/이고, 내용어는 /postman/, /brings/, /letter/가 된다.

[0216] (c) 단계에서는 상기 데이터베이스부의 언어데이터색인의 음소 단위 색인(361) 내지 음절 단위 색인(362)를 기반으로, 상기 화자내용어 테이블(615)을 Onset(초성), Necleus(중성), Coda(종성) 중 적어도 하나의 자모음 음소 단위로 분리하여 음소값행렬(620)을 생성한다. (S730) (d) 단계에서는 상기 (c) 단계에서 생성된 상기 음소값행렬(620)에 상응하여, 상기 음소 단위 색인(361)의 발화의 진행 시간값과, 발화에 따른 F0, F1, F2, F3, F4 중 적어도 하나의 Formant Frequency값과, 발화의 진폭값 중 적어도 하나를 포함하는 음성값행렬(630)을 생성한다.(S740)

[0217] 이때, 언어데이터색인(360)의 음소 단위 색인(361)은 적어도 하나의 음절(Syllable)을 가진 단어에 대해서 Onset(초성=자음), Necleus(중성= 모음), Coda(종성=자음)의 분류를 가지고 있고, 상기 음절 단위 색인(362)는 상기 분류에 따라 각 음절마다의 Waveform과 Spectrogram을 포함하고 있다. 여기서 Waveform은 발화의 시간(time)을 나타내는 x축과 발화에 따른 강도를 의미하는 진폭(amplitude)를 나타내는 y축으로 구성되며, 발화 시간에 따른 발화 진폭의 변화를 나타내는 정보이다. Spectrogram은 발화의 시간(time)을 나타내는 x축과 발화에 따른 주파수(frequency: Hz)를 나타내는 y축, 그리고 발화에 따른 강도를 의미하는 진폭(intensity)을 나타내는 z축(명암)으로 구성된다. 상기 Spectrogram에서는 각 자모음마다의 상이한 Fundamental Frequency인 F0(Hz), Formant Frequency인 F1(Hz), F2(Hz), F3(Hz)를 파악할 수 있다. F1의 경우, 모음의 저고도성(Height)과 관련이 있으며, 고모음일 경우 F1은 낮게 형성되고, 저모음일 경우 F1은 높게 형성된다. F2의 경우, 모음의 전후설성(Frontness and Backness)과 관련되어, 전설모음일 수록 높고, 후설모음일 수록 낮게 형성된다. 원순후설모음의 경우, F2가 낮게 형성되는 경향이 있다. 예를 들어, 상기 도 52에서, tr]의 모음 중, 굵은 글씨로 표시된 첫번째 모음인 **[ ]**는 Middle Front Vowel로서, 표준 발화 특징 행렬(205)에 따라, (4, 2)에 대응된다.

[0218] 도 53에서 도시된 음소 단위 색인(361)의 Frequency table의 일부 예시는 상기 언급한 F0부터 F4의 평균값 중 일부를 나타낸다. 이때 M은 성인남성(Man), W는 성인여성(Woman), C는 아동(Child)을 의미한다. 예를 들어, 상기 도 53의 모음 중, 굵은 글씨로 표시된 첫번째 모음인 **[ ]**는 Middle Front Vowel로서, 표준 발화 특징 행렬(205)에 따라, (4, 1)에 대응된다.

[0219] 이에 **[ ]**에 있어서, 성인 남성화자의 경우 F0는 127Hz, F1은 580Hz, F2는 1799Hz, F4는 3677Hz를 보인다. 이처럼, 화자에 의해 발화된 각 음소는 저마다 상이한 Frequency를 가지며, 이를 통해 각 음소는 서로 구별된다.

[0220] 이후, 상기 c에서 생성된 상기 음소값행렬(620)중 적어도 하나의 원소에, 상기 음소값행렬(620)의 원소에 대응하는 상기 (d) 단계에서 생성된 음성값행렬(630) 중 적어도 하나의 원소를 할당하여 병합한다(S750). 이를 기반으로 (f) 단계에서는 (e) 단계를 거친 적어도 하나의 원소들끼리 재조합하여, 음소단위, 단어단위, 구절단위, 문장단위, 연속발화단위 중 적어도 하나의 음성정보(640)를 생성한다. (S760)

[0221] 도면에 기재된 방법 외에도 센서부(100)의 경우 다음과 같은 것들이 포함 될 수 있다.

[0222] 1. 압력센서: MEMS 센서, Piezoelectric (압력-전압) 방식, Piezoresistive (압력-저항) 방식, Capacitive 방식, Pressure sensitive 고무 방식, Force sensing resistor (FSR) 방식, Inner particle 변형 방식, Buckling 측정 방식.

[0223] 2. 마찰 센서: 마이크로 hair array 방식, 마찰온도 측정방식.

[0224] 3. 정전기 센서: 정전기 소모 방식, 정전기 발전 방식.

[0225] 4. 전기저항 센서: 직류저항 측정방식, 교류저항 측정방식, MEMS, Lateral 전극 array 방식, Layered 전극 방식, Field Effect Transistor (FET) 방식 (Organic-FET,Metal-oxide-semiconductor-FET, Piezoelectric-oxide-semiconductor -FET 등 포함).

[0226] 5. Tunnel Effect Tactile 센서: Quantum tunnel composites 방식, Electron tunneling 방식, Electroluminescent light 방식.

[0227] 6. 열저항 센서: 열전도도 측정방식, Thermoelectric 방식.

[0228] 7. Optical 센서: light intensity 측정방식, refractive index 측정방식.

[0229] 8. Magnetism based 센서: Hall-effect 측정 방식, Magnetic flux 측정 방식.

[0230] 9. Ultrasonic based 센서: Acoustic resonance frequency 방식, Surface noise 방식, Ultrasonic emission 측정방식.

[0231] 10. 소프트 재료 센서: 고무, 파우더, 다공성 소재, 스펀지, 하이드로젤, 에어로젤, 탄소섬유, 나노탄소재료, 탄소나노튜브, 그래핀, 그래파이트, 복합재, 나노복합재, metal-고분자 복합재, ceramic-고분자 복합재, 전도성 고분자 등의 재료를 이용한 pressure, stress, 혹은 strain 측정 방식, Stimuli responsive 방식.

[0232] 11. Piezoelectric 소재 센서: Quartz, PZT (lead zirconate titanate) 등의 세라믹 소재, PVDF, PVDF copolymers, PVDF-TrFE 등의 고분자 소재, 셀룰로오스, ZnO 나노선 등의 나노소재 방식.

## 부호의 설명

[0234] 100: 센서부 110: 구강설 센서

120: 안면 센서 130: 음성취득 센서

140: 성대 센서 150: 치아 센서

200: 데이터해석부 205: 표준 발화 특징 행렬

210: 센서부의 위치 220: 발화 특징

230: 화자의 음성 240: 발화 내역 정보

250: 발화 변이 300: 데이터변환부

310: 언어 데이터 350: 데이터베이스부

360: 언어 데이터 색인

400: 통신부 500: 데이터표현부

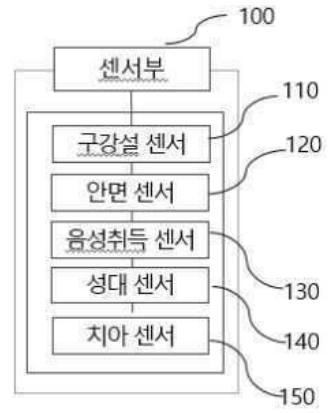
600: 텍스트정보부 610: 화자내용어

615: 화자내용어 테이블 620: 음소값행렬

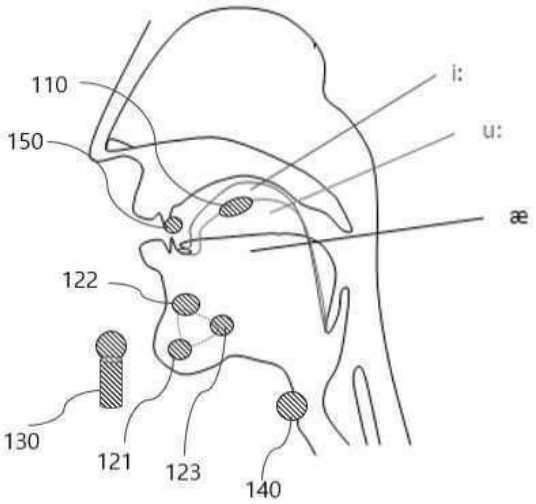
630: 음성값행렬 640: 음성정보

## 도면

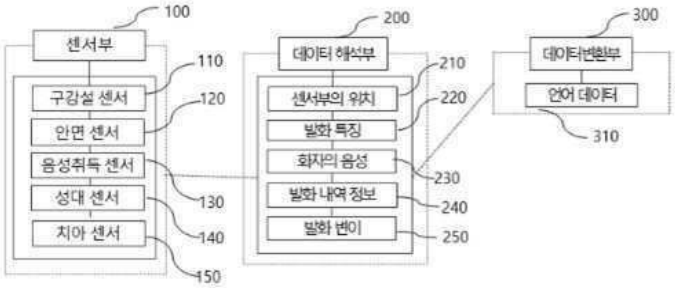
### 도면1



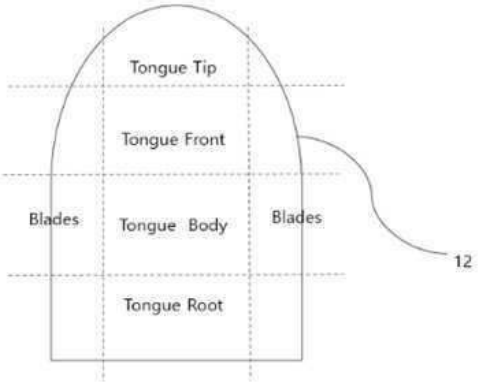
도면2


110
150
i:
u:
æ
122
130
121
123
140

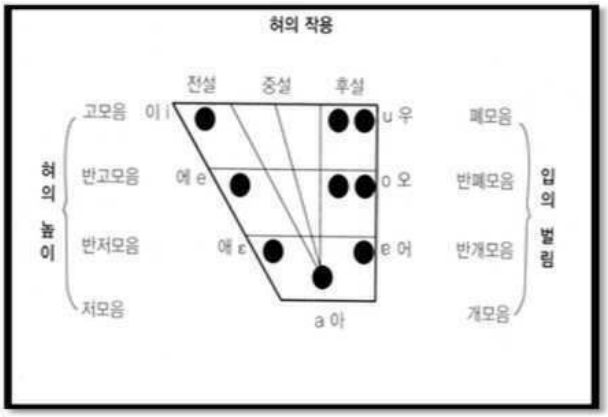
도면3


100
센서부
110
구강설 센서
120
안면 센서
음성취득 센서
130
성대 센서
140
치아 센서
150
200
데이터 해석부
210
센서부의 위치
220
발화 특징
화자의 음성
230
발화 내역 정보
240
발화 변이
250
300
데이터변환부
언어 데이터
310

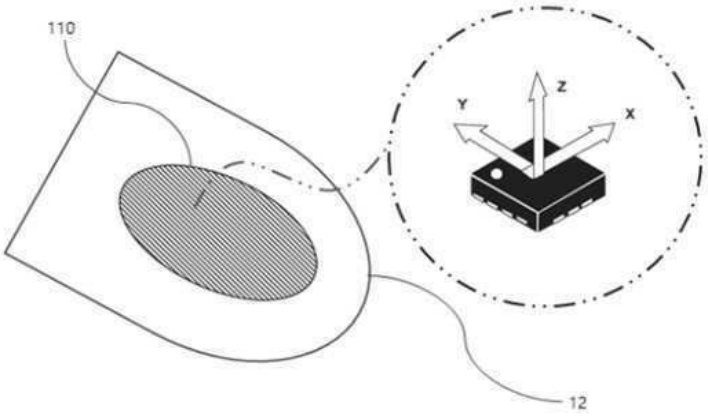
도면4


Tongue Tip
Tongue Front
Blades
Tongue Body
Blades
Tongue Root
12

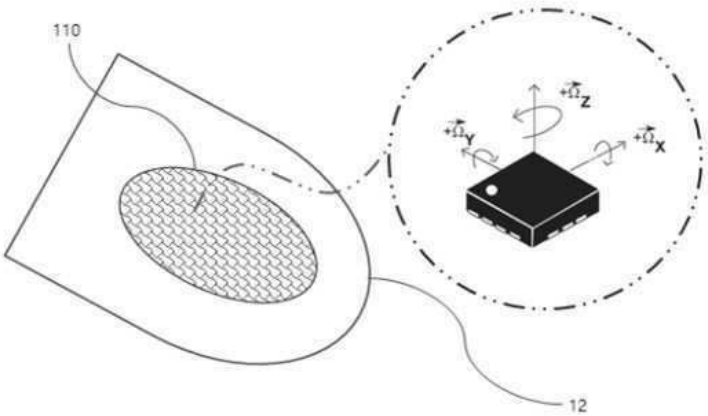
도면5


혀의 작용
전설
중설
후설
고모음
반고모음
반저모음
저모음
혀의 높이
폐모음
반폐모음
반개모음
개모음
입의 벌림
i 이
u 우
e 에
o 오
ɛ 애
ɐ 어
a 아

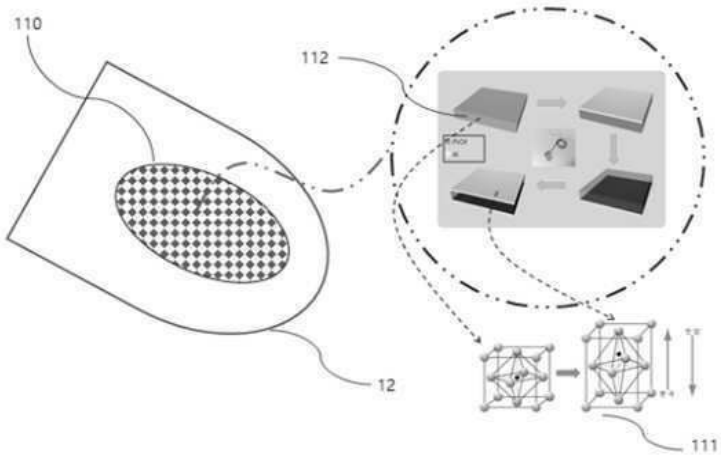
도면6


110
Y
Z
X
12

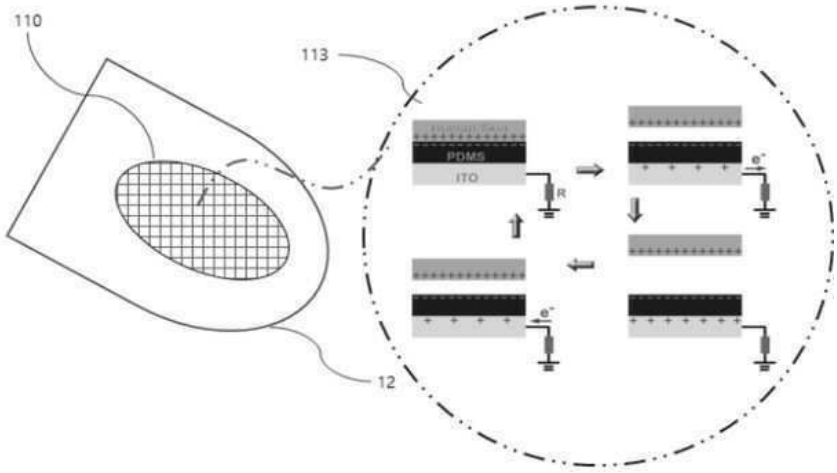
도면7


110
12

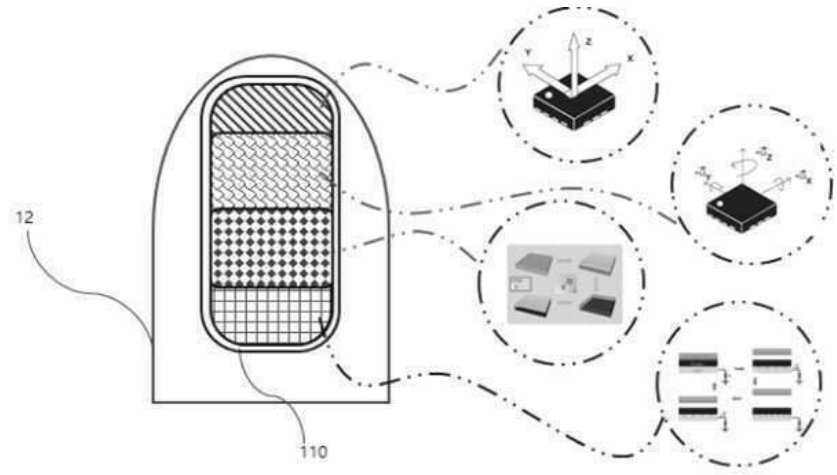
도면8



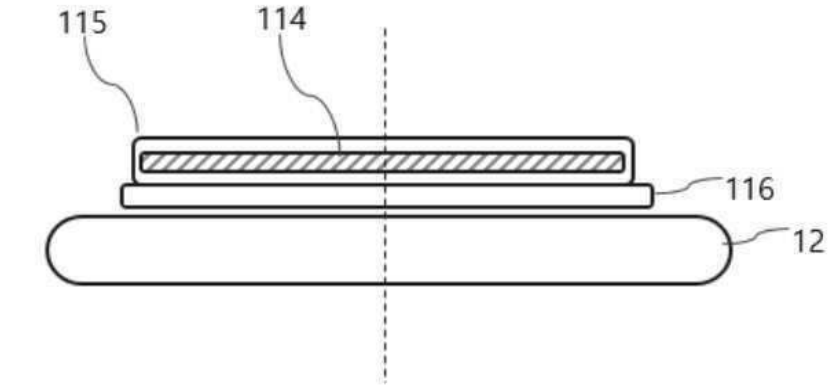

도면9



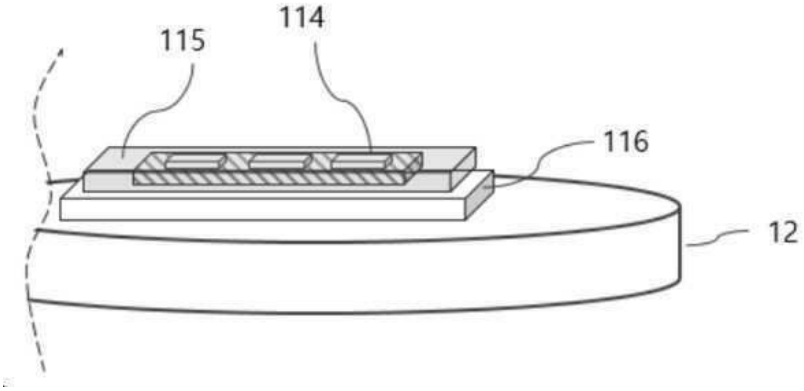

도면10



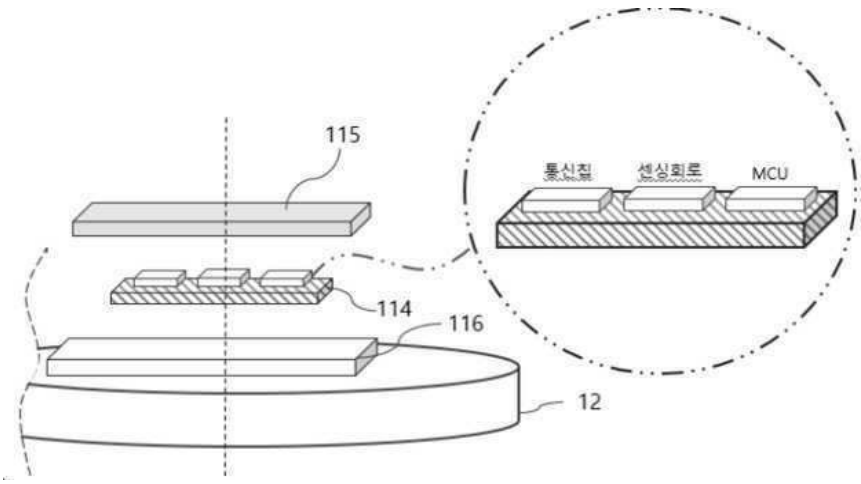
도면11




도면12




도면13



도면14

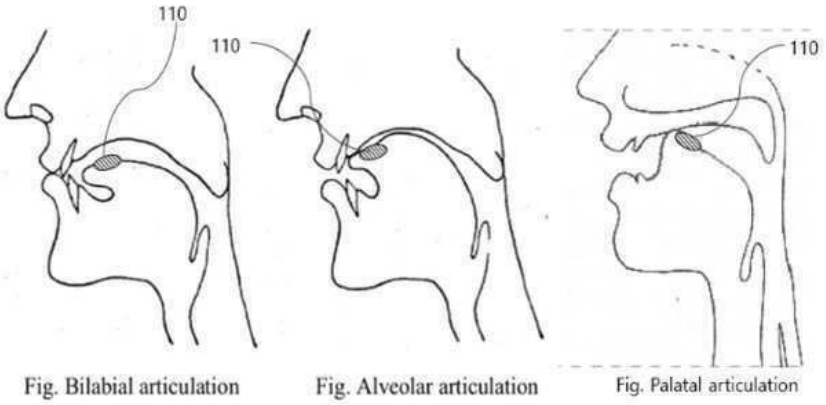
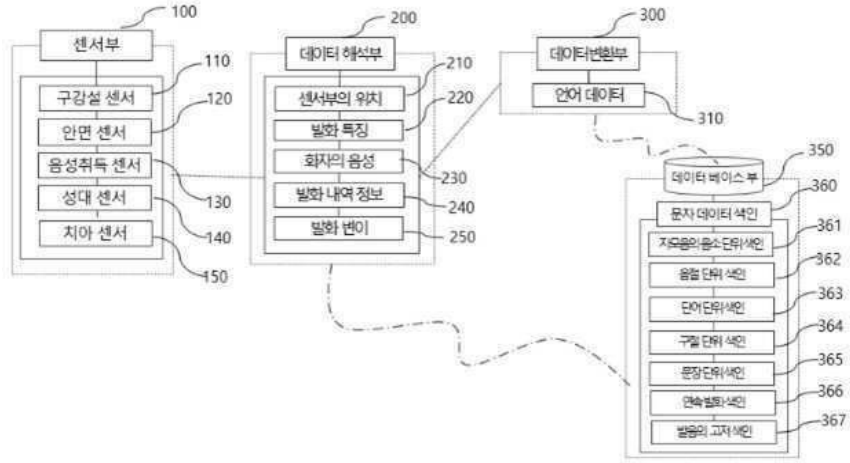
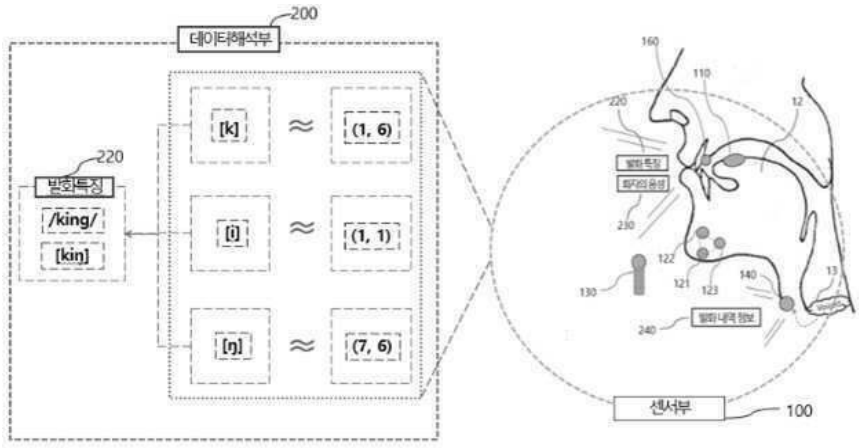



Fig. Bilabial articulation Fig. Alveolar articulation Fig. Palatal articulation

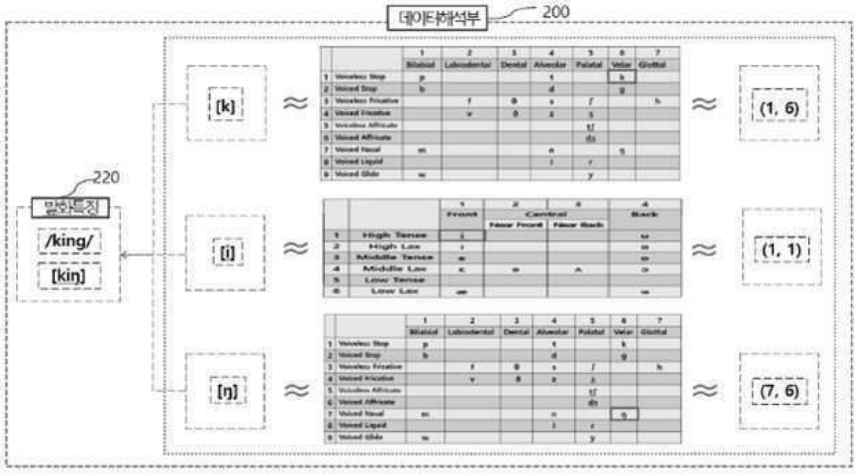
도면15




도면16



도면17




도면18

| | | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| | | Front | Central | | Back |
| | | | Near Front | Near Back | |
| 1 | High Tense | i | | | u |
| 2 | High Lax | ɪ | | | ʊ |
| 3 | Middle Tense | e | | | o |
| 4 | Middle Lax | ɛ | ə | ʌ | ɔ |
| 5 | Low Tense | | | | |
| 6 | Low Lax | æ | | | ɑ |

205

| | | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| | | Front | Central | | Back |
| | | | Near Front | Near Back | |
| 1 | High Tense | 1 | 0 | 0 | 1 |
| 2 | High Lax | 1 | 0 | 0 | 1 |
| 3 | Middle Tense | 1 | 0 | 0 | 1 |
| 4 | Middle Lax | 1 | 1 | 1 | 1 |
| 5 | Low Tense | 0 | 0 | 0 | 0 |
| 6 | Low Lax | 1 | 0 | 0 | 1 |

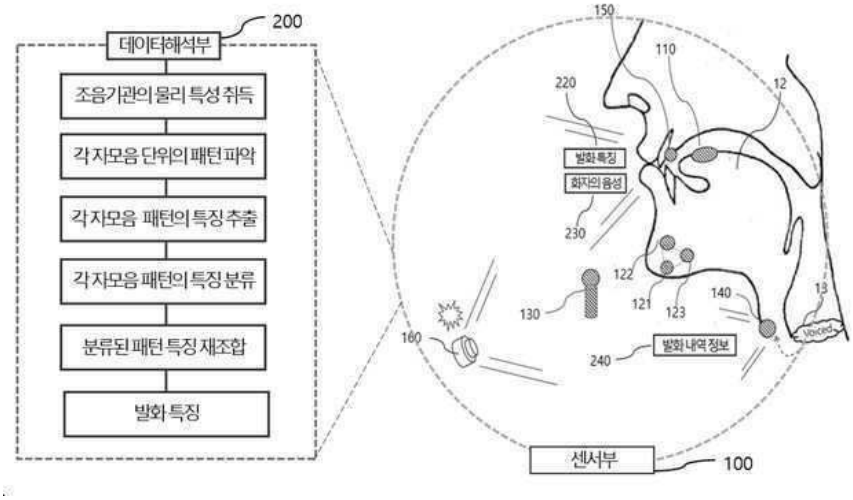
도면19

| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| | | Bilabial | Labiodental | Dental | Alveolar | Palatal | Velar | Glottal |
| 1 | Voiceless Stop | p | | | t | | k | |
| 2 | Voiced Stop | b | | | d | | g | |
| 3 | Voiceless Fricative | | f | θ | s | ʃ | | h |
| 4 | Voiced Fricative | | v | ð | z | ʒ | | |
| 5 | Voiceless Affricate | | | | | tʃ | | |
| 6 | Voiced Affricate | | | | | dʒ | | |
| 7 | Voiced Nasal | m | | | n | | ŋ | |
| 8 | Voiced Liquid | | | | l | r | | |
| 9 | Voiced Glide | w | | | | y | | |

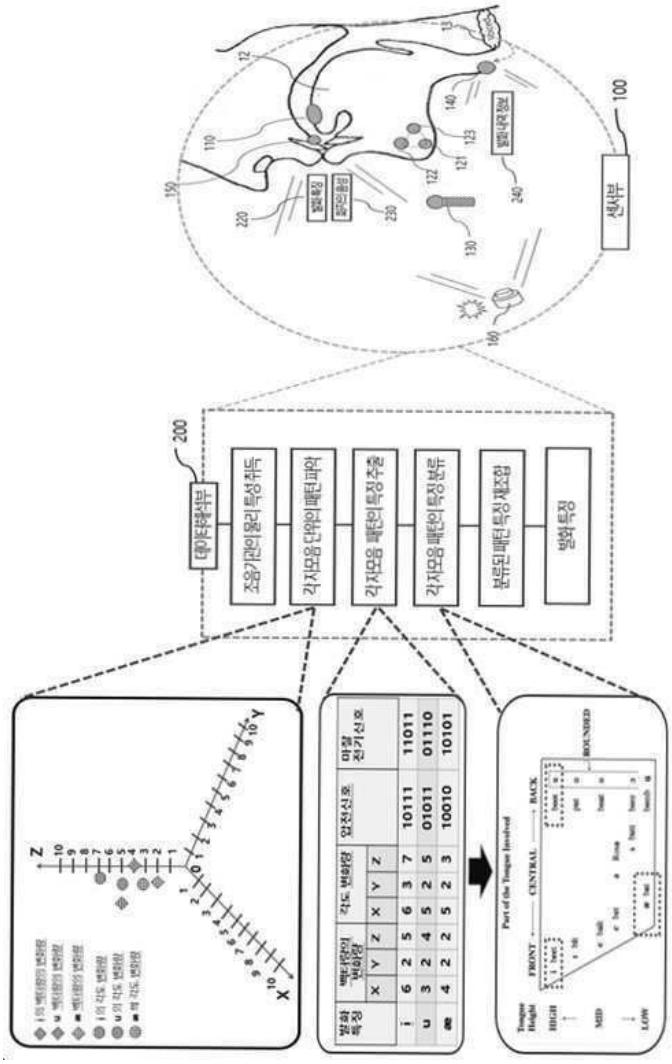
205

| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| | | Bilabial | Labiodental | Dental | Alveolar | Palatal | Velar | Glottal |
| 1 | Voiceless Stop | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| 2 | Voiced Stop | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| 3 | Voiceless Fricative | 0 | 1 | 1 | 1 | 1 | 0 | 1 |
| 4 | Voiced Fricative | 0 | 1 | 1 | 1 | 1 | 0 | 1 |
| 5 | Voiceless Affricate | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 6 | Voiced Affricate | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 7 | Voiced Nasal | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| 8 | Voiced Liquid | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| 9 | Voiced Glide | 1 | 0 | 0 | 0 | 1 | 0 | 0 |

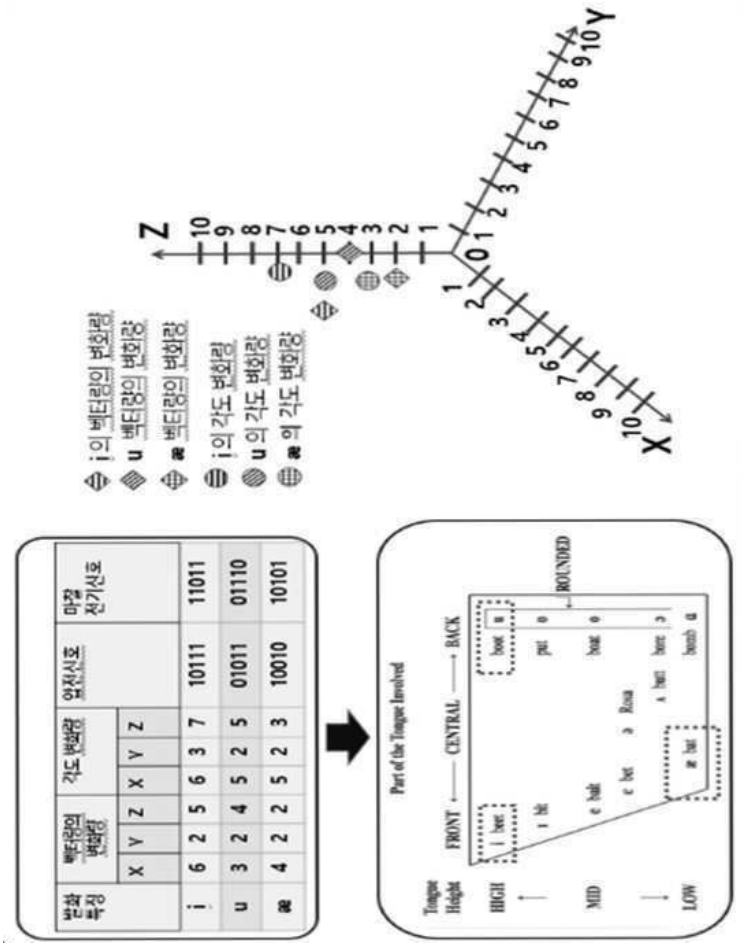
도면20


200
데이터해석부
조음기관의 물리 특성 취득
각자모음 단위의 패턴 파악
각자모음 패턴의 특징 추출
각자모음 패턴의 특징 분류
분류된 패턴 특징 재조합
발화 특징
150
110
12
220
발화 특징
화자의 음성
230
122
121
123
140
13
130
160
240
발화 내역 정보
Voiced
센서부
100

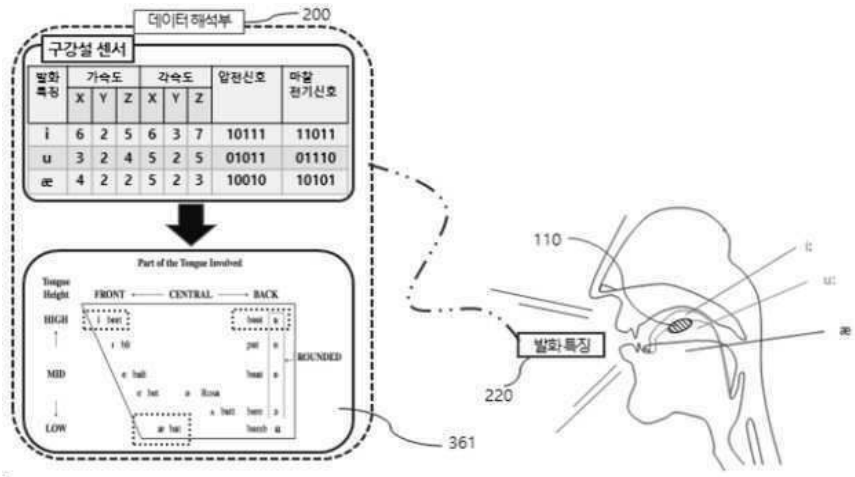
도면21


100
센서부
200
데이터해석부
조음기관의 물리 특성 취득
각자모음 단위의 패턴 파악
각자모음 패턴의 특징 추출
각자모음 패턴의 특징 분류
분류된 패턴 특징 재조합
발화 특징
발화 특징
백터량의 변화량
각도 변화량
압전신호
마찰 전기신호
i 6 2 5 6 3 7 10111 11011
u 3 2 4 5 2 5 01011 01110
æ 4 2 2 5 2 3 10010 10101
Part of the Tongue Involved
FRONT
CENTRAL
BACK
ROUNDED
HIGH
MID
LOW

도면22


Z
Y
X
i 의 벡터량의 변화량
u 벡터량의 변화량
æ 벡터량의 변화량
i 의 각도 변화량
u 의 각도 변화량
æ 의 각도 변화량
발화 특징
벡터량의 변화량
각도 변화량
압전신호
마찰 전기신호
Part of the Tongue Involved
Tongue Height
FRONT
CENTRAL
BACK
HIGH
MID
LOW
ROUNDED

도면23


데이터 해석부
200
구강설 센서
발화 특징
가속도
각속도
압전신호
마찰 전기신호
Part of the Tongue Involved
Tongue Height
FRONT
CENTRAL
BACK
HIGH
MID
LOW
ROUNDED
110
발화 특징
220
361

도면24

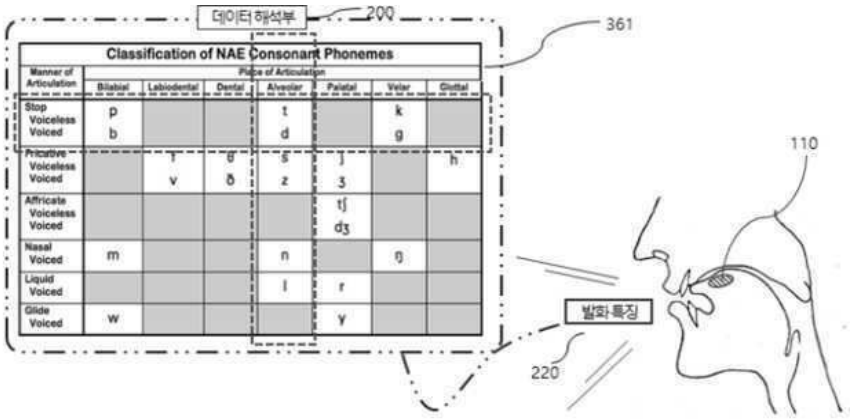
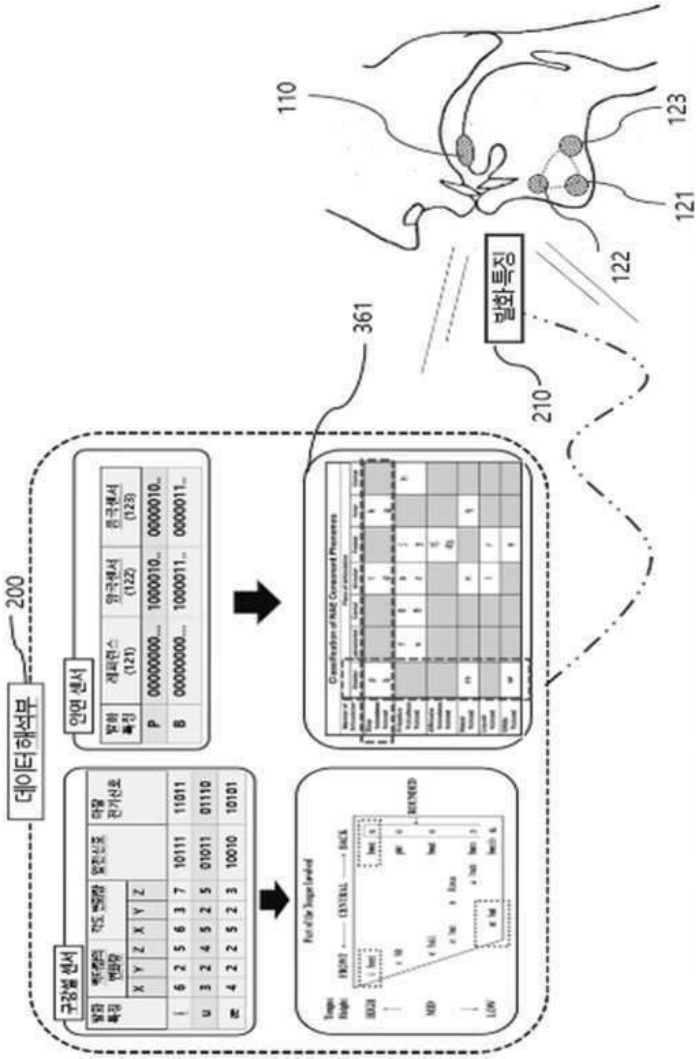
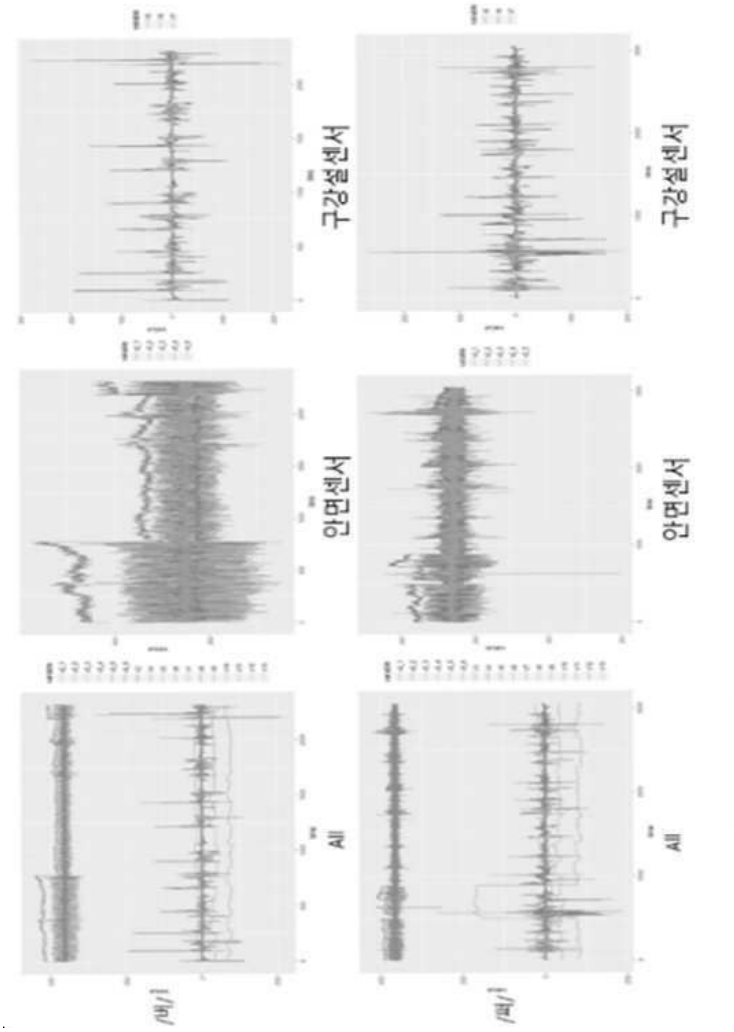



Fig. Alveolar articulation

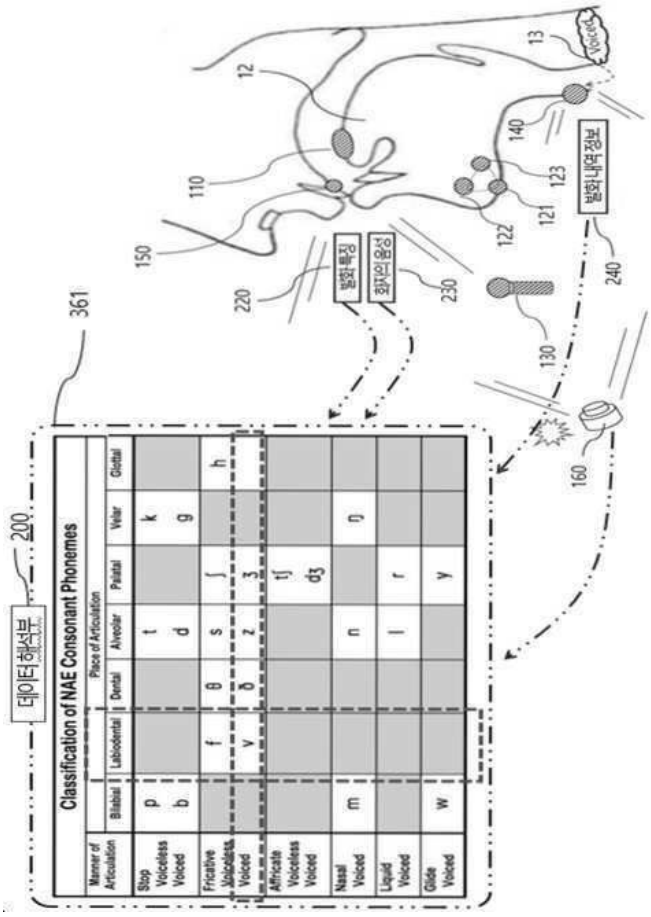
도면25



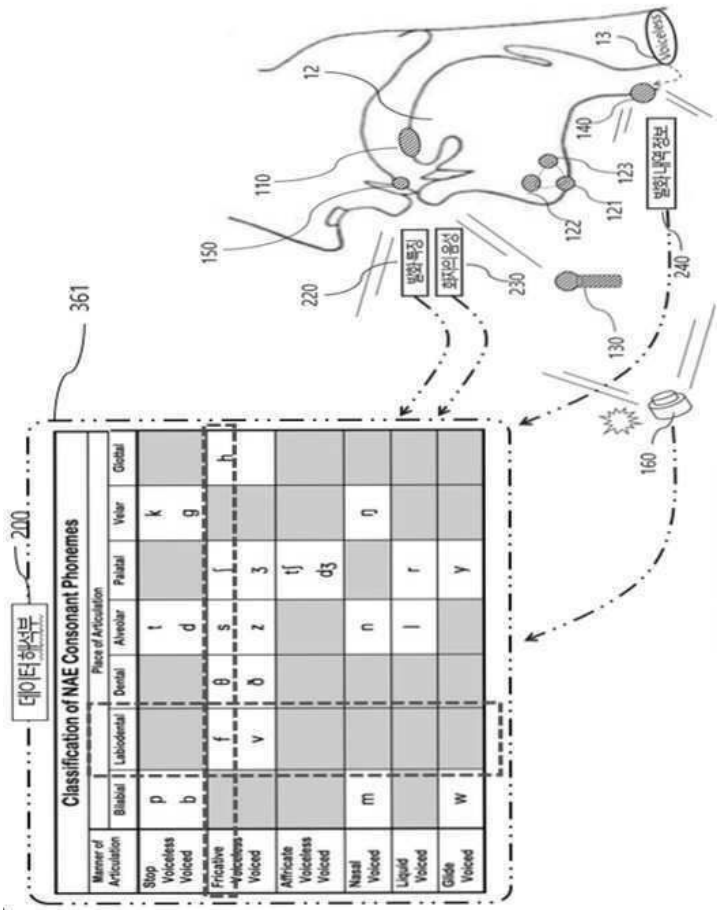
도면26



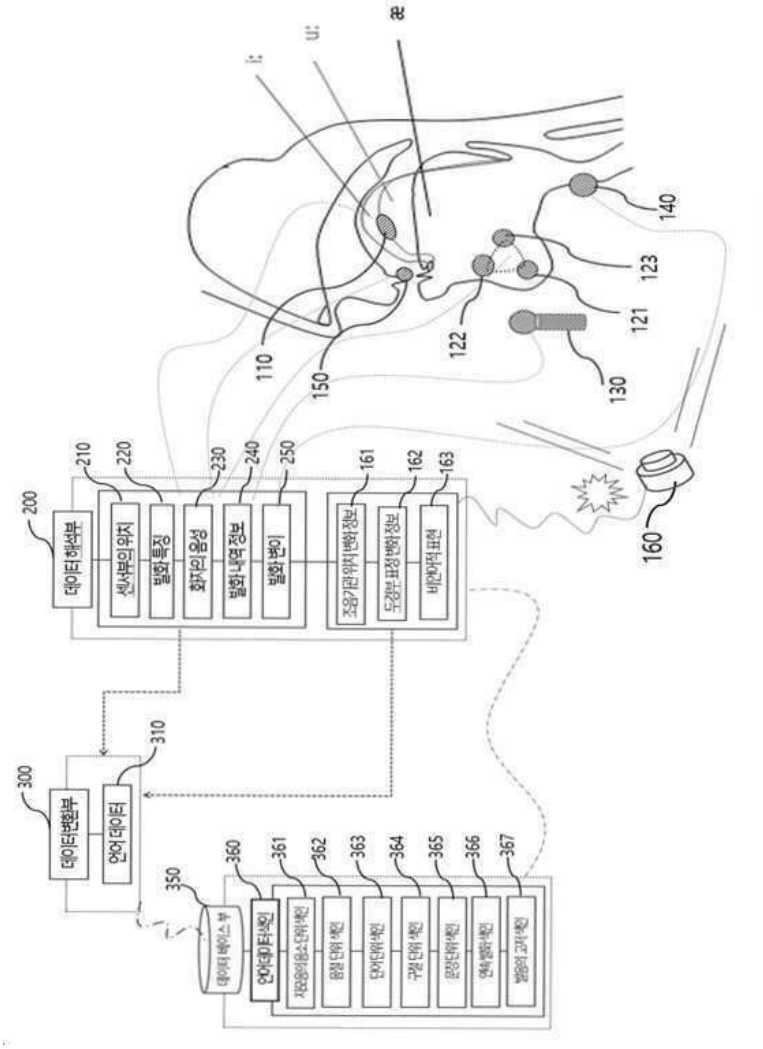
도면27


200
데이터해석부
361
Classification of NAE Consonant Phonemes
Manner of Articulation
Place of Articulation
Bilabial
Labiodental
Dental
Alveolar
Palatal
Velar
Glottal
Stop Voiceless p t k
Voiced b d g
Fricative Voiceless f θ s ʃ h
Voiced v ð z ʒ
Affricate Voiceless tʃ
Voiced dʒ
Nasal Voiced m n ŋ
Liquid Voiced l r
Glide Voiced w y
12
13
110
140
123
121
122
150
220
발화특징
화자의 음성
230
130
160
발화내역정보
240

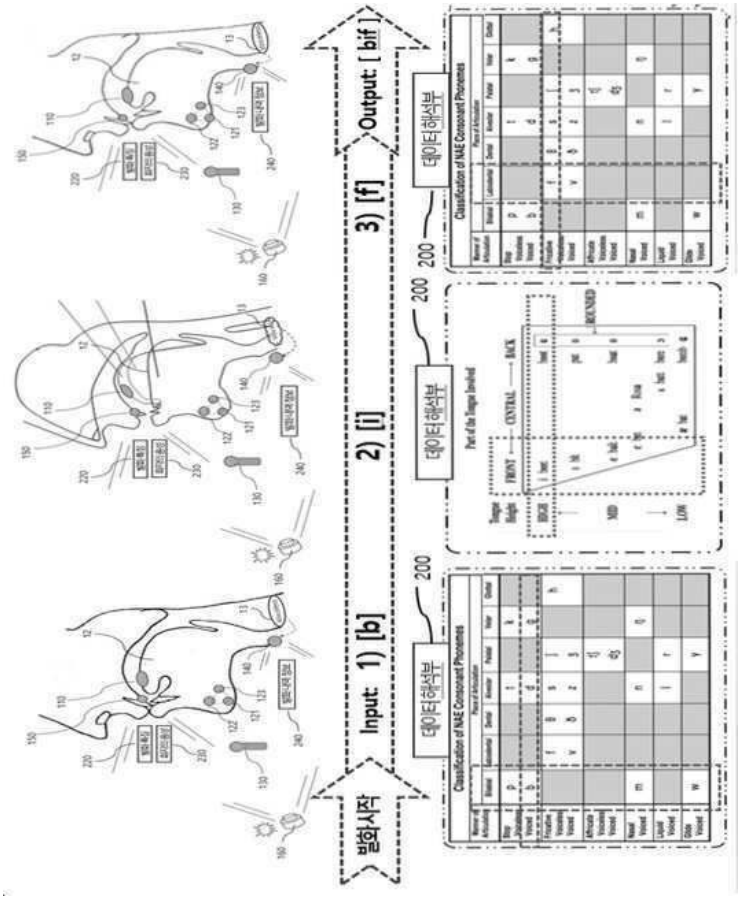
도면28


데이터해석부
200
361
Classification of NAE Consonant Phonemes
Manner of Articulation
Place of Articulation
Bilabial
Labiodental
Dental
Alveolar
Palatal
Velar
Global
Stop Voiceless p Voiced b
Fricative Voiceless f θ s ʃ h Voiced v ð z ʒ
Affricate Voiceless tʃ Voiced dʒ
Nasal Voiced m n ŋ
Liquid Voiced l r
Glide Voiced w y
12
110
150
220
발화 특징
화자의 음성
230
130
160
13
Voicelessness
140
123
121
122
발화내용 정보
240

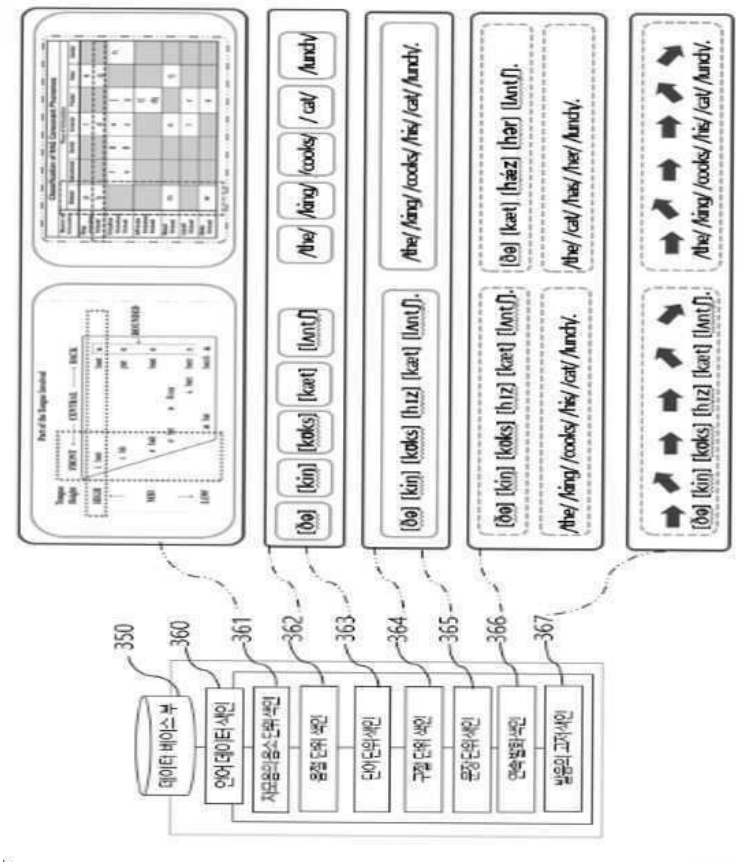
도면29


200
데이터해석부
210
센서부의 위치
220
발화특징
230
화자의 음성
240
발화 내역 정보
250
발화 변이
161
조음기관 위치 변화 정보
162
두경부 표정 변화 정보
163
비언어적 표현
300
데이터변환부
310
언어 데이터
350
데이터 베이스 부
360
언어데이터 세인
361
자연음의음소 단위 세인
362
음절 단위 세인
363
단어단위 세인
364
구절 단위 세인
365
문장단위 세인
366
연속 발화 세인
367
발음의 고저 세인
110
150
122
121
123
140
130
160

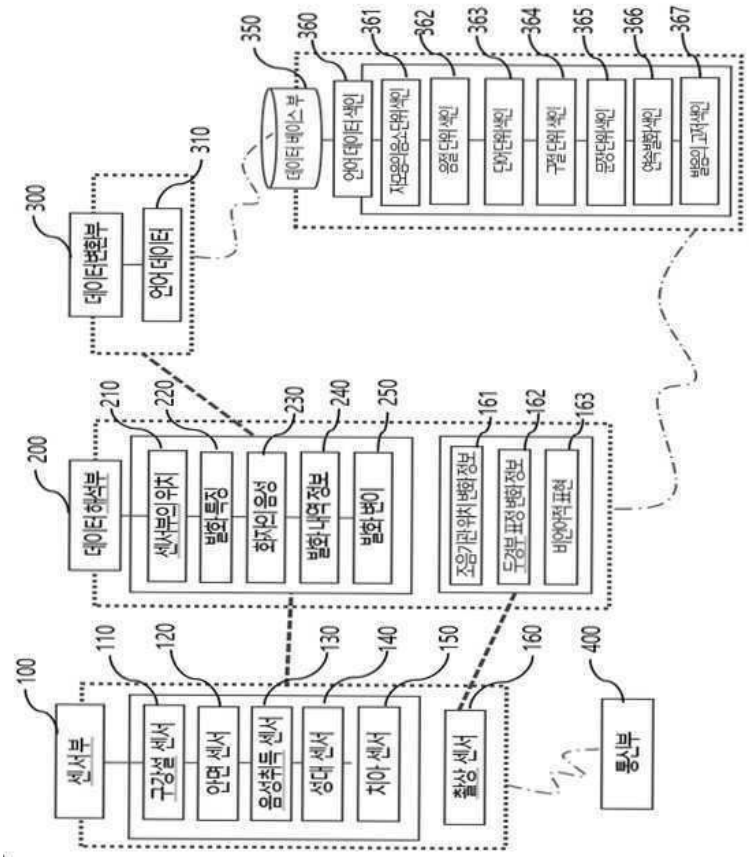
도면30



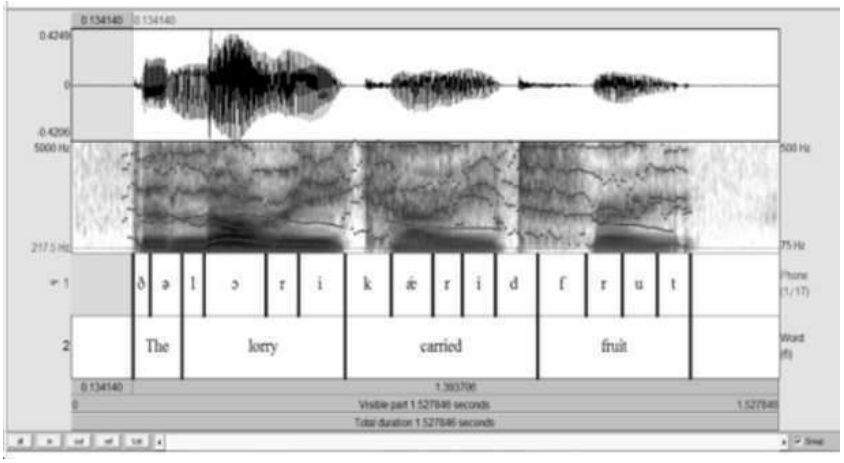
도면31


350
360
361
362
363
364
365
366
367
데이터 베이스부
언어 데이터 색인
자모음의 음소단위 색인
형태 단위 색인
단어단위색인
구절 단위 색인
문장단위색인
연속발화색인
발음의 고저색인
/the/ /king/ /cooks/ /cat/ /lunch/.
[ðə] [kiŋ] [kʊks] [kæt] [lʌnʧ].

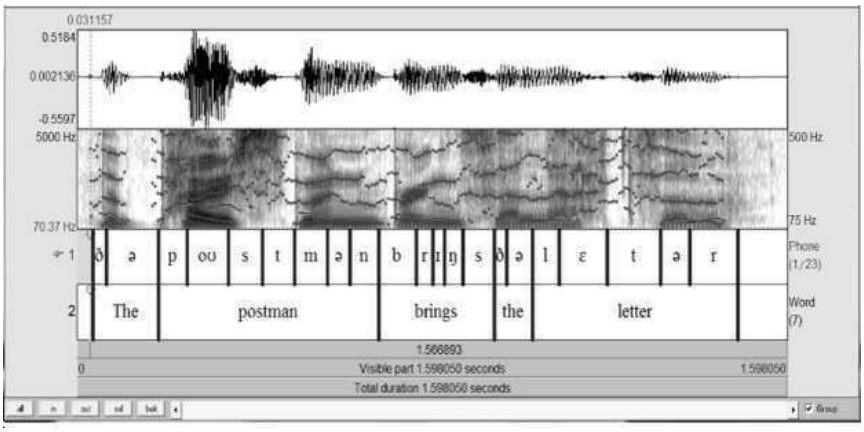
도면32


350
360
361
362
363
364
365
366
367
데이터베이스부
언어 데이터 색인
자모음/음소 단위 색인
음절 단위 색인
단어 단위 색인
구절 단위 색인
문장 단위 색인
연속발화 색인
발음의 고저색인
310
300
데이터변환부
언어 데이터
210
220
230
240
250
161
162
163
200
데이터 해석부
센서부의 위치
발화 특징
화자의 음성
발화 내역 정보
발화 변이
조음기관 위치 변화 정보
두경부 표정 변화 정보
비언어적 표현
110
120
130
140
150
160
400
100
센서부
구강설 센서
안면 센서
음성취득 센서
성대 센서
치아 센서
촬상 센서
통신부

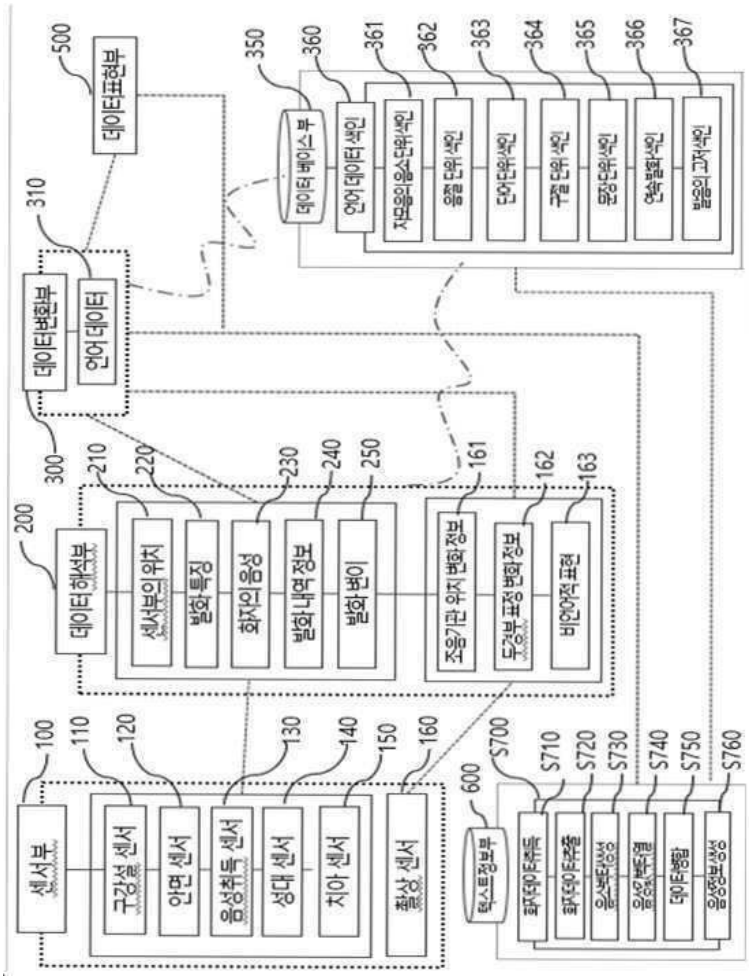
도면33


The
lorry
carried
fruit

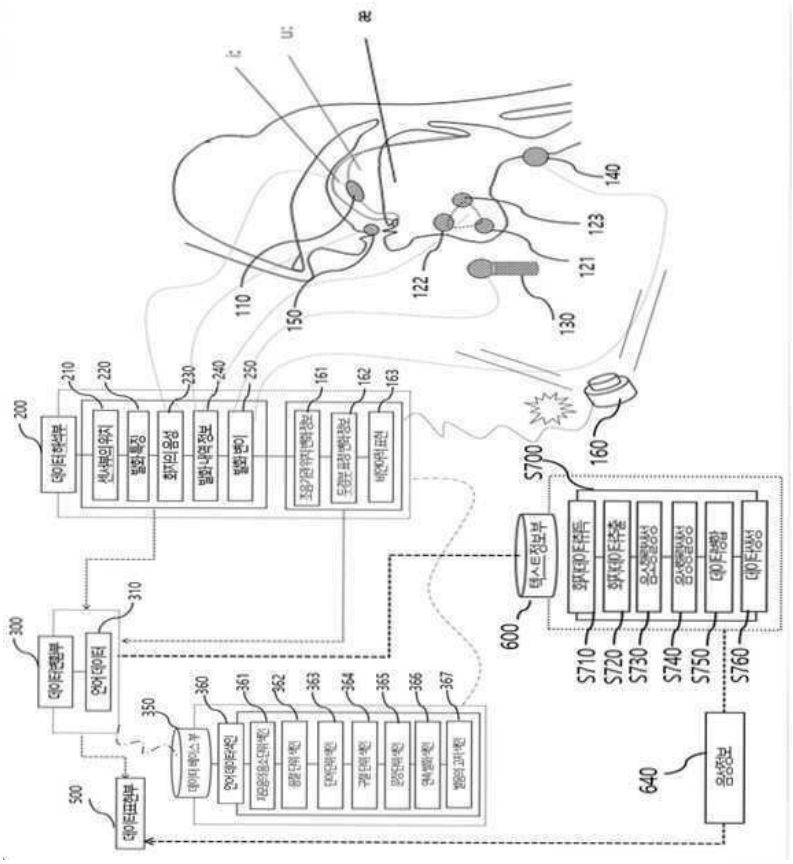
도면34


0.031157
0.5184
0.002136
-0.5597
5000 Hz
500 Hz
75 Hz
Phone (1/23)
Word (7)
ð ə p oʊ s t m æ n b r ɪ ŋ s ð ə l ɛ t ə r
The postman brings the letter
1.566893
Visible part 1.598050 seconds
Total duration 1.598050 seconds

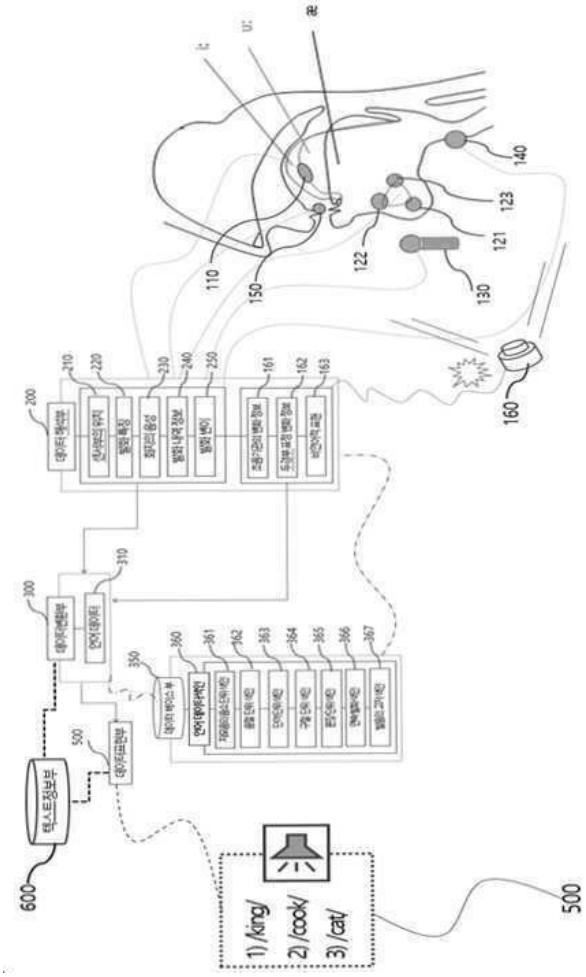
도면35


500 데이터표현부
310 데이터변환부
언어 데이터
300
200 데이터 해석부
210 센서부의 위치
220 발화 특징
230 화자의 음성
240 발화 내역 정보
250 발화 변이
161 조음기관 위치 변화정보
162 도장부 표정 변화정보
163 비언어적 표현
100 센서부
110 구강설 센서
120 안면 센서
130 음성취득 센서
140 성대 센서
150 치아 센서
160 촬상 센서
600 텍스트정보부
S700
S710
S720
S730
S740
S750
S760
350 데이터 베이스 부
360 언어 데이터 색인
361 자모음의음소단위색인
362 음절 단위색인
363 단어단위색인
364 구절 단위색인
365 문장단위색인
366 연속발화색인
367 발음의 고저색인

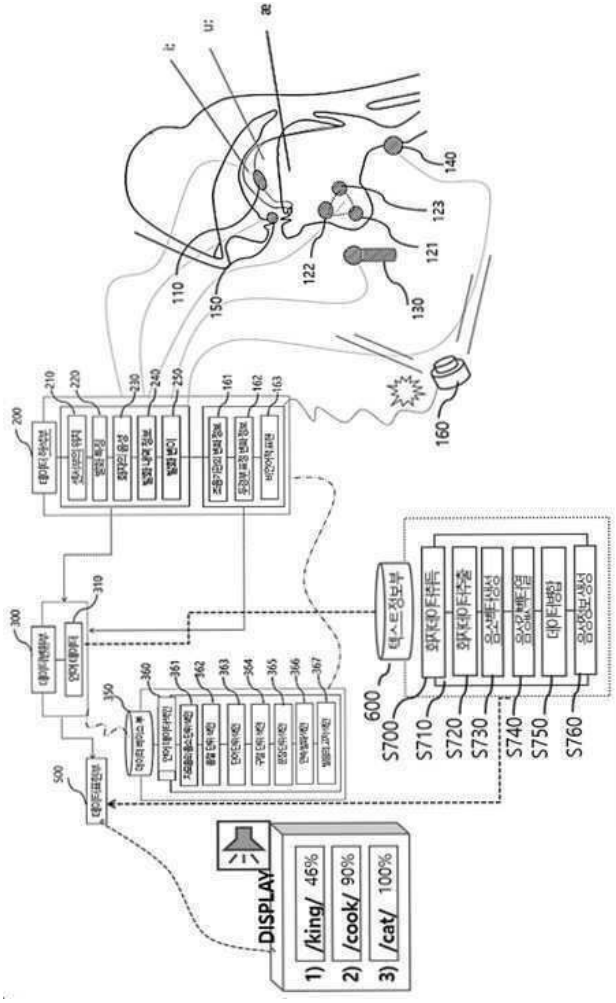
도면36


200
210
220
230
240
250
161
162
163
110
150
122
121
123
140
130
160
S700
600
S710
S720
S730
S740
S750
S760
300
310
350
360
361
362
363
364
365
366
367
500
640

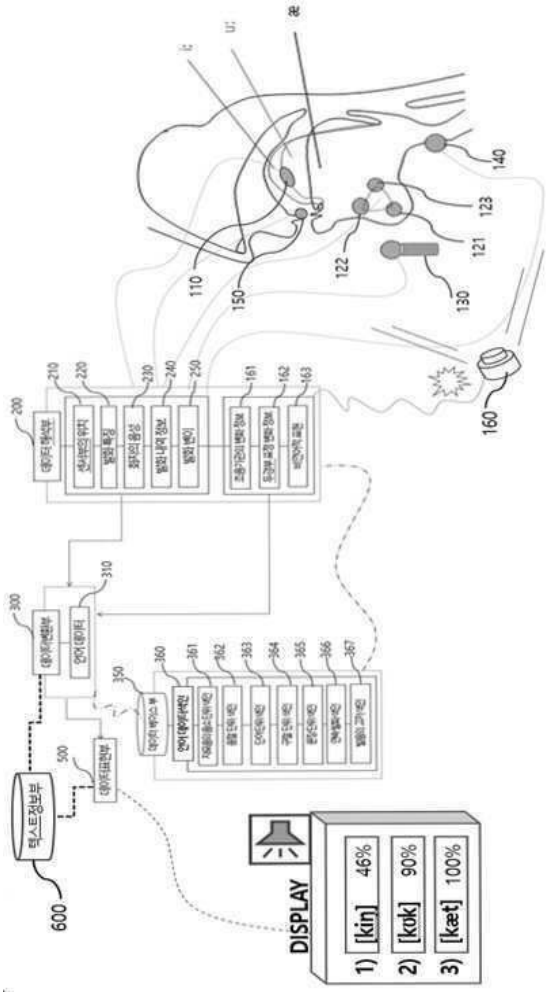
도면37


110
150
122
130
121
123
140
160
200
210
220
230
240
250
161
162
163
300
310
350
360
361
362
363
364
365
366
367
500
텍스트정보부
600
1) /king/
2) /cook/
3) /cat/
500

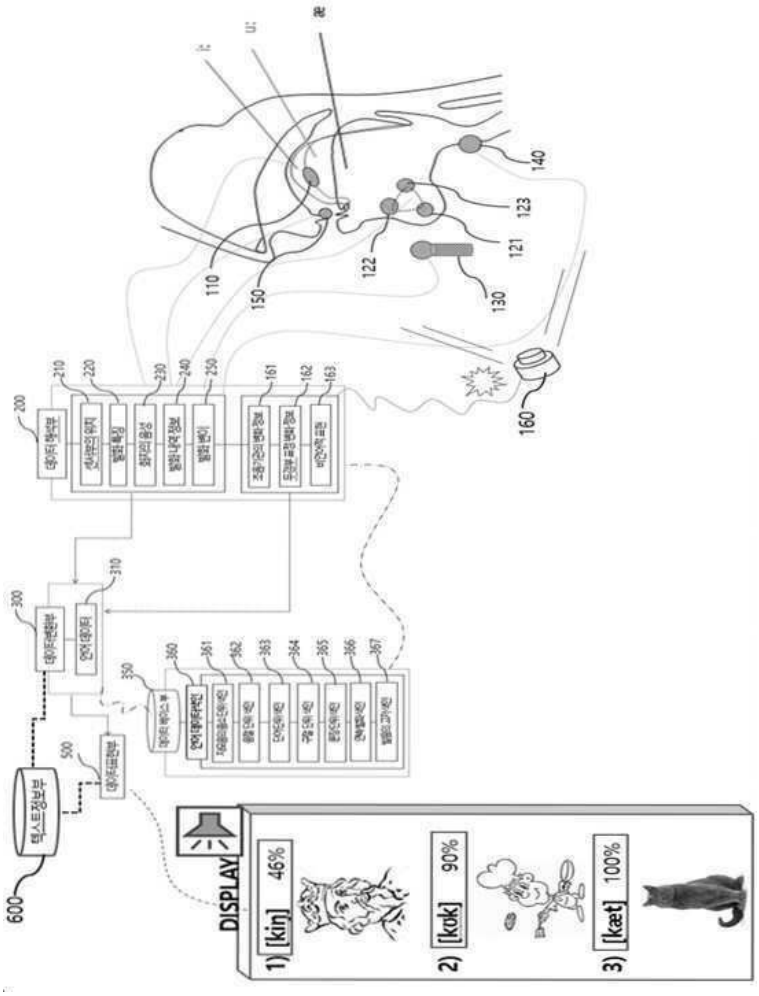
도면38


110
122
150
130
121
123
140
160
200
210
220
230
240
250
161
162
163
300
310
350
360
361
362
363
364
365
366
367
500
600
텍스트정보부
S700
S710
S720
S730
S740
S750
S760
DISPLAY
1) /king/ 46%
2) /cook/ 90%
3) /cat/ 100%

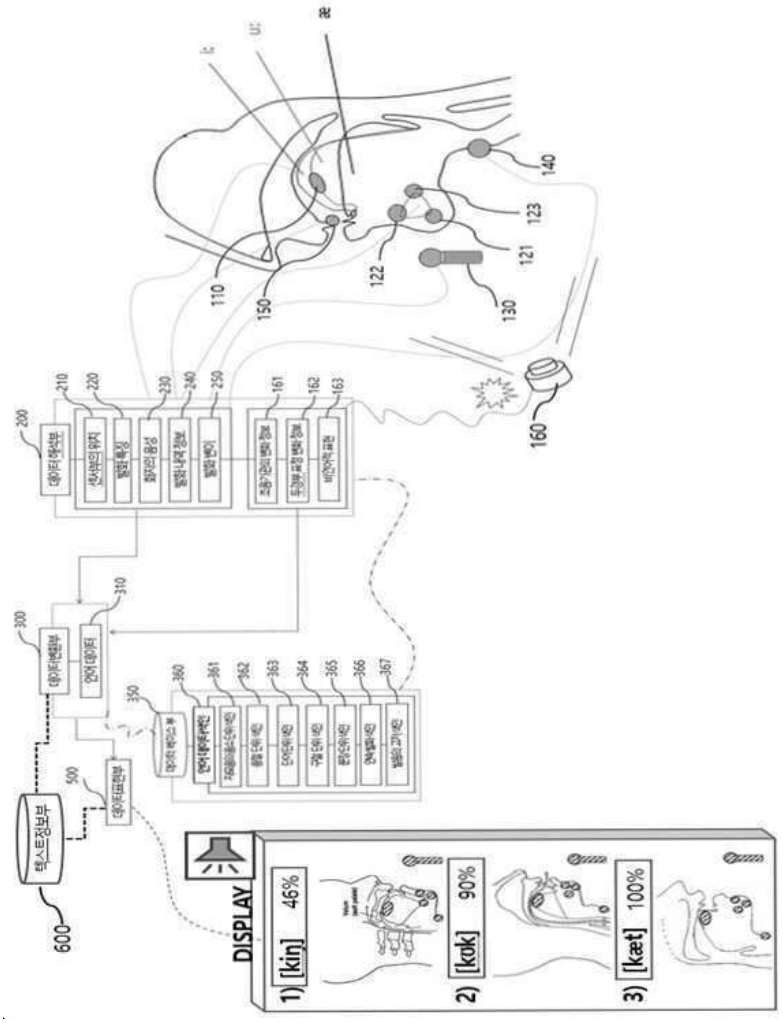
도면39


110
150
122
130
121
123
140
160
200
210
220
230
240
250
161
162
163
300
310
350
360
361
362
363
364
365
366
367
500
텍스트정보부
600
DISPLAY
1) [kin] 46%
2) [kɑk] 90%
3) [kæt] 100%

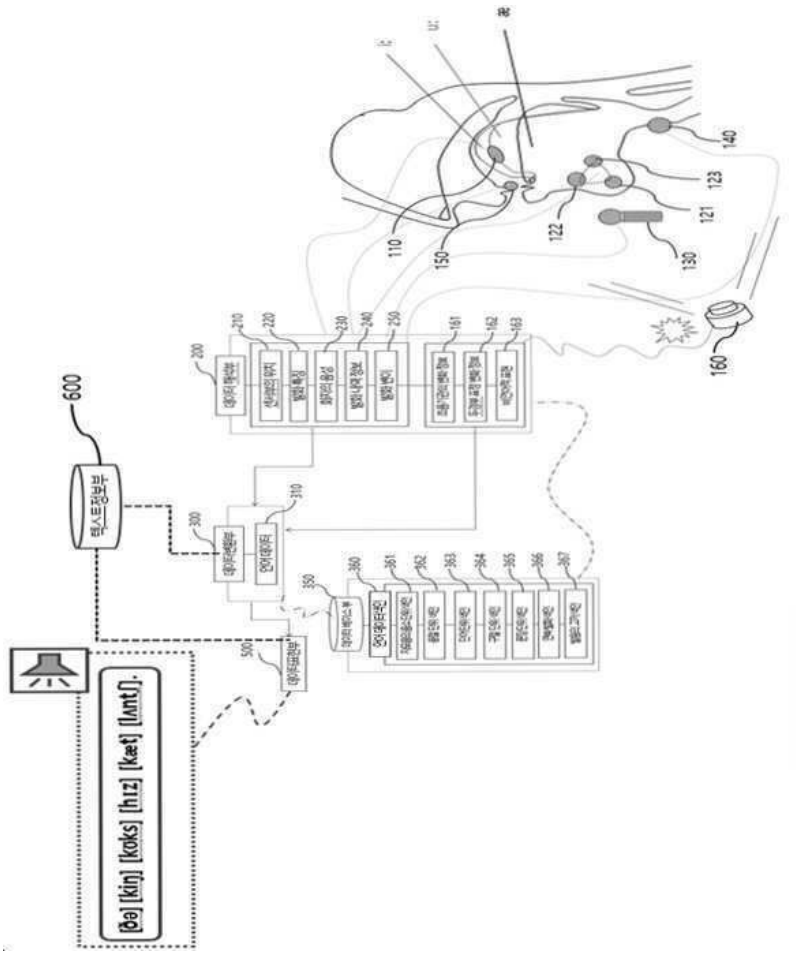
도면40



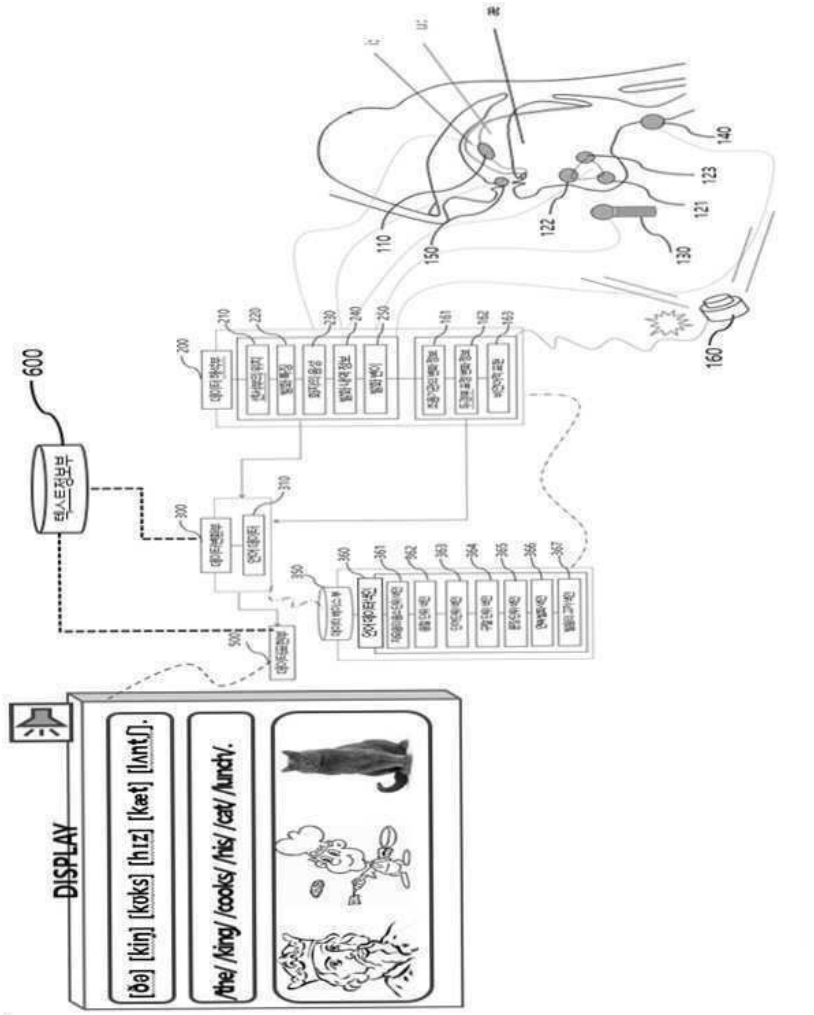
도면41



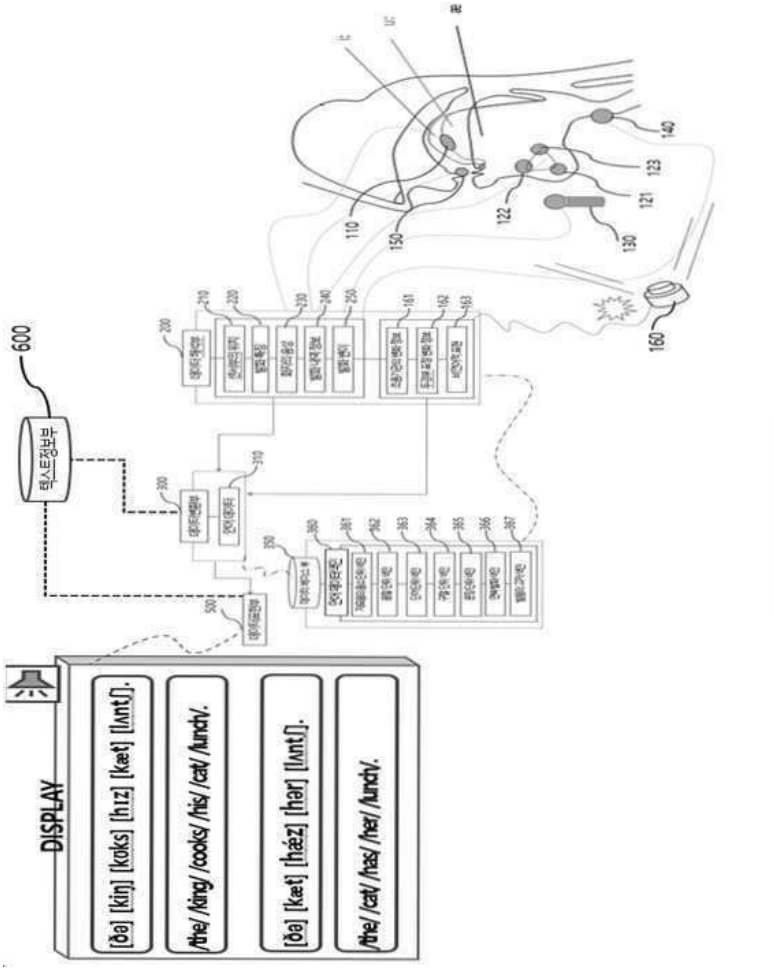
도면42



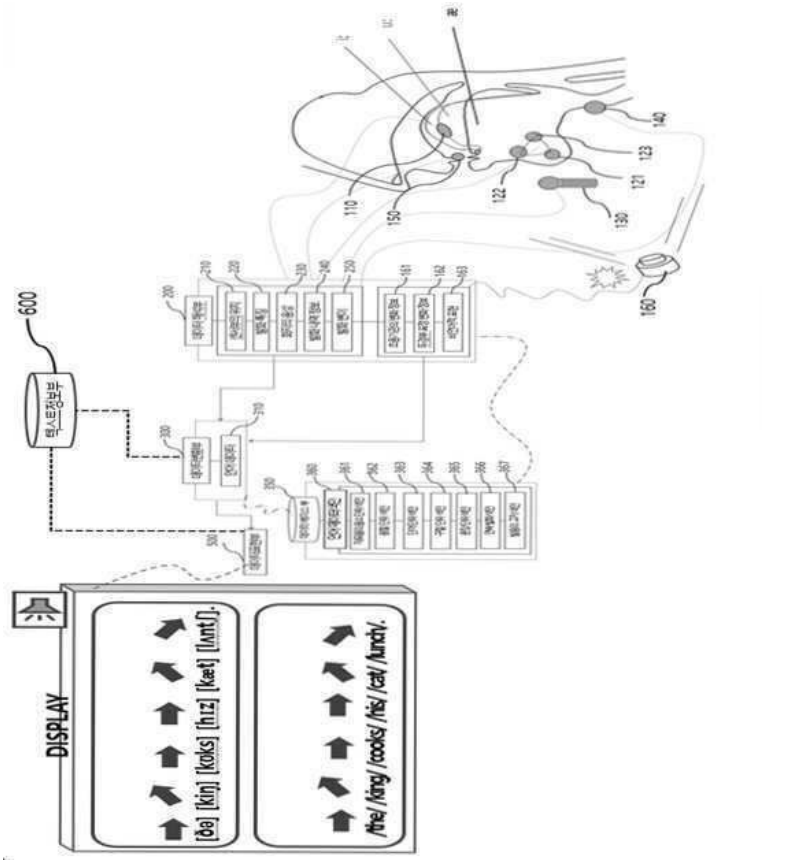
도면43


600
텍스트정보부
DISPLAY
[ðə] [kiŋ] [kʊks] [hɪz] [kæt] [lʌntʃ].
/the/ /king/ /cooks/ /his/ /cat/ /lunch/.
110
150
122
121
123
140
130
160

도면44


600
텍스트정보부
DISPLAY
[ðə] [kiŋ] [kɔks] [hɪz] [kæt] [lʌntʃ].
/the/ /king/ /cooks/ /his/ /cat/ /lunch/.
[ðə] [kæt] [hæz] [hər] [lʌntʃ].
/the/ /cat/ /has/ /her/ /lunch/.
110
121
122
123
130
140
150
160

도면45




도면46

단위: %

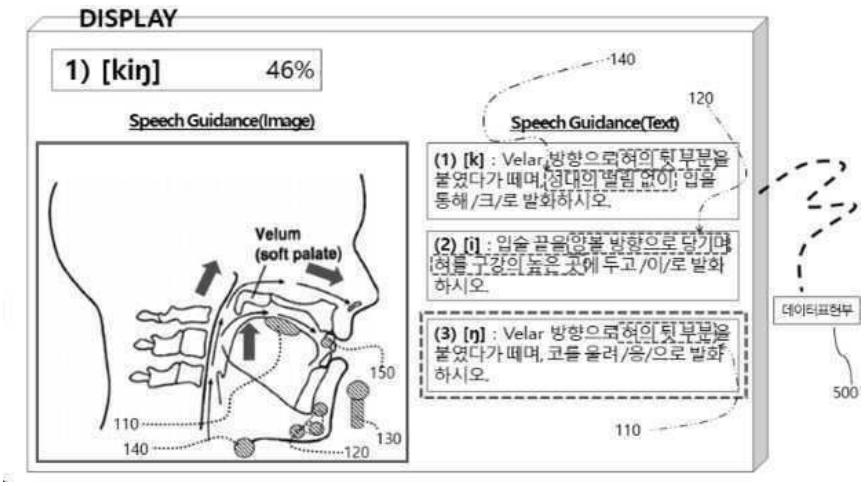
| | | | 표준 발화로 센싱된 조음기관 물리특성 | | | | |
|---|---|---|---|---|---|---|---|
| | | | 1 | 2 | 3 | 4 | 5 |
| | | | Bilabial | Alveolar | Palatal | Velar | Glottal |
| 발화된 조음기관 물리특성 | 1 | Bilabial | 80 | 5 | 3 | 5 | 7 |
| | 2 | Alveolar | 4 | 82 | 10 | 3 | 1 |
| | 3 | Palatal | 4 | 17 | 70 | 6 | 3 |
| | 4 | Velar | 3 | 5 | 9 | 81 | 2 |
| | 5 | Glottal | 9 | 4 | 4 | 8 | 75 |

도면47

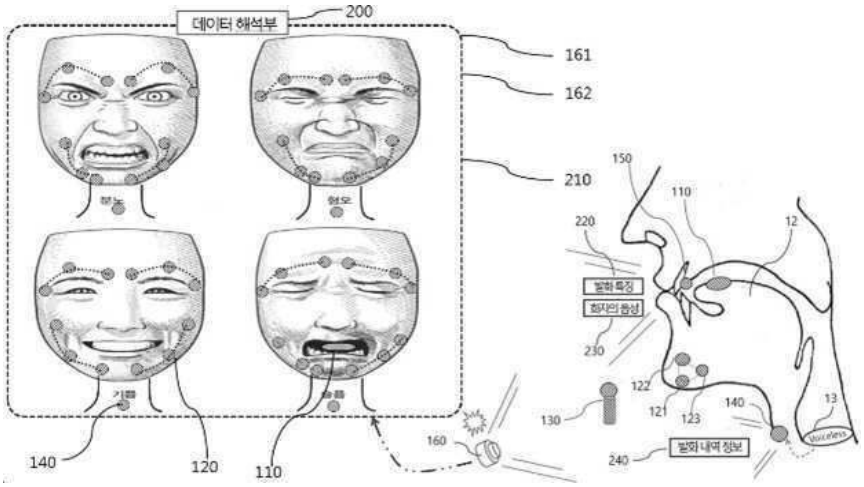
단위: %

| | | | 센싱된 조음기관 물리특성 | | | |
|---|---|---|---|---|---|---|
| | | | 1 | 2 | 3 | 4 |
| | | | True Positive | False Negative | False Positive | True Negative |
| 발화된 조음기관물리특성 | 1 | Bilabial | 80.4 | 19.6 | 15.3 | 84.7 |
| | 2 | Alveolar | 81.2 | 18.8 | 22.2 | 77.8 |
| | 3 | Palatal | 70.4 | 29.6 | 15.3 | 84.7 |
| | 4 | Velar | 81.4 | 18.6 | 19.1 | 80.9 |
| | 5 | Glottal | 74.7 | 25.3 | 21.8 | 78.2 |

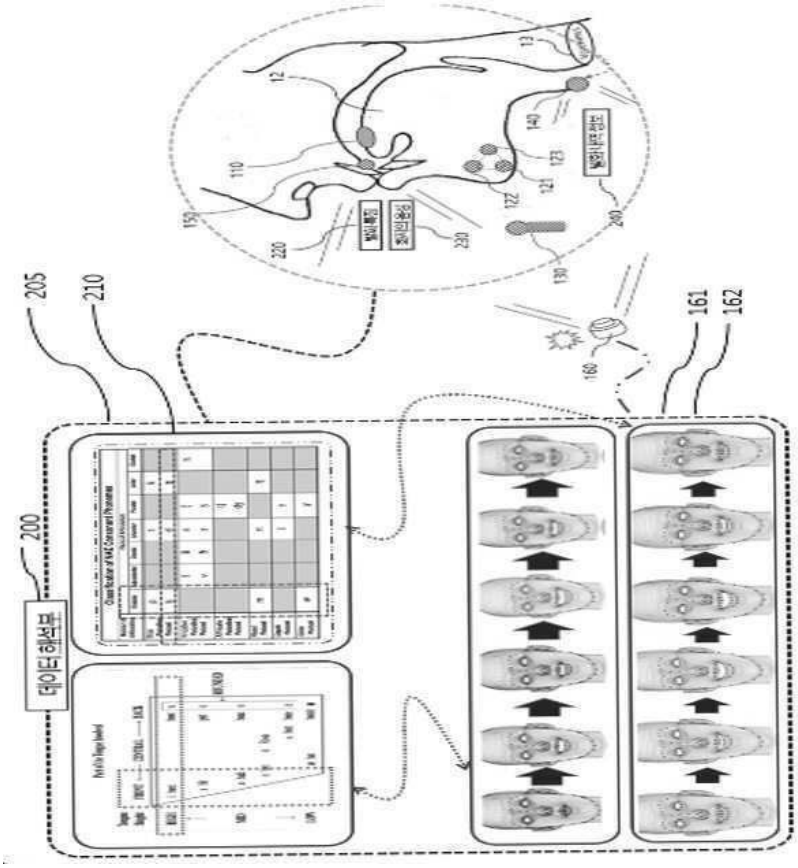
도면48



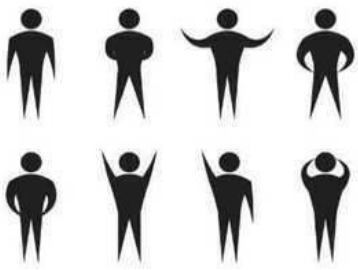

도면49

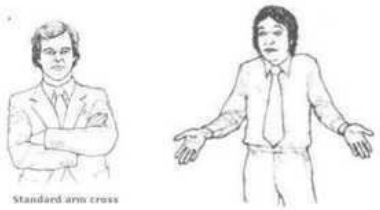


도면50


205
210
161
162
160
200
데이터베이스부

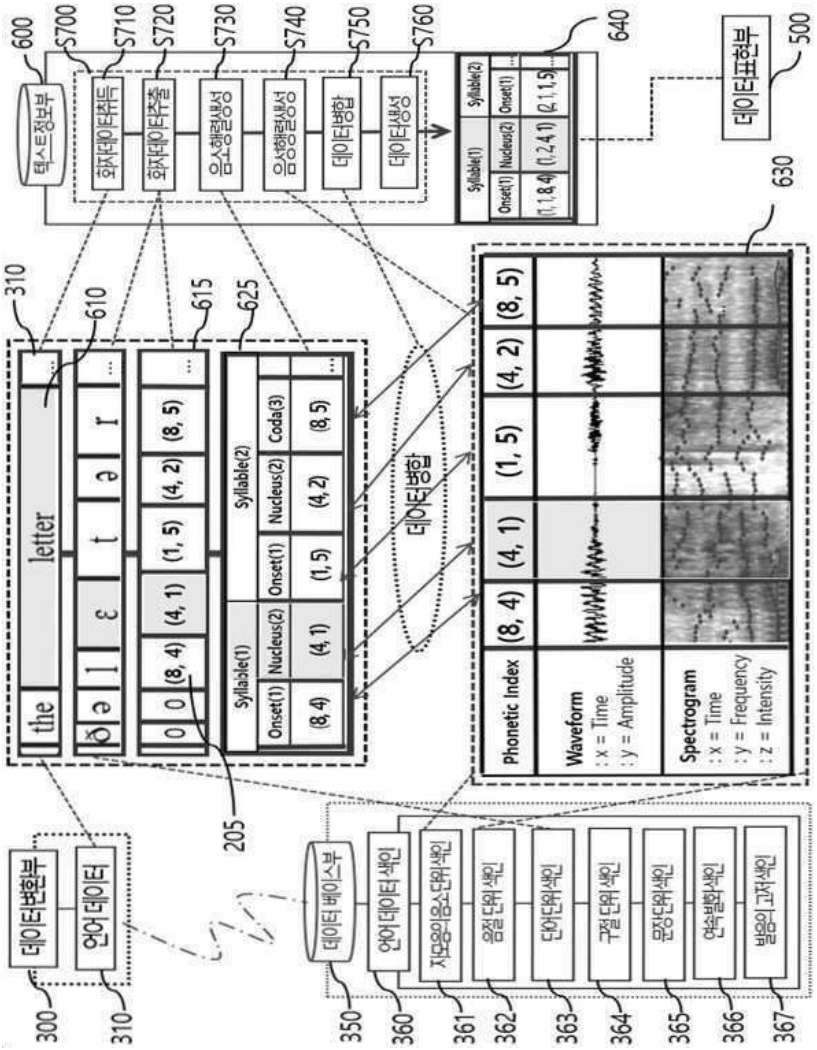
도면51




Standard arm cross


Cheek Gestures
Boredom Gesture
Interested, Evaluation Gesture

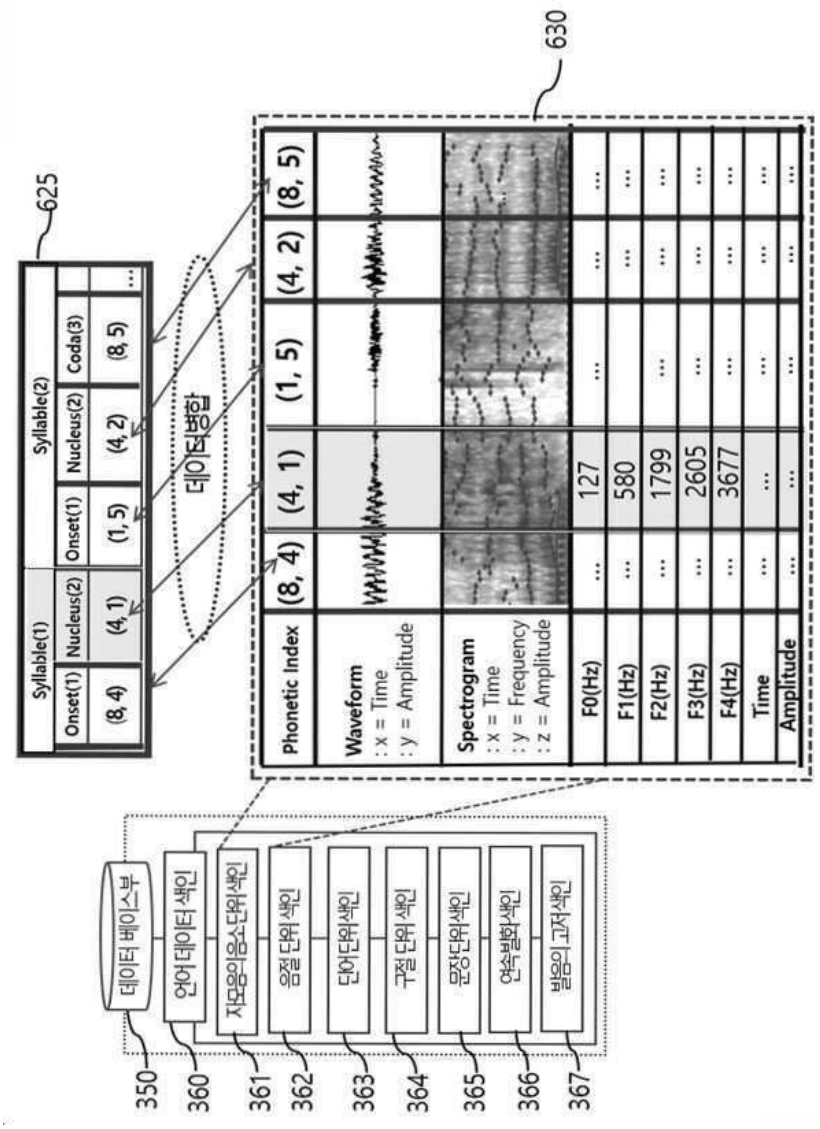
도면52


600
S700
S710
S720
S730
S740
S750
S760
640
텍스트정보부
화자데이터취득
화자데이터추출
음소해석생성
음성해석생성
데이터병합
데이터생성
데이터표현부
500
630
310
610
615
625
letter
Syllable(1)
Syllable(2)
Onset(1)
Nucleus(2)
Coda(3)
데이터병합
Phonetic Index
Waveform
: x = Time
: y = Amplitude
Spectrogram
: x = Time
: y = Frequency
: z = Intensity
205
데이터변환부
언어데이터
데이터 베이스부
언어데이터 색인
자모음/음소단위 색인
음절 단위 색인
단어단위 색인
구절 단위 색인
문장단위 색인
연속발화 색인
발음의 고저색인
300
310
350
360
361
362
363
364
365
366
367

도면53

| | | /i/ | /ɪ/ | /e/ | /ɛ/ | /æ/ | /ɑ/ | /ɔ/ | /o/ |
|---|---|---|---|---|---|---|---|---|---|
| F0 | M | 138 | 135 | 129 | 127 | 123 | 123 | 121 | 129 |
| | W | 227 | 224 | 219 | 214 | 215 | 215 | 210 | 217 |
| | C | 246 | 241 | 237 | 230 | 228 | 229 | 225 | 236 |
| F1 | M | 342 | 427 | 476 | 580 | 588 | 768 | 652 | 497 |
| | W | 437 | 483 | 536 | 731 | 669 | 936 | 781 | 555 |
| | C | 452 | 511 | 564 | 749 | 717 | 1002 | 803 | 597 |
| F2 | M | 2322 | 2034 | 2089 | 1799 | 1952 | 1333 | 997 | 910 |
| | W | 2761 | 2365 | 2530 | 2058 | 2349 | 1551 | 1136 | 1035 |
| | C | 3081 | 2552 | 2656 | 2267 | 2501 | 1688 | 1210 | 1137 |
| F3 | M | 3000 | 2684 | 2691 | 2605 | 2601 | 2522 | 2538 | 2459 |
| | W | 3372 | 3053 | 3047 | 2979 | 2972 | 2815 | 2824 | 2828 |
| | C | 3702 | 3403 | 3323 | 3310 | 3289 | 2950 | 2982 | 2987 |
| F4 | M | 3657 | 3618 | 3649 | 3677 | 3624 | 3687 | 3486 | 3384 |
| | W | 4352 | 4334 | 4319 | 4294 | 4290 | 4299 | 3923 | 3927 |
| | C | 4572 | 4575 | 4422 | 4671 | 4409 | 4307 | 3919 | 4167 |

도면54


625
630
Syllable(1)
Syllable(2)
Onset(1)
Nucleus(2)
Onset(1)
Nucleus(2)
Coda(3)
(8, 4)
(4, 1)
(1, 5)
(4, 2)
(8, 5)
데이터병합
Phonetic Index
Waveform
: x = Time
: y = Amplitude
Spectrogram
: x = Time
: y = Frequency
: z = Amplitude
F0(Hz)
F1(Hz)
F2(Hz)
F3(Hz)
F4(Hz)
Time
Amplitude
127
580
1799
2605
3677
데이터 베이스부
언어 데이터 색인
자모음/음소단위색인
음절 단위 색인
단어단위색인
구절 단위색인
문장단위색인
연속발화색인
발음의 고저색인
350
360
361
362
363
364
365
366
367