

명세서

청구범위

청구항 1

외국어 교육용 인공지능 기능을 구비한 사용자 기기에 있어서,

사용자의 음성을 입력받는 음성입력장치;

상기 사용자의 음성을 문자로 인식하여 처리하는 프로세서; 및

상기 프로세서의 처리에 따라 사용자에게 음성 반응을 제공하는 스피커를 포함하며,

상기 프로세서는,

(a) 상기 사용자의 음성의 오류가 발음상 오류인지, 또는 상기 오류가 문법상 오류인지를 판단하며,

(b) 상기 오류가 발음상 오류인 경우, 상기 오류가 있는 부분을 대체할 수 있는 표현을 포함하는 제 1 음성 반응을 상기 스피커를 통해 제공하며,

(c) 상기 오류가 문법상 오류인 경우, 상기 오류가 있는 부분에 대응하는 기 마련된 문법 설명을 포함하는 제 2 음성 반응을 상기 스피커를 통해 제공하며,

상기 사용자의 음성은 상기 사용자 기기의 동작 모드를 변경하기 위한 명령어를 포함하며,

상기 명령어에 따라 상기 프로세서는 (1) 1 문장 단위로 상기 사용자의 음성을 처리하여 상기 음성 반응을 상기 스피커에 제공하는 문답 모드, 또는 (2) 소정 시간 상기 사용자의 음성을 지속적으로 기록한 후 상기 음성 반응을 상기 스피커에 제공하는 장문 기록 모드로 선택적으로 동작하는, 사용자 기기.

청구항 2

제 1 항에 있어서,

상기 프로세서는,

상기 사용자의 음성을 문자로 인식하기 위해 사용되는 제 1 문법 모델, 및

상기 사용자의 음성을 문자로 인식한 후, 문자 인식된 언어에 존재하는 문법상 오류를 인식하기 위해 사용되는 제 2 문법 모델을 사용하는, 사용자 기기.

청구항 3

제 2 항에 있어서,

상기 제 2 문법 모델은 웹 서버의 데이터에 기반하여 학습 기반으로 각 문법상 오류 및 각 문법상 오류에 대응하는 문법 설명을 쌍으로써 갱신하도록 구성되는, 사용자 기기.

청구항 4

제 1 항에 있어서,

상기 프로세서는,

상기 오류가 없는 경우, 문자로 인식된 사용자의 언어의 내용과 대화하는 내용에 대응하는 제 3 음성 반응을 상기 스피커를 통해 제공하는, 사용자 기기.

청구항 5

삭제

청구항 6

삭제

청구항 7

삭제

청구항 8

삭제

발명의 설명

기술 분야

[0001] 본 발명은 외국어 교육용 인공지능 기능을 구비하여 사용자 음성의 문법 오류/발음 오류에 따른 음성 반응을 제공하는 사용자 기기, 및 사용자 기기의 온라인 접속을 통해 외국어 교육을 제공하는 방법에 대한 것이다.

배경 기술

[0002] 최근 인공지능(AI) 스피커의 보급으로 사용자의 음성인식을 통한 대내 가전기기들을 제어하고, 보다 발전적으로는 인공지능 캐릭터와 대화를 진행하는 서비스가 제공되고 있다.

[0003] 일부 사용자는 이러한 AI 스피커에 외국어로 명령어를 입력하여 제어함으로써 외국어 학습에도 활용하고 있다. 다만, 현재의 AI 스피커는 사용자의 외국어 명령어를 인식하여 이에 따른 기기 제어 또는 반응을 할 뿐, 보다 높은 수준의 외국어 교육에 활용되지는 못하는 상황이다.

[0004] 또한, 온라인을 통해 사용자의 음성을 인식하여 발음을 교육해 주는 서비스도 제공되고 있다. 다만, 이러한 서비스는 사용자의 음성에 대응하는 보다 정확한 발음을 제공할 뿐, 사용자의 외국어 음성에 문법적인 문제가 있는지, 발음상 문제가 있는지에 대한 구분을 제공하지 못하고 있다.

발명의 내용

해결하려는 과제

[0005] 상술한 바와 같은 문제를 해결하기 위해 본 발명의 일 실시형태에서는 사용자의 외국어 음성에 오류가 있는지, 오류가 있다면 발음상의 문제인지, 문법상의 문제인지를 효율적으로 구분하여, 이러한 구분에 따라 적절한 교육 서비스를 제공하고자 한다.

[0006] 또한, 본 발명의 일 실시형태에서는 상술한 사용자 음성의 구분에 기초하여 사용자에게 효율적인 교육 서비스를 제공하기 위한 인터페이스를 제공하고자 한다.

[0007] 본 발명의 목적은 상술한 목적에 한정되지 않으며, 이하의 설명을 통해 다양한 목적이 제시된다.

과제의 해결 수단

[0008] 상술한 바와 같은 과제를 해결하기 위한 본 발명의 일 측면에서는, 외국어 교육용 인공지능 기능을 구비한 사용자 기기에 있어서, 사용자의 음성을 입력받는 음성입력장치; 상기 사용자 음성을 문자로 인식하여 처리하는 프로세서; 및 상기 프로세서의 처리에 따라 사용자에게 음성 반응을 제공하는 스피커를 포함하며, 상기 프로세서는, (a) 상기 사용자의 음성을 문자로 인식하는 과정에서 인식되는 오류가 있는지, 상기 오류가 발음상 오류인지, 또는 상기 오류가 문법상 오류인지를 판단하며, (b) 상기 오류가 발음상 오류인 경우, 상기 오류가 있는 부분을 대체할 수 있는 표현을 포함하는 제 1 음성 반응을 상기 스피커를 통해 제공하며, (c) 상기 오류가 문법상 오류인 경우, 상기 오류가 있는 부분에 대응하는 기 마련된 문법 설명을 포함하는 제 2 음성 반응을 상기 스피커를 통해 제공하는, 사용자 기기를 제안한다.

[0009] 여기서, 상기 프로세서는, 상기 사용자의 음성을 문자로 인식하기 위해 사용되는 제 1 문법 모델, 및 상기 사용자의 음성을 문자로 인식한 후, 문자 인식된 언어에 존재하는 문법상 오류를 인식하기 위해 사용되는 제 2 문법 모델을 사용할 수 있다.

[0010] 상기 제 2 문법 모델은 웹 서버의 데이터에 기반하여 학습 기반으로 각 문법상 오류 및 각 문법상 오류에 대응

하는 문법 설명을 쌍으로써 갱신하도록 구성될 수 있다.

- [0011] 상기 프로세서는, 상기 오류가 없는 경우, 문자로 인식된 사용자의 언어의 내용과 대화하는 내용에 대응하는 제 3 음성 반응을 상기 스피커를 통해 제공할 수 있다.
- [0012] 상기 프로세서는, 상기 사용자의 음성에서 특징을 추출한 후, 상기 특징에 대응하는 후보 발음열을 검출하고, 상기 후보 발음열 중 제 1 문법 모델을 이용하여 소정 확률 이상으로 조합가능한 언어를 선택하여, 상기 사용자의 음성을 문자로 인식할 수 있다.
- [0013] 상기 프로세서는, 상기 후보 발음열 중 상기 제 1 문법 모델을 이용하여 조합가능한 언어를 선택하지 못하는 경우, 상기 오류가 발음상 오류인 것으로 판단할 수 있다.
- [0014] 상기 프로세서는, 상기 후보 발음열 중 상기 제 1 문법 모델을 이용하여 조합가능한 언어를 선택하되, 상기 후보 발음열 중 최상위 확률로 인식되는 문자와 상기 선택된 언어 사이에 차이가 소정 기준 이상인 경우, 상기 오류가 발음상 오류인 것으로 판단할 수 있다.
- [0015] 상기 프로세서는, 상기 (c) 동작에 있어서, 상기 오류가 있는 부분을 대체할 수 있는 표현에 대한 사용자의 응답을 상기 음성입력장치를 통해 수신하고, 상기 사용자의 응답에 따라 상기 사용자의 음성을 문자로 인식할 수 있다.
- [0016] 상기 프로세서는, 상기 사용자 음성 중 모국어 음성과 외국어 음성을 구분하며, 상기 사용자 음성이 상기 외국어 음성으로 판단되는 경우, 상기 (a) 내지 상기 (c)의 동작을 수행할 수 있다.
- [0017] 상기 사용자의 음성은 상기 사용자 기기의 동작 모드를 변경하기 위한 명령어를 포함할 수 있으며, 상기 명령어에 따라 상기 프로세서는 (1) 1 문장 단위로 상기 사용자의 음성을 처리하여 상기 음성 반응을 상기 스피커에 제공하는 문답 모드, 또는 (2) 소정 시간 상기 사용자의 음성을 지속적으로 기록한 후 상기 음성 반응을 상기 스피커에 제공하는 장문 기록 모드로 선택적으로 동작할 수 있다.
- [0018] 상기 사용자 기기는 IoT (Internet of Things) 방식으로 하나 이상의 태내 기기와 무선 연결될 수 있으며, 상기 사용자 음성 중 모국어 음성 또는 외국어 음성 중 하나 이상을 통해 상기 태내 기기를 제어할 수 있다.
- [0019] 한편, 상술한 문제를 해결하기 위한 본 발명의 다른 일 측면에서는 사용자 기기의 온라인 접속을 통해 외국어 교육을 제공하는 방법에 있어서, 상기 사용자 기기의 마이크를 통해 입력된 사용자의 외국어 문장 음성 발음을 인식하여 문자로 변환하고, 상기 문자로 변환된 문장의 오류를 판단하며, 상기 오류가 있는 부분이 문법상 오류인 경우, 상기 오류가 있는 부분에 대응하는 기 마련된 문법 설명을 상기 오류가 있는 부분의 강조 표시와 함께 상기 사용자 기기의 표시장치에 표시하며, 상기 오류가 있는 부분이 발음상 오류인 경우, 상기 오류가 있는 부분을 대체할 수 있는 표현을 상기 오류가 있는 부분의 강조 표시와 함께 상기 사용자 기기의 표시장치에 표시하여 추천하고, 상기 추천 표현을 포함하는 예시적인 문장의 추천 발음을 상기 사용자 기기의 스피커를 통해 제시하는 것을 포함하는, 외국어 교육 제공 방법을 제공한다.
- [0020] 상기 오류가 있는 부분을 대체할 수 있는 표현을 추천하는 것은, 복수의 표현을 추천하여 상기 사용자 기기의 표시장치에 표시하고, 상기 사용자 기기의 입력 장치를 통해 선택된 표현을 상기 추천 표현으로 결정하며, 상기 사용자의 외국어 문장 음성 발음 중 오류가 있는 부분과 상기 결정된 추천 표현 사이의 맵핑 관계를 상기 외국어 교육을 제공하는 서버의 데이터 베이스에 저장하여 인공지능 기반으로 상기 서버를 학습시키는 것을 포함할 수 있다.
- [0021] 상기 오류가 있는 부분과 상기 기 마련된 문법 설명의 대응 데이터는 인공지능 방식으로 상기 외국어 교육을 제공하는 서버의 데이터 베이스에 증가시키며, 상기 오류가 있는 부분이 문법상 오류인 것으로 판단되는 경우는, 상기 오류가 있는 부분과 대응하는 기 마련된 문법 설명이 존재하는 경우일 수 있다.

발명의 효과

- [0022] 상술한 바와 같은 본 발명의 실시예들에 따르면, 사용자의 외국어 음성에 오류가 있는지, 오류가 있다면 발음상의 문제인지, 문법상의 문제인지를 효율적으로 구분하여, 이러한 구분에 따라 적절한 교육 서비스를 제공할 수 있다.
- [0023] 또한, 본 발명의 일 실시예에 따르면, 상술한 사용자 음성의 구분에 기초하여 사용자에게 효율적인 교육 서비스를 제공하기 위한 인터페이스를 제공할 수 있다.

[0024] 본 발명에 따른 효과는 상술한 효과에 한정되지 않으며, 이하의 설명을 통해 다양한 효과가 제시된다.

도면의 간단한 설명

- [0025] 도 1은 본 발명의 일 실시예에 따른 사용자 기기의 개념을 설명하기 위한 도면이다.
- 도 2는 본 발명의 일 실시예에 따른 사용자 기기의 동작 방식을 설명하기 위한 도면이다.
- 도 3은 본 발명의 일 실시예에 따라 사용자 음성을 인식하는 과정을 구체적으로 설명하기 위한 도면이다.
- 도 4는 본 발명의 일 실시예에 따른 사용자 기기를 통해 태내 기기를 제어하는 개념을 설명하기 위한 도면이다.
- 도 5는 본 발명의 일 실시예에 따라 외국어 교육을 제공하는 방법을 설명하기 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0026] 이하 본 발명에 대한 설명을 도면을 참조하여 상세히 설명한다.
- [0027] 도 1은 본 발명의 일 실시예에 따른 사용자 기기의 개념을 설명하기 위한 도면이다.
- [0028] 도 1을 참조하면, 본 실시예에 따른 외국어 교육용 인공지능 기능을 구비한 사용자 기기는 AI 스피커(10)의 형태를 가질 수 있다. 또한, 이러한 AI 스피커(10)는 스마트폰(20a), 노트북(20b) 등 다양한 소프트웨어 처리 기능을 구비한 사용자 기기와 무선 또는 유선으로 연결될 수 있다.
- [0029] 이하의 설명에서 사용자 기기는 AI 스피커(10) 자체에 내장될 수 있으며, 이와 달리 사용자와 음성 인터페이스를 제공하기 위한 음성입력장치 및 스피커 기능은 AI 스피커(10)에, 그리고 이러한 정보를 처리하는 프로세서는 AI 스피커(10)와 연결된 다른 사용자 기기(20a 또는 20b)에 구비될 수도 있다.
- [0030] 구체적으로, 본 실시예에 따른 사용자 기기는 사용자의 음성을 입력받는 음성입력장치; 상기 사용자 음성을 문자로 인식하여 처리하는 프로세서; 및 상기 프로세서의 처리에 따라 사용자에게 음성 반응을 제공하는 스피커를 포함할 수 있다. 여기서 사용자 음성은 모국어 음성 신호일 수도, 외국어 음성 신호일 수도 있으며, 이에 대해서는 후술한다.
- [0031] 음성입력장치는 도 1에 도시된 AI 스피커(10)에 구비되는 것이 바람직하지만, 다른 사용자 기기(20a, 20b 등)에 구비될 수도 있다.
- [0032] 도 2는 본 발명의 일 실시예에 따른 사용자 기기의 동작 방식을 설명하기 위한 도면이다.
- [0033] 본 실시예에 따른 프로세서는 상기 사용자의 음성을 문자로 인식하는 과정에서 인식되는 오류가 있는지 여부를 먼저 판단할 수 있다 (S310). 만일, 오류가 있다면 프로세서는 상기 오류가 발음상 오류인지, 또는 상기 오류가 문법상 오류인지를 추가적으로 판단할 수 있다(S320).
- [0034] 만일, 상기 오류가 발음상 오류인 경우, 프로세서는 상기 오류가 있는 부분을 대체할 수 있는 표현을 포함하는 제 1 음성 반응을 상기 스피커를 통해 제공할 수 있다. 예를 들어, 사용자가 “Today, I'm gonna finish this patent specification drafting.” 이라는 외국어 문장을 의도하고 사용자 기기를 향하여 이야기하였으나, 인식된 언어에서는 “Today, I'm gonna finish this patent specticet drafting” 이라고 인지되어, 'specification'이라는 단어의 발음 문제로 판단되는 경우, 본 실시예에 따른 사용자 기기는 오류가 있는 표현을 대체할 수 있는 표현, 즉 'specification'을 포함하는 제 1 음성 반응으로서 “Do you meant 'patent specification ?'”이라는 음성 반응을 스피커를 통해 제공할 수 있다.
- [0035] 일 실시예에서는, 이러한 스피커 음성, 즉 상기 오류가 있는 부분을 대체할 수 있는 표현에 대한 사용자의 응답으로서 'Yes, you are right' 등의 응답을 상기 음성입력장치를 통해 수신하고, 상기 사용자의 응답에 따라 상기 사용자의 음성을 문자로 인식할 수 있다. 이후 본 실시예에 따른 사용자 장치는 후술하는 제 3 음성반응을 스피커를 통해 제공할 수 있다.
- [0036] 한편, 상기 오류가 문법상 오류인 경우, 상기 오류가 있는 부분에 대응하는 기 마련된 문법 설명을 포함하는 제 2 음성 반응을 상기 스피커를 통해 제공할 수 있다. 예를 들어, 사용자가 “Today, I gonna finish this patent specification drafting” 이라는 외국어 문장을 입력한 경우, 후술하는 문법 오류 모델에 기반하여 사용자 음성에 문법상 오류가 있음을 인지하고, 이에 따라 제 2 음성 반응을 제공할 수 있다. 이러한 경우, 사용자 기기는 사용자가 의도한 문장의 의도를 충분히 알 수 있기 때문에 후술하는 오류 없는 경우에 대한 제 3 음성

반응을 포함하여 제 2 음성 반응을 제공할 수 있다.

- [0037] 상술한 예에서, 사용자 기기의 프로세서는 “Good luck Aaron, go for it ! But, you'd better put 'be verb' in front of 'gonna'” 와 같은 제 2 음성 반응을 제공할 수 있다.
- [0038] 한편, 단계 S310에서 사용자의 음성에 오류가 없는 경우, 본 실시예에 따른 사용자 기기는 문자로 인식된 사용자의 언어의 내용과 대화하는 내용에 대응하는 제 3 음성 반응을 상기 스피커를 통해 제공할 수 있다. 최근 AI 기술의 발달로 본 실시예에 따른 사용자 기기는 사용자와 자연어 대화를 진행할 수 있으며, 후술하는 바와 같이 IoT (Internet of Things) 방식으로 무선 연결된 태내 가전기기들을 제어할 수도 있다.
- [0039] 도 3은 본 발명의 일 실시예에 따라 사용자 음성을 인식하는 과정을 구체적으로 설명하기 위한 도면이다.
- [0040] 도 3을 참조하면, 음성 인식을 수행하는 디바이스(100)는 특징 추출부(110), 후보 발음열 검출부(120) 및 언어 선택부(140)를 포함할 수 있다. 특징 추출부(110)는 입력된 음성 신호에 대한 특징 정보를 추출한다. 후보 발음열 검출부(120)는 추출된 특징 정보로부터 적어도 하나의 후보 발음열을 검출한다. 언어 선택부(140)는 각 후보 발음열의 출현 확률 정보에 기초하여, 음성 인식된 최종 언어를 선택한다. 또한, 언어 선택부(140)는 발음 사전(150)을 이용하여 각 후보 발음열과 대응되는 단어를 검출함으로써, 검출된 단어의 출현 확률 정보에 기초하여, 음성 인식된 최종 언어를 선택할 수 있다. 단어의 출현 확률 정보는, 음성 인식이 수행될 때, 음성 인식된 언어에서, 해당 단어가 출현할 확률 정보를 의미한다. 이하에서 디바이스(100)의 각 구성 요소에 대해 구체적으로 설명하기로 한다.
- [0041] 특징 추출부(110)는 음성 신호를 수신하면, 실제 화자가 발성한 음성 부분만을 검출하여, 음성 신호의 특징을 나타낼 수 있는 정보를 추출할 수 있다. 음성 신호의 특징을 나타내는 정보는, 예를 들면, 음성 신호가 속하는 파형(waveform)에 따라 입 모양 또는 혀의 위치를 나타낼 수 있는 정보를 포함할 수 있다. 특징 추출부(110)에 의하여 추출된 특징 정보에 기초하여 음성 신호와 대응되는 발음열이 검출될 수 있다.
- [0042] 후보 발음열 검출부(120)는 추출된 음성 신호의 특징 정보와 음향 모델(acoustic model, 130)을 이용하여 음성 신호와 매칭될 수 있는 적어도 하나의 후보 발음열(pronunciation variants)을 검출할 수 있다. 음성 신호에 따라 복수 개의 후보 발음열이 검출될 수 있다. 예를 들어, 'th', 'd' 등의 발음은 비슷하므로, 동일 음성 신호에 대하여 'th', 'd' 등의 비슷한 발음을 포함하는 복수 개의 후보 발음열이 검출될 수 있다. 후보 발음열은 단어 단위로 검출될 수 있으나, 이에 한하지 않고, 후보 발음열은 음운, 음소, 음절 등의 다양한 단위로 검출될 수 있다.
- [0043] 음향 모델(130)은 음성 신호의 특징 정보로부터 후보 발음열을 검출하기 위한 정보를 포함할 수 있다. 또한, 음향 모델(130)은 많은 양의 음성 데이터로부터 통계적인 방법을 통하여 생성될 수 있다. 예를 들면, 음향 모델(130)은 불특정 다수의 발화 데이터로부터 생성될 수도 있고, 특정 화자로부터 수집된 발화 데이터로부터 생성될 수도 있다. 따라서, 화자에 따라 음성 인식 시 개별적으로 적용될 수 있는 음향 모델(130)이 존재할 수 있다.
- [0044] 언어 선택부(140)는 발음 사전(150) 및 문법 모델(160)을 이용하여 후보 발음열 검출부(120)에서 검출된 각 후보 발음열의 출현 확률 정보를 구할 수 있다. 그리고, 언어 선택부(140)는 각 후보 발음열의 출현 확률 정보에 기초하여, 음성 인식된 최종 언어를 선택한다.
- [0045] 또한, 언어 선택부(140)는 발음 사전(150)을 이용하여 각 후보 발음열과 대응되는 단어를 구할 수 있다. 언어 선택부(140)는 각 후보 발음열과 대응되는 단어에 대한 출현 확률 값을 문법 모델(160)을 이용하여 구할 수 있다. 언어 선택부(140)는 단어에 대한 출현 확률 값이 가장 큰 후보 발음열을 최종 선택할 수 있다. 최종 선택된 발음열과 대응되는 단어가 음성 인식된 단어로 출력될 수 있다.
- [0046] 발음 사전(150)은 후보 발음열 검출부(120)에 의해 검출된 후보 발음열과 대응되는 단어(word)를 획득하기 위해 필요한 정보를 포함할 수 있다. 발음 사전(150)은 각 단어의 음운변화 현상에 따라 획득된 발음열로부터 구축될 수 있다.
- [0047] 단어의 발음은 해당 단어의 앞뒤 단어 또는 문장 상 위치, 화자의 특성 등에 따라 변경될 수 있어 일관적이지 않다. 또한, 출현 확률 값은 현재 단어가 출현할 확률 또는, 현재 단어가 특정 단어와 동시에 출현할 확률을 의미한다. 디바이스(100)는 문법 모델(160)에 포함된 단어 또는 발음열에 대한 출현 확률 값을 이용함으로써, 문맥을 고려한 음성 인식을 수행할 수 있다.
- [0048] 디바이스(100)는 발음 사전(150)을 이용하여 후보 발음열의 단어를 구하고, 문법 모델(160)을 이용하여 후보 발

음열과 대응되는 단어에 대한 출현 확률 정보를 구함으로써, 음성 인식을 수행할 수 있다. 이에 한하지 않고, 디바이스(100)는 발음 사전(150)을 통해 발음열과 대응되는 단어를 구하지 않고 후보 발음열에 대한 출현 확률 정보를 문법 모델(160)로부터 구할 수도 있다.

- [0049] 예를 들면, 한국어의 경우, 후보 발음열 검출부(120)는 '학교(hakkkyo)'라는 후보 발음열을 검출하였을 때, 언어 선택부(140)는 검출된 후보 발음열 '학교(hakkkyo)'와 대응되는 단어로, 발음 사전(150)을 이용하여 '학교'라는 단어를 구할 수 있다. 또 다른 예로, 영어의 경우, 후보 발음열 검출부(120)는 'skul'이라는 후보 발음열을 검출하였을 때, 언어 선택부(140)는 검출된 후보 발음열 'skul'과 대응되는 단어로, 발음 사전(150)을 이용하여 'school'이라는 단어를 구할 수 있다. 언어 선택부(140)는 '학교' 또는 'school'이란 단어에 대한 출현 확률 정보에 기초하여 음성 신호와 대응되는 단어를 최종적으로 선택하고 선택된 단어를 출력할 수 있다.
- [0050] 문법 모델(160)은 단어에 대한 출현 확률 정보를 포함할 수 있다. 출현 확률 정보는 단어 별로 존재할 수 있다. 디바이스(100)는 각 후보 발음열에 포함된 단어에 대한 출현 확률 정보를 문법 모델(160)로부터 획득할 수 있다.
- [0051] 예를 들어, 문법 모델(160)은 현재 단어 B가 출현하기 이전에 단어 A가 출현한 경우, 현재 단어 B가 출현할 확률에 관한 정보인 P(B|A)를 포함할 수 있다. 다시 말하면, 단어 B의 출현 확률 정보인 P(B|A)는 단어 B의 출현 전에 단어 A가 출현하는 것을 조건으로 할 수 있다. 또 다른 예로, 문법 모델(160)은 단어 B의 출현 전에 단어 A 및 C, 즉, 복수 개의 단어가 출현하는 것을 조건으로 하는 P(B|A C)를 포함할 수 있다. 다시 말하면, P(B|A C)는 단어 B의 출현 전에 단어 A와 C 모두 출현하는 것을 조건으로 할 수 있다. 또 다른 예로, 문법 모델(160)은 조건부 확률이 아닌, 단어 B에 대한 출현 확률 정보 P(B)를 포함할 수 있다. P(B)는 단어 B가 음성 인식 시 출현할 확률을 의미한다.
- [0052] 디바이스(100)는 문법 모델(160)을 이용하여 언어 선택부(140)에서 각 후보 발음열과 대응되는 단어의 출현 확률 정보에 기초하여 음성 인식된 단어를 최종 결정할 수 있다. 즉, 디바이스(100)는 출현 확률 정보가 가장 높은 단어를 음성 인식된 단어로 최종 결정할 수 있다. 언어 선택부(140)는 음성 인식된 단어를 텍스트로 출력할 수 있다.
- [0053] 한편, 본 발명의 일 실시예에 따른 언어 선택부(140)는 문법모델(160)을 이용하여 인식된 사용자 언어에 문법상 오류가 존재하는지를 문법오류모델(170)을 활용하여 수행할 수 있다.
- [0054] 즉, 본 실시예에 따른 프로세서는, 상기 사용자의 음성을 문자로 인식하기 위해 사용되는 제 1 문법 모델 (160)뿐만 아니라, 상기 사용자의 음성을 문자로 인식한 후, 문자 인식된 언어에 존재하는 문법상 오류를 인식하기 위해 사용되는 제 2 문법 모델(170)을 추가적으로 사용하는 것으로 볼 수 있다.
- [0055] 상기 제 2 문법 모델(170)은 웹 서버의 데이터에 기반하여 학습 기반으로 각 문법상 오류 및 각 문법상 오류에 대응하는 문법 설명을 쌓으로써 갱신하도록 구성될 수 있다. 예를 들어, 온라인 교육 서비스 등으로 축적된 사용자가 빈번하게 발생시키는 문법 오류(예를 들어, 시제, 단/복수 표현, 구문 등)와 이러한 문법 오류에 대한 설명을 쌓으로 갱신하도록 구성될 수 있다.
- [0056] 물론, 문법 오류와 해당 문법 오류에 대한 설명이 1대1 관계를 가지지 않을 수 있다. 예를 들어, 도 2의 제 2 음성 반응에 대한 예에서 'be going to 구문에 대한 문법 설명'이 쌓으로 연결될 수 있으며, 'be going to 구문에 대한 문법 설명'은 도 2의 예시뿐만 아니라 다양한 문장의 문법 오류와 연결될 수 있다.
- [0057] 도 3의 내용을 정리하면, 본 실시예에 따른 상기 프로세서는, 사용자의 음성에서 특징을 추출한 후(110), 상기 특징에 대응하는 후보 발음열을 검출하고(120), 상기 후보 발음열 중 제 1 문법 모델(160)을 이용하여 소정 확률 이상으로 조합가능한 언어를 선택하여(140), 상기 사용자의 음성을 문자로 인식하는 동작을 수행할 수 있다.
- [0058] 여기서, 상기 후보 발음열 중 상기 제 1 문법 모델(160)을 이용하여 조합가능한 언어를 선택하지 못하는 경우, 상기 오류가 발음상 오류인 것으로 판단할 수 있다.
- [0059] 이와 달리, 상기 프로세서는, 상기 후보 발음열 중 상기 제 1 문법 모델(160)을 이용하여 조합가능한 언어를 선택하되, 상기 후보 발음열 중 최상위 확률로 인식되는 문자와 상기 선택된 언어 사이에 차이가 소정 기준 이상인 경우, 상기 오류가 발음상 오류인 것으로 판단할 수도 있다.
- [0060] 상술한 설명에서는 사용자가 말하는 음성이 외국어인 경우를 가정하여 설명하였으나, 이에 한정할 필요는 없다.
- [0061] 예를 들어, 본 발명의 일 실시예에 따른 프로세서는, 상기 사용자 음성 중 모국어 음성과 외국어 음성을 구분할

수 있다. 프로세서가 상기 사용자 음성이 상기 외국어 음성으로 판단되는 경우, 도 2 내지 도 3과 관련하여 상술한 동작을 수행하도록 구성될 수 있다.

[0063] 한편, 본 발명의 일 실시예에서, 상기 사용자의 음성은 상기 사용자 기기의 동작 모드를 변경하기 위한 명령어를 포함할 수 있다. 상기 명령어에 따라 상기 프로세서는 (1) 1 문장 단위로 상기 사용자의 음성을 처리하여 상기 음성 반응을 상기 스피커에 제공하는 문답 모드, 또는 (2) 소정 시간 상기 사용자의 음성을 지속적으로 기록한 후 상기 음성 반응을 상기 스피커에 제공하는 장문 기록 모드로 선택적으로 동작할 수 있다.

[0065] (1) 문답 모드

[0066] 상술한 바와 달리 사용자가 사전에 휴대폰, 노트북 등 사용자 기기에 외국어 별로 해당 모드를 설명해 둘 수 있다. 대체로 해당 외국어의 초급 단계에서는 장문의 문장을 이야기하기 어려우며, 상황에 맞게 본 실시예에 따른 사용자 기기의 스피커를 통해 1문장의 질문이 제공되고, 이에 대해 사용자의 1문장 답변을 수신하여 상술한 외국어 교육 서비스를 제공할 수 있다.

[0067] 문답 모드에서는 사용자의 문법 오류와 발음 오류 중 발음 오류의 확률을 높이는 것이 바람직하나, 이에 한정될 필요는 없다.

[0069] (2) 장문 기록 모드

[0070] 상술한 바와 달리 사용자가 사전에 휴대폰, 노트북 등 사용자 기기에 외국어별로 해당 모드를 설명해 둘 수 있음은 동일하다. 대체로 해당 외국어의 고급 단계에서는 예를 들어, 일기를 쓰듯이 약 1분 가량의 사용자 speech를 수신하고, 이에 대한 일반적인 대화 (예: 상술한 제3음성 응답)를 제공하는 것이 교육상 효율적일 수 있다.

[0071] 사용자와 외국어로 자연어 대화를 하는 와중에 기 설정된 빈도로 사용자의 speech 중 문법상 개선사항을 제공해 줄 수 있다. 물론, 본 실시예에 따른 사용자 기기(예: AI 스피커)와 대화를 통해 문법 지적 빈도수를 조절할 수도 있다. 예를 들어, “Bob (스피커 캐릭터 이름), reduce the frequency of error correction” 등을 이야기하여 빈도수를 조절할 수 있다.

[0073] 한편, 본 발명의 일 실시예에서는 상술한 모드들 이외에 하기와 같은 교육 모드를 제공할 수 있다.

[0074] (3) 표현 변환

[0075] 동일한 의미를 가지는 사용자 음성 표현에 대해 조금 더 격식있는 표현을 알려주거나, 더 Casual한 표현을 알려주는 방식으로 다양한 표현을 알려줄 수 있다.

[0076] 예를 들어, 사용자가 'I'm going to finish this patent specification drafting'이라고 이야기하였을 때, 본 실시예에 따른 사용자 기기는 기 설정된 모드에 따라, 또는 사용자의 별도 문의에 따라 이를 더 Casual한 표현으로서 'I'm gonna finish this patent specification drafting'을 알려줄 수 있으며, 그 반대로 사용자의 음성 표현을 더 격식있는 표현으로 변경하여 알려줄 수도 있다.

[0077] 유사한 기능으로서 사용자의 영어 표현에 대해 모범적인 미국 영어 발음, 영국 영어 발음 등을 제시해 줄 수도 있다.

[0079] (4) 강의 모드

[0080] AI 스피커로 대표되는 본 발명의 일 실시예에 따른 사용자 기기를 통해 사용자에게 강의를 제공할 수 있다. 강의 중 사용자의 음성 답변을 요구하는 상호작용이 있을 수 있다.

[0081] 음성 강의 중 시각적인 표시가 필요한 경우 후술하는 바와 같이 IoT 방식 등 무선통신 방식으로 연결된 TV, 모니터 등을 통해 시각자료를 표시할 수 있다. 이러한 시각자료 표시에 대해 사용자가 현 시점에서 모니터/TV 등에 접근하기 어려운 경우, 이를 생략하고 진행하도록 명령할 수 있다.

[0082] 또한, 이러한 강의 모드에서, 또는 이와 별도로 본 발명의 일 실시예에 따른 사용자 기기는 사용자의 교육 스케줄링 서비스를 제공할 수 있으며, 교육 과정 별 성취도에 따라 제공되는 교육 내용을 변경하여 제안할 수도 있다.

[0084] (5) 사전 모드

[0085] 사용자가 수시로 외국어 단어, 구문의 의미를 음성으로 문의할 수 있으며, 이에 대해 본 실시예에 따른 사용자 기기는 웹에 연결되어 해당 단어, 구문의 의미/정확한 발음 등을 사용자에게 알려줄 수 있다.

- [0087] 도 4는 본 발명의 일 실시예에 따른 사용자 기기를 통해 태내 기기를 제어하는 개념을 설명하기 위한 도면이다.
- [0088] 도 4에 도시된 바와 같이, 사용자 기기(10)는 IoT 방식으로 하나 이상의 태내 기기(201, 202, 203, 204, 205, 206)와 무선 연결될 수 있으며, 사용자 기기(10)와 연결될 수 있는 태내 기기는 도 4에 도시된 휴대폰(201), 다른 AI 스피커(202), 세탁기(203), 로봇 청소기(204), 에어컨(205), 냉장고 (206) 등을 포함할 수 있으며, 이에 한정될 필요는 없다.
- [0089] 본 실시예에 따른 사용자 기기(10)는 상기 사용자 음성 중 모국어 음성 또는 외국어 음성 중 하나 이상을 통해 상기 태내 기기(201, 202, 203, 204, 205, 206)를 제어할 수 있다. 또한, 사용자의 외국어 음성을 통해 태내 기기(201, 202, 203, 204, 205, 206)를 제어하는 과정에서 도 2 내지 도 3과 관련하여 상술한 외국어 교육을 제공할 수도 있다.
- [0091] 도 5는 본 발명의 일 실시예에 따라 외국어 교육을 제공하는 방법을 설명하기 위한 도면이다.
- [0092] 도 4까지의 설명에서 사용자 기기는 AI 스피커 개념의 사용자 기기를 가정하였으나, 본 발명의 다른 실시예에 따른 사용자 기기는 사용자가 노트북, 데스크탑 등 모니터와 같은 디스플레이 수단(400)을 구비한 다른 형태의 사용자 기기를 통해서도 제공될 수 있다.
- [0093] 예를 들어, 본 실시예에 따른 사용자 기기의 온라인 접속을 통해 외국어 교육을 제공하는 방법은, 상기 사용자 기기의 마이크를 통해 입력된 사용자의 외국어 문장 음성 발음을 인식하여 문자로 변환하고, 상기 문자로 변환된 문장의 오류를 판단할 수 있다.
- [0094] 상기 오류가 있는 부분이 문법상 오류인 경우, 상기 오류가 있는 부분에 대응하는 기 마련된 문법 설명을 상기 오류가 있는 부분의 강조 표시(410)와 함께 상기 사용자 기기의 표시장치(400)에 도 5에 예시된 바와 같이 표시할 수 있다. 즉, 도 5에 도시된 예에서 'I'와 'gonna' 사이에 'be 동사'가 빠져 있는 문법상 오류가 있기 때문에 'I gonna' 부분 (410)을 오류가 있는 부분으로서 강조 표시하고, 이에 대한 반응으로서 “Good luck Aaron, go for it ! But, you'd better put 'be verb' in front of 'gonna'를 표시해 줄 수도 있다.
- [0095] 아울러, 상술한 예에서 “[be going to] 구문에 대한 문법 설명보기 (420)” 를 추가적으로 표시하고, 해당 설명보기(420)를 클릭하면, 대응하는 구체적 문법 설명 또는 문법 강의를 수강할 수 있도록 할 수 있다.
- [0096] 한편, 상기 오류가 있는 부분이 발음상 오류인 경우, 상기 오류가 있는 부분을 대체할 수 있는 표현을 상기 오류가 있는 부분의 강조 표시와 함께 상기 사용자 기기의 표시장치(400)에 표시하여 추천하고, 상기 추천 표현을 포함하는 예시적인 문장의 추천 발음을 상기 사용자 기기의 스피커를 통해 제시할 수 있다. 이를 통해 사용자에게 추천 발음을 통해 모범적인 발음에 대한 교육을 제공할 수 있다. 이러한 모범적인 발음은 단어 사이의 연음/묵음 관계를 고려하여 사용자에게 보다 유창한 발음으로서 제공해 줄 수 있다.
- [0097] 한편, 상기 오류가 있는 부분을 대체할 수 있는 표현을 추천하는 것은, 복수의 표현을 추천하여 상기 사용자 기기의 표시장치(400)에 표시하고, 상기 사용자 기기의 입력 장치를 통해 선택된 표현을 상기 추천 표현으로 결정하며, 상기 사용자의 외국어 문장 음성 발음 중 오류가 있는 부분과 상기 결정된 추천 표현 사이의 맵핑 관계를 상기 외국어 교육을 제공하는 서버의 데이터 베이스에 저장하여 인공지능 기반으로 상기 서버를 학습시키는 것을 포함할 수도 있다. 이러한 사용자와의 상호 작용으로서 본 실시예에 따른 사용자 기기는 특정 사용자에게 맞도록 스스로 진화할 수 있다.
- [0098] 한편, 상기 오류가 있는 부분과 상기 기 마련된 문법 설명의 대응 데이터는 인공지능 방식으로 상기 외국어 교육을 제공하는 서버의 데이터 베이스에 증가시키며, 상기 오류가 있는 부분이 문법상 오류인 것으로 판단되는 경우는, 상기 오류가 있는 부분과 대응하는 기 마련된 문법 설명이 존재하는 경우일 수 있다.
- [0099] 상술한 바와 같은 본 발명의 다양한 실시예들은 본 발명의 구체적 이해를 위해 제공된 것이며, 본 발명은 상술한 설명의 구체적 일례에 한정되지 않고 명세서 전체에서 제공하는 외국어 교육용 인공지능 기능을 구비한 사용자 기기 및 외국어 교육 방법에 기반하여 판단되어야 한다.

산업상 이용가능성

- [0100] 상술한 바와 같은 본 발명에 따른 외국어 교육용 인공지능 기능을 구비한 사용자 기기 및 외국어 교육 방법은 비대면 시대에 새로운 교육 방식으로서 광범위하게 활용될 수 있다.

도면
도면1

20a



20b



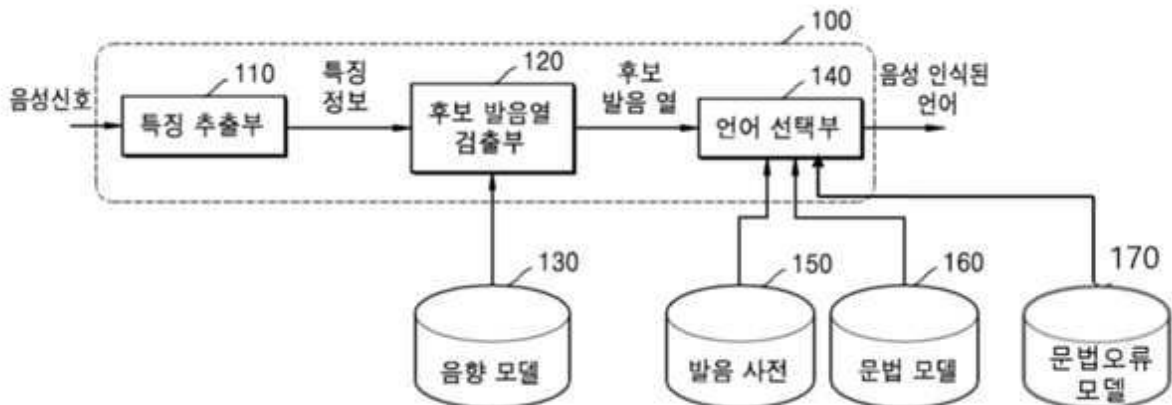
10



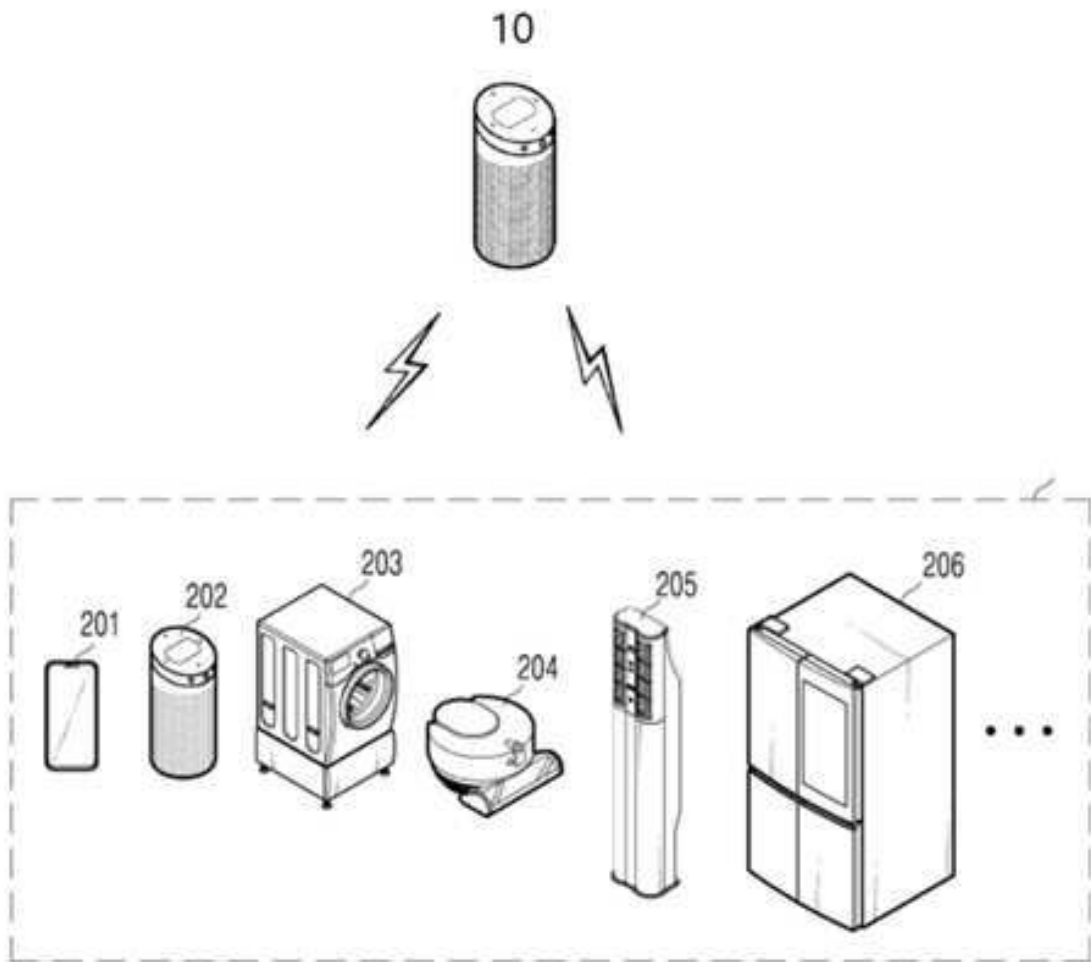
도면2



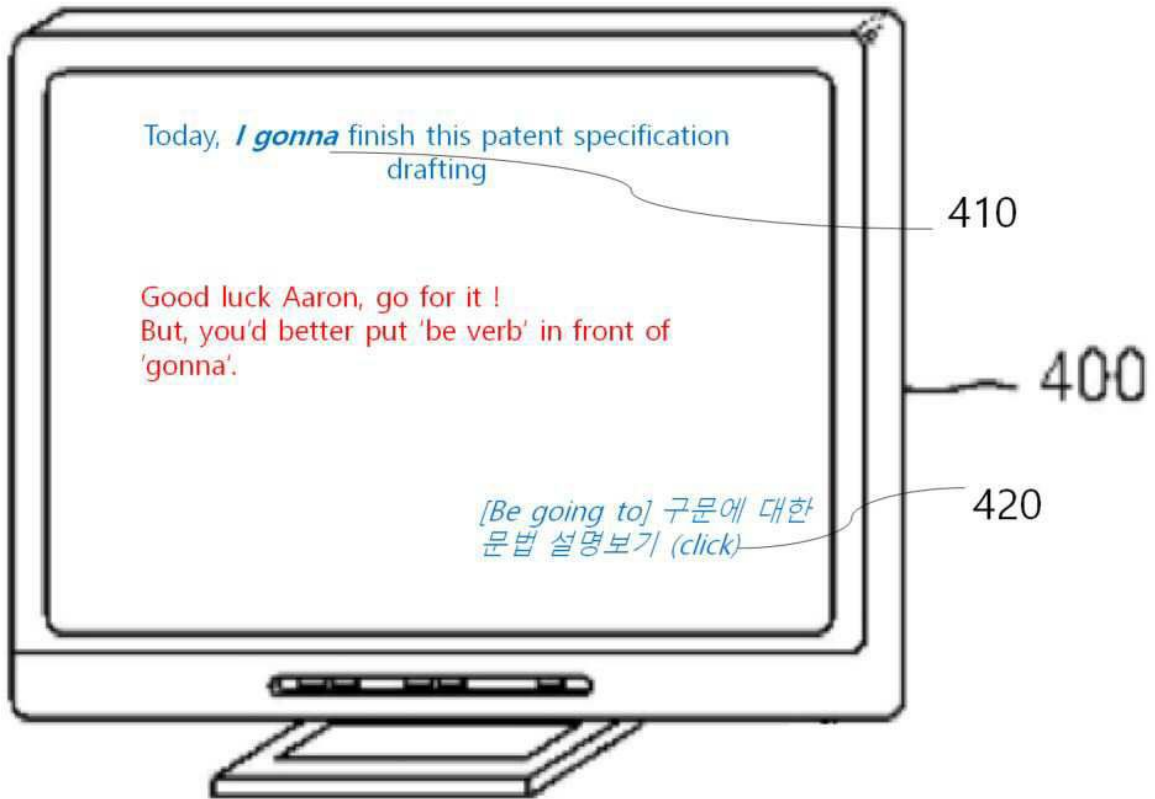
도면3



도면4



도면5



【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 청구항 1

【변경전】

외국어 교육용 인공지능 기능을 구비한 사용자 기기에 있어서,

사용자의 음성을 입력받는 음성입력장치;

상기 사용자 음성을 문자로 인식하여 처리하는 프로세서; 및

상기 프로세서의 처리에 따라 사용자에게 음성 반응을 제공하는 스피커를 포함하며,

상기 프로세서는,

(a) 상기 사용자의 음성의 오류가 발음상 오류인지, 또는 상기 오류가 문법상 오류인지를 판단하며,

(b) 상기 오류가 발음상 오류인 경우, 상기 오류가 있는 부분을 대체할 수 있는 표현을 포함하는 제 1 음성 반응을 상기 스피커를 통해 제공하며,

(c) 상기 오류가 문법상 오류인 경우, 상기 오류가 있는 부분에 대응하는 기 마련된 문법 설명을 포함하는 제 2 음성 반응을 상기 스피커를 통해 제공하며,

상기 사용자의 음성은 상기 사용자 기기의 동작 모드를 변경하기 위한 명령어를 포함하며,

상기 명령어에 따라 상기 프로세서는 (1) 1 문장 단위로 상기 사용자의 음성을 처리하여 상기 음성 반응을 상기 스피커에 제공하는 문답 모드, 또는 (2) 소정 시간 상기 사용자의 음성을 지속적으로 기록한 후 상기 음성 반응을 상기 스피커에 제공하는 장문 기록 모드로 선택적으로 동작하는, 사용자 기기.

【변경후】

외국어 교육용 인공지능 기능을 구비한 사용자 기기에 있어서,

사용자의 음성을 입력받는 음성입력장치;

상기 사용자의 음성을 문자로 인식하여 처리하는 프로세서; 및

상기 프로세서의 처리에 따라 사용자에게 음성 반응을 제공하는 스피커를 포함하며,

상기 프로세서는,

(a) 상기 사용자의 음성의 오류가 발음상 오류인지, 또는 상기 오류가 문법상 오류인지를 판단하며,

(b) 상기 오류가 발음상 오류인 경우, 상기 오류가 있는 부분을 대체할 수 있는 표현을 포함하는 제 1 음성 반응을 상기 스피커를 통해 제공하며,

(c) 상기 오류가 문법상 오류인 경우, 상기 오류가 있는 부분에 대응하는 기 마련된 문법 설명을 포함하는 제 2 음성 반응을 상기 스피커를 통해 제공하며,

상기 사용자의 음성은 상기 사용자 기기의 동작 모드를 변경하기 위한 명령어를 포함하며,

상기 명령어에 따라 상기 프로세서는 (1) 1 문장 단위로 상기 사용자의 음성을 처리하여 상기 음성 반응을 상기 스피커에 제공하는 문답 모드, 또는 (2) 소정 시간 상기 사용자의 음성을 지속적으로 기록한 후 상기 음성 반응을 상기 스피커에 제공하는 장문 기록 모드로 선택적으로 동작하는, 사용자 기기.