

발성장애아동을 위한 발성훈련시스템 설계

정은순, 김봉완, 이용주

원광대학교 컴퓨터공학과

true,tacanemo@gaebyok.wonkwang.ac.kr, yjlee@wonwms.wonkwang.ac.kr

Design of Speech-Training System for Voice Disorders Using Visual Effect

Eun-Soon Jeong, Bong-Wan Kim, Yong-Ju Lee

Dept. of Computer Engineering, Wonkwang University

요 약

본 연구는 발성장애아동을 대상으로 시각적 효과를 이용하여 발성치료 및 교육을 위한 도구 개발을 목적으로 한다. 따라서 특수아동의 발성장애에 대한 특성을 고려하여 그에 적합한 개별학습과 반복학습이 가능하도록 설계하였다. 또한 GUI와 게임적 요소를 가미하여 발성에 대한 아동의 흥미유발과 자발적 학습이 가능하도록 하였다.

1. 서 론

음성언어는 인간의 의사소통의 기본수단이며 인간 상호관계를 연결해 주는 매체이다. 언어인지과정은 청각기관의 정상적 작동으로 음을 전달하게 된다. 그러나 언어 및 청각장애자의 경우 조음기관과 공명기관을 비정상적으로 자극하거나 높은 음도와 강도를 사용하여 발성함으로써 조음 및 음성에 문제가 발생하게 된다.

컴퓨터 기반의 발성장애인을 위한 훈련시스템에 대한 연구 필요성은 강조되고 있으나 훈련시스템의 개발은 소수에 그치고 있다. 특히 발성장애인을 대상으로 하는 S/W는 개발이 활발히 이루어지지 않고 있어 외국의 언어치료 자료를 근거로 치료 및 평가하는 실정이다. 이런 외국의 언어훈련기는 우리나라 아동의 언어발달에 맞지 않은 점과 고가의 훈련시스템으로 보편화된 언어훈련기로 사용하기에 어려움이 있다.

시각적 훈련기기를 이용한 발음 훈련은 발음에 대한 시각적 피드백을 이용하여 흥미롭게 훈련과정을 거칠 수 있게되어 장애아동에게 충분한 동기유발이 주어지므로써 발음훈련의 시간을 단축시키는 효과가 있다[1].

본 연구는 다음과 같이 구성된다. 제2장은 관련연구에 대해 기술하고 제3장은 발성훈련시스템의 설계에 대해 기술하고 제4장은 결론 및 향후 연구방향에 대해 기술하였다.

2. 관련연구

2.1 발성을 위한 신호처리 이론

음성신호로부터 발성장애아를 훈련시키기 위한 파라메타를 추출하기 위해서 필요한 디지털 신호처리이론에 대해 살펴본다.

음성의 특징 파라메타에는 시간영역과 주파수 영역 파라메타가 있다. 시간 영역 파라메타에는 신호의 에너지, 크기, 영교차율, 주기 등이 있고, 주파수 영역 파라메타에는 선형예측계수, 반사계수, 포먼트 주파수, 성도 면적 함수 등이 있다. 시간영역 파라메타는 동일한 종류의 발음을 매번 발성시의 차이점 즉, 세기의 변화 및 음의 높이, 음의 빠르기 등을 잘 나타내며, 주파수 영역 파라메타는 동일한 종류의 발음시 시간이 변화하여도 잘 변화하지 않는 특징인 조음기관의 구조 및 주파수 스펙트럼 등을 잘 나타낸다[9].

2.2.1. 음의 높이(Pitch)

인간의 발성은 동일한 소리의 발음이라도 소리의 주파수의 높고 낮음에 따라서 저음과 고음이 나오게 된다. 이러한 것은 성대의 진동 주파수의 높낮이에 따라 달려있다. 따라서 발성자의 음의 높낮이를 찾기 위해서는 음성신호 데이터를 처리하여 성대의 진동 주기를 찾아내 주기가 짧은 음은 고음, 주기가 긴 음은 저음으로 구분한다. 피치를 검출하는 방법은 시간영역에서 검출하는 방법과 주파수 영역에서 검출하는 방법에 따라 여러 방법이 제안되어 있다[10].

2.2.2. 음의 세기(Loundess)

음의 세기는 허파에서 외부로 얼마나 힘차게 공기를 밀어내는가에 하는 것으로 이러한 음의 세기는 에너지를 구함으로써 쉽게 알아낼 수 있다.

2.2.3 유 · 무성음

유성음은 주기적인 성대 필스가 성도를 통해 감으로서 발생되기 때문에 유성음 각 원소마다 성대에서 고유한 공명이 일어난다. 따라서 유성음의 스펙트럼은 음소마다 고유한 공명 봉우리를 갖게 된다. 이러한 공명 봉우리를 포먼트라하며 낮은 쪽 주파수에서부터 두드러진 포먼트를 차례로 제1, 제2, 제3 포먼트 등으로 불리어 진다. 유성음의 스펙트럼에서는 보통 제1 포먼트의 주파수가 250-750Hz에 존재하며, 또한 공명현상 때문에 무성음에 비해 에너지가 크다.

무성음은 불규칙한 잡음이 성대를 자극하는 입력으로 되어 성대를 통과하는 동안 성대의 협착점에서 공명이 발생하게 되며, 무성음의 스펙트럼에서 2500Hz 근처에서 주된 공명 봉우리가 존재한다.

이러한 유·무성음의 구분을 위해서는 단구간 에너지와 영교차를 그리고 위에서 구한 피치값을 사용할 수 있다. 유성음의 경우 무성음에 비해 낮게 나타난다. 이러한 단구간 에너지와 영교차율의 임계치를 적절히 설정함으로써 유·무성음을 구분할 수 있다. 그러나 이 방법은 피치를 구하여 유·무성음을 구분하는 것보다 일반적으로 정확도가 떨어진다고 볼 수 있다. 따라서 본 논문에서는 위에서 구한 피치의 유·무성에 의해 유·무성을 구분한다.

3. 발성훈련시스템 설계

3.1 컴퓨터 기반 발성훈련 시스템(Computer Based Speech Training : CBST)

컴퓨터 기반 발성훈련시스템은 발성장애아의 발성훈련의 상태와 음성생성 문제에 대한 신뢰할만한 평가방법 및 효율적인 훈련 프로그램과 평가 결과를 충당할 수 있는 커리큘럼의 계획이 필요하다.

발성에 대한 피드백 폼은 컴퓨터 스크린에 시각적 표현으로 다양한 그래픽을 이용하여 실시간 게임, 애니메이션, 그래프 차트 등으로 표현한다[6]. 이와 같은 컴퓨터 기반 발성훈련시스템의 궁극적 목표는 발성자 자신의 환경에 맞는 최적의 발성과 음성생성을 통해 자기 감응적인 반응을 얻는 것이다.

3.2 시스템 구성

컴퓨터 기반 발성훈련시스템은 발성장애아의 음성을 아날로그에서 디지털로 변환하는 음성신호처리 모듈과 발성에 대한 음성적 특성을 추출해 내는 음성분석모듈, 그리고 발성훈련에 대한 분석 및 평가를 위한 음성인식 및 평가모듈, 발성훈련을 시각적으로 피드백함으로써 훈련효과를 높이기 위한 그래픽처리모듈, 훈련자 및 훈련결과에 대한 데이터관리를 위한 훈련관리 및 성적평가 모듈로 구성되어 있다.

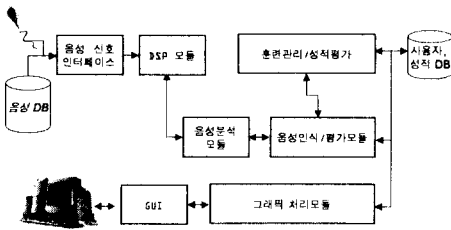


그림 1. 시스템 구성도

3.3 발성훈련모듈 구성

본 발성훈련시스템은 각 모듈별로 관리하며 크게 선행훈련, 발성인지훈련단계, 발성강화단계, 음운훈련단계로 구성된다. 첫째, 선행훈련은 발성훈련에 들어가기 전에 발성장애자의 부자연스런 발성에 대한 특정 음역대를 분석하는 모듈이다. 둘째, 발성인지단계는 발성에 대한 청각적 자기피드백을 시각적으로 구현함으로써 발성장애자가 발성에 대한 인지를 목적으로 한다. 셋째, 발성강화단계는 발성인지단계를 보완 및 강화한 단계로 발성의 유동성과 언어의 멈춤 등 음의 리듬을 훈련시켜 자연스런 발성을 유도하는데 목적이 있다.

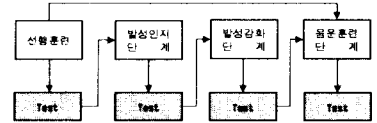


그림 2. 발성훈련 모듈

넷째, 음운훈련단계는 조음장애를 수반하는 아동을 위한 모듈로 특정음소의 왜곡, 대체 등과 같은 조음장애[4,5]와 특정음소를 발성하지 못하는 경우를 위한 훈련모듈이다.

3.3.1 선행훈련

선행훈련은 Dual-screen을 이용하여 목적음성을 제시하면 발성장애아는 다른 화면을 통해 목적음성에 match 시키는 과정으로 Waveform이나 Spectrogram을 이용하여 pitch와 loudness를 훈련한다. 이를 통해 발성장애아가 부자연스럽게 발성하는 Pitch와 loudness를 통계적으로 산출하여 다음 발성훈련과정을 결정하게 된다.

3.3.2 발성인지단계

본 훈련단계는 발성지속훈련, 소리고지훈련, 소리강도훈련으로 구성되어 있다. 발성지속훈련은 발성장애의 경우 호흡이 불규칙하여 부자연스런 발성을 하므로 발성의 유·무를 산출하여 발성을 지속시킬 수 있는 훈련모듈이다. 소리고지훈련과 소리강도훈련은 선행훈련을 기반으로 발성장애아의 장애를 수반하는 음역대를 설정하여 반복적으로 훈련시킬 수 있는 모듈이다. 또한 음역대를 낮은음도에서 높은음도까지 단계적으로 설정하여 여러 음역대를 훈련가능 하도록 설계하였다. 이러한 훈련 모듈은 발성장애아의 음성을 분석하여 그 값을 움직임 벡터로 변환하여 그래픽처리 함으로써 발성훈련 효과 및 발성에 대한 인지를 높일 수 있다.

3.3.3 발성강화훈련

본 훈련단계는 발성인지훈련단계를 보완 및 강화한 단계로 소리고지 강화모듈, 소리강도 강화모듈, 유·무성음모듈로 구성된다. 소리고지 강화모듈과 소리강도 강화모듈은 발성장애자가 최적의 음도를 기반으로 여러 음역대를 자연스럽게 발성할 수 있도록 하는 모듈로 게임적 요소를 가미하여 발성장애아가 장애물을 통과하는 과정을 통해 발성훈련하는 모듈이다. 유·무성음 훈련모듈은 발성에 있어서 유·무성음의 자각이 되지 않으면 발음·발어의 요령이 안되므로 이 훈련모듈을 통해 유·무성을 자각·인식시키는 훈련이다.

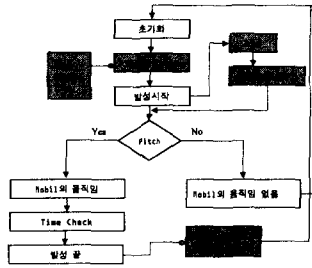


그림 3. 소리고저훈련 흐름도

다음은 소리고저 강화모듈에 대한 처리과정과 그래픽처리과정이다. 발성장애아의 음성은 음성분석모듈에서 Pitch값을 추출하게된다. 이러한 Pitch값은 움직임 벡터값으로 변환하여 Pitch 값에 따라 mobile을 움직이게 된다.

```

Procedure TPitchForm.DXTimer1Timer(Sender:
                                TObject; LagCount: Integer);
var
  Pitch:LongInt;
begin
  Pitch = SpeechAnal.GetPitch();
  DXSpriteEngine1.Move( );
  DXSpriteEngine1.Draw;
  DXDraw1.Flip;
end;
    
```

발성의 대한 청각적 피드백을 시각적 피드백으로 처리하기 위해 2차원 애니메이션이나 이미지를 이용한다.

그림 4는 DirectX를 이용한 그래픽 처리과정이다. DirectX의 DXDraw를 통해 배경 image의 생성과 실제 움직임 Mobile인 Sprite 생성을 하게 된다.

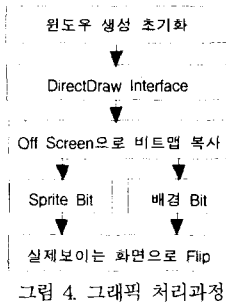


그림 4. 그래픽 처리과정

3.3.4 음운훈련단계

본 훈련단계는 단분별력테스트, 일음절훈련모듈, 이음절훈련 모듈로 구성된다.

단분별력테스트는 특정 음소를 제시하여 장애자가 이를 분별하는지를 테스트하는 과정이다. 일음절훈련모듈은 한 낱말 표현의 발달에 맞춘 것으로 핵심단어를 정확히 발성할 수 있도록 훈련시키는 모듈이다. 예로 특수아동이 '사'을 '하'으로 대치하여 발성하는 경우 '사'와 '하'와 같은 일음절을 제시하여 서로 다른 발음임을 인지시키는 모듈이다. 이음절훈련모듈은 일음절훈련과 유사 및 반대발성 성향을 갖는 이음절단어를 훈련시키는 모듈로 초성, 중성, 종성에 따라 발성이 달라지므로 이를 인지시키는 훈련 모듈이다

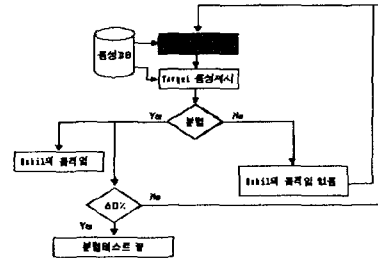


그림 5. 단어분별력테스트 흐름도

3.4 학습자 관리 및 발성훈련의 평가

발성훈련에 대한 평가방법은 발성자의 성, 나이, 정도의 특징에 대한 타당성 있는 모델을 제시하여 평가하는 방법과 발성자의 과거의 훈련과 비교하여 평가하는 방법이 있다[6]. 본 훈련시스템에서는 두 가지 평가방법을 모두 이용하여 발성훈련에 대한 분석 및 평가를 위한 것으로 훈련의 진행과정을 명확히 문서화하여 발성장애아의 발성훈련의 흐름을 파악할 수 있다. 발성의 특성이나 파라메타의 범위를 통해서 발성장애아의 발성훈련에 관한 실시간 분석과 훈련 전후를 비교함으로써 훈련의 효율성을 평가할 수 있다.

4. 결론 및 향후 연구방향

컴퓨터를 활용한 장애인용 위한 훈련시스템의 필요성은 강조되고 있으나 훈련프로그램의 개발은 소수에 그치고 있어 외국의 훈련프로그램에 의존하고 있다. 따라서 발성장애아동을 위한 발성훈련시스템을 개발하는데 그 목적이 있다.

본 논문은 선행훈련을 통해 발성장애아의 발성 level에 맞는 훈련단계를 선택할 수 있는 개별학습과 반복학습을 할 수 있도록 설계하였다. 또한 시각적 효과를 이용하여 훈련에 대한 효율성을 높이도록 하였다. 향후 연구방향은 발성장애아를 위한 발성훈련시스템 구현과 검증에 있다.

참고 문헌

- [1] 국미경, 시각적 기기를 이용한 발음훈련이 청각장애유아의 자모음 발음안정에 미치는 효과, 박사학위논문, 대구대학교 대학원, 1993
- [2] 황귀용, 음성학입문, 한신문화사, 1997
- [3] <http://serce.taegu.ac.kr/virtual/class/langu/lecture/lecture.htm>, 대구대학교 특수재활가상캠퍼스
- [4] 정옥란, 음성과 음성치료, 원미사, 1996
- [5] 한국언어병리학회편, 언어장애의 이해와 치료, 군자출판사, 1994
- [6] Charles S. Watson and Diane Kewley-Port, "Advances in computer-based speech training : Aids for the profoundly hearing impaired", Volta-Review, 1989
- [7] John Ayres, Delphi Graphics and Game programming with expored, WORDWARE, 2000
- [8] 석봉현 외, Delphi programming Bible, 영진출판사, 1998
- [9] 최일상, 성도모델링을 이용한 발성훈련기법에 관한 연구, 석사학위논문 연세대학교대학원, 1991
- [10] "Talkin.D," A Robust Algorithm for Pitch Tracking (RAPT)," in Kleijn,W.B. and Paliwal,K.K.(Eds.), Speech Coding and Elsevier, 1995