

# 말운동장애인을 위한 시-청각 단서 제공 읽기 훈련 프로그램 개발

방동혁<sup>1</sup>, 전유용<sup>1</sup>, 양동권<sup>1</sup>, 길세기<sup>1,2</sup>, 권미선<sup>3</sup>, 이상민<sup>1,2</sup>

<sup>1</sup>인하대학교 전자공학과, <sup>2</sup>인하대학교 정보전자 공동연구소, <sup>3</sup>서울아산병원 신경과

## Development of a Reading Training Software offering Visual-Auditory Cue for Patients with Motor Speech Disorder

D. H. Bang<sup>1</sup>, Y. Y. Jeon<sup>1</sup>, D. G. Yang<sup>1</sup>, S. K. Kil<sup>1,2</sup>, M. S. Kwon<sup>3</sup>, S. M. Lee<sup>1,2</sup>

<sup>1</sup>Department of Electronic Engineering, Inha University,

<sup>2</sup>Institute for Information and Electronics Research, Inha University, <sup>3</sup>Dept. of Neurology, Asan Medical Center

(Received February 18, 2008. Accepted June 13, 2008)

### Abstract

In this paper, we developed a visual-auditory cue software for reading training of motor speech disorder patients. Motor speech disorder patients can use the visual and/or auditory cues for reading training and improving their symptom. The software provides some sentences with visual-auditory cues. Our sentences used for reading training are adequately comprised on modulation training according to a professional advice in speech therapy field. To ameliorate reading skills we developed two algorithms, first one is automatically searching the starting time of speech spoken by patients and the other one is removing auditory-cue from the recorded speech that recorded at the same time. The searching of speech starting time was experimented by 10 sentences per 6 subjects in four kinds of noisy environments thus the results is that  $7.042 \pm 8.99$ [ms] error was detected. The experiment of the cancellation algorithm of auditory-cue was executed from 6 subjects with 1 syllable speech. The result takes improved the speech recognition rate  $25 \pm 9.547$ [%] between before and after cancellation of auditory-cue in speech. User satisfaction index of the developed program was estimated as good.

**Key words :** visual-auditory cue, motor speech disorder, reading training

## 1. 서론

우 리 사회는 최근 급격히 고령화 사회를 넘어 고령 사회로 진 입하고 있다. 이에 따라 고령인 질환으로 인해 발생하는 장 애의 평가 및 치료, 보조 기술들의 중요성도 높아지고 있다. 그 중 뇌졸중이나 파킨슨병과 같은 질환들의 후유증의 하나가 말운동장 애(motor speech disorder)이다. 말운동장애는 신경계의 손상으 로 말하는 기능에 필요한 근육의 운동능력이 저하되거나 협응이 안되는 마비말장애나, 음절·낱말을 배열하고 산출하는 것에 문제 가 있는 말실행증 등을 말하는데[1], 이로 인해 의사소통의 불편함 을 겪게 되므로 건강한 고령자들의 사회활동에 매우 큰 영향을 주

며, 삶의 질을 현저히 떨어뜨리는 원인이 되고 있다.

이러한 말운동장애인의 재활 및 재활보조를 위한 여러 가지 방 법들이 병원기반의 전문가들을 통해 연구되고 있다. 기존의 연구 들을 살펴보면 환자들에게 시각요소들을 이용하여 명료도를 비교 한 연구에서는 문장과 알파벳을 결합하여 환자들에게 제공한 말하 기에서 다른 요소들보다 명료도가 향상되어졌고[2], 환자들에게 시각요소들과 청각요소들의 속도를 조절하여 실험한 말하기에서 메트로놈 소리를 이용한 실험에서 가장 크게 명료도가 향상되어졌 고[3], 부가적으로 제공된 청각요소와 속도조절 청각요소를 이용 하여 환자들에게 실험한 결과 속도조절 청각요소에서 명료도의 점 수가 크게 향상되었다[4]. 위의 연구들에서는 시각요소와 청각요 소들을 이용하여 속도조절이나 부가적으로 제공하는 방법들을 환 자들에게 제공하여 명료도를 비교하는 실험을 하였다. 읽기속도 요소들을 이용하여 실험한 결과 환자들은 느린 말하기 속도에서 STI(spatiotemporal index)로 분석한 결과 시간과 공간의 변화성

Corresponding Author : 이상민

(402-751) 인천광역시 남구 용현동 253 인하대학교 하이테크 803호

Tel : +82-32-860-7420 / Fax : +82-32-868-3654

E-mail : sanglee@inha.ac.kr

이 논문은 2007년도 정부(과학기술부)의 재원으로 한국과학재단의 지원 (R01-2007-000-10801-0)과 2007년 정부(교육인적자원부)의 재원으로 한국학술진흥재단의 지원(KRF-2007-331-D00592)을 받아 수행된 연 구임.

이 줄어들었고[5], 읽기속도 요소들을 이용하여 환자들에게 실험한 결과는 느린 읽기속도에서 습관적인 읽기속도보다 말하는 기간은 증가하였고 말소리 멈춤은 줄어든 결과를 보여주었다[6]. 위의 연구들에서는 환자들의 읽기 속도를 조절하여 실험하였고 속도에 따른 읽기특성을 비교하였다. 위 연구들은 시-청각 단서들과 속도에 의해서 말운동장애인의 명료도에 영향을 준다는 연구결과들을 보여주고 있다.

본 논문에서는 개발된 시-청각 단서 제공 읽기 훈련 프로그램(이하 시-청각 단서 프로그램)에 대한 실험과 훈련 및 분석효과를 높이기 위해 개발 탑재된 말소리 개시점 찾기 알고리즘, 청각단서 제거 알고리즘에 대한 실험과 실험결과에 대하여 기술한다.

## II. 말운동장애

그림 1은 중추신경계에서 말초신경계에 이르는 말운동 시스템을 모델링한 것이다[1]. 말운동장애는 이러한 신경계의 이상을 말하는 것으로, 크게 이완형 말장애와 경직형 말장애로 나눌 수 있다.

이완형 말장애는 약화를 뜻하는 것으로 중풍이나 외상 근무력증 등과 같은 질병이 신경계에 손상을 주어 근육자체의 마비나 약화가 나타나거나, 구어산출을 조정하는 조음기관의 근육들도 손상을 받게 되고, 개별근육이 수축하는데 있어 힘이 감소되는 것을 말한다. 이완형 말장애는 주로 발화근육인 하악, 혀, 구인두 등을 직접 지배하는 최종공통경로(final common pathway)의 손상으로부터 발생하는 것으로 알려져 있다[1].

경직형 말장애는 발화근육의 과도한 긴장으로 인해 발화시 혀의 움직임이 미약하거나, 지정된 활동을 수행할 때 혀가 지나치게 느린 속도로 움직이며 다양한 활동을 하는데 혀의 운동변화가 매우 작게 되는 것을 말한다. 경직형 말장애는 주로 삼차신경, 안면신경, 미주신경, 설하신경 및 뇌신경 등을 지배하는 피질연수경로 즉 그림1에서의 직접활성화경로(direct activation pathway) 및 간접활성화경로(indirect activation pathway)의 손상으로부터 발생하는 것으로 알려져 있다[1].

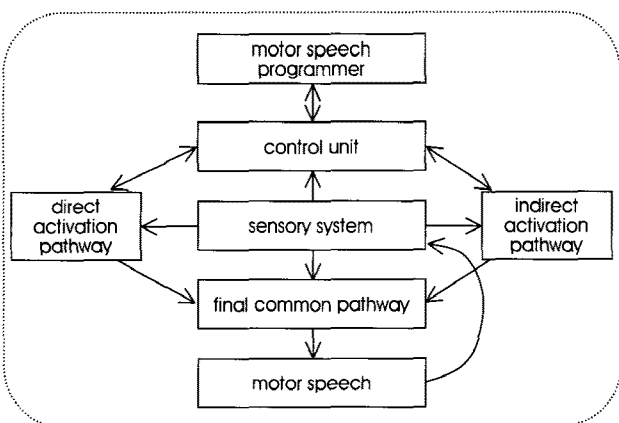


그림 1. 말운동 시스템의 모델  
Fig. 1. Model of motor speech system

말운동장애 치료 시 가장 중점을 두는 부분은 명료도, 말속도, 자연스러움 등이다[7]. 호흡에서 장애가 관찰될 경우 호흡자세를 교정하거나 호흡기능을 향상시키며, 발성단계에서는 목소리 크기와 음높이를 일정하게 유지시키는 연습을 할 수도 있다. 조음단계에서는 입술, 혀 등 발음기관의 운동능력을 증진시킴과 동시에 정확한 발음을 위한 조음훈련을 시행한다[8]. 개발된 시-청각 단서 프로그램은 말운동장애 치료 시 중점을 두는 명료도, 자연스러움 등을 개선하기 위해서 문장들을 제공하고 시-청각 단서 제공 기능을 이용하여 말운동장애인이 스스로 읽기 훈련이 가능하도록 하였다. 또한 읽기 훈련과 동시에 녹음기능을 제공하여 추후에 언어치료 전문가가 분석할 수 있도록 하였고 말소리 개시점 찾기 기능으로 말운동장애인의 말소리 개시점을 말운동장애인과 언어치료 전문가에게 제공하도록 하였다. 청각단서 제거 기능을 이용하여 녹음 파일에서 청각단서를 제거하여 언어치료 전문가가 더 정확하게 발음확인이 가능하도록 하였다.

## III. 시-청각 단서 프로그램 및 알고리즘의 개발

### A. 시-청각 단서 프로그램의 개발

본 프로그램은 말운동장애인이 스스로 읽기 연습을 하는 시-청각 단서 프로그램 부분과 말소리 개시점 찾기 알고리즘 부분 그리고 청각단서 제거 알고리즘 부분으로 나눌 수 있고 이 세 가지 부분의 필요성은 아래와 같다.

시-청각 단서 프로그램은 말운동장애인이 스스로 읽기 훈련을 할 수 있도록 시각단서, 청각단서, 초기지연시간, 주기시간 등의 환경설정이 가능하다. 말운동장애인은 환경설정을 하고 스스로 읽기 훈련을 하여 말소리의 명료도와 자연스러움을 향상시킬 수 있다.

말운동장애인이 시-청각 단서 프로그램을 통하여 읽기훈련한 말소리를 녹음하여 명료도, 말속도, 자연스러움 등을 확인할 경우에는 응용프로그램을 이용하여 전체 녹음 파일을 정밀하게 분석할 필요가 있다. 그러나 녹음된 말소리 개시점 시간만이 필요한 경우에는 응용프로그램을 이용하여 각각의 말소리 녹음파일을 확인할 필요 없이 말소리 개시점 찾기 알고리즘을 이용함으로써 자동적으로 말소리 개시점 획득이 손쉽게 가능하다. 말운동장애인에게 말소리 개시점이 정량적으로 어느 정도인지의 정보를 주기위해서 말소리 개시점을 말운동장애인에게 제공하여 말소리 개시시간을 스스로 인지가 가능하도록 하여 읽기훈련에 참고가 되도록 하였다. 또한 언어치료 전문가가 말소리 개시점 저장파일을 통하여 말소리 개시점을 분석할 수 있도록 하였다.

말운동장애인이 시-청각 단서 프로그램의 청각단서를 이용하여 읽기훈련을 할 경우에 스피커에서 발생한 청각단서와 읽기훈련 말소리가 동시에 마이크를 통하여 녹음이 되어 진다. 이 녹음 파일을 언어치료 전문가가 응용프로그램을 이용하여 말소리 녹음파일의 발음을 분석할 때 청각단서가 모음 발음의 에너지보다 상대적으로 작은 자음 발음에 들어가게 되면 청각단서의 마스킹 현상으로 인

하여 자음의 정확한 발음을 확인하기 어려워진다. 청각단서 제거 알고리즘은 자음에 들어가 있는 청각단서를 제거하여 말운동장애인의 정확한 발음을 언어치료 전문가가 확인 가능하도록 하는 역할을 한다.

1) 시-청각 단서 프로그램의 구성 및 흐름도

본 논문에서 제안하는 시-청각 단서 프로그램은 그림 2에 나타난 바와 같이 크게 나누어 시각 및 청각 단서의 제공 기능, 피험자의 프로그램 실행 목소리 녹음 기능, 말소리 개시점 찾기 기능, 청각단서 제거 기능, 사용자 정보와 훈련자료 및 읽기 시작 위치 등을 저장하는 데이터베이스 기능으로 구성되어 있으며, 프로그램 구현에는 Microsoft사의 MFC 6.0과 The MathWorks사의 MATLAB 7.1을 사용하였다.

· 시-청각 단서 제공 기능: 읽기 훈련 시 사용자에게 시각매체를 통해 문장을 제공한 후 시-청각 단서의 지연시간을 조절하고 시각 단서나 청각단서를 동시에 혹은 각각 단서 제공할 수 있도록 하였다. 시-청각 단서의 제공은 문장의 단어 부분만 선택적으로 제공하도록 하였고 띄어쓰기 공간은 단서를 제공하지 않도록 하였다. 시각단서는 파란 삼각형 모양(▲)을 사용하였으며, 청각단서는 500 Hz의 주파수와 20ms 길이의 메트로놈 응용 소리를 사용하였다.

· 녹음 기능: 녹음 기능은 말운동장애인의 말소리 명료도, 말속도, 자연스러움 등의 확인을 위해 녹음 기능을 제공한다.

· 말소리 개시점 찾기 기능: 기존에는 말운동장애인에 의해서 녹음된 파일을 응용프로그램을 이용하여 언어치료 전문가가 파일을 분석하였다. 이런 경우 비실시간성과 수동적으로 확인되는 특징을 갖는다. 읽기 개시점 찾기 프로그램은 말운동장애인이 읽기 연습을 할 경우 자신의 말소리 개시점 시간을 실시간성과 자동적으로 확인이 가능하고 이 개시점 시간을 토대로 하여 말소리의 개시점을 정량적으로 확인이 가능하다.

· 청각단서 제거 기능: 말운동장애인의 녹음 파일에서 자음에 들어가는 청각단서를 제거하는 기능을 제공한다.

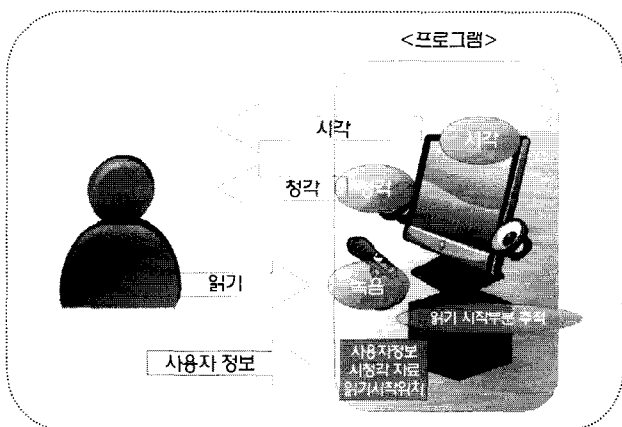


그림 2. 개발된 프로그램의 개념도  
Fig. 2. The concept of the developed program

· 데이터베이스 기능: 사용자로부터 입력된 신상정보, 전문의의 도움을 받아 작성한 읽기 훈련용 문장 자료들, 사용자의 훈련시 검출한 말소리 시작위치 정보 등을 저장하고 관리하는 기능을 제공하도록 하였다.

그림 3은 읽기 훈련의 전체적인 흐름도이다. 사용자 정보를 입력 후, 연습이 필요할 경우에는 예비 읽기 훈련문장을 이용하여 환경 설정 방법을 스스로 배울 수 있다. 본 시험에서는 자음과 모음으로 시작하는 총 16세트의 문장유형 중 한 세트(10문장)를 선택한 다음 시각 및 청각 단서의 제공여부와 녹음 파일 저장 여부 등의 옵션을 선택한다. 선택된 옵션과 함께 문장이 제공된 후 사용자의 말소리가 입력되면, 본 논문에서 제안한 읽기 개시점 찾기 알고리즘에 의해 선택된 문장 제공 후 읽기를 개시하기까지의 시간이 사용자에게 시-청각 단서 프로그램의 실행화면 우측 하단에 말소리 개시점을 제공하고 파일로 말소리 개시점을 저장한다. 선택된 문장의 읽기가 완료되면, 다음 문장의 훈련을 진행할 수 있다. 또한 한 세트의 문장 읽기가 모두 완료되면 다른 세트를 선택하여 읽기훈련을 진행할 수 있다.

2) 시-청각 단서 프로그램의 기능

그림 4는 시-청각 단서 프로그램의 기능들의 상관관계를 나타내고 있다. 피험자의 정보를 입력한 다음 환경설정으로 시각단서와

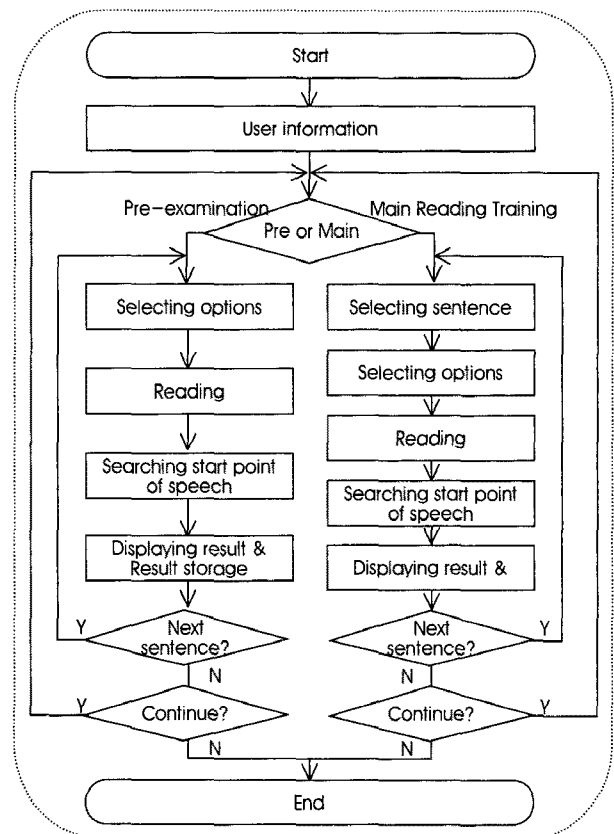


그림 3. 읽기 훈련 프로그램 흐름도  
Fig. 3. Flowchart of the reading training program

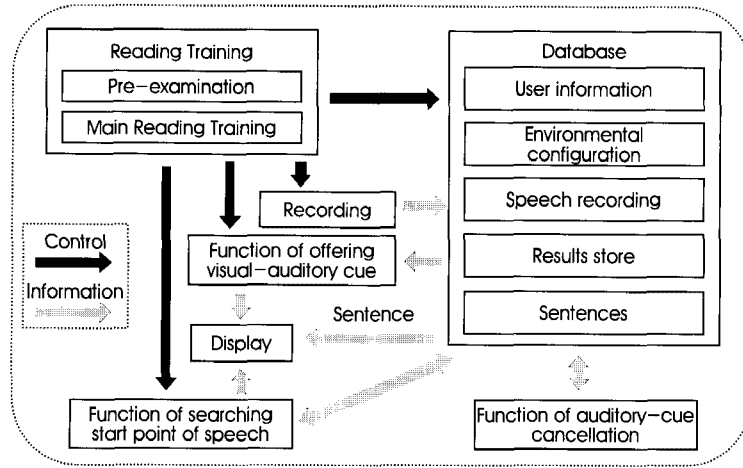


그림 4. 시청각 단서 제공 읽기 훈련 프로그램의 기능  
 Fig. 4. Function of the reading training program offering visual-auditory cue

청각단서의 설정과 초기 지연시간과 주기시간 등의 시간을 설정하고 설정된 정보를 시청각 단서 제공 기능을 이용하여 읽기훈련을 실행한다. 읽기 훈련을 할 때 말운동장애인의 말소리를 녹음한다. 말소리 녹음된 파일을 이용하여 말소리 개시점 찾기 알고리즘을 실행하여 말소리 개시점을 찾고 또한 말소리 개시점을 파일로 저장한다. 청각단서를 제거하기 위해 녹음파일에서 청각단서를 제거한다.

**B. 알고리즘 개발**

1) 정규화를 이용한 말소리 개시점 찾기 알고리즘

마이크로폰을 통해서 들어오는 말소리의 에너지 크기가 하드웨어적인 특성(사운드카드, 마이크의 감도)이나 개인차(남성, 여성, 어린, 아이)이로 발생하는 성량의 크기 차이 문제와 잡음에 의해서 발생하는 소리로 인하여 말소리의 개시점을 찾기 어려운 점이 있었다. 이러한 문제를 해결하기 위해서 녹음된 말소리의 에너지가 각기 다른 말소리 크기로 녹음되면 정량화된 말소리 크기로 변환하여 녹음된 말소리의 에너지 평균값을 동일하게 만들어준다. 녹음된 말소리의 정량화된 크기 비율 즉, CE(coefficient of expansion)를 이용하고 말소리가 없는 일정한 구간에 잡음의 평균을 취하여 RV(reference value)를 정하도록 하였다.

그림 5는 말소리 개시점 찾기 알고리즘의 흐름도이다. 알고리즘에 사용된 녹음 데이터는 다음 실행 버튼이 눌러진 후 15초간 녹음되어진 파일을 이용하여 읽기 개시점을 찾는다. 흐름도를 살펴보면 처음으로 변수를 초기화하고 절대값으로 파일의 값들을 정수로 변환한 다음 표준화시키기 위해서 1초에서 5초까지의 SA(speech average)를 0.0305(16bit 평균값 2000)의 정량화된 크기로 신장시켜준다. 정량화의 크기로 선택한 0.0305는 녹음된 파일을 조사하여 녹음된 말소리가 최대로 녹음될 수 있는 값보다 크게 선정된 평균값이다.

$$CE = 0.0305 / SA \quad (1)$$

식 (1)은 녹음된 말소리와 정량화시킨 크기의 말소리의 비율을 이용하여 CE값을 구한다. 그리고 간단한 25점 이동 평균 필터를 이용하여 간단한 필터를 한 다음 BNA (background noise

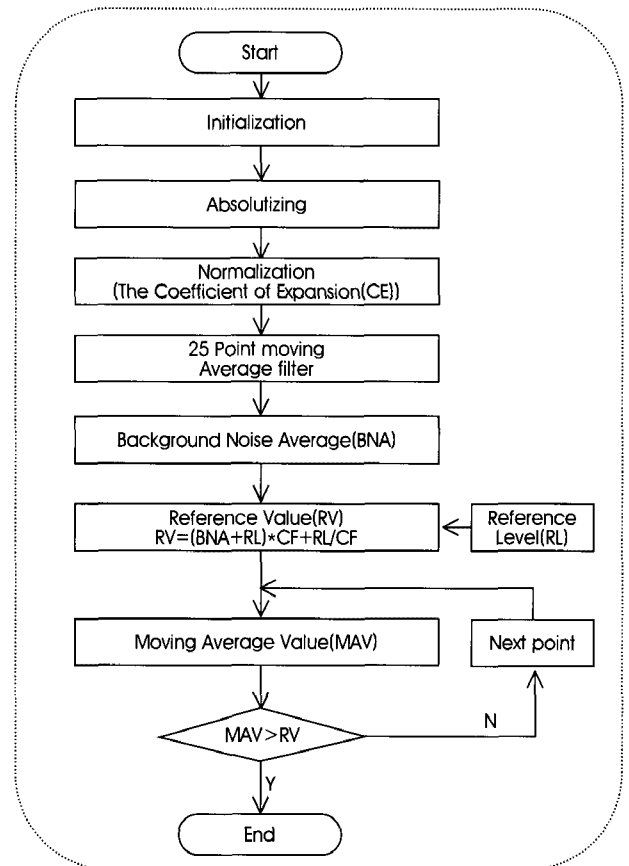


그림 5. 말소리 개시점 찾기 흐름도  
 Fig. 5. Flowchart of searching start point of speech

average value)를 0.1초에서 0.3초까지의 음성이 없는 시간을 이용하여 잡음의 평균을 구하여 BNA 변수에 저장한다.

$$RV = (BNA + RL) * CE + RL / CE \quad (2)$$

식 (2)는 정규화된 공식이다. RV는 BNA와 RL(reference level), CE를 이용하여 RV값을 구한다. 여기에서 RL은 경험에 의한 값이고 0.0046(16bit에서 300)의 정규화된 값을 사용하였다.

$$MAV = \frac{1}{110} y[n] + y[n+1] + \dots y[n+109] \quad (3)$$

식 (3)은 110개의 sample 즉, 약 10ms 간의 MAV(moving average value)을 구하여, 이미 구해진 RV가 MAV보다 큰 경우 읽기 개시점으로 인정하도록 하였다. MAV의 샘플을 10ms로 정한 이유는 10ms보다 짧은 시간이면 짧은 잡음에 의해서 영향을 받는 문제를 해결하고 10ms보다 크게 되면 신호에 따라서 오차가 커지는 문제와 계산량이 커지는 문제점이 있다. 여기에서 개시점으로 찾아진 시간은 대화상자의 화면에 표시하고 파일로 저장하게 된다.

2) 청각단서 제거 알고리즘

청각단서 제거 알고리즘은 MATLAB 7.1을 통하여 실행하였다. 청각단서는 스피커를 통해서 피험자에게 들려주고 피험자의 말소리는 마이크로폰을 통하여 녹음되어질 때 피험자의 말소리와 청각단서는 동시에 녹음된다. 녹음파일을 확인할 때 녹음되어진 청각단서가 말소리에 리듬감을 주어서 청각 단서와 말소리와의 관

계를 알 수 있는 단서를 제공해준다. 언어치료 전문가는 이 녹음 파일을 이용하여 발음의 명료도, 말속도, 자연스러움 등을 확인하게 된다. 말소리는 자음과 모음으로 이루어져 있고 모음에 비해서 자음의 말소리는 상대적으로 작은 음성에너지지를 가지고 있다. 자음의 말소리에 청각단서가 동시에 녹음이 되어지면 청각단서의 마스킹 현상에 의해서 말소리의 발음을 정확하게 분석하기 어려워진다. 자음에 들어있는 청각단서의 마스킹 현상을 제거하기 위해서 청각단서 제거 알고리즘을 사용하였다.

아래의 그림6은 청각단서 제거 알고리즘의 흐름도이다. 순수한 청각단서와 청각단서를 제거하고자하는 부분의 추출은 시간영역에서 시-청각 단서 설정 시간은 초기지연시간과 주기시간을 이용하여 추출하였다. 순수한 청각단서의 추출은 말소리가 없는 녹음 파일 뒷부분에서 청각단서 신호 3개의 평균을 이용하여 사용하였다. 이렇게 3개의 청각단서를 사용하는 이유로 청각단서가 스피커를 통해서 마이크에 녹음이 될 때 청각단서가 스피커와 공간상에서 왜곡이 되어져서 청각단서의 신호들이 달라지는 문제점이 발생하였고 이를 해결하기 위해 3개의 청각단서의 평균을 구하여 사용하였다. 알고리즘은 청각단서를 제거하고자 하는 말소리와 순수한 청각단서 3개의 평균을 구하여 추출한 시간 영역의 소리를 FFT(fast fourier transform)하여 주파수 영역으로 변환한다. FFT된 신호를 이용하여 청각단서를 빼내고 다시 IFFT(inverse FFT)하여 시간영역으로 변환하면 청각단서의 제거가 완료된다. 이렇게 청각단서를 시간영역이 아닌 주파수 영역에서 제거한 이유는 청각단서가 스피커와 공간을 통해서 마이크로폰으로 녹음되는 구간에서 기계적 특성이나 공간의 특성에 따라서 지연시간이 수[ms]에서 수백[us] 정도가 발생하는 문제점이 발생하여 주파수 영역에서 청각단서를 제거하였다.

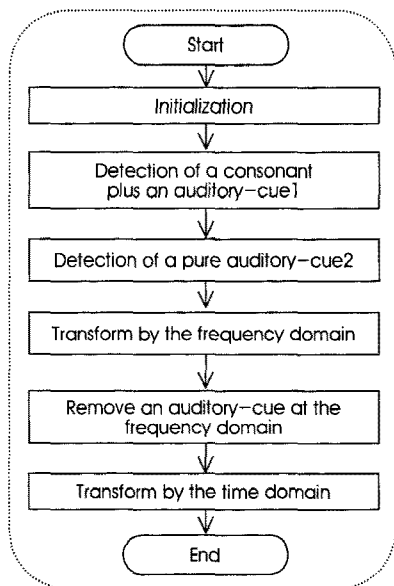


그림 6. 청각단서 제거 흐름도  
Fig. 6. Flowchart of auditory-cue cancellation

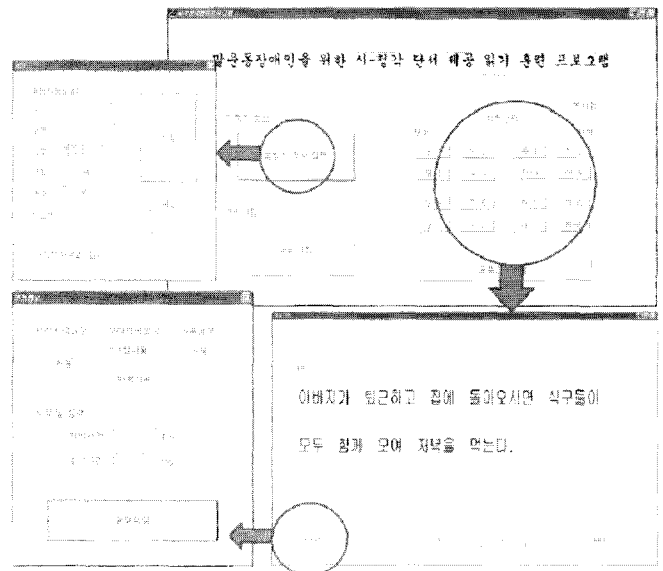


그림 7. 개발된 프로그램의 GUI 샘플들  
Fig. 7. GUI samples of the developed program

### C. 사용자 인터페이스

시-청각 단서 프로그램의 편리함을 위해 사용자 인터페이스를 Windows 다이얼로그 박스 기반으로 프로그램을 실행하였다. 환자 정보는 데이터베이스를 사용하여 신규 등록 시 입력 내용을 저장하도록 하였다. 읽기 훈련 시 시각 혹은 청각 단서를 제공하고 훈련 결과와 함께 저장하도록 하였다.

개발된 프로그램의 인터페이스는 네 가지의 폼으로 구성되어 있다. 즉, 주 대화상자 폼, 사용자의 정보 입력 폼, 시-청각 단서 및 녹음 여부 선택 폼, 읽기 훈련 실행 폼이다. 주 대화상자는 그림 7의 오른쪽 상단의 폼으로 읽기 훈련을 시작할 때 처음 보여지는 대화상자이다. 여기에서 각 훈련에 대한 시작 및 사용자 환경설정 등을 할 수 있다. 피험자 정보 입력 버튼을 클릭하면 상단부분 왼쪽의 폼이 생성되어 나타나며, 폼에 기록한 사용자의 정보를 데이터베이스에 저장할 수 있다.

읽기 훈련은 예비시험과 본시험으로 나누어 실행할 수 있는데 본 시험에 사용하는 훈련 문장 세트는 언어치료 전문가의 추천문장으로 구성하였다. 읽기 훈련을 시작할 때 그림 7 하단부분 왼쪽의 폼에서 시-청각 단서 사용여부를 선택한 후 훈련을 진행할 수 있다.

## IV. 실험 및 결과

### A. 녹음 데이터의 획득

개발된 프로그램의 정량적 검사를 위해 6명의 피험자(남성 3명, 여성 3명)를 통하여 실험하였다.

말소리 개시점 찾기 알고리즘의 실험은 4가지 잡음 환경에서 실험을 실시하였으며 실험환경은 그림 8과 같다. 실험은 의자에 앉은 상태에서 진행되었으며, 시-청각 단서는 피험자 정면의 모니터

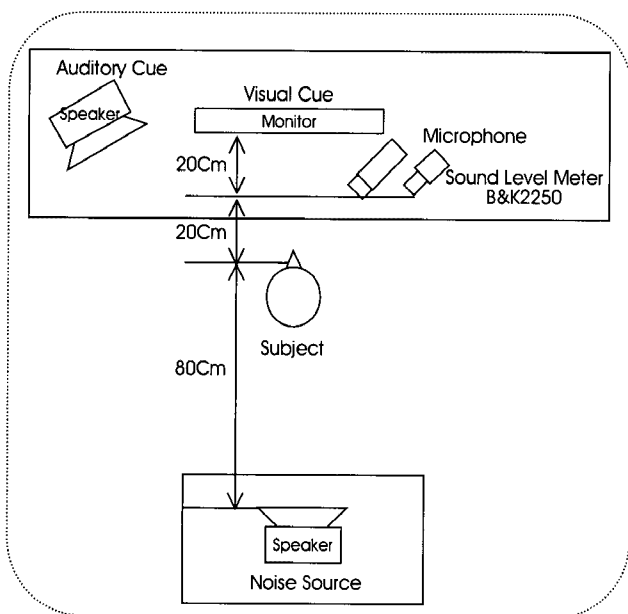


그림 8. 실험 셋업  
Fig. 8. Experimental setup

와 스피커에서 주어지고, 잡음은 정면 쪽 마이크에서 1[m] 떨어진 피험자 뒤편의 스피커에서 발생하도록 하였다. 피험자의 말소리 입력은 피험자 정면 20[cm]이내의 마이크에서 입력 받도록 하였다.

잡음환경은 실내 잡음, 백색 잡음, 자동차 실내잡음, babble 잡음의 4가지 환경을 조성하여 실험하였다. 잡음 측정은 Bruel & Kjaer사의 사운드 레벨 메타(Hand-held Analyzer Type 2250)를 이용하여 A-weighting을 주어 측정하였고 실내 잡음 환경은 43.3[dBA] 그 외의 잡음 환경들은 60[dBA]로 하여 실험하였다.

### B. 알고리즘 및 GUI 평가 방법

#### 1) 말소리 개시점 찾기 오차 평가 방법

절대시간과 말소리 개시점 찾기 알고리즘을 통하여 획득한 시간의 차이를 평가하였다. 절대시간은 응용프로그램(Cool Edit)을 통하여 개시점을 수동으로 검사하여 얻은 시간이고 말소리 개시점 찾기 알고리즘을 통하여 획득한 시간은 시-청각 단서 프로그램을 통하여 녹음된 파일을 이용하여 말소리 개시점 찾기 알고리즘으로 얻어진 시간이다.

#### 2) 청각단서 제거 결과 분석 및 평가 방법

청각단서 제거 결과 분석방법은 시각적 분석 방법과 청각단서를 제거하기 전후의 말소리 듣기 평가 방법 그리고 청각단서가 제거된 말소리를 MOS(mean opinion score) 이용하여 음질을 평가하였다. 시각적 분석방법에는 주파수영역 분석방법과 스펙트로그램 분석방법이 있는데 전자의 경우는 청각단서의 제거를 주파수 영역에서 시각적으로 보여주기 위하여 분석하였고 후자의 경우는 청각단서의 제거를 시간영역과 주파수영역에서 동시에 시각적으로 보여주기 위하여 분석하였다. 청각단서제거 말소리 듣기 평가 방법은 자음에 청각단서가 들어있는 1음절 말소리와 청각단서 제거 알고리즘으로 청각단서를 제거한 1음절 말소리를 피험자에게 듣기 실험을 실시하고 맞은 단어는 1점으로 틀린 단어는 0점으로 점수를 주어 평가하였다. 듣기 평가의 경우 학습효과를 배제하기 위해서 청각단서 전과 후의 정보를 제공하지 않았고 또한 듣기 평가 단어들의 들려주는 순서를 바꾸어서 듣기 평가를 함으로서 학습효과가 없도록 하여 실험하였다. 청각단서 제거 말소리 음질 평가 방법은 청각단서가 제거되기 전의 1음절 말소리를 1점, 순수한 1음절 말소리를 5점, 청각단서를 제거한 후 1음절 말소리 MOS를 이용하여 평가하였고 피험자에게 들려주는 순서는 청각단서가 제거되기 전, 순수한 말소리, 청각단서를 제거한 후 말소리의 순서로 들려 주어 평가하였다.

#### 3) GUI 사용 결과 평가 방법

GUI 사용 결과의 평가는 피험자로부터 시-청각 단서 프로그램을 직접 사용하도록 하여 각 항목에 대한 MOS를 주어 프로그램의 유용성을 평가하였다.

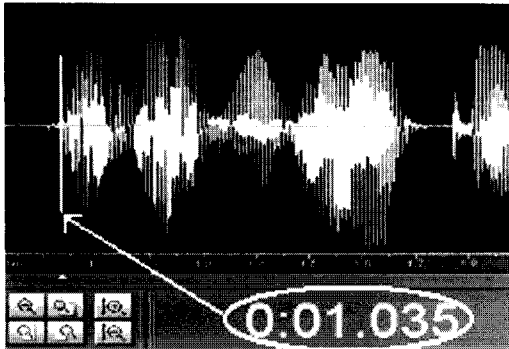


그림 9. 말소리의 개시 시간찾기의 결과  
 Fig. 9. The result of searching starting time of the speech

C. 실험 결과 분석

4가지의 잡음환경에서 피험자 당 10문장씩 읽어 개인별 40문장을 읽고 6명(남성 3명, 여성 3명)이 총 240문장을 읽어 실험을 실시하였다. 그림 9는 녹음된 신호에 대해 절대시간을 Cool Edit를 사용하여 말소리 개시점을 수동으로 구하였다. Cool Edit의 가로축은 시간축이고 세로축은 진폭축이다. 피험자가 문장 읽기를 시작한 실제 말소리 개시 시간을 표시한 화면으로 말소리의 실제 시작 시간이 1.035[sec]라는 것을 보여주고 말소리 개시점 찾기 알고리즘으로 1.03628[sec]를 찾아서 두 시간의 차이가 1.28[ms]로 검출되었다.

표 1은 절대시간과 본 프로그램이 추출한 읽기 시작 위치와의 오차를 구한 결과이다. 표에서 각 셀의 값은 잡음의 종류별로 60문장에 대한 추적 에러의 평균값과 표준편차를 구한 것이며 전체적으로는 표 1의 최하단에 표시한 것처럼 전체의 오차의 결과는  $7.042 \pm 8.99[ms]$ 의 오차결과를 보여준다.

그림 10은 청각 단서 제거 전과 후를 시각적으로 보여주고 있다. 가로축은 주파수를 나타내며 세로축은 에너지 크기이다. 그림 10(a)는 청각 단서와 음성신호가 들어 있는 부분을 주파수 영역에서 보이는 신호, 그림 10(b)는 말소리가 없는 부분의 청각 단서의 신호 3개의 합인 주파수 영역 신호, 그림 10(a)의 말소리(청각 단서

표 1. 말소리 개시점의 추출 오차의 결과  
 Table 1. Results of extraction error of a speech start point

잡음환경	잡음크기[dBA]	비교오차 [ms]	N
실내 잡음	43.3	$4.758 \pm 6.555$	60
백색 잡음	60	$5.893 \pm 7.537$	60
자동차 잡음	60	$8.935 \pm 11.126$	60
Babble noise	60	$7.545 \pm 9.604$	60
전체	-	$7.042 \pm 8.99$	240

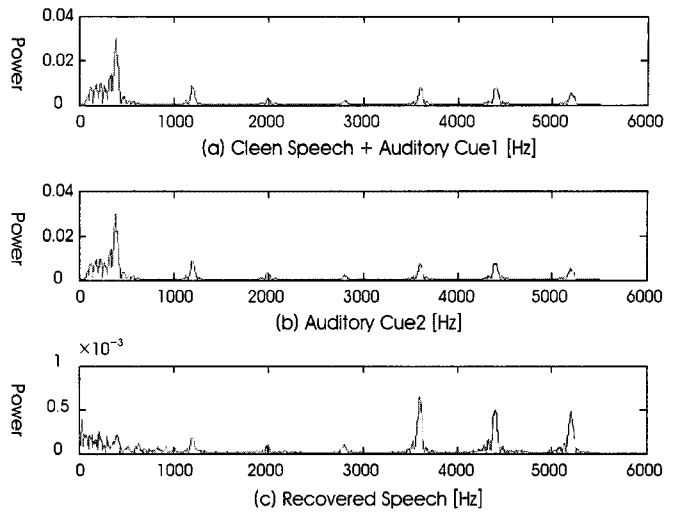


그림 10. 청각단서를 제거한 결과  
 Fig. 10. The results of removing auditory-cue

와 말소리의 신호)를 그림 10(b)의 청각 단서 신호로 제거한 결과가 그림 10(c)에 나타나 있다.

그림 11은 청각 단서 알고리즘을 통하여 청각 단서를 주파수 영역에서 제거한 결과를 스펙트로그램 상에서 확인하였다. 그림 11(a)는 깨끗한 말소리의 스펙트로그램을 나타내고 그림 11(b)는 말소리에 b1 청각 단서가 포함되어 있는 신호, 그림 11(c)는 청각 단서 제거 알고리즘을 통해서 c1에서 보이는 것과 같이 청각 단서가 제거된 결과를 보여주고 있다. 그림 11(b)와 그림 11(c)의 b1과 c1에서 청각 단서가 3000[Hz] 이하의 주파수에서는 많이 제거된 결과를 보이나 3000[Hz] 이상의 주파수에서는 육안으로 확인이 가능할 정도로 청각 단서가 제거되지 않았다. 이렇게 3000[Hz] 이상에서 스피커, 마이크, 공간상에서 왜곡이 상대적으로 크게 발생하기 때문인 것으로 판단된다.

표 2는 청각 단서 제거 전과 후 말소리 받아 적기를 통하여 정답이면 1점 오답이면 0점으로 하여 점수를 주었고 맞춘 점수의 평균을 퍼센트로 환산하였다. 사용된 1음절 단어는 병원기반의 청력검사에서 사용되는 단어를 사용하였다. 학습효과를 없애기 위해서 청각 단서 제거 전과 후의 정보를 제공하지 않고 16개의 단어들의 순서를 바꾸어서 피험자가 추측하지 못하도록 하여 실험을 하였다. 평가 결과는 신호처리전의 인식률은  $51.042 \pm 12.192\%$ 를 획득

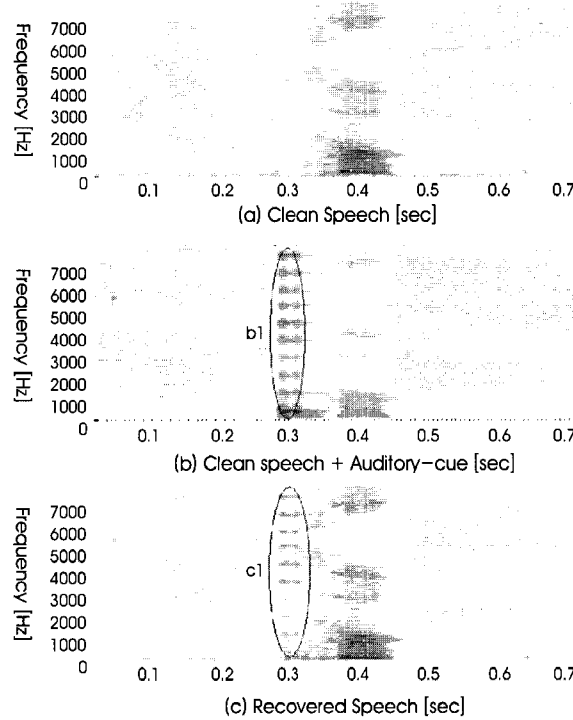


그림 11. 스펙트로그램에서 청각단서 제거의 결과  
 Fig. 11. Results of auditory-cue cancellation in spectrogram

하였고 신호처리 후의 인식률은  $76.042 \pm 5.609\%$ 를 획득하여 인식률이  $25 \pm 9.547\%$  향상되어짐을 알 수 있다. 청각단서 제거 전 듣기 시험의 오답을 분석해 보면 마스킹 현상에 의해서 오답이 발생

하는 것으로 보이고 청각단서 제거 후의 오답을 분석해 보면 3000[Hz] 이상의 주파수 성분이 남아있는 청각단서가 음의 변형을 일으키는 것으로 분석되어진다.

표 2. 청각단서 제거 전과 후 말소리의 듣기 점수

Table 2. Grade of hearing of the speech before and after cancellation on an auditory-cue

피험자	청각단서 제거 전	청각단서 제거 후	N
피험자1	10	12	16
피험자2	5	11	16
피험자3	10	12	16
피험자4	7	12	16
피험자5	7	12	16
피험자6	10	14	16
전체 정답률(%)	$51.042 \pm 12.192$	$76.042 \pm 5.609$	96

표 3. 제거된 청각단서의 말소리에 음질 점수

Table 3. Grade of the quality at the speech of an removed auditory-cue

피험자	평가(5점 만점)	N
피험자1	$4.31 \pm 0.583$	16
피험자2	$4.25 \pm 0.559$	16
피험자3	$3.94 \pm 0.658$	16
피험자4	$4.25 \pm 0.901$	16
피험자5	$3.06 \pm 0.658$	16
피험자6	$4.25 \pm 0.661$	16
전체	$4.01 \pm 0.81$	96



표 4. GUI 사용자 만족도 조사

Table 4. Investigation of user satisfaction about GUI

번호	질문 항목	평가(만족도)	N
1	GUI의 편의성	3.83±0.408	6
2	메뉴 구성 체계에 대한 만족도	3.50±0.548	6
3	조작의 용이성	3.30±0.516	6
4	프로그램 실행 오류 여부	5.00±0.000	6
총 만족도		3.92±0.775	6

표 3은 청각단서 제거 알고리즘을 통한 음질 평가의 실험은 16 세트(순수한 말소리, 청각단서가 포함된 말소리, 청각단서 제거 알고리즘으로 처리한 말소리)를 들려준다. 청각단서 제거전의 말소리(1점)와 순수한 말소리(5점)를 들려준 후 청각단서 제거 후 말소리에 대한 MOS를 평가하였다. 평가 결과는 4.01±0.81의 점수를 획득하였고 대체적으로 만족스럽다는 평가를 받았다.

본 프로그램 인터페이스에 대한 만족도를 아래의 표 4의 항목에 대해 MOS를 조사하였다. 각 항목에 대해 매우만족 5점, 매우불만족 1점을 기준으로 하여 평가하도록 하여 전체적으로 3.92±0.775의 평가를 기록, 대체적으로 만족스럽다는 평가를 받았다.

## V. 결론

본 논문에서는 객체지향 프로그램을 이용하여 말운동장애인을 위한 시-청각 단서 프로그램을 개발하였고, 건강한 사람 6명(남성 3명, 여성 3명)을 대상으로 한 실험을 통해 시-청각 프로그램의 유효성을 확인하였다. 시-청각 단서 프로그램은 사용자가 간편하게 사용할 수 있도록 GUI(대화상자)를 구성하였으며, 읽기 훈련 시 시각 단서와 청각 단서의 제공 기능 및 사용자의 말소리 개시점 시작 위치의 추출 기능과 청각단서 제거하는 기능을 가지고 있다.

개발된 프로그램의 개시점 찾기 알고리즘의 실험을 위해서 4가지(실내 잡음, 백색 잡음, 화이트 잡음, babble 잡음)의 잡음 환경을 조성한 후에 각 잡음 환경 당 10개의 문장에 대해 피험자별로 읽기 훈련 실험을 실행한 다음 응용프로그램을 이용하여 수동으로 구한 절대시간과 말소리 개시점 찾기 알고리즘을 통하여 추출된 시작 위치와의 오차를 구하는 정량적 평가 실험을 하였고 실험 결과 7.042±8.99[ms]의 오차를 획득하였다. 그리고 자음의 말소리에 청각단서가 들어갈 경우 말소리의 듣기 평가와 음질 평가를 실행하였다. 듣기 평가의 경우 1음절 모음단어에 청각단서가 들어간 말소리의 듣기 평가와 청각단서 제거 알고리즘을 통하여 청각단서를 제거한 말소리의 듣기 평가를 실행하였고 전자의 경우 51.042±12.192%를 획득하였고 후자의 경우 76.042±5.609%를 획득하

여 인식률이 25±9.547% 향상되어졌다. 또한 청각단서 제거 알고리즘의 음질 평가의 경우 청각단서를 제거한 후 피험자에게 들려주어 MOS를 측정하였고 말소리의 음질 점수는 4.01±0.81을 획득하였고 알고리즘을 통한 말소리는 전체적으로 음질이 향상되었다. 그리고 마지막으로 인터페이스에 대한 정량적 만족도 조사를 MOS를 통하여 구하였고 인터페이스에 대한 만족도는 5점 만점에 3.92±0.775점을 기록하였다. 개발된 본 프로그램은 향후 말운동장애인의 평가 및 재활에 충분한 도움이 될 수 있으리라 사료된다.

## 참고문헌

- [1] J.R. Duffy, "Motor speech disorders : Substrate, differential diagnosis, and management," St. Louis. : Mosby., 1995.
- [2] K.C. Hustad, T. Jones, S. Dailey, "Implementing speech supplementation strategies: effects on intelligibility and speech rate of individuals with chronic severe dysarthria," *J. Speech. Lang. Hear. Res.*, Vol. 46, No. 2, pp.462-474, 2003.
- [3] M.A. Pilon, K.W. McIntosh, M.H. Thaut, "Auditory vs visual speech timing cues as external rate control to enhance verbal intelligibility in mixed spastic-ataxic dysarthric speakers: a pilot study," *Brain. Inj.*, Vol. 12, No. 9, pp.793-803, 1998.
- [4] K.M. Yorkston, V.L. Hammen, D.R. Beukelman, C.D. Traynor, "The effect of rate control on the intelligibility and naturalness of dysarthric speech," *J. Speech. Hear. Disord.*, Vol. 55, No. 3, pp550-560, 1990.
- [5] M.A. McHenry, "The effect of pacing strategies on the variability of speech movement sequences in dysarthria," *J. Speech. Lang. Hear. Res.*, Jun, Vol. 46, No. 3, pp.702-710, 2003.
- [6] V.L. Hammen, K.M. Yorkston, "Speech and pause characteristics following speech rate reduction in hypokinetic dysarthria," *J. Commun. Disord.*, Vol. 29, No. 6, pp.429-444, 1996.
- [7] 김선희 외 역, "의사소통장애," 학지사, 2002.
- [8] H. Shames, N. Anderson, "Human Communication Disorders: An Introduction," Boston. : Allyn, & Bacon, 2002.