

# 만 3-5세 유아의 한국어 음성 데이터베이스 구축

## Speech Database for 3-5 years old Korean Children

유재권\*, 이경옥\*\*, 이경미\*  
덕성여자대학교 컴퓨터학과\*, 덕성여자대학교 유아교육학과\*\*

Jae-Kwon Yoo(chhj2010@duksung.ac.kr)\*, KyungOk Lee(oaklee@duksung.ac.kr)\*\*,  
Kyoung-Mi Lee(kmlee@duksung.ac.kr)\*

### 요약

유아는 만3~5세 사이에 언어 능력이 빠르게 발달하게 된다. 유아의 언어발달에 맞는 다양한 경험을 위해서는 그 시기에 맞는 콘텐츠 개발이 필요하다. 다양한 콘텐츠 개발을 위해 유아에 맞는 음성 인터페이스를 이용하는 것이 필요하지만, 한국어에서는 유아를 대상으로 한 데이터베이스가 구축이 되지 않았다. 본 논문에서는 한국어에서 만 3~5세 유아들의 객관화되고 정확한 음성 데이터 수집을 설계하기 위하여 발달 시기에 맞는 적절한 단어 선정 및 성인과 다른 유아의 행동 특성 유형을 파악하는 과정을 거쳐 음성 데이터베이스를 구축하였다. 단어의 경우 MCDI-K에서 두 단계를 걸쳐 선정하였고, 유아는 한 단어 당 세 번씩 발성하였다. 이렇게 수집된 음성 데이터는 유아별, 단어별 파일 토큰화 과정을 거쳐 데이터베이스로 구축되었다. 한국어 유아 음성 데이터베이스는 웹 페이지를 통한 기술 이전을 할 계획이며, 이를 통하여 유아들의 언어 발달에 유익한 다양한 콘텐츠 개발에 그 역할을 담당할 것으로 기대한다.

■ 중심어 : | 음성 인식 | 음성 데이터베이스 | 유아 콘텐츠 | 음성 인터페이스 |

### Abstract

Children develop their language skill rapidly between age 3 and 5. To meet the child's language development through a variety of experiences, it is necessary to develop age-appropriate contents. So it needs to develop various contents using speech interface for children, but there is no speech database of korean children. In this paper, we develop speech database of 3 to 5 years old children in korean. For collecting accurate children's speech, child education experts examine in the speech database development process. The words for database are selected from MCDI-K in two stage and children speak a word three times. Such collected speech are tokenized by child and word and stored in database. This speech database will be transferred through web and, hopefully, be the foundation of development of children-oriented contents.

■ keyword : | Speech Recognition | Speech Database | Children Contents | Speech Interface |

## I. 서론

컴퓨터 기술의 발전과 인터넷, 모바일의 보급은 학습

환경의 새로운 시대를 열고 있다. 현재 시대에 맞는 교육 환경 중 하나는 CAI(Computer-Assisted Instruction) 개념이다. CAI는 교육 환경의 접근성을 용이하게 하여

학습 시간을 단축시켜주며 개개인이 필요한 학습을 선택할 수 있도록 하여 학습 효과를 증진시킨다[1]. 이런 정보통신 기술을 이용하여 시간과 장소에 구애받지 않는 교육활동의 학습 효과의 증진은 이미 이전 연구를 통해 입증된 사실이다[2-4]. 현재 우리는 교실에서 멀티미디어를 활용하지 않은 교육은 더 이상 생각할 수 없게 되었고 최근에는 유치원 교실은 물론이고 가정마다 컴퓨터가 설치되어 있어 유아들도 누구나 쉽게 컴퓨터와 인터넷을 접할 수 있다. 이는 유아에게 주입식 교육이 아닌 다양한 형태의 실제적이고 유의미한 교육을 받을 수 있는 기반이 마련되었음을 의미한다.

유아 언어 발달은 만3~5세까지 급격하게 발달하게 된다. 유아의 언어는 읽기, 쓰기, 듣기, 말하기로 구성되지만 유아의 언어발달은 이 모든 영역들이 통합되어 발달해간다. 그러나 유아들이 자신의 의사를 조리 있게 말하는 능력은 또래관계와 더불어 사회적으로 유능한 유아가 되기 위해 필요한 요소이다. 유아의 언어 발달 능력을 향상시키기 위한 방법 중 하나로 놀이를 통한 다양한 경험을 해보는 것이다[14][15]. 유아의 언어 발달에 맞는 다양한 경험을 위해서는 그에 맞는 다양한 콘텐츠가 개발되어야 한다. 유아용 음성인식기가 개발 되면 다양한 놀이를 통해 언어 능력 발달 뿐 아니라 놀이의 목적을 통한 유아의 발달에도 도움을 준다. 이전 연구에서 음성 인터페이스를 이용한 유아용 보드게임의 개발이 있었고, 유아 음성 인터페이스는 유아의 몰입도 향상에 도움을 주어 놀이 효과를 증진시켰다[17]. 이처럼 다양한 콘텐츠 개발을 위해 유아에 맞는 음성 인터페이스 개발이 필요하지만, 현재 한국에서 유아에 맞는 음성 데이터베이스는 구축이 되지 않은 실정이다 [5].

유아의 언어 발달 능력을 향상시키기 위해서는 유아에게 맞는 음성 인터페이스를 개발이 필수적이다. 하지만 유아의 음성 데이터베이스가 구축이 되지 않은 상황에서 유아에 맞는 음성 인터페이스 구축은 어렵다. 한국어 기반의 성인 음성 데이터베이스는 구축이 되어 있지만 언어학적으로 성인과 유아는 해부학적, 생리학적 차이로 성인 음성인식기로 유아의 음성을 실험하게 되면 인식률에 큰 차이가 있다[6]. 따라서 유아에 맞는 음

성 인터페이스를 구축하기 위해 대상에 맞는 음성 데이터베이스가 필요하다.

성인과 달리 유아들의 음성 데이터를 수집하기 위해서는 먼저 유아들의 연령에 따른 단어의 수준과 발음의 차이가 있음을 인식해야 한다. 또한 유아의 데이터를 수집할 경우 성인과 달리 한 번에 단어를 녹음하는 단어의 수와 시간을 조절하여 정확한 데이터를 수집할 수 있도록 세심한 설계가 필요하다.

본 논문에서는 한국어에서 만 3~5세 유아의 음성인식에 기초한 교육 콘텐츠 개발을 위해 음성 데이터베이스를 구축하는 것을 목적으로 하였다. 이러한 목적을 달성하기 위해 II장에서는 국내외 유아 음성 데이터베이스 구축 현황을 정리하고, III장에서는 유아 음성 데이터베이스 수집의 일련의 과정을 서술했다. IV장에서는 수집된 음성 데이터베이스의 정리와 상세한 구축 과정을 기술했다. V장에서는 본 연구의 전반적인 내용을 정리하고, 향후 활용 방향에 대해 기술하고 결론을 맺는다.

## II. 배경 연구

본 장에서는 국내외 유아 음성 데이터베이스 구축 현황을 살펴본다. 국외의 경우 다양한 나라에서 유아를 포함한 음성 데이터베이스를 구축하고 관련된 연구가 진행되고 있다. 하지만 한국에서는 만3~5세 대상의 음성데이터베이스는 구축되지 않았다.

### 2.1 국외 유아 음성 데이터베이스 구축 현황

표 1. 국외 유아 음성 데이터베이스 구축 현황

음성DB의 종류	kidCC [7]	kidAM [7]	SpeeCon [8]	Childt[9]
언어	English	English	Swedish	Italian
나이	8-13세	5-10세	4-8세	7-13
화자수	11명	395명	200명	171
구축내용	400개의 문장	376 단어	숫자열, 명칭, 아동문학	아동 문학 문장

국외에서 유아 음성 데이터베이스 구축 현황은 [표 1]

과 같다. 국외에서는 미국, 스웨덴, 이탈리아에서 유아를 포함한 음성 데이터베이스가 구축되어 있었다. 우선 미국에서 구축된 음성데이터베이스는 kidCC와 kidAM이 있다[7]. KidCC는 “Go to Lotus organizer”와 같은 컴퓨터 명령어와 제어 문구를 다섯 명의 남자와 여섯 명의 여자가 참여했다. 실험을 목적으로 구축된 kidAM은 395명의 유아가 참여했고 유아가 자주 사용하는 단어 중심의 음성을 수집했고, 각각의 유아는 70단어를 발성한다.

스웨덴에서 구축된 유아를 포함한 음성데이터베이스는 SpeeCon이다[8]. 이는 성인과 유아의 음향학적 특징 차이 실험을 위해 구축된 것으로 유아의 음성 데이터베이스는 훈련과 실험에 사용된다. 데이터베이스의 구축의 목표는 소비자 어플리케이션 개발을 위한 것이므로 인터페이스가 사용될 장소로 예상 되는 공공장소 및 사무실 환경에서 수집되었다. 단어는 숫자열과 명칭 아동 문학의 문장을 포함한다. 남녀의 성비는 4:6으로 구성된다.

마지막으로 이탈리아에서 ChildIt 음성데이터베이스를 구축했다[9]. 데이터베이스 구축에 참여한 유아는 아동 문학에 나오는 58 혹은 65개의 문장을 읽는다. 이는 이탈리아 언어에서 유아의 음성 특징과 음성 인터페이스 개발을 위해 사용되고 있다.

## 2.2 국내 유아 음성 데이터베이스 구축 현황

한국에서 음성 데이터베이스를 구축하고 있는 ETRI의 공통 음성 데이터베이스 자료에 따르면 만3~5세 유아 대상의 음성 데이터베이스는 구축이 되지 않은 실정이다. ETRI는 다양한 명사와 문장으로 각종 인터페이스에 사용될 수 있는 사용 환경을 고려하여 음성 DB를 구축하고 배포하고 있다.

한국의 공동 이용을 위한 음성 언어 자원 구축 현황을 조사한 결과 초등학교 대상으로 구축된 음성 데이터베이스가 있다[5]. 인원은 500명으로 숫자, 명령어, 제어어와 단위로 구성된 고립단어 중심으로 구축되어 있다. 구축된 데이터베이스는 완구, 교육용 소프트웨어를 위한 음성 인터페이스 개발을 위한 것이다. 다양한 분야의 응용이 가능한 발성 목록의 구성은 4연 숫자 340종,

PBW 452종, 명령어와 지시어 400종, 단독숫자 41종으로 구성되어 있다. 구성된 발성 목록은 총 1,233종으로 1명의 학생이 발성하기에는 양이 많기 때문에 단독 숫자의 경우를 제외한 목록은 20개의 세트로 나누어 수집했다. 수집된 데이터의 남녀 성비는 1:1이며 1인당 발성량은 100~101단어이다. 수집 환경은 추후 음성인터페이스가 활용된 어플리케이션을 사용할 수 장소로 예상되는 사무실 또는 가정집에서 이루어졌다.

## III. 유아 음성 데이터 수집

본 장에서는 한국어에서 만 3~5세 대상의 유아 음성 데이터베이스 구축 과정을 설명한다. 유아 음성 수집 설계 및 단어 선정을 비롯한 유아 교육 기관과 유아 섭외 과정 및 음성 데이터베이스 설계 및 구축에 관한 전 과정을 소개한다.

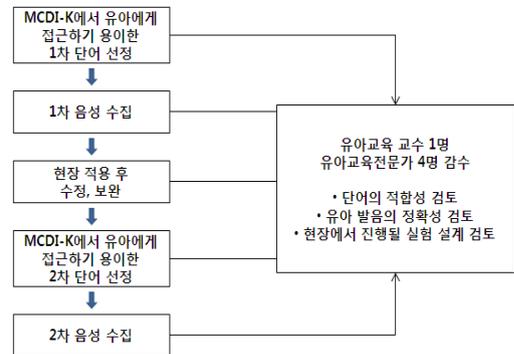


그림 1. 음성 수집 설계 과정

### 3.1 유아 음성 수집 설계

유아의 음성 수집을 위하여 [그림 1]과 같은 단계로 설계했다. 각 단계에서는 유아들의 객관화되고 정확한 음성 수집을 설계하기 위하여 유아교육전문가의 검토를 통하여 음성 수집의 단계를 진행하도록 설계 하였다.

### 3.2 유아 교육 기관 및 유아 섭외

만 3~5세 대상의 유아 음성데이터베이스 수집은 추

후 음성 인터페이스가 장착된 어플리케이션의 사용이 예상될 장소인 유치원에서 이뤄졌다.

유치원은 덕성여자대학교 아동게임연구센터를 통해 섭외되었고, 서울과 경기 지역의 유치원을 방문할 수 있었다. 방문 전에 유아 음성 데이터베이스의 수집 목적과 앞으로의 발전 가능성, 방문 날짜, 참여연구원을 기재하여 기관의 동의를 구했다. 또한 명확한 음성 데이터베이스를 수집하기 위해 유치원의 조용한 빈 교실에서 데이터를 수집하였다.

음성 데이터베이스 수집 시 유아의 성별과 나이, 거주지, 형제관계 정보도 포함하여 설계했다.

### 3.3 1차 유아 음성 수집

유아 음성 수집 설계와 기관 섭외 후 음성 시험 수집을 위해 MCDI-K(MacArthur Communicative Development Inventory - Korean)를 참고하여 1차 단어를 선정하였다. 단어 선정 완료 후 수집 환경에 대한 유아의 행동 유형을 분석하기 위해 1차 유아 음성 수집을 진행했다. 이 과정에서 녹음 시간방법에 대한 보완점을 발견하고 2차 음성 수집 방법을 수정했다. 또한 선정된 단어 중 인식이 낮은 단어를 발견하고 2차 단어 수집 과정에 반영했다.

#### 3.3.1 1차 단어 선정

국외의 유아를 포함한 음성 데이터베이스 설계 시 유아 사용될 수 있는 단어 혹은 문장으로 구성되어있다. 유아 음성 데이터베이스의 목적은 유아의 언어 발달 중 말하기 능력을 보완하기 위한 것으로 다양한 어플리케이션에 활용될 수 있어야 한다. 따라서 유아가 그 시기에 배울 수 있는 단어를 중심으로 선정되었다. 단어 선정은 한국의 영유아가 사용하는 어휘를 설계한 MCDI-K를 바탕으로 유아교육전문가를 통해 내용의 타당성을 검토받았다. 유아 대상의 다양한 체험형 게임 콘텐츠에서 사용될 수 있는 단어 개념으로 구성되어 있으며 1차 선정된 단어는 [표 2]와 같다.

표 2. 1차 단어 선정

개념	색	빨강, 파랑, 노랑, 초록, 분홍, 검정, 보라, 흰색, 주황, 하늘색
	형	동그라미, 세모, 네모, 별, 달, 하트, 삼각형, 사각형, 원
	수	영, 일, 이, 삼, 사, 오, 육, 칠, 팔, 구, 십공, 하나, 둘, 셋, 넷, 다섯, 여섯, 일곱, 여덟, 아홉, 열
	명칭 수	한 개, 두 개, 세 개, 네 개, 다섯 개, 한 마리, 두 마리, 세 마리, 네 마리, 다섯 마리, 한 명, 두 명, 세 명, 네 명, 다섯 명, 한 사람, 두 사람, 세 사람, 네 사람, 다섯 사람
	시간	먼저, 나중, 지금, 아직, 내일, 어제, 오늘, 아까
	위치	위, 아래, 오른쪽, 왼쪽, 앞, 뒤, 옆, 안, 밑
지시어	보기	확대, 축소, 회전
	실행	예, 아니오, 시작, 다시, 그만, 처음으로
단어	제목	누구일까요, 놀이터에서
	동물	여우, 사자, 코끼리, 토끼, 기린, 호랑이, 말, 오리, 돼지

#### 3.3.2 1차 유아 음성 수집 방법

이전 연구의 음성 수집 방법을 참고하여 노트북의 스크린을 통해 그림과 글자가 조합된 화면을 준비했다. 준비된 화면은 성인 진행자에 의해 유아가 혼자 한 단어 당 세 번씩 발성하도록 했다. 녹음 진행 시간은 10~15분으로 약 50개의 단어를 3회씩 150번의 발성을 했다.

선정된 단어 중에 저조한 인식을 보이는 단어가 발생했다. 단어의 예로는 수에서 영, 명칭 수, 제목에 해당되는 단어였다. 특정 숫자열 '영'은 '육'과 단어가 혼동되었으며 이는 두 단어의 'ㄱ'와 'ㅇ'가 반모음 /y/를 포함한 이중 모음이기 때문이다[6]. 명칭 수의 경우 중복되는 개, 마리, 명, 사람의 단어 때문에 인식률의 저하가 발생했다. 제목에 해당되는 단어는 다섯 개의 음절로 이루어져 다른 단어에 비해 유아가 발성하는 시간이 길었고, 같은 음절로 인해 인식에 혼동이 일어남을 확인했다. 이런 사실을 근거로 1차 음성 수집의 보완점을 발견했다.

#### 3.3.3 1차 음성 수집 보완

유아들의 집중시간으로 인해 한 번에 많은 단어를 수집할 경우 명확한 데이터 수집에 제한이 있었다. 따라서 예비연구를 통하여 한 번에 수집할 수 있는 단어를 측정하여 한 번에 수집되는 단어의 수를 제한하였다. 따라서 기존 유아 연구에서 사용한 25~40개의 단어를

5~10분 정도 수집한 것을 참고하여 35단어를 3번씩 발생하게 하여 5~7분 정도 음성을 수집하도록 2차 음성 수집 방법에 반영했다[10-13].

1차 수집 과정 중 인식을 저하의 원인이 되는 단어는 삭제하고 새로운 단어의 대체가 필요했다. 특히 명칭 수의 경우 중복되는 개, 마리, 명, 사람의 단어 때문에 인식을 저하가 일어나는 것을 확인했다. 유아의 경우 성인보다 발음이 불명확하기 때문에 중복되는 음절을 최소화하여 2차 단어 선정에 반영했다.

### 3.4 2차 유아 음성 수집 과정

1차 음성 수집 후 음성 인식기를 이용해 인식이 낮은 저조한 단어와 집중력 저하를 야기하는 수집 방법에 대한 문제를 보완하여 2차 유아 음성 수집을 진행했다.

#### 3.4.1 2차 단어 선정

단어 선정은 1차 단어 선정 방식과 동일하게 한국의 영유아가 사용하는 어휘를 설계한 MCDI-K를 바탕으로 유아교육전문가를 통해 내용을 감수 받았다. 그 후 1차 단어 선정 때의 인식을 저하의 원인이 되는 단어를 삭제하고 새로운 단어를 보완했다. 발음이 불명확한 유아에게 중복된 음절은 인식을 저하의 원인이 되기 때문에 이 부분을 염두에 두고 단어 선정을 했다. 2차로 선정된 단어는 [표 3]과 같다.

표 3. 2차 단어 선정

개념	색	빨강, 파랑, 노랑, 초록, 분홍, 검정, 보라, 흰색, 주황
	형	동그라미, 세모, 네모, 별, 달, 하트, 삼각형, 사각형, 원
	수	일, 이, 삼, 사, 오, 육, 칠, 팔, 구, 십 하나, 둘, 셋, 넷, 다섯, 여섯, 일곱, 여덟, 아홉, 열
시간	먼저, 나중, 지금, 아직, 내일, 어제, 오늘, 아까	
	위치	위, 아래, 오른쪽, 왼쪽, 앞, 뒤, 옆, 안, 밑
지시어	보기	확대, 축소, 회전
	실행	예, 아니오, 시작, 다시, 그만, 처음으로
단어	동물	여우, 사자, 코끼리, 토끼, 기린, 호랑이, 말, 오리, 돼지
	탈것	기차, 자동차, 배, 오토바이, 버스, 비행기, 트럭
	직업	간호사, 경찰관, 소방관, 우체부, 의사, 선생님
	옷	장갑, 양말, 치마, 외투, 잠바, 잠옷, 바지
	장소	병원, 유치원, 은행, 놀이터, 동물원, 학교, 수영장
장난감	로봇, 인형, 블록, 풍선, 비누방울	

#### 3.4.2 2차 유아 음성 수집 방법

2차 유아 음성 수집 환경은 [그림 2]와 같다. 국외의 이전 연구를 살펴보면 초등학생 이하 대상의 음성 수집 시 인터페이스 구축 시 사용될 장소로 예상되는 공공장소, 사무실, 학교의 컴퓨터실을 이용했다[7-9]. 본 논문에서도 데이터베이스 구축 목표에 맞게 음성 수집은 유아의 교육, 혹은 놀이를 할 수 있는 장소로 예상되는 유치원에서 진행했다. 유치원 공간은 잡음이 정제된 환경은 아니므로 조용한 빈 교실에서 녹음을 시행했다.



그림 2. 음성 수집 환경

음성 수집을 위한 장비는 노트북과 코리아 디지털의 MBL센서 KDS-1012 마이크를 사용했다. 음성 녹음을 위한 소프트웨어는 편집이 용이한 CoolEdit Pro 2.1을 사용하여 16KHz Sample Rate, Mono Channels, 16bit Resolution으로 설정하여 녹음을 진행했다.

녹음 수집 장소에는 성인 진행자와 유아 한 명이 참여하여 진행했다. 1차 음성 수집 방법을 수정하여 성인의 정확한 발음을 들려 준 후 발음을 하도록 하였다. 1차 수집 방법의 단점을 보완하여 35단어를 3번씩 발생하게 하여 5~7분 정도 음성을 수집하였다[10-13].

## IV. 유아 음성 데이터베이스 구축

만3~5세 대상의 음성데이터베이스 구축을 위한 음성 수집 과정과 동시에 유아 음성 데이터베이스 공개를 위해 음성 데이터의 토근화와 수집에 참여한 유아의 정보를 기록한다. 토근화 과정을 통해 저장될 파일은 사용자의 편의를 생각해 파일 규칙을 생성해 이름을 붙인다. 한국에서 유아 음성 데이터베이스 공개는 최초이며

이를 통해 유아 음성인식의 활발한 연구 진행의 기초가 될 것이다.

#### 4.1 유아별 음성 데이터 원본

유아 음성 데이터베이스 수집은 유아가 발성을 시작하면 녹음이 진행되기 때문에 수집의 전 과정이 녹음된다. 이 때문에 유아의 목소리 외에 진행자의 목소리와 같은 불필요한 소리가 포함되어 있다. 불필요한 데이터를 삭제하고 유아 개인의 목소리만 포함된 파일을 생성한다.

##### 4.1.1 원본 파일 이름 규칙

단어 별 파일 토큰화 전에 유아 개인의 음성 파일 원본이 필요한 사용자가 생길 수 있기 때문에 유아별 파일을 먼저 생성한다. 음성 데이터베이스 사용자의 편의를 위해 음성 데이터의 파일 이름 규칙을 만들었다[11]. 유아별 파일 이름 규칙은 [표 4]와 같다. 파일 이름 규칙을 통해 유아의 나이와 성별 등의 정보를 알 수 있다.

음성 데이터 정보를 알 수 있는 파일 이름 규칙을 살펴보면 ResD는 유아의 실 거주지를 나타내며, GraP는 조부모와의 동거 여부를 알 수 있다. Sib로 형제 수를 파악하고, Name 기호는 유아 이름의 이니셜을 사용한다. 유아의 실제 이름을 사용하지 않는 것은 유아의 정보 보호를 위해서이다. S와 Ag는 성별과 나이를 파악할 수 있도록 했고 Ext는 파일의 확장명을 표시한다.

표 4. 유아별 파일 이름 규칙

형식	ResD_Name_GraP_Sib_S_Ag_Ext
----	-----------------------------

기호	의미
ResD	Residence : {서울, 경기}
Name	Name : {ksa, ..., yjk}
GraP	Grandparents : {0, 1} * 0, 1: 조부모 동거 여부
Sib	Sibling : {1, 2, 3} * 1: 외동, 그 외: 자신 외 나머지 형제 수
S	Sex : {Male, Female}
Ag	Age : {3, 4, 5} * 나이 (만)
Ext	File Extension : {PCM}

#### 4.2 단어별 음성 데이터의 토큰화

사용자의 편의를 위해 유아가 발성한 단어별로 음성 데이터의 토큰화 작업을 실시했다. 음성 데이터의 토큰화를 위해 사용되는 소프트웨어는 음성 녹음 시 사용하는 소프트웨어와 같은 CoolEdit Pro 2.1을 이용한다. 음성 데이터의 토큰화 과정에서 음성 데이터베이스를 이용하는 사용자를 위해 원본 파일 이름 규칙과 같이 파일 이름 규칙을 부여했다.

##### 4.2.1 단어 파일 이름 규칙

단어 파일 이름 규칙은 [표 4]에서 두 가지 규칙을 추가한 [표 5]와 같다. 추가된 규칙은 발성하고 있는 단어를 직접 들어보지 않고 알 수 있는 단어 이름 WorN과 발성횟수를 의미하는 Ind를 포함한다.

표 5. 단어별 파일 이름 규칙

형식	ResD_Name_GraP_Sib_WorN_S_Ag_Ind_Ext
----	--------------------------------------

기호	의미
ResD	Residence : {서울, 경기}
Name	Name : {ksa, ..., yjk}
GraP	Grandparents : {0, 1} * 0, 1: 조부모 동거 여부
Sib	Sibling : {1, 2, 3} * 1: 외동, 그 외: 자신 외 나머지 형제 수
WorN	Word Name : {선정된 100개의 단어}
S	Sex : {Male, Female}
Ag	Age : {3, 4, 5} * 나이 (만)
Ind	Index : {1, 2, 3} * 발성 횟수
Ext	File Extension : {PCM}

#### 4.3 데이터베이스 기술이전

만3~5세 대상의 유아 음성 데이터베이스는 유아의 음성과 관련한 특징이나 인식을 연구하는 연구원들에게 웹을 통한 기술이전을 할 것이다. 기술 이전은 다음의 웹 사이트를 통해 이루어진다.

- <http://academy.duksung.ac.kr/dsimlab>

표 6. 음성 데이터베이스 구축 참여 인원

나이(만)	3	4	5	합계
남	6	7	7	20
여	7	6	7	20
합계	13	13	14	40

표 7. 음성 데이터베이스 구축 단어 수

나이(만)	3	4	5	합계
남	1,890	2,205	2,205	6,300
여	2,205	1,890	2,205	6,300
합계	4,095	4,095	4,410	12,600

웹 사이트에 방문하면 기술이전 될 예제 파일과 설명이 있다. 자료 요청을 하게 되면 [표 6]과 같이 구성된 유아 대상의 [표 7]에 명시된 데이터를 기술이전 받을 수 있다. 기술이전 되는 두 가지 형태로 유아별 원본 파일과 단어별 토큰화된 파일이다. 유아의 음성 파일과 나이, 성별, 가족 사항의 정보가 포함되어 있지만 개인의 정보 유출의 염려는 없다. 이는 유아의 음성을 연구 자료로 배포하는데 있어 필요한 기초적인 정보를 포함하고 있을 뿐 해당 유아의 이름 및 구체적인 거주지는 포함하고 있지 않기 때문이다. 음성 데이터베이스 이술 이전을 통해 관련 연구의 기초자료가 되어 활발한 연구를 기대한다.

## V. 결론

컴퓨터 기술의 발전과 함께 인터넷과 다양한 디바이스 개발로 인해 학습 환경의 새로운 시대가 열렸다. 이는 학습의 다양성과 접근성을 용이하게 하여 다양한 형태의 실질적이고 유의미한 교육을 받을 수 있는 기반이 마련 됐음을 의미한다.

본 논문은 언어발달의 중요한 시기인 만 3~5세 유아의 말하기 능력을 강화하기 위한 음성 인터페이스 구축의 기반을 마련하고자 했다. 현재 한국에서는 대부분 성인 대상의 음성 데이터베이스가 구축되어 있고, 유아 대상의 음성 데이터베이스는 없었다. 유아에게 다양한

교육 환경을 제공하기 위해서는 유아 대상의 음성 인터페이스 구축 기반으로 다양한 콘텐츠가 개발되어야 한다. 하지만 이전에 유아에 맞는 음성 데이터베이스가 필요하며 이 사실은 이전 연구를 통해 입증하였다.

향후 유아 음성 데이터베이스 구축은 음성 인터페이스 개발의 기초자료가 되어 유아의 놀이와 교육과 관련한 다양한 콘텐츠에 활용할 것이다. 성인의 경우 현재 음성 인터페이스는 여가와 관련한 체감형 게임에 사용되지만, 유아의 경우 놀이와 교육에 동시에 이용될 수 있다. 이처럼 음성을 활용한 교육과 놀이는 유아의 몰입감을 증대하여 놀이 활동에서 적극적인 모습을 발견할 수 있다. 놀이와 교육적 측면 이외에 음성인식기 측면에서 한국 유아의 음성 특징 분석을 통해 유아에게 맞는 음성인식기의 인식을 향상을 도모할 수 있다. 구축된 한국 유아 음성 데이터베이스를 기초로 지속적인 유아의 음성 데이터를 확보하고, 구축내용과 방향에 대해 관련 연구자들의 많은 참여를 기대해 본다.

## 참고 문헌

- [1] H. Deniz and H. Cakir, "Design principles for computer-assisted instruction in histology education : An exploratory study," *Journal of Science Education and Technology*, Vol.15, No.5, pp.399-408, 2006.
- [2] G. J. Hwang and H. F. Chang, "A formative assessment-based mobile learning approach to improving the learning attitudes and achievements of students," *Computers & Education*, Vol.56, No.1, pp.1023-1031, 2011.
- [3] S. Jane and C. Martyn, "E-learning and accessibility : An exploration of the potential role of generic pedagogical tools," *Computers & Education*, Vol.54, No.4, pp.1107-1116, 2010.
- [4] L. M. Leslie, C. I. Chang, S. Wang, M. E. Beier, and Y. Klisch, "Learning and motivational impacts of a multimedia science game,"

Computers & Education, Vol.57, No.1, pp.1425-1433, 2011.

- [5] 이용주, 김봉완, 김영일, 최대립, “한국의 공동이용을 위한 음성언어자원의 구축 및 보급현황”, 한국어정보학회, Vol.10, No.1, pp.81-85, 2008.
- [6] 유재권, 이경미, “한국어에서의 성인과 유아의 음성인식 비교”, 한국콘텐츠학회, Vol.11, No.5, pp.138-147, 2011.
- [7] S. Das, D. Nix, and M. Picheny, “Improvements in children’s speech recognition performance,” In: Proc. ICASSP 98, 1998.
- [8] L. Mahl, “Speech recognition and adaptation experiments on children’s speech,” KTH, Stockholm, Sweden, 2003.
- [9] M. Gerosa, D. Giuliani, and F. Brugnara, “Acoustic variability and automatic recognition of children’s speech,” Speech Communication, pp.847-869, 2007.
- [10] D. Giuliani and M. Gerosa. “Investigating recognition of children speech,” In Proc. of ICASSP, pp.137-140, 2003.
- [11] B. W. Hwang, S. Kim, and S. W. Lee, “A full-body gesture database for automatic gesture recognition,” In Proc. of IEEE Conf. on FGR, pp.243-248, 2006.
- [12] 한나영, 저소득층 밀집지역에서 소득수준에 따른 만4, 5세 유아의 수용-표현 언어발달, 어휘력, 단어읽기 비교: G시 저소득층 밀집지역을 중심으로, 건국대학교 교육대학원 석사학위논문, 2011.
- [13] 최은영, 초기 문해 기술이 유아의 단어읽기와 단어쓰기에 미치는 영향 연령 및 가정의 사회경제적 지위별 분석, 건국대학교 교육대학원 박사학위논문, 2010.
- [14] J. Nicholas and A. Geers, “Effects of early auditory experience on the spoken language of deaf children at 3 years of age,” Ear and hearing, Vol.27, No.3, pp.286-298, 2006.
- [15] J. Piaget, Play, dreams and imitation in

childhood, New York: Norton&company, 1962.

저 자 소 개

유 재 권(Jae-Kwon Yoo)

준회원



- 2009년 2월 : 덕성여자대학교 인터넷정보공학과(공학사)
- 2011년 3월 ~ 현재 : 덕성여자대학교대학원 전산정보통신학과(석사과정)

<관심분야> : 음성인식, 유아 음성 DB etc.

이 경 옥(KyungOk Lee)

정회원



- 1986년 2월 : 서울대학교 가정대학(학사)
- 1991년 6월 : 미국 쉐넬대리대학교 교육학과(교육학석사)
- 2001년 12월 : 미국 남가주대학교 교육심리학과(철학박사)

▪ 2002년 3월 ~ 현재 : 덕성여자대학교 유아교육과 교수  
<관심분야> : 아동게임, 에듀테인먼트 콘텐츠, 사회정서 발달, 연구방법

이 경 미(Kyoung-Mi Lee)

정회원



- 1993년 2월 : 덕성여자대학교 전산학과(이학사)
- 1996년 2월 : 연세대학교 전산학과(이학석사)
- 2001년 12월 : 아이오와주립대학교 전산학과(이학박사)

▪ 2003년 3월 ~ 현재 : 덕성여자대학교 컴퓨터학과 교수  
<관심분야> : 영상처리, 패턴인식, 멀티미디어, HCI etc.