SiTEC의 공동 이용을 위한 음성 코퍼스 구축 현황 및 계획

김봉완(SiTEC), 최대림(SiTEC),
김영일(SiTEC), 이광현(SiTEC), 이용주(원광대)

<차 례>

<Abstract>

Current States and Future Plans at SiTEC for Speech Corpora for Common Use

Bong-Wan Kim, Dae-Lim Choi,
Young-Il KimKwang-Hyun Lee, Yong-Ju Lee

To support speech information technology industry it is vital to create and distribute standardized speech corpora to be used for the development of products and technologies. In this article we introduce speech corpora created by Speech Information Technology & Industry Promotion Center(SiTEC) during its 1st and 2nd fiscal years (2001/5/1 - 2003/4/30) and plans for those corpora which is being created currently or will be created in near future. We introduce the corpus for car application to expand speech information technology to the field of traditional industry, the corpora for foreign languages to support exportation, the corpus for basic research for the sake of application in the industry, the corpora for common use, and others.

* Keywords: Speech Corpora

1. 서 론

한국어의 공학적인 응용을 위해서는 그 기반이 되는 요소기술로써 음성인식 및 합성으로 대표되는 음성 처리 기술과 언어 이해 및 기계 번역으로 대표되는 언어 처리 기술의 연구가 필요하다. 이러한 음성 및 언어 처리 기술의 연구를 위해 가장 먼저 확보되어야 할 것이 음성, 언어 및 각종 사전 코퍼스 등 국어 정보 베이스이다. 이들의 체계적인 조기 확보 여하에 따라 음성 및 언어 처리 연구의 성패를 좌우한다고 해도 과언이 아니다.

특히 한국어 음성을 대상으로 한 음성 코퍼스는 음성 언어 연구의 기본으로서 개발 초기부터 확보되어야 할 연구 자원이다[1]. 기관별로 자체 연구를 목적으로 개별적인 코퍼스 구축이 이루어져 왔으나, 최근에는 공동으로 사용할 수 있는 우리말 음성 언어 코퍼스의 중요성을 인식하고 체계적인 구축 및 공개에 대한 노력을 시작하였고 이에 대한 산업자원부의 정책적인 지원에 의해 음성정보기술산업지원센터(SiTEC)가 설립되어 활동 중이다.

2. SiTEC의 음성 코퍼스 구축 현황

음성 정보 기술 산업을 효과적으로 지원하기 위해서는 상품 및 기술의 개발을 위한 표준화된 음성 DB의 구축 및 보급이 필수적이라고 할 수 있다. 본 장에서는 센터에서 기술 개발 인프라 구축 사업의 일환으로 1~2차년도 사업 기간 중에 구축된 음성 코퍼스에 관한 사항을 상세히 기술한다.

구축된 음성 코퍼스는 세밀한 검증을 거쳐 2002년 8월부터 보급을 시작하여 현재 총 20종의 음성 코퍼스를 배포하고 있으며, 센터 홈페이지를 통해 자세한 사양과 샘플 데이터를 제공하고 있다[2]. 아울러 사용자가 원하는 사양의 음성 코퍼스를 별도 제작하거나 보급 중인 코퍼스 중 원하는 사양의 데이터만을 골라 재작성하는 맞춤형 음성 코퍼스 보급도 진행 중이다.

2.1. 자동차 응용을 위한 코퍼스

최근 자동차 환경에서의 음성인식 응용에 대한 관심과 수요가 많아지고 있고, 산업자원부에서도 중기 거점 과제를 통하여 자동차 환경에서의 음성인식 기술의 개발을 지원하고 있다. 일반적인 환경에서 얻어진 데이터와는 달리 자동차 내에서 얻어진 음성 정보는 자동차의 내·외부에서 비롯되는 소음 환경 요인으로 인하여 여러 위치에서 다양한 형태로 나타나게 된다. 따라서 자동차 음성 인식 기술 연구 및 응용 개발을 위해서, 다양한 주행 상태에 따른 소음 환경에서의 음성 DB 구축

이 필수적이다.

자동차 환경에서의 소음 및 음성 코퍼스의 경우 그 수집 절차, 환경 요인 등에 있어서 일반적인 경우와 달리 매우 많은 변수가 있기 때문에 1차년도(2001. 5. 1 ~ 2002. 4. 30)에는 이러한 수집 절차 및 환경 요인에 대한 연구와 분석을 위한 프로토타입 코퍼스를 구축하였다.

2차년도(2002. 5. 1 ~ 2003. 4. 30) 자동차 음성 코퍼스 구축 계획은 300명 화자, 5채널 동시 수집, 1인당 100토큰 규모였으나, 자동차 환경에서의 음성 인식에 대한 관심과 요구가 증대되면서 업체의 요구가 많아 그 규모를 400명 화자, 8채널 동시 수집, 1인당 200여 토큰 규모로 확대하였다.

자동차 응용을 위한 음성 코퍼스의 종류에 따른 수집 규모 및 자세한 사양은 다음의 <표 1>과 같다.

<표 1> 자동차 응용을 위한 음성 코퍼스

구 분	규 모	발성목록 / 수집환경
자동차 환경 단어 음성 DB (프로토타입)	- 총 1,452단어 - 100명 - 1인당 100단어 발성 - 7채널 + 1 핸즈프리 동시 수집 - 80km 주행 환경	- 자동차 컨트롤, Navigation 관련 544 단어, 단독 숫자 44종, 4연 숫자 864종 - 시속 80km의 주행 환경, 대형 차량 이용 - 표준어를 사용하는 100명의 남성 화자 - 샘플 포맷: 24kHz, 16bit linear PCM
자동차 소음 DB	- 총 270 종의 환경 - 7채널 + 1 핸즈프리 동시 수집	- 각 환경 별 데이터의 길이는 5분 - 지속적인 데이터가 최소 1분 30초 분량 정도는 포함되도록 수집 - 다른 잡음의 경우 시작점, 끝점에 대한 표시 - 샘플 포맷: 24kHz, 16bit linear PCM
대규모 자동차 환경 단어 음성 DB	- 총 1,452단어 - 400명 - 1인당 200단어 발성 - 8채널 동시 수집 - 시내, 고속도로 주행 환경	- 단독 숫자 및 4연 숫자: 900 종, 다이얼링 명령어: 63 종, 카 오디오 명령어 및 관련 단어: 306 종, 자동차 컨트롤 스위치 명령어: 120 종, 네비게이션 명령어: 57종, PDA 명령어: 121 종, 지명(군, 구 단위 이상): 446 종, 주요 도로명 (고속도로 이상): 53 종 - 성별, 지역, 연령 고려 - 설정 환경은 2환경으로 한정하여 50:50으로 하고, 1개의 공통 환경에서 데이터 수집 - 시내 주행 환경(30~60km/h), 고속도로 주행 (70~90km/h), 공통 환경(맑은 날씨/아스팔트 노면, 창문 close) - 2,000cc 승용차로 한정

2.2. 수출 지원을 위한 외국어 코퍼스

수출 지원을 외국어 음성 코퍼스로 1차년도에는 중국어 음성 코퍼스를 구축하였고, 2차년도에는 대상 언어를 확대하여 영어, 스페인어 음성 코퍼스를 구축하였다.

<표 2> 수출 지원을 위한 외국어 음성 코퍼스

구 분	규 모	발성목록 / 수집환경
중국어 음성 DB	- 300명 - 1인당 110 토큰 발성	- 음절: 421개의 음절, 중국어 성조의 모델링을 위해 경성 및 4성을 고려, PBW 단어: 2음절~4음절의 1,200 단어, 4연 숫자: 661개의 4연 숫자, 366개의 날짜 관련 단어, 문장: 각 문장 마다 최대 27개의 한자(음절)을 포함하는 400개의 문장으로 구성 - 연변 대학 지역 협력 사이트를 이용하여 북경어를 사용하는 화자로 구성 - 사무실 환경에서 SoundBlaster Live , SENNHEISER E835 Mic.를 사용하여 PC에서 수집
영어 음성 DB	- 총 1,586 단어 - 400명 - 1인당 130 토큰 발성	- 단독 숫자: 13종, 예/아니오: 2종, 알파벳: 26종, 화폐 단위: 32종, 내장 명령어: 15종, 특정 날짜 시간 표현: 49종, 응용어: 117종, 날짜: 366종, 요일: 7종, 비밀번호: 150종, 전화번호: 283종, 주 및 도시이름: 376종, 신용카드번호: 150종 - 영어를 모국어로 하는 성인 남녀를 대상으로 미국 현지 녹음 수록 - 성비는 50:50 - 연령별 분포는 20대: 73.8%, 30대: 12.5%, 40대: 13.8%로 구성
스페인어 음성 DB	- 총 1,230 단어 + 5,670문장 - 300명 - 1인당 130 토큰 발성	- 단독 숫자: 10종, 알파벳: 29종, 제어 명령어: 84종, 화폐 단위: 83종, 특정 날짜 시간 표현: 60종, 응용어: 15종, 날짜: 366종, 비밀번호: 150종, 전화번호: 283종, 신용카드 번호: 150종, 문장: 5670종 - 미국 내에서 거주하는 히스페닉계를 대상으로 미국 서남부 현지 녹음 수록 - 사무실 환경에서 SoundBlaster Live, SENNHEISER m@b40 Mic.를 사용하여 PC에서 수집(영어, 스페인어 동일) - 남자 154명, 여자 146명 - 20대: 44.3%, 30대: 29.3%, 40대: 26.3% 분포

2.3. 산업 응용을 위한 기반 기술 연구용 코퍼스

산업 응용 기반 기술 기초 연구용 코퍼스의 발성 목록은 다양한 응용과 연구에 적용하기 위하여 특정 응용에 종속되지 않은 발성 목록을 사용하는 것이 바람직하다. 따라서 이러한 목적으로 사용하기 위해 1차년도에는 한국어에서 발생할 수 있는 다양한 음운 환경 및 음절을 고려한 PRW 4,178어절을 선정하고 이를 발성 목록으로 사용하여 클린스피치 단어 음성 코퍼스를 구축하였다. 또한 수집된 음성 데이터 전량에 대하여 음운 레이블링 기준(센터 권고안)에 의해 지역 협력 사이트에서 음운 레이블링을 실시하였고 현재 레이블링 전문 인력에 의해 검증 및 수정 작업을 진행 중이다.

2차년도에는 1차년도에 수집된 클린스피치를 단어에서 문장으로 확대 구축하기 위해 형태소 분석 균형 말뭉치를 사용하여 다양한 분야에서 사용되는 문장에서 빈번히 사용되는 형태소로만 구성된 20,217문장과 이전 연구에 의해 구성된 PBS (Phonetically Balanced Sentences) 589문장을 사용하여 발성 목록을 구성하였다. 클린 스피치 낭독 문장 음성 코퍼스의 발성 목록 설계 과정은 다음과 같다.

(1) 발성 목록 선정을 위한 말뭉치의 형태 통계 분석
발성 목록 선정을 위한 모집단으로는 21세기 세종계획 형태소 분석 균형 말뭉치 1,000만어절[3]을 사용하였다. 먼저 형태소를 최소 단위로 말뭉치를 분석하고 분석된 말뭉치의 형태소를 통계 처리하여 각 형태소 유형의 빈도를 조사하였다.

(2) 형태 통계 결과의 정렬
형태소의 통계 분석 결과는 고빈도의 형태소 토큰에서 저빈도의 형태소의 토큰 순으로 정렬하여 형태소, 태그, 상대빈도, 누적빈도, 누적상대빈도, 찌프상수를 분석하였다.

(3) 문장 색인
형태소 분석 말뭉치에서 나열된 어절은 문장으로 복구하고 형태소 분석 부분은 형태 통계에서 한 문장이 갖는 최저 형태소 빈도를 찾아 문장 색인에 사용하였다.

(4) 기존 발성 문장과 비교 선정
색인된 문장은 1차년도 Dictation 낭독 음성 코퍼스 발성 목록과 비교하여 기존 발성 목록에 포함되는 문장을 삭제하였다.

(5) 발성 목록 추출
형태소 최저 빈도수로 색인된 문장에서 적당량의 문장을 추출하기 위해서 형태소 최저 빈도수의 임계치가 구해져야 한다. 이 때, 6어절 이상 25어절

이하의 문장만이 누적 문장수의 계산에 사용되었다. 최종 발성 목록은 형태소 최저 빈도수가 826 이상인 20,217 문장에 대해서 수작업에 의한 수정 및 검증을 하였으며, 이전 연구에 의해 구성된 PBS 589문장을 추가하였다.

운율 합성용 음성 코퍼스를 위한 발성 목록 선정의 모집단으로 사용된 텍스트 코퍼스는 설명문, 수필문, 사회학, 방송 3社(KBS, MBC, SBS 등)의 뉴스, 신문(조선일보, 한국일보), 경제학, 전산학, 기계학, 생물학 등의 장르별 균형 텍스트로 구성된 KAIST Taged Corpus 100만 어절[4]이고 Triphone maximization 기준을 적용하여 발성 목록을 선정하였다. 선정된 문장 발성 목록은 4,392문장이며 이 문장에 포함된 Triphone의 총 종류수는 모집단과 같이 18,025 종류이다.

Dictation용 낭독 음성 코퍼스의 발성 목록 선정을 위한 모집단으로는 KAIST에서 구축된 4,300만 어절의 KAIST Corpus를 사용하여 고빈도 어휘에 대한 분석을 수행하였다. 분석 결과 상위 고빈도 5,000어절이 전체 어절에 대해 50.6%의 coverage를 가지며 상위 10,000어절의 경우 58.4%, 20,000어절의 경우 66.2%의 coverage를 갖는 것으로 나타났다. 본 코퍼스에서는 발성 목록 선정을 위해 고빈도 5,000어절, 8,000어절 및 10,000어절을 발성 목록 선정을 위한 대상 어휘로 선정하였다. 추출된 문장의 총 수는 20,833문장으로 문장의 평균 길이는 문장 당 7.43어절이다. 또한 인식 대상 어휘에 포함되지 않은 단어가 발성된 경우에 대처하기 위한 OOV (Out of vocabulary) 테스트를 위해 다음과 같이 문장 목록을 구성하였다.

- 5K 문장 세트: 8,608 문장
  - 고빈도 5,000어절에 포함된 어휘만으로 구성된 문장 세트
- 8K-5K 문장 세트: 7,301 문장
  - 고빈도 8,000어절에 포함된 어휘만으로 구성된 문장 세트를 구성하고, 여기에서 5K 문장 세트에 포함된 문장은 중복되므로 이를 삭제한 것
- 10K-8K 문장 세트: 4,924 문장
  - 고빈도 10,000어절에 포함된 어휘만으로 구성된 문장 세트를 구성하고, 여기에서 5K 문장 세트와 8K-5K 문장 세트에 포함된 문장은 중복이므로 이를 삭제한 것

위와 같이 구성된 Dictation용 낭독 음성 코퍼스를 2차년도에 확장하기 위하여 클린스피치 낭독 문장 음성 코퍼스의 발성 목록과 동일한 총 20,000여 문장을 발성 목록으로 사용하여 총 400명의 화자를 추가하여 PC 환경 낭독 문장 음성 코퍼스를 확대 구축하였다.

<표 3> 산업 응용을 위한 기반 기술 연구용 음성 코퍼스

구 분	규 모	발성목록 / 수집환경
클린 스피치 단어 음성 DB	- PRW 4100여 어절 - 500명 - 1인당 417어절 발성 - 전량 음운 레이블링	- 한국어에서 발생 가능한 다양한 음운 환경을 고려한 PRW 4,178어절 - 전량 음운 레이블링(자동 레이블링 후, 수동으로 검증 및 수정) - DAT, AKG C414 B-ULS Mic.를 사용하여 방음실에서 녹음 - 성별, 지역, 연령 고려
클린 스피치 낭독 문장 음성 DB	- 총 20,000 여 문장 - 200명 - 1인당 100여 문장 발성	- 형태소 최저 빈도수가 826 이상인 문장(6~25어절): 20,217 문장 + PBS 589문장 - 문장 전사, 전량 자동 레이블링 - DAT, AKG C414 B-ULS Mic.를 사용하여 방음실 녹음 - SENNHEISER HMD 280 Pro Mic.
운율 합성용 낭독 문장 음성 DB	- 남녀 전문성우 각 1인 - 1인당 4,392문장 발성 - 전량 음운, 운율 레이블링	- 방음실 환경에서 전문성우를 통해 데이터 수집 - EGG signal(Laryngograph 6103) 포함
Dictation용 낭독 문장 음성 DB	- 총 20,800여 문장 - 400명 - 1인당 100여 문장 발성	- 가전 제어용 시스템의 학습 및 평가 - Andrea ANC 750 Mic - 성별, 지역, 연령 고려 - 전량 문장 전사
Dictation용 낭독 문장 음성 DB 확장	- 총 20,000여 문장 - 200명 - 1인당 100여 문장 발성	- 가전 제어용 문장 음성 DB의 확장 - Andrea ANC 750 Mic. - 성별, 지역, 연령 고려 - 전량 문장 전사
공유 숫자음 DB의 보완	- 2~3음절로 이루어진 단위 숫자 25,000종 - 500명 - 1인당 100 토큰 발성	- SENNHEISER E835S dynamic Mic. - 성별, 연령, 지역를 고려

2.4. 기타 코퍼스

완구, 교육용 S/W 등 아동용 응용에 대한 요구가 증가함에 따라 1차년도에는 아동용 음성 인식 응용을 위한 음성 코퍼스를 구축하였다.

또한 최근에 다양한 음성 정보 기술이 실생활에 적용되기에 이르렀고, 차세대 사용자 인터페이스 수단으로 부각되면서 완구, 로봇, PDA, 홈오토메이션과 같은 다양한 임베디드용 음성 인식 어플리케이션이 개발되고 있다. 2차년도에는 이처럼 다양한 기기 내장형 음성 인식용 코퍼스 구축을 위해 USB-DSP 임베디드용 음성

수집 툴킷을 이용하여 총 300명의 화자를 대상으로 기기 내장형 electret 콘덴서 마이크를 통해 수집하였다.

음성 인식 시스템의 성능에 영향을 미치는 다양한 요인 중 마이크의 음향적 특성, 마이크의 위치 및 마이크와 화자와의 거리도 매우 중요한 요인 중 하나이다. 따라서 센터에서는 이러한 다양한 변인에 따른 시험용 코퍼스의 구축을 위해 1차년도에는 마이크의 종류에 특성 변화 시험용 음성 코퍼스를 구축하였고, 2차년도에는 마이크로폰의 거리에 따른 영향을 분석하기 위한 음성 코퍼스를 구축하였다.

<표 4> 기타 코퍼스

구 분	규 모	발성 목록/수집 환경
아동용 음성 DB	- 총 1,283 토큰 - 500명 - 1인당 100단어 발성	- 4연 숫자: 340종, PBW 452어절, 명령 및 지시어: 400종, 단독 숫자 및 단위: 41종 - 서울, 경기: 50%, 충청: 10%, 영남: 25%, 호남: 15%의 지역 분포 - 초등학교 1학년~6학년의 균등 분포 - 가정집, 사무실 환경에서 SoundBlaster Live, Andrea ANC 750 Mic.를 사용하여 PC에서 수집
기기 내장형 음성 DB	- 총 4,199 토큰 - 300명 - 1인당 107~108 토큰 발성	- 단독 숫자: 12종, PDA 공통 명령어: 12종, PRW: 4,175 종 - 사무실 환경에서 USB-DSP 임베디드용 음성 수집 툴킷을 이용하여 데이터 수집 - 음성 코덱: TLV320AIC10I(General-Purpose 3V to 5.5V 16-Bit 2-KSPS DSP CODEC) - 16kHz, 16bit linear PCM - 기기 내장형 Electret 콘덴서 마이크 사용
다양한 마이크 음성 DB	- PBW_SH DB를 방음실에서 HATS를 통하여 재수집	- 마이크 종류에 따른 특성 변화 시험용 DB의 경우 8종의 마이크로폰 사용
다양한 마이크 거리 변화에 따른 음성 DB	- PBW_SH DB를 방음실에서 HATS를 통하여 재수집	- 마이크 거리에 따른 특성 변화 시험용 DB의 경우 5cm, 10cm, 20cm, 50cm, 100cm의 거리를 변화하고 2종의 마이크로 한정하여 수집

2.5. 공유 코퍼스

기존에 구축된 음성 코퍼스 중 센터를 통하여 공유 의사를 표명한 음성 코퍼스는 다음과 같다.

- PC 환경 숫자음 500명분
- PRW 전화음성 2000명분
- 숫자음 전화 음성 2000명분
- 클린스피치 PBW 70명분

- 클린스피치 PBS 20명분
- KAIST 무역 상담 코퍼스외 5종
- Web TV의 제어 명령어 인식용 음성 코퍼스

3. 3차년도 음성 코퍼스 수집 계획

센터에서는 차기년도 음성 코퍼스 구축을 위한 계획을 작성하기 위하여 음성 정보 기술 관련 전문가들을 대상으로 수요 조사를 실시하였다. 차기년도의 음성 코퍼스의 구축 계획은 조사된 결과 및 R&D Roadmap 등을 참고하여 결정하였으며 자세한 내용은 <표 5>와 같다.

센터에서는 향후 현재 구축된 음성 코퍼스의 내용과 양을 지속적으로 보완하고 확장할 계획이며 새롭게 구축하길 희망하는 음성 코퍼스는 소요 타당성 및 스펙을 예비 검토하여 프로토타입을 제작하고 그 결과를 바탕으로 관련 기관 및 연구자들과의 다각적인 논의를 통하여 보다 다양한 분야로 확장하려 하고 있다.

센터는 이러한 음성 코퍼스의 구축 내용과 방향에 대한 관련 연구자들의 많은 참여와 의견을 기대하고 있으며, 제시된 의견을 적극 반영하고자 한다.

<표 5> 3차년도 음성 코퍼스 구축 계획

명 칭	규 격	규 모
수출 지원을 위한 외국어 음성 코퍼스	• 인식용 외국어 코퍼스 - 영어: PRW 및 단문 - 중국어: 단어 및 문장 확장	영어: 400명분 중국어: 300명분
자동차 음성 코퍼스	• 실차 환경 음성 코퍼스 확장 - 발성 내용, 수록 조건, 발성자 수의 확장	5채널/400명분
산업 응용을 위한 화자인증 음성 코퍼스	• 시차별	30명분
멀티 모달 음성 코퍼스	• 음성 및 입술 영상 코퍼스	200명분
음성인식 성능 평가용 음성 코퍼스	• 실제 환경에서의 인식 성능 평가용 음성 코퍼스 • AURORA 2.0 수준	300명분
복지 응용을 위한 음성 코퍼스	• 복지 응용을 위한 장애자 및 노인 음성 프로토타입	100명분
모의 환경 음성 코퍼스	• 자동차 환경	다채널/500명분
대화체 인식용 음성 코퍼스	• 수집 분야, 수집 방법 및 정보 표기법 기초 검토 • 소규모 프로토타입 시험 제작	
한국인 외국어 음성 코퍼스	• 소요 타당성 및 스펙 예비 검토	
대화체 합성용 음성 코퍼스	• 소요 타당성 및 스펙 예비 검토	
고소음 음성 코퍼스	• 소요 타당성 및 스펙 예비 검토	
감정 음성 합성 연구용 코퍼스	• 소요 타당성 및 스펙 예비 검토	

4. 결 론

본 논문에서는 SiTEC에서 구축 중인 음성 코퍼스의 현황과 향후 계획에 관하여 보고하였다. 자동차 응용을 위한 코퍼스, 수출 지원을 위한 외국어 코퍼스, 산업 응용을 위한 기반 기술 연구용 코퍼스, 기타 코퍼스 등의 구축 내용이 소개되었다. 센터에서는 향후 현재 구축된 음성 코퍼스의 내용과 양을 지속적으로 보완하고 확장할 계획이며, 음성 코퍼스의 구축 내용과 방향에 대한 관련 연구자들의 많은 참여와 의견을 기대하고 있다.

참 고 문 헌

[1] 이용주, "음성언어코퍼스", *한국정보과학회지*, 16권, 2호, pp.41-48, 1998.2.
[2] 원광대학교 음성정보기술산업지원센터 홈페이지: http://www.sitec.or.kr
[3] 국립국어연구원, *세종계획 2001년도 국어 기초자료 구축 분과 연구 결과 보고서*, 문화관광부, 2002.
[3] 최기선, *KAIST 언어자원 2001년도판, 과학기술부 핵심 소프트웨어 과제 결과물*, http://kibs.kaist.ac.kr, 1995~2000.

접수일자: 2003년 05월 17일
게재결정: 2003년 06월 12일

▶ 김봉완(Bong-Wan Kim)
주소: 570-749 전북 익산시 신용동 344-2 원광대학교 음성정보기술산업지원센터
소속: 원광대학교 음성정보기술산업지원센터
전화: 063) 850-7452~3
FAX: 063) 850-7454
E-mail: bwkim@sitec.or.kr

▶ 최대림(Dae-Lim Choi)
주소: 570-749 전북 익산시 신용동 344-2 원광대학교 음성정보기술산업지원센터
소속: 원광대학교 음성정보기술산업지원센터 기반기술연구팀
전화: 063) 850-7452~3
FAX: 063) 850-7454
E-mail: dlchoi@sitec.or.kr

▶ 김영일(Young-Il Kim)
주소: 570-749 전북 익산시 신용동 344-2 원광대학교 음성정보기술산업지원센터
소속: 원광대학교 음성정보기술산업지원센터 기반기술연구팀
전화: 063) 850-7452~3
FAX: 063) 850-7454
E-mail: yikim@sitec.or.kr

▶ 이광현(Kwang-Hyun Lee)
주소: 570-749 전북 익산시 신용동 344-2 원광대학교 음성정보기술산업지원센터
소속: 원광대학교 음성정보기술산업지원센터 기반기술연구팀
전화: 063) 850-7452~3
FAX: 063) 850-7454
E-mail: khlee@sitec.or.kr

▶ 이용주(Young-Ju Lee)
주소: 570-749 전북 익산시 신용동 344-2 원광대학교 전기전자 및 정보공학부
소속: 원광대학교 전기전자 및 정보공학부, 음성정보기술산업지원센터
전화: 063) 850-7451
FAX: 063) 850-7454
E-mail: yjlee@wonkwang.ac.kr