2016년 한국음성학회 봄 학술대회
발표 논문집

주제: 음성언어와 인공지능
일시: 2016년 5월 20일(금) ~ 5월 21일(토)
장소: 광주과학기술원 오랑관
주최: 사단법인 한국음성학회
주관: 사단법인 한국음성학회
후원: 광주과학기술원, 한국전자통신연구원, 네이버, 한국언어재활사협회, 보이스웨어, 케이믹글로벌

한국음성학회
The Korean Society of Speech Sciences
## 2016 한국음성학회 봄 학술대회 일정표

### 5월 20일(금)

<table>
<thead>
<tr>
<th>시간</th>
<th>행사</th>
</tr>
</thead>
<tbody>
<tr>
<td>11:00-13:00</td>
<td>등록 및 중식 (광주과학기술원 오류관 로비 및 구내식당)</td>
</tr>
<tr>
<td>13:00-13:10</td>
<td>개최식 (오류관 1층 다목적홀)</td>
</tr>
<tr>
<td>13:10-14:10</td>
<td>기조발표 I (다목적홀)</td>
</tr>
<tr>
<td></td>
<td>이윤근 (한국전차통신연구원)</td>
</tr>
<tr>
<td></td>
<td>&quot;음성언어와 인공지능 (I): ETRI 연구 사례-자동통역, 대화인터페이스, DeepQA&quot;</td>
</tr>
<tr>
<td>14:10-14:30</td>
<td>휴식 및 포스터세션 준비</td>
</tr>
<tr>
<td>14:30-16:20</td>
<td>포스터발표 세션 I [24편]</td>
</tr>
<tr>
<td></td>
<td>각 포스터 1분 PPT 발표 (다목적홀)</td>
</tr>
<tr>
<td></td>
<td>포스터발표 (로비)</td>
</tr>
<tr>
<td>16:20-16:30</td>
<td>휴식</td>
</tr>
<tr>
<td>16:30-17:30</td>
<td>구두발표 11편</td>
</tr>
<tr>
<td></td>
<td>음성공학 (강의실D (203호))</td>
</tr>
<tr>
<td></td>
<td>말소리장애 및 음성의학 (다목적홀)</td>
</tr>
<tr>
<td></td>
<td>음성학 (강의실E (204호))</td>
</tr>
<tr>
<td>17:30-21:00</td>
<td>이동 및 만찬 (교외식당 ‘늘푸른 정원’)</td>
</tr>
</tbody>
</table>

### 5월 21일(토)

<table>
<thead>
<tr>
<th>시간</th>
<th>행사</th>
</tr>
</thead>
<tbody>
<tr>
<td>09:00-9:30</td>
<td>등록 및 포스터세션 준비 (오류관 로비)</td>
</tr>
<tr>
<td>9:30-10:50</td>
<td>포스터발표 세션 II [16편]</td>
</tr>
<tr>
<td></td>
<td>각 포스터 1분 PPT 발표 (다목적홀)</td>
</tr>
<tr>
<td></td>
<td>포스터발표 (로비)</td>
</tr>
<tr>
<td>10:50-11:00</td>
<td>휴식</td>
</tr>
<tr>
<td>11:00-12:00</td>
<td>기조발표 II (다목적홀)</td>
</tr>
<tr>
<td></td>
<td>김성희 (NAVER)</td>
</tr>
<tr>
<td></td>
<td>&quot;음성언어와 인공지능 (II): NAVER 연구 사례&quot;</td>
</tr>
<tr>
<td>12:00-12:30</td>
<td>스펙트로그램 리딩/우수발표자 시상 및 폐회식 (다목적홀)</td>
</tr>
<tr>
<td>12:30-13:30</td>
<td>정식 식사 (구내식당)</td>
</tr>
</tbody>
</table>

### “봄 학술대회 특별 튜토리얼 세션” (오류관 2층 강의실D (203호))

<table>
<thead>
<tr>
<th>일자</th>
<th>주제</th>
<th>강연자</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.21</td>
<td>음성학과 음성의학 전공자들 위한 신호처리기술 소개</td>
<td>김병순 (부산대)</td>
</tr>
<tr>
<td>13:30-16:40</td>
<td>음소 단위 분할 및 레이블링 소프트웨어 배포 및 사용 방법 소개</td>
<td>나민수, 정민화 (서울대)</td>
</tr>
<tr>
<td>일시</td>
<td>발표 및 내용</td>
<td>비고</td>
</tr>
<tr>
<td>------</td>
<td>-------------</td>
<td>------</td>
</tr>
<tr>
<td>첫째날 (5/20일(금))</td>
<td>한국음성학회 2016 문학술대회 프로그램</td>
<td></td>
</tr>
<tr>
<td>5.20(금)</td>
<td>11:00~13:00</td>
<td>등록, 점심시간 (구내식당)</td>
</tr>
<tr>
<td>제1부 (1층)</td>
<td>13:00~13:10</td>
<td>개회사 (한국음성학회 이호영 수석부회장)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>축사 (한국음성학회 조형우 회장)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>환영사 (광주과학기술원 문승현 총장)</td>
</tr>
<tr>
<td>제1부 기조발표 1</td>
<td></td>
<td>1부 사회: 김대옥 (경남대)</td>
</tr>
<tr>
<td></td>
<td>응성언어와 인공지능 (I): ETRI 연구 사례-자동통역, 대화인터페이스, DeepQA: 이용근 (한국전자통신연구원, 자동통역 연구기지 연구부장)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>14:10~14:30</td>
<td>휴식/포스터세션 준비</td>
</tr>
<tr>
<td>제2부 포스터발표 1</td>
<td>14:30~15:00</td>
<td>2부 진행</td>
</tr>
<tr>
<td></td>
<td></td>
<td>전체총괄: 신중원 (GIST)</td>
</tr>
<tr>
<td></td>
<td>세션시작 전 각 파티 1분별 구두 요약 PPT [총24편]</td>
<td></td>
</tr>
</tbody>
</table>
| 제2부 (1층) | 15:00~16:20 | [E1] 잡음 환경에서 DNN 기반 음성인식에 VTLP 적용 실험  
-김호신, 김영순 (부산대학교), 이성주 (한국전자통신연구원) |
| | [E2] 잡음 및 전방 환경에서 i-vector 기반 회사화 모델 전처리 성능 비교  
-이기하, 김영신, 김승훈 (부산대학교), 김성환 (대장정성) | |
| | [E3] 발성등의 기본주파수 및 발성구간 검출을 이용한 특강장권분석 기법  
-이정현, 이주현, 조진영, 박영진 (서강대학교) | |
| | | [E4] 전형 예측 문제를 통한 CONSC 기반 2-차자 화경 발화 결정 구간 검출  
-이성주, 김태환, 최호진, 한민수 (한국과학기술원) |
| | [E5] 자연어 음성인식 워크 응답터 대응법  
-이성주, 장희석, 정은, 박정규, 이용근 (한국전자통신연구원) | |
| | [E6] 개인화 음성합성 개발을 위한 소유망 발성목록 추출  
-이영기, 김태민, 홍진교, 김신호 (대덕립스) | |
| | | [E7] 한국인 영어 학습자의 발화 능숙도에 따른 음소소수 별 발화속도 차이 비교  
-나다익, 정민화 (서울대학교) | |
| | | [E8] 학습 DB 형태에 따른 TDNN 기반 i-vector 화자인식 시스템의 성능 변화  
-윤성욱, 권오목 (충북대학교) | |
| | | [E9] 말, 언어가 높은 아동의 말하기 능력을 향상시키기 위한 의사소통 훈련 방법 개발  
-김진호, 최대일, 고영환, 박연진, 이용주 (원광대학교), 이영미 (동명대학교), 권창홍 (대덕대) | |
| | | [M1] 아제노이드판도 비대, 알레르기성 비염을 가진 아동과 정상 아동의 비성도와 기타 음향 특성  
-김남숙, 성철재 (충남대학교) | |
| | | [M2] 실어증 환자의 발사 속도 및 음성특성(곤란도) 및 이에 따른 효과의 예비 연구  
-박희준, 김경수 (충남대학교), 안성득, 김상호, 김문호, 이영채, 권수현 (부산대학교) | |
| | | [M3] 증강현실과 음성인식 기능을 이용한 웨브 훈련 프로그램의 베이식게 실어증 환자의 말 명료도에 미치는 효과: 사례보고  
-신혜란, 김영현, 권순숙 (부산대학교) | |
<table>
<thead>
<tr>
<th>제3부 구두발표</th>
<th>총11편</th>
</tr>
</thead>
<tbody>
<tr>
<td>음성학 (강의실D (203호))</td>
<td>음성학 (강의실E (204호))</td>
</tr>
<tr>
<td>말소리장애 및 음성의학 (다목적홀 (112호))</td>
<td>음성학 (강의실E (204호))</td>
</tr>
<tr>
<td>화장 &amp; 사회:</td>
<td>화장 &amp; 사회:</td>
</tr>
<tr>
<td>박성희 (포스트리치)</td>
<td>김지은 (가톨릭관동대)</td>
</tr>
<tr>
<td>토론: 박행만 (서강대)</td>
<td>토론: 윤완희 (여성대)</td>
</tr>
<tr>
<td>강신미 (서경대)</td>
<td>관순봉 (부산대)</td>
</tr>
<tr>
<td>제3부</td>
<td>제3부</td>
</tr>
<tr>
<td>16:30-17:30</td>
<td>16:30-17:30</td>
</tr>
<tr>
<td>음정음표현</td>
<td>[M6] 내전행 연속성 발생장에 음성의학, 정현아 (서울대학교)</td>
</tr>
<tr>
<td>음성학, 정현아 (서울대학교)</td>
<td>음음소음: 강도 변화에 따른</td>
</tr>
<tr>
<td>Zoo Show (서울대학교)</td>
<td>음음소음: 강도 변화에 따른</td>
</tr>
<tr>
<td>[M7] 내전행 연속성 발생장에 음성의학, 정현아 (서울대학교)</td>
<td>[M6] 내전행 연속성 발생장에 음성의학, 정현아 (서울대학교)</td>
</tr>
<tr>
<td>시간</td>
<td>이동 및 만찬 (교외 식당 '늘푸른 정원')</td>
</tr>
<tr>
<td>------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>17:30~21:00</td>
<td></td>
</tr>
</tbody>
</table>

●● 동계날 (5/21일(토)): 한국음성학회 2016 봉 학술대회 프로그램

- 광주과학기술원(GIST) 오렴관

<table>
<thead>
<tr>
<th>일</th>
<th>발표 및 내용</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.21(토)</td>
<td>동록/포스터세션 준비</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>이동 &amp; (1층 로비)</th>
<th>9:30~10:50</th>
</tr>
</thead>
<tbody>
<tr>
<td>제4부 포스터발표 II</td>
<td>9:50</td>
</tr>
<tr>
<td>제4부 (1층 로비)</td>
<td>10:50</td>
</tr>
</tbody>
</table>

- [E14] 이휘백: 모델의 성능향상을 위한 빅 데이터 분석 정보 반영 -유주홍, 김희진 (한국과학기술원)

- [E15] 한국어 발음에 자동생성과 음성인식에 기반한 음성전산 지원 시스템 -남선호, 정민화 (서울대학교)


- [E17] 음성합성 시스템의 음소 지속시간 예측을 위한 딥러닝 알고리즘 비교 연구 -김현주, 김성진 (메이비주)

- [E18] 음성합성을 위한 DNN 기반 파라미터 생성 알고리즘 -박성준, 박진욱, 김준태, 한민수 (한국과학기술원)

- [E19] 쌍해 같은 환경에서 모디오 이어박스를 이용한 음성명령도 개선 -윤덕규, 남시연, 홍재훈, 최승호 (서울과학기술대학교)

- [E20] WSI을 이용한 CTC 기반 음향모델 성능 평가 -이도현, 임영규, 박호성, 김지완 (서강대학교)

- [M10] 연주가 경기 방면에 따른 대라말하기 경사 수행력 비교 - 5, 6세 아동을 위한 컴퓨터와 인쇄물 기반 경사 -오경아, 김수진 (나사렛대학교)

- [M11] 일반유치원 아동의 언어 및 말소리장애에 시청 -서운영, 김수진 (나사렛대학교)

- [M12] 말다툼화자에 대한 작업적 조언 관련 요인 분석 -유하진

<table>
<thead>
<tr>
<th>1층 로비</th>
<th>9:00~9:30</th>
</tr>
</thead>
<tbody>
<tr>
<td>일반 및 (1층 로비)</td>
<td>9:30~9:50</td>
</tr>
<tr>
<td>일반 및 (1층 로비)</td>
<td>9:50~10:50</td>
</tr>
<tr>
<td>10:50-11:00</td>
<td>휴식</td>
</tr>
<tr>
<td>1층</td>
<td>다음작</td>
</tr>
<tr>
<td>다음작</td>
<td>12:00</td>
</tr>
<tr>
<td></td>
<td>1층</td>
</tr>
<tr>
<td>다음작</td>
<td>12:30-13:00</td>
</tr>
</tbody>
</table>

---

### 한국음성학회 2016 봉 학술대회 튜토리얼 세션

- 녹주과학기술원(GIST) 오류관

<table>
<thead>
<tr>
<th>일시</th>
<th>튜토리얼 세션</th>
<th>사회</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.21일</td>
<td>13:30-16:40</td>
<td>이식재 (연세대학교)</td>
</tr>
<tr>
<td>13:30-15:30</td>
<td>음성학과 음성의학 전공자를 위한 신호처리기술 소개</td>
<td>김형순 (부산대학교)</td>
</tr>
<tr>
<td>(2층 강의실) (03호)</td>
<td>15:30-15:40</td>
<td>휴식</td>
</tr>
<tr>
<td>15:40-16:40</td>
<td>음소 단위 분할 및 레이블링 소프트웨어 배포 및 사용 방법 소개</td>
<td>나인수, 정민희 (서울대학교)</td>
</tr>
</tbody>
</table>
차 례

기조발표 I
응성언어와 인공지능 (I): ETRI 연구 사례-자동화역. 대화인터페이스, DeepQA
이윤근(한국전자통신연구원) / 3

기조발표 II
응성언어와 인공지능 (II): NAVER 연구 사례
김선희(네이버) / 23

포스터발표 I
[E1] 잠재 환경에서 DNN 기반 음성인식에 VTLP 적용 실험
김민식, 김형순 (부산대학교), 이성주 (한국전자통신연구원) / 51

[E2] 잠재 및 잔향 환경에서 i-vector 기반 화자확인 전처리 성능 비교
이지희, 김민식, 김형순 (부산대학교), 김동화 (대경대학교) / 53

[E3] 발성유의 기본주파수 및 발생균간 검출을 이용한 청각장면분석 기법
이정현, 이준민, 조지원, 박정민 (서강대학교) / 55

[E4] 선형 예측 분석을 통한 CNSC 기반 2-화자 환경 발화 전처리 구간 검출
이승형, 김재석, 최희진, 한관수 (한국과학기술원) / 57

[E5] 자연어 음성언어에 대한 음성데이터 중장방법
이성주, 강병욱, 정훈, 박정규, 이유근 (한국전자통신연구원) / 59

[E6] 개인화 음성합성 개발을 위한 소유량 발성목록 추출
이민기, 김재민, 정혁주, 김선희 (네이버 엡스) / 61

[E7] 한국인 영어 학습자의 발화 능숙도에 따른 음소영주 병 발화속도 차이 비교
나인수, 정민화 (서울대학교) / 63

[E8] 학습 DB 차이에 따른 TDNN 기반 i-vector 화자인식 시스템의 성능 변화
윤수록, 권오윤 (충북대학교) / 65

[E9] 말, 언어가 늙은 아동의 발하기능을 향상시키기 위한 의사소통 훈련 앱 개발
김진표, 최대빈, 고용완, 김용용, 이용주 (원광대학교), 이영미 (동영대학교) / 67

[M1] 아데노이드변화 비대, 알레르기성 비염을 가진 아동과 정상 아동의 비성도와 기타 음성 특성
김난숙, 성철재 (충남대학교) / 69

[M2] 설어증 환자의 동사 산출에 증강현실 콘텐츠가 미치는 효과: 예비 연구
박희준 (임용보건대학교), 안성욱, 안병강, 백인호, 김근호, 이연우, 권순복 (부산대학교) / 71

[M3] 증강현실과 음성인식 기능을 이용한 앱 훈련 프로그램이 배르니케 설어증 환자의 말 명료도에 미치는 효과: 사례보고
신혜란, 김형순, 권순복 (부산대학교) / 73
[M4] 정상성인의 모음과 자발화 종류에 따른 캡스트림과 스펙트럼 특성
최성희, 최철희 (대구가톨릭대학교) / 75

[M5] 문장유형에 따른 고기능 자폐스펙트럼장애 아동의 운율 특성
신희백, 최지은, 이윤경 (한림대학교) / 77

[M6] 언어치료전공 대학생의 언어재활사에 대한 인식 유형
이은경, 김민경 (동신대학교) / 79

[P1] L2 Effect on Fundamental Frequency: Comparison between Korean and English Produced by Monolingual and Bilingual Speakers
Soo Bin Lim, Goun Lee, Seok-Chae Rhee (Yonsei Univ.) / 81

Eun Jong Kong (Korea Aerospace Univ.), Jeffrey J. Holliday (Korea Univ.) / 83

[P3] 한국 초등학생의 L2 영어발화에서 활속 및 휙자가 발음 유창성 평가에 미치는 영향 연구
이지에, 이석재 (연세대학교) / 85

[P4] 미니언즈 음성 분석을 통한 감정인식 연구
윤광열, 이인서, 이인규, 신지영 (고려대학교), 김윤정 (평화에나주립대학교) / 87

[P5] 학습 수준에 따른 한국어 학습자의 단모음의 발화 양상 비교
권진희 (연세대학교) / 89

이시히, 이석재 (연세대학교) / 91

[P7] 발음도와 분절음 지속시간 변조에 따른 정상 성인의 말더듬 지각
이세정, 성철재 (충남대학교) / 93

[P8] 만 2-4세 아동의 단모음과 이중모음 산출 특징
송인미, 성철재 (충남대학교) / 95

[P9] 아시아권 비원어민 영어 화자 간 대화에 나타난 분절음 특성 연구
정현성 (한국교원대학교) / 97

구두발표
[음성공학]

[E10] 조음 기반의 음소 레벨 사후 학습을 이용한 한국인 영어 학습자 유창성 자동 평가
류혁수, 정민화 (서울대학교) / 101

[E11] DNN 기반의 통계 파라미터 음성합성을 위한 다(多)차자 음향모델링
김상진, 김형준 (NAVER(주)) / 103

[E12] 마비말 장애 음성인식을 위한 발음변형을 결합한 변별학습 기법
성우경, 김남길, 김홍국 (광주과학기술원) / 105

[E13] 잡음에 강한 화자인증을 위한 I-vector 후처리 기법 연구
허휘수, 양일호, 윤성헌, 유하진 (서울시립대학교) / 107

- xii -
[말소리장애 및 음성의학]
• [M7] 내전형 연축성 발성장애 음성의 과제 의준도
  심희정, 고도홍 (한림대학교) / 111

• [M8] 자폐스펙트럼장애 음성의 일레마
  이정현, 김성태 (동신대학교) / 113

• [M9] 음율·소음·강도 변화에 따른 문장 지각 능력 변화
  장선아, 장은주, 장재진 (우송대학교) / 115

[음성학]
• [P10] Individual differences in the relation between perception and production and the mechanisms of phonetic imitation
  Donghyun Kim, Meghan Clayards (McGill University) / 119

  Eon-Suk Ko (Chosun Univ.) / 121

• [P12] 음운구 초 음절 '일'의 고조 실현 연구
  유도영, 신지영 (고려대학교), 김경화 (대전성당성) / 123

• [P13] 동급화된 한국인의 L2 영어 음성코퍼스 구축
  이석재 (연세대학교) / 125

포스터발표 II
• [E14] 여취벡터 모델의 성능향상을 위한 빅 데이터 분석 정보 반영
  유주홍, 김희란 (한국과학기술원) / 129

• [E15] 한국어 발음열 자동생성과 음성인식에 기반한 음성전사 지원 시스템
  나민수, 정민화 (서울대학교) / 131

• [E16] Adaptation of Acoustic Models Using Multilinear Principal Component Analysis
  Yongwon Jeong, Hyung Soon Kim (Pusan National Univ.) / 133

• [E17] 음성합성 시스템의 음소 지속시간 예측을 위한 달리닝 알고리즘 비교 연구
  김형준, 김상진 (네이버(주)) / 135

• [E18] 음성합성을 위한 DNN 기반 파라미터 생성 알고리즘
  박성준, 박진욱, 김동태, 한인수 (한국과학기술원) / 137

• [E19] 실내 잡음 환경에서 오디오 어필레이저를 이용한 음성명료도 개선
  윤덕규, 남시연, 홍세헌, 최승호 (서울과학기술대학교) / 139

• [E20] WSI를 이용한 CTC 기반 음량모델 성능 평가
  이동현, 임민규, 박호성, 김지환 (서강대학교) / 141
[M10] 연령과 검사 방법에 따른 따라말하기 검사 수행력 비교 - 5, 6세 아동을 위한 컴퓨터와 인체물 기반 검사
    오경아, 김수진 (나사렛대학교) / 143

[M11] 일반유치원 아동의 언어 및 말소리장에 현황
    서은영, 김수진 (나사렛대학교) / 145

[M12] 말더듬과잡에 대한 직업적 조언 관련 요인 분석
    박홍주, 장혜경, 박선영 (충남대학교), 박진 (나사렛대학교) / 147

[M13] 말소리장애 아동의 언어 문제
    고유경, 김수진 (나사렛대학교) / 149

[P14] Does speech mode affect the fundamental frequency of the L2 English spoken by
    native Korean learners?
    Ye-jee Jung, Goun Lee, Seok-Chae Rhee (Yonsei Univ.) / 151

[P15] 초점어 발화에서 중국인 학습자의 한국어 폐쇄음 연구
    이선하, 오미라 (전남대학교) / 153

[P16] 한국어 이중모음 /의/의 분포와 특성
    박선우 (경명대학교) / 155

[P17] 독일어 이말 폐쇄음의 문장 위치별 파열 양상
    김소연, 이석재 (연세대학교) / 157

[P18] 제주어 강자율과 강세구 양양
    이숙향 (원광대학교) / 159
한국어 발음열 자동생성을 음성인식에 기반한 음성전사 지원 시스템

1. 서론

발화를 구성하는 단어 및 음소 등 단위에 대한 시간적 경계 분할(세그먼테이션)과 분절음의 기호 표기(레이블링)를 포함하는 음성전사는 음성의 특성분석에 앞서 공통적으로 실행되는 작업이다. 음성을 전사하는 방법은 크게 음성학적 지식을 가진 전문가에 의한 수동전사와 음성인식기술을 이용하는 자동전사[1]로 구분된다. 자동전사는 수동전사에 비해 유지비용이 적고 소요시간이 적으며 전사시행 간 일관성의 유지를 쉽다는 장점이 가진다. 그에 반해 전사시스템의 구축을 위해서는 음향모델, 발음어천 등 음성인식기의 모델과 디코더 등의 구성요소 개발과 요소의 결합을 위한 초기 비용이 크다는 단점이 있다. 본 논문에서는 음성분석에 필요한 세그먼테이션과 레이블링의 초기결과를 출력하는 음성전사 시스템의 구성요소와 기능에 대해 기술한다. 또한 연구자가 전사시스템의 구축비용 없이 음성분석을 위한 자동전사를 수행할 수 있도록 개발된 시스템을 공개한다.

2. 음성전사 지원 시스템


\[ \text{likelihood}_u = -\log \sum_{t=1}^{T} P(x_t | p_u) \]  

인덱스 t는 u의 발화시간을 나타낸다. \(x_t\)는 시간 t간간
여서 음성의 특정벡터를 가리키고 p는 해당시간에 대해 가장 높은 HMM 점수를 출력하는 음소의 인덱스이다. 강제정렬로 생성된 단어 및 음소의 시간계정정보
와 분절음의 레이블 등의 결과들은 HTK MLF 형식

시스템의 구성요소 중 음향모델과 발음열 생성과정
이 특정 언어에 종속적이다. 토론문에서는 한국인
이 발화한 한국어와 미국인 영어 발화로 각각 음향모델을
만들었고 한국어 발음열 생성기, CMU 발음사전[2] 및
영어 발음열 생성기 등을 사용하여 한국어와 영어 음
성의 단어 및 음소 단위의 시간계정 세그먼테이션과
음소 레이블링 기능을 지원한다.

4. 실험

개발된 시스템의 성능평가를 위해 한국어 연속음성
t 데이터 중 1,957 문장을 선택하고 해당 문장의 수동전
사 결과와 자동전사 결과를 비교해 오류율을 측정했
다. 정답음소와 자동전사 음소를 NIST sclite를 통해
사용하여 시간과 음소 레이블을 기준으로 정렬하고 결
과를 일치, 대치, 삽입, 삭제 등으로 분류하였다. 결과가
일치, 삽입, 삭제로 분류되는 경우의 합을 오류의 총합
으로 보고 삽입, 대치, 삭제의 총합을 전체로 보아 오
류의 비율을 계산했다. 세그먼테이션의 성능은 오류율
의 시간을 기준으로 평가하고 레이블링의 성능은 오류
레이블의 기수를 기준으로 평가했다. 평가결과는 표 1
과 같다.

표 1. 세그먼테이션 및 레이블링 성능평가

<table>
<thead>
<tr>
<th></th>
<th>일치</th>
<th>불일치</th>
<th>오류율</th>
</tr>
</thead>
<tbody>
<tr>
<td>세그먼테이션</td>
<td>8928 sec.</td>
<td>1015 sec.</td>
<td>10.5 %</td>
</tr>
<tr>
<td>레이블링</td>
<td>147,263</td>
<td>19,541</td>
<td>12.0 %</td>
</tr>
</tbody>
</table>

성능평가 결과 자동전사된 레이블 중 12%가 대치
또는 삽입, 삭제 등 오류를 보였고 시간적으로는
10.5% 구간에서 오류율을 보였다. 19,541개의 레이블링
오류중 대치오류는 12,305개, 삽입오류는 4,465개, 삽
입오류는 2,771개이다. 전체 대치오류 중 37.5%는 모음
/ə/→/ı/ 혹은 /ɛ/→/ɛ/ 유음변체에서 초성유음의 유성
음화, 초성 /s/ 탈락의 과대적응에 의한 중성 /ɛ/ 대
치 등이고 삽입오류의 54.7%는 종성 /ɣ/, /ɛ/의 삽
입이다. 원인은 /배꾸기/와 같은 불규칙 음운현상 발생,
/치음의 파어/와 같은 어절사이의 음운현상 발생 등이다.
직접적오류의 45%는 초성 /s/ 탈락의 과
대적응에 의해 발생했다. 이 중 초성유음의 유성음화,
초성 /s/ 탈락 과대적응, 종성 /ɣ/, /ɛ/의 삽입, 불
규칙/어절사이 음운현상 등은 발음열 생성과정에 반영
하여 실제 음성에서 발생할 수 있는 발음변이로 사전
에 추가함으로써 오류율 줄일 수 있을 것이다.

5. 결론

본 논문에서는 음성문서의 기초를 둔 세그먼테이
션과 레이블링을 지원하기 위한 발음열 생성과 음성인
식에 기반한 음성전사 시스템의 구성요소와 기능에 대
해 기술하였다. 제공된 시스템은 연구자가 분석 데이터
에 대한 분절음의 시간 및 레이블 정보를 축적하는 과
정에 활용되어 전사 작업의 편의성을 높일 수 있을 것
이다. 수동 전사물 자동 전사물 비교 결과 세그
먼테이션 10.5%, 레이블링 12%의 오류율을 보였다.
향후 발음열 생성과정에서 음운현상 적용을 수행함으로
써 전사지원 시스템의 성능을 높일 것이다.

프로그램 다운로드: http://slp.snu.ac.kr/program/

감사의 글
이 논문은 2015년도 정부(교육부)의 재원으로 한국
연국재단의 지원을 받아 수행된 기초연구사업(No.
2015R1D1A2A01061378)

참고문헌
터 프로그램]. 버전 2.06. http://cjseong.blog.com 또는
http://speech.cbmu.ac.kr.

dictionary. [Online]. Available http://www.speech.cs.cmu.edu/cgi-bin/cmdict

subnetwork caching model for large vocabulary continuous speech recognition. IEICE

Variation for Large Vocabulary Continuous Speech Recognition in Korean. IEICE