

2016년 한국음성학회 봄 학술대회  
발표 논문집



주제: 음성언어와 인공지능  
일시: 2016년 5월 20일(금) ~ 5월 21일(토)  
장소: 광주과학기술원 오룡관  
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)

시간	행사		
11:00~13:00	등록 및 중식 (광주과학기술원 오룡관 로비 및 구내식당)		
13:00~13:10	<b>개회식 (오룡관 1층 다목적홀)</b>		
13:10~14:10	<b>기조발표 I (다목적홀)</b>		
	이윤근 (한국전자통신연구원) “음성언어와 인공지능 (I): ETRI 연구 사례-자동통역, 대화인터페이스, DeepQA”		
14:10~14:30	휴식 및 포스터세션 준비		
14:30~16:20	<b>포스터발표 세션 I [24편]</b>		
	<b>각 포스터 1분 PPT 발표 (다목적홀)</b>		
	<b>포스터발표 (로비)</b>		
16:20~16:30	휴식		
16:30~17:30	<b>구두발표 11편</b>		
	음성공학 (강의실D (203호))	말소리장애 및 음성의학 (다목적홀)	음성학 (강의실E (204호))
17:30~21:00	이동 및 만찬 (교외식당 '늘푸른 정원')		

### ■ 5월 21일(토)

시간	행사		
9:00~9:30	등록 및 포스터세션 준비 (오룡관 로비)		
9:30~10:50	<b>포스터발표 세션 II [16편]</b>		
	<b>각 포스터 1분 PPT 발표 (다목적홀)</b>		
	<b>포스터발표 (로비)</b>		
10:50~11:00	휴식		
11:00~12:00	<b>기조발표 II (다목적홀)</b>		
	김선희 (NAVER)		
	“음성언어와 인공지능 (II): NAVER 연구 사례”		
12:00~12:30	스펙트로그램 리딩/우수발표자 시상 및 폐회식 (다목적홀)		
12:30~13:30	점심 식사 (구내식당)		

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

5.21일	음성학과 음성의학 전공자를 위한 신호처리기술 소개	김형순 (부산대)
13:30~16:40	음소 단위 분할 및 레이블링 소프트웨어 배포 및 사용 방법 소개	나민수, 정민화 (서울대)

● 첫째날 (5/20일(금)): 한국음성학회 2016 봄 학술대회 프로그램

□ 광주과학기술원(GIST) 오룡관

일시		발표 및 내용	비고
첫째날 5.20(금)			
1층 로비	11:00~ 13:00	등록, 점심식사 (구내식당)	
제1부 (1층 다목적 홀)	13:00~ 13:10	개회사 (한국음성학회 이호영 수석부회장) 축사 (한국음성학회 조철우 회장) 환영사 (광주과학기술원 문승현 총장)	1부 사회: 김재욱 (강남대)
	13:10~ 14:10	<b>제1부 기조발표 I</b> 음성언어와 인공지능 (I): ETRI 연구 사례-자동통역, 대화인터페이스, DeepQA: 이윤근 (한국전자통신연구원, 자동통역 언어지능 연구부장)	
	14:10~ 14:30	휴식/포스터세션 준비	
제2부 (1층 다목적 홀 & 1층 로비)	14:30~ 15:00	<b>제2부 포스터발표 I</b> 세션시작 전 각 포스터 1분내 구두 요약 PPT [총24편]	2부 진행 전체총괄: 신종원 (GIST)
		■[E1] 잡음 환경에서 DNN 기반 음성인식에 VTLP 적용 실험 -김민식, 김형순 (부산대학교), 이성주 (한국전자통신연구원)	진행 및 심사:  장선아 (우송대)  윤영선 (한남대)  권철홍 (대전대)
		■[E2] 잡음 및 잔향 환경에서 i-vector 기반 화자확인 전처리 성능 비교 -이가희, 김민식, 김형순 (부산대학교), 김경화 (대검찰청)	
		■[E3] 발성음의 기본주파수 및 발생구간 검출을 이용한 청각장면분석 기법 -이정환, 이준민, 조지원, 박형민 (서강대학교)	
		■[E4] 선형 예측 분석을 통한 CNSC 기반 2-화자 환경 발화 겹침 구간 검출 -이승형, 김재석, 최희진, 한민수 (한국과학기술원)	
		■[E5] 자연어 음성인식을 위한 음성데이터 증강방법 -이성주, 강병옥, 정훈, 박전규, 이윤근 (한국전자통신연구원)	
		■[E6] 개인화 음성합성 개발을 위한 소용량 발성목록 추출 -이민기, 김재민, 홍진표, 김선희 (네이버 랩스)	
		■[E7] 한국인 영어 학습자의 발화 능숙도에 따른 음소범주 별 발화속도 차이 비교 -나민수, 정민화 (서울대학교)	
		■[E8] 학습 DB 차이에 따른 TDNN 기반 i-vector 화자인식 시스템의 성능 변화 -윤성욱, 권오욱 (충북대학교)	
		■[E9] 말, 언어가 늦은 아동의 말하기 능력을 향상시키기 위한 의사소통 훈련 앱 개발 -김지호, 최대림, 고락환, 김용운, 이용주 (원광대학교), 이영미 (동명대학교)	
	■[M1] 아데노이드편도 비대, 알레르기성 비염을 가진 아동과 정상 아동의 비성도와 기타 음향 특성 -김난숙, 성철재 (충남대학교)		
	■[M2] 실어증 환자의 동사 산출에 증강현실 콘텐츠가 미치는 효과: 예비 연구 -박희준 (춘해보건대학교), 안신욱, 안병강, 배인호, 김근호, 이연우, 권순복 (부산대학교)		
	■[M3] 증강현실과 음성인식 기능을 이용한 웹 훈련 프로그램이 베르니케 실어증 환자의 말 명료도에 미치는 효과: 사례보고 -신혜란, 김형순, 권순복 (부산대학교)		

- [M4] 정상성인의 모음과 자발화 종류에 따른 캡스트럼과 스펙트럼 특성  
-최성희, 최철희 (대구가톨릭대학교)
- [M5] 문장유형에 따른 고기능 자폐스펙트럼장애 아동의 운율 특성  
-신희백, 최지은, 이윤경 (한림대학교)
- [M6] 언어치료전공 대학생의 언어재활사에 대한 인식 유형  
-이은경, 김민경 (동신대학교)
- [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.)
- [P2] An Acoustic Investigation of Affricates in Seoul Corpus  
-Eun Jong Kong (Korea Aerospace Univ.), Jeffrey J. Holliday (Korea Univ.)
- [P3] 한국 초등학생의 L2 영어발화에서 화속 및 휴지가 발음 유창성 평가에 미치는 영향 연구  
-이지예, 이석재 (연세대학교)
- [P4] 미니언즈 음성 분석을 통한 감정인식 연구  
-윤광열, 전인서, 이인구, 신지영 (고려대학교), 김윤정 (루이지애나주립대학교)
- [P5] 학습 수준에 따른 한국어 학습자의 단모음의 발화 양상 비교  
-권진희 (연세대학교)
- [P6] 한국인 L2 영어 발화에서 단어초 [p, t, k]와 [sp, st, sk]의 구분과 발화 유창성 등급과의 관련성 연구  
-이서희, 이석재 (연세대학교)
- [P7] 말속도와 분절음 지속시간 변조에 따른 정상 성인의 말더듬 지각  
-이세정, 성철재 (충남대학교)
- [P8] 만 2-4세 아동의 단모음과 이중모음 산출 특징  
-송인미, 성철재 (충남대학교)
- [P9] 아시아권 비영어권 영어 화자 간 대화에 나타난 분절음 특성 연구  
-정현성 (한국외국어대학교)

오미라  
(전남대)  
  
박한상  
(홍익대)

16:20~  
16:30

휴식

**제3부 구두발표 [총 11편]**

음성공학 (강의실D (203호))	말소리장애 및 음성의학 (다목적홀 (112호))	음성학 (강의실E (204호))
좌장 & 사회: 이근배 (POSTECH)	좌장 & 사회: 박상희 (대구사이버대)	좌장 & 사회: 김지은 (가톨릭관동대)
토론: 박형민 (서강대) 강선미 (서경대)	토론: 최철희 (대구가톨릭대) 권순복 (부산대)	토론: 윤원희 (계명대) 공은정 (항공대)
■[E10] 조음 기반의 음소 레벨 사후 확률을 이용한 한국인 영어 학습자 유창성 자동 평가 -류혁수, 정민화 (서울대학교)	■[M7] 내전형 연속성 발성장애 음성의 과제 의존도 -심희정, 고도홍 (한림대학교)	■[P10] Individual differences in the relation between perception and production and the mechanisms of phonetic imitation -Donghyun Kim, Meghan Clayards (McGill University)
■[E11] DNN 기반의 통계 파라미터 음성합성을 위한 다(多)화자 음향모델링 -김상진, 김형준 (NAVER(주))	■[M8] 자폐스펙트럼장애 음성의 딜레마 -이정현, 김성태 (동신대학교)	■[P11] Tempo-dependent effects of coda-voicing on duration in CVC words. -Eon-Suk Ko (Chosun Univ.)
■[E12] 마비말 장애 음성인식을 위한 발음변이를 결합한 변별학	■[M9] 운율-소음-강도 변화에 따른 문장 지각 능력 변화 -장선아, 장은주, 장재진 (우송대학교)	

제3부

16:30~  
17:30

	습 기법 -성우경, 김남균, 김흥국 (광주과학기술원)  ■[E13] 잡음에 강인한 화자인증을 위한 I-vector 후처리 기법 연구 -허희수, 양일호, 윤성현, 유하진 (서울시립대학교)	■[P12] 음운구 초 음절 '일'의 구조 실현 연구 -유도영, 신지영 (고려대학교), 김경화 (대검찰청)  ■[P13] 등급화된 한국인의 L2 영어 음성코퍼스 구축 -이석재 (연세대학교)
17:30~21:00	이동 및 만찬 (교외 식당 '늘푸른 정원')	

●● 둘째날 (5/21일(토)): 한국음성학회 2016 봄 학술대회 프로그램

□ □ 광주과학기술원(GIST) 오룡관

일시		발표 및 내용	비고
둘째날 5.21(토)			
1층 로비	9:00~9:30	등록/포스터세션 준비	
제4부 (1층 다목적 홀 & 1층 로비)	9:30~9:50	<b>제4부 포스터발표 II</b> 세션시작 전 각 포스터 1분내 구두 요약 PPT [총16편]	4부 진행 전체총괄: 김수진 (나사렛대)
	9:50~10:50	■[E14] 어휘벡터 모델의 성능향상을 위한 빅 데이터 분석 정보 반영 -유주홍, 김희린 (한국과학기술원)	
		■[E15] 한국어 발음열 자동생성과 음성인식에 기반한 음성전사 지원 시스템 -나민수, 정민화 (서울대학교)	진행 및 심사:
		■[E16] Adaptation of Acoustic Models Using Multilinear Principal Component Analysis -Yongwon Jeong, Hyung Soon Kim (Pusan National Univ.)	이주경 (서울시립대)
		■[E17] 음성합성 시스템의 음소 지속시간 예측을 위한 딥러닝 알고리즘 비교 연구 -김형준, 김상진 (네이버(주))	하지완 (대구대)
		■[E18] 음성합성을 위한 DNN 기반 파라미터 생성 알고리즘 -박상준, 박진욱, 김준태, 한민수 (한국과학기술원)	최성희 (대구가톨릭대)
		■[E19] 실내 잡음 환경에서 오디오 이퀄라이저를 이용한 음성명료도 개선 -윤덕규, 남시연, 홍세린, 최승호 (서울과학기술대학교)	홍기형 (성신여대)
		■[E20] WSJ을 이용한 CTC 기반 음향모델 성능 평가 -이동현, 임민규, 박호성, 김지환 (서강대학교)	유하진
		■[M10] 연령과 검사 방법에 따른 따라말하기 검사 수행력 비교 - 5, 6세 아동을 위한 컴퓨터와 인쇄물 기반 검사 -오경아, 김수진 (나사렛대학교)	
		■[M11] 일반유치원 아동의 언어 및 말소리장애 현황 -서은영, 김수진 (나사렛대학교)	
■[M12] 말더듬화자에 대한 직업적 조언 관련 요인 분석			

		-박흥주, 장혜경, 박선영 (충남대학교), 박진 (나사렛대학교) ■[M13] 말소리장애 아동의 언어 문제 -고유경, 김수진 (나사렛대학교) ■[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.) ■[P15] 초점어 발화에서 중국인 학습자의 한국어 폐쇄음 연구 -이선하, 오미라 (전남대학교) ■[P16] 한국어 이중모음 /의/의 분포와 특성 -박선우 (계명대학교) ■[P17] 독일어 어말 폐쇄음의 문장 위치별 파열 양상 -김소연, 이석재 (연세대학교) ■[P18] 제주어 강자음과 강세구 역양 -이숙향 (원광대학교)	(서울시립대)  김홍국 (GIST)
	10:50~ 11:00	휴식	
1층 다목적 홀	11:00~ 12:00	<b>제5부 기조발표 II</b> 음성언어와 인공지능 (II): NAVER 연구 사례: 김선희 (NAVER 수석연구원)	사회: 이호영 (서울대)
1층 다목적 홀	12:00~ 12:30	스펙트로그램 리딩 & 폐회식	사회: 김재욱 (강남대)  스펙트로 그램리딩: 정현성 (한국교원대)
	12:30~ 13:30	점심식사 (구내식당)	

### ●●●한국음성학회 2016 봄 학술대회 튜토리얼 세션

□ □ □ 광주과학기술원(GIST) 오룡관

일시		사회
5.21일	튜토리얼 세션	이석재
13:30~16:40		(연세대학교)
		특강자
	음성학과 음성의학 전공자를 위한 신호처리기술 소개	김형순
		(부산대학교)
(2층 강의실D (203호))	휴식	
	음소 단위 분할 및 레이블링 소프트웨어 배포 및 사용 방법 소개	나민수, 정민화
		(서울대학교)

# 차 례

## 기조발표 I

음성언어와 인공지능 (I): ETRI 연구 사례-자동통역, 대화인터페이스, DeepQA  
이윤근(한국전자통신연구원) / 3

## 기조발표 II

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

## 포스터발표 I

- [E1] 잡음 환경에서 DNN 기반 음성인식에 VTLP 적용 실험  
김민식, 김형순 (부산대학교), 이성주 (한국전자통신연구원) / 51
- [E2] 잡음 및 잔향 환경에서 i-vector 기반 화자확인 전처리 성능 비교  
이가희, 김민식, 김형순 (부산대학교), 김경화 (대검찰청) / 53
- [E3] 발성음의 기본주파수 및 발성구간 검출을 이용한 청각장면분석 기법  
이정환, 이준민, 조지원, 박형민 (서강대학교) / 55
- [E4] 선형 예측 분석을 통한 CNSC 기반 2-화자 환경 발화 겹침 구간 검출  
이승형, 김재석, 최희진, 한민수 (한국과학기술원) / 57
- [E5] 자연어 음성인식을 위한 음성데이터 증강방법  
이성주, 강병욱, 정훈, 박전규, 이윤근 (한국전자통신연구원) / 59
- [E6] 개인화 음성합성 개발을 위한 소용량 발성목록 추출  
이민기, 김재민, 홍진표, 김선희 (네이버 랩스) / 61
- [E7] 한국인 영어 학습자의 발화 능숙도에 따른 음소범주 별 발화속도 차이 비교  
나민수, 정민화 (서울대학교) / 63
- [E8] 학습 DB 차이에 따른 TDNN 기반 i-vector 화자인식 시스템의 성능 변화  
윤성욱, 권오욱 (충북대학교) / 65
- [E9] 말, 언어가 늦은 아동의 말하기 능력을 향상시키기 위한 의사소통 훈련 앱 개발  
김진호, 최대림, 고락한, 김용운, 이용주 (원광대학교), 이영미 (동명대학교) / 67
- [M1] 아데노이드편도 비대, 알레르기성 비염을 가진 아동과 정상 아동의 비성도와 기타 음향 특성  
김난숙, 성철재 (충남대학교) / 69
- [M2] 실어증 환자의 동사 산출에 증강현실 콘텐츠가 미치는 효과: 예비 연구  
박희준 (춘해보건대학교), 안신욱, 안병강, 배인호, 김근효, 이연우, 권순복 (부산대학교) / 71
- [M3] 증강현실과 음성인식 기능을 이용한 웹 훈련 프로그램이 베르니케 실어증 환자의 말 명료도에 미치는 효과: 사례보고  
신혜란, 김형순, 권순복 (부산대학교) / 73

- [M4] 정상성인의 모음과 자발화 종류에 따른 캡스트럼과 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
- [P2] An Acoustic Investigation of Affricates in Seoul Corpus  
Eun Jong Kong (Korea Aerospace Univ.), Jeffrey J. Holliday (Korea Univ.) / 83
- [P3] 한국 초등학생의 L2 영어발화에서 화속 및 휴지가 발음 유창성 평가에 미치는 영향 연구  
이지예, 이석재 (연세대학교) / 85
- [P4] 미니언즈 음성 분석을 통한 감정인식 연구  
윤광열, 전인서, 이인구, 신지영 (고려대학교), 김윤정 (루이지애나주립대학교) / 87
- [P5] 학습 수준에 따른 한국어 학습자의 단모음의 발화 양상 비교  
권진희 (연세대학교) / 89
- [P6] 한국인 L2 영어 발화에서 단어초 [p, t, k]와 [sp, st, sk]의 구분과 발화 유창성 등급과의 관련성 연구  
이서희, 이석재 (연세대학교) / 91
- [P7] 말속도와 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 확률을 이용한 한국인 영어 학습자 유창성 자동 평가  
류혁수, 정민화 (서울대학교) / 101
- [E11] DNN 기반의 통계 파라미터 음성합성을 위한 다(多)화자 음향모델링  
김상진, 김형준 (NAVER(주)) / 103
- [E12] 마비말 장애 음성인식을 위한 발음변이를 결합한 변별학습 기법  
성우경, 김남균, 김홍국 (광주과학기술원) / 105
- [E13] 잡음에 강인한 화자인증을 위한 I-vector 후처리 기법 연구  
허희수, 양일호, 윤성현, 유하진 (서울시립대학교) / 107

## [말소리장애 및 음성의학]

- [M7] 내전형 연속성 발생장애 음성의 과제 의존도

심희정, 고도홍 (한림대학교) / 111

- [M8] 자폐스펙트럼장애 음성의 딜레마

이정현, 김성태 (동신대학교) / 113

- [M9] 운율·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

- [P11] Tempo-dependent effects of coda-voicing on duration in CVC words.

Eon-Suk Ko (Chosun Univ.) / 121

- [P12] 음운구 초 음절 '일'의 고조 실현 연구

유도영, 신지영 (고려대학교), 김경화 (대검찰청) / 123

- [P13] 등급화된 한국인의 L2 영어 음성코퍼스 구축

이석재 (연세대학교) / 125

## 포스터발표 II

- [E14] 어휘벡터 모델의 성능향상을 위한 빅 데이터 분석 정보 반영

유주홍, 김희린 (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 딥러닝 알고리즘 비교 연구

김형준, 김상진 (네이버(주)) / 135

- [E18] 음성합성을 위한 DNN 기반 파라미터 생성 알고리즘

박상준, 박진욱, 김준태, 한민수 (한국과학기술원) / 137

- [E19] 실내 잡음 환경에서 오디오 이퀄라이저를 이용한 음성명료도 개선

윤덕규, 남시연, 홍세린, 최승호 (서울과학기술대학교) / 139

- [E20] WSJ을 이용한 CTC 기반 음향모델 성능 평가

이동현, 임민규, 박호성, 김지환 (서강대학교) / 141

- [M10] 연령과 검사 방법에 따른 따라말하기 검사 수행력 비교 - 5, 6세 아동을 위한 컴퓨터와 인쇄물 기반 검사  
오경아, 김수진 (나사렛대학교) / 143
- [M11] 일반유치원 아동의 언어 및 말소리장애 현황  
서은영, 김수진 (나사렛대학교) / 145
- [M12] 말더듬화자에 대한 직업적 조언 관련 요인 분석  
박흥주, 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 (계명대학교) / 155
- [P17] 독일어 어말 폐쇄음의 문장 위치별 파열 양상  
김소연, 이석재 (연세대학교) / 157
- [P18] 제주어 강자음과 강세구 역양  
이숙향 (원광대학교) / 159

# 조음 기반의 음소 레벨 사후 확률을 이용한 한국인 영어 학습자 유창성 자동 평가

류혁수<sup>1</sup>, 정민화<sup>1</sup>  
<sup>1</sup>서울대학교 언어학과

## Automatic Pronunciation Assessment of English Spoken by Korean Learners Using Phone-level Articulatory Posterior Probability

Hyuksu Ryu<sup>1</sup>, Minhwa Chung<sup>1</sup>

<sup>1</sup>Department of Linguistics, Seoul National University  
[oster01@snu.ac.kr](mailto:oster01@snu.ac.kr), [mchung@snu.ac.kr](mailto:mchung@snu.ac.kr)

### Abstract

This paper proposes novel features for automatic pronunciation assessment of English spoken by Korean learners by using articulatory Goodness-Of-Pronunciation (aGOP). For calculating aGOP, articulatory acoustic models are trained for the corresponding articulatory attributes, such as nasal, sonorant, strident, etc. For assessment modeling, we utilize 5,000 English sentences spoken by Korean learners. As a baseline, Pronunciation scores produced by various features proposed by ETS and the traditional GOP are used. In order to enhance the baseline performance, we append the proposed aGOPs as novel features to the modeling. The results show that the proposed model adding aGOPs outperforms the baseline. It is noteworthy that integrating linguistic/phonetic knowledge is useful for automatic pronunciation assessment.

### 1. 서론

본 연구는 한국인 영어학습자가 발화한 낭독체 문장에 대한 유창성을 자동으로 평가하는 환경에서 평가 성능 향상을 위해 새로운 평가자질로서 조음 자질 기반의 사후 확률을 제안하는 것을 목적으로 한다. 유창성 자동 평가에 관한 기존 연구로는 ETS에서 개발한 SpeechRater<sup>®</sup>이 있다 [1]. ETS의 연구에서는 다양한 모국어 사용자 학습자의 영어 즉흥 발화를 대상으로 유창성을 평가하기 위하여 29종의 자질을 제안한 바 있다. 하지만, ETS의 연구에서는 대부분 발화 속도와 분절음 및 휴지의 길이와 관련된 자질이 대부분을 차지하고, 학습자 발음과 관련된 자질이 드물다는 단

점이 있다. 본 연구에서는 ETS에서 제안한 자질 뿐 아니라, 학습자의 발음 특성을 반영하기 위하여 조음에 기반한 사후 확률을 새로운 자질로 제안하고자 한다.

### 2. 실험 방법

#### 2.1 코퍼스와 음향 모델

유창성 발화 평가를 위해 ETRI에서 제공한 한국인 영어학습자 낭독체 발화 코퍼스를 사용하였다. 코퍼스는 251명 학습자가 발화한 40,000 문장으로 구성되어 있으며, 평가 모델링을 위해서 5,000 문장 (4시간 49분 분량), 한국인 영어 발화 음향 모델링을 위해 35,000 문장 (34시간 40분 분량)을 사용하였다. 평가모델에 사용한 5,000 문장 가운데, 평가 모델 학습에 4,000 문장, 테스트에 1,000 문장을 사용하였다.

또한, 사후확률 계산을 위해, 원어인 음향모델을 적용하였다. TIMIT 코퍼스의 630명이 발화한 6,300문장을 이용하여 원어인 음향 모델을 학습하였다.

#### 2.2 평가 모델

한국인 영어학습자의 유창성 자동 평가를 위해 다음의 그림 1과 같은 평가 모델링 방법을 구성하였다. 그림 1에서 보는 바와 같이, 평가 모델 학습에서는 한국인 영어 학습자 음성을 한국인 학습자 음향 모델을 통해 인식한 결과를 원어인 음향모델에 대해 강제 정렬한 결과를 이용하여 사후 확률을 계산한다. 이 때, 음소 뿐 아니라, 조음 자질에 대해서도 조음 사후 확률을 계산한다. 이들 사후 확률을 이용하여 회귀 모델과 수동 평가 점수 결과를 이용하여 평가 모델을 생성한다. 이 후, 평가 모델 테스트 단계에서, 학습자 발화를 이용하여 음소 사후 확률과 조음 사후 확률을 계산하고, 이들 자질을 평가 모델에 적용하여 유창성 점수를

예측한다. 사후 확률 외에 ETS [1]에서 제시한 평가 자질 가운데, 낭독체와 관련이 있는 20개 자질을 이용한 평가모델을 베이스라인으로 사용하였다.

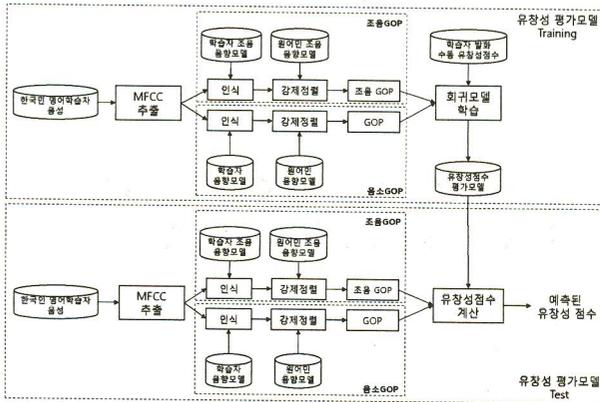


그림 1. 유창성 자동 평가 모델링

### 2.3 조음 기반 사후 확률

본 연구에서는 조음 특성에 기반한 사후확률(aGOP)을 유창성 자동 평가를 위한 자질로 제안하고 있다. 조음 기반 사후 확률은 다음과 같이 계산한다.

$$aGOP_k \equiv \frac{\log P(q_i^k | O_i) / N_i}{\log P(O_i | q_i^k) - \log(\max_j P(O_i | q_j^k))} \quad (1)$$

이 때,  $N_i$  은 음소 길이를 의미하며,  $k$ 는 조음 자질의 종류,  $O_i$ 와  $q_i^k$ 는 각각  $i$ 번째 음소의 관측값과 조음 자질  $k$ 의 값을 의미한다.

자동 평가를 위해, 음성학적 지식에 기반한 13가지 변별적 자질을 조음 자질로 도입하였다 [2]. [2]의 변별적 자질은 음성학/음운론적 지식에 의거하여 각 음소에 대해 자질의 유무를 이진값으로 변환할 수 있다. 예를 들어, voice 라는 자질에 대해 [m]은 [+voice]로, [p]는 [-voice]의 값을 할당할 수 있다.

## 3. 실험 결과

본 연구에서는 ETS에서 제안한 평가 자질 (ETS)을 베이스라인으로 삼고, 음소 기반 사후 확률 (GOP)과 15가지 조음 기반 사후 확률 (aGOP)을 추가한 평가 모델을 생성하여, 이에 대한 유창성 점수 예측 성능을 평가하였다. 성능 평가 결과는 다음의 표 1과 같다. 표 1에서 보는 바와 같이, 단일한 자질 집합을 사용하였을 경우에는 ETS 자질을 사용하는 모델이 음소 혹은 조음 사후 확률을 사용하는 모델보다 더 높은 성능을 보였다 (모델 1-3). 하지만, ETS에 조음 사후 확률 (aGOP)을 더하게 되면 (모델 6), 베이스라인보다 성능이 향상되는 것을 볼 수 있었다. 또한, 모든 종류의 자

질을 모두 포함하였을 때 (모델 7), 0.5507의 성능으로써 가장 높은 성능을 보이는 것을 볼 수 있었다.

No.	모델	상관계수
(1)	ETS	0.5434
(2)	GOP	0.3021
(3)	aGOP	0.5043
(4)	ETS+GOP	0.5456
(5)	GOP+aGOP	0.5151
(6)	ETS+aGOP	0.5492
(7)	GOP+ETS+aGOP	0.5507

표 1 자질 종류에 따른 성능 결과

## 4. 결론

본 연구에서는 한국인 화자의 영어 발화에서의 유창성 자동평가 성능 향상을 위해 조음 기반의 사후 확률이라는 새로운 자질을 제안하고자 하였다. 5,000문장의 한국인 영어 학습자 발화를 이용하여, ETS의 평가 자질과 음소 사후확률, 그리고 조음 사후 확률을 계산하고 이를 이용하여 평가 모델을 구성하였다. 그 결과, ETS의 자질만을 사용하였을 때에 비해, ETS 자질과 조음 사후 확률을 이용할 때 평가 성능이 향상됨을 관찰할 수 있었다. 또한 음소 및 조음 사후 확률을 모두 포함하였을 때 유창성 평가 성능이 가장 좋은 것을 확인할 수 있었다.

본 연구는 한국인 영어 발화의 유창성 평가를 위해 새로운 자질을 제안하고, 제안한 자질이 성능 향상에 기여함을 보였다는 점에서 그 의의를 갖는다. 향후, 본 연구를 기반으로, 조음 자질의 종류를 확장하고, 다른 언어권 학습자의 한국어 발화와 같이 학습자 모국어와 학습 언어 쌍을 확장하여 연구를 진행 할 수 있다.

## 감사의 글

본 연구는 미래창조과학부 및 정보통신기술진흥센터의 정보통신·방송 연구개발사업의 일환으로 수행하였음. [R0126-15-1117, 언어학습을 위한 자유발화형 음성대화처리 원천기술 개발]

## 참고문헌

- [1] Zechner, K., Higgins, D., Xi, X. M., and Williamson, D. M., "Automatic scoring of non-native spontaneous speech in tests of spoken English," *Speech Communication*, vol. 51, no. 10, pp. 883- 895, 2009
- [2] Hayes, B. *Introductory phonology*, Oxford: Wiley-Blackwell, 2011.