

# 2024 한국음성학회 봄 학술대회 발표 논문집



일시: 2024년 5월 18일(토)

장소: 한남대학교 56주년기념관 중회의장

주최: 사단법인 한국음성학회

주관: 사단법인 한국음성학회, 한남대학교

후원: (주)리드스피커코리아, (주)셀바스AI, (사)한국언어재활사협회

**한 국 음 성 학 회**  
**The Korean Society of Speech Sciences**

# 2024 한국음성학회 봄 학술대회 일정표

## 1) 전체 일정표

| 일정          | 발 표 및 내 용                                                                                                              |                                                            |                                                                  |
|-------------|------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------|------------------------------------------------------------------|
| 09:00~14:00 | 등록                                                                                                                     |                                                            |                                                                  |
| 09:40~09:50 | 개회식(중회의장)<br>개회식(사회: 신동진, 전주대학교)                                                                                       |                                                            |                                                                  |
| 09:50~10:40 | 특강 1(중회의장)<br>(사회: 윤영선, 한남대)<br><br>“LLM (Large Language Model) 기술 현황 및 응용”<br>구명완(서강대 컴퓨터공학과) <a href="#">[특강자료로]</a> |                                                            |                                                                  |
| 10:40~10:50 | 휴식                                                                                                                     |                                                            |                                                                  |
| 10:50~12:30 | 구두 발표 I                                                                                                                |                                                            |                                                                  |
|             | 분과 1(중회의장)<br>음성학 및 음운론<br><br>PH01, PH02, PH03, PH04<br>(4편)                                                          | 분과 2(소회의실 1)<br>음성공학<br><br>SE01, SE02, SE03, SE04<br>(4편) | 분과 3(소회의실 2)<br>말장애 및 음성의학<br><br>SD01, SD02, SD03, SD04<br>(4편) |
| 12:30~13:10 | 점심 식사                                                                                                                  |                                                            |                                                                  |
| 13:10~14:00 | 포스터 발표(로비)                                                                                                             |                                                            |                                                                  |
| 14:00~14:50 | 특강 2(중회의장)<br>(사회: 이주경, 서울시립대)<br><br>“ChatGPT를 활용한 문화와 사회성 연구”<br>김지혜(서강대 사회학과) <a href="#">[특강자료로]</a>               |                                                            |                                                                  |
| 14:50~15:00 | 휴식                                                                                                                     |                                                            |                                                                  |
| 15:00~16:15 | 구두 발표 II                                                                                                               |                                                            |                                                                  |
|             | 분과 4(중회의장)<br>음성학 및 음운론<br><br>PH05, PH06, PH07<br>(3편)                                                                | 분과 5(소회의실 1)<br>말장애 및 음성의학<br><br>SD05, SD06, SD07<br>(3편) | 분과 6(소회의실 2)<br>말장애 및 음성의학<br><br>SD08, SD09, SD10<br>(3편)       |
| 16:15~16:25 | 휴식                                                                                                                     |                                                            |                                                                  |
| 16:25~17:40 | 구두 발표 III                                                                                                              |                                                            |                                                                  |
|             | 분과 7(중회의장)<br>음성학 및 음운론<br><br>PH08, PH09, PH10<br>(3편)                                                                | 분과 8(소회의실 1)<br>음성공학<br><br>SE05, SE06, SE07<br>(3편)       | 분과 9(소회의실 2)<br>말장애 및 음성의학<br><br>SD11, SD12, SD13<br>(3편)       |
| 17:40~18:00 | 스펙트로그램 reading(중회의장)<br>우수 발표 시상 및 폐회식<br>(사회: 박상희, 대구사이버대학교)                                                          |                                                            |                                                                  |

● 포스터 발표

좌장: 고연숙(조선대), 정종민(조선대), 박정식(한국외대)

| 시간                  | 구분  | 발표 제목                                                                                                    | 발표자                                                                                                                  |
|---------------------|-----|----------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------|
| 13:10<br>~<br>14:00 | P01 | POSHA-S를 통한 직업군인의 말더듬 인식 연구<br>[논문으로]                                                                    | 차화정, 박진(가톨릭관동대)                                                                                                      |
|                     | P02 | 갑상선 수술 후 증후군의 빠른 음성회복을 위한<br>음성치료의 효과 [논문으로]                                                             | 김지성(건양대)                                                                                                             |
|                     | P03 | 청각적 자극 및 시각적 자극에 따른 말소리장애 아동의<br>단어 회상하기 과제의 수행 비교 [논문으로]                                                | 김주원, 하지완(대구대)                                                                                                        |
|                     | P04 | 아동기말실행증의 말소리 감별진단도구 마련을 위한<br>예비연구 [논문으로]                                                                | 이옥분, 박상희(대구사이버대)                                                                                                     |
|                     | P05 | 게임 닉네임의 독음 후보군 생성 방법 [논문으로]                                                                              | 김나현, 김연정(서울대),<br>이용혁, 조남현(NCSOFT)                                                                                   |
|                     | P06 | 정보 교란 기법과 스피커 인코더를 활용한 한국어<br>제로샷 음성 합성 모델 [논문으로]                                                        | 방채운(광주과학기술원),<br>전찬준(조선대),<br>김홍국(광주과학기술원)                                                                           |
|                     | P07 | 자기 지도 학습 모델 특징 결합을 이용한 적대적<br>신경망 기반 역양 음성인식 [논문으로]                                                      | 경나, 박정식(한국외대)                                                                                                        |
|                     | P08 | 음성 향상을 위한 게이트 합성곱 순환 신경망 [논문으로]                                                                          | 고형일, 나승대,<br>김명남(경북대)                                                                                                |
|                     | P09 | Pitch 정보를 활용한 Perturbation AUTOVC의 성능<br>향상 [논문으로]                                                       | 박화영(광주과학기술원),<br>전찬준(조선대),<br>김홍국(광주과학기술원)                                                                           |
|                     | P10 | 한국어 종성 유음의 발음 변이 연구 [논문으로]                                                                               | 서윤정, 드류 크로스비(고려대)                                                                                                    |
|                     | P11 | 데이터를 활용한 프랑스어 단어 수준 음절 유형 및<br>자음군 분포 분석 [논문으로]                                                          | 박소은, 김선희(서울대)                                                                                                        |
|                     | P12 | 한국인 학습자의 프랑스어 연음 오류 분석 [논문으로]                                                                            | 김민정, 김선희(서울대)                                                                                                        |
|                     | P13 | 시대별 애니메이션 캐릭터 음성에 대한 음향 및 선호도<br>분석 [논문으로]                                                               | 이선미, 조희수, 최영경,<br>김지연, 이윤하,<br>이영미(이화여대)                                                                             |
|                     | P14 | Acoustic Analyses in the Imitation of Emotional<br>Speech in Children with Typical Development<br>[논문으로] | Su Been Kim, Jung Eun<br>Kim, Soo Hyoung Cho,<br>Hyo Sun Lee, Seong Yun<br>Moon, Youngmee Lee<br>(Ewha Womans Univ.) |

# 차 례

## 특 강

|                                                  |    |
|--------------------------------------------------|----|
| <b>특강1</b> LLM (Large Language Model) 기술 현황 및 응용 |    |
| 구명완 .....                                        | 2  |
| <b>특강2</b> ChatGPT를 활용한 문화와 사회성 연구               |    |
| 김지혜 .....                                        | 50 |

## 구두 발표

|                                                                                                                                                        |    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| <b>PH01</b> 대화 상호작용의 개시자 효과와 아동의 언어 성과와의 연관성                                                                                                           |    |
| Jun Ho Chai, Jongmin Jung, and Eon-Suk Ko .....                                                                                                        | 56 |
| <b>PH02</b> Developmental Variations in English Vowel Production : A Proficiency-Based Study of Korean ESL Learners                                    |    |
| Mi Sun Park, Joo-Kyeong Lee .....                                                                                                                      | 58 |
| <b>PH03</b> Effects of Age, Gender, and Language Ability on the Acoustic Fluency and Pronunciation Judgements in English Utterances by Korean Speakers |    |
| Hyunsong Chung .....                                                                                                                                   | 61 |
| <b>PH04</b> A study on the L2 phonetic imitation: Korean learners' case of English stress assignment                                                   |    |
| Mi-Young Bae, Tae-Yeoub Jang .....                                                                                                                     | 62 |
| <b>PH05</b> 경상 방언의 내포절 의문사 음성 자극에 대한 의문사 작용역 응답 통계 분석                                                                                                  |    |
| Weonhee Yun .....                                                                                                                                      | 64 |
| <b>PH06</b> Multimodal cue usage in beginning conversational dyadic exchanges across child development                                                 |    |
| Jun Ho Chai, Barbara Zapir, 고 언 숙 .....                                                                                                                | 65 |
| <b>PH07</b> Temporal Characteristics of Polysemous Words in Spoken Thai                                                                                |    |
| Kyung-Eun Park, Tae-Yeoub Jang .....                                                                                                                   | 67 |
| <b>PH08</b> 한국어 단모음의 조음 양상과 공간 구조 : 3D 초음파 측정법을 기반으로                                                                                                   |    |
| Young-Woo Ha, Young Hwang .....                                                                                                                        | 68 |
| <b>PH09</b> 운율 위치에 따라 달라지는 서울 방언 파열음 지각                                                                                                                |    |
| So Hyunjung, Jonny Jungyun Kim .....                                                                                                                   | 70 |
| <b>PH10</b> 한국어 비유창성 유발 요인에 대한 연구                                                                                                                      |    |
| Jungyeon Hong .....                                                                                                                                    | 71 |

|                                                                                                    |    |
|----------------------------------------------------------------------------------------------------|----|
| <b>SE01</b> 한국어 음성인식을 위한 중단형 음성인식 모델의 실험적 비교                                                       |    |
| Changhan Oh, Kiyong Park, Jeomja Kang, Hwajeon Song, Wooyong Choi .....                            | 73 |
| <b>SE02</b> 짧은 발성에 강인한 ECAPA-TDNN 기반의 화자 인증 시스템                                                    |    |
| Seung-bin Kim, Chan-yeong Lim, Jungwoo Heo, Ju-ho Kim, Hyun-seo Shin, Kyo-won Koo, Ha-Jin Yu ..... | 74 |
| <b>SE03</b> Self-Instruction을 활용한 이모지 생성모델                                                         |    |
| Youngwoo Choi, Bowon Lee .....                                                                     | 75 |
| <b>SE04</b> 실시간 멀티 채널 음성 향상 모델의 출력 방식 비교 분석                                                        |    |
| Tae Han Kim, Hyung-Min Park .....                                                                  | 76 |
| <b>SE05</b> 동적 어트랙터 생성을 위한 토큰 기반 화자 분할 프레임워크                                                       |    |
| Kyo-won Koo, Jungwoo Heo, Ju-ho Kim, Hyun-seo Shin, .....                                          | 77 |
| <b>SE06</b> 한국인 아동 화자의 수준별 영어 음소 인식을 위한 Wav2Vec2 기반의 도메인 적응 기법                                     |    |
| Ji-Eun Park, Jeong-Sik Park .....                                                                  | 78 |
| <b>SE07</b> 웨어러블 청각 장치의 음성향상을 위한 청각 필터뱅크 잡음 제거 신경망                                                 |    |
| Ju Yeong Park, Seon Man Kim, Jin-Taek Seong .....                                                  | 80 |
| <b>SD01</b> 읽기장애아동과 일반아동의 동시조음 특성 비교                                                               |    |
| Sungsook Park, Cheoljae Seong .....                                                                | 81 |
| <b>SD02</b> 한국어 자음 유형에 따른 상대적기본주파수                                                                 |    |
| Jung, Ji Hang, Jaeock Kim .....                                                                    | 82 |
| <b>SD03</b> NPH 환자와 정상 노인 간 어휘인출능력 및 시선추적에 대한 비교 연구                                                |    |
| Seong-Eun Kwon, Ki-Su Park, Jang-Hyeok Yoon, Na-Yeon Kim, Ji-Ho Lee, Ji-Wan Ha .....               | 83 |
| <b>SD04</b> 경련성발성장애와 다른 음성장애의 식별을 위한 음향지표 및 말자료 개발에 대한 예비연구                                        |    |
| In Hyo Seo, Cheol Jae Seong .....                                                                  | 86 |
| <b>SD05</b> 인지-언어혼련 애플리케이션을 활용한 어휘 및 조음능력 증진                                                       |    |
| Nayeon Kim, Jiho Lee, Ki-su Park, Janghyeok Yoon, Ji-wan Ha .....                                  | 87 |
| <b>SD06</b> 롬바르드 효과가 아동과 성인의 말소리 산출에 미치는 영향: 음향학적 특성과 모음공간면적을 중심으로                                 |    |
| Yelim Jang, Jaehee Hwang, Nuri Lee, Nakyung Lee, Seeun Eum, Youngmee Lee .....                     | 88 |
| <b>SD07</b> 중학교 읽기부진 학생의 시나리오 글 읽기 유창성 특성                                                          |    |
| Jihye Park, Cheoljae Seong .....                                                                   | 89 |
| <b>SD08</b> 구개열 언어치료 - 조음동작형성법의 개발과 적용                                                             |    |
| Eun-Jeong Lee .....                                                                                | 90 |
| <b>SD09</b> 화자의 말속도, 청자의 성별, 연령과 말속도가 청자의 호감도 평가에 미치는 영향                                           |    |
| Eunvin Chun, Yoora Jeong, Heeyeon Kim, Nahyun Kim, Sujin Kim, Youngmee Lee .....                   | 91 |
| <b>SD10</b> 한국어 이중모음 분석을 위한 활음 분절 알고리즘 개발                                                          |    |
| Jiyeon Park, Cheoljae Seong .....                                                                  | 92 |
| <b>SD11</b> 단어와 터치간 시간적 일치성이 초기 언어발달에 미치는 영향: 한국 어머니와 아동의 상호작용을 통해 본 연구                            |    |
| Jongmin Jung, Suhan Kim, Eon-Suk Ko .....                                                          | 93 |

|                                                       |    |
|-------------------------------------------------------|----|
| <b>SD12</b> 장애음성 평가 및 진단을 위한 한국어 자음 spectra 에너지 분포 특성 |    |
| Donghyo Kang, Cheoljae Seong .....                    | 94 |
| <b>SD13</b> 대중가요 보컬 전공 남학생의 성종에 따른 모음 간 음역 차이         |    |
| Il-Soung Ji, Jaeock Kim .....                         | 95 |

## 포스터 발표

|                                                                                            |     |
|--------------------------------------------------------------------------------------------|-----|
| <b>P01</b> POSHA-S를 통한 직업군인의 말더듬 인식 연구                                                     |     |
| Hwajung Cha, Jin Park .....                                                                | 97  |
| <b>P02</b> 갑상선 술 후 증후군의 빠른 음성회복을 위한 음성치료의 효과                                               |     |
| Jisung, Kim .....                                                                          | 98  |
| <b>P03</b> 청각적 자극 및 시각적 자극에 따른 말소리장애 아동의 단어 회상하기 과제의 수행 비교                                 |     |
| Ju-won Kim, Ji-Wan Ha .....                                                                | 99  |
| <b>P04</b> 아동기말실행증의 말소리 감별진단도구 마련을 위한 예비연구                                                 |     |
| Okbun Lee, Sanghee Park .....                                                              | 102 |
| <b>P05</b> 게임 닉네임의 독음 후보군 생성 방법                                                            |     |
| Nahyun Kim, Yeonjeong Kim, Yonghyuk Lee, Namhyun Cho .....                                 | 104 |
| <b>P06</b> 정보 교란 기법과 스피커 인코더를 활용한 한국어 제로샷 음성 합성 모델                                         |     |
| Chae Woon Bang, Chan Jun Chun, Hong Kook Kim .....                                         | 105 |
| <b>P07</b> 자기 지도 학습 모델 특징 결합을 이용한 적대적 신경망 기반 억양 음성인식                                       |     |
| Na Geng, Jeong-Sik Park .....                                                              | 106 |
| <b>P08</b> 음성 향상을 위한 게이트 합성곱 순환 신경망                                                        |     |
| H.L. Koh, S.D. Na, M.N. Kim .....                                                          | 107 |
| <b>P09</b> Pitch 정보를 활용한 Perturbation AUTOVC의 성능 향상                                        |     |
| Hwa-Young Park, Chanjun Chun, Hong Kook Kim .....                                          | 110 |
| <b>P10</b> 한국어 종성 유음의 발음 변이 연구                                                             |     |
| YoonJeong Seo, Drew Crosby .....                                                           | 111 |
| <b>P11</b> 데이터를 활용한 프랑스어 단어 수준 음절 유형 및 자음군 분포 분석                                           |     |
| Soeun Park, Sunhee Kim .....                                                               | 112 |
| <b>P12</b> 한국인 학습자의 프랑스어 연음 오류 분석                                                          |     |
| Minjeong Kim, Sunhee Kim .....                                                             | 114 |
| <b>P13</b> 시대별 애니메이션 캐릭터 음성에 대한 음향 및 선호도 분석                                                |     |
| Sunmi Lee, Heesu Cho, Youngkyoung Choi, Jiyeon Kim, Ryunha Lee, Youngmee Lee .....         | 116 |
| <b>P14</b> 정상발달 아동의 감정발화모방에 대한 음향학적 분석                                                     |     |
| Su Been Kim, Jung Eun Kim, Soo Hyoung Cho, Hyo Sun Lee, Seong Yun Moon, Youngmee Lee ..... | 117 |

## 정보 교란 기법과 스피커 인코더를 활용한 한국어 제로샷 음성 합성 모델\*

방 채 운<sup>1</sup>, 전 찬 준<sup>2</sup>, 김 홍 국<sup>1</sup>  
<sup>1</sup>광주과학기술원, <sup>2</sup>조선대학교

### Korean Zero-Shot Speech Synthesis Model using Information Perturbation Techniques and Speaker Encoder

Chae Woon Bang<sup>1</sup>, Chan Jun Chun<sup>2</sup>, Hong Kook Kim<sup>1</sup>  
<sup>1</sup>Gwangju Institute of Science and Technology (GIST)  
<sup>2</sup>Chosun University  
 {<sup>1</sup>bang4045@gm., <sup>1</sup>hongkook@}gist.ac.kr, <sup>2</sup>cjchun@chosun.ac.kr

음성 합성이란 텍스트로부터 음성을 생성하는 기술을 말한다. 이러한 음성 합성 기술의 목적은 화자와 유사하고 잡음이 없는 음성을 생성하는 것이다. 최근에는 단일 화자를 넘어 다수의 화자에 대해 높은 성능을 보이는 음성 합성 모델이 제안되고 있다. 하지만 여전히 학습되지 않은 화자에 대해서 음성을 합성하는 것은 어려운 일이다. 또한, 학습되지 않은 화자에 대해 음성을 생성할 경우 생성된 음성의 품질과 화자 유사도가 낮다는 것이다. 이는 모델이 충분히 많은 화자 정보를 학습하지 못했기 때문에 발생한다. 따라서 본 논문에서는 효과적인 제로 샷 음성 합성을 위해 정보 교란 기법과 스피커 인코더를 활용한 한국어 제로샷 음성 합성 모델을 제안한다. 제안하는 방법은 확산형 확률 모델 기반의 음성 합성 모델인 Grad-TTS를 기반으로 한다. 먼저, 사전 학습된 스피커 인코더를 사용하여 음성 프롬프트로부터 화자 정보를 추출한다. 이를 통해 모델은 학습되지 않은 화자에 대해서 음성을 합성할 수 있도록 한다. 또한, 정보 교란 기법을 활용하여 학습 데이터에 등록된 화자뿐만 아니라 더 많은 화자 정보를 모델이 학습할 수 있도록 한다. 이는 모델이 다양한 화자 정보를 학습하여 학습되지 않은 화자에 대해 효과적인 음성 합성이 가능하게 한다. 제안된 모델의 성능을 평가하기 위해 Neural MOS와 SECS를 측정하였다. 성능 평가를 위해 남성 화자 5명, 여성 화자 5명을 뽑아 각각 10개의 음성 샘플을 생성하였으며, 평가에 사용된 화자는 모두 학습되지 않은 화자이다. 측정 결과 원본 음성의 평균 Neural MOS 점수 3.0330과 비교하여, 생성된 음성은 평균 2.8480으로 원본과 유사한 수준의 음성 품질을 보이는 것을 확인할 수 있다. 또한, 화자 유사도는 평균 0.5088로 프롬프트 음성과 유사한 음성을 생성하는 것을 확인할 수 있었다.

\* 이 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 연구개발특구진흥재단의 ‘지역의 미래를 여는 과학기술 프로젝트’ 사업(인공지능 기반 메타버스 구현을 위한 융·복합 문화 가상 스튜디오, 2022-DD-UP-0312)과 2024년도 광주과학기술원의 재원으로 “과학기술혁신사업단의 실용화 연구개발사업”의 지원을 받아 수행된 연구임.