2017년 12월 17ZS1200-17-1124P

언어장벽 없는 국가구현을 위한 자동통번역 산업경쟁력 강화 사업

Strengthening competitiveness of automatic translation industry for realizing language barrier-free Korea

ETRI 한국전자통신연구원
Electronics and Telecommunications Research Institute

인 사 말 씀

현재 세계 각국은 “과학기술의 발전만이 21세기에 선진국으로 살아남을 수 있는 길이다.” 라는 명제 하에 국운을 걸고 과학기술에 대한 인적, 물적 투자를 증가시키고 있습니다. 특히 21세기에는 지능형 지식정보화 사회가 될 것이라는 전망에서 볼 때, SW, 빅데이터, 클라우드, HCI 기술을 망라한 정보기술은 기술적, 산업적, 사회적 측면 등 모든 분야에서 그 중요성이 점점 커지고 있습니다.

이러한 지능형 지식정보 기술 중 컴퓨터와 인간사이의 가장 자연스러운 의사소통 방법인 말을 이용한 휴먼 인터페이스 기술은 지난 30여 년간 세계 각국에서 치열한 경쟁 속에 활발히 연구되어 왔습니다. 최근에는 보다 진보된 개념인 서로 다른 언어 간의 의사소통을 가능하게 해주는 자동통역 기술이 새로운 화두로 대두되었으며, 미국, 일본 등 선진국들은 이미 이 분야에 막대한 투자를 시작한 바 있습니다. 늦었지만 다행히 우리나라도 2008년 휴대형 한·영 자동통역 기술 개발을 시작으로 2012년 한/영 자동통역 대국민 서비스 실시, 2013년 한·일, 한·중까지 통역 대상 언어를 확대, 2014년 인천아시안게임, 2015년 광주유니버시아드 자동통역서비스 실시 등 활발한 연구를 진행하고 있습니다. 그러나 국내의 짧은 자동통역기술 개발 역사로 선진기술에 대비 경쟁력을 갖기 위해 음성인식, 자동번역 등 핵심기술 확보, 경쟁력 있는 연구개발 환경 구축, 시장창출을 위한 산업생태계 조성, Open R&D를 통한 인력 양성 등 노력해야 할 분야가 산재해 있는 실정입니다.

이 보고서는 이러한 관점에서 자동통역 실현에 꼭 필요하나 자동통역 기술 개발에 미흡한 부분을 강화하고 향후 자동통역 분야의 성공적인 수행에 밑거름이 되게 하자는 목표로 수행되었습니다.

2017년 12월

한국전자통신연구원 원장 이 상 훈

제 출 문

본 연구보고서는 주요사업인 "언어장벽 없는 국가구현을 위한 자동통번역 산업경쟁력 강화 사업"의 수행결과로서, 본 사업에 참여한 아래의 연구실이 작성하였습니다.

2017년 12월

사업책임자 : 책임연구원 김상훈 (음성지능연구그룹)
연구참여자 : 책임연구원 박 준 (음성지능연구그룹)
책임연구원 박상규 (지능정보연구본부)
책임연구원 김정세 (음성지능연구그룹)
책임연구원 김승희 (음성지능연구그룹)
선임연구원 김동현 (음성지능연구그룹)
선임연구원 윤 승 (음성지능연구그룹)
선임연구원 최무열 (음성지능연구그룹)
책임연구원 이영직 (음성지능연구그룹)
책임연구원 김영길 (언어지능연구그룹)
책임연구원 최승권 (언어지능연구그룹)
책임연구원 김창현 (언어지능연구그룹)
책임연구원 서영애 (언어지능연구그룹)
책임연구원 황금하 (언어지능연구그룹)
책임연구원 최미란 (언어지능연구그룹)
책임연구원 권오욱 (언어지능연구그룹)
연 구 원 이민규 (음성지능연구그룹)
연 구 원 김여정 (음성지능연구그룹)
연 구 원 이담허 (음성지능연구그룹)

요 약 문

Ⅰ. 제 목

언어장벽 없는 국가구현을 위한 자동통번역 산업경쟁력 강화 사업

Ⅱ. 연구목적 및 중요성

본 사업은 대국민 자동통번역 서비스 실시, 자동통번역 산업생태계 구축 및 개방형 R&D 추진 등을 목표로 하고 있다. 대국민 자동통번역 서비스는 한, 중, 영, 일 4개 언어를 대상으로 통역서비스를 제공하여, 2014년 아시안게임 및 2015년 광주유니버시아드 경기대회에 활용되었으며, 여기에 프랑스어, 스페인어, 독일어, 러시아어 4개 언어를 추가하여 2018년 평창 동계올림픽 지원에도 활용될 예정이다. 특히 대국민 자동통역 서비스를 통해 축적한 사용자 로그데이터는 자동통역 원천기술의 개발에 직접 활용될 예정이고, 이를 통해 최근 구글, 애플 등 다국적 기업에 의한 SW종속으로부터 국내 산업을 보호하고, 자동통번역 산업 경쟁력을 강화하고자 한다. 또한 본 사업을 통해 고품질의 자동통역 플랫폼 및 Open API를 국내 업체, 연구소, 대학 등에 지원하여 산업생태계 구축도 추진하고 있다.

Ⅲ. 연구내용 및 범위

- 한/스, 한/불 자동통역 서비스 시스템 구현 (통역률 ERR 10% 개선)
- 핸즈프리 양방향 자동통역 원천기술 개발 (2차)
- 다국어 음성인식 기술 개발(독일어, 러시아어)
- 자동통역 플랫폼 기반 Open API 개선 (4차)
- 중소기업 시제품 제작 지원
- 다국어 음성언어연구기반 조성
- 2016년 평창 올림픽 테스트 이벤트용 5개 언어 자동통역 구현

Ⅳ. 연구결과

- 세계최초 Zero-effort UI/UX 통역용 어어셋 개발 성공 (2017.8월 ISO 국제표준 예정)
- 세계수준 8개 언어 다국어 음성인식 기술 확보 (국내유일 다국어 기술 보유)
- 세계최초 발음사전 없는 음성인식 기술 및 topic LM 디코더 개발
- 실시간 방송음성 정제를 통한 대용량 학습용 DB 자동확보 기술 구현에 성공 (한국어, 독일어, 불어, 스페인어에 적용)
- 한컴인터프리 (ETRI연구소 기업 및 한컴그룹 자회사) '말랑말랑 지니톡' 30만 앱 다운로드 기록 및 MWC2016 출품 (한컴은 단말통역기, 통역박스 등 다양한 제품으로 사업 확대 추진 중)
- 평창동계올림픽 테스트 이벤트에 5개 언어 자동통역 시범서비스 적용 (2016.11~2017.3)

Ⅴ. 기대성과 및 건의

- ‘12 여수 세계엑스포, ‘14 인천 아시안게임, ‘15 광주 유니버시아드, ‘18 평창 동계올림픽에 자동통번역서비스 실시를 통해 언어장벽 없는 국가 실현 및 국격 제고
- ETRI의 자동통번역 플랫폼 및 Open API를 적극 지원하여 자동통번역 기반 새로운 BM 창조기회가 확대되고 이에 따라 대학, 중소기업, 응용SW 개발자, 수요업체(단말기, 포털사, 주력산업 제조업체 등)간의 산업생태계 조성
- ETRI와의 기술협력 및 향후 다국어 처리 기술지원을 통해 향후 3~4년 내 국내 SW기업의 기술력이 글로벌 경쟁력을 갖추게 될 것으로 기대

ABSTRACT

Ⅰ. TITLE

Strengthening competitiveness of automatic translation industry for realizing language barrier-free Korea

Ⅱ. THE OBJECTIVES

This project aims at providing automatic speech translation public service, building up the eco-system of automatic speech/text translation industry, and constructing the open R&D environment. The automatic speech translation service will include Korean, English, Chinese, and Japanese, Spanish, French, German and Russian till 2018, and support the 2018 Pyeongchang Winter Olympics. We are utilizing the user log-data from the public service in improving the automatic translation technology, and thus try to strengthen the industry competitiveness against the global enterprises like Google. We will provide the automatic translation technologies along with its base platform and open API to the domestic industries, research institutes, and universities for industrial eco-system construction.

Ⅲ. THE CONTENTS AND SCOPE OF THE STUDY

○ Providing Korean-Spanish/French automatic speech translation trial public service to public.

○ Developing the fundamental technologies for hands-free, bidirectional spontaneous speech translation (2nd phase).

○ Multilingual speech recognition technology development(German, Russian)

○ Implementation of the automatic speech translation platform and Open API (4th phase).

○ Implementation of the automatic speech translation platform and Open API.

○ Set-up of the multilingual spoken language research infrastructure.

○ Development of speech translation system (supporting 5 languages) for 2016 PyeongChang Olympics test event.

Ⅳ. RESULTS

○ Development of the new ear-set for the zero-effort UI/UX speech translation for the first time (to be fixed as ISO International Standard in August 2017.8.)

○ The state-of-the-art multilingual (8 languages) speech recognition technology

○ Development of the world's first speech recognition system without pronunciation

dictionary and the topic LM-based decoder

- automatic collection technology of the large-scale training DB by refining the real-time broadcast speech signal(Korean, German, French, and Spanish).
- Hancom-interfree, the Hancom-ETRI joint investment company, developmed the MalangMalang GenieTalk system which recorded 300,000 downloads and was displayed at MWC2016.

○ Applying the speech translation system (supporting 5 languages) to the 2016 PyeongChang Olympics test event. (2016.11~2017.3)

V. EXPECTED RESULT & PROPOSITION

○ Series of the automatic speech translation services at the 2012 Yeosu Expo, the 2014 Incheon Asian Games and the 2018 Pyeongchang Winter Olympics will make the language barrier-free nation and enhance the dignity of our homeland, Korea.

○ Service platform and its Open API will encourage the new BM developmen, which will again expedite the advent of the industrial eco-system

○ Multi-lingual translation technology cooperation between ETRI and the associated industries will make contribution to intensify the global competitiveness of the automatic translation industry

CONTENTS

Table List

Figure List

목 차

표 목차

그림 목차

1. 중요성 및 문제점

1.1. 연구개발과제의 중요성

○ 세계화의 가속화로 국가 간 인적, 물적 교류가 활발해지면서 언어 간 장벽을 허무는 자동통번역 기술의 확보가 국가 글로벌 경쟁력과 직결됨

○ 최근 Google, Apple 등 다국적 기업의 모바일용 OS 플랫폼 개발 사례와 같이 다국어 음성언어정보기술 또한 외국에 종속될 위기상황에서 국가적 차원의 지속적 투자 필요

○ 2012년 Google은 다국어 자동통역서비스 추진 발표(2010.2, 타임지). IBM은 5년 내 상용화 기술 중 자동통역을 파급효과가 가장 큰 기술로 선정 (2007년)

○ 2014년 일본정부는 2020년 동경올림픽에서 일본 IT의 부활을 천명하고, 선두기술로 자동통역서비스를 실시하기로 하고, NTT Docomo, NICT 등이 조인트 벤처기업을 설립, 매년 500억 원의 개발비를 투자하기로 함. 2020년 일본내 자동통번역 시장 10조원 예측 (출처: 일본 UFJ총합연구소, 2006)

○ IT 분야는 물론 교육, 관광, 문화, 의료 등 타 산업으로의 파급이 큰 기술

1.2. 연구개발과제 수행의 제약요인

○ 단기과제, 단위기술 개발 위주

- 단기목표 중심 기술개발 후 기술이전으로 사업이 종료함에 따라 지속가능한 음성언어 기술개발이 어려움
- 최근 SW에 공격적으로 투자를 하고 있는 Google, Apple, IBM 등 선진외국의 다국적 기업이 OS, DBMS를 독점하는 사례와 같이 음성언어처리 기술 또한 종속될 위기 상황에서 국가적인 차원의 장기적인 투자가 필요

○ 사용자 데이터 수집을 통한 선순환 기술향상 체계 부재

- 현 R&D 체계로는 실 서비스 환경 사용자 데이터 수집 불가능. 이에 따라 실 서비스를 바탕으로 자동통번역의 성능을 개선할 수 있는 조직 체계로의 전환이 시급. 또한 사용자 데이터 DB 정제 및 플랫폼 서비스를 위한 상시 인력 필요
- Google은 대용량 다국어 사용자 로그 DB 확보 및 엔진반영의 선순환 구조를 기반으로 다국어 음성인식 및 자동통번역 기술의 강자로 부상하고 있음

○ 한국시장에 국한된 기술 개발 위주

- 한국어 위주의 기술 개발로 음성언어 SW의 글로벌 산업경쟁력 확보 어려움. 아시아권 언어, 유럽어, 아랍어 등으로 확대하기 위해 장기간 다국어를 지원하는 상시의 학제간 협력 및 인력 육성 필요
- 국내 글로벌 업체(현대자동차, 삼성전자 등)에서는 다국어 음성인식 기술이 필요한 상황이지만, 다국어를 지원하는 뉘앙스 등 외산에 종속되어 있음

○ 국민 삶의 질 향상이란 국가적 미션 수행 필요

- 국가 글로벌 SW 경쟁력 강화 및 ‘12년 여수 엑스포, ‘14 인천 아시안게임, ‘18 평창 동계올림픽에 자동통번역 대국민 서비스를 통한 국격 제고 필요
- 급속한 개방과 세계화로 외국어 정보와 지식이 경쟁력이 되는 사회에서 외국어 소외 계층에게 정보접근 통로를 제공하여 외국어 격차(Language divide) 해소

1.3. 연구개발과제 수행결과 기대효과

○ ‘12 여수 엑스포, ‘14 인천 아시안게임, ‘18 평창 동계올림픽 등의 국제행사 대상 자동통번역 서비스 지원을 통한 국격 제고

○ 대국민 자동통번역 서비스 실시를 통해 언어장벽 없는 국가 실현 및 국격 제고

그림 1 자동 통번역 對국민 서비스 로드맵 (2012~2018)

○ 자동 통번역 산업생태계 조성 및 신규 서비스 창출

- ETRI의 통번역 SW 플랫폼을 중심으로 응용 SW 개발자, 수요업체(단말기, 포털사, 주력산업 제조업체 등)간의 산업생태계 조성
- 산/학/연 공동연구 활성화를 통해 ETRI 기술 인프라 공동 활용이 가능하므로 인프라가 부족한 기업의 자생력 강화 및 산업 활성화 가능
- 자동통번역 산업뿐만 아니라, 휴대폰, 자동차 등 제조업 및 교육, 관광, 문화, 의료 등 서비스 분야와 융합한 新서비스 창출·발전

○ 국내 SW 경쟁력 제고

- 공공재적 성격의 융합기술 경쟁력을 확보하고 있는 ETRI와의 기술협력 및 지원을 통해 국내

SW기업의 기술경쟁력 제고

- 출연(연)이 상대적으로 취약한 기술사업화 부문에서 출연(연) R&D성과 확산→ 국내기업 R&D지원 강화→ 자동통번역 산업생태계 구축 등 상생협력 선순환 고리 정착

○ 사교육비 경감 및 창의인재 양성

- 외국어 사교육비 부담(10조원) 대폭 경감
- 다학제간 공동 연구 등을 통한 창의 idea 확산 및 창의인재 양성

2. 현황 및 접근방법

2.1. 국내·외 현황

국외 기술현황

○ 미국방위고등연구계획국 (DARPA: Defense Advanced Research Projects Agency)는 IBM, SRI (Stanford Research Institute), BBN에 대한 지원을 통하여 GALE(Global Autonomous Language Exploitation) 프로젝트를 2006년부터 진행해 오고 있음. 영어, 중국어, 아랍어가 주 대상이며, 실제 사용할 수 있는 실용화 기술개발을 목표로 다양한 언어로 쓰인 대량의 자료를 빠른 시간 내에 통역하는 것을 목적으로 함

○ IBM은 5년 내 상용화 가능 기술로서 파급효과가 큰 자동통역을 최우선 순위로 선정. DARPA의 지원하에 MASTOR (Multilingual Automatic Speech-to-Speech Translator)라는 영/중 자동통역기를 개발. 영어-중국어 양방향 통역이 가능하고 약 3만개의 단어 인식. 여행, 긴급 의료 진단, 군의 자기 방어, 보안 상황을 대상으로 하고 있으며 노트북이나 PDA에서도 사용이 가능함. 현재 이라크에 파견된 미군을 대상으로 시험서비스 중동연구소에서 개발한 TALES (Translingual Automatic Language Exploitation System)라는 번역 엔진을 사용

○ 미국 카네기멜론대학교(CMU)에서는 2003년에 음성합성 전문기업인 셉스트럴(Cepstral), 멀티모달 테크놀로지(Multimodal Technologies), 모발 테크놀로지(Mobile Technologies) 등과 함께 양방향 자동통역시스템인 Speechalator를 개발. PDA에서 작동하고, 의료 정보를 영어와 아랍어로 상호 번역하며 번역 정확도는 약 80%임. 번역기는 두 언어의 매개체로 중간언어(Interlingua)를 이용함. CMU대학에서 창업하여 Facebook에 인수된 미국 지비고(Jibbigo)사는 스마트폰 탑재형 영어 기반 일어, 중국어, 스페인어, 아랍어, 양방향 통역기 등 다국어 통역기 개발.

○ Google은 2012년 14개 다국어를 대상으로 원격 서버기반 자동통역서비스 실시 계획을 천명하고, 현재 '구글 트랜스레이터'로 본격적인 자동통역서비스를 제공. 구글은 200억 단어 이상 규모의 병렬 말뭉치를 이용한 통계 기반의 자동번역 기술을 개발하고 있으며 현존하는 자동통역기 중 가장 많은 언어인 90개 이상 언어 간 번역을 지원

그림 2 Speechalator

- 미국 SRI (Stanford Research Institute)는 DARPA의 지원 하에 IraqComm이라는 자동통역기를 개발. 영어 및 아랍어에 대해 수만 단어를 처리하며, 자기 방어, 보안, 기초 의료 서비스 상황을 통역 대상으로 함. 2006년 미군에 의해 채택
- EU는 의회에서 행해지는 모든 연설을 포함하여, 11개 공식 언어로 문서화하는 작업에 연간 5.5억 유로를 지출하고 있으며 회원국 간의 교류에 있어서 언어장벽이 가장 심각한 문제로 대두됨. 이에 따라, EU 제6차 연구지원사업 프레임웍에서는 의회 연설문의 자동전사 및 통역/번역기술 개발을 목표로 하는 TC-STAR (Technology and Corpora for Speech to Speech Translation) 프로젝트 (2004.4 ~ 2007.3)를 지원. EU 의회에서 통역/번역 서비스의 가능성에 대한 시험 차원에서 평가를 수행함. IBM, IRST, LIMSI, UKA, UPC, RWTH 등에서 참여
- 독일 국책연구소 DFKI는 1991년부터 대화체 자동통역 시스템의 개발을 위해 휴대용 독영/독일 자동통역 시스템 VerbMobil의 개발을 주도하여 비즈니스 상담을 보조하는 구문단위 통역기술을 개발
- 일본 NEC는 2002년에 PDA에서 동작하는 일/영 여행자용 자동통역 시스템을 개발하였으며, 한국어를 포함하여 총 11개 국어를 지원하는 다국어 자동번역 기술을 개발. Bestiland라는 번역 포털 서비스를 실시함. 또, 일/중 통역 시스템도 추가 개발하여 2005년부터 나리타공항에서 시범 서비스를 실시, 이 중 일영 단방향 통역시스템은 2007년 12월 휴대폰 단말기에서 실시간으로 동작함
- 일본 NICT 연구소는 1986년부터 자동통역 기술의 실제 실용화를 목표로 연간 200억 원의 연구비를 투입하여, 일어-영어, 일어-중국어간 양방향 대화체 자동통역 시스템을 개발하고 있음. 2007년도 말부터 휴대폰 단말기에 네트워크 기반으로 일/영 통역 서비스를 제공하고 있음
- 일본 NTT, KDD, NHK, 히타치 등은 산학연간 기술 교류의 활성화를 통해 여행안내, 강의/연설, 비즈니스 회의 등 다양한 분야에서의 실용화를 시도
- 2014년 일본정부는 2020년 동경올림픽에서 일본 IT의 부활을 천명. 선두기술로 자동통역서비스를 올림픽에서 본격 실시하기로 하고, NTT Docomo, NICT 등이 조인트 벤처기업을 설립, 매년 500억 원의 개발비를 투자하기로 함
- 보이스트라(VoiceTra)는 일본 정부 주도로 “언어의 벽을 뛰어넘는 음성 커뮤니케이션 기술의 실현” 프로젝트의 일환으로 개발되었으며 2010년 6월에 최초 발표. 음성인식의 경우, 일본어, 영어, 중국어, 베트남어, 인도네시아어, 말레이시아어를 지원하고, 자동번역의 경우 이들 언어를 포함하여 대만어, 독일어, 프랑스어, 덴마크어, 네덜란드어, 이탈리아어, 스페인어, 포르투갈어, 포르투갈어, 러시아어, 아랍어, 힌디어, 태국어, 타갈로그어, 한국어 등의 21가지 언어를 지원하는 등 본격적인 다국어 자동통역기로서의 면모를 보임. 자체적으로 밝힌 바에 따르면 번역 성능은 토익 600점급에 해

그림 3 일본 NICT에서 개발한 VoiceTra

당한다고 함

○ 2014년 5월 마이크로소프트(MS)는 화상통화서비스인 스카이프 통화내용을 통역해주는 '스카이프 트랜스레이터(Skype Translator)' 서비스를 공개. 통화 내용을 실시간 통역해 음성과 자막으로 제공하는 통번역 채팅서비스로, 공개 시연에서는 영어와 독일어 사용자가 대화를 나누어도 전혀 어려움이 없도록 실시간으로 언어소통이 됨. 현재 약 40 가지 언어가 텍스트 형태로 자동번역이 됨

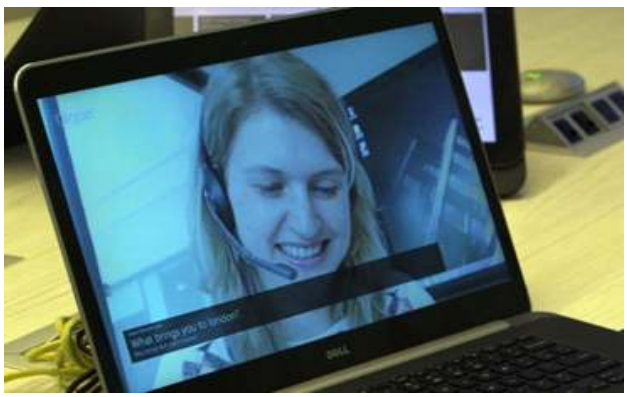


그림 4 MS-스카이프 트랜스레이터

국내 기술현황

○ 국내 핵심기술 연구는 ETRI 등 국책연구기관을 통하여 이루어지고 있으며, 한국어, 영어, 일본어, 중국어에 대해서는 경쟁력을 유지하나 다국어는 열세임

비교항목	국 내	국 외	사용자 요구수준
상용제품	ETRI 지니톡	미국 Google, 일본 NICT	-
통역 성공률	80%~85%	75%~80%	90%
통역 가능한 언어 수	3개 언어 (한, 영, 일, 중)	15개 언어 (한, 영, 일, 중, 스페인, 독일어, 프랑스 등)	20개 언어
통역 대상 어휘 수	30만 어휘	100만 어휘 이내	100만 어휘 이상
통역 대상 범위	일상, 여행 분야	무제한 분야	무제한 분야

표 1 자동통번역 국내외 기술 수준

○ 1999년 한국전자통신연구원은 국제간 자동통역 공동협력연구협의체인 C-STAR (Consortium for Speech Translation Advanced Research)의 핵심회원 기관으로 가입하여 한/영/일/프랑스 4개국간 실시간 음성언어번역 국제시연을 실시하였음. 고객이 여행사 직원과 여행계획을 상담하는 대화를 대상으로 자동통역 기술의 실현 가능성을 입증하였음. 2012년 한국어와 영어간 여행분야에서 언어 소통이 가능한 자동통역 앱 '지니톡(GenieTalk)'을 개발함. 2013년에 한국어와 일본어, 한국어와 중국어로 언어를 확장해 현재 총 4개 언어에 대해 자동통역이 가능함

○ 2001년 삼성과 히다찌 연구소가 한/일 월드컵 국제행사를 통해 1000 여개의 고정 문장에 대하여 휴대전화로 이용할 수 있는 한/일 자동통역 시범서비스 실시한 바 있음

○ 2002년 한/일 번역 업체를 중심으로 PDA 탑재 번역 제품(창신소프트의 「이지토키2002」, 시스메타의 「포켓트랜스위즈」등)이 출시되었으나, 실용적인 면에서 초보적인 수준이며, 영어 및 중국어 등에 대한 다국어 번역 서비스는 제공되고 있지 않음

○ 삼성은 2007년 휴대전화에 기본 예문 검색 기능을 탑재함. 12개의 세부 영역 중 하나를 설정하고 버튼 또는 음성으로 예문을 선택하면 대응되는 외국어 문장을 음성으로 들려줌. 미리 정해진 수백개의 문장 외에는 처리하지 못하여 활용이 제한적임

국내외 표준화 현황(또는 향후 기술 발전 추세)

○ 2007년 3월 APT ASTAP 산하 표준 전문가그룹인 A-STAR 주관 SNLPEG (Speech and Natural Language Processing Expert Group)결성. 자동통역 모듈 연동 규격 표준 작성

○ (국내) ETRI, 삼성전자, LG 등의 순으로 자동통역 기술에 대한 특허를 보유. 한/미/일/유럽을 포함한 주요국에 대비한 출원규모는 약 16.7%를 차지. 주요 출원분야는 통역서비스 향상 방법, 고속 연속음성인식, 통계기반 번역, 합성음의 발화속도 변환, 인식오류 정정방법 등이며 요소기술을 중심으로 분포

○ (국외) AT&T, IBM, Microsoft, Nuance, NEC, NICT, NTT, Hitachi, Sharp 등을 중심으로 자동통역 및 관련 요소기술에 관한 특허를 다수 보유. Sharp가 109건으로 가장 많고, IBM 85건, Microsoft 74건, NEC 64건, NICT 50건, AT&T 48건 등의 순위로 분포. Fluential, VoxTec,

BBN, SRI는 핵심특허를 보유하여 특허권을 행사하는 경우도 있음

동일, 유사내용에 대하여 국내·외 관련자들의 수행내용

○ 미국 IBM은 현재 세계최고 수준의 기술력을 보유하고 있으며, 2012년 상용화를 목표로 2007년부터 본격적으로 실용화 개발에 착수하였음. 여행, 의료, 군사 등 영역에서 영어/중국어, 영어/아랍어간 양방향 자동통역 기술을 개발하고 있으며, 영어/아랍어 자동통역기는 이라크에 파견된 미국을 대상으로 시범서비스를 실시하고 있음

○ 미국 뉘앙스는 세계 최대의 음성솔루션 업체이고 세계 주요 언어에 대한 음성인식 기술을 보유하고 있어 향후 중요 경쟁자가 될 것임. 2008년 말 애플 iPhone에 1,500 문장 수준의 단방향 일영, 일중 통역 소프트웨어를 출시. 아직은 정해진 문장만을 인식하는 수준임

○ 구글은 주요 언어간의 번역 기술을 보유하고 있고 2008년 휴대전화와 연동한 서버 방식의 연속 음성인식 기술을 선보임. 통역분야 향후 주요한 경쟁 상대임

○ 일본 국책연구소인 NICT는 1986년부터 자동통역 기술의 실상용화를 위하여 연간 200억원 규모의 연구비를 투입하여 일본어와 영어, 중국어, 독일어간 양방향 자동통역 기술을 개발하고 있음. 2008년부터 네트워크기반의 일/영, 일/중 통역 시범서비스를 실시하고 있음. 일본어와 한국어의 언어적 특성이 유사하므로 한국어 관련 기술 개발에 착수하는 경우 강력한 경쟁상대가 될 것임

경쟁국가	기관명	연구동향	Google 대비 경쟁력(%)			
			지원언어 수준	음성인식 기술수준	자동번역 기술수준	사용자 편의성
	ETRI	• 2012년 **지니톡 대국민 시범서비스 실시**. 현재 한,중,영,일 4개국 통역서비스 지원 및 스페인어, 불어 등 총 8개국 자동통역 서비스 개발계획	4개국	100	110	110
	NTT	• 일본 총무성, NICT와 NTT Docomo, Systran, FueTrek이 조인트 벤처 설립 **2015년부터 2년간 연 500억원 투입**하여 통번역 기술개발 추진	21개국	85	90	80
	Google	• Google, 2012년 **14개국 다국어를 대상으로 'Google translator' 제공** • Google의 최대 강점인 사용자 기반 빅데이터와 세계최대 컴퓨팅 파워를 결합하여 더욱더 진보된 기술이 나올 것으로 예상	14개국	100	100	100
	MS	• MS-Skype는 2014년 5월 화상통화서비스인 스카이프 통화내용을 통역해주는 **'Skype Translator' 공개**	2개국	95	95	90
	IBM	• IBM Watson연구소가 개발중인 MASTOR 자동통역시스템은 영어-아랍어간 통역. **이라크파병 군인에게 보급**하여 현지인과 통역에 사용	4개국	90	90	90
	Lexifone	• 이스라엘start-up 업체 Lexifone, 2012년 **양방향 실시간 전화 동시통역 상용서비스 실시**. 현재 세계 주요언어 8개국에 대한 대화체 통역 서비스 제공	10개국	80	80	100

그림 5 국내외 자동통역 연구동향 및 수준

동일, 유사내용과 관련하여 제안자가 이미 수행한 사업 또는 연구개발과제

○ 지식경제부 산업원천기술개발 사업의 일환으로 “휴대형 한/영 동시통역 기술 개발(2008~2011)” 사업 수행. 본 사업을 통해 스마트폰 기반 한/영 자동통역 핵심기술 확보

○ 지식경제부 산업원천기술개발 사업의 일환으로 “글로벌 소통을 위한 한/영, 한/일 동시통역 응용

SW 개발(2011~2013)” 과제에 참여기관으로 수행

○ 미래창조부 산업융합원천기술 개발 사업의 일환으로 “지식학습기반의 다국어 확장이 용이한 관광/국제행사 통역률 90%급 자동통번역 소프트웨어 원천 기술 개발 (2012~2016)” 에 참여

2.2. 핵심요소 및 접근방법

○ 對국민 자동 통번역서비스 제공

- 2012년 여수 세계엑스포 영역특화를 통한 3개 언어(한/영, 한/일) 자동통역 대국민 서비스 실시
- 2014년 인천 아시안게임용 자동통역 특화를 통한 4개 언어 (한/영, 한/중, 한/일) 자동통역 대국민 서비스 실시
- 2018년 평창 동계올림픽용 자동통역 특화를 통한 7개 언어(한/영, 한/중, 한/일, 한/스페인, 한/프랑스, 한/독일어, 한/러시아어) 자동통역 대국민 서비스 실시 계획
- 일상영역 대국민 자동통번역 시범서비스 구축 및 운용

○ 자동 통번역 산업생태계 구축

- 자동통번역 플랫폼 지원: Open API를 이용하여 자동통번역 서비스 적용 및 생태계 창출을 위한 플랫폼 지원
- 언어자원 구축 및 배포: 시범서비스를 통해 축적한 로그 DB를 가공·정제하여 기술개발에 활용할 수 있도록 업체 대상 DB 배포
- 자동통번역 Open API(Application Program Interface) 및 DB 표준화: 국내업체·연구소·대학 등의 공통 활용을 위해 API, DB 표준화 추진
- 중소기업 지원 및 공동개발: 자동통번역 산업진흥을 위한 중소기업 지원 및 국내외 업체와 협업을 통한 新비즈니스 창출

그림 6 자동 통번역 산업생태계 구축방안

○ 국내외 개방형 연구 및 인력양성 추진
  - 인문, 공학 등 학제간 융합연구를 통한 창의 연구결과 도출
  - 다국어기술 개발을 위한 국제 공동연구 지원
  - 기업체, 대학 등 센터로 인력 파견을 통해 공동연구 추진 및 창의인재 육성

2.3. 혁신성과 독창성

○ 국내외 업체에 자동통번역 공통플랫폼 및 Open API를 지원하여 실서비스 환경 대용량 로그데이터 축적을 가능하게 해줌으로서 지속적 성능 개선이 가능한 혁신적인 방법을 추진하고자 함

○ 중소기업과 공동으로 자동통번역 BM 개발 및 실용화를 추진함으로서 초기 자동통번역 기술 활용에 따른 위험부담을 줄이며 이에 따라 국내에 자동통번역 산업생태계가 조성되는데 자동통번역센터가 큰 역할을 할 것으로 보임

○ 국내에서는 한국어, 영어, 중국어, 일본어 등 아시아권 주요언어에 치우쳐 기술개발이 진행되는 한계가 있음. 이에 글로벌 시장에 진출하기 위한 유럽어, 동남아어, 중동어 등 다국어 음성언어 기반 기술을 개발함으로서 구글, 애플(뉘앙스) 등과 같이 국제적 수준의 기술확보가 가능함

1. 사업개요

가. 사업 정의

> 대국민 자동통역 서비스(지니톡)를 통해 자동통번역 산업생태계를 구축하고,
> 2018년 평창 동계올림픽 8개국 자동통역 공식 서비스 실시를 통해,
> 세계최고 수준 자동통번역 기술 확보 및 글로벌 시장 진출

나. 기술개념

"외국인과 모국어로 언어소통을 가능하게 하는 기술"

다. 경과보고

2. 사업목표

가. 최종 연구목표

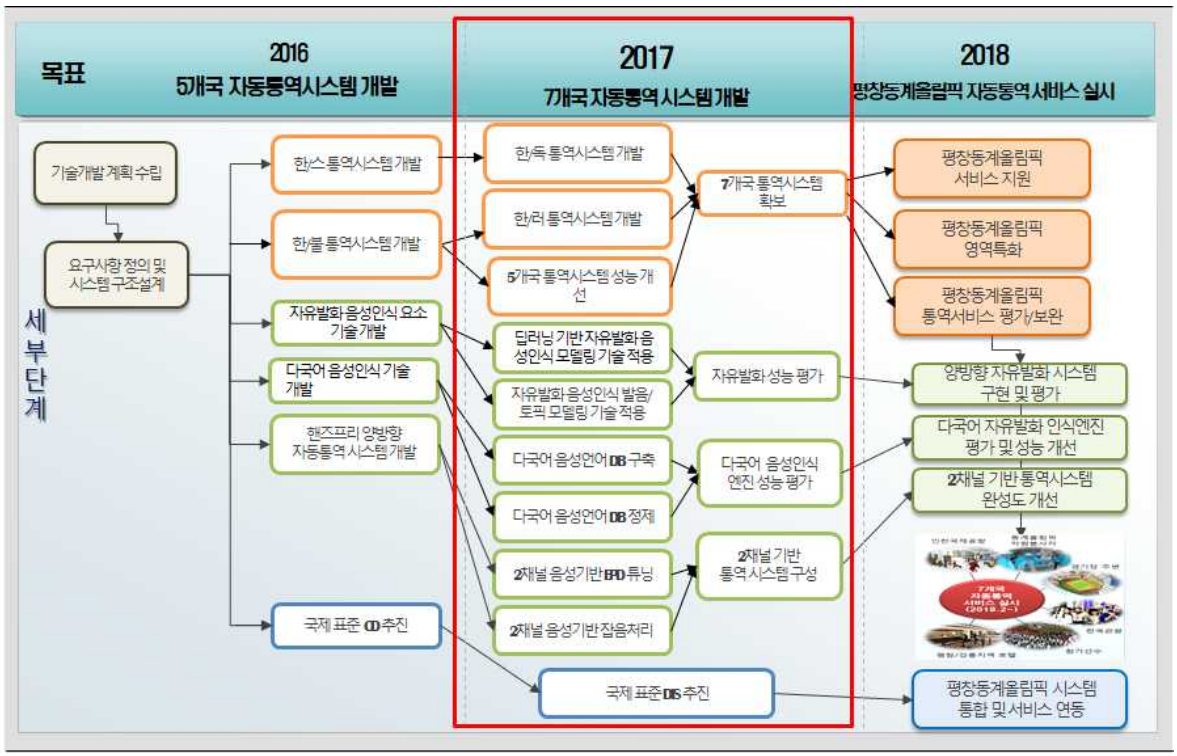
나. 단계별 연구목표

3. 추진계획 및 전략

가. 연차별 추진계획

	2015 (4차년도)	2016 (5차년도)	2017 (6차년도)	2018 (7차년도)
개발 목표	자동통역 글로벌 경쟁력 확보(1차)	자동통역 글로벌 경쟁력 확보(2차)	자동통역 글로벌 경쟁력 확보(3차)	자동통역 글로벌 시장 진출 및 평창올림픽 7개국 자동통역 서비스 실시
핵심 개발 내용	• 한/스, 한/불 자동통역 서비스 프로토타입 시스템 구현 • 자유발화 일상대화형 동시통역 원천기술 개발 (1차) • 자동통역 플랫폼 기반 Open API 구현 (3차) • 중소기업 시제품 제작 지원 (4차) • 대학, 산업체와 협업을 통한 다국어 음성언어 연구기반 조성 (3차) • 2015년 광주 유니버시아드 한, 중, 영, 일 4개국 자동통역 서비스 지원	• 한/스, 한/불 자동통역 서비스 시스템 구현 (통역률 ERR 10% 개선) • 핸즈프리 양방향 자동통역 원천기술 개발 (2차) • 다국어 음성인식 기술 개발 (독일어, 러시아어) • 단말탑재형 및 전화망 자동통역 요소기술 개발 (1차) • 자동통역 플랫폼 기반 Open API 개선 (4차) • 중소기업 시제품 제작 지원 • 다국어 음성언어연구기반 조성 • 2016년 평창 올림픽 테스트 이벤트용 5개국 자동통역 구현	• 한/독, 한/러 자동통역 프로토타입 서비스 시스템 구현 (통역률 ERR 10% 개선) • 핸즈프리 양방향 자동통역 원천기술 개발 (3차) • 다국어 음성인식 성능 개선 (8개국 대상) • 단말탑재형 및 전화망 자동통역 요소기술 개발 (2차) • 자동통역 플랫폼 Open API 지원 • 중소기업 시제품 제작 지원 • 다국어 음성언어연구기반 조성 • 2018년 평창 동계올림픽 한, 중, 영, 일, 불, 스, 독, 러 7개국 자동통역서비스 구현	• 8개국 일상대화형 동시통역 시범서비스 실시 (명칭: 지니톡sun) • 자유발화 일상대화형 동시통역 원천기술 개발 (4차) • 자동통번역 플랫폼 기반 Open API 지원 (6차) • 중소기업 시제품 제작 지원 (7차) • 대학, 산업체와 협업을 통한 다국어 음성언어 연구기반 조성 (6차) • 2018년 평창 동계올림픽 한, 중, 영, 일, 불, 스, 독, 러 7개국 자동통역서비스 실시
목표 성능	한/스 자동통역 성능 개선 (통역률 ERR 10% 달성) 한/불 자동통역 성능 개선 (통역률 ERR 10% 달성)	한/스 자동통역 성능 개선 (통역률 ERR 10% 달성) 한/불 자동통역 성능 개선 (통역률 ERR 10% 달성)	한/독 자동통역 성능 개선 (통역률 ERR 10% 달성) 한/러 자동통역 성능 개선 (통역률 ERR 10% 달성)	사용자 로그데이터 기반 8개국 자동통역 성능 개선 (통역률 ERR 10% 달성)
결과물	-스페인어, 불어 음성인식기술 -광주 유니버시아드 자동통역 서비스 지원	-6개국 다국어 자동통역시스템 구현 완료 -평창올림픽 테스트이벤트 실시	-글로벌 시장 진출을 위한 다국어 자동통역 플랫폼 구축 -8개국 평창올림픽 자동통역시범 실시	'18년 평창 동계올림픽 대국민 자동통역서비스 지원

나. 추진전략



4. 주요 성과목표 및 달성실적 (요약)

가. 성과목표 달성실적

성과지표 (주요성능 Spec)	세계최고 수준	기술개발 목표치('16)	달성 실적	목표치 산출근거	검증방법
한/독일어 통역성능 (통역률)	통역률=60% (Google)	통역률=69% (ETRI)	74.1% (한국어->독일어 통역률=69.0%와 독일어->한국어 통역률=79.3%을 평균함)	자동통역 성능개선률 측정을 통해 정량적 평가가 가능하고, 특히 세계 수준(Google) 대비 벤치마킹을 통해 상대적 우수성 입증이 가능하여 목표로 설정함	통역률(%)= 음성인식률(%) x 자동번역률(%) ※ 음성인식률(%): 20명 이상이 발성한 1,000발화 대상 단어인식률 평가 ※ 자동번역률(%): 번역 전문가 3인 이상 대상 MOS 5점 기준 주관적 평가 실시
한/러시아 통역성능 (통역률)	통역률=60% (Google)	통역률=69% (ETRI)	80.2% (한국어->러시아어 통역률=79.0%와 러시아어->한국어 통역률=81.4%을 평균함)		

나. 정량적 실적

공통지표(필수제시)					자율지표(자율제시)			
지표명			목표	실적	지표명		목표	실적
SCI 논문(건)			1	0	과학적 성과	표준화된 IF 상위 20% SCI 논문(건)	1	0
특허(건)	국내	출원	3	4 (1건 출원중)	기술적 성과	특허활용률 (기술이전건수/특허등록보유건수)	20% (10건/50건)	6% (3건/50건)
		등록	0	2		국제표준특허(건)	2	4
	국제	출원	3	2 (1건 출원중)		국제표준승인표준기고서(건)	1	2
		등록	0	1		3극 특허(건)	1	(1건 제출)
기술이전(건)			2	4	경제적 성과	연구비 대비 기술료 수입(%)	4%	25.3%
기술료(억원)			1	6.4 (부가세 제외)				

다. 정성적 실적

O 국내최초, 세계동등 수준 9개국 다국어 음성인식 기술 확보 (4/18)

※ KT 기가지니용 영, 중, 일 3개국 음성인식 기술이전 (기술료 총 9.8억원 중 본 과제 해당분 57% 반영)

O 세계최초 제로유아이(Zero UI) 자동통역 ISO 국제표준 승인 (8/22)

※ 제로유아이 자동통역 핵심기술에 대한 특허기술상 충무공상(2등) 수상 (11/22)

O 국가연구개발 우수성과 100선에 "5개국 자동통역 기술" 선정 (9/20)

O 한컴인터프리 '말랑말랑 지니톡' 50만 앱 다운로드 기록 및 스페인 MWC2017 출품

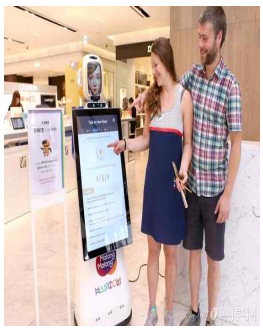
※ 과기정통부 주관 모바일 앱 어워드 공공부문 대상 수상 (11/23)

O KT 인공지능 스피커 '기가지니' 고도화를 위해 다국어 인식/번역 기술이전 체결 (9/26)

※ 기술료 총 10.78억원(부가세 포함)은 단일건으로 최근 3년내 ETRI 최고 기술료 계약임

O 올림픽조직위에 평창동계올림픽 8개국 자동통역SW 공식공급 계약 체결 (11/9)

라. 사업화 실적

사업화 기업체	사업화 내용	사업화 제품 이미지
한컴인터프리	• 2018 평창동계올림픽 8개 언어 자동통번역서비스 공식공급 계약 (2017.11.9.) • 서버형 韓-7개 언어 자동통역서비스 사업화 추진 (현재 약 50만 다운로드 기록) • 韓-3개 언어 OTG 타입 오프라인형 자동통역 사업화 추진 (인천공항, 이마트 등에서 판매 중) • 단말탑재형 하드웨어 제작 (제품명: 지니톡 페어) • 현대백화점 등 통역로봇 사업화	
KT	• 기가지니 외국어 교육서비스 상용화 예정 (2017년말) • 호텔용 기가지니 서비스 (계획) • 공항철도 티켓발매기, 서울시 사이니지 등 B2B 사업화(계획)	

참고 1) 국내최초, 세계동등 수준 9개국 다국어 음성인식 기술 확보 (4/18)

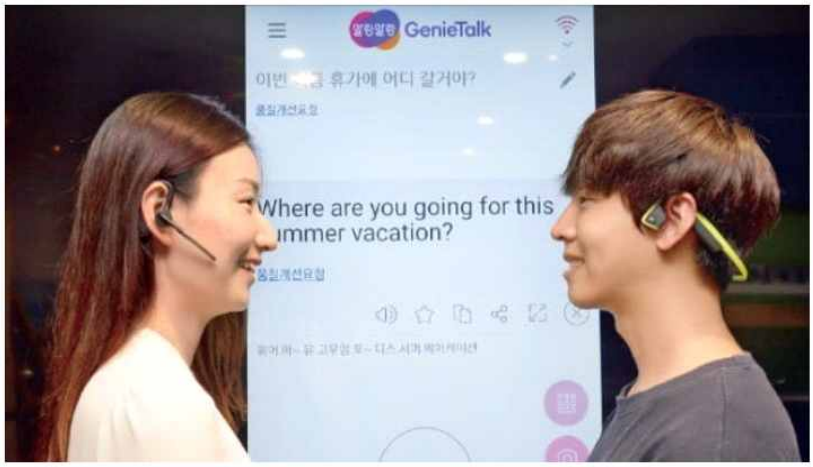
ETRI, 9개 언어 음성인식기 개발 성공

국내 연구진이 한국어, 영어, 일본어, 중국어 등 9개 언어 음성인식기 개발에 성공했다. 말을 하면 해당 언어로 바로 문자 변환이 가능하다. ET...

참고 2) 세계최초 제로유아이(Zero UI) 자동통역 ISO 국제표준 승인 (8/22)

[사이언스] "말하면 자동통역"... 한국 기술이 세계표준 됐다

입력 2017-08-28 16:31 수정 2017-08-28 16:31



한국전자통신연구원(ETRI) 연구진이 머리에 헤드셋을 착용한 채 서로 마주보고 양방향 자동통역 서비스를 시연하고 있다. 한국전자통신연구원 제공

참고 3) 국가연구개발 우수성과 100선에 "5개국 자동통역 기술" 선정 (9/20)

2017 국가연구개발 우수성과

한국전자통신연구원 김 상 훈

공동연구진 : 김영길, 박 준, 김승희, 윤 승, 김창현

성 과 명 : 평창동계올림픽 언어장벽 해소를 위한 韓-5개국 언어 (영어,중국어,일본어,스페인어,불어) 자동통역 기술

귀하의 연구성과가 「2017년 국가연구개발 우수성과 100선」으로 선정되었기에 이 증서를 드립니다.

2017년 9월 20일

과학기술정보통신부장관 유 영 (과학기술정보통신부장관인)

참고 4) 한컴인터프리 ‘말랑말랑 지니톡’ 우수성 (50만 앱 다운로드 기록)

19일(목) 국회 과학기술정보방송통신위원회 국가과학기술연구회 및 25개 출연
국정감사에서 <지니톡>과 <구글번역기>를 이용한 통역 시연 중인 송희경 의원

참고 5) 스페인 바르셀로나 MWC 2017 (2017.2월)

참고 6) KT와 다국어 인식/번역 기술이전 체결 (9/26)

◇ 지난 27일, 우면동 KT연구개발센터에서 KT와 ETRI가 '기술사업화를 위한 기술협력' MOU를 체결했다 [사진=KT]

참고 7) 올림픽조직위에 평창동계올림픽 8개국 자동통역SW 공식공급 계약 체결 (11/9)

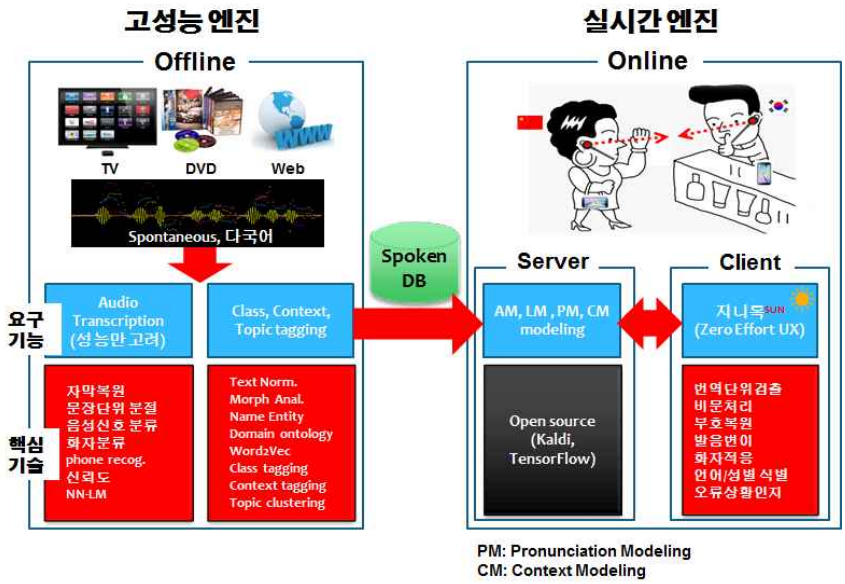
5. 세부 추진실적

가. 정성적 추진실적

1) 다국어 음성인식 기술 개발

○ 다국어 음성데이터 반자동 구축

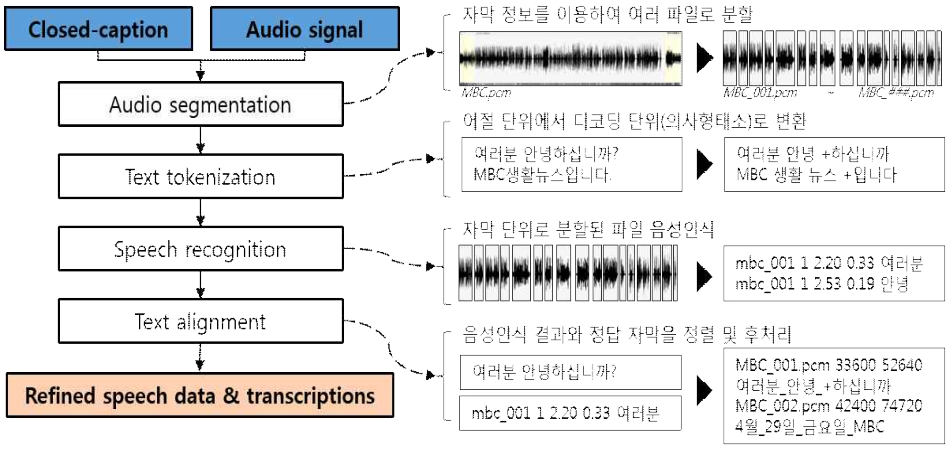
- 딥러닝(deep learning) 기반의 방법이 대두되면서 음성인식에 사용되는 성능 향상을 위한 다른 알고리듬보다 학습 데이터베이스의 보강이 성능에 많은 영향을 미치고 있다. 하지만, 다국어 음성인식을 위한 학습 음성 데이터베이스는 비교적 쉽게 구할 수 있는 한국어 데이터와는 다르게 수집에 많은 어려움이 있다. 이러한 상황에서 데이터베이스를 자동으로 생성 가능하다면 음성인식 시스템과 자동통역 시스템의 성능향상에 상당한 도움이 될 것으로 보인다.



<그림 1> 방송 미디어데이터 기반 자동 DB구축 개념도

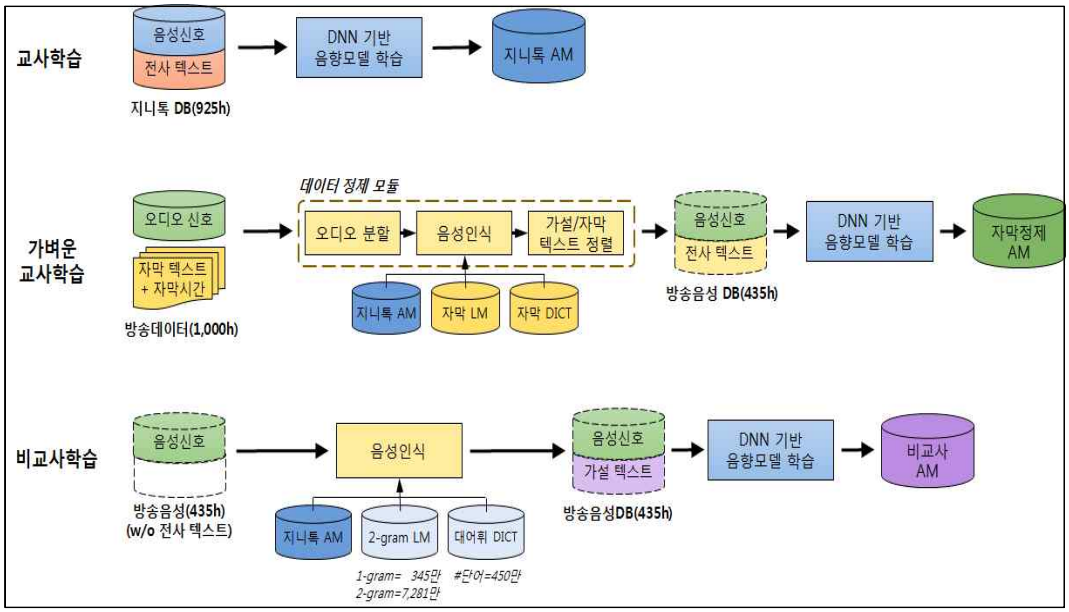
- 먼저 자막 방송 데이터를 수집하고, 수집된 자막과 방송의 음성을 시간적으로 일치 시켜 학습용 음성DB로 활용하였다. 제안된 방식은 다음의 절차를 따르며, 실험결과 자막과 음성이 100% 일치되는 경우만 가정했을 때, 전체 방송음성 분량의 약 25% 정도 음성을 DB화 가능하였다. 자동 DB 구축은 다음과 같은 순서로 이루어진다.

. **오디오 분할**: 길이가 긴 오디오 파일을 자막 시간정보를 이용하여 자막별 오디오 파일로 분할
. **텍스트 토큰화**: 자막 텍스트를 음성인식기 출력과 동일한 단위로 변환
. **음성인식**: 분할된 오디오 파일에서 음성인식기를 이용하여 각 단어의 시간 정보를 가지는 가설 텍스트 추출
. **텍스트 정렬**: 자막 및 텍스트 간의 공통 문자열을 탐색하고 이를 정제된 음성 데이터로 추출



<그림 2> 음성DB구축 자동화 방법

- 본 연구에서는 자막 기반 다국어 방송 데이터 정제 기법과 비교사 방송 데이터 정제 기법 2가지 방법을 제안하였다.
  . 자막 기반 다국어 방송 데이터 정제 기법: 본 기법은 자막 텍스트에 기반하여 음향 모델 학습에 적합한 음성 데이터베이스를 추출한다. 구체적으로는 프로그램 단위의 방송 오디오와 자막을 입력으로 하여, 음성인식 디코딩 단위와 동일하게 자막 텍스트를 정규화하고, 자막의 시간정보를 참조하여 효율적인 음성 세그먼트를 추출한다. 이후 음성인식기를 이용하여 수정된 음성 세그먼트의 가설 텍스트를 추출하고, 주어진 자막 텍스트와 정렬하여 최종 정제 결과로 추출한다.
  . 비교사 방송 데이터 정제 기법: 본 기법은 자막 기반으로 정제된 음성데이터에 대어휘 기반 음성인식을 수행하여 가설 텍스트를 추출한다. 대부분의 비교사 데이터 정제 방법은 입력된 음성데이터의 전사 텍스트를 생성하기위해 강인한 음성인식기로부터 가설 텍스트를 추출하여 음향모델 학습용 전사 텍스트로 사용한다. 반면에 본 연구에서는 다양한 어휘 목록을 이용하여 실제 발음과 유사한 발음열을 가지면서 단어의 범주에서 벗어나지 않는 대어휘 음성인식기를 사용하였다. 이렇게 구축된 음성인식기는 단어 범위 내에서 음성신호와 가장 비슷한 발음열을 출력하며, 함께 출력되는 단어별 신뢰도에 기반하여 음향모델 학습 데이터로써의 사용여부를 결정한다.



<그림 3> 데이터 정제 방법별 블록도

- 방송 데이터의 시간에 따른 정제 성능을 비교하기 위해서 2월에서 6월 사이에 방송된 한국어 방송 데이터로 1차 정제 실험을 수행하고, 약 1,000시간의 다국어 방송 데이터를 이용하여 2차 정제 실험을 수행하였다. 다른 데이터와는 다르게 영어 방송 데이터의 경우에는 지금까지 수집된 방송 분량이 미미하여 630시간의 데이터로 실험하였다. 비교사 정제 실험에서는 994시간의 한국어 방송 데이터를 정제한 결과인 435시간의 음성 데이터를 사용하였다. 제안된 정제 기법들의 성능을 평가하기 위해서 뉴스, 교양, 드라마, 어린이, 예능으로 구성된 5개의 한국어 방송 데이터에 대해서 수작업으로 평가용 데이터를 구축하였다. 다국어 방송 데이터에서는 평가용 데이터를 수작업으로 확인하기 어려워 정제 성능만을 확인하였다.

<표 1> 한국어 데이터 정제율 및 장르별 음성인식 성능 비교표

	비정제 데이터	자막시간	확장된 자막시간	교사된 데이터	제안된 방법
뉴스	27.8	85.5	86.2	79.3	86.4
교양	12.6	48.7	51.7	36.6	52.7
드라마	13.3	56.3	59.8	36.2	59.9
어린이	27.0	62.0	64.1	50.2	64.9
예능	4.7	24.6	27.8	14.2	28.9
전체	18.8	61.3	63.4	52.0	64.0
정제율(%)	100.0	17.0	16.8	-	35.0

- 제안된 시스템의 성능 평가를 위해 한국어 대용량 방송 데이터 정제실험과 다국어 방송 데이터 정제 실험, 비교사 데이터 정제 실험의 3단계로 구분하여 실험하였다. 먼저, 첫번째 실험에서는 3,000 여 시간의 한국어 방송 데이터를 이용하여 제안된 정제 방법의 효율성을 검증하였고, 얻어진 음성 데이터를 사용하여 음향모델을 학습시켜 음성인

식 성능을 확인하였다. 두 번째 실험에서는 한국어, 스페인어, 독일어, 영어로 구성된 다국어 방송 데이터를 정제하여 다국어 정제 시스템의 정제 비율을 확인하였으며, 마지막 비교사 실험에서는 앞서 추출한 435시간의 정제된 한국어 음성 데이터를 대상으로 대어휘 기반의 음성인식기를 통해 가설 텍스트를 추출하고 이들로 학습된 음향모델의 성능을 확인하였다.

<표 2> 다국어 방송 데이터 정제 성능 비교표

	한국어	스페인어	독일어	영어
원시 데이터(hr)	994	1,043	1,033	630
정제 데이터(hr)	435	376	314	100
정제율(%)	43.8	36.0	30.4	15.9

<표 3> 정제 방법에 따른 한국어 음성인식 성능 비교표

	뉴스	교양	드라마	어린이	예능	전체
교사 (925 hr)	84.3	45.1	53.5	59.4	17.9	62.3
자막정제 (435 hr)	88.8	61.3	65.2	69.0	31.8	72.2
비교사 (435 hr)	86.8	50.2	59.0	61.8	20.6	66.1

- 자동통역용 음성인식 시스템의 강인한 음향모델 학습을 위해서 다국어 방송 데이터에서 음향모델 학습용 음성 데이터를 자동 추출하는 기술을 연구하였다. 구체적으로는 자막이 존재하는 방송 데이터에서 음성 및 전사 파일로 구성된 데이터베이스를 자동으로 추출하거나, 자막이 존재하지 않는 방송 음성 데이터에서 학습용 전사 파일을 생성하는 방법을 연구하였다.
- 제안된 시스템을 이용한 대용량 데이터 정제 실험에서는 매 시간 수집되는 대량의 방송 데이터에서 음향모델 학습용 음성 데이터베이스를 정제하는데 매우 유용함을 확인하였고, 다국어 데이터 정제 실험에서는 한국어 43.8%, 스페인어 36.0%, 독일어 30.4%, 영어 15.9%의 정제 성능을 확인하였다. 마지막으로 교사된 데이터와 자막 기반으로 정제된 데이터, 비교사 데이터로 학습된 음향모델을 사용하여 음성인식 성능을 비교함으로써 비교사 정제 방법의 가능성을 확인하였다.
- 본 연구의 결과는 실제 자동통역용 음성인식 시스템의 성능향상에 도움을 줄 것으로 기대되며, 향후 방송 데이터를 보강하거나, 정제 시스템의 음성인식기를 개선한다면 더 높은 성능을 보일 것으로 예상된다. 또한, 방송 도메인의 데이터 이외에도 자동통역 시스템의 실제 사용자 로그 데이터에 적용한다면, 더욱 적합한 음향모델을 획득할 수 있으며, 비록 다양한 도메인의 데이터가 수집되지 않더라도 데이터 증강이나 도메인 적용 기법을 추가로 연구한다면 기존의 자동통역 시스템의 성능 향상에 크게 기여할 것으로 보인다.

O 대용량 다국어 문장DB 구축

- 다국어 음성인식 성능 향상을 위해 8개국 음성인식 언어모델(LM)용 문장을 확대하였다. 다국어 인식기의 LM 성능 향상을 위해 다국어 문장이 필요하다. 기존에 모은 자막 문장이 충분치 않아, 1) 자막 추가, 2) 블로그 문장 추가를 통해 그 양을 늘렸다. 이러한 문장들은 그 수집 과정에서 중복이 가능함. 양을 늘리는 과정에서 문장의 중복을 점검하여 제거하였고, 대상언어는 한.중.영.일.프.스.독.러.아랍 9개국이 된다. 수집 결과 <표 4와> 같이 중국어를 제외하고 기존 대비 2배~5배 확장되었다.

<표 4> LM용 문장코퍼스 구축 현황

(단위: 백만)

언어	1차자막 (제거)		OPUS2016 (제거)		블로그 (중복제거)		총계	
	문장	토큰	문장	토큰	문장	토큰	문장	토큰
한	82	415	-	-	278	1,768	360	2,183
영	92	638	57	386	-	-	149	1,024
중	59	395	-	-	-	-	59	395
일	12	117	-	-	62	887	74	1,004
프	44	243	46	288	1.5	41	92	572
스	25	133	88	516	1.4	42	114	691
독	15	78	13	69	1.7	41	30	188
러	72	354	29	97	15.7	343	117	794
아랍	50	205	25	127	-	-	75	332

- 한일 번역 성능 향상을 위해 한일 대역 문장 구축도 진행되었다. 동일한 드라마에 음성과 함께 부착되어 있는 한글 자막과 일본어 자막을 이용하여 한일 대역문장을 추출하였다. 이 결과는 한/일 NMT 번역기에 적용하여 한일 번역 성능 향상에 활용하였다. 파일 이름에 나타나는 드라마 이름과 회차 정보를 이용하여 파일 단위 정렬을 우선 수행하여 총 18,727 파일을 정렬하였다. 해당 파일에는 시작 시간, 끝 시간, 자막문장 형태로 나열되어 있다. 시간 정보를 이용하여 현재 보유하고 있는 한일 200만 대역 문장을 seed로 활용하여 369만 문장을 확보한 상태이며, 이를 계속 늘려나가는 중이다.

O 독일어, 러시아어, 아랍어 발음변환 규칙 구현

- 독일어는 발음이 규칙적이긴 하나, 예외발음이 많아 23만 엔트리에 대해 발음사전을 구축하였다. 이 때 발음사전에 없는 단어들에 대해서는 발음변환 규칙을 적용하였다. 발음 표기에는 총 47개 폰셋을 사용하였다. 구현된 발음변환기 성능을 측정한 결과 14.8만 발성목록에서 98% 사전 기반으로 매칭이 이루어졌고, 사전에 없는 단어의 발음을 규칙으로 생성할 경우에는 91.5% 성능을 보였다. 사전에 없는 단어의 경우는 규칙으로 생성하다 보니 다중발음의 개수가 64만 단어에서 단어당 평균 5.82개로 너무 많아, 음성인식용 단어는 64만 단어 이상은 사용할 수 없었다. 146만 단어로 늘리기 위해 발음사전에 없는 단어는 다중발음을 생성하지 않고, 하나의 발음만을 생성하게 하여 단어당 평균 2.79개로 줄일 수 있었으며, 이는 음성인식의 성능을 높일 수 있었다.
- 러시아어의 경우, 단어의 악센트를 기준으로 발음이 규칙적으로 변화하는 특징을 보인

다. 이러한 점을 반영하여 발음변환기를 구현하였으며 또한 러시아어를 표기할 때는 단어의 악센트를 거의 표시하지 않는 점을 감안하여 100만 여개의 러시아어 악센트 사전을 별도로 구축하였다. 러시아어의 경우에는 발음표기에 43개 폰셋을 사용하였다. 구현된 발음 변환기의 성능을 측정한 결과 5천 발성목록에서 95.7%가 사전에 매칭이 이루어졌고, 사전에 없는 것을 규칙으로 생성할 경우에는 90.6%의 성능을 나타냈다. 음성인식용 단어를 146만 단어에서 235만 단어로 늘려도 다중발음 과다문제는 없었다.

- 아랍어는 자음만 표기를 하고, 모음을 표기하지 않는 문제가 있다. 모음이 없다보니 규칙을 적용하기 어려웠으며, 규칙 적용시 단어당 평균 116개의 발음이 생성되어 다중발음 과다문제가 대두되었다. 이를 해결하기 위해 NNG2P(Neural Network Grapheme-to-Phoneme Conversion)를 사용하게 되었다. 모음이 표기되어 있으면 발음규칙을 적용할 수 있다. 이에, 모음이 표기된 코퍼스를 활용하여 발음변환기를 구현하여 279만 다중발음사전을 구축하였으며, 구현된 발음변환기의 성능을 측정한 결과 5천 발성목록에서 84.2%가 사전에 매칭이 이루어졌고, 사전에 없는 단어들에 대해서는 NNG2P를 이용하여 10best까지 활용하게 하면 74.1%의 성능을 나타냈다. 실제 활용에서는 5-best만 사용하였다. 발음 표기에는 총 55개 폰셋을 사용하였다.

O 독일어, 러시아어, 아랍어 등 다국어 LM 구현

- 독일어의 경우, LM용 DB를 보강하여 총 단어 개수는 641K 단어에서 1,460K 단어로 늘어났으며, g2p의 다중발음수가 너무 많아 g2p에서 사전에 없는 단어는 결과를 하나만 생성하게 하여 모델에 반영하였다. 평가셋 Nation DB에서 93.0%에서 95.0%로 성능 개선됨을 확인하였다.

<표 5> 독일어 LM용 코퍼스 현황

독일어	추가 전		**추가 후**	
구분	문장수	단어수	**문장수**	**단어수**
수집/용역	389,748	3,510,334	**389,748**	**3,510,334**
자막+블로그	25,329,657	130,278,468	**29,381,426**	**188,055,697**
뉴스(WMT)	54,619,789	821,676,352	**54,619,789**	**821,676,352**
합계	80,339,194	955,465,154	**84,390,963**	**1,013,242,383**

<표 6> 독일어 인식기 성능개선

	구분	Nation DB 평가셋 (3,750문장)
독일어	3-gram(baseline)	93.0%
	4-gram확장	94.5%
	코퍼스추가	94.8%
	단어추가 및 코퍼스추가	95.0%

- 러시아어의 경우, LM용 DB를 보강하여 총 단어 개수는 1,467lK 단어에서 2,353K 단어로 늘어났으며, 평가셋 Nation DB에서 91.5%에서 92.7%로 성능 개선됨을 확인하였다.

<표 7> 러시아어 LM용 코퍼스 현황

	추가 전		**추가 후**	
러시아어	문장수	단어수	**문장수**	**단어수**
수집/용역	225,804	1,584,679	**225,804**	**1,584,679**
자막	75,919,732	373,932,237	**101,621,506**	**450,501,600**
블로그			**15,672,244**	**341,689,986**
뉴스(WMT)	19,912,911	288,608,765	**19,912,911**	**288,608,765**
합계	96,058,447	664,125,681	**137,432,465**	**1,082,385,030**

<표 8> 러시아어 인식기 성능개선

	구분	Nation DB 평가셋 (3750문장)
러시아어	3-gram(baseline)	91.5%
	4-gram확장	92.0%
	코퍼스추가	92.6%
	단어추가 및 코퍼스추가	92.7%

- 아랍어의 경우, 구축현황은 다음과 같다.

<표 9> 아랍어 LM용 코퍼스 현황>

아랍어	문장수	단어수
수집/용역	335,814	2,040,274
자막	64,476,740	270,316,798
합계	64,812,554	272,357,072

- 스페인어 언어모델은 기존 3-gram을 4-gram으로 확장하여 적용하였다. 4-gram의 도입 결과, 개발용 평가 데이터인 플리토 셋과 휴먼미디어 셋에 대하여 복잡도(perplexity)가 각각 80에서 54로, 108에서 100으로 감소하였다. 상대적으로 플리토셋에 대한 복잡도가 크게 감소한 것은 문장 당 평균 단어 개수가 플리토셋은 9.04개로서, 휴먼미디어셋의 7.68개를 상회하여 4-gram 이 적용되는 비중이 더 커지기 때문이라 분석된다. 4-gram을 적용하기 위하여, 언어모델 크기가 대폭으로 커지는 효과를 줄이기 위하여, 인식단어사전의 단어수를 기존 688K(68만8천)개에서 고빈도 위주로 256K개로 추려내었다. 이로 인하여 OOV(사전외 단어)는 플리토셋은 6개, 휴먼미디어셋은 30개가 추가되었으나, 언어모델 크기는 기존 3-gram의 2.5G에서 1G로 크게 감축되었다. 256K 단어사전과 4-gram을 적용한 인식성능은 플리토셋에 대하여 오류감소율(ERR)이 16%, 휴먼미디어셋에서는 2.5%를 나타내었으며, 탐색공간 이미지 크기도 기존 7.64G에서 4G이하로 크게 감축되어, 4-gram 기반 언어모델 적용이 성능과 이미지 크기 측면에서 효과적임을 보여주고 있다. 아래 표x에서 위 결과의 정량적 수치를 보여주고 있다. 언어모델 중 3gram은 기존 성능을 나타내며, LM명이 256K으로 끝나는 항목은 기반코퍼스 및 절지(pruning) 기준치를 변화시켜 이미지 크기를 조정한 4gram 언어모델들이다.

<표 10> 4-gram 언어모델 적용에 따른 인식성능 및 탐색이미지 크기 변화

언어모델	이미지크기(GB)	플리토셋	휴먼미디어셋
m80.p0e-00.256K	4.00	10.25	6.75
m80.p5e-11.256K	2.72	10.45	6.79
CaNaNa.4.256K	1.39	10.45	6.95
3gram	7.64	12.49	6.92

<그림 4> 4-gram 기반 언어모델 적용에 따른 인식성능 및 탐색이미지 크기 변화도

- 프랑스어 코퍼스는 새로 2,500만 문장을 확보하여 언어모델에 추가하였으며, 기존 3-gram기반에서 4-gram기반으로 확장하였다. 코퍼스의 비중이 자막에 몰려 있어, 모든 코퍼스를 모아서 한번에 언어모델을 생성하는 대신, 기존 여행 및 관광 영역 위주 코퍼스 (ET)와 자막코퍼스를 2뭉치로 나누어 세트1(CAP1), 세트2(CAP2)로 구분하고, 3개 코퍼스 각각에 대하여 언어모델을 생성하고 이들 언어모델을 순차적으로 보간(interpolation)하였으며, 추가로 평창올림픽 관련 어휘 및 문장에 대한 언어모델을 생성하여 보간함으로써 최종 언어모델을 구축하였다. 위에서 언급한 3개 코퍼스별 언어모델의 세부 내용은 <표 11>와 같다. 표에서 보는 바와 같이 여행,관광영역의 언어모델(ET)가 자막영역의 언어모델(CAP1,CAP2)에 비하여 양이 상대적으로 적어서, ET와 CAP1을 먼저 보간하여 최적 배합비율을 구하고, 이 결과에 CAP2를 보간함으로써 ET 영역의 비중을 적정하게 설정하였다. 최종 언어모델을 적용한 결과는 개발용 평가데이터로 사용하고 있는 휴먼미디어셋(HMedia)에 대하여 기존 3-gram 언어모델 대비, 복잡도(perplexity)는 75에서 49.5로 감축하였고, 인식성능도 90.6%에서 91.62%로 오류감소율(ERR) 10%를 얻었다.

<표 11> 프랑스어 코퍼스 종류별 언어모델 세부내역

LM	복잡도	#OOV	1g	2g	3g	4g	total
ET	52.01	346	91,574	1,011,720	682,246	781,826	2,567,366
CAP1	92.4	361	351,599	6,744,940	5,545,002	7,071,269	19,712,810
CAP2	93.7	308	548,165	9,887,125	7,773,201	9,799,106	28,007,597

- 중국어 언어모델의 경우, 전년도에 4-gram확장을 수행하였고, 이를 바탕으로 올해 추가 수집한 평창올림픽 영역의 코퍼스를 반영하였다. 추가된 평창올림픽 영역 코퍼스는 한중병렬코퍼스 내 중국어 110만 문장과 평창 지역에 특화된 문틀 데이터로부터 생성한 16,000문장으로 구성되어 있다. 이를 기반으로 영역특화 언어모델을 생성하고 이를 기존 언어모델과 보간을 하여, 개발용 평가데이터 Office1, Office2에 대하여 인식성능을 측정한 결과는 아래 <표 12>와 같다. 영역특화 영역의 비중을 1%, 2%, 5%로 조정하여도 인식성능에 별다른 변화를 보이지 않은 것을 볼 때, 추가된 영역특화 문장이 평창 지역에 특화된 어휘를 포함하고 있으나 문장의 형태는 기존 코퍼스의 범위에서 벗어나지 않음을 나타내고 있다.

<표 12> 보간 비중 변화에 따른 인식성능 추이 (단위: %)

언어모델	보간 비율	Office1	Office2
CN.TLC.4.m70.n80.520K.p5e-11.pc95	95:05	94.54	93.74
CN.TLC.4.m70.n80.520K.p5e-11.pc98	98:02	94.54	93.75
CN.TLC.4.m70.n80.520K.p5e-11.pc99	99:01	94.50	93.71

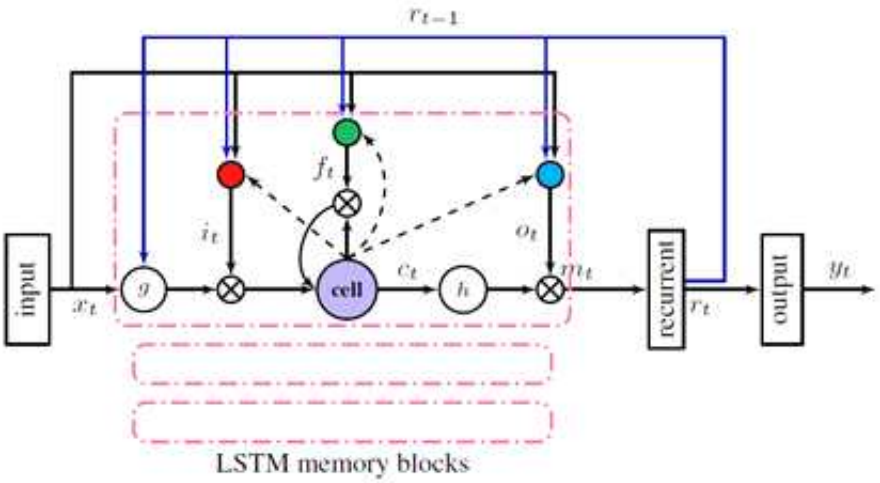
O 독일어, 러시아어, 아랍어 등 다국어 AM 구현

- RNN-LSTM(Long-Short-Term Memory Recurrent Neural Network) 음향모델은 구조적으로 튜닝할 수 있는 여러 요소를 가지고 있다. 기본적으로 학습데이터가 증가할수록 많은 매개변수를 수용하도록 구성할 수 있는데, 이와 관련된 매개변수로 senone 개수를 증가시키는 방법과 hidden layer를 증가시키는 방법, 데이터 전부가 사용되는 반복학습량인 epoch를 늘리는 방법, 그리고 각 hidden layer의 변환 매트릭스의 dimension을 증가시키는 방법이 있다. 학습 데이터의 변동에 상관없이 최적의 모델을 구성하기 위한 매개변수들이 있다. 주요 매개변수는 부분 반복인 iteration 당 사용되는 sample 개수(sample/iter)와 RNN-LSTM이 각 layer에서 연산된 결과를 다음 layer로 전달할 때 time delay 정도를 나타내는 label_delay 값이 해당될 수 있다. 위의 테이블 결과와 같이 각 매개변수를 달리하여 기본 5 layer RNN-LSTM에서 실험을 하였다. 실험에서 sample/iter 값은 기본 40000 샘플에서 20000으로 줄였을 때, 성능하락이 관측되었으나 영향은 작았다. label_delay은 기본적으로 layer를 올릴때마다 같은 time delay를 적용하는 것보다 순차적으로 time delay를 하는 것이 효과적이고, 뒤로 갈수록 delay를 증가 시키는 방법이 좀더 성능 향상에 도움됨을 알 수 있었다.

<표 13> 딥러닝 튜닝 요소

layer	epoch	sample/iter	label_delay	Valid 값	WER
5	20	40000	-1,-2,-3,-4,-5	0.722	11.2
5	20	20000	-1,-2,-3,-4,-5	0.728	11.35
5	20	20000	-1,-2,-3,-5,-7	0.725	11.15
5	20	40000	-3,-3,-3,-3,-3	0.721	11.54

- RNN-LSTM standalone 디코더를 구현하였다. 플랫폼(텐서플로우,토치,칼디)에 의존하지 않고 GPU가 아닌 CPU를 사용한다. Standalone 연산 모듈이 때문에 임베디드나 스마트폰 등의 다른 기기로의 포팅 작업이 수월하다. 기존에 사용하던 Deep Neural Network와 달리 시간 축에 대한 데이터까지 구조적으로 신경망에 훈련되는 것이 특징이다.



<그림 5> RNN-LSTM 구조

훈련 파라미터 수가 줄었음에도 인식 성능이 향상 되는 것을 표를 통해 확인할 수 있다. 연산 속도 또한 기본적으로 파라미터 수에 비례한다.

<표 14> DNN, LSTM 한국어 음향모델별 파라미터 수

	파라미터 수
DNN	15M
LSTM	13.5M

<표 15> DNN, LSTM 환경별 2,000 발화에 대한 한국어 음성인식 성능

	Clean1	Restaurant	Subway	HomeTV	Clean2
DNN	96.99	96.36	96.09	95.84	96.12
LSTM	97.34	96.95	96.5	96.4	96.75

- 이번 실험은 kaldi decoder를 이용하여 DNN과 LSTM 음향모델의 성능을 비교한 실험

이다. 실험을 위해 Librispeech evaluation 셋과, TravelGuide Office1 evalution을 이용한 비교 실험으로 성능에 주요 영향을 주는 Layer 개수와 hidden dimension을 달리하여 최적의 성능을 평가하였다. DNN의 Hidden Layer는 각각 Layer 8 (L8)과 Layer 10 (L10)를 갖고 있고, Layer Dimension은 각각 1024 (D1)과 4096 (D4)로 구성된 모델인 DNN_L8D1 과 DNN_L10D4 이다. LSTM의 Hidden Layer는 각각 Layer 5 (L5)와 Layer 6 (L6)을 갖고 있고, Layer Dimension은 각각 1024 (D1)과 2048 (D2)로 구성된 모델인 LSTM_L5D1, LSTM_L5D2, 그리고 LSTM_L6D2 이다. DNN은 layer와 dimension을 확대하여 모델크기가 10배로 커졌을 때, 기존 DNN_L8D1에 비해 성능향상이 많이 되었다. LSTM_L5D1은 확대된 DNN에 비해서 성능이 높게 측정되어 LSTM이 DNN보다 성능높이는 방법임을 알 수 있었다. LSTM은 Layer를 높이거나 Layer Dimension을 크게 하는 구조로 가져갈 때 조금씩 성능향상이 됨을 알 수 있었다.

<표 16> 영어 성능 비교 (단위: %)

	Libri eval	2620 utt	TgM eval	2000 utt
음향모델 (모델사이즈)	Libris.LM	SER	ETRI.LM	SER
DNN_L8D1 (70M)	13.47	75.15	5.42	24.9
DNN_L10D4 (818M)	11.93	72.06	5.11	23.8
LSTM_L5D1 (81M)	11.20	70.92	5.07	23.75
LSTM_L5D2 (290M)	11.02	69.96	4.88	22.75
LSTM_L6D2 (350M)	11	69.58	4.66	22.35

- ETRI decoder를 이용하여 다양한 환경의 evaluation set에 대한 DNN과 RNN-LSTM 중국어 성능을 실험하였다. DNN과 RNN은 각각 93.1 Mbyte와 89.3Mbyte로 비슷한 크기의 매개변수를 갖는 모델로 성능을 비교하였다. 각각 최적의 beam과 LM weight에 대한 성능 비교이고, DNN은 LM이 2.6일 때 최적인 반면, RNN은 2.2일 때 최적의 성능을 보였다. RNN이 TravelGuide 데이터는 짧은 clean 데이터인 office 1 (off1)을 제외한 모든 데이터 셋에서 DNN 보다 높은 성능을 보였고, 평창올림픽을 대비한 데이터 셋에 대해서는 DNN에 비해 RNN이 성능향상이 큼을 알 수 있었다.

<표 17> 중국어 비교 (단위: %)

모델(B/L)	Office1	Restaurant	Home	Subway	Office2	PyeongChang	평균
DNN_16/2.6	94.71	94.09	93.84	91.60	93.06	90.58	92.99
RNN_16/2.4	94.60	94.31	93.88	93.17	93.45	92.26	93.61
RNN_16/2.2	94.67	94.36	93.81	93.19	93.43	92.29	93.63

- 아랍어의 경우, 훈련 DB의 양이 비교적 작아서 (447시간), RNN-LSTM 모델을 쓸 경우 layer 5 를 쓰면 layer 3 일 경우보다 성능이 떨어지는 것으로 보인다. 그러나 이 경우도 기존 DNN 모델보다는 성능이 개선됨을 확인하였다.

<표 18> 아랍어 성능 (단위: %)

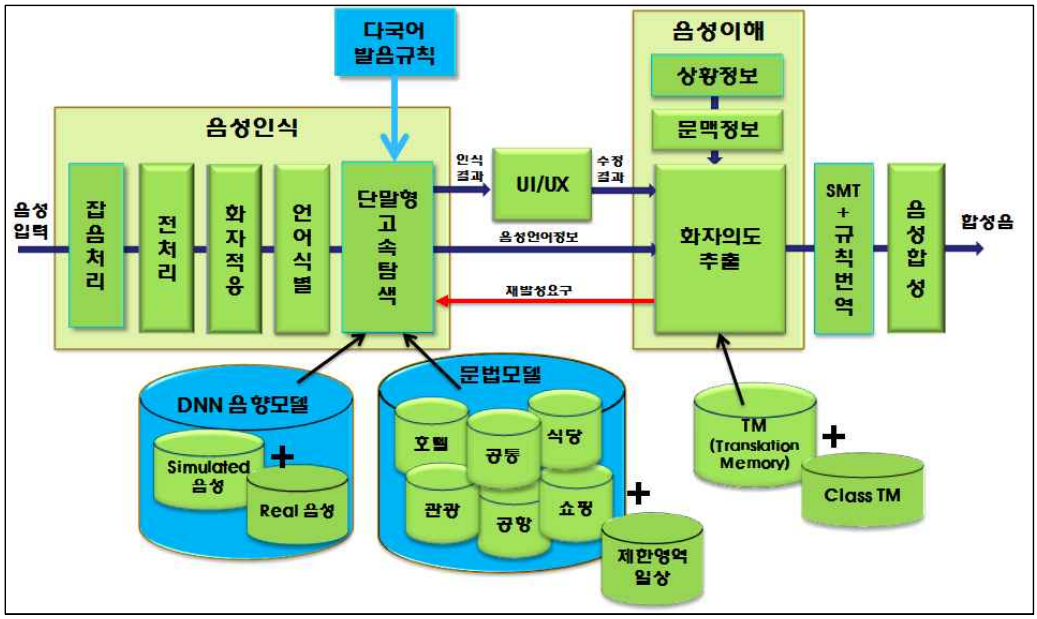
LM weight	모음을 고려한 단어인식률	모음을 제외한 단어인식률
2.3	88.3%	91.6%
2.4	88.3%	91.6%
2.5	88.3%	91.6%
2.6	88.2%	91.4%
2.7	88.1%	91.4%
2.8	88.0%	91.2%

- 일본어의 경우, DNN 대비 RNN-LSTM 적용으로 평균 ERR이 11% 향상되었다.

<표 19> 일본어 성능 비교 (단위: %)

모델(B/L)	Office1	Restaurant	Home	Subway	Office2	평균
DNN	94.5	94.0	93.5	92.7	92.3	92.99
RNN	95.0	94.8	94.1	93.5	93.2	93.61

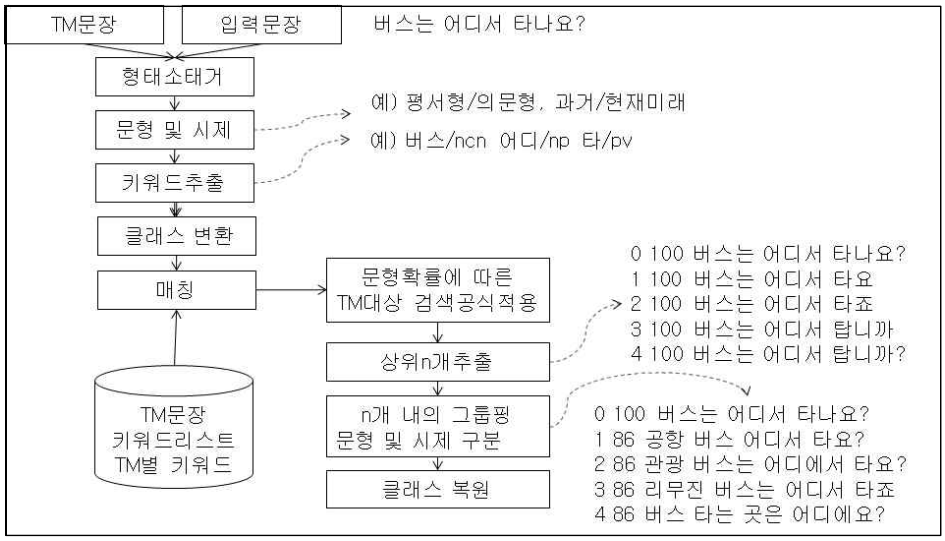
2) 한/독, 한/러, 한/아랍어 자동통역서비스 시스템 구현



<그림 6> 지니톡 자동통역시스템 구조

O 한/독일어 유사문장 검색 기능 구현

- 자동통역 성능 개선을 위해 유사문장 검색 기능을 도입하였다. 한국어와 독일어 각각의 25만 문장 쌍으로 구축된 TM(Translation Memory)을 기반으로 입력된 문장과 유사한 문장을 키워드 및 클래스 정보를 활용하여 대용량의 TM에서 검색하여 보여주는 기능이다. 러시아아의 경우, 언어적 특성 때문에 유사문장 검색 기능을 타 언어 대비 용이하지 않아 적용방안을 연구 중에 있다.



<그림 7> 유사문장 검색기 구조

O 자동통역 사용자편의성을 위한 발음읽기 기능 구현

- 발음읽기는 한국사람이 외국단어에 대해 한글로 표기된 발음으로 읽기가 가능하며, 거꾸로 외국사람이 한글로 번역된 결과를 보면서 읽는 것은 한영 발음읽기가 이미 적용되어 있다. 올해는 스/불/독/러에 대한 4개국 언어의 발음읽기를 추가 적용하였다. 한국사람이 한국어로 말을 하면, 각 나라의 언어별로 번역이 되며, 번역된 결과에 대해 해당 언어의 G2P를 적용하여, 단어별로 발음을 생성하며, 생성된 발음에 대해 IPA기호와 함께 한글발음으로 변환하여 준다.

. 스페인어 예)
입력: Siento no haber hecho tortitas esta mañana,
출력: 시엔토 노 아베르 에초 토르티타스 에스타 마냐나

. 프랑스어 예)
입력: oui je pense que c'est très beau
출력: 위 즈 팡스 크 세스 트래 브오

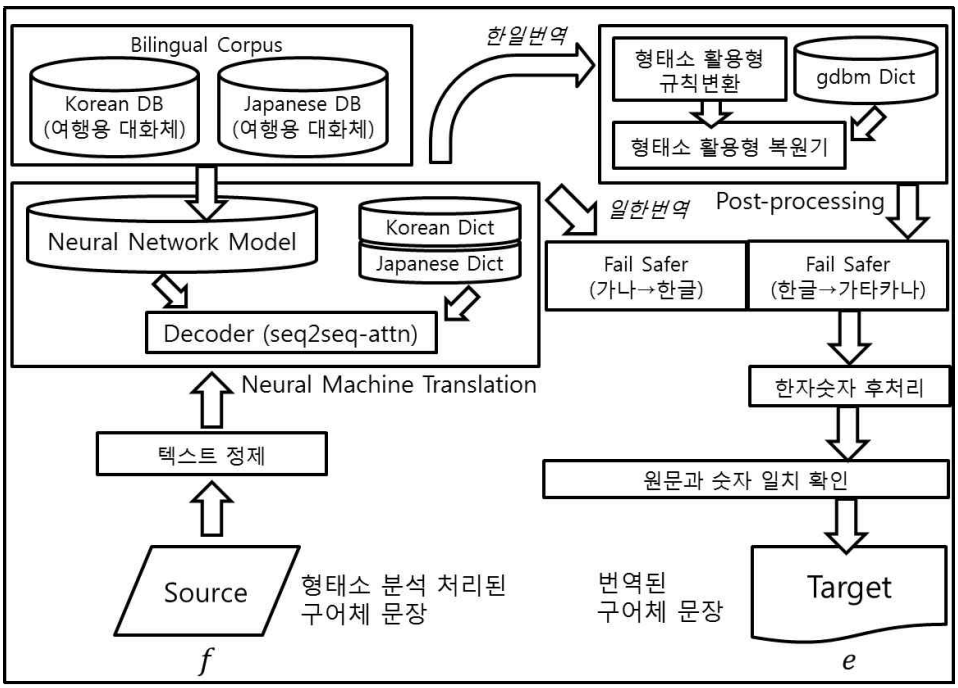
. 독일어 예)
입력: Guten tag, Herr Park!
출력: 구텐 탁 헤르 파르크

. 러시아어 예)
입력: что же помешало ожидавшемуся вами рыночному взрыву
출력: 츠토 줴 퍼미샤러 어지다프쉬무우샤 바미 릐너츠너무우 브즈릐부

O '지니톡' 자동통역 시스템에 요소기술 통합

- NMT(Neural Machine Translation) 기반 번역기 통합: 기존 Rule기반이나 SMT(Statistical Machine Translation) 방식의 번역기를 NMT 방식으로 모두 업그레이드 하였으며, 적용결과 번역률이 일본어를 제외하고 평균적으로 약 10~20%씩 대폭 향상되었다. 여기서는 한/일 번역기의 경우 적용과정을 설명한다. 한/일 여행용 대화체 bilingual corpus로

구축된 100만개 이상의 문장코퍼스를 이용하였다. 한-일 또는 일-한 번역을 수행할 때, Source f에서 텍스트 정제 (전처리)를 거친 뒤에 입력으로 받아들여, Decoder에서 Model을 불러온 다음 f에 대한 번역을 수행한다.



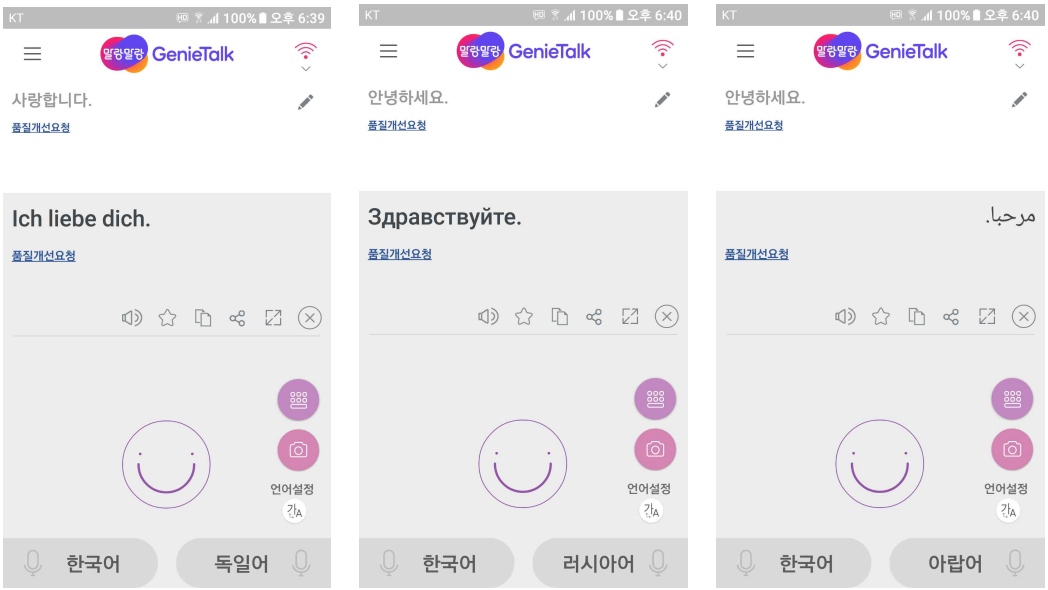
<그림 8> NMT 개발과정 구성도

만약 f가 한-일 번역 과정이면 번역된 결과는 형태소 활용형 규칙 변환 및 복원을 거친 다음 (후처리 과정), 원문과 번역문의 segment token 개수가 일치한가를 확인하여 번역문인 Target e를 출력한다. 일-한 번역도 마찬가지로 Source f에서 전처리 후 Decoder를 통하여 번역한 뒤, segment token 개수를 확인한 다음 번역문 Target e를 출력한다. 일본어의 경우, 한국어와 어순이 동일하기 때문에 기존 SMT 방식으로도 타 언어에 비해 높은 번역률을 보였으며, 이번 NMT 방식의 적용으로 좀더 향상된 성능을 확보하였다.

<표 20> 한/일 번역기 성능 비교

번역방식	번역률 (여행분야)
SMT (기존)	89.6%
NMT (신규)	93.3%

- 한-독, 한-러, 한/아랍어 자동통역 시스템은 한컴인터프리의 말랑말랑 지니톡에 <그림 3>과 같이 통합을 완료하였으며 현재 총 8개 언어에 대해 서비스를 운용중에 있다.



<그림 9> 한/독일어, 한/러시아어, 한/아랍어 자동통역 시범서비스 화면

O 자유발화를 위한 단말탑재형 음성인식 처리속도 개선

- 인식기 가장 계산량이 많은 딥러닝 계산시, 각 입력 프레임의 특징벡터를 서로 다른 CPU 코어에 할당하여 계산하였다.

<표 21> 멀티코어를 이용한 병렬화로 속도 개선

쓰레드 개수	1	2	3	4	5	6
상대 소요시간 (단말, 4 cores)	1	0.69	0.42	0.38	0.71	0.74
상대 소요시간 (서버, 4 cores)	1	0.53	0.38	-	-	-

- 4 core CPU인 경우, DNN 계산에 3 core를 할당(3 threads)하고, network search에 1 core를 할당하도록 구성하였다. 산술연산이 최적화된 솔루션에서 전체 계산 시간의 90% 이상이 메모리 읽기 단계에서 소비되고 있다. 이에 기존 32bit의 float 대신 16bit float 타입을 NEON SIMD 어셈블리로 구현하여 메모리 읽기 시간을 줄임으로써 속도를 개선하였고, 동일한 개수의 파라메터를 ½의 메모리 공간에 저장함으로써 전체 메모리 읽기의 양을 줄였다. 정밀도 하락에 따른 인식률 변화는 95.89% -> 95.84%로 매우 미미한 정도를 보인다.

<표 22> Half-precision float를 적용하여 단말탑재형 DNN 계산 모듈 속도 개선

Optimization Level	RTF	Arithmetic Error (%)
prefetch using PRFM (single precision)	0.40	7.00e-5
Half-precision Weight	0.25	4.16e-2

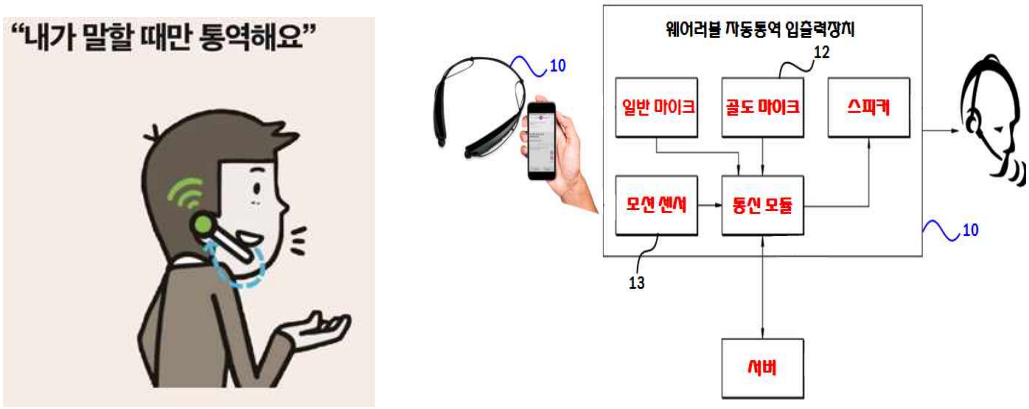
3) 자유발화 일상대화형 동시통역 원천기술 개발 (3차)

O 동시통역용 웨어러블 디바이스 구조 연구 및 국제표준화 추진 (지니톡[sun])

- 음성인식을 위해 일반적인 PC마이크, 스마트폰 마이크, 혹은 블루투스 헤드셋 등을 이용하여 버튼을 누르거나 화면을 터치하여 말을 시작해야 하는 불편함이 있을 뿐 아니라 주변 소음에 대한 오작동 방지장치가 마련되어 있지 않아 음성인식 기능의 한계로 여겨져 왔다. 이러한 문제점을 극복하고자 자연스러운 자동통역을 위하여 음성인식 과정에서, 기존의 마이크를 통한 입력 신호 데이터 뿐만 아니라, 골도 마이크(또는 귓속 내장마이크), 자이로센서 등을 이용하여 입력된 추가 신호를 통해 정확한 음성구간만을 추출하고, 간편하게 귀에 꽂아 자동통역 상황에서 키보드나 터치 등 별다른 조작없이 자연스러운 자동통역을 가능하게 하는 핸즈프리 자동통역 음성입출력 기술을 개발하였다.

〈그림 10〉 기존 스마트폰 통역대비 제안된 핸즈프리 통역의 사용자편의성 향상

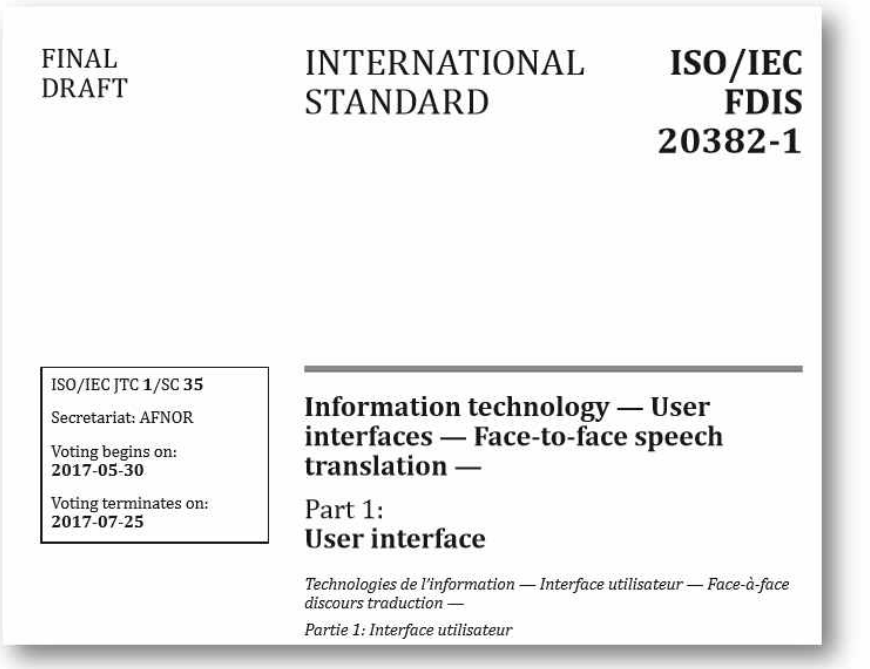
- 핸즈프리 웨어러블 자동통역 입출력장치를 이용하게 되면 통역이 필요한 상황에서 자연스러운 대화를 통한 의소소통이 가능해진다. 자동통역만을 위한 장치이기에 자동통역에 최적화된 입력도구로 필요한 상황에서 명확한 음성구간이 검출되기 때문에 연속어 음성인식에 대한 실시간 자동통역이 이루어 질 수 있는 장점이 있다. 국내 경쟁기술의 경우, 스마트폰 화면을 터치하고 음성을 입력하는 수준에 머물러 있다. 국외의 경우, 최근 출시된 구글 픽셀폰은 이어폰에 달린 터치 버튼을 누르고 말을 하는 방식을 통해 통역이 가능한데, 스마트폰 터치 방식을 이어폰 터치 방식으로 변경하여 사용자의 편의성을 조금 더 향상시킨 수준이다. 본 기술은 말을 할 때마다 버튼을 터치하거나 누르고 얘기하는 불편함이 없애고자 귓속음성이나 골도음성을 감지하여 터치 기능을 대신하기 때문에 외국인과 자동통역시 자연스러운 양방향 대화가 가능해지는 유일한 방법으로 판단된다.



<그림 11> 2채널 기반 음성끝점 검출 방법

- 지니톡[sun]형상을 구현하기 위해서는 크게 세 가지가 필요하다. 첫 번째, 귓속+외부 마이크 신호를 스마트 폰에 전달하기 위한 헤드셋 연구 개발. 두 번째로는 두 스마트폰 사이 일정한 조건(예:거리)을 만족하는 경우 자동으로 데이터를 주고받을 수 있는 채널 활성화하는 어플리케이션 개발, 마지막으로는 귓속+외부 마이크 신호를 입력받아 음성구간을 강인하게 검출하는 음성 검출 모듈 연구개발이다. 첫 번째 귓속+외부 마이크 신호를 스마트폰에 전달하기 위해 기존의 Csr 블루투스 칩셋에 Micom을 추가하여 2채널 신호를 스마트 폰으로 전달 할 수 있는 헤드셋의 소프트웨어와 하드웨어를 개발하였고, 또한 그 신호를 수신할 수 있는 스마트폰 쪽 어플리케이션 개발도 동시에 이루어졌다. 두번째로 스마트폰 간의 거리를 측정하여 자동으로 데이터를 주고 받는 채널을 활성화할수 있도록 BLE를 이용하여 실험 하였으나, 현재의 BLE는 신호 강도의 한계로 장애물이나 기타 장애 요인에 민감하게 동작한다. Bluetooth 5.0 & Anroid 8.0 OREO 패치를 통해 BLE 에 대한 이슈가 해결 될 것으로 기대하며 현재 추적 관찰 중이다. 마지막으로 귓속+외부 마이크 신호를 입력 받아 음성인식에 잘 활용 할 수 있는 방법들에 대한 연구개발을 진행하였다. 1) 음성 외 잡소리 검출 제거 (만지는 소리, 침/껌소리): 신호특성 분석 후 대처 2) 합성음이 나오는 도중 발화자가 발화시 음성구간 검출 -> 신호 특성 분석 후 대처, 3) 잡음환경에 인식이 잘되게: 귓속 음성 활용 정책 구현 등과 같은 연구 개발을 진행중이다.
- 스마트폰을 사용한 양방향 자동통역 과정에서 사용자의 불편함을 해소하기 위하여 제로유아이 기반 동시통역용 웨어러블 디바이스 표준화를 추진하였다. 핸즈프리용 이어셋(지니톡[sun]을 이용한 자동통역은 2017년 10월 ISO 국제표준 승인이 되었고, 관련 국제특허를 표준특허로 4건 반영 예정이다.

. 2016.2: CD (Committee Draft) voting 결정 (로마)
. 2016.8: DIS (Draft International Standard) voting 결정 (서울)
. 2017.2: FIS (Final International Standard) 결정 (베를린)
. 2017.8: IS (International Standard) 최종승인 (제네바)
. 2017.10: IS (International Standard) 국제표준 제정



FINAL
DRAFT

INTERNATIONAL
STANDARD

**ISO/IEC
FDIS
20382-1**

ISO/IEC JTC **1**/SC **35**
Secretariat: AFNOR
Voting begins on:
2017-05-30
Voting terminates on:
2017-07-25

Information technology — User interfaces — Face-to-face speech translation —

Part 1:
User interface

Technologies de l'information — Interface utilisateur — Face-à-face discours traduction —

Partie 1: Interface utilisateur

<그림 12> ISO 국제표준 승인

- 잡음환경에서도 강인한 음성끝점검출 및 근접시 블루투스 신호로 연결/종료, 기존 뮤직이나 전화 기능과의 사용 시나리오를 고려한 통합 성능을 개선하였다. 현재 연구소 기업인 한컴인터프리는 2018년 평창동계올림픽에 지니톡^{sun}을 적용할 계획으로 총 1,000대를 제작하기로 했으며 내년 미국 라스베가스에서 열리는 CES2017에 출품하여 핸즈프리 통역용 이어셋 기술을 국제적으로 홍보할 계획이다.

<그림 13> 제로유아이 통역용 이어셋 시제품

4) 자동통역 플랫폼 기반 Open API 개선 (5차)

한/독일어, 한/러시아어, 한/아랍어 통합을 위하여 기존에 설계되어 있던 한-영, 중, 일 스페인어, 불어 통역 플랫폼에 대상 언어인 독일어, 러시아어, 아랍어를 추가하였다. 이는 한컴인터프리의 말랑말랑 지니톡 플랫폼을 대상으로 이루어졌으며 통합 후 시범 서비스를 통해 프랑스어, 스페인어 자동 통번역 기능이 원활히 동작함을 확인할 수 있었다. 동일한 방법으로 한/스, 한/불 통합을 지니톡 플랫폼에도 실시하였으며 현재 프랑스어, 스페인어 음성 인식 기능을 Open API서버를 통해 업체에 지원이 가능한 상황이다. 금년에는 윌러스, KT 등 Open API를 제공하여 지니톡 기술의 사업화 가능성을 점검한 바 있다.

5) 중소기업 시제품 제작 지원 (6차)

<표 23> 중소기업 대상 자동통역 BM발굴 및 시제품 제작 지원 내용

지원업체	지원 내용
해보라	• 해보라는 이어셋 제조업체로 음성인식 및 통역서비스에 자사 이어셋 적용을 통해 시장을 확장하고자 함 • ETRI는 해보라 이어셋을 음성인식 및 통역성능 평가를 지원하고, 통역용 이어셋 음질 개선을 위한 시제품 제작 지원을 추진함
한컴인터프리	• 한컴인터프리는 음성인식 자동통역 사업을 주력으로 하고 있음 • 단말탑재형 자동통역, 영역특화 자동통역 BM 발굴을 공동으로 수행함 • 지니톡 페어, 지니톡 오프라인 시제품 제작

6) 대학, 산업체와 협업을 통한 다국어 음성언어 연구기반 조성 (5차)

<표 24> ETRI-대학, 업체 협력 현황

대학협력	협력 내용
충남대학교	• 프랑스어, 중국어 전공자와 발음, 문법 등 프랑스어, 중국어 음성인식기 개발에 공동 협력 추진
충북대학교	• 다국어 음성데이터와 자막을 align하여 딥러닝 학습용 음성DB를 자동으로 구축하는 기술을 공동으로 개발
부산외대	• 아랍어 전공자와 아랍어 발음, 문법 등 아랍어 음성인식기 개발에 공동 협력 추진
한국외대	• 자동통역 품질 평가 및 다국어 음성인식 오류 분석 협력
㈜플리토	• SNS업체와 공동으로 독일, 러시아 등 다국어 번역DB 및 음성DB 구축 등 협력을 통한 비즈니스 모델 구현
㈜시스트란	• 평창동계올림픽용 독, 러 번역DB 지원
휴먼미디어텍	• 다국어 음성DB 구축 협력 및 기가지니 원거리 음성DB 구축 지원

7) 2018년 평창동계올림픽 한-중,영,일,불,스,독,러,아랍어 8개국 자동통역서비스 구현

○ 평창동계올림픽용 8개국 자동통역서비스 영역 특화

- 평창동계올림픽 8개국 자동통역 서비스에 적용하기 위해 문체부와 공동으로 올림픽 용어, 경기 관련 용어, 개최지역 및 출입국 동선에 따른 음식명, 호텔명, 관광지명, 교통수단 및 역/터미널 명칭 등을 DB화 하고 있다. 올림픽 및 개최지역에서 많이 사용되는 단어를 이용하여 이동동선(입국·이동·숙박·경기관람·쇼핑·관광·식사·출국) 및 내외국인 예상접점에서 통역이 필요한 문장 및 이의 기반이 되는 문장을 수집, 정제하고 기본 문장 구조화하고, 올림픽 및 개최지역에서 많이 사용되는 한국어 문장을 8개 언어로 번역, 단어 구조화를 통해 표현력 확대를 추진하고 있다.

<표 25> 영역특화 단어 DB 구축 예

분류	한국어	번역(영어의 경우)	발음 표기(Romanize 예제)
음식명	감자떡	Rice Cake of Potato	Gamja Tteok
축제명	인제 빙어 축제	Injae Icefish Festival	Injae Bingeo Festival
경기장	쇼트트랙 경기장	Short Track Stadium	Short Track Gyeongijang
호텔명	IOC 본부 호텔	IOC Headquarters Hotel	IOC Bonbu Hotel
경기명	알파인 회전	Alpine Slalom	Alpine Hoijeon

<표 26> 영역특화 문장 DB 구축 예

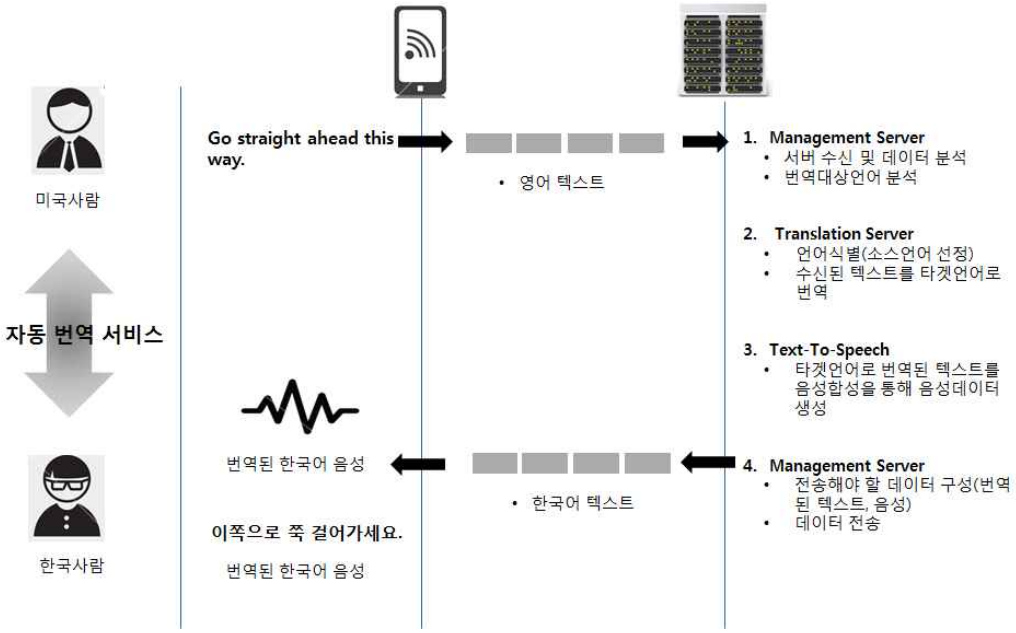
	올림픽 및 개최지역 특화 다국어 문장 콘텐츠 예제
한국어	**[음식명]**을 주문하면 얼마나 걸리죠?
영어	How long does it take for the **[음식명]** to cook?
중국어	想要订**[음식명]**需要多长时间
일본어	**[음식명]**を注文する場合、どれくらいかかりますか？
프랑스어	Combien de temps ça prend si je commande une **[음식명]** ?
스페인어	Cuanto dilata al ordenar una **[음식명]**?
독일어	Wie lange dauert es, wenn ich **[음식명]** bestelle?
러시아어	Сколько времени займет приготовление **[음식명]**?
아랍어	؟**[음식명]** إحضار سيستغرق من الوقت كم

○ 7개국 자동통번역 평창 동계올림픽 시범서비스 운용

- 테스트 이벤트 시범 서비스에 대한 자동통역 성능 평가는 크게 두 가지로 이루어진다. 첫번째로는 테스트 이벤트 기간에 원어민 패널을 동원하여 정성 평가를 실시하는 것이다. 정성 평가로는 터치 포인트별 샘플 예문을 제시한 후 터치 포인트 검증, 서비스 품질(정확도, 사용성, 안정성, 속도, 만족도), 서비스 개선 의견의 3가지 분야별 세부 평가를 실시하였다. (부록1: 테스트이벤트 평가결과)

<그림 14> 자동통번역 평창 동계올림픽 시범서비스 일정

- 평창 동계올림픽 테스트 이벤트에 적용된 자동통역 시범서비스는 다음 <그림 18>과 같은 흐름으로 제공되었다.



<그림 15> 평창동계올림픽 자동통역 서비스 흐름도

- 현재 평창동계올림픽 자원봉사자 및 강릉 등 개최지역의 택시업계 종사자, 숙박업소 종사자 대상 지니톡 사용법 교육을 완료하였다.

〈그림 16〉 평창동계올림픽 자원봉사자 대상 지니톡 교육 실시

나. 정량적 추진실적

1) 논문

번호	명 칭	게재일 (또는 제출일)	게재지명
1	A chatbot for a dialogue-based second language learning system	2017.08.25	EuroCALL 2017
2	Arabic Speech Recognition for Automatic Translation	2017.11.01	Oriental COCOSDA 2017
3	A Preliminary Study on the Pronunciation Rule of Korean Pronoun for English Speakers	2017.11.01	Oriental COCOSDA 2017
4	목적지향 대화 시스템을 위한 챗봇 연구	2017.11.30	정보처리학회 논문지

2) 특허

번호	명 칭	등록일 (또는 출원일)	국명
1	(출원)	2017.07.11	미국
2	Speech recognition system and method using incremental device-based model adaptation (등록)	2017.03.21	미국
3	(출원중)	2017.09.20	미국,일본,중국
4	(출원중)	2017.09.06	대한민국
5	(출원)	2017.08.03	대한민국
6	(출원)	2017.08.02	대한민국
7	(출원)	2017.04.12	대한민국
8	음절 다중조합 키워드 기반 문형 자동분류 방법(등록)	2017.07.25	대한민국
9	핸즈프리 자동통역을 위한 웨어러블 음성입출력 장치(등록)	2017.06.09	대한민국

3) 기술이전

번호	기술명	이전일자	이전업체	기술료(천원)
1	영/중/일 자동통역용 다국어 음성인식 및 자동번역 엔진	2017.09.26	KT	560,000
2	음성인식용 문장(대화체) 2종	2017.11.02	엘지전자(주)	48,000
3	음성인식용 단어 중 PC마이크/중가, PC헤드셋	2017.11.06	㈜맵퍼스	12,000
4	한국어 공통음성 DB	2017.11.15	삼성SDS(주)	22,320
				6.42억원 (부가세 제외)

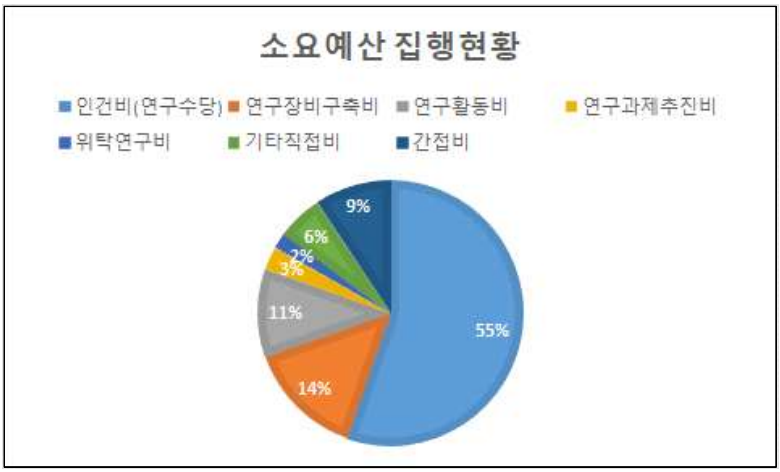
4) 표준기고서

번호	표준기고서 제목	표준화 기구
1	nformation technology — User Interface — Face-to-Face Speech Translation, Part 1. User Interface	ISO/IEC JTC1 SC35 User Interface
2	IInformation technology — User Interface — Face-to-Face Speech Translation - Part 2: System architecture and functional components	ISO/IEC JTC1 SC35 User Interface

다. 연구방법 및 추진체계의 효율적 추진여부 및 적정성(소요예산 집행의 적절성 및 전년도 평가결과·, 당해연도 중간평가 결과 지적사항에 대한 조치 및 개선실적 포함)

1) 소요예산 집행의 적절성

전체 예산 중 핵심원천기술 개발에 인건비(연구수당 포함) 55%를 투입하였고, 딥러닝서버구축 및 크라우드 소싱 등 인프라 구축에 14%를 활용하여 작년대비 인건비 비중이 높아지고 인프라 구축 등 비용은 감소함. 연구장비는 작년까지 주요 서버를 구입 완료하여, 금년에는 GPU 카드 업그레이드 등 효율화를 위한 유지보수 비용에 예산을 투입함. 특히 메년 신규 음성언어 DB구축으로 인해 고비용의 연구비 지출이 지속적으로 되는바, 금년에는 DB 자동정제를 통한 학습용DB 확보방안 연구를 착수하여 자동정제 알고리즘 개발에 인력을 투입함. 이에대한 결과로 DB 자동정제 기술을 금년에 확보한 바, 내년부터는 DB구축에 소요되는 비용 절감이 예상됨



2) 전년도 평가결과

구분	계정번호	과제명	구책임	부서명	평가점수	평가등급	순위	평가위원별 점수 및 평가의견				
								A	B	C	D	E
평가점수	16ZS1100	언어장벽 없는 국가 구	김상훈	SW·콘텐츠연구소	89.7	우수	1	87	89	90	90	94
평가의견	16ZS1100	언어장벽 없는 국가 구현을 위한 자동통번역 산업 경쟁력 강화 사업	김상훈	SW·콘텐츠연구소				○ 언어 번역 서비스는 글로벌시대에 항상 필요한 기술요소이며, 사업비 규모에 비행서 기술개발 성과 지표가 적정수준으로 판단됨. 또한 기술이전 3건, 기술료 1.26억원도 괜찮은 성과로 판단됨. ○ 시장성 부문에 있어서 평창올림픽 시범적용, 30만 앱 다운로드 등 가시적인 성과를 보인 부분도 긍정적으로 판단됨.	○ 한/스페인어(목표 69%), 한/불어(목표 68%) 통역 목표 대비 각각 76.1%, 61.8% 달성. 불어의 경우 평창올림픽 대비 평가셋 난이도 상향 조정으로 당초 목표 성능에는 미치지 못했으나 구글 시스템과 BMT 결과, 구글은 한/스 72.1%, 한/불 56.4%로 ETRI 통역시스템 성능이 구글 대비 한/스 4%, 한/불 5.4% 우수한 것으로 평가되어 당초 정한	○ 대국민 자동 통역 서비스의 산업 생태계 구축을 통해 개발된 통역 확산의 필요성이 있으며, 5차년도 목표대비 성능지표가 양호하다고 판단합니다.	○ 목표를 조과하는 우수한 성능의 기술 확보 ○ 원천 기술 확보를 고려할 때, SCI 논문의 목표 실적이 낮음 ○ 국내 특허 등록 실적 미흡 ○ App의 계산량, 메모리 용량, 반응속도 등의 지표 부족. Google App과의 비교 필요 ○ 표지의 연구비 오류 ○ 시장 점유율 확대, 사용자 증가를 위한 사업 전략 필요 ○ 글로벌 시장 진출의 구체적 전략 필요	○ 현재까지 우수한 연구결과를 체계적으로 수행하여 성취하였다고 판단함. ○ 목표를 평창동계올림픽으로 한정하지 말고 다음 global event 에 목표를 잡고 중장기적으로 세계적인 선도 기술로 발전시키기를 희망함.

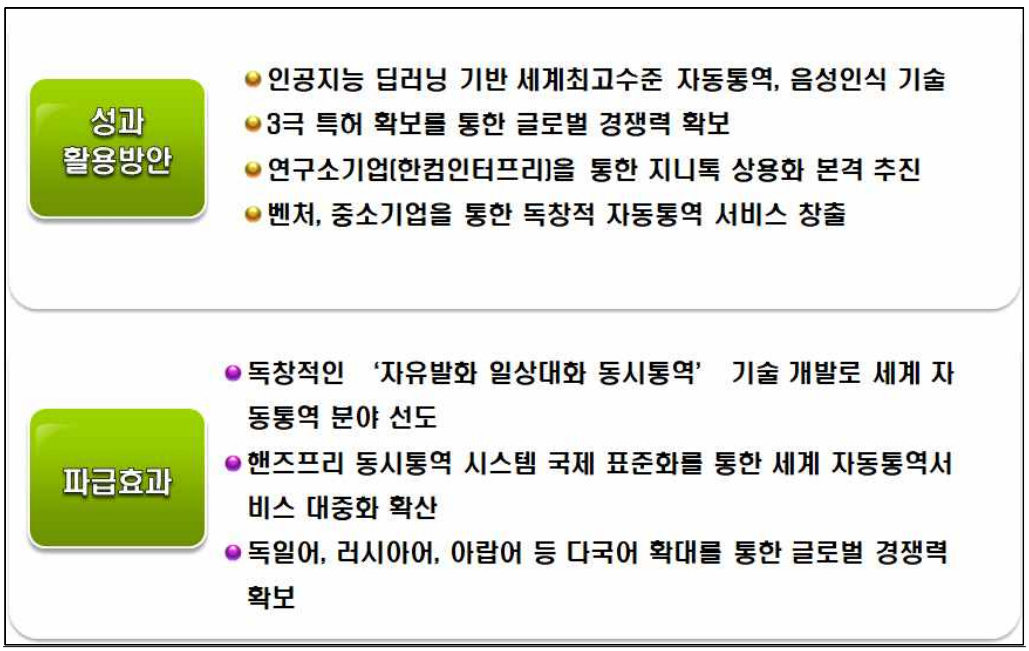
3) 전년도 평가결과 지적사항 및 조치 및 개선실적

지적사항	조치 및 개선실적
실제 상용화에 근접하게 2017년도에는 알고리즘을 더 최적화하도록 구성하였으면 함	기존 DNN 방식에서 RNN-LSTM 방식으로 음성인식 학습 알고리즘을 개선함. DNN대비 언어별 ERR 10~15% 성능 개선이 되었고, 체감성능도 확연히 좋아짐. 이 방식을 평창동계올림픽에 적용예정임 * ERR: Error Reduction Rate
원천 기술 확보를 고려할 때, SCI 논문의 목표 실적이 낮음	사업책임자 입장에서 국제특허, 국제표준 확보가 실질적으로 더 중요하기 때문에 IPR확보에 주력함. SCI논문은 1건 제출하였으나 현재 심사중임
국내 특허 등록 실적 미흡	ETRI 특허비용 절감 때문에 출원 후 등록이 늦어지고 있음. 작년 0건 대비 금년에는 2건 등록을 완료함
App의 계산량, 메모리 용량, 반응속도 등의 지표 부족. Google App과의 비교 필요	연초 6차례 동계올림픽 테스트이벤트를 통해 지니톡 앱의 완성도를 평가함. 속도, 계산량, 메모리 등 사용성에 큰 문제가 없는 것으로 평가되었음
시장 점유율 확대, 사용자 증가를 위한 사업 전략 필요	연구소기업인 한컴인터프리에서 지니톡 서버형, 단말형, 오프라인, 로봇 등 다양한 형태로 사업을 진행 중이며 언론홍보, 평창동계올림픽 공식공급사 계약 등을 통해 시장점유율 확대 및 사용자 확보에 주력하고 있음. 특히 KT 기가지니에 다국어용으로 기술이전을 완료하여 내년도부터 사용자 증가 및 시장점유율 확대가 기대됨
글로벌 시장 진출의 구체적 전략 필요	KT에 다국어 음성인식/번역 기술이전을 통해 기가지니의 글로벌 진출을 계획하고 있음. 호텔, 공항 등 1차 적용하고, 영어, 중국어. 일본어 버전 기가지니로 글로벌 시장에 진출할 계획임
성능 이외의 경쟁력 확보를 위한 새로운 차별적 전략 필요	제로유아이(Zero UI) 기술을 적용하여 외국인과 스마트폰 누름없이 자유롭게 대화가 가능한 기술을 국제표준화를 완료하였고, 현재 시제품 1000개를 제작하여 평창에서 시연 및 시범적용을 할 예정임

라. 성과활용방안 및 파급효과

모국어를 사용하는 사람간에 대화를 하듯이 외국인과 자연스런 양방향 자동통역으로 끊김없이 대화가 가능하게 해주는 세계최초 제로유아이(Zero UI) 자동통역서비스를 실현하기 위한 핵심 기술로, 본 사업에서 획득한 독창적인 음성검출 기술은 현재 국내외 경쟁기술이 없는 것으로 판단된다. 또한 본 기술은 채택된 표준을 따를 경우, 어떠한 자동통역 어플리케이션 또는 음성인식 어플리케이션에서도 손쉽게 응용이 가능하므로 스마트폰 터치 동작을 필요 없게 만들어 생산성 향상에 있어 일대 혁신을 가져올 것으로 기대된다.

본 기술과 관련한 국내외 family 특허를 다수 확보하였고, 최근 ISO 국제표준으로 승인이 된 제로유아이(Zero UI) 자동통역에서 본 발명에서 확보한 기술이 핵심이며, 2018 평창동계올림픽에 본 기술을 포함한 자동통역서비스를 세계최초로 시연하므로서 기술의 가치를 극대화 할 예정이다. 현재 국내 중견기업 및 대기업과 사업화 방안을 협의하고 있는 바, 본 기술이 적용된 웨어러블 기기가 대중화될 경우, 특허사용료(Royalty) 획득 및 실시권계약 또는 기술이전계약이 매우 긍정적으로 기대된다.



기존 상용 제품에 없던 양방향 언어소통이 가능한 통역기능이 추가되므로 기존 제품과 차별화가 가능하여 인공지능, 자동통역서비스, 웨어러블 디바이스 분야에서 시장성이 매우 유망하다고 판단된다. 특히 2017년 7월에 제로유아이(Zero UI) 자동통역 기술이 ISO 국제표준으로 최종 결정되어 이번 표준안에 핵심기술인 본 발명으로 인해 음성기반 사업 분야에서 수요창출 가능성은 매우 높을 것으로 예상된다.

2015년 기준 우리나라 출국자 1,958만 명, 입국외국인 1,335만 명 기록. 따라서 해외출입국자 3천만 명의 20%가 자동통역 서비스(이용료 1만원)를 이용할 경우, 연 600억원의 서비스 시장이 창출이 전망된다. 한/영 자동통번역의 경우 중립적 시나리오 하에, '13~'20년 누적 약 2,497억 원의 생산유발, 1,110억 원의 부가가치유발, 987명 수준의 고용유발 효과가 있고. 국내 산업 전체

의 경제적 파급효과는 ‘13~ ‘20년 누적 1.5조원의 생산유발, 6,740억 원의 부가가치유발, 5,995명의 고용유발 효과가 예상된 바 있다.
특히 2016년을 기준으로 무선 이어폰 출하량이 유선 이어폰을 넘어섰고 최근 아이폰과 구글 픽셀폰에서 이어폰 단자가 제거된 것에서 알 수 있듯이 블루투스 헤드셋이 시장의 대세로 부상하고 있다. 이에 따라 블루투스 이어폰 시장은 2017년 5,190만대에서 2019년에는 7,390만대로 확대될 것으로 예상된다.(자료=스트래티지 애널리틱스) 이렇게 치열한 블루투스 헤드셋 시장에서 본 기술은 핵심기능이 될 수 있으므로 본 기술은 시장에 빠르게 보급될 것으로 전망된다.

3. 기대성과

3.1. 기술적 측면

○ 자동통번역 기술은 대화체 음성인식, 언어번역, 음성합성 등 요소기술이 어우러진 복합기술로서, 이 요소기술들은 지능형 로봇, 텔레매틱스, 디지털 홈 등 IT성장동력산업 전 분야에서 **HCI(Human Computer Interaction)의 핵심요소**로 요구되는 원천 기술임

○ 자동통번역 세계시장 선점을 위한 국가간 기술개발 경쟁 치열한 상황임. 다국어 언어처리 기반 기술 개발을 통해 향후 유럽어, 아시아권 언어 등으로 자동통번역 및 음성인식기술을 확장할 수 있는 기반 기술이 확보됨

3.2. 경제 산업적 측면

○ 현재 자동통번역 관련 중소기업이 4~5개 정도임. 본 사업을 통해 매년 2~3개 중소기업을 지원하여 2018년에는 10여개 강소형 중소기업이 육성되고, 기존 자동통번역 업체는 글로벌 경쟁력을 가질 수 있는 중견기업으로 육성될 것으로 기대

○ 본 사업을 통해 구축한 리소스의 공동활용으로 자동통번역 기술개발 기간 단축 및 DB구축에 소요되는 비용을 절감하여 자동통번역산업의 국제경쟁력이 강화될 것으로 기대

○ 2012년의 경우, 우리나라 출국자 2,400만 명, 입국외국인 670만 명으로 추정됨. 따라서 해외출입국자 3천만 명의 20%가 자동통역 서비스(이용료 1만원)를 이용할 경우, **연 600억원의 서비스 시장이 창출**됨

○ **2020년 전세계 자동통역 시장규모는** 자동통역 단말기 6.3조원, 비즈니스용 자동통역 2.5조원, 교육용 자동통역 1.2조원 등 **총 10조원으로 예측** (출처: 일본 UFJ총합연구소, 2006)

○ 경제적 파급효과는 중립적 시나리오 하에, '13~' 20년 누적 약 2,497억 원의 생산유발, 1,110억 원의 부가가치유발, 987명 수준의 고용유발 효과가 예상

4. 향후계획

- 세계최고 중국어 음성인식기술 보유 업체인 iFlytek 대비 동등수준 중국어 인식엔진 성능 확보. 국내업체가 중국어 음성인식 시장에 진출할 경우, 경제적으로 산업적으로 큰 파급효과가 있을 것으로 예상됨
- 시스트란에 단말탑재형 자동통역기 기술이전을 통해 연말 상용화 추진
- 업체 기술지원 및 사업화 (현재 진행중인 업체: 시스트란, 헬로칫, 한글과컴퓨터, 스마트비투엠, 에버트란, SK텔링크, 국과수, LexiFone(이스라엘 통역서비스 업체) 등
- 차년도 이후에는 국내업체의 글로벌 경쟁력 확보를 위해 자유발화 일상대화형 동시통역이 가능한 세계최초 원천기술 개발에 도전할 예정이며, 2018년 평창동계올림픽에서 시험적용하도록 차년도 수행사업계획서(2017~2018)에 반영예정임

5. 부록

5.1. 번역이해도/통역성공률 평가방안

1. 평가방법

가. 평가 문장 풀(Pool)로부터 평가 대상 분야의 평균 문장길이를 고려하여 원문 선정
나. 외부용역 기관 5~7명의 전문번역가에 의한 객관적인 평가: 엔진 버전별 결과 혼합, blind 평가
다. 평가기준에 따른 각 문장별 0~4점 스코어링
라. 각 문장별로 최고, 최저 점수를 제외한 평균점수로 합산

2. 산출방법

가. 번역이해도/통역성공율 (%) = 3점이상문장수/평가문장수 x 100
나. 예: 80/100 x 100 = 80% (평가 대상 100문장 중 3점 이상인 문장 수가 80문장인 경우)

3. 평가 점수 부여 기준

점수	평가 기준
4.0	원어문의 의미가 그대로 전달된 경우
3.5	복문에서, 문장의 동사구가 정확히 전달되어 문장의 전체적인 의미의 골격이 전달되지만 동사를 제외한 1-2단어의 대역어가 잘못된 경우
3.0	문장의 동사구가 정확히 전달되어 문장의 전체적인 의미의 골격이 전달되는 경우
2.5	하나의 동사절이라도 정확히 번역되어 부분적으로 문장의 의미를 전달할 경우
2.0	하나 이상의 구가 정확히 번역되지만 전체적인 문장의 의미를 파악하기 어려운 경우
1.0	문장 중에 하나의 단어 또는 구라도 정확히 번역된 경우
0.0	번역문 출력이 안 된 경우

4. 출처:

- DARPA(Defense Advanced Research Projects Agency) 1994 adequacy test(Doyon, Taylor, and White, 1996)
- Workshop at the LREC 2002 Conference: Machine Translation Evaluation: Human Evaluators Meet Automated Metric

5.2. 음성인식률 산출 방법

1. 평가방법

가. 음향모델이나 언어모델 훈련에 사용되지 않은 음성신호 데이터를 1,500문장 이상 선정

나. 발화단위로 음성인식엔진에 입력하고, 인식 결과를 파일로 저장

다. 평가데이터 원문과 인식결과를 대응시켜 아래와 같이 음성인식률 산출

$$\text{인식률} = \frac{N - (S + D + I)}{N}$$

N: 전체 평가데이터에 포함된 단어 수

S: 원문의 단어와 달리 다른 단어로 인식된 단어의 수

D: 원문에는 존재하나 인식되지 않고 건너 뛴 단어의 수

I: 원문에는 존재하지 않으나, 인식결과에 추가로 포함된 단어의 수

2. 산출방법

가. 단어인식률의 경우, 레퍼런스와 인식결과간 단어대 단어간 비교하여 일치하는 개수를 카운트하여 전체 단어 대비 맞는 단어개수의 비율을 백분율로 측정

나. 문장인식률의 경우, 문장을 구성하는 단어가 모두 정인식 되었을 때 정인식 문장 개수를 카운트하여 전체 문장개수 대비 맞는 문장 개수의 비율을 백분율로 축정

5.3. 약어표

약어	의미
ERR	Error Reduction Rate
Open API	Open Application Program Interface
BM	Business Model
DARPA	Defense Advanced Research Projects Agency
ASP	Application Service Provider
GALE	Global Autonomous Language Exploitation
HCI	Human Computer Interaction
C-STAR	Consortium for Speech Translation Advanced Research
HMM	Hidden Markov Model
MID	Mobile Internet Device
MASTOR	Multilingual Automatic Speech-to-Speech Translator
PDA	Personal Digital Assistant
PMP	Portable Media Player
SNLP	Speech and Natural Language Processing
SRI	Stanford Research Institute
TALES	Translingual Automatic Language Exploitation System
TC-STAR	Technology and Corpora for Speech to Speech Translation
UMPC	Ultra Mobile Personal Computer
BMT	Bench Mark Test
UI	User Interface
OTG	On-The-Go
SMT	Statistical Machine Translation
LM	Language Model
TM	Translation Memory
DNN	Deep Neural Network
EPD	End-Point Detection

주 의

1. 이 연구보고서는 한국전자통신연구원의 주요사업으로 수행한 연구 결과입니다.
2. 이 보고서의 내용을 발표할 때에는 반드시 한국전자통신연구원에서 수행한 주요사업 결과임을 밝혀야 합니다.

부록 1)

- 평창올림픽 테스트이벤트 -

자동통번역서비스 정성 평가 결과보고

2017. 5. 29

Ⅰ TE 자동통번역 서비스 정성평가 개요

□ 평가 목적

T/E기간에 대회 참가자(선수, 임원, 기자, 관람객) 등을 대상으로 원활한 언어소통과 통번역 소프트웨어 품질향상을 위한 정성평가를 통해 자동통번역 서비스의 품질을 검증하고 개선사항 도출 및 보완을 통해 올림픽에 최적화된 서비스를 마련하는데 목적이 있음.

□ 평가 대회

○ 올림픽 테스트이벤트[17개] 중 6개 대회

연도	종 목	대 회	기 간 (공식연습 날짜)	비고
'16년	**스노보드**	**2017 FIS 스노보드 빅에어 월드컵**	**11.25 ~ 11.26** (23~24)	**1차**
	쇼트트랙	**2016/17 강릉 ISU 쇼트트랙 스피드 스케이팅 월드컵**	**12.16 ~ 12.18** (15)	**2차**
'17년	알파인 스키	2017 FIS 알파인 극동컵	1.16 ~ 1.17	
	크로스컨트리	**FIS 크로스컨트리 월드컵**	**2.3 ~ 2.5** (1~2)	**3차**
	노르딕복합	2017 FIS 노르딕 복합 월드컵	2.4 ~ 2.5 (3)	
	스피드스케이트	ISU 스피드 스케이팅 종목별 세계선수권대회	2.9 ~ 2.12 (6~8)	
	프리스타일스키	2017 FIS 프리스타일스키 월드컵	2.10 ~ 2.18 (7~10, 14~15)	
	스노보드	2017 FIS 스노보드 월드컵	2.12 ~ 2.19 (11, 14~15)	
	스키점프	2017 FIS 스키점프 월드컵	2.15 ~ 2.16 (14)	
	루지	FIL 루지 월드컵	2.17 ~ 2.19 (15~17)	
	컬링	2017 세계주니어 컬링 선수권대회	2.16 ~ 2.26	
	피겨 스케이팅	**2017 ISU 4대륙 피겨스케이팅 선수권대회**	**2.16 ~ 2.19** (14~15)	**4차**
	바이애슬론	IBU 바이애슬론 월드컵	3.2, 4~5 (1,3)	
	알파인 스키	2017 Audi FIS 스키 월드컵	3.4 ~ 3.5 (2~3)	
	봅슬레이/스켈레톤	**2016/17 BMW IBSF 봅슬레이/스켈레톤 월드컵**	**3.17 ~ 3.19** (13~16)	**5차**
	아이스하키	**IIHF U18 챔피언십**	**4.2 ~ 4.8**	**6차**
	아이스하키	IIHF 챔피언십(여자)	4.2 ~ 4.8	

□ 평가 내용

○ 한↔영어, 중국어, 프랑스어, 스페인어 언어쌍에 대한 Touch Point별 상황과 미션 제시를 통한 현장 만족도 조사

○ 근무지 : 경기장 내 베뉴 및 주변 주요 터치포인트

○ 평가패널 : 한국어가 가능하고 평가 대상 외국어에 대해 원어민 수준인 bilingual 패널 총 23명이 참여함 (대회 당 3~4명)

회차	종목	개최지	평가 언어쌍	패널 수
1차	스노보드	평창	한국어-영어, 중국어, 일본어, 프랑스어	4명
2차	쇼트트랙	강릉	한국어-영어, 중국어, 프랑스어, 스페인어	4명
3차	크로스컨트리	평창	한국어-영어, 중국어, 프랑스어, 스페인어	4명
4차	피겨스케이팅	강릉	한국어-영어, 중국어, 일본어, 프랑스어	4명
5차	봅슬레이/스켈레톤	평창	한국어-영어, 프랑스어, 스페인어	3명
6차	아이스하키	강릉	한국어-영어, 중국어, 프랑스어, 스페인어	4명

○ 평가자료 : 터치포인트 및 예문, 평가 설문지 **[붙임 3참조]**

○ 평가지표

- 현장 만족도 평가항목 : 터치포인트 검증, 서비스 품질(정확도, 사용성, 안정성, 속도, 만족도), 개선 의견 3가지 분야별 세부 평가 문항으로 구성

- 평가기준 : 평가 항목 별 자유롭게 기술 또는, 1~5점, 5단계로 평가

구분	정성 지표		
1. Touch Point 및 예문			
1-1 누락된 Touch Point	*추가해야 할 Touch Point 및 예문		
1-2 불필요한 Touch Point	*실제 존재하지 않는 Touch Point 및 불필요한 예문		
1-3 수정할 Touch Point	*수정되어야 할 Touch Point 및 예문		
2. 서비스 품질	**하**	**중**	**상**

2-1 정확도	1~2점	3점	4~5점
2-2 사용성	1~2점	3점	4~5점
2-3 안정성	1~2점	3점	4~5점
2-4 속도	1~2점	3점	4~5점
2-5 만족도	1~2점	3점	4~5점
3. 개선 의견			
3-1 화면 개선사항	*지니톡 화면 구성 개선 의견		
3-2 기능 개선사항	*지니톡 기능 개선 의견		
3-3 품질 개선사항	*지니톡 음성인식/번역 품질 개선 의견		

○ 평가수행

- 터치포인트 별 지니톡을 이용하여 음성인식 및 키패드 입력 등 멀티모달 방식으로 통번역 서비스 이용 후 평가 진행

○ 후속조치

- Touch Point 및 예문 검증 후 보완

- 서비스 품질 평가 결과가 낮은 항목에 대한 기능 및 품질 개선

- 지니톡 화면구성, 기능 및 통번역 품질에 대한 개선 의견 반영

Ⅱ 자동통번역 서비스 현장 만족도 조사 결과

□ 터치포인트 및 예문 보완

○ 실제 대회가 열리는 경기장 베뉴 및 주변 현장에서의 자동통번역 서비스 검증 및 설문조사를 통해 터치포인트 및 예문 목록 작성

○ 현장 확인을 통해 터치포인트(338개) 및 예문(622개) 목록 구축

○ 중복 제거 후 기 구축된 터치포인트 및 예문과 통합 예정

□ 자동통번역 서비스 설문 조사

○ 6개 대회 총 23명의 패널을 통해 자동통번역 서비스 설문 조사 수행

○ 만족도 높은 항목(4.0점 이상/5점) : UIUX 사용성, 사용방법 익히기, 기능 작동, 통번역 속도

○ 개선이 필요한 항목(4.0점 미만/5점) : 통번역 품질, 상황안내, 통번역시스템 안정성

○ 언어별 통번역 품질

- 전반적인 통번역 품질에 대한 만족도

· 평균 : 3.5점 (5점 만점)

· 영어, 중국어, 일본어 만족도가 상대적으로 높고, 프랑스어 스페인어 만족도가 상대적으로 낮으나 대체로 만족한다는 반응

- 음성인식 성능에 대한 만족도

· 평균 : 3.6점 (5점 만점)

· 모든 언어에 대해 전반적으로 만족도가 높았음.

· 고유명사 외래어 인식 및 소음 환경에서의 인식 성능 개선이 필요하다는 의견

- 번역 성능에 대한 만족도

· 평균 : 3.4점 (5점 만점)

· 중국어, 일본어에 대한 만족도가 높았음. 영어는 품질에 대한 기대치가 상대적으로 높았고, 프랑스어와 스페인어는 품질 개선이 필요하다는 의견

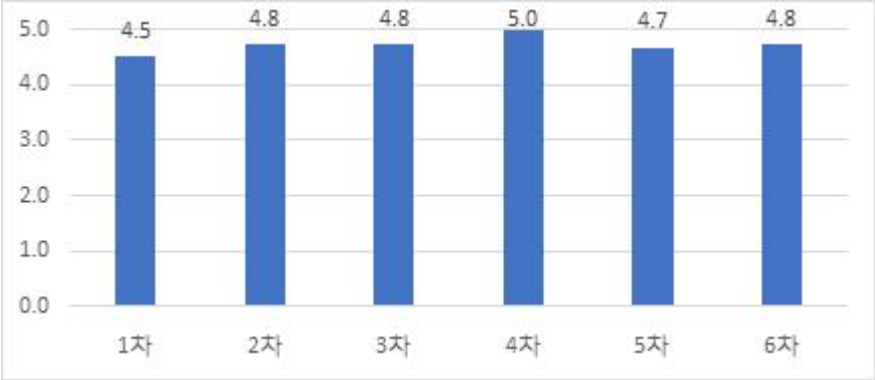
· 고유명사, 외래어, 올림픽 용어에 대한 업데이트 및 의문문/평서문 구별 등이 개선되어야 한다는 의견

○ UIUX 사용성

- 화면 문자 가독성

· 평균 : 4.7점 (5점 만점)

· 화면에 표시되는 텍스트(메뉴명 등)은 매우 쉽게 인지된다는 반응

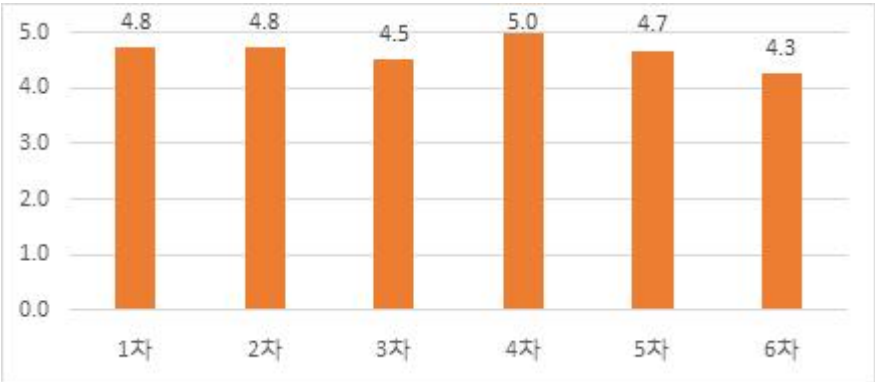


- 화면 아이콘 직관성

· 평균 : 4.7점 (5점 만점)

· 화면에 표시되는 아이콘들은 매우 쉽게 인지된다는 반응

· 언어설정 아이콘이 잘 눈에 띄지 않는다는 의견



- 원하는 메뉴 찾기

· 평균 : 4.6점 (5점 만점)

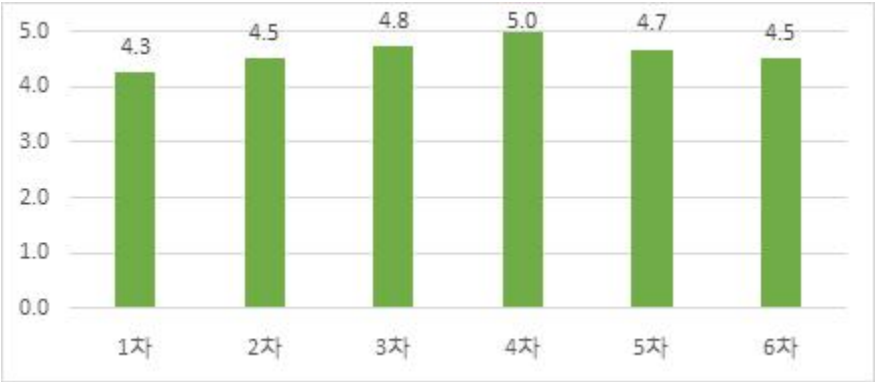
· 메뉴 구성이 심플하며 원하는 메뉴 찾기가 쉽다는 의견

· 언어설정이 편리하도록 메뉴 수정이 필요하다는 의견

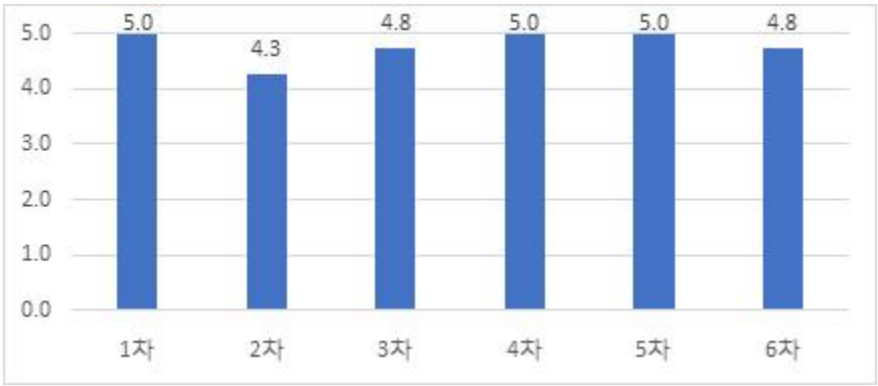
○ 기능 사용방법 익히기

- 음성인식 통역

· 평균 : 4.8점 (5점 만점)



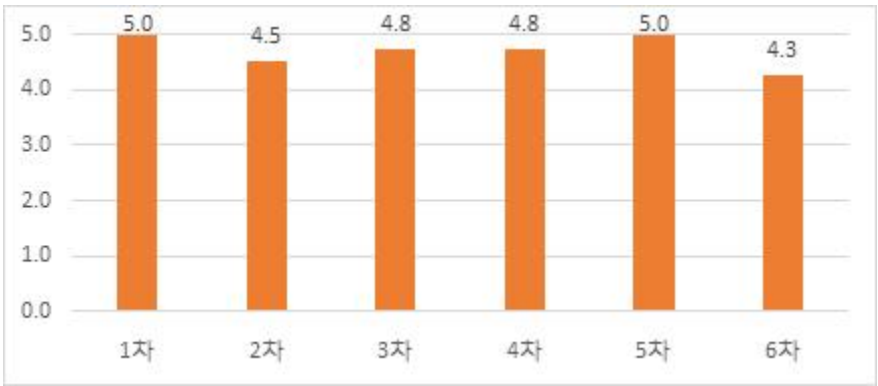
· 마이크 버튼을 터치하고 말하는 방식이 익숙하고 쉽다는 의견



- 문자입력 번역

· 평균 : 4.7점 (5점 만점)

· 문자 입력하는 방식이 음성인식보다는 번거롭지만 대체로 쉽다는 의견



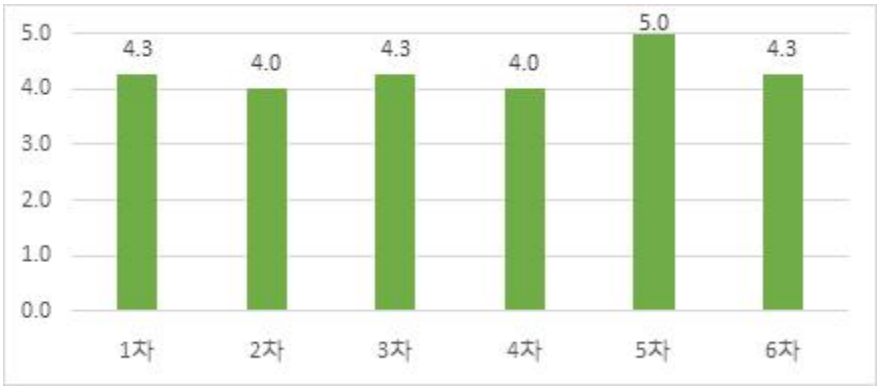
- 부가서비스

· 평균 : 4.3점 (5점 만점)

· 대화목록, 즐겨찾기, 설정 이용이 대체로 익숙하고 쉽다는 의견

· 간혹 통번역 설정 기능에 익숙하지 않은 경우에도 금방 적응함

· 대화목록 검색 및 전체 선택/공유 기능이 있으면 좋겠다는 의견



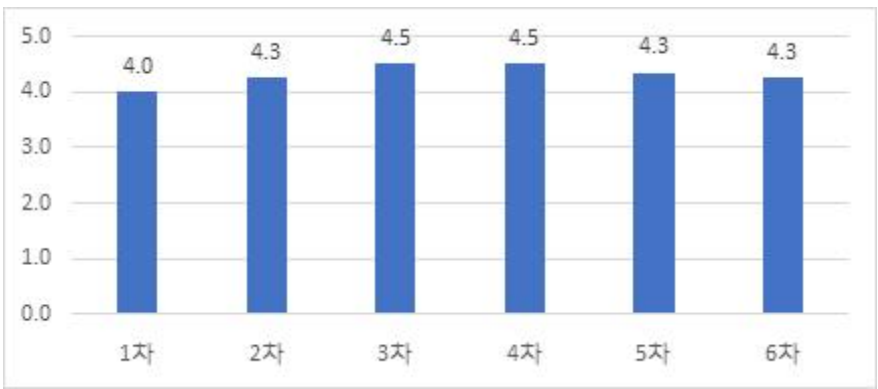
○ 화면 기능 작동

- 기능 작동

· 평균 : 4.3점 (5점 만점)

· 대체로 작동이 잘 됨.

· 입력문장 수정 후 번역 버튼을 눌렀을 때 재번역이 되지 않는 경우 있음

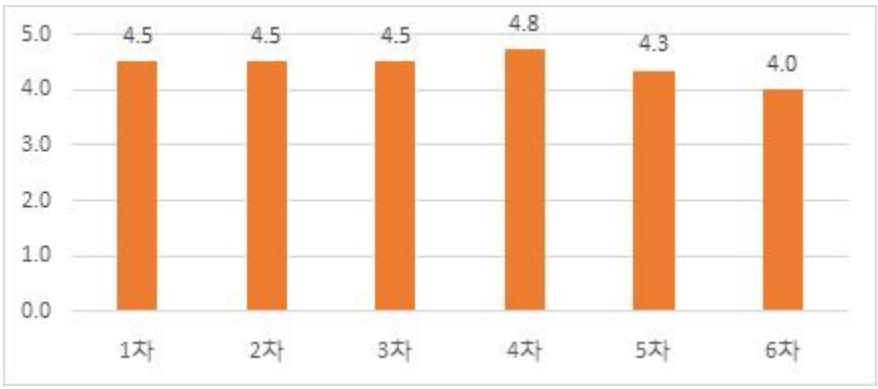


- 반응 속도

· 평균 : 4.4점 (5점 만점)

· 대체로 반응 속도가 빠름

· 간혹 번역 속도가 지연되거나, 음성인식 결과만 나오고 번역 결과가 표시되지 않는 경우가 있음



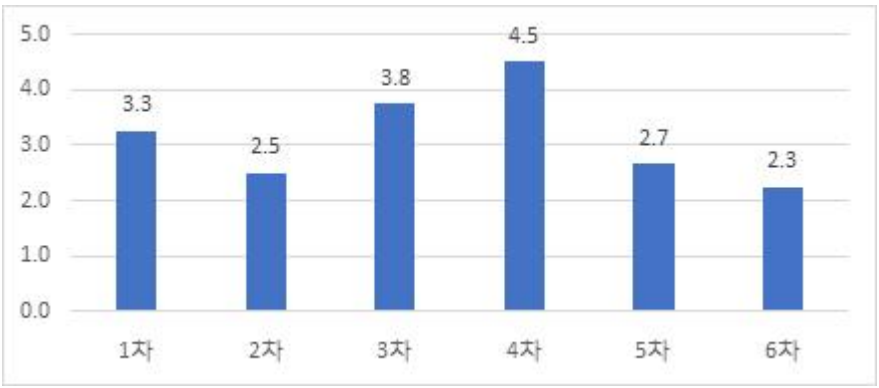
○ 상황 안내 메시지

- 서비스 이용 시 도움 정도

· 평균 : 3.2점 (5점 만점)

· 상태 안내메시지를 읽지 않는 경우가 많음

· 예외상황 발생 시 정확한 가이드를 주면 좋을 것 같다는 의견

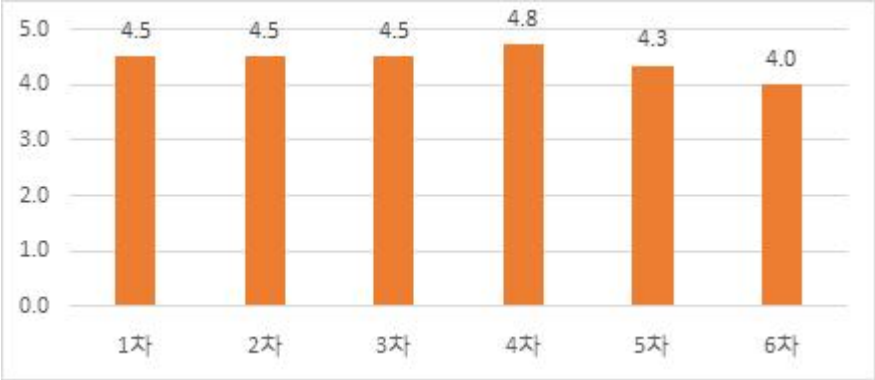


- 상황 안내 메시지 직관성

· 평균 : 3.5점 (5점 만점)

· 상태 안내메시지를 읽지 않는 경우가 많음

· 좀 더 눈에 띄게 메시지 또는 이미지를 표시해주면 좋을 것 같다는 의견



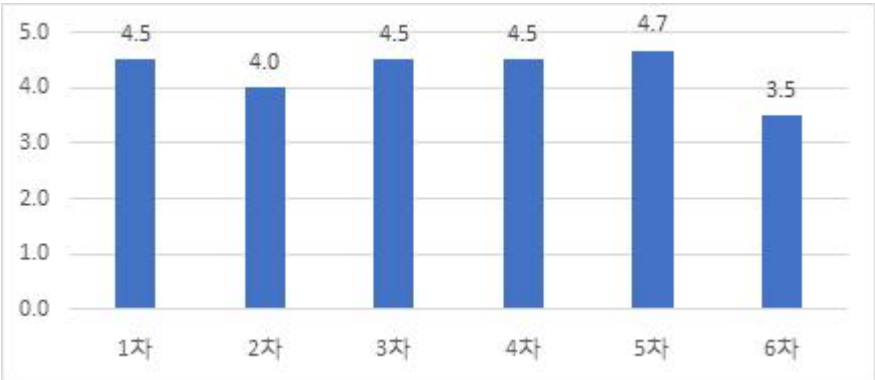
○ 통번역 속도

- 음성인식 속도

· 평균 : 4.3점 (5점 만점)

· 대체로 음성인식 속도가 매우 빠르다는 의견

· NMT(인공신경망) 적용 이후 다소 속도 저하가 느껴진다는 의견



- 번역 속도

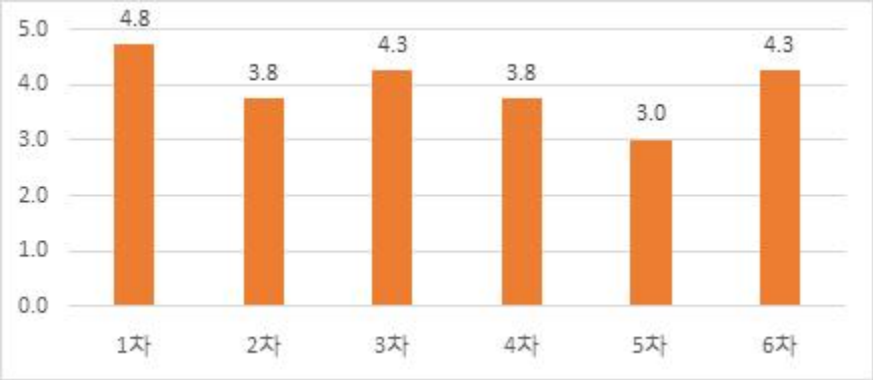
· 평균 : 4.0점 (5점 만점)

· 대체로 번역 속도가 매우 빠르다는 의견

· 간혹 음성인식 후 번역 결과가 나오지 않는 경우 속도가 느려진 것으로 느끼는 경우가 있음

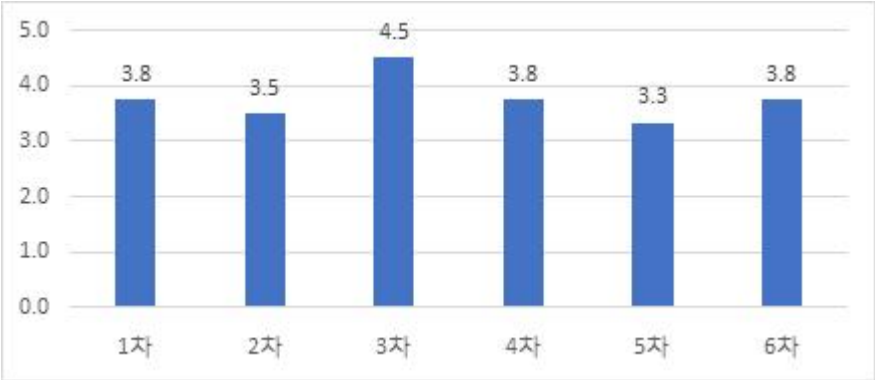
○ 통번역 안정성

- 음성인식 안정성



· 평균 : 3.8점 (5점 만점)

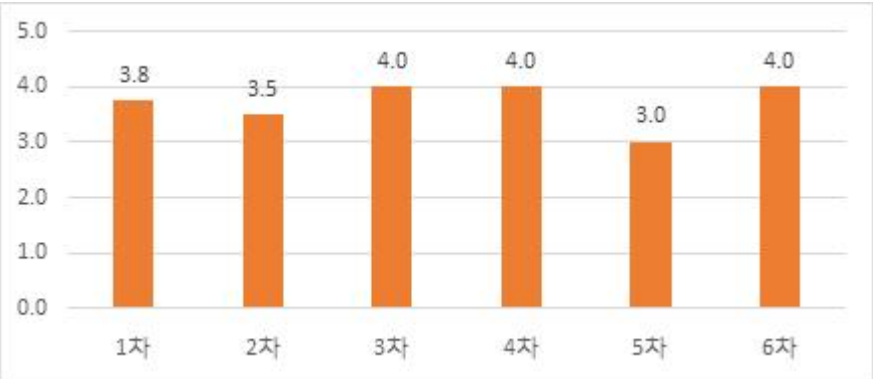
· 대체로 안정적으로 인식이 되나 가끔 음성입력이 검출 되지 않아 재시도해야 하는 경우가 있음



- 번역 안정성

· 평균 : 3.7점 (5점 만점)

· 대체로 안정적으로 번역이 되나 가끔 결과가 나오지 않는 경우 있음



Ⅲ 자동통번역 서비스 언어별 품질 평가 내용

1 언어별 주요 통번역 품질 개선 사항

구분		지니톡 장점	개선 필요 사항
공통사항		- 한국어 기반 번역은 타 번역기 대비 매우 우수함 - 소음환경에서 타 번역기 대비 음성인식이 잘 되는 편임	- 고유명사(지명, 종목명), 외래어, 약어 통번역 품질 개선 필요(지속 진행 중) - 소음 환경에서 음성인식 성능 개선 필요 - 동음 이의어를 문맥에 따라 번역 개선 - 숫자/알파벳 포함된 문장 번역 개선 - UI 사용성 개선 (상태 안내, 언어 설정 간편화 등)
언어	영어	- NMT 적용 후 번역 성능이 확연히 개선되었음	- 원어민이 사용하는 어휘 번역 품질 개선 ex) AD 등록 센터, Advil ADA access rule(장애인 정책)
	중국어	- 북경어의 경우 음성인식 및 번역 만족도가 높음 - NMT 적용 후 중문 등 복잡한 문장 번역 품질이 개선었음	- 다양한 성조, 연음에 따른 음성인식 개선 - 중문(접속사로 연결된 문장)에 대한 번역 개선 필요
	일본어	- 음성인식 및 번역 만족도가 대체로 높음	- 외래어 음성인식, 번역 개선 필요 - 숫자 음성인식 개선 필요 ※ 일본어는 현재 NMT 적용 진행 중 임
	스페인어	- 타 번역기 대비 우수한 음성인식, 번역 품질	- 영어, 중국어, 일본어에 비해 낮은 품질 - 의문문, 평서문 구분 개선 필요
	프랑스어	- 타 번역기 대비 우수한 음성인식, 번역 품질	- 영어, 중국어, 일본어에 비해 낮은 품질 - 의문문, 평서문 구분 개선 필요

2 한국어 ↔ 영어

□ 전반적인 의견

○ NMT(인공신경망 번역) 적용 이후 번역 정확도가 개선되었다는 의견이 많음. 특히 이중문장의 번역이 개선되었음

○ 지명, 올림픽 관련 용어 등 고유명사 업데이트가 필요함

○ 주변 사람 말소리 등의 잡음 환경에서의 음성인식 개선 필요함

□ 영어 음성인식

○ 영어 발화 시 고유 명사 인식 문제

- 정선, 평창, 강릉, 관동, 알펜시아, 수호랑 인식 안 됨

- wifi, advil, ice arena 인식 잘 안 됨

□ 영어 → 한국어 번역

○ 잘못된 번역

- AD card [광고 카드(AD카드)], A D card [D 카드(A D 카드)]

- stuffed animals [박제동물(봉제 인형)]

- how can I get to the ski jump stand? [스키장(스키점프대)에 가려면 어떻게 가야하나요?]

- can you stand in two lines? [줄을 두고(두 줄로) 서 주시겠어요?]

- elevators and stairs are restricted to the spectators. [관람객(엘리베이터)과 계단은 관람객들에게 제한 되어있습니다.]

- I have stuffed up nose. where can I get some medicine? [잘 모르겠어요(코가 막혔어요), 약을 어디서 구할수 있을까요?]

○ 직역 또는 어색한 번역

- Alpensia stadium [알펜시아 구장(알펜시아 스타디움)]

- grand stand [그랜드스탠드(관중석)]

- nursing room [간호방(수유실)]

- is this a mixed zone? [이 혼합된 지대(취재구역)입니까?]

- where is the venue accrediation center? [개최지 인가 센터(경기장 출입 인가 센터)가 어디에 있습니까?]

- I have a headache. [골치 아파요(두통이 있어요)]

□ 한국어 → 영어 번역

○ 고유명사 번역 문제

- 장국수, 교동 짬뽕, 켄싱턴 플로라 가든, 수호랑, 반다비

○ 직역 또는 어색한 번역

- 응원도구 [cheering tools (cheering equipment)]

- 경기 일정 표 좀 주시겠어요? [Can I have a ticket for the game?]

- 잠시 몸 수색을 하겠습니다. [I will search for it for a while.]

- 휴지통이 어디 있어요? [Where is the trash store?]

- 이쪽 게시판에 셔틀 시간 표가 붙어 있습니다. [The shuttle time ticket is attached to this side bulletin board.]

- 경기 안내 어플을 설치해 드릴까요? [Shall I install the game conduct application ?]

3 한국어 ↔ 중국어

□ 전반적인 의견

○ NMT(인공신경망 번역) 적용 이후 번역 정확도가 개선되었다는 의견이 많음

○ 번역 시 의문문/평서문 구별 개선 필요

○ 성조, 연음 변화에 따른 음성인식 개선이 필요

※ 문자 입력 번역으로 보완 가능

□ 중국어 음성인식

○ 중국어 발화 시 고유 명사 인식 문제

- 알펜시아, 와이파이, 봅슬레이, 횡계로터리, 수호랑, 반다비, 알펜시아, 아이스아레나

※ 중국어 명칭이 없는 외래어의 경우 실제 발화하는 경우가 거의 없으므로, 실제 통역 환경에서는 문제되지 않을 수 있다는 의견

○ 알파벳이 포함된 단어 인식 문제

- A 구역, D 구역

○ 성조가 같거나 비슷한 발음 오인식

- 去 → 却
- 取 → 去
- 有 → 由
- 拾 → 10
- 一 → 1
- 有雨具 → 由于剧

□ 중국어 → 한국어 번역

○ 고유명사, 외래어 번역 문제

- 跳台滑雪比赛场是在这站下车吗？ [스키 다이빙(스키점프) 경기장은 이 역에서 내리나요?]
- 雪车比赛场地是在这站下车吗？ [설거 시합 장소(슬라이딩센터)는 이 역에서 하차하는 것입니까?]
- 冰上比赛场是这站下车吗？ [컬링 경기장(빙상경기장)이 이 역에서 내리나요?]
- 需要坐摆渡车去滑板比赛场吗？ [페리(스노보드) 경기장에 셔틀버스를 타고 가야하나요?]
- 男子雪橇比赛几点开始？ [남자 배구경기는(봅슬레이 경기는) 몇시에 시작하나요?]
- 今天只有女子俯式冰橇的比赛吗？ [오늘은 여자들만 스케이트(스켈레톤)를 탈 수 있는 경기가 있나요?]
- 有可以和吉祥物照照片的地方吗？ [쏨땀(마스코트) 사진을 찍을 수 있는곳이 있나요?] : 마스코트 → 쏨땀

○ 잘못된 번역 – 다른 의미로 번역

- 在旌善会举办什么比赛？ [정선회에(정선에서) 무슨 시합을 개최합니 까?]
- 停车费多少钱？ [도중하차(주차) 요금은 얼만가요?]
- 我想领出入证。 [등록증(출입증)을 받고 싶습니다.]
- 比赛场里面有休息室吗？ [경기장 안에 라운지(휴게실)가 있나요?]
- 可以用银联卡结账吗？ [직불카드(은련카드)로 계산해도 되나요?]
- 有没有咖啡因的咖啡吗？ [카페인이 있는(없는) 커피는 없나요?]
- 女子自由滑比赛几点开始？ [여자 자유 활비새(여자 프리 경기)는 몇시에 시작합니까?]
- 2位大人1位儿童。 [2분 1분 어른 어린이.(성인 2명, 어린이 1명)]
- 包要全部打开吗？ [전부 열고 포요 니까?(가방을 모두 열어야 합니까?)]
- 这个2个这个2个打包。 [이 2 이 포장합니다.(이거 2개 이거 2개 포장해주세요)]
- 我要结账。 [체크아웃 부탁합니다.(계산하고 싶습니다.)]
- 剩下的是小费。 [팁은 팁입니다.(나머지는 팁입니다.)]

○ 직역 또는 어색한 번역 – 직역을 하여 어색한 표현으로 번역

- 从3天前开始发烧。 [3일 앞에서(전부터) 열이 나기 시작합니다.]

- 可以走到冰上比赛场吗？ [얼음 경기장(빙상경기장)까지 걸어갈 수 있나요?]

○ 숫자 번역 오류 – 숫자를 정확하게 번역하지 못하는 경우

- 我要1杯热巧克力和1杯卡布奇诺。 [핫초코 1잔 카푸치노+1(1) 잔주세요.] : 한 문장에 숫자 2건 이상 있을 경우 자주 발생함

- 男子1,500米决赛是几点开始？ [남자 1,318m(1,500m) 결승전은 몇시에 시작하나요?]

○ 의문문/평서문 번역 오류 – 의문문을 평서문으로 번역하는 경우

- 有糖浆吗？ [시럽 있어요] : 의문문 → 평서문

- 有桂皮粉吗？ [시나몬가루 있어요] : 의문문 → 평서문

□ 한국어→중국어 번역

○ 고유명사, 외래어 번역 오류 – 원문의 단어가 번역 시 누락되는 경우

- 대한민국 동쪽에 위치한 평창 강릉 정선에서 열립니다. [在大韩民国东边的平昌江陵举行。] : '정선' 누락

- 경포 해변 가 보셨어요? [去过海边吗？] : '경포' 누락

- 크로스컨트리 경기는 평창에서 열립니다. [越野赛比赛在平昌举行。] : 越野滑雪(크로스컨트리) 뒷부분 누락

○ 잘못된 번역 – 다른 의미로 번역되는 경우

- 평창 동계 올림픽 슬로건은 패션 커넥티드입니다. [平昌冬季奥运会口号时髦是连接的。] : passion → fashion : '패션' 2개 이상의 의미가 있는 단어 처리 필요

- 겉옷은 벗어 주세요. [请脱下你的夹克衫。] : 겉옷 → 잠바

- 자동 통역 어플 지니톡을 설치해 드릴까요? [给您安装自动变速器GenieTalk吗？] : 자동통역 어플 → 자동변속기

○ 이중문장 번역 오류 - 두 번째 문장의 의미가 다르게 번역되는 경우

- 봄과 가을은 시원하고 여행하기에 좋습니다. [春天和秋天是很凉爽的旅行。] : "봄과 가을은 시원한 여행이다"로 오번역

- 지금 두 알 드시고 저녁 먹고 난 뒤에 두 알 드세요. [现在两点吃,晚饭后吃两粒。] : "저녁 먹고 두 시에 먹어요"로 오번역

4 한국어 ↔ 일본어

□ 전반적인 의견

○ NMT(인공신경망 번역) 적용 이전 버전으로 테스트하였으나 대체적으로 번역이 잘 되었으며, 다른 언어에 비해 통번역 품질에 대한 만족도 높음

○ 고유명사, 외래어 번역 업데이트 필요

□ 일본어 음성인식

○ 일본어 발화 시 고유 명사 및 외래어 인식 문제

- 횡계, 반다비, 수호랑 엠블럼, 포토월, 아이스아레나, 빅에어, 코트

○ 동음이의어 인식 문제

- 並→発
- に→2
- 脱→抜
- あって→会って

○ 기타 인식 문제

- 숫자 인식이 잘 안됨 : 6, 7, 8
- 알파벳 인식이 잘 안됨 : Dゲート, A区域, B区域, A席
- 선수(先手→ 千秋)

□ 일본어 → 한국어 번역

○ 고유명사 번역 문제

- 강원도 홍보관, 옹심이

○ 외래어 번역 문제

- アイポン (아이폰 → 아이퐁),
- ワイファイ(와이파이 →와이 화이)
- アルペンシア(알펜시아 → 알펭시아)
- 龍平(용평→류우헤이)
- モバイルエプ。(모바일 앱 → 모바일에프)
- パリバゲトゥ(파리바게뜨 → 파리바게토우)
- フェイスペインティング(페이스페인팅→흐에이스페잉티)

○ 잘못된 번역 – 다른 의미로 번역되는 경우

- 明日1番早く出発するソウル行きのバスは何時ですか。→ 내일 한 번(가장) 빨리 출발하는 서울행 버스는 몇 시입니까?"

□ 한국어 → 일본어 번역

○ 잘못된 번역 – 다른 의미의 단어로 번역되는 경우

- 평창 동계 올림픽 슬로건은 패션 커넥티드입니다. [平昌冬季オリンピックのスローガンはファッションです。] : 패션(passion → fashion) : 2개 이상의 의미가 있는 단어 처리 필요, '커넥티드' 번역 누락

○ 의문문/평서문 번역 오류 – 의문문을 평서문으로 번역하는 경우

5 한국어 ↔ 프랑스어

□ 전반적인 의견

○ 타 번역기 대비 번역 품질이 좋다는 의견

○ 의문문을 평서문으로 인식하는 문제 개선 필요

○ 고유명사, 외래어 음성인식, 번역 업데이트 필요

□ 프랑스어 음성인식

○ 프랑스어 발화 시 고유 명사 및 외래어 인식 문제

- 경기도, 평창, 강릉, 정선, 파리바게트, 알펜시아, 경포, 강원랜드, 수호랑, 반다비

○ 의문문 인식 시 물음표(?) 생략되고, 평서문으로 인식되는 문제

○ 같거나 비슷한 발음 인식 문제

- je → tu
- je veux → je vous
- wifi → whippy
- je voudrai → je veux
- double → devez
- Votre → autre

○ 발화하지 않은 단어가 인식 결과에 덧붙여져 나타나는 문제

○ 알파벳이 포함된 단어 인식 문제

- AD, clase A, zone A, zone D

□ 프랑스어 → 한국어 번역

○ 잘못된 번역 – 단어가 누락되거나 번역이 안 되는 경우(주로 고유명사 또는 외래어

- puis-je avoir le schedule de bus pour aller à séoul? [서울로 가는 버스가 있나요?] : le schedule(시간표) 누락

- comment je peux y aller Alpensia sliding? [어떻게 그곳에 갈 수 있는 slid ING?] : sliding 번역 안 됨.

- je voudrais y aller ice arena. [거기에 가보고 싶어요.] : ice arena (아이스아레나) 누락.

- est-ce que je dois descendre d'ici pour aller les skis jump? [여기서 내리려면 여기서 내리면 되나요?] : skis jump(스키점프) 누락.

- est-ce que je dois descendre d'ici pour aller à court de pistes. [여기서 멀리 떨어져 있어야 하나요?] : court de pistes(쇼트트랙) 누락.

- comment je peux y aller et le stade de patinage? [그럼 어떻게 가야 하나요?] : stade de patinage(스케이트경기장) 누락

- je voudrais acheter le ticket par smart portable. [스마트 폰으로 구매하고 싶습니다.] : 'ticket(티켓)' 누락

- qu'est ce que c'est le id et pw la pour utilise de wifi? [와이 파이 아이디와 와이 파이 사용 방법은 무엇 이냐?] pw(패스워드) 누락

- quand est-ce que ça va commencer le 16éme final de la course d'homme? [대회 결승전은 언제 열리나요?] : 16강(16éme) 누락

○ 잘못된 번역 – 다른 단어 또는 의미로 번역되는 경우

- est-ce que je pourrais avoir les horaires des bus entre Pyeongchang et seoul? [휘슬러(평창)와 서울 사이의 버스 시간표를 알 수 있을까요?] :

- est-ce qu'il n'y a Paris Baguettes. [파리에는 젓가락이(파리바게뜨) 없어요]

- comment je peux y aller Alpensia glissade? [어떻게 가면 에버랜드(알펜시아 슬라이딩)에 갈 수 있나요?]

- je voudrais aller stadium d'alpensia [경기도 종합 경기장(알펜시아 스타디움)에 가려고 합니다.]

- est-ce que je peux louer siege d'auto d'enfants. [어린이용 휠체어(카시트)를 대여할 수 있는 곳이 있나요?]

- est-ce que le taxi qui pour entrer dans le jeux de stadium? [고양종합운동장(경기장)에

들어가려면 택시를 타야 하나요?] : 택시를 타고 경기장에 들어갈 수 있나요?

- je voudrais avoir un carte AD? [신용 카드(AD 카드)를 받고 싶은데요.] : AD 카드 → 신용카드

- amenez-moi d'ici? [여기서(여기로) 데려가 주시겠어요]

- est-ce que je peux refabriquer? [유레일 패스를 재발급 받을 수 있나요?] : 원문에 없는 단어

- est-ce que je peux acheter le ticket par app? [상품권은 상품권으로 살 수 있어요.(앱으로 티켓을 살 수 있나요?)]

- comment je peux aller à la stad de cross-country? [경기장에서 잠실야구장(크로스컨트리 경기장)까지 어떻게 가나요?]

- où est la stad de skeleton? [조정 경기장(스켈레톤 경기장)은 어디에 있나요?]

- c'est où arrive? [거기가(도착 지점이) 어디 죠?]

- est-ce que c'est le bêtiment est centre d'hockey? [고인돌이(이 건물이) 선학 하키 경기장 인가요?]

- je voudrais aller à zone D. [D 구역으로 가세요(가고 싶다).]

- les joueurs la bas c'est qu'elle nationalité? [선수들이 나이가 들었나요?(선수들의 국적은 어디 입니까?)]

- est-ce qu'il y a d'autres entrées? [다른 에피타이저(입구)도 있나요?] : 2개 이상의 의미가 있는 단어

○ 의문문/평서문 번역 오류 – 의문문을 평서문으로 번역하는 경우

- les compétitions de jeux d' hiver c'est? [겨울 게임 대회가 입니다.]

○ 이중문장 번역 오류

- je peux louer ici et d'apres est-ce que je peux déposer à séoul. [서울에서 렌터카와 픽업을 할 수 있나요? (여기서 빌려서 서울에서 반납해도 되나요?)]

- elle a l'air très jeune quel âge elle a? [그녀는 나이가 아주 어려 보여요. (몇 살인가요?)] 2번째 문장 번역 누락

□ 한국어 → 프랑스어 번역

○ 잘못된 번역 – 단어가 누락되거나 번역이 안 되는 경우

- 평창 올림픽은 열여덟 번째 동계 올림픽입니다. [Le jeu olympique olympique est le dix-huit olympique.] : 평창 누락, 올림픽 3번 나옴
- 크로스컨트리 경기는 평창에서 열립니다. [Le match de country country aura lieu.] : '크로스컨트리' 누락
- 요금은 10,000 원 이하로 나옵니다. [Le tarif est 10 000 wons.] : '이하로' 누락
- 아이스하키 경기는 강릉 하키 센터에서 열립니다. [Le match de 아이스하키 va s'ouvrir au centre de hockey 강릉.] : 아이스하키 강릉 번역 안 됨
- 정선에서는 알파인 경기장이 있습니다. [Il y a un stade de 알파인 à l'intérieur.] : 정선 알파인 번역 안 됨
- 동서울 터미널로 가는 버스는 몇 시에 있습니까? [À quelle heure le bus pour 동서울 part-il?]

○ 잘못된 번역 – 다른 단어 또는 의미로 번역되는 경우

- 동서울로 가는 버스가 내일 오전 6시 50분에 있습니다. [Le bus pour séoul part à 6 heures 50 demain matin.] : '동서울' → '서울'로 번역
- 대한민국 동쪽에 위치한 평창에서 열립니다. [Il s'ouvre à l'hôtel(à Pyeongchang.) de corée du sud.] 평창 누락
- 2018 년 2월 9일부터 25일까지 17일 간 열립니다. [C'est l'ouverture du 9 2 janvier de 20 18 janvier en corée du sud.]
- 평창 동계 올림픽 슬로건은 패션 커넥티드입니다. [Les jeux olympiques d'hiver 평창 sont à la mode de mode.] : passion → fashion(la mode) : '패션' 2개 이상의 의미가 있는 단어 처리 필요

○ 인칭 생략시 번역 문제

- 한국어 입력 시 주어가 생략되면, 문맥과는 달리 3인칭 → 1인칭으로 번역하는 경우가 많음.

6 한국어 ↔ 스페인어

□ 전반적인 의견

○ 의문문을 평서문으로 인식하는 문제 개선 필요

○ 고유명사, 외래어 음성인식, 번역 업데이트 필요

○ 번역 오류 패턴이 프랑스어와 유사함

□ 스페인어 음성인식

○ 스페인어 발화 시 고유 명사 및 외래어 인식 문제

- 정선, 횡계, 수호랑, 반다비, 알펜시아, 크로스컨트리, 파리바게뜨, 타이레놀, 애드빌, 라떼, 텍스리펀드, 셔틀버스, AD, ID, 푸드트럭, bobsleigh, wifi

○ 의문문 인식 시 물음표(?) 생략되고, 평서문으로 인식되는 문제

- hace mucho frío en invierno de esa zona. [그 지역의 겨울은 매우 춥다.]
- puedo recibirel horario de servicio de bus a seúl. [서울 가는 버스 시간표 받을 수 있어요.]
- hasta que horaay el bus que va a seúl. [서울 가는 버스가 몇 시까지 있을 거예요.]
- no tengo tarjeta de transporte puedo pagar efectivo. [교통 카드는 없고 현금 결제도 가능합니다.]

○ 같거나 비슷한 발음 인식 문제

- eslogan(슬로건) → es logan
- entrada → entraba로 오인식
- desde(부터) → de este(이걸로부터)
- esta (this) → está(~가 ~에 있다)

○ 인식이 잘 되지 않는 단어

- invernales (동계) 인식 안 됨.

- hay (~ 있니?) 인식 안 됨.

○ 알파벳이 포함된 단어 인식 문제

- clase A, zone D, zone A

□ 스페인어 → 한국어 번역

○ 잘못된 번역 – 단어가 누락되거나 번역이 안 되는 경우(주로 고유명사 또는 외래어

- la república de corea es él anfitrión juegos olímpicos de invierno de este año. [대한민국은 올해 동계 올림픽 anfitrión(개최) 이다.]
- décima octava olimpiadas de invierno es abierto corea del sur pyeong chang. [Décima(18번째) 올림픽은 한국의 pyeong olimpiadas 입니다.]
- ¿se puede comprar la entrada parael partido de clasificación? [Clasificación(예선 경기) 입장권을 살 수 있나요?]
- ¿se puede comprar ahora el ticket de partido final de mañana? [내일 (본선) 경기표 살 수 있나요?] : partido final(본선) 번역 안 됨
- ¿dónde queda el estadio de bobsleigh? [Bobsleigh(봅슬레이 경기장)은 어디에 있나요?]
- ¿dónde queda el estadio de skeleton? [Skeleton(스켈레톤) 경기장은 어디 입니까?]
- explícame sobre el símbolo olimpiadas invernales de pyeongchang. [평창의 동계 올림픽 símbolo(심볼)에 대해 설명해 주세요.]
- ¿habrá alguna cafetería cerca podría usar internet? [근처에 (인터넷 사용할 수 있는) 카페가 있나요?]
- ¿cuánto es el precio de bus por persona? [1 인당 (버스) 요금은 얼마 입니까?]
- este es el paradero/parada del estadio. [여기가 경기장의 (정거장) 위치 입니다.]
- quiero emitir la tarjeta de acreditación. [(인가) 카드를 발급 받고 싶습니다.]

○ 잘못된 번역 – 다른 단어 또는 의미로 번역되는 경우

- ¿dónde es el lugar de anfitrión? [예배당(개최장소)은 어디에 있나요?]

- ¿se puede comprar el boleto con la aplicación movil? [표를 가지고 있는(모바일) 앱으로 표를 살 수 있나요?] :

- ¿se puede recibir el ticket por smart phone? [스마트 폰(티켓)을 스마트 폰으로 받을 수 있나요?] : el ticket(티켓) 오번역

- hay aplicación de información de partido? [정당 정보(경기 안내 앱)가 있다.]

- ¿disculpe usted sabe donde queda la pista corta? [실례 지만, 숏 코트(쇼트트랙 경기장)가 어디 있는지 아세요?]

- ¿a qué hora empieza el partido de clasificación de mujeres de quinientos metros? [몇 시에 500 미터 여자 축구 경기가(예선전이) 시작되나요?]

- segunda semifinal. [준준준준(준준결승).]

- partido de bobsleigh de mujer esta mañana. [여자 들러리(봅슬레이)는 오늘 아침에 만난다.] : → 들러리

- necesito papel húmedo. [저는 색종이(물티슈)가 필요해요]

○ 직역 또는 어색한 번역

- ¿donde se abre el partido de patinaje de hielo? [얼음 스케이트(빙상 경기)는 어디서 열리나요?]

- ¿habrá bus que vaya al arena de hielo? [혹시 아이스쇼(아이스아레나 가는)

- arena de hielo. [얼음 모래(아이스아레나) 요]

- ¿hasta qué hora tengo que devolverme al bus? [버스로 몇 시까지 반납해야(돌아와야) 하나요?]

- ¿puedo bloquear la bicicleta? [자전거를 막아도(잠가도) 되나요?]

- ¿dónde puedo comprar peluche de mascota? [애완동물(마스코트)은 어디서 살 수 있나요?]

□ 한국어 → 스페인어 번역

○ 잘못된 번역 – 단어가 누락되거나 번역이 안 되는 경우

- 크로스컨트리 경기는 평창에서 열립니다. [El partido dividido se celebra en

pyeongchang.]

- 정선에는 알파인 경기장이 있습니다. [Hay un estadio de 알파인 en.]

- 봄과 가을은 시원하고 여행하기에 좋습니다. [La primavera y el otoño son buenas para viajar y viajar.] : viajar(여행하다) 2번 번역

- 횡계는 소고기가 유명합니다. [La carne de quebec(Hoenggye) es famosa por la carne.] : 횡계 번역 누락

○ 잘못된 번역 – 다른 단어 또는 의미로 번역되는 경우

- 아이스하키 경기는 강릉 하키 센터에서 열립니다. [El partido de hockey de hielo se celebra en el centro de hockey de mbc.] : 강릉→mbc

- 차량 출입증이 있는 차량만 출입이 가능합니다. [Sólo se puede entrar con el coche con tarjeta de embarque.] : 출입증→보딩패스(embarque)

- 경기장 안에 수유실이 있습니까? [¿Hay alguna guardería en el estadio?]: 수유실→놀이방(guardería)

- 평창 동계 올림픽 슬로건은 패션 커넥티드입니다. [El embarcadero olímpico de pyeongchang es la moda de moda.] : passion → fashion(la moda) : '패션' 2개 이상의 의미가 있는 단어 처리 필요