

AI 기반 실시간 통역 보조 시스템 개발을 위한 예비 연구: 원문의 숫자 정보 제시를 중심으로*

최 문 선
(이화여대)

1. 서론

최근 기계번역을 비롯한 자연어 기술, 언어 처리 인공지능이 급속히 발전하면서 통역사, 번역사의 업무 흐름에도 자연스럽게 이러한 기술들이 통합되고 있다. 과거 모든 업무 단계를 통번역사가 직접 수작업으로 수행하였던 방식에서 벗어나 이제는 기술을 이용하여 생산성과 효율성을 높일 수 있는 단계를 중심으로 언어 기술이 적극 활용되고 있다. 언어 산업 전문 시장조사기관인 CSA 리서치에서는 언어 기술이 통합된 번역 프로세스를 ‘증강 번역(augmented translation)’이라고 지칭하고 있다(Lommel 2020). 번역 프로세스의 중심에서 번역을 수행하는 것은 번역사이지만 번역 과정에서 번역 메모리, 적응형 인공지능 경망 기계번역, 자동 콘텐츠 강화(automated content enrichment), 지능형 용어 관리와 같은 언어 기술을 활용하는 방식으로 번역 수행 모델이 진화하고 있음을 알 수 있다.

* 본 논문은 2022년 7월 5일 한국외국어대학교 통번역연구소에서 개최한 <통번역과 기술> 학술대회에서 연구자가 발표한 내용을 기초로 작성하였다.

통역의 경우 번역에 비해 시장 규모가 작고 산업화 정도가 낮기 때문에 번역 시장처럼 체계적이고 대대적인 기술 통합이 이루어지지 않은 상태이지만 통역사 개인 수준에서 다양한 언어 기술 활용 시도가 이루어지고 있다. 일례로 팬데믹 과정에서 RSI(remote simultaneous interpreting)로 대표되는 원격통역이 급속히 확산하면서 컴퓨터 프로그램을 활용하는 통역이 또 하나의 통역 서비스 표준으로 자리 잡았다. 또한 이동 제한 조치로 인해 대면 회의 참석이 어려워지면서 사전 제작 영상에 대한 통역 수요가 증가하였는데, 통역사가 영상의 언어 정보를 처리하는 단계에서 자막 추출, 자동 번역 등의 언어 기술을 적극적으로 활용하게 되었다(편혜원 2022).

이처럼 아직 걸음마 단계이기는 하나 기술은 통역 과정에 조금씩 통합, 활용되어 가고 있다. 하루가 다르게 발전하는 언어 인공지능을 통역 과정에 더 잘 활용함으로써 인지적 한계점에서 수행되는 언어적 곡예인 동시통역을 비롯한 여러 방식(mode)의 통역 수행을 더욱 수월하게 하면서 그 품질을 더 높일 수 있을가에 대한 고민과 상상이 필요한 시점이다. 이처럼 순수하게 통역사 개인에 능력에만 의존하는 데에서 나아가 “인간의 인지적 능력의 한계를 벗어나 더 증강된 통역 역량을 갖춘 상태에서 통역하는 것”, “통역사에게 추가 능력을 ‘장착’시켜 줌으로써 더욱 높은 품질의 통역을 훨씬 수월하게 수행하는 방식”(최문선 2022: 139)을 ‘증강 통역’이라는 용어로 지칭하고자 한다. 이는 단순히 개별 기술을 통역 과정에 적용하는 시도에 그치는 것이 아니라, 기술 보조를 기본으로 하는 방식이 통역의 새로운 패러다임으로 자리 잡을 수 있다는 전망을 의미한다.

기술을 통역 과정에 접목하고자 하는 시도는 새로운 것은 아니다. 지금까지 통역과 기술은 컴퓨터 보조 통역(computer-assisted interpreting 혹은 CAI)이라는 주제로 통역학 분야에서 오랫동안 연구되어 왔다(Braun 2019). 다만 지금까지 연구 개발된 대부분의 통역 보조 도구가 수작업 기반의 용어 관리 기능에 집중되어 왔다면(Ortiz and Cavallo 2018), 최근에는 언어 기술의 발달로 통역 과정에 통합되어 실시간으로 통역을 보조하는 자동화된 기술 도구의 개발 가능성이 가시화되고 있다는 점에서 근본적인 차이가 있다.

음성인식(automatic speech recognition 혹은 ASR)은 인간의 음성 발화를 자동으로 시각화해 줄 수 있고, 여기에 자연언어처리의 기법을 적절히 활용하면

컴퓨터가 고유명사나 전문용어, 숫자와 같이 통역사 기억에 부담을 유발하는 요소를 대신 인식, 기억해줄 수 있으며, 기계번역을 통해 원문에 대한 번역문이 생성될 수도 있다. 통역사의 능력을 갖춘 언어 지능을 만드는 일이 요원한 현재로서 우리가 기술 활용을 생각할 수 있는 방향은 컴퓨터의 능력으로 통역사의 능력을 증강시켜 통역을 수행하는 것이다. 일례로 휘발성 원문을 시각적으로 고정하여 제시함으로써 통역사가 기억에 할애하는 인지 자원을 줄이고 대신 도착어 생산에 주의를 집중함으로써 도착어 품질 개선을 도모할 수 있다. 혹은 낮은 억양을 구사하는 연사의 발화를 음성인식해 주는 프로그램은 통역사의 원문 이해 과정을 수월하게 하는데 도움을 줄 수 있다(Cheung and Tianyun 2018).

기술의 도움을 받아 수행하는 언어 전이 방식이 ‘증강 통번역’으로 지칭되고 ‘컴퓨터 보조’에서 ‘컴퓨터 증강’으로 언어 서비스가 전환되는 과정(Saina 2021)에서 통역에 결정적인 역할을 할 것으로 기대되는 것이 AI를 활용한 CAI 도구이다. AI를 ‘인공지능’이라 칭할 때 인간에 맞서는 위협적 존재라는 뉘앙스가 강하지만 사실 구체적인 내용을 뜯어보면 AI는 결국 기계에 의한 자동화와 크게 다르지 않다. 다만 최근의 기술 발전을 통해 자동화의 대상과 방식이 더 ‘뚝뚝’해졌다고 볼 수 있다. 언어 AI도 마찬가지이다. 언어 기술 발달에 따라 과거에 불가능했던 서비스가 가능해지면서 음성인식, 기계번역을 비롯한 다양한 자연어처리 기술이 인간의 언어 수행을 보조하는 것이 일상화되고 있다. 이러한 언어 AI를 통역 과정에 적용하여 통역에 특화된 기술과 도구를 개발하고자 하는 것이 AI 기반 통역 보조 기술 연구이다. 지금까지의 CAI 연구와 다른 점은 통역 과정에 특화하여 개발된다는 점과 자동화를 통해 통역을 실시간으로 보조한다는 점이다. 모형 시스템을 구축하여 실시간 통역 보조 기술의 가능성을 탐색(Desmet et al. 2018; Wang and Wang 2019)하거나 기존의 음성인식, 기계번역 상용 기술을 조합하여 통역의 일부 하위 단계를 자동화한 후 통역사 입장에서 사용성과 통역 품질 면에서의 개선 정도를 연구(이주리에 2021; Cheung and Tianyun 2018)하는 등 다양한 연구 노력이 이어지고 있다.

이러한 학술연구를 발판으로 현재 해외에서는 KUDO Interpreter Assist(Fantinuoli et al. 2022), SmarTerp(Rodríguez et al. 2021), CASSIS¹⁾ 등과

1) <https://www.alexexpert.hu/cassis/> 참고.

같이 상용 RSI 애플리케이션으로의 개발을 목표로 AI 기반의 실시간 통역 보조 도구에 대한 실제 개발 사례가 발견된다. 그러나 국내에서는 사실상 AI 기반 실시간 통역 보조 도구에 대한 연구 기반이 전무한 상황이다. 구체적인 도구 개발과 함께 맞물려 있는 여러 가지 다른 문제들도 있다. 이러한 도구가 실제로 통역사가 사용할 만하며 통역 과정에 도움이 되는지(사용가능성), 통역사에게 실질적인 도움이 되기 위해서 갖추어야 하는 요건은 무엇인지(기술적 요구사항), 실제로 통역 결과 향상 효과가 있는지(통역 품질 개선 정도), 그리고 이러한 도구를 활용하는 증강 통역 능력 함양을 위한 교육훈련은 어떻게 설계되어야 하는지(교육 훈련)에 대한 질문이 제기될 수 있다. 그러나 이러한 문제에 대한 탐구를 위해서는 기술 도구 개발이 선행되어야 한다.

이러한 배경하에 본 연구에서는 AI 기반 실시간 통역 보조 시스템의 가능성을 시험하고 전술한 다양한 관련 문제에 대한 해답을 찾기 위한 기본 전제인 시스템 구축 가능성을 확인하기 위하여 프로토타입을 구축하였다. 본고를 통해 해당 프로토타입의 구축 과정과 결과물 성능을 보고하고 후속 연구 방향을 제시하고자 한다. 이어지는 2장에서는 CAI 연구의 최신 주제인 실시간 통역 보조 시스템 관련 연구개발 동향을 살펴보고, 3장에서는 AI 기반 실시간 통역 보조 시스템의 구성요소와 작동 방식, 구축된 시스템의 성능 시험 설계를 기술한다. 4장에서는 시스템 구축 결과와 성능 시험 결과를 제시한 후 이에 대해 논의하며, 5장에서 결론을 짓고 본고를 마무리하고자 한다.

2. 실시간 통역 보조 시스템 관련 연구개발 동향

통역 과정에 기술을 적용함으로써 통역 준비 과정의 효율성과 생산성을 높이고 통역 수행 및 결과물의 품질을 개선하고자 하는 노력은 꾸준히 이어져 왔다. 지난 20여 년간 특히 유럽 통역계의 주도하에 유럽 언어 중심의 CAI 도구가 개발, 상용화, 활용되어 왔다(Corpas Pastor 2018; Will 2020). 기존의 CAI 도구의 특징은 통역의 준비와 사후 관리 단계에 집중되어 있어(최문선 2022: 135) 통역이 실시간으로 진행되는 과정을 지원하지 못한다는 점이다. 예를 들어 회의 주제 관련 문서 수집(Fantinuoli 2017a), 코퍼스 기반 용어 추출(최문선

2018, 2019; Fantinuoli 2006), 코퍼스를 활용한 회의 관련 주제 탐색(Xu 2018) 등 통역의 준비 단계에서 이루어지는 언어적, 언어 외적 지식 축적 과정을 효율화하고 생산성을 높이기 위한 자동화 혹은 반자동화 접근법들에 관한 연구가 이루어졌다. 이러한 배경하에 통역 교육이나 통역 실무 현장에서 통역사들이 사용할 수 있는 상용 도구(Ortiz and Cavallo 2018: 17-18)의 개발이 이어졌으며, 대부분이 통역 과정 전, 후의 단계를 지원하는 기능을 제공한다.

그러나 최근 음성인식, 텍스트 정보 추출, 기계번역 등 다양한 자연어 처리 기술이 비약적으로 발전하면서 통역이 진행되는 과정에 실시간으로 기술이 결합되어 통역사를 보조할 수 있을 만큼의 기술 성숙이 이루어졌다. 가장 단순하게는 음성인식을 통해 연사의 음성을 텍스트화하여 시각적으로 통역사 앞에 제시할 수 있다. 이론적으로만 본다면 원문이 텍스트로 제시될 때 통역사는 동시 과제 수행의 고질적인 고충인 원문 누락을 피할 수 있다. 음성인식이라는 기술을 이용하면 순차통역의 노트테이킹 과정에서, 동시통역 시 발화중복이나 청성 시차(EVS) 등의 문제로 인해 누락되는 정보를 줄임으로써 통역 품질의 개선을 기대할 수 있다(이주리에 2021; Wang and Wang 2019). 통역 고충 중의 하나인 전문용어의 통역 문제에도 상당한 도움을 기대할 수 있다. 원문에 등장하는 전문용어를 추출하여 제시해주거나 한 걸음 더 나아가 추출된 전문용어의 도착어 대응어도 함께 제시해준다면 통역 과정에서 깊은 프로세싱을 요하지는 않지만 기억력과 즉각적 정보 인출에 많은 인지 자원이 소요되는 전문용어의 코드변환 작업을 수월하게 처리할 수 있을 것이다. 연사의 발화로부터 통역에 도움이 될 만한 정보를 시각적으로 제시함으로써 실시간으로 통역 과정을 보조하는 시스템은 통역 준비 시간을 단축하고 고도로 전문적인 분야의 통역에서 정확도를 향상하는 데 도움을 줄 수 있다.

기술 성숙도에 더하여 최근 코로나 팬데믹 이후 RSI로 대표되는 원격통역의 급속한 확산이 기술 적용에 유리한 토양을 제공하고 있다는 점 역시 중요한 요인이다. 통역 부스와 같은 물리적인 통역 장소가 아닌 컴퓨터 애플리케이션이라는 가상 공간으로 통역의 현장이 이동하면서 다양한 기술을 통역 환경에 접목하기 용이한 환경이 조성된 것이다. 이에 더하여 원격통역 환경의 특수성으로 인한 통역의 어려움이 가중됨에 따라 이러한 어려움을 경감시키고 통역 품질을 유지하기 위한 보조 기술 도입에 대한 통역사들의 기술 수용도가 높아

지는 여건도 조성되었다(Fantinuoli et al. 2022: 1). 통신 네트워크와 장비 안정성, 기기의 데이터 전송 용량 및 속도, 연사가 사용하는 마이크와 통역사가 착용하는 헤드셋 종류 등 다양한 기술적 변수로 인해 음성 수신이 불안정하기 때문에(장애리 2021: 219) 원문을 명료하게 청취하기 어렵거나 아예 음성 입력이 끊어지는 경우도 많아 통역 정확도를 유지하기가 훨씬 어렵다(최문선 2021: 305-306). 그뿐만 아니라 원격통역 환경하에서는 파트너 통역사와 물리적으로 동일한 공간에 위치하지 않는 경우도 있어 현장 부스 통역 시에 파트너 통역사에게 받는 도움을 더는 기대할 수 없게 된다(Interpretips 2020). 대면 시 비언어적 상황 정보를 바탕으로 공유할 수 있던 정보가 차단되고 통역 전후로 가능한 연사와의 의사소통도 불가능하며 연사, 청중의 반응 파악이 어려워지는 점 또한 원격통역 수행 시 추가되는 어려움이다(상우연 2020: 147). 이처럼 원격통역 환경에서는 대면 통역 방식보다 다양한 어려움이 추가적으로 발생한다. 따라서 음성인식, 자동번역 등의 첨단 언어 처리 기능을 갖춘 통역 보조 시스템을 통해 이를 해소하고자 하는 유인도 증가한다.

실시간 통역 보조 시스템은 필요한 개별 기술(모듈)을 연결하는 연계형(cascading) 접근법과, 전체 흐름을 하나의 시스템으로 구축하는 종단형(end-to-end) 접근법의 두 가지 접근법이 가능하다(Fantinuoli and Montecchio 2022: 1). 연계형의 경우 필요한 구성 기술은 음성인식을 통해 음성을 문자화하는 모듈, 문자화된 텍스트로부터 숫자나 전문용어, 고유명사 등 필요한 항목을 추출하기 위한 파싱(parsing) 모듈, 그리고 추출된 항목을 통역사에게 제시하기 위한 시각화 모듈이 필요하다(Fantinuoli and Montecchio 2022: 1). 각 자연어처리 구성 기술에 대한 연구가 오랫동안 진행되어 왔기 때문에 기계가 수행하기에 상대적으로 쉬운 과제들이며, 각 과제별 관련 데이터가 이미 방대하게 구축되어 있어서 AI 도구가 제대로 작동하기 위한 조건이 갖추어져 있다. 그러나 단계별로 수행되는 방식에서는 오류 전파(error propagation)가 발생할 수 있다는 점이 단점이다. 반면 종단형은 음성 입력에 대한 텍스트화에서 관심 항목 표시 그리고 도착어로의 번역까지 각 단계를 한 번(one shot)에 수행하는 방식으로서 하위 과제가 동시에 수행되어 오류 전파의 위험이 없고 모든 하위 과제가 텍스트 맥락(context)을 참고할 수 있으므로 결과물에서 높은 정확도를 기대할 수 있다. 각 단계를 연계할 필요가 없으므로 시스템 아키텍처가 덜 복잡하며 처

리 지연시간도 단축된다. 그러나 종단형 과제 관련 코퍼스가 구축되어 있지 않고 여러 가지 목표를 동시에 충족시켜야 하는 복잡한 과제를 수행한다는 점이 구현에 큰 어려움으로 작용한다(Rodríguez et al. 2022). 이러한 접근법을 구현하여 상용 시스템 개발도 일부에서 이루어지고 있다. 연계형 시스템으로는 미국의 KUDO Interpreter Assist(Fantinuoli et al. 2022), 종단형 시스템으로 유럽의 SmarTerp(Rodríguez et al. 2021)이 대표적 사례이다. 전자는 KUDO라는 기존의 미국 RSI 플랫폼에 추가되는 기능으로 개발되고 있으며, 후자는 자동 통역 보조 기능을 갖춘 RSI 플랫폼을 새롭게 개발하는 유럽 프로젝트이다.

지금까지의 CAI 연구 및 활용은 주로 유럽을 중심으로 이루어져 왔으며 국내에서 연구나 개발이 이루어진 CAI 도구는 전무하다. 전술한 상용화 직전 단계의 두 시스템 역시 우리나라에 거의 알려지지 않았다. 국내에서는 아직 CAI 도구를 통역 과정에 통합하는 증강 통역의 개념이 낯설고, 이를 구현하기 위한 실제 시스템 개발 및 실험 연구 등도 전혀 이루어지고 있지 않은 것으로 파악된다. 이는 국내 다른 분야에서 AI를 비롯한 첨단 기술이 빠르게 도입되고 있는 것과 매우 대조적이다. 현재 우리 사회 다양한 분야의 업무 현장과 일상생활에서도 AI가 적용될 정도로 기술 성숙도가 높아졌고 이를 통한 다양한 기회와 도전 과제가 제기되고 있다. 이러한 점은 통역 현장에도 기술 결합에 대한 압력으로 작용하고 있으며 기술 활용을 통한 증강 통역의 가능성에 대한 탐색이 필요하다.

3. 연구 설계: 숫자 정보 제시를 위한 실시간 통역 보조 시스템 프로토타입 구축

언어 기술을 통합, 적용하는 통역 보조 시스템 구축을 위해서는 컴퓨터 프로그래밍, 언어 AI에 대한 전문지식이 요구되어 본 연구 과정에서 AI 개발 전문가²⁾와 협업을 진행하였다. 통역 전문가인 연구자가 관련 선행연구 조사, 시

2) 본 연구의 기술 자문 및 개발은 에스넷시스템 Applied AI팀의 배은구 부장이 수행하였다.

시스템 구상 및 요구사항 도출과 시스템 테스트 및 성능 분석을 담당하였으며, 시스템 아키텍처 구성 및 개발은 통역 전문가와의 협의하에 AI 개발 전문가가 수행하였다.

통역 보조 시스템의 목적은 언어 기술을 이용하여 통역의 일부 하위 단계를 자동화함으로써 통역사가 경험하는 어려움을 완화 혹은 해소시키는 것이다. 켈(Gile 2009: 171)은 고유명사, 숫자, 나열, 발화 속도, 낯선 억양, 논리 불분명, 열악한 음질을 통역 시의 문제 유발 요인(problem trigger)으로 지목하였다. 이러한 문제 유발 요인이 등장하면 통역사가 가용한 인지 용량을 초과하는 상황이 발생하여 주의력 관리에 난항을 겪거나 원문의 일부분에 대해 순간적으로 주의력이 상실된다(ibid.). 특히 숫자의 경우 맥락으로 유추하기 어려워 숫자 자체를 인식, 기억하고 도착어로 전환하는 데에 많은 부하가 걸리는 작업이며, 이로 인해 뒤에 이어지는 부분에 대한 통역에서 누락이나 오류 등 부정적인 영향이 발생한다는 점이 더 큰 문제로 작용한다. 숫자로 인한 통역에서의 어려움은 교육훈련 과정 중에 있는 통역 학생들에게서 두드러지게 나타나는 문제이기도 하다. 숫자의 빠른 전환을 위해 별도의 숫자 통역 훈련을 하기도 하는 등 현재 통역 교육 과정의 교수학습, 자기훈련, 평가 등에서 숫자의 정확한 통역은 상당히 큰 비중을 차지하고 있다. 그럼에도 불구하고 다수의 실험 연구 결과에서 숫자 통역의 정확도는 30%에서 70%에 이르며, 특히 학생 통역에서 숫자 통역 오류가 높게 나타난다(Defrancq and Fantinuoli 2021: 75). 반면 부스에서 파트너의 도움을 받을 수 있는 실제 통역 상황에서는 숫자 통역 오류율이 21%에 불과한 것으로 나타난 연구 결과(Collard 2019)를 볼 때 원문의 숫자 정보를 파악, 기억, 통역하는 데 외부의 지원을 활용할 수 있는 환경에서 통역 품질의 개선을 기대할 수 있다. 이러한 배경하에 본 연구에서는 실시간 통역 보조 시스템 구축의 첫 번째 단계로서 숫자의 통역을 지원하는 통역 보조 시스템의 프로토타입을 구축하였다.

3.1 시스템 설계

통역사에게 원문 중 숫자 정보를 제시하는 방식과 관련하여 선행연구(이주리에 2021; Wang and Wang 2019) 및 연구자를 비롯한 통역사들의 통역 현장

경험에 기초할 때 범용 음성인식 프로그램을 사용하여 원문을 있는 그대로 전사(transcription)하여 제시하는 방법은 사실상 하나의 입력물(input)이 추가되는 것과 마찬가지로 그 자체로서 인지 부하를 크게 증가시킨다. 특히 실시간 전사물의 경우 마침표도 없고 단락 구분도 없이 단어가 연속적으로 나열되는 줄글 형식이므로 통역사가 필요한 정보를 찾는 과정에 시간과 인지적 처리 노력이 소요되어 즉각적 정보 인식과 처리를 요하는 통역에서 의미있는 활용을 기대하기 어렵다. 따라서 인지 부하를 최소화하기 위해서는 숫자를 추출하되 통역사가 신속히 파악할 수 있도록 구분하여 제시할 필요성이 제기된다.

추출한 정보를 통역사가 한눈에 파악할 수 있도록 제시하는 방법은 크게 두 가지로 구분할 수 있다. 우선 숫자 정보만 추출하여 최대한 간단한 방식으로 제시하는 것이다. 유럽에서 개발 중인 한 숫자 통역 지원 시스템의 사용자 인터페이스 관련 설문조사³⁾에 따르면 통역사들은 대화상자(dialogue box)같은 심플한 디스플레이로 원문에 등장하는 숫자가 차례대로 제시되는 방식을 선호하였다. 이처럼 숫자만 제시하는 디스플레이 방식은 정보 입력량을 최소화함으로써 정보 처리 부하를 낮춰주는 장점이 있다. 하지만 원문에 여러 개의 숫자가 연이어 나오거나 여러 가지 이유로 청성시차가 길어지는 경우 주변 문맥이 함께 제시되지 않으면 제시된 숫자를 통역에 활용하기 어렵다. 한편 전사 텍스트 전체를 제시할 경우 정보 입력량은 증가하나 맥락을 함께 고려함으로써 숫자 정보의 모호성을 해소할 수 있으므로(Defranq and Fantinuoli 2021: 78) 전사 텍스트 전체를 제시하되 필요한 숫자 정보의 위치를 신속히 확인할 수 있도록 숫자만 다르게 표시하는 등 구분하여 제시하는 방법도 가능하다. 이러한 장단점을 고려하여 시스템이 숫자 정보만 제시하는 방식과 원문 전체를 제시하는 두 가지 방식으로 구상하였다.

우선 음성으로 입력되는 원문 정보에서 숫자를 추출하기 위해서 음성입력을 텍스트로 전환하는 ASR 프로그램과 전사된 텍스트에서 숫자만을 추출하는 기능을 수행할 자연어처리 AI가 요구되었다. 동시통역의 조건을 고려할 때 음성의 텍스트 전환이 실시간으로 이루어져야 하므로 이것이 가능하면서 개발 및 고도화 단계에서 프로세스 및 서비스 방식의 변화에 유연하게 대응할 수 있는

3) <https://www.eabm.ugent.be/survey/> 참조.

API 방식으로 서비스를 제공할 수 있도록 구글의 Google Speech-to-Text(STT)를 선택하였다. 다음으로 텍스트에서 숫자 정보를 추출하기 위해서 텍스트 분석과 개체명 인식(named entity recognition 혹은 NER)을 위한 다양한 파이썬 기반 자연어처리 라이브러리 중 성능이 우수하고 특히 시각화 기능이 탁월한 spaCy를 이용하였다. 숫자 정보 추출을 위해 spaCy가 제공하는 총 18개의 NER 태깅 레이블 중 날짜(date), 퍼센트(percent), 화폐(money), 기수(cardinal) 등 4개의 태그를 선택하였고, 내장된 시각화 도구인 displaCy를 이용하여 태깅된 관심 항목을 서로 다른 색깔로 구분하여 표시하였다. Google Speech-to-Text API와 spaCy를 핵심 라이브러리로 하는 시스템 아키텍처를 도식화하면 다음의 <그림 1>과 같다.

<그림 1> 숫자 통역 보조 시스템 아키텍처



<그림 1>의 각 단계에 대한 설명은 다음과 같다.

- ①단계 : 마이크를 통해 입력된 음성데이터를 파일로 전환한다. (Python 라이브러리 PyAudio를 이용)
- ②단계 : PyAudio가 스트리밍된 음성을 받아 파일로 처리하면, 이 파일을 Google Speech-to-Text API에 요청(Request)으로 보낸다.
- ③단계 : Google Speech-to-Text 모델은 요청값에 대한 결과 텍스트를 반응값으로 보낸다.
- ④단계 : Google Speech-to-Text API에서 반환된 반응값의 텍스트를 spaCy

의 개체명 인식 기능을 이용하여 태깅한다. 이를 위해 spaCy의 기본 NER 태그 중 통역에 필요한 숫자와 관련된 값만 추출하여 화면에 표시한다.

⑤단계 : 웹 또는 앱 서비스를 통한 사용자 인터페이스(UI)를 개발하여 통역사에게 숫자 등 원문의 관심 정보를 제시한다.

본 연구에는 사용자 인터페이스를 갖춘 완성된 서비스를 구축하는 마지막 단계(<그림 1>의 ⑤단계)를 제외하고 4단계까지만 구현하였다. 별도의 UI가 개발되지 않았으므로 숫자 제시 화면으로는 웹 브라우저에서 코드를 실행하고 결과값을 제시해주는 Jupyter Notebook을 이용하였다.

숫자만 반환하는 방식과는 달리 원문 전체를 제시하는 방식에서는 숫자의 즉각적 위치 확인을 수월하게 만들어 주어야 했으며, spaCy의 탁월한 시각화 기능이 효과를 발휘하였다. 그러나 시각화가 문장 단위로 일괄 처리(batch processing)되다 보니 원문 발화와 시각화 사이에 문장 완성까지의 시차가 발생하는 문제가 발견되었다. 이에 따라 시차가 발생하더라도 뛰어난 시각화로 정보 확인이 용이한 완성문장 반환 방식을 그대로 구현하되, 이에 더하여 실시간 원문 참조가 가능하도록 추후 수정이나 오류 가능성이 있는 상태인 전사 중간 결과값을 실시간으로 보여주고 문장 종료와 동시에 시각화가 추가된 문장을 반환해주는 방식도 함께 구현하였다. 프로토타입 시스템으로 구현한 세 가지 방식을 요약하면 다음과 같다.

〈표 1〉 프로토타입 시스템 구현 방식

버전	구현 방식
1	숫자만 반환
2	시각화 완성 문장 반환 (spaCy의 시각화 기능을 이용하여 원문 중 숫자를 색깔로 표시)
3	전사 중간 결과 + 시각화 완성 문장 반환 (실시간 자막 + 색깔 표시 완성 문장)

3.2 시스템 성능 시험 설계

시스템의 숫자 추출 및 제시 성능을 시험하기 위하여 다양한 유형의 숫자가 다수 포함된 원문 텍스트를 선정하였다. 호주 모나슈 대학 총장의 고등교육

관련 연설문⁴⁾(Monash University 2016)과 코로나 백신 관련 보고서 중 요약문⁵⁾(Tony Blair Institute for Global Change 2021: 3)에서 숫자가 등장하는 부분을 발췌하여 아래와 같은 실험 텍스트를 준비하였다.

<그림 2> 성능 시험에 사용한 텍스트

In Australia the most recent policy manifestation of this trend was the introduction of the demand-driven system with its target of **some 40% of the 25-34 year olds** in the population having a bachelor's degree. **Since 1990** the proportion of men with a bachelor's degree more than doubled **from 9% to 24% and for women from 6.8% to 29%.**

Australian universities contributed some **1.5%** of GDP to our economy, educated **1.3 million students** and produced some **300,000 graduates** in the past year.

In **May**, Malawi destroyed almost **20,000 doses** of the AstraZeneca vaccine. In **June**, South Sudan destroyed **59,000 doses and handed back 72,000 of the 132,000 doses** it received through COVAX. The Democratic Republic of the Congo was unable to use most of **the 1.7 million** AstraZeneca doses it received under COVAX and was forced to redistribute them. Indeed, to fully account for incoming vaccine supply from COVAX and other bodies to the countries most in need of vaccines, our modelling shows that a global increase in their absorptive capacity **from 170 million doses to 410 million** doses per month will be needed by the end of **2021**.

원문에서 추출되어야 하는 숫자 항목은 <그림 2>에서 색깔로 표시한 바와 같이 총 21개로 확인되었다(40 / 25 / 34 / 1990 / 9 / 24 / 6.8 / 29 / 1.5 / 1.3 million / 300,000 / May / 20,000 / June / 59,000 / 72,000 / 132,000 / 1.7 million / 170 million / 410 million / 2021). 영어 원문에서는 숫자로 표시되지 않지만 한국어로 통역 시 숫자로 변환되어야 하는 날짜 정보인 'May', 'June'도 추출 대상에 포함하였다. 본문 중 밑줄과 볼드체로 표시한 부분은 특히 숫자가 밀집되어 통역 시 인지 부담이 증가하고 숫자 통역 오류 가능성이 높은 부분으로서, 실제 이러한 부분에서 실시간 통역 보조 시스템이 유용하게 활용될 수 있을 것으로 보았다.

성능 시험을 위해 연구자의 노트북으로 3.1절에서 기술한 바와 같이 구성된 시스템을 동작하였다. 준비된 텍스트를 연구자가 육성으로 약 120wpm의 속도

4) <그림 2>의 첫 번째 문단.

5) <그림 2>의 두 번째 문단.

로 낭독하여 노트북으로 입력하였다. 이때 실시간으로 숫자 정보가 추출되어 노트북 화면에 제시되었고, 화면 녹화를 통해 결과물을 기록하였다.

4. 결과 및 논의

4.1 시스템 구축 결과

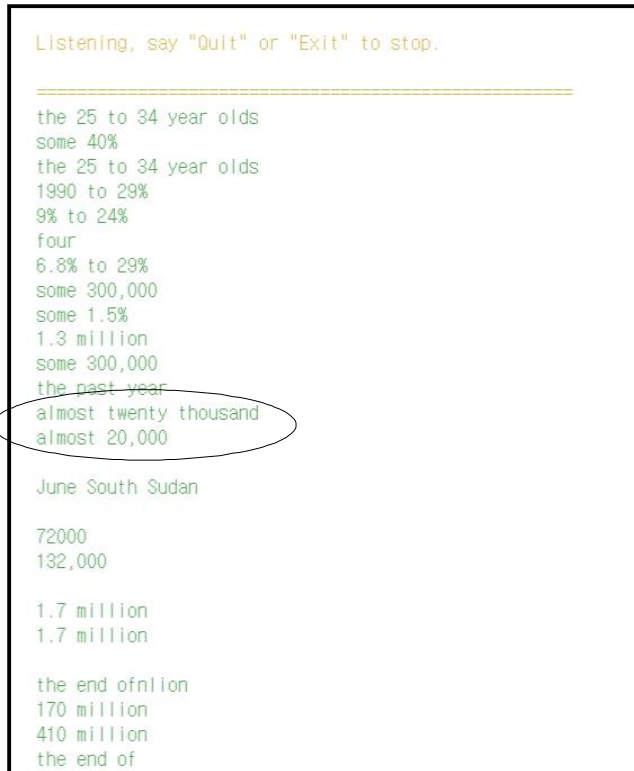
세 가지 버전으로 시스템을 구현한 결과는 다음과 같다.

4.1.1 숫자만 반환

원문 음성이 입력되는 동시에 원문 중 숫자 정보만 순서대로 위에서 아래로 대화상자에 표시되었다. 실시간으로 출력되는 과정에서는 중간 결과값이 빨간색으로 표시되었다가 한 문장 발화가 끝나면 최종 결과값이 초록색으로 표시되어서 신, 구 항목 구분이 가능했다. 연사 발화 종료 시 남은 최종 화면은 <그림 3>과 같다.

숫자만 반환하는 방식의 경우 다른 정보가 한꺼번에 제시되지 않아 통역 시 인지 부담을 최소화할 것으로 기대하였으나, 구현 과정에서 몇 가지 문제가 발견되었다. 일부 항목의 경우 실시간 출력값과 최종 결과값이 모두 화면상에 제시되어 동일한 정보가 두 번 반복되었다. 예를 들어 <그림 3>의 타원형 표시 부분을 보면 실시간 출력 시 ‘almost twenty thousand’라고 알파벳으로 표기되었다가, 해당 문장 발화가 종료되면 동일한 항목을 다시 숫자로 변환하여 ‘almost 20,000’라고 출력하였다. 또한 실시간 출력 과정에서는 먼저 출력된 값 위에 다음 값이 중첩되거나 이전 값이 삭제된 후에 다음 값이 출력되는 과정이 불안정적이었다.

〈그림 3〉 숫자만 반환된 화면



4.1.2 시각화 완성 문장 반환

다음으로 spaCy의 시각화 기능을 적용하여 완성된 문장을 반환하는 방식을 구동하였다. 새로운 문장이 시작될 때 개행이 이루어져 문장 구분 없이 연속되는 일반적인 전사 프로그램에 비하여 문장 구분이 용이하였고, 이는 관심 정보 검색이 신속히 이루어질 수 있음을 의미하였다. 주어진 문장 내에서 숫자 정보 역시 색깔 표시를 통해 즉각적으로 알아보기 쉽게 제시되었다.

〈그림 4〉 시각화 완성 문장 반환 화면

since 1990 DATE the proportion of men with a bachelor's degree more than doubled from 9% to 24% PERCENT and four CARDINAL women from 6.8% to 29% PERCENT

Australian universities contributed some 1.5% PERCENT of GDP to our economy educated 1.3 million CARDINAL students and produced some 300,000 CARDINAL graduates in the past year DATE

in May Malawi destroyed almost 20,000 CARDINAL doses of the AstraZeneca vaccine

in June South Sudan DATE destroyed 59000 Doses and hand it back 72000 of the hundred and thirty-two thousand Doses and received through Colfax

the Democratic Republic of Congo was unable to use most of the 1.7 million CARDINAL AstraZeneca doses it received under Kovacs and was forced to redistribute them

그러나 <그림 4>에서 볼 수 있듯이 이처럼 완성도 높은 처리 결과를 반환하는 버전의 경우 한 문장 발화가 끝날 때까지 기다려야 한다는 문제가 있었다. 따라서 실시간 처리가 요구되는 동시통역에서는 적용하기가 사실상 어렵고 연사 발화를 마친 후 통역하는 순차통역에의 적용 가능성이 높을 것으로 판단되었다.

4.1.3 전사 중간 결과 + 시각화 완성 문장 반환

세 번째 방식은 문장 전사 결과를 실시간으로 제시하되 해당 문장의 발화 종료 즉시 시각화가 완성된 문장(<그림 5>의 볼드체)가 이어서 제시되었다. 이 경우 중간 결과값인 실시간 전사물도 참조하면서 문장 종료 시 시각화된 문장에서 필요한 정보를 재확인할 수 있었다. 이처럼 실시간 전사 결과와 시각화 완성 문장을 모두 반환해 주는 방식은 통역 시 요구되는 즉시성을 충족시키면서도 통역사가 참조하고자 하는 정보의 맥락을 함께 제시하므로 필요한 정보 위치를 보다 신속히 찾기가 용이하다는 두 가지 장점을 제공하였다. 특히 어순 차이가 있는 언어조합의 동시통역에서 청성시차가 길어질 때에는 원문 중 숫자 정보가 즉시 제시되더라도 활용하기 어려울 수 있으므로 이때 정보의 주변 맥락, 즉 전체 문장이 함께 제시되면 놓쳐버린 숫자 정보를 환기하는 데에 도움을 줄 수 있을 것으로 예상된다.

어순 차이가 있는 언어조합의 동시통역에서 청성시차가 길어지더라도 필요한 정보를 신속히 검색하여 맥락을 고려한 의미 확정을 통해 통역 정확도가 높아질 수 있을 것으로 기대되었다.

〈그림 5〉 중간 결과값과 시각화 완성 문장 함께 반환

Australian universities contributed some 1.5% of GDP to our economy educated 1.3 million students and produced some 300,000 graduates in the past

Australian universities contributed some 1.5% PERCENT of GDP to our economy educated 1.3 million CARDINAL students and produced some 300,000 CARDINAL graduates in the past year DATE

4.2 시스템 성능 시험 결과

원문에 포함된 숫자 총 21개에 대해서 정밀도⁶⁾와 재현율⁷⁾을 산출하였다. 정밀도와 재현율은 분류 성능 평가 기준으로서 AI와 NLP를 이용하여 문서를 분석하여 원하는 정보를 반환해주는 소프트웨어의 품질 평가, 검색 엔진 등의 성능 평가 등에 사용되며 기본적으로 원하는 정보를 인출하는 능력을 평가하기 위해 적용한다. 이를 위하여 4.1절의 시스템 구동 결과 나온 출력값⁸⁾을 아래와 같이 분류하였다.

〈표 2〉 시스템 성능 시험 결과

		실제 정답(총 21개)	
		True	False
분류 결과	True	[True positive] 18개 (40 / 25 / 34 / 1990 / 9 / 24 / 6.8 / 29 / 1.5 / 1.3 million / 300,000 / 20,000 / June / 72,000 / 132,000 / 1.7 million / 170 million / 410 million)	[False positive] 4개 (four / past year / South Sudan / the end of)
	False	[True negative] 3개 (May / 59,000 / 2021)	[False negative] (해당 없음)

- 6) 정밀도(precision)란 모델이 True라고 분류한 것 중에서 실제 True인 것의 비율이며, True positive / (True positive + False positive)로 계산한다. 오류 발생 정도를 표현한다.
- 7) 재현율(recall)이란 실제 True인 것 중 모델이 True라고 예측한 것의 비율이며, True positive / (True positive + True negative)로 계산한다. 누락 발생 정도를 표현한다.
- 8) 출력값 전체는 <그림 3> 참조.

결과값 분류 결과 인출되어야 할 숫자 항목 총 21개 중 18개가 올바르게 인출되었다. 한편 인출을 의도하지 않았던 오류 항목이 4개 발생하였다. 원문의 ‘for’를 ‘four’라는 숫자 정보로 잘못 인식하여 출력되었고, ‘past year(작년)’는 한국어로 통역할 때 숫자 정보로 처리하지 않을 것이라 판단하여 인출 의도 대상(정답)에 포함하지 않았으나 spaCy의 날짜(date) 태깅 대상에 포함되어 인출되었다. ‘South Sudan’은 직전 항목인 ‘June’이 인출되면서 같이 태깅된 결과로 보이며, ‘the end of’는 숫자 정보가 아님에도 불구하고 인출되어 총 4건의 오류가 발생하였다. 반면 인출되었어야 하는 항목 중 ‘May’, ‘59,000’, ‘2021’이 누락되었다. 이에 따라 시스템 성능 시험 결과 정밀도 0.82, 재현율 0.90을 기록하였다.

4.3 논의

음성인식과 자연어처리 기술을 적용하여 통역 원문 중 숫자 정보를 추출, 제시하는 실시간 통역 보조 시스템의 구축 가능성을 확인하기 위하여 4.2절에서 기술한 바와 같이 세 가지 방식으로 시스템을 구현한 결과 현존하는 기술 구성요소의 조합만으로도 통역 현장에 바로 적용하여 사용성을 시험해 볼만한 시스템이 구성될 수 있음을 확인하였다. 그러나 숫자 정보의 정확도, 숫자가 제시되는 방식, 원문과 숫자 정보 제시 간의 지연시간 등의 장애물이 남아 있어 동시통역, 순차통역 등 통역 모드별로 최적화된 시스템을 개발하기 위해서는 앞으로 많은 개선이 요구된다.

실시간 통역 보조 시스템이 실질적으로 통역사의 인지 부담을 경감시키고 통역 품질을 개선하는 효과를 거두기 위한 기본 전제는 우수한 시스템 성능이다. 특히 실시간 통역 보조 시스템 구성요소 중 성능을 결정짓는 요소는 음성인식과 개체명 인식이다. 음성인식의 정확도(word error rate 혹은 WER)와 처리 속도(real-time factor 혹은 RTF), 개체명 인식을 통한 관심 항목 추출의 정밀도와 재현율 등이 통역 보조 시스템의 일차적인 성능 지표로 적용될 수 있다. 본 연구에서는 범용 음성인식기 중 최고 성능을 갖춘 것으로 평가되는 Google Speech-to-Text라는 음성인식기를 사용하였으며 이를 성능 시험용 텍스트에 적용한 결과 최근 새로 설립된 기관명(예: COVAX) 등 일부 고유명사를 제외하

고 거의 100%에 가까운 정확도를 보였다. 개체명 인식을 위해 spaCy가 제공하는 자연어처리 엔진을 그대로 이용한 결과 숫자 인식 역시 매우 우수한 성능을 보였다. 불필요한 오류 항목이 제시될 때 통역사 주의집중을 방해하는 치명적 결과를 초래하므로 통역 보조 시스템의 성능 지표로서는 재현율(recall)보다 정밀도(precision)가 더 의미가 있다(Fantinuoli 2017b: 30). 본고에서 구축한 시스템의 경우 앞서 4.2절에서 보고한 바와 같이 정밀도(0.82)가 재현율(0.90)보다 다소 낮아 추가적인 최적화가 요구되는 상태로 판단되었다. 이에 더하여 본 연구에서 적용된 바와 같은 통제된 음성 입력이 아니라 빠른 발화속도, 불명확한 발음이나 비표준적 억양 등 다양한 현장 상황에 대해서도 강건성을 갖추어야 성능의 안정성이 확보될 수 있다.

시스템 성능에 더하여 통역사의 수용성과 사용가능성 역시 실시간 통역 보조 시스템이 효과적으로 사용되기 위해 충족되어야 하는 조건이다. 향후 개선과 맞춤화를 통해 우수한 성능을 갖춘 통역 보조 시스템이 구축된다고 하더라도 이것이 통역사에게 과도한 인지 부담으로 작용한다면 통역사들은 실제 통역 상황에 사용하지 않을 것이기 때문이다. 통역 보조 시스템의 사용가능성을 결정짓는 가장 중요한 요소는 비용 대비 효과이다. 지금까지 기술을 활용한 용어집 구축과 활용(최문선 2018; Fantinuoli 2017a; Xu 2015), 디지털 펜과 태블릿 이용한 순차통역(Goldsmith 2018), 원문을 녹음해서 채청취함으로써 노트테이킹과 기억력 부담을 줄이고자 한 동시적 순차(Hamidi and Pöschhacker 2007; Orlando 2014) 등 컴퓨터 보조 통역의 다양한 시도들이 기술적 구현 가능성이 증명되었음에도 불구하고 지금까지 국내 통역 현장에서 이를 활용하는 통역사는 많지 않았다. 통역 현장에서의 적용이 어려웠던 가장 큰 이유는 이러한 기술 도구들이 제공하는 혜택에 비해 이를 활용하기 위한 ‘비용’이 과도하게 요구되었기 때문이다. 통역사들에게 낯선 기술적 지식을 갖출 것을 요하거나 추가적인 시간이 소요되며 무엇보다 실시간으로 이루어지는 통역 과정에 통합되지 못한 채 오히려 통역 과정을 방해하는 부담스러운 요소로 여겨졌기 때문이다. 따라서 본 연구에서처럼 기술적 구현이 확인된 실시간 통역 보조 시스템이 통역사들에 의해 수용되고 활용되어 진정한 의미의 ‘증강 통역’으로의 패러다임 전환이 가능하기 위해서는 통역사들이 해당 기술을 사용하는 비용 대비 효과가 극대화되어야 한다.

비용을 최소화한다는 것은 통역사들이 지출해야 하는 ‘인지 자원’ 면에서의 비용을 의미한다. 즉 통역사들이 통역하는 과정에서 추가적인 인지 자원이 요구되지 않거나 아주 적은 양의 추가 자원만 요구되어 효과에 의해 상쇄될만한 수준이어야 한다. 복수의 과제가 동시에 수행되어야 하는 동시통역은 그 자체 만으로도 이미 매우 높은 인지부하를 유발한다. 시버(Seeber 2007)의 동시통역 인지부하모델에 따르면 동시 수행 과제 사이의 상호작용은 새도잉이 6.5, 문장 구역이 8.8, 동시통역은 9 정도의 간섭을 일으키는 것으로 나타나 동시통역이 수행 과제 간의 간섭이 가장 심한 것을 볼 수 있다. 이에 더하여 엑셀이나 워드 파일 등 통역사가 구축한 용어집에서 용어를 검색하는 순간에는 이론적인 간섭 점수가 16.8까지 높아질 수 있다(Prandi 2017: 83). 향후 숫자를 비롯하여 전문 용어나 이에 대한 번역어를 지원하는 등 실시간으로 통역을 보조하는 CAI 도구가 통역 과정에 효과적으로 결합되는 증강 통역이 실제 효과를 거두기 위해서는 CAI 도구를 추가함으로써 발생하는 간섭의 정도가 현재의 통역 방식에서 발생하는 간섭보다 크게 높아져서는 안 된다. 이러한 측면에서 본고에서 구축한 세 가지 버전의 숫자 보조 시스템은 통역사가 컴퓨터를 직접 조작할 필요가 없는 자동화된 시스템이라는 점에서 수작업-공간적 과제(manual-spacial task)는 개입되지 않으나 필요한 정보의 위치를 시각적으로 찾기 위한 시각-공간적 과제(visual-spatial task)가 요구되는 정도를 고려하면⁹⁾ 언어적 맥락이 실시간으로 주어지는 환경에 가까운 버전 3이 가장 유리한 옵션이 될 것으로 가정할 수 있으며, 이는 향후 실제 통역사 평가를 통해 검증되어야 하는 부분이다.

5. 결론

최근까지 기술을 통역에 접목하여 효율성과 생산성을 높이려 했던 컴퓨터 보조 통역 연구와 개발 노력은 통역이 실시간으로 이루어지는 과정에서 통역사를 자동으로 지원할 수 없었다. 그러나 첨단 언어 기술이 개발되고 음성인

9) 통역 과정에서 통역사가 순간적으로 필요한 정보를 찾는데 요구되는 인지 자원의 종류에 대한 자세한 설명은 프랜디(Prandi 2017) 참조.

식, 기계번역 등 자연어처리를 위한 AI의 성능이 급속도로 향상됨에 따라 통역사가 통역 과정에서 수행하는 일부 하위 작업들을 자동화할 수 있을 정도의 수준에 다다르기 시작했다. 따라서 이러한 기술을 통역사가 활용하여 ‘증강’된 능력을 갖추고 통역한다면 지금까지와는 완전히 다른 종류의 통역 방식이 가능해질 것이다.

본 연구에서는 통역 원문에 등장하는 다양한 문제 유발 요소 중 숫자 정보의 통역에 도움을 줄 수 있는 실시간 통역 보조 시스템을 구상하고 실제 구현해 보았다. Google Speech-to-Text와 spaCy 등 음성인식과 개체명 인식 기능을 수행하는 기존 언어 AI 라이브러리를 조합하여 통역사에게 원문의 숫자 정보를 실시간으로 자동 추출하여 시각적으로 제시하는 시스템을 구축하였다. 통역사에게 가중되는 인지 부담을 최소화하기 위해 숫자만 추출하여 반환하는 시스템, 지연시간을 감수하고 숫자에 대한 시각화가 완성된 문장을 반환하는 시스템, 그리고 실시간 전사 과정을 보여준 직후 시각화가 완성된 문장을 반환하는 시스템 등 총 세 가지 버전으로 구축하였다. 숫자만 반환하는 시스템의 경우 단순한 레이아웃에 숫자 정보만 제시하므로 인지 부담을 최소화할 수 있으나 숫자가 화면상에 출력되는 과정이 불안정적이었다. 시각화 완성 문장을 반환하는 방법의 경우 최종 결과물은 가장 이상적이었으나 지연시간으로 인해 동시통역에는 적합하지 않고 순차통역에 한하여 적용할 수 있을 것으로 판단되었다. 마지막으로 실시간 자막인 중간 결과값과 시각화 완성 문장을 함께 반환하는 방법의 경우 맥락을 고려하면서도 숫자 정보의 위치 확인이 용이하게 구현되어 가장 이상적인 결과를 보였다.

본 연구는 국내에서 거의 유일하게 실시간 통역 보조 기술을 구현하고자 한 연구라는 점에서 중요한 의미를 지닌다. 그러나 성능 시험의 결과가 실제 통역 현장에서 통역사들에게 어느 정도 수용될 수 있을지에 대해서는 검증하지 못하였다는 점이 본 연구의 한계이다. 향후 실시간 통역 보조 시스템을 더욱 정교화하고 이에 대한 실제 사용자인 통역사들의 사용가능성 평가 연구가 이루어질 필요가 있다. 이를 위해서는 본 연구 범위에서 제외되었던 사용자 인터페이스 개발이 이루어져야 하며 인터페이스 개발자 등 공학 분야와의 학제 간 협력이 요구된다.

향후 증강 통역을 위한 통역 보조 시스템이 실제 현장에서 적용이 되기 위

해서는 시스템 개발과 함께 시스템 사용에 통역사들이 익숙해지는 과정이 필요하다. 실시간 통역 보조 도구의 사용 효과를 최대화하기 위해서는 주어진 시스템이 완벽하지 않더라도 통역사가 계속 사용하는 과정에서 시스템 개선을 위한 피드백을 축적해 나가고 통역 보조 도구를 사용하여 통역하는 방식에 익숙해져야 한다. 또한 통역 교육을 받는 첫 단계에서부터 기계, 즉 통역 보조 도구와 함께 통역하는 방법을 체득할 경우 통역 과정에 더 잘 통합될 수 있고 보다 높은 품질의 통역을 더욱 용이하게 수행할 수 있을 것으로 예상된다. 따라서 통역 교육 과정에 있는 초보 통역사들에게 기술을 적극적으로 활용하는 증강 통역의 방식을 교육한다면 실시간 통역 보조 도구의 활용에서 더 큰 효과를 기대할 수 있을 것이다.

본 연구에서 구현한 시스템은 통역 원문의 숫자 추출에 중점을 두었으나 향후 전문용어 추출 기능, 번역 대응어 제시 기능 등 더욱 확장된 기능을 갖춘 실시간 통역 보조 시스템이 개발된다면 의학이나 금융과 같이 전문용어 비중이 높은 회의에서 통역사의 인지 부담을 덜면서도 통역 품질을 더 높이는 데 큰 도움이 될 것이고 이는 통역사의 준비 과정과 통역 수행 과정에 대대적인 전환을 가져올 것으로 예상된다.

참고문헌

- 상우연 (2020) 「한일 원격통역 현황과 통역사 수요에 기반한 교육적 제언」, 『번역학연구』 21(5): 129-159.
- 이주리아 (2021) 「자동음성인식의 통역 활용에 관한 예비연구」, 『인문사회21』 12(5): 2407-2422.
- 장애리 (2021) 「원격통역 시대에 필요한 통역능력에 대한 소고」, 『통역과 번역』 23(1): 213-236.
- 최문선 (2018) 「전문분야 통역을 위한 용어집 구축 - 기술 도구의 적용가능성 고찰」, 『통역과 번역』 20(2): 163-191.
- 최문선 (2019) 「전문분야 통역을 위한 어휘 지식: IT 분야를 중심으로」, 『번역학연구』 20(3): 247-283.

- 최문선 (2021) 「통역 4.0: 원격통역의 규범과 실제」, 『번역학연구』 22(1): 279-312.
- 최문선 (2022) 「기계와 협업하는 인간 통역: 컴퓨터 보조 통역 (computer-assisted interpreting)의 최근 기술 동향」, 『통번역학연구』 26(2): 133-163.
- 편혜원 (2022) 「원격동시통역(RSI)의 속성과 통역사의 적응 - 슈로스버그의 전환모형을 중심으로」, 『통번역학연구』 26(2): 165-189.
- Braun, Sabine (2019) 'Technology and Interpreting', in Minako O'Hagan (ed.) *The Routledge Handbook of Translation and Technology*, New York: Routledge, <https://doi.org/10.4324/9781315311258-16>.
- Cheung, Andrew K.F. and Li Tianyun (2018) 'Automatic Speech Recognition in Simultaneous Interpreting: A New Approach to Computer-aided Interpreting', Ewha Research Institute for Translation Studies International Conference 2018, https://www.researchgate.net/publication/333457260_Automatic_speech_recognition_in_simultaneous_interpreting_A_new_approach_to_computer-aided_interpreting.
- Collard, Camille (2019) *A Corpus-based Study of Simultaneous Interpreting with Special Reference to Sex*, PhD dissertation, Ghent University.
- Corpas Pastor, Gloria (2018) 'Tools for Interpreters: The Challenges that Lie Ahead', *Current Trends in Translation Teaching and Learning E 5*: 157-182.
- Defrancq, Bart and Claudio Fantinuoli (2021) 'Automatic Speech Recognition in the Booth: Assessment of System Performance, Interpreters' Performances and Interactions in the Context of Numbers', *Target* 33(1): 73-102.
- Desmet, Bart, Mieke Vandierendonck and Bart Defrancq (2018) 'Simultaneous Interpretation of Numbers and the Impact of Technological Support', in Claudio Fantinuoli (ed.) *Interpreting and Technology*, Berlin: Language Science Press, 13-27.
- Fantinuoli, Claudio (2006) 'Specialized Corpora from the Web and Term

- Extraction for Simultaneous Interpreters’, in Marco Baroni and Silvia Bernardini (eds) *Wacky! Working Papers on the Web as Corpus*, 173-190.
- Fantinuoli, Claudio (2017a) ‘Computer-assisted Preparation in Conference Interpreting’, *Translation & Interpreting* 9(2): 24-37.
- Fantinuoli, Claudio (2017b) ‘Speech Recognition in the Interpreter Workstation’, *Proceedings of the 39th Conference Translating and the Computer*, 25-34.
- Fantinuoli, Claudio and Maddalena Montecchio (2022) ‘Defining Maximum Acceptable Latency of AI-enhanced CAI Tools’, *Computation and Language*, <https://arxiv.org/abs/2201.02792>.
- Fantinuoli, Claudio, Giulia Marchesini, David Landan and Lukas Horak (2022) ‘KUDO Interpreter Assist: Automated Real-time Support for Remote Interpretation’, *Computation and Language*, arXiv:2201.01800, <https://doi.org/10.48550/arXiv.2201.01800>.
- Gile, Daniel (2009) *Basic Concepts and Models for Interpreter and Translator Training*, revised edition, Amsterdam and Philadelphia: John Benjamins Publishing.
- Goldsmith, Joshua (2018) ‘Tablet Interpreting: Consecutive 2.0’, *Translation and Interpreting Studies* 13(3): 342-365.
- Hamidi, Miriam and Franz Pöchhacker (2007) ‘Simultaneous Consecutive Interpreting: A New Technique Put to the Test’, *Meta: Translators’ Journal* 52(2): 276-289.
- Lommel, Arle (2020, November 4) ‘Augmented Translation: Are We There Yet?’, *CSA Research Blog*, <https://csa-research.com/Blogs-Events/Blog/augmented-translation-2020>.
- Orlando, Mark (2014) ‘A Study on the Amenability of Digital Pen Technology in a Hybrid Mode of Interpreting: Consec-simul with Notes’, *Translation and Interpreting* 6(2), 9-54.
- Ortiz, Luis Eduardo Shild and Patrizia Cavallo (2018) ‘Computer-assisted

- Interpreting Tools (CAI) and Options for Automation with Automatic Speech Recognition’, *TradTerm* 32: 9-31.
- Prandi, Bianca (2017) ‘Designing a Multimethod Study on the Use of CAI Tools during Simultaneous Interpreting’, *Proceedings of the 39th Conference Translating and the Computer*, 76-88.
- Rodríguez, Susana, Roberto Gretter, Marco Matassoni, Alvaro Alonso, Oscar Corcho, Mariano Rico and Falavigna Daniele (2021) ‘SmarTerp: A CAI System to Support Simultaneous Interpreters in Real-time’, *Proceedings of the Translation and Interpreting Technology Online Conference*, 102-109.
- Saina, Francesco (2021) ‘Technology-augmented Multilingual Communication Models: New Interaction Paradigms, Shifts in the Language Services Industry, and Implications for Training Programs’, *Proceedings of the 1st Workshop on Automatic Spoken Language Translation in Real-World Settings (ASLTRW)*, 49-59.
- Seeber, Kilian (2007) ‘Thinking Outside the Cube: Modeling Language Processing Tasks in a Multiple Resource Paradigm’, *Interspeech 2007*, 1382-1385.
- Wang, Xinyu and Caiwen Wang (2019) ‘Can Computer-assisted Interpreting Tools Assist Interpreting?’, *Transletters: International Journal of Translation and Interpreting* 3: 109-139.
- Will, Martin (2020) ‘Computer Aided Interpreting (CAI) for Conference Interpreters. Concepts, Content and Prospects’, *ESSACHESS – Journal for Communication Studies* 13(25): 37-71.
- Xu, Ran (2015) *Terminology Preparation for Simultaneous Interpreters*, PhD dissertation, University of Leeds.
- Xu, Ran (2018) ‘Corpus-based Terminological Preparation for Simultaneous Interpreting’, *Interpreting* 20(1): 29-58.

<동영상 자료>

Interpretips (2020, August 10) RSI Webinar 1: An Overview of Remote Simultaneous Interpreting [Video], *YouTube*, <https://youtu.be/eeljvwNuLmE>
Rodríguez, Susana, Francesca Maria Frittella and Marco Gaido (2022, January 12) Bringing Machines & Humans Together for a Smart Interpreting: Introduction to SmarTerp AI-powered CAI Tool. TerpSummit 2022: Confidence Camp, <https://terpsummit.com>.

<인터넷 사이트>

Monash University (2016, June 8) 'Education in the Age of Disruption', <https://www.monash.edu/about/structure/senior-staff/president-and-vice-chancellor/profile/vice-chancellors-speeches/education-in-the-age-of-disruption>.
Tony Blair Institute for Global Change (2021, August) 'The Absorption-Capacity Challenge', <https://institute.global/sites/default/files/2021-08/Global%20Health%20Security%20Consortium%2C%20The%20Absorption-Capacity%20Challenge%2C%20August%202021.pdf>

[Abstract]

Prototype Development for an AI-based CAI Tool: Technological Support for Interpretation of Numbers

Moonsun Choi
(Ewha Womans University)

The technological advances in natural language processing is ushering in a new era of augmented interpreting where human interpreters perform their tasks more easily and better with the support of technology. Against this backdrop, the current study is aimed at developing and presenting the prototype of an AI-based computer-assisted interpreting (CAI) tool to support interpreters in processing numbers in real time by the automated recognition and extraction of number information from the source speech. The prototype system was built by using Google Speech-to-Text API and spaCy for automatic speech recognition (ASR) and named entity recognition (NER), respectively, as main components. During a pilot run, the system performed with the precision of 0.82 and recall of 0.90. The prototype system was configured so that its output was presented in three different layouts: only numbers presented in a dialogue box; sentences with number visualization; and real-time transcription followed immediately by fully visualized sentences. It was expected that presenting the full sentences with number visualization would be the optimal option for interpreters due to the avthe context and easy location of numbers in a given speech segment.

Keywords: computer-assisted interpreting (CAI), augmented interpreting, interpreting and technology, Google Speech-to-Text API, named entity recognition (NER)

주제어: 컴퓨터 보조 통역, 증강 통역, 통역과 기술, 구글 음성인식 API, 개체명 인식

최문선(<https://orcid.org/0000-0001-8890-8598>)

이화여자대학교 통역번역대학원 조교수

choimoonsun@ewha.ac.kr

관심분야: 통역 전략, 통역 평가, 통역 교육, 통역과 기술, 코퍼스언어학

논문 투고: 2022년 9월 4일

1차 심사 완료: 2022년 9월 27일

2차 심사 완료: 2022년 10월 11일

게재 확정: 2022년 10월 21일