

지금 이 사람

—

음성 인식을 넘어서는, 언어 인공지능을 위하여

이윤근 자동통역인공지능연구센터장을 만나다

—



답변자: 이윤근(한국전자통신연구원 자동통역인공지능연구센터장)

질문자: 권창섭(아주대학교 강사)

때: 2015. 5. 20.

곳: 한국전자통신연구원 자동통역인공지능연구센터(대전)

1990년대 후반, 휴대 전화에 대고 “우리집”을 다정하게 속삭이거나 “본부”를 다급하게 외치던 광고를 기억할 것이다. 한국에서 음성 인식 기술을 대중적으로 알린 첫 번째 장면이었다. 하지만 실제로 그 기능을 사용하는 이들을 찾아보기는 힘들었다. 버튼이 아닌 음성으로 전화를 건다는 것 자체가 낯설고 부끄러운 일이었음은 차치하고서라도, 일단 음성 인식의 정확도가 높지 않았기 때문이다.

근 20년이 지난 지금은 어떻게 달라졌을까? 언제나 기술의 진보 속도는 우리의 생각보다 빨라서 그사이 음성 인식 기술은 많이 발달했다. 또한 스마트폰의 등장으로 새로운 디바이스(device) 시장이 형성되면서 음성 인식 기술은 최근 가장 뜨거운 이슈로 떠올랐다.

음성 인식 기술이라고 하면 당신은 무엇이 가장 먼저 떠오르는가? 뉘앙스 커뮤니케이션? 애플의 시리(Siri)? 구글과 안드로이드? 우리나라 역시 이미 이들과 어깨를 나란히 할 만한 기술 수준을 갖추고 있다면? 한국전자통신연구원(ETRI)의 음성 인식 기술은 이미 정확도 90%를 웃도는 수준이며, 이를 자동 통·번역 시스템이나 더 향상된 인공지능 시스템으로 응용하기 위한 기술 연구 역시 계속되고 있다.

한국의 음성 인식 기술의 현재와 미래를 묻기 위해 대전에 있는 한국전자통신연구원(ETRI, Electronics and Telecommunications Research Institute, 이하 에트리)을 찾아 이윤근 자동통역인공지능연구센터장을 만났다.

권창섭 안녕하세요. 바쁘신 와중에 귀한 시간 내주셔서 감사합니다. 지금 센터장을 맡고 계신 곳의 정확한 이름이 ‘자동통역인공지능연구센터’이지요? 이곳에 대한 간단한 소개부터 부탁드립니다.

이윤근 기본적으로는 큰 틀에서 음성 정보와 언어 처리에 대한 일을 하는 곳입니다. 우리 센터 내에는 네 부서가 있는데, 이 부서들이 무슨 일을 하는지 말씀드리면 전체적인 이해가 쉬울 것 같네요. 첫째가 음성 처리연구실, 둘째가 언어처리 연구실, 셋째가 자동통역연구실, 넷째가 지식마이닝연구실입니다. 음성처리연구실에서는 주로 음성 인식 기술을 연구합니다. 사람의 말을 문자로 바꾸어 주고 약간의 이해를 하는 기술을 개발합니다. 언어처리 연구실은 두 영역의 연구를 하는데, 우선 자동 번역에 대한 연구를 합니다. 특정 언어의 텍스트 정보를 다른 언어로 변환하는 기술을 연구하는 것이지요. 다른 하나는 대화 처리 기술 연구인데요, 이는 음성 인식 기술과 맞물려 적용되는 분야입니다. 예전에 혹시 채팅 로봇 같은 것을 써 본 적이 있으신가요? 컴퓨터상에서 내가 어떤 말을 던지면 거기에 응답을 하는 프로그램 말입니다.

권창섭 ‘심심이’ 같은 프로그램 말씀인가요? 애플의 시리도 알고 있습니다.

이윤근 그런 프로그램의 기본이 대화 처리 기술입니다. 텍스트 입력이든 음성 입력이든 그 입력이 무슨 의미인지 이해하고 또 반응을 하도록 하는 것이 대화 처리 기술입니다.

권창섭 자동통역연구실은 어떤 연구를 하나요? 좀 전의 자동 번역 기술과 연구 영역이 겹치는 건 아닌가요?

이윤근 자동 통역 기술은 좀 전에 언급된 자동 번역 기술과 음성 인식 기술이 혼합된 것입니다. 여기에 음성 합성 기술까지 더해져 세 기술이 혼합되어야 하지요. 사람이 말을 하면 그 음성을 인식하고, 다른 언어

로 번역한 후에, 이를 그 언어로 음성 합성을 하여 출력해 내는 것이 자동 통역입니다. 음성처리연구실에서 음성 인식 기술을 연구하고, 언어 처리연구실에서 자동 번역 기술을 연구하지만, 둘이 통합된 기술은 별도로 연구해야 할 부분들이 있습니다. 그래서 자동 통역에 특화된 연구를 하는 곳이 바로 자동통역연구실이지요. 이곳에서 ‘지니톡(Genie Talk)’¹⁾이 나왔지요. 물론 자동통역연구실만의 성과라는 의미는 아닙니다. 자동통역연구실이 주관했다는 뜻이지요.

권창섭 지식마이닝연구실은 이름만으로 그 기능을 눈치 채기 어려운 데요, 어떤 일을 하는 곳인가요?

이윤근 ‘아이비엠 왓슨(IBM Watson)’이라는 슈퍼컴퓨터가 있어요. 혹시 들어 보셨나요?

권창섭 아니요. 잘 모르겠습니다.

이윤근 그럼 <제퍼디 쇼(Jeopardy Show)>라는 미국의 퀴즈 쇼는요?

권창섭 네. 그 쇼는 알고 있습니다.

이윤근 몇 년 전에 그 퀴즈 쇼에서 아이비엠(IBM)에서 만든 왓슨 컴퓨터와 인간이 대결을 했는데, 왓슨 컴퓨터가 승리를 했어요. 그래서 많이 유명해졌지요. 왓슨은 어떤 질문에 대답을 하는 데 특화되어 있는 인공 지능 컴퓨터입니다. 그런 인공 지능에 대한 연구를 하는 곳이 지식마이닝연구실입니다. 질의가 들어올 때, 그 질의의 의미를 아는 것은 물론이고, 질의의 초점도 무엇인지를 파악하여 데이터베이스에서 정보를 찾아내 답을 주는 인공 지능을 연구하는 것이지요. 그래서 그 기술의 이름이 지식 마이닝(mining)이에요.

1) 한국전자통신연구원에서 개발한 4개 국어(한국어·영어·중국어·일본어) 양방향 자동 통역 애플리케이션.

권창섭 설명을 들으니 전체적인 큰 그림도 그려집니다.

이윤근 큰 틀에서는 인공 지능, 그중에서도 언어나 음성에 기반을 둔 기술이라고 보면 되지요.

구글보다 진일보한 자동통·번역 시스템, '지니톡'

권창섭 아까 잠깐 언급하셨던 '지니톡'이 자동통역연구실 주관으로 개발해 인천 아시안게임 때 공개된 프로그램이지요?

이윤근 '지니톡'은 사실 인천 아시안게임 이전부터 개발해 왔습니다. 그러다 인천 아시안게임 때 시범 서비스를 하면서 널리 알려진 것이지요. 그 이전에도 일반 시민들에게 서비스해 왔습니다.

권창섭 2012년에 처음으로 공개되었지요?

이윤근 네. 처음에는 한국어-영어 사이의 통역만 할 수 있다가 차차 일본어와 중국어까지 확장되었습니다.

권창섭 애플리케이션이 공개된 후 대중의 반응은 어땠습니까?

이윤근 다운로드가 굉장히 많이 되었어요. 현재 220만 건 정도입니다.

권창섭 실제로 제 친구 한 명도 일본 여행 가면서 그 애플리케이션을 다운로드받더라고요.

이윤근 네. 다운로드 횟수를 보면 반응이 좋았던 것으로 보입니다. 그리고 그 애플리케이션을 사용하면 로그 데이터가 저희 서버에 저장되어서 대략적인 사용량을 파악할 수 있는데, 사용량도 많은 것으로 보입니다.

권창섭 아무래도 우리가 이전까지 가장 많이 사용하던 구글 통·번역 시스템과 비교하지 않을 수 없습니다. 이미 언론에서도 다뤄진 바 있듯

이 ‘지니톡’은 구글의 통·번역 시스템에 비해 진일보했다는 평가를 받고 있습니다. 어떤 점에서 그러한 평가를 받게 되었다고 보십니까?

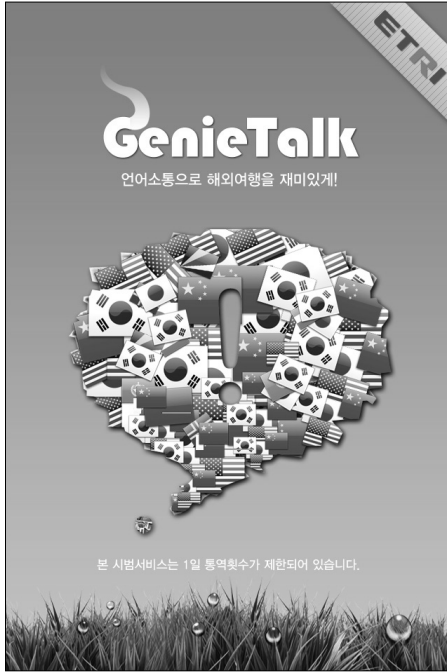
이윤근 일단은 그 둘이 기술의 형태가 다릅니다. 구글은 통계적 방식을 사용해요. 이미 구글에서는 한 언어와 다른 언어 사이에 뜻이 일치하는 대응 쌍들이 축적된 바이링구얼 코퍼스(bilingual corpus)를 상당한 정도로 축적해 놓았거든요. 이 데이터베이스 내에서 통계 처리를 하여, 입력된 문장에 가장 적절한 짝이 되는 대응 언어의 문장을 찾아내는 식으로 번역을 해 줍니다. 그런데 이것은 같은 어족 내부의 언어들, 혹은 가까운 어족 사이의 언어들끼리는 잘 맞습니다. 번역 성능이 좋죠. 인구어(印歐語, Indo-European)들끼리는 번역이 잘된다는 말입니다. 그런데 한국어와 인구어들과는 잘 되지 않죠. 어족이 다르니까요. 한편 한국어와 일본어는 또 괜찮게 번역됩니다. 즉, 다른 어족들끼리는 오류가 매우 많은 단점이 있는 거지요. 그래서 우리는 그 방식을 사용하지 않고 지식 학습 기반 자동 번역 기술을 새로 개발하여 사용합니다. 그 알고리즘을 설명 드리자면 매우 길고 어렵습니다. (웃음) 어쨌든 그래서 ‘지니톡’이 현재 구글의 시스템보다 10~15% 정도 정확도가 높다는 평가를 받고 있지요.

권창섭 영어, 중국어, 일본어 모두 그 정도 성능 우위를 점하고 있다는 말씀이신가요?

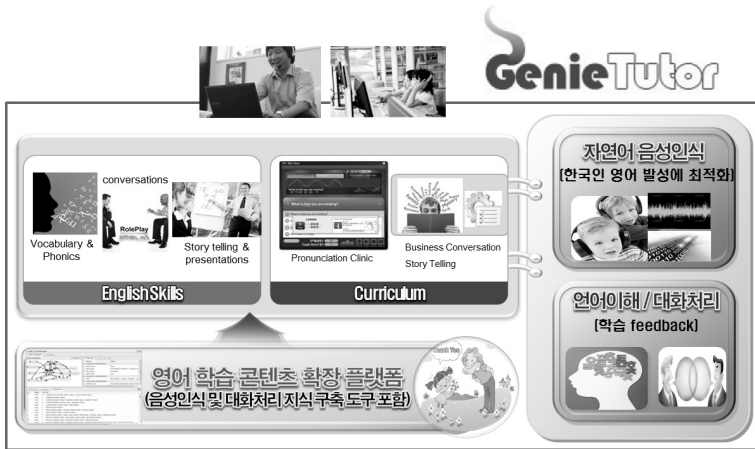
이윤근 그렇죠.

권창섭 그리고 구글은 클라이언트 서버(client server) 방식이니까 반드시 웹과 연결되어야만 사용할 수 있지 않습니까?

이윤근 구글 내부에는 임베디드 방식을 쓰는지 모르겠는데 외부로 공개된 것은 모두 클라이언트 서버 방식이니까요. 그런 점에서 ‘지니톡’은 반드시 웹에 연결되지 않아도 사용할 수 있어서 편리합니다.



[그림 1] '지니톡'의 초기 화면



[그림 2] '지니 튜터' 개념도



[그림 3] 2013년 9월 일본에서 열린 세계 최대 여행 박람회에 한일 '지니톡'을 전시했다.



[그림 4] 2'지니톡'은 2012년 스마트 콘텐츠 어워드 등 4개 부문에서 상을 받았다.

권창섭 단점이라면 어떤 것이 있습니까?

이윤근 가장 기본적으로 구글에 비해 언어 세트(set)가 적지요. 구글은 통계적 방식을 이용하기 때문에 데이터만 있으면 언어 세트 늘리기가 매우 쉽습니다. 그 대신 전반적으로 질이 낮은 것이고요. 우리가 취하는 방식인 지식 학습 기반 자동 번역은 전통적인 규칙 기반 번역 기법에 비해서는 언어 확장이 쉽지만 아무래도 전문가들의 손을 거쳐야 합니다. 그러다 보니 언어 세트를 확장하는 데 시간과 품이 더 들 수밖에 없지요. 그리고 현재는 한국어를 중심으로 한 번역 세트만 우선 개발하고 있습니다.

권창섭 현재 영어, 중국어, 일본어까지 한국어와 양방향 통·번역이 가능한데요, 차후 어떤 언어로 확장해 나갈 계획이신가요?

이윤근 현재 프랑스어와 스페인어로 확장하고 있습니다. 독일어와 러시아어도 계획 중입니다. 2018년까지 이를 완료해서 평창 동계 올림픽에서 서비스를 할 예정입니다.

권창섭 아무래도 전 세계적으로 많이 사용되는 언어부터 진행되는군요. 언어 세트를 확장하는 과정에서는 어족에 따른 난이도 차이는 없습니까? 가령 일본어나 중국어에 비해 인구어를 연구하고 시스템에 도입하는 일이 더 어렵다거나 그렇진 않은지요?

이윤근 그렇진 않습니다. 품은 대체로 비슷비슷하게 들어요.

음성 인식 기술의 어제, 오늘, 그리고 내일

권창섭 자동 통·번역 기술의 전제 조건이 될 수 있는 음성 인식 기술은 자동 통·번역 외의 다른 분야에서도 현재 수요가 늘고 있는 추세입니다. 이처럼 최근에 음성 인식 기술이 화두로 떠오른 까닭은 무엇인가요?

이윤근 무엇보다 스마트폰 때문이겠지요. 음성 인식 기술 자체의 역사는 꽤 오래됩니다. 1950년대부터 미국에서 연구가 시작되었으니까요. 상용화가 시작된 것은 1990년대 말에서 2000년대 초반 정도인데, 주로 많이 사용된 분야가 자동응답서비스(ARS, Automatic Response Service)입니다. 그런데 ARS 시장의 규모가 급속도로 커지지는 않습니다. 그래서 음성 인식 기술이 정체를 겪었지요. 그러다 다시 부흥하게 된 계기가 바로 스마트폰의 출현입니다. 가장 큰 동기 중 하나이지요.

또 세계적인 정보기술(IT) 분야의 빅 플레이어(Big Player: 큰손)들이 스마트폰에 기반을 둔 음성 인식 기술에 달려들었다는 점도 이유로 꼽을 수 있습니다. 구글, 애플, 엠에스(MS) 등이 대표적이고 아이비엠(IBM)도 참여하고 있고요. 그러면서 시장이 불타오르게 된 것이죠. 애플의 시리가 음성 인식에 이용한 인공 지능 개인 비서 서비스이고, 구글에서는 음성 인식을 통한 검색 서비스를 전 세계적으로 보급하고 있지요. 마이크로소프트도 음성 인식 기술을 오래 개발해 왔습니다만 스마트폰 시장에서는 후발 주자입니다. 이런 식으로 빅 플레이어들이 참가하다 보니 시장이 자연스레 커지고 관심도 많이 받고 활성화된 것이지요.

권창섭 한국에서 음성 인식을 연구하는 곳이 이곳 외에 또 어떤 곳들이 있습니까?

이윤근 벤처 기업들이 있긴 한데 아무래도 큰 규모의 일이다 보니 벤처 기업 독자적으로 구글이나 뉘앙스 등 음성 인식의 선두 주자들과 경쟁하기에는 역부족이지요. 그래서 보통 저희와 협력해서 사업을 진행하는 경우가 많습니다.

권창섭 한국의 빅 플레이어들은 음성 인식 연구를 하지 않나요? 네이버와 같은 대형 포털이나 삼성전자 같은 곳에서도 충분히 연구할 법하다고 생각되는데요.

이윤근 네이버나 다음도 음성 인식 기술을 가지고 있습니다. 다만, 자신들의 검색 서비스에만 특화되어 있지요.

권창섭 응용 기술까지는 나아가지 않은 것이군요?

이윤근 네. 최근의 음성 인식 기술은 거의 대화 음성 인터페이스로 확장되는데, 거기까지 연구 개발 투자를 하는 것이 쉽지 않은 형편일 겁니다. 검색을 위한 음성 인식 기술은 포털이 자체 보유하고 있습니다.

권창섭 삼성이나 엘지(LG) 등 기업들은 어떻습니까?

이윤근 엘지와 삼성도 자체 연구팀이 있긴 합니다. 그런데 삼성 스마트폰에 들어가는 모든 음성 기술이 뉘앙스에서 개발한 것이잖아요? 자체 기술만으로는 어려움이 있어서겠지요. 가장 큰 어려움이 언어 세트의 한계입니다. 기술 수준만으로는 비슷한데 뉘앙스는 이미 50여 개 언어를 지원하고 있습니다. 삼성에서 스마트폰을 만들어 세계 각국에 판매할 때 매번 다른 소프트웨어를 개발하여 탑재하는 것은 쉬운 일이 아닙니다. 그래서 상업적 경쟁력을 고려하여 이미 개발된 뉘앙스의 기술을 이용하게 된 것이지요.

권창섭 다른 기업들의 기술 수준과 비교하였을 때 에트리의 음성 인식 기술 역량은 어느 정도 경쟁력을 갖추었다고 평가할 수 있나요?

이윤근 저희도 자동 통·번역 기술을 연구하기 때문에 음성 인식 기술

에서도 다국어로 확장해야만 하지요. 원천 기술의 수준은 미국의 선두 주자들과 비교했을 때 큰 차이가 없습니다. 언어 세트만 늘리면 충분히 경쟁력이 있지요. 아까 말씀드린 대로 2018년까지 언어 세트의 수를 확장하고 차후에 아랍어나 동남아 쪽 언어까지 확장하면 추월도 할 수 있으리라 봅니다.

권창섭 아까도 잠시 언급했던 부분인데, 음성 인식의 방법론에는 크게 임베디드 방식과 클라이언트 서버 방식 두 종류가 있는 것으로 압니다. 구글은 클라이언트 서버 방식을 취하고 있는데, 뉘앙스는 어떤 방식을 사용하고 있나요?

이윤근 뉘앙스는 두 방식 모두 사용하고 있습니다. 뉘앙스는 자동 통·번역을 하는 업체는 아니고 음성 인식만 전문으로 하는 업체인데, 임베디드 방식으로도 경쟁력 있는 제품이 있고 클라이언트 서버 방식으로도 경쟁력 있는 제품을 갖추고 있습니다.

권창섭 그럼 임베디드 방식과 클라이언트 서버 방식의 장단점을 간단하게 설명해 주실 수 있을까요?

이윤근 임베디드 방식의 장점은 네트워크가 없는 곳에서도 쓸 수 있다는 장점이 있지요. ‘지니톡’이 그렇잖아요? 그러나 메모리나 중앙처리 장치(CPU)와 같은 리소스가 상대적으로 제한될 수밖에 없습니다. 게다가 애플리케이션도 상대적으로 무거운 편이고요. 서로의 장단점이 딱 반대라고 생각하시면 돼요. 클라이언트 서버 방식은 상대적으로 애플리케이션이 가벼운 편이고 리소스 제한이 덜하지만 네트워크가 연결되어 있어야만 사용할 수 있다는 단점이 있습니다.

권창섭 음성 인식이 자동 통·번역의 전제가 되듯이 음성 인식의 전제는 음성 데이터베이스의 구축일 겁니다. 이를 구축하는 데에는 어떤 방법론을 사용했는지 여쭙겠습니다.

이윤근 프로토타입을 만들기 위해서는 일정량의 데이터가 필요해요. 이런 데이터를 수집하는 전문 기관들이 존재합니다. 많은 사람들을 모아 특정한 대본을 주고 읽게 한 뒤 녹음하고, 그것을 다시 전사해 다른 곳에 공급하는 업체들이지요. 그런 업체들을 통해서 데이터를 확보하기도 합니다. 그리고 자료를 공개해서 판매하는 외국 업체들도 있는데, 거기서 구매하기도 하죠. 이런 식으로 일정량을 모아 프로토타입을 만듭니다. 그 후에는 서비스를 하면서 입력되는 각종 로그 데이터들을 이용하여 계속 성능을 높여 가는 것입니다. 저희뿐 아니라 모든 곳이 마찬가지입니다.

권창섭 음성 데이터베이스가 음성 인식에 활용될 수 있으려면 레이블링이 일정 수준 이상 되어 있어야만 할 것 같은데, 그것을 사람이 일일이 하는 것은 아닐 테고 어느 정도 자동화가 되어 있겠지요? 그 자동화의 수준은 어느 정도입니까?

이윤근 일정량은 사람이 직접 다뤄야 합니다. 음성 인식을 위해서라면 반드시 일일이 레이블링이 되어 있지 않아도 됩니다. 일정 수준 이상의 성능을 갖춘 음성 인식기라면 음성 인식기를 통과하면서 어느 정도 자동으로 레이블링되거든요. 완벽히 정확하게는 아니지만요. 하지만 많은 데이터가 입력되면 통계적으로 그런 부정확함을 가려내면서 전체적인 성능이 올라가게 되어 있습니다.

권창섭 아무래도 ‘한국어’ 음성 인식을 주로 연구하다 보면 ‘한국어’라는 언어의 개별적 특수성이 고려되지 않을까 싶습니다. 한국어의 특수성 중에 주로 어떤 특성이 연구 개발에 영향을 미친다거나 주된 고려 대상이 되는지요?

이윤근 소프트웨어를 만들 때 특정 언어와 기술의 관계가 종속적인(dependent) 경우도 있고 독자적인(independent) 경우도 있습니다. 음

성 인식은 거의 독자적이라고 보시면 돼요. 요즘은 특히 통계 기반이나 기기 학습의 방법을 많이 사용합니다. 전문가의 지식을 반영하는 것은 배제하는 편이죠. 왜냐하면 언어 확장의 걸림돌이 되거든요. 그래서 요즘은 개별 언어가 음성 인식 기술에 미치는 영향은 크게 없습니다.

권창섭 제 예상과 많이 다르네요.

이윤근 개별 언어의 특수성을 고려하는 부분은 이런 부분이죠. 텍스트가 소리로 변환될 때 자음접변이나 구개음화 등의 음운 현상이 적용되잖아요? 즉, 문자열이 발음열로 변환될 때의 규칙(Grapheme to Phoneme Conversion Rule) 같은 것은 언어마다 다르니까 고려 대상이 되지요. 그 밖에는 크게 문제될 것이 없어요. 그다음 과정들은 이미 언어 보편적으로 적용될 수 있습니다.

권창섭 한국어의 특수한 음소들, 가령 모음 ‘ㅡ’나 ‘ㅣ’ 같은 경우는 특이한 모음들 아닙니까? 그런 것들이 문제가 되지 않는지요?

이윤근 어차피 음성 인식을 위한 데이터베이스 자체가 한국어 화자들이 발음한 것들이기 때문에 매칭만 잘 되도록 설계하면 문제될 것은 없습니다. 다만, 이런 편차는 있겠네요. 음소의 수가 많은 언어가 있고 적은 언어가 있잖아요? 일본어 같은 경우는 음소의 수가 적으니까 음성 인식의 변별력은 좋은 편입니다. 그러나 또 다른 난점들이 있지요. 문자에 한자를 섞어 쓰기 때문에 텍스트와 음성 사이의 전환에 상대적인 어려움이 있습니다. 중국어는 띄어쓰기가 없다는 표기법의 문제점이 있고요. 언어마다 기술을 개발할 때 약간씩 장단점이 있긴 합니다만 그 것이 강력하게 작용하지는 않습니다.

권창섭 한국어는 상대적으로 문자와 음성을 대응시키고 전환하기가 쉬웠을 것 같습니다.

이윤근 네. 아무래도 표음문자이다 보니 쉬운 편이지요.

권창섭 같은 표음문자라도 영어는 상황이 다를 것 같은데요? 묵음이 있는 단어도 많고, 또한 문자와 음소가 일대일 대응되지 않는 경우도 많지 않습니까?

이윤근 네. 그래서 영어는 아까 언급한 문자열이 발음열로 변환될 때의 규칙 사전이 중요합니다. 한국어는 기본적으로 규칙 중심이기는 하나 형태소나 단어의 지위에 따라 음운 규칙이 달리 적용되기도 해 그런 경우들은 따로 사전 정보를 부여하기도 하지요.

권창섭 자료를 수집하는 데 간투사가 들어간다는, 버벅거린다는, 혹은 어순이 도치된 문장 등 비정형적인 문장들이 포함되어 있기도 하나요? 그리고 그런 문장들에 대한 음성 인식도 가능한 수준입니까?

이윤근 아직은 비정형적인 문장들은 잘 인식이 되지 않습니다. 사실 그것이 남은 과제 중 가장 큰 것입니다. 그런 것이 되어야 지금처럼 두 명 이상의 실제 대화도 인식되고 문자화가 될 텐데, 아직은 그런 수준이 안 되어 사용할 때 매우 정확히 이야기해야 합니다. 외국 업체들의 기술 수준도 마찬가지이고요. 현재 에트리에서 하고 있는 주요 연구가 바로 그런 비정형적인 발화나 문장들에 대한 연구입니다. 비정형 자연 언어 음성 인식이라고 할 수 있겠네요.

권창섭 워낙에 그 변이의 범위가 넓어서 아주 어려운 과제일 것 같습니다.

이윤근 네. 어려운 숙제예요.

권창섭 그런 연구를 위해서는 국어학계나 언어학계 등 학계와도 많은 협력 연구가 필요할 것 같습니다.

이윤근 네. 실제로 그렇게 하고 있습니다.

인간에 가까워진 인공 지능은 어떤 모습일까

권창섭 음성 인식 기술에서나 자동 통·번역 기술의 개발에서 향후 계획이나 중요한 과제는 무엇입니까?

이윤근 단기적인 목표가 있고 장기적인 목표가 있을 겁니다. 단기적인 목표는 세계 시장에 진출하는 것입니다. 한국어를 중심 기반으로 하고 있기 때문에 아직은 국내 시장에 머무르고 있는 형국이거든요. ‘지니톡’이란 자동 통·번역 기술이나 음성 인식 기술이나 그 수준은 세계 시장에서도 충분한 경쟁력이 있습니다. 그런데 아까도 말씀드렸듯이 언어 세트의 수가 아직은 적다는 점, 국내 업체들이 해외 시장에 잘 나가지 못하고 있다는 점 등이 겹쳐 어려움이 있습니다. 여하튼 단기적인 목표는 이러한 어려움을 극복하고 세계 시장으로 진출해서 기술 경쟁력뿐 아니라 시장 경쟁력도 높이는 것이 목표입니다. 그래서 해외의 다른 기업들이 아직 달성하지 못한 기술들을 가지고 세계 시장에 나가야 되겠죠. 가령 아까 언급된 비정형적인 문장 인식이 어려운데 이러한 문제들을 먼저 해결해서 세계 시장에서 경쟁력을 가지는 것이 일단의 목표입니다.

권창섭 궁극적으로 지향하는 바는 무엇입니까?

이윤근 궁극적인 목표는, 글썬요……, 선도적인 기업이나 우리나라 이 분야에서는 모두 마찬가지로 생각할 텐데, 이 기술은 아직 해결된 것이 아니고, 특히 음성 인식 기술은 가장 오래 연구가 지속될 분야라고 생각됩니다. 인공 지능이란 것이 여러 가지가 있겠지만, 사람이 지식을 표현하고 전달하고 축적하는 가장 기본적인 도구가 언어잖아요? 언어는 가장 중요하면서도 가장 복잡한 지식 체계이기 때문에 그것을 컴퓨터가 다루게 하는 것도 가장 어렵고 복잡한 문제이겠죠. 그리고 그것이

가능하려면 음성을 단순히 문자로 전환하는 차원을 넘어 해당 문장을 이해할 수도 있어야 하고 적절하게 피드백을 할 수 있어야 하는데, 그것도 어떻게 보면 지능이란 것의 가장 초보적인 단계 아니겠습니까? 그래서 그 단계를 넘어서 아까 언급된 지식 마이닝처럼 상식이나 전문 지식 체계 등이 갖춰지면 점점 인간에 가깝게 진화하게 되는 것이지요. 거기 가 이 분야의 최종점이라고 할 수 있습니다. 그것을 누가 달성하느냐가 문제인 거지요. 이를 달성하는 것이 우리의 장기적인 목표라고 할 수 있겠습니다.

물론 이것이 달성되었을 때 세상이 좋아질지 나빠질지는 잘 모르겠습니다. 요즘 영화를 보면 방금 말씀드린 대로 컴퓨터가 인간과 유사한 정도로 진화하였을 때의 모습을 대개 디스토피아로 그리고 있지 않습니까? 그런 것을 고려하지 않을 수는 없겠지만 과학자나 공학자로서는 달성해 내는 것이 목표일 수밖에 없지요.

권창섭 여러 생각이 많이 드는 말씀이네요. 그렇지만 이런 고민이 동반된 기술의 진보라면 인간에게 더욱 도움이 되는 방향으로 나아가는 길이 아닐까 생각합니다. 오랜 시간 좋은 말씀 감사합니다.

이윤근 감사합니다.