Journal of Korea Multimedia Society Vol. 23, No. 10, October 2020(pp. 1296-1306)
https://doi.org/10.9717/kmms.2020.23.10.1296

가상 캐릭터를 활용하여 아동의 구어 대화를 유도하는 대화형 에이전트

최지영†, 정기철††

Embodied Conversational Agent Using a Virtual Character to Induce Children's Verbal Communication

Jiyeong Choi†, Keechul Jung††

ABSTRACT

Childhood verbal communication impacts children's language skills and has a positive effect as partners use more vocabulary. But reduction in family time, caused by lowered age for private education and so on, has reduced the chance for children to speak with partners who have a proficient language skill. This vacancy was naturally occupied by the media, which has become one of the cornerstones of the growth of kids' contents. Kids contents are making various attempts to expand the breadth of services. But most contents still focus on unilateral visual information delivery yet, so there is a limit to satisfy the vacancy of conversation partners. Therefore this paper suggests an ECA(Embodied conversational agent) to induce children's spoken conversation using a virtual character frequently used in kids contents. This system is implemented by the voice bot and agent model produced using an IBM assistant and Unity. As a result of using ECA for 66 children of 5-9 years old, it showed meaningful results in terms of induction of verbal communication.

Key words: Conversational Agent, Kids Contents, Artificial Intelligence

1. 서 론

구어 상호작용, 즉 대화는 아동의 언어능력을 촉진한다[1]. 언어능력은 언어 수행뿐만 아니라 사회정서 발달과도 유의미한 상관관계를 가진 요소로[2], 발달 속도가 빠른 아동기라면 구어 상호작용에 더욱 주의를 기울일 필요가 있다. 상호작용은 자신의 언어체계와 목표 언어의 언어체계 사이의 차이를 알 수 있는 과정으로[3] 대화 상대가 알맞은 발음과 문법, 어휘를 사용할수록 올바른 언어 습득을 기대할 수 있다. 일반적인 경우, 언어 구사력이 충분하며 무리없이 아동과 지속적인 상호작용을 이어나갈 수 있는 상대는 같은 거주공간을 사용하는 보호자가 유일하다. 그러나 사교육을 시작하는 연령이 점점 낮아지고[4], 아동 사교육 시간이 증가하는[5] 등의 추세로 아동과 보호자가 분리된 시간이 늘고 있다. 이에 따라 구어 상호작용의 효과가 증폭될 환경이 줄어들었으며, 보호자와의 대화 시간은 아동 개인의 놀이 시간으로 대체되었다. 해당 과정에서 아동은 자연스레 다양한 매체와 접하게 되었고, 미디어는 아동의 언어

※ Corresponding Author : Kee chul Jung, Address: (156-743) Sangdo-ro 369, Dongjak-gu, Seoul, Korea, TEL : +82-2-828-7260, FAX : +82-2-812-5378, E-mail : kcjung@ssu.ac.kr
Receipt date : Sep. 4, 2020, Revision date : Oct. 7, 2020
Approval date : Oct. 14, 2020

† Global School of Media, College of IT, Soongsil University
(E-mail : curryisgood@naver.com)
†† Global School of Media, College of IT, Soongsil University

적·심리적 환경 조성에 큰 부분을 차지하기 시작했다. 9세 이하 자녀를 둔 부모를 대상으로 설문을 진행한 결과, 아동이 미디어 프로그램을 스스로 선택하기 시작한 나이는 만 1세에서 3세부터가 대부분(64.2%)이었으며 아동의 미디어 시청 시간은 하루 1시간 이상(64.9%)이 가장 많았다[6]. 미디어 환경은 기술적 혁신뿐 아니라 다양한 사회적, 심리적 요인들이 개입하는 장이다[7]. 그러나 현재 대다수의 미디어가 일방적인 시각 정보 전달에 집중하고 있어 대화 상대의 공석을 채우기에는 그 수가 부족한 상황이다.

상호작용성을 포함한 콘텐츠로는 대화형 에이전트를 예로 들 수 있다. 대화형 에이전트란 인간의 언어를 이해하고 의도를 분석하며 알맞은 응답을 출력하는 대화 시스템으로, 텍스트를 통해 사용자와 대화하는 챗봇이나 음성을 기반으로 대화하는 보이스봇이 이에 속한다. 이 중 그래픽 또는 물리적인 본체를 추가하여 시스템에 가시적인 형체를 부여한 것을 Embodied Conversational Agent(ECA)라고 한다. ECA는 일반적으로 보이스봇과 가상 모델을 결합해 제작되며, 비언어적 수단을 활용하여 대화 시스템을 보완할 수 있다. 관련 연구에 따르면 ECA는 사용자의 참여도를 높이고 신뢰를 주며, 전달되는 정보를 더 흥미 있고 덜 어려운 것으로 인식되게 한다고 밝혔다[8].

이와 같은 상황에서 본 논문은 미디어에 익숙한 현대 아동의 특성과 상호작용에 유리한 ECA의 장점을 활용하여 아동의 구어 대화를 유도하는 ECA 시스템을 제안한다.

2. 관련 연구

아동을 대상으로 하거나 행동을 유도하는 대화형 에이전트는 규모나 목적에 따라 다양한 형태로 개발되고 있다. 두 조건에 모두 부합하는 것은 교육 목적으로 설계된 ECA가 일반적이며, 독립적인 프로그램이나 시스템의 부가 기능으로써 활용된다.

2.1 웅진 빅박스

웅진 컴퍼스 사(社)에서 출시한 아동 대상 게임형 영어학습 플랫폼 빅박스는 '몬티와 Free Talk'라는 ECA 서비스를 제공한다. 이는 빅박스의 마스코트 캐릭터 몬티를 모델로 내세운 영어 말하기 학습 기능으로 유저들은 영어를 사용하여 몬티에게 말을 걸 수 있다.

오른쪽 하단의 마이크 버튼을 눌러 음성을 입력하면 몬티가 제스처를 취하며 반응한다. '몬티와 Free Talk'는 사전 녹음된 원어 음성과 말풍선을 통한 텍스트 응답이 답변으로 출력된다. 아직 인식이나 응답 측면에서 뛰어난 편은 아니나 계속해서 업데이트되는 중으로, 2020년 올해 9월에도 음성 인식 관련 기능이 개선되었다. 웅진 빅박스는 월 23,000원에 자유롭게 이용할 수 있으며 구글 플레이에서만 10만 회 이상의 다운로드를 달성하였다.

2.2 Rachel ECA[9]

Rachel은 서던 캘리포니아 대학교에서 개발한 자폐 아동 대상 ECA이다. 자폐 아동의 상호작용 패턴을 연구하는 것이 주된 목표이며 원활한 데이터 수집을 위해 아동의 사회·정서적인 행동을 장려하도록 설계되었다. Rachel은 WOZ(Wizard of Oz, 참가자 모르게 설계자가 시스템의 응답을 설정하는 테스팅 기법) 패러다임을 사용한 반자율 ECA로 미리 프로그래밍 된 응답을 설계자가 제어하여 출력한다. 가상 모델은 만 8살 정도의 외관을 한 인간 형태로 CSLU (Center for Spoken Language Understanding, 인간

Fig. 1. Woongjin BIGBOX 〈hello monty〉.

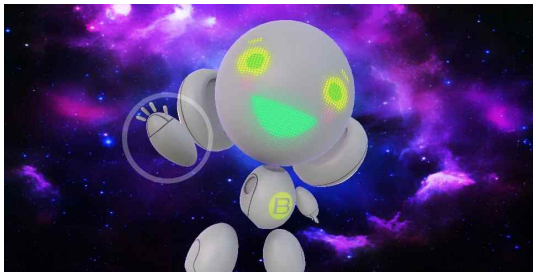


Fig. 2. Buddy.ai.

과 컴퓨터의 상호작용을 돕는 도구 모음 라이브러리)에서 제공하는 캐릭터를 활용했다. 유저에게 질문을 하거나 제시된 상황과 적합한 감정을 선택하는 감정 매칭 게임을 시작하고, 동화책을 읽다가 유저에게 발화 순서를 넘기는 등 여러 말하기 유도 기능을 포함하고 있다. 연구 결과 아동 한 명당 약 59분 정도의 상호작용 유도에 성공하였다.

2.3 Buddy.ai

Buddy.ai는 2019년 MyBuddy.ai 사(社)에서 출시한 아동 대상 AI 튜터로, 말하기 중심의 영어 교육 애플리케이션이다. 대표적인 기능 중 하나가 마스코트 캐릭터 버디와의 대화로, 유저는 버디에게 질문을 던지거나 질문에 답할 수 있으며, 버디와 하이파이브를 하는 등 터치를 통한 상호작용도 할 수 있다.

답변의 경우 유저가 입력한 음성 속 의도보다는 버디의 비언어적 출력에 집중했기에 입력 내용과 상관없이 제스처만 출력되기도 한다. 2020년 현재 Buddy.ai는 영어, 폴란드어, 러시아어, 스페인어, 터키어를 지원하며 일주일에 6.99달러(한화로 약 8,000원)를 지불하고 자유롭게 사용할 수 있다. 구글 플레이에서만 100만 회 이상의 다운로드 수를 달성하였으며 전 세계적으로 약 2만 6천 명 이상의 유료 회원을 보유하고 있다.

3. 아동 구어 대화 유도 ECA

본 논문에서 제안하는 시스템은 크게 음성 인식, 응답 선택, 가상 모델 모션 설정, 통합의 단계로 구성되어있다. 먼저 환영 응답을 출력하여 사용자를 맞이한 뒤, 음성 인식을 통해 사용자의 음성 신호를 입력받는다. Fig. 3(a)는 ECA의 음성 인식 단계로 인식률

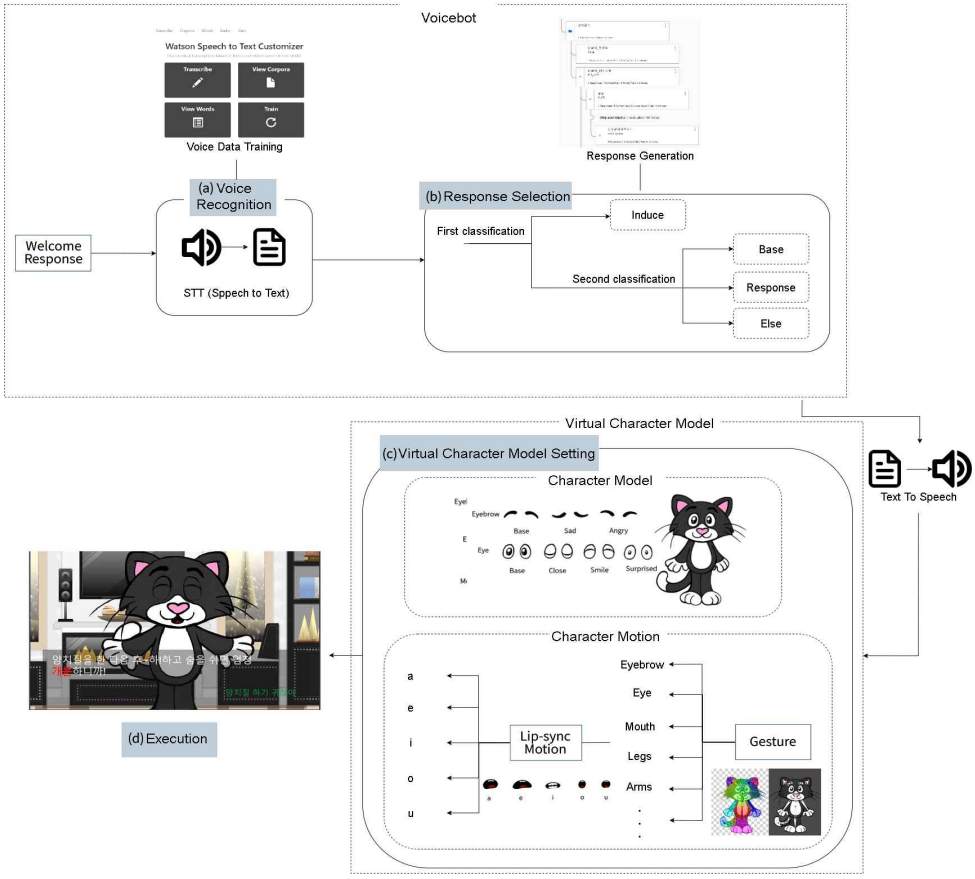


Fig. 3. ECA system architecture. (a) Voice Recognition, (b) Response Selection, (c) Virtual Character Model Setting, (d) Execution

상승을 위한 개별적인 음성 트레이닝을 사전에 실행한 상태다. 인식 단계로부터 전달된 음성 신호는 두 번의 분류 과정을 거쳐 정보값이 파악되며 해당 정보는 현재의 대화 흐름에 가장 적합한 응답을 선택하는 기준이 된다. Fig. 3(b)는 Watson Assistant를 활용하여 사전에 저장된 응답 데이터 중 출력할 응답을 탐색 및 선택하는 단계다. 본 시스템은 응답 데이터 288개가 사전에 입력되었다. 응답 선택 과정을 거쳐 출력된 결과는 음성 신호로 변환되어 가상 캐릭터 모델에 전달된다. Fig. 3(c)는 전달된 합성 음성을 토대로 립싱크 애니메이션을 포함한 캐릭터 모션을 출력하는 단계이다. 합성 음성과 캐릭터 모델의 모션이 통합되어 동시에 출력되면 ECA가 성공적으로 실행된다. Fig. 3(d)는 통합이 성공적으로 완료된 후의 실행 화면이다. 이처럼 음성을 인식하고 응답을 관리하는 보이스봇 기능과 시각적 요소를 관리하는 가상 모델 기능을 분리하여 작업한 뒤 추후에 통합하는 방식으로 제작한다. Fig. 3.은 전체적인 시스템 개념도이다.

3.1 음성 인식

시스템 실행 시 인사를 하는 환영 응답이 출력된 후 STT(Speech to Text, 음성을 텍스트로 변환) 음성 인식이 시작된다. 이때 입력된 신호를 기반으로 다음 단계가 진행되기 때문에 해당 과정에서 발생한 오류는 이후에 진행되는 모든 과정을 무의미하게 만들 가능성이 크다. 따라서 본 시스템은 인식률 향상을 위해 아동의 언어 발달 단계를 고려하여 음성 인식 모델을 설계한다. 아동의 연령이 높아질수록 대부분의 발음 오류 빈도가 줄어드는 것에 반해 이중 모음의 단모음화와 종성 자음 삽입은 만 6세까지도 꾸준히 관찰되고 있다[10]. 실제 STT를 실행한 뒤 아동과 성인에게 같은 문장을 말하도록 한 결과, 아동 화자의 인식 오류가 성인 화자에 비해 잦아 인식률에서 유의미한 차이를 보였다. 이를 보완하기 위해 5세부터 9세 아동의 발화 데이터를 수집하여 음성 데이터 트레이닝을 시행한다. IBM Watson STT를 기반으로 하여 총 16,280개의 음성 파일이 훈련에 사용되었다.

3.2 응답 선택

인식된 음성 신호를 기반으로 사전 입력된 데이터를 탐색하여 가장 적절한 응답을 출력한다. 응답 탐색에는 총 두 번의 분류 과정이 존재하며 먼저 음성 신호 입력 여부에 따라 1차 분류가 수행된다. 감지된 신호가 없을 시 유도(Induce) 카테고리에서 응답을 탐색하게 된다. 유도(Induce) 카테고리는 사용자로부터 어떠한 신호도 입력되지 않았을 때 출력되는 응답들의 집합으로, 사용자의 주의를 끌거나 대화를 유도한다. 음성이 입력되었을 경우 음성 신호 속 키워드와 의도를 분석하는 2차 분류가 진행된다. 응답(Response) 카테고리는 사용자의 발화에 반응하는 응답들의 집합으로 대부분의 응답이 이에 속한다. 기본(Base) 카테고리는 높은 우선순위를 가진 응답의 집합으로, 대화 흐름에 상관없이 조건이 충족되자마자 실행되며 욕설을 감지하거나 단어의 정의를 설명해주는 역할을 한다. 인식된 음성의 내용이 그 어떤 조건과도 일치하는 않는 경우 기타(Else) 카테고리로 넘어가 음성 신호의 재입력을 요청하는 응답이 출력된다. Watson Assistant를 활용하여 해당 과정을 관리한다.

3.3 가상 캐릭터

IBM Watson TTS(Text to Speech, 텍스트를 음성으로 변환)를 거쳐 변환된 응답은 가상 캐릭터 모델에게 전달된다. 캐릭터 모델은 전달된 음성에 따라 알맞은 모션을 출력한다.

3.3.1 모델 생성

디자인적 특징과 아동의 전반적인 선호도를 고려하여 가상 모델을 선택한다. 사실적인 모델은 정보 전달에서, 만화적으로 묘사된 모델은 커뮤니케이션 부문에 강점을 가진다[11], 본 논문은 상호작용에 중점을 두었으므로 만화적으로 묘사된 캐릭터 모델을 채택한다. 구체적인 캐릭터 디자인에 대해서는 힐드레드의 연구를 참고할 수 있다[12]. 이는 5세 전후 아동들에게 좋아하는 그림을 고르게 한 연구로, 아동들은 동물이 그려진 그림을 우선적으로 선택하였다. 실제로 동물 캐릭터는 아동이 작고 부드러운 동물을 좋아한다는 기본적인 선호와 통제감을 부여한다는 점에서 아동 대상 콘텐츠에 자주 등장한다. 꼬리, 큰 귀 등 사람 캐릭터에 비해 다양한 비언어적 요소를 가지고 있어 감정표현 수단의 폭이 넓다는 장점 또한 존재한다. 비슷한 맥락에서 각 동물에게 어울리는 형

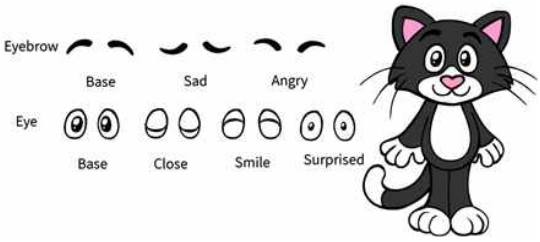


Fig. 4. Character model and emotion animation.

용사를 붙이는 실험을 시행한 결과, 대다수의 아동은 좋다, 똑똑하다와 같은 긍정적인 단어로 고양이를 선택했다[13]. 이를 미루어 보아 아동은 동물 중에서도 고양이에게 높은 흥미가 있을 것으로 추측되므로 캐릭터 모델의 생김새를 다음과 같이 설정한다.

3.3.2 캐릭터 모션 생성

Unity 2D 애니메이션 패키지에 포함된 Bone Editor를 사용하여 메쉬와 무게, 회전율을 설정한다. 눈썹을 통해 대략적인 감정을 표현하고 눈과 귀, 팔, 꼬리 모션이 이를 보조하도록 한다. 립싱크 모션은 기본, a, e, i, o, u를 나타내는 총 여섯 개의 입 모양으로 표현한다. 립싱크 모션 설정에는 SALSA2 suite를 활용하였다.

3.4 통합

unity-sdk-core-1.2.1, unity-sdk-4.7.1을 사용하여 통합 과정을 진행하였다.

UI는 다음과 같이 배치한다. 중앙에는 캐릭터 모델이 존재하며 대화창 상단에는 ECA에서 출력되는 응답 텍스트가, 하단에는 사용자의 음성을 인식한 텍스트가 나타난다. 이는 상호작용을 수월하게 만들기 위한 장치로, 음성과 동일한 내용의 텍스트를 화면에 배치하여 듣기 능력과 읽기 능력이 서로를 보조하도록 한다. 실제 이와 같은 언어 자막은 불균형한 시각적 주의를 사로잡고 정보에 대한 기억력을 상승시키는 효과가 있다[14]. 아동의 이해 여부를 추측하기 모호한 단어의 경우 자막에서 빨갛게 표시되며, 빨간 자막으로 표시된 모든 단어는 정의에 대해 질문할 수 있도록 한다. 인식된 사용자의 음성을 출력하는 텍스트는 의도대로 음성이 입력되었는가를 확인할 수 있다.

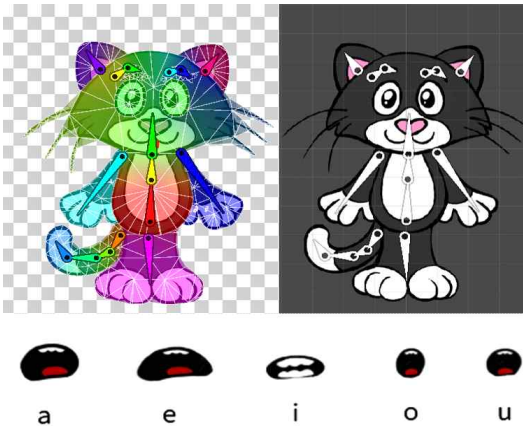


Fig. 5. Character model bone setting and lip-sync motion.

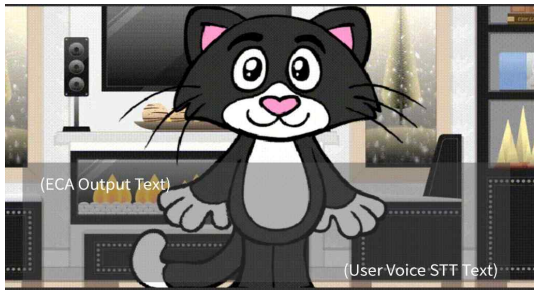


Fig. 6. ECA system UI.

3.5 실행화면

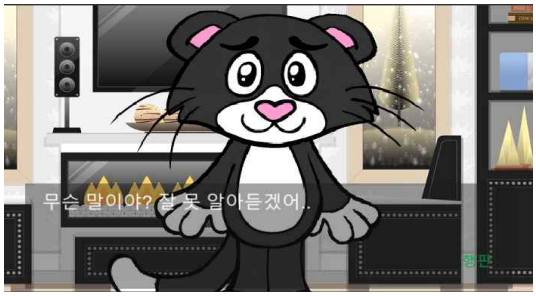
응답(Response) 카테고리는 가장 많은 응답을 가지고 있으며, 실행되는 제스처의 종류도 다양하다. Fig. 7(a)는 '양치질하기 귀찮아' 음성이 입력된 후의 결과이다. 양치질을 권유하는 응답이 출력된 다음의 화면으로 '양치질을 한 다음 후~하! 하고 숨을 쉬면 엄청 개운하니까!'라는 양치질의 이점에 대해 언급하는 응답이 출력되고 있다. 이때 '개운'이라는 단어는 정의 설명이 가능한 단어이므로 빨갛게 나타난다. 주의를 끄는 역할을 하는 유도(Induce) 카테고리는 의문형 응답을 출력한다. Fig. 7(b)는 8초 이상 아무런 음성이 인식되지 않았을 때의 화면이다. 기본(Base) 카테고리의 욕설 인식은 슬퍼하거나 화내는 제스처와 함께 출력되며, 욕설이 3회 인식되면 자동으로 프로그램이 종료된다. Fig. 7(c)는 욕설이 누적 3회 인식되어 프로그램이 종료되기 전의 실행 화면이다. 기본(Base) 카테고리의 정의 설명은 단어 또는 관용표현의 정의를 설명하는 역할로 실행 시 캐릭터 모델이

(a)Response Category

(b)Induce Category

(c) Base Category-Slang detection

(d)Base Category- Definition description



(e) Else Category

Fig. 7. Resulting images sorted by category.

책과 안경을 착용한다. 정의 설명이 가능한 모든 단어는 빨간 텍스트로 표시되며 설정된 단어로는 입이 무겁다, 개운하다, 이루어지다 등을 포함한 총 26개의 표현이 존재한다. Fig. 7(d)는 '입이 무겁다는 게 뭐야'라는 질문에 정의를 설명하고 있는 실행 화면이다. 기타(Else) 카테고리는 슬픔 제스처와 함께 출력된다. Fig. 7(e)는 실생활에서 잘 사용하지 않는 '행판'이라는 단어를 입력한 결과 화면이다.

4. 실험 및 평가

4.1 실험 대상

본 시스템의 기대 효과를 확인하기 위해 경기도 화성시에 소재한 유치원 및 초등학교에 재학 중인 5세~9세 아동 66명(유치원생 29명, 초등생 37명)을 실험 대상으로 선정하였다. 참여자의 평균 연령은 7.4세(SD=1.24)이다.

Table 1. Statistics of respondent age

Age group	Number
5	8
6	10
7	11
8	21
9	16
Total	66

4.2 예비 조사

시스템 시연 전 대화형 에이전트에 대한 기존 흥미와 흥미에 영향을 미치는 요소를 확인하기 위해 예비 조사를 시행한다. 아동에게 챗봇, 보이스봇, AI 로봇과 ECA를 포함한 대화형 에이전트 영상과 함께 간략한 설명을 한 뒤 설문지를 작성하도록 하였다. 설문지는 흥미 측정과 영향 요소 분석을 포함한 총 7개의 문항으로 구성되었다. 이때 유치원생 대상 조사는 설문지가 아닌 손을 들어 답하는 거수 방식으로 진행하였으며 설문 내용을 잘 이해하여 대답하고 있는지 수시로 확인했다. 예비 조사에 관한 연구 가설은 다음과 같다.

가설1 : 대화형 에이전트 사용 경험에 따라 대화형 에이전트에 대한 흥미에 차이가 있을 것이다.

가설 2 : 미디어 사용량에 따라 대화형 에이전트에 대한 흥미에 차이가 있을 것이다.

가설 3 : 보호자와의 대화 시간에 따라 대화형 에이전트에 대한 흥미에 차이가 있을 것이다.

크론바흐 알파값을 통해 본 설문지의 신뢰도를 계산한 결과 전체 0.82로 문항의 수정·재실험 과정 없이 결론을 도출한다. 전체 흥미 점수는 3.45점으로 양호한 수치를 보였으며 미경험자 흥미 점수(3.33)보다 경험자 흥미 점수(3.64)에서 평균적으로 더 높은 점수를 관찰할 수 있었다. 그러나 흥미 점수의 합계로 독립 표본 t 검정을 시행한 결과 대화형 에이전트 경험 여부는 흥미에 크게 관여하지 않았다. 따라서 가설1을 폐기한다.

분리가 무의미할 정도로 응답자의 수가 적은 항목들이 있었기에 다음과 같이 두 개로 분류했다. 미디어 사용량이 2시간을 넘어가는 아동(3.33)보다 그렇지 않은 아동(3.52)들이 대화형 에이전트에 더 높은 흥미를 보였으며, 그 수도 더 많았다. 흥미 점수의 합계로 독립 표본 t 검정을 시행한 결과, 미디어 사용 시간은 대화형 에이전트 흥미에 유의미한 영향($p<.05$)을 보였다. 따라서 가설2를 수용한다.

분리가 무의미할 정도로 응답자의 수가 적은 항목들이 있었기에 다섯 개의 항목을 다음과 같이 세 개로 줄여 분류했다. 세 개의 집단 중 보호자와의 대화 시간이 1시간 미만인 아동들의 대화형 에이전트 흥미(3.55)가 제일 높았으나 ANOVA 분석을 시행한 결과 보호자와의 대화 시간은 대화형 에이전트 흥미에 별다른 영향을 미치지 않았다. 따라서 가설 3을

1. 하루에 스마트폰, TV를 몇 시간 사용하나요?

① 30분보다 적게 ② 30분 ~ 1시간 ③ 1시간 ~ 2시간
④ 2시간 ~ 3시간 ⑤ 3시간보다 많이

2. 보호자(어머니, 아버지, 할아버지, 할머니 등 집안의 성인)와 하루에 어느정도 대화를 하나요?

① 30분보다 적게 ② 30분 ~ 1시간 ③ 1시간 ~ 2시간
④ 2시간 ~ 3시간 ⑤ 3시간보다 많이

3. 대화형 에이전트를 사용해 본 적이 있나요?

① 있다 ② 없다

문 항	매우 그렇다	그렇다	보통이다	아니다	매우 아니다
1. 이번 활동 주제(대화형 에이전트)는 재미있었나요?					
2. 오늘과 비슷한 활동 주제(대화형 에이전트)를 또 체험해보고 싶나요?					
3. 대화형 에이전트를 사용해보고 싶나요? 사용한 경험이 있다면 계속 사용하고 싶나요?					
4. 대화형 에이전트에 대해 관심이 있나요? 또는 관심이 생겼나요?					

Fig. 8. Images of preliminary investigation questionnaire.

폐기한다.

예비 조사를 통해 기존 흥미와 상관관계 요소를 살펴본 결과 흥미 점수는 평균 3.45의 수치를 보였으며 미디어 사용량이 대화형 에이전트 흥미에 유의미한 영향을 미친다는 사실을 확인했다.

4.3 실험

대상 아동들이 모인 공간에 ECA를 시현하며 설명한 뒤 한 명씩 5분간 ECA를 체험하도록 하였다. 해당 실험은 정확한 음성 인식을 위해 대기실과 실험실을 구분하여 이루어졌다. 다음은 ECA를 체험한 후 아동들이 작성한 설문지로 ECA에 대한 흥미(4문항), 완성도(5문항), 만족도(4문항), 대화 유도성(4문항)에 대화형 에이전트에 대한 흥미 문항(3문항)을 추가하여 총 다섯 영역 20문항으로 구성되어 있다. 응답 구성은 5단계 리커트 척도를 사용하였으며 점수가 5점에 가까울수록 긍정적으로 해석할 수 있다. 예비 조사와 동일하게 유치원생은 거수 방식으로 설문을 진행하였으며 이해 여부를 수시로 확인했다.

5. 결 론

크론바흐 알파 값을 통해 본 설문지의 신뢰도를 계산한 결과 전체 0.896으로 흥미(.78), 완성도(.65), 대화 흐름 완성도(.73), 만족도(.88), 대화 유도성(.79), 대화형 인터페이스 흥미(.85) 모두 0.6 이상의 신뢰도를 보였다. 따라서 문항의 수정·재실험 과정 없이 결론을 도출한다. 전 영역의 총점을 합산한 평균은 3.765로 전 분야의 평균 점수 모두 3.5 이상의 수치가 나왔다. 이는 ECA 시스템이 아동의 흥미를 끌고 지속적인 구어 대화를 유도하는 효과가 존재함을 보여준다.

대화형 에이전트에 대한 흥미도 또한 예비 조사 결과(3.27)와 비교해 높은 수치(3.67)로 유의미한 차이($p<.001$)가 나타났다. 예비 조사에서 대화형 에이전트 경험이 흥미에 큰 영향을 미치지 못한 것과는 반대로 ECA 체험 후 흥미도가 유의미하게 증가한 모습을 관찰할 수 있었다.

본 논문에서는 아동의 구어 대화 활동을 유도하는 ECA 시스템을 개발하고 시현하였다. 그 결과 ECA가 언어·정서 발달, 교육적 측면에서 효과를 발휘할 여지를 확인하였으며 키즈 콘텐츠로 활용할 가능성을 보였다. 특히 미디어 사용량이 하루에 두 시간 미만인 아동들에게 더 효과적일 것으로 예상되는데, 이는 미디어 사용량이 많을수록 비구어·시각적 정보전

Table 2. Differences in level of interest about conversational agent: Based on experience

Case	Number	Mean	Standard Deviation	Cronbach's Alpha	t(p)
Total	66	3.45	0.87	0.82	1.543(.129)
experienced	24	3.64	0.84		
inexperienced	42	3.33	0.87		

Table 3. Differences in level of interest about conversational agent: Based on media usage

Media usage	Number	Mean	Standard Deviation	t(p)
Total	66	3.45	0.87	1.902(.031)
Less than 2 hours	40	3.52	0.81	
More than 2 hours	26	3.33	0.75	

Table 4. Differences in level of interest about conversational agent: Based on conversation time with NOK(Next Of Kin)

Conversation Time with NOK	Number	Mean	Standard Deviation	F(p)
Total	66	3.45	0.87	0.13(.876)
Less than 1 hours	27	3.55	0.74	
1-2 hours	25	3.42	0.88	
More than 2 hours	14	3.44	0.74	

	문 항	매우 그렇다	그렇다	보통이다	아니다	매우 아니다
Degree of Interest	1. 나는 이번 체험 활동이 기대되었다.					
	2. ECA가 소개되는 모습을 보고 흥미가 생겼다.					
	3. 체험 후 ECA에 대해 관심이 생겼다.					
	4. 이번에 체험한 ECA을 더 사용해보고 싶다.					
Degree of Completeness	5. 캐릭터가 나의 말에 집중하고 있다는 느낌을 받았다.					
	6. 캐릭터가 하는 말을 어려움없이 이해할 수 있었다.					
	7. 캐릭터의 질문에 답하는 것은 어렵지 않았다					
	8. 캐릭터와의 대화는 자연스럽게 이어졌다.					
	9. 캐릭터와 다양한 대화를 할 수 있었다					
Degree of Satisfaction	10. 이번 활동은 재미있었다.					
	11. 다음에도 캐릭터와의 대화를 또 해보고 싶다.					
	12. 이번 활동에 만족한다.					
	13. 캐릭터와의 대화는 즐거웠다.					
Action Possibility	14. 나는 캐릭터와의 대화에 적극적으로 참여했다.					
	15. 캐릭터와의 대화 중 기억나는 주제가 있다.					
	16. 나는 캐릭터의 물음에 열심히 답했다.					
	17. 나는 캐릭터와의 대화에 집중했다.					
Degree of Interest about Conversational Agent	18. 나는 대화형 에이전트에 관심이 있다.					
	19. 기회가 된다면 대화형 에이전트에 대한 활동을 더 체험해보고 싶다.					
	20. 대화형 에이전트를 계속 사용해보고 싶다.					

Fig. 9. Images of official investigation questionnaire.

Table 5. Result of ECA system survey

		Mean	Standard Deviation	Cronbach's Alpha
Degree of Interest		3.67	0.93	0.78
Degree of Completeness	Appropriate Word	4.12	0.71	0.65
	Dialogue Flow	3.65	0.87	0.73
Degree of Satisfaction		3.71	0.88	0.88
Action Possibility		3.77	0.83	0.79
Degree of Interest about Conversational Agent		3.67	0.95	0.85

Table 6. Development of interest in conversational agent

Interest in Conversational Agent	Mean	SD	t(p)
Preliminary investigation	3.45	0.87	5.659(.000)
Official investigation	3.67	0.95	

달에 집중한 콘텐츠에 익숙해지기 때문이라 추측된다. 이에 대한 정확한 원인을 알기 위해서는 추가적인 연구 및 설문이 진행될 필요가 있다. 또한, 예비조사와 본 실험 사이의 흥미 변화로 미뤄보아 대화형 에이전트가 안정적인 키즈 콘텐츠로 자리 잡기 위해서는 아동에게 보다 만족감을 느끼게 할 경험 제공이 우선적으로 필요하다 해석된다. 그러나 단발성 실험 환경으로는 준비된 대화 주제 수로 지속적인 사용을 얼마나 이어갈 수 있는가를 알 수 없었으며 이에 관한 연구 및 설문이 추가로 필요하다고 사료된다.

REFERENCE

[1] R.R. Romeo, J.A. Leonard, S.T. Robinson, M.R. West, A.P. Mackey, M.L. Rowe, et al., "Beyond the 30-million-word Gap: Children's

Conversational Exposure is Associated with Language-related Brain Function," *Psychological Science*, Vol. 29, No. 5, pp. 700-710, 2018.

[2] Y. Kim, J. Kang, and J. Kim, "A Meta-analysis of the Relationship between Children's Language Ability and Socio-emotional Development," C*ommunication Sciences and Disorders,* Vol. 24, No. 2, pp. 274-287, 2019.

[3] M. Swain and S. Lapkin, "Problems in Output and the Cognitive Processes They Generate: A Step Towards Second Language Learning," *Applied Linguistics,* Vol. 16, No. 3, pp. 371-391, 1995.

[4] 7 Out of 10 are Private Tutoring before Preschool... Go to Cram School from 4.9 Years Old(2019), http://www.donga.com/news/article/all/20190415/95050050/1 (accessed January 08, 2020).

[5] Korea Children Suffers from Private Education since the Age of 2(2017), https://www.mk.co.kr/news/society/view/2017/01/20179/ (accessed January 08, 2020).

[6] Of the Children, by the Children, for the Children(2017), https://brunch.co.kr/@ntvx/ 13 (accessed June 08, 2020).

[7] W. Hwangbo and M. Wui, "A Study on Digital Media Literacy Growth Possibility for Children : Focusing on Lacan's Desire Theory," *Korea Multimedia Society*, Vol. 20, No. 2, pp. 420-428, 2017.

[8] H. Johnson, L. Nigay, and C. Roast, *People and Computers XIII: Proceedings of HCI '98,* Springer Publishers, London, 1998.

[9] E. Mower, M.P. Black, E. Flores, M. Williams, and S. Narayanan, "Rachel: Design of an Emotionally Targeted Interactive Agent for Children with Autism," *Proceedings of IEEE International Conference on Multimedia and Expo*, pp. 1-6, 2011.

[10] T. Kim and K. Baek, "A Case Study of Phonological Process of Korean Preschool Children," *Korean Language and Literature in International Context*, Vol. 49, No. 49, pp. 7-34, 2010.

[11] L. Ring, D. Utami, and T. Bickmore, "The Right Agent for the Job? The Effects of Agent Visual Appearance on Task Domain," *Lecture Notes in Computer Science*, Vol. 8637, pp. 374-384, 2014.

[12] G.H. Hildreth, "Color and Picture Choices of Young Children," *Journal of Genetic Psychology,* Vol. 49, No. 2, pp. 427-435, 1936.

[13] J. Choi and J. Baek, "The Research on the Preference about Anthropomorphic Animal Character's Visual Communication Techniques of the Picture Book," *Journal of Integrated Design Research*, Vol. 8, No. 1, pp. 129-141, 2009.

[14] S.A. Brasel and J. Gips, "Enhancing Television Advertising: Same-language Subtitles can Improve Brand Recall, Verbal Memory, and Behavioral Intent," *Journal of the Academy of Marketing Science*, Vol. 42, No. 3, pp. 322-336, 2014.

최 지 영

2016년~현재 숭실대학교 IT대학 글로벌미디어학부 전공
관심분야 : HCI, 딥러닝, 멀티미디어

정 기 철

1996년 경북대학교 컴퓨터공학과 공학석사
1999년 방문 연구원, Intelligent User Interfaces Group, DFKI(The German Research Center for Artificial Intelligence GmbH), Germany.
1999년 ETL 방문 연구원, Machine Understanding Division, Electro Technical Laboratory, Japan.
2000년 경북대학교 컴퓨터공학과 공학박사
2000년~2002년 박사후 연구원, PRIP Lab., Michgan State University, US.
2003년~현재 숭실대학교 IT대학 미디어학부 교수
2009년~2010년 SIAT (School of Interactive Art and Technology, SFU, Canada) 방문 교수
관심분야 : 인공지능, 컴퓨터비전, HCI