

얼굴형 인공지능 스피커에 대한 선호의 나이 효과*

오 성 주 황 지 현 유 지 호 한 소 원[†]
서울대학교 심리학과

본 연구에서 얼굴형 인공지능 스피커에 대한 선호도와 신뢰도가 나이에 따라 어떻게 달라지는지 살펴보았다. 인공지능 스피커의 외형은 성별(남녀)과 나이(20/60대) 네 가지 범주에서 로봇 얼굴과 사람 얼굴이 7 단계로 혼합되어 제시되었다. 참여자들 역시 이 네 가지 범주에 맞춰 모집되었고, 각 범주의 얼굴 자극에 대한 호감도와 신뢰도를 7점 척도로 평가하였다. 연구 결과, 인공지능 스피커의 외형이 사람 얼굴을 닮을수록 노인들은 호감도와 신뢰도를 높게 평가했지만, 청년들은 로봇 얼굴과 사람 얼굴이 중간 정도 섞인 형태를 가장 선호하고 신뢰하여 전체적으로 뒤집힌 U자 형태의 평가를 보였다. 네 가지 얼굴 범주에서 선호도를 평가했을 때, 참여자들은 나이와 성별과 관계없이 젊은 얼굴이 가장 높은 점수를 받았다. 그런데, 이모티콘 얼굴과 얼굴이 없이 비어있는 조건을 추가하여 물었을 때, 여전히 노인들은 사람 얼굴을 가장 선호했지만, 청년들은 상대적으로 이모티콘 얼굴과 비어있는 조건을 더 선호하였다. 종합하면, 인공지능 스피커 맥락에서, 노인들은 청년들보다 로봇 얼굴이 아닌 사람 얼굴에 높은 수용적 태도를 보였다. 홀로 사는 노인 인구가 증가하는 미래에 인공지능 스피커의 역할이 중요해짐에 따라, 노인들을 위한 인공지능 스피커의 디자인과 개발에서 본 연구 결과가 좋은 참고가 될 것으로 기대된다.

주제어 : 인공지능 스피커, 얼굴 지각, 노인, 인터페이스 디자인

* 이 논문은 서울대학교 2016년도 융·복합 연구과제 지원사업의 지원을 받아 수행된 연구임.

[†] 교신저자: 한소원, 서울대학교 심리학과, 서울시 관악구 관악로 1

연구 분야: 지각, 인지, 정서 심리학

Tel: 02-880-8618, E-mail: swhahn@snu.ac.kr

현대 사회에서 주거의 형태는 대가족 중심이 아닌 개인 중심으로 바뀌고 있다. 따라서, 개인들의 사회적 관계도 과거 물리적인 현장 중심에서 사회관계망 같은 온라인으로 바뀌고 있다. 컴퓨터와 스마트폰의 사용에 익숙한 젊은 성인들은 주거 형태의 변화에도 온라인에서 사회적 관계를 유지하며 살 수 있다. 반면, 노인들은 IT 기기에 적응을 쉽게 하지 못하여, 물리적으로도 사회적으로도 고립되어 살 가능성이 커지고 있다.

노인들의 고립적인 삶은 여러 가지 문제를 안고 있다. 첫째, 신체적 기능이 약화된 노인들은 신체 활동이 필요한 일상적인 과제들에 적절히 대처하지 못할 수 있다. 설거지, 빨래, 요리, 식사 등 집안일부터 시장, 은행, 우체국, 병원일 등 바깥 일까지 해야 할 일들은 다양하다. 둘째, 노인들의 인지 기능의 저하는 치명적인 위험을 초래할 수 있다. 예를 들어, 판단력 부족으로 보이스피싱 사기를 당할 수 있으며, 가스버너 사용에서 화재를 불러올 수 있다. 셋째, 노인들은 대인 관계의 단절로 사회적 지지를 받지 못해 외로움과 우울감에 시달릴 수 있다. 전 세계적으로 노인 인구의 증가를 고려하면, 노인들이 사회적 관계에서 단절되지 않는 방안의 마련이 절실하다. 이러한 상황에서 로봇은 고립된 노인들의 생활에 훌륭한 대안이 될 수 있다[1]. 로봇은 노인의 신체적, 인지적, 사회적 약점을 보완하여 여러 방면에서 노인들의 삶을 도울 수 있을 것이다.

이전 연구들에서, 인간을 닮은 휴머노이드 로봇은 그렇지 않은 로봇에 비교해 사회적, 감정적, 신체적, 대화 측면에서 월등한 장점이 있다고 제안되었다[2, 3, 4, 5]. 실제로 사람을 닮은 로봇이 사람을 닮지 않은 로봇보다 선호되며 사람을 닮은 로봇일수록 사람이 가지고 있는 성격 및 심리적 특징들을 더 가지고 있다고 평가받는다[6, 7, 8]. 사람들은 머리 부분에 스크린이 달린 로봇과 혈압을 측정하는 간단한 상호작용을 할 때, 스크린에 사람 얼굴이 나타날 때 사람 얼굴이 없거나 금속성 얼굴(silver face)이 나올 때 비해 로봇이 좀 더 사회적(social)이고 정감 있는(amicable) 감정을 느꼈다[9]. 자폐 아동들은 얼굴이 없는 로봇보다 얼굴을 가진 로봇과 놀기를 선호한다[10]. 얼굴만으로 이뤄진 로봇 연구에서, 목소리가 실제 사람과 닮을수록 참여자들은 더 큰 호감을 느낀다[11]. 이 연구들은 로봇이 사람과 닮은 특징을 가질수록 사용자들의 선호가 증가함을 시사한다.

그런데 이 연구들에서 로봇의 외형은 매우 단순한 범주로 표현된다. 예를 들어, 사람 얼굴 특징이 전혀 없는 기계의 모습, 동물을 닮은 모습, 사람의 얼굴 모습 등 불과 두세 범주의 외형들이 비교된다. 그렇지만, 기괴한 계곡(uncanny valley) 현상에서 밝혀졌듯이, 평가 대상들이 생물체의 한 범주 내에서 사람을 닮은 정도를 자세하게 나누어져 제시되었을 때 사람을 닮은 캐릭터 또는 로봇의 얼굴이 대체로 선호되지만, 너무 비슷한 얼굴은 정서적으로 불안을 일으킨다[12]. 그러므로 제시되는 얼굴들이 얼마나 범주적으로 다르게 묘사되느냐에 따라 결과가 달라질 수 있다. 한편, 평가자의 나이, 문화적 배경 등 평가자 요인 역시 중요하다. 중국인과 한국인은 독일인보다 사람을 닮은 로봇과 상호작용하는 것을 더 선호하고, 신뢰하고, 만족하는 경향이 있다[7]. 또한, 노인들은 젊은이들보다 로봇의 얼굴이 사람처럼 생겼을 때 더 호감을 느끼는 것으로

나타났다[13]. 이 연구들을 종합하면, 로봇의 얼굴 요인과 평가자 요인이 복합적으로 작용하여 사람을 닮은 로봇에 대한 선호가 결정된다고 볼 수 있다.

종합하면, 홀로 사는 노인들이 증가하는 지금의 추세에서, 사람을 닮은 로봇은 매우 매력 있는 대안이 될 수 있다. 이를 통해, 노인들은 최소한의 사회활동을 할 수 있고, 스스로 할 수 없는 가사에 대해 도움을 받아 좀 더 행복한 삶을 누릴 수 있을 것이다. 그런데 아직 로봇은 보통의 노인들이 구매하기에는 너무 비싸다. 또한, 아직 인간과 의사소통을 원활하게 할 정도로 높은 수준의 기능을 발휘하지 못하고 있다. 현실적으로 인공지능 스피커는 로봇 사용이 대중화되기 전까지 좋은 대안이 될 수 있다. 인공지능 스피커는 정밀한 도구적 기능은 발휘하지 못하지만, 간단한 대화를 바탕으로 간단한 정보기술 서비스를 제공하며, 몹시 나쁘지 않은 대화 상대가 될 수 있다.

최근 들어 IT 기기에 인공지능 기술을 접목하여 의사소통이 가능한 여러 인공지능 스피커들이 앞다투어 개발되고 있다. Apple의 Siri, Samsung의 Bixby, MicroSoft의 Cortana처럼 특별한 외형 없이 소프트웨어 형식으로 스마트폰이나 컴퓨터에 설치되어 작동하는 것들이 있으며, Amazon의 Echo, Google의 Home, SKT의 Nugu, KT의 GIGAGenie처럼 스피커의 형태를 갖춘 채 독립적으로 기능하는 것들이 있다. 본 연구의 목적은 인공지능 스피커에 주목하여 인공지능 스피커의 외형에 대한 노인들의 잠재적 선호를 탐색적으로 살펴보는 데 있다. 그래서 인공지능 스피커의 얼굴 특징이 어떨 때 노인들이 가장 선호하는지를 살펴보고자 한다. 이 작업은 노인들을 위한 인공지능 스피커의 제품 개발에 좋은 단서를 제공할 것으로 기대된다.

본 연구의 물음은 구체적으로 다음과 같이 정리된다. 인공지능 스피커 맥락에서 사람 얼굴이 선호되는가 로봇 얼굴이 선호되는가? 얼굴 외형을 가진 인공지능 스피커의 나이와 성별 그리고 평가자의 나이와 성별에 따라 선호도에서 차이가 있는가? 특히, 외국의 이전 연구에서 밝혀진 것처럼[13], 우리나라의 노인들도 젊은 사람들보다 사람을 닮은 로봇을 더 좋아하는지를 살펴보고자 하였다.

실 험

방 법

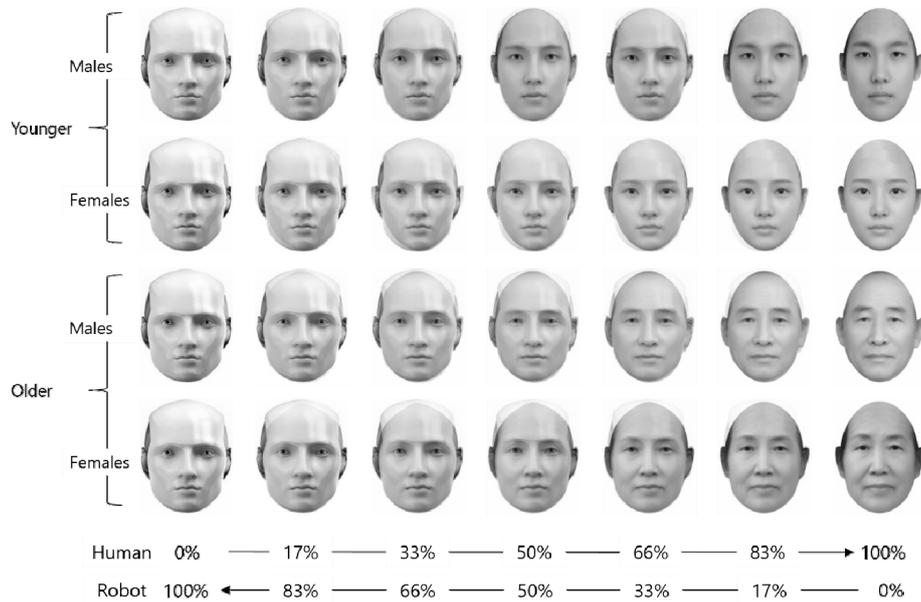
참가자

참여자들은 35명의 서울대학교에 재학 중인 대학생들(17명 여성)과 105명의 서울에 거주하는 노인들(57명 여성)이었다. 대학생들의 나이는 만 20에서 34까지였고($M=22.43$, $SD=3.13$), 노인들의

나이는 58에서 81까지였다($M=68.55$, $SD=6.14$). 참여자들 모두 스마트폰을 사용하고 있었다. 본 연구는 서울대학교 IRB 승인을 받았다.

자극 및 기구

사람 얼굴 정도(human faceness)를 조작하기 위해 하나의 로봇 얼굴과 4개의 사람 얼굴(20대 남녀 2개; 60대 남녀 2개)이 준비되었다. 사용된 로봇 얼굴은 로봇 모델(Eset Co.)이었으며 우리나라에는 잘 알려지지 않았다. 사람 얼굴 원본 4개는 각각 비슷한 나잇대와 같은 성의 5개의 얼굴이 합성되어 준비되었다. 로봇 얼굴과 사람 얼굴의 모형은 양극단에 두고 그사이를 FantaMorph 5 (Abrosoft Co.)를 이용해 각각 16.5%의 차이를 두고 각 합성되어 5장의 중간 얼굴들이 만들어졌다. 따라서, 얼굴 자극은 네 범주의 얼굴 세트이며 각 세트는 7장씩 모두 28개였다. 얼굴 자극들은 모두 중립 표정으로 정면을 바라보고 있었으며 흑백으로 제시되었다(그림 1). 각 얼굴은 가로 11cm, 세로 15cm 크기로 프린트되거나 화면에 제시되었다. 모든 참여자는 총 28장의 카드 자극을 보았으며, 청년층 남성, 청년층 여성, 노년층 남성, 노년층 여성 조건별로 7장의 카드가 섞여 동시에 제시되었다.



(그림 1) 평가에 사용된 얼굴 자극. 위부터 아래로, 20대 남성, 20대 여성, 60대 남성, 60대 여성이다. 각 세트에서 양 끝은 사람 얼굴과 로봇 얼굴의 원본이며 중간은 이 두 얼굴의 합성이다.

절차 및 설계

참여자는 연구 참여 동의서에 먼저 서명을 하였다. 인공지능 스피커에 대한 이해를 도우려고 SKT에서 출시한 인공지능 스피커 NUGU의 실물을 제시되었고, 기능에 대한 설명이 제공되었다. 또한, 참여자들이 인공지능 스피커가 얼굴의 외형을 가졌을 때를 쉽게 상상하기 위해서 NUGU에 로봇의 얼굴을 입힌 모습이 출력되어 제시되었다(그림 2). 과제에 대한 지시문은 구체적으로 다음과 같았다.

“인공지능 스피커 이미지를 이용한 카드 분류를 해보겠습니다. 최근 대화를 통해 원하는 일을 시킬 수 있는 인공지능 스피커가 출시되었습니다. 아래의 사진과 같은 원통형 모양입니다. 음악과 라디오를 듣는 미디어 기능, 집의 조명, 가스, 보일러 등을 조절하는 스마트 홈 기능, 음식 주문과 배달, 길 찾기뿐 아니라 알람, 날씨, 일정 확인 등을 도와주는 개인 비서의 역할을 합니다.”



(그림 2) SKT의 NUGU와 실험에서 제시된 얼굴형 인공지능 스피커의 예

노인 참여자 모두와 청년 참여자 35명 가운데 25명은 사진이 인쇄된 카드를 대상으로 작업을 수행하였고, 나머지 청년 10명은 컴퓨터 화면을 통해 얼굴 자극들을 제시받았고 마우스와 키보드를 이용해 과제를 수행하였다.

가장 먼저, 참여자들은 본 과제를 수행하기 전에 각 카드에 출력된 얼굴 이미지를 보고 나이를 추정하는 작업을 하였다. 이는 제작된 얼굴이 실제로 예상된 나이로 지각되는지를 확인하기 위해서였다. 두 번째로, 얼굴에 대한 호감도(likeability)에 따라 카드를 분류하는 작업을 수행하였다. 세 번째로, 얼굴에 대한 신뢰도(trustworthiness)에 따라 카드를 분류하는 작업을 수행하였다. 이때, 인공지능 스피커가 제공하는 사실적 정보에 대해 믿음이 얼마나 큰지를 구체적인 예를 들어 판단하게 하였다. 즉, 인공지능 스피커로부터 날씨, 가격, 교통 등의 정보를 받으면 그 정보

를 가장 믿을 수 있는 이미지를 1위로 가장 믿기 어려운 이미지를 7위로 순위를 정하였다. 얼굴 연구에서 호감도와 신뢰도는 유사하지만, 완전히 일치하지는 않는 개념들이다[14]. 또한, 로봇과 상호작용 맥락에서 호감도는 제품 구매와 상관이 높고 신뢰도는 제품이 주는 정보에 관한 것이므로 차별적인 성질을 가지고 있다[15].

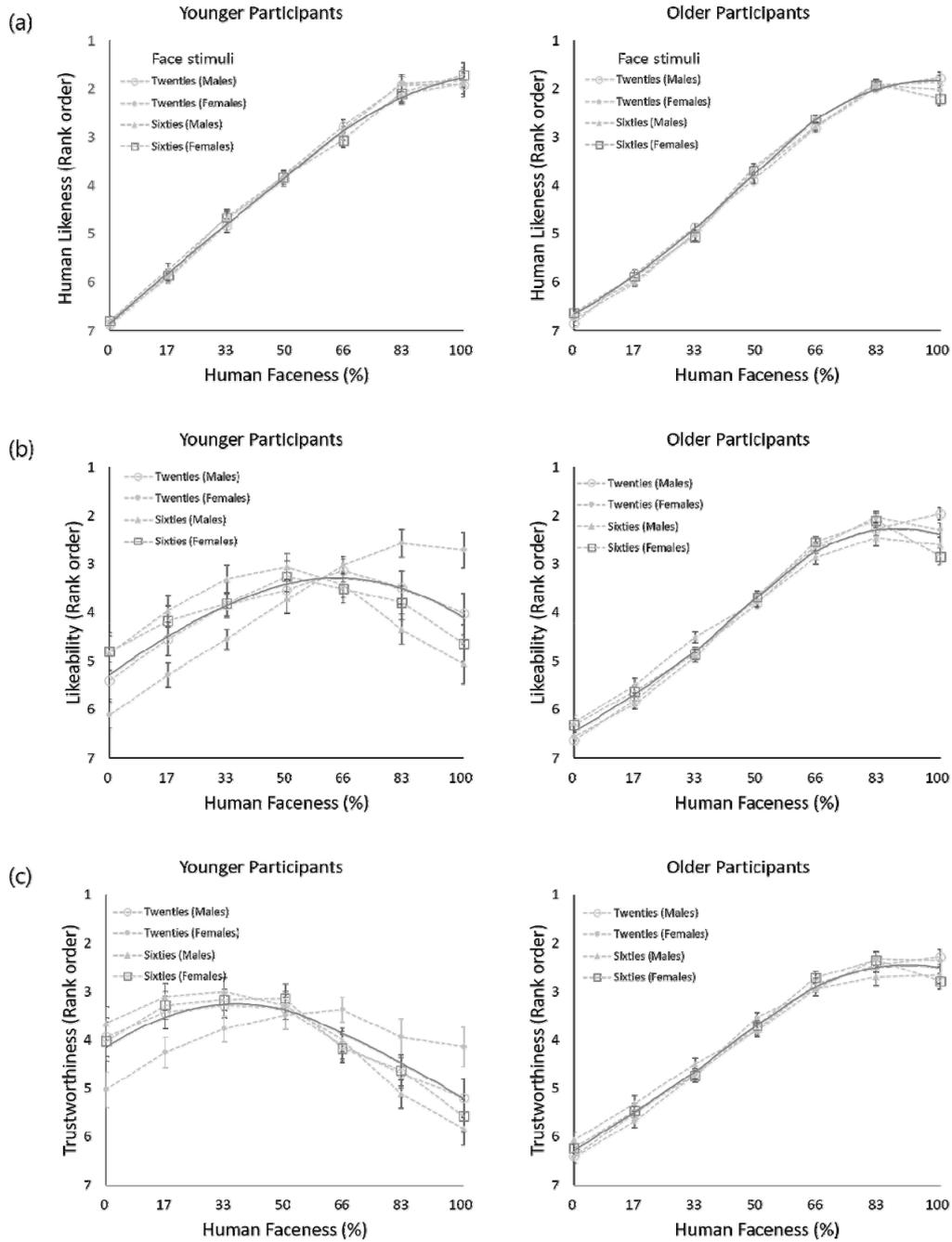
각 세트 내에서 얼굴에 대한 평가가 끝난 후, 네 범주의 얼굴을 대상으로 인공지능 스피커의 외형으로 선호하는 순서를 나열하는 작업이 이어졌다. 예를 들어, 여성 청년-남성 청년-여성 노인-남성 노인 순서가 될 수 있다. 이 과제 다음으로, 참여자들은 이 네 세트의 얼굴 이외에 단순한 이모티콘 얼굴(그림 6)과 아무것도 없는 빈 조건을 추가하여 모두 6가지 조건을 대상으로 가장 선호하는 것 하나만 고르는 과제를 수행하였다. 마지막으로, 참여자들은 인공지능 스피커를 떠나서 가정에 로봇 비서가 있다면 일상생활에서 어떤 도움을 받고 싶은지에 대해서 순위를 평가하였다. 구체적으로 5가지 범주로 항목들은 묶었는데 다음과 같았다. 범주 1: 집안일과 요리, 범주 2: 범죄와 재해 예방, 범주 3: 외출과 운동 보조, 범주 4: 대화, 여가, 그리고 취미 생활, 범주 5: 간호, 의료와 안마. 이 질문을 묻은 이유는 참여자들이 인공지능 스피커를 포함한 인공지능 장치에 대한 요구를 이해하기 위해서였다.

결 과

먼저, 얼굴 자극에 대한 지각된 나이를 분석하였다. 연구자가 의도한 바와 같이 참여자들은 20대 얼굴에 대해서 20대로 평정하였고(남성: $M=27.77$, $SD=6.68$, 여성: $M=23.21$, $SD=4.56$), 60대 얼굴은 60대로 평정하였다(남성: $M=64.20$, $SD=7.22$, 여성 자극: $M=64.01$, $SD=7.73$), $t(420)=79.80$, $p < .001$). 또한, 사람 얼굴 비율이 높을수록 실제로 사람 얼굴스러운지를 살펴보았다. 그 결과, 그림 3a에서 보듯이, 전체 참여자 집단에서 높은 상관을 보였다, $r = 0.88$, $t(3946)=114.56$, $p < .001$. 이러한 상관은 청년 참여자 집단과 노인 참여자 집단에서 비슷하였다, 각각, $r = 0.88$, $t(978)=57.09$, $p < .001$; $r = 0.88$, $t(2966)=99.29$, $p < .001$. 정리하면, 참여자들은 20대/60대 얼굴 자극을 젊다고/나이 들었다고 지각하였다. 이는 연구 목적에 맞게 얼굴 자극이 적절하게 구성되었음을 보여준다.

참여자 나이에 따른 얼굴 자극 호감도와 신뢰도

자료가 서열 척도를 통해 측정되었으므로, 각 얼굴 자극을 얼마나 선호하는지를 확인하기 위해 R 3.3.1 버전의 nparLD 패키지를 통해[15], 참여자들의 선호도를 대상으로 2x7 비모수 혼합변량분석을 시행하였다. 분석 결과, 참여자 나이의 주효과가 유의하였으나, $F_{ATS}(1, \infty)=6.04$, $p <$



(그림 3) 네 얼굴 자극 세트의 각 얼굴 자극에 대한 젊은이와 노인 참여자들의 평정. 가로축은 사람 얼굴이 섞인 비율을 의미하며, 세로축은 평정된 순위의 평균이다. (a) 얼마나 실제 사람 얼굴 같은가에 대한 평정. (b) 호감도 평정. (c) 신뢰도 평정. 점선은 자극 범주 별 평균 순위이며, 실선은 전체 평균 순위이다. Error bars indicate ± 1 SEM

.05, 효과 크기가 극히 작아(Cliff's delta=0.01) 실질적인 의미는 없는 것으로 확인되었다. 이는 모든 참여자가 자극에 대해 1위부터 7위까지 순위를 매기는 측정 방식을 고려하면 당연한 결과이다. 사람 얼굴 비율의 주효과는 유의하였으며, $F_{ATS}(6, \infty)=211.06, p < .001$, 참여자 나이와 사람 얼굴 비율의 상호작용 또한 유의하였다, $F_{ATS}(6, \infty)=36.75, p < .001$. 즉 사람 얼굴 비율에 따라 선호도가 변화하였는데, 청년과 노인 참여자들은 서로 다른 변화 경향을 보였다. 그림 3b에서 보듯이, 노인 참여자들은 사람 얼굴 비율이 증가함에 따라 더 높은 호감도를 보이는 경향을 일관되게 보인 반면, 청년 참여자들은 사람 얼굴 비율이 중상 정도로 높아졌을 때 가장 높은 호감을 느끼는 경향을 보임을 알 수 있다. 국지회귀 평활화 기법을 통해 선호도가 최상인 사람 얼굴 비율을 추정해보았을 때 노인들은 약 88%, 청년들은 약 64% 지점에서 가장 높은 호감도를 보였다. 그래프를 보면 대부분의 자극 범주가 이와 유사한 경향을 보이는 데 반해, 20대 여성 자극에 대한 청년 참여자들의 응답은 다소 다른 경향을 보인다. 그러나 일반화 가능성을 고려하여 본 연구에서는 범주별 선호도 모형을 별도로 구성하지는 않았다.

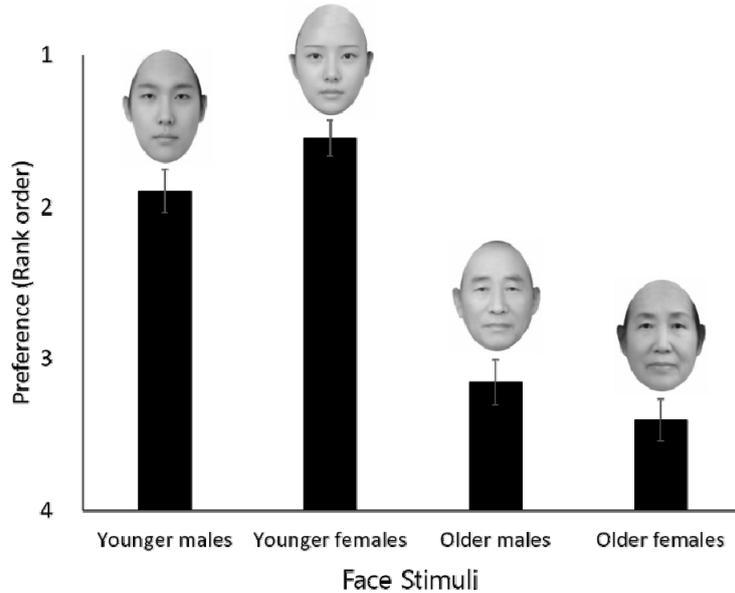
얼굴에 대한 신뢰도 평가를 분석하기 위해, 참여자들이 평정한 신뢰도에 대해 2x7 비모수 혼합변량분석을 시행하였다. 분석 결과, 참여자 나이의 주효과가 유의하였으나, $F_{ATS}(1, \infty)=19.73, p < .05$, 마찬가지로 효과 크기가 극히 미미하여(Cliff's delta = 0.005) 실질적인 의미는 없었다. 사람 얼굴 비율의 주효과는 유의하였으며, $F_{ATS}(2.12, \infty)=21.49, p < .001$, 참여자 나이와 사람 얼굴 비율의 상호작용 또한 유의하였다, $F_{ATS}(2.12, \infty)=49.20, p < .001$. 다시 말해서 사람 얼굴 비율에 따라 신뢰도가 변하였는데, 청년과 노인 참여자들은 서로 다른 변화 경향을 보였다. 그림 3c에서 보듯이, 노인 참여자들은 사람 얼굴 비율이 증가함에 따라 더 신뢰하는 경향을 일관되게 보인 반면, 청년 참여자들은 사람 얼굴 비율이 중하 정도로 높아졌을 때 가장 신뢰하는 경향을 보임을 알 수 있다. 국지회귀 평활화(LOESS Smoothing) 기법을 통해 신뢰도가 최상인 사람 얼굴 비율을 추정해보았을 때 노년층은 약 91%, 청년층은 약 36% 지점에서 가장 높은 신뢰도를 보였다. 그래프를 보면 대부분의 자극 범주가 이와 유사한 경향을 보이는 데 반해, 20대 여성 자극에 대한 청년층 참여자의 응답은 다소 다른 경향을 보인다. 그러나 일반화 가능성을 고려하여 본 연구에서는 범주별 신뢰도 모형을 별도로 구성하지는 않았다.

전반적으로 노인들은 호감도와 신뢰도를 매우 높은 상관으로 비슷하게 평가하였다($r = .99$). 반면, 청년들은 노인들보다 서로 다르게 평가하였다($r = .55$). 이는 노인들이 과제를 정확히 이해하지 못했을 수도 있고, 이해하였는데 두 개념을 통합적으로 받아들였기 때문일 수 있다.

참여자 성별에 따른 호감도와 신뢰도

참여자 나이 대신 참여자 성별에 따라 집단을 나누어 지각된 사람다움, 호감도, 신뢰도에 대해 같은 분석을 시행하였을 때는 여성 참여자가 남성 참여자보다 사람 얼굴 비율 100% 자극의

사람다움을 좀 더 높게 지각한 것을 제외하면, $H(1)=10.95, p < .001$, 유의미한 성차가 발견되지 않았다. 즉 성별이 제품의 외형에 대한 호감에 일반적으로 많은 영향을 끼침에도 불구하고, 인공지능 스피커의 외형에 대한 호감에서 사용자의 나이가 미치는 영향은 그보다 훨씬 뚜렷했다고 할 수 있다.



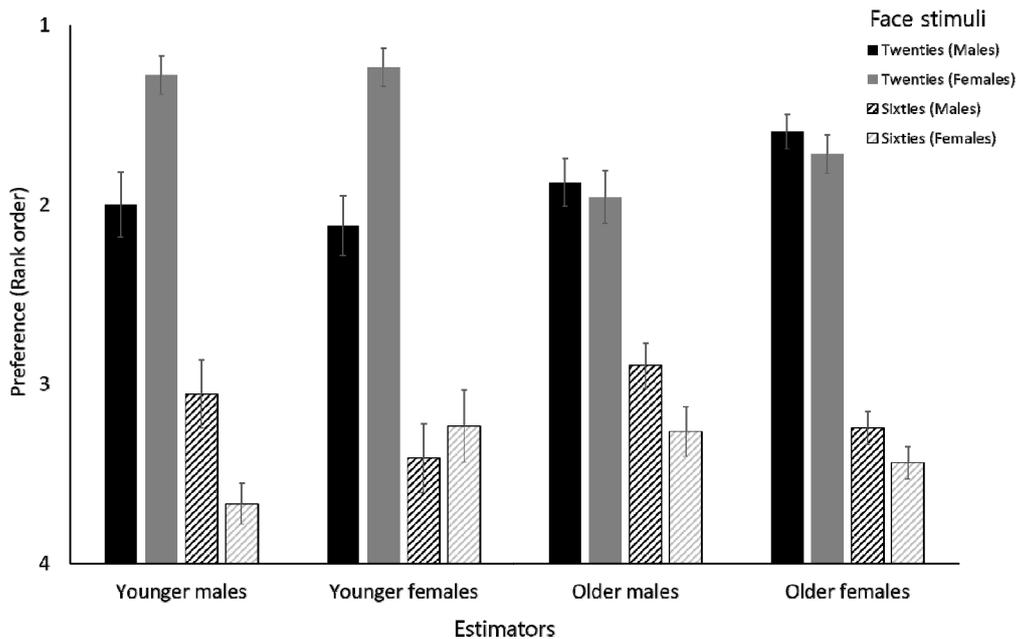
(그림 4) 네 얼굴 자극 세트에 대한 인공지능 스피커 외형으로서 호감도.
Error bars indicate ± 1 SEM

얼굴 자극 범주간 선호도

네 종류의 얼굴 자극 세트 가운데 선호도 분석하기 위해, 네 얼굴에 대한 선호 순위를 평균 내었다. 그림 4에서 보듯이, 20대 여성($M=1.55, SD = 1.39$), 20대 남성($M=1.90, SD=1.71$), 60대 남성($M=3.15, SD=1.77$), 60대 여성($M=3.40, SD=1.61$) 순이었다. 얼굴 자극의 종류와 참여자 나이의 관계를 파악하기 위해, 2(자극 나이)×2(자극 성별)×2(참여자 나이) 비모수 혼합변량분석을 시행하였다. 그 결과, 얼굴 자극의 나이에서 주효과가 나타났다, $F_{ATS}(1, \infty)=386.12, p < .001$. 이에 대해 Wilcoxon의 부호화된 순위 검정을 시행한 결과, 20대 얼굴 자극에 대한 호감도가 60대 자극에 대한 호감도에 비해 유의하게 높았다, $p < .001$. 참여자 나이의 주효과나 자극 성별의 주효과는 나타나지 않았다, 각각 $F_{ATS}(1.20, \infty)=0, p > .05$; $F_{ATS}(1, \infty)=0.34, p > .05$. 즉, 그림 5에서 볼 수 있듯이, 청년 참여자와 노인 참여자 모두 20대 얼굴 자극을 일관되게 더 선호하였다. 참여자 나이와 자극 성별의 이원상호작용효과가 나타났다, $F_{ATS}(1, \infty)=8.17, p < .01$. 또한, 얼굴 자극 나

이와 자극 성별의 이원상호작용효과도 유의미하였다, $F_{ATS}(1, \infty)=12.93, p < .001$. 이 결과들은 참여자의 나이에 따라 선호하는 얼굴 자극의 성별이 다름을 보여주는데, 청년 참여자들은 성별에 관계없이 모두 여성 얼굴 자극을 선호하였지만, 노인층에서 남녀 모두를 비슷하게 선호하였다. 그리고 참여자 나이, 자극 나이, 자극 성별의 삼원상호작용효과가 유의하였다, $F_{ATS}(1, \infty)=6.64, p < .01$.

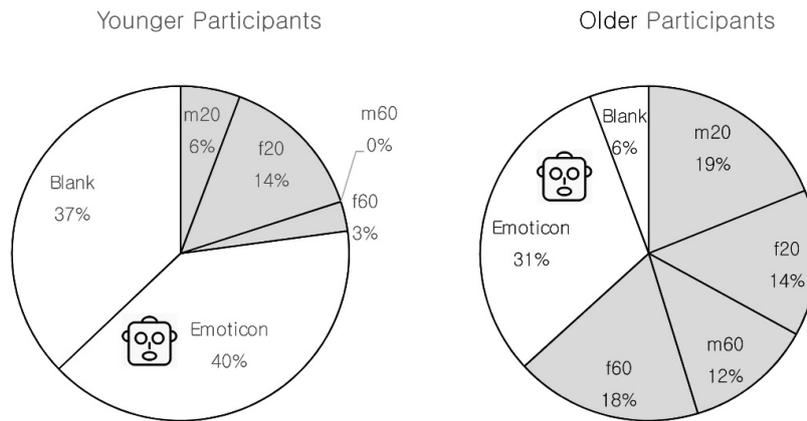
이 결과는 같은 나이대의 얼굴에 더 많은 관심과 주의를 가진다는 이전 연구들의 발견과 다르다[17, 18]. 아마도 참여자들은 인공지능 스피커의 맥락에서는 얼굴 선호 평가에서 관심과 주의보다는 얼굴 매력을 더 중요하게 생각했을 가능성이 있다. 즉, 일반적으로 평가자의 나이와 무관하게 젊은이의 얼굴이 노인의 얼굴보다 더 매력적으로 평가되는 경향이 있다[19, 20]. 따라서, 인간 사회에서 의사소통의 기회가 많은 비슷한 나이의 얼굴보다는 자기 소유로 두고서 상호작용할 인공지능 스피커의 얼굴을 고를 때는 얼굴의 매력도가 중요한 기준이었을 가능성이 있다.



(그림 5) 네 얼굴 자극 세트에 대한 선호도를 평정자들의 나이와 성별에 따른 분석 결과. Error bars indicate ± 1 SEM

다음으로, 네 가지 종류의 얼굴 이외에 아무것도 없는 빈 형태와 이모티콘 형태를 넣었을 때 선호도 차이가 분석되었다(그림 6). 인간 형태 자극(20대 남성, 20대 여성, 60대 남성, 60대 여성)과 비인간 형태 자극(이모티콘, 빈칸)의 선호 비율에 나이 차이가 있는지 확인하기 위해 카이제곱

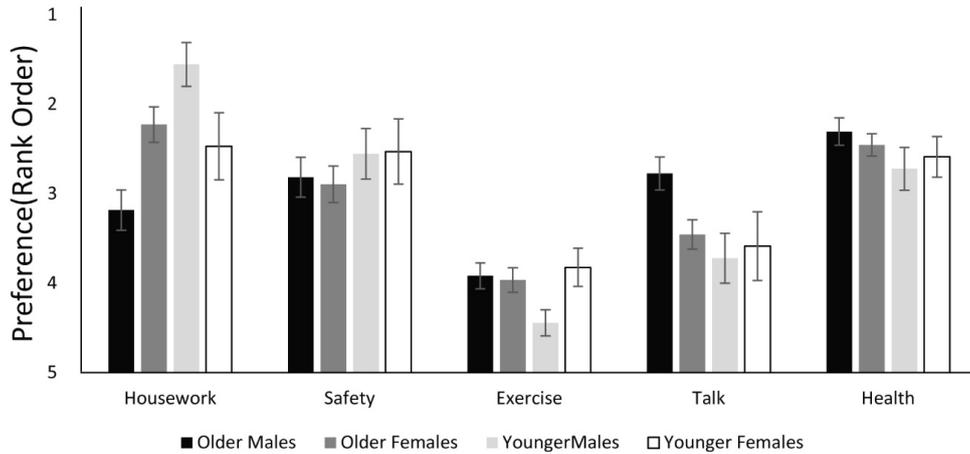
검정을 시행하였다. 그 결과 청년 참여자들은 노인 참여자들에 비해 비인간 형태 자극을 선호하는 비율이 높음이 확인되었다, (1, N=141)=17.21, $p < .001$. 흥미롭게도, 사람 얼굴과 비사람 얼굴로 나누어 보았을 때, 노인들은 인공지능 스피커의 외형이 사람이기를 원하는 비율이 63%로 높지만, 젊은이들은 불과 23%로 나타났다. 이는 노인들이 청년들보다 사람 얼굴을 압도적으로 선호함을 보여준다.



(그림 6) 네 종류의 얼굴 자극에 이모티콘 얼굴과 빈 형태를 추가하여 가장 선호하는 것을 고르라고 했을 때 참여자 나이에 따른 선호비율. 회색 음영은 사람 얼굴에 대한 비율이다. M20은 청년 남성 얼굴, f20은 청년 여성 얼굴, m60은 노인 남성 얼굴, f60은 노인 여성 얼굴을 의미한다.

로봇 비서의 필요성

로봇 비서의 역할로 선호하는 순위에 나이와 성별에 따른 차이가 있는지 확인하였다(그림 7). 2(참여자 나이)×5(역할 범주) 비모수 혼합변량분석을 실시한 결과, 역할 범주의 주효과, (3.25, ∞)=28.11, $p < .001$, 참여자 나이와 역할 범주의 상호작용효과가 유의미하였다, (3.25, ∞)=3.24, $p < .05$. 즉, 각 역할 범주마다 선호되는 정도가 달랐고, 참여자 나이 별로 선호하는 역할 범주도 서로 달랐다. 동일 연령층 내에서 성차를 확인하기 위해 Kruskal-Wallis 검정을 한 결과, 가사 범주에서는 여성 노인이 남성 노인보다 선호도가 높았고, $H(1)=9.94$, $p < .01$, 남성 청년이 여성 청년보다 선호도가 높은 것으로 분석되었다, $H(1)=3.61$, $p = .057$. 대화 범주에서는 남성 노인의 상대적 선호가 나타났다, $H(1)=7.17$, $p < .001$. 그리고 운동 범주에서는 여성 청년의 상대적 선호가 나타났다, $H(1)=4.76$, $p < .01$. 대체로 로봇에게 기대되는 역할은 참여자 사이에서 비슷했지만, 사회적 관계를 포함하는 대화에서 노인들의 기대가 높았다는 점은 노인들의 고독한 삶을 반영되었을 가능성이 있다.



(그림 7) 참여자 성별과 나이에 따른 로봇비서 역할에 대한 기대 순위의 평균. Error bars indicate ± 1 SEM

논 의

본 연구에서는 인공지능 스피커가 얼굴 모양일 때 잠재적 사용자들의 평가가 어떤지를 알아보려고 하였다. 특히, 나이에 따라 평가가 어떻게 다른지 살펴보고자 하였다. 이를 위해, 청년과 노인 참여자 모두에게 로봇 얼굴과 사람 얼굴을 합성한 자극들을 제시하여 호감도와 신뢰도를 평정하게 하였다.

자료 분석 결과, 노인들의 선호도와 신뢰도는 사람 얼굴과 비슷한 정도와 함께 선형적으로 증가하였지만, 청년들의 호감도와 신뢰도는 선형적인 관계를 갖지 않았다. 이 결과는 로봇을 대상으로 한 얼굴 선호 연구 결과와 대체로 일치한다[13]. 따라서 이 현상이 서양 문화권에서만 발견되는 것이 아니라 아시아에서도 발견되는 일반적인 것임을 보여준다. 그런데, 외국의 이전 연구에서[13], 노인들의 호감도는 이전 연구에서 ‘U’ 자형을 보였다. 즉, 양극단에 놓인 로봇과 사람 얼굴에 대해서 비슷한 정도로 호감을 느꼈고, 이 둘이 섞인 얼굴에 대해서는 매우 낮은 호감을 보였다. 반면, 본 연구에서는 얼굴 자극이 사람과 비슷할수록 호감도가 선형적으로 증가하였다. 이 두 연구 결과 차이는 아마도 합성된 얼굴의 시작점이 달랐기 때문일 수 있다. 즉, 외국의 이전 연구에서[13], 사용된 로봇 얼굴은 훨씬 더 기계의 모습을 가졌다. 반면, 본 연구에서 제시된 로봇 얼굴은 로봇 얼굴이기는 하지만 상당한 정도로 사람과 비슷하여, 이전 연구에서 로봇과 사람 얼굴이 합성된 정도와 비슷하다.

그렇다면 본 연구에서 드러난 것처럼 나이에 따라 선호도와 신뢰도가 왜 얼굴스러움 정도에 따라 다른 패턴을 보일까? 여기에는 다양한 추론들이 있을 것이다. 가장 먼저, 청년들이 일종의

기괴한 계곡을 노인들보다 더 크게 경험했을 수 있다. 노인들은 기괴한 일들과 대상들을 청년들과 비교하면 많이 경험했을 수 있다. 타인의 주점을 관찰하는 경험은 좋은 예일 것이다. 이전 연구에서, 사람을 닮은 로봇을 자주 보게 되면 점차 기괴함을 덜 느낀다는 보고가 있었다[21]. 따라서 기괴한 대상에 대한 습관화 또는 친숙성 정도가 나이에 따라 달랐기 때문에 본 연구에서 서로 다른 결과가 나왔을 수 있다. 또 다른 이유는 나이에 따라 인공지능 스피커에게 기대하는 바가 달랐기 때문일 수 있다. 본 연구에서, 인공지능 스피커의 역할을 일상생활을 도와주는 비서라고 다소 광범위하게 기술되었다. 노인 참여자들은 사회적인 상호작용을 기대하여 사람과 닮은 외형을 평가 항목에 무관하게 더 선호했지만, 젊은 참여자들은 사회적 상호작용보다는 단순한 정보 제공자와 같은 역할을 기대했기 때문에 이런 역할에는 사람스러운 얼굴이 적절하지 않다고 생각했을 수 있다. 이전 연구에서도 사회적 상호작용이 필요한 과제 맥락에서는 사람을 닮은 로봇이 선호되었지만, 보안원 역할과 같은 안전과 관련된 과제 맥락에서는 기계를 닮은 로봇이 선호되었다[6]. 그런데도 나이가 들수록 로봇에 대한 부정적인 정서를 느끼는 것이 일반적인 경향이라는 점을 생각하면[22], 본 연구 결과는 좀 더 다른 방식의 설명이 필요해 보인다.

노인 참여자들 사이에서 호감도와 신뢰도는 얼굴스러운 정도와 무관하게 비슷한 경향을 보인 반면, 청년 참여자들 사이에서 호감도와 신뢰도는 얼굴스러운 정도에 따라 조금 다른 경향을 보였다. 청년들은 IT 기기들에 노출이 잦아 이 기기들에서 볼 수 있는 인공적인 얼굴들에 대해 여러 가지 심리적 특징들을 민감하게 처리하는 반면, 노인들은 IT 기기에 덜 노출되어 로봇스러운 얼굴에 대해서 심리적 특징들을 세분화하지 못하고 통합적으로 느꼈기 때문일 수 있다. 실제로, 18~38세 사이의 젊은이들이 참여한 이전 연구에서, 참여자들은 사회성 차원과 신뢰성 차원을 별개로 평가하는 것처럼 나타났다[9]. 이 연구에서, 실제 사람과 비슷한 얼굴을 가진 로봇일수록 사회성이 높을 것이라는 평가를 받았지만, 신뢰성 평가에서는 실제 사람 얼굴은 얼굴이 없거나 금속성 얼굴(silver face)을 가진 스크린 로봇보다 차이가 없었다.

인공지능 스피커의 네 범주의 얼굴 이외에 이모티콘 얼굴과 빈 형태를 추가하여 선호를 물었을 때와 로봇 비서 역할에 대한 설문에서 노인 참여자들이 청년 참여자들보다 얼굴 외형을 더 선호하고 대화를 더 기대한다는 응답에서 앞서 제기된 두 가지 의문에 대한 설명이 어느 정도 찾을 수 있을 것이다. 만약 노인들은 청년들보다 고독감을 더 느껴 인공지능 기술이 수행적 기능보다도 사회적 상호작용의 역할을 해줄 것으로 기대한다면, 인공지능 스피커가 좀 더 사람 얼굴과 닮기를 바라는 욕구가 크고 또 그런 얼굴을 선호하고 신뢰할 것이다. 실제로 사회에 소속하고 싶은 욕구가 높은 사람일수록 인공적인 얼굴에서 사람다움을 느낄 가능성이 컸다[23]. 그러므로 노년층의 결과는 현재의 사회 소속감의 맥락에서 이해할 필요가 있다.

종합하면, 본 연구는 노인들을 위한 인공지능 스피커를 만들 때, 좀 더 실제 사람을 닮은 외형을 그리고 젊은 나이의 얼굴 외형을 적용할 것을 제안한다. 본 연구의 한계점은 인공지능 스피커의 얼굴들은 실체가 아닌 그림으로만 제시되었고, 참여자들과 간단한 상호작용 없이 평가되

었다는 점이다. 또한, 자극으로 사용된 얼굴들은 인간 사회에서 볼 수 있는 전체 얼굴을 대표하는 것과 거리가 먼 특수한 사례이다. 그러므로 실제로 인공지능 스피커가 얼굴로 대체되고 상호작용의 기회가 주어졌을 때 그리고 다양한 형태의 얼굴을 이용하여 실험하였을 때 과연 본 연구 결과와 비슷하게 나올지에 대한 검토가 필요할 것이다.

참고문헌

- [1] Frennert, S., & Östlund, B. (2014). Review: seven matters of concern of social robots and older people. *International Journal of Social Robotics*, 6(2), 299-310.
- [2] Coeckelbergh, M. (2011). You, robot: on the linguistic construction of artificial others. *AI & society*, 26(1), 61-69.
- [3] Robertson, J. (2010). Gendering humanoid robots: robo-sexism in Japan. *Body & Society*, 16(2), 1-36.
- [4] Shaw-Garlock, G. (2009). Looking forward to sociable robots. *International Journal of Social Robotics*, 1(3), 249-260.
- [5] Zhao, S. (2006). Humanoid social robots as a medium of communication. *New Media & Society*, 8(3), 401-419.
- [6] Goetz, J., Kiesler, S., & Powers, A. (2003). *Matching robot appearance and behavior to tasks to improve human-robot cooperation*. Paper presented at the Robot and Human Interactive Communication, 2003. Proceedings. ROMAN 2003. The 12th IEEE International Workshop on.
- [7] Li, D., Rau, P. P., & Li, Y. (2010). A cross-cultural study: Effect of robot appearance and task. *International Journal of Social Robotics*, 2(2), 175-186.
- [8] Syrdal, D. S., Dautenhahn, K., Woods, S. N., Walters, M. L., & Koay, K. L. (2007). *Looking Good? Appearance Preferences and Robot Personality Inferences at Zero Acquaintance*. Paper presented at the AAAI Spring Symposium: Multidisciplinary Collaboration for Socially Assistive Robotics.
- [9] Broadbent, E., Kumar, V., Li, X., Sollers 3rd, J., Stafford, R. Q., MacDonald, B. A., & Wegner, D. M. (2013). Robots with display screens: a robot with a more humanlike face display is perceived to have more mind and a better personality. *PLoS one*, 8(8), e72589.
- [10] Robins, B., Dautenhahn, K., Te Boerhorst, R., & Billard, A. (2004, September). *Robots as assistive technology—does appearance matter?* In Robot and Human Interactive Communication, 2004. ROMAN 2004. 13th IEEE International Workshop on (pp. 277-282). IEEE.
- [11] Eyssel, F., Kuchenbrandt, D., Hegel, F., & de Ruiter, L. (2012). *Activating elicited agent knowledge: How robot and user features shape the perception of social robots*. Paper presented at the RO-MAN, 2012

IEEE.

- [12] Mori, M. (1970). Bukimi no tani. *Energy*, 7(4), 33-35.
- [13] Prakash, A., & Rogers, W. A. (2013). *Younger and older adults' attitudes toward robot faces effects of task and humanoid appearance*. Paper presented at the Proceedings of the human factors and ergonomics society annual meeting.
- [14] Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, 105(32), 11087-11092.
- [15] Mathur, M. B., & Reichling, D. B. (2016). Navigating a social world with robot partners: A quantitative cartography of the Uncanny Valley. *Cognition*, 146, 22-32.
- [16] Noguchi, K., Gel, Y. R., Brunner, E., & Konietzschke, F. (2012). nparLD: an R software package for the nonparametric analysis of longitudinal data in factorial experiments.
- [17] Ebner, N. C., Gluth, S., Johnson, M. R., Raye, C. L., Mitchell, K. J., & Johnson, M. K. (2011). Medial Prefrontal Cortex Activity When Thinking About Others Depends on Their Age. *Neurocase*, 17(3), 260-269. doi:10.1080/13554794.2010.536953
- [18] Ebner, N. C., He, Y., & Johnson, M. K. (2011). Age and Emotion Affect How We Look at a Face: Visual Scan Patterns Differ for Own-Age versus Other-Age Emotional Faces. *Cognition & Emotion*, 25(6), 983-997. doi:10.1080/02699931.2010.540817
- [19] Cross, J. F., & Cross, J. (1971). Age, sex, race, and the perception of facial beauty. *Developmental Psychology*, 5(3), 433.
- [20] Palumbo, R., Adams Jr, R. B., Hess, U., Kleck, R. E., & Zebrowitz, L. (2017). Age and Gender Differences in Facial Attractiveness, but Not Emotion Resemblance, Contribute to Age and Gender Stereotypes. *Frontiers in Psychology*, 8, 1704.
- [21] Minato, T., Shimada, M., Ishiguro, H., & Itakura, S. (2004). Development of an android robot for studying human-robot interaction. *Innovations in applied artificial intelligence*, 424-434.
- [22] Broadbent, E., Stafford, R., & MacDonald, B. (2009). Acceptance of healthcare robots for the older population: review and future directions. *International journal of social robotics*, 1(4), 319-330.
- [23] Powers, K. E., Worsham, A. L., Freeman, J. B., Wheatley, T., & Heatherton, T. F. (2014). Social Connection Modulates Perceptions of Animacy. *Psychological science*, 25(10), 1943-1948.

1차 원고 접수: 2017. 06. 19

1차 심사 완료: 2017. 10. 10

2차 원고 접수: 2017. 11. 12

2차 심사 완료: 2017. 11. 24

3차 원고 접수: 2018. 02. 06

최종 게재 확정: 2018. 03. 12

(Abstract)

Age differences of preference for humanoid AI speakers

Songjoo Oh Jihyun Hwang Jiho Yew Sowon Hahn

Department of Psychology, Seoul National University

In this study, we investigated age differences of preference and trust ratings when the appearance of an artificial intelligent speaker resembles a human face. The appearance of the artificial intelligent speaker was presented in seven levels from robot face to human face. In addition, face stimuli were divided into gender (male and female) and age (20s / 60s). Participants evaluated the reliability and likability of each face stimulus on a 7-point scale. The results show that younger adults tend to prefer the face that was halfway between the robot and the human face, while older adults evaluated that the perceived reliability and likability were higher when the stimuli resembled the human face. When asked to choose the most preferred of the four face categories, all participants chose a younger face. However, with additional conditions including emoticon face and empty condition, older adults still preferred human face, while younger adults preferred emoticon face and empty condition. Taken together, older adults are more receptive to human faces than robotic faces in the context of artificial intelligence speakers. Because artificial intelligent speakers can play an important role in the elderly living alone, the present study will be a good reference in the design and development of artificial intelligent speakers for the elderly users.

Key words : artificial intelligence speaker, face perception, aging, interface design