

Comparisons of Online Assessment Platforms for Action Naming Tests in Healthy Young and Middle-aged Adults

Junyoung Shin, Jueun Kim, Jisu Song, Jae-Eun Cho, Sujin Choi, Youngmee Lee, Jee Eun Sung

Department of Communication Disorders, Ewha Womans University, Seoul, Korea

Correspondence: Jee Eun Sung, PhD
Department of Communication Disorders, Ewha Womans University, 52 Ewhayeodae-gil, Seodaemun-gu, Seoul 03760, Korea
Tel: +82-2-3277-2208
Fax: +82-2-3277-2122
E-mail: Jeesung@ewha.ac.kr

Received: March 2, 2022
Revised: May 2, 2022
Accepted: May 2, 2022

This research was partly supported by the National Research Council of Science & Technology (NST) grant by the Korea government (MSIT) (No. CAP21052-000), the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (2022R1A2C2005062) and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2022R111A4063209).

Objectives: The purpose of this study was to evaluate the validity of action naming tests when using a chatbot-based online platform. The performance of healthy young adults and healthy middle-aged adults on action naming tests was compared between online assessment platforms. **Methods:** A total of 81 participants were included in this study (40 young adults, 41 middle-aged adults). Participants were divided into two groups according to the online assessment platforms, 40 individuals (20 young adults, 20 middle-aged adults) receiving the action naming tests via Zoom and 41 individuals (20 young adults, 21 middle-aged adults) via chatbot. **Results:** The results indicated that there was no significant difference in action naming performance across online assessment platforms. In addition, young adults showed statistically higher performance compared to the middle-aged adults. There was a significant main effect of verb argument structure on action naming performance, with one- and two-place verbs having a higher correct response rate than three-place verbs. **Conclusion:** Although the younger group performed better than the middle-aged group, there were no significant differences in action naming performance between chatbot and Zoom groups, suggesting that chatbot-based action naming tests can be used interchangeably with the established online platforms. Accordingly, it is expected that we will be able to provide appropriate online language assessment services for each client in a future clinical setting.

Keywords: Aging, Action naming test, Language assessment, Online assessment platforms

고령화 문제는 우리 사회가 당면한 과제 중 하나로, 통계청에서 발표한 ‘한국의 사회동향 2020’에 따르면, 한국 인구는 세계에서 가장 빠른 속도로 고령화가 진행되고 있다. 2022년을 기준으로 한국은 2025년에 초고령 사회에 진입하여, 2045년 이전에 세계에서 노인 인구 비중이 가장 높은 국가가 되고, 2067년에는 65세 이상의 고령 인구 비중이 인구의 절반인 47%에 이르게 된다(Statistics Korea, 2019). 무서운 속도로 증가하는 노인 인구와 더불어, 한국 인구의 기대수명은 1970년 62.3세에서 2020년 83.5세로 21년가량 길어졌다(Statistics Korea, 2021). 이처럼 한국의 노인 인구가 급증하고, 노년기가 생애 주기에 비대한 비중을 차지하게 되면서, 노화가 우리 사회의 주요 담론의 대상이 되었다.

우리나라의 고령화가 빠르게 진행되면서 의사소통장애 연구 분

야에서는 노화에 따른 언어 능력의 변화에서 정상적 노화와 병리적 노화의 차이를 파악하는 연구의 필요성을 제기하고 있다(Kim, Kim, Yoo, & Kim, 2013; Kim & Sung, 2021). Kang, Kim, Seok, Cho와 Choi (2001)는 우리나라 노인들이 일상생활에서 경험하는 가장 큰 곤란이 의사소통에서의 어려움이라고 보고하였다. 실제로 노년기 의사소통 저하는 우울증을 동반한 심리적 위축, 사회활동 단절, 대인관계에서의 소외감 등을 유발하여, 궁극적으로 삶의 질을 저하시키는 심각한 요인으로 인식되고 있다(Chae, 2011). 노년층의 의사소통 능력에 관한 연구 분야에서 가장 활발히 다루고 있는 영역은 언어표현 능력으로(Kim & Kim, 2009), 노화로 인한 대표적인 변화가 바로 이름대기(naming)의 저하이다(Albert, Heller, & Milberg, 1988; Brooks, Friedman, Gibson, & Yesavage, 1993; Goulet,

Ska, & Kahn, 1994; Haug & Eggers, 1991; Kramer et al., 1999; Ramsay, Nicholas, Au, Obler, & Albert, 1999). 정상적인 노화 과정에서 이름대기 능력의 변화는 자연스러운 현상이지만(Barresi, Obler, Au, & Albert, 1999; Kim et al., 2013; Mortensen, Meyer, & Humphreys, 2006; Schwartz, 2002; Zec, Markwell, Burkett, & Larsen, 2005), 이름대기 능력의 심각한 저하는 정상 노년층으로부터 경도 인지장애(Mild Cognitive Impairment, MCI), 알츠하이머성 치매(Alzheimer's Disease, AD)와 같은 인지장애를 변별해내는 중요한 요소로 보고되기도 한다(Rosen, 1980). 따라서 노인들의 의사소통 능력의 변화가 노년기 삶의 질과 밀접한 연관이 있고, 퇴행성 신경 질환 및 인지장애와도 유의한 관련성이 있음을 고려할 때, 노화로 인한 이름대기 능력의 저하에 주목할 필요가 있다.

대면이름대기 검사(confrontation naming test)는 대상자에게 물체의 그림을 보여주고 그림에 대응하는 이름을 말하도록 함으로써 언어표현 능력을 평가하는 대표적인 이름대기 검사이다. Kim 등(2013)의 연구에서는 고령층의 언어표현 능력을 평가하는 하위 검사의 내용타당도 검증을 통해 이름대기 검사의 적절성을 확인하였고, 동사 대면이름대기 검사의 경우 노년층에 나타나는 의미론적 손상과 어휘 인출 문제를 살펴볼 수 있는 유용한 언어 평가 도구임을 시사하였다. 동사 대면이름대기 검사는 대상자가 움직임을 나타내는 그림 또는 상태 변화가 연속 그림 또는 동영상 등으로 표현된 동사 자극을 보고 해당하는 동사 어휘를 구두로 표현하는 이름대기 능력을 평가하는 검사의 유형이다. 정상 노년층은 청년층에 비해 동사 이름대기 능력이 저하되는 것으로 나타나며(Nicholas, Obler, Albert, & Goodglass, 1985), 50대 이후 점차적으로 노년층의 동사 이름대기 능력이 감퇴하는 것으로 보고되었다(Mackay, Connor, Albert, & Obler, 2002; Ramsay et al., 1999). 그러나 정상적인 노화 과정에서 연령에 따라 동사 산출 능력을 비교하거나 동사 이름대기 능력의 변화 추이를 종단적으로 살펴본 연구는 미비한 수준이다(Sung & Kwak, 2012).

이름대기 검사를 수행할 때, 이름대기의 어려움을 야기하는 요인은 다양하게 보고되고 있다(Kim & Sung, 2021; Wolf & Segal, 1992). 다수의 선행연구에서 주제로 다루고 있는 동사 이름대기 수행력의 변수는 검사 자극의 유형이다. 동사 이름대기의 검사에는 연속 그림으로 나타내는 정적 자극과(Druks, 2002; Hyun, Kim, Shin, & Seo, 2003; Zingeser & Berndt, 1990) 동영상이나 애니메이션과 같은 동적 자극이(Kim, 2006; Pashek & Tompkins, 2002; Sung, 2016) 사용되고 있다. 정적 자극의 경우, 동사의 특징인 움직임이나 상태의 변화를 보여주지 못한다는 지적도 있다(Almor et al., 2009). 국외의 신경언어장애군을 대상으로 한 선행연구들은 이

러한 지적을 뒷받침하는 결과들을 보고하기도 하고(de Almeida et al., 2021; Blankestijn-Wilmsen et al., 2017; d'Honincthun & Pillon, 2008; Schlosser et al., 2014), 반증하는 결과를 제시하기도 하였다(Berndt, Mitchum, Haendiges, & Sandson, 1997; Tranel, Manzel, Asp, & Kemmerer, 2008). Choi와 Sung(2014)의 연구에서는 그림과 애니메이션을 사용하여 정상 노년층의 자극 유형에 따른 이름대기 수행력 차이를 살펴본 결과 동적 자극인 애니메이션에서 통계적으로 더 높은 수행력이 나타났지만, 최근 정상 청년층을 대상으로 살펴본 연구결과에서는 자극 유형에 따른 수행력에서 유의한 차이가 나타나지 않았다(Choi, Jo, & Sung, 2021). 정상 성인에 대한 국내 연구결과를 종합해보면, 자극 유형의 변수는 청년층에 비해 노년층 수행력에 더 크게 작용하는 요인으로 미루어 볼 수 있다. 그러나 연령 집단에 따라 동적 자극을 활용한 동사 이름대기 수행력을 살펴본 국내 연구는 제한적이며, 정상 노화와 관련한 비교 연구도 미비한 실정이다. 따라서, 본 연구는 동적 자극인 애니메이션에서 청년층과 중노년층으로 연령을 세분화하여 동사 이름대기 수행력의 차이를 살펴보고자 한다.

동사의 구조적인 관점에서 동사 이름대기 수행력에 영향을 미치는 변수는 동사의 논항 구조(argument structure)이다. 논항은 문장 내에서 의미를 완성시키기 위해 필수적으로 요구되는 성분을 말하며, 논항의 수에 따라 1항, 2항, 3항 동사로 나누어진다. Kim과 Thompson(2000)은 미국 실어증 환자들의 동사 산출 능력이 논항별로 유의한 차이를 보이며, 논항 증가에 따라 결함이 크게 나타났다고 보고하였다. 영어권 뿐만 아니라 이탈리아, 독일, 네덜란드 등 다양한 언어권에서도 동사 논항이 증가함에 따라 동사 산출에 어려움이 나타났다는 연구결과가 보고되었다(De Bleser & Kauschke, 2003; Jonkers & Bastiaanse, 1996; Kemmerer & Tranel, 2000; Kim & Thompson, 2000; Luzzatti et al., 2002; Thompson, Lange, Schneider, & Shapiro, 1997). 국내에서도 실어증 환자와 정상 성인을 대상으로 논항 구조에 따른 동사 이름대기 수행 능력을 살펴본 것인데(Yoon & Sung, 2020), 논항별 정반응률이 유사하게 나타난 정상 집단과 달리 실어증 집단은 논항에 따른 정반응률에 차이가 있음을 알 수 있었다. 또한, 2항 동사와 3항 동사에서 유의한 차이를 보였으며, 이는 논항이 증가할수록 동사 산출이 어려웠다는 것을 의미한다. AD 집단과 정상 노인 집단을 대상으로 동사 이름대기 과제를 통한 논항 구조의 산출 능력을 살펴본 연구에서도 모든 집단이 3항 동사를 가장 어려워하였으며, AD 집단은 정상 노인 집단에 비해 3항 동사에서 통계적으로 유의미하게 저하된 수행능력을 보여 3항 동사가 정상 노인 집단과 AD 집단을 변별해주는 중요한 요소임을 밝혔다(Yi, 2014).

논항에 따른 동사 이름대기 수행력의 차이는 동사가 가진 품사적 특성과 관련이 있다. 동사는 문장의 의미를 완성하기 위해 필요한 언어적 요소인 논항을 가진 논항 구조로 이루어져 있기 때문에 명사에 비해 통사적으로 복잡한 구조를 가지며(Thompson, 2003), 이러한 복잡성에 따라 정상 노년층 및 신경언어장애 집단의 동사 산출 수행력이 달라질 수 있다(Kim & Thompson, 2000; Sung & Kwak, 2012). 특히, Sung (2016)은 동사 이름대기 수행력 산출에 영향을 미치는 요인으로 '주어(subject)+목적어(object)+서술어(verb)'의 구조를 가진 한국어의 특징을 언급하였다. 그들은 실어증 환자와 정상 노년층 집단을 대상으로 논항에 따른 동사 이름대기 수행력의 차이를 비교한 결과, 실어증 환자 집단은 1항 동사에 비해 3항 동사를 산출하는 데 어려움을 보였지만, 이외 논항 동사의 수행력에서는 유의한 차이를 보이지 않았다. 이는 주어가 자주 생략되어 2항 동사가 1항 동사로 여겨지고, 필수 논항 외의 구가 추가되어 문장이 길어지는 한국어의 특징으로 인해 논항별 수행력에 유의한 차이를 보이지 않을 수 있다고 해석하였다. 이처럼 한국어가 가지고 있는 동사의 특성으로 인해 수행력에 차이가 나타날 수 있으며, 논항 구조에 따른 동사 이름대기가 정상 노년층으로부터 장애군을 조기 선별할 수 있는 중요한 요인이 될 수 있음을 감안할 때, 동사 이름대기 검사 시 동사의 논항 구조를 고려한 수행력을 살펴보는 것이 중요하다.

최근들어 검사 자극 대면 제시 유무에 따른 언어 평가 수행력에 관한 연구가 활발하게 이루어지고 있으며(Crutchley & Campbell, 2010; Dekhtyar, Braun, Billot, Foo, & Kiran, 2020; McGill, Siegel, & Noureal, 2021), 국내에서는 Choi 등(2021)의 연구에서 정상 연령 집단을 대상으로 화상회의 기반 플랫폼을 활용하여 비대면 방식으로 동사 이름대기 검사를 실시하였고, 대면 및 비대면 방식에 따른 동사 이름대기 수행력에 유의한 차이가 없음을 보고하였다. 이러한 선행연구 결과들을 통해 기존에 대면 형태로 이루어졌던 언어 평가를 비대면 시스템으로 적용할 수 있다는 가능성을 확인할 수 있다. 그러나 또 다른 연구에서는 비대면 환경에 대한 문제점을 보고하였는데, 실험자와 피험자가 접속한 기기의 성능과 인터넷 환경에 따른 연결 불량 문제와 피험자의 즉각적인 응답과 실험자의 피드백에 어려움을 겪었다고 하였다(Moon et al., 2021). 259명의 언어재활사를 대상으로 비디오 화상회의 플랫폼을 이용한 원격 치료 및 평가에 대해 설문조사를 실시한 연구에서도(Campbell & Goldstein, 2022), 원격 서비스의 가장 큰 문제점으로 대상자 중 41.8%가 인터넷의 연결성(connectivity)을 언급하였다. 뿐만 아니라 해당 플랫폼의 음질 및 화질 문제(audiovisual technology quality)에 따른 의사소통의 불편함을 호소하면서, 전체 대상자의 81.4%가 원격 서

비스의 기술적인 문제를 개선해야 한다고 응답하였다. 이같은 원격 연결 시스템의 문제는 비대면 검사 상황에서 언어 평가의 진행을 방해하여 대면 검사와 다른 수행력을 유발하는 변수가 될 수 있다. 기존의 비대면 언어평가 플랫폼이 가지는 여러 상황적 변수들이 언어 평가 결과에 미치는 영향력을 고려할 때, 원격 시스템의 문제를 보완할 수 있는 새로운 플랫폼의 모색이 더욱 필요하다.

비대면 언어 평가 플랫폼에 대한 모색이 필요한 현 시점에서 모든 연령층의 이용률이 증가세를 보이고 있는 플랫폼은 메신저(Messenger)이다. 과학기술정보통신부에 따르면, 2020년 메신저 이용률이 전년도 대비 전체적으로 증가하였으며, 특히 70대 이용률이 87.8%로 12% 증가하여 노년층도 메신저를 더욱 적극적으로 이용하는 것으로 나타났다. 국내 노년층의 메신저 이용 현황을 기반으로, 노인 대상 의료 서비스 분야에 메신저가 활용되고 있으며, 메신저에 인공지능이 적용되어 정해진 규격이나 기계학습 과정에 따라 자동화된 대화를 진행하는 챗봇(chatbot) (Rosruen & Samanchuen, 2018) 서비스를 활용하고자 하는 연구가 국내에 활발히 이루어지고 있다(Choi & Choi, 2017; Ki, 2020; Park, 2017). 고령자를 위한 헬스케어 대화형 에이전트(Kim, So, Kang, Chung, & Kim, 2021), 챗봇을 활용한 노인우울척도 평가(Kim & Kim, 2020) 등이 대표적 예이다. 국외 연구에서는 의료기관 방문 전에 자가진단을 할 수 있는 의료용 챗봇을 제작함으로써, 챗봇을 활용한 비대면 의료 서비스의 효과성에 대해 보고하였다(Divya, Indumathi, Ishwarya, Priyasankari, & Devi, 2018). 이처럼 메신저가 장소와 연령에 제약이 적은 비대면 의사소통 창구로 인식되어, 메신저 플랫폼에 인공지능 로봇을 결합한 챗봇 서비스의 개발이 점차 다양한 분야에서 진행되고 있다는 것을 알 수 있다. 이상의 논의를 종합해 볼 때, 가속화되는 고령화 문제로 인하여 노년층 및 노년기에 진입하는 시기에 나타나는 의사소통 능력의 변화를 조기에 파악할 수 있는 언어 평가 연구가 필요하며, 코로나19 확산을 계기로 비대면 언어 재활 시스템에 대한 사회적 관심도 고조되고 있는 실정이다.

따라서, 본 연구에서는 챗봇 기반 플랫폼과 화상회의 기반 플랫폼에서 정상 청년층과 중노년층을 대상으로 동사 이름대기 검사를 실시하고, 수행력을 비교함으로써 비대면 동사 이름대기 검사의 타당성을 검증하고자 하였다. 동사 이름대기 검사의 수행력의 경우, 논항에 따른 정상 성인의 수행력 차이를 보고한 선행연구(Sung & Kwag, 2012; Thompson, 2003; Yang, 2016) 결과를 바탕으로, 전체 문항의 합계 점수와 논항별 점수를 각각 산출하여 논항별 정반응률을 살펴보고자 하였다. 이에 따른 연구 질문은 다음과 같다.

동사 이름대기 검사에서 비대면 언어 평가 플랫폼 유형(챗봇 기반 플랫폼 vs. 화상회의 기반 플랫폼)에 따른 연령 집단(청년층, 중

Table 1. ANOVA results for the level of education between young and middle-aged group by online assessment platforms

Distributed source	Sum of square	Degree of freedom	Mean square	F	p
Platform type	.248	1	.248	.102	.751
Group	.075	1	.075	.031	.862
Group x Platform type	.002	1	.002	.001	.975
Error	187.893	77	2.440		

Table 2. Descriptive information of groups

		Chatbot (N=41)	Zoom (N=40)	t	p
Age (yr)					
	Young	Mean (SD) 25.65 (4.368)	24.30 (4.281)	.987	.330
		Range 18-38	18-38		
Middle-aged					
		Mean (SD) 53.86 (2.816)	56 (4.519)	-1.832	.075
		Range 50-63	50-63		
Education (yr)					
	Young	Mean (SD) 14.50 (1.762)	14.40 (.821)	.230	.330
		Range 12-16	14-16		
	Middle-aged	Mean (SD) 14.57 (1.66)	14.45 (1.791)	.225	.826
		Range 12-16	12-16		
		t = -.134	t = -.113		
	p = .894	p = .910			

노년층) 간 동사의 논항별(1항, 2항, 3항) 정반응률은 통계적으로 유의한 차이가 있는가?

연구방법

연구대상

본 연구는 대한민국에 거주하고 있는 정상 청년층 집단 40명(남자 12명, 여자 28명)과 정상 중노년층 집단 41명(남자 18명, 여자 23명), 총 81명이다. 청년층은 만 18세에서 39세, 중노년층은 만 50세에서 65세의 연령범위 기준으로 집단을 나누었다. 중노년층의 연령 기준과 관련하여, 노화에 따른 뇌의 용적 및 기능 저하가 진행 중이고(Raz, 2005), 동사 산출 능력이 감퇴하기 시작하는 연령으로 보고된(Mackay et al., 2002) 만 50세에서 노인복지법 기준으로 노년층 진입 전 단계인 65세 사이로 절충한 중노년층을 본 연구의 대상으로 설정하였다. 본 연구의 모든 절차와 방법은 생명윤리위원회(Institutional Review Board, IRB)의 기준을 따라 사전승인을 받았으며(No. 2022-0112), 참여한 모든 대상자에게 검사자는 연구 목적 및 절차 등에 대해 설명한 후, 서면 동의를 받은 후 진행하였다.

정상 청년층과 중노년층의 공통 선정 기준은 다음과 같다. (1) 스마트폰 자판을 이용하여 문자를 입력할 수 있는 자, (2) 주관적으로 기억장애에 문제가 없다고 보고한 자, (3) 건강선별설문지(Health

Table 3. Descriptive data from neuropsychological tests in middle-aged group

	Chatbot (N=21)	Zoom (N=20)	t	p
Middle-aged group				
K-MMSE	29.52 (.602)	29.50 (.889)	.101	.920
SVLT-immediate	25.43 (4.331)	26.15 (3.660)	-.575	.569

Values are presented as mean (SD).

K-MMSE = Korean-Mini Mental State Examination (Kang, 2006); SVLT-immediate = Seoul Verbal Learning Test-immediate (Kang et al., 2012).

Screening Questionnaire; Christensen, Multhaup, Nordstrom, & Voss, 1991) 결과 인지 및 언어장애와 연관된 신경 및 정신과학적 병력이 없는 자로 보고한 사람이다. 중노년층은 추가적으로 주관적 기억감퇴 설문(Subjective Memory Complaints Questionnaire, SMCQ; Youn et al., 2009) 결과 6점 미만으로 주관적 인지장애를 호소하지 않는 자, 한국형 간이 정신상태 검사(Korean-Mini Mental State Examination, K-MMSE; Kang, 2006) 결과 연령 및 교육연수에 따른 정상범위(16%ile 이상)에 해당하는 자, 서울 언어 학습검사(Seoul Verbal Learning Test, SVLT; Kang, Jang, & Na, 2012) 결과 연령 및 교육연수에 따라 16%ile 이상의 정상범주에 해당하는 자를 선별하였다.

본 연구에 참여한 집단 및 비대면 언어 평가 플랫폼 유형 간 대상자의 연령, 교육연수 정보는 Table 1, Table 2에 제시하였으며 중노년층의 선별 검사 점수는 Table 3에 제시하였다. 비대면 언어 평가 플랫폼 유형(챗봇 기반 플랫폼, 화상회의 기반 플랫폼)에 따른 집단(중노년층, 청년층) 간, 교육연수에서 유의한 차이가 있는지 알아보기 위해 이원 혼합 분산분석(two-way mixed ANOVA)을 유의수준 .05에서 실시하였다. 그 결과, 플랫폼 유형($F_{(1,77)} = .102, p > .05$)과 집단($F_{(1,77)} = .031, p > .05$)의 주효과가 통계적으로 유의하지 않았고, 유형에 따른 집단 간 교육연수에 대한 이차상호작용이 유의하지 않았다($F_{(1,77)} = .001, p > .05$). 즉 유형에 따른 청년층의 교육연수 평균은 14.45 (SD = 1.358)로 노년층의 교육연수 평균 14.51 (SD = 1.705)과 통계적으로 유의한 차이를 보이지 않는 것으로 나타났다.

실험 자극 및 플랫폼 선정

동사 이름대기 검사

본 연구의 동사 이름대기 검사 자료는 Choi 등(2021)의 연구에서



Figure 1. Example user interface of Chatbot (left) and Zoom (right).

개발한 검사 자료를 사용하였다. 검사 문항은 동사의 논항에 따라 1항 동사 11개, 2항 동사 13개, 그리고 3항 동사 12개로 총 36개로 구성되었으며, 검사 자극은 애니메이션 파일로 제시되었다. 본 실험에 앞서 각 논항별 연습 문항 3문항을 통해 검사 방법에 대해 연습한 후, 본 문항에 대한 검사를 실시하였다. 실험에 사용한 전체 목표 동사 목록은 Appendix 1에 제시하였고, 각 문항별 검사 자극의 예시는 Appendix 2에 제시하였다.

플랫폼 선정

본 연구의 챗봇 기반 플랫폼은 메신저 이용자에게 챗봇 서비스를 제공하는 3개의 메신저 기반 플랫폼(네이버톡톡, 카카오톡, 페이스북 메신저)의 국내 사용자 현황 자료를 토대로 동사 이름대기 검사의 적용 가능성을 검토함으로써 최종적으로 선정하였다. 앱/리테일 분석 서비스 Wiseapp (2021)은 한국인 만 10세 이상 스마트폰 사용자들을 대상으로 가장 자주 사용하는 어플리케이션(application)에 대한 표본 조사를 실시하였다. 그 결과, 카카오톡과 네이버가 1, 2위를 차지하였고, 페이스북은 순위에 포함되지 않았다. 2021년 상반기 카카오톡 이용자 수는 4,566만명, 네이버 이용자 수는 4,106만명으로 나타났으나(Mobile index, 2021), 페이스북 메신저는 국내 50대 이상의 사용자 비율이 현저히 낮아서(Consumer insights, 2019). 본 연구의 중노년층 대상자에게는 적합하지 않은 플랫폼으로 판단하였다. 마지막으로 동사 이름대기 검사의 애니메이션 자극이 Graphics Interchange Format (GIF) 형식의 그래픽 파일로 제작된 가운데, 카카오톡 챗봇은 GIF 형식을 지원하지 않은 반면 네이버톡톡은 모든 형식의 이미지 파일을 지원하고 있어서 검사 자극이 적용 가능한 것으로 확인되었다. 최종적으로 챗봇 서비스를 지원하는 메신저 플랫폼의 국내 사용자 현황 및 검사 자료의 적용 가능성을 기준으로 '네이버톡톡'을 선정하였다. 화상회의 기반 플랫폼은 선행연구(Choi et al., 2021)와 같은 프로그램인 줌(Zoom)으로 선정하였다.

Table 4. Comparison of inspection methods across online platforms

Platform type	Video-conferencing program	Chatbot-based messenger
Platform	Zoom	NaverTalkTalk
Device	Laptop	Smartphone
Presentation type	Examiner provides animation stimulus to participant via screen share.	Chatbot provides animation stimulus to participant when the participant presses question buttons.
Examiner type	Person	Chatbot
Response type	Verbal response	Text-based response
Transcription type	Handwritten transcription	Automatic transcription

실험절차

연구 대상자들은 비대면 언어 평가 플랫폼에 따라 두 실험군으로 나뉘었다. 챗봇 기반 플랫폼에 참여한 실험군은 스마트폰으로 접속한 네이버톡톡(NaverTalkTalk)에서, 화상회의 기반 플랫폼에 참여한 실험군은 노트북으로 접속한 줌(Zoom)에서 검사를 실시하였다(Figure 1).

챗봇 기반 플랫폼(NaverTalkTalk)

검사자는 대상자에게 동사 이름대기 챗봇 접속용 URL 링크를 전송하였다. 모든 대상자는 스마트폰에서 해당 링크로 접속하여 검사에 참여하였다. 네이버톡톡에 접속하였을 때, 챗봇은 '지금부터 움직임에 대한 애니메이션을 보여드리겠습니다. 그림을 잘 보시고 그림 속에 있는 사람 또는 사물이 하는 행동 혹은 움직임을 채팅창에 입력해주시면 됩니다.'라는 자동 안내 메시지를 전송하여 대상자가 기본 안내에 따라 스스로 검사를 진행하도록 하였다. 대상자는 문항 버튼을 눌러서 챗봇이 제시한 애니메이션 자극을 보고 채팅으로 응답하였다. 비대면 플랫폼 간 검사 방법에 대한 정보를 Table 4 제시하였다.

화상회의 기반 플랫폼(Zoom)

검사자와 대상자는 노트북 혹은 태블릿 PC를 사용하여 화상 기반 플랫폼 줌에 접속 후 검사에 임하였다. 검사자는 구두로 '지금부터 움직임에 대한 애니메이션을 보여드리겠습니다. 그림을 잘 보시고 그림 속에 나타내는 행동이나 움직임 및 동작을 말로 표현해 주시면 됩니다.'라고 안내한 후 검사를 시작하였다. 줌의 화면 공유 기능을 이용하여 검사자가 검사 자극을 대상자에게 보여주고 대상자는 구두로 응답하였다.

자료분석

실험이 종료된 이후 언어병리학전공 석사과정생 4명과 언어병리

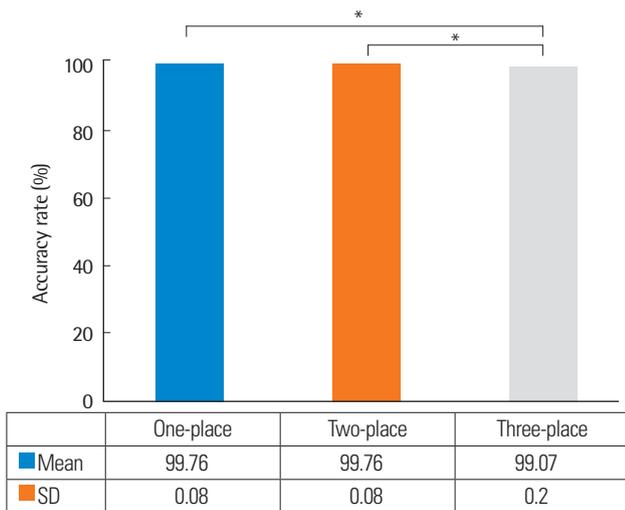


Figure 2. Result of post-hoc analysis on the main effects of the accuracy rate by the arguments.
**p* < .05.

학전공 박사 수료생 1명이 녹화 자료를 통해 대상자의 반응을 전사하였고, 아래의 검사 채점 기준에 따라 수행력을 평가하였다. 동사 이름대기의 검사에 대한 수행력은 정반응 문항 1점, 오반응 문항 0점으로 채점하였다. 목표 동사를 산출할 경우 정반응으로 처리하였고, 그 외의 (1) 목표 동사와 의미적으로 관련된 어휘(예: ‘주다’ → ‘건네다’), (2) 목표 동사의 어간을 취하는 복합동사 산출(예: ‘올다’ → ‘올고 있네’), (3) 지역적 방언(예: ‘떼다’ → ‘따다’)을 산출한 경우도 정반응으로 간주하였다. 그 외 (1) 의미적으로 관련되지 않은 어휘(예: ‘열다’ → ‘밀다’), (2) 상위어 사용(예: ‘꼬집다’ → ‘만지다’), (3) 명사 + ‘-히’ + ‘다’ 사용(예: ‘버리다’ → ‘빼기하다’)과 같은 반응은 오반응으로 간주하였다.

논항별로 1항 11개, 2항 동사 13개, 3항 동사 12개로 논항에 따른 정반응한 문항의 수를 각 논항별 문항 수로 나누어 백배수한 값으로 정반응률(%)을 산출하였다.

$$\text{정반응률(}\%) = \frac{(\text{논항별 정반응 문항 수})}{(\text{각 논항별 전체 문항 수})} \times 100$$

자료의 통계적 처리

동사 이름대기 검사에서 비대면 언어 평가 플랫폼 유형에 따른 집단 간 논항별 정반응률의 유의성을 검증하기 위하여 PASW (PASW statistics 25.0, SPSS Inc.)를 사용하여, 삼원혼합 분산분석 (three-way mixed ANOVA)을 유의 수준 .05에서 실시하였다. 이때, 비대면 언어 평가 플랫폼 유형(챗봇 기반 플랫폼 vs. 화상회의 기반 플랫폼)과 집단(청년층 vs. 중노년층)을 개체 간 요인으로, 논항 (1항, 2항, 3항)을 개체 내 요인으로 설정한 후, 정반응률을 종속변

Table 5. Comparison of accuracy rate (%) between young and middle-aged group by online assessment platforms

	Chatbot (N=41)		Zoom (N=40)	
	Young	Middle-aged	Young	Middle-aged
One-place	99.861 (.621)	99.603 (.996)	100 (.00)	99.853 (1.017)
Two-place	99.861 (.621)	99.338 (1.212)	99.861 (.621)	100 (.00)
Three-place	99.853 (1.017)	98.677 (2.259)	99.583 (1.017)	98.472 (2.623)

Values are presented as mean (SD).

Table 6. ANOVA results from the accuracy rate on action naming tests between young and middle-aged group by online assessment platforms

Distributed source	Sum of square	Degree of freedom	Mean square	F	p	η^2_p
Between factor						
Group	15.953	1	15.953	10.008	.002**	.115
Platform type	.558	1	.558	.350	.556	.005
Group×Platform type	.149	1	.149	.094	.760	.001
Error	122.737	77	1.594			
Within factor						
Argument	25.291	1.484	17.037	8.099	.002**	.095
Ar×Group	7.684	1.484	5.177	2.461	.105	.031
Ar×Platform type	1.939	1.484	1.306	.621	.493	.008
Ar×Group×Platform type	2.404	1.484	1.620	.770	.430	.010
Error	240.440	114.303	2.104			

Ar=Argument.
***p* < .01.

수로 설정하였다. 반복측정 분산분석에서는 반복 측정되는 변량의 동질성 가정을 확인하기 Mauchly의 구형성(Sphericity) 검정을 실시하고, 가정이 충족될 경우 구형성 가정값을 사용하고, 그렇지 않은 경우에는 Greenhouse-Geisser의 수정값을 사용한다. 본 연구에서는 개체 내 요인인 논항에 대한 구형성 가정이 충족되지 않아, Greenhouse-Geisser로 수정된 자유도와 F값을 보고하였다.

연구결과

비대면 언어 평가 플랫폼 유형 및 논항에 따른 집단 간 동사 이름대기 정반응률에 차이가 통계적으로 유의한지 살펴보기 위하여 삼원혼합 분산분석(three-way mixed ANOVA)을 실시하였다. 이에 대한 기술통계 결과 및 논항별 정반응률 그래프는 Table 5, Figure 2와 같고, 삼원혼합 분산분석 결과는 Table 6과 같다.

분석 결과, 집단에 따른 주효과 통계적으로 유의하였다($F_{(1,77)} = 10.008, p < .01$). 즉, 청년층 집단의 평균 정반응률이 99.792% (SD = .115)로 중노년층 집단 평균 99.279% (SD = .114)에 비해 유의하게

높은 것으로 나타났다.

뿐만 아니라 논항에 따른 주효과가 통계적으로 유의하였다($F_{(1,484, 114,303)} = 8.099, p < .01$). 이에 따라 LSD 사후검정을 통하여 주효과의 비교를 실시한 결과, 1항 평균 정반응률은 99.762% ($SD = .087$), 3항 평균 정반응률은 99.079% ($SD = .209$)로 1항이 3항보다 정반응률이 유의하게 높은 것으로 나타났다. 또한 2항 평균 정반응률은 99.765% ($SD = 0.84$)로 3항 평균 정반응률과 비교하였을 때, 유의하게 높게 나타났다. 그러나 1항과 2항의 평균 정반응률에서는 통계적으로 유의한 차이가 없었다.

그 밖에 비대면 언어 평가 플랫폼 유형($F_{(1,77)} = .350, p > .05$)에 따른 주효과는 통계적으로 유의하지 않았다. 또한, 이차 및 삼차 상호작용이 유의하지 않았다. 즉, 비대면 언어 평가 플랫폼 유형에 따른 집단 간 동사 논항별 정반응률에는 차이가 없었다.

논의 및 결론

본 연구에서는 고령화 문제와 코로나(COVID-19) 확산으로 인한 비대면 언어 평가 시스템의 수요가 증가함에 따라, 챗봇 플랫폼을 활용한 비대면 동사 이름대기 검사의 유효성에 대해 살펴보고자 하였다. 이에 따라 동사 이름대기 검사와 비대면 플랫폼을 선정하여, 비대면 언어평가 플랫폼 유형(챗봇 기반 플랫폼 vs. 화상회의 기반 플랫폼)에 따른 연령 집단(청년층 vs. 중노년층) 간 논항별(1항 vs. 2항 vs. 3항) 수행력을 비교하였다.

본 연구의 결과를 하나씩 살펴보면, 비대면 언어 평가 플랫폼 유형에 따른 동사 이름대기의 정반응률은 통계적으로 유의하지 않았다. 이러한 결과는 자극 제시 방법에 따라 언어 평가 수행력에 유의한 차이를 보이지 않는다는 선행연구의 결과와 일치한다(Choi et al., 2021; Dekhtyar et al., 2020; Theodoros et al., 2008). 즉, 동사 이름대기 검사를 챗봇 기반 플랫폼과 화상회의 기반 플랫폼에서 상호교환적으로 사용할 수 있다는 가능성을 시사하는 것과 동시에, 두 가지 비대면 플랫폼에서 모두 동사 이름대기 수행력을 신뢰도 있게 측정할 수 있다는 결과로도 해석할 수 있다. 그러나 본 연구는 정상 청년층과 중노년층만을 대상으로 확인한 결과이기 때문에, 추후 연령 범위를 확대하여 세부적으로 살펴볼 필요가 있다. 또한, 비대면 플랫폼 간 대상자 반응 산출 방식에 있어서 챗봇 기반 플랫폼은 쓰기로, 화상회의 기반 플랫폼은 구어로 산출하는 차이가 있다. 따라서, 비대면 동사 이름대기 검사를 상호교환적으로 사용하기 위해서는 쓰기와 구어 산출 방법에 따른 기제에 대해 고려한 추가 연구가 진행되어야 하겠다. 하지만 연구 성과 면에서, 본 연구가 동사 이름대기를 통해 국내에서 시도하지 않았던 두 가지 비대면

플랫폼을 비교했다는 데에 큰 의의가 있다.

동사의 논항에 따른 정반응률의 차이는 통계적으로 유의하였는데, 논항별 정반응률은 2항, 1항, 3항 순으로 높게 나타났고, 특히 1항과 3항, 2항과 3항이 유의한 차이를 보인 반면에, 1항과 2항 동사 간의 정반응률에는 유의한 차이가 없었다. 이러한 결과는 논항에 따른 동사 이름대기 수행력에 차이를 보인다는 선행연구들과 일치한다(De Bleser & Kauschke, 2003; Jonkers & Bastiaanse, 1996; Kemmerer & Tranel, 2000; Kim & Thompson, 2000; Luzzatti et al., 2002; Thompson et al., 1997). 정상 청년층을 대상으로 zoom을 이용하여 비대면 동사 이름대기 검사를 실시한 최근 연구(Choi et al., 2021)의 결과, 2항, 1항, 3항 순으로 높은 수행력을 보였다. Yoon과 Sung (2020) 연구에서 실어증 환자와 정상 성인이 공통적으로 2항, 1항, 3항 순서로 높은 정반응률을 보인 결과와도 일치하였다. 또 다른 연구에서는 1항 동사와 2항 동사에서 유의한 차이가 나타나지 않아, 본 연구결과와 동일하였다(Kim, 2006; Sung, 2016; Yang, 2016). 본 연구에서 정상 집단을 대상으로 실시하였기 때문에, 논항별 수행력 차이를 일반화하기에는 한계가 존재할 수 있겠지만, 동사 구조적 측면에서 3항에서의 가장 낮은 정반응률이 나타난 본 연구의 결과는 논항이 증가할수록 동사 산출이 어렵다는 다수의 선행연구의 결과를 추가 확인한 유의미한 결과라고 할 수 있겠다.

다음으로 연령 집단에 따른 정반응률의 차이가 통계적으로 유의하게 나타났다. 동사 이름대기 검사에서 청년층이 중노년층에 비해 더 높은 수행력을 보였는데, 이는 연령이 증가함에 따라 이름대기 능력은 저하되고, 특히 청년층에 비해 노년층의 이름대기 능력이 더 저조하다는 다수의 선행연구의 결과와 일치하는 결과이다(Barresi et al., 1999; Mortensen et al., 2006; Ramsay et al., 1999; Sung & Kim, 2011). 그러나 노화에 따른 이름대기 저하와 관련하여 본 연구의 결과를 해석할 때, 두 정상 집단의 동사 이름대기 평균 정반응률이 모두 99% 이상의 높은 수치를 나타내고, 평균 점수로 부터 자료가 얼마나 분포되어 있는지를 알 수 있는 표준편차가 작은 수치임을 감안해야 할 것으로 판단된다.

본 연구의 결과를 종합해보면, 비대면 언어 평가 플랫폼 유형에 따른 집단 간 동사 이름대기 검사 수행력에 유의한 차이가 나타나지 않았다. 즉, 동사 이름대기 검사는 줌(Zoom)을 활용하여 검사자가 직접 자극을 제시하거나 검사자가 아닌 챗봇이 자극을 제시하는 방법 중 어떠한 방법으로 제시해도 검사 수행력에는 차이가 없었다. 그러므로 동사 이름대기 검사는 대상자의 상황이나 환경에 따라 가장 적절한 비대면 언어 평가 플랫폼을 선정하여 검사를 수행할 수 있기 때문에 임상적 활용 가치가 우수하다고 사료된다.

본 연구에서는 대면 유무에 따라 수행력을 비교한 다수의 선행

연구와 달리, 화상회의 기반 비대면 플랫폼과 ‘챗봇’이라는 새로운 비대면 플랫폼 간의 비교를 시도하였다. 특히, 비대면 플랫폼 선정 과정에서 연령대별로 플랫폼 사용 현황을 고려하였고, 서로 다른 인터페이스의 플랫폼 간 타당성 비교를 처음 시도하였다는 점에서 큰 의의가 있다. 또한, 본 연구결과에서 챗봇 기반 동사 이름대기 검사의 타당성을 확보하고, 추후 언어 평가 서비스의 접근성에 제약이 있는 대상자에게 적절한 비대면 서비스를 제공할 수 있는 온라인 언어 평가의 유효성을 확인하여 임상적 활용도가 높을 것으로 기대된다.

본 연구에서의 제한점 및 제언은 다음과 같다. 먼저 정상 청년층과 중노년층만을 대상으로 진행하였고 집단별 대상자 표본 수가 약 20명이기 때문에, 본 연구의 결과를 모든 연령대의 대상자들에게 적용하고, 노화에 따른 동사 이름대기 능력의 저하로 일반화하기에는 어려움이 있을 수 있다. 따라서, 연령 범위와 대상자 수를 확대하여 비대면 플랫폼 유형에 따른 동사 논항별 수행력 비교 연구가 이루어져야 할 것이다. 또한 본 연구에서는 검사 시 대상자의 반응 시간에 대해 고려하지 않았다. 추후 연구에서 이러한 요인을 고려한다면 연령 범위에 따른 이름대기 능력에 대한 더욱 정밀한 측정이 가능할 것으로 기대한다. 한편, 챗봇 기반 동사 이름대기 검사는 대상자가 검사 문항별로 버튼을 눌러서 애니메이션 자극을 직접 불러오는 형태로 진행되었다. 향후에는 챗봇이 대상자의 응답에 반응하여 자동으로 다음 검사 문항을 제시할 수 있는 인공지능 기술이 챗봇 시스템에 적용된다면 스마트폰 인터페이스가 익숙하지 않은 고령자도 챗봇의 자동화된 진행에 따라 원활하게 검사를 수행할 수 있을 것으로 기대한다. 다음으로 챗봇 기반 동사 이름대기 검사 시 대상자의 응답에 오타는 오반응으로 처리하지 않았다. 실제 이름대기 능력 저하를 겪는 노년층 혹은 환자군의 음소착어 발생 빈도를 고려할 때, 챗봇 기반 플랫폼을 통한 검사를 진행할 시 오타와 같은 수행 오류와 음소착어를 구분하여 채점할 수 있는 지침이 마련되어야 할 필요가 있다. 마지막으로 이러한 세부 오류 유형을 챗봇 코딩 시 적용한다면 자동화된 채점을 통해 효율적이고 보다 질적인 부분을 살펴볼 수 있을 것으로 사료된다.

REFERENCES

- Albert, M. S., Heller, H. S., & Milberg, W. (1988). Changes in naming ability with age. *Psychology & Aging*, 3(2), 173-178.
- Almor, A., Aronoff, J. M., MacDonald, M. C., Gonnerman, L. M., Kempler, D., Hintiryan, H., Hayes, U. L., Arunachalam, S., & Andersen, E. S. (2009). A common mechanism in verb and noun naming deficits in Alzheimer's patients. *Brain & Language*, 111(1), 8-19.
- Barresi, B. A., Obler, L. K., Au, R., & Albert, M. L. (1999). Language related factors influencing naming in adulthood. In H. E. Hamilton (Ed.), *Language and communication in old age: multidisciplinary perspective* (pp. 76-89). NY: Taylor & Francis, Inc.
- Berndt, R. S., Mitchum, C. C., Haendiges, A. N., & Sandson, J. (1997). Verb retrieval in aphasia. 1. characterizing single word impairments. *Brain & Language*, 56(1), 68-106.
- Blankstijn-Wilmsen, J., Damen, I., Voorbraak-Timmerman, V., Hurkmans, J., Brouwer de Koning, J., Pross, A., & Jonkers, R. (2017). The effect of static versus dynamic depictions of actions in verb and sentence production in aphasia. *Aphasiology*, 31(10), 1166-1182.
- Brooks III, J. O., Friedman, L., Gibson, J. M., & Yesavage, J. A. (1993). Spontaneous mnemonic strategies used by older and younger adults to remember proper names. *Memory*, 1(4), 393-407.
- Campbell, D. R., & Goldstein, H. (2022). Evolution of telehealth technology, evaluations, and therapy: effects of the COVID-19 pandemic on pediatric speech-language pathology services. *American Journal of Speech-Language Pathology*, 31(1), 271-286.
- Chae, S. (2011). A review of age-related hearing loss. *Journal of the Korean Medical Association*, 54(9), 908-909.
- Choi, S., & Sung, J. E. (2014). Task-specific and argument structure effects on verb production in normal elderly adults: animation vs. picture comparisons. *Journal of Rehabilitation Research*, 18(4), 279-293.
- Choi, S., & Choi, Y. (2017). Analysis on the conversational commerce service interface of the AI chat-bot based on mobile messenger apps. *Proceedings of the 2017 HCI Korea conference*, 237-240.
- Choi, S., Jo, E., & Sung, J. E. (2021). Preliminary study on the action naming test: online vs. offline comparisons by presentation type. *Journal of Speech-Language & Hearing Disorders*, 30(2), 87-97.
- Christensen, K. J., Multhaup, K. S., Nordstrom, S., & Voss, K. (1991). A cognitive battery for dementia: development and measurement characteristics. *Psychological Assessment: A Journal of Consulting & Clinical Psychology*, 3(2), 168-174.
- Consumer Insights. (2019). *Mobile communication planning survey in the first half of 2019*. Retrieved from https://www.consumerinsight.co.kr/voc_view.aspx?no=2962&id=ins02_list&PageNo=1&schFlag=1.
- Crutchley, S., & Campbell, M. (2010). Telespeech therapy pilot project: stakeholder satisfaction. *International Journal of Telerehabilitation*, 2(1), 23.
- de Almeida, R. G., Mobayyen, F., Antal, C., Kehayia, E., Nair, V. P., &

- Schwartz, G. (2021). Category-specific verb-semantic deficits in Alzheimer's disease: evidence from static and dynamic action naming. *Cognitive Neuropsychology*, 38(1), 1-26.
- De Bleser, R., & Kauschke, C. (2003). Acquisition and loss of nouns and verbs: parallel or divergent patterns?. *Journal of Neurolinguistics*, 16(2-3), 213-229.
- Dekhtyar, M., Braun, E. J., Billot, A., Foo, L., & Kiran, S. (2020). Videoconferencing administration of the Western Aphasia Battery-Revised: feasibility and validity. *American Journal of Speech-Language Pathology*, 29(2), 673-687.
- d'Honinckun, P., & Pillon, A. (2008). Verb comprehension and naming in frontotemporal degeneration: the role of the static depiction of actions. *Cortex*, 44(7), 834-847.
- Divya, S., Indumathi, V., Ishwarya, S., Priyasankari, M., & Devi, S. K. (2018). A self-diagnosis medical chatbot using artificial intelligence. *Journal of Web Development & Web Designing*, 3(1), 1-7.
- Druks, J. (2002). Verbs and nouns: a review of the literature. *Journal of Neurolinguistics*, 15(3-5), 289-315.
- Goulet, P., Ska, B., & Kahn, H. J. (1994). Is there a decline in picture naming with advancing age?. *Journal of Speech and Hearing Research*, 37, 629-644.
- Haug, H., & Eggers, R. (1991). Morphometry of the human cortex cerebri and corpus stratum during aging. *Neurobiology of Aging*, 12(4), 336-338.
- Hyun, J. M., Kim, H. H., Shin, J. C., & Seo, S. G. (2003). Retrieval of nouns and verbs in Broca's and Wernicke's aphasia. *Korean Journal of Communication & Disorders*, 8(3), 171-187.
- Jonkers, R., & Bastiaanse, Y. R. M. (1996). The influence of instrumentality and transitivity on action naming in Broca's and anomia aphasia. *Brain and Language*, 55, 37-39.
- Kang, S. K., Kim, D. Y., Seok, D. I., Cho, H. J., & Choi, K. H. (2001). Studies on communication disorders in the elderly to improve their quality of life. *Journal of the Research Institute for Special Education & Rehabilitation Science*, 40, 109-134.
- Kang, Y. (2006). A normative study of the Korean Mini-Mental State Examination (K-MMSE) in the elderly. *Korean Journal of Psychology: General*, 25(2), 1-12.
- Kang, Y. U., Jang, S. M., & Na, D. L. (2012). *Seoul neuropsychological screening battery (SNSB-II)*. Seoul: Human Brain Research & Consulting Co.
- Kemmerer, D., & Tranel, D. (2000). Verb retrieval in brain-damaged subjects: 1. Analysis of stimulus, lexical, and conceptual factors. *Brain & Language*, 73(3), 347-392.
- Ki, D. H. (2020). Medistaff: medical network service. *Proceedings of the Korea IT Service Society Conference*, 354-356.
- Kim, B., So, M., Kang, M., Chung, H., & Kim, J. (2021). A study on conversational agents in healthcare for the older adults: focusing on onboarding scenarios. *Proceedings of HCI Korea Conference*, 653-659.
- Kim, J., & Kim, I. (2020). Usability assessment of FHIR-based Geriatric Depression Scale questionnaire using chatbot. *Journal of KIISE*, 47(7), 650-654.
- Kim, J. W., & Kim, H. H. (2009). The characteristics of Korean morpheme production in Alzheimer's patients. *Journal of Speech-language & Hearing Disorders*, 18(4), 75-90.
- Kim, K. Y. (2006). *Verb production and argument structures in aphasics* (Master's thesis). Yonsei University, Seoul, Korea.
- Kim, M., & Thompson, C. K. (2000). Patterns of comprehension and production of nouns and verbs in agrammatism: Implication for lexical organization. *Brain & Language*, 74(1), 1-25.
- Kim, S. H., & Sung, J. E. (2021). Age-related differences in confrontation naming ability according to visual stimulus valence. *Communication Sciences & Disorders*, 26(3), 616-629.
- Kim, H. H., Kim, S. R., Yoo, H. J., & Kim, J. W. (2013). Content validity of speech-language assessment items for the elderly. *Journal of Rehabilitation Research*, 17(3), 347-373.
- Kramer, A. F., Hahn, S., Cohen, N. J., Banich, M. T., McAuley, E., Harrison, C. R., Chason, J., Vakil, E., Bardell, L., Boileau, R. A., & Colcombe, A. (1999). Ageing fitness and neurocognitive function. *Nature*, 400(6743), 418-419.
- Luzzatti, C., Raggi, R., Zonca, G., Pistarini, C., Contardi, A., & Pinna, G. D. (2002). Verb-noun double dissociation in aphasic lexical impairments: the role of word frequency and imageability. *Brain & Language*, 81(1-3), 432-444.
- MacKay, A. J., Connor, L. T., Albert, M. L., & Obler, L. K. (2002). Noun and verb retrieval in healthy aging. *Journal of the International Neuropsychological Society*, 8(6), 764-770.
- McGill, M., Siegel, J., & Noureal, N. (2021). A preliminary comparison of in-person and telepractice evaluations of stuttering. *American Journal of Speech-Language Pathology*, 30(4), 1737-1749.
- Mobile index. (2021). *First half mobile app landscape analysis 2021*. Retrieved from <https://www.mobileindex.com/mi-chart/top-100/top-apps>.
- Moon, S. J. E., Dabbs, A. D., Hergenroeder, A. L., Vendetti, M. L., Jones, K. B., Willey, B. M., Morrell, M. R., & Imes, C. C. (2021). Considerations for assessing physical function and physical activity in clinical trials during the

- COVID-19 pandemic. *Contemporary Clinical Trials*, 105, 106407.
- Mortensen, L., Meyer, A. S., & Humphreys, G. W. (2006). Age-related effects on speech production: a review. *Language & Cognitive Processes*, 21(13), 238-290.
- Nicholas, M., Obler, L., Albert, M., & Goodglass, H. (1985). Lexical retrieval in healthy aging. *Cortex*, 21(4), 595-606.
- Park, D. (2017). A study on conversational public administration service of the chatbot based on artificial intelligence. *Journal of Korea Multimedia Society*, 20(8), 1347-1356.
- Pashek, G. V., & Tompkins, C. A. (2002). Context and word class influences on lexical retrieval in aphasia. *Aphasiology*, 16(3), 261-286.
- Ramsay, C. B., Nicholas, M., Au, R., Obler, L. K., & Albert, M. L. (1999). Verb naming in normal aging. *Applied Neuropsychology*, 6(2), 57-67.
- Raz, N. (2005). The aging brain observe *in vivo*: differential changes and their modifiers. In R. Cabeza, L. Nyberg, & D. Park (Eds.), *Cognitive neuroscience of aging* (pp.19-57). New York: Oxford University Press.
- Rosen, W. G. (1980). Verbal fluency in aging and dementia. *Journal of Clinical & Experimental Neuropsychology*, 2(2), 135-146.
- Rosruen, N., & Samanchuen, T. (2018). Chatbot utilization for medical consultant system. *Proceeding of the 2018 3rd Technology Innovation Management and Engineering Science International Conference (TIMES-iCON)*, 1-5.
- Schlosser, R. W., Koul, R., Shane, H., Sorce, J., Brock, K., Harmon, A., Moerlein, D., & Hearn, E. (2014). Effects of animation on naming and identification across two graphic symbol sets representing verbs and prepositions. *Journal of Speech, Language, & Hearing Research*, 57(5), 1779-1791.
- Schwartz, B. L. (2002). The phenomenology of naturally-occurring tip-of-the-tongue states: a diary study. In S. Shohov (Ed.), *Advances in psychology research* (pp. 71-84). Nova Science Publishers: New York.
- Statistics Korea. (2019). *Elderly person statistics 2019*. Retrieved from <http://www.kostat.go.kr>.
- Statistics Korea. (2021). *Lifetable 2020*. Retrieved from <http://www.kostat.go.kr>.
- Sung, J. E. (2016). The effects of verb argument complexity on verb production in persons with aphasia: evidence from a subject-object-verb language. *Journal of Psycholinguistic Research*, 45(2), 287-305.
- Sung, J. E., & Kim, J. K. (2011). Differential performance on generative naming and typicality rating between young and older adults: a preliminary study to develop a semantic complexity-based naming treatment for persons with neurogenic communication disorders. *Journal of Speech & Hearing Disorders*, 20(1), 31-49.
- Sung, J. E., & Kwak, E. J. (2012). Age-related verb naming abilities depending on the argument structures. *Korean Journal of Communication & Disorders*, 17(4), 550-564.
- Theodoros, D., Hill, A., Russell, T., Ward, E., & Wootton, R. (2008). Assessing acquired language disorders in adults via the Internet. *Telemedicine & e-Health*, 14(6), 552-559.
- Thompson, C. K. (2003). Unaccusative verb production in agrammatic aphasia: the argument structure complexity hypothesis. *Journal of Neurolinguistics*, 16(2), 151-167.
- Thompson, C. K., Lange, K. L., Schneider, S. L., & Shapiro, L. P. (1997). Agrammatic and non-brain-damaged subjects' verb and verb argument structure production. *Aphasiology*, 11(4-5), 473-490.
- Tranel, D., Manzel, K., Asp, E., & Kemmerer, D. (2008). Naming dynamic and static actions: neuropsychological evidence. *Journal of Physiology-Paris*, 102(1-3), 80-94.
- Wiseapp Retail Goods. (2021). *Most frequently used apps by Koreans*. Retrieved from https://mobile.newsis.com/viewhtml?ar_id=NISX20211019_0001618674.
- Wolf, M., & Segal, D. (1992). Word finding and reading in the developmental dyslexias. *Topics in Language Disorders*, 13(1), 51-65.
- Yang, Y. M. (2016). *Study on the relationship between verb processing and argument structure in patients with Broca's aphasia* (Master's thesis). Daegu University, Gyeongbuk, Korea.
- Yi, Y. (2014). *A study on production ability of verb argument structure in patients with Alzheimer's disease: focused on verb naming task* (Master's thesis). Chungnam National University, Daejeon, Korea.
- Yoon, H., & Sung, J. E. (2020). Effects of verb argument structure and the types of presentation modality on verb production in individuals with aphasia using a verb-final language. *Communication Sciences & Disorders*, 25(2), 399-410.
- Youn, J. C., Kim, K. W., Lee, D. Y., Jhoo, J. H., Lee, S. B., Park, J. H., . . . Woo, J. I. (2009). Development of the subjective memory complaints questionnaire. *Dementia & Geriatric Cognitive Disorders*, 27(4), 310-317.
- Zec, R. F., Markwell, S. J., Burkett, N. R., & Larsen, D. L. (2005). A longitudinal study of confrontation naming in the "normal" elderly. *Journal of the International Neuropsychological Society*, 11(6), 716-726.
- Zingeser, L. B., & Berndt, R. S. (1990). Retrieval of nouns and verbs in agrammatism and anomia. *Brain & Language*, 39(1), 14-32.

국문초록

비대면 언어 평가 플랫폼 유형에 따른 청년층 및 중노년층 간 동사 이름대기 수행력 비교

신준영 · 김주은 · 송지수 · 조재은 · 최수진 · 이영미 · 성지은

이화여자대학교 대학원 언어병리학과

배경 및 목적: 본 연구에서는 챗봇 기반 비대면 동사 이름대기 검사의 타당성을 확인하기 위해 비대면 언어 평가 플랫폼 유형(챗봇 기반 플랫폼 VS 화상회의 기반 플랫폼)에 따른 정상 청년층과 정상 중노년층의 동사 논항별 수행력을 비교하였다. **방법:** 본 연구는 만 18-39세 청년층 40명과 만 50-65세 중노년층 41명 총 81명을 대상으로 하였다. 동사 이름대기 검사는 청년층 20명, 중노년층 20명은 화상회의 기반 플랫폼에서 실시하였고, 청년층 20명, 중노년층 21명은 챗봇 기반 플랫폼에서 실시하였다. **결과:** 비대면 언어 평가 플랫폼 유형에 따른 동사 이름대기의 정반응률은 유의하지 않았고, 청년층의 수행력이 중노년층에 비해 통계적으로 유의하게 높았다. 또한, 동사 논항에 따른 정반응률의 차이가 유의하였으며, 1항과 3항, 2항과 3항 동사 간의 차이가 유의하게 나타났다. **논의 및 결론:** 청년층이 중노년층에 비해 동사 이름대기 수행력이 더 높았지만, 챗봇과 줌에 따른 수행력 차이는 나타나지 않았다. 이러한 결과로, 본 연구에서 사용한 비대면 동사 이름대기 검사는 대상자의 상황에 따라 기존의 비대면 방법과 상호교환적으로 사용될 수 있기 때문에 임상에서의 활용도가 높을 것으로 기대한다.

핵심어: 노화, 동사 이름대기, 비대면 언어 평가 플랫폼

본 연구는 2022년도 정부(과학기술정보통신부)의 재원으로 국가과학기술연구회 창의형 융합연구사업(No. CAP21052-000)의 지원 및 정부(과학기술정보통신부)의 재원으로 한국연구재단(No. 2022R1A2C2005062)의 지원, 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업(No. NRF-2022R111A4063209).

참고문헌

- 강수균, 김동연, 석동일, 조홍중, 최경희 (2001). 노인성 질환자의 삶의 질 제고를 위한 의사소통 문제 연구, 특수교육재활과학연구, 40(2), 109-134.
- 강연옥 (2006). K-MMSE (Korean-Mini Mental State Examination)의 노인규준연구. 한국심리학회지: 일반, 25(2), 1-12.
- 강연옥, 장승민, 나덕렬 (2012). 서울신경심리검사 2판. 서울: 휴브알앤씨.
- 기동훈 (2020). 환자정보 보호를 위한 의사들 전용 보안메신저 플랫폼. 한국IT서비스학회 학술대회 논문집, 354-356.
- 김가영 (2006). 실어증 환자의 동사 산출과 논항구조 간의 관계. 연세대학교 대학원 석사학위논문.
- 김보희, 소민경, 강무석, 정한나, 김진우 (2021). 고령자를 위한 헬스케어 대화형 에이전트 연구: 온보딩 시나리오를 중심으로. 한국 HCI학회 학술대회, 653-659.
- 김신희, 성지은 (2021). 시각적 자극의 정서가(Emotional Valence)에 따른 노년층의 대면이름대기 능력. Communication Sciences and Disorders, 26(3), 616-629.
- 김정완, 김향희 (2009). 알츠하이머성 치매 환자의 한국어 형태소 산출 특성. 언어치료연구, 18(4), 75-90.
- 김지웅, 김일곤 (2020). 챗봇을 이용한 FHIR 기반 노인우울척도 설문지의 사용성평가. 정보과학회논문지, 47(7), 650-654.
- 김향희, 김수련, 유현지, 김정완 (2013). 고령층의 말·언어 능력 평가도구의 하위 검사에 대한 내용타당도. 재활복지, 17(3), 347-373
- 모바일 인덱스 (2021). 상반기 모바일 앱 랜드스케이프 분석. 서울: 모바일 인덱스.
- 박동아 (2017). 인공지능 기반 대화형 공공 행정 챗봇 서비스에 관한 연구. 멀티미디어학회논문지, 20(8), 1347-1356.
- 성지은, 곽은정 (2012). 연령 및 동사 논항 구조에 따른 애니메이션을 활용한 동사 이름대기 과제 수행력 차이. 언어청각장애연구, 17(4), 550-564.
- 성지은, 김진경 (2011). 청년층과 노년층 간 생성이름대기에서 나타난 범주별 전형성 비교: 의미복잡성 이론에 근거한 이름대기 치료 프로토콜 개발 예비 연구, 언어치료연구, 20(1), 31-50.

- 양유미 (2016). 브로카실어증 환자의 동사 처리와 논항구조 간의 관계. 대구대학교 대학원 석사학위논문.
- 와이즈앱 리테일 굿즈 (2021). 한국인이 가장 자주 사용하는 앱. 성남: 와이즈앱 리테일 굿즈.
- 윤혜수, 성지은 (2020). 자극 제시 유형 및 동사 논항구조에 따른 실어증 환자의 동사 이름대기 특징. *Communication Sciences & Disorders*, 25(2), 399-410.
- 이용숙 (2014). 알츠하이머 치매 환자의 논항구조 산출능력: 동사 이름대기 과제를 중심으로. 충남대학교 대학원 석사학위논문.
- 채성원 (2011). 노화성 난청에 대한 고찰. *대한의사협회지*, 54(9), 908-909.
- 최수민, 최용순 (2017). 모바일 메신저 기반 인공지능 챗봇의 상품 주문결제 인터페이스 연구. 한국 HCI 학회 학술대회, 237-240.
- 최수진, 성지은 (2014). 과제 제시 유형 및 논항 수에 따른 노년층의 동사 이름대기 수행력 차이: 애니메이션과 그림 비교. *재활복지*, 18(4), 279-293.
- 최수진, 조은하, 성지은 (2021). 비대면 동사 이름대기 검사 도구 개발을 위한 기초 연구: 자극 제시 유형에 따른 비교. *언어치료연구*, 30(2), 87-97.
- 컨슈머인사이트 (2019). *컨슈머인사이트 2019년도 상반기 이동통신 기획조사*. 서울: 컨슈머인사이트.
- 통계청 (2019). *2019 고령자통계*. 대전: 통계청.
- 통계청 (2021). *생명표 2021*. 대전: 통계청.
- 현정문, 김향희, 신지철, 서상규 (2003). 베르니케실어증과 브로카실어증 환자의 명사와 동사 인출 비교. *언어청각장애연구*, 8(3), 171-187.

ORCID

신준영(제1저자, 대학원생 <https://orcid.org/0000-0003-4198-6753>); 김주은(공동저자, 대학원생 <https://orcid.org/0000-0001-5625-0226>); 송지수(공동저자, 대학원생 <https://orcid.org/0000-0003-4489-0917>); 조재은(공동저자, 대학원생 <https://orcid.org/0000-0002-6543-6812>); 최수진(공동저자, 대학원생 <https://orcid.org/0000-0003-1158-5940>); 이영미(공동저자, 교수 <https://orcid.org/0000-0003-1809-5944>); 성지은(교신저자, 교수 <https://orcid.org/0000-0002-1734-0058>)

Appendix 1. 전체 목표 동사 목록

논항	동사 목록
1항(N=11)	짚다, 날다, 웃다, 기다, 울다, 걷다, 자다, 눅다, 깨지다, 피다, 내리다
2항(N=13)	마시다, 불다, 신다, 앓다, 열다, 차다, 닦다, 굽다, 입다, 자르다, 깎다, 찢다, 꼬집다
3항(N=12)	메다, 떼다, 뽑다, 심다, 버리다, 던지다, 걸다, 업다, 따다, 뿌리다, 붙이다, 쌓다

Appendix 2. 논항별 애니메이션 자극 예시

