

The Impact of Personalized Book Reading Activities Using AI-generated Stories on Vocabulary Acquisition in Preschool Children

Jae-Eun Cho, Dongsun Yim

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

Correspondence: Dongsun Yim, PhD
Department of Communication Disorders,
Ewha Womans University, 52 Ewhayeodae-gil,
Seodaemun-gu, Seoul 03760, Korea
Tel: +82-2-3277-6720
Fax: +82-2-3277-2122
E-mail: sunyim@ewha.ac.kr

Received: July 5, 2024
Revised: August 18, 2024
Accepted: August 18, 2024

This work was supported by the Ministry of Science and ICT of the Republic of Korea and the National Research Foundation of Korea (NRF-2022R1A2C1005268).

Objectives: This study aimed to evaluate the effectiveness of book reading interventions, employing both traditional and AI-generated stories, on typically developing preschool children's vocabulary acquisition. **Methods:** A total of fifteen children aged 3 to 6 participated in this study. All participants were divided into two groups: one reading traditional stories and the other AI-generated ones. Over three weeks, participants engaged in eight structured book reading sessions, including pre-reading, main reading, and post-reading vocabulary activities. Pre- and post-intervention evaluation assessments measured receptive and expressive vocabulary skills for 16 target words. A Wilcoxon Signed-Rank Test was conducted to compare the vocabulary accuracy within each group, and a Mann-Whitney U-Test was conducted to compare the vocabulary accuracy differences between the two groups. **Results:** Statistical analyses revealed significant improvements in both groups' receptive and expressive vocabulary accuracy post-intervention, with no significant differences between the traditional and AI-generated story groups. **Conclusion:** AI-generated stories were found to effectively enhance preschool children's vocabulary skills comparably to traditional methods. This suggests the feasibility of utilizing AI to create personalized stories for effective vocabulary interventions in language therapy settings. Overall, the study highlights the potential of AI-driven approaches in overcoming traditional method limitations, offering innovative avenues for enhancing children's vocabulary development.

Keywords: AI-generated stories, Book reading interventions, Vocabulary acquisition, Preschoolers

학령 전 아동의 어휘 중재 방법은 목표 어휘의 문맥적 정보와 정의를 명확히 전달하고, 다양한 맥락에서 반복적으로 노출시켜 어휘 이해를 깊이 확장하는 것을 특징으로 한다(Kelley, 2017). 맥락 안에서 어휘를 학습할 경우, 구문이나 의미 규칙과 같은 언어적 특성뿐만 아니라, 어휘를 자연스럽게 올바르게 사용하는 방법을 습득할 수 있다(Amirian & Momeni, 2012). 여러 선행연구는 어휘를 명시적(explicit)으로 학습하는 방식과 맥락 안에서(in-context) 학습하는 방식을 결합하였을 때 효과적인 학습이 이루어지는 것을 확인하였다(Jenkins, Matlock, & Slocum, 1989; Marulis & Neuman, 2013; National Reading Panel, 2000). Coyne, McCoach와 Kapp (2007)은 맥락을 통한 어휘 중재와 명시적 어휘 중재가 학령

전 일반 아동의 어휘 학습에 미치는 효과를 비교하였는데, 책 속의 문맥을 통해서만 목표 어휘에 노출되었던 집단에 비해 목표 어휘에 대하여 이야기를 나누거나 어휘를 직접 사용해보는 등 명시적 학습을 함께 진행한 집단의 학습 효과가 높은 것으로 확인되었다. 즉, 효과적인 어휘 중재를 위해서는 명시적 방식으로 새로운 어휘를 학습하고 맥락 안에서 자연스럽게 목표 어휘를 노출하는 것이 중요하다고 할 수 있다.

5-12세 아동 대상 어휘 학습법에 대한 메타 분석 연구를 진행한 Stahl과 Fairbanks (1986)는 효과적인 어휘 학습법의 특징으로 다음 세 가지를 제시하였다. (1) 목표 어휘의 정의와 문맥에 대한 정보를 제공한다. (2) 목표 어휘의 의미를 심층적으로(in-depth) 제시한

다. (3) 학습자가 목표 어휘에 반복적으로 노출되도록 한다. Graves, Fernández, Gomez와 Schmidhuber (2006) 또한 효과적인 어휘 교수법에 대해 다음 세 가지 특징을 제시하였다. (1) 학습자가 목표 어휘를 여러 맥락에서 검토하고, 복습하고 상기하도록 한다. (2) 어휘의 의미에 대해 학습자가 적극적으로 참여하여 논의한다. (3) 각각의 어휘를 가르치고, 어휘에 대하여 논의하며 학습하는 데 충분한 시간을 할애한다. 이와 관련하여 Beck, McKeown과 Kucan (2002)이 제시한 아동 대상 어휘 중재 방법은 다음과 같다. 학습자가 새로운 어휘를 처음 접하면 먼저 학습자가 이해할 수 있는 수준의 난이도로 어휘에 대한 정의를 설명한다. 이때, 어휘의 정의는 사전적 정의를 따르는 대신 아동이 일상생활에서 접하는 표현을 사용하여 목표 어휘의 의미와 사용법을 풀어서 설명한다(Beck et al., 2002). 예를 들어, 표준국어대사전에서 확인할 수 있는 ‘저리다’의 정의는 ‘뼈마디나 몸의 일부가 오래 눌러서 피가 잘 통하지 못하여 감각이 둔하고 아리다’이다. 그러나 ‘감각’, ‘둔하다’, ‘아리다’와 같은 표현은 만 4-5세의 미취학 아동에게 낯선 표현일 수 있다. 어휘 노출 강도에 따른 책임기 어휘 중재가 일반 아동 및 단순언어장애 아동의 어휘 학습에 미치는 영향에 대하여 연구한 Park과 Yim (2019)은 어휘에 대한 아동의 이해를 돕기 위하여 ‘저리다’를 ‘몸의 어떤 부분이 오래 눌러서 피가 잘 통하지 못하여 느낌이 이상한 것’으로 정의하였다. 어휘의 의미를 심층적으로 학습할 수 있는 방법은 동의어와 반의어 익히기, 목표 어휘를 포함하는 문장 만들기, 학습자가 기존에 알고 있던 어휘와의 유사점 및 차이점에 대해 논의하기 등을 예로 들 수 있다(Stahl & Nagy, 2006). 단어연상(word association) 활동은 새로 학습한 어휘의 의미를 심층적으로 학습할 수 있는 방법 중 하나이다. 학령기 단순언어장애 아동의 어휘 중재 관련 연구를 진행한 Steele과 Mills (2011)는 아동은 목표 어휘와 유의어를 짝짓는 활동을 통해 목표 어휘의 의미를 익힐 수 있으며, 목표 어휘와 유의어가 각각 어떤 문맥에서 활용될 수 있는지에 대해 아동과 논의하며 어휘의 의미를 확장할 수 있다고 설명하였다. 아동이 동일한 어휘에 여러 차례 노출될 경우, 해당 어휘를 보다 쉽게 습득할 수 있으며(Leung & Pikulski, 1990; Penno, Wilkinson, & Moore, 2002), 아동은 어휘가 반복될수록 의미적, 음운적 표상을 보다 견고하게 구축한다(McGregor, Friedman, Reilly, & Newman, 2002). Tomasello와 Haberl (2003)는 공동주목하기 중 상호작용 대상자에 의한 어휘 사용 빈도가 높을수록 영유아의 초기 어휘 습득에 긍정적 영향을 미친다는 점을 강조하였으며, Naigles와 Hoff-Ginsberg (1998)는 특정 어휘의 총 노출 빈도, 발화 말미에 사용된 빈도, 그리고 다양한 구문 프레임에서의 사용 빈도가 영유아의 초기 동사 어휘 습득과 관련이 있다고 보고하였다. 단순언어지연 아동과 정상

아동의 빠른우연학습(QUIL) 효과 차이를 비교한 Rice, Oetting, Marquis, Bode와 Pae (1994)는 초기 어휘 이해에 있어 단순언어지연 아동과 정상발달 아동 모두 어휘 노출 빈도가 높아짐에 따라 효과적인 어휘 학습이 이루어지는 것을 확인하였다. 특히 단순언어지연 아동의 경우 특정 어휘에 반복적으로 노출될수록 정상발달 아동의 어휘 학습 능력에 상응하는 어휘력을 갖출 가능성이 높은 것으로 나타났다. Stahl과 Fairbanks (1986)와 Graves 등(2006)에서 공통적으로 제시하는 바와 같이, 아동의 어휘 학습을 효과적으로 돕기 위해서는 목표 어휘에 대한 정보를 유의어나 예문과 같은 다양한 맥락에서 제시하고, 해당 정보를 아동이 충분히 습득할 수 있도록 반복하여 노출시켜주는 것이 중요하다.

언어치료 임상에서 활용할 수 있는 다양한 어휘 중재법 중 성인과 함께하는 책임기 활동이 학령전 아동의 어휘력 향상에 효과적인 것으로 알려져있다(Noble et al., 2019; Whitehurst et al., 1988). 아동은 동화 속 삽화와 글의 맥락을 통해 자연스럽게 어휘에 노출되며, 책임기 활동을 통해 어휘력을 증진시킬 수 있다(Mol, Bus, & de Jong, 2009; Storket et al., 2017; Towson, Akemoglu, Watkins, & Zeng, 2021). Patterson (2002)은 21-27개월 스페인어-영어 이중언어 아동을 대상으로 특정 어휘가 책임기 활동을 통해 반복 노출되었을 때와 텔레비전 영상물을 통해 반복 노출되었을 때의 아동의 어휘 습득 효과를 비교하였다. 그 결과, 책임기를 통한 어휘 노출은 아동의 표현어휘력과 유의한 정적 상관을 보였으며, 영상물을 통한 어휘 노출은 아동의 표현어휘력과 유의한 관계가 없는 것으로 나타났다. 이는 책임기가 언어 학습을 위한 효과적인 수단이라는 여러 선행연구의 결과를 뒷받침한다(Mol et al., 2009; Mol & Bus, 2011). Dickinson 등(2019)은 학령전 아동을 대상으로 책임기 활동과 놀이를 통한 명시적 학습을 결합하였을 때 나타나는 어휘 학습 효과에 대해 연구하였다. 그 결과, 책임기 활동만 실시한 집단은 수용어휘력이 유의하게 증가하였으며, 책임기 활동과 함께 어휘를 명시적으로 학습할 수 있는 놀이 활동을 실시한 집단은 수용어휘력과 표현어휘력이 모두 유의하게 증가하였다. 이렇듯, 여러 선행연구에서는 책임기를 통한 어휘 중재의 중요성을 강조하였으며, 실제로 단순언어지연 아동과 정상발달 아동 모두 책 속의 문맥을 통해 성인과 상호작용하며 어휘를 학습할 때 효과적인 어휘 학습이 이루어진다는 것을 알 수 있다. 그러나, 아동의 언어능력은 기질이나 인지능력 등 개인의 내재적 요인(Paradis, 2011)과 양육자와의 상호작용 방식, 언어에 대한 질적 및 양적 노출과 같은 외재적 요인(Anderson, Graham, Prime, Jenkins, & Madigan, 2021)에 의해 각기 다른 모습으로 나타나기 때문에, 어휘 중재 시 아동이 개별적으로 필요로 하는 목표 어휘의 특성이나 범주 또한 아동에 따라 다양하게 나타난

다(Attig & Weinert, 2020; Fernald & Marchman, 2012; Weisleder & Fernald, 2013). 언어재활사는 특정 아동을 대상으로 중재 목표 어휘를 선정한 후, 어휘를 효과적으로 중재할 수 있는 적절한 활동을 선정하여 회기를 구성한다. 학령 전 아동을 위한 일반적인 어휘 중재 활동의 예시로는 장난감을 활용한 놀이, 지시따르기나 카드놀이를 비롯한 게임 형태의 활동, 책읽기 활동 등이 포함된다(Kim, 2020). 놀이나 게임, 또는 그림카드를 활용하는 방식의 경우, 언어재활사는 대상 아동이 학습을 필요로 하는 특정 어휘에 맞게 활동을 수정 및 보완하여 중재 기법을 적용하고 회기에 활용할 수 있다(Lee et al., 2024). 즉, 시중에서 구매할 수 있는 언어치료 교구를 변형하여 활용하거나, 특정 아동을 위한 중재 자료를 손쉽게 제작하는 것이 가능하다. 반면, 책을 활용하여 어휘를 중재할 경우에는 책 속에 미리 주어진 텍스트 안에서 목표 어휘를 선정해야 하므로 아동의 개별 특성이나 어휘발달 수준을 바탕으로 선정한 개별화된 목표 어휘를 다루는 데에는 현실적인 어려움이 따른다. 또한, 책 속의 제한된 문맥 안에서만 반복하며 어휘를 학습하기 때문에 다양한 문맥 안에서 쓰임새를 익히기 어렵다는 제한점이 있다(Lee et al., 2024). 따라서, 책읽기 중재를 통해 아동의 개별 목표 어휘를 효과적으로 중재할 수 있도록 특정 어휘를 포함하는 개별 맞춤형 이야기를 제작하는 방안을 모색할 필요가 있다.

인공지능(Artificial Intelligence, AI)은 인간의 인지 기능을 모방하고 수행하는 기술로, 머신러닝, 자연어 처리(Natural Language Processing, NLP), 머신비전 등을 기반으로 다양한 분야에서 적용되고 있다(Saini, 2023). 아동 발달을 추적하고 분석하는 디지털 도구도 발전하고 있으며, 대표적으로 TalkBetter (Hwang et al., 2014), TalkLIME (Song, Kim, Kim, Park, & Yim, 2016), Language Environment Analysis (LENA; Gilkerson et al., 2017) (LENA) 등이 있다. 특히, LENA는 소형 녹음기와 클라우드 기술을 활용해 아동의 언어 발달 지표를 분석하는 도구로, 이를 통해 부모교육 프로그램의 효용성을 평가하고 아동 언어 발달과 환경 간의 상관관계를 분석하는 다양한 연구가 진행되고 있다(Duncan et al., 2023; Ganek, Smyth, Nixon, & Eriks-Brophy, 2018; Weil & Middleton, 2010).

대형 언어 모델(Large Language Models, LLMs)과 인공지능의 발달은 이야기 생성의 자동화를 촉진하였다(Shakeri, Neustaedter, & DiPaola, 2021). 아동 도서를 맞춤형으로 제작하는 People in Books는 가족 구성원의 얼굴을 이야기책 삽화에 삽입하였고(Follmer, Ballagas, Raffle, Spasojevic, & Ishii, 2012), TellTable (Cao, Lindley, Helmes, & Sellen, 2010)은 주변 환경의 사물과 관련된 정보를 활용하여 아동이 직접 이야기를 제작할 수 있도록 하였다. 이외에도 Zarei, Chu, Quek, Rao와 Brown (2017)은 아동이 직

접 아바타를 이야기의 캐릭터로 구현할 수 있도록 하는 등 아동 발달 영역에서의 인공지능 융합은 개별 맞춤형 이야기책의 생성을 가능하게 하였다. Lee 등(2024)은 2-5세 일반 아동의 일상 발화를 분석하여 개별 아동의 학습 목표 어휘를 선정한 후, GPT-4를 활용해 목표 어휘를 포함한 맞춤형 이야기책을 제작하였다. 이후 4주 동안 가정에서 아동과 주 양육자가 책읽기 활동을 진행하도록 한 후, 목표 어휘와 AI-생성동화에 대한 양육자의 의견을 조사하였다. 그 결과, 양육자들은 AI-생성동화가 아동의 개별 어휘 발달 특성을 잘 반영하였고, 책읽기 활동을 통해 아동이 새로운 어휘를 습득하고 실제로 사용할 수 있었다고 응답하였다.

인공지능을 활용하면 언어재활사는 대형언어모델에 적절한 프롬프트를 제시하여 중재에 필요한 교재나 교구를 손쉽게 창작할 수 있다. 어휘 중재 시 맞춤형 동화를 제작하여 활용할 수 있다는 것은 개별 아동이 필요로 하는 특정 어휘를 포함시키기 어려웠던 기존 책읽기 중재의 한계를 극복할 수 있다는 가능성을 의미한다. Lee 등(2024)은 아동의 개별 어휘력을 반영한 맞춤형 AI 동화를 생성하고, 이를 실제 아동의 가정에서 읽게 하여 AI 동화가 아동의 어휘 습득에 미치는 영향을 확인하였다. 그러나, 어휘 중재 자료로서 AI 동화의 효용성을 보다 객관적으로 검증하기 위해서는 AI 동화를 활용한 어휘 중재 연구를 진행하고, 이를 기존의 전통적 책읽기 어휘 중재와 비교하여 분석할 필요가 있다. 따라서 본 연구에서는 대형언어모델을 활용하여 특정 목표 어휘 학습을 위한 개별 맞춤형 동화를 제작하고, 해당 동화를 활용한 책읽기 어휘 중재가 실제 아동의 어휘 학습에 유의미한 영향을 미치는지 분석하여 AI-생성 동화의 효용성을 탐구하고자 한다.

연구방법

연구대상

본 연구의 대상은 만 3-6세 정상발달 아동 15명으로, 유치원, 소셜네트워크서비스(SNS), 포털 사이트의 커뮤니티에 연구 안내문을 배포하여 참여자를 모집하였다. 선별검사 및 연구 참여는 법정 대리인의 동의하에 진행하였으며, 연구대상 선정 기준은 다음과 같다. (1) 수용·표현어휘력 검사(Receptive & Expressive Vocabulary Test, REVT; Kim, Hong, Kim, Jang, & Lee, 2009) 결과, 수용어휘력과 표현어휘력이 모두 10%ile (-1.25 SD) 이상인 아동, (2) 한국판 카우프만 간편지능검사2 (Korean Kaufman Brief Intelligence Test, Second Edition, KBIT-2; Moon, 2019) 또는 한국판 카우프만 아동 지능검사 2 (Kaufman Assessment Battery for Korean Children, Second Edition, KABC-II; Moon, 2014)의 비언어성 지능지수가 85

점(-1 SD) 이상인 아동, (3) 지능, 감각 및 정서적 결함을 보이지 않고, 신경학적 관련 장애 이력이 없으며, 구강구조 및 기능에 결함을 보이지 않는 것으로 보고된 아동. AI-생성동화의 어휘 학습 증진 효과를 확인하기 위해 모든 참여자는 기존동화 책임기 집단과 AI-생성동화 책임기 집단 중 하나의 집단에 무작위로 배치되었다. 집단 간 연령의 동질성 검증을 위해 독립표본 t -검정을 실시한 결과, 생활연령($t = .059, p = .441$), KBIT-2/KABC-II ($t = .240, p = .165$), REVT-R ($t = .957, p = .788$), REVT-E ($t = .798, p = .709$) 모두 유의한 차이를 보이지 않았다. 각 집단별 생활연령, 비언어성 지능, 수용어휘력, 표현어휘력 점수 및 독립표본 t -검정 결과는 Table 1에 제시하였다.

연구도구

기존동화

어휘 노출 강도에 따른 단순언어장애 아동의 어휘 학습 효과를 확인한 Park과 Yim (2019)의 연구와 동일한 도서를 본 연구의 기존동화 8권으로 선정하였다. 선행연구에서 참고한 도서 선정 기준은 (1) 다채로운 삽화가 포함되어 있고, (2) 삽화나 내용에 새로운 어휘가 포함되며, (3) 문장의 길이가 적절하고, (4) 이야기 장르이며, (5) 아동의 연령에 적절한 도서이다(Hargrave & Sénéchal, 2000). 선행 연구에서는 목표 어휘의 노출 정도를 통제하기 위해 각각의 목표 어휘가 동화 안에서 단 1회만 사용될 수 있도록 해당 어휘를 유사

어로 변경하거나 불필요한 문장을 삭제하는 방식으로 동화를 일부 수정하였다. 본 연구에서는 선행연구에서 수정한 동화를 그대로 사용하였다.

AI-생성동화

본 연구의 AI-생성동화는 기존동화 8권의 중심 내용을 바탕으로 GPT-4가 새로 생성한 동화를 의미한다. AI 동화 생성은 포항공대 컴퓨터공학과 인간중심 지능형 시스템 연구실과 이화여대 언어병리학과 아동언어연구실이 공동 개발한 아동 개별 맞춤형 동화 생성 시스템을 활용하여 제작하였다. 시스템은 로컬 Linux서버에서 GPT-4 (OpenAI, 2023) 및 DeepL (DeepL, 2023) 클라우드 API와 함께 Python으로 구현하였다. GPT-4는 한국어를 지원하지 않지만, 동화 생성 작업을 한국어로 진행할 경우 동음이의어와 같은 어휘의 특성을 반영하기 어렵고(e.g., 목표 어휘 ‘금(crack)’을 ‘금(gold)’으로 인식), 목표 어휘가 문맥에 부적절하거나 부자연스러운 형태로 포함되었다. 따라서 동화를 영어로 생성하고, 결과물에 대한 한국어 번역 단계와 전문가 검토 작업 단계를 추가로 진행하는 방식을 택하였다. 각각의 AI-생성동화는 글의 총 길이가 기존 동화의 80-120% 수준이며, 기존동화의 대화체 문장 수에 준하는 정도의 대화체를 포함하도록 통제하였다. 또한, 기존 동화와 마찬가지로 모든 목표 어휘가 이야기 안에서 1회만 사용되도록 별도의 수정 과정을 거쳤다. AI 동화 제작과정은 Figure 1에 제시하였다.

Table 1. Demographic information on participants and results of independent sample t -test

| Characteristic | Traditional stories (N=7) | | AI-generated stories (N=8) | | t | p |
|----------------|---------------------------|-------|----------------------------|-------|------|------|
| | Mean | SD | Mean | SD | | |
| Age (mo) | 61.14 | 10.21 | 60.88 | 7.43 | .059 | .441 |
| KBIT-2 | 114.57 | 25.15 | 111.88 | 18.30 | .240 | .165 |
| KABC-II | | | | | | |
| REVT-R | 64.29 | 14.21 | 57.50 | 13.25 | .957 | .788 |
| REVT-E | 76.57 | 21.20 | 69.38 | 13.36 | .798 | .709 |

KBIT-2=Korean Kaufman brief intelligence test, Second edition (Moon, 2019); KABC-II=Kaufman assessment battery for Korean children, Second edition (Moon, 2014); REVT-R=Receptive & expressive vocabulary test–receptive (Kim et al., 2009); REVT-E=Receptive & expressive vocabulary test–expressive (Kim et al., 2009).

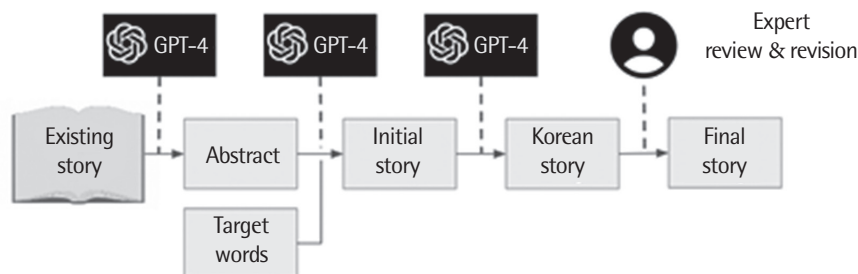


Figure 1. AI story generation process.

AI-동화 생성 단계는 (1) 이야기 틀 생성 단계, (2) 목표 어휘를 포함하는 영문 초안 생성 단계, (3) 초안 검토 및 수정사항 반영 단계, (4) 한국어 번역본 생성 단계, (5) 국어교육 전문가 검증 단계, (6) 전자문서 형식 구현 단계에 따른다. 각 단계별 세부 작업 내용은 Appendix 1에 제시하였다.

자동화 시스템을 통해 생성된 한국어 번역본이 대상 아동의 연령과 연구의 목적에 적절한지에 대해 언어학적 측면에서 검증하는 단계를 실시하였으며, 검증에는 1급 중등학교 정교사 자격을 소지한 25년 이상 경력의 국어교육 전문가가 참여하였다. 연구자는 국어교육 전문가에게 본 연구의 AI-동화 생성 단계에 대해 충분히 설명하였으며 국문 번역본 검토 및 수정 가이드라인에 따라 결과물을 검토 및 수정하도록 안내하였다. 전문가 검증 단계에서 참고한 국문 번역본 검토 및 수정 기준은 Appendix 2에 제시하였다. 연구자는 시스템을 통해 생성한 한국어 번역본과 각 동화 안에 포함되어야 할 목표 어휘 목록을 국어교육 전문가에게 전달하였으며, 국어교육 전문가는 검토 및 수정 기준에 따라 결과물이 아동의 어휘 중재에 활용되기에 적절한지에 대해 언어학적 측면을 바탕으로 검증하였다. GPT-4의 국문 번역본에서는 어색한 번역투, 대상자 연령에 적절하지 않은 어휘, 지나치게 긴 문장 사용과 같은 부적절한 항목이 관찰되었다. 전문가 검증 절차를 거친 결과물은 연구자의 최종 검토 후 책읽기 중재 자료로 활용되었다. 국어교육 전문가의 번역본 검토 및 수정작업 예시는 Table 2에 제시하였다.

사진 자료

사진 자료는 선행연구(Kim, Chae, & Yim, 2020; Park & Yim, 2019)에서 사용한 자료와 동일한 사진을 사용하였으며, 사전-사후 표현어휘 검사용, 사전-사후 수용어휘 검사용, 책읽기 활동용으로 총 세 가지 용도로 구분된다. 사진 선정 과정은 다음과 같다.

사전-사후 표현어휘 검사용

목표 어휘를 나타내기에 적합하다고 판단되는 사진을 1차로 선정한다. 이후, 언어병리학 석사과정 이상 10인에게 사진을 제시한 후 사진을 보고 연상되는 단어를 모두 기재하도록 한다. 이때, 목표

어휘, 유의어, 정의에 적절한 응답을 하는 경우 1점을 부여하고, 합산 점수가 5점 이상인 사진을 표현어휘 검사용 사진으로 선정한다. 사진은 전자기기의 스크린을 통해 아동에게 보여주기에 용이하도록 Google Presentation Slides에 삽입하여 제작하였고, 책읽기 중재 사전-사후 표현어휘 검사용 사진으로 사용하였다.

사전-사후 수용어휘 검사용

목표 어휘 사진 한 장과 다른 의미를 가진 사진 세 장을 오답용 사진으로 선정한다. 오답용 사진과 목표 어휘 사진으로 구성된 네 개의 항목 중 정답을 고르도록 하고, 해당 구성이 적절한지에 대해 언어병리학 석사과정 이상 2인에게 Likert 5점 척도로 평가하도록 한다. 평균 점수가 4점 이상인 사진을 수용어휘 검사용 자료로 선정한다.

책읽기 활동용

책읽기 활동에 사용할 사진은 목표 어휘를 포함한 문장이 잘 묘사할 수 있는 사진으로 선정한다. 문장과 사진이 적절한지에 대해 언어병리학을 전공한 박사 1인이 Likert 5점 척도로 평가하여 4점 이상 획득한 자료를 중재에 사용한다.

연구절차

본 연구는 AI-생성동화를 활용한 책읽기 중재와 기존동화를 활용한 책읽기 중재가 아동의 어휘 학습에 미치는 영향을 살펴보고, 중재 활동 조건에 따른 차이가 나타나는지 확인하기 위한 목적으로 중재 사전-사후 검사 설계를 적용하였다. 연구 절차는 선별 검사, 목표 어휘 선정을 위한 사전 평가, 책읽기 중재 8회기, 사후 평가의 순서로 진행하였다.

선별 검사

수용·표현어휘력 검사(REVT; Kim et al., 2009)를 실시하여 어휘력을 평가하였고, 한국판 카우프만 간편지능검사2 (KBIT-2; Moon, 2019)를 실시하여 비언어성 지능을 평가하였다. 만 4세 미만의 대상자는 한국판 카우프만 아동 지능검사2 (KABC-II, Moon,

Table 2. Example of reviewing and revising a Korean translation of AI-generated stories

| GPT-4 Korean translation | Expert review and revision | Revised contents |
|--|--|--|
| 구지구지는 등에 금 ^a 이 가는 이상한 모습을 가지고 있었고, 다른 오리 새끼들보다 훨씬 더 단단한 살갗 ^a 을 가지고 있었어요. 구지구지의 납작한 ^a 주둥이와 힘센 손은 오리 새끼로서는 매우 이례적이었습니다. | 구지구지는 등에 이상한 금 ^a 이 가 있었고, 다른 새끼 오리들보다 훨씬 더 단단한 살갗 ^a 을 가지고 있었어요. 그리고 납작한 ^a 주둥이와 힘센 손은 다른 새끼 오리들과는 정말 달랐어요. | <ul style="list-style-type: none"> • Correction of awkwardly translated sentences • Adjustment of inappropriately complex vocabulary • Remove repetitive subjects and connect naturally with conjunctions |

^atarget word.

2014)로 비언어성 지능을 평가하였다. 비언어성 지능 검사의 표준 점수가 85점 미만인 참여 아동 2인과 수용어휘력 또는 표현어휘력이 10%ile 미만인 아동 1인, 그리고 집중력 저하 문제로 검사 참여에 어려움을 보인 아동 1인의 데이터는 분석에서 제외하였다.

사전 평가

목표 어휘 선정 방식은 선행연구 Park과 Yim (2019)을 참고하였다. 먼저, 책읽기 활동에 사용할 기존동화 여덟 권에서 명사 23개, 동사 18개, 형용사 19개, 총 60개의 후보 어휘를 선정하였다. 이후, 후보 어휘 60개에 대한 수용 및 표현 어휘 검사를 실시하여 각각의 아동이 이해 및 표현에 어려움을 보이는 어휘 16개를 해당 아동의 최종 목표 어휘로 선정하였다. 검사는 수용·표현어휘력검사(REVT; Kim et al., 2009)와 동일한 방식으로 진행하였으며, 목표 어휘 선정 기준은 다음과 같다. (가) 수용어휘와 표현어휘 모두 오반응한 문항을 우선 순위로 선정한다. (나) 수용어휘는 정반응하였으나 표현어휘에서 오반응한 문항을 선정한다. (다) 수용어휘에서 정반응하였으나 표현어휘에서 오반응한 문항이 여러개인 경우, “(목표 어휘)가 뭐야?”라는 연구자의 질문에 아동이 어휘의 개념 설명에 어려움을 보이는 문항을 선정한다.

책읽기 중재

책읽기 중재 시, 어휘 노출 빈도에 따른 단순언어지연 아동의 어휘 학습 효과를 확인한 선행연구에 따르면, 아동은 목표 어휘가 36번 노출되었을 때 학습 효과가 가장 높은 것으로 나타났다(Storkel et al., 2017). 이에 본 연구에서는 대상 아동이 중재 기간 동안 목표 어휘에 총 36회 노출될 수 있도록 회기를 구성하였다. 아동은 주 3회, 20분씩, 대면 또는 비대면으로 총 8회의 중재 회기에 참여하였으며, 각 아동에게는 네 권의 동화가 무작위로 배정되었다. 아동의 연령과 집중력을 고려하여 매 회기 두 권의 동화를 읽어주었으며, 목표 어휘는 각 동화별로 4개씩 하루에 총 8개를 중재하였다. 네 권의 동화는 2회기당 한 번씩 중재에 사용하여, 8회기 동안 총 네 번 반복하여 읽었으며, 목표 어휘는 동화를 네 번 반복하여 읽는 동안 각 9회씩 총 36회 노출하였다. 아동이 책읽기 도중 대화를 시도하거나 이야기 또는 어휘와 관련하여 질문을 할 경우, 연구자는 아동과 자연스럽게 대화할 수 있지만, 목표 어휘를 노출하거나 목표 어휘 학습에 도움이 되는 단서를 제공하지 않도록 유의하여 부수적인 학습 효과를 최소화하였다. 책읽기 중재는 책읽기 전 어휘학습 단계, 책읽기 단계, 책읽기 후 어휘학습 단계로 구성하였다. 책읽기 전 어휘학습 단계에서 연구자는 아동에게 “우리, 책을 읽기 전에 책에 나올 단어들을 먼저 알아보자.”라고 이야기한 후 목표 어휘를 묘사

하는 사진을 제시하였다. 이어서, 유의어와 어휘의 정의를 설명해주었다. 이후, 새로운 그림 자료를 제시하며 목표 어휘가 다른 맥락에서 어떻게 사용되는지 설명하였다. 책읽기 단계에서 목표 어휘가 나오면 연구자는 본문을 읽은 후 이어서 목표 어휘의 유의어를 언급하였다. 책을 아동에게 모두 읽어준 후, 연구자는 “책에 나왔던 단어들이 다른 데에서 어떻게 쓰이는지 알아보자.”라고 말한 후, 앞서 설명한 맥락 사진을 제시하면서 목표 어휘가 포함된 문장을 읽어주었다. 이어서 목표 어휘에 맞는 사진을 제시하며 목표 어휘의 유의어와 정의를 다시 한 번 설명하였다. 두 집단의 중재 회기 일정과 중재스크립트 예시는 Appendix 3과 Appendix 4에 제시하였다.

사후 평가

책읽기를 통한 어휘 학습 효과를 확인하기 위해 책읽기 활동 8회를 모두 완료한 후 사전검사와 동일한 방법으로 사후 평가를 진행하였다. 사전 평가에서는 목표 어휘를 선정하기 위해 60개의 후보 어휘를 모두 검사하였지만, 사후 평가에서는 아동별 목표 어휘 16개 문항에 대해서만 검사를 진행하였다.

자료분석

표준화 검사도구

표준화 검사도구인 한국판 카우프만 간편지능검사2 (KBIT-2; Moon, 2019), 한국판 카우프만 아동 지능검사2 (KABC-II, Moon, 2014), 한국어 수용·표현어휘력 검사(REVT; Kim et al., 2009)의 분석은 검사도구의 지침서에 따라 실시하였다. 수용·표현어휘력 검사는 통계적 처리 시 원점수로 기입하였다.

책읽기 활동 목표 어휘 학습 효과

책읽기 활동 목표 어휘 16개를 대상으로 수용어휘력과 표현어휘력 평가를 실시하였다. 아동의 반응은 기록지에 모두 기록하였으며, 정반응한 항목에는 1점, 오반응한 항목에는 0점을 부여하였다. 평가는 책읽기 활동 전과 활동 후 각 1회씩 총 2회 실시하였다.

통계적 처리

AI-생성동화를 활용한 책읽기 활동과 기존동화를 활용한 책읽기 활동이 아동의 어휘 학습에 미치는 효과를 확인하기 위하여 비모수통계 방법을 사용하였다. Wilcoxon 부호-순위 검정(Wilcoxon Signed-Rank Test)을 실시하여 어휘 평가 시점(사전 vs. 사후)에 따른 각 집단의 수용 및 표현어휘 정반응률을 비교하였으며, 맨-휘트니 U-검정(Mann-Whitney U-Test)을 실시하여 책읽기 중재 활동 조건(기존동화 vs. AI-생성동화)에 따른 집단 간 사전-사후 수용 및

표현어휘 정반응률을 확인하였다. 모든 통계적 처리는 IBM SPSS Statistics 29를 사용하였다.

신뢰도

목표 어휘에 대한 수용 및 표현어휘 평가의 신뢰도를 검증하기 위하여 검사자 간 신뢰도를 산출하였다. 목표 어휘에 대한 사전-사후 평가는 연구자가 직접 실시하였으며, 신뢰도 검증을 위해 연구에 참여하지 않은 언어병리학과 석사학위 소지자 1인을 제2 평가자로 선정하였다. 제2 평가자는 제1 평가자가 실시한 평가의 녹음파일과 기록지 사본을 사용하여 독립적으로 채점하였다. 각 문항에 대하여 아동이 정반응한 문항은 1점, 오반응한 문항은 0점을 부여하였다. 평가자 간 신뢰도는 두 평가자가 채점한 점수 중 서로 일치한 점수를 전체 문항 수로 나누고 곱하기 100을 하여 구하였다. 그 결과, 평가신뢰도는 100% 일치하는 것으로 나타났다.

중재 충실도

중재 충실도 검증을 위해 언어병리학과 석사학위 소지자 1인이 연구자의 녹음된 중재 음성과 중재 스크립트를 대조하여 5개의 평가 문항에 대해 4점 척도로 점수를 부여하였다. 중재 충실도 평가 문항은 다음과 같다. (1) 책읽기와 어휘 학습 활동은 적절한 언어를 사용하여 진행되었는가? (2) 책읽기 활동 시 연구자는 스크립트를

올바르게 읽었는가? (3) 어휘 학습 활동 시 연구자는 스크립트를 올바르게 읽었는가? (4) 목표 어휘의 노출 빈도를 회기당 9회로 일정하게 유지하였는가? (5) 스크립트 이외의 발화에서 목표 어휘 학습에 도움이 될 만한 단서를 최소화하였는가?

그 결과, 연구자의 중재 충실도는 98.75%로 나타났다.

연구결과

평가 시점(사전 vs. 사후)에 따른 기존동화 집단과 AI-생성동화 집단의 수용 및 표현어휘 정반응률

기존동화를 활용한 책읽기 중재가 아동의 수용 및 표현어휘 학습에 미치는 영향을 확인하기 위해 일측근 부호-순위 검정을 실시하여 집단 내 사전-사후 수용 및 표현어휘의 정반응률을 비교하였다. 기존동화 집단의 사전-사후 수용 및 표현어휘 정반응률에 대한 기술통계량 및 일측근 부호-순위 검정 결과는 Table 3과 같다.

분석 결과, 기존동화를 활용한 집단의 사후 수용어휘 정반응률($M=94.57, SD=9.27$)이 사전 수용어휘 정반응률($M=18.86, SD=24.02$)에 비해 유의하게 높은 것으로 나타났다($Z=-2.384, p<.05$). 표현어휘의 경우, 사후 표현어휘 정반응률($M=55.43, SD=22.69$)이 사전 표현어휘 정반응률($M=.00, SD=.00$)에 비해 유의하게 높은 것으로 나타났다($Z=-2.371, p<.05$). 기존동화 집단 아동의 사

Table 3. Descriptive statistics and Wilcoxon signed-rank test results for pre-post vocabulary accuracy for traditional stories group

| N=7 | Pre | | Post | | Z | p |
|-----------------------|-------|-------|-------|-------|--------|-------|
| | Mean | SD | Mean | SD | | |
| Receptive vocabulary | 18.86 | 24.02 | 94.57 | 9.27 | -2.384 | .017* |
| Expressive vocabulary | .00 | .00 | 55.43 | 22.69 | -2.371 | .018* |

* $p<.05$.

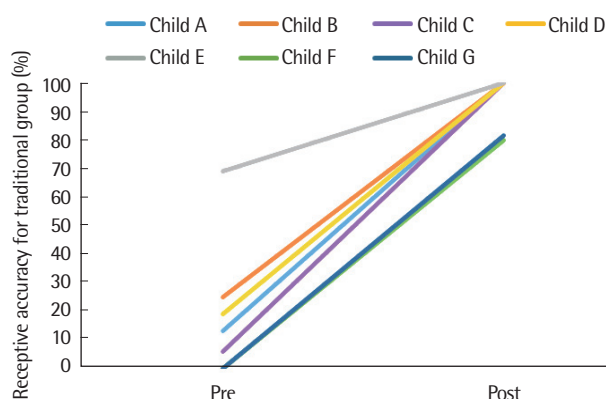


Figure 2. Receptive vocabulary pre-post test accuracy for traditional story group.

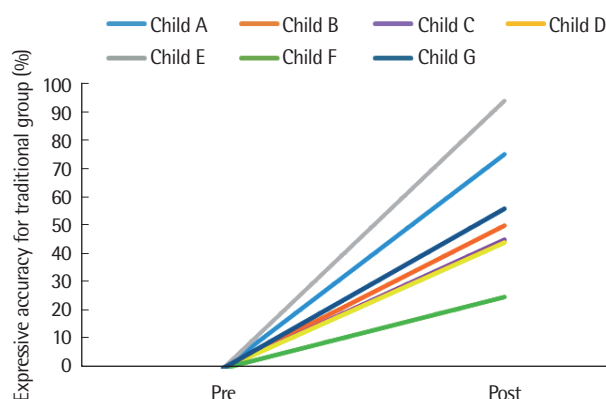


Figure 3. Expressive vocabulary pre-post test accuracy for traditional story group.

전-사후 수용어휘 정반응률에 대한 결과는 Figure 2와 같으며, 사전-사후 표현어휘 정반응률에 대한 결과는 Figure 3과 같다.

AI-생성동화를 활용한 책읽기 중재가 아동의 수용 및 표현어휘 학습에 미치는 영향을 확인하기 위해 월콕슨 부호-순위 검정을 실시하여 집단 내 사전-사후 수용 및 표현어휘의 정반응률을 비교하였다. AI-생성동화 집단의 사전-사후 수용 및 표현어휘 정반응률에 대한 기술통계량 및 월콕슨 부호-순위 검정 결과는 Table 4와 같다.

분석 결과, AI-생성동화를 사용한 집단의 사후 수용어휘 정반응률($M=90.75$, $SD=8.89$)이 사전 수용어휘 정반응률($M=9.50$, $SD=13.57$)에 비해 유의하게 높은 것으로 나타났다($Z=-2.527$, $p<.05$). 표현어휘의 경우, 사후 표현어휘 정반응률($M=47.88$, $SD=33.07$)이 사전 표현어휘 정반응률($M=.00$, $SD=.00$)에 비해 유의하게 높았다($Z=-2.536$, $p<.05$). AI-생성동화 집단 아동의 사전-사후 수용어휘 정반응률에 대한 결과는 Figure 4와 같으며, 사전-사후 표현어휘 정반응률에 대한 결과는 Figure 5와 같다.

책읽기 중재 활동 조건(기존동화 vs. AI-생성동화)에 따른 사전-사후 수용 및 표현어휘 정반응률

책읽기 중재 활동 조건(기존동화 vs. AI-생성동화)에 따른 집단 간 사전 수용 및 표현어휘 정반응률에 유의한 차이가 있는지 확인하기 위해 맨-휘트니 U -검정을 실시하였다. 집단 간 사전 수용 및 표현어휘 정반응률에 대한 기술통계량 및 맨-휘트니 U -검정 결과는 Table 5에 제시하였다.

사전 수용어휘 정반응률에 대한 분석 결과, 기존동화 집단의 사전 수용어휘 정반응률($M=18.86$, $SD=24.02$)과 AI-생성동화의 사전 수용어휘 정반응률($M=9.50$, $SD=13.57$) 간 차이가 통계적으로 유의하지 않았다($Z=-.899$, $p>.05$). 사전 표현어휘 정반응률에 대한 분석 결과, 기존동화 집단의 사전 표현어휘 정반응률($M=.00$, $SD=.00$)과 AI-생성동화의 사전 표현어휘 정반응률($M=.00$, $SD=.00$) 간 차이가 통계적으로 유의하지 않았다($Z=.000$, $p>.05$). 집단 간 사전 수용어휘 및 표현어휘 정반응률에 대한 그래프는

Table 4. Descriptive statistics and Wilcoxon signed-rank test results for pre-post vocabulary accuracy for AI-generated stories group

| N=7 | Pre | | Post | | Z | p |
|-----------------------|------|-------|-------|-------|--------|-------|
| | Mean | SD | Mean | SD | | |
| Receptive vocabulary | 9.50 | 13.57 | 90.75 | 8.89 | -2.527 | .012* |
| Expressive vocabulary | .00 | .00 | 47.88 | 33.07 | -2.536 | .011* |

* $p<.05$.

Table 5. Descriptive statistics and Mann-Whitney U -Test results for pre-test receptive and expressive vocabulary accuracy by group

| | Traditional stories (N=7) | | AI-generated stories (N=8) | | Z | p |
|-----------------------|---------------------------|-------|----------------------------|-------|-------|-------|
| | Mean | SD | Mean | SD | | |
| Receptive vocabulary | 18.86 | 24.02 | 9.50 | 13.57 | -.899 | .369 |
| Expressive vocabulary | .00 | .00 | .00 | .00 | .000 | 1.000 |

SR= Subject relative clauses; OR= Object relative clauses.

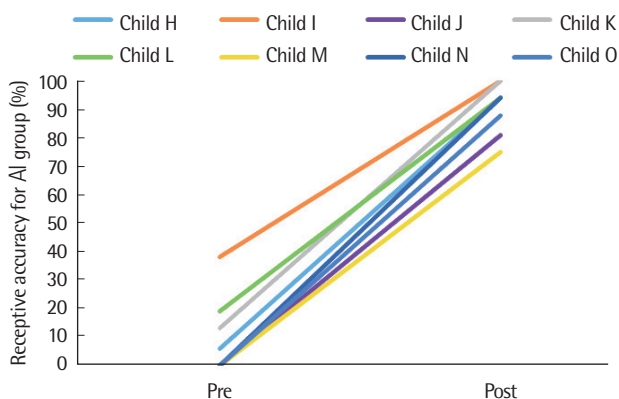


Figure 4. Receptive vocabulary pre-post test accuracy for AI-generated story group.

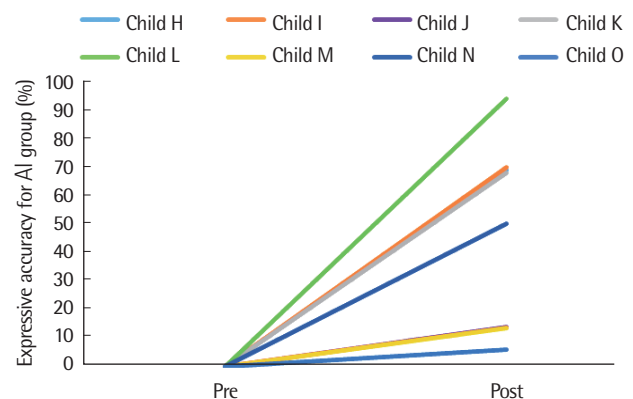


Figure 5. Expressive vocabulary pre-post test accuracy for AI-generated story group.

Table 6. Descriptive statistics and Mann-Whitney *U*-Test results for post-test receptive and expressive vocabulary accuracy by group

| | Traditional stories (N = 7) | | AI-generated stories (N = 8) | | Z | p |
|-----------------------|-----------------------------|-------|------------------------------|-------|--------|------|
| | Mean | SD | Mean | SD | | |
| Receptive vocabulary | 94.57 | 9.27 | 90.75 | 8.89 | -1.230 | .219 |
| Expressive vocabulary | 55.43 | 22.69 | 47.88 | 33.07 | -.350 | .727 |

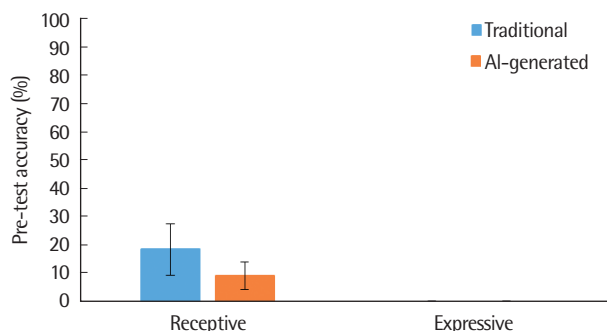
**Figure 6.** Receptive and expressive vocabulary pre test accuracy for both groups.

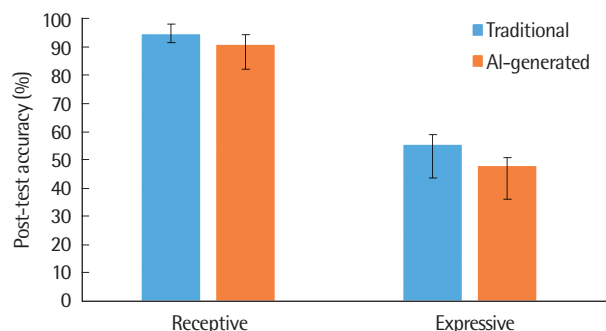
Figure 6에 제시하였다.

책읽기 중재 활동 조건(기존동화 vs. AI-생성동화)에 따른 집단 간 사후 수용 및 표현어휘 정반응률에 유의한 차이가 있는지 확인하기 위해 맨-휘트니 *U*-검정을 실시하였다. 집단 간 사후 수용 및 표현어휘 정반응률에 대한 기술통계량 및 맨-휘트니 *U*-검정 결과는 Table 6에 제시하였다.

사후 수용어휘 정반응률에 대한 분석 결과, 기존동화 집단의 사후 수용어휘 정반응률($M = 94.57$, $SD = 9.27$)과 AI-생성동화의 사후 수용어휘 정반응률($M = 90.75$, $SD = 8.89$) 간 차이가 통계적으로 유의하지 않았다($Z = -1.230$, $p > .05$). 사후 표현어휘 정반응률에 대한 분석 결과, 기존동화 집단의 사후 표현어휘 정반응률($M = 55.43$, $SD = 22.69$)과 AI-생성동화의 사후 표현어휘 정반응률($M = 47.88$, $SD = 33.07$) 간 차이가 통계적으로 유의하지 않았다($Z = -.350$, $p > .05$). 집단 간 사후 수용어휘 및 표현어휘 정반응률에 대한 그래프는 Figure 7에 제시하였다.

논의 및 결론

본 연구는 AI-생성동화와 기존동화를 활용한 책읽기 중재를 통해 아동의 어휘 습득 효과를 확인하고 어휘 중재 도구로서의 AI-생성동화의 효용성을 확인하고자 하였다. 이에 만 3-6세의 아동 15명을 AI-생성동화 집단과 기존동화 집단으로 나누어 3주간 책읽기 활동을 실시하고, 어휘 중재의 효과를 확인하기 위하여 중재 전과 후에 목표어휘에 대한 수용 및 표현어휘력 평가를 진행하였다.

**Figure 7.** Receptive and expressive vocabulary post test accuracy for both groups.

기존동화를 활용한 책읽기 중재의 효과를 확인하기 위해 집단 내 사전-사후 수용어휘 및 표현어휘의 정반응률을 분석하였다. 그 결과, 사후 수용어휘 정반응률이 사전 수용어휘 정반응률에 비해 유의하게 높았다. 표현어휘의 경우, 사후 표현어휘 정반응률이 사전 표현어휘 정반응률에 비해 유의하게 높았다. 이러한 연구결과는 책읽기를 통한 어휘 중재가 아동의 어휘 학습에 효과적이라는 선행연구들의 결과와 일치한다(Mol et al., 2009; Mol & Bus, 2011). 기존 집단 동화의 사후 표현어휘 정반응률($M = 55.43$, $SD = 22.69$)은 사후 수용어휘 정반응률($M = 94.57$, $SD = 9.27$)에 비해 대상자 간 편차가 큰 것으로 나타났다. 어휘 표현 시 음운을 조작하고 산출해 내는 고차원적인 언어 처리 과정이 요구된다는 점을 고려한다면, 표현어휘 결과에서 확인할 수 있는 대상자 간 편차는 아동의 집행 기능 및 기질과 같은 개인의 내적 요소 차이에 기인하는 것으로 추측할 수 있다(Vallotton, & Ayoub, 2011; Yang & Yim, 2018). 효과적인 어휘 중재를 위해 본 연구에서는 선행연구에서 제시하는 바와 같이 목표 어휘가 포함된 이야기 속에서 문맥에 따른 어휘의 사용법을 익히고, 책읽기 전과 후에 목표 어휘를 명시적으로 학습하는 기회를 제공하였다. 또한 Stahl과 Fairbanks (1986)와 Graves 등 (2006)에서 설명하는 효과적인 중재의 특징에 따라 어휘의 정의를 설명해주고, 어휘의 유의어와 활용 문맥에 대해 학습하여 어휘 지식을 심층적으로 넓혔으며, 또한 명시적 학습과 문맥에서의 학습을 결합한 방식을 통해 각 회기 내에서 목표 어휘가 반복되어 노출되도록 설정하였다. 연구자는 미리 중재 스크립트를 준비하였으며 스크립트에 따라 중재에 충실히 임하였다. 체계적으로 구조화된 어

휘 학습 활동 안에서 아동은 주어진 목표 어휘를 효과적으로 학습할 수 있었다. 아동 대상 책임기 어휘 중재 시 이를 참고하여 회기를 구성한다면 어휘 학습 효과를 기대할 수 있을 것이다.

AI-생성동화를 활용한 책임기 활동의 중재 효과를 확인하기 위해 집단 내 사전-사후 수용어휘 및 표현어휘 정반응률을 분석하였다. 그 결과, 사후 수용어휘 정반응률이 사전 수용어휘 정반응률에 비해 유의미하게 높았으며, 사후 표현어휘 정반응률이 사전 표현어휘 정반응률에 비해 유의미하게 증가하였다. 이러한 결과는 본 연구에서 기존동화 집단을 통해 확인한 결과와 일치하며, AI-생성동화를 활용한 책임기 활동이 아동의 어휘력 향상에 기여할 수 있음을 보여준 Lee 등(2024)의 연구와도 일관성을 보인다. 이는 AI-생성동화를 활용한 책임기 중재가 아동의 어휘력 증진에 효과적인 수 있음을 시사한다. 그러나 본 연구의 AI-생성동화 중재 효과에 대한 결과를 해석할 시에는 AI-생성동화 제작 시 OpenAI사의 GPT-3.5와 GPT-4를 기반으로 하는 ChatGPT 시스템을 그대로 사용하지 않았다는 점을 고려할 필요가 있다. 본 연구에서 사용한 AI동화 생성 시스템은 언어병리학 전문가들의 엄격한 검증을 거쳐 선정한 기준에 따라 이야기를 생성하도록 구조화되어 있으며, 이를 바탕으로 컴퓨터공학 전문 엔지니어들이 로컬 서버에서 본 연구 디자인에 적합한 프롬프트를 적용하여 시스템을 구현하였다. 이뿐만 아니라, 중요한 언어 발달 시기에 있는 실제 아동을 대상으로 하는 연구인 만큼, 동화의 내용이 연구 대상자에게 적절하며 언어학적 측면에서 오류를 보이지 않는지 검증하는 절차를 포함하였다. 즉, 본 연구에서는 AI-생성동화 제작을 인공지능에 100% 의존하는 대신 인간이 동화 생성 작업에 일부 참여하였고, 사전에 구조화한 중재 스크립트에 따라 통제된 상황에서 연구자와 아동이 1:1 책임기 활동을 실시하였기 때문에 이를 통한 어휘 중재 효과가 유의한 것은 어느정도 예견되었던 결과라고 할 수 있다. 이는 ChatGPT를 사용하여 동화를 생성하고 어휘 중재를 실시하였을 때 그 결과가 본 연구의 결과와 일치하지 않을 수 있다는 가능성을 내포하므로 결과 해석 시 이를 반드시 고려해야 한다.

책임기 중재 활동 조건(기존동화 vs. AI-생성동화)에 따른 사전-사후 수용 및 표현어휘력을 비교하기 위해 집단 간 사전-사후 수용 및 표현어휘의 정반응률을 분석하였다. 연구결과, 두 집단 간 사전 수용어휘 정반응률에 유의한 차이가 나타나지 않았으며, 사전 표현어휘 정반응률에 유의한 차이가 나타나지 않았다. 즉, 사전 수용 및 표현어휘 정반응률은 책임기 중재 조건에 따른 차이가 없는 것으로 확인되었다. 집단 간 사후 수용 및 표현어휘의 정반응률을 분석한 결과, 집단 간 사후 수용어휘 정반응률의 차이가 유의하지 않았으며, 사후 표현어휘 정반응률의 차이가 유의하지 않았다. 이는

스크립트 기반의 어휘 중재 시, AI-생성동화를 활용한 책임기 어휘 중재가 기존동화를 활용하는 전통적 방식의 책임기 어휘 중재에 준하는 수준으로 아동의 어휘력 증진에 긍정적 영향을 미친다는 점을 시사한다. 전통적 방식의 책임기 중재는 책 속에 주어진 제한된 문맥 안에서만 어휘를 학습하기 때문에 다양한 문맥 안에서 어휘가 어떤 형태로 쓰이는지에 대해 학습하기 어렵다는 제한점이 있다. 반면 대형언어모델 인공지능을 활용한다면 언어재활사는 특정 어휘를 포함하는 동화를 빠르고 쉽게 창작할 수 있다(Kasneeci et al., 2023). 이는 아동 개개인의 어휘능력을 반영하여 선정한 개별 목표 어휘를 포함하는 동화 생성이 가능하다는 것을 의미한다. 또한 인공지능을 활용하면 특정 어휘를 포함하는 비슷한 내용의 동화를 얼마든지 새롭게 생성할 수 있기 때문에 아동은 하나의 중심 내용을 공유하는 여러 개의 동화를 반복적으로 읽으며 새로운 문맥 안에서 목표 어휘의 다양한 쓰임새를 학습할 수 있다. 본 연구의 결과는 구조화된 스크립트 기반의 책임기 중재 방식이 아동의 어휘력을 효과적으로 향상시킬 수 있음을 보여주며, 또한 AI-생성동화가 기존 동화에 준하는 어휘 학습 중재 도구로서 활용될 수 있음을 시사한다. 이는 아동언어발달과 인공지능 기술의 융합 가능성을 제시하며, 아동 언어 분야의 새로운 연구 방향을 제시한다는 점에서 중요한 의의를 갖는다.

본 연구의 제한점은 다음과 같다. 첫째, AI동화 생성 작업을 영어로 진행한 후 최종 단계에서 한국어로 번역하는 과정을 추가하였다. 이는 GPT-4의 한국어 성능이 영어에 비해 다소 제한적이기 때문에 선택한 방법이다. 한국어로 이야기를 생성하도록 프롬프트를 제시하였을 때 GPT-4의 결과물은 동음이의어 번역의 오류가 빈번하였으며, 한국어의 고유한 특징인 조사 및 연결어미의 사용을 자연스럽게 반영하지 못하였다. 그러나 대형언어모델과 인공지능이 빠른 속도로 발전하고 있다는 점을 고려한다면 이러한 문제는 시간이 지남에 따라 점차 축소될 것으로 예상된다. 둘째, GPT-4가 AI-동화를 한국어로 번역하는 과정에서 발생하는 형태론적, 의미론적 문제를 보완하기 위하여 국어전문가의 검증을 거치는 과정을 진행하였다. 인공지능 기술이 진화함에 따라 대형언어모델의 소수 언어 기능이 향상될 것으로 기대되지만, 결과물이 각 언어의 특성을 적절히 반영하고 사용자의 목적에 부합하는지 여부를 검증하기 위해서는 인간이 일부 동화 제작 과정에 개입할 필요가 있다. 따라서 효과적인 어휘 중재 도구 개발 시스템을 구축하기 위해서는 자동화 프로세스에 인간이 어느 정도로 개입할지에 대해 지속적으로 연구하고 검토해나가야 할 것이다. 셋째, 본 연구의 대상 아동들은 특정 60개 어휘에 대한 사전 평가를 진행하였으며, 연구자는 평가 결과를 바탕으로 60개 중 16개의 어휘를 목표 어휘로 선정하였다.

따라서 아동의 어휘 수준에 따라 대상자 간 목표 어휘의 개별화를 부분적으로 실현하였지만, 아동 고유의 발달적, 환경적 특성에 따라 형성된 개별 어휘능력을 반영하지 못하였다는 제한점이 있다. 추후 연구를 통해서 대상 아동의 실제 어휘능력을 면밀히 분석하여 실질적인 학습이 필요한 독립적인 목표 어휘를 선택하고, 인공지능을 활용하여 이를 효과적으로 증재하는 방안을 모색할 필요가 있다. 마지막으로, 본 연구의 대상자 수는 15명으로, 연구의 결과가 책임기를 통한 아동의 일반적인 어휘 학습능력을 반영한다고 보기에는 한계가 있다. 연구결과의 일반화를 위해서는 추후 보다 많은 수의 아동을 대상으로 연구가 진행되어야 할 것이다. Chat GPT (OpenAI, 2022)가 제안하는 언어 증재 활동의 임상적 유용성을 확인한 Du와 Juefei-Xu (2023)의 연구에 따르면, 언어재활사는 ChatGPT의 번역 기능을 활용할 때 결과물이 대상자의 문화와 언어에 부합하는지 민감하게 고려해야 하며, ChatGPT가 제공하는 정보 활용 시 해당 정보의 출처 및 신뢰성에 대한 주의가 필요하다고 강조하였다. 생성형 AI의 효율성과 신뢰성을 보장하기 위해서는 대규모의 데이터가 확보되어야 하며 고품질의 데이터 레이블을 통해 알고리즘의 정확성과 견고성을 향상시켜야 한다(Hughes, Zhu, & Bednarz, 2021). 더불어, AI 모델 훈련에 사용되는 데이터에는 민감한 정보가 포함될 수 있으므로, 데이터 보안을 강화하여 개인 정보가 무단으로 유출되거나 오용되지 않도록 하는 것이 중요하다(Yu & Guo, 2023). 본 연구의 결과는 대형언어모델이 아동의 어휘 증진을 위한 증재 도구로서 유용할 수 있다는 가능성을 보여주었다. 그러나 생성형 AI가 언어병리학 분야에서 널리 활용되기 위해서는 향후 연구에서 정보의 신뢰성뿐만 아니라 개인 정보와 지적재산권 보호와 같은 안정성 측면도 신중히 고려되어야 할 것이다.

REFERENCES

- Anderson, N. J., Graham, S. A., Prime, H., Jenkins, J. M., & Madigan, S. (2021). Linking quality and quantity of parental linguistic input to child language skills: a meta-analysis. *Child Development*, 92(2), 484-501.
- Amirian, S. M. R., & Momeni, S. (2012). Definition-based versus contextualized vocabulary learning. *Theory & Practice in Language Studies*, 2(11), 2302-2307.
- Attig, M., & Weinert, S. (2020). What impacts early language skills? Effects of social disparities and different process characteristics of the home learning environment in the first 2 years. *Frontiers in Psychology*, 11, 557751.
- Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life: robust vocabulary instruction*. The Guilford Press.
- Cao, X., Lindley, S. E., Helmes, J., & Sellen, A. (2010). Telling the whole story: anticipation, inspiration and reputation in a field deployment of TellTable. *Proceedings of the 2010 ACM conference on Computer Supported Cooperative Work*, 251-260.
- Chen, C. Y. (2003). *Guji Guji*. Seoul: Yerin dang.
- Coyne, M., McCoach, D. B., & Kapp, S. (2007). Vocabulary intervention for kindergarten students: comparing extended instruction to embedded instruction and incidental exposure. *Learning Disability Quarterly*, 30(2), 74-88.
- DeepL. (2023). *DeepL API*. Retrieved from <https://www.deepl.com/en/docs-api>.
- Dickinson, D. K., Collins, M. F., Nesbitt, K., Toub, T. S., Hassinger-Das, B., Hadley, E. B., Hirsh-Pasek, K., & Golinkoff, R. M. (2019). Effects of teacher-delivered book reading and play on vocabulary learning and self-regulation among low-income preschool children. *Journal of Cognition and Development*, 20(2), 136-164.
- Du, Y., & Juefei-Xu, F. (2023). Generative AI for therapy? Opportunities and barriers for ChatGPT in speech-language therapy. *Proceedings of the International Conference on Learning Representations*, 1-6.
- Duncan, R. J., Anderson, K. L., King, Y. A., Finders, J. K., Schmitt, S. A., & Purpura, D. J. (2023). Predictors of preschool language environments and their relations to children's vocabulary. *Infant & Child Development*, 32(1), e2381.
- Fernald, A., & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Development*, 83(1), 203-222.
- Follmer, S., Ballagas, R., Raffle, H., Spasojevic, M., & Ishii, H. (2012). People in books: using a FlashCam to become part of an interactive book for connected reading. *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, 685-694.
- Ganek, H., Smyth, R., Nixon, S., & Eriks-Brophy, A. (2018). Using the Language ENvironment Analysis (LENA) system to investigate cultural differences in conversational turn count. *Journal of Speech, Language, & Hearing Research*, 61(9), 2246-2258.
- Gilkerson, J., Richards, J. A., Warren, S. F., Montgomery, J. K., Greenwood, C. R., Kimbrough Oller, D., Hansen, J. H. L., & Paul, T. D. (2017). Mapping the early language environment using all-day recordings and automated analysis. *American Journal of Speech-Language Pathology*, 26(2), 248-265.
- Graves, A., Fernández, S., Gomez, F., & Schmidhuber, J. (2006). Connectionist temporal classification: labelling unsegmented sequence data with re-

- current neural networks. *Proceedings of the 23rd International Conference on Machine Learning*, 369-376.
- Hargrave, A. C., & Sénéchal, M. (2000). A book reading intervention with preschool children who have limited vocabularies: the benefits of regular reading and dialogic reading. *Early Childhood Research Quarterly*, 15(1), 75-90.
- Hughes, R. T., Zhu, L., & Bednarz, T. (2021). Generative adversarial networks- enabled human-artificial intelligence collaborative applications for creative and design industries: a systematic review of current approaches and trends. *Frontiers in Artificial Intelligence*, 4, 604234.
- Hwang, I., Yoo, C., Hwang, C., Yim, D., Lee, Y., Min, C., Kim, J., & Song, J. (2014). TalkBetter: family-driven mobile intervention care for children with language delay. *Proceedings of the 17th ACM conference on Computer Supported Cooperative Work & Social Computing*, 1283-1296.
- Jenkins, J., Matlock, B., & Slocum, T. (1989). Two approaches to vocabulary instruction: the teaching of individual word meanings and practice in deriving word meaning from context. *Reading Research Quarterly*, 24(2), 215-235.
- Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., & Hüllermeier, E. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning & Individual Differences*, 103, 102274.
- Kelley, E. S. (2017). Measuring explicit word learning of preschool children: a development study. *American Journal of Speech-Language Pathology*, 26(3), 961-971.
- Kim, S. Y., Chae, M., & Yim, D. (2020). The effect of word exposure intensity, vocabulary skills, and working memory on preschool children's expressive word learning during book reading intervention. *Journal of Speech-Language & Hearing Disorders*, 33(3), 89-113.
- Kim, Y. (2020). *Assessment and treatment of language disorders in children*. Seoul: Hakjisa.
- Kim, Y. T., Hong, G. H., Kim, K. H., Jang, H. S., & Lee, J. Y. (2009). *Receptive & expressive vocabulary test (REVT)*. Seoul: Seoul Community Rehabilitation Center.
- Lee, J., Yoon, S., Lee, K., Jeong, E., Cho, J. E., Park, W., Yim, D., & Hwang, I. (2024). Open sesame? Open salami! Personalizing vocabulary assessment-intervention for children via pervasive profiling and bespoke storybook generation. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1-32.
- Leung, C. B., & Pikulski, J. J. (1990). Incidental learning of word meanings by kindergarten and first-grade children through repeated read aloud events. *National Reading Conference Yearbook*, 39, 231-239.
- Marulis, L. M., & Neuman, S. B. (2013). How vocabulary interventions affect young children at risk: a meta-analytic review. *Journal of Research on Educational Effectiveness*, 6(3), 223-262.
- McGregor, K. K., Friedman, R. M., Reilly, R. M., & Newman, R. M. (2002). Semantic representation and naming in young children. *Journal of Speech, Language, & Hearing Research*, 45(2), 332-346.
- Mol, S. E., & Bus, A. G. (2011). To read or not to read: a meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin*, 137(2), 267-296.
- Mol, S. E., Bus, A. G., & de Jong, M. T. (2009). Interactive book reading in early education: a tool to stimulate print knowledge as well as oral language. *Review of Educational Research*, 79(2), 979-1007.
- Moon, S. B. (2014). *Kaufman assessment battery for Korean children second edition*. Seoul: Hakjisa.
- Moon, S. B. (2019). *Korean Kaufman brief intelligence test-II (KBIT-II)*. Seoul: Hakjisa.
- Naigles, L. R., & Hoff-Ginsberg, E. (1998). Does frequency count? Parental input and the acquisition of vocabulary. *Journal of Child Language* 25, 95-120.
- National Reading Panel. (2000). *Teaching children to read: an evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (National Institute of Health Pub. No. 00-4769). Washington, DC: National Institute of Child Health and Human Development.
- Noble, C., Sala, G., Peter, M., Lingwood, J., Rowland, C., Gobet, F., & Pine, J. (2019). The impact of shared book reading on children's language skills: a meta-analysis. *Educational Research Review*, 28, 100290.
- OpenAI. (2022). *ChatGPT*. Retrieved from <https://www.openai.com/chatgpt>.
- OpenAI. (2023). *GPT-4 technical report (Report No. 2303.08774)*. Retrieved from <https://arxiv.org/abs/2303.08774>.
- Paradis, J. (2011). Individual differences in child English second language acquisition: comparing child-internal and child-external factors. *Linguistic Approaches to Bilingualism*, 1(3), 213-237.
- Park, S., & Yim, D. (2019). The effect of word exposure intensity during storybook reading on the vocabulary learning of children with specific language impairment. *Journal of Speech-Language & Hearing Disorders*, 28(4), 123-136.
- Patterson, J. (2002). Relationships of expressive vocabulary to frequency of reading and television experience among bilingual toddlers. *Applied Psy-*

- cholingistics*, 23(4), 493-508.
- Penno, J. F., Wilkinson, I. A., & Moore, D. W. (2002). Vocabulary acquisition from teacher explanation and repeated listening to stories: do they overcome the Matthew effect? *Journal of Educational Psychology*, 94(1), 23-33.
- Rice, M. L., Oetting, J. B., Marquis, J., Bode, J., & Pae, S. (1994). Frequency of input effects on word comprehension of children with specific language impairment. *Journal of Speech, Language, & Hearing Research*, 37(1), 106-122.
- Saini, N. (2023). Research paper on artificial intelligence & its applications. *International Journal for Research Trends & Innovation*, 8(4), 356-360.
- Shakeri, H., Neustaedter, C., & DiPaola, S. (2021). Saga: collaborative storytelling with gpt-3. *Proceedings of the 2021 Conference on Computer Supported Cooperative Work and Social Computing*, 163-166.
- Song, S., Kim, S., Kim, J., Park, W., & Yim, D. (2016). TalkLIME: mobile system intervention to improve parent-child interaction for children with language delay. *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 304-315.
- Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: a model-based meta-analysis. *Review of Educational Research*, 56(1), 72-110.
- Stahl, S. A., & Nagy, W. E. (2006). *Teaching word meanings*. Lawrence Erlbaum; Mahwah, NJ.
- Steele, S. C., & Mills, M. T. (2011). Vocabulary intervention for school-age children with language impairment: a review of evidence and good practice. *Child Language Teaching & Therapy*, 27(3), 354-370.
- Storkel, H. L., Voelmle, K., Fierro, V., Flake, K., Fleming, K. K., & Romine, R. S. (2017). Interactive book reading to accelerate word learning by kindergarten children with specific language impairment: identifying an adequate intensity and variation in treatment response. *Language, Speech, & Hearing Services in Schools*, 48(1), 16-30.
- Tomasello, M., & Haberl, K. (2003). Understanding attention: 12- and 18-month-olds know what is new for other persons. *Developmental Psychology*, 39(5), 906-912.
- Towson, J. A., Akemoglu, Y., Watkins, L., & Zeng, S. (2021). Shared interactive book reading interventions for young children with disabilities: a systematic review. *American Journal of Speech-Language Pathology*, 30(6), 2700-2715.
- Vallotton, C., & Ayoub, C. (2011). Use your words: the role of language in the development of toddlers' self-regulation. *Early Childhood Research Quarterly*, 26(2), 169-181.
- Weil, L. W., & Middleton, L. (2010). Use of the LENA tool to evaluate the effectiveness of a parent intervention program. *Perspectives on Language Learning & Education*, 17(3), 108-111.
- Weisleder, A., & Fernald, A. (2013). Talking to children matters: early language experience strengthens processing and builds vocabulary. *Psychological Science*, 24(11), 2143-2152.
- Whitehurst, G. J., Falco, F. L., Lonigan, C. J., Fischel, J. E., DeBaryshe, B. D., Valdez-Menchaca, M. C., & Caulfield, M. (1988). Accelerating language development through picture book reading. *Developmental Psychology*, 24(4), 552-559.
- Yang, Y., & Yim, D. (2018). The role of executive function for vocabulary acquisition and word learning in preschool-age children with and without vocabulary delay. *Communication Sciences and Disorders*, 23(1), 43-59.
- Yu, H., & Guo, Y. (2023). Generative artificial intelligence empowers educational reform: current status, issues, and prospects. *Frontiers in Education*, 8, 1183162.
- Zarei, N., Chu, S. L., Quek, F., Rao, N. J., & Brown, S. A. (2020). Investigating the effects of self-avatars and story-relevant avatars on children's creative storytelling. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1-11.

Appendix 1. AI 동화 생성 단계별 세부 작업 내용

| 세부 작업 내용 | |
|----------|--|
| 단계 1 | GPT-4에게 기존동화의 전체 텍스트를 세 문장으로 요약하여 '이야기 틀'을 생성하도록 지시(query)한다. 이 단계를 통해 기존동화의 큰 흐름을 제외한 모든 세부 내용을 제거한다. '이야기 틀' 생성은 기존동화의 영문 원서 또는 영문 번역본을 사용하였다. |
| 단계 2 | '단계 1'에서 생성된 '이야기 틀'을 활용하여 특정 목표어휘를 포함하는 이야기를 생성하도록 지시(query)한다. 하나의 동화는 16개의 목표어휘 중 4개의 어휘를 포함하며, 각각의 동화에 포함되는 어휘는 기존동화와 동일하도록 하였다. |
| 단계 3 | '단계 2'에서 생성된 영문 초안을 수정하거나 보완할 필요가 있다고 판단될 경우, 시스템의 툴킷(tool kit)을 사용하여 갈등을 추가하거나 대화를 추가한다. 목표어휘가 모두 적절히 반영되었는지 검토 후 어휘가 부적절한 형태로 포함되거나 누락된 경우 목표어휘를 재반영하여 수정한다. |
| 단계 4 | 생성한 영문 동화를 DeepL API (DeepL, 2023)을 사용하여 한국어로 번역한다. |
| 단계 5 | 국문 번역본을 검토 및 수정하여 최종본을 제작한다. 이 단계에서는 문법이 올바른지, 내용이 만 4-5세 연령에 적합한지, 문맥과 어조가 적절하고 자연스러운지에 대해 25년 이상 경력의 중등 국어교육 전문가가 검증한다. |
| 단계 6 | 기존동화와 동일한 산화를 AI-생성동화의 이야기 흐름에 맞도록 재배열한다. 삽화는 Google Presentation Slides를 사용하여 아동에게 제시할 수 있는 형태로 작업한다. |

Appendix 2. 국문 번역본 검토 및 수정 기준

| 구분 | 기준 |
|-----------|---|
| 부적합한 어휘 | <ul style="list-style-type: none"> 어휘의 난이도가 아동의 연령에 적절하지 않은 경우, 유의어 사전, 배경지식을 활용하여 적절한 유의어로 대체한다. 어색한 번역투의 표현은 유의어사전, 배경지식을 활용하여 자연스러운 유의어로 대체한다. 일상적이지 않은 외래어는 문맥을 해치지 않는 적절한 단어로 대체한다. |
| 부적합한 문장 | <ul style="list-style-type: none"> 문맥, 어순, 어휘 조합 등이 부적절하여 해석이 난해한 문장은 직접 대체 문장을 작성하거나 삭제한다. |
| 문장 길이 | <ul style="list-style-type: none"> 이어진 문장, 안긴문장과 같이 길이가 긴 문장은 분할하여 문장 길이를 조절한다. |
| 내용의 인과성 | <ul style="list-style-type: none"> 전체적인 이야기 흐름과 어울리지 않는 문장은 삭제한다. 필요한 문장이지만 내용의 인과성이 낮은 경우, 세부 내용을 추가하거나 적절한 표현으로 대체한다. |
| 부적합한 종결어미 | <ul style="list-style-type: none"> 종결 어미의 문체가 혼용되거나 부자연스러운 경우 평서문 비격식 높임체를 사용하여 자연스럽게 수정한다. |
| 대명사 | <ul style="list-style-type: none"> 어색한 대명사의 사용은 대상의 이름으로 대체한다. |

Appendix 3. 기존동화 및 AI-생성동화 집단 중재 회기 일정

| 회기 | 소요시간 | 동화 | 중재어휘 | 노출빈도 |
|-----|------|--------------|--------------------------------|------|
| 1회기 | 20분 | 동화 A 동화 B | 어휘 a, b, c, d 어휘 e, f, g, h | 9회 |
| 2회기 | 20분 | 동화 C 동화 D | 어휘 i, j, k, l 어휘 m, n, o, p | 9회 |
| 3회기 | 20분 | 동화 A 동화 B | 어휘 a, b, c, d 어휘 e, f, g, h | 9회 |
| 4회기 | 20분 | 동화 C 동화 D | 어휘 i, j, k, l 어휘 m, n, o, p | 9회 |
| 5회기 | 20분 | 동화 A 동화 B | 어휘 a, b, c, d 어휘 e, f, g, h | 9회 |
| 6회기 | 20분 | 동화 C 동화 D | 어휘 i, j, k, l 어휘 m, n, o, p | 9회 |
| 7회기 | 20분 | 동화 A 동화 B | 어휘 a, b, c, d 어휘 e, f, g, h | 9회 |
| 8회기 | 20분 | 동화 C 동화 D | 어휘 i, j, k, l 어휘 m, n, o, p | 9회 |

Appendix 4. 중재스크립트 예시

책 제목: 악어오리 구지구지 (천즈위엔, 2003)

1. 책 읽기 전: 문맥 + 유의어 + 정의

“우리 책 읽기 전에 책에 나올 단어들을 먼저 알아볼까?”

- 문맥 말하기 (e.g., “농부아저씨가 비가 안 와서 금이 간 땅을 보고 슬퍼하고 있어.”) *그림 제시
- 유의어 말하기 (e.g., “금은 틈 같은 거야”) *그림 제시
- 정의 말하기 (e.g., “금은 물건이나 벽이 살짝 갈라져 생긴 가느다란 선이야”)

| 번호 | 단어목록 | 유의어 말하기 | 정의 말하기 | 문맥 말하기 |
|----|------|--------------------|---|------------------------------------|
| 1 | 금 | 금은 틈 같은 거야 | 금은 벽이 살짝 갈라져 생긴 가느다란 선이야. (그림 가리키며) 이걸 금이라고 해 | 농부아저씨가 비가 안 와서 금이 간 땅을 보고 슬퍼하고 있어. |
| 2 | 괴상하다 | 괴상하다는 이상하다와 같은 말이야 | 괴상하다는 보통과 다르게 특이하고 징그럽단 뜻이야. (그림 가리키며) 이런 걸 괴상하다고 해 | 동물원에 갔다가 다리만 줄무늬인 괴상한 당나귀를 봤어. |
| 3 | 납작하다 | 납작하다는 평평하다와 같은 말이야 | 납작하다는 두께가 얇고 판판하고 넓다는 뜻이야. (그림 가리키며) 이런 걸 납작하다고 해 | 아빠가 호떡을 꼭 눌렀더니 납작해졌어. |
| 4 | 살갓 | 살갓은 피부 같은 거야 | 살갓은 우리 몸에 있는 살가죽이야. (그림 가리키며) 이런 걸 살갓이라고 해 | 뜨거운 햇빛 때문에 아이의 살갓이 빨갛게 변해버렸어. |
| 5 | 움켜잡다 | 움켜잡다는 쥐다와 같은 말이야 | 움켜잡다는 손가락을 접어서 힘 있게 꼭 잡는 거야. (그림 가리키며) 이렇게 하는 게 움켜잡는 거야 | 엄마가 손가락을 내밀었더니 아기가 꼭 움켜잡았어. |
| 6 | 갈다 | 갈다는 다듬다와 같은 말이야 | 갈다는 무언가를 날카롭게 만들려고 다른 물건에 대고 문지른다는 뜻이야. (그림 가리키며) 이렇게 하는 걸 갈다라고 해 | 종이가 잘 안 잘려서 가위를 갈았어. |
| 7 | 장하다 | 장하다는 대단하다와 같은 말이야 | 장하다는 정말 훌륭해서 칭찬할만하다는 뜻이야. 이럴 때 (그림 가리키기) 부모님이 장하다고 말해 | 축구 경기에서 우승하더니 우리나라 선수들 정말 장하다. |

2. 책 읽기: 본문 + 유의어 + 정의

“자 이제 책을 읽어보자. 우리 앞에 했던 단어들이 나오나 한번 볼까?”

- 본문 읽기 (e.g., “어느 날, 동지 속의 알이 하나씩 금이 가기 시작했어요.”)
- 유의어 말하기 (e.g., “금은 틈 같은 거야”) *그림 제시
- 정의 말하기 (e.g., “금은 물건이나 벽이 살짝 갈라져 생긴 가느다란 선이야”)

| 번호 | 단어목록 | 본문읽기 | 유의어 말하기 |
|----|------|--------------------------------------|--------------------|
| 1 | 금 | 어느 날, 동지 속의 알이 하나씩 금이 가기 시작했어요 | 금은 틈 같은 거야 |
| 2 | 괴상하다 | 마지막 네 번째 오리는 온몸이 온통 푸르스름한 게 아주 괴상했어요 | 괴상하다는 이상하다와 같은 말이야 |
| 3 | 납작하다 | 넌 털도 없고, 납작한 부리랑 오동통한 다리도 없잖아 | 납작하다는 평평하다와 같은 말이야 |
| 4 | 살갓 | 네 살갓은 푸르스름하고 발톱과 이빨은 뾰족하잖아 | 살갓은 피부 같은 거야 |
| 5 | 움켜잡다 | 날카로운 발톱은 오리를 꼭 움켜잡을 수 있어 | 움켜잡다는 쥐다와 같은 말이야 |
| 6 | 갈다 | 날카로운 이빨을 열심히 갈고 있었답니다 | 갈다는 다듬다와 같은 말이야 |
| 7 | 장하다 | 무서운 악어를 물리친 구지구지는 정말 장해요 | 장하다는 대단하다와 같은 말이야 |

3. 책 읽기 후: 문맥 + 유의어 + 정의

“우리 책에 나왔던 단어들이 다른 데 어떻게 쓰이는지 볼까?”

- 문맥 말하기 (e.g., “농부아저씨가 비가 안 와서 금이 간 땅을 보고 슬퍼하고 있어.”) *그림 제시
- 유의어 말하기 (e.g., “금은 틈 같은 거야”) *그림 제시
- 정의 말하기 (e.g., “금은 물건이나 벽이 살짝 갈라져 생긴 가느다란 선이야”)

| 번호 | 단어목록 | 문맥 말하기 | 유의어 말하기 | 정의 말하기 |
|----|------|------------------------------------|--------------------|---|
| 1 | 금 | 농부아저씨가 비가 안 와서 금이 간 땅을 보고 슬퍼하고 있어. | 금은 틈 같은 거야 | 금은 벽이 살짝 갈라져 생긴 가느다란 선이야. |
| 2 | 괴상하다 | 동물원에 갔다가 다리만 줄무늬인 괴상한 당나귀를 봤어. | 괴상하다는 이상하다와 같은 말이야 | 괴상하다는 보통과 다르게 특이하고 징그럽단 뜻이야. |
| 3 | 납작하다 | 아빠가 호떡을 꼭 눌렀더니 납작해졌어. | 납작하다는 평평하다와 같은 말이야 | 납작하다는 두께가 얇고 판판하고 넓다는 뜻이야. |
| 4 | 살갓 | 뜨거운 햇빛 때문에 아이의 살갓이 빨갛게 변해버렸어. | 살갓은 피부 같은 거야 | 살갓은 우리 몸에 있는 살가죽이야. |
| 5 | 움켜잡다 | 엄마가 손가락을 내밀었더니 아기가 꼭 움켜잡았어. | 움켜잡다는 쥐다와 같은 말이야 | 움켜잡다는 손가락을 접어서 힘 있게 꼭 잡는 거야. |
| 6 | 갈다 | 종이가 잘 안 잘려서 가위를 갈았어. | 갈다는 다듬다와 같은 말이야 | 갈다는 무언가를 날카롭게 만들려고 다른 물건에 대고 문지른다는 뜻이야. |
| 7 | 장하다 | 축구 경기에서 우승하더니 우리나라 선수들 정말 장하다. | 장하다는 대단하다와 같은 말이야 | 장하다는 정말 훌륭해서 칭찬할 만하다는 뜻이야. |

Appendix 5. 기존 동화와 AI-생성동화 예시

1) 기존동화 예시

제목: 악어오리 구지구지

목표어휘: 금, 괴상하다, 움켜잡다, 갈다

동그런 알 하나가 굴러왔어요. 어느 날, 둥지 속의 알이 하나씩 **금**이 가기 시작했어요. 맨 먼저 껍질을 깨고 푸른 점이 난 아기오리가 나왔어요. 두 번째로 나온 아기오리는 줄무늬가 있었어요. 세 번째 아기오리는 노란 색깔이었지요. 마지막 네 번째 오리는 온몸이 온통 푸르스름한 게 아주 **괴상했어요**. 이 오리는 자꾸자꾸 '구옥구옥' 하고 중얼거리는 거예요. 그래서 엄마오리는 구지구지라고 이름을 지었습니다. 그러던 어느 날이었어요. 호숫가에 구지구지와 똑같이 생긴 동물들이 나타났어요. 바로 악어였어요. "저 멍청한 악어 좀 봐! 저 녀석은 맛난 오리들처럼 겹잡아?" "난 악어가 아녜요! 오리라고요, 오리!" 구지구지가 소리쳤어요. "뭐라고? 네 꼴 좀 봐? 넌 털도 없고, 납작한 부리랑 오동통한 다리도 없잖아." "네 살갗은 푸르스름하고 발톱과 이빨은 뾰족하잖아!" "넌 우리랑 똑같은 악어야!" 날카로운 발톱은 오리를 꼭 **움켜잡**을 수 있어. "악어들이 귀찮말을 했어요." "내일 오리들한테 다이빙을 하자고 꾀어서 이 다리 위로 데리고 와. 우리는 저 아래에서 입을 벌리고 있을게. 알겠지?" 악어들은 이 말을 남기고는 스르르 사라졌어요. "어찌지? 못된 악어들이 우리 가족을 잡아먹으려고 해." 구지구지는 바위에 앉아 한참을 생각했어요. "아하! 바로 그거야!" 뭔가를 생각해낸 구지구지는 가벼운 마음으로 집으로 돌아갔어요. 그 날 밤, 세 마리 악어는 맛있는 오리들을 떠올리며, 날카로운 이빨을 열심히 **갈**고 있었답니다. 다음 날, 구지구지는 정말로 오리들을 데리고 다리 위로 왔어요. 악어들은 다리 밑에서 입을 찹 벌리고 있었어요. 어서 오리가 푹푹 떨어지기만을 기다렸지요. 아, 그런데 밑으로 떨어진 건 오리들이 아니었어요. 커다랗고 딱딱한 돌덩어리였어요. 팍, 와삭……. 악어들의 이빨은 몽땅 부러져 버렸지요. "으악! 악어 살려!" 세 마리 악어는 엉엉 울부짖으며 도망갔어요. 구지구지가 오리들을 무사히 구해냈어요. "만세! 구지구지 만세!" 무서운 악어를 물리친 구지구지는 정말 장해요. 악어오리 구지구지는 오리 가족들과 언제나 행복했어요.

2) AI-생성동화 예시

제목: 악어오리 구지구지

목표어휘: 금, 괴상하다, 움켜잡다, 갈다

옛날 어느 연못에 어미 오리와 네 마리의 아기 오리가 살고 있었어요. 그런데 아기 오리들의 모습이 조금 달랐어요. 아기 오리 중 세 마리는 자그마하고 푹신푹신한 평범한 모습이었지만, 넷째인 구지구지는 조금 **괴상했어요**. 구지구지는 몸집이 더 크고 피부가 거칠거칠하고, 부리 대신에 긴 주둥이가 있었지요. 그래서 구지구지는 오리가 아니라 악어처럼 보였답니다. 그러던 어느 날, 오리 가족은 연못가 땅바닥에 길게 난 **금**을 발견했어요. 구지구지는 그게 뭔지 궁금해서 알아보기로 했어요. 구지구지가 숲 속으로 들어가자 무서운 악어 세 마리가 숨어 있었어요. 그건 바로 악어들이 지나간 꼬리 자국이었던 거예요. 악어들은 배가 몹시 고파 보였고, 구지구지를 보자 이렇게 말했어요. "넌 오리가 아니라 우리같은 악어야. 우리를 도와주면 진짜 악어가 되는 법을 가르쳐 줄게." 악어들은 오리 가족이 연못에서 수영을 하고 있을 때, 아기 오리를 한 마리씩 **움켜잡**을 계획이라고 말했어요. 구지구지는 너무 놀랍고 무서웠어요. 하지만 사랑하는 오리 가족을 지키기 위해서 정신을 차리고 빨리 방법을 찾아야 했어요. 구지구지는 아기 오리만 한 돌멩이들을 찾아서 한쪽 끝을 뾰족하게 **갈**기 시작했어요. 그런 다음 기다리고 있는 악어들에게 웃으면서 다가갔어요. "맛있는 아기 오리들을 잡아 왔어요." 구지구지는 한쪽 끝이 뾰족한 돌멩이를 내밀었고, 악어들은 그게 오리인 줄 알고 입을 찹 벌렸어요. 그러자 구지구지는 재빨리 악어들의 입 속으로 돌멩이를 던졌어요. 악어들은 돌멩이에 맞아 이빨이 부러졌고 너무 아파 도망가버리고 말았어요. 그 후로 연못은 다시 평화로워졌답니다. 오리 가족들은 가족을 지킨 구지구지를 자랑스러워했어요. 모습은 조금 달랐지만 구지구지는 더욱 사랑받는 오리가 되었지요.

국문초록

AI-생성동화를 활용한 개별 맞춤형 책읽기 활동이 미취학 아동의 어휘 학습에 미치는 영향

조재은 · 임동선

이화여자대학교 대학원 언어병리학과

배경 및 목적: 본 연구는 AI-생성동화와 기존동화를 활용한 책읽기 중재를 통해 아동의 어휘 습득 효과를 확인하고 어휘 중재 도구로서의 AI-생성동화의 효용성을 확인하고자 하였다. **방법:** 총 15명의 3-6세 일반 아동이 연구에 참여하였으며, 모든 참가자는 기존동화 집단과 AI-생성동화 집단으로 나뉘었다. 참가자들은 책읽기 전 활동, 책읽기 본 활동 및 책읽기 후 활동을 포함하는 구조화된 책 읽기 세션에 3주에 걸쳐 총 8회 참여했다. 어휘력 증진 여부를 확인하기 위해 책읽기 활동 전-후에 16개의 목표 어휘에 대한 수용 및 표현 어휘력을 측정했다. 각 집단 내 어휘 정확도를 비교하기 위해 Wilcoxon 부호 순위 검정을 실시하였고, 집단 간 어휘 정확도 차이를 비교하기 위해 Mann-Whitney U-검정을 실시하였다. **결과:** 두 그룹 모두 후 중재 후의 수용 및 표현 어휘 정확도가 유의하게 향상되었으며, 두 그룹 간에 유의한 차이가 나타나지 않았다. **논의 및 결론:** AI-생성동화를 활용한 책읽기 어휘 중재는 기존동화를 활용하는 전통적 방식에 준하는 수준으로 아동의 어휘력 증진에 긍정적 영향을 미치는 것으로 확인되었다. 본 연구는 어휘 학습 중재 도구로서 AI-생성동화의 유용성을 입증하였으며, 아동의 언어 발달과 인공지능 기술의 융합에 대한 새로운 가능성을 제시한다는 점에서 중요한 의의를 갖는다.

핵심어: AI-생성동화, 책읽기, 어휘 중재

본 연구는 2024년도 대한민국 과학기술정보통신부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2022R1A2C1005268).

참고문헌

- 김신영, 채미선, 임동선 (2020). 책읽기 중재 시 어휘 노출 빈도 및 학령전기 아동들의 어휘력, 작업기억 능력이 표현어휘 학습에 미치는 영향. *언어치료 연구*, 33(3), 89-113.
- 김영태 (2020). *아동언어장애의 진단 및 치료*. 서울: 학지사.
- 김영태, 홍경훈, 김경희, 장혜성, 이주연 (2009). *수용·표현어휘력검사(Receptive & Expressive Vocabulary Test, REVT)*. 서울: 서울장애인종합복지관.
- 문수백 (2014). *한국판 카우프만 아동 지능검사*. 서울: 학지사.
- 문수백 (2019). *한국판 카우프만 간편지능검사 2*. 서울: 학지사.
- 박수연, 임동선 (2019). 어휘 노출 강도에 따른 책읽기를 활용한 어휘 중재가 단순언어장애 아동의 어휘 학습에 미치는 효과. *언어치료연구*, 28(4), 123-136.
- 양윤희, 임동선 (2018). 학령전기 어휘발달 지연 및 또래 아동의 어휘습득을 위한 집행기능의 역할. *Communication Sciences and Disorders*, 23(1), 43-59.
- 천즈위엔 (2003). *악어오리 구지구지*. 서울: 예림당.

ORCID

조재은(제1저자, 대학원생 <https://orcid.org/0000-0002-6543-6812>); 임동선(교신저자, 교수 <https://orcid.org/0000-0001-8254-9504>)