



## 초기 문해력 측정 검사 도구 설계

심영택<sup>1</sup>, 한정혜<sup>2</sup>

### 《 요 약 》

초기 문해력은 발생적 문해력이 완성되는 시점으로 문자 인식하기, 따라 읽기, 유창하게 읽기 등의 독서 활동을 하는 시기이다. 그런데 미숙한 독자는 읽기 기술과 단어 지식이 부족하기에 읽기 발달 격차로 읽기 매튜 효과가 발생하게 된다. 본고는 이러한 독자 발달과 읽기 발달의 특성을 고려하여, 짧은 시간 안에 미숙한 독자를 적절하게 선별할 수 있는 초기 문해력 진단 도구를 설계하여, 현장 적합성과 활용 빈도를 높이고자 하였다. 먼저 기존 연구를 통해 초기 문해력 측정 도구가 갖추어야 하는 이론적인 틀과 맥락을 고찰하고, 인쇄물 인식 능력, 음운 인식 능력, 단어 읽기, 어휘 능력, 읽기 유창성 등 다섯 영역으로 구분한 후, 현장 교사의 도움으로 초등학교 저학년 학생들 28명에 문항별 검증을 거쳐 31개 문항으로 구성된 1차 도구를 설계하였다. 그리고 프로토타입 측정 도구를 통해 초등학교 1학년 350명의 데이터를 수집하고 분석하였다. 주목할 점은 인쇄물 인식 영역과 어휘 능력 영역에서 부진 학생들의 정답률이 크게 떨어지는 것을 확인할 수 있었으며, 음운 인식과 단어 읽기의 경우에는 다소 비슷한 정답률을 나타내었다. 평균 소요 시간의 경우, 부진 학생들이 문항에 대한 이해를 하는 과정에서 상당히 많은 시간을 소요하는 것으로 나타났다. 데이터 분석 결과를 토대로 문해력 전문가와 함께 영역별 난이도 조절을 위한 문항 수정을 거침으로써, 초기 문해력 진단을 위한 35개 문항 설계를 완료하였다.

**주제어** : 초기 문해력, 검사 도구 설계, 인쇄물 인식, 음운 인식, 단어 읽기, 어휘 능력, 읽기 유창성

\* 이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2023R1A2C1006289).

1. 청주교육대학교, 교수, shimyt@cje.ac.kr(제1저자)  
2. 청주교육대학교, 교수, hanjh@cje.ac.kr(교신저자)

## I. 연구 필요성 및 목적

초기 문해력은 발생적 문해력이 완성되는 시점으로 이는 천경록(2020)의 독자 발달과 읽기 발달의 단계로 볼 때, 유아 독자와 어린이 독자 사이에 해당하며, ‘읽기 맹아기’와 ‘읽기 입문기’ 발달에 걸쳐 일어난다.<sup>1)</sup> 이 시기의 주요 독서 양상과 활동은 문자 인식하기(문자 지각, 해독 지각), 따라 읽기(그림책, 글자책, 낱말 카드 등), 띄어 읽기(소리 내어 읽기, 해독 완성, 기초 기능), 유창하게 읽기(낭독, 묵독을 시작) 등이다.

M. Wolf(2007)는 초기 문해력 단계의 독자를 ‘예비 독자’(emerging prereader)로 명명하고, 말놀이를 통한 ‘인지적 발달’, 동화책 읽어주기를 통한 ‘정서적 발달’, 멈춤 표시, 시리얼 상자, 아이 자신의 이름, 형제자매나 친구들의 이름 등 아이가 살고 있는 환경 안에서 흔히 볼 수 있는 ‘환경적 문자 읽기’, 문자 이름을 이용하여 단어 철자를 쓰는 ‘창안된 철자 쓰기’(invented spelling), 동요를 이용한 문장 속 단어와 음절, 음소를 파악하는 ‘음소 인지 과정’ 등을 초점화하면서 자소와 음소의 변별 능력을 위한 교수법을 강조한 바 있다.

한편, 윤혜경(1997)의 연구에 의하면, 글을 읽는 행위란 시각적 상징과 청각적 상징 사이에 존재하는 대응 관계로 파악하고, 글자를 소리로 전환하는 읽기 행위를 ‘인지적’ 작업으로 보고 있다. 그리고 자소-음소의 대응에 의한 읽기를 ‘자소 읽기’로, 문자 단어-음성 단어 간 대응에 의한 읽기를 ‘단어 읽기’로, 글자-음절 간의 대응에 의한 읽기를 ‘글자 읽기’로 명명한 바 있다. 부연하자면 읽기란 물리적인 현상으로서 시각적인 형태(글자)와 청각적인 형태(소리)를 이미지나 상징으로 ‘표상’하는 과정이며, 인간은 이러한 표상들을 주고받으며 의사소통하고, 스스로 자기 자신을 인간 ‘답게’ 만들어 가는 과정으로 보고 있다.<sup>2)</sup>

그런데 읽기 발달 격차와 읽기 매튜 효과에 관한 Stanovich(1986) 연구에 의하면, 미숙한 독자(poor readers)는 인쇄물의 목적을 이해하지 못하며, 아동기 내내 활자 경험에 덜 노출되고, 그에 따라 인지 발달 자체에도 부정적인 영향을 받게 되는 경우가 빈번하다고 한다. 미숙한 독자는 단어 재인뿐만 아니라, 이해력 자체도 부족하기에 읽기 기술이 연습을 통해 향상되지 않는다 면, 다양한 어휘나 개념에 접근할 기회는 애당초 기대할 수 없다. 따라서 학년이 올라갈수록 능숙한 독자와 미숙한 독자 사이에 읽기 능력 격차가 점점 더 커지게 된다.

1) Scarborough(1998)와 McCardle(2001) 등의 연구에 의하면, 초등학교 입학 이전 유치원 시기에 인지-언어적 기술과 관련된 변인으로 잘 설계된 선별 검사를 사용할 경우 읽기 성취에 대한 80% 이상의 정확도를 가지고, 아동의 미래 읽기 성취 정도를 예상할 수 있다고 본다(박현정, 2010).

2) 이러한 시각적 상징 능력을 M. Wolf(2007)에서는 ‘패턴 불변성’(pattern invariance)이라는 개념으로 설명하고 있는데, 이 능력은 사람들에게 지각 패턴의 표상을 기억 속에 저장해 두었다가 새로운 것(문자 학습)을 배워야 하는 상황이 발생할 때 꺼내서 적용한다고 한다.

또한 Scarborough(2001)의 연구에 의하면, 형식적인 읽기 지도가 시작되는 초등학교 입학 시에 읽기 능력이 상위권에 속하는 아동은 학교생활을 하는 동안 읽기 능력이 뒤쳐질 확률은 5~10% 정도이나, 이 시점에 읽기 능력이 뒤떨어지는 아동이 지체된 상태로 계속 남을 확률은 65~75%나 된다고 한다(박현정, 2010:2). Juell(1988)도 읽기 기술이 부족한 아동은 또래를 따라 잡기 위해서 극심한 어려움을 겪게 되며, 1학년 때 읽기를 잘 못하는 아동이 4학년이 되어 평균 수준의 읽기 능력을 지닐 가능성은 13%에 불과하다고 주장하였다.

이러한 주장들을 한마디로 요약하자면 학령기 초기, 즉 초등 저학년 시기에 읽기 부진은 문해력 발달에 결정적인 영향을 끼친다는 사실이다. 그래서 미국의 경우, 2001년 이후 초기 문해력이 부족한 학생들을 조기에 선별·진단하고자 소리 분절, 단어 읽기, 읽기 유창성 등 6개 항목으로 구성된 DIBELS(Dynamic Indicators for Basic Early Literacy Skills)를 개발하였는데, 이 도구는 5분 정도의 소요 시간이 걸리는 일대일 대량 측정 방식이다. 한편 읽기 결함을 자동으로 선별하기 위해 컴퓨터 기반 자동 측정 도구 CARE(Carolina Automated Reading Evaluation)도 개발하였는데, 이는 글자 식별, 문장 식별, 읽기 유창성 등 16개의 다양한 하위 항목을 통해 읽기 능력을 평가하며 평균 30분 내외 소요된다. Hoskins et al. (2021)은 DIBELS와 CARE를 비교하여 정확성과 신뢰성에 유의미한 상관관계가 있음을, 그리고 학교에서 적용 가능성을 제시한 바 있다.

한편, 국내의 경우 대표적 문해력 진단 도구는 <표 1>과 같이 7개의 지필 검사 도구와 2개의 웹 기반 검사 도구들이 있다. 지필 검사 도구들은 모두 유료이며 몇몇 검사 도구는 고가의 가격으로 책정되어 있다. 웹 기반 진단 도구는 무료 접근이 가능하다는 장점이 있지만, 컴퓨터를 켜놓고 전문가가 옆에서 개인별로 직접 관찰하며 점수를 채점하는 방식으로 진행된다. 평가 시간이 길거나 평가 방법과 결과 처리가 다소 번잡한 도구도 있다. 예를 들어, 청주교대 문해력지원센터에서 사용하는 지필 검사의 경우, 교사 6명이 85명의 학생을 상대할 때, 교사 1인당 약 10분씩, 총 140분 동안 검사를 해야 하며, 그 결과를 별도의 기록지에 분석·정리하는 시간 역시 교사 1인당 약 104분이 추가 소요된다고 한다.<sup>3)</sup> 이뿐 아니라 학교 간 문해력 격차가 심할 경우 검사 시간이 더 길어지며, 원활한 검사가 이루어지려면 학교 내에 별도의 공간을 확보해야 하며, 그 공간으로 학생을 인솔하는 교사가 최소한 2인 이상이 필요하며, 이 모든 과정을 거친 뒤 검사 결과를 종합하여 관리자와 담임, 기초학력 전담 교사와 함께 지도 대상 아동을 선정하는 문제 등이 있다.

3) 2024년 청주교대 문해력지원센터에서 실제로 실시한 검사(이예지 파견교사)를 토대로 한 결과임.

〈표 1〉 국내 문해력 진단 도구

방식	검사명	연령(세)	특징
지필식	기초학습능력검사 NISE-BACT	4-13	읽기와 쓰기 사전·사후검사 실시
	한국어 읽기 검사 KOLAR	7-12	핵심 검사와 상세 검사로 이원화
	읽기 성취 및 읽기 인지처리능력 검사 RA-RCP	7-12	인지 처리능력 평가
	종합학습능력검사 CLT	3-15	난독증 및 난산증 조기 선별 진단
	언어 문해해결력 검사 TOPS	5-12	언어사용의 논리적 사고력 측정
	구문 의미 이해력 검사 KOSECT	3-9	구문 의미 이해 판별에 많이 적용
	한국어 단어, 읽기 이해력 진단검사 K-TOLLS	5-8	난독증 및 학습장애 판별시 활용
웹기반	웰리미	6-12	캐릭터 스토리텔링 실시
	한글또박또박	6-12	평가 후 보충 교재 연계

음성과 시선 등의 멀티모달 초기 문해력 진단 도구는 이러한 독자 발달과 읽기 발달의 특성뿐만 아니라 이를 편리하고 짧은 시간에 정확하게 평가할 수 있도록 설계해야 정확도와 현장에서의 활용 빈도를 높일 수 있을 것이다. 즉, 기존 지필식 방식에서는 평가자가 직접 판단하는 발음을 객관적으로 채집하여 평가할 수 있으며, 기존의 진단에서는 어려운 시선 처리, 웹 기반 방식의 한계도 극복할 방안을 모색해야 할 것이다.

본고에서는 이러한 점들을 염두에 두면서 읽기와 쓰기를 시작하는 일종의 ‘인지-언어적 두뇌 혁명’이 일어나고 있는 아이들의 초기 문해력 수준이 어떠한지, 검사 영역별 그 문해력의 상태가 어떠한지를 파악하고, 아동이 집중할 수 있는 시간 내에 교사가 다수의 학생을 동시에 진단·분석하는 도구를 설계하고 개발하고자 한다.

## II. 연구 방법

### 1. 연구 절차

본 연구는 현장 교사가 비교적 짧은 시간 안에 다수의 초등학교 저학년 학생들을 대상으로 초기 문해력을 측정하는 도구를 설계하고자 <표 2>와 같은 절차로 진행되었다.

<표 2> 연구 절차



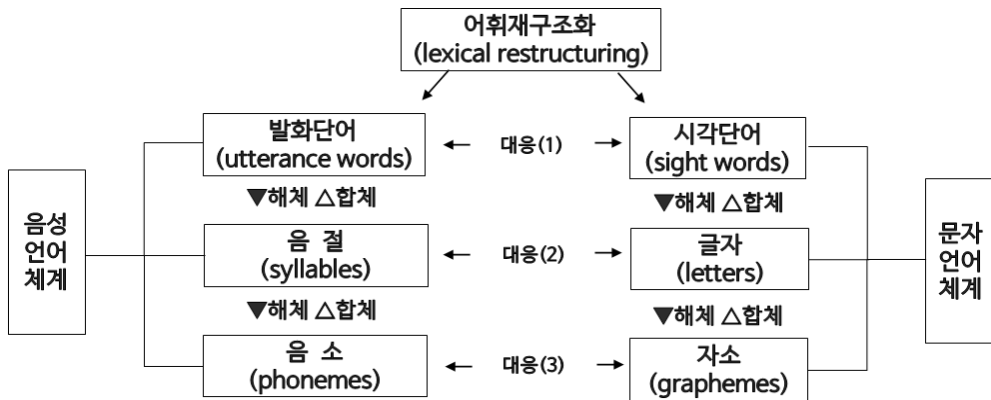
첫째, 국내외 문헌 및 기존 측정 도구 분석을 통하여 초기 문해력 측정 도구가 갖추어야 하는 이론적인 틀과 맥락을 설계하였다. 먼저 국어과 교육과정을 분석하고 그 특징 몇 가지를 이 설계에 반영하고자 하였다.<sup>4)</sup> 또한 파일럿 분석과 현장 적용 분석을 위한 데이터 채집에 대한 교내 생명연구윤리(IRB) 승인 심의용 계획서를 작성하였다. 둘째, 현재 한국의 교육 현실과 교사 1인이 검사해야 할 웹 기반의 상황 맥락을 고려하여 초기 문해력 측정 도구 모형을 설계하였다. 셋째, 모형의 5가지 세부 영역별(인쇄물 인식 능력 검사, 음운 인식 능력 검사, 단어 읽기 능력 검사, 어휘 능력 검사, 읽기 유창성 검사) 문항을 개발하고 태블릿용 프로토타입 측정 도구를 개발하였다. 연구 참여자는 연구책임자, 공동연구진 문해력 전문가, 초등 파견교사 연구원, 초등 저학년 현장 교사 2인이다. 현장 교사 2인 중 남교사(총 경력 19년)는 저학년 경력이 8년이며, 여교사

4) 2022 국어과 교육과정 중 초등학교 1~2학년 읽기 영역에서 강조하고 있는 ‘읽기의 기초’ 내용 요소는 세 가지인데, 글자와 단어 읽기, 문장 및 짧은 글 소리 내어 읽기, 알맞게 띄어 읽기이다.

(총 경력 17년)는 저학년 경력이 5년으로, 이들을 중심으로 포커스 그룹 인터뷰(Focus Group Interview)를 진행하여 문항에 대하여 난이도나 구성 등에 대한 교육적 경험과 의견을 반영하였다. 문항에 대한 아동의 반응이나 선호도는 두 교사의 반 학생들이 문항을 풀 때 보여준 반응과 교사의 관찰 의견을 반영함으로써 측정 도구의 타당성과 신뢰성을 확보하고자 하였다. 넷째, 개발된 프로토타입을 포커스 그룹에 속한 2인의 현장 교사의 초등학교 저학년 학생들 28명에 적용하여, 문항의 이해도와 난이도 분포 등을 고려하여 1차 도구 개발을 완료하였다. 다섯째, 개발된 1차 도구에 대하여 초등학교 저학년 350명의 데이터를 확보하여 분석함으로써 현장 교사들과 함께 영역별 난이도 조절을 위한 문항 수정을 하여 최종 초기 문해력 진단 도구 설계를 완료하였다.

## 2. 연구 방법

초기 문해력 진단 도구를 설계하기까지 연구 방법을 소개하면 다음과 같다. 첫째, 어휘 재구조화하는 작업을 연구의 시발점으로 삼았다([그림 1] 참조<sup>5)</sup>). 즉 일상적인 문장과 발화 속에 다양한 형태로 숨겨져 있는 ‘단어’를 핵심 단위로 삼고, 이들 단어를 각각 해체(解體)하거나 합체(合體)하는 능력을 살펴보거나, 발화 단어와 시각 단어 사이의 대응, 음절 마디와 글자 마디 사이의 대응, 음소(音素)와 자소(字素) 사이의 대응 능력을 살펴보고자 하였다.



[그림 1] 발화 단어와 시각 단어의 어휘 재구조화

5) 이 그림은 윤혜경(1997)의 ‘한글 읽기 발달의 한 모형’을 심영택이 부분적으로 수정한 것이다.

둘째, 설계 완료된 진단 도구 후보 문항은 아이들의 집중력을 고려하여 가능한 10분 이내 패드용으로 파일럿 측정이 가능하게끔 구성하였으며, 파일 형태로 간단히 저장할 수 있는 앱 형태로 개발하였다. 그리고 파일럿 적용 분석을 통하여 발견된 몇 가지 문제점은 다음과 같이 수정하였다. 아이들이 이해하기 힘든 진단 문항은 설명과 함께 예시 애니메이션을 보여주어, 앱 형태에서 쉽게 따라 할 수 있도록 구성하였다. 설명하는 목소리는 초등 저학년 학생의 선호도를 현장 교사로부터 자문받아 Tycast에서 제공하는 친근한 중성적 남자 아동 목소리로 하고, 그 내용을 TTS(Text to Speech)로 작업하였다. 삽화의 경우, 디지털 도구로 작업을 진행하되, 아이들의 시선 쏠림에 영향을 주지 않게끔 복잡하거나 현란하지 않은 일반적인 그림 풍으로 그려 줄 것은 그림 동화 전문작가에게 요청하였다.

셋째, 진단 도구를 적용하여 데이터를 채집한 과정은 다음과 같다. 먼저 IRB의 승인을 받은 후, 청주 시내 일반적인 학군의 초등학교를 중심으로 학년당 4개 학급 이하의 작은 규모 학교 2개교(1~2학년 28명)와 학년당 7개 학급 이상의 큰 규모 학교 3개교(1학년 332명)를 섭외하였다. 집단 측정 과정에서 잡음으로 인해 음성데이터가 없는 학생과 결측 데이터를 가진 학생 10명을 삭제하여, 최종 350명(남 165명, 여 185명)의 데이터를 확보하였다. 학교별 상황에 따라 여분의 교실이 있는 경우, 특정 시간대에 담임 관찰하에 1회에 1~6명씩 나누어 개별 또는 모둠별로 측정하거나, 연구보조원 2명의 지원을 받아 학급 전체 인원을 대상으로 한꺼번에 측정하였다.

마지막으로 채집된 데이터는 다음과 같이 분석하였다. 연구에 참여한 일반 아동 집단과 사전에 교사가 초기 문해력 부진아로 분류한 집단 모두 기존 문해력 진단 도구(정종성, 엄훈, 2018)에서 사용한 정답률과 소요 시간에 대한 평가분포를 분석하였다. 향후 앱 기반 문해력 진단 도구로서 채집할 데이터는 정답률, 소요 시간, 음성데이터 특징, 시선 데이터(경로, 멈춤 등) 등 다양하게 고려될 수 있으나 문항 난이도 조절을 위해서 정답률을 우선 고려했다. 초기 문해력 전문가와 현장 교사들이 각 문항의 난이도 조정, 문항 내용 수정, 응답 항목의 변경 등을 통하여 최종적으로 설계한 초기 문해력 진단 도구의 연구 결과는 다음과 같다.<sup>6)</sup>

6) 본고는 1차년도 연구로, 초기 문해력 측정 검사를 통해, 읽기 부진아를 선별하기 위한 것이 목적으로 하며(도구 목적 측면), 이를 토대로 2차년도에 초기 문해력 부진아 분석에 초점을 맞추고자 한다(도구 활용 측면).

### III. 연구 결과

#### 1. 인쇄물 인식

초기 문해력 개별화 수업에서 빈번하게 제기되는 교사들의 질문은 인쇄물(printed text)을 보는 것의 의미와 인쇄물을 보는 방법에 대한 것이다. 질문의 기저에는 인간의 언어 능력과 인지 능력이 어떻게 ‘보편적’으로 발달하는지에 대한 본질적인 의문과 더불어, 아이들의 문해력(읽기 기능과 태도)을 신장 또는 변화시키고자 하는 교사의 교육적인 열정이 잠재되어 있다.

M. Wolf(2007: 27~30)는 <책 읽는 뇌>에서 어떤 인쇄물을 보고 무의미한 형태와 유의미한 형태(문자나 단어)로 구별하는 능력은 이 시각적 특화를 더 발달시키는 능력, 즉 특화의 특화(specialization within a specialization)를 가능케 하는 내재적인 역량이 잠재되어 있기 때문이라고 본다. 그리고 P. J. Schwanenflugel과 N. F. Knapp(2016)에 의하면 전반적으로 인쇄물에 더 많이 노출된 (학령기 이전) 아동일수록 문자언어 텍스트의 목적과 기능을 더 잘 이해하는 경향이 있다고 한다. 한편, M. M. Clay(2016: 48)에 의하면, 읽기는 훑어보기(looking)에서 시작되고 훑어보기를 멈추는 순간 끝나는 듯하지만, 우리의 뇌는 스냅샷(snapshot)으로 인쇄물의 정보를 처리하지 않기에 초기 문해력 수업에서 ‘인쇄물을 보는 방법’을 학습해야 한다고 주장한다.

본고는 이와 같은 인쇄물의 의미와 학습해야 할 내용을 고려하면서 초등학교 저학년 문해력 기초 진단 도구에서 점검해야 할 항목을 다음 <표 2>와 같이 추출하였다. 단어를 그림으로 인식하기, 인쇄물의 특징인 표지와 제목 등에 대한 인식, 인쇄물에서 주목해야 할 위치인 행과 열 인식 등 총 4개 문항으로 구성하였다.

<표 2> 인쇄물 인식 능력 검사

하위 검사 영역	문항	선택	비고
단어를 그림으로 인식하기	글자와 그림의 유사성	‘나’와 비슷한 그림	1, 2차 도구
인쇄물의 특징 인식하기	제목 찾기	제목, 지은이, 출판사	
	둘째 줄, 첫 글자	줄글 제시(동화책)	2차 도구
	뒤집어진 글자 찾기	산, 공, 강	

## 2. 음운 인식: 자소-음소 대응 과제를 중심으로

음운 인식 능력은 음절을 ‘해체’하여 음소로 분리하거나 음소를 ‘합체’하여 음절을 만드는 음소 조작 능력을 의미한다. 개인 내적인 이러한 인지적 과제를 처리하는 음운 인식 능력은 읽기 학습 부진을 판별하는 데 가장 중요한 요인임에는 틀림이 없다. Stanovich(1993)는 이 능력을 학령기 아동의 읽기 능력 발달을 예측해 주는 핵심 요인이라고 하며, Meyer(2000)는 음운 인식과 직접적으로 관련된 음소 분리 능력은 유치원과 초등학교 저학년 시기의 읽기 능력을 가장 잘 예측해 준다고 한다(김자경, 2005).

본고에서 이러한 음운 인식 결핍 모델을 중시하여 초기 문해력 진단 도구 설계를 한 까닭은 다음과 같다. 일반 학생 집단은 연구 도구 개발의 원리에서 설명한 바와 같이, 자소-음소 대응이 잘 형성되어 있어 글을 읽을 때, 단어를 음절로, 음절을 다시 자소로 분해하는 과정이 ‘거의 자동으로’ 이루어진다. 즉 시각적으로 배열된(자소) 단어를 청각적으로 표상(음소)하는 과정에서 별도의 인지적 자원(cognitive resources)을 사용하지 않는다. M. Wolf(2007)는 이러한 ‘인지적 자동성’(cognitive automaticity)으로 인해, 이러한 표상들을 자동에 가까운 속도로 인지하게 되어, 사람들이 읽고 쓸 때 끊임없이 확대되는 정신적 프로세스에 훨씬 더 많은 시간을 할애할 수 있다고 한다.

반면, 읽기 학습 부진아 집단은 ‘불완전한’ 자소-음소 대응 지식으로 인해, 어떤 단어의 시각적 배열(visual array)에 따른 표상을 분석하는 작업에, 또는 그 배열에 따른 청각적 표상을 부여하는 작업에 자신의 한정된 인지적 용량(cognitive capacity)을 다 소진해 버린다고 한다.<sup>7)</sup> 읽기 학습 부진아 집단을 이러한 음운 인식 결핍 이론으로 설명하고자 하는 Stanovich(1993)에 의하면, 음운처리 과정은 일반적인 인지 능력과는 별도로 독립된 모듈이며, 거의 자동으로 작동하는 모듈로 보고 있다(원호택, 송종용, 1998).

초기 문해력 진단 도구를 설계함에 있어 음운 인식과 관련된 문항들을 여타 문항에 비해 상대적으로 많이 배정한 까닭은 이러한 음운 인식 결핍의 특성 때문이다. 음운 인식 정답률이 높은 문항은 교체하거나 수정하여 다음 <표 3>과 같이 구성하였다. 낱자의 소리 말하기에 대한 인식, 낱자의 이름에 대한 인식, 낱자의 소리 듣기와 이름 듣기, 소리와 문자의 대응 인식, 첫소리·가

7) 예를 들면, 읽기 장애 집단은 단어 앞 부분에 있는 자소만을 처리하고, 나머지 부분은 단어의 시각적 형태와 같은 맥락 정보를 이용한다고 하며, 이러한 맥락 정보를 사용하기 위해 작업 기억 용량을 더 많이 사용한다고 한다(원호택, 송종용, 1998). 이러한 연구를 재해석하면, 읽기 장애 집단은 엄밀하게 말하면, 맥락 정보를 잘못 사용하고 있으며, 그 맥락 정보도 단어의 형태적(도형적) 특징을 (시각적으로) 보고, 추측하는 활동을 하고 있다. 말하자면 ‘독해(understanding)’가 아니라 ‘해독(decoding)’을 하는데 한정된 인지적 자원(용량)을 다 쓰고 있는 셈이다.

운뎃소리·끝소리에 대한 인식 등 총 15문항으로 구성하였다.

〈표 3〉 음운 인식 능력

하위 검사 문항	자소·음소	선택	비고	
낱자 소리인식 낱자 이름인식	낱자의 소리 각각 말하기	ㄱ, ㄷ, ㄴ	1차 도구	
		(3개 문항) ㄷ, ㄱ, ㄴ	2차 도구	
	낱자의 이름 각각 말하기	ㄷ, ㅁ, ㅅ, ㅇ	1차 도구	
		(4개 문항) ㅅ, ㅎ, ㅌ, ㅃ	2차 도구	
	낱자의 소리나 이름, 듣고 일치하는 자모 찾기		ㄷ, ㅎ, ㅍ, ㅡ	1차 도구
			ㅍ, ㅎ, ㅍ, ㅡ	2차 도구
ㄴ, ㄴ, ㄴ, ㄴ			1, 2차 도구	
[ㄴ ㄴ] [ㄷ ㄴ] [ㄴ ㄴ] [ㄴ ㄴ]				
강, 공, 경				
소리와 문자의 대응 인식	[싱] 소리 찾기	굴, 싱, 천, 탑	1차 도구	
	‘싱’ 글자 인식	성, 싱, 상, 장	2차 도구	
	[꽃잎] 소리 찾기	꼬침, 끈넙, 꼬입	1, 2차 도구	
첫소리, 가운뎃소리, 끝소리 인식	[곰]의 첫소리 바꾸기	몸, 감, 곰	1, 2차 도구	
	[물]의 끝소리 탈락	바, 무, 소	1, 2차 도구	

### 3. 단어 읽기: 음절과 글자의 대응

우리의 귀와 눈은 일종의 모니터 역할을 하면서 지금 들려오는 소리나 보고 있는 문자들에 주의 집중하면서 모니터링한다. 그리고 기존의 표상(음절/음소)이나 이미지(글자/자소)를 떠올리면서 ‘그 어떤’ 소리나 문자를 탐색한다. 동일성 여부는 물리적인 속성이나 시각적인 특색과 같은 자질(資質)들을 교차점검하면서 이루어진다. 만약 그 자질들이 ‘익숙함/친숙함’으로 인식되는 경우, 동일/유사한 정보로 처리되고, ‘낯설/어색함’으로 처리되는 경우, 새로운 정보(표상/이미지)로 매치시킨다.

이 과정에서 물리적인 소리(음절/음소)와 시각적인 문자(글자/자소) 사이에 일치하는 현상이 발생하게 되면, 인지 시스템에 큰 문제가 발생하지 않는다. 예를 들어, 초등학교 저학년 아이들이 ‘강아지’라는 단어를 처음 보고 있는 상황에서 교사가 이 단어를 [강아지]라고 발음하였다고 가정해 보자. 대부분 아이들은 이 단어가 세 글자로 되어 있고, 이 단어는 세 음절로 발음할 수 있음을 인식할 것이다. 하지만 인지/재인 과정에서 문제가 되는 것은 물리적인 소리와 시각적인 문자 사이에 불일치 현상, 즉 대응 관계가 복잡하거나 불분명한 경우이다. 예를 들어 ‘꽃잎’이

라는 단어를 아이들이 처음 보았을 때, 누군가 이 단어를 [꼐넛]이라고 발음할 것이다. 이 단어의 글자 마디는 ‘꼐’와 ‘넛’으로 나뉘는데, 그 소리마디(음절)는 [꼐]와 [넛]이다. ‘강아지’와 [강아지]로 처리했던 기존의 인지 시스템으로는 처리하기 힘든 문제 상황이 발생하게 된다.

초기 문해력 진단 도구 개발에서 추출한 다음과 같은 항목, 즉 소리(음절)와 문자(글자)의 대응 관계는 단어 읽기 능력과 밀접한 관계가 있다. 그 항목은 다음 <표 4>와 같다. 음절 수와 글자 수의 일치 인식, 음절 위치 교체, 음절 탈락과 추가에 따른 단어 인식 등 총 5문항으로 구성하였다.

<표 4> 단어 읽기 능력 검사: 음절과 글자의 대응

하위 검사 문항	음절-글자	선택	비고
음절, 글자 수 인식	소리내어 읽기	같이	2차 도구
	‘상자’ 음절 수와 글자 수	물, 가위, 동화책	1,2차 도구
음절, 글자 변화 인식	‘박수’ 음절 위치 교체	바퀴, 수박, 이사	1,2차 도구
	‘눈물’에서 [물] 음절 탈락	남, 눈, 굴	1,2차 도구
	‘새우’에서 음절 추가	새우등, 새우뺨, 새우잠	2차 도구

#### 4. 어휘 능력: 끝말잇기, 동의어, 상·하위어, 반의어, 연상어

어휘란 어떤 일정한 범위 안에서 쓰이는 단어의 수효(數效)를 의미한다. 초기 문해력 진단 검사에서 어휘 능력 검사는 그 범위를 명확히 규정하기 힘들겠지만, 일반적인 상황에서 적절한 단어를 어느 정도 사용할 수 있는지, 그 능력을 예측하고자 하는 작업이다. 여기서 쟁점은 일반적인 상황 설정과 단어 사용의 적절성에 있다. 미국 독서위원회에서 어휘를 ‘효과적인 의사소통을 위해서 반드시 알아야 하는 단어들(words)을 가리키는’ 것으로 정의하고, 아이들이 말하거나 듣기를 하면서 사용하거나 인식하는 구두 어휘(oral vocabulary)와 인쇄물에서 사용하거나 인식하는 읽기 어휘(reading vocabulary)로 구분한 바 있다(P. J. Schwanenflugel & N. F. Knapp, 2016).

초기 문해력에서 이러한 구분이 중요한 까닭은 아이들이 읽을 수는 없지만, 알고 있는 단어가 생각보다 많을 수 있기 때문이다. 일반적인 어휘 검사(語彙檢査)는 국어교육에 필요한 자료를 얻기 위해 개인이나 집단이 알고 있는 어휘의 양(量)이나 경향을 측정하고자 한다. 하지만 학령

기가 시작되는 초기 문해력에서 어휘 검사는 단순 말놀이 활동이나 의미의 비교/대조 등을 통해, 기본적/기초적인 구두 어휘와 읽기 어휘의 수준을 진단하고 점검하는 일이 중요하다.

본고는 이와 같은 어휘 능력 검사의 범주와 교육적인 의미 등을 고려하여 정답률이 매우 높은 기존 문항 3개를 신규문항 1개로 교체하여 다음 <표 5><sup>8)</sup>와 같이 지시 대상 인식, 대등한 단어 범주 인식, 상하의 단어 범주 인식 등 총 7문항으로 구성하였다.

〈표 5〉 어휘 능력

하위 검사 문항	의미관계	선택	비고
지시 대상	‘기린’ 해독 및 낱말 재인	[기린], [돼지], [나비] 그림	1, 2차 도구
	음성 듣고 낱말 재인	[돼지], [토끼], [나비] 그림	
	‘나팔’과 ‘꽃’의 합성어	나팔관, 나팔꽃, 나팔수	1차 도구
	단어 정확하게 듣기	노래	
	단어 소리내어 읽기	학교	
		개미-미끄럼틀(끝말잇기)	개구리, 꿀림, 틀니
대등 범주	달걀(동의어)	채소, 계란, 고기	1, 2차 도구
	추워요(반의어)	시원해요, 더워요, 차가워요	
	여름(연상어)	매미, 선풍기, 눈사람	
상하 범주	아빠, 엄마, 동생, 나	학교, 동네, 가족	

## 5. 읽기 유창성: 읽기 유창성의 성격과 발달적 특성

읽기 유창성(reading fluency)은 글을 소리내어 읽되, 정확하게 발음하고(정확성), 속도를 조절하며 읽고(자동성), 상황이나 맥락을 고려하며 리듬감(운율성)을 살려 읽는 것으로 개념화할 수 있다. Schwanenflugel & Knapp(2016)은 이러한 세 가지 읽기 유창성을 다음과 같이 부연 설명하고 있다. 읽기 맥락에서 정확성(accuracy)은 단어를 올바르게 읽을 수 있는 것으로, 까다로운 다음절이나 다형태소로 된 단어나 어렵거나 모호한 뜻을 지닌 단어를 해독할 수 있는 기능과 관련된 능력인데, 이는 물리적인 음성의 표준 발음 문제이다. 그리고 자동성(automaticity)은

8) 어휘 능력에서 다루고 있는 ‘노래’, ‘학교’와 같은 문항은 향후 ‘단어 읽기 능력’(음절과 글자 인식) 영역으로 이동해서 다루고자 한다.

별다른 노력을 들이지 않은 채 글을 소리 내어 읽는 과정을 빠르게 수행하는 것으로, 작업 기억을 단어 해독과 통사적/의미적 관계 처리에 사용하는 능력인데, 이는 인지적 자원의 배분 차원 문제이다. 끝으로 운율성(prosody)은 말의 자연스러운 분절을 고려하며 글을 소리 내어 읽는 것으로, 중요한 정보 이해와 자신의 감정을 강세와 리듬, 숨 고르기 등으로 표현하는 능력인데, 이는 의도적인 의미 전달의 문제이다.

읽기 유창성은 읽기 기능의 발달적 특성에 비추어 볼 때 폐쇄적인 해독 행위와 개방적인 독해 활동의 중간 지점에 있다고 볼 수 있다. 그리고 학령기 초기에 구두 읽기 유창성(oral reading fluency)의 변화가 실제로 이루어지며 중학교 입학할 무렵에는 그 수준의 변화가 거의 일어나지 않는다. Pikulski & Chard(2005)에 의하면 읽기 유창성은 글자나 단어의 해독(decoding) 단계에서 문장이나 문단의 독해(comprehension) 단계로 넘어가는 ‘다리 역할’을 한다고 한다(엄훈 외, 2022).

본고는 이와 같은 읽기 유창성의 성격과 발달적 특성을 고려하여 정답률이 매우 높은 기존 문항을 대체하고 초등학교 저학년 문해력 기초 진단 도구에서 점검해야 할 항목을 다음 <표 6>과 같이 단문 읽기, 적절한 문장 순서, 긴 글 읽기(음성, 시선 추적) 총 4문항을 구성하였다.

**<표 6> 읽기 유창성 영역**

하위 검사	시간 제한(정확성, 속도성)	비고
단문 읽기	하늘이 파랗습니다.	1차 도구 문항
	마침내 비가 ( ) 내리기 시작했습니다.	2차 수정 문항
문장 순서	나는 밥을 많이 먹었다.	2차 수정 문항
긴 글 읽기 (음성/ 시선추적)	아기 코끼리 가족은 모두 함박웃음을 터뜨렸습니다.	2차 수정 문항 (음성)
	<아기코끼리 방귀 소리>	1, 2차 도구 문항 (시선추적)

## 6. 초기 문해력 진단 도구 개발

앞에서 살펴본 프로토타입 모형 영역별(3개, 14개, 3개, 9개, 1개)로 총 30문항에 대해서 초등학교 1학년과 2학년 각 1개 학급씩 총 28명(남 12명, 여 16명)에게 파일럿 텍스트를 실시하였다.<sup>9)</sup> 실험 측정시 아이들이 지시문을 제대로 이해하는지, 응답하는데 인터페이스 문제는 없는

9) 31번 시선 추적 문항은 논의의 범위와 분량상 추후 2차년도 연구에서 자세히 분석할 예정이다.

지, 진단 과정에서의 문제 발생 상황 등 진단 문항 외의 다양한 항목에 대해서 담임교사와 연구진의 관찰기록이 수반되었다.

〈표 7〉은 개발된 영역별 문항 30개에 대해서 학생들의 정답률과 평균 응답 소요 시간이다. 여기에는 담임교사가 문해력 지도가 필요하다고 판단하는 부진 학생 3명의 데이터도 포함되어 있다. 문항별 정답률을 보면, 인쇄물 인식과 어휘 능력은 변별력이 있는 문항이 적절히 있으나, 음운 인식과 단어 읽기의 경우 정답률이 매우 높아 변별력을 높이기 위해 재조정이 필요한 것으로 나타났다. 그러나 전반적으로 부진 학생의 수가 3명으로 매우 적어 정답률을 직접 비교하는 것은 다소 어려움이 있다고 판단하였다. 이에 평균 소요 시간을 살펴본 결과 ‘어휘 능력’ 영역의 4번, 25번, 26번, 27번 문항은 전체 학생과 부진 학생의 평균 소요 시간에 큰 차이가 있는 것으로 나타났다. 이는 이들 문항이 전체 학생과 부진 학생을 판별하는데 중요한 역할을 하고 있음을 시사한다.

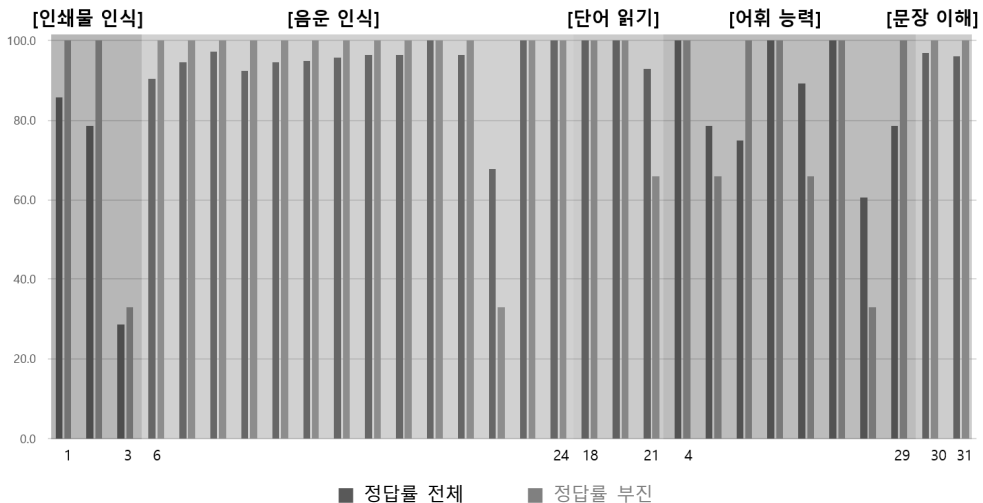
〈표 7〉 1차 28명(3명 부진 포함) 30개 문항에 대한 파일럿 테스트 결과

영역	문항	문항 형태	정답률		평균 소요 시간	
			전체 (N=28)	부진 (N=3)	전체 (N=28)	부진 (N=3)
인쇄물 인식	1	선택형	85.7	100	1.32	2.5
	2	선택형	78.6	100	1.89	1.9
	3	선택형	28.6	33	3.79	4.25
음운 인식	6	녹음	90.3	100	-	-
	7	녹음	94.6	100	-	-
	8	녹음	97.2	100	-	-
	9	녹음	92.3	100	-	-
	10	녹음	94.6	100	-	-
	11	녹음	94.9	100	-	-
	12	녹음	95.7	100	-	-
	13	선택형	96.4	100	0.64	0.25
	14	선택형	96.4	100	0.29	0.3
	15	선택형	100.0	100	0.18	0.35
	16	선택형	96.4	100	0.68	0.2
	22	선택형	67.9	33	0.00	0
	23	선택형	100.0	100	0.68	0.45
	24	선택형	100.0	100	0.00	0
단어 읽기	18	선택형	100.0	100	0.25	1.05
	19	선택형	100.0	100	0.29	0.45
	21	선택형	92.9	66	0.82	2.3
어휘 능력	4	선택형	100.0	100	0.25	0.6
	5	선택형	78.6	66	1.86	1.2
	17	선택형	75.0	100	2.11	1.35

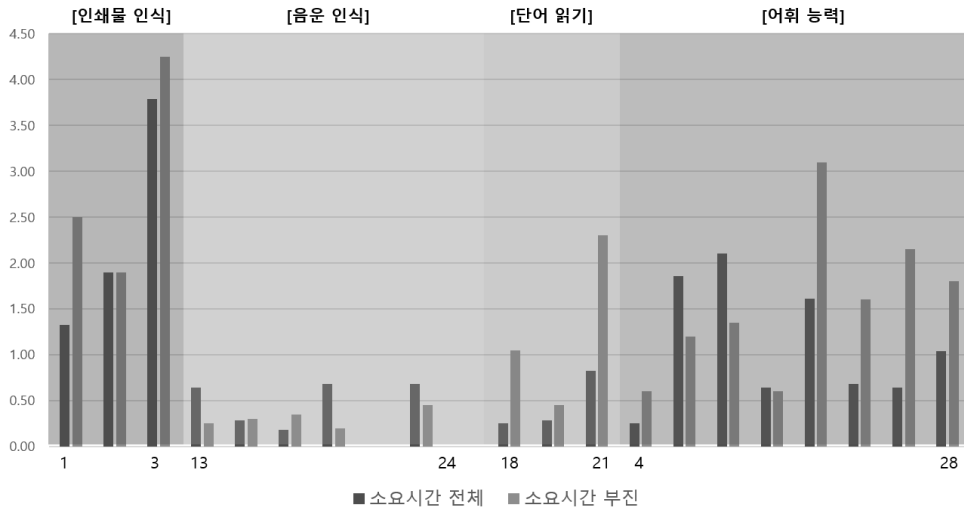
영역	문항	문항 형태	정답률		평균 소요 시간	
			전체 (N=28)	부진 (N=3)	전체 (N=28)	부진 (N=3)
	20	선택형	100.0	100	0.64	0.6
	25	선택형	89.3	66	1.61	3.1
	26	선택형	100.0	100	0.68	1.6
	27	선택형	60.7	33	0.64	2.15
	28	선택형	78.6	100	1.04	1.8
	29	녹음	96.9	100	-	-
문장 이해	30	녹음	96.0	100	-	-

그리고 문해력 진단 평가 후 <표 7>의 선택형 21문항 × 2종류(정답점수, 응답반응속도) × 28명 = 1,176개의 데이터, <표 7>의 음성 녹음으로 9문항 × 28명 = 252개의 데이터를 수집하여 분석하였다.

[그림 2], [그림3]은 개발된 영역별 문항에 대해서 학생들의 정답률과 평균 응답 소요 시간을 비교하는 그래프이다. 음운 인식 영역의 앞 문항들은 정답률이 높게 분포하고 있어 난이도 조정이 필요하기도 하지만, 학생의 긍정적 자기효능감을 위한 난이도 유지도 고려해야 한다. 그런데 어휘 능력 영역에서 부진 학생들의 정답률은 상대적으로 크게 떨어지는 것을 확인할 수 있다. 한편, 평균 소요 시간의 경우 대부분 문항에서 부진 학생들이 상당히 시간을 많이 소요하는 데, 이는 문항에 대한 이해가 부족해서 생기는 현상으로 보인다.



[그림 2] 1차 파일럿 테스트 정답률 비교



[그림 3] 1차 파일럿 테스트 평균 소요 시간 비교

〈표 8〉은 2차 적용까지 마친 후 학생들의 정답률과 평균 응답 소요 시간으로 전체 학생과 1차 파일럿 및 2차 적용에 참여한 부진 학생 25명의 결과와 비교하여 제시하였다.

먼저 부진 학생들의 문항별 정답률이 60% 이하인 문항은 1번(그림글자), 3번(책 표지), 27번(달걀 동의어), 28번(여름 관련 단어)인데, 1차 파일럿 적용 연구와 비슷하게 인쇄물 인식과 어휘 능력 영역에 해당한다. 인쇄물 인식과 어휘 능력에서 두 집단의 차이가 확실하게 나타났으며, 음운 인식과 단어 읽기의 경우에는 다소 비슷한 정답률을 나타내었다.

한편 평균 소요 시간에서는 특이한 점이 관찰되었는데 부진 학생들의 정답률이 매우 낮은 ‘어휘 능력’ 영역에서는 평균 소요 시간이 높게 나타났으나, 부진 학생들의 정답률이 매우 높은 ‘음운 인식’ 영역과 ‘단어 읽기’ 영역에서는 평균 소요 시간이 비슷하거나 아니면 오히려 50% 이상 짧은 문항도 있었다. 이러한 결과는 문해력 부진 학생들의 경우 이해가 잘 되지 않는 어휘 영역의 문항에서는 문제와 지문, 보기를 이해하는데 상당한 시간을 소요하는 반면, 본인들이 이해가 되었다고 판단되는 문제들은 깊은 고민 없이 정답을 선택하기 때문인 것으로 파악된다. 그리고 일반 학생에 비해 부진 학생의 소요 시간이 40초 이상 지연되는 문항으로는 인쇄물 인식 영역인 3번과 어휘 영역인 25번, 27번, 28번이 관찰되었다.

그 밖에 특이한 사항으로는 22번 문항(〈표 3〉 음운 인식 영역의 첫소리, 가운데소리, 끝소리 인식-‘꿈’의 첫소리를 바꾸기 문항)의 경우 문해력 부진 학생들의 정답률은 20%밖에 되지 않는 데도 불구하고 응답 소요 시간은 0초로 기록되었다. 일반학생의 소요 시간도 0초로 기록되어

있는데, 이 원인은 예시 문제를 보고 나서 제시된 22번 문항의 문제를 미리 이해했다고 생각하고 문제가 음성으로 제시되고 있는 중에 미리 정답이라고 생각되는 답을 계속해서 클릭했기 때문으로 추정된다.

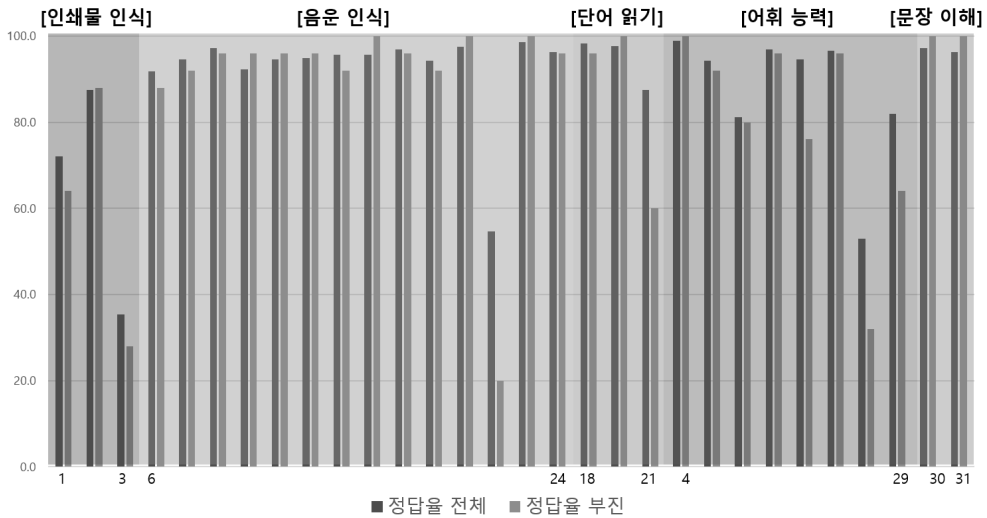
1차 파일럿 테스트 결과 개발된 문항이 문해력 부진 학생을 판별할 수 있는 인공지능 학습데이터로 사용할 수 있겠다는 연구진의 판단에 따라 초등학교 1학년 4개 학교의 31학급 360명에게 2차로 적용하여 결측치를 제외하고 총 350명의 데이터를 수집하였다. 여기에는 파일럿 집단 28명이 수정된 도구로 재응시한 데이터가 포함되어 있는데, <표 8>과 같이 수정된 도구가 문항 또는 순서가 변경되었으며 정답을 알려주지도 않았기 때문에 재측정해도 차이가 없다고 판단했기 때문이다. 따라서 분석 대상 350명 중에는 담임이 문해력이 떨어져 수업을 따라가기 힘들다고 판단하는 25명의 문해력 부진 학생이 포함되어 있었으며,  $350명 \times 21문항 \times 2종류(정답점수, 응답반응속도) = 14,700$ 개의 데이터와 음성 녹음으로  $350명 \times 9문항 = 3,150$ 개의 데이터이다. 이때 담임교사의 주관적 판단 기준에 의한 부진 학생의 데이터가 부진 학생으로 라벨링되는 것은 아니고, 설계된 문항에 대한 데이터 분포를 맥락적으로 확인하고자 함이다.

<표 8> 2차 적용 결과

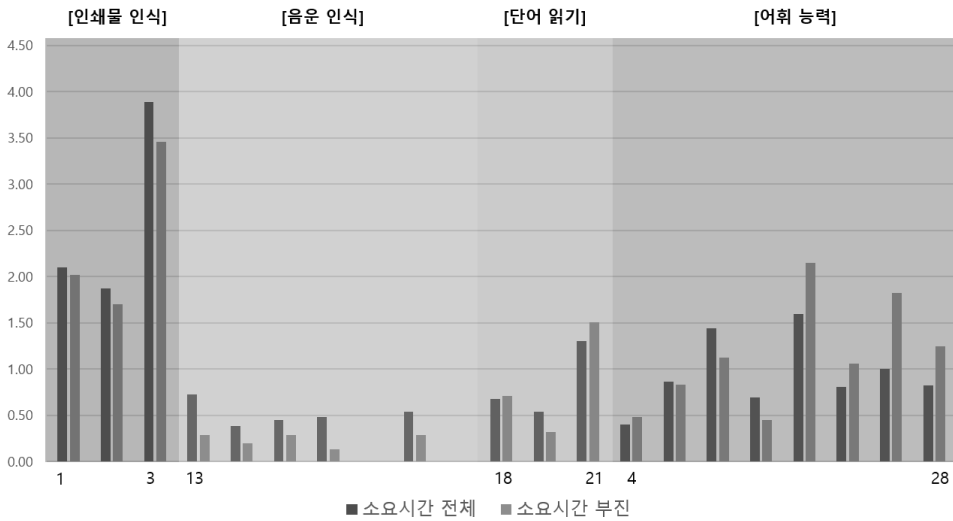
영역	문항	문항 형태	정답률		평균 소요 시간	
			전체 (N=350)	부진 (N=25)	전체 (N=350)	부진 (N=25)
인쇄물 인식	1	선택형	72.0	64	2.10	2.02
	2	선택형	87.4	88	1.87	1.70
	3	선택형	35.4	28	3.89	3.46
음운 인식	6	녹음	91.7	88	-	-
	7	녹음	94.6	92	-	-
	8	녹음	97.2	96	-	-
	9	녹음	92.3	96	-	-
	10	녹음	94.6	96	-	-
	11	녹음	94.9	96	-	-
	12	녹음	95.7	92	-	-
	13	선택형	95.7	100	0.72	0.29
	14	선택형	96.9	96	0.38	0.19
	15	선택형	94.3	92	0.45	0.29
	16	선택형	97.4	100	0.48	0.13

영역	문항	문항 형태	정답률		평균 소요 시간	
			전체 (N=350)	부진 (N=25)	전체 (N=350)	부진 (N=25)
	22	선택형	54.6	20	0.00	0.00
	23	선택형	98.6	100	0.54	0.29
	24	선택형	96.3	96	0.00	0.00
단어 읽기	18	선택형	98.3	96	0.68	0.70
	19	선택형	97.7	100	0.53	0.32
	21	선택형	87.4	60	1.30	1.50
어휘 능력	4	선택형	98.9	100	0.40	0.48
	5	선택형	94.3	92	0.86	0.83
	17	선택형	81.1	80	1.44	1.12
	20	선택형	96.9	96	0.69	0.45
	25	선택형	94.6	76	1.59	2.14
	26	선택형	96.6	96	0.81	1.06
	27	선택형	52.9	32	1.00	1.82
	28	선택형	82.0	64	0.82	1.25
	29	녹음	97.2	100	-	-
문장 이해	30	녹음, 시선	96.3	100	-	-

[그림 4], [그림 5]는 1, 2차 결과를 종합한 학생들의 정답률과 평균 응답 소요 시간을 그래프로 나타낸 자료이다. 앞서 서술하였듯 ‘인쇄물 인식’ 및 ‘어휘 능력’ 영역에서 문해력 부진 학생들의 정답률이 낮게 나타나는 경향이 관찰되었다. 평균 소요 시간은 ‘어휘 능력’ 부분에서 일반 학생들에 비해 상당히 많은 시간을 소요하고 있는 것이 확인되나, 그 외에는 대부분 더 적은 시간을 문제 풀이에 사용하고 있는 것으로 나타났다.



[그림 4] 1, 2차 적용 전체 정답률 비교



[그림 5] 1, 2차 적용 전체 소요 시간 비교

1차 적용 시에는 학생수가 28명으로 4번에 나누어 소규모 측정이 이루어졌기 때문에 연구자의 통제 아래, 비교적 학생들이 성실하게 진단에 참여했다고 판단되며, 2차 때에는 적게는 5명~17명씩 이루어졌기 때문에 학생에 따라서 문항별 집중도가 흩뜨려진 데이터도 포함될 가능성이 크고 실제로 10명의 결측 데이터도 발생했다. 따라서 정답률과 소요 시간이 뚜렷하게 역관

계를 보이지 않는 것으로 판단된다. 전체 실험에 참가한 350명 학생 데이터에 대하여 문항 1개당 1점의 동일 가중 점수로 채점한 전체 득점 점수와 응답시간을 표로 정리한 결과는 <표 9>와 같다.

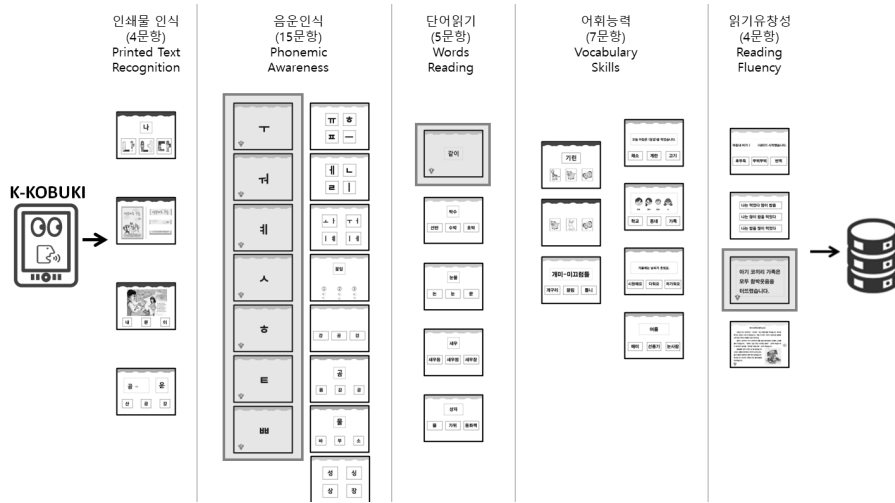
〈표 9〉 문해력 검사 평균 점수 및 평균 응답시간

영역		득점(점)		응답시간(초)	
		M	SD	M	SD
선택형 문항 (21개)	전체 학생 (N = 350)	18.1	1.93	20.5	12.94
	부진 학생 (N = 25)	16.7	1.90	25.0	13.58
녹음형 문항 (9개)	전체 학생 (N = 350)	8.6	0.92	-	-
	부진 학생 (N = 25)	8.3	1.59	-	-

선택형 문항의 최고 득점은 20점이고 최저 득점은 10점이었다. 문해력 부진 학생의 평균은 16.7점으로 전체 학생의 평균 18.1보다 1.4점 낮은 것으로 나타났다. 한편 녹음형 문항에서의 점수는 전체 학생 8.6, 부진 학생 8.3으로 큰 차이가 없었으나 표준편차가 각각 0.92, 1.59로 나타나 부진 학생의 경우엔 학생에 따라 편차가 큰 것으로 나타났다.

한편 선택형 문항의 응답시간은 가장 빠르게 응답한 학생이 0초였고, 해당 학생의 전체 점수는 29점(30점 만점)으로 매우 높은 수준으로 나타났다. 반면 응답시간이 가장 느렸던 학생은 95초를 소요하였으며, 전체 점수는 18점으로 문해력 부진 학생 중 하위 1%에 해당하였다.

이러한 결과를 바탕으로 문해력 전문가 1인, 현장 교사 3인, 문해력과 인공지능융합 연구자 1인이 타당도 협의를 거쳐 [그림 4]와 같이 초기 문해력 측정을 위한 검사 도구를 수정하여 제안한다. 부진아들도 모두 맞춘 문항으로 13번, 16번, 23번, 19번이 있는데, 이를 모두 난이도를 조절하면 부진아의 검사에 좌절감을 느낄 수 있기에 13번과 16번 문항의 난이도만 상향 조정하였다. 최종적으로 개발된 문항은 인쇄물 인식 4문항, 음운 인식 15문항, 단어 읽기 5문항, 어휘 능력 7문항, 읽기 유창성 4문항 총 35문항으로 구성되어 있다. 이 중 선택형 문항은 22개, 녹음형 문항은 9개(파란 문항), 시선 추적 문항은 1개로 구성하였다. 수정된 문항은 차후 학교에 적용하여 분석에 활용하고자 한다.



[그림 4] 초기 문해력 진단 도구 설계

## IV. 결론 및 제언

### 1. 결론

초기 문해력은 발생적 문해력이 완성되는 시점으로 문자 인식하기, 따라 읽기, 띄어 읽기, 유창하게 읽기 등의 독서 활동을 하는 시기이다. 그런데 미숙한 독자(poor reader)는 독립적인 읽기 기술 자체뿐만 아니라 단어 지식 자체도 부족하기에 읽기 발달 격차로 인해 읽기 매튜 효과, 즉 읽기의 부익부빈익빈(富益富貧益貧) 현상이 발생하게 된다. 본고의 초기 문해력 진단 도구는 이러한 독자 발달과 읽기 발달의 특성뿐만 아니라, 짧은 시간에 편리하게 미숙한 독자를 선별할 수 있게끔 설계하여, 현장 적합성과 활용 빈도를 높이고자 하였고, AI디지털 교과서의 도입을 고려하여 지필식과 컴퓨터 기반의 문제점 극복 방안을 모색하고자 하였다.

먼저 국내외 문헌 및 기존 측정 도구 분석을 통하여 초기 문해력 측정 도구가 갖추어야 하는 이론적인 틀과 맥락을 고찰하고 문해력 측정 도구 영역을 설계하였다. 모형에 제시된 다섯 영역은 인쇄물 인식 능력, 음운 인식 능력, 단어 읽기, 어휘 능력, 읽기 유창성이며, 영역별 검사 문항을 개발하고 페드용 프로토타입 측정 도구를 개발하였다. 개발된 프로토타입을 포커스 그룹에 속한 현장 교사 2인이 속한 초등학교 저학년 10명에 문항별 적용하여, 문항의 이해도와 난이

도 분포 등을 고려하여 1차 도구를 개발하였다. 이어서 문해력 전문가와 현장 교사들이 함께 영역별 난이도 조절로 문항 수정을 하였다.

초기 문해력 진단 도구의 28명에 대한 파일럿 테스트 결과, 문항별 정답률을 보면, 인쇄물 인식과 어휘 능력은 변별력이 있는 문항이 적절히 있으나, 음운 인식과 단어 읽기의 경우 정답률이 매우 높아 변별력을 높이기 위해 재조정이 일부 필요한 것으로 나타났다. 주목할 점은 어휘 능력 영역에서 부진 학생들의 정답률이 크게 떨어지는 것을 확인할 수 있었으며, 평균 소요 시간의 경우 대부분의 문항에서 부진 학생들이 문항에 대한 이해를 하는 과정에서 상당히 시간을 많이 소요하는 것으로 나타났다. 파일럿 테스트 결과 개발된 문항으로 초등학교 1학년 4개 학교의 31학급 총 350명의 자료를 수집하였다. 문항별 정답률을 보면 인쇄물 인식과 어휘 능력에서는 두 집단의 차이가 확실하게 나타났으며, 음운 인식과 단어 읽기의 경우에는 비슷한 정답률을 나타내었다. 1, 2차의 결과를 종합하여 문항 1개당 1점의 점수로 채점하여 전체 득점 점수와 응답시간을 표로 정리한 결과, 선택형 문항의 최고 득점은 20점이고 최저 득점은 10점이었다. 문해력 부진 학생의 평균은 16.7점으로 전체 학생의 평균 18.1보다 1.4점 낮은 것으로 나타났다. 한편 녹음형 문항에서의 점수는 전체 학생 8.6, 부진 학생 8.3으로 큰 차이가 없었으나 표준편차가 각각 0.92, 1.59로 나타나 부진 학생의 경우엔 학생에 따라 편차가 큰 것으로 나타났다.

## 2. 제언

컴퓨터 기반 중심의 초기 문해력 진단 도구는 속도성, 정확성, 용이성과 같은 장점을 지니고 있음에도 불구하고 다음과 같은 단점/약점을 극복해야 한다. 첫째는 집단을 대상으로 이 도구를 사용하기에 문해력이 부진한 아이들이 겪는 인지적 어려움이 무엇인지 구체적으로 파악하기 힘들다. 즉, 특정 문자를 시각적으로 인지하는 것이 왜 어려운지, 어떤 소리는 들리는데 왜 다른 소리가 들리지 않는지, 다음에 어떤 글자가 나오는 것을 예측하는 것이 힘든지, 무의미 단어를 읽는 것을 왜 어려워하는지 등이다. 범박하게 요약하자면 소리-글자에 대한 모니터링하기, 탐색하기, 교차점검하기, 매치시키기와 같은 언어적 정보처리 과정에 대한 총체적 시스템이 어떻게 작동되고 있는지 파악하기 힘들다.

둘째, 진단 검사와 교수 방법을 곧바로 연결할 수 없다는 점이다. 인지-언어적인 특성을 이용한 초기 문해력 진단 검사에서 중요한 것은 정보처리 과정 시스템이 어디서 오작동을 일으키는 지 파악하는 일이다. 그리고 그 오작동을 바로 잡는 적절한 교수 방법/기법을 구안하거나 설계하는 일이다. 하지만 컴퓨터 기반은 비대면의 검사 방식이기에 제시된 문항에 부적절한 답을 한

경우, 시·청각적인 오단서, 문법적·의미적인 오단서를 파악하기가 불가능에 가깝다. 즉, 평가의 객관성을 확보할 수 있다는 장점은 있으나, 그 아이에게 적합한 교수 방법의 적절성을 확보하기 힘들다는 단점이 있다.

끝으로 추후 이들 문항의 비중과 점수를 각각 하하(1점), 중하(2점), 중중(3점)을 부가하고, 이들 문항 전체를 다 맞춘 경우와 그렇지 않은 경우를 세분하여, 최종적으로 구분 점수를 부가하여 인공지능 모델링을 하는 방안도 모색할 수 있을 것이다. 이는 초등학교 저학년 아동의 문해력 수준을 영역별로 구체적으로 비교하고, 그 이후의 변화를 추적하는 지표로 사용할 수 있을 것이다.

※ 논문 투고일: 2024. 7. 1.    ※ 논문 수정일: 2024. 8. 14.    ※ 게재 확정일: 2024. 8. 30.

### 〈참고문헌〉

- 교육부(2022). **국어과 교육과정**, 교육부 고시 제2022-33호 [별책 5].
- 김자경(2005). 초등학교 학습장애 판별 준거에 관한 논의: 불일치 준거와 대안적인 방안을 중심으로. **특수아동교육연구**, 7(4), 257-278.
- 김중훈, 이승미, 정평강, 송푸름(2019). 교육과정 기반의 초등학교 1학년 한글 해득 진단 도구 문항 개발 연구. **교육과정평가연구**, 22(2), p.31-60.
- 박순경(2018). **한글 해득 웹 진단 도구(한글 토박 토박) 개발 연구**.
- 박현정(2010). **읽기학습부진 위험군 아동의 조기선별 및 조기개입의 효과 검증**. 이화여자대학교 대학원. 박사학위청구논문.
- 엄훈, 염은열, 김미혜, 박지희, 진영준(2022). 초기 문해력 교육 읽기 따라잡기로 시작해요. 서울: 사회평론아카데미.
- 원호택, 송중용.(1998). 읽기장애의 정보처리적 접근: 독해를 중심으로. **심리학의 연구문제**, (5), 53-70.
- 윤혜경(1997). **아동의 한글 읽기발달에 관한 연구: 자소-음소 대응규칙의 터득을 중심으로**. 부산대학교 대학원 박사학위청구논문.
- 이경화, 김상한, 김명순, 이수진, 최규홍, 이경남, 강서희, 박혜림(2018). 웹 기반 한글 해득 표준화 진단 검사도구 개발 연구. **청람 어문 교육**, 67, 63-88.
- 이승미, 김중훈, 정평강, 송푸름(2018). 웹기반 CR-CBA 한글해득수준진단검사의 측정학적 적합성. **교육과정평가연구**, 21(4), 125-149.
- 정종성, 엄훈(2018). 초기 문해력 측정을 위한 2학년용 검사 도구 개발. **독서연구**, 49, 9-28.
- 천경록(2020). 독서 발달과 독자 발달의 단계에 대한 고찰. **국어교육학연구**, 55(3), 313-340.
- Clay, M. M.(2016). *Literacy Lessons Designed for Individuals*. Second Edition. The Marie Clay Literacy Trust.
- Han, J., & Shim, Y. (2023). *Inclusive Design of a Tool to Screen Literacy of Lower Grade Elementary School Students*. 2023 IEEE International Conference on e-Business Engineering (ICEBE), 178-180.
- Hoskins, W. H., Hobbs, W. I., Eason, M. J., Decker, S., & Tang, J. (2021). The design and implementation of the Carolina Automated Reading Evaluation for reading deficit screening. *Computers in Human Behavior Reports*, 4, 100123. Available online 24 July 2021.
- Juel, C.(1988). Learning to read and write: A longitudinal study of fifty-fourchildren from first through fourth years. *Journal of Educational Psychology*, 80(4), 437-447.
- McArdle A., Pattwell D., Vasilaki A., Griffiths RD & Jackson MJ. (2001). *Contractile activity-induced oxidative stress: cellular origin and adaptive responses*. Am J Physiol Cell Physiol, 280(3), 621-627. <https://doi.org/10.1152/ajpcell.2001.280.3.C621>
- Ruddell, R.B., & Unrau, N.J. (2004). *Theoretical Models and Processes of Reading*. Fifth Edition. International Reading Association.
- Scarborough, H. S.(1998). Early Identification of Children at Risk for Reading Disabilities: Phonological Awareness and Some Other Promising Predictors. In B. K. Shapiro, P. J. Accardo, & A. J. Capute (Eds.), *Specific Reading Disability: A View of the Spectrum*

- (pp. 75-119). Timonium. MD: York Press.
- Schwanenflugel-Paula J. & Knapp-Nancy F.(2016). **독서교육의 이론과 실재를 위한 독서심리학**. 서혁, 윤준채, 이소라, 류수경 역(2021). 서울: 사회평론아카데미.
- Stanovich, K. (1986). Matthew Effects in Reading: Some Consequences of Individual Differences in the Acquisition of Literacy. *Reading Research Quarterly*. 21(4), 360-407.
- Wolf, M. (2007). Proust and the Squid: The Story and Science of the Reading Brain. 이희수 옮김(2009). **책 읽는 뇌**. 살림.

〈Abstract〉

---

## Designing an Early Literacy Assessment Tool

---

Shim, Youngtaek<sup>1</sup>, Han, Jeonghye<sup>2</sup>

Early literacy refers to the period during which foundational literacy skills are being developed, including recognizing letters, reading aloud, and reading fluently. However, struggling readers often lack the reading skills and knowledge, which leads to the Matthew effect in the reading of development due to reading skill gaps. Considering the developmental characteristics of such readers and their reading development, an early literacy assessment tool was designed to allow for a quick and accurate evaluations, and was suitable for educational settings and can be used frequently. First, a theoretical framework and context for the early literacy assessment tools were examined through a review of the existing research. The assessment was divided into five domains: printed text recognition, phonemic awareness, word reading, vocabulary skills, and reading fluency. With the assistance of the field teachers, the initial tool design, which consisted of 31 items was completed by validating the items with 28 lower elementary school students. The data from 332 first-grade students was collected and analyzed using the prototype measurement tool. Notably, the struggling students showed significantly lower accuracy rates in the printed text recognition and vocabulary domains, while their phonemic awareness and word reading displayed relatively similar accuracy rates to other students. In terms of the average response time, the struggling students took considerably longer to comprehend the items. Based on the data analysis results, the item modifications were made in collaboration with the literacy experts to adjust the difficulty level of each domain. The design of the diagnostic tool composed of 35 items, was completed, the aim of which was to help develop the early literacy deficit prediction model, by using data collected through a tablet-based application.

**Keywords** : Early Literacy, Assessment Tool Design, Printed Text Recognition, Phonemic Awareness, Word Reading, Vocabulary Skills, Reading Fluency

---

1. Professor, Cheongju National University of Education, shimyt@cje.ac.kr (First Author)

2. Professor, Cheongju National University of Education, hanjh@cje.ac.kr (Corresponding Author)