

# 2023년 묵자-점자 병렬말뭉치 구축

- PART 1. 묵자-점자 병렬말뭉치 구축
  - PART 2. 점역 소프트웨어 고도화
- 사업수행기관 | (사)한국시각장애인연합회





## 제 출 문

국립국어원장 귀하

국립국어원과 체결한 보조사업자 계약에 따라 '2023년 묵자-점자  
병렬말뭉치 구축' 사업 최종 보고서를 작성하여 제출합니다.

■ 사업기간: 2023년 8월 14일~2024년 2월 29일

2024년 2월

사업수행기관 (사)한국시각장애인연합회 말뭉치사업단

사 업 총 괄

이 연 주

사업수행인력

배진희, 정지혜, 천사랑

## 목 차 ...

### I. 목자-점자 병렬말뭉치 구축

#### 제1장 서론

1. 사업 배경 .....	3
2. 사업 목적 및 내용 범위 .....	4
3. 사업 방법 .....	4

#### 제2장 목자-점자 병렬말뭉치 데이터 구축 체계

1. 말뭉치 데이터 구축 단계 .....	7
2. 데이터 수집 .....	8
1) 원천 데이터 선정 .....	8
(1) 웹 기반 관리시스템 .....	9
2) 원천 데이터 수집 .....	13
(1) 말뭉치 분석 .....	13
(2) 1차 변환 목자 텍스트 파일 .....	16
(3) 목자 추출 시나리오 조건 .....	16
(4) 2차 변환 목자 파일 .....	18
3) 카테고리 구성 및 파일 명명 규칙 .....	19
4) 최종 적합 파일 유형 .....	19
3. 데이터 정제 .....	20
1) 부적합 목자 데이터 삭제 .....	20
2) 데이터 정제 .....	20
4. 데이터 가공 .....	21
1) MTPE .....	21
2) 3차 변환 목자-점자 말뭉치 파일 .....	22
(1) JAVA 객체 변환 .....	22
(2) 데이터베이스 등록 .....	23

5. 데이터 검사 .....	24
1) 검수 지침 .....	24
(1) 사이트 접속 및 로그인 .....	24
(2) 내비게이션 구성 .....	25
(3) A/B그룹 검수 화면 .....	26
(4) 적합/오류/보류 버튼 체크 .....	26
(5) 오류/보류 사유 기재 .....	26
(6) 오류/보류 점형 교정(A그룹) .....	27
(7) 작업 이력 보기 .....	28
(8) 작업 이력 보기 상세 .....	29
(9) C그룹 검수 화면 .....	32
(10) 적합/오류/보류 버튼 체크 및 사유 기재 .....	33
(11) 작업 이력 보기 .....	33

### 제3장 목자-점자 병렬말뭉치 검수 및 통계 현황

1. 검수 방법 .....	37
1) 검수 참여자 .....	37
2) 검수 단계 .....	39
(1) 1차 검수 단계 .....	40
(2) 2차 검수 단계 .....	41
3) 검수 교육 절차 .....	43
4) 자료 처리 .....	44
2. 조사 결과 .....	45
1) 1차 검수 데이터 통계 현황 .....	45
2) 2차 검수 데이터 통계 현황 .....	51
3) 최종 데이터 통계 현황 .....	54

## 제4장 묵자-점자 병렬말뭉치 적합/오류/보류 사례

1. 오류/보류 주요 유형 .....	57
2. 적합 사례(예시) .....	59
1) 시나리오 1: 한글 + (영어 + ... + ) .....	59
2) 시나리오 2: 영어 + (한글 + ... + ) .....	61
3) 시나리오 3: 한글 + (숫자 + ... + ) .....	63
4) 시나리오 4: 숫자 + (한글 + ... + ) .....	65
5) 시나리오 5: 영어 + (숫자 + ... + ) .....	67
3. 데이터 주요 오류/보류 사례 .....	68
1) 1급 점자표 오류 .....	68
2) 알파벳 약자/약어 오류 .....	71
3) 문장 부호 오류 .....	75
4) 단위/연산부호 오류 .....	78
5) 기타 부호 오류 .....	81
6) 영어 대문자 표기 오류 .....	83
7) 제2외국어/한자 표기 .....	86
8) 기타 .....	88
4. 데이터 주요 오류/보류 처리 .....	90

## 제5장 결론

1. 결론 .....	93
-------------	----

## II. 점역 소프트웨어 고도화

### 제6장 서론

1. 사업 배경 .....	97
2. 사업 목표 .....	99
3. 사업 방법 .....	100
1) 연구진 구성 .....	100
2) 문헌 조사 .....	100
3) 소프트웨어 기능 분석 .....	101
4) 연구·개발진 협의체 운영 .....	101
5) 객체지향 및 CBD 방식의 소프트웨어 개발 .....	101

### 제7장 점사랑 기능 개선을 위한 분석 및 설계

1. 점사랑 4.0 점역 엔진 및 편집기 기능 분석 .....	105
1) 점사랑 4.0 점역 엔진 기능 분석 .....	105
(1) 한글 점역 오류 분석 .....	105
(2) 영어 점역 오류 분석 .....	106
(3) 수식 점역 오류 분석 .....	108
2) 점사랑 4.0 편집기 기능 분석 .....	109
(1) 표 서식 기능 분석 .....	109
(2) 문서 최적화 기능 분석 .....	110
2. 점사랑 5.0 개발 방향 및 기능 개선 사항 .....	112
1) 점사랑 5.0 개발 방향 .....	112
2) 점사랑 5.0 기능 개선 사항 .....	112
(1) 열기 .....	113
(2) 점역/역점역 .....	113
(3) 문서 최적화 .....	114
(4) 표 .....	115
(5) 점자판 .....	115
(6) 점자 그림 .....	116

(7) 수식 .....	116
(8) 악보 점역 .....	117
(9) 점역 코드 .....	118
(10) 정렬 .....	118
(11) 스타일 설정 .....	118
(12) 템플릿 관리 .....	118

## 제8장 축각 그래픽 소프트웨어

1. 해외 축각 그래픽 소프트웨어 기능 분석 .....	121
1) QuickTac의 주요 특징 .....	121
(1) NimPro와 상호 운용 .....	121
(2) QuickTac의 파일 형식 .....	122
(3) 표준 그래픽 파일을 배경으로 활용 .....	122
(4) 점자 출력기의 호환성 .....	122
(5) 직관적인 사용자 가이드 .....	122
(6) 간편한 그리기 도구 .....	123
2) QuickTac의 메뉴 구성 .....	123
(1) 파일 메뉴 .....	123
(2) 편집 메뉴 .....	123
(3) 보기 메뉴 .....	123
(4) 도구 메뉴 .....	123
(5) 도움말 메뉴 .....	124
3) QuickTac의 동작 .....	125
(1) 파일 형식 .....	125
(2) DBT로 파일 불러오기 .....	125
(3) 배경 이미지 사용 .....	125
(4) 페이지 설정 및 점자 출력기 구성 .....	126
(5) 점자 출력기에 대한 참고 사항 .....	126
2. 점사량 5.0 축각 그래픽 기능 설계 .....	127
1) 축각 그래픽 소프트웨어의 주요 기능 .....	127
2) 축각 그래픽 소프트웨어의 기능 설계 .....	128

(1) 파일 메뉴 .....	128
(2) 편집 메뉴 .....	128
(3) 텍스트 메뉴 .....	129
(4) 이미지/도형 메뉴 .....	130
(5) 보기 메뉴 .....	131
(6) 설정 메뉴 .....	131
(7) 도움말 메뉴 .....	131

## 제9장 악보 점역 소프트웨어

1. 해외 악보 점역 소프트웨어 기능 분석 .....	135
1) GOODFEEL의 주요 특징 및 기능 .....	135
(1) GOODFEEL의 프로그램 구성 .....	135
(2) 각 구성 요소의 기능 .....	136
(3) 점역 옵션 .....	136
(4) 점역 처리 과정 .....	137
2) GOODFEEL의 메뉴 구성 .....	137
(1) File(파일) 메뉴 .....	137
(2) Options(옵션) 메뉴 .....	138
(3) Help(도움말) 메뉴 .....	138
2. 점사랑 5.0 악보 점역 기능 설계 .....	139
1) 악보 점역 소프트웨어의 주요 기능 .....	139
2) 악보 점역 소프트웨어의 기능 설계 .....	139
(1) 파일 메뉴 .....	140
(2) 옵션 메뉴 .....	141
(3) 도움말 메뉴 .....	141

## 제10장 소프트웨어 활용법에 대한 온라인 콘텐츠 개발

1. 온라인 콘텐츠의 필요성 .....	145
2. 온라인 콘텐츠의 구성 .....	145

1) 구성 방향 .....	145
2) 구성안 .....	145
(1) 1~4차시: 점사랑 5.0 소개 .....	146
(2) 5~8차시: 기본 편집 작업 .....	146
(3) 9~12차시: 고급 편집 기능 .....	147

## 제11장 점역 소프트웨어의 검증 및 평가

1. 점역 소프트웨어의 점역 결과 검증 .....	151
1) 점역 엔진 검증을 위한 데이터베이스 구축 .....	151
2) 데이터베이스 자료 입력 규칙 .....	151
(1) 공통 사항 .....	152
(2) 한글 점자 규정 .....	152
(3) 수학 점자 규정 .....	152
3) 점역 엔진 검증 프로그램 업데이트 .....	153
4) 점역 엔진의 검증 결과 .....	154
(1) 한글 모드의 점역 검증 결과 .....	154
(2) 영어 모드의 점역 검증 결과 .....	155
(3) 수학 모드의 점역 검증 결과 .....	156
5) 검증 결과에 대한 반영 .....	157
2. 점역 소프트웨어의 사용자 평가 .....	158
1) 사용자 평가단의 구성 .....	158
2) 사용자 평가 영역의 구성 .....	158
3) 사용자 평가 결과 .....	159
(1) 공통 영역의 평가 결과 .....	159
(2) 접근성 영역의 평가 결과 .....	161
(3) 점사랑 메뉴 영역의 평가 결과 .....	162
(4) 점역 코드 영역의 평가 결과 .....	166
(5) 촉각 그래픽 메뉴 영역의 평가 결과 .....	168
3. 검증 및 평가에 대한 시사점 .....	169

제12장 결론 및 제언

1. 결론 ..... 173

2. 제언 ..... 176

참고문헌

## 표 목 차 ...

〈표 2-1〉 데이터 구축 단계(세부 절차 설명) -----	7
〈표 2-2〉 데이터 수집 단계 -----	8
〈표 2-3〉 신문 말뭉치 JSON 파일 예시 -----	14
〈표 2-4〉 온라인 대화 말뭉치 JSON 파일 예시 -----	15
〈표 2-5〉 1차 변환 단계 예시 -----	16
〈표 2-6〉 시나리오 조건 -----	17
〈표 2-7〉 말뭉치 종류별 추출 문장 수 -----	18
〈표 2-8〉 시나리오별 추출 문장 수 -----	18
〈표 2-9〉 데이터 카테고리 -----	19
〈표 2-10〉 최종 데이터 적합 엑셀 파일 유형 -----	19
〈표 2-11〉 데이터 정제 수행 과정 -----	20
〈표 2-12〉 가공 세부 절차 -----	21
〈표 2-13〉 가공 작성 예시 -----	22
〈표 2-14〉 데이터 검사 세부 절차 -----	24
〈표 3-1〉 검수 참여자 정보(64명) -----	37
〈표 3-2〉 1차 검수자(A그룹, C그룹) 정보(37명) -----	40
〈표 3-3〉 2차 검수자(B그룹) 정보(27명) -----	42
〈표 3-4〉 검수 일정 -----	43
〈표 3-5〉 A그룹 검수 데이터 통계 현황 -----	45
〈표 3-6〉 C그룹 검수 데이터 통계 현황 -----	48
〈표 3-7〉 B그룹 검수 데이터 통계 현황 -----	51
〈표 3-8〉 최종 데이터 통계 현황 -----	54
〈표 4-1〉 A그룹, B그룹 데이터 오류/보류 주요 유형 -----	57
〈표 4-2〉 시나리오 1(한글+(영어)) 적합 사례 ① -----	59
〈표 4-3〉 시나리오 1(한글+(영어)) 적합 사례 ② -----	60
〈표 4-4〉 시나리오 2(영어+(한글)) 적합 사례 ① -----	61
〈표 4-5〉 시나리오 2(영어+(한글)) 적합 사례 ② -----	62
〈표 4-6〉 시나리오 3(한글+(숫자)) 적합 사례 ① -----	63
〈표 4-7〉 시나리오 3(한글+(숫자)) 적합 사례 ② -----	64
〈표 4-8〉 시나리오 4(숫자+(한글)) 적합 사례 ① -----	65
〈표 4-9〉 시나리오 4(숫자+(한글)) 적합 사례 ② -----	66

〈표 4-10〉 시나리오 5(영어+(숫자)) 적합 사례	67
〈표 4-11〉 알파벳 R 1급 점자표 오류 사례	68
〈표 4-12〉 알파벳 B 1급 점자표 오류 사례	69
〈표 4-13〉 알파벳 H, E, R 1급 점자표 오류 사례	70
〈표 4-14〉 AS 약자 오류	71
〈표 4-15〉 EDM[Electronic Dance Music(전자 무용 음악)] 오류	72
〈표 4-16〉 MOU[Memorandum of Understanding(양해각서)] 오류	73
〈표 4-17〉 AND 오류	74
〈표 4-18〉 다른 문장부호 표기 오류	75
〈표 4-19〉 물결표 오류	76
〈표 4-20〉 홑화살괄호 오류	77
〈표 4-21〉 로마자가 아닌 단위 부호 점역 오류	78
〈표 4-22〉 부피 단위(리터(ℓ)) 유니코드 점역 오류	79
〈표 4-23〉 곱셈기호 알파벳 표기 오류	80
〈표 4-24〉 화학기호 점역 오류/보류	80
〈표 4-25〉 연산 기호 점형 오류 ①	81
〈표 4-26〉 연산 기호 점형 오류 ②	82
〈표 4-27〉 영어 대문자 표기 오류 ①	83
〈표 4-28〉 영어 대문자 표기 오류 ②	84
〈표 4-29〉 영어 대문자 표기 오류/보류	85
〈표 4-30〉 영어 외 외국어 사례 ①	86
〈표 4-31〉 영어 외 외국어 사례 ②	87
〈표 4-32〉 한글 점자 규정 외 기타 기호 보류 ①	88
〈표 4-33〉 한글 점자 규정 외 기타 기호 보류 ②	88
〈표 4-34〉 이모티콘 문장 기호 보류 ①	89
〈표 4-35〉 이모티콘 문장 기호 보류 ②	89
〈표 5-1〉 JSON 파일 작성 예시	94
〈표 7-1〉 한글 점역 오류 사례 분석	105
〈표 7-2〉 영어 점역 오류 사례 분석	106
〈표 7-3〉 수식 점역 오류 사례 분석 ①	108
〈표 7-4〉 수식 점역 오류 사례 분석 ②	108
〈표 7-5〉 점사량 5.0 기능 개선 사항	113
〈표 8-1〉 파일 메뉴 기능 설계	128
〈표 8-2〉 편집 메뉴 기능 설계	129

〈표 8-3〉 텍스트 메뉴 기능 설계	129
〈표 8-4〉 이미지/도형 메뉴 기능 설계	130
〈표 8-5〉 보기 메뉴 기능 설계	131
〈표 8-6〉 설정 메뉴 기능 설계	131
〈표 8-7〉 도움말 메뉴 기능 설계	131
〈표 9-1〉 파일 메뉴 기능 설계	140
〈표 9-2〉 옵션 메뉴 기능 설계	141
〈표 11-1〉 점역 엔진 검증을 위한 데이터베이스 구성	151
〈표 11-2〉 한글 모드 점역의 오류 유형	155
〈표 11-3〉 영어 모드 점역의 오류 유형	155
〈표 11-4〉 수학 모드 점역의 오류 유형	156
〈표 11-5〉 사용자 평가단의 구성	158
〈표 11-6〉 사용자 평가 영역의 구성	159
〈표 11-7〉 공통 영역의 평가 결과	160
〈표 11-8〉 접근성 영역의 평가 결과	161
〈표 11-9〉 파일 메뉴의 평가 결과	162
〈표 11-10〉 편집 메뉴의 평가 결과	163
〈표 11-11〉 보기 메뉴의 평가 결과	164
〈표 11-12〉 삽입 메뉴의 평가 결과	164
〈표 11-13〉 문서 메뉴의 평가 결과	165
〈표 11-14〉 설정 메뉴의 평가 결과	166
〈표 11-15〉 점역 코드의 평가 결과	166
〈표 11-16〉 촉각 그래픽 메뉴의 평가 결과	168

## 그림 목 차 ...

[그림 2-1] 말뚝치 웹 기반 관리시스템 메인 화면	9
[그림 2-2] 말뚝치 검수자 검수 화면	10
[그림 2-3] 말뚝치 검수자 교정 화면	10
[그림 2-4] 말뚝치 검수자 작업 내역 보기 화면	11
[그림 2-5] 기존 오류·보류 말뚝치 문장 재판정 화면	11
[그림 2-6] 말뚝치 웹 기반 관리시스템 검수 관리 화면	12
[그림 2-7] 말뚝치 웹 기반 관리시스템 문장 관리 화면	12
[그림 2-8] 말뚝치 웹 기반 관리시스템 검수 문장 비교 화면	13
[그림 2-9] JAVA 객체 변환 화면	23
[그림 2-10] 데이터베이스 예시	23
[그림 2-11] 말뚝치 관리시스템 사이트 접속 로그인 화면	25
[그림 2-12] 말뚝치 관리시스템 내비게이션 화면	25
[그림 2-13] 말뚝치 관리시스템 A/B그룹 목자-점자 검수 화면	26
[그림 2-14] 적합/오류/보류 버튼 및 오류/보류 버튼 클릭 시 텍스트 박스 화면	27
[그림 2-15] 오류/보류 검수 화면에서 수정 전 선택 화면	28
[그림 2-16] 점자 점형 교정 새 창 화면	28
[그림 2-17] 작업 이력 보기 화면	29
[그림 2-18] 작업 이력 보기 클릭 후 화면	29
[그림 2-19] 주차별 검수 문장 분류 메뉴	30
[그림 2-20] 판정별 검수 문장 분류 메뉴	30
[그림 2-21] 수정 이력별 검수 문장 분류 메뉴	31
[그림 2-22] 작업 이력 화면 내 수정 시 화면	31
[그림 2-23] B그룹이 문장 판정하여 수정 불가능한 화면	32
[그림 2-24] 말뚝치 관리시스템 C그룹 목자-점자 검수 화면	32
[그림 2-25] C그룹 작업 이력 보기 화면	33
[그림 2-26] C그룹 판정 상태 콤보박스	33
[그림 3-1] A그룹 검수 적합/오류/보류 문장 수	46
[그림 3-2] A그룹 검수 주차별 검수 문장 수 및 누적 검수율	46
[그림 3-3] A그룹 검수 주차별 적합/오류/보류 문장 수	47
[그림 3-4] A그룹 검수 주차별 최종 적합 문장 수 및 적합률	47
[그림 3-5] C그룹 검수 적합/오류/보류 문장 수	49
[그림 3-6] C그룹 검수 주차별 검수 문장 수 및 누적 검수율	49

[그림 3-7] C그룹 검수 주차별 적합/오류/보류 문장 수	50
[그림 3-8] C그룹 검수 주차별 최종 적합 문장 수 및 적합률	50
[그림 3-9] B그룹 검수 적합/오류/보류 문장 수	52
[그림 3-10] B그룹 검수 주차별 검수 문장 수 및 누적 검수율	52
[그림 3-11] B그룹 검수 주차별 적합/오류/보류 문장 수	53
[그림 3-12] B그룹 검수 주차별 최종 적합 문장 수 및 적합률	53
[그림 4-1] A그룹 데이터 오류/보류 주요 유형	58
[그림 4-2] B그룹 데이터 오류/보류 주요 유형	58
[그림 4-3] 검수자 소통 자유게시판	90
[그림 7-1] 점역 옵션 설정 메뉴의 신규 규정 선택 화면	114
[그림 7-2] 문서 최적화 메뉴의 자동 교열 기능 화면	115
[그림 7-3] 점사랑 수식편집기 화면	116
[그림 7-4] 점사랑에서 악보 삽입 시 악보 선택 화면	117
[그림 7-5] 선택한 악보의 자동 점역 결과 화면	117
[그림 8-1] QuickTac 프로그램 화면	121
[그림 8-2] QuickTac의 도형 생성 구현 화면	124
[그림 8-3] 점사랑 5.0 측각 그래픽의 도형 생성 구현 화면	127
[그림 8-4] 점자 텍스트가 포함된 도형 생성 구현 화면	130
[그림 9-1] GOODFEEL 프로그램 화면	136
[그림 9-2] GOODFEEL의 점역 옵션 설정 화면	138
[그림 9-3] Lime 프로그램의 한글 입력 화면	140
[그림 10-1] 동영상 콘텐츠 1차시 화면	148
[그림 11-1] 점역 엔진 검증 프로그램의 화면 구성	153
[그림 11-2] 한글 모드의 점역 검증 결과 화면	154
[그림 11-3] 수학 모드의 점역 검증 결과 화면	156
[그림 11-4] 점자 종합정보 누리집의 점사랑 문의 게시판 화면	157

2023년 목자-점자 병렬말뭉치 구축  
(1. 목자-점자 병렬말뭉치 구축)

# 제 1 장

## 서 론

1. 사업 배경
2. 사업 목적 및 내용 범위
3. 사업 방법





## 1 사업 배경

점자는 그 자체가 디지털 매체이므로 디지털 기술과 접목하는데 매우 유리한 매체이다. 디지털 기술이 확산되면 점자가 굳이 필요하지 않을 수도 있다고 예상하는 사람도 일부 있었지만 이러한 예상과는 달리 점자와 접목시킨 정보통신 기술, 다양한 응용 소프트웨어가 지속적으로 개발·보급되고 있다. 각종 점자 번역 소프트웨어와 점자 정보 단말기를 접목한 응용 소프트웨어, 인터넷 웹 환경에서 사용할 수 있는 전자 점자 및 스마트 기기용 애플리케이션 등 종류와 분야도 확대되고 있다.

점역 소프트웨어를 사용해 점자 자료를 제작하면 생산성을 크게 높일 수 있다. 먼저 어문 중심 자료인 경우 점 입력보다 속도가 빠른 워드 입력을 통해 내용을 컴퓨터 파일로 제작한 다음 점역 소프트웨어를 사용해 점자 자료로 변환하므로 제작 기간을 현저히 줄일 수 있고, 기본적으로 워드 입력이 가능한 사람이면 누구나 점자 자료 제작을 할 수 있으므로 필요한 인력을 쉽게 충원할 수 있다. 또한 점자 자료 편집 과정 역시 디지털화되므로 내용의 수정이나 변경이 쉽고 점자 자료 형태 역시 적은 노력으로 변경할 수 있어 사용자의 필요에 맞는 자료 제공이 가능하다.

21세기 지식정보화 사회에서 점자가 독서와 교육 및 고용 등을 위한 정보 접근 매체에만 국한될 필요는 없다. 시각장애인이 일상생활을 영위하고 사회 활동에 참여하며 문화생활을 향유하는 데 자유롭고 편리하게 점자를 사용할 수 있어야 한다.

이를 해결하고자 일상생활에서 시각장애인의 점자 사용 및 학습 편의를 도모하기 위한 점자 정보화 기반을 구축하는 것이 이 사업의 목적이다. 목자-점자 병렬말뭉치 구축으로 점자 사용 환경 개선 방안을 제시하고자 한다.

본 사업을 통해 문자언어 사용에 어려움을 겪는 시각장애인의 불편을 해소할 실질적인 방안을 제시할 수 있다. 이는 지식정보화 시대를 살아가는 시각장애인들이 점자를 쉽고 편리하게 사용할 수 있는 환경을 조성하고, 비장애인과 통합된 사회 구성원으로서 문자 생활을 영위하는 데 기여할 것이다.

## 2 사업 목적 및 내용 범위

이 사업의 목적은 말뭉치 구축 및 안정된 점역 소프트웨어를 개발하여 일상생활에서 시각장애인의 점자 사용 및 학습 편의를 증진하고, 향후 점자 친화적인 생활환경 조성을 위한 개선 방안을 제시하는 것이다. 이러한 사업 목적을 달성하기 위해 설정한 사업 목표는 다음과 같다.

첫째, 목자-점자 병렬말뭉치 데이터 구축 방안을 제시한다. 한국어·영어 혼용 자료(목자-점자 일대다 대응 자료 포함) 확보, 목자-점자 대응 문장 총 십만 쌍(백만 어절 이상)을 구축하였다.

둘째, 목자-점자 병렬말뭉치 검수 및 통계 방안을 제시한다. 말뭉치 구축 전체 과정을 체계적으로 진행하기 위한 웹 기반 관리시스템 활용 방안을 제시하였다. 오류 방지를 위해 데이터 집계 특성에 따른 통계 분석을 시행하였다.

셋째, 목자-점자 병렬말뭉치 오류 및 보류 사례를 제시한다. 검수 과정에서 발견된 오류와 보류 사례는 점자 규정과 연동되어야 하는 부분이 많아 말뭉치 사업 전반에 점자 규범 정비 및 연구 위원회의 활동과 연계가 필요하다. 이를 위해 점자 규정 및 관련 지침 준수 여부를 분석하는 데 기초 자료로 활용할 뿐 아니라 사업 효과를 평가하는 실증적인 자료를 확보하고자 하였다.

## 3 사업 방법

이 사업의 목표를 달성하기 위해 사용한 사업 방법은 다음과 같다.

첫째, 말뭉치 대상 데이터 자료 및 구축 방식이다. 저작권 문제를 해결하고, 개인정보, 상품정보, 기업정보 등 민감정보 비식별화를 위해 국립국어원 말뭉치 신문 분야에서 1개 이상의 단어가 중복되지 않은 십만 문장(백만 어절 이상)을 추출하였다. 또한 시나리오 5종 분야로 세분화하여 자료를 검출하였다. 1개 이상의 영어 단어가 포함된 문장을 추출하였고, 통일 영어 점자 규정에서 정하고 있는 규정 적용 사례 위주로 추출하였다.

둘째, 목자-점자 병렬말뭉치 구축 체계 및 검수 방안이다. 목자-점자 병렬말뭉치 구축을 위한 웹 기반 관리시스템과 구축된 데이터를 활용하는 방법을 제시하였다. 신문과 온라인 대화를 데이터셋으로 구축하여 향후 점역 엔진 학습 데이터로써 활용될 수 있게 하는 것을 목적으로 구축하였다. 고품질의 목자-점자 병렬말뭉치 데이터셋을 구축하기 위한 데이터 수집, 정제, 가공, 검사 단계까지 고려해야 할 점을 제시하였다.

2023년 목자-점자 병렬말뭉치 구축  
(1. 목자-점자 병렬말뭉치 구축)

## 제 2 장

# 목자-점자 병렬말뭉치 데이터 구축 체계

1. 말뭉치 데이터 구축 단계
2. 데이터 수집
3. 데이터 정제
4. 데이터 가공
5. 데이터 검사





## 1 말뭉치 데이터 구축 단계

목자-점자 병렬말뭉치 구축을 위한 프로세스와 데이터를 각 단계별로 잘 수행하기 위해서 크기는 5단계로 나누어 서술하였다. 고품질의 목자-점자 병렬말뭉치 데이터세트를 구축하기 위한 데이터 수집, 정제, 가공, 검사 단계까지의 고려해야 할 점과 이에 대한 예시를 설명한다.

표 2-1 데이터 구축 단계(세부 절차 설명)

구축 단계	설 명
1단계 수집	<ul style="list-style-type: none"> <li>- 원천 데이터 선정: 원천 데이터를 수집하는 대상 선정(저작권 확인, 문장 형식 검토)</li> <li>- 원천 데이터 수집: 텍스트(문장) 형태의 원천 데이터 수집</li> </ul>
2단계 정제	<ul style="list-style-type: none"> <li>- 부적합 데이터 삭제: 수집 과정에서 형식 오류가 생긴 문장 삭제</li> <li>- 데이터 정제: 목적에 맞게 데이터 수정</li> </ul>
3단계 가공	<ul style="list-style-type: none"> <li>- MTPE 또는 점역: 기계 번역을 활용하여 1차 점역으로 1차 점역물 생성</li> </ul>
4단계 검사	<ul style="list-style-type: none"> <li>- 1차 검수: 전체 목자-점자 문장 평가</li> <li>- 2차 검수: 1차 적합 문장 평가/검증을 위해 별도 검수자가 교차 검수</li> <li>- 말뭉치 교정: 기존 구축 목자-점자 병렬말뭉치 오류·보류 문장 재판정 및 오류 문장 교정</li> <li>- 검수 지침 필요</li> </ul>

※ MTPE(Machine Translation Post-Editing): 기계 번역 포스트 에디팅, 기계가 자동으로 번역한 점역문을 사람이 검수하는 과정

## 2 데이터 수집

수집은 크게 원천 데이터 선정 및 원천 데이터 수집 2단계로 나뉜다.

표 2-2 데이터 수집 단계

세부 절차	작업
1. 원천 데이터 선정	<ul style="list-style-type: none"> <li>- 저작권 확인</li> <li>- 문장 적절성: 분야, 길이, 분량</li> <li>- 기술 문제 검토: 수집 작업 적절성</li> </ul>
2. 원천 데이터 수집	<ul style="list-style-type: none"> <li>- 수집 방법 결정</li> <li>- 수집 기준 설정: 문장 분절, 길이, 시나리오 조건 선정, 기호 처리 등</li> <li>- 메타데이터 결정</li> <li>- 1차 변환: JSON에서 목자 문장 추출 과정</li> <li>- 2차 변환: 정의된 시나리오 5종에 해당하는 목자 문장 추출 과정</li> </ul>

※ JSON(Java Script Object Notation): 데이터를 저장하거나 전송할 때 많이 사용되는 경량의 DATA 교환 형식으로 Javascript에서 객체를 만들 때 사용하는 표현식을 의미

### 1) 원천 데이터 선정

원천 데이터를 선정할 때 가장 먼저 저작권 문제와 개인정보, 상품정보, 기업정보 등 민감정보 비식별화에 맞춰 데이터 분야를 정했다. 적절한 문장이 있는 데이터원을 국립국어원 온라인의 ‘모두의 말뭉치’ 누리집에 탑재된 말뭉치 원천 데이터(JSON: Java Script Object Notation)를 제공받아 저작권 문제를 해결하고, 개인정보, 상품정보, 기업정보 등 민감정보 비식별화를 확인했다.

국립국어원에서 제공받은 원천 데이터 자료는 국회 회의록 말뭉치 2021, 맞춤법 교정 말뭉치 2022, 신문 말뭉치 2022, 온라인 대화 말뭉치 2022 등 총 4개 원천 데이터였다. 본 사업에서는 신문 말뭉치 2022, 온라인 대화 말뭉치 2022를 원천 데이터 자료로 선정하였다.

해당 데이터원에서 문장의 길이, 분량을 얻을 수 있는지 확인하고, 부족할 경우 추가 데이터원을 선정하였다. 이때 분량은 구축 목표량의 1.2배 이상 확보하여 이후 가공 단계에서의 어려움을 줄이고자 하였다.

선정 시 수집 기술도 같이 검토되었다. 예를 들면, 원천 데이터가 JSON의 형태로서 해당 파일을 가공하거나, 추출된 한글·영어 혼용 말뭉치와 이를 점역한 점자 문장을 병렬말뭉치로 구축하여 관리하도록 웹 기반 관리시스템을 도입하였다.

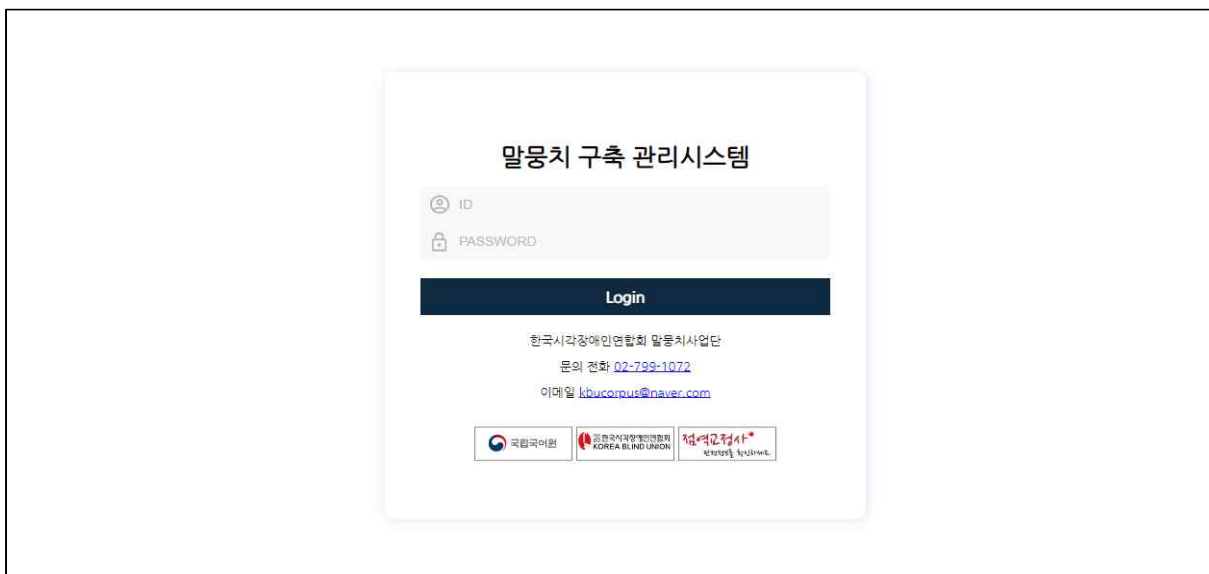
기존 구축 목자-점자 병렬말뭉치 오류·보류 문장의 경우 2021~2022년 목자-점자 병렬말뭉치 구축 결과물을 이미 보유하고 있어 2021~2022년 목자-점자 병렬말뭉치 오류·보류 해당 자료로 선정하였고, 마찬가지로 병렬말뭉치로 구축하여 관리하도록 웹 기반 관리시스템을 도입하였다.

## (1) 웹 기반 관리시스템

### ① 검수 진행

웹 기반 관리시스템의 사용자는 관리자, 검수자 두 가지 종류로 이루어진다. 관리자는 국립국어원 말뭉치 파일(JSON 형식)을 내려받아 시스템에 등록한다. 말뭉치 등록 시 점역 엔진(㈜ACNS 자사 서버 기반 점역 엔진)을 사용하여 추출된 문장과 점역 결과가 함께 저장된다. 등록된 말뭉치는 특수 기호, 영문, 숫자 등의 포함 여부 등의 정보와 함께 저장되고, 문장 수 정보도 함께 저장된다.

관리자는 검수자의 계정을 생성하거나 삭제·변경할 수 있으며, 검수자 계정 생성 시 작업량 균등 분배를 위해 처리해야 할 문장 수를 지정한다. 검수자는 관리시스템에 접속하여 자신에게 할당된 문장을 적합/오류/보류로 처리한다. 이때 검수자에게 할당된 문장은 미리 할당하는 방식이 아니라 자동으로 시스템에서 배정하는 방식으로, 접속 시점에 아직 검수되지 않은 문장이 자동 배정된다.



**그림 2-1** 말뭉치 웹 기반 관리시스템 메인 화면

추출된 목자-점자 병렬말뭉치 데이터는 웹 기반의 시스템을 통하여 통합 관리되며 문제 풀이 형식 기반의 시스템을 통하여 검수자에게 제공된다. 검수자에게 할당된 데이터는 웹 기반의 검수 시스템을 통하여 검수되고 있으며 검수자의 작업 기록을 남겨 작업의 신뢰도를 높이고 있다.

검수자는 웹 기반 관리시스템에 자신의 계정으로 접속하여 자신에게 할당된 문장을 문제 풀이 형식으로 검수를 진행한다. 이때 웹 기반 관리시스템은 검수자에게 검수 진행도를 표시하여 검수자 스스로 작업량을 조절하고, 무리한 일정 내에 처리되지 않도록 한다. 웹 기반 관리시스템은 시각장애인 검수자를 고려하여 웹 접근성 기준에 맞추어 개발되었다.

그림 2-2 말뭉치 검수자 검수 화면

그림 2-3 말뭉치 검수자 교정 화면

검수자의 검수 결과는 주 단위로 관리되며 검수자 그룹 교차 검수를 통하여 검수의 신뢰도를 높일 수 있도록 처리하였다. 또한 검수자 스스로도 이전 검수 결과를 조회하고, 검수 결과를 수정할 수 있도록 내역 보기 기능을 제공하고 있다. 기존 구축 목자-점자 병렬말뭉치 오류·보류 문장 재판정 및 오류 문장 교정 또한 동일한 방법으로 진행한다.

그림 2-4 말뭉치 검수자 작업 내역 보기 화면

그림 2-5 기존 오류·보류 말뭉치 문장 재판정 화면

## ② 검수 관리

검수의 신뢰도를 높이하고자 검수자들의 작업 현황을 파악하여 너무 높은 검수 오류율과 떨어지는 작업 속도 등을 체크하고, 원활한 검수 작업이 진행될 수 있도록 구축하였다. 또한 검수자 회원 관리 시스템을 도입하여 작업 시작 중에 개인 사정으로 작업을 더 이상 진행하지 못하는 검수자나 추가되는 검수자를 손쉽게 관리할 수 있도록 하였다.

말뭉치 내역 현황

검수작업

2023년 말뭉치 사업 ▾ 2차 검수자(B) ▾

주차

전체 ▾

전체 처리 현황

엑셀다운로드

주차	기간	총 검수수 (135,000문장)	누적 검수율(%)	검수결과									총 어절수	총 문장수
				적합				오류	보류	계	오류율(%)	보류율(%)		
				적합	적합율(%)	어절수	문장수							
1 주차	20231030 ~ 20231105	3	0%	3	100%	50	3	0	0	3	0%	0%	50	3
2 주차	20231106 ~ 20231112	12729	9.43%	12,173	95.63%	236,797	12,173	433	123	12729	3.4%	0.97%	247,009	12729
3 주차	20231113 ~ 20231119	13537	10.03%	13,027	96.23%	266,268	13,027	389	121	13537	2.87%	0.89%	277,431	13537
4 주차	20231120 ~ 20231126	13926	10.32%	13,637	97.92%	274,313	13,637	210	79	13926	1.51%	0.57%	280,459	13926
5 주차	20231127 ~ 20231203	13303	9.85%	12,909	97.04%	262,531	12,909	249	145	13303	1.87%	1.09%	270,610	13303
6 주차	20231204 ~ 20231210	13932	10.32%	13,518	97.03%	277,357	13,518	332	82	13932	2.38%	0.59%	286,096	13932
7 주차	20231211 ~ 20231217	13849	10.26%	13,505	97.52%	270,937	13,505	281	63	13849	2.03%	0.45%	277,989	13849
8 주차	20231218 ~ 20231224	13158	9.75%	12,908	98.1%	277,928	12,908	212	38	13158	1.61%	0.29%	283,535	13158
9 주차	20231225 ~ 20231231	13798	10.22%	13,554	98.23%	278,468	13,554	205	39	13798	1.49%	0.28%	283,728	13798
10 주차	20240101 ~ 20240107	14493	10.74%	14,244	98.28%	291,828	14,244	204	45	14493	1.41%	0.31%	297,266	14493
11 주차	20240108 ~ 20240114	12272	9.09%	12,031	98.04%	249,476	12,031	204	37	12272	1.66%	0.3%	254,579	12272
합계		135,000	100.00%	131,509	97.41%	2,685,953	131,509	2,719	772	135,000	2.01%	0.57%	2,758,752	135,000

그림 2-6 말뭉치 웹 기반 관리시스템 검수 관리 화면

## ③ 문장 관리

웹 기반 관리시스템에서는 현재까지 검수된 문장들의 목록을 조회할 수 있다. 이 자료는 데이터베이스에 저장되어 있으며 JSON, Excel 등 다양한 포맷으로 제공할 수 있다.

말뭉치 검수데이터 내역					
정렬기준: 처리일시(오름차순) ▼					
보정작업 액셀 다운로드					
말뭉치 액셀 다운로드					
말뭉치 JSON 다운로드					
No.	ID	검수자	사유	사나리오	내용
83539	<a href="#">KBUICORPUS0083528</a>	kbu23a1273 한혜림	결과: 적합 사유: ()	한글(영어)	문 대통령은 이날 오후 자신의 페이스북 등 사회관계망서비스(SNS)에 '6·10민주항쟁과 관련한 '영웅, 민주인권기념관'을 착공합니다'란 제목의 글을 올렸다.
83538	<a href="#">KBUICORPUS0083527</a>	kbu23a1273 한혜림	결과: 적합 사유: ()	한글(영어)	도는 10일 논산시청 회의실에서 논산시, 환경부, 롯데건설, 현대차증권과 '논산시 수소경제도시 전환을 위한 업무 협약(MOU)'을 했다.
83537	<a href="#">KBUICORPUS0083526</a>	kbu23a1273 한혜림	결과: 적합 사유: ()	한글(영어)	주요 7개국(G7) 정상들이 내년까지 최소 10억회(트스) 분량의 코로나19 백신을 전세계에 공급, 팬데믹(전염병의 세계적 대유행)을 끝내겠다고 밝힐 계획이다.
국내 일반지주회사들이 총 55조원이 넘는 현금 및 현금성자산을 갖고 있는 것으로 나타났다. 개정된 공경거래법에 따라 일반지주회사들은 연말부터 기업형벤처캐피탈(VC)을 보유할 수 있어 활발한 벤처투자가 기대된다는 분석이다.					

그림 2-7 말뭉치 웹 기반 관리시스템 문장 관리 화면

## ④ 검수 문장 비교

검수자 간 교차 검수 시 서로 판정이 다를 경우 판정 결과를 조회할 수 있다. 비교 조회 시 말뭉치 원본과 점형을 수정한 버전을 비교할 수 있다.

검수 결과 비교 조회				
검수작업	2023년 말뭉치 사업	비교 검수 차수	기준 차수 : 1차 검수자(A) / 비교 차수 : 2차 검수자(B)	조회
검수 결과 비교 내역				
No.	검수자ID	판정결과	사유	내용
1	기준 : kbu23a5797 비교 : kbu23b5982	기준 : 오류 비교 : 적합	기준 : 문장번호 오류로문 일부 다듬어줄 피아싱 비교 :	원제 방: 5층 옥탑(일베X) / 1.5층(약10평) / 반려전X(키우는중) / 역 앞에서 5분 / 물음선(인터넷, 에어컨, 냉장고, 불박이웃장(한쪽면이 다) / 월세 1000/ 55 / 매달 약 65정도 나감 원본 : 판정 : 이사갈 방: 5층 (일베X) / 2층(약12평) / 반려전X(키우는중) / 역 앞에서 15분 / 물음선(에어컨, 인터넷, 냉장고, 세탁기X(냉장고, 세탁기는 넣어준다)고 함), 화장) / 전세 2.4억 / 매달 약 75 내외 로 나올거같음. 원본 : 판정 :
2	기준 : kbu23a5670 비교 : kbu23b5158	기준 : 오류 비교 : 적합	기준 : 기타(직접입력)임 표 후 띄어쓰기 수정 비교 :	원본 : 판정 :
조회 : 4,682건				

그림 2-8 말뭉치 웹 기반 관리시스템 검수 문장 비교 화면

## 2) 원천 데이터 수집

원천 데이터가 선정되면 데이터를 수집하게 된다. 먼저 원천 데이터 형태에 따라 수집 방법을 결정한다. 예를 들어 JSON 파일에서 Java 기반의 Batch 프로그램(Batch Processing: 컴퓨터 프로그램 흐름에 따라 순차적으로 자료를 처리하는 방식)으로 목자 문장을 추출하는 방법을 결정했으며, 수집에 필요한 기준은 문장의 길이와 여러 가지 기호의 통일 방법 등이 있다. 데이터 구축 목적에 따라 분야별 메타데이터를 정의하였다. 이러한 기준에 맞춰 메타데이터와 함께 수집된 문장은 정제 전 데이터베이스에 저장된다. 단, 기존 구축 목자-점자 병렬말뭉치의 오류·보류 문장은 원천 데이터 수집 과정을 생략한다.

### (1) 말뭉치 분석

언어의 정확성, 영향력, 사용 빈도 등을 고려하여 국립국어원의 신문 말뭉치 2022와 온라인 대화 말뭉치 2022를 기본 데이터 문장으로 선정하였다. 원문 말뭉치 분량은 신문 말뭉치 2022는 2021년에 생산된 신문 기사 978,342건, 온라인 대화 말뭉치 2022는 2022년에 생산된 게시판(네이버 카페, 네이버판, 다음 카페, 디시인사이드, 루리웹 등) 게시글 8,282건과 누리 소통망(인스타그램, 페이스북) 게시글 296,892건이다.

신문 말뭉치 2022와 온라인 대화 말뭉치 2022 파일 기존 구조는 JSON으로 처리되어 있으며 신문 말뭉치는 <표 2-3>, 온라인 대화 말뭉치는 <표 2-4>와 같은 구조를 가진다.

표 2-3 신문 말뭉치 JSON 파일 예시

```

{
  "id": "NZRW22000000001",
  "metadata": {
    "title": "국립국어원 신문 말뭉치 NZRW22000000001",
    "creator": "국립국어원",
    "distributor": "국립국어원",
    "year": "2022",
    "category": "신문 > 기타",
    "annotation_level": [
      "원시"
    ],
    "sampling": "본문 전체"
  },
  "document": [
    {
      "id": "NZRW22000000001.15",
      "metadata": {
        "title": "미디어오늘 2021년 기사",
        "author": "정민경 기자",
        "publisher": "미디어오늘",
        "date": "20210108",
        "topic": "사회",
        "original_topic": "사회"
      },
      "paragraph": [
        {
          "id": "NZRW22000000001.15.2",
          "form": "지난 6일 오후부터 내린 폭설 여파로 택배나 배달 업무가 차질을 빚는 가운데, 신문 배달도 난항을 겪고 있다."
        },
        .....
      ]
    },
    .....
  ]
}

```

표 2-4 온라인 대화 말뭉치 JSON 파일 예시

```

{
  "id": "EPRW2200000744",
  "metadata": {
    "title": "국립국어원 온라인 게시 자료 원시 말뭉치 EPRW2200000744",
    "creator": "국립국어원",
    "distributor": "국립국어원",
    "year": "2022",
    "category": "웹 > 게시판",
    "annotation_level": [
      "원시"
    ],
    "sampling": "게시자 모집 후 무작위 추출"
  },
  "document": [
    {
      "id": "EPRW2200000744.1",
      "metadata": {
        "title": "다이어트",
        "author": "writer744",
        "publisher": "네이트판",
        "date": "20220829",
        "topic": "생활/쇼핑_일상생활",
        "crawl_date": "20220906 15:09:05",
        "url": "https://pann.nate.com/talk/367421445"
      },
      "paragraph": [
        {
          "id": "EPRW2200000744.1.1",
          "form": "다이어트 하는 방법 좀 알려주세요"
        }
        .....
      ]
    }
    .....
  ]
}

```

말뭉치 파일 제목과 작성자, 생성 연도 등의 메타 데이터가 헤더 형식으로 들어 있고, JSON 파일 속성에 1개의 신문 기사 또는 게시글을 적절한 크기로 나눈 문장의 배열로 구성되어 있다.

## (2) 1차 변환 목자 텍스트 파일

분석이 끝난 전체 저장된 말뭉치 문장을 대량의 기계 점역이 가능한 형태의 목자 문장 텍스트 파일로 1차 변환한다. 목자 문장 데이터 추출은 Java 기반의 Batch 프로그램으로 작업하였다. 카테고리 출처 등의 기본 정보는 원본 국립국어원 말뭉치에서 확인이 가능하기 때문에 문장 식별자(ex: NZRW2200000001.15.2)와 문장(ex: 지난 6일 오후부터 내린 폭설 여파로 택배나 배달 업무가 차질을 빚는 가운데, 신문 배달도 난항을 겪고 있다.)을 단일 텍스트 파일로 저장한다.

이 과정에서 가운뎃점 등 눈으로 보기에는 동일하나 유니코드에서 다르게 처리되는 경우 전처리를 수행하여 기계 점역에서 오류가 발생하지 않도록 처리 한다. 저장 포맷은 1차 변환하여 중간단계로 활용하는 용도로서 단순 텍스트 파일로 저장한다.

문장 식별자와 문장으로 구성된 1차 변환 예시는 <표 2-5>과 같다.

표 2-5 1차 변환 단계 예시

NZRW2200000001.15.1:폭설로 신문 배달 차질에 “사과드린다” NZRW2200000001.15.2:지난 6일 오후부터 내린 폭설 여파로 택배나 배달 업무가 차질을 빚는 ..... ...
--

각각의 정보는 라인 단위로 구분되며 [문장식별자]:[문장]의 구조를 가진다. 이 정보는 향후 대량의 기계 점역을 위한 2차 변환 단계를 거치게 된다.

## (3) 목자 추출 시나리오 조건

2차 변환 전 적용할 시나리오 21종 조건에서 말뭉치 목자 추출 시나리오는 다음과 같다. 영어는 알파벳 a부터 z까지, 그리스 문자, 로마 숫자를 포함하고, 대소문자는 상관없다. 다음에 열거할 조건에서 플러스(+) 부분에 스페이스(공백) 한 칸 또는 두 칸이 들어가는 경우도 포함한다. 다음 조건에서 ...으로 표시한 자리에 문장 부호 또는 기타 기호(동그라미 숫자, 네모 숫자 포함)가 들어가는 경우도 포함한다.

- ① 한글 + (영어 + ... + )
- ② 영어 + (한글 + ... + )
- ③ 한글 + (숫자 + ... + ) → 열고 닫는 괄호가 반드시 쌍으로 존재할 것.
- ④ 숫자 + (한글 + ... + )
- ⑤ 영어 + (숫자 + ... + )

- ⑥ 한글 + (숫자 + ... + ) + 영어 → ③의 조건에 포함되어 삭제
- ⑦ 한글 + 숫자
- ⑧ 영어 + 숫자 → ⑦과 ⑧은 경우의 수가 많으므로 추출조건을 더 상세히 제시
- ⑨ 한글 + 영어 + 숫자
- ⑩ 한글 + 숫자 + 영어
- ⑪ 영어 + 한글 + 숫자
- ⑫ 영어 + 숫자 + 한글
- ⑬ 숫자 + 영어 + 한글
- ⑭ 숫자 + 한글 + 영어
- ⑮ 한글 + 문장 부호 + 영어
- ⑯ 영어 + 문장 부호 + 한글
- \* ⑮와 ⑯은 플러스(+) 부분을 붙여 쓴 경우임. 즉, 괄호나 빈칸이 들어가는 경우는 제외
- ⑰ 한글 + 영어 철자 나열 + 한글
- ⑱ 한글 문장 내에 로마자 있는 것
- ⑲ 한글 + 가운뎃점 의미로 사용되는 아래아 + 한글
- ⑳ 화폐 단위가 포함된 문장
- ㉑ 문장 부호 이외의 점자 규정에 포함되어 있는 부호

표 2-6 시나리오 조건

구 분	대상 조건(21종)	선정 조건(5종)
시나리오 조건	한글(영어), 영어(한글), 한글(숫자), 숫자(한글), 영어(숫자)+한글, 한글(숫자)+영어, 한글+숫자, 영어+숫자, 한글+영어+숫자, 한글+숫자+영어, 영어+한글+숫자, 영어+숫자+한글, 숫자+영어+한글, 숫자+한글+영어, 한글+문장 부호+영어, 영어+문장 부호+한글, 한글+ 영어 철자 나열 +한글, 한글 문장 내에 로마자 있는 것, 한글+가운뎃점 의미로 사용되는 아래아+한글, 화폐 단위, 문장 부호 이외의 점자 규정에 포함되어 있는 부호 등	한글 + (영어), 영어 + (한글), 한글 + (숫자), 숫자 + (한글), 영어 + (숫자)

## (4) 2차 변환 목자 파일

2차 변환은 기계 점역 전 미리 정의된 시나리오 5종에 해당하는 목자 문장 135,000개를 추출해내는 과정이다. 이때 너무 짧거나 긴 문장은 제거하기 위해 어절 수를 10어절 이상 30어절 미만으로 조건을 체크해서 추출했다. 선정된 시나리오 5종에 따른 말뭉치 종류별 및 시나리오별 추출 문장 수는 각 <표 2-7>, <표 2-8>과 같다.

표 2-7 말뭉치 종류별 추출 문장 수

말뭉치 JSON 종류	말뭉치 JSON 파일명	카테고리	문장 수
온라인 대화 말뭉치 2022 (국립국어원 온라인 게시 자료 원시 말뭉치)	EPRW2200000715 ~ EPRW2200000780	웹 > 게시판	14
	ESRW2200000002 ~ ESRW2200000704	웹 > 누리소통망	1,468
신문 말뭉치 2022 (국립국어원 신문 말뭉치)	NIRW2200000001 ~ NIRW2200000002	신문 > 인터넷 기반 신문	23,073
	NLRW2200000001 ~ NLRW2200000017	신문 > 지역 종합지	52,978
	NPRW2200000001 ~ NPRW2200000004	신문 > 전문지	57,467
총합			135,000

표 2-8 시나리오별 추출 문장 수

시나리오	추출 방법	비율(%)	추출 문장 수(개)
한글 + (영문) 형태 포함	정규표현식	70.51	95,187
영문 + (한글) 형태 포함	정규표현식	9.13	12,326
한글 + (숫자) 형태 포함	정규표현식	14.76	19,924
숫자 + (한글) 형태 포함	정규표현식	0.74	997
영문 + (숫자) 형태 포함	정규표현식	4.86	6,566
합계		100	135,000

### 3) 카테고리 구성 및 파일 명명 규칙

목자-점자 병렬말뭉치는 대분류 아래의 대분류와 중분류로 구분하였다. 카테고리는 <표 2-9>와 같다. 카테고리 중에서 중분류, 소분류를 다음과 같이 넣는다.

표 2-9 데이터 카테고리

대분류	중분류	도큐먼트 메타데이터
1. 신문	가. 전국 종합지	title(도서명), author(저자), publisher(출판사), date(작성연도), original topic(주제)
	나. 지역 종합지	
	다. 전문지	
	라. 인터넷 기반 신문	
	마. 기타	
2. 온라인 대화	가. 게시판	
	나. 누리소통망	

### 4) 최종 적합 파일 유형

최종 데이터 적합 엑셀 파일은 <표 2-10>과 같이 메타데이터와 카테고리를 포함한다.

표 2-10 최종 데이터 적합 엑셀 파일 유형

dataset	메타데이터	
	language_info	parallel
국립국어원 목자 - 점자 병렬말뭉치	source_language(원자료 언어, kl), target_language(번역 대상 언어, kb)	id(목자-점자 병렬말뭉치 문장 고유 ID), original_id(원문 말뭉치 문장 고유 ID), source(목자 문장), target(점자 문장), revision1(검수된 점자 문장)

### 3 데이터 정제

본 정제 단계에서는 수집된 목자 데이터에 대하여 가공 전, 작업에 용이한 형태로 목자 데이터를 교열하거나, 교열이 어려운 문장을 제외하는 등의 과정을 수행한다.

표 2-11 데이터 정제 수행 과정

세부 절차	작업
1. 부적합 데이터 삭제	<ul style="list-style-type: none"> <li>- 분절 오류 문장 삭제</li> <li>- 고유명사 처리 기준 미달 문장 삭제(개인정보, 미등록어 과다 등)</li> <li>- 심각한 비문 삭제</li> </ul>
2. 데이터 정제	<ul style="list-style-type: none"> <li>- 점형에 부적합한 기호 처리</li> <li>- 문장 종료 처리(분리, 병합, 마침표 등)</li> <li>- 문장 오타 주석 오류 확인</li> </ul>

#### 1) 부적합 목자 데이터 삭제

수집한 목자 데이터 문장의 효율적인 가공을 위해 구축 대상에 적합하지 않은 목자 문장을 먼저 삭제한다. 삭제해야 하는 목자 문장은 대표적으로 문장이 강제로 분절되었을 때 오류가 생긴 경우와 공개적으로 사용하기에 적절하지 않은 개인정보가 있는 경우, 가공이 어려운 미등록어가 한 문장 내에 많은 경우를 예로 들 수 있다. 국문이 비문일 때에도 데이터 가공 단계에서의 오류가 발생하므로 정제 단계에서 해당 목자 문장들을 삭제한다.

#### 2) 데이터 정제

부적합한 목자 문장을 제외하고 나면 일반적으로 중복된 목자 문장을 삭제하고, 점형에 부적합한 기호를 교체하거나 삭제하여 문장 부호 오류, 오타 등을 확인하는 정제 과정을 거친다. 정제된 결과인 목자 문장을 웹 기반 관리시스템에 업데이트한다.

## 4 데이터 가공

가공 방법은 기계 번역 결과를 활용하는 기계 번역 후 편집(MTPE)으로 정한다. 이를 검수하는 과정은 반드시 후행되어야 함을 유의한다.

표 2-12 가공 세부 절차

세부 절차	작업
1. MTPE	<ul style="list-style-type: none"> <li>- 적합한 기계 번역(MT: Machine Translation) 선택</li> <li>- 기계 번역 후 오류 수정</li> <li>- 기계 번역 품질이 낮은 문장은 분류하여 사람이 점역</li> <li>- 3차 변환: 최종 형태로 정리한 묵자 문장을 기계 점역</li> </ul>

### 1) MTPE

MTPE 방법을 활용하여 가공하는 경우 분야별 특성에 따라 알맞은 기계 번역을 선택해야 한다. 기계 번역의 결과를 놓고 사람이 원문과 대조하여 점형문의 오류를 확인한다. 오류를 수정하는 것 뿐 아니라, 이후 구축된 데이터로 NMT(Neural Machine Translation) 학습 등의 활용을 고려하고, 형식적 일관성 등에 대한 작업 지침을 마련하여 이에 맞춰 검수한다. 기계 점역에 따라 몇몇 문장은 심각한 오류를 나타내는 경우가 있는데 이때는 사람이 점역하거나, 정제된 문장 수가 충분하다면 이를 보류 및 삭제하는 등 효과적인 프로세스를 만들어 가공한다.

- 묵자-점자 문장 한 쌍으로 구성
- 문자 세트: 유니코드 상의 한글 음절을 낱자와 점자로 매칭한 테이블
- 말뭉치 파일은 JSON 등 기계가 읽을 수 있는(Machine Readable) 형태여야 함

표 2-13 가공 작성 예시

```

{
  "source": {
    "text": "페트병 커피와 캔 커피는 계절에 따라 상호 보완적이다. 편의점 CU 커피 담당 MD(상품기획자)는  

    “겨울철에는 따뜻한 음료 수요가 많아 온장고에 넣을 수 있는 캔커피가 인기를 끌고, 여름에는 가정에서 냉장  

    보관했다가 바로 꺼내먹을 수 있는 페트 커피가 대세”라고 말했다.",
    "text_language": "korean",
    "category": "normal",
  },
  "braille_translation_result": {
    "unicode": " ",
    "ttb_option": "none",
    "characters_per_line": "32"
  }
}

```

## 2) 3차 변환 목자-점자 말뭉치 파일

3차 변환은 최종 형태로 정리한 말뭉치 목자 문장을 기계 점역한다. 기계 점역에 사용된 소프트웨어는 웹 기반 관리시스템과 연계된 (주)ACNS의 점자 기계 번역 솔루션(Docu Braille V1.0)을 사용하였다. 점자 기계 번역 솔루션은 API와 연계하여 점자 유니코드 값으로 실시간 점역하여 사용자에게 보여주며 데이터베이스의 점자 데이터 항목에 반영하는 구조를 가지고 있다.

### (1) JAVA 객체 변환

선정된 목자 문장파일 135,000개는 JAVA에서 처리 가능한 JSON Object 객체 형태로 1차 변환 처리 한다. 그리고 가공된 JSON Object에서 목자 대상이 되는 문서 값을 추출하여 JSON Array객체로 2차 변환한다. JSON Array객체에서 “Metadata”, “Paragraph” 데이터를 추출하는 과정에서 Metadata는 JSON Object객체로, Paragraph는 JSON Array객체로 3차 변환한다.

#	이름	데이터 유형	길이/설정	부호 없음	NULL 허용	0으로 채움	기본값	코멘트
1	year	VARCHAR	4	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	기본값 없음	년도
2	brl_id	VARCHAR	20	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	기본값 없음	검수ID (코드+일련번호)
3	brl_ser	INT	11	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	기본값 없음	검수일련번호
4	process_ser	TINYINT	4	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	'1'	검수작업자수
5	json_id	VARCHAR	14	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL	검수작업파일 및 메타정보 테이블의 JSON ID 값
6	subcategory	VARCHAR	50	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL	종류
7	title	VARCHAR	200	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL	도서명
8	author	VARCHAR	100	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL	저자
9	publisher	VARCHAR	50	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL	출판사
10	write_date	VARCHAR	8	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL	작성년도
11	original_topic	VARCHAR	200	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL	주제
12	original_id	VARCHAR	30	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL	원본 paragraph id
13	scenario	TINYINT	4	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL	시나리오(정규식에 의한 분류값) CD010
14	form	TEXT		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	기본값 없음	문장
15	original_braille	TEXT		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	기본값 없음	기계점역점자
16	word_count	SMALLINT	6	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL	어절 수
17	form_length	SMALLINT	6	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL	문장 길이
18	target_yn	VARCHAR	1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL	작업대상여부
19	job_userid	VARCHAR	20	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL	작업자ID (검수 중 인 작업자ID)
20	reg_date	DATETIME		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL	등록일시
21	reg_userid	VARCHAR	20	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	NULL	등록사용자ID

그림 2-9 JAVA 객체 변환 화면

## (2) 데이터베이스 등록

위의 과정을 통하여 생성된 2개의 파일의 목록은 아래와 같다.

- Metadata 객체: title(도서명), author(저자), publisher(출판사), data(작성일), original\_topic(주제) 등 항목으로 데이터베이스에 저장

- Paragraph 객체: ID(원본 paragraph ID), form(목자) 항목으로 데이터베이스에 저장

위의 2가지 자료를 토대로 <그림 2-10>과 같은 데이터베이스에 해당 자료를 입력한다.

brl_id	json_id	subcategory	title	author	publisher	write_date	original_id	original_topic	scenario	word_count	form_length	form	original_braille
BRUCORPUS00000000	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	CBS노점뉴스 특약식...	노점뉴스	20210101	NRW2200000001.1.2	지역,경제지역,정치지역,문화...	1	16	80	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000001	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	CBS노점뉴스 장영주...	노점뉴스	20210101	NRW2200000001.2.12	국회,국회,국회,국회,국회...	2	15	74	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000002	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	CBS노점뉴스 장영주...	노점뉴스	20210101	NRW2200000001.3.12	국회,국회,국회,국회,국회...	1	11	53	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000003	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	CBS노점뉴스 장영주...	노점뉴스	20210101	NRW2200000001.3.3	국회,국회,국회,국회,국회...	2	20	109	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000004	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	CBS노점뉴스 장영주...	노점뉴스	20210101	NRW2200000001.4.2	IT,국회,IT,국회,국회...	2	22	97	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000005	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	CBS노점뉴스 장영주...	노점뉴스	20210101	NRW2200000001.4.7	IT,국회,IT,국회,국회...	1	23	101	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000006	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	출신CBS 이상훈 기자	노점뉴스	20210101	NRW2200000001.5.9	지역,경제지역,정치지역,문화...	3	29	132	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000007	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	출신CBS 이상훈 기자	노점뉴스	20210101	NRW2200000001.5.13	지역,경제지역,정치지역,문화...	3	23	92	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000008	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	출신CBS 이상훈 기자	노점뉴스	20210101	NRW2200000001.5.14	지역,경제지역,정치지역,문화...	3	23	114	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000009	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	출신CBS 최희정 기자	노점뉴스	20210101	NRW2200000001.6.3	사회,사건,사건,사건,사건...	5	18	82	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000010	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	CBS노점뉴스 장영주...	노점뉴스	20210101	NRW2200000001.9.3	국회,국회,국회,국회,국회...	2	23	115	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000011	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	CBS노점뉴스 장영주...	노점뉴스	20210101	NRW2200000001.9.8	국회,국회,국회,국회,국회...	2	19	88	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000012	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	한글CBS 송승환 기자	노점뉴스	20210101	NRW2200000001.11.3	사회,사건,사건,사건,사건...	5	16	71	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000013	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	한글CBS 송승환 기자	노점뉴스	20210101	NRW2200000001.11.4	사회,사건,사건,사건,사건...	5	25	99	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000014	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	한글CBS 송승환 기자	노점뉴스	20210101	NRW2200000001.11.12	사회,사건,사건,사건,사건...	2	20	109	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000015	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	CBS노점뉴스 장영주...	노점뉴스	20210101	NRW2200000001.12.2	국회,국회,국회,국회,국회...	3	10	55	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000016	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	CBS노점뉴스 장영주...	노점뉴스	20210101	NRW2200000001.14.8	IT,국회,IT,국회,국회...	1	30	126	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000017	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	채우CBS 김대기 기자	노점뉴스	20210101	NRW2200000001.22.6	지역,경제지역,정치지역,문화...	3	13	82	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000018	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	CBS노점뉴스 특약식...	노점뉴스	20210101	NRW2200000001.33.2	지역,경제지역,정치지역,문화...	1	11	55	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000019	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	CBS노점뉴스 장영주...	노점뉴스	20210102	NRW2200000001.46.5	국회,국회,국회,국회,국회...	1	21	124	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000020	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	CBS노점뉴스 장영주...	노점뉴스	20210102	NRW2200000001.46.5	국회,국회,국회,국회,국회...	1	22	119	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000021	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	CBS노점뉴스 장영주...	노점뉴스	20210102	NRW2200000001.46.6	국회,국회,국회,국회,국회...	1	11	60	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000022	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	CBS노점뉴스 장영주...	노점뉴스	20210102	NRW2200000001.74.2	국회,국회,국회,국회,국회...	3	10	45	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000023	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	CBS노점뉴스 장영주...	노점뉴스	20210102	NRW2200000001.74.9	국회,국회,국회,국회,국회...	1	16	107	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000024	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	CBS노점뉴스 장영주...	노점뉴스	20210103	NRW2200000001.75.2	사회,사건,사건,사건,사건...	1	28	117	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000025	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	CBS노점뉴스 장영주...	노점뉴스	20210103	NRW2200000001.77.6	경제,문화지역,정치지역,문화...	3	28	128	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000026	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	CBS노점뉴스 장영주...	노점뉴스	20210103	NRW2200000001.80.3	국회,국회,국회,국회,국회...	1	26	150	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000027	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	출신CBS 김대기 기자	노점뉴스	20210103	NRW2200000001.84.4	지역,경제지역,정치지역,문화...	1	15	75	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000028	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	CBS노점뉴스 장영주...	노점뉴스	20210103	NRW2200000001.87.2	문화,문화,문화,문화,문화...	3	23	107	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...
BRUCORPUS00000029	NRW2200000001	신문 > 인터넷 기반 신문	노점뉴스 2021년 기사	CBS노점뉴스 장영주...	노점뉴스	20210103	NRW2200000001.89.3	국회,국회,국회,국회,국회...	1	29	120	한글 교정용 부호인, 정치, 문화, 경제...	한글 교정용 부호인, 정치, 문화, 경제...

그림 2-10 데이터베이스 예시

## 5 데이터 검사

본 검사 단계에서는 검수자가 가공 결과물을 확인하여 적합/오류/보류를 판정하고 지침에 맞지 않는 경우에 이를 적절하게 기재한다. 이때 오류/보류 문장은 점형을 바르게 교정하여 적합 문장 수량 증가 및 점형 수정 예시 결과물을 구축한다. 작업 품질에 따라 검수 교차 작업을 실시하여 품질을 확보하였다. 기존 구축 목자-점자 병렬말뭉치 오류·보류 문장은 해당 연도 구축 시 오류/보류 판정을 받았으므로 적합/오류/보류 판정보다 점형 교정에 의미를 두고 데이터를 검사한다.

표 2-14 데이터 검사 세부 절차

세부 절차	작업	해당 검수자 그룹
1차 검수	<ul style="list-style-type: none"> <li>- 검수 단계 설정: 1단계</li> <li>- 지침에 따른 점형 문장 적합/오류/보류 판정</li> <li>- 오류/보류 문장은 올바른 점형으로 교정</li> </ul>	A그룹 C그룹
2차 검수	<ul style="list-style-type: none"> <li>- 검수 단계 설정: 2단계</li> <li>- 1차 검수된 적합 문장의 적합/오류/보류 판정</li> </ul>	B그룹

### 1) 검수 지침

사람이 직접 검수할 때에는 일관성이 일정 수준 확보된 형태로 작업을 하는 것이 중요하다. 이를 위한 검수 지침은 아래와 같다.

#### (1) 사이트 접속 및 로그인

- 말뭉치 웹 기반 관리시스템 <http://www.kbucorpus.kr> 사이트에 접속
- 아이디와 패스워드를 입력하여 로그인
- 아이디는 kbu23a0000 형식이며 하단의 규칙에 따라 본인 아이디가 생성됨
- kbu23은 모두 동일하며, 본인이 속한 그룹에 따라 a/b/c를 뒤에 붙임
- 마지막 숫자 4개는 본인 핸드폰 번호 마지막 뒷자리임(ex. kbu23a2222, kbu23b1111)
- 비밀번호는 보안 사항으로 검수자별 개별 안내
- Chrome 프로그램이 실행되는 기종의 한소네의 경우, 점자를 읽을 순 있으나 A/C그룹의 점자 교정 작업은 한소네 소프트웨어 문제로 진행 불가

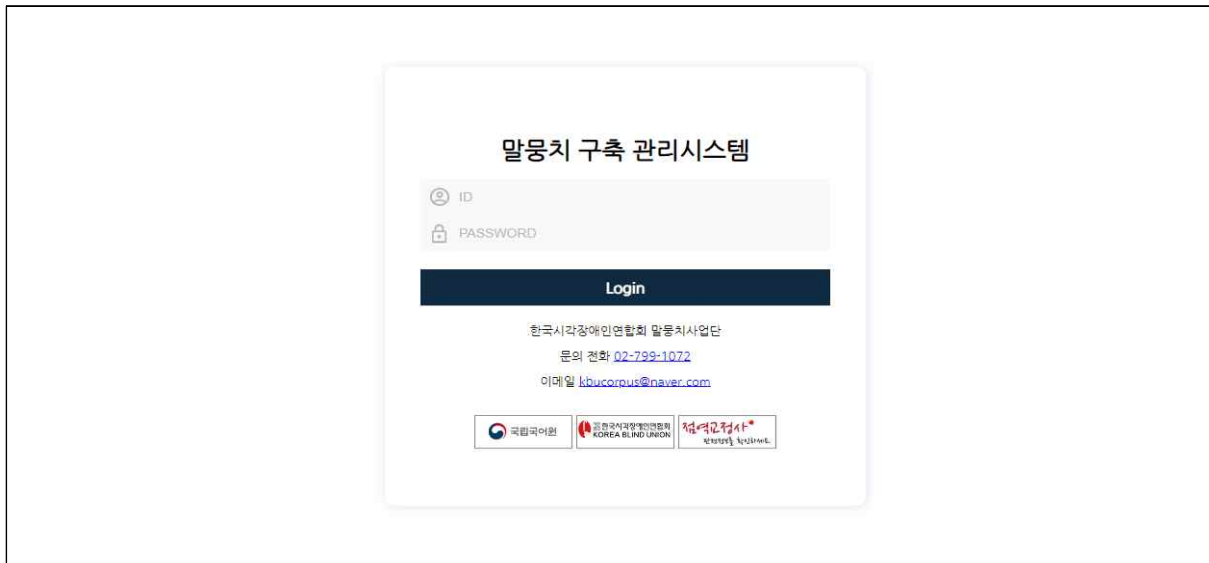


그림 2-11 말뭉치 관리시스템 사이트 접속 로그인 화면

## (2) 내비게이션 구성

- 검수: 목자-점자 병렬말뭉치 검수 화면
- 공지사항: 말뭉치사업단에서 공지사항을 올리는 페이지
- 게시판: 검수자가 직접 작성하고 답변할 수 있는 다용도 게시판 페이지
- 비밀번호: 비밀번호 변경 가능(새 창 열림)
- 로그아웃: 관리시스템 사이트에서 계정 종료
- Alt+M키를 통해 내비게이션 메뉴 비활성화 가능



그림 2-12 말뭉치 관리시스템 내비게이션 화면

### (3) A/B그룹 검수 화면

- 화면에는 1개 문장의 목자-점자만 나오며 윗줄에는 목자, 아랫줄에는 점자가 표기됨
- 목자 상단에는 전체 처리한 문장 수, 전체 처리율, 주간 처리율이 표기되어 있음
- 센스리더 사용자는 점자 영역에 초점 접근 후 Home키 입력 후 방향키로 이동하면 아스키 코드(ex. 135점, 24점) 형태로 음성 출력됨
- 비시각의 경우, 점자 폰트 가독성 및 편의를 위해 3종의 점자 폰트로 변경 가능

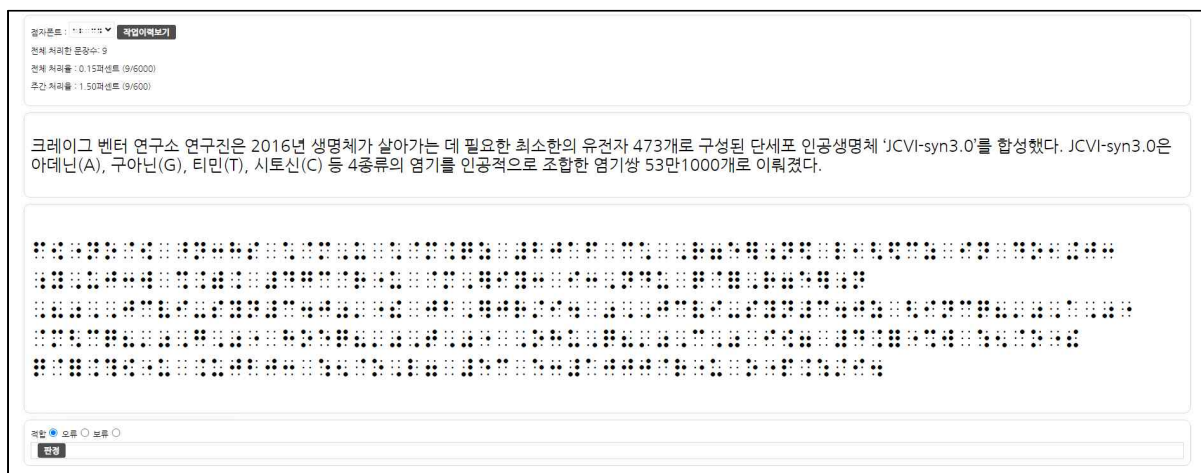


그림 2-13 말뭉치 관리시스템 A/B그룹 목자-점자 검수 화면

### (4) 적합/오류/보류 버튼 체크

- 목자-점자 일치 여부에 따라 말뭉치 하단에 있는 적합/오류/보류 라디오 버튼 체크
- 적합: 목자와 점자 점역이 규정과 일치하며 모두 올바르게 점역된 경우
- 오류: 목자와 점자 점역이 규정과 다르거나 목자 내용이 바르게 점역되지 않은 경우
- 보류: 한글 점자 규정에 정의되지 않은 내용/기호가 있거나 목자 내용이 제2외국어, 수식, 화학 공식, 악보 등과 같이 한글/영어 규정이 아닌 경우

### (5) 오류/보류 사유 기재

- 적합한 경우 바로 적합 라디오 버튼 체크 후 판정 버튼 선택
- 오류/보류의 경우 콤보박스에서 해당 사유 선택 후 텍스트 박스에 틀린 단어 기재
- A그룹: 점자 점형 수정 후 오류/보류 선택은 그대로 두고 판정 버튼 선택
- B그룹: 사유 선택&틀린 단어 기재 후 바로 판정 버튼 선택

- 오류/보류 시 틀린 단어를 텍스트 박스에 기재해야 판정 버튼이 활성화되며 다음 문장으로 넘어갈 수 있음
- 한 문장에 오류와 보류가 혼재되어 있는 경우, 오류 버튼을 누른 후 오류 사유와 보류 사유 모두 작성
- 틀린 사항이 여러 개인 경우, 기타(직접 입력) 사유 콤보박스 선택 후 틀린 단어 모두 작성
- 목자 원문에 띄어쓰기가 잘못되어 있고, 점자를 확인하였을 때 그대로 점역되어 있다면 '적합'이나, 띄어쓰기로 한글 점자 규정에 어긋난다면 '오류'임
- A그룹은 점자 점형 수정 후 적합으로 판정을 변경하지 말고 처음 판정 그대로 진행

점자문서: 1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 20. 21. 22. 23. 24. 25. 26. 27. 28. 29. 30. 31. 32. 33. 34. 35. 36. 37. 38. 39. 40. 41. 42. 43. 44. 45. 46. 47. 48. 49. 50. 51. 52. 53. 54. 55. 56. 57. 58. 59. 60. 61. 62. 63. 64. 65. 66. 67. 68. 69. 70. 71. 72. 73. 74. 75. 76. 77. 78. 79. 80. 81. 82. 83. 84. 85. 86. 87. 88. 89. 90. 91. 92. 93. 94. 95. 96. 97. 98. 99. 100. 101. 102. 103. 104. 105. 106. 107. 108. 109. 110. 111. 112. 113. 114. 115. 116. 117. 118. 119. 120. 121. 122. 123. 124. 125. 126. 127. 128. 129. 130. 131. 132. 133. 134. 135. 136. 137. 138. 139. 140. 141. 142. 143. 144. 145. 146. 147. 148. 149. 150. 151. 152. 153. 154. 155. 156. 157. 158. 159. 160. 161. 162. 163. 164. 165. 166. 167. 168. 169. 170. 171. 172. 173. 174. 175. 176. 177. 178. 179. 180. 181. 182. 183. 184. 185. 186. 187. 188. 189. 190. 191. 192. 193. 194. 195. 196. 197. 198. 199. 200. 201. 202. 203. 204. 205. 206. 207. 208. 209. 210. 211. 212. 213. 214. 215. 216. 217. 218. 219. 220. 221. 222. 223. 224. 225. 226. 227. 228. 229. 230. 231. 232. 233. 234. 235. 236. 237. 238. 239. 240. 241. 242. 243. 244. 245. 246. 247. 248. 249. 250. 251. 252. 253. 254. 255. 256. 257. 258. 259. 260. 261. 262. 263. 264. 265. 266. 267. 268. 269. 270. 271. 272. 273. 274. 275. 276. 277. 278. 279. 280. 281. 282. 283. 284. 285. 286. 287. 288. 289. 290. 291. 292. 293. 294. 295. 296. 297. 298. 299. 300. 301. 302. 303. 304. 305. 306. 307. 308. 309. 310. 311. 312. 313. 314. 315. 316. 317. 318. 319. 320. 321. 322. 323. 324. 325. 326. 327. 328. 329. 330. 331. 332. 333. 334. 335. 336. 337. 338. 339. 340. 341. 342. 343. 344. 345. 346. 347. 348. 349. 350. 351. 352. 353. 354. 355. 356. 357. 358. 359. 360. 361. 362. 363. 364. 365. 366. 367. 368. 369. 370. 371. 372. 373. 374. 375. 376. 377. 378. 379. 380. 381. 382. 383. 384. 385. 386. 387. 388. 389. 390. 391. 392. 393. 394. 395. 396. 397. 398. 399. 400. 401. 402. 403. 404. 405. 406. 407. 408. 409. 410. 411. 412. 413. 414. 415. 416. 417. 418. 419. 420. 421. 422. 423. 424. 425. 426. 427. 428. 429. 430. 431. 432. 433. 434. 435. 436. 437. 438. 439. 440. 441. 442. 443. 444. 445. 446. 447. 448. 449. 450. 451. 452. 453. 454. 455. 456. 457. 458. 459. 460. 461. 462. 463. 464. 465. 466. 467. 468. 469. 470. 471. 472. 473. 474. 475. 476. 477. 478. 479. 480. 481. 482. 483. 484. 485. 486. 487. 488. 489. 490. 491. 492. 493. 494. 495. 496. 497. 498. 499. 500. 501. 502. 503. 504. 505. 506. 507. 508. 509. 510. 511. 512. 513. 514. 515. 516. 517. 518. 519. 520. 521. 522. 523. 524. 525. 526. 527. 528. 529. 530. 531. 532. 533. 534. 535. 536. 537. 538. 539. 540. 541. 542. 543. 544. 545. 546. 547. 548. 549. 550. 551. 552. 553. 554. 555. 556. 557. 558. 559. 560. 561. 562. 563. 564. 565. 566. 567. 568. 569. 570. 571. 572. 573. 574. 575. 576. 577. 578. 579. 580. 581. 582. 583. 584. 585. 586. 587. 588. 589. 590. 591. 592. 593. 594. 595. 596. 597. 598. 599. 600. 601. 602. 603. 604. 605. 606. 607. 608. 609. 610. 611. 612. 613. 614. 615. 616. 617. 618. 619. 620. 621. 622. 623. 624. 625. 626. 627. 628. 629. 630. 631. 632. 633. 634. 635. 636. 637. 638. 639. 640. 641. 642. 643. 644. 645. 646. 647. 648. 649. 650. 651. 652. 653. 654. 655. 656. 657. 658. 659. 660. 661. 662. 663. 664. 665. 666. 667. 668. 669. 670. 671. 672. 673. 674. 675. 676. 677. 678. 679. 680. 681. 682. 683. 684. 685. 686. 687. 688. 689. 690. 691. 692. 693. 694. 695. 696. 697. 698. 699. 700. 701. 702. 703. 704. 705. 706. 707. 708. 709. 710. 711. 712. 713. 714. 715. 716. 717. 718. 719. 720. 721. 722. 723. 724. 725. 726. 727. 728. 729. 730. 731. 732. 733. 734. 735. 736. 737. 738. 739. 740. 741. 742. 743. 744. 745. 746. 747. 748. 749. 750. 751. 752. 753. 754. 755. 756. 757. 758. 759. 760. 761. 762. 763. 764. 765. 766. 767. 768. 769. 770. 771. 772. 773. 774. 775. 776. 777. 778. 779. 780. 781. 782. 783. 784. 785. 786. 787. 788. 789. 790. 791. 792. 793. 794. 795. 796. 797. 798. 799. 800. 801. 802. 803. 804. 805. 806. 807. 808. 809. 810. 811. 812. 813. 814. 815. 816. 817. 818. 819. 820. 821. 822. 823. 824. 825. 826. 827. 828. 829. 830. 831. 832. 833. 834. 835. 836. 837. 838. 839. 840. 841. 842. 843. 844. 845. 846. 847. 848. 849. 850. 851. 852. 853. 854. 855. 856. 857. 858. 859. 860. 861. 862. 863. 864. 865. 866. 867. 868. 869. 870. 871. 872. 873. 874. 875. 876. 877. 878. 879. 880. 881. 882. 883. 884. 885. 886. 887. 888. 889. 890. 891. 892. 893. 894. 895. 896. 897. 898. 899. 900. 901. 902. 903. 904. 905. 906. 907. 908. 909. 910. 911. 912. 913. 914. 915. 916. 917. 918. 919. 920. 921. 922. 923. 924. 925. 926. 927. 928. 929. 930. 931. 932. 933. 934. 935. 936. 937. 938. 939. 940. 941. 942. 943. 944. 945. 946. 947. 948. 949. 950. 951. 952. 953. 954. 955. 956. 957. 958. 959. 960. 961. 962. 963. 964. 965. 966. 967. 968. 969. 970. 971. 972. 973. 974. 975. 976. 977. 978. 979. 980. 981. 982. 983. 984. 985. 986. 987. 988. 989. 990. 991. 992. 993. 994. 995. 996. 997. 998. 999. 1000. 1001. 1002. 1003. 1004. 1005. 1006. 1007. 1008. 1009. 1010. 1011. 1012. 1013. 1014. 1015. 1016. 1017. 1018. 1019. 1020. 1021. 1022. 1023. 1024. 1025. 1026. 1027. 1028. 1029. 1030. 1031. 1032. 1033. 1034. 1035. 1036. 1037. 1038. 1039. 1040. 1041. 1042. 1043. 1044. 1045. 1046. 1047. 1048. 1049. 1050. 1051. 1052. 1053. 1054. 1055. 1056. 1057. 1058. 1059. 1060. 1061. 1062. 1063. 1064. 1065. 1066. 1067. 1068. 1069. 1070. 1071. 1072. 1073. 1074. 1075. 1076. 1077. 1078. 1079. 1080. 1081. 1082. 1083. 1084. 1085. 1086. 1087. 1088. 1089. 1090. 1091. 1092. 1093. 1094. 1095. 1096. 1097. 1098. 1099. 1100. 1101. 1102. 1103. 1104. 1105. 1106. 1107. 1108. 1109. 1110. 1111. 1112. 1113. 1114. 1115. 1116. 1117. 1118. 1119. 1120. 1121. 1122. 1123. 1124. 1125. 1126. 1127. 1128. 1129. 1130. 1131. 1132. 1133. 1134. 1135. 1136. 1137. 1138. 1139. 1140. 1141. 1142. 1143. 1144. 1145. 1146. 1147. 1148. 1149. 1150. 1151. 1152. 1153. 1154. 1155. 1156. 1157. 1158. 1159. 1160. 1161. 1162. 1163. 1164. 1165. 1166. 1167. 1168. 1169. 1170. 1171. 1172. 1173. 1174. 1175. 1176. 1177. 1178. 1179. 1180. 1181. 1182. 1183. 1184. 1185. 1186. 1187. 1188. 1189. 1190. 1191. 1192. 1193. 1194. 1195. 1196. 1197. 1198. 1199. 1200. 1201. 1202. 1203. 1204. 1205. 1206. 1207. 1208. 1209. 1210. 1211. 1212. 1213. 1214. 1215. 1216. 1217. 1218. 1219. 1220. 1221. 1222. 1223. 1224. 1225. 1226. 1227. 1228. 1229. 1230. 1231. 1232. 1233. 1234. 1235. 1236. 1237. 1238. 1239. 1240. 1241. 1242. 1243. 1244. 1245. 1246. 1247. 1248. 1249. 1250. 1251. 1252. 1253. 1254. 1255. 1256. 1257. 1258. 1259. 1260. 1261. 1262. 1263. 1264. 1265. 1266. 1267. 1268. 1269. 1270. 1271. 1272. 1273. 1274. 1275. 1276. 1277. 1278. 1279. 1280. 1281. 1282. 1283. 1284. 1285. 1286. 1287. 1288. 1289. 1290. 1291. 1292. 1293. 1294. 1295. 1296. 1297. 1298. 1299. 1300. 1301. 1302. 1303. 1304. 1305. 1306. 1307. 1308. 1309. 1310. 1311. 1312. 1313. 1314. 1315. 1316. 1317. 1318. 1319. 1320. 1321. 1322. 1323. 1324. 1325. 1326. 1327. 1328. 1329. 1330. 1331. 1332. 1333. 1334. 1335. 1336. 1337. 1338. 1339. 1340. 1341. 1342. 1343. 1344. 1345. 1346. 1347. 1348. 1349. 1350. 1351. 1352. 1353. 1354. 1355. 1356. 1357. 1358. 1359. 1360. 1361. 1362. 1363. 1364. 1365. 1366. 1367. 1368. 1369. 1370. 1371. 1372. 1373. 1374. 1375. 1376. 1377. 1378. 1379. 1380. 1381. 1382. 1383. 1384. 1385. 1386. 1387. 1388. 1389. 1390. 1391. 1392. 1393. 1394. 1395. 1396. 1397. 1398. 1399. 1400. 1401. 1402. 1403. 1404. 1405. 1406. 1407. 1408. 1409. 1410. 1411. 1412. 1413. 1414. 1415. 1416. 1417. 1418. 1419. 1420. 1421. 1422. 1423. 1424. 1425. 1426. 1427. 1428. 1429. 1430. 1431. 1432. 1433. 1434. 1435. 1436. 1437. 1438. 1439. 1440. 1441. 1442. 1443. 1444. 1445. 1446. 1447. 1448. 1449. 1450. 1451. 1452. 1453. 1454. 1455. 1456. 1457. 1458. 1459. 1460. 1461. 1462. 1463. 1464. 1465. 1466. 1467. 1468. 1469. 1470. 1471. 1472. 1473. 1474. 1475. 1476. 1477. 1478. 1479. 1480. 1481. 1482. 1483. 1484. 1485. 1486. 1487. 1488. 1489. 1490. 1491. 1492. 1493. 1494. 1495. 1496. 1497. 1498. 1499. 1500. 1501. 1502. 1503. 1504. 1505. 1506. 1507. 1508. 1509. 1510. 1511. 1512. 1513. 1514. 1515. 1516. 1517. 1518. 1519. 1520. 1521. 1522. 1523. 1524. 1525. 1526. 1527. 1528. 1529. 1530. 1531. 1532. 1533. 1534. 1535. 1536. 1537. 1538. 1539. 1540. 1541. 1542. 1543. 1544. 1545. 1546. 1547. 1548. 1549. 1550. 1551. 1552. 1553. 1554. 1555. 1556. 1557. 1558. 1559. 1560. 1561. 1562. 1563. 1564. 1565. 1566. 1567. 1568. 1569. 1570. 1571. 1572. 1573. 1574. 1575. 1576. 1577. 1578. 1579. 1580. 1581. 1582. 1583. 1584. 1585. 1586. 1587. 1588. 1589. 1590. 1591. 1592. 1593. 1594. 1595. 1596. 1597. 1598. 1599. 1600. 1601. 1602. 1603. 1604. 1605. 1606. 1607. 1608. 1609. 1610. 1611. 1612. 1613. 1614. 1615. 1616. 1617. 1618. 1619. 1620. 1621. 1622. 1623. 1624. 1625. 1626. 1627. 1628. 1629. 1630. 1631. 1632. 1633. 1634. 1635. 1636. 1637. 1638. 1639. 1640. 1641. 1642. 1643. 1644. 1645. 1646. 1647. 1648. 1649. 1650. 1651. 1652. 1653. 1654. 1655. 1656. 1657. 1658. 1659. 1660. 1661. 1662. 1663. 1664. 1665. 1666. 1667. 1668. 1669. 1670. 1671. 1672. 1673. 1674. 1675. 1676. 1677. 1678. 1679. 1680. 1681. 1682. 1683. 1684. 1685. 1686. 1687. 1688. 1689. 1690. 1691. 1692. 1693. 1694. 1695. 1696. 1697. 1698. 1699. 1700. 1701. 1702. 1703. 1704. 1705. 1706. 1707. 1708. 1709. 1710. 1711. 1712. 1713. 1714. 1715. 1716. 1717. 1718. 1719. 1720. 1721. 1722. 1723. 1724. 1725. 1726. 1727. 1728. 1729. 1730. 1731. 1732. 1733. 1734. 1735. 1736. 1737. 1738. 1739. 1740. 1741. 1742. 1743. 1744. 1745. 1746. 1747. 1748. 1749. 1750. 1751. 1752. 1753. 1754. 1755. 1756. 1757. 1758. 1759. 1760. 1761. 1762. 1763. 1764. 1765. 1766. 1767. 1768. 1769. 1770. 1771. 1772. 1773. 1774. 1775. 1776. 1777. 1778. 1779. 1780. 1781. 1782. 1783. 1784. 1785. 1786. 1787. 1788. 1789. 1790. 1791. 1792. 1793. 1794. 1795. 1796. 1797. 1798. 1799. 1800. 1801. 1802. 1803. 1804. 1805. 1806. 1807. 1808. 1809. 1810. 1811. 1812. 1813. 1814. 1815. 1816. 1817. 1818. 1819. 1820. 1821. 1822. 1823. 1824. 1825. 1826. 1827. 1828. 1829. 1830. 1831. 1832. 1833. 1834. 1835. 1836. 1837. 1838. 1839. 1840. 1841. 1842. 1843. 1844. 1845. 1846. 1847. 1848. 1849. 1850. 1851. 1852. 1853. 1854. 1855. 1856. 1857. 1858. 1859. 1860. 1861. 1862. 1863. 1864. 1865. 1866. 1867. 1868. 1869. 1870. 1871. 1872. 1873. 1874. 1875. 1876. 1877. 1878. 1879. 1880. 1881. 1882. 1883. 1884. 1885. 1886. 1887. 1888. 1889. 1890. 1891. 1892. 1893. 1894. 1895. 1896. 1897. 1898. 1899. 1900. 1901. 1902. 1903. 1904. 1905. 1906. 1907. 1908. 1909. 1910. 1911. 1912. 1913. 1914. 1915. 1916. 1917. 1918. 1919. 1920. 1921. 1922. 1923. 1924. 1925. 1926. 1927. 1928. 1929. 1930. 1931. 1932. 1933. 1934. 1935. 1936. 1937. 1938. 1939. 1940. 1941. 1942. 1943. 1944. 1945. 1946. 1947. 1948. 1949. 1950. 1951. 1952. 1953. 1954. 1955. 1956. 1957. 1958. 1959. 1960. 1961. 1962. 1963. 1964. 1965. 1966. 1967. 1968. 1969. 1970. 1971. 1972. 1973. 1974. 1975. 1976. 1977. 1978. 1979. 1980. 1981. 1982. 1983. 1984. 1985. 1986. 1987. 1988. 1989. 1990. 1991. 1992. 1993. 1994. 1995. 1996. 1997. 1998. 1999. 2000. 2001. 2002. 2003. 2004. 2005. 2006. 2007. 2008. 2009. 2010. 2011. 2012. 2013. 2014. 2015. 2016. 2017. 2018. 2019. 2020. 2021. 2022. 2023. 2024. 2025. 2026. 2027. 2028. 2029. 2030. 2031. 2032. 2033. 2034. 2035. 2036. 2037. 2038. 2039. 2040. 2041. 2042. 2043. 2044. 2045. 2046. 2047. 2048. 2049. 2050. 2051. 2052. 2053. 2054. 2055. 2056. 2057. 2058. 2059. 2060. 2061. 2062. 2063. 2064. 2065. 2066. 2067. 2068. 2069. 2070. 2071. 2072. 2073. 2074. 2075. 2076. 2077. 2078. 2079. 2080. 2081. 2082. 2083. 2084. 2085. 2086. 2087. 2088. 2089. 2090. 2091. 2092. 2093. 2094. 2095. 2096. 2097. 2098. 2099. 2100. 2101. 2102. 2103. 2104. 2105. 2106. 2107. 2108. 2109. 2110. 2111. 2112. 2113. 2114. 2115. 2116. 2117. 2118. 2119. 2120. 2121. 2122. 2123. 2124. 2125. 2126. 2127. 2128. 2129. 2130. 2131. 2132. 2133. 2134. 2135. 2136. 2137. 2138. 2139. 2140. 2141. 2142. 2143. 2144. 2145. 2146. 2147. 2148. 2149. 2150. 2151. 2152. 2153. 2154. 2155. 2156. 2157. 2158. 2159. 2160. 2161. 2162. 2163. 2164. 2165. 2166. 2167. 2168. 2169. 2170. 2171. 2172. 2173. 2174. 2175. 2176. 2177. 2178. 2179. 2180. 2181

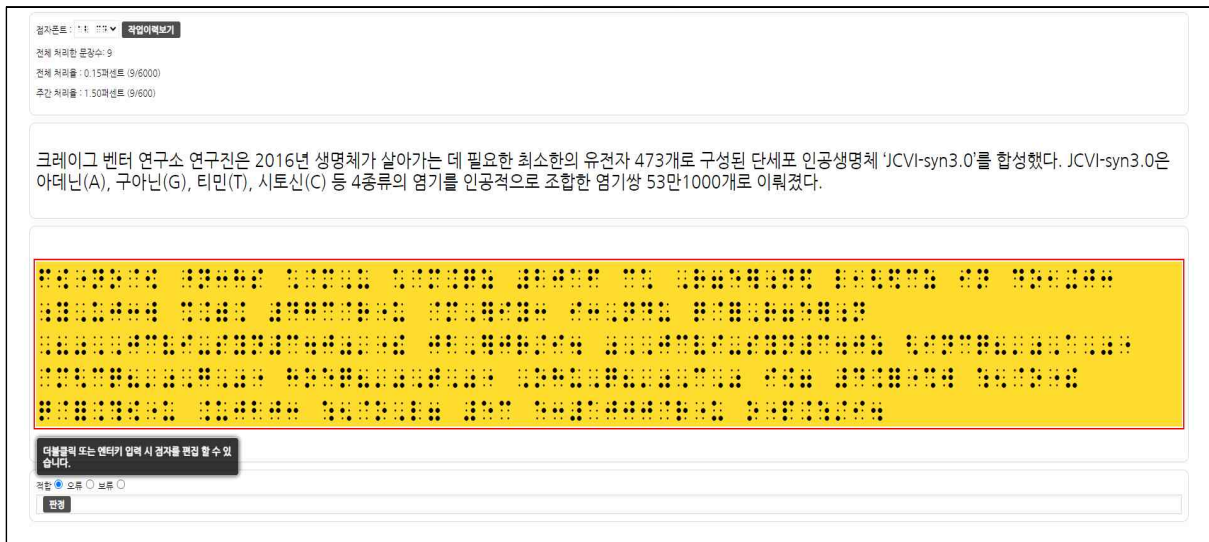


그림 2-15 오류/보류 검수 화면에서 수정 전 선택 화면



그림 2-16 점자 점형 교정 새 창 화면

### (7) 작업 이력 보기

- 본인이 검수한 문장은 작업 이력 보기 버튼 클릭 후 리스트에서 확인 및 수정 가능
- 화면에 최대 10개의 문장이 보이며, 더보기 버튼 선택 시 10개 문장씩 추가로 보임



그림 2-17 작업 이력 보기 화면



그림 2-18 작업 이력 보기 클릭 후 화면

### (8) 작업 이력 보기 상세

- 주차별/판정별 선택 확인 가능
- 특정 주차 내역 또는 특정 판정 내역만 확인이 가능하며 해당 버튼 클릭
- 주차별: 전체보기, 1주차(10.30~), 2주차(11.6~), 3주차(11.13~), 4주차(11.20~), 5주차(11.27~), 6주차(12.4~), 7주차(12.11~), 8주차(12.18~), 9주차(12.25~), 10주차(1.1~), 11주차(1.8~)
- A/C그룹은 1~10주차, B그룹은 2~11주차에 작업 이력이 나타남
- 판정별: 전체보기, 적합, 오류, 보류
- 판정 내역 옆에 소괄호로 해당 판정 숫자가 기재되어 있음(ex: 적합(513))



그림 2-19 주차별 검수 문장 분류 메뉴



그림 2-20 판정별 검수 문장 분류 메뉴

- 수정 이력별 내역 확인 가능
- 판정 수정한 문장은 수정 이력에 기록되며, 메뉴 클릭 후 확인 가능
- 판정을 수정하지 않은 문장은 미수정에 기록되며, 메뉴 클릭 후 확인 가능
- 판정과 마찬가지로 옆에 소괄호로 해당 판정 숫자가 기재되어 있음(ex: 수정 이력(29))

그림 2-21 수정 이력별 검수 문장 분류 메뉴

- 작업 이력 화면 내 수정 가능
- 작업 이력 화면에서 수정 희망 시 검수 방법과 동일하게 판정 후 수정 가능
- 적합 → 오류/보류: 콤보박스에서 해당 사유 선택 후 텍스트 박스에 틀린 단어 기재 뒤 수정하기 눌러야 변경 적용
- 오류/보류 → 적합: 적합 라디오 버튼 선택 후 수정하기 눌러야 변경 적용
- '수정하기' 클릭 시 '수정하시겠습니까?' 팝업이 뜨며 확인/취소 버튼 활성화
- A그룹의 경우, B그룹이 해당 문장을 판정했으면 수정 불가하며, 판정 버튼이 '2차 검수가 처리되어 수정할 수 없습니다' 텍스트로 변경

그림 2-22 작업 이력 화면 내 수정 시 화면



그림 2-23 B그룹이 문장 판정하여 수정 불가능 화면

### (9) C그룹 검수 화면

- 화면에는 A/B그룹과 동일하게 윗줄에는 목자, 아랫줄에는 점자가 표기됨
- 목자 상단에는 본인이 받은 적합/오류/보류 문장 수가 표기되어 있음
- C그룹은 A/B그룹의 작업 이력 보기 버튼 대신 상단에 주차 내역이 있으며, 메뉴 건너뛰기 버튼 선택 시 바로 문장으로 이동
- 센스리더 사용자는 점자 영역에 초점 접근 후 Home키 입력 후 방향키로 이동하면 아스키 코드(ex. 135점, 24점) 형태로 음성 출력됨
- 비시각의 경우, 점자 폰트 가독성 및 편의를 위해 3종의 점자 폰트로 변경 가능



그림 2-24 말뭉치 관리시스템 C그룹 목자-점자 검수 화면

## (10) 적합/오류/보류 버튼 체크 및 사유 기재

- 판정 기준 및 오류/보류 사유 기재, 점형 교정 방법은 A/B그룹과 동일
- C그룹은 점자 점형 수정 후 판정에서 적합으로 변경 후 수정하기 버튼 선택

## (11) 작업 이력 보기

- C그룹은 A/B그룹의 작업 이력 보기 버튼 대신 상단에 주차 내역이 있으며, 메뉴 건너뛰기 버튼 선택 시 바로 문장으로 이동
- A/B그룹과 달리 주차별/판정별/수정 이력별 선택 확인은 없으나 메뉴 건너뛰기 버튼 우측에 판정별 선택 콤보박스로 선택 확인 가능
- 작업 이력 화면 내 수정은 A/B그룹과 동일

그림 2-25 C그룹 작업 이력 보기 화면

그림 2-26 C그룹 판정 상태 콤보박스



2023년 목자-점자 병렬말뭉치 구축  
(1. 목자-점자 병렬말뭉치 구축)

# 제 3 장

## 목자-점자 병렬말뭉치 검수 및 통계 현황

1. 검수 방법
2. 조사 결과





# 1 검수 방법

## 1) 검수 참여자

본 사업의 목적을 위한 검수 참여자의 선정은 목적적 표집(Purposeful Sampling)을 통해 이루어졌다. 검수 참여자는 다양한 시각장애 관련 단체에 소속되어 있거나 점자 편집 업무를 담당하고 있는 점역·교정사 중에서 본 사업의 목적에 맞는 대표자들을 선정하였다.

검수 참여자의 선정 과정에서 점역·교정사 자격증 보유 현황, 취득 과목 여부를 고려하여 검수자 64명을 선정하였다. 본 사업에 참여한 점역·교정사의 구성은 시각장애인 점역·교정사 17명, 비장애인 점역·교정사 47명이었다.

검수 참여자의 점역·교정사 자격 급수를 보면, 1급 소지자가 총 54명(84.4%)으로 가장 많았고, 자격 유형은 국어 외에 영어 자격 소지자가 63명(98.4%)으로 가장 많았으며, 그 다음으로는 수학·과학·컴퓨터, 음악, 일본어의 순이었다.

검수자는 A/B/C 3그룹으로 나뉘 목자와 점역된 점자 문장을 비교하여 점역이 바르게 되었는지 내용을 확인하고, 문장의 오류와 보류 내용으로 서술하였다. A그룹은 점역 판정 후 오류/보류 문장을 점자로 바르게 교정하였다. B그룹은 A그룹이 검수한 검수 결과가 정확한지 문장을 재판정하여 교차 검수하였다. C그룹은 2021~2022년에 구축한 목자-점자 병렬말뭉치 오류/보류 문장을 점자로 바르게 교정하였다.

구체적인 검수자 정보는 <표 3-1>과 같다.

표 3-1 검수 참여자 정보(64명)

연번	이 름	그룹	할당 검수 수(문장)	장애 구분	자격증 급수	취득 과목
1	공 *	A그룹	5,000	중증	1급	국어/영어/수과컴/음악/일본어
2	김 * 윤	A그룹	5,000	-	1급	국어/영어/수과컴
3	김 * 경	A그룹	4,500	-	1급	국어/영어/수과컴
4	김 * 연	A그룹	5,000	-	1급	국어/영어/수과컴
5	김 * 기	A그룹	5,500	-	1급	국어/영어/수과컴/음악
6	김 * 희	A그룹	5,500	-	1급	국어/영어/수과컴
7	김 * 란	A그룹	4,000	경증	1급	국어/영어/수과컴
8	박 * 연	A그룹	5,000	-	1급	국어/영어/음악

연번	이 름	그룹	할당 검수 수(문장)	장애 구분	자격증 급수	취득 과목
9	박 * 영	A그룹	5,000	중증	1급	국어/영어/수과컴
10	봉 * 용	A그룹	5,000	중증	1급	국어/영어/음악
11	서 * 옥	A그룹	5,000	-	1급	국어/영어/수과컴
12	서 * 희	A그룹	5,000	-	1급	국어/영어/수과컴/일본어
13	신 * 비	A그룹	5,000	-	1급	국어/영어/수과컴
14	유 * 연	A그룹	5,000	-	1급	국어/영어/수과컴
15	윤 * 희	A그룹	5,000	-	1급	국어/영어/수과컴
16	이 * 연	A그룹	4,000	-	1급	국어/영어/일본어
17	이 * 정	A그룹	5,000	-	1급	국어/영어/수과컴
18	이 * 정	A그룹	5,000	-	1급	국어/영어/수과컴
19	임 * 성	A그룹	5,500	-	1급	국어/영어/음악
20	장 * 진	A그룹	5,000	-	2급	국어/영어
21	전 * 희	A그룹	5,500	-	2급	국어/수과컴
22	전 * 인	A그룹	5,000	-	1급	국어/영어/수과컴
23	정 * 정	A그룹	5,000	-	2급	국어/영어
24	조 * 희	A그룹	5,500	-	2급	국어/영어
25	카 * * * * 카	A그룹	5,000	-	1급	국어/영어/일본어
26	하 * 리	A그룹	5,000	중증	1급	국어/영어/음악/일본어
27	한 * 란	A그룹	5,000	-	1급	국어/영어/수과컴/음악
28	강 * 혜	B그룹	5,000	-	1급	국어/영어/수과컴
29	강 * 원	B그룹	5,000	-	1급	국어/영어/음악
30	고 * 진	B그룹	5,000	중증	1급	국어/영어/일본어
31	김 * 준	B그룹	5,000	중증	1급	국어/영어/수과컴
32	김 * 혜	B그룹	5,000	-	1급	국어/영어/음악
33	박 * 화	B그룹	5,000	-	1급	국어/영어/수과컴
34	박 * 규	B그룹	4,000	중증	1급	국어/영어/수과컴
35	박 * 재	B그룹	5,000	중증	1급	국어/영어/음악
36	박 * 호	B그룹	5,000	-	2급	국어/영어
37	박 * 미	B그룹	5,000	-	2급	국어/영어
38	박 * 정	B그룹	4,000	중증	1급	국어/영어/수과컴

연번	이 름	그룹	할당 검수 수(문장)	장애 구분	자격증 급수	취득 과목
39	박 * 조	B그룹	5,000	-	2급	국어/영어
40	양 * 린	B그룹	5,500	-	1급	국어/영어/수과컴
41	윤 * 정	B그룹	5,000	중증	1급	국어/영어/수과컴
42	이 * 내	B그룹	5,000	-	2급	국어/영어
43	이 * 진	B그룹	5,000	중증	1급	국어/영어/수과컴
44	이 * 연	B그룹	5,000	-	1급	국어/영어/수과컴
45	이 * 미	B그룹	5,500	-	1급	국어/영어/음악
46	이 * 찬	B그룹	5,000	중증	1급	국어/영어/음악
47	전 * 민	B그룹	5,000	-	1급	국어/영어/음악
48	정 * 영	B그룹	5,000	중증	1급	국어/영어/음악
49	정 * 수	B그룹	5,000	-	2급	국어/영어
50	조 * 영	B그룹	5,000	-	1급	국어/영어/수과컴
51	조 * 영	B그룹	5,000	-	1급	국어/영어/음악
52	진 * 라	B그룹	5,500	-	1급	국어/영어/수과컴/음악
53	한 * 경	B그룹	5,000	중증	2급	국어/영어
54	한 * 진	B그룹	5,500	-	1급	국어/영어/수과컴
55	김 * 정	C그룹	5,000	-	1급	국어/영어/수과컴/음악
56	김 * 영	C그룹	5,000	-	1급	국어/영어/수과컴
57	박 * 연	C그룹	5,000	-	1급	국어/영어/수과컴/음악/일본어
58	신 * 규	C그룹	5,000	-	1급	국어/영어/수과컴
59	안 * 진	C그룹	5,000	-	1급	국어/영어/수과컴
60	오 * 리	C그룹	5,500	-	1급	국어/영어/수과컴/일본어
61	이 * 비	C그룹	5,500	-	1급	국어/영어/음악
62	하 * 섭	C그룹	5,000	중증	1급	국어/영어/수과컴/음악/일본어
63	하 * 총	C그룹	5,000	-	1급	국어/영어/수과컴
64	한 * 수	C그룹	4,000	중증	1급	국어/영어/수과컴

## 2) 검수 단계

이 사업에서는 1단계로 목자와 점역된 점자 문장을 비교하여 점역이 바르게 되었는지 내용을 확인하는 1차 검수, 2단계에서 다른 검수자의 적합 검수 문장을 확인하는 2차 검수를 수행하였다.

1차 검수 단계는 A그룹과 C그룹, 2차 검수는 B그룹이 작업하였다. 검수 결과는 주 단위로 관리되며 재확인 검수를 통하여 검수 신뢰도를 높일 수 있도록 처리하였다.

### (1) 1차 검수 단계

목자-점자 병렬말뭉치 구축을 위한 1차 검수는 두 가지로 나누어 기록하도록 구성하였다.

첫째, 검수 결과, 적합/오류/보류 여부를 결정하도록 되어 있다. 즉 적합/오류/보류 중 하나를 선택하여야 한다. 적합은 검수하였을 때 다 맞게 점역되어 있고, 오류는 목자의 내용이 바르게 점역되지 않은 경우이고, 보류는 한글 점자 규정에 정의되지 않은 내용/기호가 있거나 목자 내용이 제2외국어, 수식, 화학 공식, 악보 등과 같이 한글/영어 규정에 해당하지 않는 경우이다.

둘째, 오류/보류의 경우 오류/보류라고 판단한 내용을 표기하고 점역이 잘못된 부분을 바르게 교정하도록 되어 있다. 예를 들어 영어 단어 다음에 로마자 종료표가 빠짐, 영어 단어 앞에 1급 점자 표 오기, 한글 점자 규정에 나와 있지 않은 글머리 기호, 한글 점자 규정에서 정의되지 않은 점형 등 어떤 내용에서 점역이 잘못됐는지 해당 내용을 표기하도록 되어 있다. 표기 후 잘못 점역된 부분은 정확한 점형으로 교정 후 다음 검수 문장을 진행하도록 되어 있다.

1차 검수 작업하는 A그룹, C그룹 검수자 정보는 <표 3-2>와 같다.

**표 3-2** 1차 검수자(A그룹, C그룹) 정보(37명)

연번	이름	그룹	할당 검수 수(문장)	장애 구분	자격증 급수	취득 과목
1	공 * *	A그룹	5,000	중증	1급	국어/영어/수과컴/음악/일본어
2	김 * 윤	A그룹	5,000	-	1급	국어/영어/수과컴
3	김 * 경	A그룹	4,500	-	1급	국어/영어/수과컴
4	김 * 연	A그룹	5,000	-	1급	국어/영어/수과컴
5	김 * 기	A그룹	5,500	-	1급	국어/영어/수과컴/음악
6	김 * 희	A그룹	5,500	-	1급	국어/영어/수과컴
7	김 * 란	A그룹	4,000	경증	1급	국어/영어/수과컴
8	박 * 연	A그룹	5,000	-	1급	국어/영어/음악
9	박 * 영	A그룹	5,000	중증	1급	국어/영어/수과컴
10	봉 * 용	A그룹	5,000	중증	1급	국어/영어/음악
11	서 * 옥	A그룹	5,000	-	1급	국어/영어/수과컴
12	서 * 희	A그룹	5,000	-	1급	국어/영어/수과컴/일본어
13	신 * 비	A그룹	5,000	-	1급	국어/영어/수과컴

연번	이 름	그룹	할당 검수 수(문장)	장애 구분	자격증 급수	취득 과목
14	유 * 연	A그룹	5,000	-	1급	국어/영어/수과컴
15	윤 * 희	A그룹	5,000	-	1급	국어/영어/수과컴
16	이 * 연	A그룹	4,000	-	1급	국어/영어/일본어
17	이 * 정	A그룹	5,000	-	1급	국어/영어/수과컴
18	이 * 정	A그룹	5,000	-	1급	국어/영어/수과컴
19	임 * 성	A그룹	5,500	-	1급	국어/영어/음악
20	장 * 진	A그룹	5,000	-	2급	국어/영어
21	전 * 희	A그룹	5,500	-	2급	국어/수과컴
22	전 * 인	A그룹	5,000	-	1급	국어/영어/수과컴
23	정 * 정	A그룹	5,000	-	2급	국어/영어
24	조 * 희	A그룹	5,500	-	2급	국어/영어
25	카 * * * * 카	A그룹	5,000	-	1급	국어/영어/일본어
26	하 * 리	A그룹	5,000	중증	1급	국어/영어/음악/일본어
27	한 * 란	A그룹	5,000	-	1급	국어/영어/수과컴/음악
28	김 * 정	C그룹	5,000	-	1급	국어/영어/수과컴/음악
29	김 * 영	C그룹	5,000	-	1급	국어/영어/수과컴
30	박 * 연	C그룹	5,000	-	1급	국어/영어/수과컴/음악/일본어
31	신 * 규	C그룹	5,000	-	1급	국어/영어/수과컴
32	안 * 진	C그룹	5,000	-	1급	국어/영어/수과컴
33	오 * 리	C그룹	5,500	-	1급	국어/영어/수과컴/일본어
34	이 * 비	C그룹	5,500	-	1급	국어/영어/음악
35	하 * 섭	C그룹	5,000	중증	1급	국어/영어/수과컴/음악/일본어
36	하 * 총	C그룹	5,000	-	1급	국어/영어/수과컴
37	한 * 수	C그룹	4,000	중증	1급	국어/영어/수과컴

## (2) 2차 검수 단계

목자-점자 병렬말뭉치 구축을 위한 2차 검수는 1차 검수에서 판정된 135,000문장을 할당해 검수하였다. 검수 방법은 이전과 동일하게 적합/오류/보류를 결정하는 방법을 선택하였다. C그룹이 검수한 기존 구축 목자-점자 병렬말뭉치 오류·보류 5,000문장은 해당 연도 구축 시 이미 오류/보류 판정을 받았으므로 점형 교정에 의미를 두고 2차 검수를 진행하지 않았다.

2차 검수 작업하는 B그룹 검수자 정보는 <표 3-3>과 같다.

표 3-3 2차 검수자(B그룹) 정보(27명)

연번	이 름	그룹	할당 검수 수(문장)	장애 구분	자격증 급수	취득 과목
1	강 * 혜	B그룹	5,000	-	1급	국어/영어/수과컴
2	강 * 원	B그룹	5,000	-	1급	국어/영어/음악
3	고 * 진	B그룹	5,000	중증	1급	국어/영어/일본어
4	김 * 준	B그룹	5,000	중증	1급	국어/영어/수과컴
5	김 * 혜	B그룹	5,000	-	1급	국어/영어/음악
6	박 * 화	B그룹	5,000	-	1급	국어/영어/수과컴
7	박 * 규	B그룹	4,000	중증	1급	국어/영어/수과컴
8	박 * 재	B그룹	5,000	중증	1급	국어/영어/음악
9	박 * 호	B그룹	5,000	-	2급	국어/영어
10	박 * 미	B그룹	5,000	-	2급	국어/영어
11	박 * 정	B그룹	4,000	중증	1급	국어/영어/수과컴
12	박 * 조	B그룹	5,000	-	2급	국어/영어
13	양 * 린	B그룹	5,500	-	1급	국어/영어/수과컴
14	윤 * 정	B그룹	5,000	중증	1급	국어/영어/수과컴
15	이 * 내	B그룹	5,000	-	2급	국어/영어
16	이 * 진	B그룹	5,000	중증	1급	국어/영어/수과컴
17	이 * 연	B그룹	5,000	-	1급	국어/영어/수과컴
18	이 * 미	B그룹	5,500	-	1급	국어/영어/음악
19	이 * 찬	B그룹	5,000	중증	1급	국어/영어/음악
20	전 * 민	B그룹	5,000	-	1급	국어/영어/음악
21	정 * 영	B그룹	5,000	중증	1급	국어/영어/음악
22	정 * 수	B그룹	5,000	-	2급	국어/영어
23	조 * 영	B그룹	5,000	-	1급	국어/영어/수과컴
24	조 * 영	B그룹	5,000	-	1급	국어/영어/음악
25	진 * 라	B그룹	5,500	-	1급	국어/영어/수과컴/음악
26	한 * 경	B그룹	5,000	중증	2급	국어/영어
27	한 * 진	B그룹	5,500	-	1급	국어/영어/수과컴

### 3) 검수 교육 절차

본 사업에서는 목자와 점역된 점자 문장을 비교하여 점역이 바르게 되었는지 내용을 확인하는 검수자를 모집하여 참여자 교육을 실시하였다. 검수 참여자 교육은 대면 사전교육으로 이론 교육과 실습 교육으로 구성하였다. 이론 교육은 2023년 10월 20일 오후 2시부터 오후 5시까지 3시간 동안 실시하였다.

검수 참여자 이론 교육 내용은 첫째, 국립국어원의 ‘2023년 목자-점자 병렬말뭉치 구축’ 사업을 소개하였다. 둘째, 목자-점자 병렬말뭉치 구축의 목적인 말뭉치 십만 문장(백만 어절 이상) 구축과 웹 기반 관리시스템 등을 안내하였다. 주요 내용은 목자와 점역된 점자 문장을 비교하여 점역이 바르게 되었는지 내용을 확인하고, 웹 기반 관리시스템에서 판정하기 위한 검수 실습 등으로 진행하였다. 셋째, 변경될 한글 점자 규정에 맞춰 검수 및 교정을 진행해야 하므로 개정 한국 점자 규정(문화체육관광부 고시 제2024-0005호, 이하 “개정 한국 점자 규정”이라 함) 주요 내용을 교육하였다.

목자-점자 병렬말뭉치 구축을 위한 검수 작업은 우선 1차 검수는 2023년 10월 30일부터 2024년 1월 7일까지 약 10주 동안 이루어진다. 2차 검수는 2023년 11월 6일부터 2024년 1월 14일까지 약 10주 동안 이루어진다.

구체적 1차 검수 일정과 2차 검수 일정은 <표 3-4>와 같다.

표 3-4 검수 일정

구 분	날 짜	내 용
1주	10월 30일 ~ 11월 5일	- A/C그룹 검수자 작업 시작 - 누적 검수량: A/C그룹 500문장
2주	11월 6일~11월 12일	- B그룹 검수자 작업 시작 - 누적 검수량: A/C그룹 1,000문장, B그룹 500문장
3주	11월 13일~11월 19일	- 누적 검수량: A/C그룹 1,500문장, B그룹 1,000문장
4주	11월 20일~11월 26일	- 누적 검수량: A/C그룹 2,000문장, B그룹 1,500문장
5주	11월 27일~12월 3일	- 누적 검수량: A/C그룹 2,500문장, B그룹 2,000문장
6주	12월 4일~12월 10일	- 누적 검수량: A/C그룹 3,000문장, B그룹 2,500문장
7주	12월 11일~12월 17일	- 누적 검수량: A/C그룹 3,500문장, B그룹 3,000문장
8주	12월 18일~12월 24일	- 누적 검수량: A/C그룹 4,000문장, B그룹 3,500문장
9주	12월 25일~12월 31일	- 누적 검수량: A/C그룹 4,500문장, B그룹 4,000문장
10주	1월 1일~1월 7일	- A/C그룹 검수자 작업 완료 - 누적 검수량: A/C그룹 5,000문장, B그룹 4,500문장
11주	1월 8일~1월 14일	- B그룹 검수자 작업 완료 - 누적 검수량: B그룹 5,000문장

#### 4) 자료 처리

검수자는 웹 기반 관리시스템에 접속하여 자신에게 할당된 문장을 처리하고, 이때 할당된 문장은 미리 할당하는 방식이 아니라 자동으로 시스템에서 배정하는 방식으로 접속 시점에 아직 검수되지 않은 문장이 자동으로 배정되게 하였다. 각각의 검수자가 처리한 총 검수 수, 누적 검수율, 오류 수, 적합 수, 보류 수, 오류율, 총 어절 수, 총 문장 수 등은 시스템에 기록되어 관리된다. 검수자는 관리자가 지정한 문장 수가 될 때까지 시스템에 접속할 때마다 처리할 문장을 지정받는다. 검수자별 처리 현황과 전체 통계 현황 진행률(처리 문장/전체 문장)을 실시간/시각적으로 확인할 수 있다.

웹 기반 관리시스템에서 검수한 목자-점자 병렬말뭉치는 엑셀 파일 형식으로 변환할 수 있다. 코딩 값은 해당 항목에 '개수'로 표기하였으며, 각 검수 결과별 적합 수, 오류 수, 보류 수 값을 합산하여 전체 검수 현황 결과를 산출했다. 같은 방식으로 적합 검수별로 어절 수, 문장 수 값을 분류하여 세부 검수별 어절 현황, 문장 현황을 산출했다. 오류 및 보류의 사유별로도 세분화하였다.

## 2 조사 결과

### 1) 1차 검수 데이터 통계 현황

1차 검수 단계인 A그룹, C그룹 검수 데이터 구축 결과는 각각 <표 3-5>, <표 3-6>과 같다. A그룹 검수 데이터의 검수 결과(1주~10주: 10월 30일~1월 7일 기준), 총 검수 수는 전체 135,000문장으로 누적 검수율은 100%로 조사되었다.

적합 수는 전체 135,000문장 중 128,682문장으로 적합률은 95.32%로 조사되었다. 적합 어절 수는 2,626,617개이다. 오류 수는 전체 135,000문장 중 5,307문장으로 오류율은 3.93%로 조사되었고, 보류 수는 1,011문장인 0.75%로 낮게 조사되었다.

표 3-5 A그룹 검수 데이터 통계 현황

주차	기 간	총 검수 수 (135,000 문장)	주차 검수율 (%)	검수 결과								총 어절 수
				적합		오류		보류		계		
				적합	어절 수	적합률 (%)	오류	오류율 (%)	보류		보류율 (%)	
1주	10월 30일 ~11월 5일	12,623	9.35	11,734	227,956	92.96	746	5.91	143	1.13	12,623	244,754
2주	11월 6일 ~11월 12일	13,103	9.71	12,274	251,270	93.67	757	5.78	72	0.55	13,103	268,926
3주	11월 13일 ~11월 19일	13,052	9.67	12,396	248,702	94.97	510	3.91	146	1.12	13,052	262,622
4주	11월 20일 ~11월 26일	13,936	10.32	13,010	263,826	93.36	719	5.16	207	1.49	13,936	282,705
5주	11월 27일 ~12월 3일	13,794	10.22	13,009	267,588	94.31	628	4.55	157	1.14	13,794	284,277
6주	12월 4일 ~12월 10일	13,959	10.34	13,383	266,751	95.87	461	3.3	115	0.82	13,959	278,600
7주	12월 11일 ~12월 17일	13,794	10.22	13,349	287,815	96.77	392	2.84	53	0.38	13,794	297,736
8주	12월 18일 ~12월 24일	13,794	10.22	13,448	276,211	97.49	312	2.26	34	0.25	13,794	283,648
9주	12월 25일 ~12월 31일	13,891	10.29	13,482	275,692	97.06	365	2.63	44	0.32	13,891	284,741
10주	1월 1일 ~1월 7일	13,054	9.66	12,597	260,806	96.5	417	3.19	40	0.31	13,054	270,743
합 계		135,000	100	128,682	2,626,617	95.32	5,307	3.93	1,011	0.75	135,000	2,758,752



그림 3-1 A그룹 검수 적합/오류/보류 문장 수



그림 3-2 A그룹 검수 주차별 검수 문장 수 및 누적 검수율

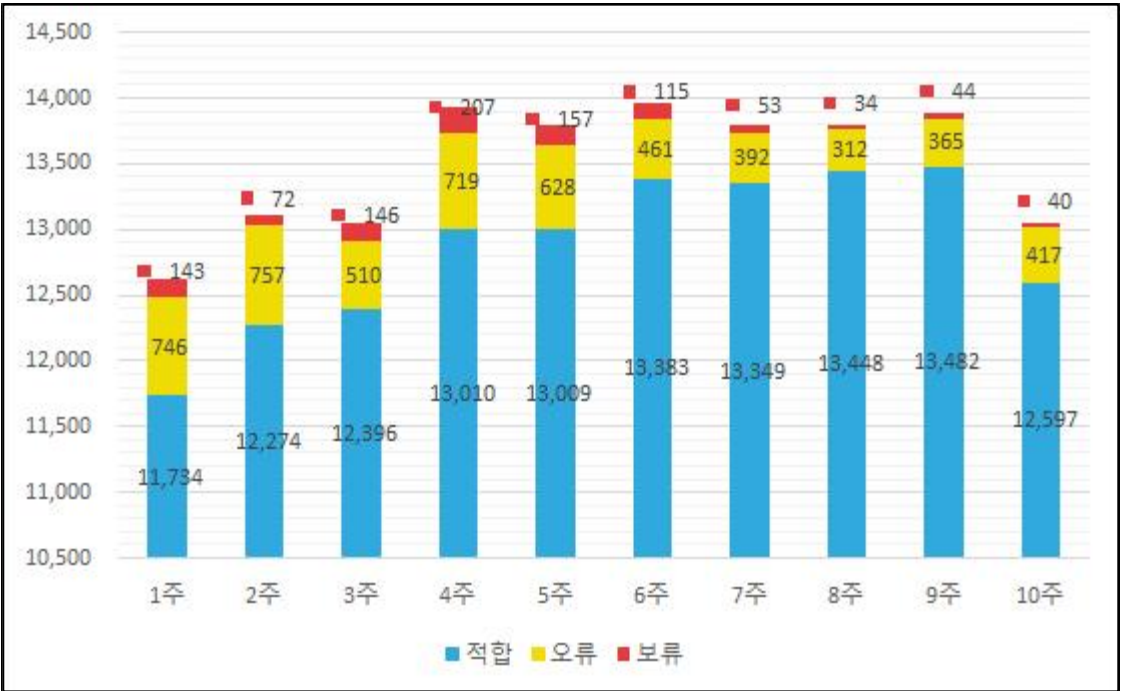


그림 3-3 A그룹 검수 주차별 적합/오류/보류 문장 수

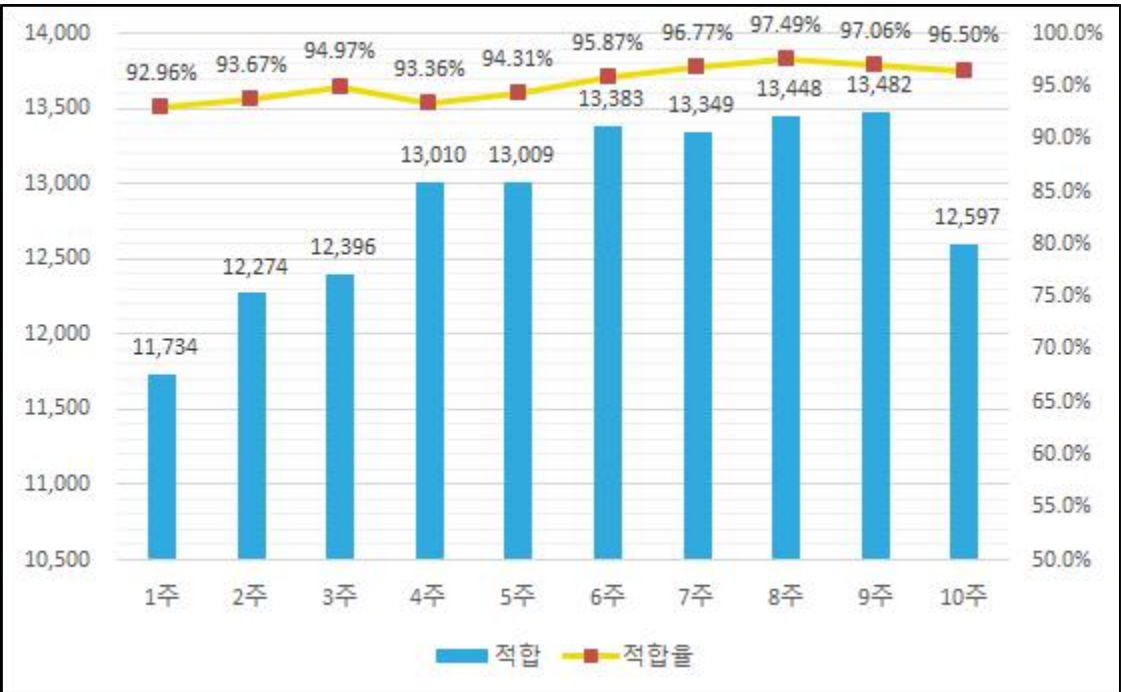


그림 3-4 A그룹 검수 주차별 최종 적합 문장 수 및 적합률

C그룹 검수 데이터의 검수 결과(1주~10주: 10월 30일~1월 7일 기준), 총 검수 수는 전체 50,000문장으로 누적 검수율은 100%로 조사되었다.

적합 수는 전체 50,000문장 중 46,990문장으로 적합률은 93.98%로 조사되었다. 적합 어절 수는 1,446,905개이다. 오류 수는 전체 50,000문장 중 0문장으로 오류율은 0%로 조사되었고, 보류 수는 3,010문장인 6.02%로 낮게 조사되었다.

표 3-6 C그룹 검수 데이터 통계 현황

주차	기 간	총 검수 수 (135,000 문장)	주차 검수율 (%)	검수 결과								총 어절 수
				적합			오류		보류		계	
				적합	어절 수	적합률 (%)	오류	오류율 (%)	보류	보류율 (%)		
1주	10월 30일 ~11월 5일	4,953	9.91	4,672	101,748	94.33	0	0	281	5.67	4,953	110,178
2주	11월 6일 ~11월 12일	4,814	9.63	4,462	96,934	92.69	0	0	352	7.31	4,814	105,172
3주	11월 13일 ~11월 19일	4,953	9.91	4,601	101,127	92.89	0	0	352	7.11	4,953	109,945
4주	11월 20일 ~11월 26일	5,515	11.03	5,146	113,207	93.31	0	0	369	6.69	5,515	121,305
5주	11월 27일 ~12월 3일	5,209	10.42	4,851	119,400	93.13	0	0	358	6.87	5,209	128,619
6주	12월 4일 ~12월 10일	4,977	9.95	4,750	189,610	95.44	0	0	227	4.56	4,977	202,525
7주	12월 11일 ~12월 17일	5,724	11.45	5,426	232,864	94.79	0	0	298	5.21	5,724	254,032
8주	12월 18일 ~12월 24일	5,063	10.13	4,729	175,889	93.40	0	0	334	6.60	5,063	198,574
9주	12월 25일 ~12월 31일	4,451	8.90	4,252	160,702	95.53	0	0	199	4.47	4,451	170,504
10주	1월 1일 ~1월 7일	4,341	8.68	4,101	155,424	94.47	0	0	240	5.53	4,341	170,320
합 계		50,000	100	46,990	1,446,905	93.98	0	0	3,010	6.02	50,000	1,571,174



그림 3-5 C그룹 검수 적합/오류/보류 문장 수

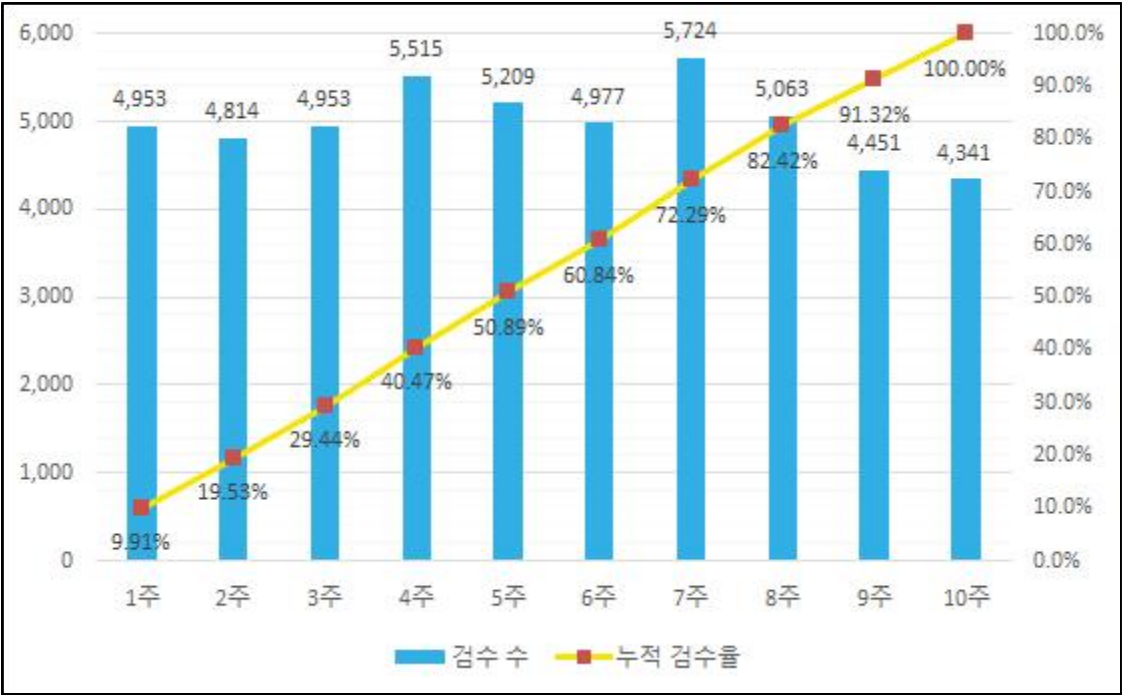


그림 3-6 C그룹 검수 주차별 검수 문장 수 및 누적 검수율

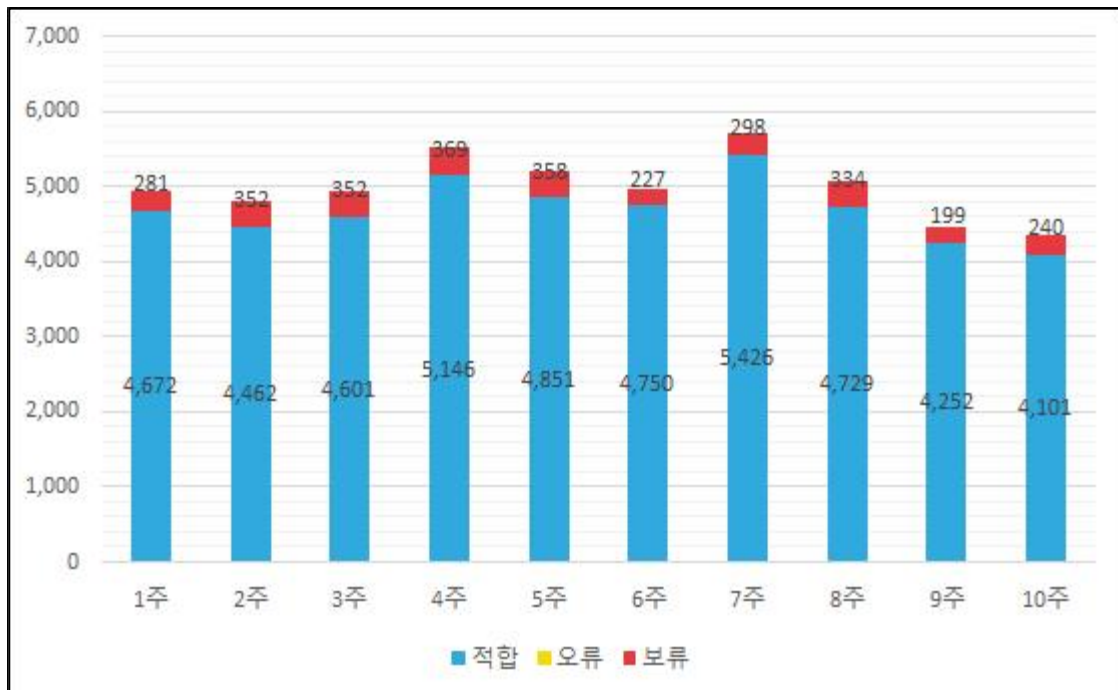


그림 3-7 C그룹 검수 주차별 적합/오류/보류 문장 수

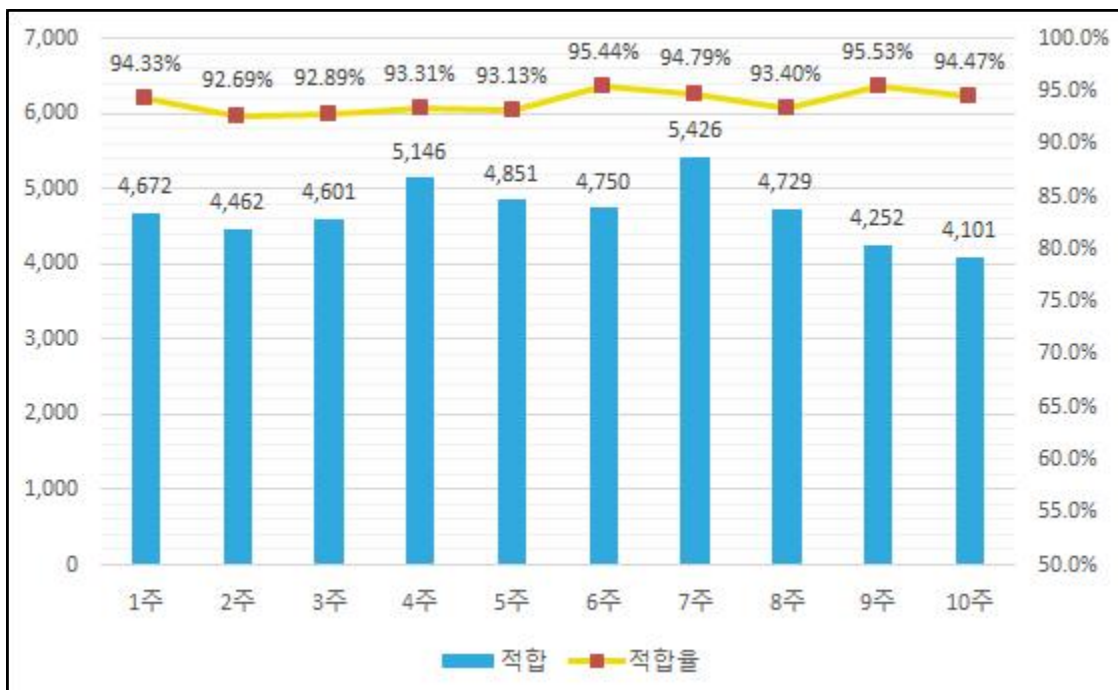


그림 3-8 C그룹 검수 주차별 최종 적합 문장 수 및 적합률

## 2) 2차 검수 데이터 통계 현황

2차 검수 데이터 통계 현황을 조사하였다. 2차 검수 단계인 B그룹 검수 데이터 구축 결과는 <표 3-7>과 같다. B그룹 검수 데이터의 검수 결과(1주~10주: 11월 6일~1월 14일 기준), 총 검수 수는 전체 135,000문장으로 누적 검수율은 100%로 조사되었다.

적합 수는 전체 135,000문장 중 131,509문장으로 적합률은 97.41%로 조사되었다. 적합 어절 수는 2,685,953개이다. 오류 수는 전체 135,000문장 중 2,719문장으로 오류율은 2.01%로 조사되었고, 보류 수는 772문장인 0.57%로 낮게 조사되었다.

표 3-7 B그룹 검수 데이터 통계 현황

주차	기 간	총 검수 수 (135,000 문장)	주차 검수율 (%)	검수 결과								총 어절 수
				적합		오류		보류		계		
				적합	어절 수	적합률 (%)	오류	오류율 (%)	보류		보류율 (%)	
1주	11월 6일 ~11월 12일	12,732	9.43	12,176	236,847	95.63	433	3.4	123	0.97	12,732	247,059
2주	11월 13일 ~11월 19일	13,537	10.03	13,027	266,268	96.23	389	2.87	121	0.89	13,537	277,431
3주	11월 20일 ~11월 26일	13,926	10.32	13,637	274,313	97.92	210	1.51	79	0.57	13,926	280,459
4주	11월 27일 ~12월 3일	13,303	9.85	12,909	262,531	97.04	249	1.87	145	1.09	13,303	270,610
5주	12월 4일 ~12월 10일	13,932	10.32	13,518	277,357	97.03	332	2.38	82	0.59	13,932	286,096
6주	12월 11일 ~12월 17일	13,849	10.26	13,505	270,937	97.52	281	2.03	63	0.45	13,849	277,989
7주	12월 18일 ~12월 24일	13,158	9.75	12,908	277,928	98.1	212	1.61	38	0.29	13,158	283,535
8주	12월 25일 ~12월 31일	13,798	10.22	13,554	278,468	98.23	205	1.49	39	0.28	13,798	283,728
9주	1월 1일 ~1월 7일	14,493	10.74	14,244	291,828	98.28	204	1.41	45	0.31	14,493	297,266
10주	1월 8일 ~1월 15일	12,272	9.09	12,031	249,476	98.04	204	1.66	37	0.3	12,272	254,579
합 계		135,000	100	131,509	2,685,953	97.41	2,719	2.01	772	0.57	135,000	2,758,752



그림 3-9 B그룹 검수 적합/오류/보류 문장 수

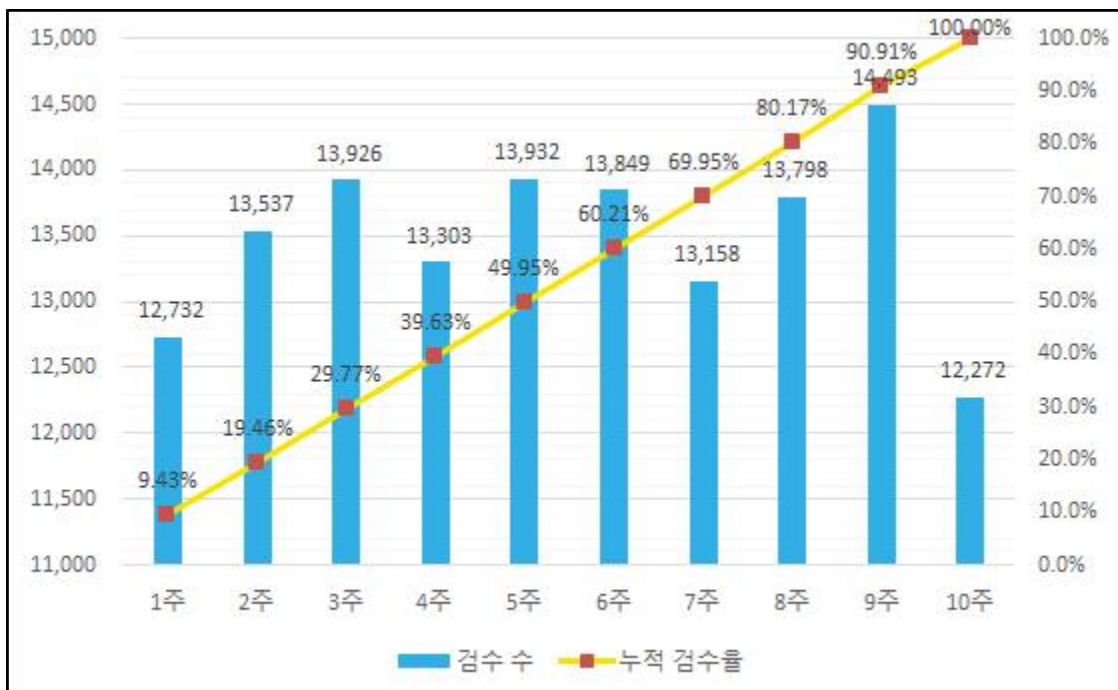


그림 3-10 B그룹 검수 주차별 검수 문장 수 및 누적 검수율



그림 3-11 B그룹 검수 주차별 적합/오류/보류 문장 수

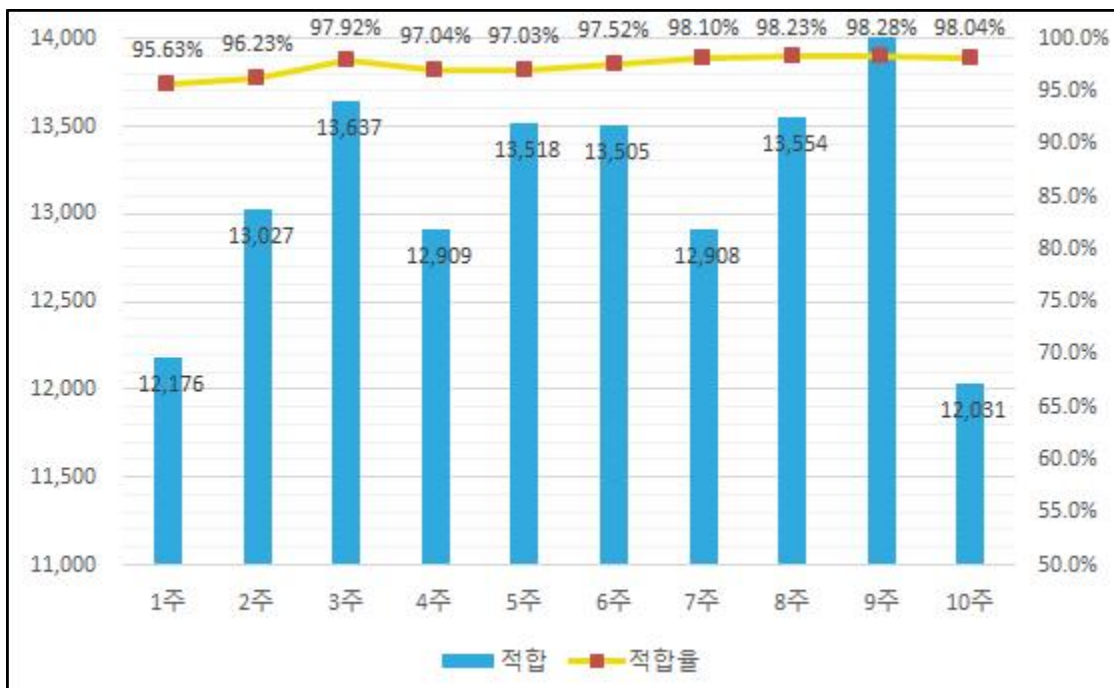


그림 3-12 B그룹 검수 주차별 최종 적합 문장 수 및 적합률

### 3) 최종 데이터 통계 현황

최종 데이터 검수 통계 현황을 정리하였다. 목자-점자 병렬말뭉치 적합 문장의 정확도를 높이기 위해 A그룹과 B그룹에서 중복 적합 판정을 받은 문장을 최종 데이터로 추출하였다. 최종 검수 데이터 통계 현황은 <표 3-8>과 같다.

A그룹 검수 데이터의 검수 결과(1주~10주), 총 검수 수는 전체 135,000문장이고, 적합 수는 128,682문장으로 적합률은 95.32%로 조사되었다. 적합 어절 수는 2,626,617개이다. 오류 수는 5,307문장으로 오류율은 3.93%로 조사되었고, 보류 수는 1,011문장인 0.75%로 조사되었다.

B그룹 검수 데이터의 검수 결과(1주~10주), 총 검수 수는 전체 135,000문장이고, 적합 수는 131,509문장으로 적합률은 97.41%로 조사되었다. 적합 어절 수는 2,685,953개이다. 오류 수는 2,719문장으로 오류율은 2.01%로 조사되었고, 보류 수는 772문장인 0.57%로 조사되었다.

A그룹과 B그룹에서 중복 적합 판정을 받은 문장을 추출한 결과, 총 검수 수는 전체 135,000문장이며, 적합 수는 126,693문장으로 적합률은 93.85%로 조사되었다. 적합 어절 수는 2,584,694개이다.

C그룹 검수 데이터의 검수 결과(1주~10주), 총 검수 수는 전체 50,000문장이고, 적합 수는 46,990문장으로 적합률은 93.98%로 조사되었다. 적합 어절 수는 1,446,905개이다. 오류 수는 0문장으로 오류율은 0%로 조사되었고, 보류 수는 3,010문장인 6.02%로 조사되었다.

실제 말뭉치 JSON 파일을 구축하여 제출한 최종 적합 문장 수는 A그룹과 B그룹 중복 적합 판정 수 126,693문장(2,584,694 어절)과 C그룹 최종 적합 문장 수 46,990문장(1,446,905 어절)을 합산한 173,683문장(4,031,599 어절)이다.

표 3-8 최종 데이터 통계 현황

주차	기 간	총 검수 수	주차 검수율 (%)	검수 결과								총 어절 수
				적합		오류			보류		계	
				적합	어절 수	적합률 (%)	오류	오류율 (%)	보류	보류율 (%)		
A그룹 & B그룹	10월 30일 ~1월 14일	135,000	100	126,693	2,584,694	93.85	-	-	-	-	135,000	2,758,752
C그룹	10월 30일 ~1월 7일	50,000	100	46,990	1,446,905	93.98	0	0	3,010	6.02	50,000	1,571,174
합 계		185,000	100	173,683	4,031,599	93.88	-	-	-	-	185,000	4,329,926

2023년 목자-점자 병렬말뭉치 구축  
(1. 목자-점자 병렬말뭉치 구축)

# 제 4 장

## 목자-점자 병렬말뭉치 적합/오류/보류 사례

1. 오류/보류 주요 유형
2. 적합 사례(예시)
3. 데이터 주요 오류/보류 사례
4. 데이터 주요 오류/보류 처리





# 1 오류/보류 주요 유형

목자-점자 병렬말뭉치를 구축하는 A그룹과 B그룹 검수 데이터 중 오류/보류 주요 유형 현황을 조사하였다. A그룹, B그룹 오류/보류 주요 유형은 <표 4-1>과 같다.

A그룹 검수 데이터의 경우(1주~10주: 10월 30일~1월 7일 기준), 오류 수는 5,307문장, 보류 수는 1,011문장으로 총 오류/보류 수는 6,318문장이다. 오류/보류 주요 유형에서 알파벳 약자/약어 오류가 2,263건으로 가장 많으며 전체 오류/보류의 35.82%를 차지하였다. 그 다음으로 기타 오류가 1,248건으로 19.75%, 문장 부호 오류가 1,130건으로 17.89%를 차지하였다.

B그룹 검수 데이터의 검수 결과(1주~10주: 11월 6일~1월 14일 기준), 오류 수는 2,719문장, 보류 수는 772문장으로 총 오류/보류 수는 3,491문장이다. 오류/보류 주요 유형에서 알파벳 약자/약어 오류가 1,085건으로 가장 많으며 전체 오류/보류의 31.08%를 차지하였다. 그 다음으로 기타 오류가 684건으로 19.59%, 문장 부호 오류가 667건으로 19.11%를 차지하였다.

표 4-1 A그룹, B그룹 데이터 오류/보류 주요 유형

A그룹			B그룹		
오류/보류 사유	비율(%)	건수(개)	오류/보류 사유	비율(%)	건수(개)
1급 점자표 오류	3.80	240	1급 점자표 오류	8.94	312
알파벳 약자/약어 오류	35.82	2,263	알파벳 약자/약어 오류	31.08	1,085
문장 부호 오류	17.89	1,130	문장 부호 오류	19.11	667
단위/연산기호 오류	12.25	774	단위/연산기호 오류	7.68	268
영어 대문자 표기 오류	5.84	369	영어 대문자 표기 오류	8.11	283
제2외국어/한자 표기	0.05	3	제2외국어/한자 표기	0.17	6
로마자 시작표/종료표 오류	4.61	291	로마자 시작표/종료표 오류	5.33	186
기타	19.75	1,248	기타	19.59	684
합 계		6,318	합 계		3,491

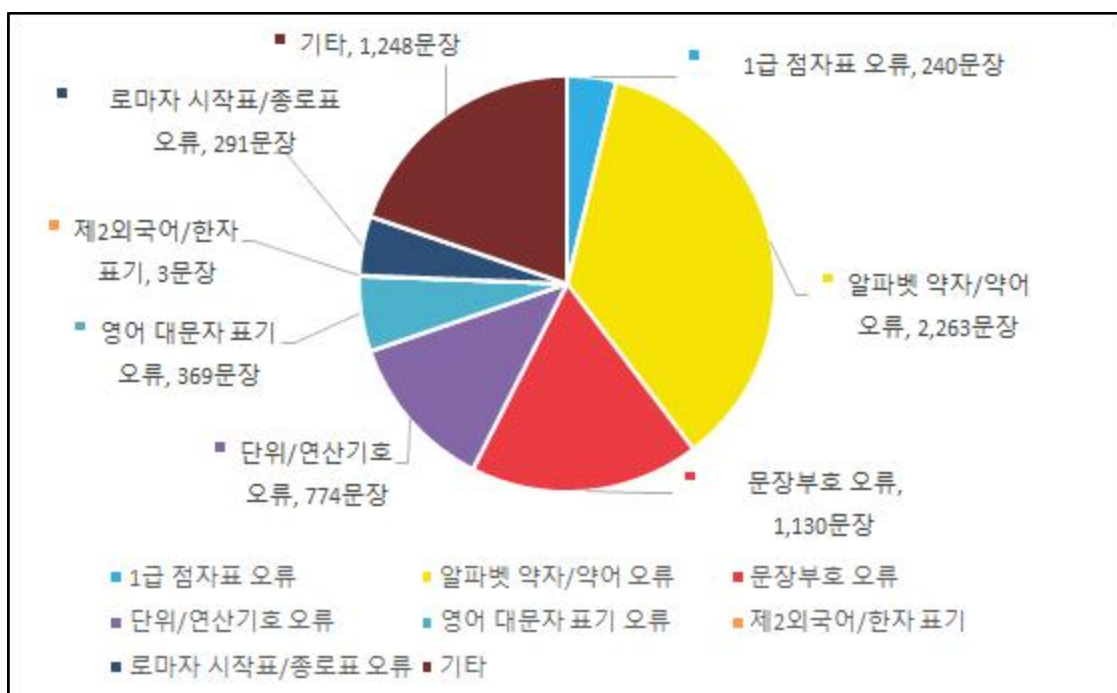


그림 4-1 A그룹 데이터 오류/보류 주요 유형

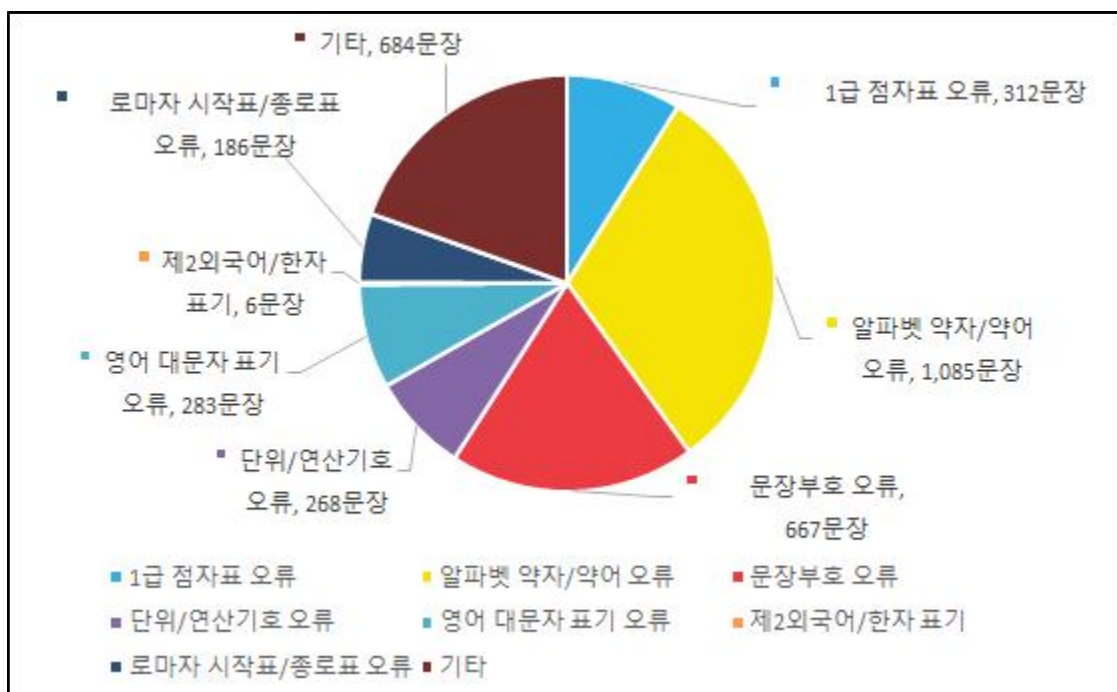


그림 4-2 B그룹 데이터 오류/보류 주요 유형

## 2 적합 사례(예시)

1) 시나리오 1: 한글 + (영어 + ... + )

- ‘한글’ 단어 뒤에 괄호가 있고, 그 속에 ‘영어’ 단어가 포함되어 있는 문장
- 점자에서 괄호는 한글 점자 규정에 따라야 하며, 영어 단어 앞에 로마자 시작표가 있는지 확인해야 함

**표 4-2** 시나리오 1(한글+(영어)) 적합 사례 ①

## 시나리오 1(한글+(영어))

[illegible]

**표 4-3** 시나리오 1(한글+(영어)) 적합 사례 ②

## 시나리오 1(한글+(영어))

Category	국립국어원 목자-점자 병렬말뭉치/2. 온라인 대화/나. 누리소통망
Metadata	<ul style="list-style-type: none"> <li>* title: N/A</li> <li>* author(저자): writer335</li> <li>* publisher(출판사): 인스타그램</li> <li>* date(작성연도): 20210512</li> <li>* original_topic(주제): 문화/예술_공연/전시/박람회</li> </ul>
Form	유테코(UECO)는 각종 전시회, 국제회의, 문화공연 등 다양한 콘텐츠와 시민들의 휴식과 여가를 즐길 수 있는 복합비즈니스 공간입니다!
Braille	



**표 4-5** 시나리오 2(영어+(한글)) 적합 사례 ②

## 시나리오 2(영어+(한글))

[illegible]

**표 4-6** 시나리오 3(한글+(숫자)) 적합 사례 ①

시나리오 3(한글+(숫자))	
Category	국립국어원 목자-점자 병렬말뭉치/1. 신문/나. 지역 종합지
Metadata	* title: 부산일보 2021년 기사 * author(저자): 조영미 * publisher(출판사): 부산일보 * date(작성연도): 20210524 * original_topic(주제): 문화,영화 문화,전시_공연 문화,미술_건축
Form	<p>특별전에서 상영하는 단편 5편도 만만하지 않은 주제 의식을 담았다. '생존자 로카야(2019)'는 아프가니스탄 영화로 자살 폭탄 테러에서 살아남은 12살 로카야가 극심한 트라우마를 겪는 와중에 열띤 취재 대상이 되면서 겪는 일을 그렸다.</p>
Braille	

**표 4-7** 시나리오 3(한글+(숫자)) 적합 사례 ②

### 시나리오 3(한글+(숫자))

[illegible]

**표 4-8** 시나리오 4(숫자+(한글)) 적합 사례 ①

Category	국립국어원 목자-점자 병렬말뭉치/1. 신문/라. 인터넷 기반 신문
Metadata	<ul style="list-style-type: none"> <li>* title: 뉴스핌 2021년 기사</li> <li>* author(저자): 김경민 기자</li> <li>* publisher(출판사): 뉴스핌</li> <li>* date(작성연도): 20210906</li> <li>* original_topic(주제): IT_과학&gt;과학, 경제&gt;자동차, 경제&gt;무역</li> </ul>
Form	특히 관계사인 아이디언스가 개발 중인 표적항암제 신약후보물질인 'IDX-1197(베 나다파립)'과 관련해 글로벌 기준에 부합하는 고도화된 품질관리 체계를 확립한다는 계획이다. IDX-1197은 현재 한국과 미국에서 관련 임상이 진행되고 있다.
Braille	

표 4-9 시나리오 4(숫자+(한글)) 적합 사례 ②

시나리오 4(숫자+(한글))	
Category	국립국어원 목자-점자 병렬말뭉치/2. 온라인 대화/가. 게시판
Metadata	<ul style="list-style-type: none"> <li>* title: 캠퍼스 정보 참여 인원 사전 체크 진행합니다!</li> <li>* author(저자): writer740</li> <li>* publisher(출판사): 네이버카페</li> <li>* date(작성연도): 20220526</li> <li>* original_topic(주제): 생활/쇼핑_가족/연애/모임</li> </ul>
Form	7/8(금) ~ 7/9(토): 자울 캠핑/차박 및 캠퍼스 이벤트 행사 진행
Braille	<pre> 1. 캠퍼스 정보 참여 인원 사전 체크 진행합니다! 2. author(저자): writer740 3. publisher(출판사): 네이버카페 4. date(작성연도): 20220526 5. original_topic(주제): 생활/쇼핑_가족/연애/모임 </pre>

5) 시나리오 5: 영어 + (숫자 + ... + )

- 영어 뒤에 괄호가 있고, 그 속에 숫자가 포함되어 있는 문장

표 4-10 시나리오 5(영어+(숫자)) 적합 사례

시나리오 5(영어+(숫자))	
Category	국립국어원 목자-점자 병렬말뭉치/1. 신문/다. 전문지
Metadata	<div>* title: 머니투데이 2021년 기사</div> <div>* author(저자): 이정원 기자 </div> <div>* publisher(출판사): 머니투데이</div> <div>* date(작성연도): 20211006</div> <div>* original_topic(주제): 사회,사건_사고 사회,여성 사회,미디어</div>
Form	6일 인천지법에 따르면 전날 살인 혐의로 1심에서 징역 14년을 선고받은 A(24)씨와 검찰이 법원에 항소장을 제출했다.
Braille	<div>6일 인천지법에 따르면 전날 살인 혐의로 1심에서 징역 14년을 선고받은 A(24)씨와 검찰이 법원에 항소장을 제출했다.</div>

### 3 데이터 주요 오류/보류 사례

#### 1) 1급 점자표 오류

- 한글 문장 중 단독으로 쓰인 로마자의 경우 로마자 시작표 뒤에 적는 첫 번째 로마자에는 1급 점자표를 적지 않으나 사용한 경우
- 수정 방법: 로마자 시작표와 함께 적는 첫 번째 로마자에는 1급 점자표 삭제

표 4-11 알파벳 R 1급 점자표 오류 사례

#### 시나리오 1(한글+(영어))

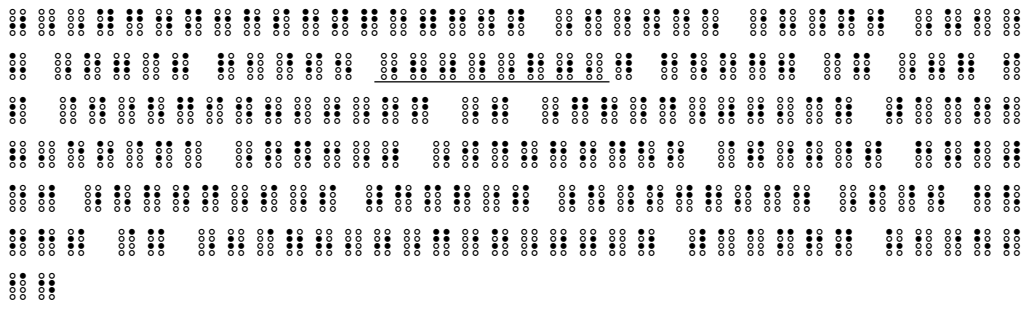


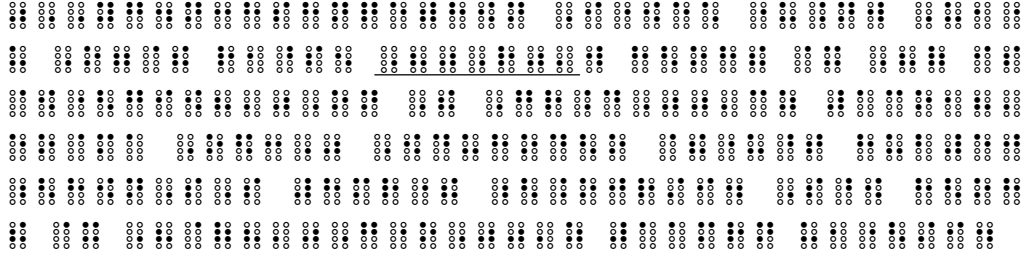

Form	YG엔터테인먼트에 따르면 로제의 솔로 싱글 앨범 'R'의 타이틀 곡 '온 더 그라운드 (On The Ground)'는 13일(미국 시간) 스포티파이 글로벌 톱50 차트에서서 8위를 차지했다. 서브 타이틀 곡 '곤(Gone)'은 11위에 올랐다.
Braille	
오점역 설명	단독으로 쓰인 로마자 R 앞에  과  을 함께 적음
Braille 수정	
수정 설명	로마자 R 앞에 있는  삭제

표 4-12 알파벳 B 1급 점자표 오류 사례

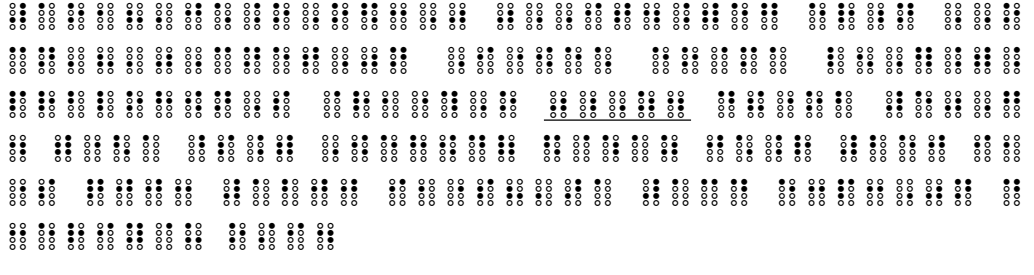


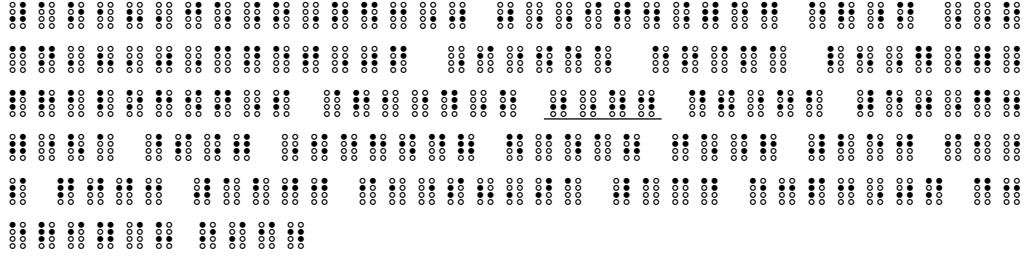

시나리오 5(영어+(숫자))	
Form	1일(현지시간) IT전문 매체 씨넷(Cnet)에 따르면 미국 삼성전자 웹사이트에서 갤럭시 Z 플립 5G 언락 버전 스마트폰 가격은 현재 250달러 인하한 1200달러(약 134만원)에 판매되고 있다.
Braille	
오점역 설명	단독으로 쓰인 로마자 Z 앞에  과  을 함께 적음
Braille 수정	
수정 설명	로마자 Z 앞에 있는  삭제

표 4-13 알파벳 H, E, R 1급 점자표 오류 사례

## 시나리오 1(한글+(영어))

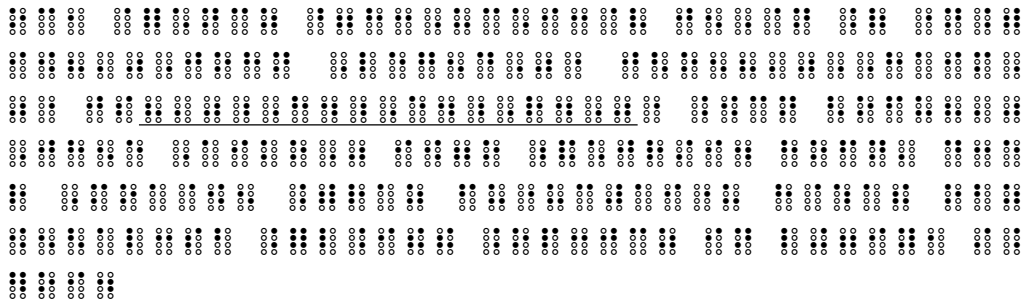


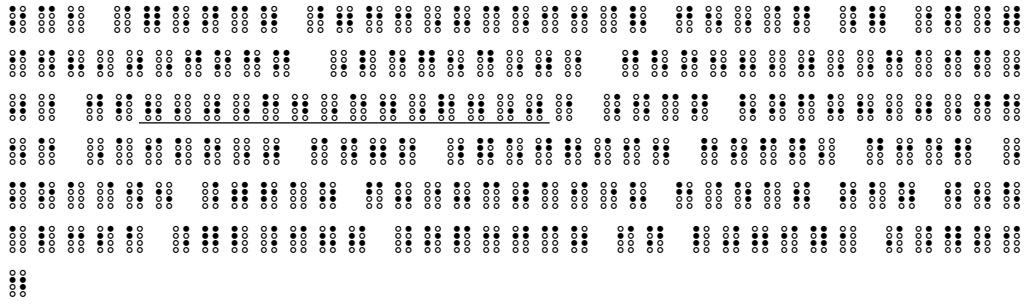

Form	이날 공연에는 방탄소년단과 함께 존 레전드(John Legend), 하임(HAIM), 허 (H.E.R.), 즈네 아이코(Jhene Aiko) 등이 출연했다. 티켓 판매 수익금 전액은 코로 나19로 타격을 입은 음악산업 종사자들을 지원하는 데 사용될 예정이다.
Braille	
오점역 설명	단독으로 쓰인 로마자 H, E, R 앞에  과  을 함께 적음
Braille 수정	
수정 설명	로마자 H, E, R 앞에 있는  삭제



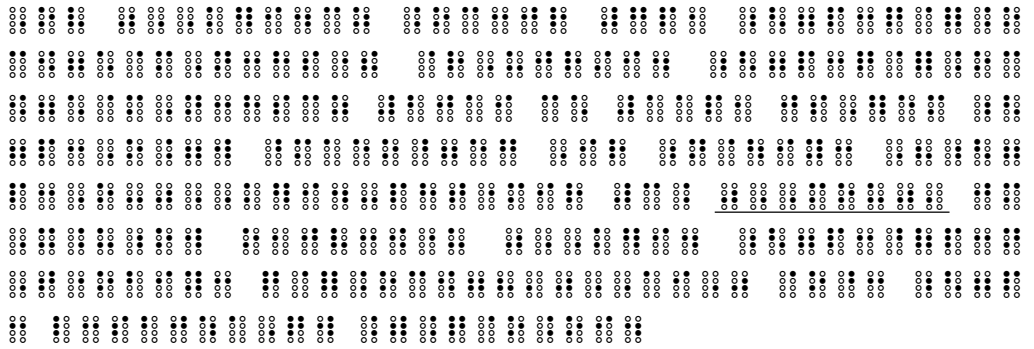
표 4-15 EDM[Electronic Dance Music(전자 무용 음악)] 오류

## 시나리오 1(한글+(영어))

Form	CMF 코드뮤직페스티벌은 현실과 메타버스를 연결하는 실물(physical)과 디지털(digital)이 연결된 Phygital NFTMarketplace COLLEXX(컬렉스)가 만든 HIPHOP & EDM 페스티벌이자 제주 내 최초 대형 뮤직페스티벌!
Braille	<pre> CMF 코드뮤직페스티벌은 현실과 메타버스를 연결하는 실물(physical)과 디지털(digital)이 연결된 Phygital NFTMarketplace COLLEXX(컬렉스)가 만든 HIPHOP &amp; EDM 페스티벌이자 제주 내 최초 대형 뮤직페스티벌! </pre>
오점역 설명	이-디-엠으로 발음하는 EDM에서 ED약자를 사용하여 ⠠EDM으로 오점역
Braille 수정	<pre> CMF 코드뮤직페스티벌은 현실과 메타버스를 연결하는 실물(physical)과 디지털(digital)이 연결된 Phygital NFTMarketplace COLLEXX(컬렉스)가 만든 HIPHOP &amp; EDM 페스티벌이자 제주 내 최초 대형 뮤직페스티벌! </pre>
수정 설명	E와 D를 정자로 풀어서 EDM를 ⠠EDM으로 수정

표 4-16 MOU[Memorandum of Understanding(양해각서)] 오류

## 시나리오 1(한글+(영어))

Form	<p>시와 KERI는 지난해 7월 창원인공지능연구센터를 개소했다. 창원인공지능연구센터는 2019년 11월 허성무 창원시장의 북미방문 때 체결된 ‘창원시-KERI-위털루대 3자 MOU’ 후속조치의 일환으로 KERI 창원본원에 설립된 인공지능(AI) 기반 창원 산업 혁신의 전진기지다.</p>
Braille	
오점역 설명	<p>엠-오-유로 발음하는 MOU에서 OU 약자를 사용하여 로 오점역</p>
Braille 수정	
수정 설명	<p>O와 U를 정자로 풀어서 OU를 로 수정</p>

## 시나리오 1(한글+(영어))

Form	<p>순천향대전안병원에 따르면 대장암센터는 SNS ‘밴드(BAND)’에 ‘순천향대전안병원 대장암팀’이라는 모임방을 개설하고 대장암수술 환자, 가족들과 실시간 온라인 상담을 진행하고 있다.</p>
Braille	
오점역 설명	<p>밴드로 발음하는 BAND에서 AND 약자를 사용하지 않아 로 오점역</p>
Braille 수정	
수정 설명	<p>A와 N와 D를 AND로 묶어 로 수정</p>

## 3) 문장 부호 오류

- 목자 원본에 적힌 띄어쓰기와 상관없이 한글 점자 규정에서 정하는 문장 부호의 띄어쓰기 방법대로 점역해야 함
- 목자 원본에서 작은따옴표나 큰따옴표 등의 문장 부호를 입력할 때 생김새가 비슷하지만 문장 부호가 아닌 다른 문자를 입력하는 경우가 종종 있으며, 이 경우 점역 프로그램으로 변환 시 해당 문장 부호가 아닌 문자가 점역됨
- 개정 한국 점자 규정에서 물결표, 화살괄호 점형이 변경되었으므로 이에 맞춰 점형이 변형되어야 하나 적용되지 않은 경우

표 4-18 다른 문장부호 표기 오류

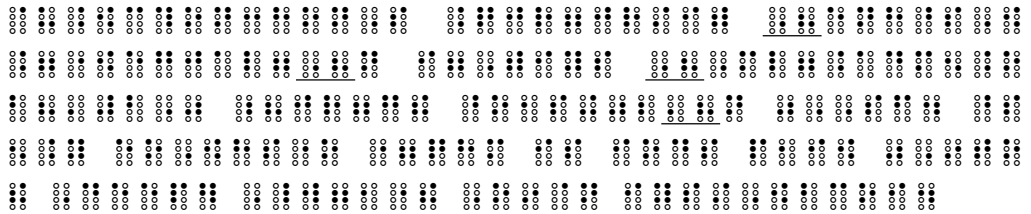
시나리오 3(한글+(숫자))	
Form	고즈넉이엔티에서 진행했던 '메타버스 장르문학상'에 당선되어 '수상작품집(2) 『행운을 빌어요』'에 SF 범죄 미스터리 <인투 더 디퍼 월드 Into the Deeper World>를 실게 되었습니다.
Braille	
오점역 설명	목자에서 문장부호가 모점으로 표기했으나 작은따옴표 자리에 위치해 점자에서 문장 부호가 작은따옴표 점형으로 표기됨
Form 수정	고즈넉이엔티에서 진행했던 '메타버스 장르문학상'에 당선되어 '수상작품집(2) 『행운을 빌어요』'에 SF 범죄 미스터리 <인투 더 디퍼 월드 Into the Deeper World>를 실게 되었습니다.
수정 설명	목자에서 모점 문장부호를 작은따옴표 문장부호로 수정

표 4-19 물결표 오류

## 시나리오 1(한글+(영어))

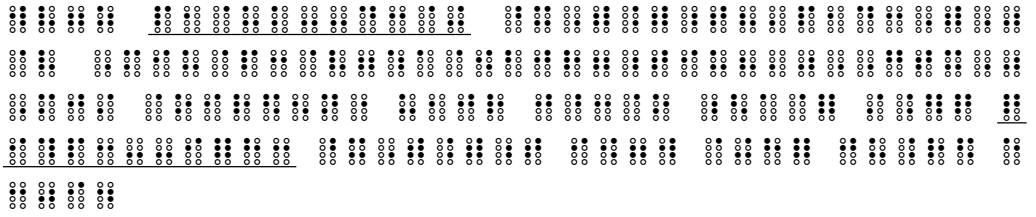


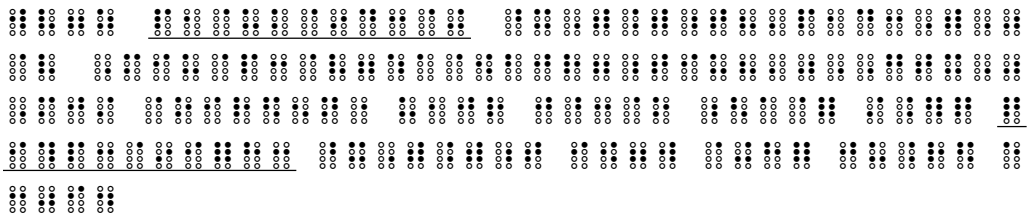


Form	의왕은 월곶~판교 복선전철(월판선)과 수도권광역급행철도(GTX) 수혜 기대감에, 올해 하반기 착공 예정인 인덕원~동탄 복선전철 등의 교통 호재가 많다.
Braille	
오점역 설명	개정 한국 점자 규정에서 물결표 점형이  에서  로 변경됐으나 이전 규정 점형으로 오점역
Braille 수정	
수정 설명	 점형을  로 수정

표 4-20      홀화살괄호 오류

시나리오 2(영어+(한글))

Form	사진을 해시태그와 함께 개인 SNS(소셜네트워크서비스)에 인증하면 추첨을 통해 오리지널 굿즈인 체이스의 경찰관 모자부터 귀여운 퍼피 구조대원들이 담긴 백팩, <퍼피 구조대> 원서 풀 세트 등을 증정할 예정이다.
Braille	<div>⠠사⠠진⠠을⠠</div> <div>⠠해⠠시⠠태⠠그⠠와⠠</div> <div>⠠함⠠께⠠</div> <div>⠠개⠠인⠠</div> <div>⠠SNS(소셜네트워크서비스)에 인증하면 추첨을 통해 오리지널 굿즈인 체이스의 경찰관 모자부터 귀여운 퍼피 구조대원들이 담긴 백팩, &lt;퍼피 구조대&gt; 원서 풀 세트 등을 증정할 예정이다.</div>
오점역 설명	개정 한국 점자 규정에서 홀화살괄호 점형이 <div>⠠(</div> , <div>⠠)</div> 에서 <div>⠠(</div> , <div>⠠)</div> 로 변경됐으나 이전 규정 점형으로 오점역
Braille 수정	<div>⠠사⠠진⠠을⠠</div> <div>⠠해⠠시⠠태⠠그⠠와⠠</div> <div>⠠함⠠께⠠</div> <div>⠠개⠠인⠠</div> <div>⠠SNS(소셜네트워크서비스)에 인증하면 추첨을 통해 오리지널 굿즈인 체이스의 경찰관 모자부터 귀여운 퍼피 구조대원들이 담긴 백팩, &lt;퍼피 구조대&gt; 원서 풀 세트 등을 증정할 예정이다.</div>
수정 설명	<div>⠠(</div> , <div>⠠)</div> 점형을 <div>⠠(</div> , <div>⠠)</div> 로 수정

## 4) 단위/연산기호 오류

- 알파벳으로 표기된 단위부호가 부호가 아닌 알파벳으로 점역되는 경우

표 4-21 로마자가 아닌 단위 부호 점역 오류

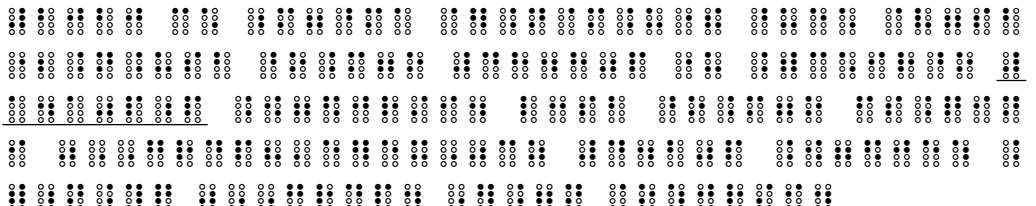
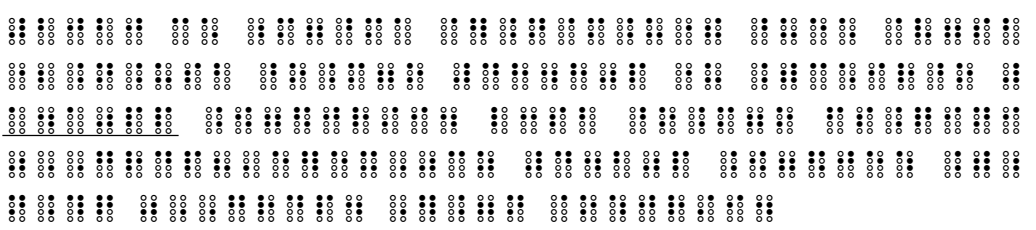
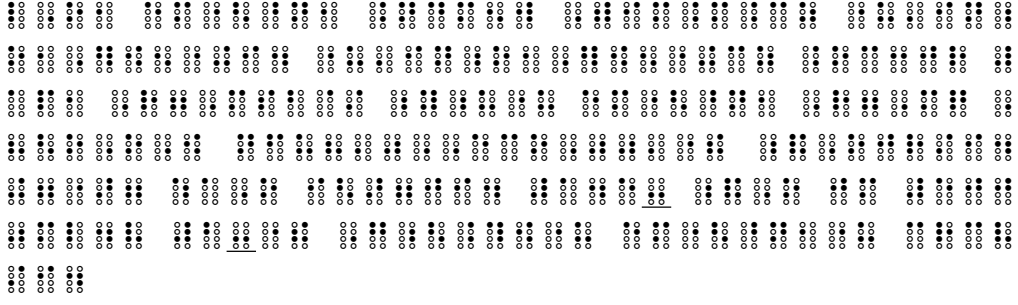

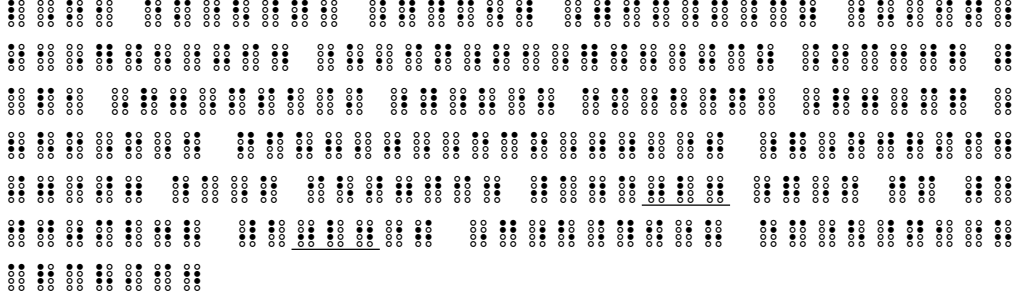


시나리오 2(영어+(한글))	
Form	2020년 충북 경제구조를 보면 광업·제조업 비중이 48.8%로 전년대비 1.1%p 증가했다. 산업 비중이 커진데다 GRDP(명목)는 4.2% 증가하며 전체적인 GRDP 성장에 기여했다.
Braille	
오점역 설명	단위 부호 %p(퍼센트 포인트)의 p를 단위의 한 부분이 아닌 알파벳으로 오판단하여 p 앞에 1급 점자표 1가 붙어 있음
Braille 수정	
수정 설명	1급 점자표 1 삭제

표 4-22 부피 단위(리터(ℓ)) 유니코드 점역 오류

시나리오 1(한글+(영어))

Form	사실 무라벨 제품의 선두주자는 롯데칠성음료다. 롯데칠성음료는 지난해 1월 생수업 계 최초로 무라벨 생수인 ‘아이시스 에코(ECO)’를 출시했다. 용량 역시 다양하다. 1.5ℓ 출시 후 500mℓ와 2ℓ를 순차적으로 무라벨로 내놨다.
Braille	
오점역 설명	목자에서 부피의 단위 리터(영문 소문자 엘) 점형을  으로 오점역
Braille 수정	
수정 설명	 점형을  로 수정

시나리오 3(한글+(숫자))

표 4-24 화학기호 점역 오류/보류





### 시나리오 1(한글+(영어))



- 목자 원본에서 뱀셈표(마이너스 부호)가 붙임표로 점역되는 경우

표 4-26 연산 기호 점형 오류 ②

## 시나리오 1(한글+(영어))

Form	3월 소비자물가지수(CPI)는 전년 동기 대비 0.4% 상승해 올 들어 처음 마이너스에서 플러스 전환에 성공했다. CPI 역시 전달 증가율인 <u>-0.2%</u> 와 시장 전망치인 0.3%를 모두 상회했다.
Braille	
오점역 설명	뱀셈표(  )가 붙임표(  )로 오점역
Braille 수정	
수정 설명	 점형을  로 수정

6) 영어 대문자 표기 오류

- 영어 단어 내에서 대문자를 표기할 때 대문자 단어표와 대문자 종료표를 사용하는 경우가 있고, 대문자 글자표를 사용하는 경우가 있음을 구분해야 함

표 4-27 영어 대문자 표기 오류 ①

시나리오 1(한글+(영어))	
Form	시는 협약을 통해 내년부터 정보통신기술(ICT)에 기반한 도어 투 도어(DOOR TO DOOR) 방식의 ‘착한셔틀 시범사업’을 추진해 중증장애인 근로자의 이동권을 증진하고, 안전한 근로 생활환경을 제공한다는 방침이다.
Braille	<div> <div> </div> <div> </div> </div>
오점역 설명	에서 대문자가 종료 시 뒤에 문장 부호가 있으면 대문자 종료표를 사용할 필요가 없으나 가 적용됨
Braille 수정	<div> <div> </div> <div> </div> </div>
수정 설명	대문자 종료표 삭제

표 4-28 영어 대문자 표기 오류 ②

### 시나리오 1(한글+(영어))

[illegible]

- 한글 점자 규정과 통일 영어 점자 규정이 서로 충돌하거나 또는 통일 영어 점자 규정에서 규정되어 있지 않아 표기 지침을 선택하기 어려운 경우

표 4-29 영어 대문자 표기 오류/보류

시나리오 1(한글+(영어))	
Form	이후 E·S·G 요인들이 국가신용등급에 미치는 영향을 고려해 최종적으로 ESG 신용영향점수(CIS)를 5개 등급으로 평가했다.
Braille	
보류 이유	통일 영어 점자 규정에서 가운뎃점에 대한 표기 규정이 되어있지 않아 S, G에도 각각 로마자표와 대문자표를 사용해야 하는지, 아니면 E앞에만 로마자표와 대문자표를 써도 되는지 여부를 결정하기 어려움

## 7) 제2외국어/한자 표기

- 프랑스어, 러시아어 등 외국어 점역 시 통일 영어 점자 내 변음 부호(제2외국어)를 적용하는데 검수자의 한계가 있어 제2외국어는 심화 범위로 판단 후 보류
- 한자의 경우 점역으로 표기할 때 소리를 나타내는 음만 표기할지, 동일한 음으로 오역을 방지하기 위해 음과 훈을 같이 사용할지 논의가 필요하여 보류

표 4-30 영어 외 외국어 사례 ①

시나리오 1(한글+(영어))	
Form	<p>피부미용 의료기기 전문기업 (주)이루다는 자사의 마이크로니들 이용 고주파 전기수술 기 시크릿알에프(Secret RF)가 브라질 위생 감시국 (ANVISA-Agência Nacional de Vigilância Sanitária) 인증을 획득하고 브라질 현지 기업인 메드시스템즈(MedSystems)와 공급 계약을 마쳤다고 3일 밝혔다.</p>
Braille	
보류 이유	ANVISA-Agência Nacional de Vigilância Sanitária에 표기된 변음 부호는 제2외국어이므로 어떻게 점역해야 할지 어려움이 있음

비밀번호 1(숫자) <한글>

시나리오 4(숫자+한글)



- ^^, \_-, ^^; 와 같이 문장 부호가 감정 이모티콘으로 사용된 경우 점자 규정에 기재된 부호 점형으로 기재해야 하는지 점형이 정해지지 않았기에 오류가 아닌 보류로 판단

표 4-34 이모티콘 문장 기호 보류 ①

## 시나리오 3(한글+(숫자))

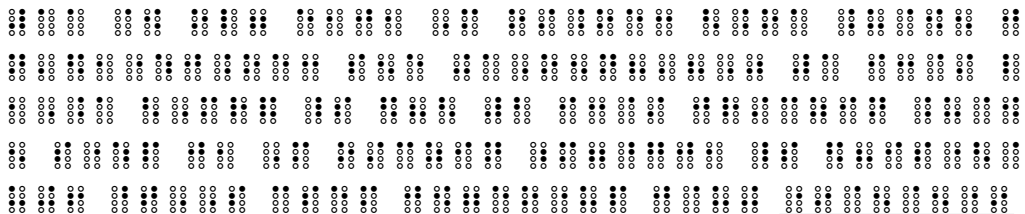
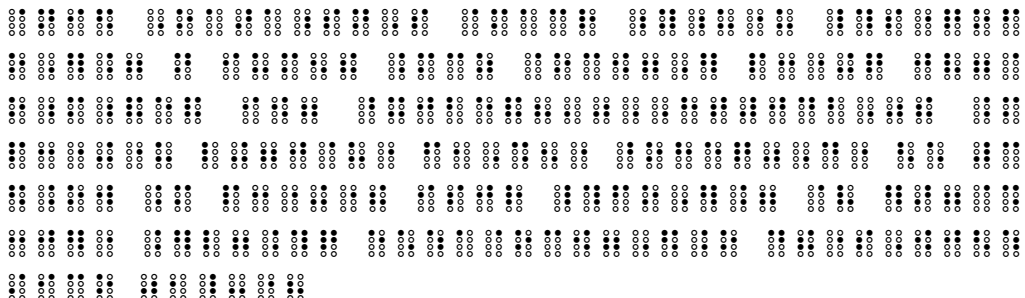
Form	그 이상 말할 수 있지만 아직 조금 헛갈려하지만 엄마 욕심에(?) 1단계 살짝 아쉬울 것 같아 2단계 했는데 조금 어려워 할 수 있는데 충분할 것 같더라고요! 벌써 너무 흥미로워 하고요 *^^*
Braille	
보류 이유	*^^* 이모티콘 표기를 어떻게 점역해야 할지 어려움이 있음

표 4-35 이모티콘 문장 기호 보류 ②

## 시나리오 1(한글+(영어))

Form	게다가 식약처에서 국내 최초로 #어린이키성장 에 도움을 주는 기능성 원료인 항기 추출물 등 복합물(HT042)을 주 원료로 사용하고, 칼슘, 비타민D, 아연 3가지의 주 원료를 더해 #키성장 과 영양균형, 정상적인 면역기능까지 플러스하였다니 >_<
Braille	
보류 이유	>_< 이모티콘 표기를 어떻게 점역해야 할지 어려움이 있음

## 4 데이터 주요 오류/보류 처리

오류/보류 주요 유형을 조사 후 각 유형에 맞춰 오류/보류 데이터 처리를 진행하였다. 1급 점자표 오류, 알파벳 약자/약어 오류, 문장 부호 오류, 영어 대문자 표기 오류와 같이 개정 한국 점자 규정에 점형 규정이 정해져 있는 유형은 판정 후 올바른 점형으로 교정하여 적합 데이터로 변환하였다. 단위/연산기호 오류, 기타 부호 오류, 제2외국어/한자 표기와 같이 개정 한국 점자 규정에 점형 규정이 정해지지 않은 유형은 교정 처리를 할 수 없어 그대로 오류/보류 데이터로 처리하였다.

검수 시 검색엔진 오류가 자주 발생하는 점형은 웹 기반 관리시스템 점역 엔진(㈜ACNS 자사 서버 기반 점역 엔진) 업데이트를 통해 점형을 수정하여 적합률을 올리고, 2회(5주차, 6주차) 업데이트를 진행하였다. 그 외에도 웹 기반 관리시스템 공지사항을 통해 업데이트 패치 소식, 자주 묻는 질문 등을 올려 정확도를 올리하고자 하였다. 또한 검수자 64명이 자유롭게 질문하고 대답할 수 있는 자유게시판을 신설하여 검수자 간 소통과 정보 공유를 통해 검수 정확도와 적합률을 제고하고자 하였다.

소통게시판
24번글 제목: 알파벳 단어약자 관련 문의 (댓글 수: 3) 작성자: 이한찬(kbu23b3823) 등록일시: 2024-01-03 15:58:36
23번글 제목: SDGs NSAIDs BBMA's 등 두문자어s 형태의 경우 문의입니다. (댓글 수: 1) 작성자: 이연진(kbu23b0523) 등록일시: 2023-12-31 21:25:37
22번글 제목: 단위관련 문의입니다. (댓글 수: 1) 작성자: 전상희(kbu23a1534) 등록일시: 2023-12-06 23:13:12
21번글 제목: 화학 원소 표기에 관한 정확한 지침 부탁드립니다. (댓글 수: 2) 작성자: 전상희(kbu23a1534) 등록일시: 2023-12-06 22:33:56
20번글 제목: 1급점자표 문의 드립니다. (댓글 수: 4) 작성자: 이주연(kbu23b5982) 등록일시: 2023-12-05 16:59:51
19번글 제목: >표시는 어떻게 되는건지 궁금합니다. (댓글 수: 2) 작성자: 박현조(kbu23b0933) 등록일시: 2023-12-04 17:36:39
18번글 제목: 문장검수 중 질문이 있습니다. (댓글 수: 1) 작성자: 이연진(kbu23b0523) 등록일시: 2023-12-03 13:17:01
17번글 제목: 약어 관련 문의입니다. (댓글 수: 2) 작성자: 이한미(kbu23b5497) 등록일시: 2023-11-28 23:50:06
16번글 제목: 원소 기호 표기 (댓글 수: 1) 작성자: 강미혜(kbu23b7877) 등록일시: 2023-11-28 12:39:54
15번글 제목: 한글문장내에서 영어문장 괄호사이에서 대문자 구결표를 사용할 수 있는지 문의드립니다. (댓글 수: 1) 작성자: 이주연(kbu23b5982) 등록일시: 2023-11-16 17:31:05
14번글 제목: 개정 65항 화폐단위 표기 (댓글 수: 1) 작성자: 검수교정 001(group-a-001) 등록일시: 2023-11-16 14:17:45

그림 4-3 검수자 소통 자유게시판

2023년 목자-점자 병렬말뭉치 구축  
(1. 목자-점자 병렬말뭉치 구축)

# 제 5 장

## 결 론

### 1. 결 론





## 1 결 론

이 사업의 목적은 병렬말뭉치 구축 및 안정된 점역 소프트웨어를 개발하여 일상생활에서 시각장애인  
의 점자 사용 및 학습 편의를 증진하고, 향후 점자 친화적인 생활환경 조성을 위한 개선 방안을  
제시하는 것이다. 이러한 사업 목적을 달성하기 위해 설정한 사업 목표는 다음과 같다.

첫째, 목자-점자 병렬말뭉치 데이터 구축 방안을 제시하였다. 한국어·영어 혼용 자료(목자-점자  
일대다 대응 자료 포함) 확보, 목자-점자 대응 문장 총 십만 쌍(백만 어절 이상)을 구축하였다.

둘째, 목자-점자 병렬말뭉치 검수 및 통계 방안을 제시하였다. 말뭉치 구축 전체 과정을 체계적으  
로 진행하기 위한 웹 기반 관리시스템 활용 방안을 제시하였다. 오류 방지를 위해 데이터 집계 특성  
에 따른 통계 분석을 시행하였다.

셋째, 목자-점자 병렬말뭉치 오류 및 보류 사례를 제시하였다. 검수 과정에서 발견된 오류 사례와  
보류 사례는 점자 규정과 연동되어야 하는 부분이 많아서 말뭉치 사업 전반에 점자 규범 정비 및  
연구 위원회의 활동과 연계가 필요하다. 이를 위해 점자 규정 및 관련 지침 준수 여부를 분석하는  
데 기초 자료로 활용할 뿐만 아니라 사업 효과를 평가할 수 있는 실증적인 자료를 확보하였다.

이 사업의 결과를 기초로 결론을 제시하면 다음과 같다.

첫째, 목자-점자 병렬말뭉치 구축 체계 및 검수 방안이다. 목자-점자 병렬말뭉치 구축을 위한 웹  
기반 관리시스템과 이 데이터를 활용하는 방법을 제시하였다. 신문과 온라인 대화를 데이터셋으로  
구축하여 향후 점역 엔진 학습 데이터로써 활용될 수 있게 하는 것을 목적으로 구축하였다. 고품질  
의 목자-점자 병렬말뭉치 데이터셋을 구축하기 위한 데이터 수집, 정제, 가공, 검사 단계까지의  
고려해야 할 점과 이에 대한 예시를 설명하였다. 구축 과정은 말뭉치 웹 기반 관리시스템을 통해  
효율적으로 구축 및 관리하였다.

둘째, 검수자 64명에게 각자 그룹에 맞춰 문장을 할당하였고(A/B그룹 각 135,000문장, C그룹  
50,000문장), 그룹 특성에 맞춰 검수를 진행하였다. A그룹 적합 수는 128,682문장으로 적합 검수  
율은 95.32%, 적합 어절 수는 2,626,617개이다. B그룹 적합 수는 131,509문장으로 적합 검수율  
은 97.41%, 적합 어절 수는 2,685,953개이다. A그룹과 B그룹 중복 적합 수는 126,693문장으로  
적합률은 93.85%, 적합 어절 수는 2,584,694개이다. C그룹 적합 수는 46,990문장으로 적합 검수  
율은 93.98%, 적합 어절 수는 1,446,905개이다. 실제 말뭉치 JSON 파일을 구축하여 제출한 최종  
적합 문장 수는 A그룹과 B그룹 중복 적합 판정 수 126,693문장(2,584,694 어절)과 C그룹 최종  
적합 문장 수 46,990문장(1,446,905 어절)을 합산한 총 173,683문장, 어절 4,031,599개이다.

넷째, 이 사업에서는 말뚝치 JSON 파일을 구축하여 제출할 예정이다. JSON 파일의 작성 예시를 제시하면 <표 5-1>과 같다.

**표 5-1** JSON 파일 작성 예시

```
{
  "language_info": {
    "source_language": "kl",
    "target_language": "br"
  },
  "parallel": [
    {
      "id": "KBUCORPUS0134974",
      "original_id": "NPRW2200000004.46144.2",
      "source": "DL이앤씨(옛 대림산업)는 인공지능(AI) 설계기술을 개발해 특허출원을 완료했다고 13일 밝혔다. A가 아파트 환경을 분석해 30분 만에 약 1000 건의 지하주차장 설계안을 자동으로 만들어 내는 기술이다.",
      "target": ""
    }
  ],
  "revision": {
    "revision1": ""
  }
}
```

2023년 목자-점자 병렬말뭉치 구축  
(II. 점역 소프트웨어 고도화)

# 제 6 장

## 서론

1. 사업 배경
2. 사업 목표
3. 사업 방법





## 1 사업 배경

국립국어원은 지난 2022년 묵자-점자 병렬말뭉치 구축 사업의 일환으로 점역 소프트웨어 고도화 사업을 수행하기 위해 한국시각장애인연합회를 보조사업자로 선정하였다. 이 사업은 2022년 8월부터 2023년 4월까지 진행되었다.

그리고 그 결과 다음과 같은 성과를 달성하였다.

- 묵자-점자 문서 편집 및 점역·역점역 기능 강화: 점역 코드 도입을 통한 서식 적용, 표 서식 입력 및 점역 기능, 수식 입력 및 점역 기능 등을 새로이 개발하였다. 또, 묵자 편집기 내에 점자 서식을 유지할 수 있는 점자판 기능, 한글(HWP)에서 작성된 표, 수식, 글자 속성, 문단 서식 점역 기능 등을 추가하였다.
- 점역 엔진 검증 프로그램 개발: 점역 결과의 정확도를 높이기 위해 점자 규정 및 점자 규정 해설서의 예문과 점형을 데이터베이스로 구축하고, 이를 점역 엔진 검증 프로그램과 연동하여 점역 결과의 오류를 신속하게 검출할 수 있도록 하였다. 또, 묵자-점자 병렬말뭉치 데이터베이스를 점역 엔진 검증 프로그램과 연동하여 점역의 오류 사항을 효과적으로 검출할 수 있는 토대를 마련하였다.
- 점역 엔진 API 개발: 한글, 영어, 구두점 등이 포함된 텍스트에 대한 점역 및 역점역 기능을 별도로 개발할 필요 없이 API를 이용하여 쉽게 구현할 수 있는 기반을 마련하였다.

국립국어원은 2022년 사업에 이어 점자 정보화 기반을 확충하고 점자 사용 환경을 개선하여 시각장애인의 정보 접근성을 향상하기 위한 2023년 묵자-점자 병렬말뭉치 구축 사업의 일환인 점역·역점역 소프트웨어 개발 및 고도화 사업을 추진하였고, 이 사업은 다음과 같은 내용으로 구성되어 있다.

- 점역·역점역 및 점자 문서 편집 소프트웨어 고도화, 점역 핵심 코드 공개
- 최신 개정 한국 점자 규정을 반영한 점역·역점역 엔진 개발
- 음악 점자 규정을 반영한 악보 점역 및 편집 소프트웨어 구현
- 점역 소프트웨어와 연동된 촉각 그래픽 소프트웨어 구현
- 점역·역점역 소프트웨어 활용법 및 점자 출판물 제작 방법에 대한 온라인 콘텐츠 개발

2022년 점역·역점역 소프트웨어 개발 및 고도화 사업의 결과로 개발된 점사랑 4.0의 검증 결과 및 본 사업을 위해 수행한 점사랑 4.0 기능 개선을 위한 분석 내용을 바탕으로 본 사업의 필요성 및 방향을 정리하면 다음과 같다.

- 최신 개정 한국 점자 규정 반영: 개정 한국 점자 규정(문화체육관광부 고시 제2024-0005 호, 이하 “개정 한국 점자 규정”이라 함)이 2024년 1월 29일에 고시됨에 따라 사용자들의 요구에 대응하기 위해 점사랑 소프트웨어에서 이 규정을 빠르게 지원할 필요가 있다.
- 프로그램 안정성 확보: 2022년 점역 소프트웨어 고도화 사업 최종 보고서에 점역 소프트웨어의 유지 보수 및 지속적인 기능 개선의 필요성을 언급하였다. 그리고 사용자 평가 결과 및 배포 후 수집된 사용자 의견에 따라 프로그램 안정성과 점역 정확도에 대한 개선 요구에 대응해야 한다.
- 표 서식 자동화 및 수식 점역 기능 보완: 점사랑 4.0에서 표 서식을 점역하는 기능을 구현 하였으나, 표 제목, 행 제목, 열 제목 등의 표 속성과, 행과 열을 풀어 점역하는 표 등의 서식 기능은 아직 구현되지 않거나 기능이 부족하여 사용자들의 요구를 만족시키지 못하고 있다. 또한 수식, 옛 글자 점역 기능에서 나타나는 오류들의 보완이 필요하다.
- 촉각 그래픽 소프트웨어 개발: 현재 사용되는 촉각 그래픽 생성 프로그램은 해외에서 개발된 고가의 프로그램이거나 특정 점자 프린터만 지원하는 등 사용자의 접근에 많은 제약이 있다. 점사랑과 연동되고 편집과 점자 인쇄가 가능한 촉각 그래픽 소프트웨어의 개발이 필요하다.
- 악보 점역 소프트웨어 개발: 2023년 목자-점자 병렬말뭉치 구축 점역 소프트웨어 고도화 사업 내용은 음악 점자 규정을 반영한 악보 점역 및 편집 소프트웨어를 구현하는 작업을 포함하고 있다. 기존에 국내에서 개발된 악보 점역 소프트웨어로는 하상장애인복지관에서 2012년에 발표한 '칸타빌레'가 있다. 이 소프트웨어는 범용 악보 편집 프로그램인 Muscore에서 입력한 자료를 불러와 점역하는 방식으로 구현되어 있으나, 많은 점역 오류를 발생시키고 있어 교열 및 편집에 원활하게 활용할 수 없다는 문제점이 있다.
- 온라인 콘텐츠 개발: 점사랑의 현장 활용도를 높이려면 점역·역점역 소프트웨어 활용법 및 점자 출판물 제작 방법에 대한 교육 콘텐츠의 개발 및 보급이 필요하다.

## 2 사업 목표

2023년 점역·역점역 소프트웨어 고도화 사업의 목표는 다음과 같이 설정하였다.

- 안정성 및 정확도 향상: 연구, 개발진 협의회를 통해 소프트웨어 기능 분석을 수행하고, 기존 점사랑의 문제점을 도출하여 소프트웨어 설계에 반영한다. 또, 사용자 평가 및 피드백을 바탕으로 점사랑의 안정성과 점역 정확도를 향상시키고, 지속적인 유지보수 및 기능 개선을 수행한다.
- 개정 한국 점자 규정 적용: 개정 한국 점자 규정을 신속하게 점사랑 소프트웨어에 반영하여 점자 사용자들의 요구에 부응하는 소프트웨어 지원을 제공한다.
- 표 서식 및 수식 점역 기능 개선: 표 서식과 수식 점역 기능의 자동화 및 보완을 통해 사용자의 요구를 충족시킬 수 있는 기능을 개발한다.
- 축각 그래픽 소프트웨어 개발: 점사랑 소프트웨어와 연동되는 축각 그래픽 소프트웨어를 개발하여 축각 자료의 출력 기능을 강화함으로써 축각 자료 제작 및 보급을 지원한다.
- 악보 점역 소프트웨어 개발: 음악 점자 규정을 반영한 악보 편집 및 점역 소프트웨어를 구현하여 점자 악보 자료의 제작 및 편집 작업을 용이하게 한다.
- 온라인 교육 콘텐츠 개발 및 보급: 점사랑 소프트웨어 활용법 및 점자 출판물 제작 방법에 관한 온라인 교육 콘텐츠를 개발하고, 이를 널리 보급하여 현장에서의 점사랑 활용도를 높인다.

이러한 사업 목표들을 수행함으로써 점자 정보화 기반을 강화하고 시각장애인의 정보 접근성을 개선하는 데 기여할 수 있을 것으로 기대된다.

### 3 사업 방법

본 사업은 다음과 같은 방법으로 수행하였다.

#### 1) 연구진 구성

점역 소프트웨어 설계 및 개발 관련 전문 인력 3명, 1급 점역·교정사 2명 등 총 5명으로 연구진을 구성하였다. 연구진의 역할은 다음과 같다.

- 점역 소프트웨어 요구 사항 분석 및 기능 정의
- 이미 개발된 점역 소프트웨어(점사랑)의 오류 및 개선 사항 분석
- 국내외 악보 점역 소프트웨어(Cantabile, GOODFEEL 등)의 기능 및 UI 분석
- 국외 촉각 그래픽 소프트웨어(QuickTac, TactileView 등)의 기능 및 UI 분석
- 악보 점역 및 촉각 그래픽 소프트웨어 기능 및 UI 설계
- 소프트웨어 개발 업체와 협력하여 테스트 및 안정화 작업 수행
- 점자 출판물 제작 방법 온라인 콘텐츠 자문 및 검수

#### 2) 문헌 조사

다음과 같은 문헌 조사를 통해 다양한 정보를 수집, 분석하여 소프트웨어 설계의 기초 자료로 활용하였다.

- 규정 연구: 새로이 발표된 한국 점자 규정 개정안을 2020 한국 점자 규정과 비교, 분석하여 점역 소프트웨어에 적용할 수 있도록 하였다.
- 지침 연구: 시각장애학생용 대체 교과서 및 교수·학습자료 제작 지침을 분석하여 원본 자료의 정렬 형태를 유지하는 표 및 원본 자료의 정렬 형태를 수정하는 표를 서식 자동화 기능을 통해 구현할 수 있도록 하였다.
- 기타 문헌 조사: 국제음성기호 점자 일람표를 조사하여 음성 기호 점역 기능을 보완할 수 있는 방안을 마련하고, 유니코드 일람표를 조사하여 자동 교열 기능을 보완할 수 있는 방안을 마련하였다.

### 3) 소프트웨어 기능 분석

2023년 점역 고도화 사업 내용에 포함되어 있는 촉각 그래픽 및 악보 점역 소프트웨어 설계 및 개발을 위해 국외 촉각 그래픽 소프트웨어 QuickTac 및 악보 점역 소프트웨어 GOODFEEL의 기능을 조사, 분석하여 사용자 인터페이스와 기능을 설계하고 이를 점사랑에 통합할 수 있도록 하였다.

### 4) 연구·개발진 협의체 운영

위에 제시한 조사, 비교, 분석을 원활하게 수행하기 위해 2주마다 연구·개발진 협의체를 운영하고, 이를 통해 점사랑 보안을 위한 문제점과 해결 방안을 도출하며, 소프트웨어 설계, 개발과 관련된 협의를 진행하였다. 2024년 2월 현재 총 9차의 개발협의체 정기 온라인 회의를 개최하였으며, 세부적인 의견 교환을 위한 협의체 메일링리스트를 만들어 주요 개발 이슈들을 상시 논의하였다.

### 5) 객체지향 및 CBD 방식의 소프트웨어 개발

소프트웨어 개발 방법은 객체지향 및 CBD 방식을 채택하였다. 점역 소프트웨어 개발 시 구현했던 소스 및 바이너리를 효과적으로 재사용할 수 있도록 산출물에 따라 객체 및 컴포넌트로 구분하여 개발하고 이를 점역 엔진 API 등에 반영하였다.



2023년 목자-점자 병렬말뭉치 구축  
(II. 점역 소프트웨어 고도화)

# 제 7 장

## 점사랑 기능 개선을 위한 분석 및 설계

1. 점사랑 4.0 점역 엔진 및 편집기 기능 분석
2. 점사랑 5.0 개발 방향 및 기능 개선 사항





본 장에서는 점사랑 4.0의 기능 보완을 위해 먼저 각 영역별 점역 엔진의 문제점 및 편집기의 서식 자동화와 관련된 사항을 분석하였다. 또한 이를 통해 도출된 점사랑 5.0의 개발 방향 및 기능 개선 사항을 소개하였다.

## 1 점사랑 4.0 점역 엔진 및 편집기 기능 분석

## 1) 점사랑 4.0 점역 엔진 기능 분석

점역 엔진 기능 분석은 2022년 고도화 사업 수행 과정에서 개발한 검증 프로그램을 활용하여 진행하였다. 검증 데이터는 점사랑 4.0의 점역 엔진을 검증하는 데 활용한, 점자 규정 및 점자 규정 해설서의 예문을 중심으로 구축한 데이터를 사용하였다. 이러한 오류들은 수행사에 상시 전달, 수정되어 점사랑 5.0에 반영하였다.

### (1) 한글 점역 오류 분석

한글 점역 오류는 띄어쓰기 오류, 동국정운식 표기 오류, 로마자 표기나 로마자 종료표 오류 등의 유형이 있었으며 구체적인 사례는 <표 7-1>과 같다.

**표 7-1 한글 점역 오류 사례 분석**

[illegible]

검증 데이터	검증용 정답	점사량 점역 결과	오류 내용
고외 씨 무려 창송취죽	[정답 이미지]	[결과 이미지]	동국정운식 표기법이 기본 값으로 설정되어 있어 발생하는 이음과 옛이음 받침 오류
썰나모 썩여디고 갬대에	[정답 이미지]	[결과 이미지]	
방 들뜰쫓	[정답 이미지]	[결과 이미지]	
루이 XVI세	[정답 이미지]	[결과 이미지]	로마자 대문자표를 반복하여 표기하는 오류
CD-ROM은 오디오용으로 개발된 지름 12cm의 광디스크에 각종 데이터가 미리 기록되어 있는 장치이다.	[정답 이미지]	[결과 이미지]	단위 기호 종료표 생략 오류
그리스어 알파벳의 첫 자와 끝 자는 α와 ω다	[정답 이미지]	[결과 이미지]	

## (2) 영어 점역 오류 분석

영어 점역 오류 분석 결과 합자표, 1급 점자 기호, 특수 문자 오류 등의 유형이 있었으며, 구체적인 사례는 〈표 7-2〉와 같다.

표 7-2 영어 점역 오류 사례 분석

[illegible]

[illegible]

### (3) 수식 점역 오류 분석

수식 점역 오류 분석은 실제 수학 점자 사용 환경을 충분히 반영하기 위해 중·고등학교 수학 교과서 등 다양한 자료를 참조하여 예제를 작성하였다. 분석 결과 띄어쓰기, 묶음 괄호 오류 유형 등이 나타났으며 구체적인 사례 분석 예시는 아래 <표 7-3>과 <표 7-4>에 제시하였다.

**표 7-3** 수식 점역 오류 사례 분석 ①

예제 수식

모든 실수  $x$ 에 대하여  $f(-x)=-f(x)$ 인 다항함수  $f(x)$ 가  $f(-1)=2$ ,  
 $\lim_{x \rightarrow -1} \frac{f(1)-f(-x)}{x^2-1}=3$ 을 만족시킬 때,  $\lim_{x \rightarrow -1} \frac{\{f(x)\}^2-4}{x+1}$ 의 값은?

검증용 정답

오류 내용

점사랑 점역 결과

오류 내용

오류 내용

$x$ ,  $f(x)$ , 수식 사이 한글 띄어쓰기 오류, 분수식에서 묶음 괄호 오류

표 7-4 수식 점역 오류 사례 분석 ②

[illegible]

## 2) 점사랑 4.0 편집기 기능 분석

### (1) 표 서식 기능 분석

「시각장애학생용 대체 교과서 및 교수·학습자료 제작 지침」에서는 표의 형식을 크게 「원본 자료의 정렬 형태를 유지하는 표」와 「원본 자료의 정렬 형태를 수정하는 표」로 구분하고 있다. 이 중 원본 자료의 형태를 유지하는 표는 점사랑 4.0에 서식 자동화가 적용되어 있으나, 원본 자료의 정렬 형태를 수정하는 표는 점사랑 4.0에는 아직 적용되어 있지 않다. 정렬 형태를 수정하는 표에는 여러 종류가 있는데, 점사랑 5.0에서는 기술적인 한계로 그 중 「열 항목을 풀어 점역하는 표」만을 표 서식 자동화에 추가하기로 하였다.

열 항목을 풀어 점역하는 표의 서식 자동화 적용 방안은 다음과 같다.

- ① 표를 점역할 경우 기본적으로 표의 각 행 그룹을 하나의 문단 단위로 구분하며 각 셀 항목 간 구분은 점역 형식 옵션 중 ‘열 항목 풀어 점역’, ‘열 항목 세로선으로 구분’에 따라 구분한다.
- ② 점역 형식 옵션은 ‘원본 정렬 형태를 유지하는 표’, ‘열 항목을 풀어 점역하는 표(빈칸 구분)’, ‘열 항목을 풀어 점역하는 표(세로선 구분)’로 구성한다.
- ③ ‘원본 정렬 형태를 유지하는 표’ 옵션의 경우 행의 너비가 ‘줄 당 칸 수(32칸)’ 이하일 경우엔 표의 정렬 형태대로 ‘행(좌-우)’, ‘열(위-아래)’의 배열을 맞춰 정렬하고, 행의 첫 셀은 들여쓰기 없이 1칸부터 배치한다. 각 열의 너비는 해당 열의 가장 긴 셀을 기준으로 결정한다. 즉 1열에서 가장 긴 셀의 너비가 6칸, 2열의 가장 긴 셀의 너비가 10칸, 3열의 가장 긴 셀의 너비가 8칸일 경우 1열은 3~8, 2열은 11~20, 3열은 23~30 구간의 너비를 갖는 배열로 배치된다.
- ④ 열 항목 사이가 ③과 같은 경우로 적용되어 열 항목 사이에 다섯 칸 이상의 빈칸이 있을 때 열 항목 사이를 연결해 주는 유도점인 ∷(5점)을, 앞에 빈칸 하나와 뒤에 빈칸 2개를 두고, 그 사이 잔여 칸에 유도점을 채운다. 다만, 열 제목 사이에는 유도점을 적지 않는다.
- ⑤ ‘원본 정렬 형태를 유지하는 표’의 옵션을 선택했더라도 표의 행 최대 너비가 ‘줄 당 칸 수(32칸)’보다 클 경우 ‘열 항목’ 구분은 ⑥항의 옵션으로 구분한다. 이 경우 행의 첫 번째 셀의 내용을 표기할 때는 3칸(들여쓰기 2칸) 위치부터 배치한다.
- ⑥ ‘열 항목을 풀어 점역하는 표(빈칸 구분)’ 옵션을 선택한 경우 표의 행 너비와 상관없이 무조건 ‘열 항목’ 구분은 ‘빈칸 2개’로 구분하며, 행의 첫 셀은 2칸 들여 쓴 3칸부터 배치한다.
- ⑦ ‘열 항목을 풀어 점역하는 표(세로선 구분)’ 옵션을 선택한 경우 행의 너비와 상관없이 ‘열 항목’ 사이의 구분은 무조건 ‘∷∷∷∷∷’로 구분한다(‘∷∷∷∷’ 앞뒤에 빈칸이 하나씩 포함됨). 이 경우 행의 첫 셀은 2칸 들여 쓴 3칸부터 배치하며, 세로선은 열 제목에도 표기한다.

- ⑧ 빈 셀의 경우 온표 두 개(⠠ ⠠)로 표기한다.
- ⑨ 열 제목 줄(1행)과 항목(2행) 사이의 구분선은 1칸부터 줄의 마지막 칸(기본값은 32칸)까지 ⠠(5점)으로 채워 적는다.
- ⑩ 표의 첫 번째 열 항목이 '행 제목' 속성(예로, <th>태그)을 가진 경우 해당 열의 '행 제목'의 문자열 끝에 '쌍점(⠠ ⠠)' 문자를 삽입한다.
- ⑪ 표 제목(caption)은 표의 위 테두리의 이전 줄에서 4칸 들여쓰기하여 5칸 위치부터 표기한다. 외부에서 읽어 들인 문서 또는 점사랑에서 작성한 표 제목이 표 안에 있더라도 점역할 경우 항상 제목을 표 시작 이전 줄에 표기한다. (HWP 문서의 경우 제목이 '위쪽', '우측면'에 배치된 경우 점사랑에서는 모두 위쪽 제목으로 적용한다.)
- ⑫ 표 제목(caption)이 표 하단부에 적용된 경우 항상 표 하단부 밖 다음 줄의 4칸 들여쓰기하여 5칸 위치부터 표기한다. 외부에서 읽어 들인 문서 또는 점사랑에서 작성한 표 설명(주석, 출처 등)이 표 안에 있더라도 점역할 경우 제목을 표 하단부 바깥 다음 줄에 표기한다.
- ⑬ 표의 첫 번째 셀에 대각선이 표시된 경우 대각선이 우상향 또는 우하향 대각선에 상관없이 대각선 구분을 점역 시 '⠠ ⠠'으로 점역한다. 풀어쓰는 순서는 행의 제목을 먼저 표기 후 '⠠ ⠠'을 삽입하고 이어 열 제목을 표기한다.

## (2) 문서 최적화 기능 분석

문서 최적화는 문단 재구성, 자동 교열, 기존 코드 변환 등을 하위 기능으로 구성하고 있다.

- ① '문단 재구성'은 '설정' 메뉴의 '문단 재구성 설정' 옵션에 준하여 문단을 재구성한다. '문단 재구성'은 원본 문서가 점자 문서 또는 텍스트 문서인 경우 모두 사용할 수 있다. 기존에 점역된 점자 문서의 문단을 새로이 구성하려고 할 때 사용하며 이미 서식화된 점자 문서(BRF, BRL 등)의 머리말, 꼬리말, 쪽 번호, 빈 줄, 빈칸 등을 삭제할 수 있다. 원본 텍스트 문서가 스캔 후 OCR인식 등에 의해 만들어진 경우, 전체 문서의 문단 구성이, 편집이 많이 요구되어 점역하기에 적합하지 않을 경우 문단을 빠르게 재구성할 수 있다.
- ② 자동 교열은 기존의 '유효성 검사'의 명칭을 변경한 것으로, 목자 문서 또는 점자 문서에서 오류를 자주 발생시키는 항목을 검사하고, 오류가 의심되는 항목을 목록으로 출력하여, 출력된 목록에서 해당 항목으로 이동할 수 있는 기능을 제공한다.
  - 유니코드 검사: 구두점 등과 혼용하여 쓰이면서 점역 오류를 일으키는 기호들을 점역에 적합한 기호로 대체할 수 있도록 오류 기호들을 찾아 준다.

- 문단 들여쓰기 검사: 문단의 시작이 홀수의 빈칸으로 되어 있는 들여쓰기 위치를 찾아 준다.
  - 페이지 검사: 목차 페이지 순서에 오류가 있는 위치를 찾아 준다.
  - 문장 부호 띄어쓰기 검사: 문장 부호의 앞이나 뒤에 부적절한 빈칸이 있는 위치를 찾아준다.
  - 쌍으로 쓰이는 기호 검사: 쌍으로 쓰이는 문장 부호, 표 및 글상자 테두리 오류가 있는 위치를 찾아 준다.
- ③ 기존 코드 변환: 점사랑 3.0을 비롯한 국내 점역 소프트웨어에서 사용되는 코드를 점사랑 코드로 변환하는 기능이다. 외부 편집기에서 점역 설정을 위해 삽입하는 상용 점역코드인 @@ff, @@FF, @@p, ^#, #...#^ 등의 4종의 코드를 점사랑 내부 점역코드 중 호환되는 코드로 변환한다.

## 2 점사랑 5.0 개발 방향 및 개선 사항

### 1) 점사랑 5.0 개발 방향

점사랑 4.0의 기능 분석을 통해 다양한 시사점이 도출되었다. 이러한 시사점을 근거로 점사랑 5.0의 개발 방향을 다음과 같이 제시하였고, 추가 기능과 개선 사항들을 정리해 점사랑 5.0에 구현하였다.

- 개정 한국 점자 규정을 반영하고 점역 결과의 정확도를 높인다.
- 표, 문단 속성, 글자 속성 등 서식 자동화 기능을 강화한다.
- 수식 입력기, 점자판 등 사용자 경험 측면에서 직관성을 높인다.
- 점역 코드의 사용 규칙 등 사용자 인터페이스 측면에서 일관성을 유지한다.
- 유니코드 점역 지원 등 시대 변화에 능동적으로 대응한다.
- 한자, 기타 외국어 점역 지원 등 현장의 사용자 요구를 반영한다.
- 축각 그래픽 생성, 악보 점역 지원을 통해 다목적, 다기능 통합 도구로서의 점역 기능을 최신화한다.
- 오류 수정 및 보안 강화 등 프로그램의 안정성을 높이고 내실화한다.

### 2) 점사랑 5.0 기능 개선 사항

점사랑 5.0의 편집 및 점역 기능은 점사랑 4.0의 기능과 사용자 인터페이스에 기반을 두고 기능 개선을 진행하였다. 점역 기능에서 가장 중점을 두고 추진한 계획은 개정 한국 점자 규정의 반영이었다. 2024년 4월 출시를 앞둔 점사랑 5.0은 개정된 한국 점자 규정을 최초로 적용한 점역 엔진을 개발하는 성과를 이루었다. 이 밖에도 수식, 축각 자료, 점자 악보 등을 통합하는 문서 제작을 위해 많은 부분을 개선하였다.

다음 <표 7-5>에 점사랑 5.0의 메뉴 구조를 간략히 정리하고, 새로운 기능 및 개선 사항을 소개하였다.

표 7-5 점사랑 5.0 기능 개선 사항

주 메뉴	메뉴 항목	기능 개선 항목
파일	새 문서, 열기, 닫기, 최근문서, 저장, 다른 이름으로 저장, 선택된 영역 저장, 점역/역점역, 점자 인쇄 미리보기, 점자 인쇄, 목자 인쇄, 방석 인쇄, 종료	열기, 점역/역점역
편집	되돌리기, 다시실행, 잘라내기, 복사, 붙여넣기, 삭제, 전체 선택, 찾기, 계속 찾기, 이전으로 계속 찾기, 바꾸기, 이동, 코드 찾기, 강제 줄 바꿈, 점형 입력, 점자 편집, 문서 최적화	문서 최적화
보기	조판부호, 도구 모음, 도구 모음 크기, 유니코드 점형, 줄번역, 상태표시줄, 줄 번호, 코드보기	
삽입	라인 입력, 테두리 입력, 점자 차례, 표, 점자판, 점자 그림, 수식 입력, 악보 입력, 코드 목록, 코드 삽입, 특수문자, 한자	표, 점자판, 점자 그림, 수식 입력, 악보 입력, 코드 목록, 코드 삽입
문서	강조 표시, 첨자, 정렬, 단락 속성, 쪽 나누기, 페이지 번호, 점역자 주, 꼬리말/머리말, 워드랩, 점역 모드, 스타일	강조 표시, 첨자, 정렬, 단락 속성
설정	기본 환경 설정, 점역 옵션 설정, 점자 프린터 설정, 문단 재구성 설정, 점자 정보 단말기 설정, 6점키 설정, 스타일 설정, 템플릿 관리, 편집창 설정, 단축어 설정, 사용자 정의 점역	스타일 설정, 템플릿 관리
창	다음창, 계단식, 바둑판식, 분할 창	

### (1) 열기

외부 편집기 파일(HWP 등)을 읽어 들일 때 아래 사항들을 점역에 적합한 서식으로 변환하는 기능을 추가하고 안정화하였다.

- ① 문자 속성: 밑줄, 진하게, 위첨자, 아래첨자, 기울임꼴, 점자 폰트
- ② 문단 속성: 들여쓰기, 왼쪽/가운데/오른쪽 정렬
- ③ 표 속성: '원본의 형태를 유지하는 표'/'열 항목을 풀어 점역하는 표' 형태로 점역이 가능하도록 표의 자료를 읽어 들임
- ④ 수식: 한글 프로그램 수식 입력기로 제작된 수식

### (2) 점역/역점역

- ① 최신 점자 규정에 기반한 점역 및 역점역 엔진을 구현하였다.
  - 개정 한국 점자 규정(문화체육관광부 고시 제2024-0005호)
  - 통일 영어 점자 규정

- 시각장애 학생용 대체 교과서 제작 지침 및 적용 방안 연구(국립특수교육원, 2021)
- 점자 도서 제작 지침(국립장애인도서관, 2022)

② 외국어 점역 기능을 확대하였다.

- 영어 점역 기능에 추가하여 알파벳 문자를 사용하는 독일어, 스페인어, 프랑스어 점역 기능 구현

③ 유니코드 점형을 지원하였다.

- "유니코드 점자\_음성 처리 실용화 연구(국립국어원, 2014)"에 제시되어 있는 유니코드 점형 지원

점역 옵션 설정	
한글 점자(U):	2020 구 규정
한글 워드랩(O):	2020 구 규정
줄 당 칸 수(C):	32
쪽 당 줄 수(L):	26
인쇄 방식 설정(P):	양면
쪽 매김 방식(T):	홀수쪽 출력
점자 쪽 번호 출력 위치(B):	아래에서 오른쪽
목자 쪽 번호 출력 위치(R):	아래에서 왼쪽
머리말/꼬리말 정렬(G):	머리말/꼬리말 중앙
품프드에 의한 쪽 구분(F):	사용함
쪽 당 첫 줄 빈 줄 처리(N):	유지
원본 문서 들여쓰기(I):	0
CR코드 대체 문자(A):	삭제
연속 빈 칸 삭제(W):	사용 안 함
연속 빈 줄 삭제(S):	사용 안 함
문단 끝 빈 칸 삭제(M):	사용 안 함
문단 점역 기준(J):	한글 위주
한글 점역 방식(K):	약자
영문 점역 방식(E):	통일영어점자 2급
한자 점역 방식(H):	음

확인(O) 취소 기본값(D)

그림 7-1 점역 옵션 설정 메뉴의 신규 규정 선택 화면

### (3) 문서 최적화

① 기존의 '상용코드 점역 코드로 전환' 메뉴를 '문서 최적화' 메뉴로 변경하였다.

② '문서 최적화' 메뉴의 하위 항목으로 다음 기능을 추가하였다.

- 문단 재구성: [파일]의 '문단 재구성' 메뉴를 '문서 최적화' 하위 항목으로 이동한 것으로, [설정]의 '문단 재구성 설정'에 따라 편집창에 열려 있는 목자/점자 문서의 문단 서식을 재구성하는 기능

- 자동교열: '유효성검사'의 명칭을 변경한 것으로, 목자 문서 또는 점자 문서에서 오류가 자주 발생하는 항목을 검사하고, 오류가 의심되는 항목을 리스트로 출력하여, 출력된 리스트에서 해당 항목으로 이동할 수 있는 기능 추가
- 기존 코드 변환: '상용 코드 점역 코드로 전환'의 명칭을 변경한 것으로, 기존 점역 소프트웨어(점사랑 3.0, 브레일서울 등)의 점역 코드가 포함되어 있는 텍스트를, 개발되어 배포될 '점사랑' 점역 코드로 변환하는 기능

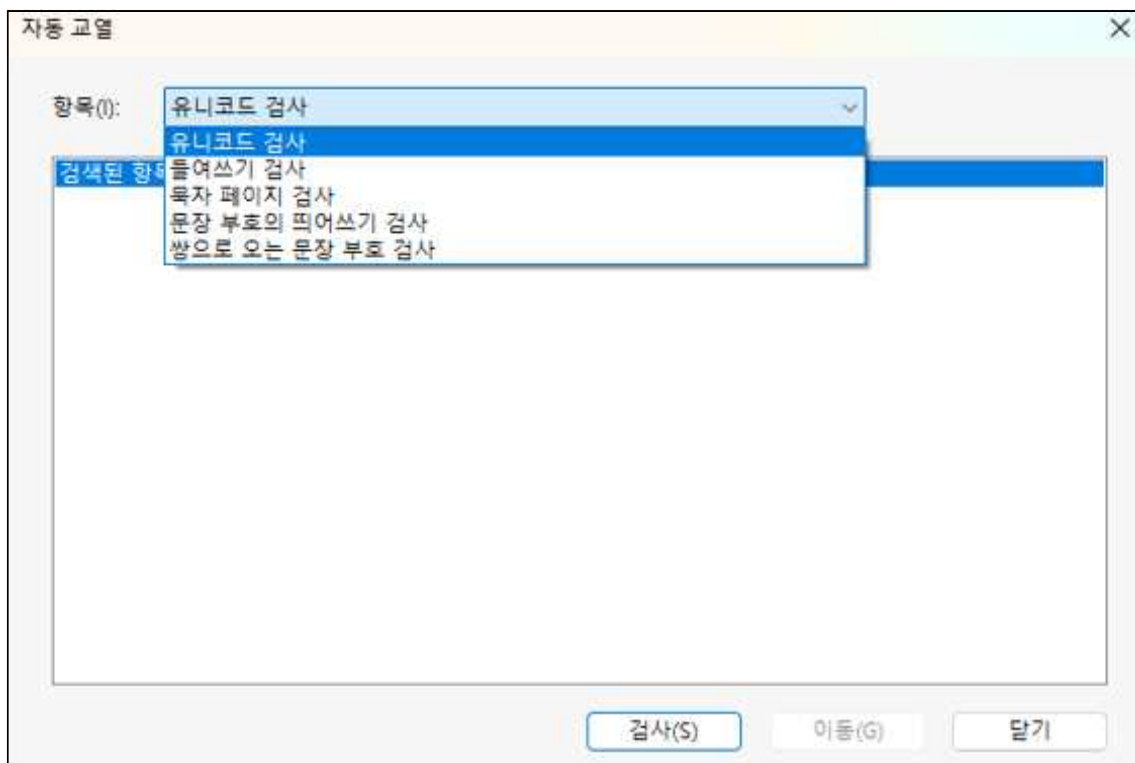


그림 7-2 문서 최적화 메뉴의 자동 교열 기능 화면

#### (4) 표

표 생성 기능을 강화하고, 편집 기능을 안정화하였다.('표 서식 기능 분석' 내용 참조)

#### (5) 점자판

- ① 시각적 인식도를 높이기 위해 점자 인쇄물과 유사하도록 줄 사이와 칸 사이의 정렬을 변경하였다.
- ② 줄 단위 및 칸 단위 편집 및 이동 기능을 추가하였다.
  - '점자판' 편집 기능에 입력 시 커서 이동 방향 전환 기능(왼쪽에서 오른쪽으로/위에서 아래로)

③ 점자 페이지 내 점자판 배치 지정 기능을 추가하였다.

- 점자판으로 작성한 내용이 두 페이지에 걸쳐 인쇄될 때 점자판으로 입력한 내용이 한 페이지에 인쇄되도록 조정

## (6) 점자 그림

축각 그래픽 편집과 점자 인쇄가 가능한 축각 그래픽 소프트웨어를 구현하여 점사랑 5.0에 통합하였다. ("제8장 축각 그래픽 소프트웨어" 참조)

## (7) 수식

편집기에 수식 입력을 클릭하면 대화 상자를 거치지 않고 바로 수식 편집창으로 전환하도록 변경하였다.

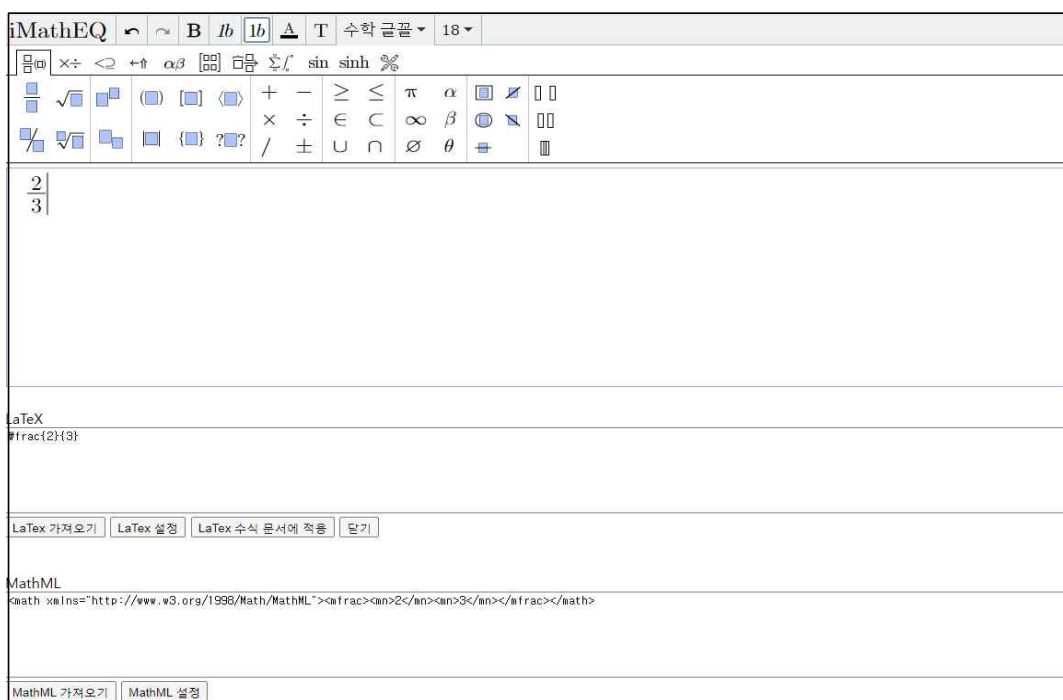


그림 7-3 점사랑 수식 편집기 화면

## (8) 악보 점역

음악 점자 규정을 반영한 악보 편집 및 점역 소프트웨어를 구현하여 점사랑 5.0에 통합하였다.

("제9장 악보 점역 소프트웨어" 참조)

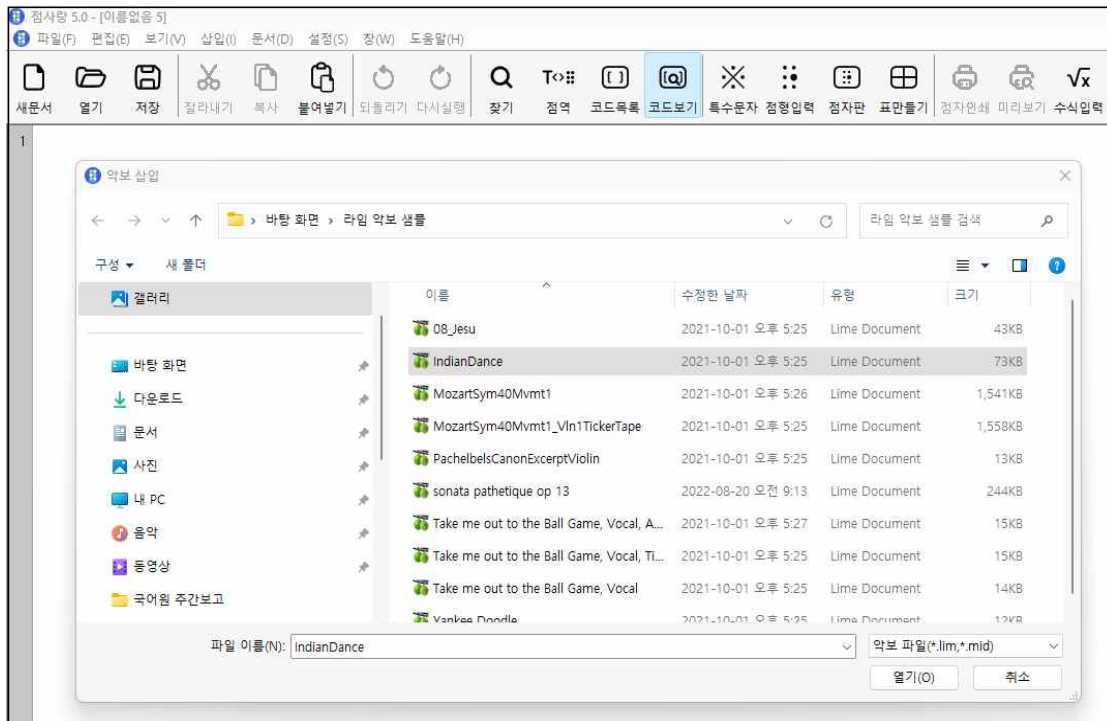


그림 7-4 점사랑에서 악보 삽입 시 악보 선택 화면

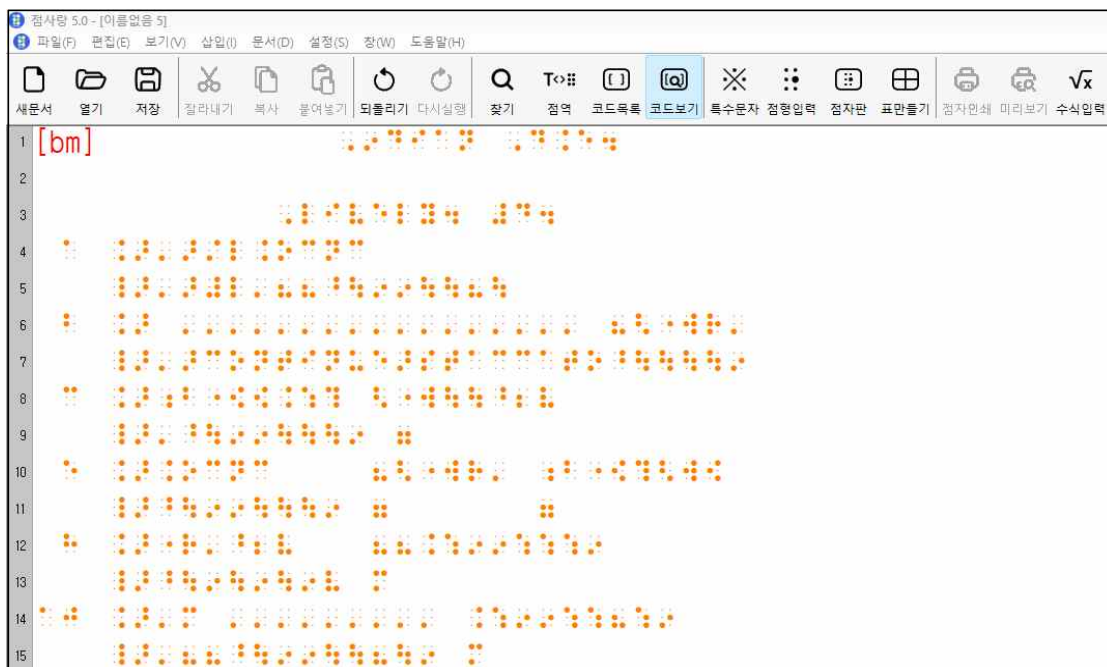


그림 7-5 선택한 악보의 자동 점역 결과 화면

### (9) 점역 코드

다음의 속성으로만 표현되고 있던 기능을 점역 코드로도 표현할 수 있도록 개선하였다.

- ① 강조 표시(진하게, 기울임, 밑줄)
- ② 첨자(위첨자, 아래첨자, 없음)
- ③ 문단 정렬(왼쪽/가운데/오른쪽 정렬)
- ④ 들여쓰기
- ⑤ 문단 여백
- ⑥ 강제 줄바꿈

### (10) 정렬

점자 문서의 경우 '왼쪽/가운데/오른쪽 정렬'의 화면 표현을, 디스플레이 너비를 기준으로 정렬하는 대신에 점자 인쇄물과 유사하게 정렬하도록 기능을 개선하였다.

### (11) 스타일 설정

스타일의 단일 코드를 사용할 수 있도록 기능을 개선하였다. 점사랑 4.0에서는 '시작 코드'와 '종료 코드'로 쌍을 이루는 코드만 등록이 가능하다는 제한점이 있었다.

### (12) 템플릿 관리

단락 속성, 워드랩, 점역 옵션, 사용자 생성 스타일, 사용자 정의 점역을 템플릿으로 저장할 수 있도록 기능을 보완하였다.

2023년 목자-점자 병렬말뭉치 구축  
(II. 점역 소프트웨어 고도화)

# 제 8 장

## 촉각 그래픽 소프트웨어

1. 해외 촉각 그래픽 소프트웨어 기능 분석
2. 점사랑 5.0 촉각 그래픽 기능 설계





## 1 해외 촉각 그래픽 소프트웨어 기능 분석

QuickTac은 Duxbury Systems에서 제공하는 무료 촉각 그래픽 소프트웨어로, Duxbury Braille Translator(DBT)로 불러올 수 있는 촉각 그래픽을 만들 수 있다. 점사량 5.0에 포함될 촉각 프로그램의 설계에 참고하기 위해 QuickTac의 주요 특징 및 메뉴 구성을 중심으로 관련 기능을 분석하였다.

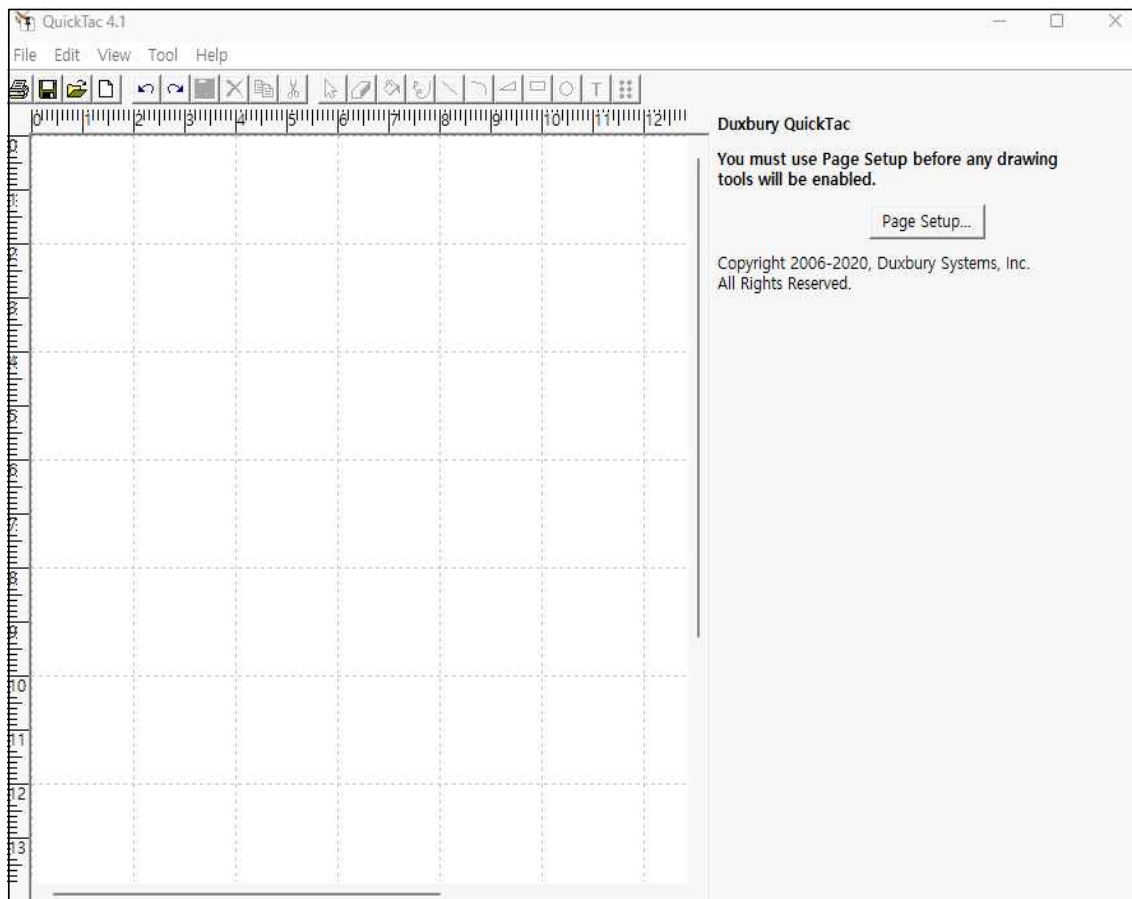


그림 8-1 QuickTac 프로그램 화면

### 1) QuickTac의 주요 특징

QuickTac의 주요 특징은 다음과 같다.

#### (1) NimPro와 상호 운용

QuickTac은 점자 자료 제작을 위해 교과서 및 교육 자료(NIMAS 문서)를 만드는 Duxbury의

프로그램인 NimPro와 상호 운용된다. QuickTac은 NimPro의 작동에 통합되어 있고, Trace 명령을 사용하여 축각 그래픽을 만들면 NimPro가 QuickTac을 호출하게 된다. 이렇게 하면 QuickTac 작업의 "단일 작업" 모드가 설정되고, 이 모드에서 QuickTac은 파일 새로 만들기 및 파일 열기와 같은 기능을 비활성화하여 선택한 이미지 작업만을 수행할 수 있도록 기능을 제한한다. 물론 NimPro 외부에서 QuickTac을 사용하면 이러한 제한이 없다.

## (2) QuickTac의 파일 형식

QuickTac은 두 가지 파일 형식에 대해 생성하고 편집할 수 있다. 하나는 축각 이미지의 일반 편집용인 "dxg" 형식이고 다른 하나는 DBT로 내보내기 위한 "sig" 형식이다.

## (3) 표준 그래픽 파일을 배경으로 활용

QuickTac에서 표준 그래픽 파일(\*.jpg, \*.gif, \*.bmp, \*.png)은 불러올 수 없다. 대신 시각적 이미지를 캔버스의 배경 이미지로 사용하여 시각적 이미지를 축각적인 그래픽으로 변환하는 데 도움을 받을 수 있다.

## (4) 점자 출력기의 호환성

각 세션을 시작할 때에는 점자 출력기의 종류를 선택해야 하는데, 파일 메뉴의 페이지 설정을 통해 이 작업을 수행할 수 있다. 점자 출력기의 각 모델은 텍스트와 그래픽을 약간씩 다르게 처리하기 때문에 원하는 점자 출력기 종류를 지정하지 않고 축각적인 그래픽을 만들고자 하면 예기치 않은 문제가 발생할 수 있다. 따라서 QuickTac의 도구 메뉴는 점자 출력기의 종류를 설정할 때까지 활성화되지 않는다. 점자 출력기 종류를 지정하지 않거나 포괄적인 점자 출력기(Generic Embosser)를 지정하여 QuickTac을 사용할 수 있지만 그래픽에 축각 텍스트나 점자 레이블이 포함된 경우 대부분 간격 문제가 발생하게 된다.

## (5) 직관적인 사용자 가이드

텍스트 프로그램과 달리 QuickTac의 복사 및 붙여넣기 기능은 점을 잘라내거나 복사한 부분을 배치하는 데 도움이 되는 시각적인 클립보드를 제공한다. 또한 QuickTac 창의 오른쪽 부분에 클립보드 자료를 그래픽의 올바른 지점으로 이동하는 방법을 안내하는 지침이 나타난다.

## (6) 간편한 그리기 도구

QuickTac에는 메뉴나 아이콘에는 없지만 중요한 그리기 기능이 하나 있는데, 바로 단일 점 편집 기능이다. 어떤 도구를 선택했는지에 관계없이 그리기 창의 아무 곳이나 Shift 키를 누른 상태에서 클릭하면 단일 점 모드를 켜거나 끌 수 있다. 또한 특정 그림 파일을 패턴으로 사용할 수 있는 기능을 제공한다. 패턴으로 사용하려는 그림의 아무 영역을 선택하고, 편집 메뉴에서 채우기로 저장(Save as Fill)을 선택하면 나중에 특정 이미지를 채우는 데 필요한 파일을 저장할 수 있다. QuickTac은 촉각 그래픽 도구로 여러 장점을 가지고 있지만, 시각장애인이 사용하도록 설계되지는 않았다.

## 2) QuickTac의 메뉴 구성

QuickTac의 메뉴는 크게 파일, 편집, 보기, 도구, 도움말로 구성되어 있다.

### (1) 파일 메뉴

전체 파일에 영향을 미치는 작업 메뉴이다. 이 메뉴의 선택 항목들을 이해하려면 .dxg와 .sig 파일 형식을 이해해야 한다. QuickTac의 기본 파일 형식은 .dxg이다. 그러나 .dxg 파일을 텍스베리로 가져오기 위해서는 해당 형식을 .sig 파일로 변환해야 한다.

### (2) 편집 메뉴

편집 메뉴는 일반적인 잘라내기와 붙여넣기 작업으로 그림 편집을 용이하게 한다. 메뉴의 선택 항목은 아래에 나열되어 있지만 메뉴에는 나타나지 않는 ‘단일 점 편집’ 기능도 있다.

### (3) 보기 메뉴

보기 메뉴를 사용하면 QuickTac 창의 그리기 영역에서 사물이 보이는 방식을 제어할 수 있다. 종종 보기 메뉴는 작동 모드를 설정한다. 해당 모드가 실행되면 메뉴에서 해당 항목 옆에 확인 표시가 나타난다.

### (4) 도구 메뉴

도구 메뉴는 촉각 그래픽 이미지를 만들기 위해 점을 찍는 다양한 방법을 제공한다. QuickTac에서 수행하는 대부분의 작업은 도구 메뉴와 편집 메뉴에서 수행하게 된다. 단, 파일 메뉴에서 페이지 설정을 사용하여 점자용 QuickTac을 구성할 때까지 도구 메뉴의 도구를 사용할 수 없다. 메뉴에 있는 대부분의 도구는 도구를 선택하면 이미지 영역 오른쪽 창에 자체 단계별 지침이 표시된다.

도구 메뉴의 그리기 도구는 도구 모음에서도 찾을 수 있다. 선과 곡선을 그리는 도구에는 선 두께와 선 패턴에 대한 설정이 있다. 사방이 닫힌 기하학적 도형을 만드는 도구에는 선 두께, 선 패턴, 채우기에 대한 설정이 있다.

#### (5) 도움말 메뉴

도움말 메뉴는 온라인 설명서와 유용한 기타 정보 소스를 제공한다.

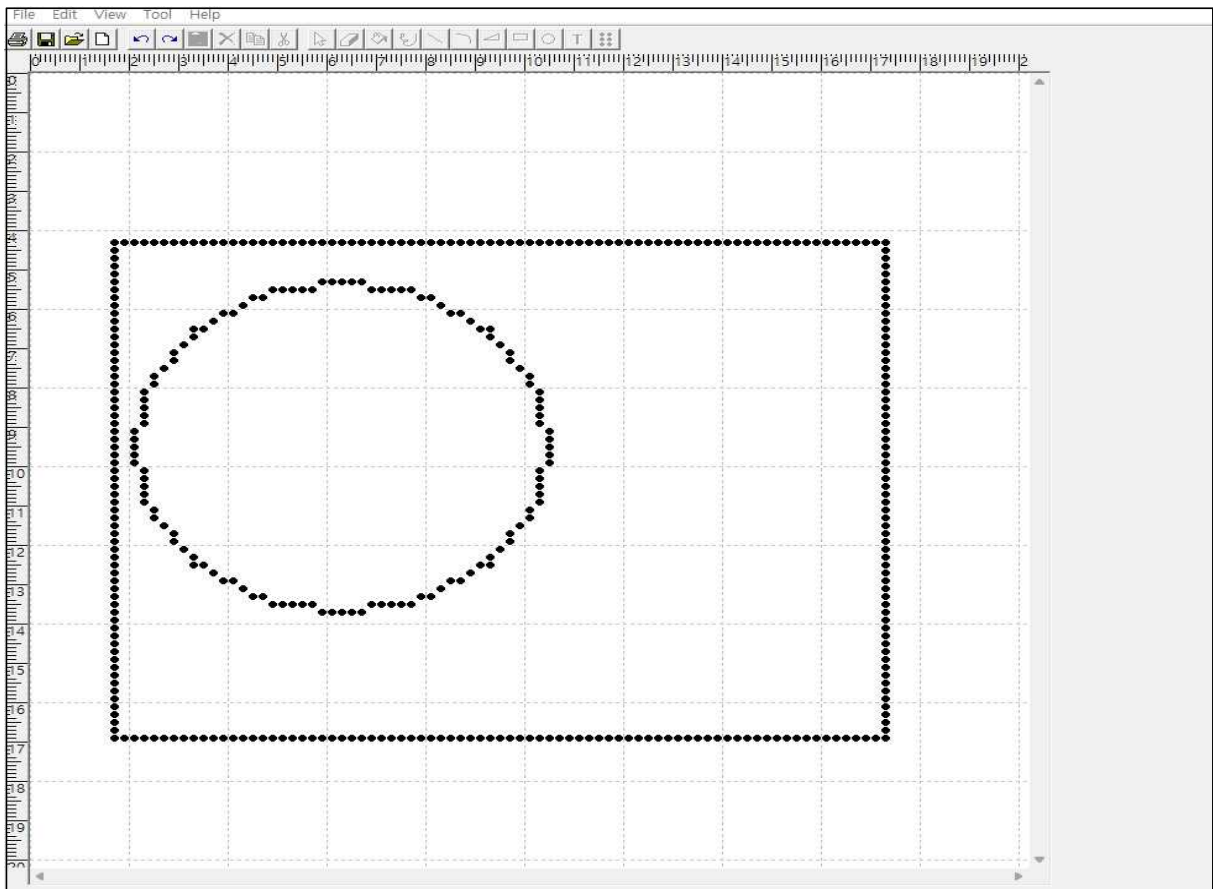


그림 8-2 QuickTac의 도형 생성 구현 화면

### 3) QuickTac의 동작

QuickTac의 기능을 효율적으로 사용하기 위해서는 다음의 개념 및 지침의 이해가 필요하다.

#### (1) 파일 형식

QuickTac의 기본 파일 형식은 일반적인 편집에 사용되는 \*.dxg 파일이다. 그러나 \*.dxg 파일은 Duxbury Braille Translator로 불러올 수 없다. DBT의 경우 해당 형식을 \*.sig 파일로 변환해야 한다.

파일 작업을 할 때에는 \*.dxg 형식으로 자유롭게 불러오고 저장할 수 있지만, 완료되면 \*.sig 파일로 저장하여 DBT로 불러와야 한다.

DBT의 향후 업데이트에서 .dxg 파일을 불러오는 것이 허용될 수 있겠으나, 현재 DBT용 QuickTac을 사용하려면 두 개의 서로 다른 파일을 관리해야 한다. \*.dxg 파일을 생성하거나 수정하려면 파일 메뉴에서 저장 옵션을 사용해야 한다. \*.sig 형식은 점자 출력기용 그래픽의 일반적인 파일 형식이다. DBT로 불러오려면 \*.sig 파일이 필요하다. \*.sig 파일을 만들거나 수정하려면 파일 메뉴에서 다른 이름으로 저장 옵션을 사용하면 된다.

#### (2) DBT로 파일 불러오기

촉각 그래픽을 만든 후 DBT로 해당 파일을 가져오려면 2단계 과정이 필요하다.

- ① QuickTac에서: QuickTac으로 그린 후 파일을 \*.sig 파일로 저장한다. QuickTac의 파일 메뉴로 이동한 후, 다른 이름으로 저장을 사용하여 \*.sig 파일로 저장한다.
- ② DBT에서: 이미지를 만들고 싶은 곳에 커서를 위치한다. DBT 레이아웃 메뉴에서 Picture를 선택한 후, 추가(Add)를 선택한다. 파일 열기 대화 상자가 나타나면, QuickTac에서 생성한 \*.sig 파일을 찾아 열기 버튼을 선택한다.

#### (3) 배경 이미지 사용

QuickTac을 사용하여 처음부터 새로운 촉각 그래픽을 만들 수 있지만 기존 시각적 이미지를 이용하여 촉각 그래픽을 만들 수도 있다. 예를 들어 주요 지질학적 특징이 표시된 유타주의 지도 이미지가 있을 때, 산맥, 큰 호수 및 국립공원의 위치를 나타내는 촉각 버전을 렌더링할 수 있다. 페이지 설정 대화 상자의 컨트롤을 사용하여 촉각 렌더링을 만들기 위한 이미지를 선택하고 크기를 조정하고 위치를 지정할 수 있다. 페이지 설정 대화 상자를 닫으면 유타주의 지도 이미지가 그리기 캔버스의 배경으로 표시되어 촉각 그래픽을 그리는 데 도움이 된다. 보기 메뉴에서 배경 이미지의 출력을 On/Off 하면서 촉각 그래픽 버전만 확인하고 작업을 할 수도 있다.

#### (4) 페이지 설정 및 점자 출력기 구성

파일 메뉴의 페이지 설정 기능을 사용하여 점자 출력기를 구성하고 그림의 크기와 위치를 설정하는 과정이 필요하다. Duxbury DBT에서 문서의 페이지 크기와 레이아웃을 설정하려면 문서를 특정 점자 출력기로 설정해야 하는 것과 같이, QuickTac을 사용할 때도 동일하게 작업을 수행해야 한다. 즉, 완성된 촉각 그래픽을 보낼 점자 출력기 및 용지 크기를 선택한다. 이를 통해 그림에 사용할 수 있는 최대 크기가 결정된다.

QuickTac의 설정 과정은 DBT와 유사하다. 사용자가 점자 출력기 선택을 위한 대화 상자를 호출하는 새로 만들기 버튼을 선택하면, 페이지 설정 대화 상자가 출력된다. 해당 대화 상자의 일반 탭이 나타나고 첫 번째 컨트롤은 사용하려는 항목을 선택할 수 있는 점자 출력기의 드롭다운 목록이다. 점자 출력기 모델 드롭다운 목록을 열고 점자 출력기 종류를 선택한다. 이후 용지 구성을 설정한 다음 필요한 경우 배경 이미지 위치를 설정할 수도 있다.

점자 출력기 및 배경 이미지를 선택한 후의 페이지 설정 대화 상자에서는 점 해상도, 용지 크기, 페이지 방향 선택기 및 페이지 여백을 채울 필드에 대한 컨트롤이 있다. 또한 결과 점의 간격, 점의 이미지 크기, 이미지 크기에 대한 인치 또는 밀리미터의 값이 있다.

선택 버튼을 클릭하면 읽어 들인 배경 이미지의 썸네일을 표시하는 작은 창이 나타난다. 이 창에서 바깥쪽 검은색 직사각형은 최대 페이지 크기와 비율을 나타낸다. 안쪽 파란색 사각형을 조작하여 점자 출력기 페이지에서 배경 이미지의 크기와 위치를 조정할 수 있다. 배경 이미지를 배치한 후 확인을 누르면 기본 QuickTac 화면으로 돌아가게 된다. 배경 배치와 일치하도록 창이 자동으로 새로 고쳐지지 않은 경우 창 크기를 전환하여(예: "페이지에 맞추기" 보기에서 100% 보기로 또는 그 반대로 전환하여) 항목을 제자리에 가져올 수 있다.

#### (5) 점자 출력기에 대한 참고 사항

QuickTac에서 직접 촉각 그래픽을 점자 출력기로 출력할 수 있다. 그래픽을 \*.sig 파일로 저장하여 DBT로 불러올 수도 있다.

특정 점자 출력기에 따라 텍스트 모드(일반 점자)와 촉각 그래픽 모드의 점 간격이 다를 수 있다. 별도의 그래픽 모드를 지원하는 일부 점자 출력기의 경우 동일한 페이지에 텍스트와 촉각 그래픽이 있는 경우 생성된 점자에 비표준 점자 간격이 있을 수 있다. 각 점자 출력기 모델은 촉각 그래픽을 처리하는 것과 관련하여 차이가 있으므로 테스트 과정이 필요할 수 있다.

## 2 점사랑 5.0 축각 그래픽 기능 설계

### 1) 축각 그래픽 소프트웨어의 주요 기능

점사랑 5.0 내에 포함될 축각 그래픽 소프트웨어의 주요 기능은 다음과 같다.

- 축각 그래픽 생성 및 편집 기능을 구현한다.
- 표준 그래픽 파일(\*.jpg, \*.gif, \*.bmp, \*.png)을 불러와 텍타일 이미지로 제작하기 위한 바탕 그림으로 사용하는 기능을 지원한다.
- 점역 소프트웨어(점사랑)와 연동하여 출력 기능을 지원한다.
- 텍스트와 그래픽 처리 방식이 다른 점자 프린터의 특성을 고려하여 텍스트와 이미지를 순차적으로 인쇄(인덱스 계열 프린트 지원 필수)하는 기능을 지원한다.
- 직선, 곡선, 삼각형, 사각형, 원 등 그리기 도구를 지원한다.

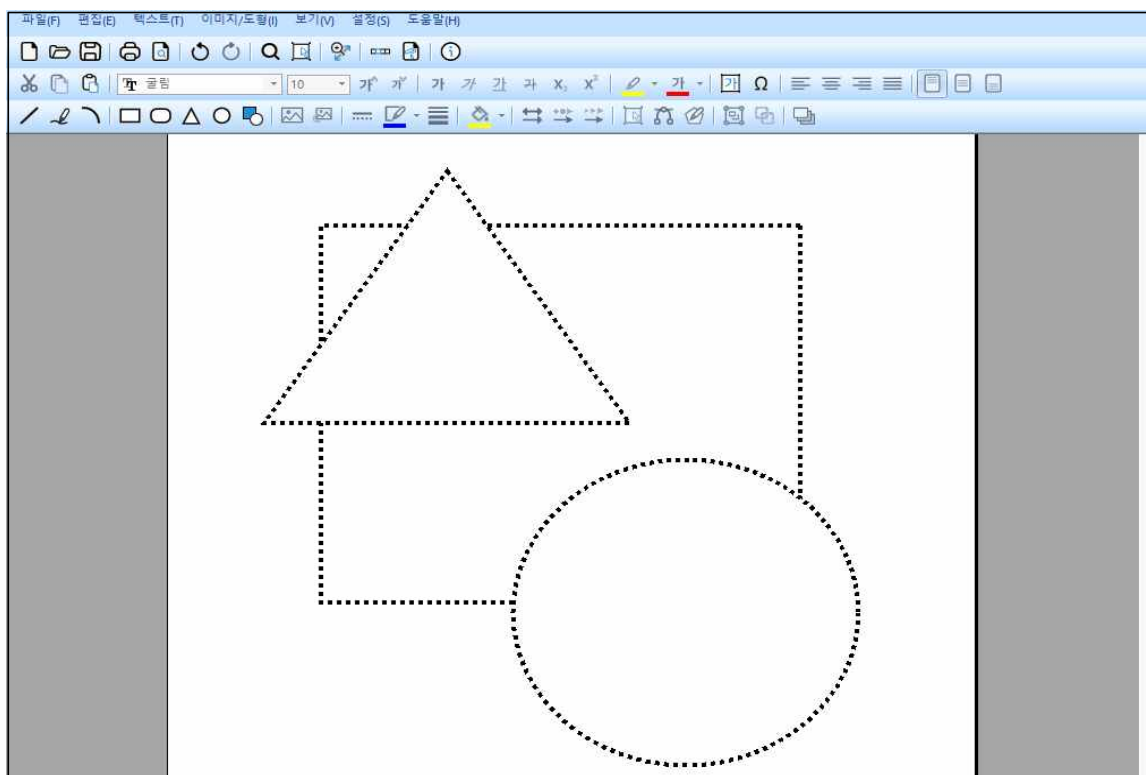


그림 8-3 점사랑 5.0 축각 그래픽의 도형 생성 구현 화면

## 2) 촉각 그래픽 소프트웨어의 기능 설계

촉각 그래픽 소프트웨어의 메뉴는 크게 ‘파일’, ‘편집’, ‘텍스트’, ‘이미지/도형’, ‘보기’, ‘설정’, ‘도움말’로 구성하였다.

### (1) 파일 메뉴

파일 메뉴에 포함될 주요 기능은 새로 만들기, 열기, 저장, 다른 이름으로 저장, 점자 인쇄 미리보기, 점자 인쇄, 목자 인쇄 미리 보기, 목자 인쇄, 최근에 사용한 파일, 끝내기 등이다. 각 기능의 주요 특징은 <표 8-1>과 같다.

표 8-1 파일 메뉴 기능 설계

메뉴	기능	설명
파일	새로 만들기	- 현재 문서와 별도로 새로운 촉각 그래픽 프로그램 창을 실행
	열기	- 저장된 bgfx 파일을 대화 상자를 호출하여 선택
	저장	- 현재 편집창의 문서를 파일로 저장 - 파일명이 지정되지 않은 경우 ‘저장’ 대화 상자가 호출
	다른 이름으로 저장	- 현재 편집창의 문서를 다른 이름, 또는 다른 파일 형식으로 저장
	점자 인쇄 미리보기	- 점자로 인쇄될 내용을 미리 확인
	점자 인쇄	- 편집중인 문서를 점자 프린터로 인쇄
	목자 인쇄 미리 보기	- 목자로 인쇄될 내용을 미리 확인
	목자 인쇄	- 편집 중인 문서의 내용을 일반 프린터로 인쇄
	최근에 사용한 파일	- 최근에 편집했던 문서의 목록을 표시 - 기본 목록 개수는 ‘10’이며, 개수 설정은 기본 환경 설정 메뉴에서 지정
	끝내기	- 촉각 그래픽 프로그램을 종료 - 문서가 변경이 있는 경우 저장 여부를 묻는 대화 상자를 호출

### (2) 편집 메뉴

편집 메뉴에 포함될 주요 기능은 되돌리기, 다시 실행, 잘라내기, 복사, 붙여넣기, 선택하여 붙여넣기, 지우기, 모두 선택, 찾기, 바꾸기, 다시 찾기 등이다. 각 기능의 주요 특징은 <표 8-2>와 같다.

표 8-2 편집 메뉴 기능 설계

메뉴	기능	설명
편집	되돌리기	- 편집된 문서의 상태를 이전 상태로 되돌리기
	다시 실행	- 되돌리기 실행을 취소
	잘라내기	- 블록으로 선택된 영역의 내용을 잘라내어 클립보드에 저장
	복사	- 블록으로 선택된 영역의 내용을 복사하여 클립보드에 저장
	붙여넣기	- 잘라내기 또는 복사한 내용을 커서의 위치에 삽입
	선택하여 붙여넣기	- 클립보드의 내용 중 텍스트, 이미지 등을 구분하여 붙여넣기 실행
	지우기	- 선택된 영역 또는 개체를 삭제
	모두 선택	- 문서 전체의 내용을 선택
	찾기	- 문서 내에서 특정 문자열을 찾아 그 위치로 커서를 이동 - 찾기 대화 상자가 수행되고 이후 F3키를 이용하여 계속 찾기 지원
	바꾸기	- 문서에서 특정 문자열을 찾아 다른 문자열로 변경
	다시 찾기	- 찾기 대화 상자를 통해 최근 검색한 문자열과 검색 방향을 기준으로 반복 찾기를 수행

### (3) 텍스트 메뉴

텍스트 메뉴에 포함될 주요 기능은 글상자 삽입, 기호, 굵게, 기울임꼴, 밑줄, 취소선, 아래 첨자, 위 첨자, 글자 모양, 글자 크기, 가로 맞춤, 세로 맞춤 등이다. 각 기능의 주요 특징은 <표 8-3>과 같다.

표 8-3 텍스트 메뉴 기능 설계

메뉴	기능	설명
텍스트	글상자 삽입	- 마우스 커서를 드래그하여 글상자의 크기를 정한 후 마우스 클릭을 해제하여 글상자 생성 후 텍스트 입력
	기호	- 한글 자모, 그리스어, 원문자, 괄호 문자 등의 특수문자 입력
	굵게	- 선택한 글자를 굵게 변경하거나 새로 입력하는 글자를 굵게 표시
	기울임꼴	- 선택한 글자를 기울임 꼴로 변경하거나 글자를 기울임 꼴로 입력
	밑줄	- 선택한 글자 아래에 밑줄을 표시하거나 밑줄이 표시된 글자를 입력
	취소선	- 선택한 글자 중간에 취소선을 표시하거나 취소선 표시된 글자를 입력
	아래 첨자	- 선택한 글자를 아래 첨자로 변경하거나 아래 첨자를 입력
	위 첨자	- 선택한 글자를 위 첨자로 변경하거나 위 첨자를 입력
	글자 모양	- 글자의 세부 모양을 설정하는 대화 상자 실행
	글자 크기	- 글자 크기를 설정하는 대화 상자 실행
	가로 맞춤	- 현재 설정된 가로 맞춤을 체크, 글상자 내부 행의 가로 맞춤을 지정
	세로 맞춤	- 현재 설정된 세로 맞춤을 체크, 글상자 내부 행의 세로 맞춤을 지정

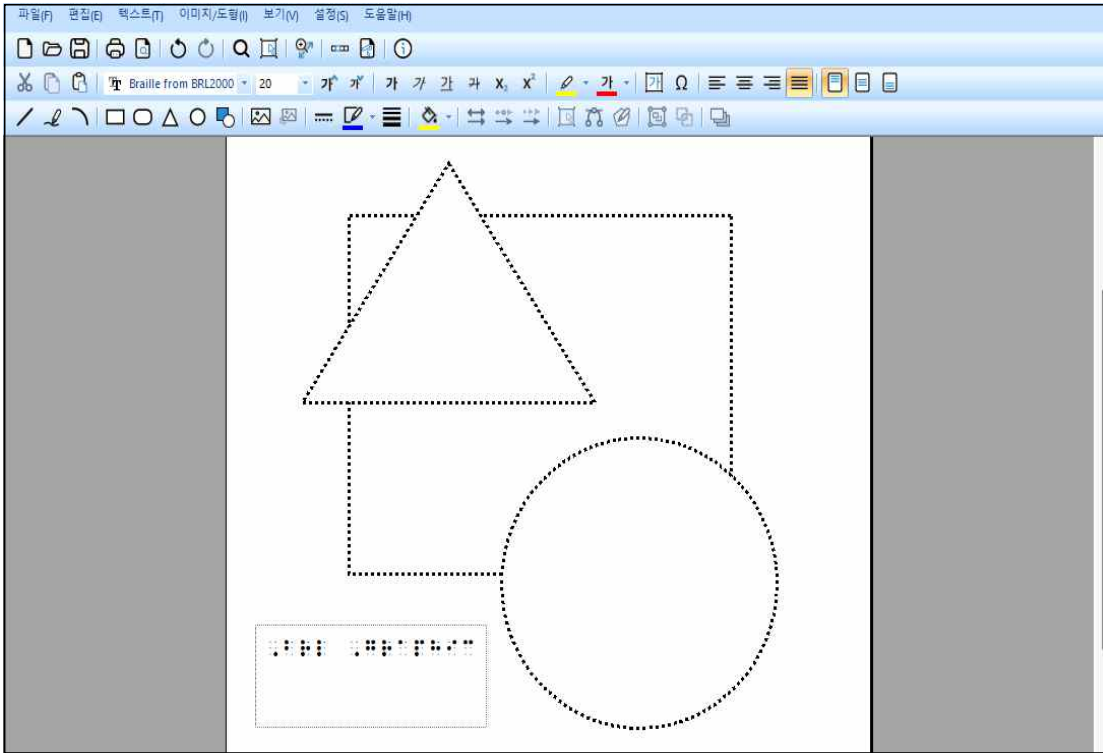


그림 8-4 점자 텍스트가 포함된 도형 생성 구현 화면

(4) 이미지/도형 메뉴

이미지/도형 메뉴에 포함될 주요 기능은 선 삽입, 도형 삽입, 이미지 삽입, 이미지 교체, 크기 변경, 곡률 변경, 다각형 편집, 틀 순서 변경, 개체 묶기/풀기, 고치기 등이다. 각 기능의 주요 특징은 <표 8-4>와 같다.

표 8-4 이미지/도형 메뉴 기능 설계

메뉴	기능	설명
이미지/도형	선 삽입	- 직선, 호, 곡선, 다각선, 자유 그리기 등의 선 삽입
	도형 삽입	- 사각형, 삼각형, 직각삼각형, 마름모, 평행사변형, 오각형, 원 등의 도형을 삽입
	이미지 삽입	- 이미지 파일을 선택할 수 있는 대화 상자 실행
	이미지 교체	- 이미지 교체를 위한 이미지 파일을 선택할 수 있는 대화 상자 실행
	크기 변경	- 선택된 개체의 가로세로 크기를 변경
	곡률 변경	- 사각형, 원형사각형의 곡률을 변경
	다각형 편집	- 기존 선 및 다각형 도형의 점의 위치를 변경할 수 있는 모드로 진입
	틀 순서 변경	- 선택한 개체의 앞뒤 순서를 변경
	개체 묶기/풀기	- 선택된 두 개 이상의 개체를 하나의 그룹으로 묶기 - 그룹으로 묶인 개체를 그룹 해제
	고치기	- 틀 또는 도형의 너비, 높이 등의 속성을 변경

### (5) 보기 메뉴

보기 메뉴에 포함될 주요 기능은 화면 확대/축소, 빈칸/탭 보기, 이미지 보기, 투명선 보기, 격자, 도구 막대, 프로그램 스타일 등이다. 각 기능의 주요 특징은 <표 8-5>와 같다.

**표 8-5** 보기 메뉴 기능 설계

메뉴	기능	설명
보기	화면 확대/축소	- 화면의 크기를 일정 비율로 확대하거나 축소
	빈칸/탭 보기	- 빈칸 또는 탭과 같은 서식 부호를 화면에 표시하거나 숨김
	이미지 보기	- 이미지를 화면에 표시하거나 숨김
	투명선 보기	- 투명선을 화면에 표시하거나 숨김
	격자	- 격자를 화면에 표시하거나 숨김
	도구 막대	- 화면에 도구 막대(툴바)를 표시하거나 숨김
	프로그램 스타일	- 원도 테마에 맞게 화면 UI를 변경

### (6) 설정 메뉴

설정 메뉴에 포함될 주요 기능은 편집 용지, 페이지 배경, 점자 프린터 설정, 인쇄 설정 등이다. 각 기능의 주요 특징은 <표 8-6>과 같다.

**표 8-6** 설정 메뉴 기능 설계

메뉴	기능	설명
설정	편집 용지	- 편집 용지 대화 상자에서는 용지의 종류, 방향, 제책, 여백 등을 지정 - 사용자 정의에서는 폭과 길이를 직접 입력하여 종이의 크기를 설정
	페이지 배경	- 축각 그래픽 프로그램 화면의 배경색을 지정
	점자 프린터 설정	- 점자 프린터의 종류를 선택하고 속성 등을 변경하는 대화 상자 실행
	인쇄 설정	- 묵자 프린터의 종류를 선택하고, 속성 등을 변경하는 대화 상자 실행

### (7) 도움말 메뉴

도움말 메뉴에 포함될 주요 기능은 축각 그래픽 도움말, 축각 그래픽 정보, 점자 종합정보 누리집 등이다. 각 기능의 주요 특징은 <표 8-7>과 같다.

**표 8-7** 도움말 메뉴 기능 설계

메뉴	기능	설명
도움말	축각 그래픽 도움말	- 축각 그래픽 프로그램 도움말 원도를 실행
	축각 그래픽 정보	- 축각 그래픽 프로그램 버전 정보가 표시된 대화 상자 실행
	점자 종합정보 누리집	- 점자 종합정보 누리집 사이트로 이동



2023년 목자-점자 병렬말뭉치 구축  
(II. 점역 소프트웨어 고도화)

# 제 9 장

## 악보 점역 소프트웨어

1. 해외 악보 점역 소프트웨어 기능 분석
2. 점사랑 5.0 악보 점역 기능 설계





## 1 해외 악보 점역 소프트웨어 기능 분석

GOODFEEL은 미국의 Dancing Dots사에서 1997년에 출시한 현재 전 세계에서 가장 발달한 악보 점역 소프트웨어이다. 함께 제공되는 Lime이라는 프로그램으로 편집한 묵자 악보를 점자 악보로 자동 점역하는 기능을 가지고 있고, 또한 Finale, Sibelius 등의 악보 편집 프로그램에서 작성한 MusicXML 형식의 악보를 불러와 점역할 수도 있다. 음악 점역 소프트웨어 개발이라는 큰 과제를 두고 가장 먼저 해야 할 일은 이 GOODFEEL 프로그램에 대한 연구와 분석이었다.

하지만 풀 패키지를 구매할 경우 \$1,595(2024년 2월 환율 기준 약 217만원)의 높은 가격으로 개인이 구매하기에는 그 부담이 매우 크고, 한글이 전혀 지원되지 않는 등의 단점도 있다.

### 1) GOODFEEL의 주요 특징 및 기능

GOODFEEL의 주요 특징은 다음과 같다.

#### (1) GOODFEEL의 프로그램 구성

GOODFEEL은 묵자 악보를 점자 음악, 언어 및 청각 음악 신호를 포함한 접근 가능한 악보 형식으로 자동화하는 소프트웨어 제품군이다. 시각장애인 사용자도 비시각장애인과 마찬가지로 악보를 읽고 쓸 수 있는 환경을 제공한다. 인쇄/확대 인쇄 표기법과 점자 악보를 생성하고, 음성 및 음악 신호로 악보를 표시하는 기능을 가지고 있는데, 이것은 하나의 프로세스로 동시에 진행되며 동기화된다. 사용자는 이 중 자신이 선호하는 모드를 선택할 수 있다.

이 제품군에는 4개의 소프트웨어 타이틀이 포함되어 있다.

- ① SharpEye: 묵자 악보 스캔용
- ② Lime: 묵자 악보 입력 및 편집용
- ③ Lime Aloud: (선택 사항) "악보 읽어주기" 기능을 통해 JAWS 사용자 및 NVDA와 같은 다른 스크린리더가 Lime에 접근할 수 있도록 하는 스크립트
- ④ GOODFEEL: Lime의 묵자 악보를 해당 점자 악보로 자동 변환



그림 9-1 GOODFEEL 프로그램 화면

### (2) 각 구성 요소의 기능

Lime의 악보 편집 기능은 음악 표기를 위한 일종의 워드 프로세서라고 생각하면 된다. Lime에 악보의 음표와 관련 정보를 입력하고 Lime의 파일 메뉴에서 “Launch GOODFEEL”을 선택하기만 하면 자동으로 악보가 생성된다.

GOODFEEL로 악보를 점자로 변환하기 위해서는 잘 만들어진 Lime의 악보가 필요하다. 메모와 기타 주석을 Lime의 편집기에 직접 입력할 수 있고, MusicXML 또는 NIFF 인터체인지 형식으로 저장된 악보를 Lime으로 가져올 수 있다. SharpEye 또는 기타 음악 OCR 어플리케이션을 사용하여 MusicXML 또는 NIFF를 통해 Lime으로 전달할 수 있는 묵자 악보를 스캔할 수 있다.

### (3) 점역 옵션

GOODFEEL은 사용자에게 제목, 박자, 가사 및 리허설 마커 등을 약자 또는 정자의 점자로 점역되는 옵션 등 많은 점역 옵션을 제공한다. GOODFEEL 4 버전은 최초로 UEB(Unified English Braille, 통일 영어 점자 규정)에 대한 지원을 포함하고 있다. Options > 점역 옵션 메뉴 항목을 사용하여 검토할 수 있다.

#### (4) 점역 처리 과정

- ① Intermediate(중간) 대화 상자: Lime 파일을 GOODFEEL에 불러오면 몇 가지 일반적인 오류가 있는지 검사되고 오류가 발견되면 GOODFEEL은 대화 상자의 Report 탭에 이 오류를 나열한다. 스크린리더 사용자는 Tab 키를 한 번 눌러 오류 목록에 들어가 화살표를 눌러 내용을 읽을 수 있다.
- ② 프로세스 버튼을 탭하거나 ALT+P를 눌러 처리한다. 프로세스 버튼(왼쪽 아래)은 점자 악보를 생성하고 점자를 보낼 위치를 묻는 대화 상자를 표시한다.
- ③ Braille Destination(점자 목적지) 대화 상자: GOODFEEL이 음악을 점자로 변환하면 다음을 수행할 수 있는 Braille Destination 대화 상자가 나타난다.
  - 점자 출력 버튼으로 점자를 출력한다.
  - 점자 편집기(Braille Editor) 버튼으로 점자를 편집하거나 본다.
- ④ BrailleView에 대한 참고 사항(Braille Editor): BrailleView는 별도의 응용 프로그램이다. 그 주요 목적은 사용자가 점자 출력을 점자 프린터에 보내기 전에 검토할 수 있도록 하는 것이다.

## 2) GOODFEEL의 메뉴 구성

GOODFEEL의 메뉴 구성은 크게 File(파일), Options(옵션), Help(도움말)로 구성되어 있다.

#### (1) File(파일) 메뉴

- ① Load(불러오기): 이 메뉴 항목을 통해 Lime과 MIDI 파일을 GOODFEEL에 로드할 수 있다. 현재 GOODFEEL은 Lime과 MIDI 파일만 인식하고 지원한다. MusicXML과 NIFF 파일은 Lime에 가져온 후 GOODFEEL에 전달할 수 있다. 이 메뉴 항목을 선택하면 GOODFEEL을 열 파일을 지정할 수 있는 파일 선택 대화 상자가 열린다.
- ② Open Braille File(점자 파일 열기): GOODFEEL은 이전에 점역된 모든 음악 점자 파일을 "GOODFEEL Output" 폴더에 저장한다. 이 메뉴 항목은 해당 파일의 보기, 편집 및 점자 인쇄를 용이하게 한다. 이 메뉴 항목을 선택하면 파일 열기 대화 상자 창이 표시된다. 파일을 선택한 후 GOODFEEL은 Braille Destination 대화 상자를 표시한다.
- ③ Open Output Folder(출력 폴더 열기): 이 항목을 선택하면 이전에 점역된 점자 파일들이 있는 창이 열린다. 파일을 선택해서 더블클릭하면 BrailleView에서 해당 파일을 확인할 수 있다.
- ④ Exit(종료하기): 이 항목을 선택하면 GOODFEEL이 종료된다.

## (2) Options(옵션) 메뉴

- ① Transcription Options(점역 옵션): 이 메뉴에서 점자 출력 형식을 제어할 수 있다.

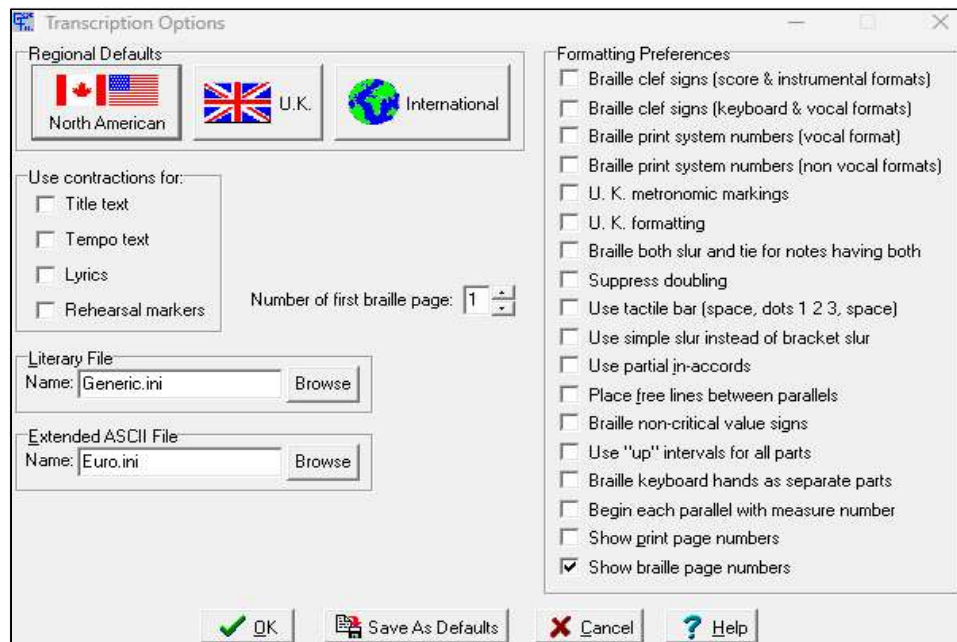


그림 9-2 GOODFEEL의 점역 옵션 설정 화면

- ② Automation Options(자동화 옵션): 이 옵션은 음악을 점역하는 다음 단계에서 GOODFEEL이 진행하는 방법을 제어한다.
- ③ Embosser Setup(점자 프린터 설정): 이 메뉴는 GOODFEEL의 점자 출력을 제어하는 데 사용된다.
- ④ Lime Options(라임 옵션): Lime 메뉴는 “Options(옵션)”으로만 구성되어 있다.
- Options(옵션): 사용자가 Lime 파일을 읽을 때 사용할 표기법 환경을 설정할 수 있다.
- ⑤ Title Page Options(제목 페이지 옵션): 제목 페이지 정보 메뉴를 사용하면 점역하기 전에 개별 제목 페이지 구성 요소를 변경할 수 있다.

## (3) Help(도움말) 메뉴

- ① Help(도움말): GOODFEEL 사용에 관한 도움말을 제공한다.
- ② www.dancingdots.com: Dancing Dots사의 홈페이지로 이동한다.
- ③ About GOODFEEL...(GOODFEEL에 대하여...): 사용 중인 GOODFEEL에 대한 기본 정보(버전, 시리얼 넘버, 연락처 등)를 확인할 수 있다.

## 2 점사랑 5.0 악보 점역 기능 설계

본 절에서는 GOODFEEL을 통해 분석한 결과를 중심으로 점사랑 5.0에 포함될 악보 점역 기능을 설계하였다. 점사랑 5.0의 악보 점역 기능은 GOODFEEL을 만든 Dancing Dots사와의 라이선스 계약을 통해 GOODFEEL의 악보 점역 엔진 및 Lime의 편집 기능을 연동하여 사용한다. 점사랑 5.0 사용자는 추가 비용 지불 또는 기간 제한 없이 제공된 Lime 및 GOODFEEL 프로그램을 한국어 윈도우 운영체제에서 사용할 수 있다.

### 1) 악보 점역 소프트웨어의 주요 기능

점사랑 5.0 내에 포함될 악보 점역 소프트웨어의 주요 기능은 다음과 같다.

- 악보 편집 및 점역을 통합한 소프트웨어를 구현한다.
- GUI 기반의 악보 제작 및 편집 기능을 지원한다.
- 점역 소프트웨어(점사랑)와 연동하여 출력 기능을 지원한다.
- 점역을 위한 MusicXML 파일 형식을 지원한다.
- 건반악기, 관현악, 성악 등의 악보를 점역하는 기능을 지원한다.
- 제목 텍스트, 박자 표시, 가사 및 리허설 마커를 점역하는 옵션을 포함한다.
- 사업 종료 이후에도 점사랑과 연동하여 지속적으로 사용 가능한 소프트웨어를 구현한다.

### 2) 악보 점역 소프트웨어의 기능 설계

악보 점역 소프트웨어의 메뉴는 크게 ‘파일’, ‘옵션’, ‘도움말’로 구성하며, 기본적으로 GOODFEEL의 메뉴 구성과 크게 다르지 않게 구성하였다. 가장 큰 차이점은 한글 옵션의 추가이다. 편집에 관련한 메뉴는 Lime의 기능 및 사용자 인터페이스를 준용한다. 2024년 2월 27일 현재 Lime 메뉴의 한글화 작업 중이며, 최종 출시될 사용자 배포용 버전은 Lime에서 한글 메뉴를 지원할 예정이다.



그림 9-3 Lime 프로그램의 한글 입력 화면

## (1) 파일 메뉴

파일 메뉴에 포함될 주요 기능은 원본 불러오기, 점자본 불러오기, 종료 등이다. 각 기능에 대한 설명은 다음과 같다.

표 9-1 파일 메뉴 기능 설계

영역	기능	설명
파일	원본 불러오기	<ul style="list-style-type: none"> <li>- 악보 점역을 위해 다음 형식의 파일을 불러오는 기능 <ul style="list-style-type: none"> <li>• 라임(.lim) 형식</li> <li>• MusicXML(.mxml) 형식</li> </ul> </li> <li>- '자동화 옵션' 설정 및 'Input File checker' 결과에 따라 다음 작업 중에서 수행 <ul style="list-style-type: none"> <li>• 'Intermediate dialog' 대화 상자 열기</li> <li>• 이전 GOODFEEL 출력 파일을 덮어쓸지 묻기</li> <li>• 음악 점자 인쇄</li> <li>• 점자 출력 대상 묻기</li> <li>• 닫기</li> </ul> </li> </ul>
	점자본 불러오기	<ul style="list-style-type: none"> <li>- 점자 인쇄하거나 점자 편집기에 출력하기 위해 점역된 악보 파일을 불러오는 기능</li> </ul>
	종료	<ul style="list-style-type: none"> <li>- 악보 점역 기능 종료</li> </ul>

## (2) 옵션 메뉴

옵션 메뉴는 점역 옵션, 자동화 옵션, 점자 프린터 설정, Lime 옵션, 제목 페이지 정보 등으로 구성되어 있다.

표 9-2 옵션 메뉴 기능 설계

영역	기능	설명
선택사항	점역 옵션	<ul style="list-style-type: none"> <li>- 점자 악보로 점역할 때 다음 사항을 점자 서식에 반영</li> <li>• 한글/영어 악자 사용/사용안함(제목, 빠르기, 가사, 리허설 마커 등)</li> <li>• 페이지 표시/안함(점자, 목자)</li> <li>• 음자리표, 시스템 번호, 음정표 표시 방법 등</li> </ul>
	자동화 옵션	<ul style="list-style-type: none"> <li>- 원본 파일을 불러와 점역하기 전에 'input file checker'에서 발견한 경고나 오류 내용을 확인하는 방법 설정(확인하지 않음, 오류 확인, 경고 확인, 모두 확인)</li> <li>- 점역 결과 출력 장치 기본값 설정(점자 프린터, 점자 편집기, 파일 저장)</li> </ul>
	점자 프린터 설정	<ul style="list-style-type: none"> <li>- 점자 프린터에 대한 다음 값들을 설정하는 기능</li> <li>• 줄당 칸 수</li> <li>• 페이지당 줄 수</li> <li>• 프린터 종류(원도/포트에 연결된 점자 프린터)</li> <li>• 양면/단면 인쇄</li> <li>• 인쇄 종료 후 폼피드 추가</li> <li>• 기본값으로 저장</li> </ul>
	Lime 옵션	<ul style="list-style-type: none"> <li>- Lime에서 정의한 악보 표기 문맥을 점역하기 위해 지정하는 기능</li> </ul>
	제목 페이지 정보	<ul style="list-style-type: none"> <li>- 제목 페이지 내용과 점역 방법을 다음과 같이 지정하는 기능</li> <li>• Lime으로부터 제목 페이지 내용 불러오기</li> <li>• 제목 페이지의 일부 내용을 직접 입력(점역자, 후원자 등)</li> <li>• 각 성부마다 점역</li> <li>• 처음에만 점역</li> <li>• 점역하지 않음</li> </ul>

## (3) 도움말 메뉴

도움말 메뉴를 선택하면 .CHM 형식의 도움말을 제공한다.



2023년 목자-점자 병렬말뭉치 구축  
(II. 점역 소프트웨어 고도화)

# 제 10 장

## 소프트웨어 활용법에 대한 온라인 콘텐츠 개발

1. 온라인 콘텐츠의 필요성
2. 온라인 콘텐츠의 구성





본 장에서는 소프트웨어 활용법에 대한 동영상 형태의 온라인(유튜브 등) 콘텐츠 개발의 필요성과 그 구성에 대해 제시하였다.

## 1 온라인 콘텐츠의 필요성

최근 많은 종류의 교육 콘텐츠들이 온라인 서비스가 가능한 형태로 제작되고, 여러 온라인 플랫폼을 통해 서비스되는 추세이다. 여러 소프트웨어의 활용법 또한 온라인을 통해 강의 동영상으로 송출되고 있으며, 사용자의 입장에서 언제 어디서나 접근이 용이하고 원하는 만큼 반복 학습이 가능한 온라인 콘텐츠는 이제 당연한 서비스로 인식되고 있다.

이에 점사랑 5.0도 그런 사용자의 요구를 반영한 온라인 콘텐츠를 개발 및 보급하여 프로그램의 접근성과 사용자의 편의성을 높이는 환경을 조성하고자 한다.

## 2 온라인 콘텐츠의 구성

### 1) 구성 방향

- 점사랑의 활용도를 높이고, 완성도 높은 점자 도서를 제작할 수 있는 능력을 배양하는 내용으로 구성한다.
- 점자 초보자도 쉽게 접근할 수 있는 프로그램 활용 방법을 제시한다.
- 난이도와 우선순위를 반영한 커리큘럼에 따라 3개의 카테고리로 나누고, 총 12차시로 구성한다.
- 각 콘텐츠는 10분 내외의 동영상 형태로 제작한다.

### 2) 구성안

각 카테고리별 콘텐츠는 사용자의 숙련도와 점사랑 5.0의 특징, 새로운 기능들을 고려하여 다음과 같이 구성하였다. 구성안의 세부 내용은 제작 과정 중 일부 변경될 수도 있다.

## (1) 1~4차시: 점사랑 5.0 소개

### ① 점자 출판물 제작을 위한 기초 지식 및 점사랑의 주요 특징 소개

- 2020/2023 한글 점자 규정 지원, UEB 지원
- 다양한 편집용 점자코드 지원, 표 및 점자판 자체 생성/점역 지원
- 한글&워드 문서 포맷 임포트, 글꼴/문단/표 서식 점역
- 수학 수식 편집기 지원(온/오프라인용 편집기)
- 음악 점자 점역, 축각 그래픽 제작/출력
- 옛 글자, 유니코드(점자, 한자 및 특수기호) 점역 지원

### ② 소프트웨어 다운로드 및 설치 방법

- GOODFEEL, Lime을 포함한 3단계 설치 과정 설명
- 설치 후 도움말의 업데이트를 통해 업그레이드 지원
- 온라인 업데이트 시에는 2, 3단계 설치 불필요 언급

### ③ 점사랑의 화면 및 메뉴 구성

- 메뉴, 도구바, 상태창, 점역줄, 줄 번호 중심으로 간략한 해당 창 특성 설명
- 새문서 종류 설명 필요 BLT, BLB 편집창 특성 및 설명

### ④ 지원되는 파일 형식 및 유형별 파일의 특징

- BLT vs BLB & BRF vs BRL 특성 중심으로 설명
- 기타 외부 지원 파일 형식 간략 설명

## (2) 5~8차시: 기본 편집 작업

### ① 새 문서 선택, 저장, 점역하기

- 새 문서 BLT 문서 선택하여 문서 작성하고 저장하기
- 작성한 문서를 대상으로 기본 상태로 점역하기
- 점역된 BLB 편집창과 BLT 원본창 간 전환하기

### ② 기본적인 편집 기능 활용

- BLT문서에서 편집 메뉴의 복사, 삭제, 붙여넣기, 찾기, 바꾸기 중심으로 설명

## ③ 쪽/장 바꾸기, 들여쓰기/내어쓰기, 정렬 관련 편집

- BLT 문서의 문서 메뉴를 중심으로 설명
- 삽입된 코드 보기 온/오프 특성 및 상태창 정보 활용법 설명

## ④ 목차 만들기

- BLT 문서에서 삽입 메뉴의 점자 차례 삽입 기능을 중심으로
- 편집창에서 해당 영역 블록 설정하여 코드 목록(F5)의 목차 코드 및 목차 생성 코드 삽입
- BLT vs BLB 편집창에서 목차 생성(toc 코드) 적용 특성 설명

## ⑤ 머리말/꼬리말, 페이지 번호 설정하기

- 문서 메뉴의 해당 기능을 중심으로 설명하고 코드 목록의 관련 코드는 병행하여 간략히 설명

## ⑥ 역점역용 점자 문서 편집하기

- BLB 문서에서 입력 모드 전환하며 역점역용 문서 편집하기(편집 메뉴의 점자 편집 중심으로)
- 스크린리더 상태창 정보 확인하며 입력

## ⑦ 점역 옵션 및 기타 환경 설정

- 기본적인 옵션 설정 중심으로 선별적 설명

## (3) 9~12차시: 고급 편집 기능

## ① 표 점역을 위한 편집 방법

- BLT 문서에서 표 만들기, 행/열 제목, 표 제목 설정
- 표 행/열 추가/삭제, 행/여열 바꾸기 편집

## ② 수식 입력 및 편집

- 온라인 수식 편집기 호출, 수식 삽입
- 한글 문서에서 수식 클립보드로 복사하여 붙여넣기 및 오프라인 수식 삽입 관련 간략 설명

## ③ 코드 리스트 활용 및 코드 보기 방식 설정

- 코드 목록(F5) vs 코드 삽입(F9) 간략 설명과 입력 예시
- 코드 보기 온/오프에 따른 BLB 편집창 렌더링 온/오프 및 코드 찾기 관련성 설명

#### ④ 스타일 활용하기

- 즐겨쓰는 코드 파라미터를 스타일로 등록하기(설정 메뉴의 스타일 설정 중심으로)
- 내가 설정한 스타일 삽입(F6)하며 문서 편집하기
- 코드 목록(F5)과 스타일 삽입(F6) 차이 이해하기

#### ⑤ 목자 및 점자 인쇄하기

- 출력 옵션 이해(전체 출력, 현재 쪽 출력, 원하는 쪽만 출력 등)

#### ⑥ 외부 편집기(HWP)에서의 편집 및 점사랑에서 불러오기

- 점사랑 점역 코드 외부 문서에서 입력하여 가져오기 [[점역코드]] 형식 중심으로

#### ⑦ 유효성검사 및 기타 점사랑 활용 방안

- 일반 외부 점역 코드(@@ff, @@FF, @@p, 미정의 점형) 내부 점역 코드로 변환하기
- 자동 교열 간략 설명

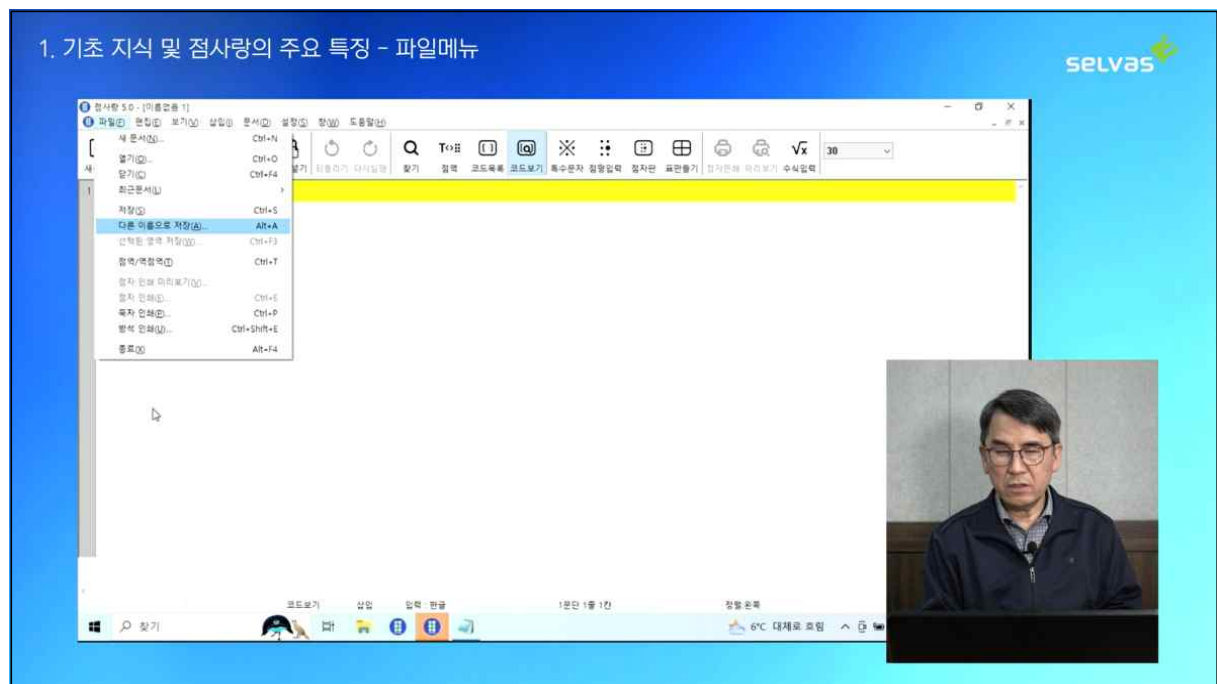


그림 10-1 동영상 콘텐츠 1차시 화면

2023년 목자-점자 병렬말뭉치 구축  
(II. 점역 소프트웨어 고도화)

# 제 11 장

## 점역 소프트웨어의 검증 및 평가

1. 점역 소프트웨어의 점역 결과 검증
2. 점역 소프트웨어의 사용자 평가
3. 검증 및 평가에 대한 시사점





## 1 점역 소프트웨어의 점역 결과 검증

점역 결과에 대한 정확도를 높이기 위해서는 다양한 사례 수집뿐만 아니라 다수 사용자에게 의한 오랜 기간의 오류 검출 과정이 필요하다. 이러한 이유로 점역 소프트웨어 개발에 있어 점역 결과의 신뢰성을 확보하는 작업은 매우 어려운 과정이다.

본 사업에서는 이러한 문제를 짧은 개발 기간 내에 해결하기 위해 2022년 점역 소프트웨어 고도화 사업에서 개발한 점역 엔진 검증 프로그램을 업데이트하고, 개정 한국 점자 규정의 예문으로 데이터베이스를 구축하여 점사랑 5.0의 점역 엔진을 검증하였다. 또한 10만 문장의 목자-점자 병렬말뭉치 데이터베이스를 사용하여 다양한 사례에 대한 검증 과정을 추가하여 점역 결과에 대한 정확도를 높이는 데 활용하였다.

### 1) 점역 엔진 검증을 위한 데이터베이스 구축

점역 엔진을 검증하기 위해 목자-점자 병렬말뭉치 데이터베이스 외에 별도의 점자 규정 예문을 데이터베이스로 구축하였다.

점역 엔진 검증을 위한 데이터베이스에서 각 레코드를 구성하는 필드는 <표 11-1>과 같다.

표 11-1 점역 엔진 검증을 위한 데이터베이스 구성

	필드명	자료형	설명
1	번호(no)	정수형	레코드의 일련번호
2	고시년도(ryear)	정수형	2020, 2023 등의 고시년도
3	점역모드(mode)	정수형	한글, 영어, 수학(LaTeX)
4	참조자료(refer)	정수형	규정
5	조항(prov)	문자열	규정의 해당 조항 번호
6	목자(prntext)	문자열	점역을 위한 목자 텍스트
7	점자(brltext)	문자열	목자 텍스트를 점역한 유니코드 점형
8	비고(other)	문자열	기타 참고할 내용

### 2) 데이터베이스 자료 입력 규칙

데이터베이스의 자료 입력은 점자 규정의 모든 내용을 입력하는 것을 원칙으로 하되, 데이터베이스 내 입력이 불가능한 서식이 포함된 예문은 제외하였다.

### (1) 공통 사항

- ① 한글 점자 규정 및 수학 점자 규정 각 항목별 예문을 입력한다.
- ② 하나의 단어나 복합어 같은 짧은 예문의 경우 여러 개를 하나의 셀에 입력하되, 단어 사이에 두 칸을 띄운다.
- ③ 같은 항목의 짧은 단어라도 예외나 붙임 등 따로 구분해야 하는 예문의 경우 별도로 입력한다.
- ④ 줄이 바뀔 때 나타나는 점역 규정의 특징을 보여주는 항목은 줄이 변경되는 해당 글자 앞에서 Enter를 입력하고, 검증 프로그램에서는 이를 빈칸으로 인식하여 처리한다.
- ⑤ 텍스트로 입력할 수 없는 수식은 한컴 스크립트를 적용하여 입력하고, 점사랑에서 LaTeX로 변환하여 입력한다.
- ⑥ 기타 문맥 및 점역사의 선택에 따라 점역 결과가 달라질 수 있는 예문은 생략한다.

### (2) 한글 점자 규정

- ① 옛 글자와 함께 나오는 방점(거성, 상성) 항목은 사용되는 유니코드가 명확하지 않기 때문에 생략한다.
- ② 49항 [붙임] 항목처럼 점역을 할 때만 사용되는 대체 기호나 설명의 경우 생략한다.
- ③ 66항 ‘점역자 주’처럼 점역 엔진의 테스트와 무관한 내용의 규정은 생략한다.
- ④ 19항 [붙임]의 동국정운식 한자음 표기에 사용되는 음가 없는 받침 ‘ㅇ’은 목자의 모양만으로는 옛 글자 여부가 명확하지 않기 때문에 생략한다.

### (3) 수학 점자 규정

- ① 수식의 목자는 한컴프로그램 수식 입력기를 통한 한컴 스크립트 문법으로 입력한다.
- ② 한컴 스크립트 문법으로 입력한 수식을 클립보드로 복사하고, 점사랑 5.0에서 클립보드 한글 수식 가져오기를 통해 LaTeX 수식으로 변환하여 데이터베이스에 입력한다.
- ③ 수식 입력기에서 지원하지 않거나 다른 문자와 혼동이 되는 수학 기호는 입력을 생략한다.

### 3) 점역 엔진 검증 프로그램 업데이트

점역 엔진의 정확도를 검증하기 위해 2022년 점역 소프트웨어 고도화 사업에서 개발한 PC용 점역 엔진 검증 프로그램을 업데이트하고, 개정 한국 점자 규정(2024)의 내용으로 구축된 데이터베이스와 묵자-점자 병렬말뭉치 데이터베이스를 사용하였다. 본 프로그램은 데이터베이스와 연동하여 점자 규정 준수 여부를 자동으로 검증하고 오류 사항을 출력하는 프로그램이다.

PC용 점역 엔진 검증 프로그램의 주요 기능은 다음과 같다.

- 검증할 점역 대상을 ‘한글 모드’, ‘영어 모드’, ‘수학 모드’ 등으로 구분하여 검증을 실행한다.
- PC에 저장된 엑셀 파일을 호출하여 점역 결과에 대한 검증을 실행한다.
- 엑셀 파일에 저장된 예문을 자체 점역 엔진으로 점역하고, 엑셀 파일 내 점형과 비교하여 오류를 검출한다.
- 점역 결과를 비교할 때 문장 앞뒤의 공백 삭제를 선택할 수 있는 기능을 제공한다.
- 검색 결과의 전체 또는 오류에 해당하는 레코드만을 화면에 출력하고 파일로 저장할 수 있는 기능을 제공한다.

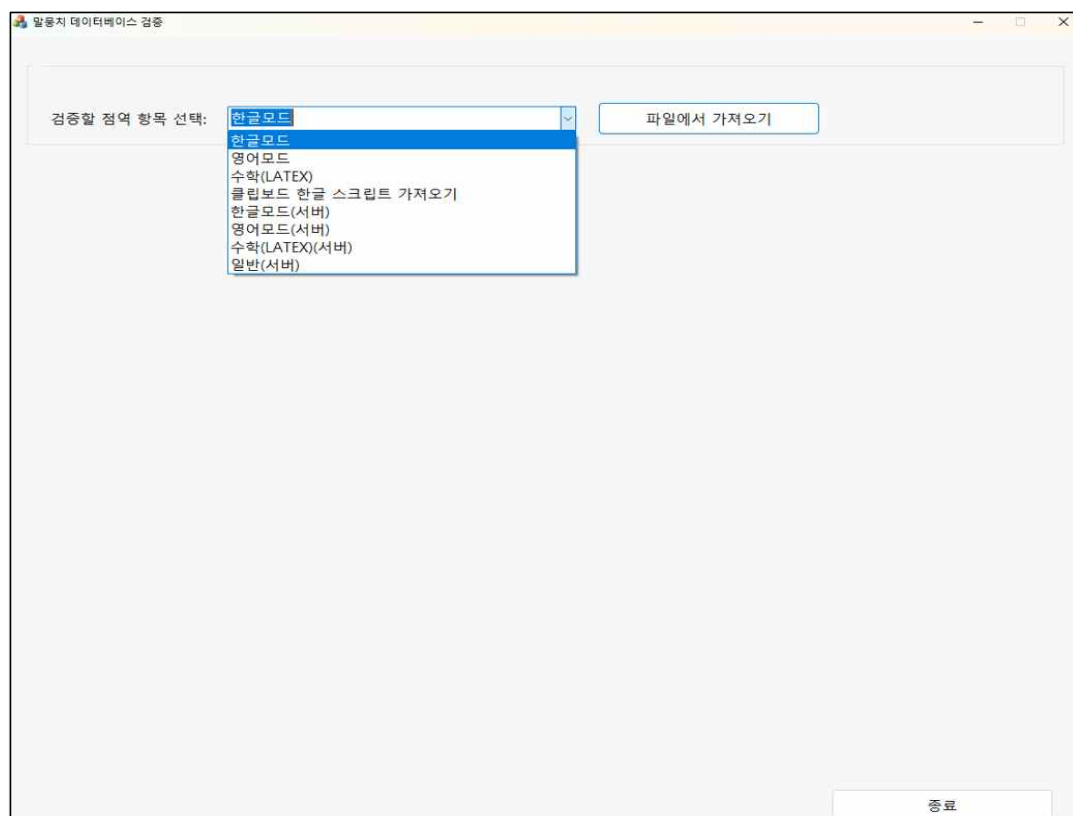


그림 11-1 점역 엔진 검증 프로그램의 화면 구성

#### 4) 점역 엔진의 검증 결과

업데이트한 PC용 점역 엔진 검증 프로그램과 구축된 개정 한국 점자 규정(2024), 묵자-점자 병렬 말뭉치 데이터베이스를 이용하여 점역 엔진의 정확도를 검증하고 이를 점역 소프트웨어 개발에 반영하였다. 본 검증 프로그램은 2024년 1, 2월에 집중적으로 사용되었으며, 점역 결과의 다양한 오류를 검출하여 점역 엔진을 개선하는 데 활용하였다.

다음에서 제시하고 있는 검증 결과는 2024년 2월 26일 업데이트 버전으로 검출한 오류 내용이며, 여기서 발생한 오류는 이후 하자보수 기간에 수정, 보완할 예정이다.

##### (1) 한글 모드의 점역 검증 결과

검증 프로그램을 이용한 한글 모드의 점역 결과는 검증 데이터 363개 중 적합 338개, 오류 25개, 적합률 93%였다. 한글 모드 점역에서 검출된 주된 오류의 유형은 <표 11-2>와 같다.

말뭉치 데이터베이스 검증

검증할 점역 항목 선택: 한글모드 파일에서 가져오기

검증 결과: 한글모드, 검증 데이터 363개 중, 정답 338, 오답 25, 정답률 93%입니다.

번호	검증 데이터	결과	점사랑 점역 결과	검증용 정답
1	거리 너비 두더지	True	거리 너비 두더지	거리 너비 두더지
2	리코더 미소 보리	True	리코더 미소 보리	리코더 미소 보리
3	셔츠 저고리 추수	True	셔츠 저고리 추수	셔츠 저고리 추수
4	커피 터무니 피리 호수	True	커피 터무니 피리 호수	커피 터무니 피리 호수
5	아버지 야구 어머니	True	아버지 야구 어머니	아버지 야구 어머니
6	여우 오이 요리	True	여우 오이 요리	여우 오이 요리
7	우유 유리 으스스	True	우유 유리 으스스	우유 유리 으스스
8	이모 피아노 도오미 귀시아	True	이모 피아노 도오미 귀시아	이모 피아노 도오미 귀시아

☐ 틀린 것만 보기    ☐ 앞/뒤 스페이스 제거

관련 정보: 고시년도: 2023, 참고자료: 규정, 조항: 1항

원문: 거리 너비 두더지

원문 점역 답자: 거리 너비 두더지

점사랑 점역 결과: 거리 너비 두더지

검증 결과 저장하기 종료

그림 11-2 한글 모드의 점역 검증 결과 화면



검증 프로그램을 이용한 수학 모드의 점역 결과는 검증 데이터 285개 중 적합 267개, 오류 18개, 적합률 93%였다. 수학 모드 점역에서 검출된 주된 오류의 유형은 <표 11-4>와 같다.

**표 11-4** 수학 모드 점역의 오류 유형

검증용 데이터	검증용 점역 결과	점사량 점역 결과	오류 내용
$-5 < x < -2$	<div> <math>-5 &lt; x &lt; -2</math> </div>	<div> <math>-5 &lt; x &lt; -2</math> </div>	개정 전 점자 규정 적용 오류
$(\text{원의 둘레}) = (\text{반지름}) \times 2 \times 3.14$	<div> <math>(\text{원의 둘레}) = (\text{반지름}) \times 2 \times 3.14</math> </div>	<div> <math>(\text{원의 둘레}) = (\text{반지름}) \times 2 \times 3.14</math> </div>	수학 괄호 미사용 오류
$0.7 \dot{\phantom{x}} 3 \dot{\phantom{x}} 9$	<div> <math>0.7 \dot{\phantom{x}} 3 \dot{\phantom{x}} 9</math> </div>	<div> <math>0.7 \dot{\phantom{x}} 3 \dot{\phantom{x}} 9</math> </div>	순환마디 표시 점형 미적용 오류
$x \not < y$	<div> <math>x \not &lt; y</math> </div>	<div> <math>x \not &lt; y</math> </div>	불필요한 빈칸 삽입 오류
$\sqrt{x^2}$ 은 $ x $ 이다.	<div> <math>\sqrt{x^2}</math>은 <math> x </math>이다. </div>	<div> <math>\sqrt{x^2}</math>은 <math> x </math>이다. </div>	수식 앞뒤 2칸 띄어쓰기 미적용 오류
$\sqrt[n]{a}$	<div> <math>\sqrt[n]{a}</math> </div>	<div> <math>\sqrt[n]{a}</math> </div>	묵음 괄호 미사용 오류



**그림 11-3** 수학 모드의 점역 검증 결과 화면

#### 4) 검증 결과에 대한 반영

점역 엔진 검증 프로그램에서 검출된 오류의 유형으로는 점형 오류, 띄어쓰기 오류, 규칙 미준수 등 점자 규정을 준수하지 않은 경우가 대부분이었다. 점자 규정에서 명확하게 제시되지 않은 유니코드 문자 및 일부 규칙은 해설서의 예시를 기준으로 하였다. 검출된 오류를 분석하여 점역 모드에 따라 점사랑 5.0에 반영하였다.

검증 결과 나타난 오류들은 계속해서 수정 및 업데이트 진행 중이며, 출시 이후 발견되는 오류 수정 사항은 사업종료 후 1년간의 하자보수 기간 동안 점사랑 업데이트 메뉴를 통해 온라인으로 제공한다. 또한, 국립국어원 점자 종합정보 누리집에 ‘점사랑 문의하기’ 게시판을 운영하고, 사용자의 의견을 수렴해 점역 오류 및 프로그램 버그 등도 꾸준히 보완해 나갈 예정이다.

번호	제목	첨부	답변상태	등록일	조회
22	<a href="#">blb 상태에서 점역 시 페이지 번호 오류 및 개선 요청</a>		답변	2024-02-27	6
21	<a href="#">blb 파일에서 목차 편집 시 오류</a>		답변	2024-02-15	13
20	<a href="#">점사랑4.0사용법</a>		답변	2024-02-06	25
19	<a href="#">점역기능 문의 (목차 안내선 코드)</a>		답변	2024-01-22	22
18	<a href="#">점역 관련 문의</a>		답변	2024-01-10	21
17	<a href="#">옛글자 점역</a>		답변	2023-12-04	50
16	<a href="#">점사랑 프로그램 변경 요청드립니다.</a>		답변	2023-11-27	76
15	 <a href="#">키보드로 점자 복사</a>		답변	2023-11-27	11
14	<a href="#">blb 파일과 brf 파일에서 페이지 설정이 다르게 작동합니다.</a>		답변	2023-11-26	21
13	 <a href="#">점사랑4.0 페이지 번호 관련 문의</a>		답변	2023-11-01	1

그림 11-4 점자 종합정보 누리집의 점사랑 문의 게시판 화면

## 2 점역 소프트웨어의 사용자 평가

2024년 2월 8일부터 15일까지 8일간 점사랑 5.0의 사용자 평가를 실시하였다. 이번 평가에 사용한 버전은 5.0.0.30 베타 버전이며, 평가 결과 부적합으로 제시된 사항들과 전반적인 평가 등은 2월 27일부터 3월 15일까지 약 2주간에 걸쳐 프로그램에 반영하여 업데이트할 예정이다.

### 1) 사용자 평가단의 구성

사용자 평가단은 1급 점역·교정사 11명(시각장애인 3명, 비장애인 8명)으로 구성하였다. 시각장애인 점역·교정사는 모두 스크린리더를 사용하여 점역 소프트웨어에 접근할 수 있는 사용자로서 접근성 영역을 주로 평가하였고, 비장애인 점역·교정사는 점역 소프트웨어의 메뉴에 따라 영역을 구분하여 평가하였다. 자세한 사용자 평가단의 구성 내용은 <표 11-5>와 같다.

표 11-5 사용자 평가단의 구성

번호	자격급수	취득과목	장애등급	평가 영역
1	1급	한/영/수/음/일		공통, 점사랑
2	1급	한/영/수/음	시각1급	공통, 접근성
3	1급	한/영/수/음		공통, 점사랑, 점역 코드, 그래픽
4	1급	한/영/수/음	시각1급	공통, 접근성
5	1급	한/영/수/일		공통, 음악
6	1급	한/영/수		공통, 점사랑, 점역 코드, 그래픽
7	1급	한/영/수/일		공통, 점사랑
8	1급	한/영/수/음		공통, 음악
9	1급	한/영/수/음		공통, 점역 코드
10	1급	한/영/수	시각1급	공통, 접근성, 점사랑, 점역 코드, 그래픽
11	1급	한/영/수		공통, 점사랑, 점역 코드, 그래픽

### 2) 사용자 평가 영역의 구성

사용자 평가는 크게 공통 영역, 접근성 영역, 점사랑 메뉴 영역, 점역 코드 목록, 촉각 그래픽 메뉴 영역으로 구분하였다. 또한 점역의 모든 과정을 순차적으로 진행하면서 제작 과정에서 발생하는 문제점 및 개선점을 자유롭게 작성할 수 있도록 평가 과정에서 활용할 테스트 체크리스트를 평가단 전원에게 배포하여 작성하게 하였다.

공통 영역에서는 점사랑 5.0 소프트웨어에 대한 포괄적인 질문을 제시함으로써 소프트웨어 전반에 대한 긍정, 부정 평가가 이루어지도록 했으며, 문항 수는 21개로 구성하였다. 이 영역은 문항의 특성을 고려하여 다른 영역과는 달리 평가단 전원을 대상으로 실시하였다.

접근성 영역에서는 스크린리더 사용자의 접근성을 확인할 수 있도록 하였으며 문항 수는 24개로 구성하였다. 이 영역은 문항의 특성을 고려하여 시각장애인 점역·교정사 3명만을 대상으로 실시하였다.

점사랑 5.0 메뉴 영역에서는 파일, 편집, 보기, 삽입, 문서, 설정 등의 영역으로 구분하여 문항을 제시하였으며 점사랑 5.0에서 새로이 추가된 악보 점역 기능 관련 문항은 삽입 영역에 포함하여 제시하였다. 총 문항 수는 210개로 구성하였고, 각 메뉴별 기능에 따라 질문을 구체화함으로써 평가 결과를 프로그램 개발에 쉽게 반영할 수 있도록 하였다. 이 영역은 비장애인 점역·교정사 5명과 시각장애인 점역·교정사 1명이 참여하였다.

점역 코드 영역에서는 점사랑에서 사용되는 점역 코드가 바르게 동작하는지 확인하기 위해 총 41종의 점역 코드와 각 코드에 대한 설명을 제시하였다. 이 영역은 비장애인 점역·교정사 4명과 시각장애인 점역·교정사 1명이 참여하였다.

축각 그래픽 메뉴 영역에서는 새로이 추가된 축각 그래픽 프로그램의 파일, 편집, 텍스트, 삽입, 보기, 설정 등의 메뉴 영역으로 구분하여 문항을 제시하였으며 총 문항 수는 90개로 구성하였고, 각 메뉴별 기능에 따라 질문을 구체화함으로써 평가 결과를 프로그램 개발에 쉽게 반영할 수 있도록 하였다. 이 영역은 비장애인 점역·교정사 3명이 참여하였다.

사용자 평가 영역 구성에 대한 자세한 내용은 <표 11-6>과 같다.

표 11-6 사용자 평가 영역의 구성

영역	문항 수	참여자
공통	21	11명
접근성	24	3명
점사랑	210	6명
점역 코드	41	5명
그래픽	90	4명

### 3) 사용자 평가 결과

#### (1) 공통 영역의 평가 결과

공통 영역에서는 점사랑 5.0에 대한 세부적인 기능 평가보다는 소프트웨어 전반에 대해 평가하되, ‘적합’, ‘일부 부적합’, ‘부적합’으로 구분하여 평가하도록 하였다. 공통 영역 평가는 평가단 11명이 참여하였다.

평가 결과를 살펴보면, 프로그램의 안정성, 목·점자 편집 기능, 화면 인터페이스, 키보드 및 마우스를 활용한 편집 기능, 단축키 및 기타 설정 등 기존의 점역 소프트웨어가 제공했던 기본 기능은 대체로 양호하게 나타났으나, 표나 수식 등에 대한 평가에서는 보완할 부분이 많은 것으로 나타났다. 또한, 점역 결과에 대한 문항에서 상대적으로 일부 부적합 평가가 높게 나타난 것은 개정 한국 점자 규정을 반영하면서 한글과 영어가 혼합된 문장에서 나타나는 기호 처리 문제 및 수식에서 칸 띄우기 오류 및 줄바꿈 위치에서의 처리 오류가 발생한 것으로 개선이 필요하다. 또한 점자 인쇄에 대한 문항에서는 무응답의 비율이 높게 나타났는데, 이는 검증 당시 축각 그래픽 인쇄 기능이 활성화되어 있지 않은 상태였고, 평가자에 따라 인쇄 환경이 적절하게 구성되지 못했기 때문으로 보인다.

공통 영역에 대한 평가 결과는 <표 11-7>과 같다.

표 11-7 공통 영역의 평가 결과

번호	항목	평가 결과			
		적합	일부 부적합	부적합	무응답
1	프로그램이 안정적으로 설치, 실행, 동작, 종료되는가?	10	1		
2	파일의 생성, 저장, 불러오기가 안정적으로 동작하는가?	11			
3	편집창에서 키보드 입력, 단축키를 이용한 메뉴 호출 등이 원활한가?	11			
4	편집창에서 키보드를 이용한 글자, 단어, 줄, 문단, 쪽 단위 등 커서 이동이 원활한가?	10	1		
5	편집창에서 마우스를 이용한 문서 편집이 원활한가?	8			3
6	표 편집기는 안정적으로 동작하고, 사용하기에 편리한가?	4	5	1	1
7	점자판 기능은 안정적으로 동작하고, 사용하기에 편리한가?	7	2		2
8	수식입력기는 안정적으로 동작하고, 사용하기에 편리한가?	2	7	1	1
9	축각 그래픽 편집기는 안정적으로 동작하고, 사용하기에 편리한가?	1	1	2	7
10	악보 변환과 편집 기능은 안정적으로 동작하고, 사용하기에 편리한가?	3		2	6
11	한글, 영어는 점자 규정에 맞게 점역되는가?	3	7	1	
12	표는 점역 지침에 맞게 점역되는가?	6	3		2
13	수식은 점자 규정에 맞게 점역되는가?	3	8		
14	악보는 점자 규정에 맞게 점역되는가?	4	1		6
15	점자 인쇄물은 작성한 내용 및 서식에 따라 출력되는가?	2	1		8
16	축각 그래픽 인쇄물은 그래픽 편집기로 작성한 내용과 일치하는가?				11
17	목자 인쇄물의 목자 폰트와 점자 폰트는 바르게 출력되는가?	3			8
18	점역 코드는 정의된 대로 동작하는가?	6	5		
19	메뉴, 화면 구성 등 사용자 인터페이스는 편리한가?	10	1		
20	도구 모음 항목 구성은 적절한가?	11			1
21	설정 메뉴 항목들의 설정값은 바르게 적용되는가?	6	4	1	

## (2) 접근성 영역의 평가 결과

접근성 영역에서는 스크린리더를 통해 점사랑 5.0을 사용하는 시각장애인 사용자의 접근성 및 편의성에 대해 평가하되, ‘적합’, ‘일부 부적합’, ‘부적합’으로 구분하여 평가하도록 하였다. 접근성 영역의 평가는 시각장애인 점역·교정사 3명이 참여하였다. 평가 결과를 살펴보면, 대체로 접근성에 대한 평가가 양호했으나, 편집기 내에서 블록 선택 기능에 대한 음성 지원, 점자 정보 단말기에 대한 지원 등에서 부적합 평가가 제시되었다. 다만 이러한 접근성의 문제점은 점사랑 5.0 소프트웨어 자체적으로 해결할 수 없는 기술적인 부분들이 있어 스크린리더 제조업체의 지원을 통한 개선이 필요하다.

접근성 영역에 대한 평가 결과는 <표 11-8>과 같다.

표 11-8 접근성 영역의 평가 결과

번호	항목	평가 결과			
		적합	일부 부적합	부적합	무응답
1	모든 메뉴(하위 메뉴 포함)에 키보드 접근이 가능한가?	2	1		
2	토글 메뉴의 선택/해제 상태를 음성으로 확인할 수 있는가?	3			
3	대화 상자의 모든 구성 요소에 키보드 접근이 가능한가?	3			
4	대화 상자의 구성 요소의 컨트롤 종류를 음성 출력하는가?	3			
5	대화 상자의 구성 요소의 이름(용도)을 음성 출력하는가?	3			
6	대화 상자에서 체크박스 등과 같은 선택 가능한 컨트롤에 대해 선택 상태를 음성 출력하는가?	3			
7	목자 편집창에서 모든 내용에 키보드와 음성으로 접근할 수 있는가?	2	1		
8	목자 편집창에서 코드 보기를 토글로 동작시켜 확인할 수 있는가?	2			1
9	목자 편집창에서 블록으로 선택한 내용을 확인할 수 있는가?	1		1	1
10	목자 편집창에서 점형 입력 시 음성 출력이 가능한가?	2		1	
11	목자 편집창에서 표 작성 및 셀 이동 등이 가능한가?	2			1
12	목자 편집창에서 점자판 작성 및 셀 이동 등이 가능한가?	1			2
13	점자 편집창에서 모든 내용에 키보드와 음성으로 접근할 수 있는가?	2	1		
14	점자 편집창에서 코드 보기를 토글로 동작시켜 확인할 수 있는가?	2			1
15	점자 편집창에서 블록으로 선택한 내용을 확인할 수 있는가?	1		1	1
16	점자 편집창에서 점형 입력 시 음성 출력이 가능한가?	2		1	
17	점자 정보 단말기로 편집창, 메뉴, 대화 상자 컨트롤에 접근할 수 있는가? ('설정 메뉴'의 '점자 정보 단말기 설정' 참조)		1	2	
18	찾기/바꾸기, 이동 기능을 음성 출력으로 실행할 수 있는가?	1		2	
19	코드 찾기 기능을 음성 출력으로 실행할 수 있는가?	2		1	
20	상태 표시줄에 접근이 가능한가?	2		1	
21	줄 번호 선택 시 음성 출력하는가?	2		1	
22	수식 입력기에 접근이 가능한가?	2		1	
23	수식입력기로 목자 편집창에 입력된 내용에 접근이 가능한가?	2		1	
24	생성된 점자 차례에 접근이 가능한가?	2			1

### (3) 점사랑 메뉴 영역의 평가 결과

점사랑 메뉴 영역에서는 파일, 편집, 보기, 삽입, 문서, 설정 등의 각 메뉴를 중심으로 세부 기능을 평가하되, ‘적합’, ‘일부 부적합’, ‘부적합’으로 구분하여 평가하도록 하였다. 점사랑 메뉴 영역의 평가는 비장애인 점역·교정사 5명과 시각장애인 점역·교정사 1명이 참여하였다.

점사랑 메뉴 영역의 문항은 각 메뉴별 기능에 따라 질문을 구체화함으로써 평가 결과를 프로그램 개발에 쉽게 반영할 수 있도록 하였다. 여기에서는 각 메뉴에 따른 평가 결과 중 각 문항의 세부 내용은 생략하고, ‘적합’, ‘일부 부적합’, ‘부적합’의 비율과 부적합으로 평가된 주요 내용만을 제시하였다.

#### ① 파일 메뉴의 평가 결과

파일 메뉴의 평가 문항은 총 35개이며, 부적합 결과로 제시된 것으로는 목자 편집창에서 점자인쇄 문제, 목자 인쇄 대화 상자에서 일부 옵션이 동작하지 않는 오류 등이 있었다.

파일 메뉴의 평가 결과는 <표 11-9>와 같다.

표 11-9 파일 메뉴의 평가 결과

번호	항목	문항 수	평가 결과			
			적합	일부 부적합	부적합	무응답
1	새 문서	2	12			
2	열기	4	23	1		
3	닫기	2	12			
4	최근 문서	2	9	2	1	
5	저장	5	30			
6	다른 이름으로 저장	2	12			
7	선택된 영역 저장	2	10		2	
8	점역/역점역	2	12			
9	점자 인쇄 미리보기	2	12			
10	점자 인쇄	4	9		5	10
11	목자 인쇄	4	16		1	7
12	방식 인쇄	1	2		1	3
13	종료	3	18			
합계		35	177	3	10	20

### ② 편집 메뉴의 평가 결과

편집 메뉴의 평가 문항은 총 43개이며, 부적합 결과로 제시된 것으로는 바꾸기에서 엔터 및 탭 문자 찾기 오류, 바꿀 문자열이 하이라이트 되지 않는 문제, 문서 끝이나 문서 처음에서 계속 찾기가 되지 않는 문제 등이 있었다.

편집 메뉴의 평가 결과는 <표 11-10>과 같다.

**표 11-10** 편집 메뉴의 평가 결과

번호	항목	문항 수	평가 결과			
			적합	일부 부적합	부적합	무응답
1	공통	1	6			
2	되돌리기/다시 실행	1	6			
3	잘라내기/복사/붙여넣기/삭제/전체 선택	1	6			
4	찾기	9	50	3		1
5	바꾸기	9	45	4	4	1
6	이동	5	24		5	1
7	코드 찾기	5	28		1	1
8	강제 줄바꿈	2	8	3	1	
9	점형 입력	3	15	1	1	1
10	점자 편집	2	11	1		
11	문서 최적화	5	26	3	1	
합계		43	225	15	13	5

### ③ 보기 메뉴의 평가 결과

보기 메뉴의 평가 문항은 총 15개이며, 보기 메뉴의 평가는 대체로 적합하다는 의견이 많았다. 부적합 결과로 평가된 것으로는 상태 표시줄이 해제되지 않는 문제, 점역 코드에 커서가 위치하면 줄 전체가 하이라이트 되는 문제가 있었다.

보기 메뉴의 평가 결과는 <표 11-11>과 같다.

표 11-11 보기 메뉴의 평가 결과

번호	항목	문항 수	평가 결과			
			적합	일부 부적합	부적합	무응답
1	조판 부호	3	15			3
2	도구 모음	1	5			1
3	도구 모음 크기	1	5			1
4	유니코드 점형	1	5			1
5	줄 번역	2	9	1		2
6	상태표시줄	2	11		1	
7	줄 번호	2	12			
8	점역 코드 보기	3	15		2	1
합계		15	77	1	3	9

## ④ 삽입 메뉴의 평가 결과

삽입 메뉴의 평가 문항은 총 40개이며, 부적합 결과로 제시된 것으로는 점사랑 수식 편집기에서 수식 복사, 편집 시 발생하는 오류, 악보 입력에서 점역 옵션을 설정하지 못하는 문제, 병합된 셀이 있는, hwp로 작성된 표를 불러오지 못하는 문제 등이 있었다.

삽입 메뉴의 평가 결과는 <표 11-12>와 같다.

표 11-12 삽입 메뉴의 평가 결과

번호	항목	문항 수	평가 결과			
			적합	일부 부적합	부적합	무응답
1	라인 입력	1	5			1
2	테두리 입력	1	5			1
3	점자 차례	2	5	1		6
4	표	9	47	5	2	
5	점자판	8	38	3		7
6	점자 그림	1	2	2		2
7	수식 입력	5	17	4	9	
8	악보 입력	4	10			14
9	코드 목록	5	27	1		2
10	코드 삽입	2	11		1	
11	특수문자	1	6			
12	한자	1	6			
합계		40	179	16	12	33

### ⑤ 문서 메뉴의 평가 결과

문서 메뉴의 평가 문항은 총 21개이며, 부적합 결과로 제시된 것으로는 쪽 나누기에서 페이지 첫 줄에 빈 줄이 삽입되는 문제, 워드랩에서 단어 단위 줄바꿈이 되지 않는 문제 등이 있었다.

문서 메뉴의 평가 결과는 <표 11-13>과 같다.

**표 11-13** 문서 메뉴의 평가 결과

번호	항목	문항 수	평가 결과			
			적합	일부 부적합	부적합	무응답
1	강조 표시	2	9	1		2
2	첨자	1	4	1		1
3	정렬	2	10			2
4	단락 속성	2	10			2
5	쪽 나누기	3	16		2	
6	페이지 번호	3	17	1		
7	점역자 주	1	6			
8	꼬리말/머리말	3	18			
9	워드랩	1	6			
10	점역 모드	1	6			
11	스타일	2	7	1		4
합계		21	109	4	2	11

### ⑥ 설정 메뉴의 평가 결과

설정 메뉴의 평가 문항은 총 49개이며, 부적합 결과로 제시된 것으로는, 기본환경 설정과 점역옵션 설정에서 일부 설정값이 저장되지 않는 문제, 사용자 정의 점역에서 정의된 한글 및 영어 문자가 바뀌어 점역되는 문제가 있었다.

설정 메뉴의 평가 결과는 <표 11-14>와 같다.

표 11-14 설정 메뉴의 평가 결과

번호	항목	문항 수	평가 결과			
			적합	일부 부적합	부적합	무응답
1	기본 환경 설정	5	26		2	2
2	점역 옵션 설정	14	68	4	6	6
3	점자 프린터 설정	2	5			7
4	문단 재구성 설정	3	14	1		3
5	점자 정보 단말기 설정	2	2			10
6	6점키 설정	3	18			
7	스타일 설정	3	15			3
8	템플릿 관리	2	10			2
9	편집창 설정	2	12			
10	단축어 설정	6	35			1
11	사용자 정의 점역	7	39	3		
합계		49	244	8	8	34

#### (4) 점역 코드 영역의 평가 결과

점역 코드 영역에서는 점사랑에서 사용되는 점역 코드가 바르게 동작하는지 확인하기 위해 총 41종의 점역 코드와 각 코드에 대한 설명을 제시하고 기능을 평가하되, ‘적합’, ‘일부 부적합’, ‘부적합’으로 구분하여 평가하도록 하였다. 이 영역의 평가는 비장애인 점역·교정사 4명과 시각장애인 점역·교정사 1명이 참여하였다.

총 41개의 점역 코드에 대한 기능 평가에서 부적합 결과로 제시된 의견으로는 프랑스어 점역 오류, 내어쓰기 설정값 적용 오류 등이 있었다.

점역 코드의 평가 결과는 <표 11-15>와 같다.

표 11-15 점역 코드의 평가 결과

번호	항목	평가 결과			
		적합	일부 부적합	부적합	무응답
1	[a] ... [/a] (점자 아스키 코드)	3	0	0	2
2	[ak:N] ... [/ak] (옛 한글 모드 지정)	2	0	1	2
3	[al:T,N] (문단의 정렬 지정)	5	0	0	0
4	[b] ... [/b] (숫자 점형 입력)	5	0	0	0
5	[bd] ... [/bd] (굵게)	5	0	0	0
6	[bl:N] (점자 라인 지정)	5	0	0	0
7	[bp:N,T] (점자 페이지 번호 지정)	4	1	0	0
8	[bx:N] ... [/bx] (점자 테두리 지정)	5	0	0	0

번호	항목	평가 결과			
		적합	일부 부적합	부적합	무응답
9	[cmp:N] (공백 삭제)	5	0	0	0
10	[cn:N] ... [/cn] (목차 코드 지정)	3	0	2	0
11	[cp:N,L] (조건부 페이지 넘김)	2	1	1	1
12	[dw] (영어 약자 사용 중지)	4	1	0	0
13	[eg:N] (영어 1급, 2급 지정)	4	0	0	1
14	[em] (영어 모드 지정)	5	0	0	0
15	[fl:L, G] ... [/fl] (외국어 점형 지정)	1	0	2	2
16	[ft] ... [/ft] (꼬리말 지정)	5	0	0	0
17	[hd] ... [/hd] (머리말 지정)	5	0	0	0
18	[hj:N] (한자 점형 지정)	3	1	0	1
19	[in:N] (들여쓰기 및 내어쓰기)	2	2	1	0
20	[it] ... [/it] (기울임)	4	0	1	0
21	[kg:N] (한글의 정자, 약자 지정)	5	0	0	0
22	[km] (한글 모드 지정)	4	1	0	0
23	[l] (강제 줄바꿈)	5	0	0	0
24	[ln:N] (줄번호 표시)	4	1	0	0
25	[mg:L,R] (문단 여백 지정)	5	0	0	0
26	[mm] ... [/mm] (수식 모드 지정)	4	1	0	0
27	[p] (새 문단 시작)	5	0	0	0
28	[pg] (강제 쪽 바꿈)	5	0	0	0
29	[pp] (강제 장 바꿈)	5	0	0	0
30	[pr] ... [/pr] (발음 기호 지정)	2	0	0	3
31	[rp:N,T] (목차 참조 페이지 지정)	4	0	0	1
32	[rp] (목차 참조 페이지 + 1 자동 증가)	4	0	0	1
33	[sc] ... [/sc] (수식 스크립트 지정)	4	0	0	1
34	[sfh:A,B,C,D] (꼬리말 / 머리말 위치 지정)	5	0	0	0
35	[spl:A,B,C,D] (페이지 번호 위치 지정)	5	0	0	0
36	[sub] ... [sub] (아래 첨자)	5	0	0	0
37	[sup] ... [sup] (위 첨자)	4	0	1	0
38	[tn] ... [/tn] (점역자 주)	5	0	0	0
39	[toc:R,B,N] (목차 페이지 생성)	3	0	1	1
40	[ul] ... [/ul] (밑줄)	4	0	1	0
41	[wr:T,N] (한글 줄바꿈 방식 지정)	3	1	1	0
합계		167	10	12	16

### (5) 촉각 그래픽 메뉴 영역의 평가 결과

촉각 그래픽 메뉴 영역에서는 파일, 편집, 텍스트, 삽입, 보기, 설정 등의 각 메뉴를 중심으로 세부 기능을 평가하되, ‘적합’, ‘일부 부적합’, ‘부적합’으로 구분하여 평가하도록 하였다. 촉각 그래픽 메뉴 영역의 평가는 비장애인 점역·교정사 3명이 참여하였다.

촉각 그래픽 메뉴 영역의 문항은 각 메뉴별 기능에 따라 질문을 구체화함으로써 평가 결과를 프로그램 개발에 쉽게 반영할 수 있도록 하였다. 여기에서는 각 메뉴에 따른 평가 결과 중 각 문항의 세부 내용은 생략하고, ‘적합’, ‘일부 부적합’, ‘부적합’의 비율과 부적합으로 평가된 주요 내용만을 제시하였다.

촉각 그래픽 메뉴의 평가 문항은 총 87개이며, 부적합 결과로 제시된 것으로는 촉각 그래픽 점자인쇄 메뉴가 활성화되지 않는 문제, 바꾸기에서 점형이 입력되지 않는 문제, 삽입한 그림이 깨져 나타나는 문제 등이 있었다.

촉각 그래픽 메뉴의 평가 결과는 <표 11-16>과 같다.

**표 11-16** 촉각 그래픽 메뉴의 평가 결과

번호	항목	문항 수	평가 결과			
			적합	일부 부적합	부적합	무응답
1	파일 메뉴	24	58	1	4	9
2	편집 메뉴	17	46	2	3	
3	텍스트 메뉴	13	38	1		
4	삽입 메뉴	16	46	2		
5	보기 메뉴	10	29	1		
6	설정 메뉴	7	18	3		
합계		87	235	10	7	9

### 3 검증 및 평가에 대한 시사점

점역 소프트웨어 점사량 5.0의 점역 결과 검증 및 사용자 평가를 통해 나타난 시사점은 다음과 같다.

첫째, 점역 엔진의 정확도를 높이는 데 있어 데이터베이스를 구축하고 점역 엔진 검증을 위한 전용 프로그램을 개발하여 활용하는 것은 매우 효과적이다. 기존에는 점역상의 오류를 찾아내기 위해 다수의 사용자가 장기간 테스트하는 과정이 필요했다. 그러나 본 사업에서 구축한 다양한 예문을 포함한 데이터베이스와 PC용 검증 프로그램을 연동하여 점역의 오류를 검출한 결과, 점역 엔진의 오류를 신속하게 파악하고 점사량 5.0에 개정 한국 점자 규정을 반영할 수 있었다.

둘째, 변경된 점자 규정을 신속하게 반영하고 점역 결과의 정확도를 개선하여 점역 엔진의 성능을 향상시키기 위해서는 지속적인 데이터베이스 업데이트가 필요하다. 향후 개정 한국 점자 규정 해설서가 출간되는 경우 그 예문을 추가해야 하고, 또 규정과 해설서의 예문만으로는 점역 결과의 정확도를 완전히 담보할 수 없으므로 점역의 특성을 고려한 다양한 예문을 데이터베이스에 축적하여 점역 엔진을 검증하는 데 지속적으로 활용할 계획이다.

셋째, 점역 소프트웨어의 완성도를 높이기 위해서는 새로 도입된 기능에 대하여 충분한 검증 과정이 필요하다. 사용자 평가 결과 부적합으로 판정된 항목 중 다수가 기능이 향상된 표와 수식 입력기 및 새롭게 도입된 점자 악보와 축각 그래픽 등에서 지적되었다. 이는 새롭게 도입되는 기능으로, 사용자에게 따른 다양한 요구사항, 합의된 인터페이스 도출의 어려움, 짧은 테스트 과정 등이 원인으로 추정된다. 현재 향상된 표 점역 기능에서는 형태가 단순한 표를 대상으로 정렬 형태를 유지하는 표와 열 항목을 풀어 점역하는 표로 점역할 수 있으나 아직 셀이 병합된 표 등 좀 더 다양한 형태의 표를 점역 지침에 따라 자동으로 점역하지 못하고 있다. 또한 수식 입력기에서도 시각적으로 수식을 신속하게 확인할 수 있는 사용자 인터페이스 강화, 한컴 스크립트 호환성 증대, 수학 점자 규정의 충분한 반영 등이 이루어져야 한다.

새롭게 도입된 점자 악보 점역 기능에서는 다양한 점역 옵션을 적용하여 점자 악보를 점역할 수 있도록 기능을 보완하고, 축각 그래픽 프로그램에서는 점자 프린터 종류 및 용지 규격에 따라 그림판이 설정되어 보다 정교한 축각 그래픽을 제작할 수 있는 기능을 추가하여야 한다.

넷째, 시각장애인 사용자를 위한 접근성 개선은 스크린리더 제조업체와의 협력이 필요하다. 시각장애인 사용자는 주로 스크린리더를 통해 점역 소프트웨어를 사용하게 되는데, 이때 필요한 음성 지원이 원활하지 못한 경우 점역 소프트웨어를 충분히 활용할 수가 없게 된다. 다만, 이러한 음성 지원은 점역 소프트웨어의 개발만으로는 해결할 수 없으며, 스크린리더와의 정보 교환에 대한 적절한 프로토콜이 필요하다. 특히 단순한 텍스트로 처리할 수 없는 표, 점자판, 악보, 축각 그래픽 등의 기능에서는 스크린리더의 지원이 반드시 필요한 상황이다. 그러므로 스크린리더 제조업체와의 협력을 통해 시각장애인 사용자의 편의에 맞는 적합한 음성 지원 방식을 구현할 필요가 있다.

2023년 목자-점자 병렬말뭉치 구축  
(II. 점역 소프트웨어 고도화)

# 제 12 장

## 결론 및 제언

1. 결론
2. 제언





## 1 결론

국립국어원과 한국시각장애인연합회는 2021년부터 묵자-점자 병렬말뭉치 구축 사업을 이어오고 있다. 그 사업의 일환인 점역 소프트웨어 개발 및 고도화 사업은 점자 정보화 기반을 강화하고 시각장애인의 정보 접근성을 개선하는 데 그 목적을 두고 있으며, 2023년 점역 소프트웨어 고도화 사업의 목표는 다음과 같이 설정하였다.

- 점사랑 4.0 안정성 및 정확도 향상
- 개정 한국 점자 규정 적용
- 표 서식 및 수식 점역 기능 개선
- 촉각 그래픽 소프트웨어 개발
- 악보 점역 소프트웨어 개발
- 온라인 교육 콘텐츠 개발 및 보급

본 사업의 결과를 기초로 사업 목표별 결론은 다음과 같다.

첫째, 점사랑의 안정성과 점역 정확도를 향상시키기 위해 정기적인 연구, 개발진 협의회를 실시하였고, 이를 통해 소프트웨어 기능 분석 및 점사랑 4.0의 문제점을 도출하여 소프트웨어 설계에 반영하였다. 또, 사용자 평가 및 피드백을 바탕으로 점사랑의 안정성과 점역 정확도를 향상시키는 데 반영하였다.

이러한 과정을 통해 개선된 점사랑 5.0의 주요 특징은 다음과 같다.

- hwp 파일 서식 자동화: hwp 파일을 읽어 들일 때 문자 속성, 문단 속성, 표 서식, 수식 등을 점역에 적합한 서식으로 변환하는 기능을 추가하고 안정화하였다.
- 문서 최적화 기능: 문단 재구성('문단 재구성 설정'에 따라 묵자/점자 문서의 문단 서식을 재구성하는 기능), 자동교열(묵자 문서 또는 점자 문서에서 오류가 자주 발생하는 항목을 검사하는 기능), 기존 코드 변환(기존 점역 소프트웨어(점사랑 3.0, 브레일서울 등)의 점역 코드를 점사랑 5.0 점역 코드로 변환하는 기능) 등을 추가 및 안정화하였다.
- 점자판 기능 개선: 시각적 인식도를 높이기 위해 점자 인쇄물과 유사하도록 줄 사이와 칸 사이의 정렬을 변경하였고, 효율적인 입력을 위해 줄 단위 및 칸 단위 편집 및 이동 기능을 추가하였다.

- 점역 코드 안정화: 점사랑 4.0에서 속성으로만 표현되었던 강조 표시, 첨자, 문단 정렬, 들여쓰기, 문단 여백, 강제 줄바꿈 등의 점역 코드를 ‘점역 코드 보기’를 통해 화면에 표시 여부를 지정할 수 있도록 개선하였다.
- 스타일 설정: 점사랑 4.0에서는 ‘시작 코드’와 ‘종료 코드’로 쌍을 이루는 코드만 등록이 가능하다는 제한점이 있었으나, 스타일의 단일 코드를 사용할 수 있도록 기능을 개선하였다.
- 템플릿 관리: 단락 속성, 워드랩, 점역 옵션, 사용자 정의 점역을 템플릿으로 저장할 수 있도록 기능을 보완하였다.
- 프로그램 다중 실행: 점사랑 프로그램을 2개 이상 실행하여 편집할 수 있도록 기능을 개선하였다.

둘째, 최신 개정 한국 점자 규정을 점사랑 소프트웨어에 반영하였다. 다만, 당분간 상황에 따라 이전 한국 점자 규정과 개정 한국 점자 규정이 모두 필요할 수 있으므로 ‘점역 옵션 설정’을 통해 사용자가 점자 규정을 선택하여 점역할 수 있도록 기능을 제공하였다. 이로써 개정된 점자 규정을 신속하게 현장에 적용할 수 있는 기반이 마련되었으며, 이전과 달리 점자 규정이 개정되었음에도 개정된 점자 규정을 적용한 점역 소프트웨어가 수년 간 개발되지 못해 발생했던 현장의 어려움도 해소할 수 있게 되었다.

셋째, 표 서식과 수식 점역 기능의 자동화 및 보완으로 사용자의 요구를 반영하였다. 표의 경우 ‘시각장애 학생용 대체 교과서 제작 지침 및 적용 방안 연구(국립특수교육원, 2021)’에서 제시하고 있는 표 점역 지침에 따라 ‘원본 정렬 형태를 유지하는 표’, ‘열 항목을 풀어 점역하는 표(빈칸 구분)’, ‘열 항목을 풀어 점역하는 표(세로선 구분)’ 등으로 구분하였고, 제목 셀 선택 유무에 따라 점역 기능을 자동화하였다. 또한 수식 입력기의 경우, 편집 작업의 효율성을 위해 편집기에 수식으로 표시되어 있는 부분을 클릭하면 대화 상자를 거치지 않고 바로 수식 편집창으로 전환되도록 개선하였다.

넷째, 점사랑 소프트웨어와 연동되는 축각 그래픽 소프트웨어를 개발함으로써 축각 자료 제작 및 보급을 지원할 수 있도록 하였다.

개발된 축각 그래픽 소프트웨어의 주요 기능은 다음과 같다.

- 축각 그래픽 생성 및 편집 기능 구현
- 표준 그래픽 파일(\*.jpg, \*.gif, \*.bmp, \*.png)을 불러와 텍타일 이미지로 제작하기 위한 바탕 그림으로 사용하는 기능 지원

- 점역 소프트웨어(점사랑)와 연동한 출력 기능 지원
- 텍스트와 그래픽 처리 방식이 다른 점자 프린터의 특성을 고려하여 텍스트와 이미지를 순차적으로 인쇄(인덱스 계열 프린트 지원)하는 기능 지원
- 직선, 곡선, 삼각형, 사각형, 원 등 그리기 도구 지원

다섯째, 서양 음악 점자 규정을 반영하고 점사랑 소프트웨어와 연동되는 악보 편집 및 점역 소프트웨어를 점사랑 5.0에 포함하였다. 본 사업에서는 짧은 개발 기간을 고려하여 직접 악보 편집 및 점역 소프트웨어를 개발하지 않고 전 세계적으로 가장 많이 사용되고 있는 Dancing Dots사에서 개발한 GOODFEEL 프로그램의 악보 편집 및 점역 기능을 이용하여 점사랑 5.0과 연동될 수 있도록 기능을 지원하였다.

점사랑 5.0 내에 포함된 악보 편집 및 점역 소프트웨어의 주요 기능은 다음과 같다.

- 악보 편집 및 점역을 통합한 소프트웨어 구현
- GUI 기반의 악보 제작 및 편집 기능 지원
- 점역 소프트웨어(점사랑)와 연동하여 출력 기능 지원
- 점역을 위한 Lime, MusicXML, NIFF 등의 파일 형식 지원
- 건반악기, 관현악, 성악 등의 악보를 점역하는 기능 지원
- 제목 텍스트, 박자 표시, 가사 및 리허설 마커를 점역하는 옵션 지원
- 사업 종료 이후에도 점사랑과 연동하여 지속적으로 사용 가능한 라이선스 확보

여섯째, 점사랑 소프트웨어 활용법 및 점자 출판물 제작 방법에 관한 온라인 교육 콘텐츠를 제작하였다. 기존의 점자 관련 교육 콘텐츠가 주로 점자 규정만을 학습 내용으로 다루었던 것과 달리 본 사업에서 제작한 콘텐츠는 점사랑 소프트웨어를 중심으로 점자 출판물 제작 과정을 다루고 있다. 본 사업에서 제작된 콘텐츠를 현장에 보급함으로써 점역·교정사의 점사랑 활용도를 높이고 점자 출판물 제작 과정의 효율성을 높일 수 있는 기반이 마련되었다.

## 2 제언

본 사업에서는 점자 출판물 제작 시 해외 소프트웨어에 의존하거나 수작업으로 진행할 수밖에 없던 비효율적인 제작 과정의 문제를 해결하기 위해 다양한 기능 개선을 시도하였다. 즉, 촉각 자료 제작 때문에 일부 페이지만을 해외 소프트웨어에 의존하여 발생했던 비효율적인 제작 과정의 문제를 극복하기 위해 점사랑 5.0과 연동되는 촉각 그래픽 소프트웨어를 개발하였다. 그리고 점사랑과 연동되는 악보 자동 점역 기능을 개발함으로써 소수의 음악 점역사만이 수작업으로 작업했던 악보 점역 방법을 개선하였다. 그러나 점자 출판물 제작에 있어 아직도 해결되어야 할 과제는 많이 남아 있으며, 점자 출판물 제작의 효율성을 높이기 위한 연구가 지속되어야 한다.

점자 출판물 관련 시각장애인의 정보 접근권 향상을 위해 본 사업의 결과를 바탕으로 삼아 점역 소프트웨어의 미래 개선 방향에 대한 제언을 한다면 다음과 같다.

첫째, 점역 소프트웨어의 유지보수 및 안정화 과정이 필요하다. 점사랑 5.0은 짧은 개발 기간 내에 많은 기능이 추가 및 개선되었고, 개정된 한국 점자 규정을 반영하였다. 그러므로 개발 기간 내에 미처 발견되지 못한 프로그램의 안정성 및 점역 정확도의 문제점이 사업 종료 이후에 나타날 수 있을 것으로 예상된다. 점역 소프트웨어가 현장에서 적절하게 활용되기 위해서는 무엇보다 프로그램의 안정성 및 점역 정확도가 중요하다. 이러한 상황을 고려할 때 점역 소프트웨어의 유지보수 및 안정화 과정이 지속적으로 이루어질 필요가 있다.

둘째, 다양한 외국어 점역 기능이 개발되어야 한다. 본 사업에서는 독일어, 스페인어, 프랑스어 등의 점역 기능이 개발되었다. 그러나 이외에도 외국어 교육에 대한 시각장애인의 다양한 요구가 계속되고 있는 실정이다. 특히 중국어, 일본어, 에스페란토 등은 시각장애인의 요구가 많음에도 아직 국내에 적합한 점역 소프트웨어가 없어 해당 점역 기능이 시급히 개발될 필요가 있다.

셋째, 데이지(DAISY) 파일에 대한 점역 기능이 개발되어야 한다. 현재 국립장애인도서관, 국립특수교육원 등에서는 시각장애인이 요구하는 경우 데이지 파일 또는 점역 파일로 대체 자료를 제작하고 있다. 특히 같은 자료에 대하여 2가지 대체 자료, 즉 데이지 파일과 점역 파일 모두 제작하는 경우도 빈번하게 발생하고 있다. 이러한 경우 각각 별도의 제작 과정을 거치기 때문에 제작 과정이 비효율적이다. 이미 제작된 데이지 파일의 구조를 분석하고 이를 활용하여 점역 코드와 연계할 수 있다면 점자 출판물 제작 과정을 자동화할 수 있어 대체 자료 제작 과정을 단축시킬 수 있다.

넷째, PDF 파일에 대한 점역 기능이 개발되어야 한다. 최근 국내 공공기관 등에서는 배포용 전자 문서로 PDF 파일을 많이 사용하고 있다. 사실상 표준으로 자리 잡은 정보공개용 PDF 포맷 문서는 파일에서 텍스트 정보를 추출하여 점자 번역은 가능하지만 서식 측면에서는 많은 제약이 있어 점자 출판물 제작에 많은 어려움을 야기하고 있다. 또한 확장자는 PDF로 동일하지만, PDF 문서 포맷은 다양한 유형이 존재하기 때문에 점역 자동화를 구현하기 더욱 어려운 실정이다. 접근성을 고려하여 제작된 PDF 표준은 PDF/UA(Universal Accessibility)이다. 그러나 일반적인 PDF 파일을 PDF/UA로 제작하는 작업은 전문 지식과 수작업이 요구되기 때문에 완벽한 PDF/UA 표준 파일이 제공되는 경우는 거의 드문 상황이다. 그러므로 PDF 파일에 대한 점역 기능을 지원하기 위해서는 소프트웨어 개발뿐만 아니라 제작자들의 인식 개선도 함께 이루어져야 한다.

다섯째, 교열 작업의 자동화 기능이 필요하다. 대체 자료 제작 시 원본 파일이 제공되지 않는 경우, 문서 스캔 및 OCR 등을 이용한 문서 인식 기능을 활용하여 최초의 텍스트 파일을 만들게 된다. 다만 이러한 파일은 문서 인식 기능이 완벽하지 않기 때문에 오류가 많아 상당 시간의 교열 과정을 필수적으로 거치게 된다. 최근 대규모 언어 모델(LLM) 등을 활용하여 개발된 ChatGPT와 같은 인공지능 소프트웨어는 문맥에 맞게 텍스트의 오타를 수정해 주는 기능을 제공하고 있다. 이러한 인공지능 소프트웨어의 기능을 점자 출판물 제작 과정에서 교열 자동화에 활용할 수 있다면 점자 출판물 제작 과정을 단축시킬 수 있을 것으로 기대된다.

여섯째, 촉각 그래픽 소프트웨어의 보완 및 기능 개선이 필요하다. 본 사업에서 개발된 촉각 그래픽 소프트웨어는 일부 점자 프린터만을 지원하고 있다. 다양한 점자 프린터 지원을 통해 현장 활용도를 높여야 하고, 선의 두께, 선의 패턴, 곡률 등의 정밀도가 높은 고품질의 촉각 자료 제작을 위해 기능 개선이 필요하다.

## 참고문헌

- 국립국어원 점자 종합정보 누리집(2022). 점사랑 4 매뉴얼. [korean.go.kr/braille](http://korean.go.kr/braille)
- 개정 참여자 및 기관(2022). 개정 점자 도서 제작 지침. 서울: 국립장애인도서관.
- 김영일, 강성주, 송지숙, 이민진, 허병훈 (2021). 시각장애학생용 대체 교과서 제작 지침 및 적용 방안 연구. 아산: 국립특수교육원.
- 김정호, 박성수, 이윤택(2014). 유니코드 점자·음성 처리 실용화 연구. 서울: 국립국어원.
- Braille Authority of North America (2010). Guidelines and standards for tactile graphics. [www.brailleauthority.org](http://www.brailleauthority.org)
- Braille Authority of North America (2016). Braille formats: Principles of print-to-braille transcription. [www.brailleauthority.org](http://www.brailleauthority.org)
- Duxbury Systems, Inc (2023). Duxbury DBT 12.7 Help File. [www.duxburysystems.com](http://www.duxburysystems.com)
- Duxbury Systems, Inc (2023). QuickTac 4.1 manual. [www.duxburysystems.com](http://www.duxburysystems.com)
- Thinkable (2022). TactileView user manual. [thinkable.nl](http://thinkable.nl)
- Dancing Dots (2023). GOODFEEL 4.0 Help File. [www.dancingdots.com](http://www.dancingdots.com)
- Dancing Dots (2023). GOODFEEL 4 Tutorial. [www.dancingdots.com](http://www.dancingdots.com)

---

#### 사업 총괄(PM)

이연주(한국시각장애인연합회 사무총장)

#### 사업 수행 인력(PL)

배진희(한국시각장애인연합회 말뚝치사업단 직원)

정지혜(한국시각장애인연합회 말뚝치사업단 직원)

#### 사업 수행 인력

천사랑(한국시각장애인연합회 말뚝치사업단 직원)

---

## 2023년 목자-점자 병렬말뚝치 구축 사업 최종 보고서

---

2024년 2월 29일 인쇄

2024년 2월 29일 발행

| 발행인 | 장소원

| 발행처 | 국립국어원

서울시 강서구 금남화로 154

전화 02-2669-9775, 전송 02-2669-9737

| 인쇄처 | 경성문화사

TEL: 02) 786-2999

---

※ 이 책은 국립국어원의 용역비로 수행한 '2023년 목자-점자 병렬말뚝치 구축' 사업 결과를 발간한 것입니다.

