

국립국어원 2018-01-19

발간등록번호
11-1371028-000723-01

2018년 한국수어 말뭉치 연구 및 구축

제 출 문

국립국어원장 귀하

국립국어원과 체결한 연구 용역 계약에 따라 ‘2018년 한국수어 말뭉치 연구 및 구축’에 관한 연구 보고서를 작성하여 제출합니다.

■ 사업 기간: 2018년 6월 8일 ~ 2018년 12월 12일

2018년 12월 12일

연구 책임자: 윤석민(전북대학교)

연구 기관: 전북대학교 산학협력단

연구 책임자: 윤석민

공동 연구원: 이영재, 이현숙, 고아라, 강진주,
최화니, 김성준

연구 보조원: 나두호, 서일권, 노유리, 이재준

협력 연구원: 변강석, 이은영, 박정민, 김다해,
김효정

<연구진>

책임 연구원	윤석민(전북대학교 국어국문학과 교수)
공동 연구원	이영재(전주선화학교 교사)
	이현숙(광주에바다농아교회 담임목사)
	고아라(전북대학교 국어국문학과 시간강사)
	강진주(전북대학교 국어국문학과 시간강사)
	최화니(전북대학교 국어국문학과 시간강사)
	김성준(전북대학교 국어국문학과 시간강사)
연구 보조원	나두호(광주광역시수어통역센터 교육팀장)
	서일권(대구대학교 사회복지학과 학부과정 재학)
	노유리(한국복지대학교 광고홍보과 학부과정 재학)
	이재준(전북대학교 국어국문학과 학부과정 재학)
협력 연구원	변강석(강남대학교)
	이은영(서울수화전문교육원)
	박정민(한국복지대학교)
	김다해(한국복지대학교)
	김효정(한국복지대학교)

2018년 한국수어 말뭉치 연구 및 구축

본 연구의 목적은 지역별 농인이 참여하는 한국수어 말뭉치 자료를 수집함과 동시에 기존 한국수어 말뭉치 자료의 오류를 바로잡고 기본 전사를 실시하고 한국수어 말뭉치 전사자 및 연구자의 양성을 위한 말뭉치 활용 워크숍을 개최함으로써, 언어로서의 한국수어 말뭉치 연구의 확대에 기여하고 지속적이고 체계적인 한국수어 말뭉치 연구의 토대를 마련하는 데에 있다. 연구 결과를 요약하면 다음과 같다.

첫째, 2015년도 한국수어 말뭉치의 시범 구축에 이어 두 번째로 한국수어 말뭉치를 구축함으로써 한국수어 자료를 확대하였다. 특히 한국수어 말뭉치 자료의 지역 균형성을 확보하기 위해 강원 지역, 전남 지역, 전북 지역, 경남 지역의 총 네 곳의 지역을 선정하여 언어 자료를 수집하였다. 총 16쌍의 언어 제공자를 대상으로 촬영을 진행하였으나 언어 제공자 선정 기준에 따라 최종적으로는 언어 제공자 14쌍의 원시 말뭉치 자료를 구축하였다. 연구의 목적에 따라 언어 제공자 질문지를 수정하고, 지역별 유도 자료를 제작한 뒤, 사회자를 선정하고 교육한 이후에 원시 말뭉치 촬영이 진행되었다. 원시 말뭉치 자료는 지역별로 강원 지역 11시간 21분 20초, 전남 지역 11시간 28분 22초, 전북 지역 12시간 50분 40초, 경남 지역 5시간 57분 15초 분량의 촬영이 이루어졌으며, 결과적으로 41시간 12초 분량의 원시 말뭉치 자료를 수집하였다. 이후 수집된 원시 말뭉치 영상을 가공하고 분절한 뒤 1차 번역을 실시하는 과정을 거쳐 최종적으로 전사 영상을 구축하였다.

둘째, 2015년 수집한 언어 자료(동영상) 중에서 6개의 영상, 총 재생시간 82분 54초 분량의 영상에 대해 우세손과 비우세손에 대한 번역 및 기본 전사를 완료하였다. 이 과정에서 전사자들의 역량 강화와 전사 일치도를 높이기 위한 전사 교육 또한 수시로 이루어졌다. 2015년에 수집된 언어 자료의 경우 번역이 되어 있지 않았기 때문에, 전사에 앞서 번역 작업이 우선적으로 이루어졌다. 번역 총렬까지 입력된 전사 영상은, 전사 품질을 일정 부분 유지하기 위해 기존에 전사 경험이 있는 전사팀A(북지대)에서 50분 분량을 담당하고, 전사팀B(전북대)에서 30분 분량을 담당하는 것으로 결정하여 전사 작업을 하였다. 전사팀B(전북대)의 경우, 공동전사를 기본 원칙으로 하여 매주 2회 이상 전사 회

의를 진행하였고 특히 연구진 전원이 모이는 정기 회의에서 전사 결과를 검토하고 전사 내용을 확정함으로써 전사의 질을 확보하고자 하였다. 전사 결과 발생한 토큰 수는 총 13,302개이며, 새로운 타입은 201개 생성하였다.

셋째, 한국수어 말뭉치의 구축 및 연구 성과를 연구자와 일반에게 공개하여 말뭉치 구축의 필요성 및 관심을 불러일으키기 위해 연구자 및 일반을 대상으로 한 워크숍을 기획하고, 한국수어 말뭉치 전문 전사자 양성을 위한 전사 교육을 실시하였다. 워크숍 및 전사 교육은 총 4차례에 걸쳐 이루어졌으며, 이를 통해 한국수어 말뭉치 구축의 중요성과 필요성에 대한 인식을 확산하고 원시 말뭉치 구축과 전사 말뭉치 구축을 위한 실질적인 능력을 확보할 수 있었다. 특히, 한국수어 말뭉치 구축에 참여하기를 원하는 농인 및 언어 연구자의 인력 풀을 넓힐 수 있는 계기가 되었다.

넷째, 향후 한국수어 말뭉치 구축의 효율성 확보를 위해, 2017년 개정된 한국수어 말뭉치 구축 지침을 다시 정비하였다. 주요 개정 내용은 언어 자료 수집 부분 중 언어 제공자 관리를 위한 넘버링 및 지역 구분과 관련된 내용을 추가하여 제시하고, 언어 자료 수집 과정에서 촬영 환경 조성에 대한 지침을 효율적인 방향으로 수정하고 보완한 것이다. 또한 전사 과정에서 발견된 문제점들을 바탕으로 글로스 확정 방법을 정비하고, 한국어 글로스에서 동음이의어로 나타나는 문제에 대한 해결방법을 보다 정교화하여 제시하였다. 그 외에 기존에 존재하던 타입에 대한 정비를 실시하여, 중복 타입의 삭제 및 새로운 타입의 추가 등을 제안하였다.

다섯째, 최종적인 연구 결과를 바탕으로 자문위원회의 검토를 받아 필요한 부분에 반영하였으며, 한국수어 말뭉치 연구에 대한 다양한 의견을 접할 수 있었다. 농인들의 참여율을 더 늘려야 한다는 의견에서부터 전사 전문 인력 양성에 대한 의견을 비롯하여 한국수어 말뭉치 연구의 발전을 위한 제언이 많았다. 더불어 본 연구팀의 연구 성과에 대해서도 긍정적인 평가를 보였다.

주요어: 한국수어, 한국수어 말뭉치, 전사, 지역 균형적 말뭉치 수집, 한국수어 연구

<Abstract>

Research and Construction of Corpus of Korean Sign Language in 2018

This study aim to collect data on the Korean Sign Language(KSL) corpus involving deafs living in different regions and It will correct errors in the KSL corpus, which is already construct. We also wanted to contribute to the expansion of KSL research by annotating and holding a workshop to foster KSL researchers. The results of the study are summarized as follows:

First, for the second time since Test construction of KSL corpus in 2015, the KSL corpus was constructed to expand the KSL data. In particular, language data from Gangwon Province, South Jeolla Province, North Jeolla Province, and South Gyeongsang Province were collected for the balanced regionalization of the KSL corpus data. A total of 16 pairs of language providers were filmed and 14 pairs of raw corpus data were built.

The questionnaire for the language provider was modified and the guidance material was modified to suit local characteristics. The selected host was trained to shoot a raw corpus. The original corpus data were taken in 11 hours, 21 minutes, 20 seconds in Gangwon Province, 11 hours, 28 minutes, 22 seconds in South Jeolla Province, 12 hours, 50 minutes, 40 seconds in North Jeolla Province, and 5 hours, 57 minutes and 15 seconds in South Gyeongsang Province. The corpus was then edited and translated into a first-time to contrsuct the corpus.

Second, we have chosen 6 corpus videos to translate and annotate them. The amount is 82 minutes and 54 seconds. In doing so, we frequently conducted annotation training to improve the quality of the annotations. Of the total amount, the Annotation Team A(KNUW) annotated 50 minutes, the Annotation Team B (JBNU) annotated 30

minutes, and the Annotation Team B (JBNU) held two meetings a week to improve the quality of the annotations. When we finished commenting, we found 13,302 tokens and 201 new types.

Third, we conducted workshops with other researchers and the public to spread the results of the research and construction of the KSL corpus. We also frequently trained to train professional KSL corpus annotators. Workshop and commenting education were conducted four times, and through this, we tried to spread awareness about the importance and necessity of construction a KSL corpus. And through the workshop, we was able to develop the practical skills needed to construct a KSL corpus. In particular, we secured Deaf and linguistics researchers who wanted to participate in the construction of the KSL and KSL corpus.

Fourth, the guidelines for the construction of the KSL corpus were revised again to construct the KSL corpus in the future. It presented additional details that distinguish areas from numbering for managing language providers. We has created a way to effectively film the filming environment of the Korean language corpus. We also revised the Gloss confirmation method on issues arising from the annotation of the KSL corpus and refined guidelines to resolve the homonym problem of the Gloss name. In addition, the existing type was refined to erase the duplicates and to add the newly discovered type.

Fifth, the results of the study were reviewed by the advisory committee and reflected. The main idea was that deaf children needed to be more involved and that they should be trained to specialize in annotating. In addition, the research team gave a positive evaluation of the results.

Keyword: Korean Sign Language(KSL), KSL Corpus, annotation, Constructing a Regional Balanced KSL Corpus, KSL Research

차 례

<국문 초록>	i
<Abstract>	iii

I. 서론

1. 연구 필요성과 목적	1
2. 연구 내용	2
3. 연구 특징	3
4. 기대 효과	4

II. 한국수어 말뭉치 구축: 강원/경남/전남/전북

1. 조사 지역 선정	5
2. 언어 제공자 모집 및 선정	6
2.1. 언어 제공자 모집	6
2.2. 언어 제공자 선정 및 통계	8
3. 언어 제공자 질문지	10
3.1. 언어 제공자 질문지 수정·보완	10
3.2. 질문지 작성의 문제점	19
4. 유도 자료 및 설명 영상 제작	21
4.1. 유도 자료 제작	21
4.2. 유도 자료 설명 영상 제작	25

차 례

5. 사회자 선정 및 교육	28
5.1. 사회자 선정	28
5.2. 사회자 교육	29
6. 원시 말뭉치 수집을 위한 촬영	30
6.1. 촬영 환경 조성	30
6.2. 외부 촬영실 구축(강원도)	32
6.3. 촬영의 실제	34
7. 추가 촬영	36
8. 원시 말뭉치 자료 처리	36
8.1. 한국수어 말뭉치 구축 자료 분석	36
8.2. 원시 말뭉치 1차 번역 자료 구축	38
9. 한국수어 원시 말뭉치 구축 지침 정비 및 제언	39
9.1. 한국수어 원시 말뭉치 구축 지침 정비	39
9.2. 한국수어 원시 말뭉치 구축 제언	43

Ⅲ. 한국수어 말뭉치 전사

1. 한국수어 말뭉치 전사 기본 사항	47
2. 한국수어 말뭉치 전사 과정	48
2.1. 전사 영상 선정	49
2.2. 기본 번역 및 검수	50
2.3. 전사 작업	54
2.4. 전사 결과	57

차 례

3. 한국수어 말뭉치 전사 지침 정비 및 제언	97
3.1. 한국수어 말뭉치 전사 지침 정비	97
3.2. 한국수어 말뭉치 전사 제언	101

IV. 한국수어 말뭉치 활용 워크숍 및 전사 교육

1. 한국수어 말뭉치 활용 워크숍의 의의	113
2. 한국수어 말뭉치 활용 워크숍 개최	114
2.1. 제1차 워크숍 “한국수어 말뭉치 구축의 필요성과 전사의 기초”	114
2.2. 제2차 워크숍 “한국수어 자료 통합 지원 시스템의 활용과 실제”	114
2.3. 제3차 워크숍 “한국수어 말뭉치 구축과 활용”	116
2.4. 제4차 워크숍 “한국수어 말뭉치 구축의 실제”	117

V. 결론 및 제언

1. 연구 결과 요약	119
1.1. 한국수어 말뭉치 구축	119
1.2. 한국수어 말뭉치 전사	120
1.3. 한국수어 말뭉치 활용 워크숍 및 전사 교육	121
1.4. 한국수어 말뭉치 구축 지침 정비 및 타입 정비	121
1.5. 자문위원회 의견	122

차 례

2. 향후 연구 방향 제언	123
2.1. 연구 체계의 효율성 확보	123
2.2. 인력 보충	124
2.3. 예산 확보	124
2.4. 대학 및 관련 기관과의 협조 체제 구축	125
2.5. 교육 저변 확대	125
 참고 문헌	 127
 <부록 1> 2018 한국수어 자료 수집을 위한 질문지 및 동의서	 129
<부록 2> 한국수어 말뭉치 구축 지침(2018년 개정)	141

표 차례

<표 1> 말뚱치 수집 대상 지역	5
<표 2> 선정된 언어 제공자 정보	8
<표 3> 사회자 명단	28
<표 4> 2018년도 말뚱치 영상 촬영 목록	34
<표 5> 추가 촬영 내역	36
<표 6> 전체 촬영 영상 중 유도 자료별 촬영 시간	37
<표 7> 유도 자료별 촬영 시간	37
<표 8> 한국수어 말뚱치 구축 지역 구분	40
<표 9> 2018년 한국수어 말뚱치 전사 영상 목록	50
<표 10> 2018년 한국수어 말뚱치 전사 일정표	56
<표 11> 2018년 한국수어 말뚱치 전사 과제별 세부 사항	57
<표 12> 2018년 한국수어 말뚱치 전사 신규 타입 목록	58
<표 13> [가만있다1] 타입 정보(수정 전)	110
<표 14 > [가만있다1] 타입 정보(수정 후)	110
<표 15> 2018년 워크숍 개최 실적	113
<표 16> 제1차 워크숍 프로그램 목차	114
<표 17> 제2차 워크숍 프로그램 목차	115
<표 18> 제3차 워크숍 프로그램 목차	116
<표 19> 제4차 워크숍 프로그램 목차	117

그림 차례

<그림 1> 언어 제공자 연령별 그래프	9
<그림 2> 언어 제공자 성별 분포 그래프	9
<그림 3> 수정 전 - 2017년 구축 지침에 따른 언어 제공자 질문지	11
<그림 4> 수정 후 - 2018년 한국수어 말뭉치 연구 및 구축 사업팀의 언어 제공자 질문지	12
<그림 5> 수정 전 - 2017년 구축 지침에 따른 언어 제공자 질문지	12
<그림 6> 수정 후 - 2018년 한국수어 말뭉치 연구 및 구축 사업팀의 언어 제공자 질문지	13
<그림 7> 수정 전 - 2017년 구축 지침에 따른 언어 제공자 질문지	13
<그림 8> 수정 후 - 2018년 한국수어 말뭉치 연구 및 구축 사업팀의 언어 제공자 질문지	14
<그림 9> 수정 전 - 2017년 구축 지침에 따른 언어 제공자 질문지	14
<그림 10> 수정 후 - 2018년 한국수어 말뭉치 연구 및 구축 사업팀의 언어 제공자 질문지	14
<그림 11> 수정 전 - 2017년 구축 지침에 따른 언어 제공자 질문지	15
<그림 12> 수정 후 - 2018년 한국수어 말뭉치 연구 및 구축 사업팀의 언어 제공자 질문지	15
<그림 13> 수정 전 - 2017년 구축 지침에 따른 언어 제공자 질문지	15
<그림 14> 수정 후 - 2018년 한국수어 말뭉치 연구 및 구축 사업팀의 언어 제공자 질문지	16
<그림 15> 수정 전 - 2017년 구축 지침에 따른 언어 제공자 질문지	16
<그림 16> 수정 후 - 2018년 한국수어 말뭉치 연구 및 구축 사업팀의 언어 제공자 질문지	17
<그림 17> 수정 전 - 2017년 구축 지침에 따른 언어 제공자 질문지	18

그림 차례

<그림 18> 수정 후 - 2018년 한국수어 말뭉치 연구 및 구축 사업팀의 언어 제공자 질문지	18
<그림 19> 추가적으로 논의가 필요한 질문지	19
<그림 20> 2018년 언어 제공자 질문지; 직장정보	21
<그림 21> 유도 자료 G-특정 사건 묘사 - 공통으로 제시되는 부분	25
<그림 22> G-특정 사건 묘사 - 지역별로 다르게 제시되는 부분	25
<그림 23> J-농사회 행사 - 지역별로 다르게 제시되는 부분	25
<그림 24> 1차 연습 촬영	27
<그림 25> 3차 본 촬영	27
<그림 26> 촬영 예시	30
<그림 27> 2018년 촬영장 사진	31
<그림 28> 2015년 촬영장 사진	31
<그림 29> 2018년 구축 영상 예시	31
<그림 30> 2015년 구축 영상 예시	31
<그림 31> 강원도 원주 촬영 영상	33
<그림 32> 강원도 춘천 촬영 영상	33
<그림 33> 잘못된 타입 선택 예시 [무엇1]	64
<그림 34> 잘못된 타입 선택 예시 [지시1호]+[counting hand]	65
<그림 35> 잘못된 타입 선택 예시 [승용차가다1]+[승용차가다1]	65
<그림 36> 잘못된 타입 선택 예시 [ㄷㅏ]	66
<그림 37> 잘못된 타입 선택 예시 [맞다1]	67

그림 차례

<그림 38> 잘못된 타입 선택 예시 [엄마1]	67
<그림 39> 잘못된 타입 선택 예시 [항상1A]	68
<그림 40> 잘못된 타입 선택 예시 [정말로1]	68
<그림 41> 잘못된 타입 선택 예시 [둘러보다1]	69
<그림 42> 잘못된 타입 선택 예시 [운전2]	70
<그림 43> 잘못된 타입 선택 예시 [걷다1]	70
<그림 44> 잘못된 타입 선택 예시 [외국2]	71
<그림 45> 잘못된 타입 선택 예시 [마주서다1]	72
<그림 46> 잘못된 타입 선택 예시 [깨끗하다1]+[구형1]	72
<그림 47> 잘못된 타입 선택 예시 [영점1]	73
<그림 48> 잘못된 타입 선택 예시 [명칭이1]	74
<그림 49> 잘못된 타입 선택 예시 [뽀1]	74
<그림 50> 잘못된 타입 선택 예시 [수다4]	75
<그림 51> 잘못된 타입 선택 예시 [간접1]	76
<그림 52> 잘못된 타입 선택 예시 [빛1]	76
<그림 53> 잘못된 타입 선택 예시 [안되다1]	77
<그림 54> 잘못된 타입 선택 예시 [새롭다1]	77
<그림 55> 잘못된 타입 선택 예시 [떠나다2]	78
<그림 56> 잘못된 타입 선택 예시 [실험1]	78
<그림 57> 잘못된 타입 선택 예시 [어른1]	79
<그림 58> 타입 생성 오류 예시 [빨강1]+[켜다1]+[파랑1]+[켜다1]	80

그림 차례

<그림 59> 타입 생성 오류 예시 [오월1]+[아래1]	80
<그림 60> 타입 생성 오류 예시 [비행기1]+[구형1A_면]	81
<그림 61> 타입 생성 오류 예시 [화나다3]	82
<그림 62> 타입 생성 오류 예시 [치다1]	82
<그림 63> 타입 생성 오류 예시 [메르스2]	83
<그림 64> 타입 생성 오류 예시 [설레다1]	84
<그림 65> 타입 생성 오류 예시 [야금야금쓰다1]	84
<그림 66> 타입 생성 오류 예시 [전화하다1]+[구형1]	85
<그림 67> 타입 생성 오류 예시 [무례1]	86
<그림 68> 타입 생성 오류 예시 [글1]	86
<그림 69> 타입 생성 오류 예시 [~대로1]	87
<그림 70> 타입 생성 오류 예시 [훌륭하다2]	87
<그림 71> 수어와 제스처의 구분 예시 [~하자2]	88
<그림 72> 수어와 제스처의 구분 예시 [많다5]	89
<그림 73> 수어와 제스처의 구분 예시 [수다2]	89
<그림 74> 수어와 제스처의 구분 예시 [웃걷다1]	90
<그림 75> 수어와 제스처의 구분 예시 [제스처]	90
<그림 76> 수어와 제스처의 구분 예시 [대회2]+[대회2]+[대회2]	91
<그림 77> 수어와 제스처의 구분 예시 [쓰다1]	92
<그림 78> 고정적 어휘와 생산적 어휘의 구분 예시 [구형1A_면]	92
<그림 79> 고정적 어휘와 생산적 어휘의 구분 예시 [구형1_면]	93

그림 차례

<그림 80> 고정적 어휘와 생산적 어휘의 구분 예시 [이형1]+[일형3]	94
<그림 81> 고정적 어휘와 생산적 어휘의 구분 예시 [지시1호]+[사람4]	94
<그림 82> 토큰 누락 예시 [허리1]	95
<그림 83> 분절 오류 예시 [교통1]	96
<그림 84> 분절 오류 예시 [왜1]+[왜1]+[왜1]	96
<그림 85> 기존 ID 글로스의 구조	97
<그림 86> 2018 수정 ID 글로스의 구조	98
<그림 87> 제1차 워크숍 행사 사진	114
<그림 88> 제2차 워크숍 행사 사진	115
<그림 89> 제3차 워크숍 행사 사진	116
<그림 90> 제4차 워크숍 행사 사진	117

I. 서론

1. 연구 필요성과 목적

한국수화언어법의 제정으로 한국수어는 한국어로서의 지위를 제도적으로 확보한 상태이다. 그러나 연구 대상으로서의 인식이 희박하고 연구 자료가 제대로 보급되지 않아 언어로서의 연구 가치가 확립되지 않은 상태이다. 이에 따라 국립국어원에서는 법령에 의거하여 2013년부터 한국수어 말뭉치의 연구와 구축 사업을 연차 사업으로 진행하고 있다. 기존에 진행된 한국수어 말뭉치 연구는 말뭉치 구축의 필요성을 확인하고 구축 지침을 제정하여 적용한 초기 단계의 연구로서 국내에서 한국수어 연구의 중요한 전환점이라고 볼 수 있다.

한국수어 말뭉치는 아직 초기단계이기 때문에 연구의 확대와 지속성의 확보가 중요하다. 본 연구는 한국수어 말뭉치 연구의 초기 확장 단계에 해당하는 연구로서 다음과 같은 필요성에 의해 진행되었다.

○ 언어로서의 한국수어 연구의 필요성

한국수어는, 음성언어와 다른 또 하나의 언어이며 한국 농인의 제1 언어이다. 그렇기 때문에 음성 한국어 수준의 언어학적 연구가 필요하다. 그러나 언어학자들에게는 한국수어에 대한 인식이 희박하여 언어학적 연구가 지엽적으로 이루어지고 있었을 뿐이다. 한국수화언어법의 제정되었기 때문에 한국수어가 연구는 제도적 뒷받침을 마련하게 되었고 이에 따라 연구 저변 확대를 위해 한국수어 말뭉치 구축과 연구가 필요하다.

○ 수어 연구를 위한 말뭉치 구축의 필요성

한국수어는 음성언어와 달리 습득과 학습 환경이 매우 열악하여 많은 변이가 나타나므로 다양한 수어 자료들을 광범위하게 기록하여 말뭉치로 제공해야 한다. 그리고 한국수어를 언어로서 체계적으로 연구하기 위해서는 연구자에게 연구 자료를 제공해야 할 필요가 있다. 따라서 한국수어 말뭉치의 구축이 필요하다.

○ 수어 사전 편찬을 위한 말뭉치 구축의 필요성

‘한국수화 사전’(2005년 초판, 2007년 개정판)은 수어의 지역별, 계층별 다양성을 반영하지 못하는 등 여러 가지 문제점이 있다. 그 이유는 음성 한국어의 표준국어대사전과는 다르게 전반적인 실제 언어 자료를 토대로 만들어진 것이 아니기 때문이다. 이 때문에 ‘한국수화 사전’은 한국수어의 사전으로서의 대표성을 확보하지 못한 상태이다. 따라서 한국수어의 실제 언어 자료를 토대로 구축한 말뭉치를 이용해 대표성을 갖춘 한국 수어 사전을 편찬할 필요가 있다.

○ 장기적이고 지속적인 말뭉치 구축의 필요성

국립국어원은 2015년 ‘한국수어 말뭉치 기반 조사 및 시범 구축’을 시작으로 2016년 ‘한국수어 자료 통합 지원 시스템 구축’, 2017년 ‘한국수어 말뭉치 연구 및 구축’을 통해 한국수어 말뭉치 구축의 기반을 닦고 연구를 진행하고 있다. 한국수어 말뭉치는 한국 농인들의 언어 기본권을 위해 그 필요성이 확실한바 일시적인 연구로 그칠 것이 아니라 지속적이고 장기적인 계획 속에서 수어 말뭉치 구축이 진행되어야 한다.

○ 지역 균형적인 말뭉치 구축의 필요성

2015년 연구에서 시범 구축한 원시 말뭉치의 전사가 아직 다 이루어지지 않았고, 전사 규칙 또한 재정비가 필요한 상황이다. 아울러 현재 구축된 말뭉치 자료는 서울 지역 농인들을 대상으로 수집한 것으로 다른 지역의 수어는 반영하지 못하므로, 서울 이외의 다른 지역에서도 수어 자료를 수집하고 지역 균형적인 말뭉치를 구축할 필요가 있다.

○ 말뭉치 구축 및 분석을 위한 전문가 양성 및 교육의 필요성

타당하고 합리적인 말뭉치의 구축 및 분석을 위해서는 수어 연구자와 말뭉치 전사자의 양성이 필수적이다. 이에 따라 말뭉치 활용 워크숍을 개최하여 교육 기회를 제공하고 연구자들 간의 교류의 장을 마련할 필요가 있다.

이에 따라 본 연구에서는 다음의 연구 목표를 설정하였다.

첫째, 지역 균형적 말뭉치의 구축을 위해 지역별 농인이 참여하는 한국수어 말뭉치 자료를 수집하고 정리한다.

둘째, 구축한 말뭉치를 수어 연구에 활용하기 위해 기존 한국수어 말뭉치 자료의 오류를 바로잡고 기본 전사를 실시한다.

셋째, 한국수어 말뭉치 전사자 및 연구자의 양성을 위한 말뭉치 활용 워크숍을 개최한다.

2. 연구 내용

본 연구는 위와 같은 연구의 필요성과 목적에 따라 한국수어 말뭉치에 대한 연구와 구축을 진행한다. 구체적인 연구의 내용은 다음과 같다.

○ 지역별 농인이 참여하는 한국수어 말뭉치 구축

- 한국수어 말뭉치 구축 지침에 따라 한국수어 언어 자료 수집을 위한 촬영을 진행한다.
- 한국수어 언어 자료 수집은 전북, 전남, 강원, 경남 네 지역의 농인 30명을 대상으로 진행한다.
- 수집된 원시 말뭉치 자료를 편집 가공한다.
- 구축된 원시 말뭉치에 대한 1차 번역을 진행한다.
- 다양한 지역의 한국수어 말뭉치를 구축하기 위한 구축 지침 보완한다.

○ 기존 한국수어 말뭉치 자료 기본 전사

- 2015년 국립국어원 사업을 통해 구축된 말뭉치에 대한 기본 전사 실시(80분)
- 한국수어 말뭉치 전사 지침의 보완 및 정교화
- 한국수어 말뭉치 구축 및 전사 인력 실무 교육

○ 한국수어 말뭉치 활용 워크숍 개최

- 한국수어 말뭉치 전사자 및 연구자를 위한 말뭉치 활용 워크숍 개최 (2회 이상)

3. 연구 특징

본 연구의 특징은 다음과 같다.

○ 정교하고 체계적인 말뭉치 구축을 위한 토대 마련

- 기존 말뭉치 구축 사업에서 수집된 자료의 지역적 한계를 극복하고 지역 균형적인 말뭉치를 구축할 수 있다.
- 기존 말뭉치 구축 사업에서 제공한 전사 지침 등의 문제를 수정하여 보다 정교하고 체계적인 말뭉치를 구축할 수 있다.
- 장기적으로 이어져야 할 말뭉치 구축 사업의 토대를 마련하여 향후 말뭉치 구축과 수어 연구에 기여한다.

○ 수어 연구 경험이 있는 언어학 연구자를 중심으로 연구진 구성

- 2012년 ‘농인의 국어능력 향상을 위한 기초연구’, 2013년 ‘농인의 문해 교육 실태 기초 연구’, 2014년 ‘수어기반 문해 교육 지원 체계 조사 및 설계’를 수행한 경험이 있는 연구자들이 참여하여 시행착오를 줄이고 효율적인 연구를 진행한다.
- 언어학 전공자들을 중심으로 연구진을 구성하여 언어학적인 관점에서 과학적이고 체계적인 연구를 실시한다.
- 한국어 말뭉치 구축 사업(세종 계획)을 수행한 경험이 있는 연구자들의 참여로 언어학적 관점이 반영된 수어 말뭉치 구축이 가능하다.

- 농인 당사자와 한국수어에 능통한 연구진 참여
- 언어학을 전공하는 농인 당사자 및 수어를 모어 수준으로 구사하는 연구진이 참여하여 연구의 신뢰성을 높인다.
- 농사회와의 강력한 사회 관계망을 갖춘 연구진이 참여하여 언어 제공자 모집 등 여러 부문에서 농사회 협조의 유도가 용이하다.

4. 기대 효과

이러한 연구 내용과 특징들을 바탕으로 본 연구에서 기대하는 효과는 다음과 같다.

- 한국수어 연구자들의 연구 기반 마련
- 한국수어 말뭉치 구축을 통해 연구자들이 보다 과학적이고 체계적인 수어 연구를 할 수 있는 토대를 마련할 수 있다.
- 농인의 수어 교육 기초 자료 제공
- 구축된 말뭉치는 농학교의 한국수어 교과서 및 교사용 지도서, 농인 학습용 수어 자료 제작 등에 기초가 되는 자료로 제공될 수 있다.
- 지역 균형적 한국수어 말뭉치 구축
- 서울·경기 지역에 편향되어 있는 기구축 말뭉치의 한계를 극복하고 다양한 지역의 특색 있는 말뭉치를 구축할 수 있다.
- 청인의 수어 교육 기준과 수준별 교육 자료 제작
- 구축된 말뭉치를 바탕으로 청인을 대상으로 한 수어 교육이나 한국수화(수어) 통역사 자격시험 등에서 활용할 수 있는 수준별 학습 및 평가 자료를 제작할 수 있다.
- 한국수어 말뭉치 연구 및 분석 전문 인력 양성
- 말뭉치 활용 설명회를 개최하여 관심 있는 여러 연구자들에게 연수 기회를 제공하고 말뭉치 연구 및 분석에 필요한 전문 인력을 양성할 수 있다.
- 한국수어 사전 편찬
- 조사 자료의 타당성과 신뢰도를 확보하여 구축된 말뭉치를 기반으로 하여 각종 수어 사전을 편찬할 수 있다.

Ⅱ. 한국수어 말뭉치 구축: 강원/경남/전남/전북

2018년 한국수어 말뭉치 연구 및 구축 사업의 첫 번째 목표는 최소 3개 지역, 총 40시간 이상의 말뭉치 자료를 구축하는 것이다. 이를 위해 지난 2015년부터 진행된 국립국어원의 한국수어 말뭉치 구축 사업의 결과 및 2017년 한국수어 말뭉치 구축 지침을 토대로 하여 한국수어 말뭉치를 구축하였으며, 보다 넓은 범위의 다양한 말뭉치 자료를 수집하고자 하였다.

1. 조사 지역 선정

2017년까지 시행되었던 일련의 한국수어 말뭉치 구축 사업을 통해 모아진 수어 언어 자료들은 서울 지역의 수어에만 한정되어 있었다. 그로 인해 특정 지역 내에서의 수어의 쓰임을 파악하거나 지역 간 차이를 보이는 수어를 포착하기는 불가능했다. 따라서 2018년 한국수어 말뭉치 연구 및 구축 사업에서는 서울 이외의 다른 지역들을 선정하여 말뭉치를 구축함으로써, 궁극적으로 한국수어 말뭉치 자료의 지역 균형을 확보하고자 하였다.

지역 수어를 수집하기 위해서는 해당 지역의 농사회 및 농인 관련 단체의 협조가 반드시 필요하다. 그렇기 때문에 구축 지역을 선정하는 데 있어서 적극적인 협조 의사를 밝힌 기관들이 포함된 지역들을 우선순위에 두고 논의함과 동시에 연구팀에서 쉽게 접촉할 수 있는 지역들도 함께 고려하였다. 그 결과 최종적으로 말뭉치 구축을 진행할 지역으로 네 곳의 행정 구역을 선정하였으며, 선정된 지역은 아래와 같다.

연번	구축 지역	언어 제공자
1	전북 지역	5쌍
2	전남 지역 ¹⁾	4쌍
3	강원 지역	4쌍
4	경남 지역	2쌍

<표 1> 말뭉치 수집 대상 지역

이 중 전라남북도의 경우, 현 연구팀 소속 연구진들의 활동 인근 지역으로 실질적

1) 이때 전라남도와 광주광역시를 구분하지 않고 기재하였다. 왜냐하면 전라남도의 다른 시군구와 광주광역시는 지리적으로 가까워(전라남도의 가운데에 광주광역시가 자리하고 있음) 서울과 경기도, 인천광역시를 묶어 수도권이라고 하는 것처럼 하나의 생활권으로 묶을 수 있다고 보았기 때문이다. 광주광역시에 자택이 있으나 나주나 순천 등에 직장이 있어 출퇴근을 하는 경우, 혹은 그 반대의 경우가 실제로 존재한다.

으로 접촉이 가능한 농인 커뮤니티 등이 다양하게 존재한다는 이점이 있었다. 이들 지역은 같은 호남권이라는 지역적 공통점이 존재함과 동시에, 행정 구역상 전라남도 와 전라북도로 구분이 되기 때문에 서로 간의 지역 수어를 비교·대조하기에도 적절하 다고 판단하였다. 덧붙여 거리적인 측면에서도 가깝기 때문에 촬영 등의 연구 진행에 용이하다는 이점도 있다.

반면에 강원 지역은 지역 농아인협회에서 적극적으로 참여 및 협조 의사를 알려진 점이 결정적으로 작용하였다. 또한 앞서 선정된 두 지역이 남부 지방인 것을 감안했 을 때, 중부 지방인 강원 지역을 선정함으로써 지역 균형적 연구라는 목적이 더욱 부 각될 수 있을 것으로 보았다.

경남 지역의 경우 호남 지역과의 수어 비교군으로서 선정하였다. 그러한 까닭에 이 번 연구에서는 많은 언어 제공자 쌍을 모집하지는 않았으나, 추후 말뭉치 구축 확대 가 이루어져야 하는 지역이다. 본 사업이 지역 균형적 말뭉치 구축의 시발점인 것을 고려할 때, 지역 수어 말뭉치 구축의 기반으로 삼을 수 있다는 데에 의의를 두고자 한다.

2. 언어 제공자 모집 및 선정

2.1. 언어 제공자 모집

1) 모집 방법

2017년 개정 한국수어 말뭉치 구축 지침에 따르면, 언어 제공자의 모집은 세 가지 방법을 이용하도록 제시되어 있다. 첫째, 지역의 수화통역센터에 공문을 발송하여 모 집하는 방법, 둘째, 농인 커뮤니티의 페이스북과 같은 누리소통망에 공고를 하는 방 법, 셋째, 농사회의 리더를 통해 개별적으로 모집하는 방법이다.

그러나 2017년의 개정 지침 중 ‘누리소통망 공고’의 형태로 모집을 하는 경우, 모 집 기간 내에 충분한 언어 제공자 확보가 되지 못할 경우 말뭉치 구축이 지연될 수 있다는 문제점이 있어서 논외로 하였다. 따라서 유관 기관의 협조를 받거나, 연구팀 자체적으로 농사회를 통해 개별적으로 모집하는 방법 등을 적절히 분배하여 이용하기 로 결정하였다. 본 사업팀에서 언어 제공자를 모집한 방법은 다음과 같다.

- ① 농아인협회와의 협의를 통한 예비 명단 확보
- ② 농사회 커뮤니티를 통한 직접 모집

가장 먼저 지역 농아인협회와의 협의를 통해 언어 제공자 예비 명단을 제공받았다. 특히나 연구팀이 직접 접촉하기에 지리적으로 먼 거리에 있는 강원 지역의 경우 협회

의 지원이 많은 도움이 되었다. 그와 동시에 지역의 농학교 졸업생 모임이나 수어통역센터 등의 농사회 커뮤니티를 통해 연구팀이 직접 언어 제공자를 모집하였다.

2018년 한국수어 말뭉치 연구 및 구축팀이 가지는 장점 중 한 가지는 농인들이 연구에 직접적으로 참여한다는 점이다. 특히 조사 지역으로 선정한 네 곳의 행정 구역 중 근거리 지역인 전북, 전남, 경남 지역의 경우, 연구팀 소속의 농인과 수화통역사를 중심으로 언어 제공자 모집이 이루어졌다. 그 과정에서 후술할 언어 제공자 기본 조건이 잘 갖추어졌는지를 더욱 면밀하게 살필 수 있었다.

2) 언어 제공자 기본 조건

언어 제공자가 말뭉치 구축에 적합한지를 판단하기 위해 연구진 회의를 통해 기본 조건을 충족하는지를 점검하였다. 2017년 개정 한국수어 말뭉치 구축 지침에서 제시한 언어 제공자의 기본 조건은 아래와 같다(국립국어원, 2017: 207).

- ① 19세 이상 성인 농인
- ② 가장 편하고 많이 사용하는 의사소통 방법이 한국수어임
- ③ 현재 언어 자료 수집 지역에 10년 이상 거주하고 있음
- ④ 일주일에 3번 이상 농인을 만남
- ⑤ 농학교를 졸업함

본 사업팀에서도 위의 기본 조건에 대부분 동의하는 바이다. 그에 따라 전체적인 맥락은 따르되, 일부 항목에 대해서는 표현 및 내용을 재정비하였다. 그중에서 ②번의 ‘가장 편하고 많이 사용하는 의사소통 방법’이라는 표현은 상당히 모호하다. 따라서 주된 의사소통을 한국수어로 사용하는 농인이라는 의미에서 ‘한국수어를 모국어로 사용하는 농인’으로 수정하였다. ③번의 ‘현재 언어 자료 수집 지역에 10년 이상 거주하고 있음’의 경우, 반드시 같은 도시권역이 아니더라도 도내에서 거주지를 옮긴 경우까지 포함하기로 하였다. ④번 항목의 ‘일주일에 3번 이상 농인을 만남’에 대해서는 단순히 만남의 횟수만으로 한국수어의 유창함을 논할 수 있는지의 문제가 제기되었다. 그에 따라 본 사업팀은 만남의 횟수만을 제한하기보다는 얼마나 농사회에서 활발하게 활동하고 있는지, 농사회 내에서의 해당 농인의 성향이나 활동 범위를 함께 고려하는 것으로 결정하였다. ⑤번 조건인 ‘농학교를 졸업함’의 경우, 농학교가 한국수어를 습득하는 데 최상의 교육 장소이면서 한국수어를 사용하는 데에 최적의 장소라는 이유로 기존의 조건을 유지하고자 하였다²⁾.

이에 따라 수정된 언어 제공자 기본 조건은 다음과 같다.

2) 그러나 이 부분에 대해 회의적인 견해도 더러 있었다. 이와 관련된 이야기는 9.2. 한국수어 원시 말뭉치 구축 제언 부분에서 다시 하도록 한다.

- ① 19세 이상 성인 농인
- ② 한국수어를 모국어(제1언어)로 사용하는 농인
- ③ 농학교를 졸업한 농인
- ④ 조사 지역으로 선정된 행정 구역 내에서 최근 10년 이상 지낸 농인
- ⑤ 농사회에서의 교류가 활발한 농인

2.2. 언어 제공자 선정 및 통계

상술한 과정들을 통해 연구팀이 직접 모집하였거나, 협회를 통해 확보한 언어 제공자 예비 명단으로부터 기본 조건을 충족한 경우 최종적으로 언어 제공자로 선정하였다. 그 외에 다음과 같은 경우도 언어 제공자 자격 기준에 적합하다고 판단되어 언어 제공자로 선정하였다.

- 농사회(농아인협회 및 농인커뮤니티)에서 (지역) 수어 사용과 관련하여 추천을 받은 농인

이에 따라 당초 총 30명의 언어 제공자를 선정하였으며, 이들을 친분, 연령, 성별, 지역 등을 고려하여 2인 1조로 구성하여 총 15쌍의 언어 제공자 쌍을 확정하였다. 그러나 모든 촬영이 끝난 이후 검토하는 과정에서 1쌍이 언어 제공자의 조건에 맞지 않는다는 문제가 제기되었고, 1쌍을 추가로 선정하게 되었다.³⁾ 최종 언어 제공자 명단은 다음과 같다.

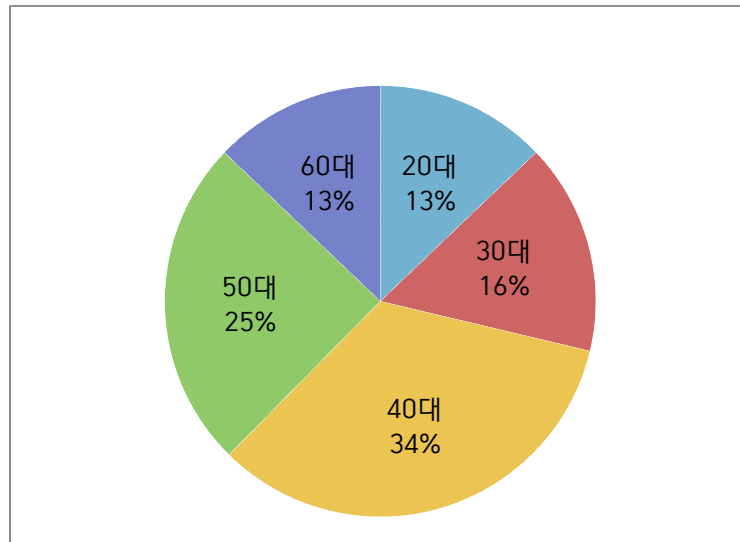
지역	성명	성별	출생년도	지역	성명	성별	출생년도
전북	김다O	여	1986	전남	양영O	여	1968
전북	신은O	여	1981	전남	이진O	여	1971
전북	양정O	남	1954	강원	정광O	여	1978
전북	강승O	여	1971	강원	곽순O	여	1963
전북	노대O	남	1989	강원	김남O	남	1985
전북	강성O	남	1995	강원	구자O	여	1975
전북	이상O	남	1972	강원	김청O	여	1968
전북	최명O	여	1977	강원	최경O	여	1964
전북	이혜O	여	1981	강원	한옥O	여	1964
전북	전미O	여	1955	강원	이인O	여	1970
전남	정영O	남	1954	경상	손경O	여	1979

3) 이에 대한 내용은 7. 추가 촬영 부분에서 자세히 다루도록 한다.

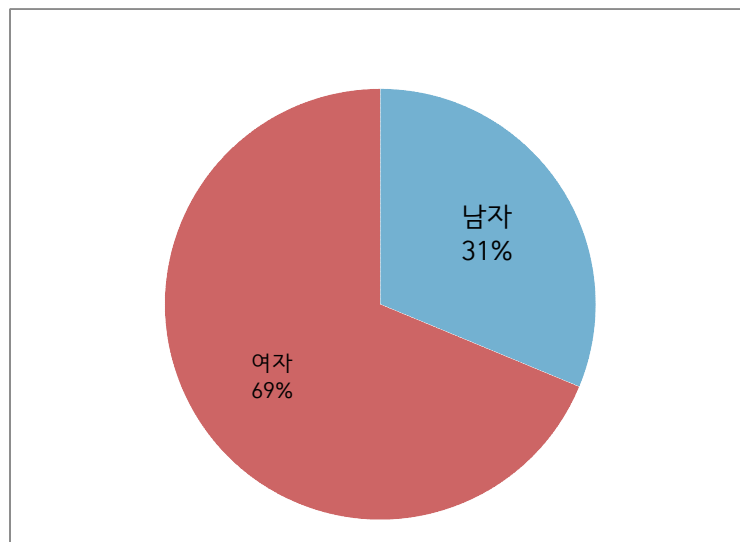
전남	최인O	남	1965	경상	이옥O	여	1960
전남	정영O	남	1969	경상	조수O	남	1958
전남	이나O	여	1970	경상	김현O	남	1960
전남	이보O	여	1990	전북	김승O	여	1978
전남	허성O	여	1994	전북	이경O	여	1977

<표 2> 선정된 언어 제공자 정보

언어 제공자들의 성별과 연령별 통계를 그래프로 보이면 다음 그림과 같다.



<그림 1> 언어 제공자 연령별 그래프



<그림 2> 언어 제공자 성별 분포 그래프

언어 제공자들은 남성보다 여성이 많았으며, 연령별로는 40대와 50대가 가장 큰 비중을 차지했다. 남성보다 여성의 비중이 높은 이유는 남성들에 비해서 여성들이 대화

하는 자리에 거부감을 덜 느끼기 때문에 섭외 과정에서 조금 더 긍정적으로 받아들일 수 있었던 것으로 이해할 수 있다. 또한 촬영 요일이나 시간대를 감안할 때 남성의 경우 직장 문제로 시간을 내기 힘들었던 측면도 있다.

연령대별로 살펴봤을 때 40-50대의 언어 제공자들이 많았는데, 이는 지역에서 조금 더 오래 지낸 중장년층 계열이 청년층에 비해 지역 수어를 더 능숙하게 사용하기 때문인 것으로 보인다. 다만 60대의 언어 제공자가 적은 것은 아마도 ‘농학교를 졸업함’이라는 조건 때문인 것으로 파악할 수 있는데, 당시 시기상 지역 내에 농학교가 없었거나, 농학교가 있더라도 형편상 농학교에 진학하지 못했을 가능성이 크다.

실제 한국어 음성 언어의 경우에도 지역어 조사 시 고연령층이 지역 방언을 잘 알고 있는 경우가 많다. 이를 통해 수어의 경우도 마찬가지로 연령대가 높을수록 지역의 특성을 반영한 수어를 많이 알고 있으며, 지역적 특색이 포함된 수어를 산출할 가능성이 높을 것으로 예상할 수 있다. 현 연구팀에서는 연령대와 성별을 고려한 사회언어학적 연구까지 나아가지는 않았지만, 이러한 자료를 토대로 앞으로 한국수어의 사회언어학적 연구도 이루어질 수 있을 것으로 기대한다.

3. 언어 제공자 질문지

언어 제공자의 메타 데이터를 수집하기 위해서는 질문지를 작성하는 것이 필요하다. 질문지는 개인 정보, 가족 정보, 사용 언어에 관한 정보, 학교 및 직업에 관한 정보, 농사회 활동 등에 대한 내용으로 구성되어 있다. 질문지를 확정하고 배포, 작성하는 데에는 아래와 같은 단계를 거쳤다.

- 1) 2017년 사업팀의 지침을 토대로 질문지 수정·보완
- 2) 언어 제공자들에게 질문지 배포·사전 작성 요청 → 원활히 이루어지지 않음⁴⁾
- 3) 촬영 전 언어 제공자들과 함께 질문지 작성
- 4) 메타 데이터 등록

3.1. 언어 제공자 질문지 수정·보완

수집된 언어 제공자의 메타 데이터는 한국수어 자료 통합 지원 시스템에 등록한다. 기존 연구팀과의 일관성을 위해 2017년 한국수어 말뭉치 구축팀의 질문지를 최대한 활용하고자 하였다. 다만 질문이나 답변란을 명시적으로 표시해야 하거나 부가적인 정보가 더 필요하다고 생각되는 부분, 서식을 통일해야 하는 부분들에 대해서는 연구진 회의를 통해 일부 내용을 수정하였다. 전체 질문지의 내용은 <부록>에 제시하며,

4) 이에 대한 내용은 3.2. 질문지 작성의 문제점 부분에서 다루도록 한다.

아래에는 수정된 내용과 함께 추후 논의가 필요한 부분에 대해서만 간략히 제시하도록 한다.⁵⁾

1) 질문이나 답변란이 좀 더 명시적으로 표시되어야 하는 부분

2017년도의 자료 수집을 위한 질문지에서 가장 보완이 필요한 부분이 바로 이 부분이었다. 질문의 의도를 파악하기 힘들거나 답변을 적을 때 작성자로 하여금 고민을 많이 하게 만들 것으로 예측되는 부분이 있었기에, 질문의 내용과 답변란의 구성, 척도 등을 적절하게 수정하였다.

특히나 척도가 제대로 표시되지 않아 작성에 어려움을 겪을 수 있는 부분이 눈에 띄었다.

28. 본인의 한국어 능력 정도

	좋다				나쁘다	
읽기	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
쓰기	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

29. 외국수어 사용 여부

미국수어	<input type="checkbox"/>	매우 잘 한다	<input type="checkbox"/>	잘 한다	<input type="checkbox"/>	보통	<input type="checkbox"/>	잘 못한다	<input type="checkbox"/>	약간
일본수어	<input type="checkbox"/>	매우 잘 한다	<input type="checkbox"/>	잘 한다	<input type="checkbox"/>	보통	<input type="checkbox"/>	잘 못한다	<input type="checkbox"/>	약간
국제수어	<input type="checkbox"/>	매우 잘 한다	<input type="checkbox"/>	잘 한다	<input type="checkbox"/>	보통	<input type="checkbox"/>	잘 못한다	<input type="checkbox"/>	약간
_____	<input type="checkbox"/>	매우 잘 한다	<input type="checkbox"/>	잘 한다	<input type="checkbox"/>	보통	<input type="checkbox"/>	잘 못한다	<input type="checkbox"/>	약간

<그림 3> 수정 전 - 2017년 구축 지침에 따른 언어 제공자 질문지

위 그림에서 28번 질문인 한국어 능력 정도는 ‘좋다’와 ‘나쁘다’의 척도보다는 ‘잘 한다’와 ‘못한다’로 구분하는 것이 훨씬 자연스럽게 느껴진다. 또한 중간 단계의 척도가 제시되어 있지 않아서 어느 정도 잘하고 못하는지를 체크하기에 어려움이 있다.

29번의 외국 수어 사용 여부 또한 척도가 애매하다. ‘잘 못한다’와 ‘약간’의 정도 차이를 구분해 내기 모호하며, 외국 수어를 전혀 모르는 경우도 반영되어 있지 않다. 이러한 의견을 종합하여 다음과 같이 수정하였다.

5) 아래의 질문지 이미지에서 수정 전 오류 항목은 붉은색으로, 수정 후 항목에는 파란색으로 표시하였다. 이는 수정된 부분을 명시적으로 보여주고자 한 것으로, 실제로 언어 제공자에게 배포한 최종 질문지에는 이와 같은 표시는 되어 있지 않다.

28. 본인의 한국어 능력 정도 *척도 명시적으로 수정함

읽기 ☐ 아주 잘한다 ☐ 잘한다 ☐ 보통 ☐ 조금 한다 ☐ 전혀 못한다
쓰기 ☐ 아주 잘한다 ☐ 잘한다 ☐ 보통 ☐ 조금 한다 ☐ 전혀 못한다

29. 외국수어 사용 여부 *척도 명시적으로 수정함

미국수어 ☐ 아주 잘한다 ☐ 잘한다 ☐ 보통 ☐ 조금 한다 ☐ 전혀 못한다
일본수어 ☐ 아주 잘한다 ☐ 잘한다 ☐ 보통 ☐ 조금 한다 ☐ 전혀 못한다
국제수어 ☐ 아주 잘한다 ☐ 잘한다 ☐ 보통 ☐ 조금 한다 ☐ 전혀 못한다
_____ ☐ 아주 잘한다 ☐ 잘한다 ☐ 보통 ☐ 조금 한다 ☐ 전혀 못한다

<그림 4> 수정 후 - 2018년 한국수어 말뭉치 연구 및 구축 사업팀의 언어
제공자 질문지

질문에 대한 답변으로 5단계의 척도를 명시적으로 제시하여, 작성자들이 좀 더 직
관적으로 답변을 선택할 수 있는 효과를 준다.

다음은 질문 및 답변란이 모호하게 작성되어 혼란을 야기할 수 있는 부분이다.

30. 출신 학교

출신학교1: _____ (제일 오래 다녔던 학교 이름을 적어주세요)

주된 의사소통 방법:

☐ 한국수어 ☐ 한국어대응수어+말 ☐ 한국수어 + 한국어 ☐ 구화
☐ 수화통역사와 함께 ☐ 기타 _____

출신학교2: _____ (제일 오래 다녔던 학교 이름을 적어주세요)

주된 의사소통 방법:

☐ 한국수어 ☐ 한국어대응수어+말 ☐ 한국수어 + 한국어 ☐ 구화
☐ 수화통역사와 함께 ☐ 기타 _____

<그림 5> 수정 전 - 2017년 구축 지침에 따른 언어 제공자 질문지

보다시피 2017년 질문지의 학교 정보에서는 학교를 다닌 기간을 전혀 알 수 없고,
무엇보다 학교 정보 오른쪽에 있는 ‘제일 오래 다녔던 학교를 적어주세요’라는 부분이
요구하는 정보를 알아채기 힘들다. 이에 우리는 농인들이 해당 질문에 대해 어느 정
도로 정보를 제공했는지를 파악하기 위해서 한국수어 자료 통합 지원 시스템에 기등
록된 언어 제공자들의 정보를 열람하였다. 해당 질문에 대한 답변으로는, 농학교에서
초중등과정을 졸업한 농인이 있는가 하면, 농아유치원부터 시작해서 농학교를 거쳐
대학교까지 졸업한 농인도 있었다. 이러한 정보를 질문지에 반영하여 다음과 같이 질

문과 답변란을 수정하였다.

30. 출신학교 *농아유치원부터 대학교까지 적을 수 있도록 학교 항목 5개로 늘림
*답변란 명시적으로 수정

출신학교1: _____살부터 _____살까지 _____(학교이름)

주된 의사소통 방법:

- ☐ 한국수어 ☐ 한국어대응수어+말 ☐ 한국수어 + 한국어 ☐ 구화
☐ 수화통역사와 함께 ☐ 기타 _____

출신학교2: _____살부터 _____살까지 _____(학교이름)

주된 의사소통 방법:

- ☐ 한국수어 ☐ 한국어대응수어+말 ☐ 한국수어 + 한국어 ☐ 구화
☐ 수화통역사와 함께 ☐ 기타 _____

<그림 6> 수정 후 - 2018년 한국수어 말뭉치 연구 및 구축 사업팀의
언어 제공자 질문지

농인들 중에는 나이가 든 이후에 평생교육원이나 방송통신대학교를 졸업한 사람들이 있다는 것을 감안하여, 학교 항목을 5개로 늘려 고등교육과정까지 충분히 적을 수 있게 하였고, 언제부터 언제까지 어떤 학교에 다녔는지를 묻는 것으로 수정하였다.
언어 제공자들 간의 관계를 명확하게 파악하기 힘든 경우도 존재했다.

언어제공자 A와 언어제공자 B

- 1) ● 우리는 서로 매우 잘 안다
● 우리는 서로 잘 안다
● 우리는 과거에 만난 적이 있다
● 우리는 오늘 처음 만났다
- 2) 우리는 서로 친척 관계이거나 결혼을 통한 친척 관계이다. ☐ 예 ☐ 아니오
(‘예’라고 답한 경우) 어떤 친척 관계인지(예, 조카 등)

<그림 7> 수정 전 - 2017년 구축 지침에 따른 언어 제공자 질문지

위의 질문에서 체크 박스가 제대로 표시되지 않은 것은 차치하고, 2)의 내용을 살펴 보면 언어 제공자들이 친인척 관계인 경우에만 관계를 표시하게끔 되어 있었다. 그러나 실제로 친인척 관계의 농인들끼리 언어 제공자 쌍으로 묶일 확률은 매우 적으며, 왜 친인척 관계에 대해서만 답변을 작성하는지 파악할 수가 없었다. 언어 제공자 쌍은 선후배 관계나 직장 동료, 혹은 동네 주민일 경우, 전혀 모르는 사람들로 구성되기도 한다. 따라서 서로의 관계를 보다 구체적으로 적을 수 있도록 하였다.

언어제공자 A와 언어제공자 B의 관계

- 1) ☐ 우리는 서로 매우 잘 안다
☐ 우리는 서로 잘 안다
☐ 우리는 과거에 만난 적이 있다
☐ 우리는 오늘 처음 만났다

* 서로의 관계를 명시적으로 표시할 수 있도록 수정

- 2) 만약 서로가 아는 관계라면 어떤 관계인지 구체적으로 작성해 주세요.

(예: 형, 조카, 친구, 선배 등) _____

<그림 8> 수정 후 - 2018년 한국수어 말뭉치 연구 및 구축 사업팀의 언어
제공자 질문지

이 외에도 아래의 항목들에서 조금 더 명시적인 표현이 사용될 필요가 있었다.

[37번에서 '예'라고 대답한 경우] 마지막 직장에 관한 다음 질문에 답하시오

38. 마지막 직장 정보

나는 _____년부터 _____로써(직종) 일했다.

41. 한국수어 강사 활동 정도

[40번에서 '예'라고 답한 경우] ☐ 자주 ☐ 가끔 ☐ 드물게

<그림 9> 수정 전 - 2017년 구축 지침에 따른 언어 제공자 질문지

[37번에서 '예'라고 대답한 경우] 마지막 직장에 관한 다음 질문에 답하시오

38. 마지막 직장 정보 *답변란 명시적으로 수정

직종(직장) _____ 근무 기간 _____년부터 _____년까지

41. 한국수어 강사 활동 정도 *기간을 알 수 있도록 척도 수정

[40번에서 '예'라고 답한 경우]

☐ 1년 이하 ☐ 1~2년 ☐ 3~4년 ☐ 5년 이상

<그림 10> 수정 후 - 2018년 한국수어 말뭉치 연구 및 구축 사업팀의 언어
제공자 질문지

2) 추가적인 정보가 더 필요한 부분

추가적인 정보가 필요한 부분은 주로 답변란에서 나타났다. 제시된 답변 항목만으

로 충분한 정보를 얻기 힘든 경우들이 존재했고, 답변 선택지 중 누락된 항목이 존재하는 경우도 나타났다.

아래에 몇 가지 예를 제시한다.

15. 형제 및 자매의 청력상태/나이

- ① ☐ 농인 ☐ 난청인 ☐ 청인
 ② ☐ 농인 ☐ 난청인 ☐ 청인
 ③ ☐ 농인 ☐ 난청인 ☐ 청인
 ④ ☐ 농인 ☐ 난청인 ☐ 청인
 ⑤ ☐ 형제 및 자매 없음

<그림 11> 수정 전 - 2017년 구축 지침에 따른 언어 제공자 질문지

위의 그림은 15번 형제 및 자매의 청력상태와 나이를 알고자 하는 질문 항목이다. 그런데 질문란에 '나이'라는 표시가 있음에도 불구하고 답변란에 '나이'를 기재할 수 없는 항목이 없다는 것이 문제가 되었다. 따라서 괄호 정보를 추가하여 형제 및 자매와의 나이 차이를 적도록 하였다.

15. 형제 및 자매의 청력상태/나이 *나이 정보가 없어서 괄호 정보 추가함

- ① ☐ 농인 ☐ 난청인 ☐ 청인 (나보다 나이가 _____살 ☐ 많다 ☐ 적다)
 ② ☐ 농인 ☐ 난청인 ☐ 청인 (나보다 나이가 _____살 ☐ 많다 ☐ 적다)
 ③ ☐ 농인 ☐ 난청인 ☐ 청인 (나보다 나이가 _____살 ☐ 많다 ☐ 적다)
 ④ ☐ 농인 ☐ 난청인 ☐ 청인 (나보다 나이가 _____살 ☐ 많다 ☐ 적다)
 ⑤ ☐ 형제 및 자매 없음

<그림 12> 수정 후 - 2018년 한국수어 말뭉치 연구 및 구축 사업팀의 언어 제공자 질문지

한편 다음과 같은 경우도 있었다.

27. 농(난청)인과의 정기적인 만남 유무: ☐ 예 ☐ 아니오

<그림 13> 수정 전 - 2017년 구축 지침에 따른 언어 제공자 질문지

27번 질문은 언어 제공자가 농사회에서 어느 정도로 활발하게 활동하고 있는지를

파악하고자 하는 질문이다. 그런데 단순히 만남의 유무만 제시되어 있어 해당 언어 제공자의 활동성을 파악하기에는 다소 빈약하다. 또한 ‘정기적인 만남’이라는 것이 ‘일 년에 한 번’인지, ‘한 달에 한 번’인지, ‘매주 한 번’인지도 가늠할 수 없다. 본 연구팀에서는 ‘한 주에 몇 회 가량 농민들과 어울리는지’⁶⁾의 부가 정보를 적을 수 있도록 하여 언어 제공자가 농사회에서 얼마만큼 활동을 하는지를 대략적으로나마 파악하고자 하였다.

27. 농(난청)인과의 정기적인 만남 유무: ☐ 예 (주 ____회 만남) ☐ 아니오

<그림 14> 수정 후 - 2018년 한국수어 말뭉치 연구 및 구축 사업팀의 언어 제공자 질문지

아래는 답변의 선택지 중 누락된 항목이 존재하는 경우들이다.

9. 아버지의 청력상태: ☐ 농인 ☐ 청인

10. 어머니의 청력상태: ☐ 농인 ☐ 청인

11. 부모와의 동거 유무: ☐ 예 ☐ 아니오

12. 부모 외의 양육자: (11번 질문에 ‘아니오’라고 답한 경우)

☐ 할아버지 ☐ 할머니 ☐ 친척 ☐ 기타

17. 가족 및 친척의 청력상태

할머니 ☐ 농인 ☐ 난청인

할아버지 ☐ 농인 ☐ 난청인

_____ ☐ 농인 ☐ 난청인

_____ ☐ 농인 ☐ 난청인

<그림 15> 수정 전 - 2017년 구축 지침에 따른 언어 제공자 질문지

그림의 내용을 살펴보면, 부모의 청력 상태는 ‘농인’과 ‘청인’만으로 구분하고, 조부모의 청력 상태는 ‘농인’과 ‘난청인’만으로 구분하는 오류가 있다. 이 경우, 질문지 작

6) 앞서 언어 제공자 기본 조건에서도 서술하였듯 주당 만남 횟수가 언어 제공자의 활동성을 그대로 보여주는 척도라고는 볼 수 없다. 다만 상술했듯 부가적인 정보로 이용하고자 하였고, 그 외에 해당 농인이 얼마나 활동적인지는 성격이나 농사회 활동 범위 등을 물음으로써 다각적으로 파악하고자 하였다.

성 시 선택지에 해당 항목이 없어 혼란을 느끼는 언어 제공자들이 있을 수 있다. 따라서 ‘농인, 난청인, 청인’의 항목을 모두 제시하여 언어 제공자의 혼란을 덜 수 있도록 하였다. 그 밖에도 12번 질문에 대한 답변으로 ‘없음’ 항목을 추가하고, 선택지 외의 정보가 나올 경우를 대비하며 부가 정보를 작성할 수 있는 칸을 삽입하였다.

9. 아버지의 청력 상태: ☐ 농인 ☐ 난청인 ☐ 청인 *난청인 정보 추가

10. 어머니의 청력 상태: ☐ 농인 ☐ 난청인 ☐ 청인 *난청인 정보 추가

11. 부모와의 동거 유무: ☐ 예 ☐ 아니오

12. 부모 외의 양육자 * 없음 항목 추가 * 이하 기타 항목에 추가정보 작성란(밑줄) 삽입

☐ 할아버지 ☐ 할머니 ☐ 친척 ☐ 없음 ☐ 기타 _____

17. 가족 및 친척의 청력상태 * 청인 정보 추가

할머니 ☐ 농인 ☐ 난청인 ☐ 청인

할아버지 ☐ 농인 ☐ 난청인 ☐ 청인

_____ ☐ 농인 ☐ 난청인 ☐ 청인

_____ ☐ 농인 ☐ 난청인 ☐ 청인

<그림 16> 수정 후 - 2018년 한국수어 말뭉치 연구 및 구축 사업팀의 언어
제공자 질문지

수정 이후의 내용을 보면 선택지가 일관되게 제시되어 있는 것을 알 수 있다. 이를 통해 언어 제공자들이 질문지를 더 원활하게 작성할 수 있으며, 후술할 서식의 측면에서도 훨씬 정연함을 보인다.

3) 서식이 통일되지 않은 부분

질문지 중간에 서식이 통일되지 않아 작성에 어려움을 느낄 수 있는 부분들이 존재했다. 특히나 아래에 제시한 부분이 특징적이다.

5. 현재 사는 곳: _____ 거주 기간(예:08-현재): _____

6. 현재까지 살았던 장소들

_____년부터 _____년까지 거주지 _____(지역명)

_____년부터 _____년까지 거주지 _____(지역명)

<그림 17> 수정 전 - 2017년 구축 지침에 따른 언어 제공자 질문지

5번 질문은 현재 거주 지역에 관한 질문이며, 6번 질문은 현재까지 거주했던 지역들을 묻는 질문이다. 이들은 모두 지역 수어를 포착하기 위한 연구에서 중요하게 다뤄질 수 있는 질문들이다.

그러나 요구하는 정보가 동일한 범주인 것에 반해서 5번의 답변 서식과 6번의 답변 서식이 달라서 언어 제공자들이 작성하는 데에 혼란을 느낄 수 있는 측면이 있다. 이에 본 연구팀은 6번의 답변 서식을 고려하여 5번의 답변 서식을 수정하였다.

5. 현재 사는 곳 * 6번을 고려하여 서식 통일함

_____년부터 현재까지 거주지 _____(지역명)

6. 현재까지 살았던 장소들

_____년부터 _____년까지 거주지 _____(지역명)

<그림 18> 수정 후 - 2018년 한국수어 말뭉치 연구 및 구축 사업팀의 언어 제공자 질문지

이 외에 오타나 일련번호 등의 오류 교정, 기타 서식의 일치 등을 살펴보았고, 최종적으로 완성된 질문지는 본 보고서의 <부록>에 제시하도록 한다.

4) 추후 연구에서 논의가 더 필요한 부분

말뭉치 구축 사업이 이번에 처음으로 하는 사업이 아니기 때문에 현 연구팀에서 선불리 수정하기 어려운 부분들이 존재했다. 왜냐하면 이미 한국수어 자료 통합 지원 시스템에 기존 언어 제공자들의 메타 데이터가 등록이 되어 있기 때문에, 이번 언어 제공자들의 메타 데이터도 기존의 것들과 형식적으로 크게 어긋나지 않게 하는 수준에서만 오류를 교정해야만 했다.

다만 현 연구팀에서 수정하지 못한 것들 중에 논의가 더 필요한 부분들이 보인다. 바로 다음과 같은 부분들이다.

13. 아버지와 주된 의사소통 방법

- ☐ 한국수어 ☐ 한국어대응수어(+말) ☐ 제스처 ☐ 말 ☐ 필담
☐ 홈싸인 ☐ 기타 _____

30. 출신학교

출신학교1: _____ 살부터 _____ 살까지 _____ (학교이름)

주된 의사소통 방법:

- ☐ 한국수어 ☐ 한국어대응수어+말 ☐ 한국수어 + 한국어 ☐ 구화
☐ 수화통역사와 함께 ☐ 기타 _____

<그림 19> 추가적으로 논의가 필요한 질문지

‘주된 의사소통 방법’에 대한 답변 선택란의 체계가 서로 다르다는 것이다. 가족과의 의사소통 방법에는 ‘한국어 대응 수어(+말)’이라고 되어 있지만, 출신 학교에서의 의사소통 방법에는 ‘한국어 대응 수어+말’로 표시되어 있다. 괄호 안의 정보가 어떤 의미인지 파악하기 어려운 측면이 있다. 또한 출신 학교에서의 의사소통 방법 중에 ‘한국수어+한국어’가 함께 있는 것도 어색하다. ‘+’ 표시를 통해 정보를 복합적으로 나타내기보다는 ‘복수 선택’을 할 수 있도록 안내하는 방법을 사용하도록 제안하는 바이다.

그리고 ‘한국수어+한국어’에서 이때의 한국어는 구화인지 필담인지도 알 수 없다. ‘홈싸인’이 정확히 무엇을 의미하는지도 파악하기 힘들다. 이런 경우 문해력이 다소 낮은 농인 언어 제공자들이 잘못된 정보를 선택할 수도 있기에 보다 신중한 선택지를 제시하는 것이 바람직하겠다. 추후 이러한 정보들에 대한 논의가 더 필요할 것으로 보인다.

3.2. 질문지 작성의 문제점

원시 말뭉치를 구축하기에 앞서, 최종 수정된 질문지를 통해 언어 제공자들의 메타 데이터들을 수집하였다. 이 과정에서의 몇 가지 문제점을 발견하였다. 차기 연구에서 다음과 같은 사항들에 주의를 기울여야 하겠다.

1) 질문지의 사전 미작성 문제

사전에 언어 제공자들에게 질문지를 배포하고 작성해 오도록 부탁하였으나, 개인정

보 기재에 대한 부담감 때문인지 질문지의 사전 작성이 완벽하게 이루어지지 않은 부분들이 있었다. 그런 까닭으로 불가피하게 촬영 현장에 도착한 이후에 질문지를 작성하게 되었는데, 개인정보에 대해 절대로 오용하지 않겠다는 설명이 충분히 이루어진 이후에는 언어 제공자들도 편안하게 질문지를 작성할 수 있었다.

문해 능력이 다소 낮은 농인들이 있을 경우를 대비하여, 질문지를 작성할 때에는 농인 사회자, 청인 연구진들이 함께 모여서 질문지 작성을 도왔다. 잘 이해가 가지 않거나 어떻게 적어야 할지 고민하는 언어 제공자들에게 작성 방식을 설명해줌과 동시에 중간 중간 이야기를 나눔으로써 친근하고 편안한 분위기를 연출하고자 했다.

2) 질문지 자체의 오류

청인의 관점에서 질문지를 바라보고 수정하다 보니, 미처 포착하지 못한 자잘한 오류들이 있었다.⁷⁾ 대표적으로 아래와 같은 항목들이다.

① 농인과의 정기적인 만남

- 문제점: ‘주 n회’로 표시하게 되어 있으나, 하루에 여러 번 만나는 경우
- 대 안: ‘주 n일’로 표시하는 방안. 또는 척도를 제시해주는 방안

② 직장 퇴직 여부

- 문제점: ‘예/아니오’로 체크하게 되어 있는데, 무직자가 ‘아니오’를 체크하는 경우
- 대 안: 현 직장 여부와 직장 정보, 퇴직 여부를 모두 쓰는 방안

‘농인과의 정기적인 만남’을 묻는 질문에 주 20회라는 결과가 나오기도 했다. 하루에 여러 번 만나는 것을 모두 셈한 경우이다. 경우에 따라 하루에 농인을 몇 번을 만나든 1회로 간주하는 농인들도 있었다. 이런 점을 반영하여 추후에는 주 n일, 또는 척도를 제시해 주는 방안을 고려해 볼 수 있다.

다음으로는 ‘직장 퇴직 여부’를 표시하는 항목이었다. 실제 질문지 항목에서 ‘직장 정보’와 관련된 항목은 아래와 같다.

7) 2~3회차 촬영에서 이러한 문제점을 발견하게 되었는데, 이미 작성이 끝난 설문지이기 때문에 중간에 내용을 바꿀 수는 없다고 판단하고 마지막 회차 촬영까지 질문지의 내용은 그대로 유지하였다. 차기 사업을 진행한다면 해당 내용을 반영하여 질문지를 수정해야 하겠다.

36. 직장에서의 작업 형태

☐ 팀에서 일함 ☐ 혼자 일함

37. 직장 퇴직 여부: ☐ 예 ☐ 아니오

[37번에서 '예'라고 대답한 경우] 마지막 직장에 관한 다음 질문에 답하십시오

38. 마지막 직장 정보

직종(직장) _____ 근무 기간 _____년부터_____년까지

<그림 20> 2018년 언어 제공자 질문지; 직장정보

‘직장 재직 여부’를 먼저 묻는 것이 아니라 ‘직장 퇴직 여부’를 먼저 묻고 있다. 그로 인해 ‘무직자’인 경우에도 ‘아니오’에 체크하는 경우가 있었다. 이에 ‘직장 퇴직 정보’에서 얻을 수 있는 정보보다는 직장을 다니거나 다녔었는지를 묻는 것이 더 유효한 정보라고 판단하여, 추후에는 ‘현재 직장 재직 여부와 직종, 퇴직 여부’를 함께 묻고자 한다.

4. 유도 자료 및 설명 영상 제작

수어 언어를 연구하거나 한국수어 말뭉치를 구축하기 위해서는 최대한 자연스러운 상황에서 다양한 수어 어휘나 표현 등을 수집하여야 한다. 이때 다양한 수어 언어 자료를 이끌어내기 위한 그림이나 과제, 질문 등을 유도 자료라고 한다. 그러나 언어 제공자들에게 단순히 유도 자료만 보여준다고 해서 서로 대화를 해나가는 것은 아니다. 유도 자료만 보아서는 그것이 어떤 내용인지 잘 이해하지 못할 수도 있다. 그렇기 때문에 수어로 유도 자료의 내용에 대해서 설명하는 설명 동영상이 필수적으로 제작되어야 한다. 유도 자료 및 설명 영상은 정기 모임마다 수시로 논의하여 수정하였다.

4.1. 유도 자료 제작

1) 유도 자료에 대한 이해 교육 및 수정 논의

한국수어 말뭉치 구축을 위한 촬영에서 유도 자료가 중요한 역할을 하는 만큼, 본격적인 유도 자료 제작에 앞서 연구진 전원을 대상으로 한 수어교육 및 상시교육을 통해 유도 자료에 대한 이해 교육을 가장 먼저 실시하였다. 우선 2017년 개정 한국수

어 말뭉치 구축 지침을 토대로 유도 자료의 목적과 필요성, 종류에 대해서 살펴보면
서 유도 자료에 대한 기본적인 이해를 선행 작업으로 삼았다. 그와 동시에 유도 자료
의 내용을 어떻게 얼마만큼 수정할 것인가, 혹은 수정하지 않을 것인가에 관한 논의
를 진행하였다.

결론적으로 2018년 한국수어 말뭉치 연구 및 구축을 위한 유도 자료는 전반적으로
2015년 한국수어 말뭉치 구축팀이 사용한 유도 자료를 따르되, 연구진 회의를 통해
변경이 필요하다고 생각되는 항목과 지역 수어 채집을 위한 일부 내용만을 수정하는
것으로 결정하였다. 기존 연구팀의 유도 자료를 따르는 까닭으로는 추후 이루어질 지
역 간 수어의 비교를 위해서이다. 동일한 주제로 이야기했을 때 얻어지는 수어 자료
인 만큼, 같은 어휘이지만 수형이나 수동이 다른 풍부한 표현을 얻을 수 있다는 이점
이 있고, 그것이 지역 간 수어의 비교로까지 이어질 수 있기 때문이다.

2) 2015년 제작된 유도 자료 검토

기존에 제작되어 사용되었던 유도 자료의 내용을 대부분 따르기로 결정한 만큼, 해
당 유도 자료의 내용을 면밀히 분석하는 것을 우선 과제로 삼았다. 2017년 유도 자료
의 목록은 다음과 같다.

- A-수어 이름 말하기
- B-준비해 온 농담
- C-다시 이야기하기(애니메이션)
- D-수어 보고 다시 이야기(넥타이)
- E-토론
- F-표지판
- G-특정 사건 묘사
- H-주제 영역
- I-개별 수어의 다양한 표현
- J-농사회 행사
- k-그림 이야기 설명(흥부와 놀부)
- L-옛날 수어와 현재 수어
- M-농학교 이야기

이들 목록 중에서 ‘E-토론’, ‘G-특정 사건 묘사’, ‘J-농사회 행사’의 내용은 현 연구
팀에서 수정해야 할 부분으로 선정하였다. 왜냐하면 ‘J-농사회 행사’의 경우 서울 지
역의 포스터만 제시되어 있었기 때문에, 해당 지역의 실제 농사회 행사 포스터를 제
시하는 것이 좋겠다는 의견이 있었다. ‘E-토론’과 ‘G-특정 사건 묘사’의 경우에는 세

부 주제가 너무 오래된 사건이라는 지적이 있었다. 아래에 해당 과제의 세부 주제를 제시한다.

- E-토론; ① 1종 면허 농인
② 농학교 농인 교사
③ 한국수어 통역사 협회 설립
④ 메르스

- G-특정 사건 묘사; ① 1974년 육영수 여사 피살 사건
② 1995년 삼풍백화점 붕괴
③ 2002년 월드컵 스페인전
④ 2008년 남대문 화재 사건
⑤ 2009년 노무현 대통령 자살 사건
⑥ 2014년 세월호 침몰

이 중 ‘E-토론’ 영역에서 ‘한국수어 통역사 협회 설립’의 경우, 해당 유도 자료에서 설명하고 있는 주제와 기구축된 촬영 영상들에서 언어 제공자들이 이야기하고 있는 내용이 다르다는 것에 착안하여, 다수의 언어 제공자들이 이해하고 이야기하는 방향으로 초점을 바꾸기로 하였다. 다음으로 ‘메르스’는 2015년에 유행한 호흡기 질병으로 당시 감염자가 발견되면서 사회적 파장을 불러 일으켰던 질병이지만, 종식 이후로는 큰 이슈가 되지 못했다고 판단하여 현 시점에 어울리는 다른 토론 주제로 대체하기로 하였다.

‘G-특정 사건 묘사’ 영역에 제시된 여섯 가지의 사건들은 모두 대한민국의 역사에서 중요하게 기록될 만한 사실들이지만, 현 시점인 2018년의 입장에서 보았을 때 10년 이상이 흐른 오래된 사건들이 대다수였다. 따라서 가장 최근에 일어난 ‘세월호 침몰’을 제외하고 전면적으로 사건을 교체하기로 결정하였다.

3) 유도 자료 수정·제작

2018년 한국수어 말뭉치 연구 및 구축팀에서는 기존 한국수어 자료 통합 지원 시스템에 등록된 유도 자료를 바탕으로 새로운 유도 자료를 제작하였는데, 상술한 ‘E-토론’, ‘G-특정 사건 묘사’, ‘J-농사회 행사’의 영역들에서 수정이 이루어졌다. 그들 중 특히나 ‘E-토론’, ‘G-특정 사건 묘사’의 경우, 청인들의 시각에서만 이슈가 되는 것이 아닌 농인들의 시각에서 이슈가 된 사건들을 찾는 것이 담화 소재로서 중요하기 때문에, 사건을 선정하는 과정에서 연구팀 내 농인과 청인들 간의 많은 의견 교류가 있었다.

‘J-농사회 행사’의 경우, 해당 지역의 농사회 행사 포스터를 최대한 찾아서 제시하는 것으로 결정되었다. ‘E-토론’, ‘G-특정 사건 묘사’ 영역에서는 다음과 같은 세부

주제가 선정되었다.⁸⁾

- E-토론; ① 1종 면허 농인
② 농학교 농인 교사
③ 수어통역사 센터 분리
④ 폭염

- G-특정 사건 묘사⁹⁾; ① 2011년 가슴기 살균제 사건
② 2014년 세월호 침몰
③ 2018년 평창 동계 올림픽
④ 2018년 남북 정상회담
⑤ 2018년 러시아 월드컵
⑥ 전북/전남 지역 - 1980년 광주 민주화 운동
강원 지역 - 2018년 고성 산불
경남 지역 - 2017년 포항 지진

‘E-토론’에서 ‘한국수어 통역사 협회 설립’의 경우, 다수의 언어 제공자들이 접근했던 ‘수어통역사 센터 분리’에 관한 주제로 초점을 변경하였다. 또한 앞서 시의성에 맞지 않다고 생각했던 ‘메르스’를 올 여름 꾸준히 이슈가 되었던 ‘폭염’ 문제로 교체하였다.

대대적으로 세부 주제를 수정하기로 결정했던 ‘G-특정 사건 묘사’의 경우는 현재성과 지역성을 고르게 반영하고자 노력했다. 이때 여섯 가지의 사건 중 다섯 가지는 세월호를 포함하여 비교적 최근에 이슈가 된 사건들로 대체하였고, 나머지 한 가지는 지역성을 띠는 사건이 반드시 들어갈 수 있도록 하였다.

전라 지역의 광주 민주화 운동은 매우 오래된 사건이긴 하지만 이 지역에서는 상징적인 의미가 크기 때문에 현재성에 조금 어긋나는 부분이 있더라도 감안하고 선정하였다. 강원 지역의 경우 산불이 자주 나는 지역으로, 2018년 봄에도 고성에 큰 산불이 난 적이 있기에 지역 농인들이 할 이야기가 많을 것으로 예상하였다. 경남 지역의 포항 지진 같은 경우, 2017년 수능이 미뤄질 만큼의 사회적 여파가 컸던 자연재해로서 이야기하기에 좋을 것으로 판단하여 선정하였다. 이와 같이 유도 자료에 지역성을 반영한 이유는 자신 주변의 지역과 관련된 이야기를 함으로써 조금 더 적극적으로 이야기를 진행할 수 있다는 장점이 있다는 것과 더불어 해당 지역만의 특징적인 수어들을 포착하고자 하는 본 사업팀의 의도이다.

이상으로 수정된 유도 자료를 보면 아래와 같다.¹⁰⁾

8) 이전 유도 자료와 바뀐 부분은 굵은 글씨와 밑줄로 표시하였음.

9) 시간 순서로 기재하였음.

10) E-토론의 경우, 그림 자료로 주어지는 것이 아니기 때문에 제시하지 않았다.



<그림 21> 유도 자료 G-특정 사건 묘사 - 공통으로 제시되는 부분



<그림 22> G-특정 사건 묘사 - 지역별로 다르게 제시되는 부분



<그림 23> J-농사회 행사 - 지역별로 다르게 제시되는 부분

4.2. 유도 자료 설명 영상 제작

2018년 한국수어 말뭉치 연구 및 구축팀에서는 기존의 유도 자료 설명 영상을 참고하여, 사회자를 대상으로 유도 자료 설명 영상을 다시 촬영하는 것을 목표로 하였다.

1) 기존 유도 자료 설명 영상 검토 및 대본 작성

2017년 연구팀의 유도 자료 설명 영상을 먼저 검토함으로써, 어떤 방식으로 유도 자료를 설명하는지를 파악하였다. 일반적으로 각각의 과제에 대해 사례를 들어서 유도 자료의 내용을 설명하거나 세부 주제를 설명하는 방식이 사용되었다.

현 연구팀에서는 우선 기존의 설명 영상에 대해 연구팀 내 수화통역사 및 농인들이 직접 설명 영상을 듣고 충분히 이해할 수 있는 내용인지 판단하는 작업을 실시하였다. 이후 영상의 내용을 문서화하여 함께 읽어보면서 설명이 부족한 부분을 표시함으로써, 추후 제작할 유도 자료 영상의 대본 형식으로 작성하였다.

유도 자료 설명 영상 대본의 일부를 제시하면 다음과 같다.

<기존 유도 자료 설명 영상 F-표지판 번역>

표지판 | 표지판. 뭐 그림이나 이런 모양 길거리나 차도나 인도에 있는데요. 이것은 어떤 그림일까요. 전 세계 어디에나 이런 것들이 있죠. 그런데 그림에 답이 쓰여 있지 않아요. 그것들에 대해서 편안하게 이야기를 나누고 표지판의 의미에 대해서도 대화를 나누기 바랍니다.

유도 자료 'F-표지판'의 경우 유도 자료 설명 영상을 본 뒤에 표지판 이미지를 제시하여 언어 제공자들에게 대화를 하도록 유도하는 자료이다. 그런데 기존의 유도 자료 설명 영상은 '뭐, 이런, 이것, 그것'과 같은 지시어를 많이 사용하여 설명하기 때문에, 아직 표지판 이미지를 보지 못한 입장에서는 다소 정보가 부족한 상태에서 설명 영상을 보게 되어 이해가 힘든 부분들이 생길 수 있다. 반면 현 연구팀의 유도 자료 설명 영상 대본을 보면 '표지판'이라는 단어를 지속적으로 언급함으로써 추후에 나올 내용이 표지판이라고 언어 제공자들이 파악하기 쉽게 설명하고 있는 것을 알 수 있다.

<유도 자료 설명 영상(나두O) F-표지판 대본>

표지판 | 표지판이란,

그것은 무엇인가요? 표지판은 교통과 사람들이 다니는 거리 어디서나 잘 보입니다. 원형표지판이 있고 삼각표지판도 있습니다. 표지판이 무엇인가요? 전 세계를 살펴보니 일제히 모두 표지판이 있지만 좀 다른 점이 있습니다. 표지판이 무엇인지 편안하게 대화하고 무슨 뜻인지, 어떤 원인인지에 대해 자세히 A언어 제공자와 B언어 제공자가 대화합니다.

이와 같이 대본을 작성하여 연습함으로써 유도 자료에 대한 이해도를 더 높일 수 있으며, 어떤 방법으로 설명하는 것이 효과적인지 연구진들끼리 지속적으로 피드백을 할 수 있었다.

2) 유도 자료 설명 영상 촬영

유도 자료 설명 영상의 대본이 완성됨에 따라 유도 자료 설명 영상 촬영을 진행하였다. 촬영은 크게 4회에 걸쳐 이루어졌으며, 일정은 다음과 같다.

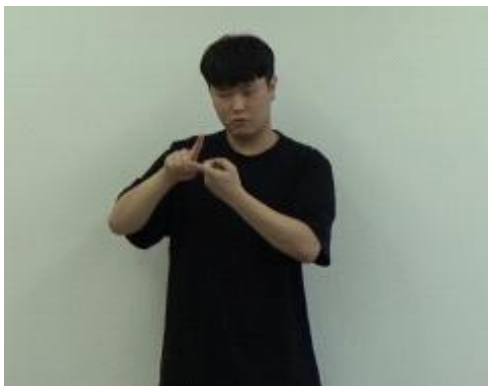
- ① 1차 연습 촬영일 (2018년 7월 14일) - 서일O 연습 촬영
- ② 2차 연습 촬영일 (2018년 7월 30일) - 서일O, 나두O 연습 촬영
- ③ 3차 본 촬영일 (2018년 8월 7일) - 서일O 촬영 완성
- ④ 4차 본 촬영일 (2018년 8월 11일) - 노유O, 나두O 촬영 완성

각 회차별 촬영이 이루어진 뒤에는 연구진 전원이 영상을 돌려보면서 수형이나 수동이 잘 전달되는지, 표정은 괜찮은지, 화면 구도는 괜찮은지를 논의하여 수정할 사항을 제안하였다. 또한 촬영 영상을 대본으로 작성하여 부족한 설명을 어떻게 보충할 것인지에 대한 논의도 함께 이루어졌다.

1차 연습 촬영은 7월 중순경, 사회자 서일O을 대상으로 이루어졌다. 카메라 앞에 서는 것이 처음인지라 다소 어색한 자세와 표정을 고쳐야 한다는 의견이 많았으며, 기존의 설명 영상에 의존하여 내용을 그대로 흉내 내려고 하는 태도에 대한 지적 또한 있었다. 이후 서일O의 설명 영상에 대한 대본을 토대로 하여 설명 내용을 일부 윤색하고, 모든 사회자들에게 연습을 철저히 할 것을 당부하는 선에서 촬영을 마무리하였다.

2차 연습 촬영은 7월 말에 이루어졌다. 사회자 나두O, 서일O을 대상으로 촬영을 진행하였다. 나두O의 경우 카메라 앞에 서 본 경험이 많아서인지 크게 문제되는 부분은 없었다. 서일O의 경우, 1차 촬영에 비해서는 많이 발전한 모습을 보였기 때문에 3차에서 최종 촬영을 하기로 하였다.

3차 본 촬영일은 언어 제공자들과의 촬영 일정을 조율하는 과정 중에서 이루어졌으며, 서일O의 최종 설명 영상을 촬영하였다. 마지막으로 4차 촬영일은 개인 일정상 유도 과제 최종 설명 영상을 촬영하지 못한 사회자 나두O와 노유O의 촬영이 있었다.



<그림 24> 1차 연습 촬영



<그림 25> 3차 본 촬영

결론적으로 유도 자료 설명 영상을 사회자 3인이 모두 찍게 된 것인데, 이는 단순히 사회자의 역할 및 유도 자료에 대한 이해도를 높이기 위한 도구로서의 결과물만은 아니다. 지역별, 성별, 연령별 언어 제공자의 특성에 맞추어서 설명 영상을 제공하기 위한 목적이 있었다. 또한 조금 더 익숙한 얼굴의 사회자가 등장하는 설명 화면을 제시하는 것이 궁극적으로 언어 제공자의 긴장도를 완화시킬 수 있다고 보았기 때문이다.

5. 사회자 선정 및 교육

5.1. 사회자 선정

수어라는 언어 자료의 특성상 말뭉치 수집 시 최대한 수어 언어 자료를 다양하고 풍부하게 유도해내기 위해서는 숙련된 농인 사회자가 필요하다. 2018년 한국수어 말뭉치 연구 및 구축팀에서는 1인이 아닌 3인의 사회자를 선정하였다.

한 쌍의 언어 제공자를 대상으로 촬영, 즉 원시 말뭉치를 구축하는 데 걸리는 시간은 3시간 정도로, 이에 따라 사회자가 느끼는 신체적·정신적 피로도는 매우 크다고 할 수 있다. 촬영 중간에 휴식 시간이 있기는 하지만 10여 분 정도로 매우 짧은 시간이고, 그마저도 사회자들은 언어 제공자들과 촬영 중 포착된 새로운 수어에 관한 이야기를 하거나 촬영 분위기를 매끄럽게 만들기 위해 대화를 시도하기 마련이다. 하루에 두 쌍의 언어 제공자를 대상으로 촬영을 하는 경우도 존재하며, 동시에 여러 팀을 촬영하게 되는 경우도 있을 수 있다.

따라서 1인의 사회자를 선정하였을 경우 사회자가 느끼는 신체적·정신적 부담감을 감소시키고, 동시 촬영을 하게 되는 상황 등을 고려하여 최종적으로 3인의 사회자를 선정하게 되었다. 최종 선정된 사회자 명단은 다음과 같다.

연번	성명	성별	연령	비고
1	나두O	남	50대	농인
2	노유O	여	20대	농인
3	서일O	남	20대	농인

<표 3> 사회자 명단

사회자는 모두 농인으로, 50대 남성과 20대 남성, 20대 여성이다. 사회자의 나이와

성별에 차이를 둬으로써 언어 제공자들의 특성에 따라 다각적으로 대응할 수 있게 하였다. 또한 사회자가 다수이기 때문에 촬영 일정이 변동되더라도 3인 중 1인은 참석하여 사회자의 역할을 할 수 있다는 장점이 있다.

5.2. 사회자 교육

기본적으로 사회자의 역할은 언어 제공자들에게 유도 자료를 제시하고, 언어 제공자에게서 많은 수어 언어 자료를 이끌어내는 것이다. 이러한 과정에서 언어 제공자가 유도 자료를 이해하지 못하거나 대화가 자연스럽게 진행되지 않을 수 있는데, 그때 언어 제공자의 대화를 촉진하고 유도 자료에 대한 내용을 보충 설명하는 것이 사회자의 또 다른 역할이다. 따라서 사회자 역할을 잘 수행할 수 있도록 교육하는 것이 중요하다.

사회자를 위한 교육 내용은 다음과 같다.

- ① 한국수어 말뭉치 교육
- ② 유도 자료의 내용 및 목적 교육
- ③ 대화 촉진 전략 및 유도 자료별 균형적 대화 조절 전략 교육
- ④ 특징적인 지역 수어 산출 유도 교육

농인들의 경우 한국수어에 대해서는 잘 알고 있지만 말뭉치와 같은 일반언어학적인 내용에 대해서는 다소 이해가 부족한 편이다. 그리하여 상시교육을 통해 말뭉치가 무엇인지, 말뭉치가 왜 중요한지, 말뭉치 구축의 의의가 무엇인지를 깨닫게 하는 것이 사업 착수와 동시에 진행되었다.

유도 자료와 관련된 교육은 이미 이전 단계에서 유도 자료를 분석하고 설명 영상을 제작하는 과정에서 유도 자료의 순서나 내용에 대한 이해를 어느 정도 마친 상태이기 때문에 매우 수월하게 이루어졌다.

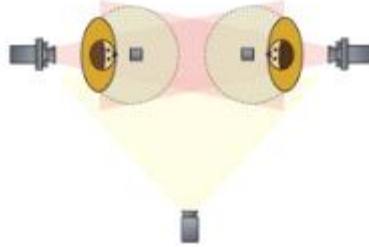
그 외에 언어 제공자들이 긴장하여 대화가 단절되었을 경우 자연스럽게 대화를 이끄는 전략을 포함하여, 한 명의 언어 제공자에게만 또는 하나의 유도 자료로만 대화가 치우치지 않게 조절하는 전략에 대해 교육하였다. 더불어 지역 균형적 말뭉치 구축을 위해 특색 있는 지역 수어를 산출하기 위한 전략도 함께 학습하였다.

6. 원시 말뚝치 수집을 위한 촬영

6.1. 촬영 환경 조성

선정된 언어 제공자를 대상으로 원시 말뚝치 영상을 촬영했다. 촬영은 2017년도 한 국수어 말뚝치 구축 지침에 의하여 환경을 조성하여 진행하였다. 지침의 촬영 환경은 다음과 같다.

- 면적은 32m^2 이상
- 외부 빛을 차단할 수 있는 공간
- 3대 이상의 카메라와 그림자가 생기지 않는 수준의 조명 설치가 가능한 공간



<그림 26> 촬영 예시

1) 촬영 공간

본 사업을 위해 전북대학교의 협조로 전북대학교 학군단 406호에 촬영실을 구축했다. 촬영실의 실측 면적은 29m^2 로 이는 지침에서 제시하고 있는 32m^2 의 기준에 미달하지만 촬영장을 구축하고 테스트 촬영을 진행한 결과 촬영 공간의 면적보다 얻어지는 결과물의 품질이 더 중요한 결정 근거가 된다고 판단하여 지침상의 면적보다 작은 공간에서 촬영을 진행했다.

2) 촬영 배경(스크린)

촬영 환경 구축은 언어 제공자와 사회자 뒤편에 크로마키 스크린을 거치대를 사용하여 설치했다. 색은 지난 사업과 동일하게 청색 계열로 설치했다. 지난 사업은 촬영장 구축 시 스크린 천을 벽면에 부착했다. 이는 구축된 촬영장을 구축이 끝날 때까지 해체할 필요가 없이 진행할 수 있었기 때문이다. 그러나 본 사업에서는 전북 지역 이외에 타 지역 촬영이 예정되어 있었기 때문에 최초 기자재 구입 단계에서 촬영실의 이동을 감안하여 부착형이 아닌 거치형으로 크로마키 스크린을 구비했다.

아울러 거치형 크로마키 스크린을 이용함으로써 얻을 수 있는 이득은 촬영 가능 면적의 축소이다. 앞서 1)의 촬영 공간에서 언급한 바와 마찬가지로 지침 규정보다는

촬영 결과물이 중요하기 때문에 길이 3m 길이의 크로마키 스크린 3대로 ‘ㄷ’자 모양의 촬영 공간을 구성하였고 예비 촬영을 통해 촬영 결과물을 검토했다. 이를 통해 기존 지침보다도 작은 공간에서도 촬영을 진행할 수 있다는 결과를 얻을 수 있었다.



<그림 27> 2018년 촬영장 사진



<그림 28> 2015년 촬영장 사진

3) 캠코더

촬영을 위한 카메라는 HD 이상 화질과 크기, 배터리 지속 시간, 조작성 등을 고려하여 Sony의 CX-450모델을 채택했다. 무엇보다 캠코더의 크기를 고려했다. 그 이유는 실제 촬영 때 캠코더를 완전히 감출 수 없기 때문에 부담스럽지 않은 크기로 된 캠코더를 활용해야 위화감을 줄일 수 있을 것이라는 판단에 의해서였다.

캠코더는 크로마키 스크린 거치봉 위 2m 높이에 클립으로 거치하여 앉아있는 언어 제공자를 대각선 위에서 촬영하도록 설치했다. 2m 위에 캠코더가 위치하기 때문에 언어 제공자들 사이에 자연스럽게 대화가 진행된다면 캠코더를 의식하지 못하고 촬영이 진행될 수 있도록 했다. 아울러 CX-450 모델은 원격으로 실시간 보기(Live view)를 지원하기 때문에 촬영 실황을 스마트폰이나 PC로 모니터링하면서 촬영을 진행하여 촬영 중 돌발 상황을 대비하였다.

4) 조명

조명의 경우 2017년도 지침에 의거하여 5개의 조명을 구비하였으나 전북대학교 촬영실의 경우 천장에 설치된 LED 전등이 광원 지향성 없이 설치되어 있었고 이후 강원도 촬영에서도 충분한 광량을 확보할 수 있는 조명이 구비되어 있어 촬영 시 조명을 사용하지 않았다. 이는 충분한 촬영 환경이 구성된다면 촬영에 위화감을 줄 수 있는 인위적 요소를 최대한 배제하는 편이 바람직할 것이라는 논의에 의한 것이다.



<그림 29> 2018년 구축 영상 예시



<그림 30> 2015년 구축 영상 예시

위 촬영 결과물을 보면 2018년 구축 영상은 별도 조명을 설치하지 않은 것이고 2015년 구축 영상은 조명을 설치한 영상이다. 두 영상을 비교한 결과 결과물로서의 차이는 없다.

6.2. 외부 촬영실 구축(강원도)

본 사업은 기획 단계부터 여러 지역의 수어 말뭉치 구축을 목표로 했었고 이에 따라 크로마키 스크린을 거치형으로 구비하고 유도 자료 제시용 PC도 노트북을 사용하는 등 촬영 기자재의 이동성을 고려했다.

수어 말뭉치 수집 지역은, 전남, 전북, 경남, 강원 총 네 지역이었다. 언어 제공자의 편의를 위해서는 각 지역 촬영 때마다 촬영실을 구축하는 편이 유리하나 그럼에도 기존 구축된 전북대 촬영실을 최대한 활용하는 것이 촬영 결과물의 균질성을 확보하는 데에 유리하기 때문에 전남, 전북, 경남 언어 제공자 11쌍은 전북대학교에 구축된 촬영실에서 촬영을 진행했다. 그러나 강원도의 경우는 이동 거리가 왕복 600km가 넘기 때문에 전북대학교 촬영장에서 언어 제공자의 촬영이 불가능했다. 이 때문에 강원도는 해당 지역에 촬영실을 구축하여 진행했다.

1) 장소 선정을 위한 답사

8월 6일 한국농아인협회 강원도지부와의 사전 협의 때 촬영실을 구축하기 위한 장소답사를 진행하였다. 답사 장소는 한국농아인협회 강원도협회 대회의실과 춘천시에 위치한 강원도시청자미디어센터였다. 강원도 촬영은 일정상 문제로 동시에 두 팀을 촬영해야 했기 때문에 공간 구성에 대해 논의를 거쳐 촬영실을 한국농아인협회 강원도협회 대회의실로 결정했다. 시청자미디어센터 촬영실은 촬영 기자재와 조명등이 모두 구비되어 있었으나 촬영 시 농인들이 위화감을 느낄 수 있는 우려가 있어 평소 농인들에게 익숙한 공간인 회의실로 공간을 결정했다.

강원도 촬영 일정 협의 이후 언어 제공자의 일정상의 문제로 원주에서 한 팀을 촬영해야 했다. 원주는 강원도협회의 협조로 원주시지부의 교육실을 빌려 진행했다.

2) 실제 외부 촬영실 구축

강원도에서 총 4쌍을 촬영했다. 장소별로 강원도농아인협회 원주시지부 교육실에서 1쌍, 한국농아인협회 강원도협회 회의실에서 3쌍을 촬영했다. 촬영은 8월 17일부터 8월 18일까지 이틀에 걸쳐 진행됐고 8월 17일은 원주, 8월 18일은 춘천에서 촬영을 진행했다.



<그림 31> 강원도 원주 촬영 영상

위는 원주시지부에 구축한 촬영실과 촬영한 영상이다. 촬영실을 구축한 원주시지부 교육실은 공간 문제로 조명을 추가로 설치할 수 없어 자연광과 교육실 내부의 전등을 활용해서 촬영을 진행했다. 예비 촬영 결과 수어와 언어 제공자의 표정을 확인하는데 문제가 없어 촬영을 진행했다.



<그림 32> 강원도 춘천 촬영 영상

위는 춘천시에 위치한 한국농아인협회 강원도협회에서 촬영한 영상이다. 춘천 촬영의 경우 언어 제공자 일정과 촬영 일정 때문에 동시에 두 팀을 촬영해야했다. 이에 따라 크로마키 스크린을 2면 더 구비하여 회의실 내부에 촬영실을 H자 모양으로 배치해서 촬영 공간이 등을 지는 형태로 배치했다. 추가로 필요한 캠코더 3기는 전북대학교 인문대학의 인문역량강화사업단(CORE)에서 대여하여 동일한 기종으로 촬영을 진행했다.

수어 말뭉치 촬영은 언어 제공자의 수어를 촬영하기 때문에 서로간의 소리 간섭이 없는 촬영이라는 점을 감안하여 동시 촬영을 진행하였고 위와 같은 촬영 결과물을 얻을 수 있었다.

6.3. 촬영의 실제

본 사업은 총 30인의 언어 제공자를 대상으로 하여 총 15쌍의 영상을 촬영했다. 촬영은 2018년 8월 10일부터 9월 8일까지 약 한 달간 진행됐고 한 팀을 촬영할 때 평균적으로 4시간 정도가 소요됐다.

촬영은 한 제공자 쌍당 180분의 촬영 시간을 확보하는 것을 기본적인 목표로 하여 13개의 유도 자료를 모두 소화할 수 있도록 진행했다. 이를 위해 사회자로 하여금 13번까지 유도 자료를 가급적 모두 진행하되 180분의 촬영 시간을 맞추기 위해 시간 안배에 신경을 써 언어 제공자의 대화를 진행하도록 했다.

이를 통해 진행된 촬영 내역은 다음과 같다.

연 번	촬영일	촬영 장소	언어 제공자		사회자	촬영 시간	언어 제공자 쌍 번호
			A	B			
1	2018.08.10.	전북대학교 촬영실	언어 제공자63	언어 제공자64	서일O	3:05:15	전북 지역 1번 쌍
2	2018.08.13.	전북대학교 촬영실			서일O		
3	2018.08.16.	전북대학교 촬영실			노유O		
4	2018.08.17.	원주시지부 교육실	언어 제공자69	언어 제공자70	나두O	2:52:55	강원 지역 1번 쌍
5	2018.08.18.	강원도협회 회의실	언어 제공자71	언어 제공자72	나두O	2:37:04	강원 지역 2번 쌍
6	2018.08.18.	강원도협회 회의실	언어 제공자73	언어 제공자74	노유O	3:00:11	강원 지역 3번 쌍
7	2018.08.18.	강원도협회 회의실	언어 제공자75	언어 제공자76	서일O	2:51:10	강원 지역 4번 쌍
8	2018.08.25.	전북대학교 촬영실	언어 제공자77	언어 제공자78	노유O	2:59:48	전북 지역 4번 쌍
9	2018.08.25.	전북대학교 촬영실	언어 제공자79	언어 제공자80	나두O	2:43:19	전남 지역 4번 쌍
10	2018.08.28.	전북대학교 촬영실	언어 제공자81	언어 제공자82	서일O	3:42:41	전북 지역 5번 쌍
11	2018.08.30.	전북대학교 촬영실	언어 제공자83	언어 제공자84	노유O	3:14:54	전남 지역 2번 쌍

12	2018.08.31.	전북대학교 촬영실	언어 제공자85	언어 제공자86	노유O	3:01:13	경남 지역 1번 쌍
13	2018.08.31.	전북대학교 촬영실	언어 제공자87	언어 제공자88	서일O	2:56:02	경남 지역 2번 쌍
14	2018.09.08.	전북대학교 촬영실	언어 제공자89	언어 제공자90	서일O	2:35:00	전남 지역 3번 쌍
15	2018.09.08.	전북대학교 촬영실	언어 제공자91	언어 제공자92	나두O	2:55:09	전남 지역 4번 쌍

<표 4> 2018년도 말뭉치 영상 촬영 목록

총 15쌍의 촬영을 진행하였고 촬영 시 사회자, 언어 제공자A, 언어 제공자B로 나누어 촬영을 진행하였고 중간에 1회의 휴식 시간이 있었기 때문에 전체 촬영 결과물은 90개의 파일이다. 파일을 720p화질로 촬영하였고 전체 용량은 160기가바이트였다.

촬영 과정에서 정확한 시작과 끝을 맞추지 않고 촬영 개시 이전에 녹화를 시작하고 촬영이 끝나고 정리가 된 뒤에 녹화를 끝냈기 때문에 촬영 결과물 원본 파일을 만들어내기 위해 세 파일의 시간대를 동기화시키고 영상 전후와 중간 쉬는 시간을 잘라냈다. 해당 작업은 외부 전문가에게 의뢰하여 진행하였고 프로그램은 Adobe사의 Premiere CC 2017을 사용했다.

위 작업을 완료한 파일은 총 45개로 전체 시간은 44시간 49분 45초였다. 그런데 해당 과정을 진행하던 중 최초 저장 단계에서 파일 1개가 소실된 것을 확인했다. 소실된 파일은 전북지역 3번 쌍의 강성O의 영상으로 소실 경위를 파악한 결과 캠코더 외장메모리로 파일이 저장되고 이 외장메모리를 PC에 인식시켜 파일을 저장했는데 옮기는 중 메모리에 이상이 생겨 손상된 파일인 채로 저장되었다는 것을 확인했다.

촬영 초기에 발생한 일이고 바로 다음날 강원도 촬영이 예정되어 있어 확인 절차가 미흡한 상태에서 파일 원본이 삭제되었다. 해당 문제를 파악한 뒤 처리 방안을 논의했고 결정 사항은 아래와 같다.

- ① 전체 촬영 시간이 41시간 28분 13초로 계획된 구축 시간인 40시간을 초과하기 때문에 추가 촬영 없이 해당 쌍의 영상을 무효 처리한다.
- ② 실제 촬영은 진행됐고 상대방 언어 제공자와 사회자 촬영 영상이 존재하기 때문에 촬영수당은 정상적으로 지급한다.

처리 방안 결정 단계에서 동일 언어 제공자를 재촬영하는 것은 이미 촬영을 통해 유도 자료와 촬영 절차를 숙지했기 때문에 최초 촬영 때와는 다른 가공된 언어 자료로서의 수어를 사용할 가능성이 제기되어 배제하였다. 아울러 본 사안을 통해 캠코더 메모리 카드를 여벌로 준비하여 파일 이동 이후에도 원본을 일정 기간 동안 보존하는 방향으로 촬영 파일 보관 방식을 개선하였다.

7. 추가 촬영

총 15쌍을 대상으로 진행한 촬영 내용과 언어 제공자 질문지를 검토하는 과정에서 언어 제공자 중 언어 제공자 조건에 부합하지 않는 언어 제공자를 확인하였다. 이를 원시 말뭉치 구축 내역에 포함할 것인지에 대한 논의를 진행하였다.

해당 촬영 결과물은 2018년 8월 13일 촬영분으로, 이 중 언어 제공자 66번이 농학교를 다니지 않았다는 문제가 제기되었다. 해당 언어제공자는 농인 가족이 많아 자연스러운 농식 수어를 사용하고 있을뿐더러, 지역 커뮤니티에서도 활발한 활동을 하고 있기 때문에 언어 제공자로서 부족함이 없다는 의견들도 있었다. 그러나 농학교 졸업 여부가 언어 제공자 선정 조건에 중요한 요소로 작용한다는 점과 기존에 이미 구축된 지침을 준수하는 것이 언어 자료와 말뭉치의 객관성과 신뢰성을 확보하는 방법이라는 판단하에 해당 촬영분을 폐기하고 추가 촬영을 계획했다.

폐기한 촬영분이 전북 지역 촬영이기 때문에 전북 지역 언어 제공자를 2명 추가 섭외하여 2018년 10월 6일 촬영을 진행했다. 섭외 절차는 농인 커뮤니티를 통한 섭외로 이전에 촬영한 언어 제공자의 추천을 통해 선정했다. 아울러 지침상의 언어 제공자 선정 기준을 검토하여 전부 적용했고 언어 제공자 질문지를 수령한 결과 내용에 이상이 없었다.

연 번	촬영일	촬영장소	언어 제공자		사회자	촬영시간	언어 제공자 쌍 번호
			A	B			
16	2018.10.06.	전북대학교 촬영실	언어 제공자93	언어 제공자94	노유O	03:02:56	전북 지역 6번 쌍

<표 5> 추가 촬영 내역

말뭉치와 언어 제공자 번호는 파일 누락과 언어 제공자 재선정을 표시하기 위해 기존에서 이어지는 번호로 설정했다. 결과적으로 본 연구를 통해 총 16쌍의 농인 언어 제공자를 확보하여 촬영을 진행하였고 이 중 말뭉치로서 구축되는 한국수어 언어 제공자는 총 14쌍이다. 전체 촬영 시간은 41시간 12초이다.

8. 원시 말뭉치 자료 처리

8.1. 한국수어 말뭉치 구축 자료 분석

본 연구를 통해 2018년에 구축한 말뭉치 영상은 1차 번역과 향후 전사를 위해 유도 자료별로 분절하여 관리된다. 아래는 총 14개의 원시 말뭉치 영상을 유도 자료별

로 분절하여 촬영된 시간을 정리한 내용이다.

번호	1번	2번	3번	4번	5번	6번	7번	8번	9번	10번	11번	12번	13번
전북 1번 쌍	0:00:46	0:05:08	0:06:31	0:00:50	0:26:14	0:12:57	0:22:46	0:50:27	0:18:15	0:02:44	0:01:20	0:08:52	0:07:14
강원 1번 쌍	0:01:45	0:03:54	0:19:37	0:02:36	0:26:07	0:12:33	0:18:07	0:14:27	0:11:17	0:04:29	0:06:03	0:09:44	0:16:48
강원 2번 쌍	0:01:14	0:04:32	0:16:07	0:00:58	0:21:27	0:09:34	0:11:57	0:11:32	0:05:50	0:08:41	0:05:54	0:07:06	0:30:32
강원 3번 쌍	0:00:57	0:04:39	0:16:31	0:02:39	0:16:30	0:13:25	0:16:19	0:34:19	0:05:35				
강원 4번 쌍	0:00:39	0:09:47	0:09:33	0:00:47	0:15:22	0:07:51	0:20:45	0:32:10	0:08:07	0:08:24	0:05:06	0:08:42	0:17:05
전북 4번 쌍	0:01:03	0:09:15	0:12:30	0:00:55	0:19:16	0:15:39	0:26:07	0:40:43	0:03:47	0:02:41	0:02:31	0:05:00	0:21:25
전남 1번 쌍	0:01:48	0:04:22	0:14:26	0:01:54	0:19:43	0:19:09	0:15:38	0:05:21	0:07:15	0:05:51	0:01:57	0:16:11	0:18:15
전북 5번 쌍	0:01:17	0:10:59	0:09:36	0:01:35	1:25:46	0:10:17	0:11:52	0:16:01	0:10:42	0:07:32	0:04:33	0:09:36	0:12:32
전남 2번 쌍	0:02:04	0:09:28	0:17:14	0:03:03	0:38:40	0:20:54	0:27:39	0:20:26	0:05:17	0:05:04	0:06:00	0:05:54	0:14:04
경남 1번 쌍	0:01:08	0:15:56	0:10:23	0:01:07	0:38:33	0:07:48	0:18:59	0:35:31	0:06:15	0:07:27	0:02:40	0:06:06	0:06:19
경남 2번 쌍	00:03:51	0:04:29	0:12:31	0:01:04	0:56:27	0:18:29	0:12:42	0:17:35	0:05:53	0:02:46	0:03:48	0:04:56	0:12:33
전남 3번 쌍	0:00:41	00:03:28	00:09:10	00:16:44	00:01:01	00:27:56	00:14:08	00:40:26	00:03:00	00:07:55	00:04:32	00:01:23	00:06:07
전남 4번 쌍	0:00:45	0:09:10	0:11:54	0:01:06	0:20:06	0:20:43	0:21:05	0:23:33	0:09:25	0:09:53	0:04:15	0:09:40	0:08:16
전북 6번 쌍	0:00:41	0:07:04	0:10:42	0:00:54	0:31:58	0:08:47	0:22:22	0:26:46	0:05:23	0:04:56	0:03:45	0:11:14	0:26:57

<표 6> 전체 촬영 영상 중 유도 자료별 촬영 시간

위 시간은 원본에서 유도 자료가 제시되는 시간과 유도 자료와 관련 없는 대화들을 제외하고 편집된 시간이다. 이를 각 유도 자료별 시간의 총 합과 평균 시간으로 정리하면 다음과 같다.

유도 자료	전체 시간	평균 시간
1번(수어 이름 말하기)	0:14:48	0:01:08
2번(농담)	1:42:11	0:07:18
3번(Tweety&sylvester)	2:56:45	0:12:37
4번(넥타이 다시 이야기)	0:36:12	0:02:35
5번(토론)	6:57:10	0:29:48
6번(표지판)	3:26:02	0:14:43
7번(사건)	4:20:26	0:18:36
8번(주제)	6:09:17	0:26:23
9번(개별 수어의 다양한 표현)	1:46:01	0:07:34
10번(농사회 행사)	1:18:23	0:06:02
11번(흥부 놀부)	0:52:24	0:04:02
12번(옛날 수어)	1:44:24	0:08:02
13번(농학교)	3:18:07	0:15:14

<표 7> 유도 자료별 촬영 시간

이를 통해 확인할 수 있는 점은 5번 유도 자료인 토론이 가장 긴 언어 자료를 유도했고 주제영역, 특정 사건 순으로 많은 언어 자료를 유도해냈다. 그리고 농사회 행사 유도 자료로 생각보다 긴 언어 자료를 유도할 수 없다는 점을 보면 이번 연구에 참여한 언어 제공자는 농사회에서 진행되는 공식적인 행사에 참여율이 낮다고 유추할 수 있다.

언어 제공자 표본의 수가 작고 변수가 될 수 있는 요인들이 많기 때문에 이 결과만 가지고 유도 자료의 효용성 전체를 판단할 수는 없지만 대체적으로 자연스러운 대화를 유도하는 유도 자료들이 많은 언어 자료를 이끌어낼 수 있다는 점을 확인할 수 있다. 그리고 언어 자료 유도 시간이 짧은 편인 유도 자료를 중심으로 유도 자료 수정 방안을 고려하는 편이 향후 말뭉치 수집에 더 큰 도움이 될 것이란 점을 알 수 있다.

8.2. 원시 말뭉치 1차 번역 자료 구축

본 연구는 기획단계에서 연구 기간 동안 촬영한 영상의 1차 번역 자료를 연구 결과물과 함께 제공하는 것을 그 목표로 삼았다. 그 이유는 다음과 같다.

- ① 한국수어가 지역별로 그 쓰임에 차이가 있다.
- ② 한국수어는 상황 의존적 언어이기 때문에 대화 상황에 참여하지 않은 제3자가 대화 내용을 온전히 이해하기란 쉽지 않다.
- ③ 구축된 말뭉치 번역이 원래 의미와 다르게 해석될 경우 전체 말뭉치 신뢰도에 문제가 생긴다.

언어 제공자 한 쌍이 대화를 진행한 경우 실제 대화 참여자는 언어 제공자 두 명이지만 사회자가 대화 주제를 주도하거나 직접 대화에 참여하기 때문에 사회자 역시 간접적인 대화 참여자로 볼 수 있다. 본 연구는 이와 같은 사회자의 역할을 주목하여 원시 말뭉치 촬영 자료에 대한 1차 번역을 사회자를 통해 진행했다.

1차 번역은 13개의 유도 자료로 분절된 전사영상을 Elan을 이용하여 번역을 진행하였다. 번역이 1차 번역에 해당하기는 하지만 번역의 단위는 전사 영상의 번역과 동일한 수준으로 수어 화자의 담화 단위 정도로 분절하여 번역을 진행했다. 아울러 전사 영상의 번역과 마찬가지로 번역자와 감수자가 역할을 나누어 번역된 자료를 감수하여 최종 번역본을 제작해서 번역 결과물의 신뢰성을 확보했다.

9. 한국수어 원시 말뭉치 구축 지침 정비 및 제언

9.1. 한국수어 원시 말뭉치 구축 지침 정비

원시 말뭉치 구축은 구축 지역 선정, 언어 제공자 선정, 언어 제공자 질문지 수정, 유도 자료 제작, 사회자 선정 및 교육, 원시 말뭉치 촬영, 원시 말뭉치 자료 가공 및 한국수어 자료 통합 지원 시스템 탑재의 과정을 거쳐 이루어졌다. 이러한 과정에서 본 연구팀이 겪은 원시 말뭉치 구축의 문제점을 제시하고 해결 방안을 제안하고자 한다.

1) 언어 제공자 넘버링(numbering) 및 언어 제공자 지역 구분

본 사업은 2015년 이후 두 번째로 한국수어 말뭉치 영상을 구축하였다. 2015년의 한국수어 말뭉치는 서울과 경기 지역의 농인들을 대상으로 구축되었는데, 이때 말뭉치 영상의 경우, ‘서울 지역 1번 쌍’과 같은 이름을 부여받았다. 그러나 2018년에 이루어진 본 사업의 경우, 지역수어를 구축하는 것을 목표로 하였기 때문에 반드시 ‘OO 지역’이라는 명칭을 확정해야만 했다. 이때의 지역 명칭을 행정 구역에 따를 것인지 등에 대한 논의가 필요했다. 2015년 구축된 촬영 영상의 경우 서울과 경기 지역을 묶어 ‘서울 지역’이라고 명명하였다. 이에 우리는 지리적으로 인접하여 같은 생활권이라고 여겨질 수 있는 지역을 하나로 묶고자 한다. 즉, 광역시의 경우 인접한 도 단위의 행정 구역에 포함하는 것으로 지역 구분에 대한 지침을 새로 추가하였다. 또한 앞으로의 연구에서 말뭉치 구축이 더욱 확대될 것을 고려하여 언어 제공자 넘버링에 관한 내용을 지침 또한 제시하였다. 이러한 지침은 언어 제공자를 더욱 효율적으로 관리할 수 있다는 측면에서 의의가 있다고 할 수 있겠다. 이에 따라 추가된 지침을 제시하면 다음과 같다.

<수정 전>

해당 지침 없음

<수정 후>

- ① 언어 제공자는 촬영된 순서에 따라 번호를 매긴다. 동일한 쌍에서는 언어 제공자A가 앞선 번호를 받는다.
- ② 언어 제공자 쌍은 촬영 지역 별로 순서에 따라 번호를 매긴다. 언어 제공자 쌍의 지역 구분은 다음과 같다.

지역명	포함 지역
서울	서울, 경기, 인천 등 수도권
강원	강원도 전체
충북	충청북도 전체
충남	충청남도 및 세종, 대전 지역
경북	경상북도 및 대구 지역
경남	경상남도 및 부산, 울산 지역
전북	전라북도 전체
전남	전라남도 및 광주 지역
제주	제주도 지역 전체

<표 8> 한국수어 말뭉치 구축 지역 구분

2) 촬영실 공간 구성의 기본 전제

기존의 2017년 촬영실 구축 지침에서 고정된 촬영실을 기점으로 연구진과 언어 제공자들의 동선을 조정했다는 점은 균질성 있는 한국수어 말뭉치 자료의 구축이라는 점에서 당연히 고려되어야 하는 부분이다. 그러나 향후 어떠한 환경에서 한국수어 영상을 촬영하게 될지 알 수 없다. 2018년부터 말뭉치 구축 지역이 확대됨과 동시에 지역 수어 말뭉치의 수집이 이루어지면서 거점 촬영실 외에 이동식 촬영실의 구축이 요구된 것이 바로 그것이다. 그렇기에 다양한 촬영 환경에 대비할 수 있도록 새로운 구축 지침이 필요하다. 특히 한국수어 원시 말뭉치 구축은 동영상 형태로 구축되기 때문에 촬영 환경과 저장에 큰 영향을 미친다. 이에 따라 본 사업에서는 구축 지침의 일부를 수정하여 말뭉치 수집 지역의 확대에 대응하고자 했다.

기본적으로 원시 말뭉치 구축에서 중요하게 여겨야 할 부분은 언어 제공자와 촬영 결과물의 품질이다. 이에 맞추어 원시 말뭉치 촬영 환경과 관련된 모든 구축 지침이 구성되어야 한다. 따라서 본 사업팀에서는 촬영 환경 조성에 적용되어야 할 기본 전제를 먼저 설정하였다. 추가된 지침 내용인, 촬영실 공간 구성의 기본 전제는 아래와 같다.

<수정 전>

해당 지침 없음

<수정 후>

- ① 고정된 기준보다 예비 촬영을 통한 촬영 결과물을 토대로 촬영실을 구축해야 한다.
- ② 수형, 수위, 수동을 간섭 없이 확인할 수 있어야 하고 언어 제공자의 표정과 입 모양도 확인이 가능해야 한다.
- ③ 언어 제공자로 하여금 위화감을 느낄 수 있는 요소는 촬영 결과물에 간섭하지 않는 범위 내에서 최대한 제외시킨다.
- ④ 언어 제공자에게 편의를 제공할 수 있도록 상황에 맞춰 촬영실 환경을 가변적으로 조절해야 한다.

위의 전제들은 결과적으로 촬영에 고정된 환경 기준을 제시하는 것보다 예비 촬영을 통한 결과물을 토대로 촬영 환경을 구축해야 하고 언어 제공자를 우선적으로 고려해야 한다는 점을 말한다.

3) 촬영실 구축

위에서 새로 설정한 촬영실 구축의 기본 전제에 따라 세부적인 지침을 새롭게 설정하고자 하였다. 중요한 것은 장소의 크기 등을 제한하기보다는 지침만으로 전체적인 구축 환경을 조성할 수 있도록 하는 것이었다. 특히 촬영 결과물과 직결되는 항목들을 자세히 하여 향후 연구에 도움이 될 수 있도록 하였다. 기존 지침에서는 촬영실 조건과 촬영 시스템의 내용이 한데 묶여 있었으나 본 사업팀에서는 촬영실 구축과 촬영 유의사항으로 내용을 분리하여 촬영실 구축에 대한 지침만을 따로 제시하였다. 이에 따라 수정된 지침은 아래와 같다.

<수정 전>

- ① 촬영실의 크기는 32m²(5.3m x 6m) 이상이어야 한다.
- ② 촬영실 창문은 빛을 차단할 수 있는 커튼이나 암막이 필요하다.
- ③ 촬영 중 청인은 촬영실에 들어갈 수 없으므로 촬영 감독과 보조 인력이 촬영을 준비하고 모니터할 수 있는 별도의 공간이 필요하다.
- ④ 언어 제공자들이 오기 편하고 연구진들이 24시간 접근이 가능한 곳이어야 한다.
- ⑤ 두 언어 제공자를 각각 촬영할 수 있는 카메라 2대와 두 언어 제공자와 사회자를 다 촬영할 수 있는 카메라 1대 등 총 3대 이상의 카메라가 필요하다.
- ⑥ 카메라의 메모리 용량은 64GB 이상으로 하는 것이 안전하다.
- ⑦ 카메라의 전원 연결이 중간에 끊어지지 않도록 유의한다.

- ⑧ 카메라는 언어 제공자들이 방해를 받지 않고 대화를 할 수 있도록 최대한 눈에 띄이지 않게 설치한다.
- ⑨ 촬영을 위한 조명, 유도 자료 제시를 위한 컴퓨터와 모니터의 설치는 다음 <그림 67>과 같이 한다.

<수정 후>

- ① **촬영 배경:** 언어 제공자의 수어를 잘 포착할 수 있도록 촬영 배경을 설치한다. 구성한다. 배경은 크로마키 스크린을 벽면에 부착하거나 별도 장치를 통해 거치한다. 색은 청색계열로 한다.
- ② **캠코더:** 720p 이상의 해상도로 촬영될 수 있어야 하며 호환의 문제가 없도록 MP4 포맷으로 촬영해야 한다. 그리고 두 언어 제공자와 전경을 찍을 수 있도록 세 대의 캠코더를 설치하여 촬영을 진행하며 언어 제공자에게 큰 위화감을 주지 않도록 캠코더의 크기에 유의해야 한다.
- ③ **조명:** 촬영 영상을 확인했을 때 그림자가 생기지 않도록 조명을 설치하되 예비 촬영을 통해 유동적으로 설치한다. 조명은 800w 이상의 출력이어야 하며 조명 광의 지향성이 없도록 소프트 박스나 반사판을 활용해야한다. 조명이 촬영 결과물의 품질을 올리는 데에 기여하기는 하지만 언어 제공자에게 위화감을 줄 수 있기 때문에 최소한의 조명만을 설치해야 한다.
- ④ **사회자용 PC:** 언어 제공자에게 유도 자료를 제시하기 위해 사회자용 PC를 설치한다. 두 언어 제공자에게 각각 다른 화면을 보여줄 수 있도록 두 대의 피시를 사회자가 조작할 수 있는 환경을 조성해야 한다.
- ⑤ **유도 자료 제시용 모니터:** 언어 제공자에게 유도 자료를 보여주는 화면을 두 언어 제공자 사이에 각각 설치하고 화면이 서로의 대화를 방해하지 않는 최대 높이에 설치한다. 동일한 화면이 제시될 수 있도록 모니터는 같은 모델을 설치해야 한다.

4) 촬영실 구성의 유의사항

촬영실 조건과 촬영 시스템의 내용이 함께 제시되어 있던 기존 지침에서 촬영 유의사항과 관련된 내용을 분리하였다. 촬영 시 주의해야 할 사항들을 정리하여 추가로 제시하였다. 수정된 지침을 보이면 다음과 같다.

<수정 전>

- ⑥ 카메라의 메모리 용량은 64GB 이상으로 하는 것이 안전하다.
- ⑦ 카메라의 전원 연결이 중간에 끊어지지 않도록 유의한다.

<수정 후>

- ① 거점 촬영실은 접근성이 좋은 장소로 선정해야 한다.
- ② 촬영이 4시간 이상 진행되기 때문에 캠코더의 용량과 배터리 방전에 주의해야 한다.
- ③ 촬영 장비는 가급적 언어 제공자에게 위화감을 주지 않도록 설치해야 한다.
- ④ 촬영 결과물의 균질성이 확보되는 범위 안에서 최대한 언어 제공자에게 편의가 제공될 수 있어야 한다.

9.2. 한국수어 원시 말뭉치 구축 제언

1) 언어 제공자 모집 기준 제언

현재 지침에 수록된 언어 제공자 모집 기준을 일정 부분 조정할 필요가 있다. 본 연구의 언어 제공자 모집 기준을 다시 살펴보면 다음과 같다.

- ① 19세 이상 성인 농인
- ② 한국수어를 모국어(제1언어)로 사용하는 농인
- ③ 농학교를 졸업한 농인
- ④ 조사 지역으로 선정된 행정 구역 내에서 최근 10년 이상 지낸 농인
- ⑤ 농사회에서의 교류가 활발한 농인

그러나 ③번의 ‘농학교를 졸업한 농인’의 경우 재고의 여지가 있다. 앞서도 기술하였듯 농학교라는 공간이 한국수어를 습득하고 사용하는 데에 적절한 공간이라는 것은 부정할 수 없다. 그러나 그러한 입장에서 접근한다면 사실은 농학교에 다녔는지의 여부보다 기숙사 생활을 했는지의 여부를 더 우선시해야 한다. 왜냐하면 ‘농학교를 졸업함’이라는 항목이 내포하고 있는 의미는 궁극적으로는 ‘한국수어를 일상에서 얼마나 많이 접하고 사용해 왔는가’라는 질문과 일치하기 때문이다.

기숙사에서 생활하지 않았거나 농학교에 다니지 않았다고 하더라도, 어린 시절부터 농부모나 농형제 가족으로부터 지속적으로 한국수어에 노출되고 한국수어를 습득한 농인들의 경우, 한국수어를 매우 자연스럽게 습득하고 유창하게 구사할 수 있다. 게다가 고연령대 농인의 경우, 경제적인 사정이나 지역 내 농학교의 부재로 인해 농학교를 경험하지 못한 농인들이 있을 수도 있다.¹¹⁾ 이들이 지역 농사회에서 충분히 한국수어를 익혔음에도 불구하고 언어 제공자로 선정되지 못한다면 아주 중요한 수어 언어 자료를 놓치는 셈이 된다. 본 사업팀에서 당초 언어 제공자로 선정했으나 추후 검토 과정에서 제외시키게 된 언어 제공자 ‘양정O’ 씨와 같은 경우가 바로 이런 경우이

11) 단적인 예로 전북 전주시의 선화학교는 1975년도에 설립되었다.

다. 비록 학력은 무학이지만, 농인 형제들이 많은 탓에 농인 형제들과 어울리며 자연스럽게 한국수어를 익히고 구사해 왔다.

따라서 자연수어를 사용한다고 판단이 되는 상황이라면, 충분히 언어 제공자의 조건을 갖추고 있는 것으로 볼 수 있으며, ‘농학교 졸업 여부’ 등은 상황에 따라 유연하게 적용할 수 있어야 할 것이다. 최종적으로 다음과 같은 언어 제공자 모집 기준을 제시할 수 있다.

- ① 19세 이상 성인 농인
- ② 한국수어를 모국어(제1언어)로 사용하는 농인
- ③ 자연수어를 사용하는 농인
 - 농학교를 졸업한 농인
 - 농부모 및 농형제 등을 통해 어릴 때부터 한국수어에 노출되고 한국수어를 습득한 농인
- ④ 조사 지역으로 선정된 행정 구역 내에서 최근 10년 이상 지낸 농인
- ⑤ 농사회에서의 교류가 활발한 농인

위와 같이 모집 기준을 조정한 뒤, 모집한 예비 언어 제공자들 중에서 농학교 기숙사 생활을 하였거나, 초등학교 이전에 한국수어를 습득한 경우 등을 우선적으로 선발하도록 하는 부가적인 조건을 추가함으로써 언어 제공자를 최종 확정하는 방안을 제안하는 바이다.

덧붙여 지역수어 사용자를 가려내기 위한 ④번 항목의 경우도, 지역에 거주하는 젊은 세대의 경우 일자리나 학업을 위해 타지나 대도시로 나가는 경우가 많기 때문에 ‘조사 지역으로 선정된 행정 구역 내에서 최근 10년 이상 지낸 농인’의 조건에 부합하는 경우를 찾기가 어려운 것이 현실이다. 그런 만큼 거주 기간에 대한 제한을 완화하고 다른 부차적인 조건을 통해 보완하는 방법이 필요할 것으로 보인다.

2) 말뭉치 구축 시간 제언

현재 연구 기간이 단기간인 것에 비해 말뭉치 구축 시간이나 촬영 쌍이 다소 많아 연구자들에게 부담이 된다. 2015년의 말뭉치 구축 과정에서도 18일 동안 30쌍의 촬영을 진행하기에 어려움이 있었다고 밝힌 바 있다(국립국어원, 2017:121). 이에 연구 결과의 질적 제고를 위해서라도 연구 기간 대비 말뭉치 구축 시간을 줄일 필요가 있다.

특히나 본 사업팀은 수어 말뭉치 구축 사업에 처음으로 참여한 까닭에, 장비를 구비하고 장소를 마련하는 것에서부터 시작하여 유도 자료를 제작하고 사회자를 교육하는 등의 기본적인 체제를 확립하는 데에만 한 달 이상이 소요되었다. 거기에 언어 제공자를 모집하고 제반 사항들을 조율하고 준비하는 시간을 포함하여, 사업 시작 두

달만에 겨우 원시 말뭉치 구축, 즉 촬영을 시작할 수 있었다.

촬영은 1쌍 당 3시간여를 진행하는데, 그마저도 시작 전 준비 시간과 대화 중간 쉬는 시간 등을 포함하면 여유롭게 4~5시간 정도를 잡아야 한다. 다시 말해 하루에 많이 찍어봐야 2쌍을 찍는 것이 겨우인 것이다. 거기에 더해 농인들도 직업이 있기 때문에 평일에 촬영을 진행하는 것보다 주말에 촬영을 진행하는 것을 더욱 선호한다. 현실적으로 촬영할 수 있는 날짜가 굉장히 제한적이고, 촬영에 소요되는 기간은 그만큼 더 길어질 수밖에 없다. 그에 더해 원시 말뭉치를 후처리하여 분절하고 1차 번역에 걸리는 시간은 촬영에 소요되는 시간만큼, 혹은 그보다 더 많이 소요된다.

다시 말해 언어 제공자 15쌍, 총 40시간의 말뭉치를 구축하는 것은 단기간의 사업에서 수행하기에는 다소 무리인 양이라고 판단이 된다. 한국수어 말뭉치를 구축하는 일은 매우 중요한 일이며 현재의 연구 내용이 후속 사업에 상당한 영향을 미칠 수 있는 만큼 급하게 자료를 구축하기보다는 시간적인 여유를 가지고 최대한 질 좋은 말뭉치를 구축하는 방향으로 나아가야 할 것이다.

Ⅲ. 한국수어 말뭉치 전사

한국수어 말뭉치 자료는 2015년부터 구축되기 시작하여 2017년까지 약 300분의 기본 전사가 진행된 상태이다. 그리고 본 사업을 통해 추가적으로 약 80분의 전사를 진행하였다. 말뭉치 구축은 2017년도에 개정된 한국수어 말뭉치 전사 지침을 기반으로 진행하였으며, 전사자 자질 향상을 위한 워크숍과 상시 교육을 병행하였다.

1. 한국수어 말뭉치 전사 기본 사항

2017년도 한국수어 말뭉치 전사 지침에서 제시한 전사 기본 사항은 다음과 같다(국립국어원, 2017: 216-217).

- 1) 전사를 위해 엘란(Elan) 프로그램을 사용한다.
- 2) 전사를 위한 전사자 교육을 실시한다.
- 3) 전사의 일치도를 높이기 위해 공동 전사를 원칙으로 한다. 공동 전사가 어려운 경우 최소한 처음 5회 이상은 공동 전사를 실시한다.
- 4) 전사의 일치도를 높이기 위해 주 1회 이상의 전사 회의를 실시한다.
- 5) 전사 토큰 타입 매칭(token-type matching) 일치도를 높이기 위해서 한국수어 자료 통합 지원 시스템을 사용한다.
- 6) 언어 자료 영상에서 2명의 언어 제공자가 다 찍힌 영상(full shot)의 왼편 언어 제공자는 A, 오른편 언어 제공자는 B로 이름 붙인다.
- 7) eaf 파일(엘란(Elan) 프로그램)에 전사를 마친 파일명은 시스템에서 다운 받은 파일명을 그대로 유지해야 한다.

본 사업에서도 위의 기본 사항을 바탕으로 전사를 진행하되, 전사자 교육과 전사 회의 항목을 보다 강화하였다. 대부분의 전사자가 ELAN 프로그램을 활용한 수어 전사 방법을 알지 못하므로 전사 역량 강화를 위한 교육의 확대가 꼭 필요하기 때문이다. 전년도 사업에서는 총 3회에 걸쳐 전사자 교육을 실시한 바 있는데¹²⁾, 본 사업에서는 말뭉치 활용 워크숍과 전사자 상시 교육을 합하여 총 5회 이상의 교육을 실시하기로 결정하였다. 말뭉치 활용 워크숍은 연구자와 일반을 대상으로 기획되는 것으로서 총 2회 이상 실시하며, 되도록 모든 전사자가 참여하여 전사의 기초를 배울 수 있게 하였다. 또한 필요에 따라 전사자 상시 교육을 추가적으로 실시하여 부족한 부분을 메울 수 있도록 하였다.

12) 2017년 6월 21일부터 23일까지 사업에 대한 소개, ELAN 사용법, 전사 지침 등에 대한 교육이 실시되었다. 자세한 내용은 국립국어원(2007: 51-52) 참조.

이와 더불어 전사 일치도를 높이기 위한 전사 회의도 주 1회에서 주 2회 이상으로 확대 실시하기로 결정하였다. 기본적으로는 연구자 전원이 참여하는 매주 정기 모임에서 전사 회의를 함께 진행하되, 수어 능력이 부족한 청인 전사자들을 중심으로 추가 회의를 주 1회 실시하도록 하였다.

이에 따라 수정된 전사 기본 사항은 다음과 같다.

- 1) 전사를 위해 엘란(Elan) 프로그램을 사용한다.
- 2) 전사를 위한 전사자 교육을 실시한다. 전사자 교육은 말뭉치 활용 워크숍과 전사자 상시 교육을 합하여 총 5회 이상 실시한다.
- 3) 전사의 일치도를 높이기 위해 공동 전사를 원칙으로 한다. 공동 전사가 어려운 경우 최소한 처음 5회 이상은 공동 전사를 실시한다.
- 4) 전사의 일치도를 높이기 위해 주 2회 이상의 전사 회의를 실시한다.
- 5) 전사 토큰 타입 매칭(token-type matching) 일치도를 높이기 위해서 한국수어 자료 통합 지원 시스템을 사용한다.
- 6) 언어 자료 영상에서 2명의 언어 제공자가 다 찍힌 영상(full shot)의 원편 언어 제공자는 A, 오른편 언어 제공자는 B로 이름 붙인다.
- 7) eaf 파일(엘란(Elan) 프로그램)에 전사를 마친 파일명은 시스템에서 다운 받은 파일명을 그대로 유지해야 한다.

2. 한국수어 말뭉치 전사 과정

본 연구에서는 2015년도 사업을 통해 구축된 언어 자료 중 80분 분량에 대한 기본 전사를 진행하였다. 총 80분 중 50분은 전년도 연구팀(전사팀A, 복지대)의 협조를 받아 외주 형태로 진행되었다. 이는 기존 연구와의 원활한 연계와 더불어 구축되는 말뭉치의 균질성을 확보하기 위함이다. 그리고 남은 30분의 분량은 본 연구팀에서(전사팀B, 전북대) 자체적으로 전사를 실시하였다.

말뭉치 전사의 전 과정에서 전사자의 역량 강화를 위한 교육이 지속적으로 병행되었다. 전사 교육은 말뭉치 활용 워크숍과 전사자 상시 교육의 두 갈래로 이루어졌는데, 특히 매주 정기 회의에서 전사자 상시 교육을 실시함으로써 전사자의 자질 향상을 위해 노력하였다. 상시 교육의 대략적인 내용은 다음과 같다.

- ① 한국수어 교육
- ② 말뭉치 언어학 교육
- ③ 전사 지침 교육
- ④ ELAN 사용법 교육
- ⑤ 한국수어 자료 통합 지원 시스템 사용법 교육

이 중 한국수어 교육은 본격적인 연구가 시행되기 이전인 2018년 5월경부터 매주 꾸준히 진행되었다. 수어 능력이 부족한 청인 연구원들을 대상으로 농인 연구원들이 번갈아 가며 교육을 담당하였다. 또한 말뭉치 언어학과 전사 지침 교육은 연구 책임자(윤석민)를 중심으로 필요한 자료를 강독하고 토론하는 스터디 형식으로 진행되었다. 그리고 ELAN 및 한국수어 자료 통합 지원 시스템 사용법 교육은 워크숍 이후 시스템 관리자(김성준)가 주축이 되어 진행하였다. 관련 안전이 발생할 때마다 2017년 한국수어 자료 통합 지원 시스템 고도화 사업 부록에 있는 사용자 매뉴얼을 토대로 토론을 실시하였다.

이와 같은 교육하에 이루어진 말뭉치 전사는 크게 전사 영상 선정, 기본 번역 및 검수, 전사 작업의 절차로 진행되었다. 자세한 내용을 제시하면 다음과 같다.

2.1. 전사 영상 선정

본격적인 전사를 수행하기 위한 첫 번째 단계는 80분 분량의 전사 영상을 선정하는 것이었다. 기구축 수어 말뭉치 중 전사되지 않은 영상의 개수는 총 108개로, 이들 중 전사할 영상을 선정하기 위한 기준이 필요했다. 기준 마련을 위한 논의는 크게 통일성과 다양성 중 어느 쪽을 중시할 것인가의 측면에서 이루어졌다. 각 방식의 장점 및 단점은 다음과 같다.

① 통일성 중시

- 동일한 언어 제공자 쌍 혹은 동일한 유도 자료를 기준으로 전사 영상을 통일하는 방식
- 장점: 전사의 편의성, 정확성 증진
- 단점: 상대적으로 다양한 수어 사용 양상을 살피기 어려움

② 다양성 중시

- 다양한 언어 제공자와 다양한 유도 자료를 포함하여 전사 영상을 선정하는 방식
- 장점: 상대적으로 다양한 수어 사용 양상을 살필 수 있음
- 단점: 전사의 어려움, 오랜 시간 소요

이 중 본 사업에서는 두 가지 측면을 모두 고려하되 다양성에 보다 초점을 두고 여러 분야의 자료를 전사하기로 하였다. 언어 제공자와 유도 자료를 통일하는 것도 의미가 있으나 이번 전사의 목적은 궁극적으로 연구팀이 다양한 수어 자료를 접하고 전사 경험을 축적하는 데 있기 때문이다. 따라서 최대한 다양한 유형의 자료를 전사하기로 결정하였다.

이러한 선정 기준에 의거하여 최종적으로 선정된 영상은 다음과 같다.

연번	과제명	재생시간	번역	전사
1	서울 지역 19번 짱 E-토론	00:13:30	서일권, 이영재	복지대
2	서울 지역 19번 짱 J-농사회 행사	00:11:37	나두호, 이영재	복지대
3	서울 지역 1번 짱 E-토론	00:17:03	나두호, 노유리, 이영재	복지대
4	서울 지역 29번 짱 M-농학교 경험 이야기	00:09:16	노유리, 이영재	복지대
5	서울 지역 29번 짱 H-주제 영역	00:20:46	나두호, 이영재	전북대
6	서울 지역 24번 짱 J-농사회 행사	00:09:42	나두호, 이영재	전북대

<표 9> 2018년 한국수어 말뭉치 전사 영상 목록

전사 영상의 개수는 총 6개이며 전체 재생시간은 82분 54초이다. 또한 전사 영상의 언어 제공자 짱은 총 4쌍(1번 짱, 19번 짱, 24번 짱, 29번 짱), 유도 자료는 총 4가지(토론, 주제 영역, 농사회 행사, 농학교 경험)로 언어 제공자와 유도 자료 모두 다양성을 고려하여 선정하였다. 전체 영상의 번역 및 검수는 이영재, 나두호, 노유리, 서일권이 실시하며 이중 연번 1-4번 영상은 전사팀A(복지대)에 전사를 의뢰했다. 전사팀 B(전북대)에서 모든 전사를 전담하는 영상은 연번 5-6번이다.

2.2. 기본 번역 및 검수

전사할 영상을 선정한 뒤에는 각 영상의 기본 번역 및 검수를 실시하였다. 이는 다음의 4단계로 진행되었다.

- 1) 업무 분담 및 초벌 번역
- 2) 초벌 번역 검토 및 수정
- 3) ELAN 번역 총렬 입력
- 4) 번역 검수 작업

각 단계별로 진행 일정 및 논의 내용을 상세화하면 다음과 같다.

1) 업무 분담 및 초벌 번역

농인 연구자들(이영재, 나두호, 노유리, 서일권)이 전사 영상을 각자 분담하여 초벌 번역을 수행하였다. 연구자들이 아직 ELAN 프로그램에 익숙지 않은 상태였기 때문에 우선 한글 프로그램으로 번역을 실시하였다. 언어 제공자 A, B가 발화한 내용과 발화

시간을 각자의 방식으로 적어 보는 작업이었다. 번역 과정에서 드러나는 문제점들은 추후 논의를 통해 시정하기로 하고 우선 자유로운 방식으로 번역하도록 하였다.

2) 초벌 번역 검토 및 수정

초벌 번역이 끝난 뒤 연구원 전체가 모여 번역 파일을 함께 살펴보고 문제점 및 개선 사항을 점검하는 시간을 가졌다.

초벌 번역에서 드러난 첫 번째 문제점은 발화 시간의 표기 문제였다. 다소 긴 내용을 한꺼번에 이어서 번역하면서 언어 제공자의 발화 시간을 상세하게 파악하기 힘든 경우가 있었다. 이를테면 다음과 같은 경우이다.

<서울 지역 19번 쌍 E-토론 초벌 번역 사례>

A(0:11): 네 서울지역입니다. 한 곳은 서울농학교, 서울애화학교, 서울삼성학교, 서울구화학교 이렇게 네 곳이 있습니다. 하지만 한 곳에 농인 선생님이 없으시고, 서울농학교에는 한 명이 계시고요. 서울애화학교는 3~4명 정도 계시지만 알고 보니 완농이 아니었고 구화 반 농인 반 이런 분들이 계셨고, 서울삼성학교는 자세히 잘 모르기 때문에 넘어가겠습니다. 이제 교육에 대해서 얘기할 건데요. 대부분 농학교를 보면 선생님은 거의 청인이 많고 농인이 적죠? 하지만 제 개인적으로는 청인이 반 농인이 반 이렇게 나눠서 교육하면 좋을 것 같습니다. 왜냐하면 선생님들이 교육하는 데에 제한이 있기 때문입니다. 예를 들면 청인 선생님들이 교육하는 데에 말로 가르치시거나 칠판에다 쓰기 때문에 농학생들이 공부하는 데 어려움이 많습니다. 반면 농선생님은 가르치실 때 수화로 통해서 수업하는 데에 어려움이 없긴 하나 농선생님이 너무 많으면 저율이 수평이 안 될까봐 서로 맞출 필요가 있을 것 같습니다.

위의 예문은 언어 제공자 A가 11초부터(0:11) 발화한 내용을 번역한 것이다. 이처럼 길게 이어지는 단락을 번역할 때 언어 제공자를 기준으로 발화 시간을 한 번만 표시하게 되면 번역본을 바탕으로 영상을 이해하는 데 어려움이 발생할 수 있다. 또한 차후 번역한 내용을 ELAN에 입력할 때에는 맥락에 따라 더 짧은 단위로 끊어서 입력해야 할 수 있으므로 발화 시간을 보다 상세하게 표기하기로 하였다. 예를 들면 다음과 같은 번역이 바람직하다고 볼 수 있다.

<서울 지역 29번 쌍 M-농학교 경험 이야기 초벌 번역 사례>

B(0:23-0:25): 제일 좋았던 추억은 초등학교 때였어요.

B(0:25-0:26): 잊지 못할 추억이었죠.

B(0:26-0:29): 친구들이랑 엄청(잘) 친해요.

이처럼 동일한 언어 제공자 B의 발화가 이어지는 경우에도 번역본을 문장 단위로 잘게 쪼개어 발화 시작 시간과 끝 시간을 함께 표시해 주는 것이 바람직하다. 반드시 문장 단위가 아니더라도 전체적인 맥락을 고려하면서 발화 시간을 최대한 상세하게 표기하도록 하였다.

다음으로 번역 방식의 문제도 있었다. 번역본마다 지나친 의역으로 일부 내용이 누락된 경우와, 지나친 직역으로 내용 파악이 어려운 경우가 공존하였다. 이에 본 연구에서는 전년도 연구에서 채택한 자유 번역¹³⁾ 형태를 바탕으로 번역 방식을 통일하였다. 직역을 기본으로 하되 목표 언어인 한국어로도 충분히 이해가 가능하도록 문법 형태소를 적절하게 첨가하고 필요에 따라 어휘, 어순 등도 자연스럽게 수정하기로 하였다. 그리고 맥락상 생략된 부분이나 의역이 필요한 부분에는 괄호를 이용하여 필요한 내용을 삽입하도록 하였다. 이러한 방식에 따라 번역본을 수정한 사례를 제시하면 다음과 같다.

<서울 지역 24번 쌍 J-농사회 행사 번역 수정 사례>

① 수정 전

A(0:47-0:57): 나 고3 학생 마지막 기념 정장 입고 다 모여 고3 다함께 모여 사진
나 준 거 없다 사진 준 거 없다

② 수정 후

A(0:47-0:57): 제가 고3 학생일 때 마지막 기념으로 정장 입고 다 모여서, 고3 때
다함께 모여서 (찍은) 사진을 저에게 주지 않았습니다.

또한 번역 과정에서 주의를 기울인 것은 차별적 용어의 사용 여부였다. 농인 연구자들이 번역을 수행하는 과정에서 특별한 의식 없이 ‘일반’, ‘일반인’, ‘일반 선생님’ 등의 용어를 ‘농인’에 대비되는 용어로 사용하는 경우가 있었다. 이는 해당 용어의 차별적 함의를 미처 인식하지 못한 결과이므로 다음과 같이 바람직한 용어로 수정하였다.

13) 2017년도 연구에서는 정밀 번역과 자유 번역 중 자유 번역 방식을 채택한 바 있다. 두 번역 형태의 특징은 다음과 같다(국립국어원, 2017:238-239).

① 자유 번역(Free Translation)

- 목표 언어인 한국어에서 자연스럽게 실행되는 텍스트를 선호한다.
- 음성 언어 번역에 의해 부분적으로 결정되므로 문장의 길이가 상대적으로 길다.
- 목표 언어에 적합한 지시 표현을 사용한다.
- 언어학자 또는 다른 사용자들이 수어 담화의 내용에 대한 이해를 하는 데에 도움이 된다.

② 정밀 번역(Narrow Translation)

- 비교적 짧은 절 또는 문장의 부분에서 수어 텍스트에 가까운 형태이다.
- 목표어인 한국어로 매끄럽게 번역할 필요는 없다.
- 수어 담화의 구조를 이해하는 데 도움을 주고, 기계 번역 알고리즘을 학습하고 테스트하는 데 유용하다.

<서울 지역 19번 쌍 E-토론 번역 수정 사례>

① 수정 전

A(0:11): 대부분 농학교를 보면 선생님은 거의 일반인이 많고 농인이 적죠? 하지만 제 개인적으로는 일반인이 반 농인이 반 이렇게 나눠서 교육하면 좋을 것 같습니다. 왜냐하면 선생님들이 교육하는 데에 제한이 있기 때문입니다. 예를 들면 일반 선생님들이 교육하는 데에 말로 가르치시거나 칠판에다 쓰기 때문에 농학생들이 공부하는 데 어려움이 많습니다.

① 수정 후

A(0:11): 대부분 농학교를 보면 선생님은 거의 청인이 많고 농인이 적죠? 하지만 제 개인적으로는 청인이 반 농인이 반 이렇게 나눠서 교육하면 좋을 것 같습니다. 습니다. 왜냐하면 선생님들이 교육하는 데에 제한이 있기 때문입니다. 예를 들면 청인 선생님들이 교육하는 데에 말로 가르치시거나 칠판에다 쓰기 때문에 농학생들이 공부하는 데 어려움이 많습니다.

그밖에도 언어 제공자 A, B를 남/녀, 오른쪽/왼쪽 등으로 표기하는 등의 사소한 차이를 통일하고, 논의한 내용을 바탕으로 초벌 번역본을 각자 수정하였다.

3) ELAN 번역 총렬 입력

농인 연구원들이 초벌 번역본을 최종 수정한 뒤에는 그 결과물을 청인 연구원들(강진주, 김성준, 최화니)이 ELAN 프로그램으로 입력하는 작업을 수행하였다. 이를 통해 연구자들이 ELAN 활용 방법을 습득하고 수어 분절에 대한 감각을 익힐 수 있도록 하였다. 이와 더불어 농인 연구원들도 각자 번역한 한글 파일의 일부를 ELAN에 옮기는 연습을 함께하였다. 추후 수지 글로스 입력과 같이 더욱 정밀한 전사 작업을 수행할 때는 농인 연구원들의 역할이 매우 중요하기 때문이다. 각종 워크숍과 상시 교육을 통해 배운 ELAN 사용법을 연구자 전원이 꾸준히 연습하도록 독려하였다.

번역 총렬 입력 작업은 개별적으로 이루어지지 않고 연구원들이 3인 이상 모인 상태에서 실시되었다. 수어를 모국어로 사용하는 농인 연구원과 달리 청인 연구원들은 수어에 대한 직관이 부족하므로 입력 과정에서 지속적인 도움이 필요했기 때문이다. 따라서 모든 자료의 입력은 수어 통역사(이현숙)와 농인 연구원들과의 협업을 통해 진행되었다.

또한 입력 과정에서 드러나는 문제점들은 그 자리에서 함께 논의하여 해결하였다. 첫 번째로 드러난 문제점은 연구원들이 번역의 시작점과 끝점을 잡지 못하는 것이었다. 번역 총렬의 입력은 단어 하나하나에 글로스를 붙이는 작업이 아님에도 불구하고 개별 단어의 분절 지점에 집착하여 좀처럼 입력 속도가 나지 않는 경우가 있었다. 이에 본 연구에서는 수어 단어의 분절 원칙¹⁴⁾을 고려하되, 번역 총렬에서의 분절이 단

14) 2017년도 연구에서 제시한 수어 단어의 분절 원칙은 다음과 같다(국립국어원, 2017: 221).

어가 아닌 텍스트 단위의 분절임을 감안하여 보다 유연하게 전사하도록 하였다. 더불어 여러 연구원들이 함께 특정 영상을 반복 재생하면서 대략적인 시작점과 끝점이 어디인지를 살피고 학습할 수 있게 하였다.

또 다른 문제점은 분절 단위의 설정이었다. 앞서 초벌 번역본을 최대한 잘게 쪼개는 방식으로 수정하였는데, 어느 경우에는 지나치게 짧게 끊다 보니 의미 파악이 제대로 되지 않는 문제가 발생하였다. 따라서 본 연구에서는 맥락을 고려하여 분절 단위를 설정하기로 하였다. 문장 단위가 아닌 담화 단위로, 전체적인 맥락을 해치지 않는 선에서 분절 작업을 수행하였다.

4) 번역 검수 작업

ELAN 프로그램으로 번역 총렬 입력을 마친 뒤에는 최종적인 검수가 이루어졌다. 여러 명의 농인 연구자들이 개별적으로 번역을 진행하였기 때문에 번역 지침이 있었음에도 여전히 자료별 통일성이 부족한 상태였다. 또한 지역별, 개인별 차이가 큰 수어의 특성 때문에 수어를 모국어로 사용하는 농인들도 영상을 정확하게 번역하기 어려워 공백으로 남겨둔 부분도 일부 있었다.

따라서 최종적으로 전사 책임자(이영재)를 중심으로 검수 작업을 실시하였다. 파일 전체를 꼼꼼히 훑어보면서 잘못된 부분은 수정하고 부족한 부분은 채워 넣는 작업을 수행하였다.

2.3. 전사 작업

기본 번역 및 검수가 완료된 이후에는 본격적인 전사 작업을 실시하였다. 이는 번역 총렬 이외의 나머지 총렬(수지 글로스 총렬 및 코멘트 총렬)을 입력하는 작업을 말한다. 전사팀A(복지대)가 총 50분, 전사팀B(전북대)가 총 30분 분량의 전사 작업을 실시하였다.

모든 전사 작업은 2017년도에 구축된 한국수어 말뭉치 전사 지침에 따라 진행되었다(국립국어원, 2017: 219-249). 특히 전사팀B(전북대)는 전사 작업을 처음으로 수행하기 때문에 전사 지침을 숙지하는 한편으로 말뭉치 활용 워크숍을 통해 전사팀A(복지대)로부터 여러 가지 전사 노하우를 배우고, 전사 방식을 최대한 균일하게 맞추고자 노력하였다. 그 과정에서 일부 전사 지침은 실제 전사 과정에서 다소 축소되기도 하였다.

-
- ① 이동이 없는 수어는 수형이 보이기 시작하는 첫 번째 프레임이 수어의 시작이다.
 - ② 이동이 있는 수어 단어의 경우는 손이 움직이기 시작하는 첫 번째 프레임이 수어의 시작이다.
 - ③ 이동도 수형의 변화도 없이 이전 수어에서 새로운 수어가 나타나는 경우는 수형이 변화하기 시작하는 첫 번째 프레임이 수어의 시작이다.
 - ④ 수어 단어의 끝은 수형이 변화하기 시작하는 첫 번째 프레임의 바로 이전 프레임이다.
 - ⑤ 이동이 시작되는 첫 번째 프레임의 바로 이전 프레임이 이전 수어 단어의 끝이다.

전사 과정에서 간소화된 지침은 형태론적 수정과 변이형의 표기이다. 2015년 연구에서는 한국수어 말뭉치 전사를 위해 기본형과 형태론적 수정, 변이형의 표기 지침을 별도로 마련하였으나, 전사 작업 도중 세 가지를 더 이상 구분하지 않고 하나의 타입으로 묶어서 전사하는 방향으로 규칙을 수정한 바 있다(국립국어원, 2015: 101). 이는 이 세 가지 형태를 구분하기 위해 몇 개월 이상의 오랜 교육 기간이 필요하다는 점, 정밀한 지침 도출을 위해서는 보다 충분한 전사량을 확보해야 한다는 점 때문이었다. 따라서 전사팀A도 전사의 정확성을 높이기 위해 기존의 방식을 취하여 전사하기로 하였다. 기구축된 타입에서 기본형, 형태론적 수정, 변이형이 구분된 경우에는 그에 맞게 구분하여 전사하되, 새로운 타입을 만들어야 하는 경우에만 형태론적 수정과 변이형을 고려하지 않고 기본형으로 구축·전사하였다. 그러나 추후에는 이와 관련된 교육을 확대하고 지침을 더욱 상세화할 필요가 있으므로, 관련 지침은 삭제하지 않고 일부분을 수정 보완하여 제시하기로 결정하였다.¹⁵⁾

이와 더불어 2017년 전사 지침에 새롭게 포함된 입 움직임 전사 여부에 대한 논의도 있었다. 2017년 연구에서는 입 움직임을 마우딩(Mouthings: 구어 단어를 조음할 때와 같이 입모양을 만드는 것)과 마우스 제스처(Mouth gestures: 구어 단어 조음 외의 모든 입 움직임)로 구분하고 이를 전사하기 위한 정밀 전사 규칙을 마련하였으나(국립국어원, 2017: 239-249), 실질적인 정밀 전사는 수행되지 못한 바 있다. 본 연구에서 이 지침에 따라 정밀 전사를 추가로 수행할 필요성에 대한 논의가 있었으나, 결론적으로는 전사하지 않기로 하였다.

입 움직임 전사의 가장 큰 문제점은 기준의 모호함이다. 마우스 제스처는 혀와 이가 보이느냐 보이지 않느냐 등에 따라 전사 기호가 구분되는데, 여기에는 주관적인 판단이 개입될 수 있다. 입을 ‘가장 작게 벌림’, ‘중간 정도로 벌림’, ‘가장 크게 벌림’, ‘다물’, ‘더 세게 다물’ 등의 등급 지표를 객관적으로 파악하여 정확하게 기술하는 데에는 많은 어려움이 따를 수 있다.

또 다른 문제는 동작의 다양성이다. 마우스 제스처는 개인별 차이가 매우 크다. 예컨대 ‘먹다’라는 기본적인 수어를 할 때에도 입을 벌리는 사람이 있는가 하면 다무는 사람도 있다. 이처럼 개인별로 무척 다양하게 나타나는 동작들을 유형화하기가 쉽지 않을 것으로 예상된다.

따라서 입 움직임을 전사는 차기 사업으로 제안하고자 한다. 이를 위해서는 보다 정밀한 전사 지침이 마련되어야 함은 물론이고 각 연구원들의 역량 증진 또한 필요하다.

한편 전사팀B(전북대)에서 전사의 정확성을 높이기 위해 주의를 기울인 부분은 전면적인 공동 전사를 수행하였다는 점이다. 기존 전사 지침은 기본적으로 교육과 일치도 검사 이후로는 공동 전사를 원칙으로 하고 있지만, 예외를 두어 개별 전사를 허용하고 이를 보완하기 위한 전사 회의를 수시로 진행했다.

전사의 효율성을 위해서는 다수의 인원이 여러 영상을 전사하는 것이 유리하다. 그

15) 자세한 내용은 3.1. 한국수어 말뭉치 전사 지침 정비 참조.

그러나 전사 지침에서 원칙적으로 공동 전사 규정을 둔 이유는 다수가 하나의 영상을 전사하는 것이 전사의 질이 가장 높기 때문일 것이다. 따라서 전사팀B(전북대)에서는 공동 전사만을 진행했다. 구체적으로는 개별 전사자가 동일한 파일의 동일한 부분을 전사하고, 전사 회의를 통해 파일을 결합하는 방식으로 진행했다. 이는 결국 ‘개별 전사+공동 전사’이기는 하지만 전사 회의 결과로 도출된 내용만을 전사 내용으로 확정했기 때문에 원칙적인 공동 전사에서 벗어나지 않았다.

많은 분량의 전사를 진행할 때는 시간과 노력이 많이 소요되는 공동 전사를 진행하기 어려울 수 있다. 그러나 전사팀B(전북대)가 맡은 30분의 분량은 정해진 시간 내에 소화가 가능한 분량이고 연구원들의 전사 경력이 부족하기 때문에, 전사의 신뢰성을 위해 전면적 공동 전사를 기본 원칙으로 삼아 전사를 진행하였다.

본격적인 전사 작업은 말뭉치 영상 촬영이 완료된 시점인 10월 10일부터 진행하였다. 매주 2회씩 전사 회의를 진행하여 청인 연구진끼리 전사 진행 상황을 검토하고 점검한 뒤, 전체 연구진이 모이는 정기 모임을 통해 전사 내용 전체를 검토하고 확정된 내용을 토대로 전사 작업을 실시하였다.

일시	참여 대상	작업 내용	전사 대상 파일
2018.10.10.	청인 연구진	Elan 사용법 및 분절 지점 논의	서울 지역 29번 쌍 H-주제 영역
2018.10.13.	전체 연구진	Elan 추가 교육 및 청인 연구진 전사 분량 분배	서울 지역 29번 쌍 H-주제 영역
2018.10.17.	청인 연구진	제1차 초벌 전사	서울 지역 29번 쌍 H-주제 영역
2018.10.20.	전체 연구진	전사 내용 확인	서울 지역 29번 쌍 H-주제 영역
2018.10.24.	청인 연구진	제2차 초벌 전사	서울 지역 29번 쌍 H-주제 영역
2018.10.27.	전체 연구진	전사 내용 확인	서울 지역 29번 쌍 H-주제 영역
2018.10.31.	청인 연구진	제3차 초벌 전사	서울 지역 29번 쌍 H-주제 영역
2018.11.03.	전체 연구진	전사 내용 확인	서울 지역 29번 쌍 H-주제 영역
2018.11.09.	청인 연구진	제4차 초벌 전사	서울 지역 29번 쌍 H-주제 영역
2018.11.10.	전체 연구진	전사 내용 확인	서울 지역 29번 쌍 H-주제 영역
2018.11.16.	청인 연구진	제5차 초벌 전사	서울 지역 29번 쌍 H-주제 영역

2018.11.17.	전체 연구진	전체 내용 확인 및 종합	서울 지역 29번 쌍 H-주제 영역
2018.11.23.	청인 연구진	새로운 전사 분량 확인 및 제1차 초별전사	서울 지역 24번 쌍 J-농사회 행사
2018.11.24.	전체 연구진	전사 내용 확인	서울 지역 24번 쌍 J-농사회 행사
2018.11.29.	청인 연구진	제2차 초별전사	서울 지역 24번 쌍 J-농사회 행사
2018.12.1.	전체 연구진	전사 내용 확인	서울 지역 24번 쌍 J-농사회 행사
2018.12.7.	청인 연구진	제3차 초별 전사	서울 지역 24번 쌍 J-농사회 행사
2018.12.8.	전체 연구진	전사 내용 확인 및 종합	서울 지역 24번 쌍 J-농사회 행사

<표 10> 2018년 한국수어 말뭉치 전사 일정표

위 표는 30분 분량에 대한 전사를 위해 진행된 전사 회의 일정표이다. 청인 전사자는 모두 같은 부분을 전사했고 그 결과를 청인 전사 회의를 통해 공유하고 논의했다. 위 회의에서 논의된 구체적인 내용 및 전사 결과는 다음 장에서 다룰 것이다.

또한 전체 80분 중 50분에 해당하는 부분은 전사팀A(복지대)의 협조를 받아 진행했다. 번역 총렬은 앞서 살핀 것과 마찬가지로 전사팀B(전북대)에서 입력하고 이후 수지 글로스와 코멘트 총렬의 전사를 외주 형태로 의뢰했다. 전사 의뢰와 결과 수령은 모두 한국수어 자료 통합 지원 시스템을 통해 이루어졌고 작업이 완료된 전사파일은 전북대 팀에서 다시 검수하여 최종 검토를 실시했다.

2.4. 전사 결과

본 연구에서는 총 6개 파일, 82분 54분 분량의 영상에 대해 번역 총렬과 수지 글로스 총렬, 코멘트 총렬의 기본 전사를 완료하였다. 우세손을 기준으로 산출한 총 발생 토큰 수는 13,302개이며, 새로운 타입 수는 총 199개이다. 각 과제별 세부 정보는 다음과 같다.

연번	과제명	영상 시간	토큰 수	새 타입 수
1	서울 지역 19번 쌍 E-토론	00:13:30	1,783	11
2	서울 지역 19번 쌍 J-농사회 행사	00:11:37	1,653	14
3	서울 지역 1번 쌍 E-토론	00:17:03	2,699	21
4	서울 지역 29번 쌍 M-농학교 경험 이야기	00:09:16	1,534	8

5	서울 지역 29번 쌍 H-주제 영역	00:20:46	3,835	112
6	서울 지역 24번 쌍 J-농사회 행사	00:09:42	1,798	33
총 6개 파일		82분 54초	13,302	199

<표 11> 2018년 한국수어 말뭉치 전사 과제별 세부 사항

2.4.1 신규 타입 목록

한국수어 말뭉치 전사는 앞서 밝힌 바와 같이 아직 초기 단계이다. 그렇기 때문에 전사를 진행하는 과정에서 새로운 타입이 지속적으로 발견되었다. 전사를 진행하는 과정에서 199개의 새로운 타입을 발견하였으며, 그밖에도 기구축된 타입을 검토하고 문제점을 수정하는 과정에서 2개의 타입을 추가로 등록하였다¹⁶⁾. 결과적으로 새롭게 시스템에 등록한 타입 목록은 총 201개이다.

<표 12> 2018년 한국수어 말뭉치 전사 신규 타입 목록

연번	타입명	새타입 출현 말뭉치 영상	연번	타입명	새타입 출현 말뭉치 영상
1	~하자2	서울 지역 29번 쌍 M-농학교 경험 이야기	102	스위스1	서울 지역 29번 쌍 H-주제 영역
2	CCTV1	서울 지역 29번 쌍 H-주제 영역	103	스카우트1	서울 지역 24번 쌍 J-농사회 행사
3	counting hand 더하다	서울 지역 29번 쌍 H-주제 영역	104	실험1	서울 지역 24번 쌍 J-농사회 행사
4	가속1	서울 지역 29번 쌍 H-주제 영역	105	심하다2	서울 지역 29번 쌍 H-주제 영역
5	가자1	서울 지역 29번 쌍 H-주제 영역	106	십년후1	서울 지역 1번 쌍 E-토론
6	가져오다 5	서울 지역 29번 쌍 H-주제 영역	107	십이년1	서울 지역 29번 쌍 M-농학교 경험 이야기
7	가중1	서울 지역 29번 쌍 H-주제 영역	108	쌓다3	서울 지역 19번 쌍 J-농사회 행사
8	각자1	서울 지역 19번 쌍 J-농사회 행사	109	아침4	서울 지역 29번 쌍 H-주제 영역
9	간접1	서울 지역 24번 쌍 J-농사회 행사	110	알다4	서울 지역 29번 쌍 H-주제 영역
10	감속1	서울 지역 29번 쌍 H-주제 영역	111	야금야금 쓰다1	서울 지역 29번 쌍 H-주제 영역
11	감싸다1	서울 지역 29번 쌍	112	어디2	서울 지역 19번 쌍

16) 아래 표에 제시한 신규 타입 목록 중에서 200번 [본인1], 201번 [피곤하다1]이 여기에 해당한다. 이에 대한 내용은 3.2. 한국수어 말뭉치 전사 제언 중 3.2.1. 타입 정비 제언 참조.

		H-주제 영역			E-토론
12	같다4	서울 지역 19번 째 J-농사회 행사	113	어떻게1	서울 지역 29번 째 H-주제 영역
13	거꾸로1	서울 지역 29번 째 H-주제 영역	114	어른1	서울 지역 24번 째 J-농사회 행사
14	공공칠1	서울 지역 29번 째 H-주제 영역	115	오다2	서울 지역 24번 째 J-농사회 행사
15	공부2	서울 지역 29번 째 M-농학교 경험 이야기	116	어이없다1	서울 지역 29번 째 H-주제 영역
16	괴팍하다 2	서울 지역 29번 째 H-주제 영역	117	엄매이다1	서울 지역 1번 째 E-토론
17	구정1	서울 지역 29번 째 H-주제 영역	118	엄격하다2	서울 지역 29번 째 H-주제 영역
18	권리1	서울 지역 29번 째 H-주제 영역	119	엎다3	서울 지역 24번 째 J-농사회 행사
19	기록2	서울 지역 29번 째 H-주제 영역	120	에펠탑1	서울 지역 29번 째 H-주제 영역
20	기억2	서울 지역 24번 째 J-농사회 행사	121	여권1	서울 지역 29번 째 H-주제 영역
21	기준1	서울 지역 19번 째 E-토론	122	여왕1	서울 지역 29번 째 H-주제 영역
22	길2	서울 지역 29번 째 H-주제 영역	123	열다7	서울 지역 29번 째 H-주제 영역
23	너그럽다 1	서울 지역 19번 째 J-농사회 행사	124	열다8	서울 지역 29번 째 H-주제 영역
24	넘기다1	서울 지역 24번 째 J-농사회 행사	125	옛새1	서울 지역 29번 째 H-주제 영역
25	넘다2	서울 지역 29번 째 H-주제 영역	126	영점1	서울 지역 29번 째 H-주제 영역
26	노랑2	서울 지역 29번 째 H-주제 영역	127	오가다1	서울 지역 29번 째 H-주제 영역
27	놀라다8	서울 지역 29번 째 H-주제 영역	128	오십년1	서울 지역 19번 째 J-농사회 행사
28	눈치2	서울 지역 29번 째 H-주제 영역	129	오케이3	서울 지역 29번 째 H-주제 영역
29	눈치3	서울 지역 29번 째 H-주제 영역	130	오케이3	서울 지역 29번 째 H-주제 영역
30	놀리다1	서울 지역 24번 째 J-농사회 행사	131	오토바이 가다1	서울 지역 29번 째 H-주제 영역
31	단독1	서울 지역 24번 째 J-농사회 행사	132	웁기다7	서울 지역 29번 째 H-주제 영역

32	단체로가 다1	서울 지역 24번 째 J-농사회 행사	133	웁기다7	서울 지역 29번 째 H-주제 영역
33	달려가다 1	서울 지역 29번 째 H-주제 영역	134	웃겂다1	서울 지역 29번 째 H-주제 영역
34	달리다3	서울 지역 24번 째 J-농사회 행사	135	외국1	서울 지역 29번 째 H-주제 영역
35	달리다4	서울 지역 24번 째 J-농사회 행사	136	웁하다2	서울 지역 24번 째 J-농사회 행사
36	대의원1	서울 지역 29번 째 H-주제 영역	137	위장1	서울 지역 29번 째 H-주제 영역
37	대회1	서울 지역 1번 째 E-토론	138	윽혹3	서울 지역 29번 째 H-주제 영역
38	대회2	서울 지역 24번 째 J-농사회 행사	139	웁백1	서울 지역 29번 째 M-농학교 경험 이야기
39	던지다2	서울 지역 19번 째 J-농사회 행사	140	웁십년1	서울 지역 29번 째 H-주제 영역
40	도장3	서울 지역 24번 째 J-농사회 행사	141	웁십오1	서울 지역 1번 째 E-토론
41	독일1	서울 지역 29번 째 H-주제 영역	142	으깨다1	서울 지역 29번 째 H-주제 영역
42	돈주다3	서울 지역 29번 째 H-주제 영역	143	음주운전1	서울 지역 29번 째 H-주제 영역
43	두번째2	서울 지역 19번 째 E-토론	144	이십팔1	서울 지역 29번 째 H-주제 영역
44	듣지못하 다1	서울 지역 19번 째 E-토론	145	이주일1	서울 지역 29번 째 H-주제 영역
45	떠나다2	서울 지역 24번 째 J-농사회 행사	146	익숙하다1	서울 지역 29번 째 H-주제 영역
46	뜻밖에2	서울 지역 29번 째 H-주제 영역	147	인격1	서울 지역 19번 째 E-토론
47	리터1	서울 지역 1번 째 E-토론	148	인생1	서울 지역 1번 째 E-토론
48	마주보다 2	서울 지역 24번 째 J-농사회 행사	149	인쇄1	서울 지역 24번 째 J-농사회 행사
49	마주서다 1	서울 지역 29번 째 H-주제 영역	150	인터뷰1	서울 지역 29번 째 H-주제 영역
50	막다2	서울 지역 29번 째 H-주제 영역	151	일형3_네 모	서울 지역 29번 째 H-주제 영역
51	만큼1	서울 지역 29번 째 H-주제 영역	152	작다5	서울 지역 1번 째 E-토론
52	맞지않다	서울 지역 1번 째	153	잡동사니1	서울 지역 24번 째

	3	E-토론			J-농사회 행사
53	멍청이1	서울 지역 29번 짱 H-주제 영역	154	장애3	서울 지역 24번 짱 J-농사회 행사
54	메르스2	서울 지역 29번 짱 H-주제 영역	155	절하다1	서울 지역 29번 짱 H-주제 영역
55	메르스3	서울 지역 29번 짱 H-주제 영역	156	정말로1	서울 지역 29번 짱 H-주제 영역
56	면목없다 1	서울 지역 24번 짱 J-농사회 행사	157	정신차리 다2	서울 지역 29번 짱 H-주제 영역
57	모델1	서울 지역 29번 짱 H-주제 영역	158	제주도1	서울 지역 24번 짱 J-농사회 행사
58	모르다2	서울 지역 19번 짱 J-농사회 행사	159	졸리다1	서울 지역 1번 짱 E-토론
59	목차1	서울 지역 1번 짱 E-토론	160	종합3	서울 지역 24번 짱 J-농사회 행사
60	무릎1	서울 지역 29번 짱 H-주제 영역	161	중복2	서울 지역 19번 짱 E-토론
61	무역1	서울 지역 1번 짱 E-토론	162	진주1	서울 지역 29번 짱 H-주제 영역
62	무조건3	서울 지역 29번 짱 H-주제 영역	163	진주2	서울 지역 29번 짱 H-주제 영역
63	미리1	서울 지역 29번 짱 H-주제 영역	164	찌그러지 다1	서울 지역 29번 짱 H-주제 영역
64	밀알선교 단1	서울 지역 29번 짱 H-주제 영역	165	참다3	서울 지역 29번 짱 H-주제 영역
65	바퀴1	서울 지역 29번 짱 H-주제 영역	166	창문내리 다1	서울 지역 29번 짱 H-주제 영역
66	반반하다 1	서울 지역 29번 짱 H-주제 영역	167	철저하다1	서울 지역 1번 짱 E-토론
67	반하다1	서울 지역 29번 짱 H-주제 영역	168	촉박하다1	서울 지역 24번 짱 J-농사회 행사
68	브레이크 밟다1	서울 지역 29번 짱 H-주제 영역	169	최고2	서울 지역 29번 짱 H-주제 영역
69	방귀1	서울 지역 29번 짱 M-농학교 경험 이야기	170	추돌1	서울 지역 29번 짱 H-주제 영역
70	배신2	서울 지역 29번 짱 H-주제 영역	171	추석1	서울 지역 29번 짱 H-주제 영역
71	번개1	서울 지역 29번 짱 H-주제 영역	172	충고	서울 지역 1번 짱 E-토론
72	베트남1	서울 지역 29번 짱 H-주제 영역	173	충돌1	서울 지역 29번 짱 H-주제 영역

73	변명1	서울 지역 29번 째 H-주제 영역	174	치다1	서울 지역 29번 째 H-주제 영역
74	복지1	서울 지역 29번 째 H-주제 영역	175	치다2	서울 지역 29번 째 H-주제 영역
75	복수하다 2	서울 지역 24번 째 J-농사회 행사	176	칠년전1	서울 지역 19번 째 J-농사회 행사
76	부채춤1	서울 지역 29번 째 H-주제 영역	177	칠월1	서울 지역 29번 째 H-주제 영역
77	부채춤2	서울 지역 29번 째 H-주제 영역	178	택시2	서울 지역 1번 째 E-토론
78	부터2	서울 지역 19번 째 E-토론	179	터키1	서울 지역 29번 째 H-주제 영역
79	비교2	서울 지역 19번 째 E-토론	180	트럭2	서울 지역 29번 째 H-주제 영역
80	비교3	서울 지역 29번 째 H-주제 영역	181	파악하다3	서울 지역 19번 째 J-농사회 행사
81	비위1	서울 지역 1번 째 E-토론	182	파악하다4	서울 지역 29번 째 M-농학교 경험 이야기
82	비키다1	서울 지역 29번 째 H-주제 영역	183	파악하다4	서울 지역 29번 째 M-농학교 경험 이야기
83	빌붙다1	서울 지역 1번 째 E-토론	184	팔년전1	서울 지역 19번 째 J-농사회 행사
84	사단법인 1	서울 지역 1번 째 E-토론	185	퍼지다2	서울 지역 24번 째 J-농사회 행사
85	사십분1	서울 지역 29번 째 M-농학교 경험 이야기	186	평균1	서울 지역 24번 째 J-농사회 행사
86	사춘기1	서울 지역 29번 째 H-주제 영역	187	평등1	서울 지역 1번 째 E-토론
87	살빼다1	서울 지역 24번 째 J-농사회 행사	188	평생1	서울 지역 24번 째 J-농사회 행사
88	삼박사일 1	서울 지역 29번 째 H-주제 영역	189	포크레인1	서울 지역 1번 째 E-토론
89	삼십구1	서울 지역 29번 째 H-주제 영역	190	한살 많다1	서울 지역 19번 째 J-농사회 행사
90	삼십사1	서울 지역 29번 째 H-주제 영역	191	할말이없 다3	서울 지역 24번 째 J-농사회 행사
91	삼십시간	서울 지역 1번 째 E-토론	192	힘담1	서울 지역 19번 째 J-농사회 행사
92	삼십칠1	서울 지역 29번 째 H-주제 영역	193	혈압1	서울 지역 29번 째 H-주제 영역
93	생각없다	서울 지역 24번 째	194	화나다3	서울 지역 29번 째

	1	J-농사회 행사			H-주제 영역
94	서투르다 2	서울 지역 19번 쌍 J-농사회 행사	195	환갑잔치1	서울 지역 29번 쌍 H-주제 영역
95	설레다1	서울 지역 29번 쌍 H-주제 영역	196	환기1	서울 지역 29번 쌍 H-주제 영역
96	세번째2	서울 지역 1번 쌍 E-토론	197	횡령2	서울 지역 19번 쌍 E-토론
97	소진하다 1	서울 지역 29번 쌍 H-주제 영역	198	횡령3	서울 지역 19번 쌍 E-토론
98	수다2	서울 지역 19번 쌍 E-토론	199	후원1	서울 지역 19번 쌍 J-농사회 행사
99	수다3	서울 지역 29번 쌍 H-주제 영역	200	본인1	
100	수다4	서울 지역 24번 쌍 J-농사회 행사	201	피곤하다1	
101	수입1	서울 지역 29번 쌍 H-주제 영역			

2.4.2. 전사 품질 개선

본 연구팀에서는 전사 작업 초기에 여러 가지 문제에 직면하였다. 수어에 대한 직관이 없는 청인 전사자들은 어떤 것이 수어이고 어떤 것이 제스처인지 구분하는 일에서부터 혼란을 겪었고, 한국수어 자료 통합 지원 시스템에 등록된 타입을 검색하는 데에서도 많은 어려움이 있었다. 직관이 있는 농인 전사자들 사이에서도 어떤 타입을 선택하여 전사해야 할 것인지에 대해 의견이 분분하기도 하였다.

이에 청인 전사자들은 수차례의 공동 전사와 소모임을 통해 전사 역량을 기르는 데 주력하였고, 농인 전사자들과의 정기 회의를 통해 서로 의견을 교환하였다. 전사 책임자가 개별적으로 전사 결과물을 확인하고 수정하는 것이 아니라, 모든 전사자들이 함께 모여 문제 사항을 검토하고 해결하는 방식으로 전사 품질을 개선하였다. 청인 전사자들은 결정된 사안을 반영하여 스스로 전사 오류를 수정하였으며, 이를 통해 전사에 대한 감각을 익혔다.

여기에서는 전사 회의 과정에서 주요하게 나타난 오류와 전사 품질 개선 사례를 제시한다.

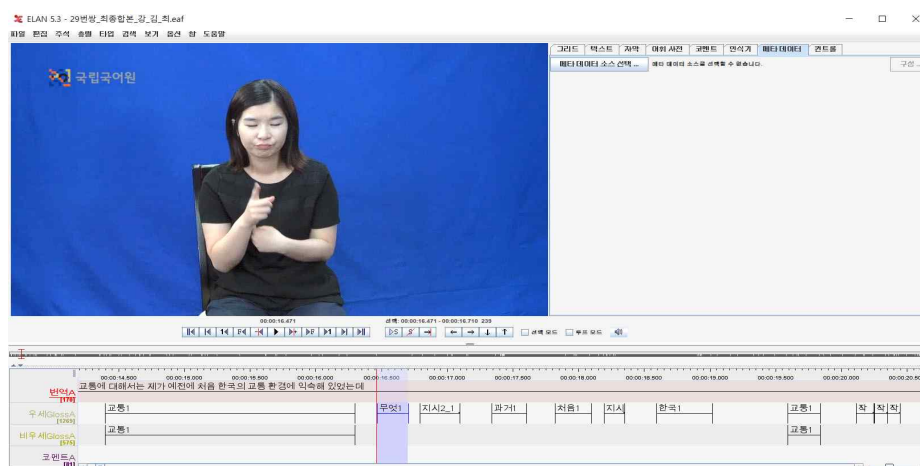
1) 잘못된 타입 선택

가장 많이 나타난 문제는 수형이 비슷한 수어를 혼동하여 잘못된 타입으로 전사하는 오류였다. 하나의 수어 단어를 제대로 인식하지 못하고 둘로 나누어 전사하는 경

우도 많았다. 그러한 사례를 제시하면 다음과 같다.

[무엇1]

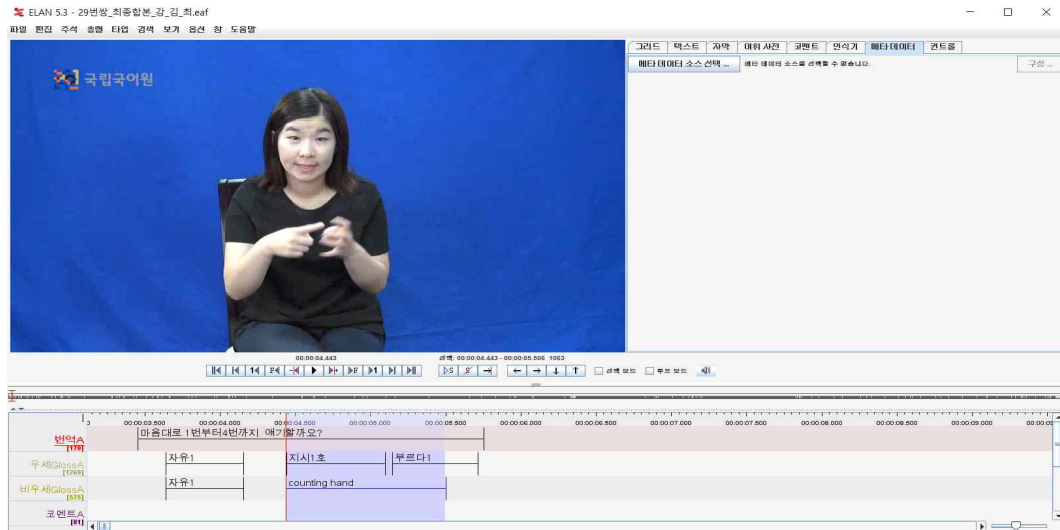
- 토큰 맥락: [교통1]+[무엇1]+[지시2_1A]+[과거1]+[처음1]+[지시2_1A]+[한국1]+[교통1]+[작다1]+[익숙하다1]
- 번역: 교통에 대해서는 제가 예전에 처음 한국의 교통 환경에 익숙해 있었는데
- 전사 오류: [일1]
- 개선 내용: [무엇1]과 유사한 수형의 [일1]로 전사하는 오류가 있었다. 그러나 해당 수어는 [일1]과 달리 좌우로 손을 흔드는 동작이 있으므로 [무엇1]로 수정하였다.



<그림 33> 잘못된 타입 선택 예시 [무엇1]

[지시1호]+[counting hand]

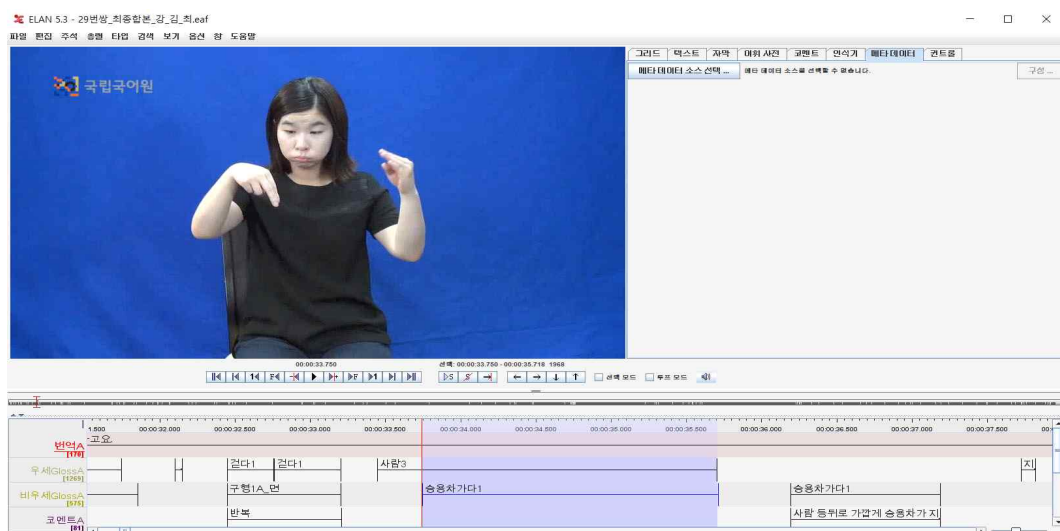
- 토큰 맥락: [자유1]+[지시1호]+[counting hand]+[부르다1]
- 번역: 마음대로 1번부터 4번까지 얘기할까요?
- 전사 오류: [사계절1]
- 개선 내용: 동작상 유사하게 보이는 [사계절1]로 전사하는 오류가 있었다. 그러나 해당 수어는 우세손이 반복적으로 움직이므로 한 번만 움직이는 [사계절1]과 동일한 타입으로 처리할 수 없다. 따라서 이는 생산적 어휘로 보고 [지시1호]+[counting hand]로 수정하였다.



<그림 34> 잘못된 타입 선택 예시 [지시1호]+[counting hand]

[승용차가다1]+[승용차가다1]

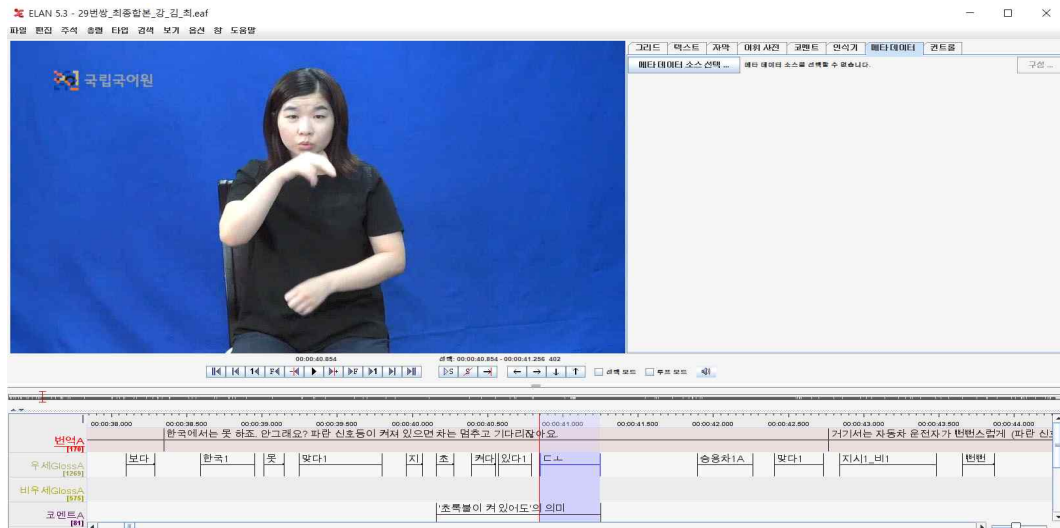
- 토큰 맥락: [제스처1]+[충격1]+[예3]+[지시1_비1]+[초록1A]+[켜다1]+[있다1]+[지시1_1]+[걷다1]+[사람3]+[승용차가다1]+[승용차가다1]+[지시1_1]+[보다7]
- 번역: 엄청 놀랐던 것은, 예를 들어 파란 신호등이 켜져 있어서 횡단보도를 건너 가고 있었는데 차가 바로 내 뒤에서 바짝 붙어 지나갔더라고요.
- 전사 오류: [승용차1A]
- 개선 내용: 승용차가 사람 뒤에 바짝 붙는 동작을 [승용차1A]로 단순 전사하는 오류가 있었다. 또한 동작이 둘로 나누어지는 것을 분절하지 않고 하나로 이어서 전사하기도 하였다. 그러나 이는 승용차의 각기 다른 움직임을 표현하는 수어로 보고 [승용차가다1]+[승용차가다1]로 수정하였다.



<그림 35> 잘못된 타입 선택 예시 [승용차가다1]+[승용차가다1]

[ㄷㅅ]

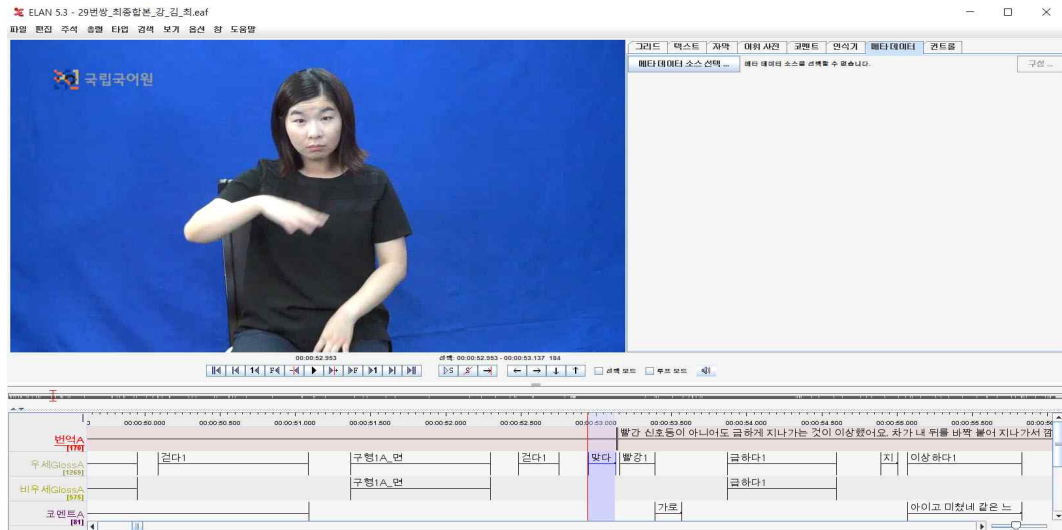
- 토큰 맥락: [한국1]+[못하다1]+[맞다1]+[지시1_비1A]+[초록1A]+[켜다1]+[있다1]+[ㄷㅅ]+[승용차1A]+[맞다1]
- 번역: 한국에서는 못하죠. 안 그래요? 파란 신호등이 켜져 있으면 차는 멈추고 기다리잖아요.
- 전사 오류: [사람3]+[오다1]
- 개선 내용: 지문자 ‘ㄷ’을 [사람3]으로, ‘ㅅ’를 [오다1]로 전사하는 오류가 있었다. 그러나 이는 보조사 ‘도’를 나타내는 것으로 인식되어 [ㄷㅅ]로 수정하였다.



<그림 36> 잘못된 타입 선택 예시 [ㄷㅅ]

[맞다1]

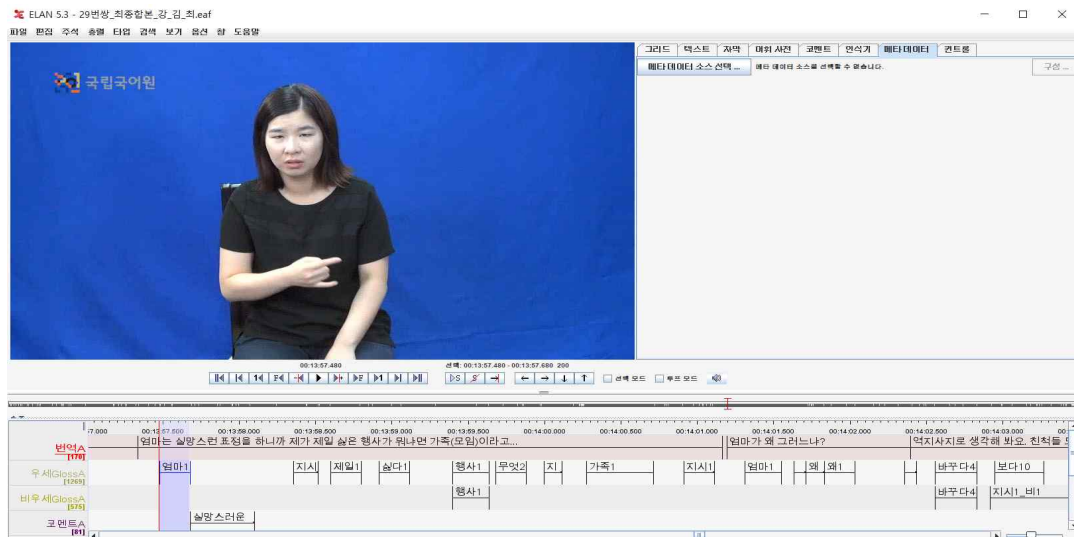
- 토큰 맥락: [초록1A]+[켜다1]+[사람1]+[걷다1]+[구형4]+[켜다1]+[걷다1]+[구형1A_면]+[걷다1]+[맞다1]
- 번역: 파란 신호등이 켜질 때 횡단보도를 건너가는 게 당연한데
- 전사 오류: [기다리다1]
- 개선 내용: [맞다1]의 수형이 다소 기울어져 나타난 것을 [기다리다1]로 인식하고 잘못 전사한 오류가 발생하여 올바르게 수정하였다.



<그림 37> 잘못된 타입 선택 예시 [맞다1]

[엄마1]

- 토큰 맥락: [엄마1]+[지시1_1A]+[제일1]+[싫다1]+[행사1]+[무엇2]+[지시1_1]+[가족1]+[지시1_비1]
- 번역: 엄마는 실망스런 표정을 하니까 제가 제일 싫은 행사가 뭐냐면
- 전사 오류: [지시1_비1B]
- 개선 내용: [엄마1]을 낮은 수위로 발화한 것을 지시로 인식하고 잘못 전사한 오류가 발생하여 올바르게 수정하였다.

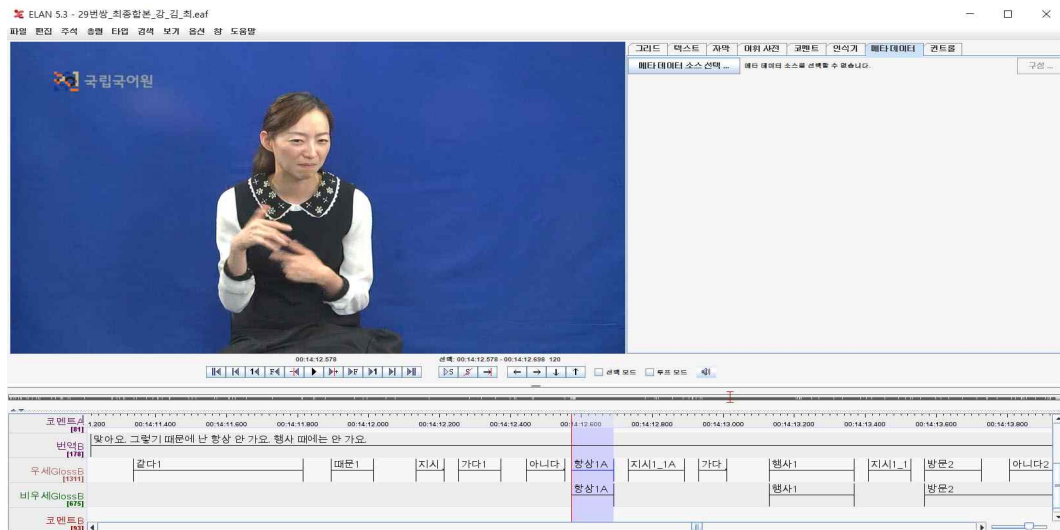


<그림 38> 잘못된 타입 선택 예시 [엄마1]

[항상1A]

- 토큰 맥락: [같다1]+[때문1]+[지시1_1A]+[가다1]+[아니다2]+[항상1A]+[지시1_1A]+[행사1]+[지시1_1]+[방문2]+[아니다2]

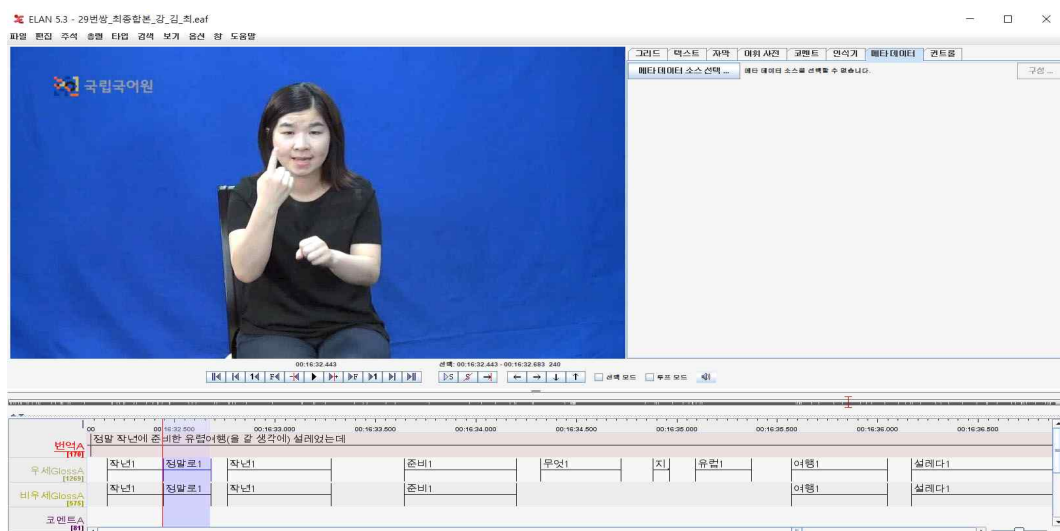
- 번역: 맞아요. 그렇기 때문에 난 항상 안 가요. 행사 때에는 안 가요.
- 전사 오류: [아니다1]
- 개선 내용: 동작상으로 [아니다1]에 가깝게 발화되어 잘못 인식한 오류가 있었다. 그러나 맥락상 [항상1A]가 더 적합한 것으로 판단하여 올바르게 수정하였다.



<그림 39> 잘못된 타입 선택 예시 [항상1A]

[정말로1]

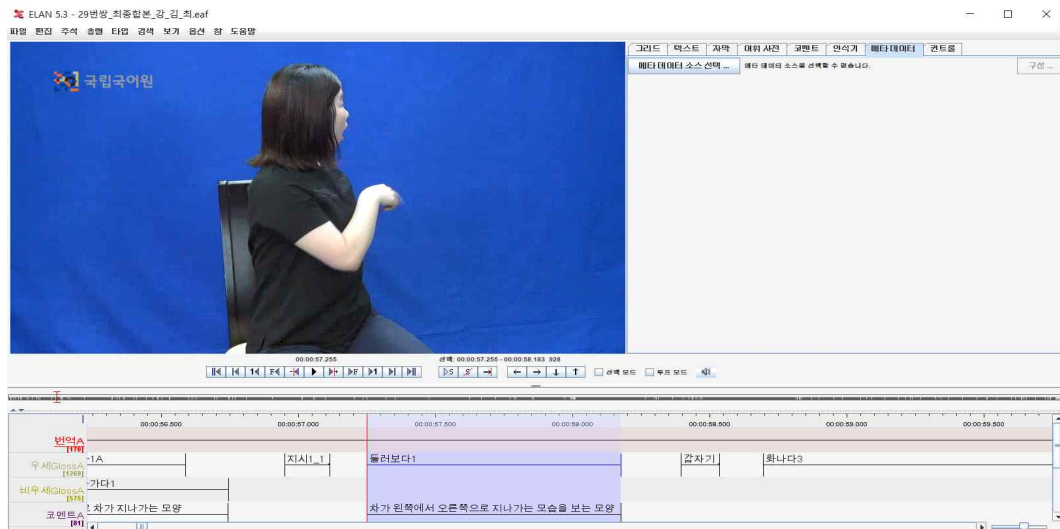
- 토큰 맥락: [작년1]+[정말로1]+[작년1]+[준비1]+[무엇1]+[유럽1]+[여행1]+[설레다1]
- 번역: 정말 작년에 준비한 유럽 여행(을 갈 생각에) 설레었는데
- 전사 오류: [맞다1]
- 개선 내용: 비슷한 수형의 [맞다1]로 전사한 경우가 있었다. 그러나 해당 수어는 '맞다'와 달리 명백하게 부사적 쓰임을 갖는 수어이므로 별도의 타입으로 구분해야 한다고 판단하였다. 따라서 [정말로1]을 새로운 타입으로 선정하여 전사하였다.



<그림 40> 잘못된 타입 선택 예시 [정말로1]

[돌러보다1]

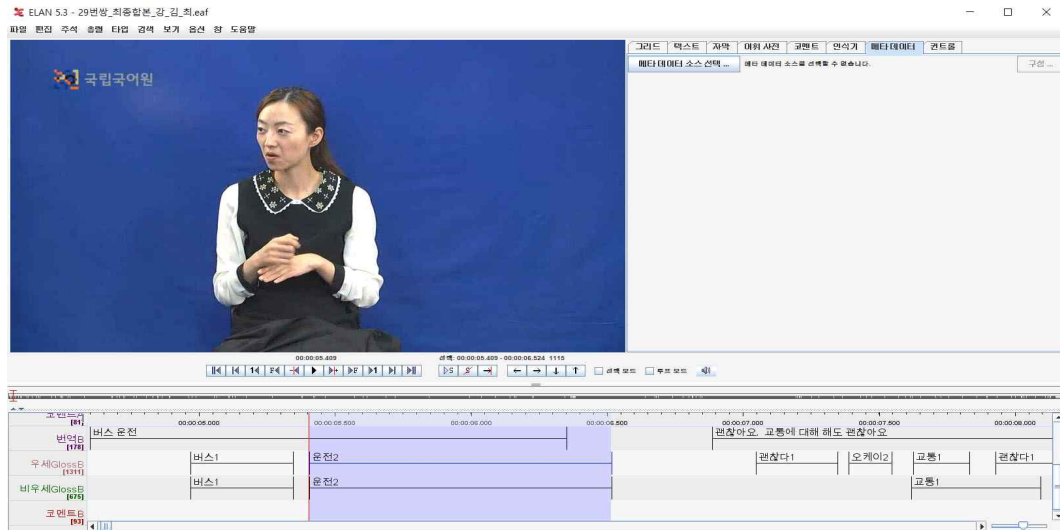
- 토큰 맥락: [빨강1]+[급하다1]+[지시1_1]+[이상하다1]+[승용차1A]+[승용차가다1]+[지시1_1A]+[돌러보다1]+[갑자기2]+[화나다3]
- 번역: 빨간 신호등이어도 급하게 지나가는 것이 이상했어요. 차가 내 뒤를 바짝 붙어 지나가서 깜짝 놀랐기도 했지만 한편으로는 화가 났어요.
- 전사 오류: [지시2_비1] / [보다16]
- 개선 내용: 차를 바라보는 동작을 지시로 인식하여 [지시2_비1]로 전사하는 오류가 있었다. 또한 수형이 유사한 [보다16]으로 전사하는 오류도 있었다. 그러나 [보다16]은 해당 수어와 수동이 다르므로 동일한 수어로 처리할 수 없다. 해당 수어는 동작상 [돌러보다1]과 일치하지만, 맥락상 ‘주위를 이리저리 두루 살펴보는’ 것이 아니라 단순히 차가 지나가는 동작을 ‘보는’ 것을 표현하고 있으므로 코멘트에 맥락 정보를 추가로 입력하기로 결정하였다(‘차가 왼쪽에서 오른쪽으로 지나가는 모습을 보는 모양’).



<그림 41> 잘못된 타입 선택 예시 [돌러보다1]

[운전2]

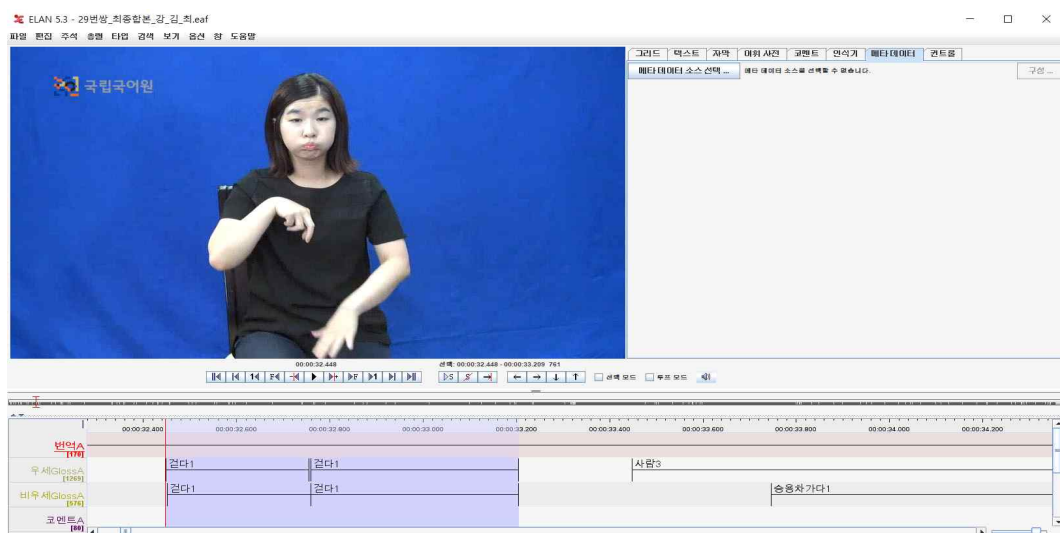
- 토큰 맥락: [버스1]+[운전2]
- 번역: 버스 운전
- 전사 오류: [운전2]+[구형1A_면]
- 개선 내용: 비우세손의 위치가 [운전2]와 달라서 우세손은 [운전2]로, 비우세손은 [구형1A_면]으로 분리하여 전사하는 오류가 있었다. 그러나 비우세손도 [운전2]를 구성하는 수어 동작으로 볼 수 있으므로 모두 [운전2]로 수정하였다.



<그림 42> 잘못된 타입 선택 예시 [운전2]

[걷다1]

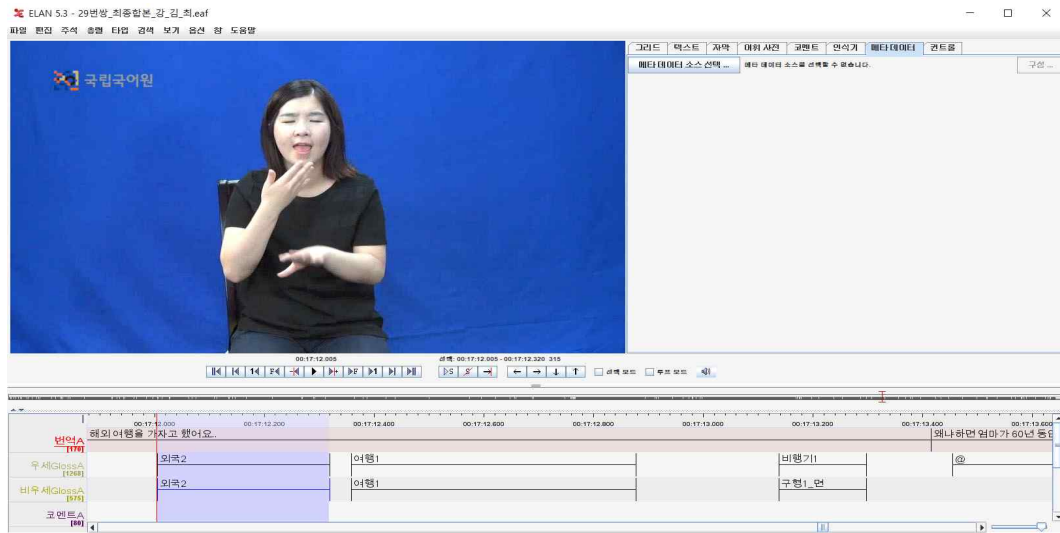
- 토큰 맥락: [제스처]+[충격1]+[예3]+[지시1_비1]+[초록1A]+[켜다1]+[있다1]+[지시1_1]+[걷다1]+[사람3]+[승용차가다1]+[승용차가다1]+[지시1_1]+[보다기]
- 번역: 엄청 놀랐던 것은, 예를 들어 파란 신호등이 켜져 있어서 횡단보도를 건너 가고 있었는데 차가 바로 내 뒤에서 바짝 붙어 지나갔더라고요.
- 전사 오류: [걷다1]+[구형1A_면]
- 개선 내용: 해당 수어의 비우세손을 [구형1A_면]으로 전사하는 오류가 있었다. 실제로 [걷다1]의 타입 영상에는 우세손의 동작만 있으나, 이는 비우세손까지 모두 사용하는 기본 동작의 변이형으로 보인다. 따라서 이를 나누어 전사하지 않고 모두 [걷다1]로 전사하기로 결정하였다.



<그림 43> 잘못된 타입 선택 예시 [걷다1]

[외국2]

- 토큰 맥락: [외국2]+[여행1]+[비행기1]+[구형1_면]
- 번역: 해외여행을 가자고 했어요.
- 전사 오류: [다르다1]+[나라1B]
- 개선 내용: 외국(다른 나라)을 표현하는 수어 동작을 [다르다1]와 [나라1B]로 분리하여 전사한 경우가 있었다. 그러나 이는 하나의 수어 단어라는 판단하에 [외국2]를 새로운 타입으로 선정하여 전사하였다.

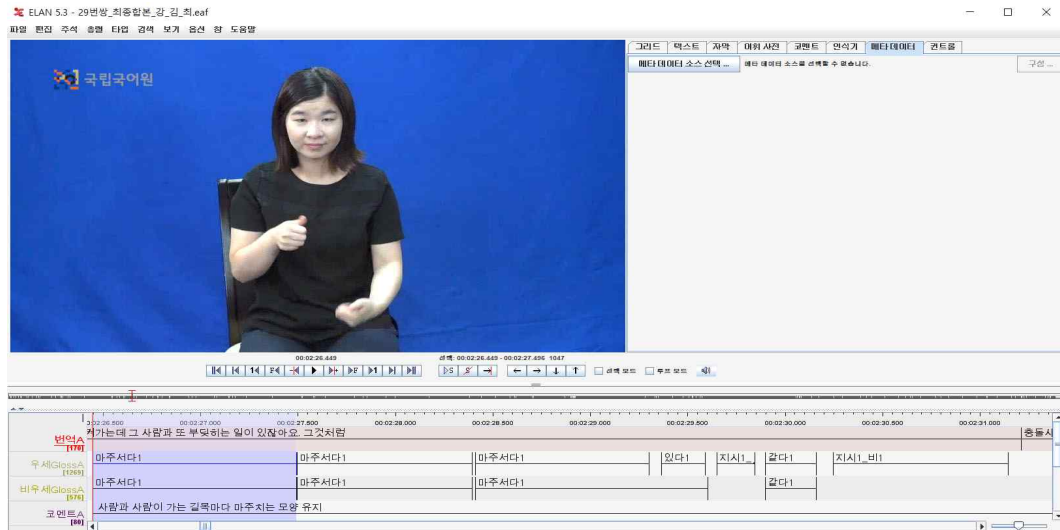


<그림 44> 잘못된 타입 선택 예시 [외국2]

[마주서다1]

- 토큰 맥락: [지하철1]+[지시1_1A]+[내리다2]+[걸다1]+[마주서다1]+[마주서다1]+[마주서다1]+[있다1]
- 번역: 지하철에서 내려 지나가는데 반대쪽에서 오는 사람을 비켜간다고 갔는데 그 사람에게 막히고 방향을 돌려 다른 쪽으로 비켜 가는데 그 사람과 또 부딪히는 일이 있잖아요.
- 전사 오류: [사람4] / [맞붙다1]
- 개선 방향: 번역 부분의 ‘반대쪽에서 오는 사람을 비켜간다고 갔는데 그 사람에게 막히고 방향을 돌려 다른 쪽으로 비켜 가는데 그 사람과 또 부딪히는 일이 있잖아요.’라는 부분이 모두 하나의 수어 동작으로 설명된 경우인데, 우세손과 비우세손 모두 [사람4]의 수형으로 마주 부딪치는 동작이었기 때문에 처음에 [사람4]로 전사하는 오류가 있었다. 유사한 수동과 수형을 띄는 [맞붙다1]과의 차이에 대해서도 논의가 있었다. 그러나 [맞붙다1]은 시합이나 대결을 할 때 싸우기 위해 서로 달려드는 모습이고, 해당 영상에서의 수어 동작은 가는 길목마다 앞사람과 마주치는 모습으로 참여 농인 연구진들 모두 두 동작을 다른 수어, 다른 의미로 봐야한다는 의견이 있었다. 최종적으로 [마주서다1] 타입을 새로 만들어 매치하였

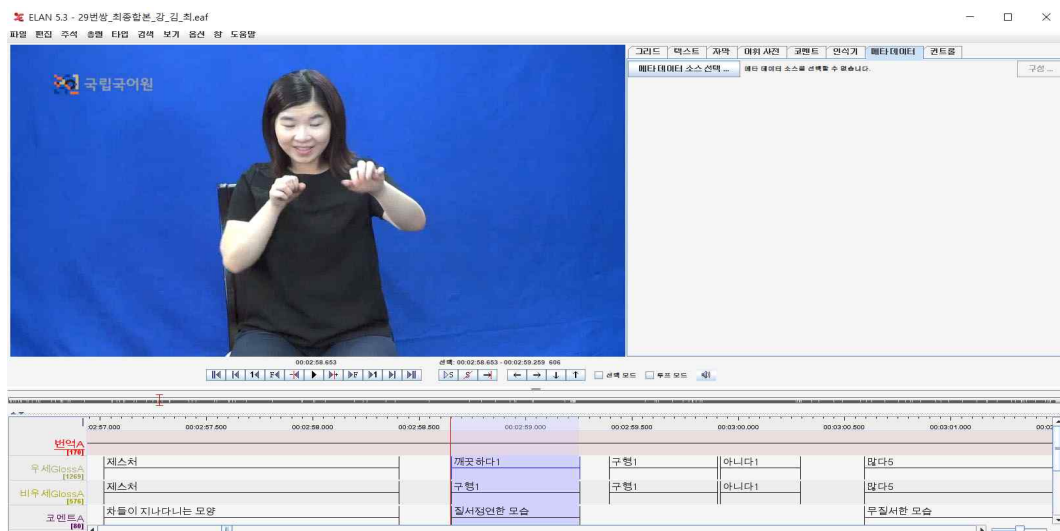
다.



<그림 45> 잘못된 타입 선택 예시 [마주서다1]

[깨끗하다1]+[구형1]

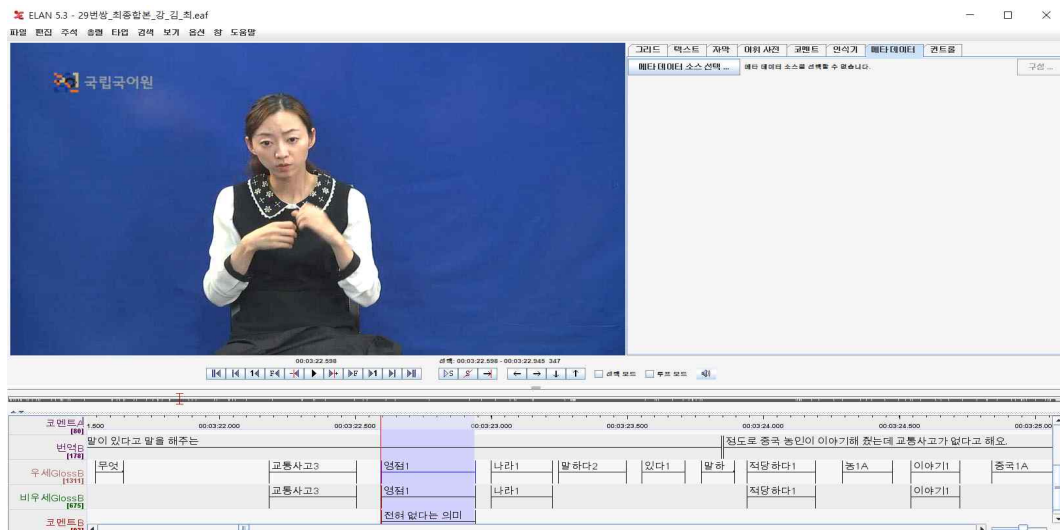
- 토큰 맥락: [가다1A]+[제스처]+[깨끗하다1]+[구형1]+[구형1]+[아니다1]
- 번역: 자동차들이 질서 있게 움직이는 것이 아니라
- 전사 오류: [찍다2]
- 개선 방향: 우세손의 수형과 수동이 [찍다2]의 우세손 동작과 유사하다고 판단하여 잘못 전사하는 오류가 있었다. 그러나 비우세손의 동작까지 고려했을 때 [찍다2]와 동작이 다를 뿐 아니라, 전체적인 번역의 맥락을 고려했을 때도 ‘질서’라는 의미가 ‘찍다’로 표현될 수 없다고 보았다. 차량들이 질서정연하게 가는 모습을 표현하고자 한 것이기 때문에 우세손은 [깨끗하다1]로, 비우세손은 [구형1]로 수정하였다.



<그림 46> 잘못된 타입 선택 예시 [깨끗하다1]+[구형1]

[영점1]

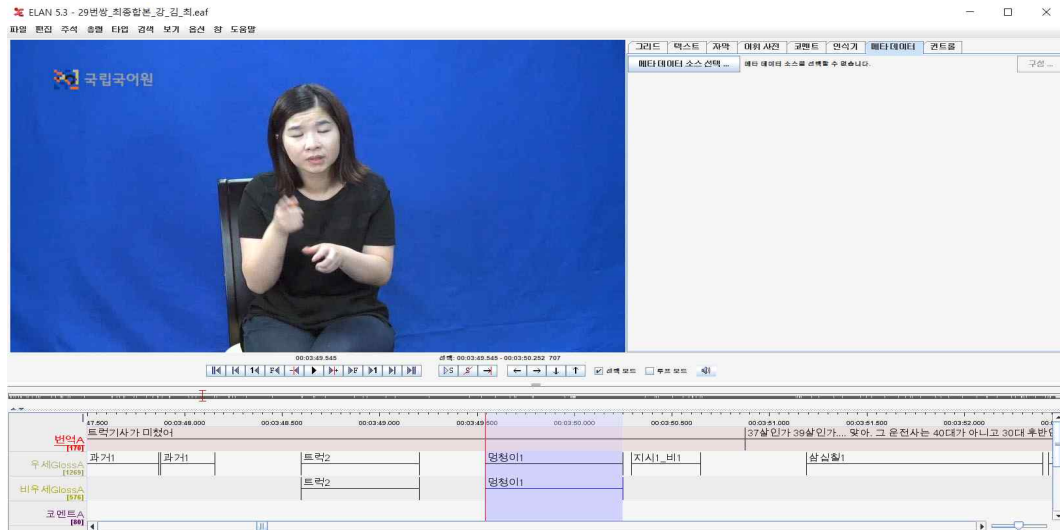
- 토큰 맥락: [교통사고3]+[영점1]+[나라1]+[말하다2]+[있다1]
- 번역: 교통사고가 없는 나라라는 말이 있다고
- 전사 오류: [없다1]+[없다4]
- 개선 방향: 우세손으로는 [없다1]을, 비우세손으로는 [없다4]를 나타내는 동작이 있었다. ‘없다’라는 같은 의미를 나타내는 동작이 우세손과 비우세손에서 서로 나타난다는 점이 특이했지만, 수형이 분명하고 맥락상으로도 ‘없다’에 해당하는 의미가 와야 하므로, 해당 글로스로 전사하였다. 그러나 이것은 ‘0점’을 나타내는 일종의 약어로 쓰인다는 농인 연구진들의 의견이 있었고, 따라서 ‘영점’이라는 새 타입을 만들어 매치하였다.



<그림 47> 잘못된 타입 선택 예시 [영점1]

[명청이1]

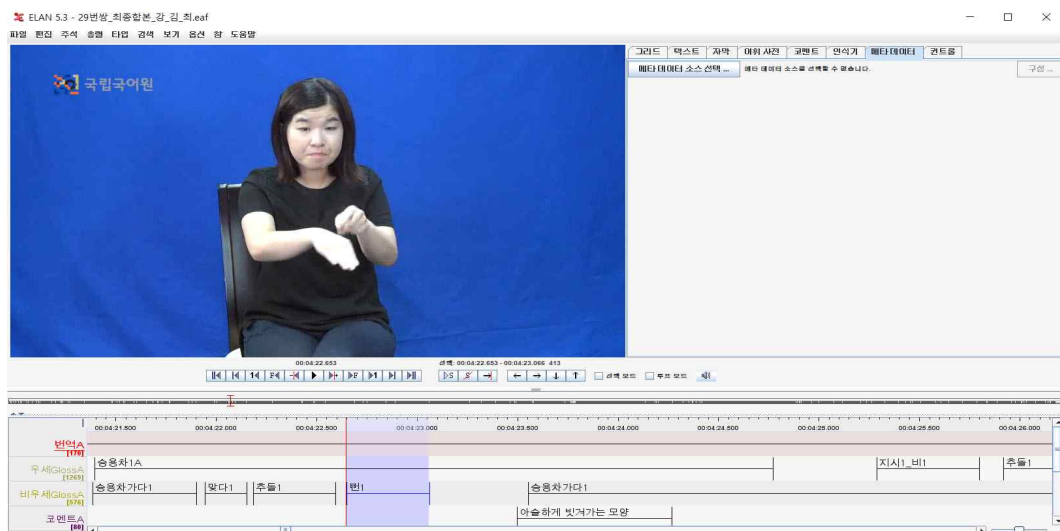
- 토큰 맥락: [과거1]+[과거1]+[트럭2]+[명청이1]+[지시1_비1]
- 번역: 예전에 그 트럭기사는 미쳤어.
- 전사 오류: [생각1]+[비상식1]
- 개선 방향: 맥락상 ‘미치다, 머리가 어떻게 되다, 바보이다, 무모했다’라는 의미인데, 해당 타입과는 동작이 달라서 어떤 타입인지 찾을 수 없었다. 결국 동작을 쪼개어 [생각1]과 [비상식1]로 나누어 전사하는 오류가 있었다. 그러나 농인 연구자들이 유의어들을 더 찾아보고 주위 다른 농인들에게 해당 표현에 대해 물어본 결과, 수어 사전에 있는 ‘명청이, 명청하다’와 동작이 같은 것을 포착하였다. 따라서 [명청이1]을 새로운 타입으로 만들어 매치하였다.



<그림 48> 잘못된 타입 선택 예시 [멍청이1]

[뻔1]

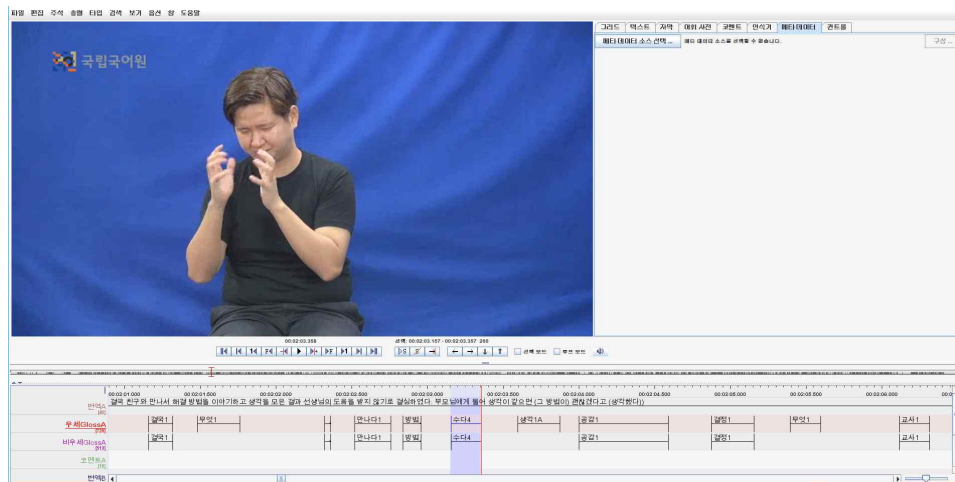
- 토큰 맥락: [승용차가다1]+[맞다1]+[추돌1]+[뻔1]+[승용차가다1]
- 번역: 추돌 사고가 날 듯 아슬아슬하게
- 전사 오류: [말하다1]
- 개선 방향: 수위가 입 아래쪽이었기 때문에 [말하다1]로 오인하여 전사하는 오류가 있었다. 그러나 [뻔1]은 턱을, [말하다1]은 입을 가리킨다는 측면에서 수위에 차이가 있다. 또한 해당 부분 맥락을 고려하였을 때도 [뻔1]이 더 자연스럽게 때문에, [뻔1]로 수정하였다.



<그림 49> 잘못된 타입 선택 예시 [뻔1]

[수다4]

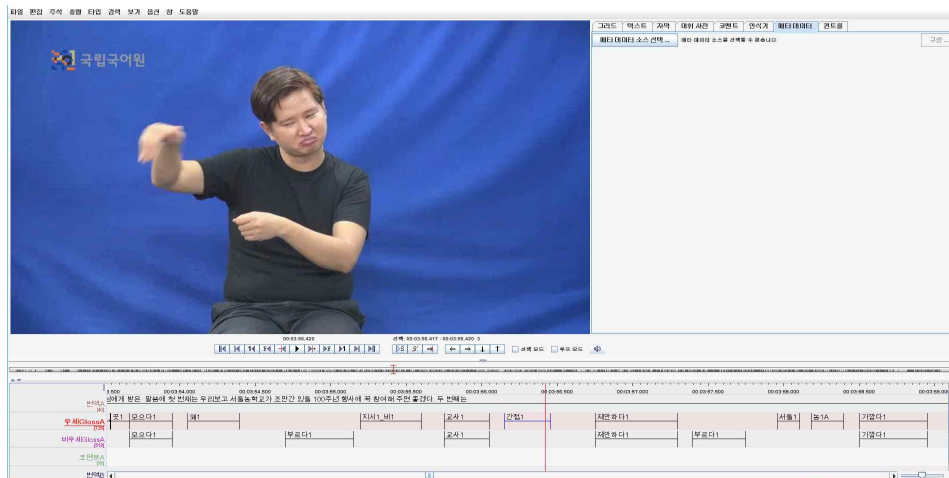
- 토큰 맥락: [친구1]+[만나다1]+[방법1]+[수다4]+[생각1A]+[공감1]
- 번역: 결국 친구와 만나서 해결 방법을 이야기하고 생각을 모은 결과
- 전사 오류: [모으다3]
- 개선 내용: 기존의 [모으다1]의 타입과 수형이 비슷하여 [모으다3]이라는 새로운 타입으로 전사하였으나, 점검 결과 해당 동작은 ‘친구와 이야기를 하는 것’을 의미하므로 [수다4]라는 새로운 타입을 설정하는 것이 타당하다고 보았다.



<그림 50> 잘못된 타입 선택 예시 [수다4]

[간접1]

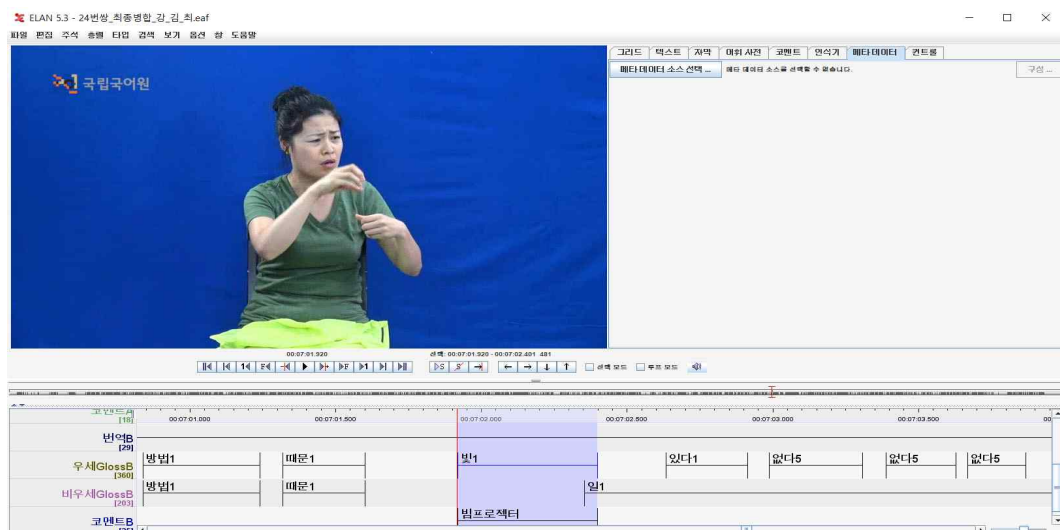
- 토큰 맥락: [부르다1]+[지시1_비1]+[교사1]+[간접1]+[제안하다1]
- 번역: (그 친구는) 선생님에게 받은 말씀에
- 전사 오류: [부르다1]
- 개선 내용: 수동이 유사하다고 판단하여 [부르다1]로 전사하는 오류가 있었다. 그러나 선생님께서 하신 말씀을 친구가 대신 우리들에게 전달하는 맥락에서 나온 동작으로 ‘간접적으로’라는 의미를 가진 수어이기 때문에 [간접1]이라는 새로운 타입으로 수정하였다.



<그림 51> 잘못된 타입 선택 예시 [간접1]

[빛1]

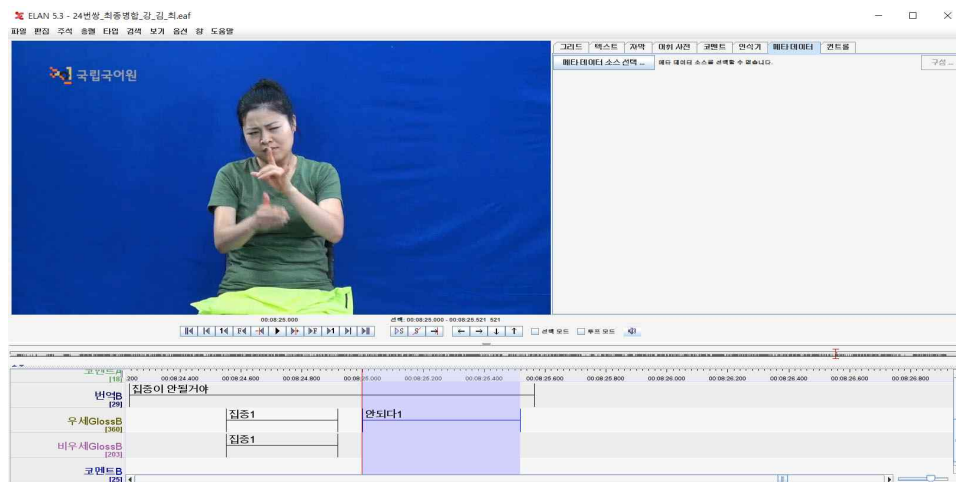
- 토큰 맥락: [쓰다1]+[완전1]+[결국1]+[쓰다1]+[시간2]+[아끼다1]+[방법1]+[때문1]+[빛1]+[일1]+[있다1]+[없다5]+[없다5]+[없다5]
- 번역: 필기를 다 끝내고 나중에 필기를 하는 것이 시간을 절약하는 방법이기 때문에 빔 프로젝터가 하나 있고 다른 교실에는 없어요.
- 전사 오류: [켜다1] / [그냥1]
- 개선 내용: [빛1]과 동작상·의미상 유사성을 지닌 [켜다1]로 전사하거나, 혹은 유사한 동작의 [그냥1]로 전사하는 오류가 있었다. 그러나 이는 [켜다1], [그냥1]과 구별되는 수어라는 판단하에 [빛1]로 수정하였다.



<그림 52> 잘못된 타입 선택 예시 [빛1]

[안되다1]

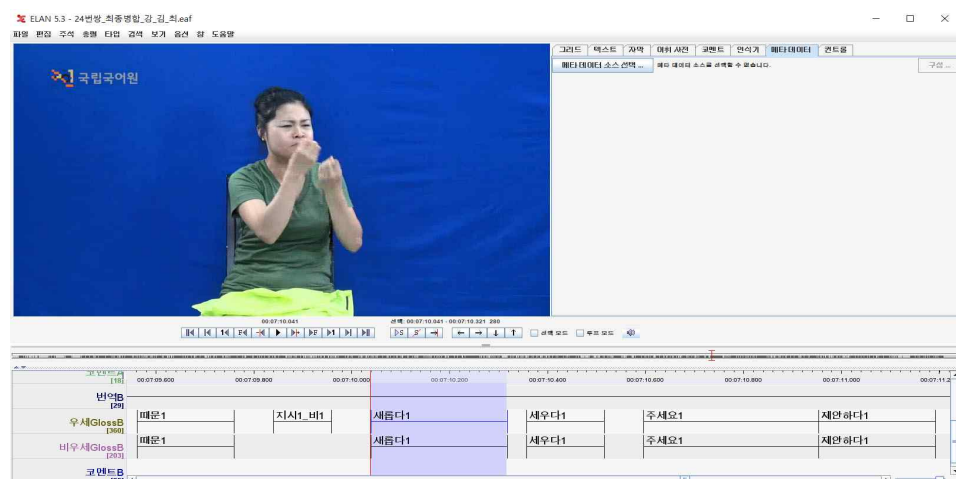
- 토큰 맥락: [집중1]+[안되다1]
- 번역: 집중이 안 될 거야.
- 전사 오류: [불가능3]
- 개선 내용: [안되다1]을 동작상·의미상 유사성 때문에 [불가능3]으로 전사하는 오류가 있었고 이러한 오류가 전반적으로 나타났기 때문에 전체 검토를 통해 수정하였다.



<그림 53> 잘못된 타입 선택 예시 [안되다1]

[새롭다1]

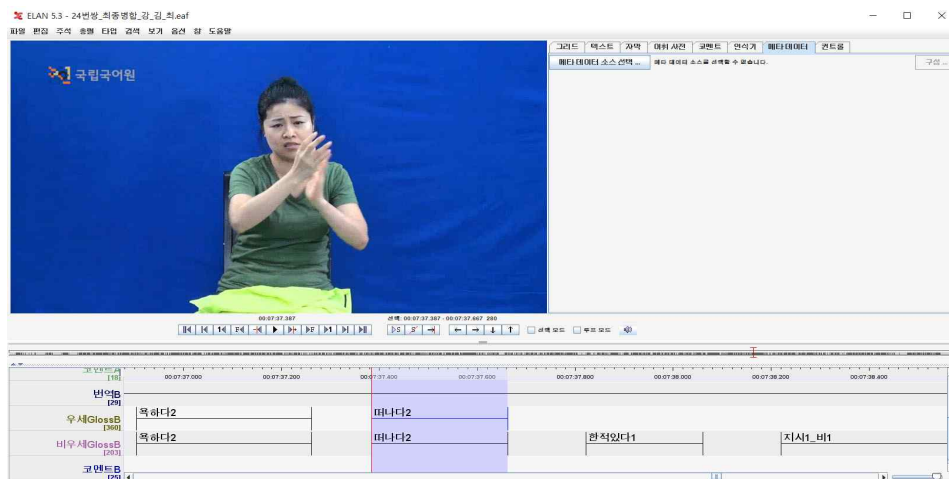
- 토큰 맥락: [새롭다1]+[세우다1]+[주세요1]+[제안하다1]
- 번역: 새로 설치해 달라고 했는데
- 전사 오류: [놀라다4]
- 개선 내용: [새롭다1]과 [놀라다4]의 동일성 문제가 제기되었으나 구별되는 단어
이므로 두 타입 모두 병치하는 것으로 결정하였다.



<그림 54> 잘못된 타입 선택 예시 [새롭다1]

[떠나다2]

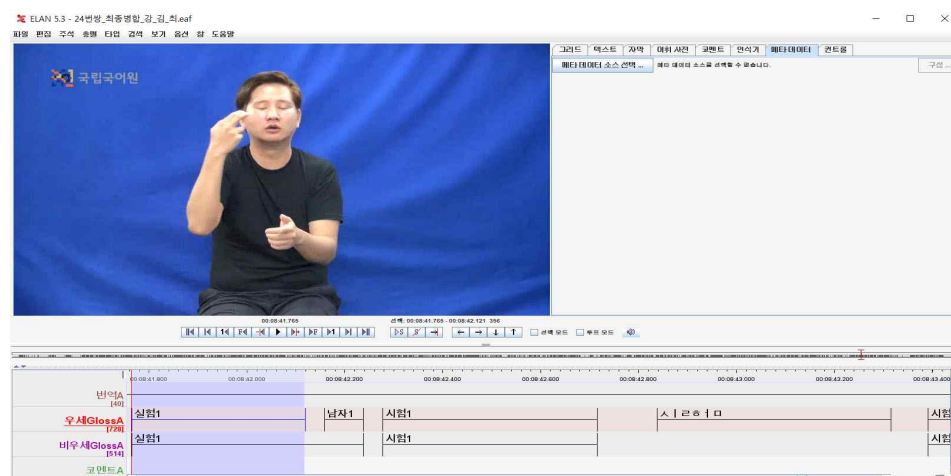
- 토큰 맥락: [고등학교삼학년1]+[졸업1]+[욕하다2]+[떠나다2]+[한적있다1]+[지시1_비1]
- 번역: 고3 졸업하고 나가면서 학교를 향하여 욕하고 간 적이 있었어요.
- 전사 오류: [떠나다1]
- 개선 내용: 해당 수어는 [떠나다1]과 우세손은 동일하지만 비우세손 동작이 다르다. 따라서 새로운 타입 [떠나다2]를 만들어 전사하기로 결정하였다.



<그림 55> 잘못된 타입 선택 예시 [떠나다2]

[실험1]

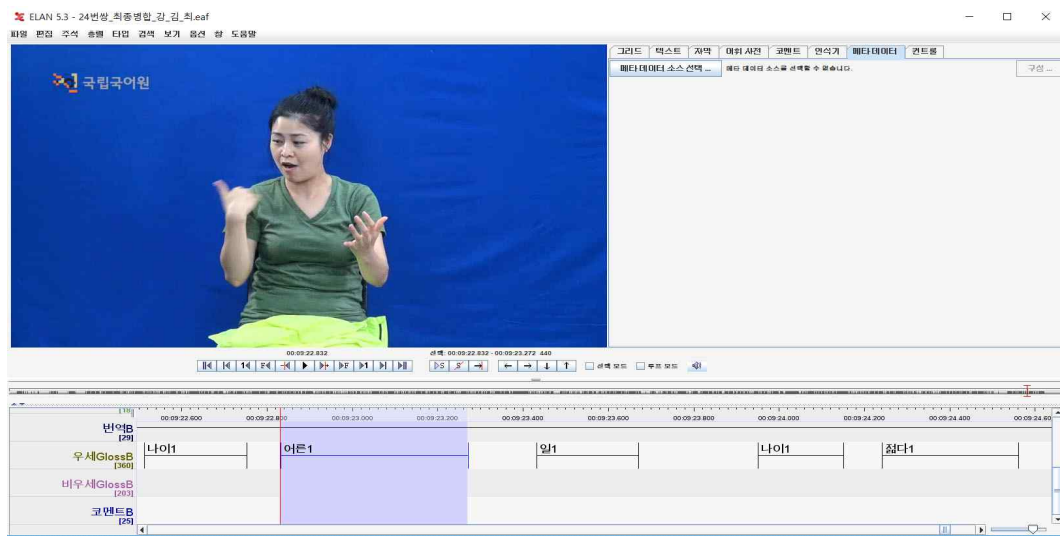
- 토큰 맥락: [실험1]+[남자1]+[시험1]+[스 | 르ㅎ | 모]
- 번역: 선생님은 실험하고 있다고 하니
- 전사 오류: [농담2]+[시험1] / [SN]+[시험1]
- 개선 내용: [시험1] 앞에 짧게 들어가는 수어 동작을 [농담2]와 [SN]으로 혼동하였으나 해당 수어는 분리할 수 없는 하나의 타입이라는 판단하에, [실험1]이라는 새로운 타입을 만들어 전사하기로 결정하였다.



<그림 56> 잘못된 타입 선택 예시 [실험1]

[어른1]

- 토큰 맥락: [나이1]+[어른1]+[일1]+[나이1]+[젊다1]+[대화1A]+[가능1]+[사람1A]
- 번역: 나이가 어른인 사람 한 명과 대화가 가능한 젊은 사람 몇 명이
- 전사 오류: [자라다4]+[사람1]
- 개선 내용: ‘어른’이라는 의미의 수어를 [자라다4]와 [사람1]로 분절하여 전사하는 오류가 있었다. 그러나 이는 별개의 두 단어라기보다 하나의 수어 단어라는 판단 하에 [어른1]이라는 타입을 새로 만들어 전사하기로 결정하였다.



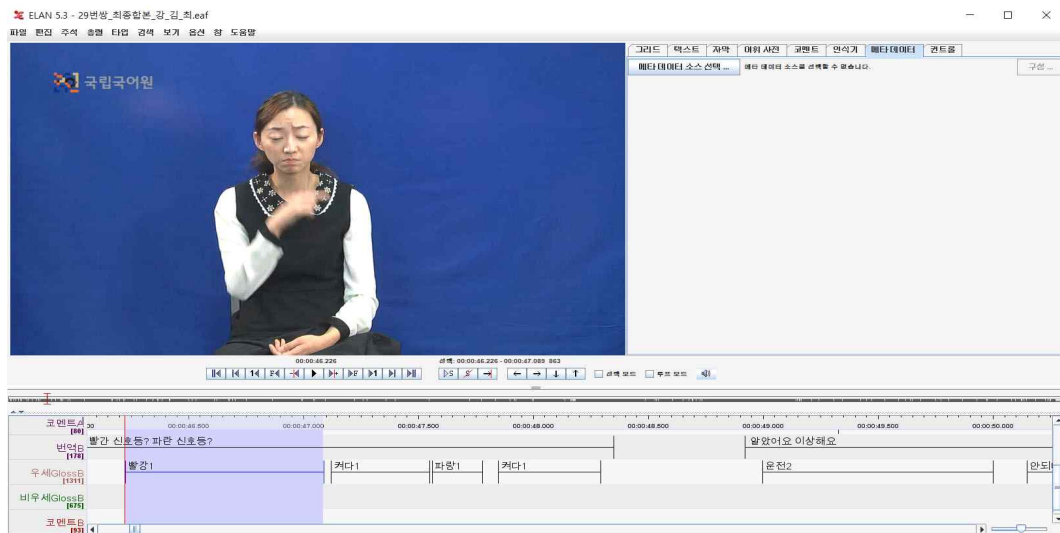
<그림 57> 잘못된 타입 선택 예시 [어른1]

2) 타입 생성 오류

새로운 타입을 생성할 때에도 오류가 있었다. 굳이 새로운 타입을 만들 필요 없이 기존 타입으로 전사할 수 있는 경우, 타입명이 적절하지 않은 경우 등이다. 그러한 사례를 제시하면 다음과 같다.

[빨강1]+[켜다1]+[파랑1]+[켜다1]

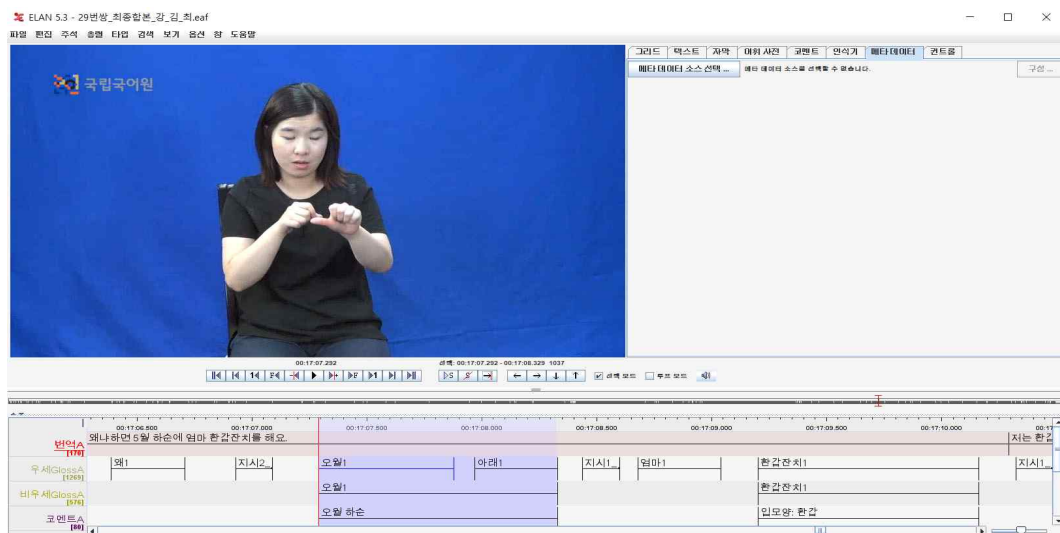
- 토큰 맥락: [빨강1]+[켜다1]+[파랑1]+[켜다1]
- 번역: 빨간 신호등? 파란 신호등?
- 전사 오류: [빨간불1]+[파란불1]
- 개선 내용: ‘빨간불’과 ‘파란불’은 국어에서 새로운 의미를 갖는 합성어이므로 [빨간불1], [파란불1]이라는 새로운 타입을 만들어야 한다는 의견이 있었다. 그러나 이는 농인에게 하나의 단어로 인식되기보다 별도의 두 단어로 인식된다는 판단 하에 각각 분절하여 전사하기로 결정하였다.



<그림 58> 타입 생성 오류 예시 [빨강1]+[켜다1]+[파랑1]+[켜다1]

[오월1]+[아래1]

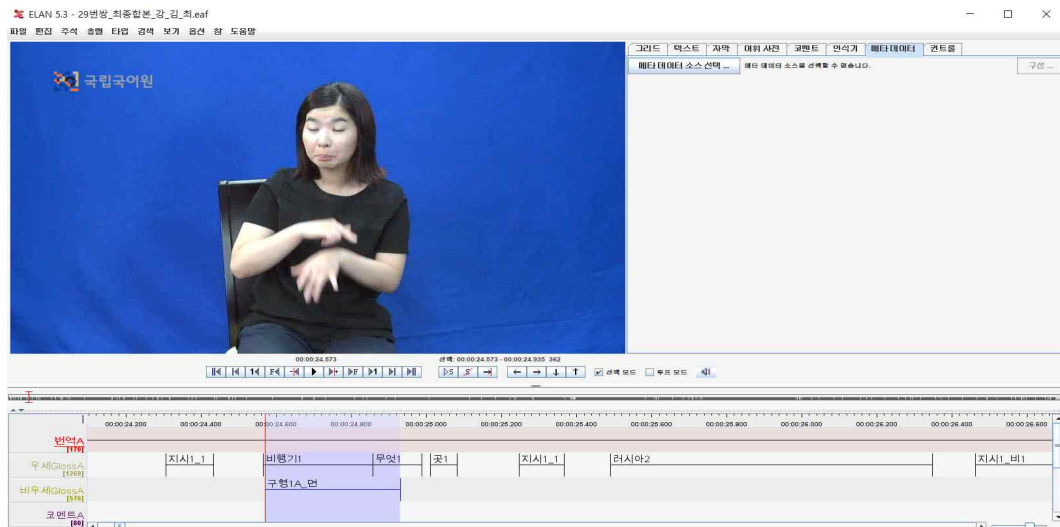
- 토큰 맥락: [왜1]+[지시2_1B]+[오월1]+[아래1]+[지시1_1]+[엄마1]+[환갑잔치1]
- 번역: 왜냐하면 5월 하순에 엄마 환갑잔치를 해요.
- 전사 오류: [오월1]+[하순1]
- 개선 내용: [하순1]을 새로운 타입으로 선정하여 전사할 수 있다는 의견이 있었다. 그러나 이는 [아래1]로 충분히 전사 가능한 것으로 판단하였다.



<그림 59> 타입 생성 오류 예시 [오월1]+[아래1]

[비행기1]+[구형1A_면]

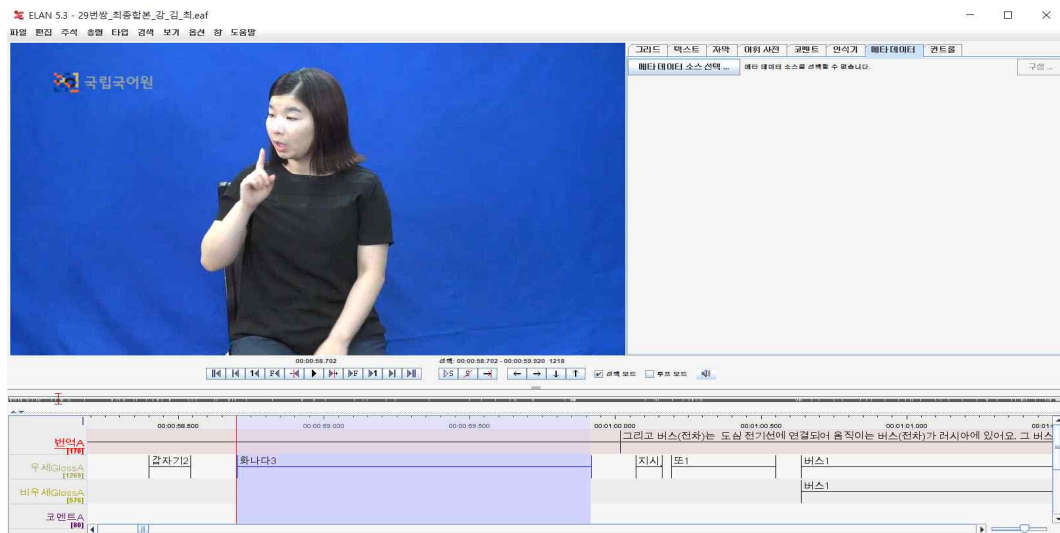
- 토큰 맥락: [비행기1]+[지시1_1]+[비행기1]+[구형1A_면]+[곳1]+[지시1_1]+[러시아2]+[지시1_비1]+[러시아2]
- 번역: 비행기 타고 도착해 보니 그 곳이 러시아였습니다.
- 전사 오류: [비행기가다1]
- 개선 내용: 비행기의 이착륙을 보여주는 해당 수어를 [비행기가다1]라는 새로운 타입을 선정하여 전사해야 한다는 의견이 있었다. 이는 기존에 [버스1]과 [버스가다1], [승용차1]과 [승용차가다1] 등이 별도의 타입으로 선정되어 있는 것을 고려한 것이었다. 그러나 기존 전사 결과물을 보면 비행기가 착륙하는 것을 [비행기1]로 전사한 기록이 있고(서울지역 24번 쌍 M-농학교 경험 이야기 전사, 00:06:06.190), 해당 수어가 [비행기1]과 별개의 타입으로 보이지 않는다는 판단하에 [비행기1]+[구형1A_면]으로 수정하였다.



<그림 60> 타입 생성 오류 예시 [비행기1]+[구형1A_면]

[화나다3]

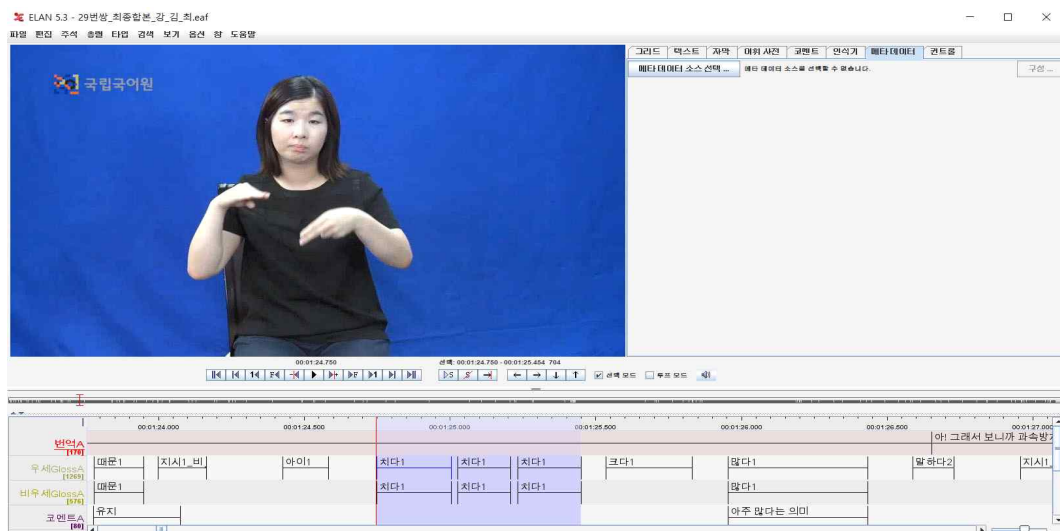
- 토큰 맥락: [빨강1]+[급하다1]+[지시1_1]+[이상하다1]+[승용차1A]+[승용차가다1]+[지시1_1A]+[둘러보다1]+[갑자기2]+[화나다3]
- 번역: 빨간 신호등이어도 급하게 지나가는 것이 이상했어요. 차가 내 뒤를 바짝 붙어 지나가서 깜짝 놀랐기도 했지만 한편으로는 화가 났어요.
- 전사 오류: [열받다1] / [혈압1]
- 개선 내용: 해당 수어가 타입 목록에 없으므로 [열받다1] 혹은 [혈압1] 타입을 새로 만들 수 있다는 의견이 있었다. 그런데 해당 수어는 기등록된 [화나다1], [화나다2] 타입과 의미상 차이가 없다고 판단하여 [화나다3]을 새로운 타입으로 올리기로 결정하였다.



<그림 61> 타입 생성 오류 예시 [화나다3]

[치다1]

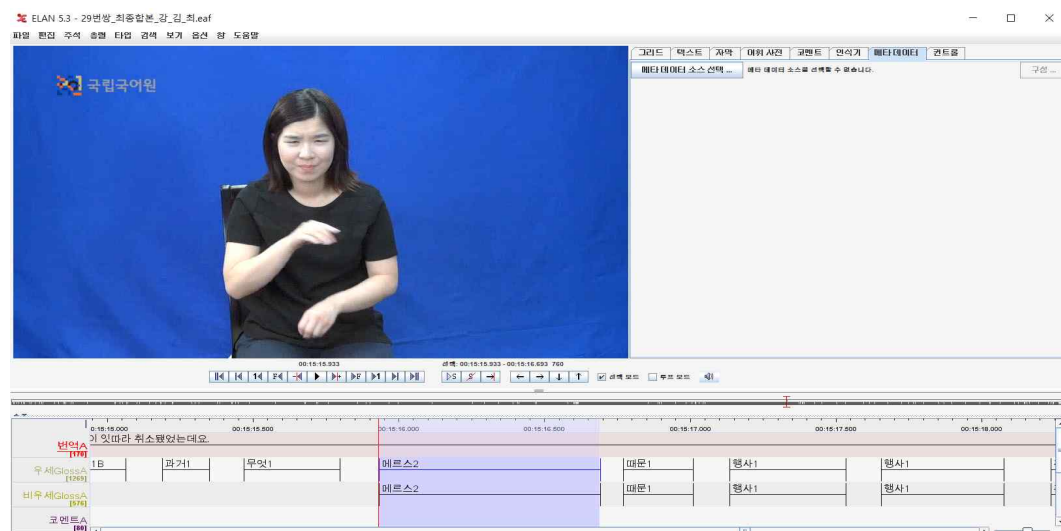
- 토큰 맥락: [지시1_1]+[부르다1]+[왜1]+[지시1_비1]+[러시아2]+[급하다1]+[성격1]+[때문1]+[지시1_비1]+[아이1]+[치다1]+[치다1]+[치다1]+[크다1]+[많다1]
- 번역: 왜 그러느냐고 물어봤는데 러시아 사람들의 조급한 성격 때문에 버스가 너무 빠르게 다녀 어린이들이 버스에 치여 죽는 사고가 많이 생기고 있다고 했습니다.
- 전사 오류: [승용차가다1]+[사람3] / [교통사고4]
- 개선 내용: 승용차가 사람을 치는 모습을 [승용차가다1]와 [사람3]으로 분리하여 전사하는 경우가 있었다. 또는 [교통사고4]를 새로운 타입으로 만들어야 한다는 의견도 있었다. 그러나 ‘교통사고’의 경우 교통 기관끼리의 충돌을 내포하는 경우가 많으므로 [치다1]을 새로운 타입으로 선정하여 전사하기로 결정하였다.



<그림 62> 타입 생성 오류 예시 [치다1]

[메르스2]

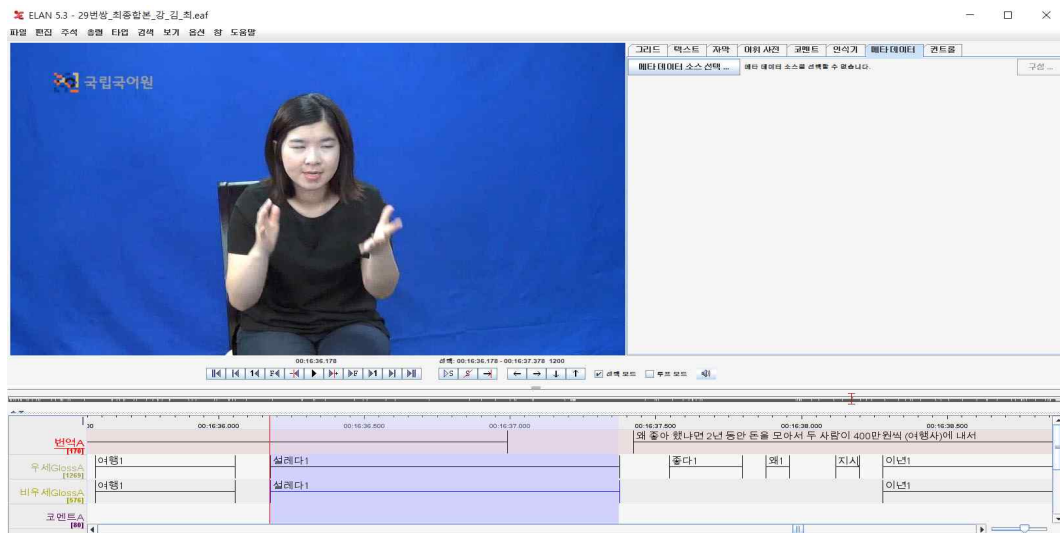
- 토큰 맥락: [다행1]+[무엇1B]+[과거1]+[무엇1]+[메르스2]+[때문1]+[행사1]+[행사1]+[취소1]+[취소1]
- 번역: 다행인 게 메르스 전염병 때문에 여러 행사들이 잇따라 취소됐었는데요.
- 전사 오류: [메르스2]+[전염1] / [낙타1]
- 개선 내용: 메르스를 나타내는 수어 동작에서 '전염'을 분리하여 전사한 경우가 있었다. 또한 해당 수어가 낙타의 혹을 표현하였다는 점에서 메르스가 아닌 [낙타1]로 전사하려는 시도가 있었다. 그러나 해당 수어는 메르스를 표현하는 하나의 단어로 인식할 수 있다는 판단하에 [메르스2]를 새로운 타입으로 선정하여 전사하였다.



<그림 63> 타입 생성 오류 예시 [메르스2]

[설레다1]

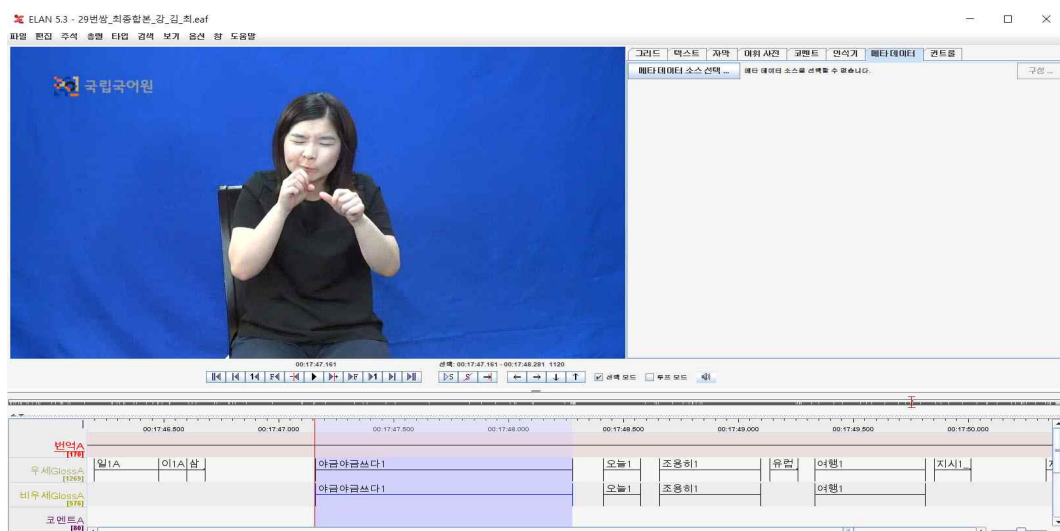
- 토큰 맥락: [작년1]+[정말로1]+[작년1]+[준비1]+[무엇1]+[유럽1]+[여행1]+[설레다1]
- 번역: 정말 작년에 준비한 유럽 여행(을 갈 생각에) 설레었는데
- 전사 오류: [간절1]
- 개선 내용: 동작상 유사한 [간절1]을 새로운 타입으로 선정하여 전사하는 오류가 있었다. 그러나 이는 여행을 기다리는 설렘을 표현하는 수어이므로 [설레다1]로 타입명을 교체하여 전사하였다.



<그림 64> 타입 생성 오류 예시 [설레다1]

[야금야금쓰다1]

- 토큰 맥락: [지시1호]+[counting hand]+[일1A]+[이1A]+[삼1]+[야금야금쓰다1]
- 번역: 휴가일수 16일 중 1~2일씩 조금씩 아껴 쓰면서
- 전사 오류: [뿔다3] / [야금야금1] / [야금야금하다1] / [조금씩아껴쓰다1]
- 개선 내용: 해당 수어를 [뿔다3]의 변이형으로 인식하여 전사하는 경우가 있었으나 이는 동작과 의미가 모두 다른 별개의 수어이며, 적절한 타입이 없어 새로운 타입을 만들기로 결정하였다. ‘조금씩 야금야금 아껴 쓰다’라는 의미의 수어이므로 [야금야금1], [야금야금하다1], [조금씩아껴쓰다1] 등의 타입명이 제안되었는데, 그 중 비교적 간결하면서도 의미를 가장 정확하게 표현할 수 있는 [야금야금쓰다1]로 타입명을 최종 결정하였다.



<그림 65> 타입 생성 오류 예시 [야금야금쓰다1]

[전화하다1]+[구형1]

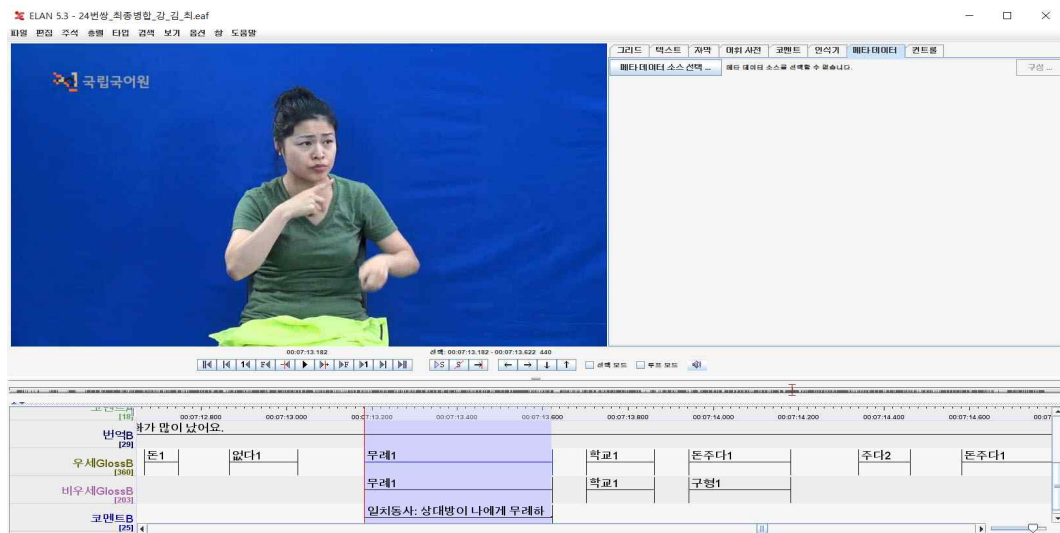
- 토큰 맥락: [일하다1]+[전화하다1]+[구형1]+[듣다1]+[없다1]
- 번역: 일을 하며 지냈는데 연락이 오지 않았어요.
- 전사 오류: [연락하다1]
- 개선 내용: [전화하다1] 타입의 경우 수화기를 귀에 가져다 대는 동작이 보이는데, 해당 동작의 경우 수위에 차이가 있어 [연락하다1]이라는 새로운 타입을 만들어 전사하였다. 그러나 전화기 모양을 의미하는 기본 수형에 차이가 없고, 기존 전사 결과물에서도 번역 '연락하다'에 대응하는 수어 동작이 해당 동작과 동일하며 [전화하다1]로 전사가 되어 있었음을 고려하여 [전화하다1]로 전사하기로 결정하였다.



<그림 66> 타입 생성 오류 예시 [전화하다1]+[구형1]

[무례1]

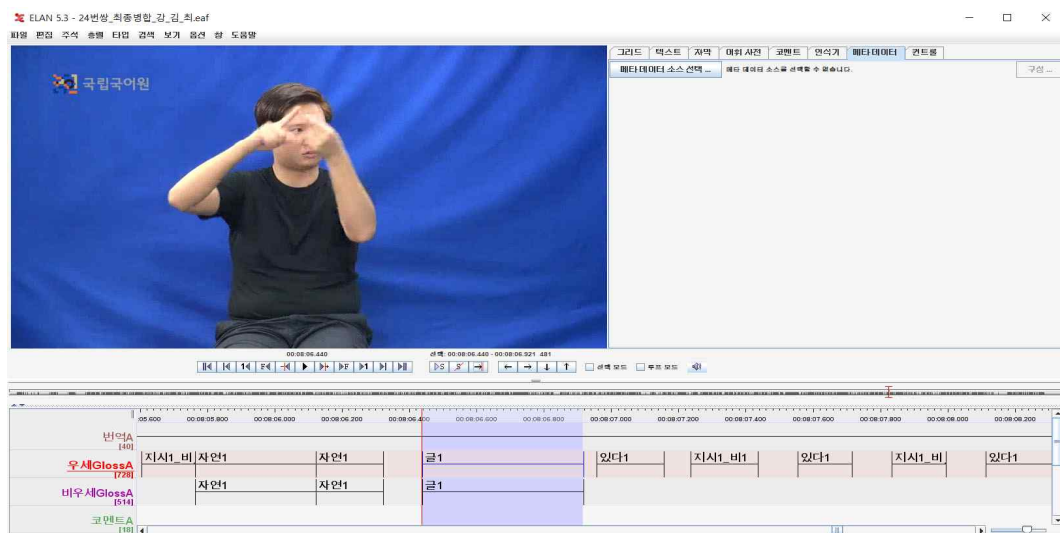
- 토큰 맥락: [돈1]+[없다1]+[지시2_1]+[돈1]+[없다1]+[돈1]+[없다1]+[무례1]+[학교1]+[돈주다1]+[구형1]
- 번역: 선생님은 돈이 없다고, 돈이 없다고 하시면서 우리보고 학교에 돈을 내라고 하셔서
- 전사 오류: [###] / [까불다1] / [굴다1]
- 개선 내용: 해당 수어에 대응하는 적절한 타입을 찾지 못하여 [###]로 남겨 두거나, 혹은 [까불다1]이나 [굴다1]이라는 타입을 새로 만들어야 한다는 의견이 있었다. 그러나 이는 '상대방이 나에게 무례하게 말하다'는 의미로 기존에 구축된 타입 [무례1]과 동일한 수어라는 판단하게 새 타입을 만들지 않고 [무례1]로 전사하기로 결정하였다.



<그림 67> 타입 생성 오류 예시 [무레1]

[글1]

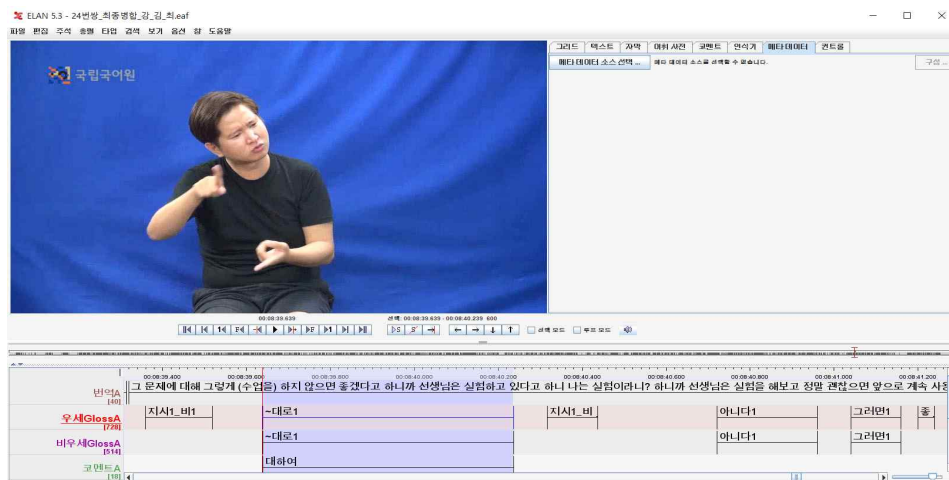
- 토큰 맥락: [지시1_비1]+[자연1]+[자연1]+[글1]+[있다1]
- 번역: 자동으로 자막이 나오니
- 전사 오류: [자막1]
- 개선 내용: 번역 맥락에 의존하여 [자막1]이라는 타입을 새로 만들어 전사해야 한다는 의견이 있었다. 그러나 이는 기존에 등록된 타입 [글1]과 동일한 수어라는 판단하에 새 타입을 만들지 않고 [글1]로 전사하기로 결정하였다.



<그림 68> 타입 생성 오류 예시 [글1]

[~대로1]

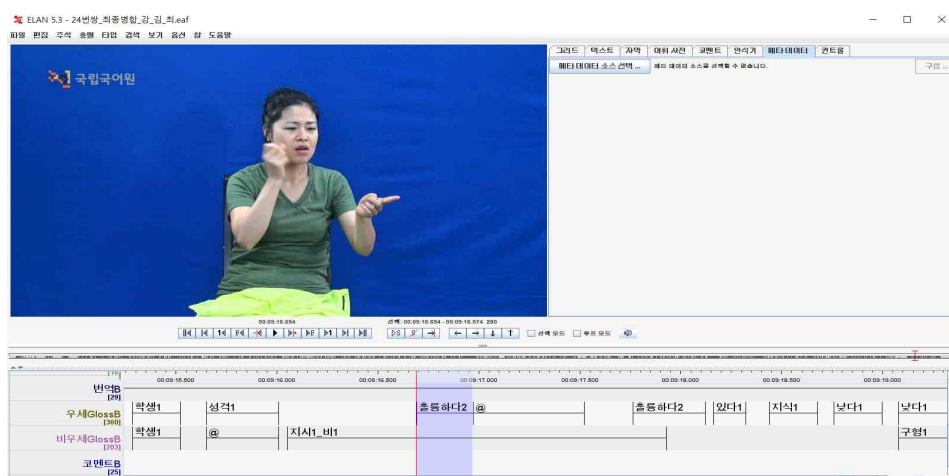
- 토큰 맥락: [지시1_비1]+[~대로1]+[지시1_비1A]+[아니다1]+[그러면1]+[좋다1]
- 번역: 그 문제에 대해 그렇게 (수업을) 하지 않으면 좋겠다고 하니까
- 전사 오류: [대하다1]
- 개선 내용: 한국수어사전에서는 ‘대로’와 ‘대하다’가 동형어가 아니기 때문에 기존 타입 [~대로1]과 구분하여 [대하다1]을 만들 수도 있었으나 수형과 그 쓰임을 볼 때 동일 타입으로 판단하였다.



<그림 69> 타입 생성 오류 예시 [~대로1]

[훌륭하다2]

- 토큰 맥락: [학생1]+[성격1]+[훌륭하다2]+[@]+[훌륭하다2]+[있다]+[지식1]+[낮다1]
- 번역: 학교에서 학생들이 말하는 거... 머리가 좋고 나쁘고에 상관없이
- 전사 오류: [똑똑하다3]
- 개선 내용: 해당 수어를 의미 맥락상 새 타입 [똑똑하다3]으로 만들 수도 있었으나 [훌륭하다2]와 동일한 수어라 판단하여 새 타입을 만들지 않는 것으로 결정했다.



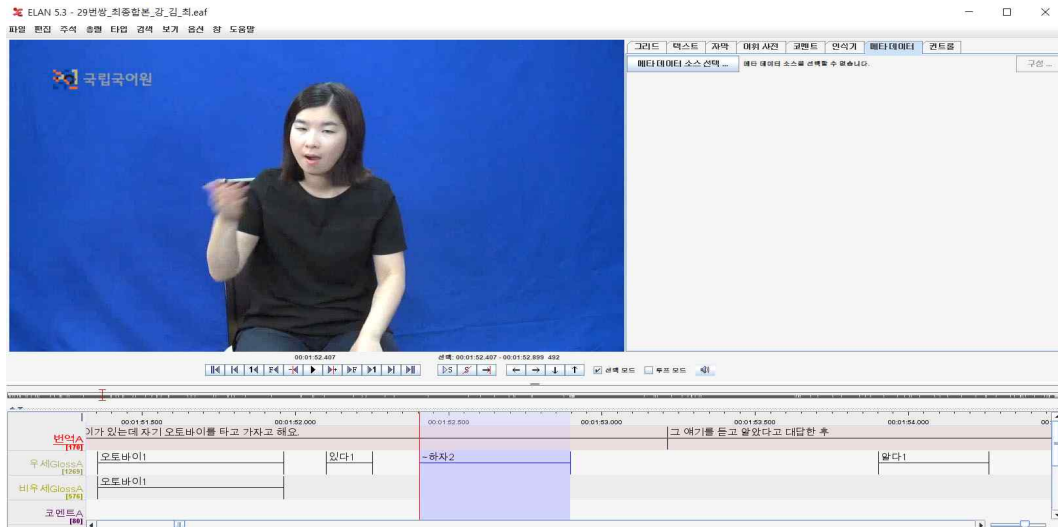
<그림 70> 타입 생성 오류 예시 [훌륭하다2]

3) 수어와 제스처의 구분

수어와 제스처를 구분하지 못하는 문제도 많이 나타났다. 명백한 수어 단어를 단순한 제스처로 오해하거나, 반대로 제스처를 수어 단어로 오해하여 잘못 전사하는 오류였다. 그러한 사례는 다음과 같다.

[~하자2]

- 토큰 맥락: [오토바이1]+[있다1]+[~하자2]
- 번역: 자기 오토바이가 있는데 오토바이를 타고 가자고 해요.
- 전사 오류: [제스처]
- 개선 내용: 청인들도 대강의 의미를 짐작할 수 있다는 측면에서 [제스처]라고 전사하는 오류가 있었다. 그러나 수어에서도 ‘청유형’의 의미를 지닌 문장을 표현할 수 있다. 그에 따라 ‘청유 표현’의 [~하자2]로 수정하였다.



<그림 71> 수어와 제스처의 구분 예시 [~하자2]

[많다5]

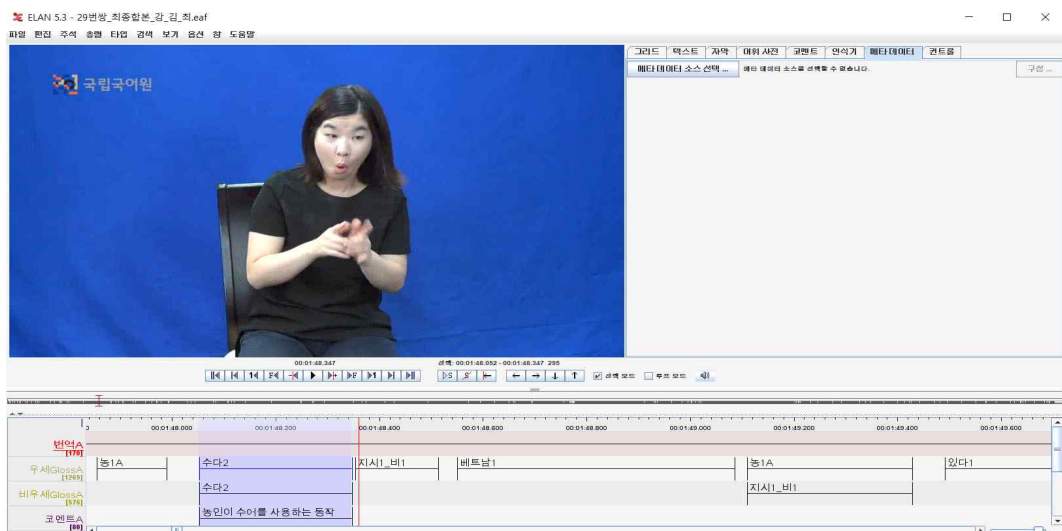
- 토큰 맥락: [많다5]+[승용차1A]+[많다5]+[지시2_1B]+[보다7]+[놀라다8]
- 번역: 서로 얹히고 설켜서 가는 모습의 사진을 보고 놀랐어요.
- 전사 오류: [제스처]
- 개선 내용: 차량들이 무질서하게 지나가는 모습을 나타낸 [제스처]라고 판단하는 오류가 있었다. 그러나 단순 제스처와는 다르게 수형이 정확하게 움직이고 있다는 부분에서 수어로 볼 수 있다는 지적이 있었다. 따라서 ‘우글우글하고 복잡스럽다’는 의미가 들어있는 [많다5]로 수정하였다.



<그림 72> 수어와 제스처의 구분 예시 [말다5]

[수다2]

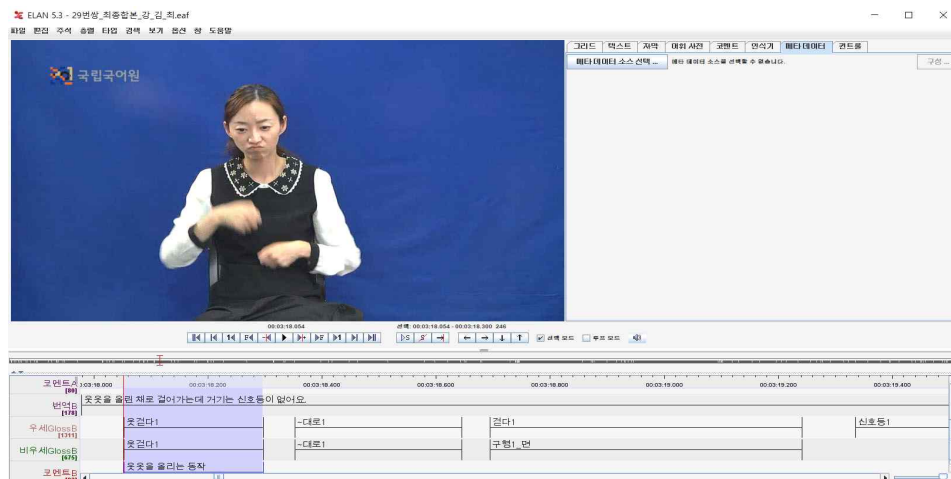
- 토큰 맥락: [농1A]+[수다2]+[지시1]+[베트남1]
- 번역: 수어를 사용하는 베트남 농인
- 전사 오류: [제스처] / [###]
- 개선 내용: 해당 동작을 특별한 의미가 없는 [제스처]로 인식하여 전사하거나, 적절한 글로스를 찾아내지 못해 모르는 수어인 [###]로 표시하는 오류가 있었다. 그러나 이는 하나의 수어 단어로서 기존에 구축된 타입 [수다2]에 해당하는 것으로 보고 올바르게 수정하였다.



<그림 73> 수어와 제스처의 구분 예시 [수다2]

[웃걷다1]

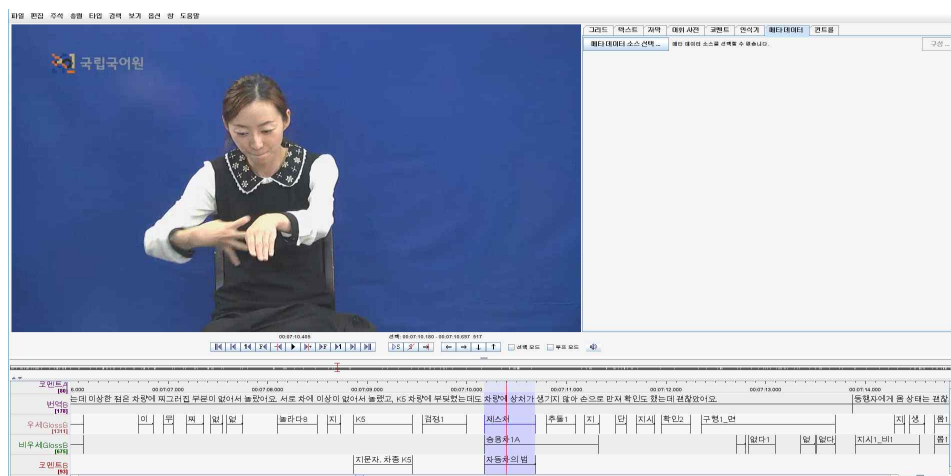
- 토큰 맥락: [웃걷다1]+[~대로1]+[걷다1]
- 번역: 웃웃을 올린 채로 걸어가는데
- 전사 오류: [제스처] / [벗다3]
- 개선 내용: 수어 동작을 봤을 때 [벗다3]과 구별되고 수어 화자의 직관상 수어로 인식되기 때문에 [제스처]가 아닌 새 타입 [웃걷다1]로 결정하였다.



<그림 74> 수어와 제스처의 구분 예시 [웃걷다1]

[제스처]

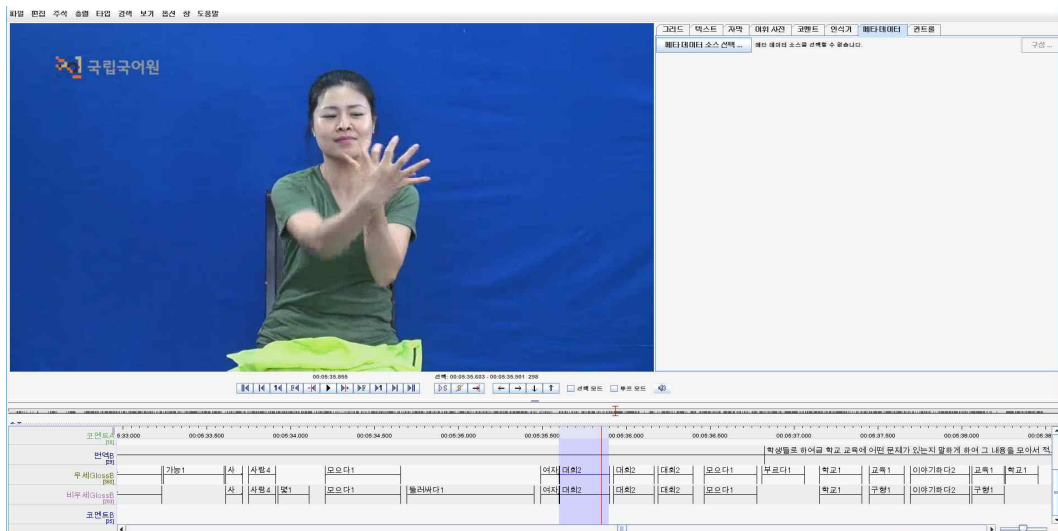
- 토큰 맥락: [검정1]+[제스처]+[승용차1A]+[추돌1]+[지시1_비1]+[단단하다1]
- 번역: K5 차량에 부딪혔는데도
- 전사 오류: [굵히다1]
- 개선 내용: 차에 부딪혀서 굵혔다고 해석하여 새 타입 [굵히다1]로 전사하는 경우가 있었으나 해당 맥락은 차량에 굵혔다는 것이 아니라 충돌 부위가 차량의 범퍼라는 것을 가리키는 동작이기 때문에 이를 [제스처]로 판단하였다.



<그림 75> 수어와 제스처의 구분 예시 [제스처]

[대회2]+[대회2]+[대회2]

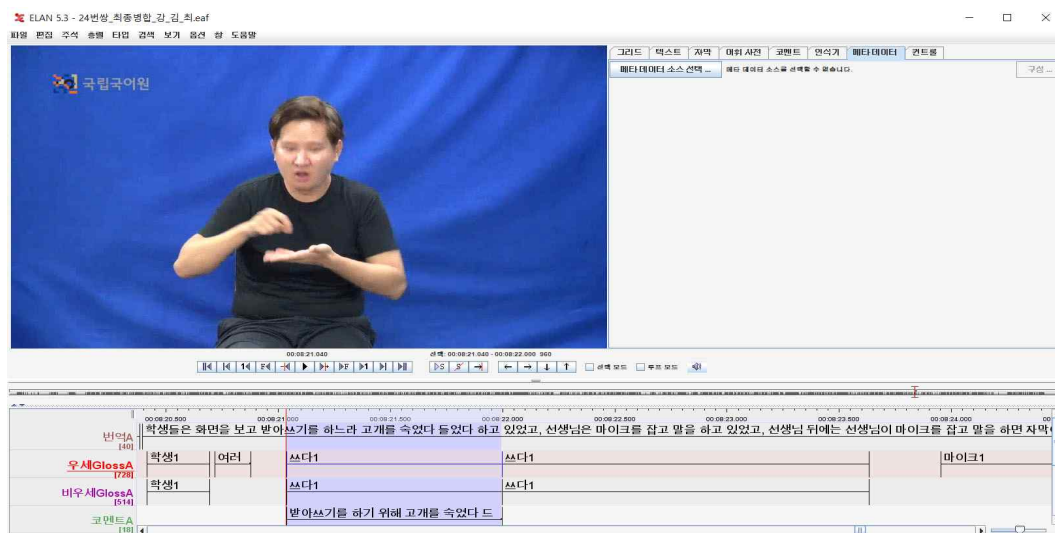
- 토큰 맥락: [대화1A]+[가능1]+[사람1]+[사람1]+[몇1]+[모으다1]+[둘러싸다1]+[여자1]+
[대회2]+[대회2]+[대회2]+[모으다1]
- 번역: 대화가 가능한 젊은 사람 몇 명이 학생들을 앞에다 모아 놓고
- 전사 오류: [제스처]
- 개선 내용: 사람들이 많이 모여 있음을 뜻하는 [제스처]로 판단하여 전사하였으나, 이는 많은 사람이 늘어서 있는 대회를 나타내는 수어 동작이기 때문에 [대회2]라는 새로운 타입을 생성하여 오류를 바로잡았다.



<그림 76> 수어와 제스처의 구분 예시 [대회2]+[대회2]+[대회2]

[쓰다1]

- 토큰 맥락: [학생1]+[여러가지1]+[쓰다1]+[쓰다1]
- 번역: 학생들은 화면을 보고 받아쓰기를 하느라 고개를 숙였다 들었다 하고 있었고
- 전사 오류: [제스처]
- 개선 내용: 학생들이 받아쓰기를 하기 위해 고개를 숙였다 드는 동작을 [제스처]로 전사하는 오류가 있었다. 그러나 이때 언어 제공자의 우세손과 비우세손이 [쓰다1]을 나타내고 있으므로, [쓰다1]로 전사하되 코멘트에 ‘받아쓰기를 하기 위해 고개를 숙였다 드는 동작’이라는 설명을 추가하기로 결정하였다.



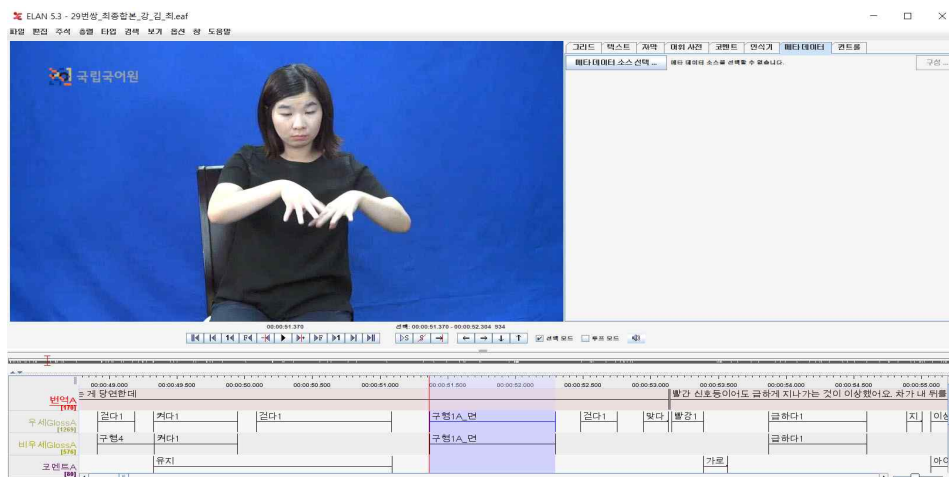
<그림 77> 수어와 제스처의 구분 예시 [쓰다1]

4) 고정적 어휘와 생산적 어휘의 구분

고정적 어휘와 생산적 어휘를 구분하지 못하여 발생한 오류들도 있었다. 그 사례를 제시하면 다음과 같다.

[구형1A_면]

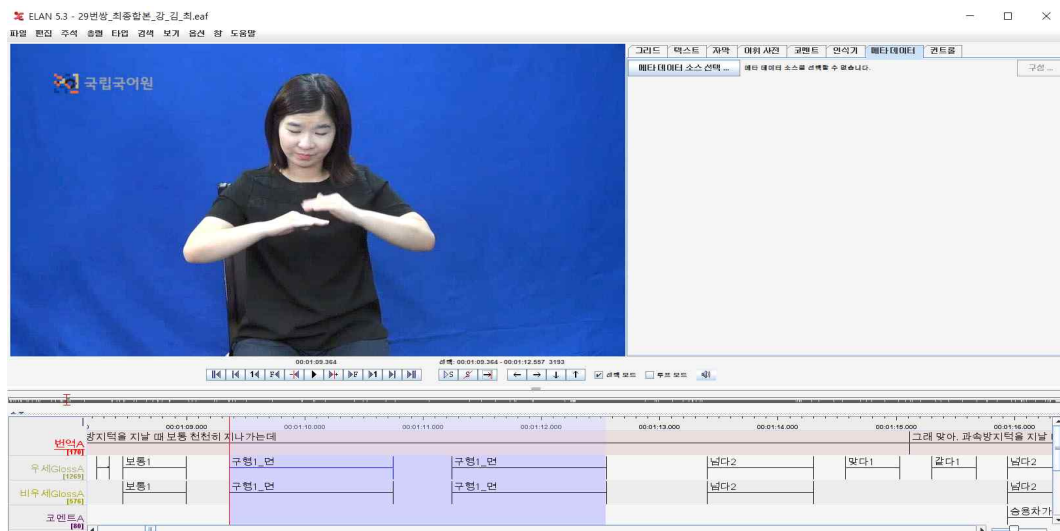
- 토큰 맥락: [초록1A]+[켜다1]+[사람1]+[걷다1]+[구형4]+[켜다1]+[걷다1]+[구형1A_면]+[걷다1]+[맞다1]
- 번역: 파란 신호등이 켜질 때 횡단보도를 건너가는 게 당연한데
- 전사 오류: [횡단보도1]
- 개선 내용: 위 수어를 새 타입 [횡단보도1]로 전사하는 경우가 있었으나 맥락에서 의미가 파악되기 때문에 [구형1A_면]으로 전사하였다.



<그림 78> 고정적 어휘와 생산적 어휘의 구분 예시 [구형1A_면]

[구형1_면]

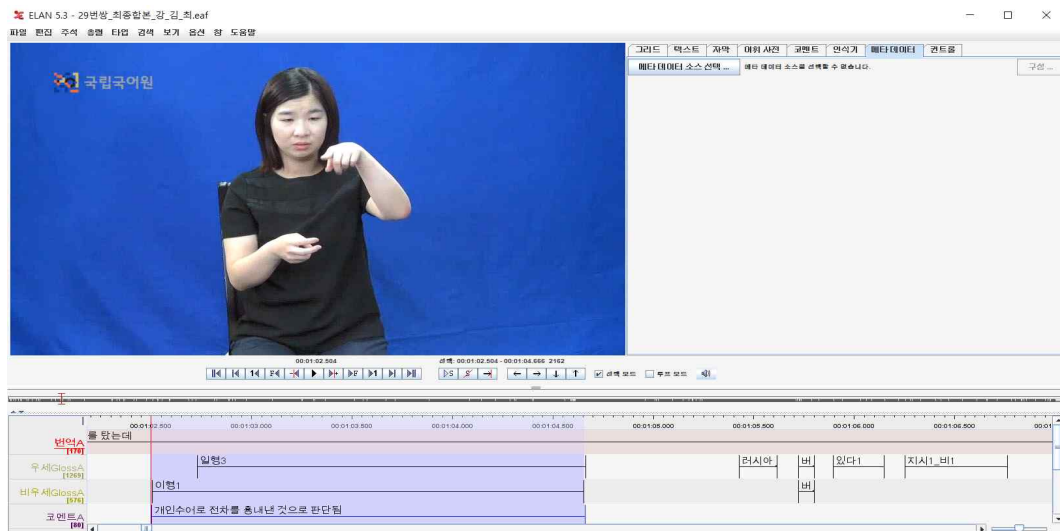
- 토큰 맥락: [지시1_비1]+[보통1]+[구형1_면]+[구형1_면]+[넘다2]+[맞다1]
- 번역: 보통 과속방지턱을 지날 때 보통 천천히 지나가는데
- 전사 오류: [과속방지턱1]
- 개선 내용: 과속방지턱의 모습을 표현하는 동작을 [과속방지턱1]이라는 새로운 타입으로 만들어 전사해야 한다는 의견이 있었다. 그러나 이는 아직 고정된 어휘라기보다 생산적 어휘 단계에 있는 것으로 판단하여 [구형1_면]으로 전사하기로 결정하였다.



<그림 79> 고정적 어휘와 생산적 어휘의 구분 예시 [구형1_면]

[이형1]+[일형3]

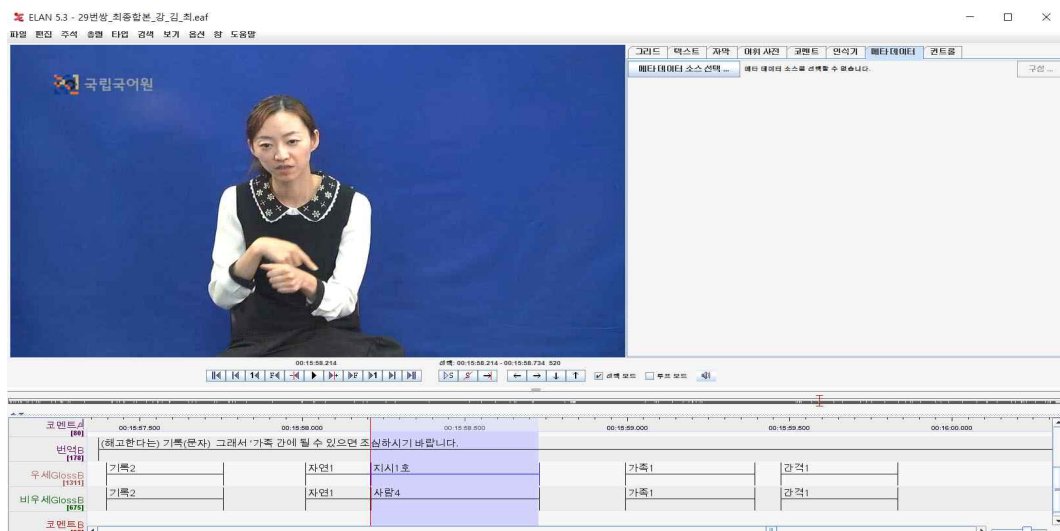
- 토큰 맥락: [지시1_1]+[또1]+[버스1]+[이형1]+[일형3]+[러시아2]+[버스1]+[있다1]+[지시1_비1]+[타다1]
- 번역: 그리고 버스(전차)는 도심 전기선에 연결되어 움직이는 버스(전차)가 러시아에 있어요. 그 버스를 탔는데
- 전사 오류: [전차1]
- 개선 내용: 전차의 모양을 나타내는 동작을 [전차1] 타입을 새로 만들어 전사하는 오류가 있었다. 그러나 이는 고정된 어휘라기보다 전차의 전기선이 연결되는 모양을 흉내 내는 동작이라는 판단하에 새로운 타입을 선정하지 않고 [이형1]+[일형3]으로 전사하였다.



<그림 80> 고정적 어휘와 생산적 어휘의 구분 예시 [이형1]+[일형3]

[지시1호]+[사람4]

- 토큰 맥락: [지시1호]+[사람4]+[가족1]+[간격1]+[가능1]+[그러면1]+[조심1]+[부탁1]
- 번역: 가족 간에 될 수 있으면 조심하시기 바랍니다.
- 전사 오류: [주위1]
- 개선 내용: 유사한 동작의 [주위1]을 새로운 타입으로 선정하여 전사한 경우가 있었다. 그러나 이는 ‘너희’를 지시하는 생산적 어휘로 판단하여 [지시1호]+[사람4]로 처리하였다.



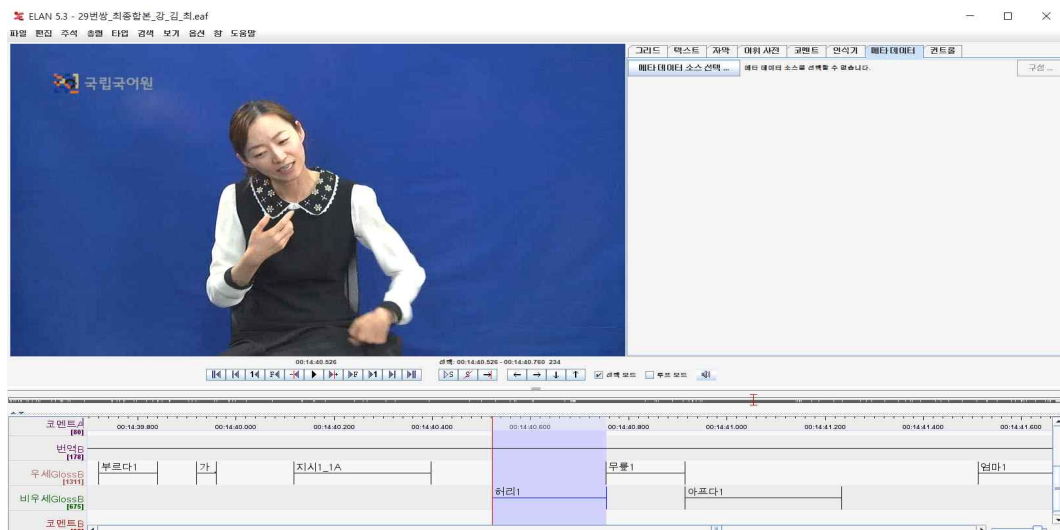
<그림 81> 고정적 어휘와 생산적 어휘의 구분 예시 [지시1호]+[사람4]

5) 토큰 누락

빠르게 지나가는 수어 동작을 인지하지 못하여 토큰 누락이 발생하기도 하였다. 그러한 사례는 다음과 같다.

[허리1]

- 토큰 맥락: [엄마1]+[부르다1]+[가자2]+[지시1_1A]+[허리1]+[무릎1]+[아프다1]+[엄마1]+[가지1]+[변명1]+[무릎1]
- 번역: 엄마가 허리, 무릎이 아프신 척 연기를 하시면서(나랑 같이 가자고 하시는 거지).
- 전사 오류: [허리1] 누락
- 개선 내용: 비우세손으로 허리 아래쪽을 짚는 동작을 인식하지 못하여 토큰 누락이 발생하였다. 수위가 낮기는 하나 허리를 지시한 것으로 판단하여 올바르게 수정하였다.



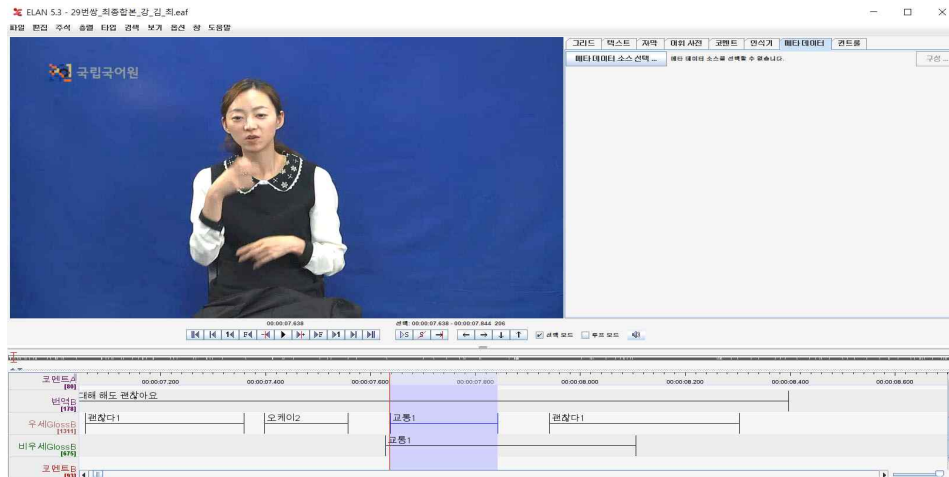
<그림 82> 토큰 누락 예시 [허리1]

6) 분절 오류

수어의 시작점과 끝점을 잘못 인식하여 분절 오류가 발생하는 경우도 있었다. 또한 반복 수어에 대한 인식이 부족하여 하나의 수어 동작을 여러 번 분절하거나, 혹은 여러 개의 수어 동작을 분절하지 않고 하나로 전사하는 경우도 있었다. 그 사례를 제시하면 다음과 같다.

[교통1]

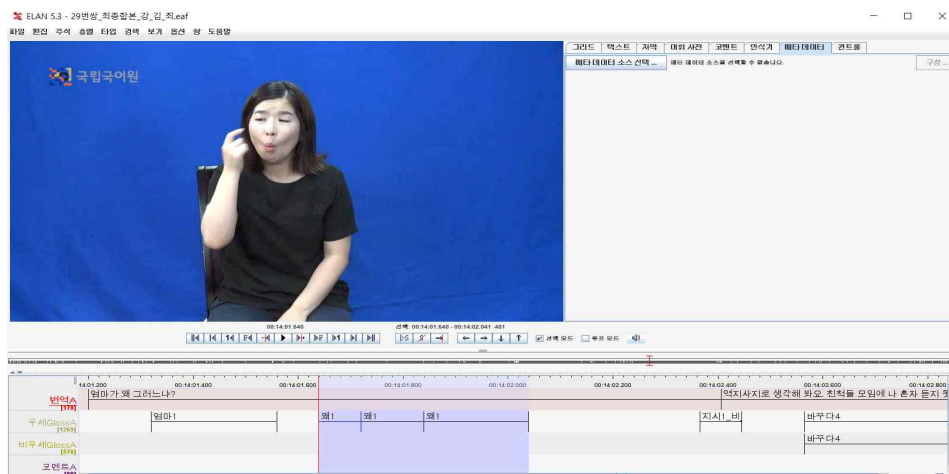
- 토큰 맥락: [괜찮다1]+[오케이2]+[교통1]+[괜찮다1]
- 번역: 괜찮아요. 교통에 대해 해도 괜찮아요.
- 전사 오류: [교통1]의 분절 오류
- 설명: [교통1]의 수형이 달라진 이후에도 수위가 일정하여 일종의 동작 유지로 보고 길게 분절하는 오류가 있었다. 그러나 수형이 달라진 이후에는 [교통1]을 나타내는 것으로 볼 수 없으므로 수형이 변하기 이전 프레임에서 분절하였다.



<그림 83> 분절 오류 예시 [교통1]

[왜1]+[왜1]+[왜1]

- 토큰 맥락: [엄마1]+[왜1]+[왜1]+[왜1]
- 번역: 엄마가 왜 그러느냐?
- 전사 오류: [왜1]
- 개선 내용: [왜1]라는 수어 동작을 한 단위로 보고 분절하지 않는 오류가 있었다. 그러나 이는 [왜1]을 세 번 반복한 것으로 보고 재분절하였다.



<그림 84> 분절 오류 예시 [왜1]+[왜1]+[왜1]

3. 한국수어 말뭉치 전사 지침 정비 및 제언

3.1. 한국수어 말뭉치 전사 지침 정비

2017년도 한국수어 말뭉치 전사 지침을 바탕으로 전사를 진행하는 과정에서 몇 가지 문제점이 발견되었다. 이에 전사 지침의 일부 사항을 다음과 같이 수정 보완하고자 한다.

1) ID 글로스의 구조

기존 전사 지침에서는 기본형과 형태론적 수정, 변이형을 표시하기 위해 숫자와 알파벳을 사용하고 있다. 같은 글로스를 사용하는 다른 타입의 수어일 경우에는 ID 글로스 바로 뒤에 숫자 1, 2, 3 등으로 번호를 부여하며, 형태론적 수정도 마찬가지로 ID 글로스 뒤에 1, 2, 3 등으로 번호를 붙여 나타낸다. 이처럼 표기 방식이 동일하기 때문에 혼란이 발생할 여지가 있다. 예를 들어 [주다11]은 ‘주다’라는 글로스를 사용하는 열한 번째 수어일 수도 있고, [주다1]의 형태론적인 수정일 수도 있다.¹⁷⁾ 이 두 가지를 명확히 구분하기 위하여 본 연구에서는 형태론적 수정에 -(언더바)를 붙이기로 결정하였다. 이에 따라 해당 전사 지침을 수정하면 다음과 같다.

<수정 전>

- ⑥ 형태론적인 수정은 해당 타입의 ID 글로스 다음에 1, 2, 3 등으로 나타낸다. 변이형은 해당 타입의 ID 글로스 다음에 알파벳 A, B, C 등으로 나타낸다.
- ⑦ 연구 목적에 따라 형태론적인 수정과 변이를 다 나타내지 않고 기본형으로 전사할 수 있다.
- ⑧ ID 글로스의 구조는 <그림 85>와 같다.



<그림 85> 기존 ID 글로스의 구조

17) 타입 영상을 보면, [주다11]은 형태론적 수정이지만, [없다11]은 ‘없다’라는 글로스를 사용하는 열한 번째 수어인 것으로 확인할 수 있다.

<수정 후>

- ⑥ 형태론적인 수정은 해당 타입의 ID 글로스 다음에 _1, _2, _3 등으로 나타낸다.
변이형은 해당 타입의 ID 글로스 다음에 알파벳 A, B, C 등으로 나타낸다.
- ⑦ 연구 목적에 따라 형태론적인 수정과 변이를 다 나타내지 않고 기본형으로 전사할 수 있다.
- ⑧ ID 글로스의 구조는 <그림 86>과 같다.



<그림 86> 2018 수정 ID 글로스의 구조

2) 글로스 확정 방법

글로스는 한국어 단어를 사용하여 붙이는 것이 기본 원칙이다. 여기에는 기본적으로 표준국어대사전에 표제어로 등재된 규범적 어휘를 활용해야 한다. 그런데 기존에 구축된 타입 목록을 보면 국어의 어문 규범에 맞지 않는 표기 사례들이 상당수 존재하고 있다. 따라서 이러한 문제를 수정함과 동시에 전사 지침에도 이와 관련된 항목을 명시할 필요가 있다.

다만 어떤 경우에는 표제어를 글로스로 사용하는 것이 적합하지 않을 가능성이 있으므로 이러한 예외까지 고려하여 지침을 작성해야 한다. 예를 들어 새로운 타입 중 [야금야금쓰다]는 ‘조금씩 야금야금 아껴 쓰다’라는 의미를 갖는 수어로서 표준국어대사전에 등재된 어휘 중에는 ‘야금야금하다’에 가깝다고 볼 수 있다. 그러나 ‘야금야금하다’라는 단어는 수어 화자들에게 비교적 익숙지 않으므로 글로스의 이해를 저해할 수 있으며, 상대적으로 자주 쓰이는 단어인 ‘야금야금’은 ‘쓰다’의 의미를 내포하지 않는다는 단점이 있다. 이러한 경우에는 표준국어대사전에 등재된 어휘가 아니더라도, 수어 화자들의 언어 인식을 반영하고 정확하게 수어 단어를 대표할 수 있는 ‘야금야금쓰다’ 등의 구 구성으로 글로스를 붙일 수 있다.

또한 [CCTV]와 같은 외래어 타입의 경우에도 수어 화자들의 언어 의식을 반영하여 영문자로 표기할 수 있다. 이는 표준국어대사전에 ‘시시티브이’로 등재되어 있지만 한글 표기보다는 ‘CCTV’와 같은 표기가 일반적으로 쓰이기 때문이다. 이러한 관점에 따라 해당 전사 지침을 수정하면 다음과 같다.

<수정 전>

- ① 글로스란 수어에 붙이는 이름(label)으로, 숫자나 알파벳 등 다양하게 붙일 수도 있으나 그 이름에 해당하는 수어가 머릿속에 떠오르기 쉽게 붙이는 것이 좋다. 따라서 여기서는 한국어를 사용하여 이름을 붙인다. 그러나 여기에 붙인 한국어 단어는 의미의 번역과는 다른 것이다.
- ② 글로스 선택 시 하나의 수어가 여러 개의 한국어 단어로 번역이 될 때는 가능한 많은 사람이 가장 쉽게 해당 수어를 떠올릴 수 있는 한국어 단어를 선택한다.
예) 꼼꼼하다, 자세하다, 구체적 → [꼼꼼하다]

<수정 후>

- ① 글로스란 수어에 붙이는 이름(label)으로, 숫자나 알파벳 등 다양하게 붙일 수도 있으나 그 이름에 해당하는 수어가 머릿속에 떠오르기 쉽게 붙이는 것이 좋다. 따라서 여기서는 한국어를 사용하여 이름을 붙인다. 그러나 여기에 붙인 한국어 단어는 의미의 번역과는 다른 것이다.
- ② 글로스는 표준국어대사전에 등재된 어휘를 기본으로 하되 수어 화자들의 언어 의식을 반영하여 익숙하고 알기 쉬운 표현으로 선택한다.
예) 야금야금하다 → [야금야금쓰다]
예) 시시티브이 → [CCTV]
- ③ 글로스 선택 시 하나의 수어가 여러 개의 한국어 단어로 번역이 될 때는 가능한 많은 사람이 가장 쉽게 해당 수어를 떠올릴 수 있는 한국어 단어를 선택한다.
예) 꼼꼼하다, 자세하다, 구체적 → [꼼꼼하다]

3) 동음이의어 글로스 구분

기존 전사 지침에서는 글로스로 사용한 한국어 단어가 동음이의어일 경우, 그 수어가 쉽게 유추될 수 있도록 글로스에 의미를 추가하도록 하고 있다. 이를테면 ‘싸다’를 [물건싸다], [값싸다]로 구분하여 글로스를 붙이는 방식이다. 그런데 한국어에는 글로스로 사용 가능한 동음이의어가 매우 많기 때문에 이에 대한 규정을 보다 구체화할 필요가 있다. 이에 본 연구에서는 다음의 지침을 추가하기로 결정하였다.

첫째로, 글로스에 의미를 추가할 때에는 한국어의 어순에 맞게 한다는 조건을 덧붙인다. 기존에 구축된 타입 중에는 [걷다동물], [굽다고기]처럼 한국어 어순에 맞지 않아 부자연스러운 글로스가 발견된다. 그러나 수어에 매겨진 글로스가 한국어이기 때문에 한국어의 문법에 맞게 결정하는 것이 타당하며, 글로스 부여 원칙 중 ‘머릿속에 떠오르기 쉽게’ 글로스를 붙인다는 기술 또한 한국어의 문법을 전제한 규정이라고 볼 수 있다. 그러므로 [동물걷다], [고기굽다]처럼 한국어 어순에 맞게 글로스를 붙인다는 기준을 추가할 필요가 있다.

둘째로, 비슷한 의미를 가진 다른 글로스로 교체하는 방식을 지침에 추가한다. 예를

들어 표준국어대사전에서 동음이의어 관계인 예06(윗사람의 부름에 대답하거나 묻는 말에 긍정하여 대답할 때 쓰는 말)과 예08(어떤 사실을 설명하거나 증명하기 위하여 내세워 보이는 대표적인 것)을 글로스로 사용할 경우, 예06에 해당하는 수어는 [예]로, 예08에 해당하는 수어는 비슷한 의미의 표현인 [보기]로 교체하여 글로스를 붙일 수 있다.

셋째로, 글로스의 품사를 교체하는 방법도 있다. 예를 들어 표준국어대사전에서 동음이의어 관계인 일01(무엇을 이루거나 적절한 대가를 받기 위하여 어떤 장소에서 일정한 시간 동안 몸을 움직이거나 머리를 쓰는 활동)과 일05(자연수의 맨 처음 수)를 글로스로 사용할 경우, 명사형으로만 쓰이는 일05와 달리 일01은 ‘일하다’와 같은 동사형으로도 나타날 수 있으므로 각각 [일]과 [일하다]로 글로스를 붙일 수 있다.

이러한 방식에 의거하여 해당 전사 지침을 수정하면 다음과 같다.

<수정 전>

- ⑦ 글로스로 사용한 한국어 단어가 동음이의어인 경우는 그 수어가 쉽게 유추될 수 있도록 글로스에 의미를 추가한다. 예를 들어 글로스가 ‘물건을 안에 넣고 보이지 않게 씌워 가리거나 둘러 말다.(표준국어 대사전 싸다01)’인 경우 [물건싸다 <그림 93>]로, ‘물건 값이나 사람 또는 물건을 쓰는 데 드는 비용이 보통 보다 낮다(표준국어 대사전, 싸다05).’인 경우는 [값싸다 <그림 94>]로 글로스를 붙인다.

<수정 후>

- ⑦ 글로스로 사용한 한국어 단어가 동음이의어인 경우는 타입의 혼동을 방지하기 위해 다음과 같이 구분하여 글로스를 부여한다.
 - 글로스에 의미를 추가한다. 이때 의미를 추가하여 구 구성이 되는 경우에는 한국어 어순에 맞도록 한다.
예) [걷다](표준국어대사전 걷다02)와 [옷걷다](표준국어대사전 걷다03)
 - 비슷한 의미를 지닌 다른 글로스로 교체한다.
예) [예](윗사람의 부름에 대답하거나 묻는 말에 긍정하여 대답할 때 쓰는 말)와 [보기](어떤 사실을 설명하거나 증명하기 위하여 내세워 보이는 대표적인 것)
 - 글로스의 품사를 교체한다.
예) [일](자연수의 맨 처음 수)과 [일하다](무엇을 이루거나 적절한 대가를 받기 위하여 어떤 장소에서 일정한 시간 동안 몸을 움직이거나 머리를 쓰다)

4) 제스처

기존 전사 지침에서 제스처는 코멘트 층열에 입력하도록 되어 있다. 그러나 기구축 전사 결과물을 살펴보면 제스처는 [제스처]라는 타입으로 등록되어 있으며, 수지 글로

스 층렬에 입력되어 있는 것을 볼 수 있다. 실제로 코멘트 층렬에는 자유로운 텍스트를 입력할 수 있으므로 [제스처] 타입을 단순히 연결하기보다는 해당 제스처의 구체적인 내용을 설명하는 것이 더욱 적절해 보인다. 그러므로 다음과 같이 해당 전사 지침을 수정하였다.

<수정 전>

나. 제스처는 코멘트 층렬에 [제스처]로 입력한다.

<수정 후>

나. 제스처는 수지 글로스 층렬에 [제스처]로 입력하고 코멘트 층렬에 구체적인 내용을 설명한다.

3.2. 한국수어 말뭉치 전사 제언

3.2.1. 타입 정비 제언

본 연구에서는 전사를 수행하는 과정에서 기존에 구축된 타입 목록들을 꾸준히 살펴보고 검토하였으며, 그 과정에서 몇 가지 문제를 발견하였다. 이들은 보다 정밀한 검토를 통해 수정·보완이 필요할 것으로 보인다. 검토가 필요한 항목은 크게 중복 타입 삭제, 삭제 타입 복원, 타입 영상 수정, 타입명 수정 부문으로 나눌 수 있다.

1) 중복 타입 삭제

[놀라다기] / [충격2]

- 문제점: [놀라다기]과 [충격2]가 동일한 수어이다.
- 논의 내용: 서로 다른 타입인 [놀라다기]과 [충격2]가 같은 동작을 나타내고 있으며, [놀라다기]의 타입 설명에도 ‘충격2에 있다(동일하다)’는 설명이 있다.
- 해결 내용: [놀라다기]를 삭제해야 한다.

[여섯달1] / [육개월1]

- 문제점: [여섯달1]과 [육개월1]이 동일한 수어이다.
- 논의 내용: 서로 다른 타입인 [여섯달1]과 [육개월1]이 같은 동작을 나타내고 있으며, [여섯달1]에서는 ‘육개월’이라는 입모양이 보인다.
- 해결 내용: [여섯달1]을 삭제해야 한다.

2) 삭제 타입 복원

[본인1]

- 문제점: [본인1]은 지시와 구별되는 타입이다.
- 논의 내용: 기존 타입 목록 중 [본인1]이 삭제된 기록이 있으나(삭제 일시: 2017. 12. 1. 21:15:53) 이는 단순히 자신을 지시하는 것과는 구별되는 하나의 단어이므로 새로운 타입으로 다시 등록할 수 있다.
- 해결 내용: [본인1]을 복원해야 한다.¹⁸⁾

3) 타입 영상 수정

[이전1]

- 문제점: [이전1]의 영상이 잘못되었다.
- 논의 내용: 기존의 [이전1] 타입 영상에는 ‘미리’와 ‘이전’이라는 수어 동작이 모두 포함되어 있다. 그러나 이들은 비우세손의 형태가 서로 다른 별개의 수어 단어이므로 별도의 타입으로 분리해야 한다. 현재 [미리1]을 새로운 타입으로 등록하였으며, [이전1]의 영상도 적절하게 교체할 필요가 있다.
- 해결 내용: [이전1]의 영상을 교체해야 한다.

[약하다1]

- 문제점: [약하다1]의 영상이 잘못되었다.
- 논의 내용: 기존의 [약하다1] 타입 영상은 ‘약하다’가 아닌 ‘피곤하다’에 해당한다. ‘피곤하다’와 ‘약하다’는 서로 다른 수위에서 발화되는 별개의 단어이므로 별도의 타입으로 등록해야 한다. 현재 [피곤하다1]을 새로운 타입으로 등록하였으며, [약하다1]의 영상도 적절하게 교체할 필요가 있다.
- 해결 내용: [약하다1]의 영상을 교체해야 한다.

4) 타입명 수정

① 의미 오류 수정

[외국1]

- 문제점: [외국1]의 글로스가 잘못되었다.
- 논의 내용: [외국1]로 글로스된 수어 동작은 본래 ‘외국’이 아닌 ‘미국’을 의미한다.
- 해결 내용: [외국1]을 [미국2]로 변경해야 한다.

② 타입 분리

18) 현재 타입 삭제 목록에 있는 영상을 재촬영하여 새로운 타입으로 등록하였다.

[예1], [예2] / [예3]

- 문제점: [예1], [예2]와 [예3]이 이형이의어임에도 불구하고 동일한 글로스를 사용하고 있다.
- 논 의 내용: [예1]과 [예2]는 ‘네(윗사람의 부름에 대답하거나 묻는 말에 긍정하여 대답할 때 쓰는 말)’의 의미를 갖는 데 비해 [예3]은 ‘보기(어떤 사실을 설명하거나 증명하기 위하여 내세워 보이는 대표적인 것)’의 의미를 가지고 있다. 이들을 같은 글로스로 처리한 것은 대응하는 한국어가 동음이의어이기 때문에 발생한 오류로 보인다.
- 해결 내용: [예3]을 [보기1]로 수정해야 한다.

[병1] / [병2]

- 문제점: [병1]과 [병2]가 수어에서 이형이의어임에도 불구하고 같은 글로스를 사용하고 있다.
- 논 의 내용: [병1]은 ‘질병’을 의미하고, [병2]는 ‘액체나 가루를 담는 데에 쓰는 목과 아가리가 좁은 그릇’을 뜻한다. 이는 ‘병’이라는 단어가 한국어에서 동음이의어로 처리되기 때문에 발생한 오류로 보이며, 수어에서는 이형이의어로 쓰이므로 다른 글로스를 부여하는 것이 옳다.
- 해결 내용: [병1]을 [질병1]로 수정해야 한다.

[배1] / [배2]

- 문제점: [배1]과 [배2]가 수어에서 이형이의어임에도 불구하고 같은 글로스를 사용하고 있다.
- 논 의 내용: [배1]은 ‘사람이나 짐 따위를 싣고 물 위로 떠다니도록 나무나 쇠 따위로 만든 물건’을 뜻하며, [배2]는 ‘사람이나 동물의 몸에서 위장, 창자, 콩팥 따위의 내장이 들어 있는 곳으로 가슴과 엉덩이 사이의 부위’를 나타내는 말로, [배1]과 [배2]는 서로 다른 의미의 단어이다. 이는 한국어에서 동음이의어로 처리되기 때문에 발생한 오류로 볼 수 있으며, 수어에서는 이형이의어로 쓰이므로 같은 글로스로 처리해서는 안 된다.
- 해결 내용: [배1]을 [선박1]로 수정해야 한다.

[팔1] / [팔2], [팔3]

- 문제점: [팔1], 그리고 [팔2]와 [팔3]이 수어에서 이형이의어임에도 불구하고 같은 글로스를 사용하고 있다.
- 논 의 내용: [팔1]은 ‘숫자 8’의 의미이며, [팔2]와 [팔3]은 ‘사람의 어깨와 손목 사이의 부분에 해당하는 신체 부위’를 의미한다. 서로 다른 의미의 수어에 같은 글로스를 부여한 것은 오류이다.
- 해결 내용: [팔1]은 다른 숫자들의 글로스를 고려하여 [팔1]로 유지하고, [팔2], [팔

3]을 각각 [팔뚝1], [팔뚝2]로 수정해야 한다.

③ 비규범적 표기 수정

[천정작업1]

- 문제점: [천정작업1]에서 비표준어를 사용하고 있다.
- 논의 내용: 건축물에서 지붕의 안쪽, 또는 건물에서 편평한 위쪽 면을 가리키는 말의 표준어는 ‘천장’이다.
- 해결 내용: 타입명을 [천장작업1]으로 수정해야 한다.

[목메달다1] / [목메달다2]

- 문제점: 글로스에 비규범적 표기를 사용하였다.
- 논의 내용: ‘목메달다’는 ‘목매달다’의 잘못이다.
- 해결 내용: 타입명을 [목매달다1], [목매달다2]로 수정해야 한다.

[초코렛1]

- 문제점: 글로스에 비규범적 표기를 사용하였다.
- 논의 내용: ‘초코렛’은 ‘초콜릿(chocolate)’의 잘못이다.
- 해결 내용: 타입명을 [초콜릿1]로 수정해야 한다.

[싸이렌1]

- 문제점: 글로스에 비규범적 표기를 사용하였다.
- 논의 내용: ‘싸이렌’은 ‘사이렌(siren)’의 잘못이다.
- 해결 내용: 타입명을 [사이렌1]로 수정해야 한다.

[스킨십1]

- 문제점: 글로스에 비규범적 표기를 사용하였다.
- 논의 내용: ‘스킨십’은 ‘스킨십(skinship)’의 잘못이다.
- 해결 내용: 타입명을 [스킨십1]로 수정해야 한다.

[센치미터1]

- 문제점: 글로스에 비규범적 표기를 사용하였다.
- 논의 내용: ‘센치미터’는 ‘센티미터(centimeter)’의 잘못이다.
- 해결 내용: 타입명을 [센티미터1]로 수정해야 한다.

[브레지어1] / [브레지어2]

- 문제점: 글로스에 비규범적 표기를 사용하였다.

- 논의 내용: ‘브래지어’는 ‘브래지어(brassiere)’의 잘못이다.
- 해결 내용: 타입명을 [브래지어1], [브래지어2]로 수정해야 한다.

[푸켓1]

- 문제점: 글로스에 비규범적 표기를 사용하였다.
- 논의 내용: ‘푸켓’은 ‘푸켓(Phuket)’의 잘못이다.
- 해결 내용: 타입명을 [푸켓1]로 수정해야 한다.

④ 한국어의 어순에 맞는 글로스 부여

[걷다동물1]

- 문제점: ‘걷다-동물’이라는 표현이 한국어의 어순 구조에 비추어 봤을 때 부자연스럽다.
- 논의 내용: 수어로 이루어진 문장 발화 자체는 한국어의 문법을 그대로 따르지 않기도 하지만, 수어에 매겨진 글로스는 한국어이기 때문에 한국어의 문법에 맞추어 나타나야 하는 것이 타당하다. 또한 글로스 부여 원칙에서 머릿속에 떠오르기 쉽게 붙인다는 것은 한국어의 문법을 전제한 것이다.
- 해결 내용: 타입명을 [동물걷다1]로 수정해야 한다.

[굽다고기1]

- 문제점: ‘굽다-고기’라는 표현이 한국어의 어순 구조에 비추어 봤을 때 부자연스럽다.
- 논의 내용: 수어로 이루어진 문장 발화 자체는 한국어의 문법을 그대로 따르지 않기도 하지만, 수어에 매겨진 글로스는 한국어이기 때문에 한국어의 문법에 맞추어 나타나야 하는 것이 타당하다. 또한 글로스 부여 원칙에서 머릿속에 떠오르기 쉽게 붙인다는 것은 한국어의 문법을 전제한 것이다.
- 해결 내용: 타입명을 [고기굽다1]로 수정해야 한다.

⑤ 글로스의 구조 수정

[주다11] 외 형태론적 수정

- 문제점: [주다11] 외 형태론적 수정의 표기 방식이 모호하다
- 논의 내용: [주다11]은 [주다1]의 형태론적 수정을 나타낸 것이다. 그러나 이러한 표기는 ‘주다’의 열한 번째 타입과 구분되지 않으므로 혼동을 일으킬 수 있다. 기본형과 형태론적 수정을 보다 명확히 구분할 수 있도록 글로스의 구조를 변경해야 한다.
- 해결 내용: 형태론적 수정은 번호 앞에 _(언더바)를 붙여 수정할 수 있다. 수정 가

능한 타입 목록은 다음과 같다.

- [뇌물11] → [뇌물1_1]
- [뇌물11A] → [뇌물1_1A]
- [막대기12] → [막대기1_2]
- [말하다11] → [말하다1_1]
- [사주다11] → [사주다1_1]
- [선배11] → [선배1_1]
- [시키다11] → [시키다1_1]
- [연결11] → [연결1_1]
- [주다11] → [주다1_1]
- [주다11A] → [주다1_1A]
- [주다11B] → [주다1_1B]
- [죽다11] → [죽다1_1]
- [촬영11A] → [촬영1_1A]

[입원1] 외 신체 지시

- 문제점: [입원1]의 표기 방식이 모호하다.
- 논의 내용: 현재 [입원1]은 ‘입 주위를 동그랗게 가리키는 것’을 의미한다. 그러나 이 경우 ‘병원에 들어가다’라는 의미의 ‘입원’과 혼동을 일으킬 수 있다.¹⁹⁾²⁰⁾
- 해결 내용: [입원1]을 [입2]로 수정할 수 있다. 수정 가능한 총 타입 목록은 다음과 같다.
 - [눈원1] → [눈2]
 - [머리원1] → [머리2]
 - [볼원1] → [볼2]
 - [입원1] → [입2]
 - [코원1] → [코2]
 - [턱원1] → [턱2]

⑥ 기타: 재고가 필요한 타입명

[둘1] / [네사람1]

- 문제점: 글로스 체계에 통일성이 부족하다.
- 논의 내용: ‘두 명’을 뜻하는 수어의 글로스는 [둘1], ‘네 명’을 뜻하는 수어의 글로스는 [네사람1]으로 되어 있다.

19) 실제로 [입원1]이 이미 존재한다는 이유로 ‘병원에 들어가다’라는 의미의 수어에는 [입원하다1]이라는 글로스를 부여하였다.

20) 서울 지역 24번 쌍 E-토론에 나오는 [입원1] 토큰의 일부는 ‘병원에 들어가다’라는 의미의 수어에 연결되어 있는 등 적절하지 못한 글로스로 인한 전사 오류가 발견되었다.

- 해결 방안: 단위명사를 활용하여 [두명1], [네명1]으로 통일할 수 있다.

[카톡보내다1] / [카톡받다1]

- 문제점: 글로스 체계에 통일성이 부족하다.
- 논의 내용: ‘주다’와 ‘받다’는 형태론적 수정 관계로서 모두 [주다]라는 하나의 글로스로 대표된다. 그런데 ‘카톡을 주고받는’ 수어의 경우, [카톡보내다1]와 [카톡받다1] 두 타입이 모두 존재한다.
- 해결 방안: [카톡받다1]을 삭제할 수 있다.

[캡모자1]

- 문제점: 기본형의 글로스가 없는 상태에서 의미가 덧붙은 글로스를 사용하고 있다.
- 논의 내용: 현재 타입 목록에 ‘모자’와 관련된 타입은 [캡모자1]이 유일하다. 이 경우 종류가 다른 모자의 수어가 나올 때마다 ‘--모자1’이라는 새로운 글로스를 부여해야 하는 부담이 생긴다.
- 해결 내용: [캡모자1]을 [모자1]로 수정할 수 있다.

[하지마1]

- 문제점: 글로스의 명칭이 적절하지 않다.
- 논의 내용: 이미 [금지1]이라는 타입이 존재하는 시점에서 ‘하지마’와 ‘금지’의 의미 차이를 깨닫기 어렵기 때문에, 하나의 글로스로 통일할 필요가 있다. 글로스 부여 원칙에 의거하여 머릿속에 쉽게 떠오를 수 있는 단어를 부여하는 것이 좋다면, ‘하다’와 ‘말다’가 결합한 ‘하지마’보다는 ‘금지’라는 단어가 훨씬 직관적이고 명료하다.
- 해결 내용: [하지마1]을 [금지2]로 수정할 수 있다.

3.2.2. 한국수어 자료 통합 지원 시스템 정비 제언

장기적으로 농인들은 물론 청인 연구자들도 한국수어 자료 통합 지원 시스템을 쉽게 활용할 수 있는 방안을 구축해야 한다. 시스템의 활용성을 높이기 위해서는 타입 관련 정보를 수정, 보완할 필요가 있다.

1) 표준국어대사전과의 연계

타입명을 선정할 때에는 표준국어대사전에 표제어로 등재된 어휘와 규범적 표기를 적극 활용해야 한다. 또한 타입과 관련된 여러 가지 정보를 시스템에 제시할 때에도 표준국어대사전과의 연계가 필요하다. 비규범적 표기를 사용하면 기술의 정확성이 떨어질 뿐만 아니라 시스템 활용성도 떨어질 수밖에 없다.

2) 동형어 제시

타입 설명 항목에 동형어 항목을 새로 추가하고, 동형어로도 타입 검색을 할 수 있도록 시스템을 정비할 필요가 있다. 하나의 수어 단어가 여러 개의 한국어 단어와 대응하는 경우가 많기 때문에, 여러 가지 의미를 갖는 단일한 타입명을 검색하기가 어렵기 때문이다. 예를 들어 수어에서 ‘작다, 적다, 조금, 약간, 약소하다, 자그마하다’는 모두 [작다]라는 타입으로 대표되는데, 이를 찾아내기가 쉽지 않다. 현재 한국수어사전에 동형어가 모두 제시되어 있는 것처럼, 시스템에서도 동형어를 제시해주는 것이 바람직하다.

3) 타입 수형 사진 제시

현재 시스템상에 타입 설명 영상은 모두 올라와 있으나 수형 사진은 거의 없다. 그러나 사소한 수형의 차이로 구분되는 수어 단어들이 있기 때문에 되도록이면 수형 사진을 함께 올리는 것이 좋다. 현재 한국수어사전에는 수형에 대한 설명이 제시되어 있는데, 한국수어 자료 통합 지원 시스템에서는 수형 사진이 그 역할을 대신할 수 있다. 또한 전사팀에서 기본 수어로 업로드한 지시류 수어 단어의 경우에도 혼란이 없도록 수형 사진을 모두 제시하는 것이 바람직하다.

4) 타입 설명 확대

현재 타입 설명 방식에는 특별한 체계가 없고 불필요한 설명들이 많이 삽입되어 있다. 이들 중 유의미한 것만 남기고 불필요한 것은 삭제하는 조치가 필요하다.

현재 시스템상에 올라와 있는 타입 설명의 유형은 대체로 다음과 같다.

가. 잉여적 설명

- 타입명을 그대로 제시하거나 큰 의미 없는 설명을 제시하는 경우

예 [예3]: 예3 / [혈체어2]: 혈체어2입니다. / [두살많다1]: 두살많다 표현입니다.

나. 수정 내역

- 타입 수정 전후의 내용을 제시하는 경우

예 [제스처]: 제스처->제스처 / [작다4]: 작다4 -> 수어사용에서 작다, 조금, 적다를 함께 사용한다고 보고 [조금2]에서 [작다4]로 수정함.

다. 개인적 메모

- 전사자가 개인적으로 메모를 남긴 경우(타입 선정 이유, 삭제 제언 등)

예 [목메달다2]: 수형이달라서 올립니다. / [놀라다7]: 충격2에있습니다.

라. 관련어 제시

- 유의어, 동형어를 함께 제시하는 경우

예 [외국1]: 청색, Jeans, 청바지, 외국, / [일년후1]: 일 년 후, 내년

마. 보충 설명

- 수어의 특징이나 사용 환경 등, 유의미한 기술을 보충하는 경우

예 [택시2]: 택시 옛날 수화 / [지시2그림]: 표지판에서 자주 나오는 수형

이들 중 (가)는 잉여적인 설명이므로 삭제해도 무방하며, (나)의 경우에도 수정 내역은 시스템에서 별도로 관리할 수 있기 때문에 불필요하다. 그리고 (다)와 같은 개인적 메모는 반드시 전자 책임자와 시스템 관리자가 확인할 필요가 있다. 개인 전자자가 등록한 타입은 반드시 철저한 검토를 거쳐야 하며, 검토 후에 그와 관련된 메모는 그대로 남겨둘 필요 없이 삭제해야 한다.

결과적으로 타입 설명에서 의미 있는 유형은 (라)와 (마)이다. (라)의 경우 동형어 항목을 별도로 두어 제시할 수 있으며, (마)는 가장 바람직한 유형으로서 이러한 설명만을 남기는 것이 바람직하다.

그 밖에도 타입 설명에 들어갈 수 있는 정보는 다음과 같다.

가. 본말 정보

- 한자어, 외래어는 원어 정보를 함께 제시할 수 있다.

예 맹[盲], 푸켓[Phuket]

- 표준국어대사전에 등재되지 않은 신조어나 준말의 경우, 간략한 설명을 제시할 수 있다.


예 [카톡보내다1]: 카카오톡(KakaoTalk; 스마트폰용 무료 통화 및 메신저 응용 프로그램) 메시지를 보낸다.

나. 기저 이미지

- 기저 이미지가 다양한 경우에는 그에 대한 설명을 추가할 수 있다.

예 ‘핸드폰’을 나타내는 수어: [핸드폰2]는 ‘폴더형 핸드폰을 나타내는 동작’, [핸드폰3]은 ‘슬라이드형 핸드폰을 나타내는 동작’, [핸드폰4]는 ‘스마트폰을 나타내는 동작’ 등

이와 같은 제언을 바탕으로 한국수어 자료 통합 지원 시스템을 정비할 수 있다. 예를 들어 [가만있다1]은 다음과 같이 수정할 수 있다.

타입 범주	수지 타입 범주
타입명	가만있다1
타입 파일 종류	영상
타입 구분	직접 등록
타입 영상	
타입 수형 사진	
타입 설명	가만있다1
등록자	시스템 관리자
등록일시	2017. 1. 10. 17:16:54

<표 13> [가만있다1] 타입 정보(수정 전)

타입 범주	수지 타입 범주
타입명	가만있다1
타입 파일 종류	영상
타입 구분	직접 등록
타입 영상	
타입 수형 사진	
타입 설명	불상의 모습을 나타내는 동작
동형어	멈추다, 멎다
등록자	시스템 관리자
등록일시	2017. 1. 10. 17:16:54

<표 14 > [가만있다1] 타입 정보(수정 후)

3.2.3. 전사 정비 제언

현재 말뭉치 전사 지침과 타입 목록들을 재정비한 후, 기구축 말뭉치의 전체적인 수정 및 새로운 말뭉치 전사가 함께 이루어져야 한다. 2017년 연구를 통해 일부 지침이 수정되었지만 전사 과정에서 새로운 문제들이 꾸준히 발견되기 때문에 지속적인 검토가 필요하다. 또한 기존에 구축된 말뭉치가 수정된 지침을 반영하지 못하고 있으므로 혼란을 초래할 수 있다. 그러므로 언어학적 관점에서 기존의 전사 작업을 정비하고, 앞으로의 작업을 계획할 필요가 있다.

1) 비수지 층렬 전사

이번 말뭉치 전사는 번역 층렬과 수지 글로스 층렬, 코멘트 층렬에 한하여 이루어졌다. 그러나 앞으로는 비수지 층렬의 전사도 이루어져야 한다. 또한 2017년 연구에서 제안하였던 입 움직임의 전사도 이루어질 수 있다. 이를 위해서는 전사 지침부터 다시 한 번 상세하게 검토할 필요가 있다. 현재 비수지 층렬에 대한 전사 지침이 미비한 상태이므로 이 부분을 보다 구체화하는 작업이 필요하다.

2) 언어학적 검토

언어학적 특성에 기반하여 타입 검토와 지침 검토, 전사 작업이 이루어져야 한다. 수어와 한국어의 언어학적 특성을 함께 이해하고, 농인 연구자와 청인 연구자들 모두가 활용할 수 있는 말뭉치를 구축할 필요가 있다.

3) 전사 시간 확보

현재 전사 시간은 1분당 3시간으로 배정되어 있다. 그러나 이는 정밀한 전사를 수행하기에는 턱없이 부족한 시간이다. 전사의 질을 보장하기 위해서는 1분당 최소 5시간 이상의 전사 시간이 필요하다.

IV. 한국수어 말뭉치 활용 워크숍 및 전사 교육

1. 한국수어 말뭉치 활용 워크숍의 의의

한국수어 말뭉치 연구 및 구축 사업은 한국수어 말뭉치의 연구와 구축만을 그 목적으로 제한하지 않는다. 본 연구와 기존 연구를 통해 축적된 한국수어와 한국수어 말뭉치의 구축 및 연구 성과를 연구자와 일반에게 공개하여 말뭉치 구축의 필요성을 알리고 한국수어에 대한 관심을 불러일으키기 위해 연구자 및 일반을 대상으로 한 워크숍을 기획하고 실행했다.

한국수어 말뭉치 구축 사업을 통해 진행된 워크숍은 워크숍마다 성격이 조금씩 차이가 나지만 기본적으로 워크숍을 통해 한국수어나 한국수어 말뭉치를 연구하는 연구자들을 초청하여 연구진의 전사 능력 및 한국수어와 한국수어 말뭉치에 대한 이해 수준의 향상에 그 초점을 두었다. 아울러 다양한 홍보 방식으로 일반인과 학생들의 참여를 유도하고자 노력했다.

또한 우리나라에서 한국수어를 전문적으로 연구하는 사람들은 많지 않기 때문에 이러한 워크숍을 통해 동일 계열의 주제를 다루는 연구자들끼리 학술적, 인적 교류를 하는 자리를 마련하여 전체적인 연구 능력과 결과물의 질을 함양할 수 있다. 아울러 비전문가와 타 분야 전문가들에게 한국수어가 가진 연구 주제로서의 가치를 선전했다는 점에서 본 연구를 통해 진행된 워크숍에 의의가 있다고 볼 수 있다.

다음은 본 연구에서 진행한 워크숍 개최 실적이다.

행사명	일시	장소	주제
제1차 워크숍	2018년 6월 16일 14시-16시	전북, 전북대학교	한국수어 말뭉치 구축의 필요성과 전사의 기초
제2차 워크숍	2018년 7월 27일 14시-16시	전북, 전북대학교	한국수어 자료 통합 지원 시스템의 활용과 실제
제3차 워크숍	2018년 8월 2일 10시-13시	전북, 전북대학교	한국수어 말뭉치 구축과 활용
제4차 워크숍	2018년 9월 1일 10시-16시	경기, 한국복지대학교	한국수어 말뭉치 구축의 실제

<표 15> 2018년 워크숍 개최 실적

2. 한국수어 말뭉치 활용 워크숍 개최

2.1. 제1차 워크숍 “한국수어 말뭉치 구축의 필요성과 전사의 기초”

제1차 워크숍은 6월 16일 14시부터 진행됐으며 나사렛대학교 고인경의 ELAN 교육 자료집을 토대로 ELAN 프로그램을 실제 구동해보고 단축키 등 프로그램 사용법을 시연 및 실습해보는 내용으로 진행되었다.

실제 연구 내용에 착수하기 이전의 워크숍으로서 전북대학교 내에서 진행하기로 하여 행사 일주일 전에 현수막을 광장에 게시하여 홍보를 진행하였으며 행사에는 연구 팀과 함께 대학원생이 참여하였다. 해당 워크숍 목차는 아래와 같다.

내용	발표자
ELAN 프로그램의 기본 사항 및 구동 방법	발표자료: 고인경(나사렛대학교) 진행: 김성준(전북대학교)
종합 토론	

<표 16> 제1차 워크숍 프로그램 목차



<그림 87> 제1차 워크숍 행사 사진

2.2. 제2차 워크숍 “한국수어 자료 통합 지원 시스템의 활용과 실제”

제2차 워크숍은 7월 27일 진행됐으며 한국수어 자료 통합 지원 시스템을 개발한

다이퀘스트의 박종윤 차장이 교육을 진행했다. 교육의 내용은 시스템의 기본 메뉴의 설명과 더불어 시스템 내의 각 권한이 갖는 역할과 작동 범위에 관한 것이었다. 이를 통해 전사 작업의 분배와 전사 자료의 업로드, 타입과 토큰 관리 등의 기능을 함께 실습했다. 아울러 한국수어 말뭉치 구축을 위해 개정된 ELAN 프로그램의 타입 검색 기능을 활용한 전사 방법에 대한 내용을 실습했다.

내용	발표자
한국수어 자료 통합 지원 시스템의 실재	박종윤 (다이퀘스트)
종합 토론	

<표 17> 제2차 워크숍 프로그램 목차



<그림 88> 제2차 워크숍 행사 사진

2.3. 제3차 워크숍 “한국수어 말뭉치 구축과 활용”

제3차 워크숍은 8월 2일에 전년도 연구 사업을 수행한 한국복지대의 전사팀을 초청하여 워크숍을 진행하였다. 홍보 현수막을 행사 일주일 전에 외부에 게시하고 전북 지역 내의 농인 커뮤니티를 통해 농인들의 참여를 유도하여 전북대학교 학부생 및 대학원생, 전북 지역 거주 농인 등이 워크숍에 참석하였다. 진행된 워크숍의 발표 내용은 아래와 같다.

내용	발표자
수어 글로스의 기본 개념	원성옥 (한국복지대)
수어 분절 규칙과 전사 규칙	변강석 (강남대학교)
전사의 실제 (ELAN/한국수어자료통합지원시스템)	이은영 (서울수화전문교육원)

<표 18> 제3차 워크숍 프로그램 목차



<그림 89> 제3차 워크숍 행사 사진

2.4. 제4차 워크숍 “한국수어 말뭉치 구축의 실제”

제4차 워크숍은 9월 1일 한국복지대에서 진행됐으며 말뭉치 구축의 필요성과 실재를 주제로 수어 말뭉치의 활용에 대한 이론적 내용과 함께 지역 균형 말뭉치 구축을 위한 강원 지역 말뭉치 구축 사례 소개, ELAN을 이용한 말뭉치 분절 및 글로싱 실습 교육을 진행하였다. 행사는 경기도 평택시 한국복지대학교 인학관에서 진행됐고 외부에 일주일 전 현수막을 게시하고 전북대학교 내에도 게시하여 홍보를 진행했다. 진행된 프로그램은 아래와 같다.

내용	발표자
수어 말뭉치의 구축 목적과 활용 의의	원성옥 (한국복지대)
수어 말뭉치 구축의 사례	강진주 (전북대학교)
수어 말뭉치 전사 프로그램(ELAN)의 실제	변강석 (강남대학교)
수어 말뭉치 전사의 실제	이은영 (서울수화전문교육원)

<표 19> 제4차 워크숍 프로그램 목차



<그림 90> 제4차 워크숍 행사 사진

V. 결론 및 제언

1. 연구 결과 요약

한국수어 말뭉치 연구는 2015년 서울 지역을 중심으로 하는 말뭉치 시범 구축 사업을 시작으로, 2016년에는 한국수어 말뭉치 구축을 위한 통합 지원 시스템을 구축하고, 2017년에는 기구축 한국수어 말뭉치의 정비와 전사가 이루어졌다. 본 사업에서는 한국수어 말뭉치의 구축을 지역 균형적으로 확대함과 동시에, 기구축된 한국수어 말뭉치의 전사를 진행하고, 한국수어 말뭉치의 활용 및 전사자 양성을 위한 워크숍을 개최하였다. 그 과정에서 기구축된 한국수어 말뭉치의 타입을 정비하고 원시 말뭉치 구축 지침과 전사 지침을 보완하는 작업도 함께 실시하였다. 연구 결과를 요약하면 다음과 같다.

1.1. 한국수어 말뭉치 구축

본 사업에서는 2015년도 한국수어 말뭉치의 시범 구축에 이어 두 번째로 한국수어 말뭉치를 구축함으로써 한국수어 자료를 확대하였다. 구축은 ‘지역 선정 - 언어 제공자 선정 - 언어 제공자 질문지 수정 - 유도 자료 제작 - 사회자 선정 및 교육 - 원시 말뭉치 촬영 - 원시 말뭉치 자료 가공 및 한국수어 자료 통합 지원 시스템 탑재’의 순서로 이루어졌다. 이에 따른 결과는 다음과 같다.

첫째, 지역 선정에서는 한국수어 말뭉치 자료의 지역 균형성을 확보하기 위해 강원 지역, 전남 지역, 전북 지역, 경남 지역의 총 네 곳의 지역을 선정하여 한국수어 말뭉치를 확대 구축하였다.

둘째, 언어 제공자 선정에서는 최종적으로 14쌍의 언어 제공자가 선정되었다. 지역별로는 강원 지역 4쌍, 전남 지역 4쌍, 전북 지역 4쌍, 경남 지역 2쌍이다. 총 32명, 16쌍의 언어 제공자를 대상으로 촬영을 진행하였으나 언어 제공자 선정 기준에 따라 최종적으로 언어 제공자 14쌍의 원시 말뭉치 자료를 구축하였다.

셋째, 언어 제공자 질문지 수정에서는 언어 제공자의 메타 데이터를 수집하기 위한 질문지의 내용을 보다 직관적으로 파악할 수 있도록 표현이나 질문의 방향 등을 수정하였다.

넷째, 유도 자료 제작에서는 지역 수어의 포착과 더불어 풍부한 한국수어의 채집을 위해 기존에 작성된 유도 자료의 내용을 일부 수정하였다. 일부 주제의 경우, 언어 제공자들이 더욱 적극적으로 이야기를 진행해 나갈 수 있도록 현재성을 띤 주제들로 대체하였고, 지역 수어를 이끌어 내기 위해 지역성을 가미한 주제를 제시하였다.

다섯째, 사회자 선정 및 교육에서는 언어 제공자들의 대화를 돕기 위한 사회자를

기존 1명에서 3명으로 확대 선정하여 교육하였다. 사회자의 나이와 성별에 차이를 둬으로써 언어 제공자들의 특성에 따라 다각적으로 대응할 수 있게 함과 동시에 사회자의 피로도를 경감시키고자 하였다.

여섯째, 원시 말뭉치 촬영에서는 지역별로 강원 지역 11시간 21분 20초, 전남 지역 11시간 28분 22초, 전북 지역 12시간 50분 40초, 경남 지역 5시간 57분 15초 분량의 촬영이 이루어졌으며, 총 41시간 12초 분량의 원시 말뭉치 자료가 구축되었다.

일곱째, 원시 말뭉치 자료 가공 및 한국수어 자료 통합 지원 시스템 탑재에서는 수집된 원시 말뭉치 영상의 시작점과 끝점을 맞추는 등의 가공을 거쳐 유도 자료별로 분절하였으며, 1차 번역을 실시하였다. 1차 번역의 경우, 사회자에 의해 초별 번역이 이루어졌고, 수화통역사나 농통역사 등 수어 전문가 집단에 의해 검토를 진행한 뒤, 최종적으로 연구팀 소속의 농인 연구원이자 전사 책임자(이영재)가 검수하였다. 그리고 이러한 과정을 통해 구축된 한국수어 말뭉치 자료는 한국수어 자료 통합 지원 시스템에 업로드하였다.

1.2. 한국수어 말뭉치 전사

본 사업에서는 2015년 수집한 언어 자료(동영상) 중에서 일부 영상을 선정하여 우세손과 비우세손에 대한 번역 및 기본 전사를 완료하였다. 전사는 ‘전사자 교육 - 전사 영상 선정 - 전사 영상 번역 - 전사 영상 구분 - 전사 점검 - 전사 결과 확립’의 순서로 이루어졌다. 이에 따른 결과는 다음과 같다.

첫째, 전사자 교육에서는 전사에 앞서 전사자들의 역량 강화 목적으로 말뭉치 교육과 전사 교육이 실시되었다. 전사를 진행하는 중에도 이러한 교육들은 상시로 이루어졌다.

둘째, 전사 영상 선정에서는 2015년 수집한 언어 자료(동영상) 중 6개 영상, 총 재생시간 82분 54초 분량을 선정하였다. 이때 통일성과 다양성을 고려하여 전사 영상을 선정하였다.

셋째, 전사 영상 번역에서는 전사에 들어가기 이전에 선정한 영상에 대한 번역 작업을 실시하였다. 번역은 농인 연구진에 의한 초별 번역이 우선적으로 이루어졌으며, 그 후 청인 연구진들이 해당 번역을 엘란(ELAN)에 입력하였다. 마지막으로 전사 책임자(이영재)를 중심으로 번역의 통일성을 확보하고 내용을 교정하는 검수 작업을 진행하였다.

넷째, 전사 영상 구분에서는 전사 수준을 일정 부분 유지하기 위해 기존에 전사 경험이 있는 전사팀A(복지대)에서 50분 분량을 담당하고, 전사팀B(전북대)에서 30분 분량을 담당하는 것으로 결정하여 배분하였다.

다섯째, 전사 점검에서는 전사팀B(전북대)의 경우, 매주 2회 이상 전사 회의를 진행하였다. 우선 청인 연구진끼리 모여 서로의 전사 결과를 상호 점검하였고, 이후 전체

연구진이 모이는 전사 회의를 통해 전사 내용 전체를 검토하고 확정하여 전사의 질을 확보하였다.

여섯째, 전사 결과, 발생한 토큰 수는 총 13,302개이며, 새로운 타입은 201개 생성하였다.

1.3. 한국수어 말뭉치 활용 워크숍 및 전사 교육

본 사업에서는 한국수어 말뭉치의 구축 및 연구 성과를 연구자와 일반에게 공개하여 말뭉치 구축의 필요성 및 관심을 불러일으키기 위해 연구자 및 일반을 대상으로 한 워크숍을 기획하고, 한국수어 말뭉치 전문 전사자 양성을 위한 전사 교육을 실시하였다. 워크숍 진행 내역은 다음과 같다.

첫째, 제1차 워크숍은 ‘한국수어 말뭉치 구축의 필요성과 전사의 기초’라는 주제로 2018년 6월 16일 전라북도 전주시 소재의 전북대학교에서 개최되었다.

둘째, 제2차 워크숍은 ‘한국수어 자료 통합 지원 시스템의 활용과 실제’라는 주제로 2018년 7월 27일 전라북도 전주시 소재의 전북대학교에서 개최되었다.

셋째, 제3차 워크숍은 ‘한국수어 말뭉치 구축과 활용’이라는 주제로 2018년 8월 2일에 전라북도 전주시 소재의 전북대학교에서 개최되었다.

넷째, 제4차 워크숍은 ‘한국수어 말뭉치 구축의 실제’라는 주제로 2018년 9월 1일에 경기도 소재의 한국복지대학교에서 개최되었다.

이를 통해 한국수어 말뭉치 구축의 중요성과 필요성에 대한 인식을 확산하고 원시 말뭉치 구축과 전사 말뭉치 구축을 위한 실질적인 능력을 확보할 수 있었다. 특히, 한국수어 말뭉치 구축에 참여하기를 원하는 농인 및 언어 연구자의 인력풀을 넓힐 수 있는 계기가 되었다.

1.4. 한국수어 말뭉치 구축 지침 정비 및 타입 정비

본 사업에서는 향후 한국수어 말뭉치 구축의 효율성 확보를 위해, 2017년 개정한 한국수어 말뭉치 구축 지침을 다시 정비하였다. 주요 정비 내용은 다음과 같다.

첫째, 지침의 정비는 2015년도에 개발하고 2016년, 2017년에 개정된 한국수어 말뭉치 구축 지침을 토대로 하였으며, 2018년도에 말뭉치 구축을 진행하면서 수정하거나 보완한 내용을 추가하는 한편, 가독성을 위해 몇몇 내용들을 개조식 표현으로 수정하였다.

둘째, 언어 제공자 부분에서 언어 제공자 쌍의 넘버링(numbering) 및 지역 구분과 관련된 내용을 추가하여, 추후 자료가 일관성 있게 관리될 수 있도록 하였다.

셋째, 언어 자료 수집 부분에서 촬영 환경을 중심으로 하여 내용을 수정 보완하였다. 특히 언어 제공자들을 좀 더 배려하고, 촬영 결과물의 품질을 우선시하는 방향으로 촬영실을 효율적으로 구축하기 위한 내용으로 수정 보완하였다.

넷째, 전사 과정에서 발견된 문제점들을 바탕으로 글로스의 구조 및 글로스 확정 방법에 대한 내용을 수정하여 제시하였다. 그에 더하여 한국어 글로스에서 동음이의어로 나타나지만 수어에서는 이형이의어로 나타나는 경우에 대해 그 해결방안을 보다 정교화하여 제시하였다.

다섯째, 기존 연구에서 등록한 타입에 대한 정비를 실시하여, 중복 타입의 삭제 및 새로운 타입의 추가 등을 제안하였다.

1.5. 자문위원회 의견

본 사업에서는 최종적인 연구 결과를 바탕으로 자문위원회의 검토를 받았다. 검토 결과 중에서, 표현의 제고 등 수정이 필요한 부분에 대해서는 자문위원회의 의견을 받아들이는 방향으로 진행하였다. 그 외에 자문위원들의 의견을 수렴하여 제시하면 다음과 같다.

첫째, 한국수어 말뭉치 연구에 농인들의 적극적인 참여를 촉구하는 의견이 있었다.

- 한국수어를 모국어로 사용하는 농인이 대거 참여하여 초벌 번역을 할 수 있도록 실시하되, 최종적으로는 학문적 수준이 높은 농인들이 정리를 할 수 있었으면 한다.

둘째, 전사 전문 인력을 양성하고 효율적으로 관리할 필요성이 있다는 의견이 있었다.

- 전사 전문 인력을 양성함에 있어 전사자 간 일치도 문제가 크기 때문에 이를 해결하기 위한 방안을 다각도로 찾아야 한다. 전문 전사팀을 새로 구성하여 국내에서 수행되고 있는 말뭉치 전사 사업을 병합하여 운영하는 것도 좋은 방안일 것으로 생각한다.
- 번역과 전사 품질을 위해 전사 인력을 국립국어원에서 관리하는 것이 적절할 것으로 생각된다. 번역은 외부 인원이 교육이나 회의를 통해 진행할 수도 있겠지만 전사의 경우는 국립국어원에서 채용하고 체계적으로 관리하는 것이 효과적일 것이다.

셋째, 한국수어 말뭉치 연구에 대한 필요성과 지속성을 언급하여 국립국어원의 지원 확대를 강조하기도 하였다.

- 선진국 독일 영국 네덜란드 등 외국 수어 말뭉치 사업을 살펴본다면 대부분 최소 3년 이상으로 충분한 연구 기간이 주어져서 양적으로나 질적으로나 손색이 없는 수어말뭉치 사업이 수행되고 있다고 한다. 이에 국립국어원이 단기간의 연구 과제를 매년마다 시행하기보다 2년 또는 3년 주기로 연구 프로젝트 발주 사업을 수

행하는 것이 신빙성 있는 수어 말뭉치 사업 구축에 큰 기여할 수 있으리라 생각한다.

- 전사에 투입되는 예산을 좀 더 증대하고, 시간적 여유를 가지고 심층적인 연구를 수행할 수 있도록 다양한 지원을 강화하여야 할 것이다.

넷째, 한국수어 말뭉치 연구의 발전을 위해 추후 말뭉치 연구의 보완 사항을 지적하기도 하였다.

- 워크숍 개최 장소는 가급적이면 연1회 정도는 참가자들의 이동 편의를 위한 동선을 고려하여 지역이나 장소를 선정한다면 보다 많은 인원이 참여할 것으로 기대된다.
- 수어 수집의 효율성을 높이기 위해 다양한 각도에서 촬영하도록 하여 입체적으로 볼 수 있도록 해야 할 것이다.
- 사라져 가는 과거 농식 수어 보존을 위해서 농인 어르신들의 수어를 수집할 수 있는 말뭉치 사업이 진행되었으면 한다.
- 한국수어를 모국어로 사용하는 50세 이상 농인 또는 구화를 전혀 사용하지 않는 수어 위주의 농인을 다수 선정하였으면 한다.

다섯째, 본 사업팀의 연구 성과에 대한 긍정적인 평가도 눈에 띄었다.

- 지역 균형적 말뭉치 구축을 위해 한국수어 언어 자료 수집 지역을 전북, 전남, 강원, 경남 네 지역으로 확대한 것이 매우 의미가 크다.
- 유도 자료 제작에서 언어 제공자의 지역을 고려하여 지역과 관련된 이야기를 적극적으로 진행할 수 있도록 하고, 언어 제공자의 지역과 관련된 주제를 선정하여 해당 지역수어를 포착하고자 한 것이 매우 의미가 크다.
- 의미 오류, 비규범적 표기 등의 문제를 바르게 표기한 타입명 수정이 돋보인다.
- 시스템 정비 후 타입 정보를 추가하여 타입 수형 사진, 타입 설명, 동형어를 제시함으로써 앞으로 수어 연구자와 수어 사용자 모두에게 말뭉치의 활용도를 높일 수 있는 상당히 가치 있는 자료가 될 수 있을 것이다.
- 부적격 언어 제공자를 발견하고 새로 선정하여 추가 촬영을 신속하게 진행한 점은 책임감과 연구 목적에 맞게 연구를 잘 수행한 예시로 볼 수 있다.

2. 향후 연구 방향 제언

한국수어 말뭉치 구축은 장기적으로 이루어져야 하는 사업이다. 따라서 연구진들이 지속적으로 참여 가능한 기반을 갖추는 것이 중요하다. 이에 본 연구 결과를 바탕으로 하여 다음과 같이 제언하고자 한다.

2.1. 연구 체계의 효율성 확보

본 사업의 목표는 말뚝치 구축과 전사, 워크숍을 수행하는 것이었다. 그러나 현실적으로 세 가지 층위의 일을 동시에 수행하는 것은 상당히 비효율적일 뿐만 아니라, 심층적인 연구를 수행하기에도 시간적으로 매우 부족하다. 또한 사업을 진행하는 연구진에게도 부담이 되는 일이다. 따라서 추후의 사업은 충분한 시간적인 여유를 가지고 진행해야 하며, 말뚝치 구축, 전사, 워크숍을 별도로 나누어 수행하는 것이 더욱 효과적일 것이다.

2.2. 인력 보충

1) 전사 인력 보충

본 사업에서는 전사팀을 A팀과 B팀으로 나누어 전사 작업을 진행하였다. 전북대 전사팀에서는 총 30분 분량을 담당하였는데, 세 명의 청인 연구진들이 초별 전사를 전담하고, 그 후에 정기 회의에서 농인 연구진들과 함께 전사 내용을 검토하는 방식으로 전사 작업이 이루어졌다. 전사자 한 명당 약 10분 정도의 분량이 배당되었다.

그러나 세 명의 청인 연구진들이 애초에 전사만을 담당하는 것이 아니라 말뚝치 구축부터 워크숍에 이르기까지 많은 작업을 동시에 수행해야 했기 때문에, 전사에 집중할 수 있는 시간적인 여유가 부족했다. 또한 전사 내용을 검토하는 데에도 초별 전사만큼이나 많은 시간이 소요되었다. 그러므로 절대적인 전사 인력을 보충할 필요가 있으며, 보다 더 효율적인 전사가 될 수 있도록 전사만을 담당하는 전문 전사자를 양성해야 할 것이다.

2) 촬영 인력 보충

본 사업에서는 총 4개 지역, 15쌍의 언어 제공자를 대상으로 말뚝치를 구축하였다. 그러나 연구 기간 전부를 촬영에만 쏟을 수 없는 것이 현실이다. 따라서 촬영의 효율성을 높이기 위해 촬영장을 복수로 구축하는 방안이 필요한데, 그에 따라 촬영 인력 또한 보충할 필요가 있다.

또한 연구진들이 촬영에 익숙하지 않아 연구 초반 말뚝치 구축 시 캠코더 조작에 어려움을 겪거나, 이동식 촬영장을 구축하는 데 시간을 오래 할애하는 등의 문제가 발생하였다. 그러므로 촬영 인력의 전문성을 확보하는 것도 매우 중요한 부분이다.

2.3. 예산 확보

1) 전사 시간 확대

현재 전사 작업에 대해서 1분당 3시간의 작업 시간이 책정되어 있는데 이는 실질적으로 매우 짧은 시간이다. 전사의 질을 더욱 향상하기 위해서는 최소 5시간으로 수정하는 것이 바람직하다. 이를 위해 전사에 투입되는 예산을 추가로 확보해야 한다.

2) 촬영 기반 확충

한 쌍의 언어 제공자와 사회자를 촬영하기 위해서는 최소 3대의 캠코더와 3면의 스크린, 5대의 조명, 그리고 일정 크기 이상의 촬영 공간이 필요하다. 그러나 촬영의 효율성을 높이기 위해서는 한 번에 한 쌍만을 촬영하기보다는 여러 쌍을 함께 촬영하는 것이 좋다. 따라서 촬영 기기 및 장소의 확충이 필요하며, 관련 예산을 확보할 필요가 있다.

2.4. 대학 및 관련 기관과의 협조 체제 구축

본 사업팀의 경우, 전북대학교에서 회의실과 촬영장을 구축할 수 있는 공간을 지원받아 업무를 수행하였다. 또한 말뭉치 구축을 위한 언어 제공자를 선정할 때를 비롯하여 외부 촬영을 진행할 때에도 강원 지역 농아인협회의 적극적인 협조를 받아 비교적 수월하게 연구를 진행할 수 있었다. 추후 연구에서도 여러 대학 및 관련 기관과의 긴밀한 협조 체제를 구축하는 것이 중요하다.

2.5. 교육 저변 확대

다른 학문에 비해 현재 한국수어 및 한국수어 말뭉치를 전문적으로 다루는 워크숍이나 교육의 장이 많이 부족한 실정이다. 그러나 본 사업의 워크숍을 통해 한국수어에 관심을 갖는 농인들과 청인들이 많이 있음을 확인할 수 있었다. 따라서 대중을 위한 수어 교육을 좀 더 확대해야 하며, 한국수어 및 한국수어 말뭉치 전문가를 양성하기 위한 목적으로서의 워크숍 또한 더욱 활성화되어야 할 것이다.

참고 문헌

- 국립국어원(2015), 한국수어 말뭉치 기반 조사 및 시범 구축.
국립국어원(2016), 한국수어 자료 통합 지원 시스템 구축.
국립국어원(2017), 한국수어 말뭉치 연구 및 구축.
국립국어원(2017), 한국수어전문용어사전 구축.
국립국어원, 사단법인 한국농아인협회(2014), 한국수어 코퍼스 구축을 위한 기초 연구.
국립국어원, 한국농아인협회(2013), 농인의 국어능력향상을 위한 기초연구.
국립국어원, 한국농아인협회(2014), 농인을 위한 문화 정보 및 문해 교육 지원 체계 기초 조사.
국립국어원, 한국농아인협회(2014), 농인의 문해 교육 실태 기초 연구.
Thomas Hanke, Lutz König, Sven Wagner, Silke Matthes(2010). 2010' 4th Workshop on the Representation and Processing of Sign Languages. DGS Corpus & Dicta-Sign: The Hamburg Studio Setup.
Onno Crasborn, Inge Zwitterlood(2008). 2008' 3rd workshop on the representation and Processing of Sign Language. The Corpus NGT: an online corpus for professionals and laymen.

웹사이트

- CMDI <http://www.clarin.eu/content/component-metadata>
http://media.dwds.de/clarin/userguide/text/metadata_CMDI.xhtml
DGS Korpus Projekt 독일수화 코퍼스 프로젝트 www.dgs-korpus.de
EU Project Visicast 아바타 개발 프로젝트 www.visicast.co.uk
IMDI http://corpus1.mpi.nl/ds/imdi_browser

<부록 1> 2018 한국수어 자료 수집을 위한 질문지 및 동의서

자료 수집을 위한 질문지

개인 정보

1. 이름: _____

2. 생년월일: _____

3. 주소: _____

4. 성별: ☐ 여성 ☐ 남성

5. 현재 사는 곳

_____년부터 현재까지 거주지 _____(지역명)

6. 현재까지 살았던 장소들

_____년부터 _____년까지 거주지 _____(지역명)

_____년부터 _____년까지 거주지 _____(지역명)

_____년부터 _____년까지 거주지 _____(지역명)

_____년부터 _____년까지 거주지 _____(지역명)

7. 청력상태/청력손실 시기

☐ 농 (☐ 태어날 때부터 ☐ ____살 이후부터)

☐ 난청 (☐ 태어날 때부터 ☐ ____살 이후부터)

8. 청력보장구 착용 여부

☐ 인공와우 시술을 받았다

☐ 규칙적으로 보청기를 사용한다

☐ 가끔 보청기를 사용한다

☐ 보청기 등의 보장구를 사용하지 않는다

가족 정보

9. 아버지의 청력 상태: ☐ 농인 ☐ 난청인 ☐ 청인

10. 어머니의 청력 상태: ☐ 농인 ☐ 난청인 ☐ 청인

11. 부모와의 동거 유무: ☐ 예 ☐ 아니오

12. 부모 외의 양육자

☐ 할아버지 ☐ 할머니 ☐ 친척 ☐ 없음 ☐ 기타 _____

13. 아버지와 주된 의사소통 방법

☐ 한국수어 ☐ 한국어 대응 수어(+말) ☐ 제스처 ☐ 말 ☐ 필담
☐ 홈싸인 ☐ 기타 _____

14. 어머니와 주된 의사소통 방법

☐ 한국수어 ☐ 한국어 대응 수어+말 ☐ 제스처 ☐ 말 ☐ 필담
☐ 홈싸인 ☐ 기타 _____

15. 형제 및 자매의 청력상태/나이

- ① ☐ 농인 ☐ 난청인 ☐ 청인 (나보다 나이가 _____살 ☐ 많다 ☐ 적다)
② ☐ 농인 ☐ 난청인 ☐ 청인 (나보다 나이가 _____살 ☐ 많다 ☐ 적다)
③ ☐ 농인 ☐ 난청인 ☐ 청인 (나보다 나이가 _____살 ☐ 많다 ☐ 적다)
④ ☐ 농인 ☐ 난청인 ☐ 청인 (나보다 나이가 _____살 ☐ 많다 ☐ 적다)
⑤ ☐ 형제 및 자매 없음

16. 형제 및 자매와 주된 의사소통 방법

☐ 한국수어 ☐ 한국어 대응 수어+말 ☐ 제스처 ☐ 말 ☐ 필담
☐ 홈싸인 ☐ 기타 _____

17. 가족 및 친척의 청력상태

할머니 ☐ 농인 ☐ 난청인 ☐ 청인
할아버지 ☐ 농인 ☐ 난청인 ☐ 청인
_____ ☐ 농인 ☐ 난청인 ☐ 청인
_____ ☐ 농인 ☐ 난청인 ☐ 청인

18. 배우자의 청력상태

☐ 농인 ☐ 난청인 ☐ 청인 ☐ 배우자 없음

19. 배우자와의 주된 의사소통 방법

☐ 한국수어 ☐ 한국어 대응 수어+말 ☐ 제스처 ☐ 말 ☐ 필담
☐ 홈싸인 ☐ 기타 _____

20. 자녀의 청력상태

☐ 농인 (명)
☐ 난청인 (명)
☐ 청인 (명)
☐ 자녀 없음

21. 자녀와의 주된 의사소통 방법

☐ 한국수어 ☐ 한국어 대응 수어+말 ☐ 제스처 ☐ 말 ☐ 필담
☐ 홈싸인 ☐ 기타 _____

언어

22. 언어 제공자(본인)의 주된 의사소통 방법

- ☐ 한국수어 ☐ 한국어 대응 수어+말 ☐ 제스처 ☐ 말 ☐ 필담
☐ 홈싸인 ☐ 기타 _____

23. 한국수어 습득 시기: ____살 때부터

24. 한국수어 습득 장소

- ☐ 농학교
☐ 농인 교회
☐ 농아인협회
☐ 기타 _____

25. 한글 지문자(지화) 사용 유무

- ☐ 자주 사용 ☐ 가끔 사용 ☐ 거의 사용하지 않음 ☐ 전혀 사용하지 않음

26. 지역수어 사용: 지역명 _____

27. 농(난청)인과의 정기적인 만남 유무: ☐ 예 (주 ____회 만남) ☐ 아니오

28. 본인의 한국어 능력 정도

- | | | | | | |
|----|---------------------------------|------------------------------|-----------------------------|--------------------------------|---------------------------------|
| 읽기 | <input type="checkbox"/> 아주 잘한다 | <input type="checkbox"/> 잘한다 | <input type="checkbox"/> 보통 | <input type="checkbox"/> 조금 한다 | <input type="checkbox"/> 전혀 못한다 |
| 쓰기 | <input type="checkbox"/> 아주 잘한다 | <input type="checkbox"/> 잘한다 | <input type="checkbox"/> 보통 | <input type="checkbox"/> 조금 한다 | <input type="checkbox"/> 전혀 못한다 |

29. 외국 수어 사용 여부

- | | | | | | |
|-------|---------------------------------|------------------------------|-----------------------------|--------------------------------|---------------------------------|
| 미국수어 | <input type="checkbox"/> 아주 잘한다 | <input type="checkbox"/> 잘한다 | <input type="checkbox"/> 보통 | <input type="checkbox"/> 조금 한다 | <input type="checkbox"/> 전혀 못한다 |
| 일본수어 | <input type="checkbox"/> 아주 잘한다 | <input type="checkbox"/> 잘한다 | <input type="checkbox"/> 보통 | <input type="checkbox"/> 조금 한다 | <input type="checkbox"/> 전혀 못한다 |
| 국제수어 | <input type="checkbox"/> 아주 잘한다 | <input type="checkbox"/> 잘한다 | <input type="checkbox"/> 보통 | <input type="checkbox"/> 조금 한다 | <input type="checkbox"/> 전혀 못한다 |
| _____ | <input type="checkbox"/> 아주 잘한다 | <input type="checkbox"/> 잘한다 | <input type="checkbox"/> 보통 | <input type="checkbox"/> 조금 한다 | <input type="checkbox"/> 전혀 못한다 |

학교 및 직업

30. 출신 학교

출신 학교1: _____살부터 _____살까지 _____(학교이름)

주된 의사소통 방법:

- ☐ 한국수어 ☐ 한국어 대응 수어+말 ☐ 한국수어 + 한국어 ☐ 구화
☐ 수화통역사와 함께 ☐ 기타 _____

출신 학교2: _____살부터 _____살까지 _____(학교이름)

주된 의사소통 방법:

- ☐ 한국수어 ☐ 한국어 대응 수어+말 ☐ 한국수어 + 한국어 ☐ 구화
☐ 수화통역사와 함께 ☐ 기타 _____

출신 학교3: _____살부터 _____살까지 _____(학교이름)

주된 의사소통 방법:

- ☐ 한국수어 ☐ 한국어 대응 수어+말 ☐ 한국수어 + 한국어 ☐ 구화
☐ 수화통역사와 함께 ☐ 기타 _____

출신 학교4: _____살부터 _____살까지 _____(학교이름)

주된 의사소통 방법:

- ☐ 한국수어 ☐ 한국어 대응 수어+말 ☐ 한국수어 + 한국어 ☐ 구화
☐ 수화통역사와 함께 ☐ 기타 _____

출신 학교5: _____살부터 _____살까지 _____(학교이름)

주된 의사소통 방법:

- ☐ 한국수어 ☐ 한국어 대응 수어+말 ☐ 한국수어 + 한국어 ☐ 구화
☐ 수화통역사와 함께 ☐ 기타 _____

31. 학교 기숙사 입사 유무: ☐ 예 ☐ 아니오

32. 학교 기숙사 입사에 대한 추가 정보 [31번에서 '예'라고 답한 경우]

- ☐ 기숙사에서 생활하고, 주말에 집에 규칙적으로 갔다
☐ 기숙사에서 생활하고, 방학 때만 집에 갔다

33. 최종 학력

- ☐ 초등학교 졸업 ☐ 중학교 졸업
☐ 고등학교 졸업 ☐ 전공과 졸업
☐ 전문대학 졸업 ☐ 4년제 대학교 졸업
☐ 대학원 석사 졸업 ☐ 대학원 박사 졸업
☐ 기타 _____

34. 직업훈련 기관명/지역

기관명 _____ 지역 _____

35. 대학명/전공

대학명 _____ 전공 _____

36. 직장에서 작업 형태

- ☐ 팀에서 일함 ☐ 혼자 일함

37. 직장 퇴직 여부: ☐ 예 ☐ 아니오

[37번에서 '예'라고 대답한 경우] 마지막 직장에 관한 다음 질문에 답하십시오

38. 마지막 직장 정보

직종(직장) _____ 근무 기간 _____년부터_____년까지

39. 취미, 또는 흥미가 있는 영역(예: 축구, 자동차, 요리, 독서, 등산 등)

농사회에서의 활동

40. 한국수어 강사 경험

현재 한국수어를 가르치고 있다 ☐ 예 ☐ 아니오
과거에 한국수어를 가르쳤다 ☐ 예 ☐ 아니오

41. 한국수어 강사 활동 정도 [40번에서 '예'라고 답한 경우]

☐ 1년 이하 ☐ 1~2년 ☐ 3~4년 ☐ 5년 이상

42. 수어 예술 분야 활동에서 수어 사용 여부(예: 수어 예술 분야 - 연극, 시, 퍼포먼스 등)

☐ 수어를 사용한다 ☐ 수어를 사용하지 않는다

43. 우세손

	오른손 사용	왼손 사용	가끔 왼손 사용
글쓰기	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
양치질하기	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
칼질하기	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
수어할 때	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

44. 마지막으로 하고 싶은 말

날짜: 2018년 _____월 _____일

서명: _____

영상 촬영, 배포, 판권 소유 및 자료 활용 동의서

한국수어 말뭉치 연구 및 구축을 위한 영상 촬영에 참가해 주신 여러분께 진심으로 감사의 말씀을 드립니다. 여러분의 참여는 한국 수어의 위상을 높이고 수어에 대한 연구와 응용 분야를 활성화하는 데 큰 역할을 할 것입니다. 영상 촬영 및 자료 활용을 위해 다음 사항에 대한 동의를 받고자 하니 도움을 부탁드립니다. 감사합니다.

본인(이하, 을)은 2018년 한국수어 말뭉치 연구 및 구축 사업팀(이하, 갑)이 본인을 대상으로 하여 수행하는 모든 영상 촬영에 대해 동의하며, 촬영된 영상 및 음성과 그 가공물을 한국수어 말뭉치 연구 및 구축에 사용하는 모든 권리를 갑에게 위임한다. 또한 을은 갑이 본인의 영상과 음성의 전체 또는 일부에 대해 녹화, 가공, 저장하고 배포할 권리 일체를 갑에게 부여한다.

을은 촬영된 영상과 음성이 한국수어 말뭉치 구축과 학술적 목적을 위해 활용될 수 있으며, 다양한 방법으로 가공 및 변형될 수 있음에도 동의한다. 이에 따라 을은 해당 영상의 판권 및 소유권을 주장하지 않는다는 점에 대해 동의하며, 영상과 음성 및 그 가공물에 대한 모든 소유권 및 저작권 등이 갑에게 있음을 인정한다.

을은 위의 내용에 따라 추후 초상권과 관련하여 갑이 속해 있는 개인 또는 단체, 기구 또는 인쇄매체 및 언론매체를 상대로 일체의 소송을 제기하지 않을 것에 대해서도 동의한다.

을은 자신의 영상이 촬영된 후 1주일의 경과하기 전에 을에 대한 촬영 내용 중 지극히 개인적인 것으로 공개되지 않기를 원하는 내용에 대하여 부분 삭제를 요청할 수 있다.

(위 내용을 확인하였으며 동의합니다.)

자료 제공자(을)

성명: _____

주소: (우편번호 _____) _____

전화: _____ E-mail: _____

서명: _____ 날짜: 2018년 _____월 _____일

2018년 한국수어 말뭉치 연구 및 구축 사업팀 귀중

언어 제공자 및 사회자 관계에 관한 질문지

자료 수집 날짜: _____

사회자: _____

언어 제공자 A: _____

언어 제공자 B: _____

언어 제공자 A와 언어 제공자 B의 관계

1) ☐ 우리는 서로 매우 잘 안다

1) ☐ 우리는 서로 잘 안다

1) ☐ 우리는 과거에 만난 적이 있다

1) ☐ 우리는 오늘 처음 만났다

2) 만약 서로가 아는 관계라면 어떤 관계인지 구체적으로 작성해 주세요.

2) (예: 형, 조카, 친구, 선배 등) _____

사회자와 언어 제공자 A의 관계

1) ☐ 우리는 서로 매우 잘 안다

1) ☐ 우리는 서로 잘 안다

1) ☐ 우리는 과거에 만난 적이 있다

1) ☐ 우리는 오늘 처음 만났다

2) 만약 서로가 아는 관계라면 어떤 관계인지 구체적으로 작성해 주세요.

2) (예: 형, 조카, 친구, 선배 등) _____

사회자와 언어 제공자 B의 관계

1) ☐ 우리는 서로 매우 잘 안다

1) ☐ 우리는 서로 잘 안다

1) ☐ 우리는 과거에 만난 적이 있다

1) ☐ 우리는 오늘 처음 만났다

2) 만약 서로가 아는 관계라면 어떤 관계인지 구체적으로 작성해 주세요.

2) (예: 형, 조카, 친구, 선배 등) _____

<부록 2> 한국수어 말뭉치 구축 지침(2018년 개정)

2018년 개정 한국수어 말뭉치 구축 지침

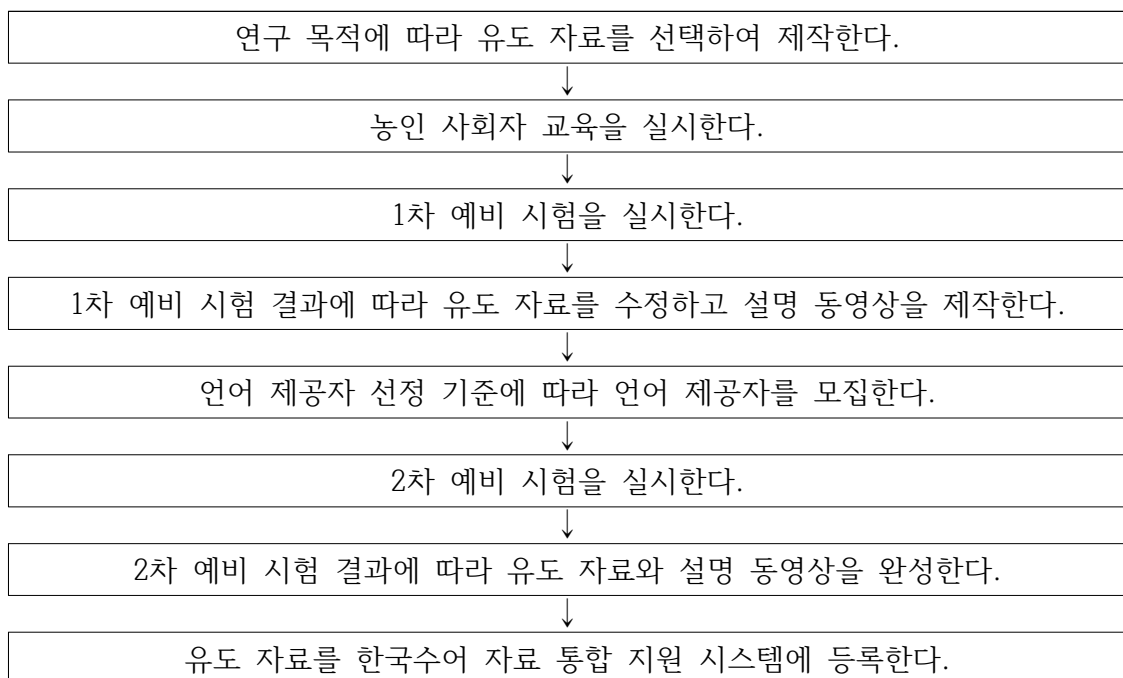
1. 개요

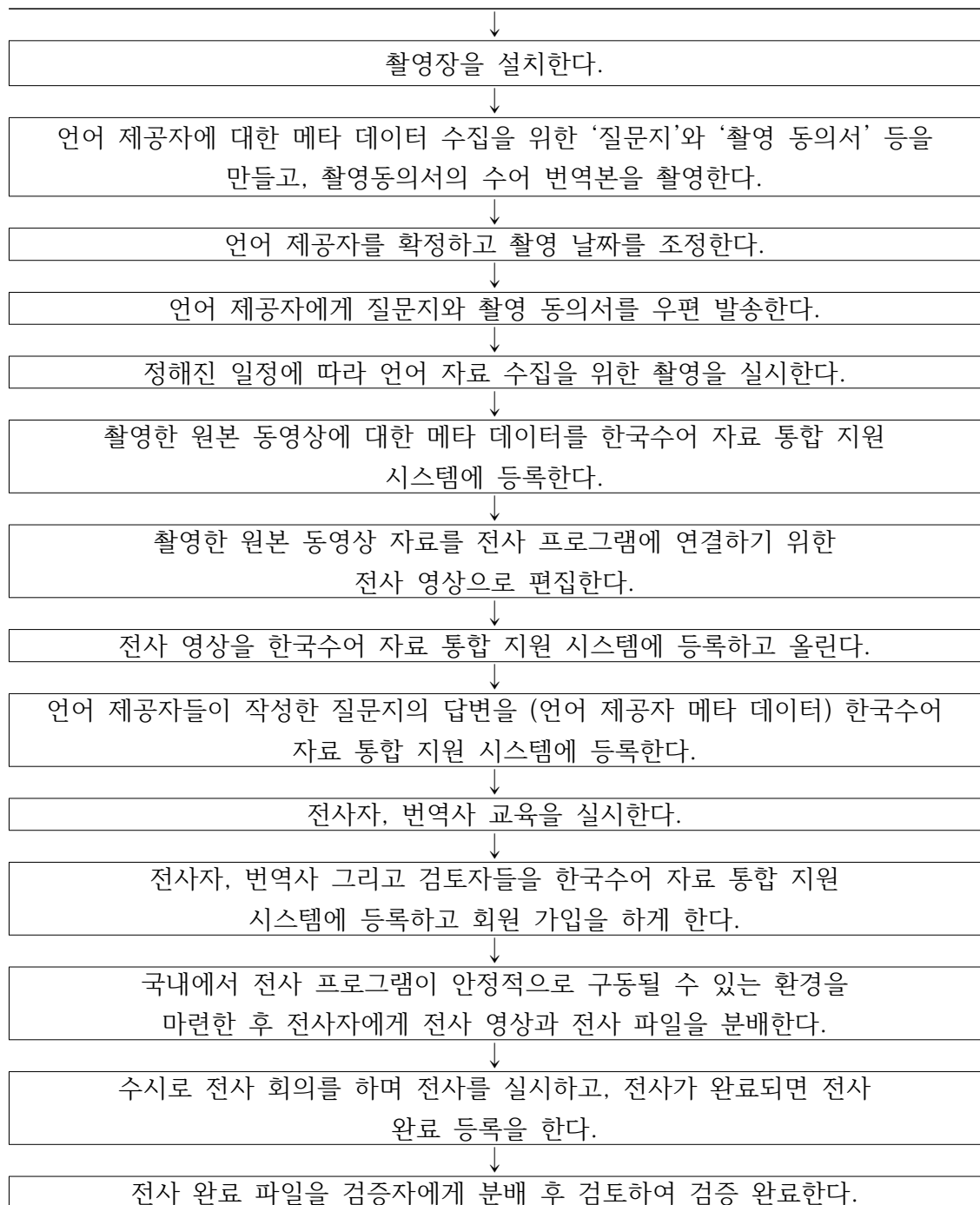
본 지침서는 2017년 개발한 한국수어 자료 통합 지원 시스템을 기반으로 한 한국수어 말뭉치를 구축을 위한 지침서이다. 본 지침서는 2017년도에 진행된 한국수어 말뭉치 연구 및 구축 사업의 후속 사업으로 2018년도에 진행한 ‘2018년 한국수어 말뭉치 연구 및 구축’의 결과를 반영하여 수정·보완 하였다.

본 지침서에서는 전반적인 구축 절차를 표로 제시한 후, 언어 자료 수집에 필요한 유도 자료의 제작에서부터 언어 제공자 모집 및 선정, 언어 자료 수집을 위한 동영상 촬영과 언어 자료의 전사에 필요한 절차 및 한국수어 자료 통합 지원 시스템에 관한 등록 방법을 제공한다. 한국수어 자료 통합 지원 시스템은 수어 말뭉치 관리, 전사 템플릿 관리, 타입 관리를 통해 전사에 필요한 정보를 관리한다. 또한 전사 관리를 통해 전사 작업을 분배하고 엘란(Elan)을 통해 전사한 결과물을 시스템에 구축하여 자료를 통합한다. 통합한 전사 자료는 한국수어 말뭉치 연구에 활용할 수 있도록 제공한다.

2. 한국수어 말뭉치 구축 절차

한국수어 말뭉치를 구축하기 위한 절차는 다음과 같다.





3. 한국수어 말뭉치 수집 지침

3.1. 유도 자료

3.1.1. 유도 자료의 목적 및 필요성

수어를 연구하거나 한국수어 말뭉치를 구축하기 위해서는 자연스러운 상황에서 다양한 수어 어휘와 표현 등을 수집하여야 한다. 그러나 수어 자료를 수집할 때는 음성 언어 자료를 수집할 때와는 달리 동영상을 촬영해야 하기 때문에 언어 제공자들이 자연스러운 상황에서처럼 대화가 이루어지지 않아 다양한 언어 자료를 수집하기가 어렵다. 따라서 다양한 어휘와 표현을 끌어낼 수 있는 그림이나 과제 또는 질문 등을 만들어 제공할 필요가 있다. 이때 제공하는 과제나 그림을 유도 자료라 한다.

이러한 유도 자료는 음성 언어와 관계없이 수어를 끌어낼 수 있으며, 수어 산출(output)을 조절할 수 있다는 장점이 있다. 즉, 주제 영역이나 담화 양식 등을 조절하므로 연구 목적에 맞는 수어 자료를 이끌어낼 수 있다. 그러나 수어를 사용하는 농인들도 한국어 사용 환경에 노출되어 있으므로 음성 언어도 사용한다. 이 때 음성 언어는 주로 한국어 문법을 따르는 수어인 유사 수어 체계(Signed Korean)와 함께 사용된다. 수어 데이터를 모집할 때는 음성 언어의 영향을 최대한 줄이는 것이 중요한 점이다. 그러므로 글자를 피하고 시각적인 유도 자료를 사용하는 것이 좋다.

3.1.2. 유도 자료의 종류

홍성은(2010)에 의하면 유도 자료의 형태는 다음의 5가지 종류가 있다. 한국수어 말뭉치 구축이나 연구의 목적에 따라 다음의 유도 자료 형태 중 선택하여 유도 자료를 제작한다. 각 유도 자료 형태에 따른 특성과 장단점은 다음과 같다.

1) 언어 입력: 문자 언어와 수어

① 문자 언어: 수어는 문자가 없기 때문에 음성 언어의 문법 체계를 따르는 글을 말한다. 그림과는 달리, 연관된 사물이 상세하게 표기될 수 있다는 사실이 일반적인 장점이다. 그러나 문자 언어는 농인 언어 제공자에게 음성 언어 영향을 미치고, 언어 제공자들이 한국수어(KSL) 대신 유사 수어 체계(Signed Korean)를 사용한다는 큰 위험성이 있다. 단어 목록표를 사용한다면 개별 용어를 물어볼 수 있다는 게 장점이겠지만 맥락 없이 애매한 다의어 표현으로 잘못된 답을 찾게 될 수도 있다. 그리고 단어를 통해 얻어진 수어 데이터는 자연스러운 데이터가 아니고, 맥락을 고려하지 않은 단순한 번역일 수 있다는 단점이 있다.

② 수어: 직접 수어를 주고 언어 제공자로부터 수어를 끌어내는 유도 자료를 말한다. 농인 언어 제공자가 자신에게 제일 편한 언어로 말할 수 있다는 것이 큰 장점이다. 그러나 수어를 직접적으로 보여줄 경우 언어 제공자가 자기의 수어보다는 제시된 수어를 사용할 수 있다는 단점이 있다. 수어를 유도 자료로 사용할 때는 현장에서 바로 제시하기보다 미리 제작된 수어 영상물을 보여주는 것이 신뢰성이 높고 조절이 가능하다. 수어 영상물의 장점은 현장에서 일어날 수 있는 그 어떠한 실수도 없이 언어 제공자들에게 수시로 상영될 수 있다는 것이고, 단점은 어떠한 상호 작용도 일어날 수 없다는 것이다.

2) 그림, 사진, 만화

① 그림: 움직임이 없는 모든 형태의 그림을 의미한다. 그림은 시각적인 자극으로서 농인 언어 제공자들에게 매력적인 유도 자료이다. 또한 그림은 간단하게 만들 수 있고 어떤 그림 스타일이 선정되는지에 따라 그림의 세부 사항 정도가 변형될 수 있다. 그림이 상세할수록 사물은 좀 더 세부적으로 묘사될 수 있지만, 그림이 자세할수록 언어 제공자들이 상상할 수 있는 공간은 더 좁아지기 때문에 풍부한 상세함이 곧 약점이 될 수 있다. 개별 그림을 사용하는 경우 서로 연관성이 없는 그림에 관한 것으로 종종 사람, 동물, 사물 혹은 어떠한 사건, 행동, 활동 등이 그려져 있다. 서로 연관성이 없는 그림 사용의 장점은 언어 제공자가 그림에 대해 설명을 해야 할 때, 모든 것을 세부적으로 정확하게 말해야 한다는 점이다. 그러나 단점은 그림들이 일관성이 없어 보이기 때문에 언어 제공자들의 진술이 꽤 부자연스럽다는 것이다. 연관성이 있는 그림이야기는 대부분 그 내용이 시간적인 순서대로 전개되는 연속적인 장면들로 만들어진 그림이야기를 의미한다. 그림이야기는 복잡하고 긴 내용을 시각적으로 보여주는 장점도 있지만, 이야기 내용이 그림으로 축소되면서 부분적으로 오해를 일으킬 수도 있다는 단점이 있다.

② 사진: 실제와 가장 가깝고 세부적이라는 큰 장점을 가지고 있다. 다른 한편으로는 사물에 대해서 추상화하는 것이 난해하므로, 사진이 가지고 있는 세부적인 면이 바로 단점이 될 수도 있다. 또한 한 사물을 찍은 사진이 그 분류에 속하는 사물의 총칭(예: 상위어)을 의미하는지, 정확한 부류(예: 하위어)를 나타내는지, 언어 제공자들에게는 종종 명확하지 않다. 예를 들어, 너도밤나무를 찍은 사진일 경우 이 사진을 보고 “너도밤나무”라고 대답할 수도 있지만, 이것의 총칭으로 “나무”라고 할 수도 있다는 것이다. 이 외에도, 뚜렷하고 세부적인 사진 속에는 원래 사진의 주체와 상관없는 배경들 또한 자세하게 보이기 때문에, 언어 제공자들이 사진에 집중하는 데 방해받을 수 있다. 또 저작권 문제가 발생할 수도 있다는 단점도 있다.

③ 만화: 딱딱하지 않은 비공식적인 분위기를 만들어 낼 수 있기 때문에, 가장 자연스럽고 편안하게 데이터를 얻어낼 수 있다. 유명한 만화 인물을 보여 줄 경우 언어 제공자들이 바로 알아보는 효과가 있을 수 있다. 그러나, 만화 속에서만 사용하는 규

칙이나 표현법이 있기 때문에 만화에 익숙하지 않은 사람들의 경우에는 내용을 이해하는 데 어려움이 따를 수 있다. 또한 유명한 만화는 대부분 저작권 문제가 있기 때문에 사용 시 꼭 유의해야 한다.

3) 동영상과 애니메이션

① 동영상: 연속적으로 움직이는 그림이나 형태를 보여주는 것을 의미한다. 그림과 마찬가지로 동영상도 시각적 자료이기 때문에, 수어 자료를 얻어내는 데 매우 적합하다. 또한 동영상은 언어 제공자에게 매우 매력적이고 재미있는 자극이 될 수 있으므로 유도 자료로서 언어 제공자에게 긍정적인 효과를 낼 수 있다. 한편, 영상물은 기술 장비가 필요하다는 불편한 점이 있다. 영상물은 사진과 마찬가지로 사물, 사람, 사건을 현실적으로 묘사할 수 있는 장점을 가진 반면, 일반 영화나 영화의 한 장면을 자극으로 제공할 경우 저작권에 대한 문제가 있다는 것이 단점이다.

② 애니메이션은 주인공들이 사람과 직접적인 연관이 없기 때문에, 성과 인종에 있어 중립적인 위치를 유지할 수 있다. 그러나 애니메이션을 직접 제작하는 데 매우 많은 비용이 들어간다는 단점이 있다.

4) 주제, 동화와 우화

① 주제: 언어 제공자들이 특정 주제에 대해 이야기를 나누게 하여 언어 자료를 수집하는 것이다. 이때, 언어 제공자들에게 스스로 주제를 찾거나 결정을 하게 하면 그들이 선택한 주제에 대해 이야기를 할 수 있기 때문에 언어 자료가 매우 자연스럽게 되는 큰 장점이 있다. 자유 주제로 이야기할 때, 언어 제공자들에게 이야기 전개와 표현에 있어 충분한 시간이 주어지기 때문에, 그들은 자신이 실험 상황에 있다는 것을 잊어버릴 확률이 매우 높다. 그러나 유도된 언어 데이터를 서로 비교하는 것이 거의 불가능하다는 단점이 있을 수 있다. 자유 주제가 아닌 토론 주제로 언어 자료를 유도할 경우 찬반론으로 대립적인 토론이 가능한 주제는 언어 제공자 개개인이 자신들이 개인적인 생각을 좀 더 많이 그리고 자주 표현할 수 있다는 장점을 가지고 있다. 그러나 언어 제공자들이 그들의 개인적인 생각을 다른 사람들 앞에서 이야기하는 것을 부끄러워할 수도 있는데, 이 같은 경우는 언어 자료를 유도하기가 어렵다. 또한 모든 언어 제공자들이 거의 같은 의견을 가지고 있을 경우에도 토론다운 토론을 이끌어 내는 것이 힘들다.

② 동화나 우화: 잘 알려진 동화나 우화의 내용을 그림이나 동영상을 통해 제시하는 것을 의미한다. 이미 잘 알고 있는 내용의 동화나 우화를 유도 자료로 제시하면 문자 언어 자극과 같은 특별한 도움 없이 확실한 내용을 물어볼 수 있다는 장점이 있다. 특히 다른 나라에서 유도된 데이터를 비교할 수 있는 장점이 있다. 다만 유명한 동화와 우화가 문화에 따라 다를 수 있고, 농인들이 언어 상황에 따라 청인들이 알고

있는 동화를 잘 모를 수도 있다는 점을 고려하여야 한다.

5) 놀이

언어 제공자들이 서로에게 이야기할 수 있게 하는 또 다른 방법은 놀이이다. 놀이는 언어 제공자들에게 대화를 나눌 수 있게 동기를 부여하여 서로 자연스럽게 대화를 나누게 한다. 놀이는 언어 제공자들에게 카메라를 의식하지 않도록 하므로 자연스럽게 언어 자료를 유도할 수 있게 한다. 반면에, 승패가 있는 놀이는 언어 제공자들이 서로 지지 않으려 하기 때문에 그들에게 스트레스를 줄 수도 있다.

3.1.3 유도 자료 제시 방법

- ① 기본적으로 유도 자료는 모니터 화면을 통해 슬라이드 쇼의 형태로 제시되어야 한다.
- ② 유도 자료를 언어 제공자에게 동일한 방식으로 제공하기 위해 유도 자료 설명 영상을 제작해야 한다.
- ③ 유도 자료 설명 영상을 시청하게 한 뒤 유도 자료를 제시하여 유도 자료를 올바르게 이해하도록 해야 한다.
- ④ 한 언어 제공자에게 유도 자료를 제시할 때 다른 언어 제공자에게는 검은 슬라이드를 제시하여 유도 자료 제시로 인한 촬영의 지장이 없도록 해야 한다.
- ⑤ 사회자는 유도 자료 설명 영상과 유도 자료가 제시된 뒤 추가적인 설명을 할 수 있다.

3.1.4 유도 자료 등록

유도 자료로 제작하여 사용한 특정 유도 과제는 한국수어 자료 통합 지원 시스템에 등록한다.

유도 과제 등록

1

유도 과제명

유도 과제 개요

2

유도 과제 파일 추가

파일종류	파일붙임	파일설명	삭제
수어설명영상 ▼	0KB / 100MB <div></div> <div>선택</div> <div>삭제</div>	<div></div>	<div>삭제</div>
과제영상 ▼	0KB / 100MB <div></div> <div>선택</div> <div>삭제</div>	<div></div>	<div>삭제</div>
과제사진 ▼	0KB / 10MB <div></div> <div>선택</div> <div>삭제</div>	<div></div>	<div>삭제</div>

3

등록

목록

<그림 1> 유도 자료 등록

가. 유도 자료 등록 화면으로 유도 자료명, 유도 자료 개요 등록 양식을 제공한다.

나. 유도 자료 파일 등록 영역으로 유도 자료 파일 추가 버튼 클릭 시 유도 자료 파일 붙임 양식이 생성된다.

- 파일 종류는 수어 설명 영상, 과제 영상, 과제 사진 중 선택한다.
- 파일 붙임은 선택 버튼을 클릭하여 파일을 첨부하고, 삭제가 필요할 경우에는 삭제 버튼을 클릭하여 첨부한 파일을 삭제한다.
- 파일 설명은 유도 자료 파일에 대한 파일 설명을 등록한다.
- 삭제 버튼 클릭 시 해당 유도 자료 파일 양식을 삭제한다.

유도 자료 파일 버튼 영역으로 등록 버튼 클릭 시 유도 자료 등록이 처리된다. 목록 버튼 클릭 시 유도 자료 목록으로 이동한다.

3.2. 언어 제공자 모집 및 선정

3.2.1 언어 제공자 기본 조건

- ① 19세 이상 성인 농인
- ② 한국수어를 모국어(제1언어)로 사용하는 농인
- ③ 농학교를 졸업한 농인
- ④ 조사 지역으로 선정된 행정 구역 내에서 최근 10년 이상 지낸 농인

- 149 -

- ⑤ 농사회에서의 교류가 활발한 농인

3.2.2. 언어 제공자 모집 방법

- ① 각 지역 농아인협회를 통한 모집(협조 공문 활용)
- ② 누리소통망(SNS)을 이용한 온라인 광고
- ③ 지역 농사회 커뮤니티를 통한 직접 접촉

3.2.3. 언어 제공자 선정 시 유의사항

- ① 언어 제공자 기본 조건 충족 여부 확인
- ② 다음의 우선 선발 조건을 적용한 언어 제공자에게 선순위 부여
 - 농부모, 농형제, 또는 농배우자가 있음
 - 초등학교 이전에 한국수어를 습득하였음
- ③ 언어 제공자 간의 연령, 성별, 친밀도를 고려한 언어 제공자 쌍 선정
- ④ 촬영 장소 및 촬영 일자 확정

3.2.4. 언어 제공자에 대한 자료 수집 및 촬영 동의서

- ① 언어 제공자 정보 확보를 위한 질문지 및 촬영 동의서 작성(부록 참조)
- ② 촬영된 말뭉치 영상의 초상권과 저작권에 관련된 “영상 촬영, 배포, 판권 소유 및 자료 활용 동의서” 및 “개인 정보 수집·이용·제공 및 조회 동의서” 작성
- ③ 질문지, 동의서는 언어 제공자가 촬영일 이전에 작성할 수 있도록 사전 전달
- ④ 촬영 당일 언어 제공자와 사회자 관계에 대한 질문지(부록 9 참조)를 작성

3.2.5. 언어 제공자 넘버링(numbering) 및 기초 자료 탑재

1) 언어 제공자 넘버링 기본 규칙

- ① 언어 제공자는 촬영된 순서에 따라 번호를 매긴다. 동일한 쌍에서는 언어 제공자A가 앞선 번호를 받는다.
- ② 언어 제공자 쌍은 촬영 지역 별로 순서에 따라 번호를 매긴다. 언어 제공자 쌍의 지역 구분은 다음과 같다.

지역명	포함 지역
서울	서울, 경기, 인천 등 수도권
강원	강원도 전체
충북	충청북도 전체
충남	충청남도 및 세종, 대전 지역
경북	경상북도 및 대구 지역
경남	경상남도 및 부산, 울산 지역
전북	전라북도 전체
전남	전라남도 및 광주 지역
제주	제주도 전체

2) 언어 제공자 기초 자료 탑재

한국수어 말뭉치 자료 통합 지원 시스템에 수거한 질문지와 동의서를 기반으로 한 언어 제공자 기초 자료를 업로드한다. 언어 제공자 정보를 주제별로 구분된 시스템 화면에 탑재한다.

1 언어 제공자 등록

개인 정보

가족 정보

언어

학교 및 직업

농인 공동체 안에서의 활동(예, 농인 클럽 내에서)

2

12. 아버지

☐ 농인
☐ 난청인
☐ 청인
☐ CODA(농부모 청인자녀)

13. 어머니

☐ 농인
☐ 난청인
☐ 청인
☐ CODA(농부모 청인자녀)

14. 나는 부모님과 함께 3세부터 7세까지 살았다

☐ 예
☐ 아니요

15. 나는 다른 양육자의 돌봄을 받았다

☐ 이돌 양육중 남성은
☐ 농인
☐ 난청인
☐ 청인
☐ CODA(농부모 청인자녀)

☐ 이돌 양육중 여성은
☐ 농인
☐ 난청인
☐ 청인
☐ CODA(농부모 청인자녀)

16. 어머니와의 주된 의사소통 방법

☐ 한국수어(KSL)
☐ 유사수어체계(Signed Korean)
☐ 몸짓이나 제스처
☐ 말(음성 언어)

☐ 필담(문자 언어)
☐ 흠씨인
☐ 기타

17. 아버지와와의 주된 의사소통 방법

☐ 한국수어(KSL)
☐ 유사수어체계(Signed Korean)
☐ 몸짓이나 제스처
☐ 말(음성 언어)

☐ 필담(문자 언어)
☐ 흠씨인
☐ 기타

18. 형제 자매

추가

☐ 농인
☐ 난청인
☐ 청인

살
☐ 나보다 나이가 많다
☐ 나보다 나이가 어리다

삭제

19. 형제 및 자매들과의 주된 의사소통 방법

☐ 한국수어(KSL)
☐ 유사수어체계(Signed Korean)
☐ 몸짓이나 제스처
☐ 말(음성 언어)

☐ 필담(문자 언어)
☐ 흠씨인
☐ 기타

20. 가족이나 친척들 중 농인이나 난청인이 있습니까?

추가

할머니

☐ 농인
☐ 난청인

삭제

21. 배우자

☐ 농인
☐ 난청인
☐ 청인
☐ CODA(농부모 청인자녀)
☐ 배우자 없음

22. 자녀수

농인

선택하세요

명

난청인

선택하세요

명

청인

선택하세요

명

23. 배우자와의 주된 의사소통 방법

☐ 한국수어(KSL)
☐ 유사수어체계(Signed Korean)
☐ 몸짓이나 제스처
☐ 말(음성 언어)

☐ 필담(문자 언어)
☐ 흠씨인
☐ 기타

24. 자녀와의 주된 의사소통 방법

☐ 한국수어(KSL)
☐ 유사수어체계(Signed Korean)
☐ 몸짓이나 제스처
☐ 말(음성 언어)

☐ 필담(문자 언어)
☐ 흠씨인
☐ 기타

3

이전

다음

목록

<그림 2> 언어 제공자 등록창

3.2.6. 촬영 설계 시 유의점

- ① 언어 제공자와 지속적으로 연락을 유지하여 촬영이 취소되지 않도록 한다.
- ② 질문지와 동의서는 촬영 장소에도 비치하여 언어 제공자가 관련서류를 미리 작성해오지 않더라도 촬영 장소에서 바로 작성할 수 있도록 한다.
- ③ 전체 촬영은 촬영 대상 시간보다 1.5배 정도 소요되기 때문에 촬영 시간 선택을 소요시간과 중간 휴식시간을 감안하여 선정한다.
- ④ 언어 제공자와 사회자 간의 친밀도 향상을 위해 촬영 전 예비 모임을 진행한다.

3.3. 언어 자료 수집

3.3.1. 촬영실 구축

한국수어 말뭉치의 구축은 동영상 촬영을 통해 진행되기 때문에 음성 말뭉치보다 촬영의 변수가 많고 언어 제공자에게도 위화감을 줄 수밖에 없다. 이 때문에 말뭉치 영상이 촬영되는 공간인 촬영실의 구축에 각별한 주의를 기울일 필요가 있다. 그리고 한국수어 말뭉치의 수집 범위가 확대됨에 따라 거점 촬영실 이외에 이동식 촬영실을 구성해야 한다.

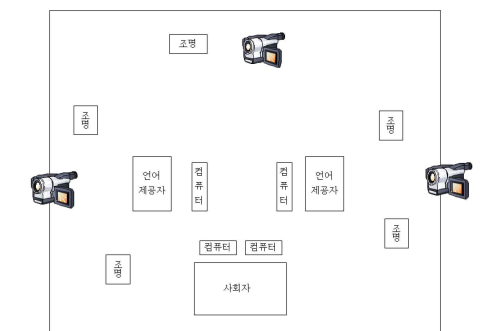
1) 촬영실 공간 구성의 기본 전제

- ① 고정된 기준보다 예비 촬영을 통한 촬영 결과물을 토대로 촬영실을 구축해야 한다.
- ② 수형, 수위, 수동을 간섭 없이 확인할 수 있어야 하고 언어 제공자의 표정과 입 모양도 확인이 가능해야 한다.
- ③ 언어 제공자로 하여금 위화감을 느낄 수 있는 요소는 촬영 결과물에 간섭하지 않는 범위 내에서 최대한 제외시킨다.
- ④ 언어 제공자에게 편의를 제공할 수 있도록 상황에 맞춰 촬영실 환경을 가변적으로 조절해야 한다.

2) 촬영실 세부 구성

- ① **촬영 배경:** 언어 제공자의 수어를 잘 포착할 수 있도록 촬영 배경을 설치한다. 구성한다. 배경은 크로마키 스크린을 벽면에 부착하거나 별도 장치를 통해 거치한다. 색은 청색 계열로 한다.

- ② **캠코더**: 720p 이상의 해상도로 촬영될 수 있어야 하며 호환의 문제가 없도록 MP4 포맷으로 촬영해야 한다. 그리고 두 언어 제공자와 전경을 찍을 수 있도록 세 대의 캠코더를 설치하여 촬영을 진행하며 언어 제공자에게 큰 위화감을 주지 않도록 캠코더의 크기에 유의해야 한다.
- ③ **조명**: 촬영 영상을 확인했을 때 그림자가 생기지 않도록 조명을 설치하되 예비 촬영을 통해 유동적으로 설치한다. 조명은 800w 이상의 출력이어야 하며 조명 광의 지향성이 없도록 소프트 박스나 반사판을 활용해야 한다. 조명이 촬영 결과물의 품질을 올리는 데에 기여하기는 하지만 언어 제공자에게 위화감을 줄 수 있기 때문에 최소한의 조명만을 설치해야 한다.
- ④ **사회자용 PC**: 언어 제공자에게 유도 자료를 제시하기 위해 사회자용 PC를 설치한다. 두 언어 제공자에게 각각 다른 화면을 보여줄 수 있도록 두 대의 PC를 사회자가 조작할 수 있는 환경을 조성해야 한다.
- ⑤ **유도 자료 제시용 모니터**: 언어 제공자에게 유도 자료를 보여주는 화면을 두 언어 제공자 사이에 각각 설치하고 화면이 서로의 대화를 방해하지 않는 최대 높이에 설치한다. 동일한 규격의 화면이 제시될 수 있도록 모니터는 같은 모델을 설치해야 한다.



<그림 3> 촬영실 모습



<그림 4> 실제 촬영실 구성

3) 촬영실 구성의 유의사항

- ① 거점 촬영실은 접근성이 좋은 장소로 선정해야 한다.
- ② 촬영이 4시간 이상 진행되기 때문에 캠코더의 용량과 배터리 방전에 주의해야 한다.
- ③ 촬영 장비는 가급적 언어 제공자에게 위화감을 주지 않도록 설치해야 한다.
- ④ 촬영 결과물의 균질성이 확보되는 범위 안에서 최대한 언어 제공자에게 편의가 제공될 수 있어야 한다.

3.3.2. 사회자 교육

1) 사회자의 역할

- ① 말뭉치 영상 촬영 전반을 진행한다.
- ② 언어 제공자에게 유도 자료를 제시하고 필요시 추가 설명을 한다.
- ③ 과제 소요 시간을 조절한다.
- ④ 대화가 자연스럽게 진행될 수 있는 분위기를 조성한다.
- ⑤ 언어 제공자의 대화를 중재하고 촉진한다.
- ⑥ 촬영 일지를 기록한다.

2) 사회자 교육

- ① 한국수어 말뭉치와 유도 자료의 내용 및 목적에 대한 이해
- ② 언어 제공자가 동영상 설명 자료를 이해하지 못할 시 필요한 보충 설명 방법
- ③ 언어 제공자가 자연스러운 분위기에서 편하게 대화할 수 있는 분위기 조성 방법
- ④ 언어 제공자의 대화가 단절되었을 때 필요한 대화 촉진 전략
- ⑤ 두 언어 제공자의 수어가 균형 있게 산출될 수 있도록 조절하는 전략
- ⑥ 각 과제의 소요 시간을 균형 있게 조절하는 전략
- ⑦ 그 외 사회자가 지켜야 할 태도

3.3.3. 실제 촬영

1) 촬영 인원의 구성 및 역할

- ① 사회자: 촬영 전반에 대한 진행
- ② 촬영 감독: 카메라와 조명 상태를 점검하여 촬영 결과물이 잘 얻어질 수 있도록 조정하며 촬영된 영상 데이터를 관리
- ③ 촬영 보조원: 언어 제공자 안내, 사회자 및 촬영 감독 보조

2) 촬영 진행 순서

- ① 촬영 전 언어 제공자가 작성한 질문지와 동의서를 수령한다.
- ② 언어 제공자와 사회자 간의 친밀감 형성을 위해 간단한 다과를 하며 시간을 갖는다.
- ③ 촬영실 기자재를 점검하고 촬영 준비를 한다.
- ④ 촬영이 진행될 때는 촬영 공간에 촬영 인원을 제외하고는 접근할 수 없게 한다.

- ⑤ 사회자는 촬영 시작 지점과 과제의 시작 지점에서 종을 친다.
- ⑥ 촬영이 끝난 후 촬영된 영상에 문제가 없는지를 점검하고, 3대의 카메라로 촬영한 영상을 폴더별로 구분하여 저장함과 동시에 백업 영상을 만든다.
- ⑦ 촬영 일지를 기록한다.
- ⑧ 검수를 마친 원본 영상은 한국수어 자료 통합 지원 시스템에 등록한다.
- ⑨ 원본 영상 등록 창은 <그림 5>와 같고, 등록 방법은 다음과 같다.

원본 영상 등록

1

사업명

2015년 한국수어 시범 말뚝치 ▼

제목

11번째 원본 영상

저작자

국립국어원

제작일자

20150701

촬영장소

스튜디오

확장자

MTS ▼

해상도

1280X720 ▼

파일 총 용량

10 G

영상시간

01 시 12 분 50 초

수어장르

토론 ▼

유도 자료 종류

이미지 ▼

2

카메라

카메라추가

카메라 기종	카메라명	파일명	프레임수	화면화질	연결된층위	층위	삭제
			25 ▼	HD ▼			

연결

삭제

3

언어제공자

언어제공자추가

이름	성별	출생년도	지역	삭제
최봉규	남성	1989년	서울	

삭제

4

등록

목록

<그림 5> 원본 영상 정보 등록

- 가. 원본 영상 정보 메타 정보 등록 영역으로 사업명, 제목, 저작자, 제작일자, 촬영장소, 확장자, 해상도, 파일 총 용량, 영상 시간, 수어장르, 유도 자료 종류로 메타 정보를 등록한다.
- 나. 카메라 정보 등록 영역으로 카메라 기종 카메라명, 파일명, 프레임 수, 화면 화질, 연결된 층위를 등록한다.
 - 카메라 추가 버튼 클릭 시 카메라 정보 등록 양식이 추가된다.
 - 연결 버튼 클릭 시 층위 연결 팝업을 통해 층위를 연결한다.
 - 삭제 버튼 클릭 시 해당 카메라 정보 등록 양식을 삭제한다.
- 다. 언어 제공자 등록 영역으로 언어 제공자 검색 팝업을 통해 언어 제공자를 검색 후 추가한다.
- 라. 원본 영상 정보 버튼 영역으로 등록, 목록 버튼을 제공한다.

- 등록 버튼 클릭 시 원본 영상 정보를 등록 처리한다.
- 목록 버튼 클릭 시 원본 영상 정보 목록으로 이동한다.

3) 유의사항

- ① 사회자는 언어 제공자들이 서로 하고자 하는 말을 자유롭게 할 수 있도록 배려한다.
- ② 촬영 감독은 카메라 등 촬영 환경에서 문제가 생길 때 즉시 확인하여 조치해야 하므로 농인이 할 것을 권장한다.
- ③ 촬영한 동영상은 촬영을 마친 후 바로 문제가 있는지를 확인하여 오류가 발견될 시 보완을 위한 조치를 신속하게 진행해야 한다.

4. 자료 처리 지침

4.1. 동영상 편집 절차 및 시스템 탑재

4.1.1 동영상 편집 절차

- ① 영상 편집은 Adobe사의 Premiere CC를 활용한다.
- ② 촬영된 원본 영상의 시작점과 끝점을 동기화 시켜 전체 원본 영상을 제작한다. 동기화 지점은 촬영 시작 시 사회자가 울린 종을 기준으로 한다.
- ③ 유도 자료별로 동영상을 분절하여 번역 및 전사를 위한 2차 파일을 제작한다. 제작 시 엘란(Elan)에서 작업이 용이하도록 MPEG1 코덱에 720p 해상도, 비트율 3000 이하로 인코딩한다.
- ④ 유도 자료별로 분절된 영상의 파일명은 “언어 제공자 쌍-유도 자료 이름-언어 제공자”으로 한다. 언어 제공자는 전경을 촬영한 사회자 영상 기준 왼쪽의 언어 제공자를 언어 제공자A로 하고 오른쪽을 언어 제공자B로 한다.

4.1.2 편집 동영상 검수 및 등록

편집이 완료된 동영상은 다음의 절차에 따라 검수 확인 절차를 거친다.

- ① 과제별로 오른쪽, 왼쪽, 중앙 카메라 촬영 파일이 모두 있는지 확인한다.
- ② 같은 과제의 영상에서 언어 제공자A, 언어 제공자B, 사회자 3개의 동영상의 재생 시간이 동일한지 확인한다.
- ③ 엘란(Elan) 프로그램인 eaf 파일에서 언어 제공자A, 언어 제공자B, 사회자 총 3

개 동영상 파일을 불러와 다음의 사항을 확인한다.

- 앞뒤에 여유가 있는가?
- 파일명과 영상의 내용이 맞는가?

- ④ 문제가 있는 영상을 발견하면 기록하여 편집자와 함께 수정하도록 한다.
- ⑤ 편집 검수 완료 동영상은 전사 영상이 되며, 이를 한국수어 자료 통합 지원 시스템에 등록한다.
- ⑥ 전사 영상은 <그림 6>과 같이 시스템에 등록한다.(자세한 내용은 ‘한국수어 자료 통합 지원 시스템 사용 설명서’ 참고)

전사 영상 등록

1

제목

원본 영상 정보

2

5번째 원본 영상

연결

3

카메라명	영상품질	원본파일명	붙임파일
왼쪽	선택하세요 ▼	KSLSL5XLT.MTS	0KB / 1200MB
오른쪽	선택하세요 ▼	KSLSL5XRT.MTS	0KB / 1200MB

4

유도 과제

유도 과제명

등록자

등록일

삭제

과제 번호 1: 옛날 수어와 현재 수어

시스템 관리자

2016. 11. 29.

삭제

유도 과제 추가

5

등록

목록

<그림 6> 전사 영상 등록

5. 전사 지침

5.1.개요

5.1.1. 기본 사항

- ① 전사를 위해 엘란(Elan) 프로그램을 사용한다.
- ② 전사를 위한 전사자 교육을 실시한다.
- ③ 전사의 일치도를 높이기 위해 공동 전사를 원칙으로 한다. 공동 전사가 어려운 경우 최소한 처음 5회 이상은 공동 전사를 실시한다.
- ④ 전사의 일치도를 높이기 위해 주 1회 이상의 전사 회의를 실시한다.
- ⑤ 전사 토큰 타입 매칭(token-type matching) 일치도를 높이기 위해 한국수어 자

료 통합 지원 시스템을 사용한다.

- ① 언어 자료 영상에서 2명의 언어 제공자가 각각 찍힌 영상(full shot)의 왼편 언어 제공자는 A, 오른편 언어 제공자는 B이다.
- ② eaf 파일(엘란(Elan) 프로그램)에 전사를 마친 파일명은 한국수어 자료 통합 지원 시스템에서 내려 받은 파일명과 동일해야 하므로 절대 수정하지 않도록 한다.

5.1.2. 엘란(Elan)의 특성 및 예시

엘란(Elan)(EUDICO Linguistic Annotator)은 비디오 및 음성 자료에 대해 전사를 입력하고, 편집하고, 시각화하고, 검색할 수 있는 전사 도구이다. 이 프로그램은 멀티미디어 기록물에 대한 주석 작업과 탐색의 기술적 기초를 제공하기 위해 네덜란드 막스플랑크 심리언어학 연구소에서 개발하였다. 엘란(Elan)은 일반 언어, 수어, 제스처의 분석을 위해 특별히 설계된 프로그램이지만, 다양한 영상·음성 자료를 대상으로 주석 작업과 분석, 문서화를 통해 미디어 말뭉치를 구축하려는 경우에도 유용하게 사용될 수 있다.



<그림 7> 엘란(Elan) 전사 예

위의 그림은 엘란(Elan)으로 한국수어 동영상을 전사한 예를 보인 것이다. 엘란(Elan) 창에는 많은 정보 메뉴들이 있는데, 좌측 상단에는 수어 화자 영상 화면이 배치되어 있고 오른쪽 위에는 인식 정보와 메타자료 메뉴가 있다. 아래 부분에는 입력된 주석과 타임라인 창이 함께 제시된다. 각 주석은 층렬에 따라 입력되어 있는데, 시간 간격에 따라 정렬되어 있다. 모든 메뉴나 창들은 서로 연동되고, 한 시점에서 동일한 수어 영상에 관한 사항을 나타낸다. 이는 한 창에서 어느 시점의 주석이나 영상을 클릭하면 모든 다른 창들도 해당 시점으로 이동한다는 것을 의미한다.

5.2 전사자 품질 관리

- ① 전사 교육을 받은 전사자에게 동일한 영상을 제시하여 시범 전사를 하게 한다.
- ② 전사 일치도를 확인하기 위해 연구진이 동일한 영상을 직접 전사하여 검토 기준을 정하고 각 전사자의 전사 결과와 비교하여 그 신뢰도를 측정한다.
- ③ 신뢰도를 측정할 때 엘란(Elan) 프로그램에 있는 전사 신뢰도 측정 기능(Calculation of Inter Annotator Reliability), 또는 신뢰도 측정을 할 수 있는 다른 프로그램을 사용한다.
- ④ 신뢰도 측정은 분절 일치도와 토큰 일치도를 바탕으로 한다.
- ⑤ 신뢰도 측정 후 신뢰도 검사에서 일치도가 낮게 나타난 토큰의 특징을 분석하여 그 결과를 토대로 추가 한 차례의 전사교육을 실시한다. 신뢰도 측정 후 신뢰도 검사에서 분절 일치도가 낮게 나왔을 경우 분절에 대한 재교육을 함께 실시한다.
- ⑥ 재교육을 실시한 후 본 전사에 들어간다.
- ⑦ 품질 관리를 위해 신뢰도 측정을 본 전사 기간에도 실시한다.

5.3. 기본 전사 규칙

5.3.1 기본 원칙

- ① 주석 파일에 글로스를 붙인다는 것은 모든 의미 있는 손동작에 이름을 붙이는 것을 말한다. 이때 이름은 우리말(한국어) 단어를 가져와 사용한다. 즉, 글로스를 붙인다는 것은 수어 형태에 한국어로 된 이름을 붙이는 것이다.
- ② 글로스 부여 원칙은 수어의 형식과 의미, 기저 이미지를 보아 결정한다.
- ③ 글로스는 수어 화자와 손마다 분리된 층렬을 할당한다. 따라서 총 4개의 층렬이 필요하다. 예) 우세 글로스A(왼쪽 언어 제공자의 우세손), 비우세 글로스A(왼쪽 언어 제공자의 비우세손), 우세 글로스B(오른쪽 언어 제공자의 우세손), 비우세 글로스B(오른쪽 언어 제공자의 비우세손)
- ④ 하나의 형태에는 항상 같은 하나의 글로스를 할당한다. 이때 글로스는 문맥과는 독립적이므로 의미의 번역과는 차이가 있다. 따라서 번역이 되지 않도록 유의하여야 한다.

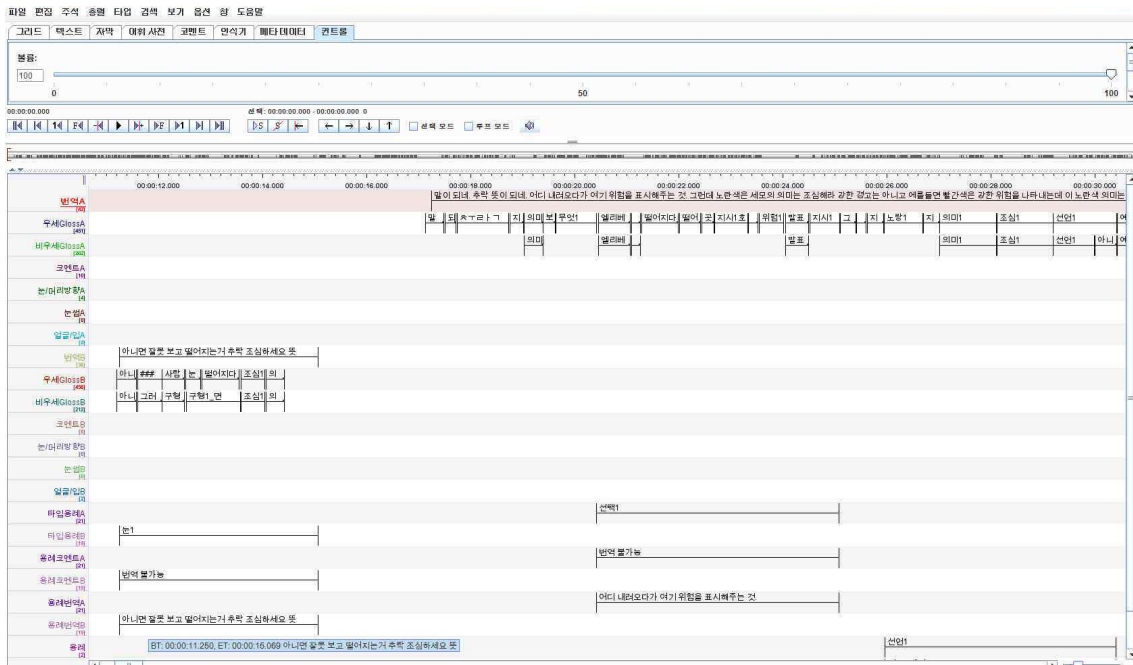
5.3.2 층렬

- ① 수어 데이터를 체계적으로 전사할 때는 전사 정보를 여러 층렬로 나누어 입력한

다.

② 기본적인 층렬은 다음과 같다.

- 번역 층렬
- 수지 글로스 층렬
- 비수지 층렬
- 코멘트 층렬



<그림 8> 기본적인 층렬

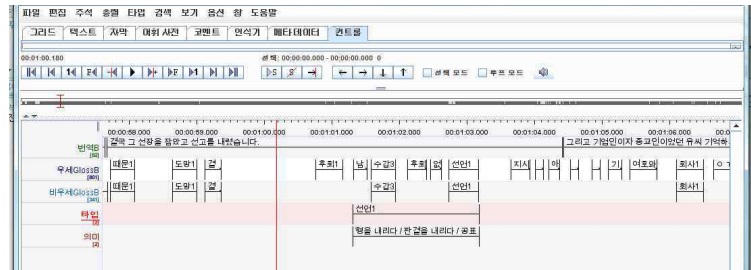
③ 수지 글로스 층렬은 우세손과 비우세손으로 나눈다.

④ 비수지 층렬은 연구 목적에 따라 다양하게 나눌 수 있지만 전사의 편리를 위해서 눈과 머리 방향, 눈썹, 얼굴과 입 등으로 나누는 것이 좋다. 눈과 머리의 방향은 주로 함께 움직이며 문법적인 기능을 나타내므로 하나의 층렬을 할당하고, 눈썹은 의문문, 화제화 등 문장의 형태를 나타내므로 또 다른 층렬을 할당한다.

⑤ 수어 화자들은 수지 없이도 비수지 기호를 사용하여 의미 전달을 한다. 따라서 비수지 기호 층렬을 따로 입력하지 않고 수지 층렬만을 입력할 때에도 다음의 경우는 입력을 해야 한다.

- 수어 없이 긍정의 의미로 고개를 끄덕인 경우 코멘트 층렬에 ‘끄덕임’으로 입력한다.
- 수어 없이 부정의 의미로 고개를 흔든 경우 ‘가로저음’으로 입력한다.

⑥ 수지 글로스 층렬 아래에 의미 층렬을 둘 수 있다. 즉, <그림 9>와 같이 타입 층렬과 의미 층렬을 만들고 타입 층렬에는 맥락에 따라 다양한 의미로 사용되는 해당 수어의 글로스를 입력하고, 의미 층렬에는 맥락에서 결정된 의미를 입력한다. 이렇게 기록된 의미 정보들은 수어 사전 제작 시에 활용할 수 있다.



<그림 9> 의미 총렬

5.3.3 수어 단어의 분절

- ① 이동이 없는 수어는 수형이 보이기 시작하는 첫 번째 프레임 이 수어의 시작이다.
- ② 이동이 있는 수어 단어의 경우는 손이 움직이기 시작하는 첫 번째 프레임이 수어의 시작이다.
- ③ 이동도 수형의 변화도 없이 이전 수어에서 새로운 수어가 나타나는 경우는 수형이 변화하기 시작하는 첫 번째 프레임이 수어의 시작이다.
- ④ 수어 단어의 끝은 수형이 변화하기 시작하는 첫 번째 프레임의 바로 이전 프레임이다.
- ⑤ 이동이 시작되는 첫 번째 프레임의 바로 이전 프레임이 이전 수어 단어의 끝이다.

5.3.4 타입

- ① 전사 시 글로스를 붙인 개별적인 수어 어휘 하나하나를 토큰이라 하고, 이러한 토큰들 중에서 같은 형태의 토큰을 묶은 것을 타입이라 하는데, 이는 실제적이고 개별적인 것이 아니 라 추상적인 개념이다.
- ② 어휘는 형태(form)와 의미(meaning)로 이루어지는데, 도상성이 강한 수어 어휘에서는 기저 이미지(underlying image)도 고려되어야 한다. 즉, 같은 형태, 같은 의미 또는 같은 그림일 경우 같은 타입으로 분류(예: [소개 <그림 10>])하였다. 형태는 같지만 기저 이미지가 다른 경우²¹⁾, 또는 핵심 의미는 같지만 기저 이미지가 다른 경우는 다른 타입(예: [먹다1 <그림 11>], [먹다2 <그림 12>], [먹다3 <그림 13>])으로 분류한다.

21) 형태는 같지만 기저 이미지가 다른 경우의 예로 독일수어 [우유]와 [경쟁]이 있다. [우유]의 기저 이미지는 소의 젖을 짜는 그림이나 [경쟁]은 두 사람이 경쟁을 하는 그림이다. 한국수어의 기 전사에서 는 이와 같은 예가 아직 발견되지 않았다.



<그림 10> [소개]



<그림 11> [먹다1]



<그림 12> [먹다2]



<그림 13> [먹다3]

- ③ 기저 이미지(underlying image)란 수어 화자들의 머릿속에 그려지는 공통적인 이미지로 도상성이 강한 수어 어휘 기저에 자리하고 있는 이미지이다. 그 이미지는 수어의 의미 한 부분을 직접적으로 표현(예: [궁궐 <그림 14>])하기도 하고, 수어의 의미와 간접적으로 연관(예: [중국 <그림 15>])되기도 한다(Langer, G.König, S.Konrad, R.2007). 즉, [궁궐]이란 수어는 궁궐 지붕의 모양을 직접적으로 표현하고 있으며, [중국]이라는 수어는 중국을 대표할 수 있는 전통적인 옷 모양을 나타낸 것이다.



<그림 14> [궁궐]



<그림 15> [중국]

- ④ 수어 어휘를 나타내는 타입은 기본형, 변이, 형태론적인 수정으로 나눌 수 있다.
- 기본형은 관련 수어들 중에서 제일 단순하고 가장 흔히 사용되는 수어이며 맥락 없이도 독립적으로 사용될 수 있는 형태이다 (Mandel,1997).
 - 형태론적인 수정은 수어의 형태를 확장하거나 정보를 추가하기 위해 수어의 한 형태가 바뀐 것으로, 도상성이 강한 수어에서 나타난다. 이는 하나의 기저 이미지 형태에 의미를 확장하거나 상세화하기 위하여 형태가 수정되

는 경우를 말한다. 예) [공 <그림 16>], [큰 공 <그림 17>]



<그림 16> [공]



<그림 17> [큰공]

- 변이형(variant)은 형태가 아주 작고 미세하게 다르나 기저 이미지와 의미가 같은 수어로, 서로 바꾸어 사용하여도 무방한 것이다. Langer, G./König, S./Konrad, R. (2007)에 의하면 수어의 의미에서는 차이가 없으나 형태의 관습적인 부분에 작은 변화가 있을 때를 변이로 본다고 하였다. 예) [없다1], [없다1A] / [사다1], [사다1A], [사다1B] / [후배1], [후배1A]등이 있다. 그런데 여기서 나타나는 형태의 변동은 기저 이미지로 설명되지 않는다는 것이 중요하다. 그러나 기저 이미지가 다른 경우는 별개의 타입으로 보아야 한다. 즉, 변이는 미묘하게 바뀌는 형태들이 똑같은 그림을 바탕으로 하고 있으며, 이것은 주로 수형에서 흔하게 나타난다.

예) 시골



<그림 18> [시골]



<그림 19> [시골]

- ⑤ 같은 글로스를 사용하는 다른 타입의 수어일 경우는 ID 글로스 바로 뒤에 숫자 1, 2, 3, 등으로 나타낸다. ID 글로스 뒤에 하나의 숫자만 붙어있는 경우 해당 타입의 기본형을 의미한다.

예) 없다1 <그림 20>, 없다2 <그림 21>, 없다3 <그림 22>



<그림 20> [없다1]



<그림 21> [없다2]



<그림 22> [없다3]

- ⑥ 형태론적인 수정은 해당 타입의 ID 글로스 다음에 _1, _2, _3 등으로 나타낸다.
변이형은 해당 타입의 ID 글로스 다음에 알파벳 A, B, C 등으로 나타낸다.
- ⑦ 연구 목적에 따라 형태론적인 수정과 변이를 다 나타내지 않고 기본형으로 전사할 수 있다.
- ⑧ ID 글로스의 구조는 <그림 23>와 같다.



<그림 23> ID 글로스의 구조

- ⑨ 한국어로는 두 개의 단어이지만 수어로는 하나의 수어인 경우는 가장 중립적이고 가장 쉽게 떠올릴 수 있는 한국어 단어로 글로스를 부여한다.
예) 도움을 주다, 도움을 받다 → [돕다]

예) 주다, 받다 → [주다]

5.3.5 글로스(ID glosses)

1) 글로스(ID glosses) 확정 방법

- ① 글로스란 수어에 붙이는 이름(label)으로, 숫자나 알파벳 등 다양하게 붙일 수도 있으나 그 이름에 해당하는 수어가 머릿속에 떠오르기 쉽게 붙이는 것이 좋다. 따라서 여기서는 한국어를 사용하여 이름을 붙인다. 그러나 여기에 붙인 한국어 단어는 의미의 번역과는 다른 것이다.
- ② 글로스는 표준국어대사전에 등재된 어휘를 기본으로 하되 수어 화자들의 언어 의식을 반영하여 익숙하고 알기 쉬운 표현으로 선택한다.
예) 야금야금하다 → [야금야금쓰다]
예) 시시티브이 → [CCTV]
- ③ 글로스 선택 시 하나의 수어가 여러 개의 한국어 단어로 번역이 될 때는 가능한 많은 사람이 가장 쉽게 해당 수어를 떠올릴 수 있는 한국어 단어를 선택한다.
예) 꼼꼼하다, 자세하다, 구체적 → [꼼꼼하다]



<그림 24> [꼼꼼하다]

- ④ 문법적 표시는 가장 중립적인 것, 즉 무표적인 것으로 한다. 즉, 동사나 형용사의 경우 기본형으로 글로스를 붙인다.
예) 잡다, 잡는다, 잡았다, 잡겠다 → [잡다]
예) 예쁜, 예쁘고, 예뻐다 → [예쁘다]
- ⑤ 글로스는 하나의 단어로 하는 것을 원칙으로 한다. 단 하나의 단어로 의미가 불충분한 경우 [버스가다]처럼 띄어쓰기 없이 붙여 쓴다. <그림 25>는 [버스], <그림 26>는 [버스가다]로 글로싱한 수어이다.



<그림 25> [버스]



<그림 26> [버스가다]

- ⑥ 수어 동의어(같은 의미이지만 형태가 다른 수어)로 글로스가 같은 경우는 글로스 뒤에 1, 2와 같은 숫자를 표시하여 [없다1<그림 27>], [없다2 <그림 28>], [없다3 <그림 29>]과 같이 글로스를 붙인다.



<그림 27> [없다1]



<그림 28> [없다2]



<그림 29> [없다3]



- ⑦ 글로스로 사용한 한국어 단어가 동음이의어인 경우는 타입의 혼동을 방지하기 위해 다음과 같이 구분하여 글로스를 부여한다.

- 글로스에 의미를 추가한다. 이때 의미를 추가하여 구 구성이 되는 경우에는 한국어 어순에 맞도록 한다.

예) [걷다](표준국어대사전 걷다02)와 [웃걷다](표준국어대사전 걷다03)



<그림 30> [걷다]



<그림 31> [웃걷다]

- 비슷한 의미를 지닌 다른 글로스로 교체한다.

예) [예](윗사람의 부름에 대답하거나 묻는 말에 긍정하여 대답할 때 쓰는 말)와 [보기](어떤 사실을 설명하거나 증명하기 위하여 내세워 보이는 대표적인 것)



<그림 32> [예]



<그림 33> [보기]

- 글로스의 품사를 교체한다.

예) [일](자연수의 맨 처음 수)과 [일하다](무엇을 이루거나 적절한 대가를 받기 위하여 어떤 장소에서 일정한 시간 동안 몸을 움직이거나 머리를 쓰다)



<그림 34> [일]



<그림 35> [일하다]

2) 수와 단위

- ① 수를 나타내는 수어는 한국어 단어로 전사한다. 즉 1은 [일1]로, 2는 [이1], 10은 <그림 36>과 같이 [십1] 또는 [십2] 등으로 글로스를 붙인다.



<그림 36> [십1], [십2]

② 첫 번째, 두 번째와 같은 서수는 [첫번째1], [두번째1] 등으로 글로스를 붙인다.

③ 1년, 2분과 같이 수가 수어에 혼입되는 경우는 다음과 같이 글로스를 붙인다.

예) 1년 → [일년1] <그림 37>

예) 2분 → [이분1] <그림 38>



<그림 37> [일년1]



<그림 38> [이분1]

④ 비우세손에 나타나는 ‘counting hand(손으로 세기)’는 펼쳐진 손가락의 수와 상관없이 모두 [counting hand <그림 39>]로 전사하고, 비우세손의 ‘counting hand’와 함께 나타나는 우세손의 수어는 해당 수어로 글로스를 붙인다.



<그림 39> [counting hand]

⑤ 알파벳 지문자로 나타내는 단위 수어는 한국어로 글로스를 붙인다.

예) ‘mm’ → [밀리미터1]

3) 명사와 동사

- ① 명사와 동사의 형태가 같은 수어는 이동이 없는 경우는 명사로 글로스하고, 이동이 있는 경우는 동사로 글로스를 붙인다. 예를 들어 <그림 40>처럼 동일한 수어 형태에서 움직임이 없으면 [버스1]로, 움직임이 있는 경우는 [버스가다1]로 글로스를 붙인다.



<그림 40> [버스1], [버스가다1]

- ② 일치동사 중 <그림 41>의 ‘야단치다’처럼 비우세손이 나타나지 않아도 되는 경우에는 다음과 같이 우세손과 비우세손 글로스를 따로 붙인다.

예) 우세 글로스: [야단치다1]

비우세 글로스: [사람4]



<그림 41> [야단치다]

- * 일치동사란 [돕다], [가르치다]라는 수어처럼 동사 속에 행위자(→행위 주체)와 행위(→대상)를 받는 자가 포함되어 나타나는 동사이다.

4) 합성어

- ① 둘 이상의 수어가 결합된 형태이나 명확하게 하나의 단어로 인지되는 경우에는 하나의 글로스로 주석한다. 예를 들어 <그림 42>의 ‘손+돈’으로 된 수어는 두 수어를 묶어 분절한 후에 [수당1]로 글로스를 붙인다.



<그림 42> [수당1]

- ② 둘 이상의 수어가 결합된 상태에서 실제로 하나의 단어로 합성된 상태인지 의심스러운 경우는 두 개의 수어로 분리하여 각각 글로스를 붙인다. 예를 들어 <그림 43>의 분홍이라는 수어는 [빨강1], [약하다1]로 글로스를 붙인다.



<그림 43> [빨강1], [약하다1]

- ③ 둘 이상의 수어의 연쇄가 고정되어 하나의 합성어라고 인정하는 기준은 다음과 같다.

- 두 수어의 조합이 특정한 의미를 갖는다.
- 조합된 수어의 의미가 예측되지 않을 수 있다.
- 조합된 두 수어 사이에 다른 수어나 신호를 삽입할 수 없다.
- 두 수어의 결합 시에 나타난다고 증명된 음운론적 특성이 적용된다.

- ④ Valli & Lucas(2000)가 제시한 합성어에서 나타나는 특성은 다음과 같다.

- 첫 번째 수어의 마지막과 다음 수어의 처음 사이에 움직임(M)이 들어간다. 그런데 여기서 주의할 점은 앞 수어의 마지막과 뒤 수어의 처음 사이에 첨가된 움직임(M)은 다음의 움직임에 동화가 된다는 것이다.

예를 들어 우리 자료의 [오가다 <그림 45>]라는 수어에서도 [오다]와 [가다 <그림 44>]라는 두 수어가 결합될 때 두 수어를 연결하는 움직임(M)이 첨가되면서 수어 화자들은 하나의 단어로 인식하였다.



<그림 44> [오다], [가다]



<그림 45> [오가다]

- 두 단어가 결합할 때 움직임 사이의 접촉이 없는 정지(H)동작이 생략된다. 예를 들어 [안산]이라는 지명의 수어는 [편안하다]라는 수어와 [산]이라는 수어가 결합하여 이루어진 복합어인데, [편안하다]에서 마지막 정지 동작(H)이 탈락하면서 [산]이라는 수어의 첫 움직임(M)과 [편안하다]라는 수어의 마지막 사이에 다른 움직임(M)이 첨가되었다(<그림 46>, <그림 47> 참조).



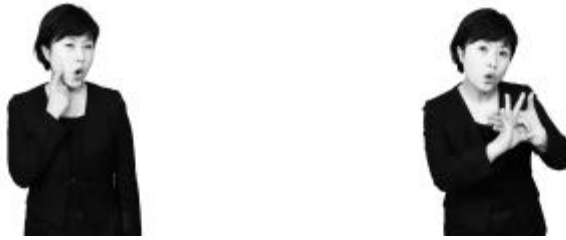
<그림 46> [편안하다], [산]



<그림 47> [안산]

- 동화 현상이 나타난다. 즉, 어떤 한 분절이 가까이에 있는 다른 분절의 특

성을 갖게 되는 현상이 나타난다. 예를 들어 [만약]이라는 수어는 [거짓]+[면]이라는 수어로 이루어진 합성어인데, [거짓]이라는 수어를 할 때 뒤의 [-면]이라는 수어의 수형에 동화되어 1지만 편 수형이 1지와 5지로 동그라미를 만들고, 2, 3, 4지를 편 [-면]이라는 수어의 수형과 같아지므로 역행 동화 현상이 일어난다. (<그림 48>, <그림 49> 참조)



<그림 48> [거짓], [-면]



<그림 49> [만약]











5) 지시(가리키기)



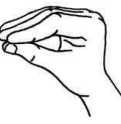





- ① 검지만 편 수형(1형)인 경우 [지시1]로 글로스를 붙인다.
- ② 손가락을 다 편 수형(9형)인 경우 [지시2]로 글로스를 붙인다.
- ③ 자신을 지칭하는 경우는 수형에 따라 [지시1] 또는 [지시2] 뒤에 [-1], 타인을 지칭하는 경우는 [-비1]를 붙여 글로스를 붙인다. 수어에서는 2인칭과 3인칭을 나타낼 때 손의 지시에서는 차이가 없고 눈 응시(eye gaze)와 맥락으로 나타내므로, 손의 글로스에서는 2인칭과 3인칭을 구분하여 글로스를 붙이지 않고 한 타입으로 보아 [-비1]로 글로스를 붙인다.
- ④ [지시1] 또한 [지시2]가 움직임이 있을 경우 글로스에 [호]를 붙인다. 움직임은 주로 반원으로 나타난다.
- ⑤ [지시1] 또한 [지시2]가 선을 따라 그리는 경우 아니면 하나의 모양/형태를 손끝으로 표현하는 경우 글로스에 [그림]을 붙인다.
- ⑥ 지시는 총 8개의 타입으로 나뉜다.
[지시1_1], [지시1_비1], [지시1호], [지시1그림]
[지시2_1], [지시2_비1], [지시2호], [지시2그림]
- ⑦ 한국수어에서 신체 부위를 의미하는 수어는 그 부위를 지시하며 나타난다. 따라

서 수어 화자가 팔을 가리키며 ‘팔’이라는 수어를 사용할 경우 글로스 규칙을 적용하여 [팔1]로 글로스를 붙인다. 신체부위를 가리키는 수어를 [지시1_비1] 아니면 [지시2_비1]로 전사하지 않는다.

6) 생산적 수어(productive sign)

생산적 수어는 관습적인 의미를 가지지 않고 맥락 속에서 자연적으로 만들어진 수어로 맥락 안에서 이해되는 어휘이다. 그러므로 한 단어로 번역 또는 한 단어로 글로스를 정하기 어렵다. 그러나 이 수어들은 능숙한 수어 화자들 사이에서 이해가 가능하다. 그리고 이 수어들은 단순히 사물을 수형이나 수동으로 나타내는 것이 아니라, 생산적 수어에서 나타나는 체계적인 규칙이 적용된다. 이에 전사 시에는 수형에 기준을 두어 글로스를 정한다. 글로스의 예는 다음과 같다.

	타입명	수형	설명	예시
1	일형1		1지만 편 형태로 얇고 긴 물체 또는 그림을 나타냄	
2	육형1		1지와 5지의 손가락이 퍼진 상태로 하나의 형태를 표현	
3	육형2		1지와 5지의 손가락이 퍼져 있으나 마디가 구부러진 상태로 수어를 시작하는 것	
4	구형1		1, 2, 3, 4, 5지를 모두 펴고 손가락 사이의 벌어짐이 없는 손으로 납작하고 평평한 면 아니면 구획을 나누어 줄 때 선의 의미를 나타내는 것	
5	구형1A		1, 2, 3, 4, 5지를 모두 펴고 손가락 사이가 벌어진 손으로 납작하고 평평하며 아주 넓은 면 아니면 여러 개의 선을 나타내는 것	

6	구형2		1, 2, 3, 4, 5지를 모두 편 손이지만 마디가 구부러진 형태	
7	구형3		5지와 1, 2, 3, 4지 손가락의 접촉이 있으며 손가락의 마디가 구부러지지 않는 것	
8	구형4		5지와 나머지 손가락 끝의 접촉이 있고 '0'형으로 보이는 것	
9	십형1		1지와 5지의 손가락 끝이 동그랗게 붙어 있으며 나머지 세 손가락은 모두 펴져 있는 것	

7) 토큰으로 처리하지 않아야 할 수어들(Non-tokens)

수어 말뭉치 자료에서는 실제 사람들의 발화 형태를 반영하는 것이 아닌 형태들이 포함될 수 있다(Gabriele Langer 외, 2016). 즉, 대화 대상자의 발화를 그대로 반복하거나 따라하는 것과 같은 것은 실제 수어의 사용 빈도에 포함시켜서는 안 된다. 따라서 이러한 수어는 토큰으로 처리하지 않고 이를 비토큰(non-token)이라 한다.

- ① 메타 언어적 언급은 토큰으로 처리하지 않는다.
- ② 손으로는 수어를 하면서 비수지로 부정을 표현하는 수어는 토큰으로 처리하지 않는다. 예를 들면 어떤 ‘특정 수어’를 하면서 이렇게 수어를 하지 않는다는 의미로 머리를 흔들면 이때의 ‘특정 수어’는 토큰이 아니다.
- ③ 대화에서 이미 앞에서 발화한 수어를 그대로 모방해서 반복하거나 유도 자료에서 제시된 수어를 따라 한 것은 토큰으로 처리하지 않는다. 이는 이미 앞에서 서술된 수어를 반복한 것이기 때문에 사용 빈도에 포함되어서는 안 되기 때문이다.

8) 기타

- ① 지문자에 글로스를 붙일 때는 자음과 모음을 분리하여 글로스를 붙인다. 예) ‘학교’ → [ㅎㅏ ㄱㅑ ㄱㅑ ㅓ]
- ② 이름 수어인 경우는 대상자의 성과 이름 앞에 [SN]을 붙인다. 수어 이름이지만 그 사람의 한국어 이름을 모르는 경우는 [SN]으로만 글로스를 붙인다.

- ③ 수어가 앞뒤 수어의 영향을 받아 음운 변동이 나타난 경우에는 수어의 원형에 해당하는 글로스를 달고 코멘트 층렬에 ‘음운변동’으로 주석을 단다.
 - ④ 수어가 완전히 조음되지 않은 경우는 글로스 대신에 [%]로 주석을 입력하여 거짓 시작임을 나타낸다.
 - ⑤ 전사자가 수어로 인지했지만 무슨 수어인지 모르는 경우는 글로스 대신에 [###]로 주석을 단다. 그 후에 온라인 문의공간에 문의를 하거나 전사 회의에 참석하여 의문을 해결하게 되면 ID 글로스로 바꿔 입력한다. 전사자들이 모두 모여 논의를 한 후에도 어떤 수어인지 명확하지 않을 경우에는 [###]를 그대로 둔다.
 - ⑥ 수어가 영상 화면 바깥에서 조음이 된다는지 하여 보이지 않는 경우에도 [###]로 입력한다.
 - ⑦ 수어 화자가 실수로 수어를 잘못 표현했고, 그 수어가 사용되지 않는 수어일 때 [@]로 입력한다.
 - ⑧ 코멘트 층렬에는 언어학적 유형 등 자유로운 텍스트를 입력할 수 있다. 그러나 코멘트 층렬은 너무 길게 입력하지 않는 것이 좋다.
 - ⑨ 수어 화자가 미국수어를 하는 경우 [미국수화]로 입력하고, 국제수어를 할 경우 [국제수화]로 입력한다. 그리고 그 수어의 의미는 코멘트 층렬에 입력한다.
 - ⑩ 제스처
- 가. 다음과 같은 기준에 의해 수어와 제스처를 구분한다.
- 수어가 아니다.
 - 수어를 모르는 사람이 봐도 그 의미를 이해할 수 있다.
 - 수어를 모르는 청인도 사용하는 몸짓이다.
- 나. 제스처는 수지 글로스 층렬에 [제스처]로 입력하고, 코멘트 층렬에 구체적인 내용을 설명한다.

5.3.6 번역

전사 자료를 공개할 경우, 수어를 모르는 사람들이 자료에 접근할 수 있도록 번역을 제공한다. 번역은 목적에 따라 형태를 달리하는데 네덜란드 수어 말뭉치(NGT) 주석 규약(Onno Crash -born 외, 2015)에 따르면 같은 층렬 세트에 대한 자유 번역과 정밀(또는 문자적) 번역을 서로 구별한다. 즉, 목적에 따라 자유 번역과 정밀 번역으로 번역을 제공할 수 있다. 본 말뭉치에서 채택하고 있는 번역 방식은 자유 번역이다. 다음은 두 번역 형태의 특징이다.

① 자유 번역(Free Translation)

- 목표 언어인 한국어에서 자연스럽게 실행되는 텍스트를 선호한다.
- 음성 언어 번역에 의해 부분적으로 결정되므로 문장의 길이가 상대적으로

길다.

- 목표 언어에 적합한 지시 표현을 사용한다.
- 언어학자 또는 다른 사용자들이 수어 담화의 내용에 대한 이해를 하는 데에 도움이 된다.

② 정밀 번역(Narrow Translation)

- 비교적 짧은 절 또는 문장의 부분에서 수어 텍스트에 가까운 형태이다.
- 목표어인 한국어로 매끄럽게 번역할 필요는 없다.
- 수어 담화의 구조를 이해하는 데 도움을 주고, 기계 번역 알고리즘을 학습하고 테스트하는 데 유용하다.

5.3.7. 정밀 전사 규칙

2017년도 연구 결과 제작된 2017년도 한국수어 말뭉치 구축 지침 중 입 움직임에 대한 정밀 전사 규칙을 다음과 같이 제시한다.

1) 입 움직임 전사의 기본 원칙

입의 움직임은 2개의 범주 즉, 마우딩(Mouthings)과 마우스 제스처(Mouth gestures)로 나눌 수 있다. 마우딩은 구어를 조음할 때처럼 입의 모양을 만드는 것을 말하며, 마우스 제스처는 구어의 단어를 조음하는 것 외의 모든 입의 움직임을 의미한다. 마우딩은 구어 단어 전체 또는 한 부분을 조음하는 것처럼 보이며, 때로는 하나의 수어 단어 이상에서 지속되기도 한다. 그러나 마우스 제스처는 구어 단어의 조음과는 관련이 없는 입의 움직임이므로 음성 언어로부터 온 것이 아니며, 마우딩은 음성 언어로부터 왔거나 빌려온 것이다(Crasborn, 2016).

입 움직임 전사의 기본 원칙은 다음과 같다.

- ① 마우스 제스처는 타입을 추가하며, 입 모양은 별도의 타입이 없다.
- ② 마우스 제스처와 관련 전사는 2개의 층렬로, 마우스 제스처 글로스와 마우스 제스처 분류(class) 층렬로 나누어 기입한다.
- ③ 마우딩 관련 전사는 4개의 층렬로, 마우딩을 기입하는 층렬 외에 스프레딩, 마우딩의 완성도, 마우딩의 분류 층렬로 나누어 기입한다.
- ④ 입의 움직임을 동반한 수지 글로스의 분류를 기입하는 층렬과 입의 움직임과 관련된 코멘트를 기입할 층렬을 만든다. 그리하여 각 언어 제공자별로 8개의 층렬을 만들어 전사한다.
- ⑤ 2015년에 진행된 한국수어 말뭉치 기반 조사 및 시범 구축 사업에서 번역과 수지 글로스 전사가 된 파일들로 입 움직임에 대한 전사를 시작할 때는 템플릿을 사용하면 일단 2명의 언어 제공자에 대한 전사를 위한 층렬이 각각 8개씩 총 16개의 층렬이 생성된다. 따라서 입 움직임에 관련된 전사 층렬은 이 8개의 층렬

다음에 추가한다.

2) 총렬

각 언어 제공자에 대한 기본 총렬인 ‘번역’, ‘우세 글로스’, ‘눈/머리 방향’, ‘눈썹’, ‘얼굴/입’, ‘코멘트(comment)’ 총렬 다음에 입 움직임과 관련한 8개의 총렬을 추가한다. 추가한 총렬은 마우딩(Mouthing) 관련 총렬로 ‘마우딩’, ‘스프레딩(Spreading)’, ‘마우딩 완성도’, ‘마우딩 분류’가 있으며, 마우스 제스처(Mouth Gesture) 관련 총렬로 ‘마우스 제스처’, ‘마우스 제스처 분류’ 총렬이 있다. 그 외에 ‘수어 분류’ 총렬과 ‘M 코멘트’ 총렬 등 8개의 총렬이 2명의 언어 제공자에게 각각 할당되어 총 16개의 총렬을 추가한다. 총렬의 이름은 띄어쓰기 없이 입력하며 추가된 총렬의 속성은 다음과 같다.

- ① 마우딩(Mouthing): 한국어 단어를 조음하는 것처럼 보이는 입의 움직임을 해당 단어의 철자법에 따라 한글로 입력하는 총렬이다.
- ② 스프레딩(Spreading): 마우딩이 하나의 수어를 넘어서 얼마나 지속되는지를 입력하는 총렬이다.
- ③ 마우딩 완성도: 마우딩으로 나타난 구어 단어의 완성도를 입력하는 총렬이다.
- ④ 마우딩 분류: 마우딩을 특성에 따라 분류하고 이를 입력하는 총렬이다.
- ⑤ 마우스 제스처(MouthGesture): 구어 단어를 발음하는 것 외의 모든 입의 움직임에 대한 글로스를 입력하는 총렬이다.
- ⑥ 마우스 제스처 분류: 마우스 제스처를 특성에 따라 분류하고 이를 입력하는 총렬이다.
- ⑦ 수어 분류(Class) : 마우딩 또는 마우스 제스처와 함께 나타난 수어의 품사 중에 명사와 동사를 구분하여 입력하는 총렬이다.
- ⑧ M 코멘트: 입의 움직임과 관련한 전사 시에 생성된 총렬에 담을 수 없는 다양한 의견 등을 자유롭게 입력하는 총렬이다.

3) 분절

- ① 수어가 시작되고 끝나는 지점과 상관없이 마우딩과 마우스 제스처가 시작하고 끝나는 지점을 분절의 기준으로 삼는다.
- ② 마우딩 또는 마우스 제스처의 마지막 모양이 의미 없이 오랜 시간 유지되는 경우는 분절에 포함하지 않는다.
- ③ 마우딩 또는 마우스 제스처 관련 속성의 총렬은 해당 마우딩 또는 마우스 제스처의 분절 프레임과 동일하게 분절한다. 즉, 스프레딩, 마우딩 완성도, 마우딩 분류 총렬은 마우딩 총렬 분절과 마우스 제스처 분류는 마우스 제스처 총렬의

분절과 동일하게 분절하다. M코멘트 층렬 역시 해당 마우딩 또는 마우스 제스처 분절과 동일하게 분절한다.

4) 마우딩 관련 층렬의 전사

(1) 마우딩

- ① 구어의 단어를 조음하는 대로 입력하되, 조음점이 보이지 않는 경우는 마우딩을 하면서 하고 있는 수어를 참조해서 한글 맞춤법에 맞추어 입력한다. 예를 들어 ‘고기’와 ‘공기’, ‘오이’의 입 모양은 구분할 수가 없으나, 해당 분절의 관련 수어가 [고기]일 경우는 [고기]로 관련 수어가 [공기]일 경우는 [공기]로 입력한다.
- ② 마우딩이 단어의 일부분만 나타난 경우는 입 모양으로 표현된 부분을 입력하고 추론이 가능한 나머지 부분은 괄호를 하고 입력한다. 예를 들어 수어 [조심]을 하면서 입 모양은 ‘조’만 나타난 경우에 [조(심)]으로 주석을 입력한다.
- ③ 수어와 입 모양이 다른 경우에는 입 모양이 나타난 그대로 주석을 입력한다. 예를 들어 [일본]이라는 수어를 하면서 입의 움직임은 ‘일반’으로 나타난 경우, [일반]으로 입력하고 M 코멘트 층렬에 [일본]이라고 입력한다.

(2) 스프레딩(Spreading)

- ① 마우딩은 그 지속 시간이 길어 하나 이상의 수어에서 나타나기도 한다. 이러한 현상을 스프레딩이라고 하는데, 이 층렬은 이러한 스프레딩 현상을 확인하기 위한 층렬이다.
- ② 마우딩이 지속되는 시간을 분절한 후에 다음의 특성에 따라 오른쪽, 왼쪽, 양쪽, 일치, 부분 등으로 입력한다.
 - 오른쪽: 마우딩의 시작점이 의미와 관련된 수어의 시작점과 같으나, 아무 관련이 없는 다음 수어까지 지속되는 경우이다. 즉, 마우딩의 지속 시간이 해당 수어의 지속 시간보다 길어 다음 수어까지 덮는다. 예를 들어 [학교]라는 수어를 하면서 마우딩으로 ‘학교’를 하는데, 수어 [학교] 다음 수어인 [지시] 수어를 할 때까지 ‘학교’의 조음이 지속되는 경우에 [오른쪽]이라고 입력한다.
 - 왼쪽: 마우딩이 의미와 관련된 해당 수어보다 먼저 나타나는 경우이다. 예를 들어, [장소]라는 수어와 함께 마우딩 ‘학교’가 시작되고 이 입 모양이 수어 [학교]가 나타날 때까지 지속되는 경우에 [왼쪽]이라고 입력한다.
 - 양쪽: 마우딩이 의미와 연관된 해당 수어보다 먼저 나타나서 다음 수어를 할 때까지 지속되어 늦게 끝나는 경우이다. 예를 들어 [장소] [학교] [지시] 순으로 수어를 하는데, [장소]라는 수어를 하면서 ‘학교’라는 마우딩이 시작되어 [지시]수어를 할 때까지 지속되는 경우, [양쪽]이라고 입력한다.

- 일치: 마우딩의 시작점과 끝점이 의미와 연관된 해당 수어의 시작점과 끝점과 같은 경우에는 [일치]라고 입력한다.
- 부분: 마우딩이 의미와 연관된 해당 수어와 함께 나타나지만 지속되는 구간이 수어보다 짧은 경우이다. 예를 들어 ‘학교’ 마우딩이 [학교]수어와 함께 나타나긴 하지만 수어가 끝나기 이전에 입 모양이 먼저 끝나는 경우에는 [부분]이라고 입력한다.
- ③ 스프레딩 현상은 오른쪽, 왼쪽, 양쪽으로 나타날 수 있다. 이러한 스프레딩을 결정하는 시간 기준은 ‘스프레딩이 되었다고 보이는 부분이 10프레임 이상일 때’ 또는 ‘스프레딩이 앞 혹은 뒤 수어의 50% 이상을 덮을 때’이다. 즉, 오른쪽 스프레딩은 마우딩이 해당 수어와 함께 시작되어 다음 수어의 10 프레임 이상 혹은 다음 수어 지속 시간의 50% 이상까지 지속될 때이다.

(3) 마우딩 완성도

- ① 마우딩은 해당 구어 단어의 일부분만 나타나기도 한다. 이 총렬은 마우딩이 해당 단어의 모든 음절이 나타났는지 부분적으로 나타났는지에 대한 완성도를 확인하는 총렬이다.
- ② 마우딩 완성도는 다음의 특성에 따라 완, 앞, 뒤, 앞뒤 등으로 입력한다.
 - 완: 구어 단어의 모든 음절을 조음한 경우이다. 예를 들어 ‘어린이’ ‘조심’ 등 모든 음절이 나타난 경우 [완]으로 입력한다.
 - 앞: 구어 단어의 첫 음절만 조음한 경우이다. 예를 들어 ‘조심’에서 ‘조’ 음절만 나타난 경우 [앞]으로 입력한다.
 - 뒤: 구어 단어의 마지막 음절만 조음한 경우이다. 예를 들어 ‘조심’에서 ‘심’ 음절만 나타난 경우 [뒤]로 입력한다.
 - 앞뒤: 구어 단어의 첫 음절과 마지막 음절만 조음한 경우이다. 예를 들어 ‘자동차’에서 ‘동’ 음절은 탈락되고 ‘자차’라는 음절만 나타난 경우 [앞뒤]로 입력한다.

(4) 마우딩 분류



- ① 마우딩의 기능 및 특성에 따른 분류를 입력하는 총렬이다.
- ② 마우딩의 분류는 다음의 특성에 따라 M, M-add, M-spec로 입력한다.
 - M: 일반적인 마우딩으로 [학교]라는 수어를 하면서 ‘학교’라는 마우딩이 나타나는 경우이다.
 - M-add: 수어에 해당하는 구어 단어의 마우딩이 나타나는 것이 아니라 다른 마우딩이 나타나면서 그 마우딩이 부가 정보를 제공해 주는 경우이다. 예를 들어 [학교]라는 수어를 하면서 마우딩은 ‘맞다’라고 하므로 “학교가 맞다.”라는 정보를 제공한다.




- M-spec: 수어와 함께 나타나는 마우딩이 해당 수어의 의미를 구체화시키는 경우이다. 예를 들어 [새]라는 수어를 하면서 마우딩은 ‘비둘기’라고 하므로 [새]라는 수어의 의미를 구체화시킨다.
- M-solo: 수어 없이 마우딩만 나타나는 경우이다.

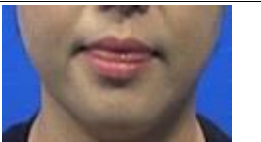
5) 마우스 제스처 관련 층렬

(1) 마우스 제스처

- ① 마우딩 이외의 모든 입의 움직임을 특정 기호를 사용하여 입력한다.
- ② 입의 움직임을 나타내는 기호는 컴퓨터 자판에서 선택이 가능한 기호 중 최대한 도상성을 살린 기호로 정하였다. 가장 기본적인 기호는 다음과 같다.

연 번	입 모양	기호	사진
1	입 모양에 관계없이 가장 작게 벌림	O1	
2	입 모양에 관계없이 중간 정도로 벌림	O2	
3	입 모양에 관계없이 가장 크게 벌림	O3	
4	입을 모으면서 입술을 내뺌	B	
5	입이 좌우로 벌어지고 이가 보임	H1	
6	입이 좌우로 크게 벌어지고 이가 보임	H2	
7	입을 다물	-1	
8	입을 더 세게 다물	-2	

9	양쪽 입 꼬리가 약간 내려감	^1	
10	양쪽 입 꼬리가 많이 내려감	^2	
11	왼쪽 입 꼬리만 올라감	<	
12	오른쪽 입 꼬리만 올라감	>	
13	왼쪽 볼에 바람을 넣음	(-	
14	오른쪽 볼에 바람을 넣음	-)	
15	양 볼에 바람을 넣음	(-)	
16	양 볼에 바람을 뺌)-(
17	볼 안에서 혀로 볼을 밀어냄	o>	
18	아랫니만 보임	M	

19	윗니만 보임	W	
20	지속 또는 반복적으로 움직임	~	
21	양악을 엇갈리게 움직임	Z	
22	양 볼의 바람을 내뿜음	(=)	
23	입 안쪽에서 혀를 약간 다물	C	
24	내민 혀를 입 안으로 집어넣음	UJ	

- ③ 입의 모양은 기본 기호를 조합하여 입력한다. 예를 들어 ‘입술을 내밀고 있는데 이가 보이는 경우’는 [BH]로, ‘입을 다 물고 양쪽 입꼬리가 약간 내려간 경우’는 [-^1]로, ‘이가 보이는 상태로 왼쪽 입꼬리만 올라 간 경우’는 [<H]로, ‘입이 약간 벌어지고 혀가 조금 나온 경우’는 [O1U1]로 ‘혀를 여러 번 내미는 경우’는 [U1~]로 입력한다.
- ④ 연속적으로 일어나는 입의 움직임은 움직임의 순서대로 기호를 조합하여 입력한다. 예를 들어 ‘입을 약간 벌리고 혀를 내밀었다가 집어넣으면서 바람을 내뿜는 경우’는 [O1UJ(=)]로, ‘입을 다물었다가 이가 보이게 양쪽으로 벌리면서 바람을 내뿜는 경우’는 [-H1(=)]로 입력한다.

(2) 마우스 제스처 분류

- ① 마우스 제스처의 특성에 따른 분류를 입력하는 총렬이다.
- ② 마우스 제스처 분류는 다음의 특성에 따라 A, E, 4, W로 입력한다.
- A: 입의 움직임이 손으로 하는 수어의 의미를 꾸미거나 그 의미 정보에 대한 부가적인 정보를 추가해 주므로 문장 안에서 부사어(Adverbial)와 같은 역할을 하는 경우이다. 예를 들어 [바람/불다]라는 수어를 하면서 입으로 내는 바람에 의해 바람의 정도를 표현한다.
 - E: 손의 움직임과 함께 조음소로 작용하는 입의 움직임으로 이 입의 움직임만으

로는 추가적이거나 독립적인 의미는 없는(Empty) 경우이다. 특히 이 입의 움직임이 없는 경우 농인들은 그 수어가 자연스럽지 않다고 느낀다. 예를 들면 [알다]라는 수어를 하면서 혀를 움직이는 동작, [있다]라는 수어를 하면서 입을 벌리는 것 등이 이 분류에 속한다. 또한 [끝]이라는 수어를 할 때 우세손 끝이 비우세손바닥에 닿는 순간 ‘혀끝을 살짝 이로 무는’ 입의 움직임처럼 손의 움직임을 그대로 복사(Eco)를 하는 경우도 이 분류에 속한다. 이 분류에 속하는 입의 움직임은 [가능하다]라는 수어에서처럼 손의 움직임은 없이 ‘입을 닫았다가 열면서 바람을 순식간에 내는(파)는 입의 움직임만으로도 의미를 나타내기도 한다.

- 4: 입이 실제 입의 역할을 하는 경우이다. 예를 들면 [먹다] [씹다] [외치다]와 같은 수어를 하면서 실제 먹거나 씹거나 외치는 입의 모양을 하는 것이다.
- W: 입의 모양이 독립적으로 나타나는 것이 아니라, 전체 얼굴 표정(Whole-face)의 한 부분으로 나타나는 경우이다. 이 분류의 입 움직임은 주로 감정을 표현한다. 예를 들면 [징그럽다]라는 하는 수어를 하면서 나타나는 입의 움직임은 얼굴 전체의 움직임과 함께 나타나므로 이 분류에 속한다.

6) 수어의 분류(Class) 관련 층렬

- ① 수어의 어떤 품사와 입 모양 또는 입 제스처가 함께 나타나는지, 품사의 다른 빈도는 어떠한지 등을 관찰하기 위하여 입 모양 또는 입 제스처가 나타난 경우 함께 나타난 수어의 품사를 별도로 기록한다. 단, 아직 수어의 품사에 대한 연구가 없고 이를 나눌 수 있는 명확한 기준이 없기 때문에 명사와 동사만을 구분하여 입력한다. 동사의 경우 단순 동사, 일치 동사, 공간 동사로 구분하여 입력한다. 입 모양 또는 입 제스처와 함께 나타난 수어의 품사 중에 명사와 동사를 구분하여 입력하며 동사는 단순 동사, 일치 동사, 공간 동사로 분류하여 주석한다.

7) 기타

- ① 수어와 함께 나타나는 마우스 제스처 중에 독립적으로 사용될 수 있으며 명확한 의미를 가진 것들이 있다. 이러한 마우스 제스처가 수어 없이 독립적으로 나타난 경우 ‘M 코멘트’층렬에 [!!!]를 입력한다.
예를 들어 [가능하다]라는 수어의 경우 입술의 긴장을 유지한 채로 입을 닫았다가 여는 마우스 제스처가 동반되지만, 이 마우스 제스처가 수어 없이 홀로 나타났을 때 ‘M 코멘트’층렬에 [!!!]를 입력한다.
- ② 입의 움직임이 있으나 입의 일부 또는 전체가 가려져 판단하기 어려운 경우는 ‘M 코멘트’ 층렬에 [@@]를 입력한다.
- ③ 화자가 어떤 의도를 가지고 입을 움직였다가 주저하거나 도로 입을 다물기도 한다. 이때는 해당 티어에 주석 대신에 [%]를 입력하여 거짓 시작임을 나타낸다.

연구 책임자 윤석민

공동 연구원 이영재, 이현숙, 고아라, 강진주, 최화니, 김성준

연구 보조원 나두호, 서일권, 노유리, 이재준

협력 연구원 변강석, 이은영, 박정민, 김다해, 김효정

담당 연구원 이수미(학예연구사)
김민정(연구원)

발행인: 소강춘

발행처: 국립국어원

서울시 강서구 금낭화로 154

전화 02-2669-9775, 전송 02-2669-9737

인쇄일: 2018년 12월 12일

발행일: 2018년 12월 12일

인 쇄: 온누리

※ 이 책은 국립국어원의 용역비로 수행한 ‘2018년 한국수어 말뭉치 연구 및
구축’의 결과물을 발간한 것입니다.