

국립국어원 2020-01-15

발간등록번호

11-1371028-000822-01

2020년 한국수어 말뭉치 구축 및 연구

연구 책임자 | 강창욱



국립국어원

제 출 문

국립국어원장 귀하

국립국어원과 체결한 연구용역 계약에 따라 「2020년 한국수어
말뭉치 구축 및 연구」에 관한 연구보고서를 작성하여 제출합니다.

■ 사업 기간: 2020년 6월 10일 ~ 2020년 11월 27일

2020년 11월 27일

연구 책임자 : 강창욱(강남대학교)

연구수행기관	강남대학교 산학협력단
연구 책임자	강창욱
공동 연구원	원성욱, 홍성은, 변강석, 이은영, 송미연, 남민우
연구 보조원	명혜진, 홍장미, 이은주

2020년 한국수어 말뭉치 구축 및 연구

- 연구 기관 : 강남대학교 산학협력단
- 연구 기간 : 2020년 6월 10일 ~ 11월 27일 (6개월)
- 연구진 : 총 10명

구분	이름	소속
책임 연구원	강창욱	강남대학교 중등특수교육과 교수
공동 연구원	원성옥	한국복지대학교 한국수어교원과 교수
	홍성은	(주) 에스이랩 연구위원
	변강석	강남대학교 외래교수
	이은영	서울수어전문교육원 강사
	송미연	강남대학교 특수교육재활연구소 연구원
	남민우	서울 동작관악특수교육지원센터 교육공무원
연구 보조원	명혜진	강남대학교 수어통역사
	홍장미	한국복지대학교 외래강사
	이은주	강남대학교 산학협력단 연구원

<국문 요약>

2020년 한국수어 말뭉치 구축 및 연구

이 연구는 한국수어의 언어적 지위 확보와 발전을 위한 연구 기반이 되는 원시 수어 말뭉치 추가 구축 및 국내외 수어말뭉치 연구를 분석하여 국내 수어 말뭉치의 현 위치를 파악하고, 향후 체계적으로 수어말뭉치 구축 및 활용을 위한 중장기 발전방향을 도출하는 데 목적이 있다. 이 연구는 균형성과 대표성을 갖춘 지역 수어 자료를 추가 확보함으로써 수어 사용자의 다양한 수어 형태를 확인할 수 있으며 전년도에 수집된 경남지역 수어 주석을 통해 주석 작업을 완료한 다른 지역의 수어들과 비교할 수 있는 연구 지평을 넓혔다. 한국수어 말뭉치 구축 중장기 발전계획 수립은 향후 한국수어의 언어적 특성을 체계적으로 밝히고 수어 연구 활성화를 위한 실천 방향과 기구축된 한국수어 말뭉치 공개를 체계적으로 접근할 수 있는 기반이 조성될 것으로 기대된다.

‘한국수어 말뭉치 구축’은 2018년도 원시 말뭉치 수집 지역 중 4쌍이 수집된 전남지역(광주 포함)의 농인을 대상으로 15쌍 50시간 이상의 수어 촬영 및 편집을 목표로 한 ‘원시 수어말뭉치 수집’ 과제와 80분 분량의 경남지역 수어 영상을 주석 파일로 작업하는 ‘한국수어 말뭉치 주석’ 과제를 중심으로 추진하였으며, ‘한국수어 말뭉치 연구’는 산재한 국내 수어 자료 실태 분석과 국내외 수어 말뭉치 실태 분석을 통해 ‘한국수어 말뭉치 중장기 발전 계획’ 수립을 목표로 추진하였다.

‘한국수어 원시 말뭉치 수집’은 한국수어 말뭉치 자료의 지역 균형성을 확보하기 위하여 2018년 수행된 원시 말뭉치 수집에서 4쌍을 수집한 전남지역(광주 포함)에서 총 19쌍 38명의 언어제공자로부터 총 59시간 53분 분량의 영상을 촬영하였으며, 편집 및 검증 과정(3회)을 거쳐 최종적으로 49시간 24분 분량의 완성된 수어 자료를 확보하였다. 완성된 영상은 「한국수어 자료 통합지원 시스템」에 탑재하여 원시 수어 말뭉치를 확대·구축하였다.

‘한국수어 말뭉치 주석’은 2018년도와 2019년에 걸쳐 수집되어 주석 데이터로 만들어지지 않은 경남지역 말뭉치 수어 자료에 대해 실시하였다. ‘과제 J(농사회 행사)’ 4개(42분 47초)와 ‘과제M(농학교 경험)’ 3개(37분 32초) 모두 80분 19초 분량의 수어 자료를 기본 주석과 의미 주석을 달았다. 기본 주석에서는 두 손에서 나타나는 총 토큰 수가 13,673개였다. 이 중 [%](수어를 완전히 조어하지 않았을 때 입력 기호), [@](수어 화자가 실수로 잘못 표현했거나 사용되지 않는 수어일 때 입력 기호), [###](주석자가 수어로 인지했지만 무슨 수어인지 모르거나 영상화면 바

끝에서 조음되어 수어가 보이지 않을 때 입력 기호)을 제외한 실질 토큰 수는 13,250개였다. 실질 토큰 중 [지시] 타입을 제외한 모든 토큰에 의미 주석을 달았다. 의미 주석은 해당 토큰을 기록할 때 문장 맥락 안에서의 의미를 중심으로 대역어 형태로 기술하는 것을 말한다. 의미 주석 결과, 총 11,766 토큰에 기록된 의미는 대부분 타입명에 상응하는 의미로 기록되었다. 고정된 수어 타입 [이상하다1]은 비수지 및 문맥에 따라 ‘호기심’, ‘궁금하다’의 의미로 쓰였으며, 일반적으로 ‘괜찮다’, ‘나쁘지 않다’, ‘무방하다’ 등으로 사용되는 타입 [괜찮다1]은 ‘할 수 있다’는 의미로도 쓰이는 사실이 확인되었다. 또한 경남지역 수어 자료의 주석 작업을 통해 새롭게 등록된 타입은 [상주다1], [도전1], [엘지1] 등 모두 160개였다.

‘한국수어 말뭉치 중장기 발전 계획’은 한국수어의 수집과 보존, 체계적인 연구를 목적으로 2015년 국내 최초로 구축된 한국수어 말뭉치가 수어 자료로서의 활용 효용성이 있는 만큼 체계적으로 운영되고 관리될 필요에 따라 그동안의 성과와 해외 말뭉치 현황 분석을 토대로 수립하였다. 한국수어 말뭉치 성과 분석에서는 2019년까지 5개 지역 110명의 언어제공자를 대상으로 촬영한 약 167시간 분량의 언어 자료가 각 과제별로 분할, 편집되어 총 4,278개, 137시간 56분 12초 분량의 원시 수어 말뭉치 영상이 시스템에 탑재되었으며, 이들 수어 영상 중 주석이 완성되어 데이터화한 것은 총 20시간 13분 분량, 총 토큰 수 154,172개로, 구축된 수어 영상의 약 14.6%에 해당하는 자료였다. 해외 수어 말뭉치 현황을 분석한 결과, 장기간의 프로젝트로 지역을 대표할 수 있을 만큼의 충분한 수의 언어제공자로부터 수어 영상을 수집하고 있으며, 지속적이고 다양한 층위의 주석 작업을 통해 수어 말뭉치를 기반으로 수어의 언어적 특성을 밝히는 다양한 연구가 진행되고 있었다. 특히 나라별로 공개 범위와 방법은 상이하나 수어 말뭉치 구축과 함께 수어 말뭉치 공개 준비 작업이 동시에 이루어지고 있음이 확인되었다.

한국수어 말뭉치 구축 사업은 한국수어의 지역별 다양성 확인 및 한국수어의 형태 분석이 가능할 수 있는 강점이 있다. 또한 한국수화언어법의 제정 이후 변화하는 사회 제도와 코로나 19로 인한 공공수어 통역의 확대로 수어에 대한 관심이 고조되고 있는 만큼 향후 5년 동안 전국 단위의 수어 자료 수집과 번역, 주석되지 않은 방대한 양의 수집된 수어 영상의 주석 데이터화, 대중과 연구자 대상 수어말뭉치 공개 사업을 집중적으로 추진해 나가야 하며, 6년차부터는 한국수어와 한국어 간 비교를 위한 병렬 말뭉치와 수어학습자 말뭉치 등 다양한 종류의 말뭉치로 확대 구축해야 할 것이다.

주요어: 한국수어, 한국수어 연구, 한국수어 말뭉치, 주석

<영문 요약>

Building and Research of the Korean Sign Language (KSL) Corpus

The aim purpose of this study is to secure the linguistic status of the Korean Sign Language (KSL) by collecting additional data to the KSL Corpus, analyzing domestic and foreign sign language corpora in order to identify the current position of the KSL Corpus and to systematically derive the mid- to long-term development plan for the building and use of the KSL Corpus. By collecting additional balanced and representative sign language data, it is possible to confirm various forms of sign language types. The annotation of the data collected in Gyeongnam in the previous year can be compared to the annotation of data from other areas. This step will broaden the research horizon of the KSL Corpus. The establishment of a mid- to long-term development plan for the KSL Corpus is expected to reveal the linguistic characteristics of KSL, give new impulses in sign language research and provide a systematical approach concerning the access of the KSL Corpus.

This research project consists of three main parts. The first part is the additional data collection in the area Jeonnam (including Gwangju). This area is one of the areas where data was collected in 2018. At that time data from 4 pairs of deaf informants was collected. This year we collected additional data from 15 pairs of deaf informants which results in about 50 hours of data. The second part contains the annotation of 80 minutes of sign language data from Gyeongnam. The third part consists of a status analysis of the scattered domestic sign language data and of the current KSL Corpus for the mid- to long-term development plan.

The final data collection in the area of Jeonnam (including Gwangju) consists of data (59 h 53 min) from 19 pairs of informants (38 informants). This data was edited and checked 3 times and resulted in 49 hours 24 minutes. The film data was uploaded at the Integrative System for Korean Sign Language Resources.

The data from Gyeongnam which was collected in 2018 and 2019 was annotated this year. We annotated four files (42min 27 sec) of task J (deaf events) and 3 files (37min 32 sec) of task M (experience in deaf school) in respect of basic

annotation and semantic annotation. There are 13,673 tokens in the basic annotation for both hands. If you subtract the annotations with [%] (symbol used to mark unfinished articulation of signs, false stars), [@] (symbol used for slip of the hands or signs that don't exist), and [###] (symbol used for signs the annotators don't know or signs which are out of the video frame), the actual number of tokens is 13,250. All tokens except the index signs were also annotated for meaning. Meaning annotation refers to token in the form of a spoken word centering on the meaning within the context of a sentence. An analysis of the meaning annotation showed that the meaning of 11,766 token corresponded mostly to the name of the type form. Considering the nonmanual signals the sign type STRANGE1 can mean in context 'curiosity' or 'curious'. And the common sign type OKAY1, which is usually used for 'it's okay', 'not bad', 'alright' etc. can also mean 'able to'. Through the annotation work on sign language data from Gyeongnam area 160 new types were registered, including REWARD1, CHALLENGE1 and LG1.

The KSL Corpus mid- to long-term development plan is based on the analysis of the KSL Corpus which was built in 2015 for the purpose of collecting, preserving KSL and stimulating systematic sign language research. The analysis showed the need of a systematic administration of the KSL corpus data and considered the current situation of the overseas corpus. The analysis of the current KSL Corpus showed that until 2019 data from 110 informants was collected in 5 areas. This amounts to 167 hours of film data which results in 137 h 56 min and 12 sec after being edited in 4,278 files. This material was uploaded on the Integrative System for Korean Sign Language Resources. The annotated data is 20 h 13min long (14.6%) and includes 154,172 tokens. The analysis of the overseas long-term sign language corpus projects shows that sign language data is collected from sufficient number of informants to represent the region. Also the multi-tiers annotation is done continuously and various studies are being conducted to reveal the linguistic characteristics of sign language. Although the scope and method of corpus data access differs by country, the various corpus projects have in common that the building of the corpus and the preparation of the data access are being carried out simultaneously.

The KSL Corpus is able to identify the diverse form of KSL by region. The changes of the social system after the enactment of the Korean Sign Language Act as well as the expansion of public sign language translation due to the covid-19 crisis have caused an increasing interest in sign language. Therefore over

the next five years it is necessary to intensively focus on collecting and translating sign language data nationwide, annotating the large amount of sign language data and open the corpus to the public and researchers. From the 6th year it is necessary to expand into various types of corpus such as parallel corpus for comparison between sign language and spoken language by using for example materials such as video sign books or videos from the Deaf Broadcast.

Keywords: Korean Sign Language, Sign language research, sign language corpus, annotation

목 차

I. 서론	3
1. 연구의 필요성과 목적	3
1) 연구 필요성	3
2) 연구 목적	6
2. 연구 방법	7
1) 연구 내용	7
2) 연구 방법	7
3) 연구 절차	9
4) 과업별 추진 일정	10
II. 한국수어 원시 말뭉치 수집	13
1. 언어 자료 수집 절차	13
2. 유도 자료 제작	14
1) 유도 자료 제작 지침	14
2) 유도 자료 내용 수정	24
3. 언어제공자 모집 및 선정	30
1) 조사 지역 선정	30
2) 언어제공자 선정 조건	31
3) 언어제공자 모집	31
4) 언어제공자 기본 사항	32
5) 언어제공자 질문지	45
6) 촬영 동의	45
4. 사회자 교육	46
1) 사회자의 역할	46

2) 사회자 선정 및 교육	46
3) 유도과제 설명 동영상 제작	48
5. 언어 수집 촬영	48
1) 촬영실 조건 및 촬영 시스템 지침	48
2) 촬영장 설치	49
3) 언어제공자 촬영	52
4) 촬영 지침 개선 사항	54
5) 촬영 지침	55
6) 촬영 시간	56
7) 자료 수집을 위한 촬영 시 고려 사항	58
6. 영상 편집	59
1) 편집 지침	59
2) 영상 편집 내용 검수	63
7. 시스템 탑재	65
1) 언어제공자 정보 등록	65
2) 유도 과제 정보 등록	67
3) 원본 및 주석 영상 등록	69
III. 한국수어 말뭉치 주석	77
1. 주석 범위 및 진행 절차	77
1) 주석 범위	77
2) 주석 진행 절차	78
2. 주석자 및 검증자 교육	79
1) 주석자 교육	79
2) 검증자 교육	81
3) 교육 내용	81
3. 기본 주석 및 의미 주석 실시	83
1) 주석 총렬	83

2) 기본 주석	84
3) 의미 주석	86
4. 기본 주석 및 의미 주석 검토	89
1) 검토 과정	89
2) 분절 관련 검토 내용	91
3) 타입명 관련 검토 내용	95
5. 타입 등록	97
6. 수어 말뭉치 주석 규약	99
 IV. 한국수어 말뭉치 구축을 위한 중장기 발전 계획	103
1. 한국수어 말뭉치 추진 배경 및 경과	103
2. 2015-2019 추진 성과	104
1) 추진 성과	105
2) 한계 및 시사점	118
3. 한국수어 자료 실태	122
1) 해외 수어 자료 현황	123
2) 국내 한국수어 자료 현황	125
3) 국내 수어 자료 실태	132
4. 해외 수어 말뭉치 구축 탐색	141
1) 원시 수어 말뭉치 수집	141
2) 주석 달기	145
3) 수어 말뭉치 활용 연구	149
4) 해외 말뭉치 구축 탐색의 시사점	152
5. 해외 수어 말뭉치 공개 방법	158
1) 독일	158
2) 네덜란드	166
3) 호주	169
4) 영국	171

5) 미국	173
6) 해외 수어 말뚱치 공개의 시사점	174
6. 한국수어 말뚱치 구축을 위한 중장기 계획	177
1) 중장기 계획 수립 배경	177
2) 한국수어 말뚱치 구축 SWOT 분석과 대응 전략	178
3) 중장기 계획 수립	183
4) 한국수어 말뚱치 중장기 계획 소요 예산	204
5) 연도별 추진 계획	204
V. 요약 및 제언	215
1. 요약	215
2. 활용 방안 및 제언	221
참고문헌	224
부록	229
[부록 1] 한국수어 말뚱치 구축 가이드	229
[부록 2] 한국수어 말뚱치 주석 규약	233
[부록 3] 촬영 동의서	265
[부록 4] 언어제공자 질문지	270
[부록 5] 언어제공자와 사회자 관계에 관한 질문	277
[부록 6] 사회자 일지	279
[부록 7] 후보 타입 등록 목록	280
[부록 8] 한국수어 말뚱치 수집 시간 현황(2015~2019)	290
[부록 9] 한국수어 말뚱치 기본(의미) 주석 현황(2015~2019)	292
[부록 10] 한국수어 말뚱치 번역 시간 현황(2015~2019)	293

표 목 차

〈표 II-1〉 연도별 ‘E-토론’ 주제 비교	25
〈표 II-2〉 연도별 ‘G-특정사건’ 주제 비교	27
〈표 II-3〉 ‘H-주제 영역’ 과제	28
〈표 II-4〉 언어제공자 선정 조건	31
〈표 II-5〉 언어제공자의 일반적인 사항	32
〈표 II-6〉 전남지역 말뭉치 전체 촬영 시간	57
〈표 II-7〉 전남지역 말뭉치 과제별 편집 시간	57
〈표 III-1〉 기본 주석과 의미 주석의 범위	77
〈표 III-2〉 과제별 기본 주석 토큰 수	84
〈표 III-3〉 과제별 기본 주석 총렬의 [%], [#], [@] 토큰 수	85
〈표 III-4〉 타입별 토큰 수	86
〈표 III-5〉 과제별 고정된 수어의 의미 주석 토큰 수	86
〈표 III-6〉 한국수어 말뭉치 주석 규약의 구성	99
〈표 IV-1〉 한국수어 원시 말뭉치 수집 양적 성과	107
〈표 IV-2〉 한국수어 원시 말뭉치 수집 질적 성과	108
〈표 IV-3〉 한국수어 원시 말뭉치 지역 및 과제별 수집 현황	109
〈표 IV-4〉 한국수어 말뭉치 주석의 양적 성과	113
〈표 IV-5〉 한국수어 말뭉치 주석의 질적 성과	113
〈표 IV-6〉 한국수어 원시 말뭉치 주석 시간 및 토큰 수 현황	115
〈표 IV-7〉 한국수어 말뭉치 활용 사례 연구 내용	119
〈표 IV-8〉 한국농아방송 2019년 유형별 영상 목록	126
〈표 IV-9〉 국립장애인도서관 수어 영상 도서 주제별 제작 현황	128
〈표 IV-10〉 국립장애인도서관 수어 영상 도서 2019년 제작 현황	129
〈표 IV-11〉 국립장애인도서관 한국수어영상자료 목록 예시	136

〈표 IV-12〉 호주 말뚱치의 주석 층위 및 유형	146
〈표 IV-13〉 해외 수어 말뚱치의 수어 자료 수집 현황	154
〈표 IV-14〉 해외 수어 말뚱치의 수어 자료 주석 및 활용 현황	155
〈표 IV-15〉 해외 수어 말뚱치의 공개 현황	176
〈표 IV-16〉 해외 수어 말뚱치의 구축 현황	181
〈표 IV-17〉 해외 수어 말뚱치 구축 연도별 추진 계획	205
〈표 IV-18〉 해외 수어 말뚱치 구축 소요 예산	206

그 림 목 차

[그림 Ⅰ-1] 연구 절차	9
[그림 Ⅱ-1] 언어 자료 수집 절차	14
[그림 Ⅱ-2] 수어 이름 말하기	16
[그림 Ⅱ-3] 준비해 온 농담 말하기	16
[그림 Ⅱ-4] ‘다시 이야기하기’ 과제 제시용 ‘Tweety & Sylvester’	17
[그림 Ⅱ-5] ‘수어 보고 다시 이야기하기’ 과제 제시용 ‘나라별 넥타이 고르는 법’	18
[그림 Ⅱ-6] ‘표지판’ 과제의 예	19
[그림 Ⅱ-7] ‘특정 사건’ 과제의 예	20
[그림 Ⅱ-8] ‘주제 영역’ 과제의 예	21
[그림 Ⅱ-9] ‘개별 수어의 다양한 표현’ 과제의 예	22
[그림 Ⅱ-10] ‘농사회 행사’ 과제의 예	22
[그림 Ⅱ-11] ‘그림 이야기 보고 설명하기’ 과제의 예	23
[그림 Ⅱ-12] 2020년도 ‘G-특정사건’ 주제별 이미지	26
[그림 Ⅱ-13] 2020년도 ‘H-주제 영역’ 이미지	28
[그림 Ⅱ-14] 전남지역에서 사용하는 개별 수어 유도 자료	28
[그림 Ⅱ-15] 전남지역에서 개최된 행사 포스터	29
[그림 Ⅱ-16] 사회자의 유도자료 설명 동영상 예시 화면	29
[그림 Ⅱ-17] 한국수어 말뭉치 언어자료 수집 현황	30
[그림 Ⅱ-18] 언어제공자 모집 및 선정 절차	32
[그림 Ⅱ-19] 언어제공자의 연령 분포	33
[그림 Ⅱ-20] 언어제공자의 성별	34
[그림 Ⅱ-21] 언어제공자의 거주 지역 현황	34
[그림 Ⅱ-22] 언어제공자의 청력상태/청력손실 시기	35
[그림 Ⅱ-23] 언어제공자의 청각보조기기 착용 여부	35

[그림 II-24] 언어제공자 아버지의 농인 여부	36
[그림 II-25] 언어제공자 어머니의 농인 여부	36
[그림 II-26] 언어제공자와 부모와의 동거 여부	37
[그림 II-27] 언어제공자와 어머니와의 주된 의사소통 방법	37
[그림 II-28] 언어제공자와 아버지와의 주된 의사소통 방법	38
[그림 II-29] 언어제공자와 형제 및 자매들과의 주된 의사소통 방법	39
[그림 II-30] 언어제공자의 배우자의 농인 여부	39
[그림 II-31] 언어제공자와 배우자와의 주된 의사소통 방법	40
[그림 II-32] 언어제공자와 자녀와의 주된 의사소통 방법	40
[그림 II-33] 언어제공자(본인)의 주된 의사소통 방법	41
[그림 II-34] 언어제공자의 한국수어 습득 시기	42
[그림 II-35] 언어제공자의 한국수어 습득 장소	42
[그림 II-36] 언어제공자의 한글지문자(지화) 사용 유무	43
[그림 II-37] 언어제공자의 농(난청)인과의 정기적인 만남 여부	43
[그림 II-38] 언어제공자의 최종 학력	44
[그림 II-39] 언어제공자의 재직 여부	44
[그림 II-40] 언어제공자의 직장에서의 주된 의사소통 방법	45
[그림 II-41] 사회자 교육 및 유도자료 설명 동영상 촬영 연습	47
[그림 II-42] 촬영실 모습	49
[그림 II-43] 촬영실 전체 배치도	50
[그림 II-44] 언어제공자 촬영실 배치도	50
[그림 II-45] 언어제공자 모니터 공간	51
[그림 II-46] 언어제공자 휴식 공간	52
[그림 III-1] 주석 진행 절차	78
[그림 III-2] 제스처 이동 단계의 예 ‘한국수어 [가족]’	82
[그림 III-3] 분절 기준에 대한 설명 영상 QR 코드	82
[그림 III-4] 엘란 프로그램의 기본 주석 총렬 구성	83
[그림 III-5] 엘란 프로그램을 통한 주석 기록의 예	84

[그림 III-6] 주석 검토 과정	90
[그림 III-7] 「한국수어자료통합지원시스템」 내 ‘타입후보’ 목록	98
[그림 IV-1] 한국수어 말뭉치 추진 배경	103
[그림 IV-2] 한국수어 말뭉치 구축 경과	104
[그림 IV-3] 한국농아방송 제작 과정	128
[그림 IV-4] 국립장애인도서관 수어 영상 도서 동화 예시	130
[그림 IV-5] 수어 영상 도서 제작 과정	131
[그림 IV-6] 한국농아방송 연도별 영상 수	132
[그림 IV-7] 뉴스 진행자 성별 비율	133
[그림 IV-8] 뉴스 진행자 연령 비율	133
[그림 IV-9] 뉴스 진행자 역할 비율	134
[그림 IV-10] 뉴스 진행자 출신학교 비율	134
[그림 IV-11] 국립장애인도서관 수어 영상 도서 접속 경로	137
[그림 IV-12] 독일 수어 말뭉치(DGS-KORPUS) 누리집 접속 화면	158
[그림 IV-13] 일반공개(MY DGS)와 연구자 공개(MY DGS-annotated)	159
[그림 IV-14] 일반공개(MY DGS) 접속 화면	160
[그림 IV-15] 연구자 공개(MY DGS-annotated) 접속 화면	161
[그림 IV-16] 지역별 언어제공자 연령대별 현황	162
[그림 IV-17] 말뭉치 과제별 영상 및 주석 파일 저장 화면	163
[그림 IV-18] 파일명 ‘dgskorpus_ber_01’의 예시 화면	164
[그림 IV-19] ‘타입’ 메뉴 접속 화면	165
[그림 IV-20] 타입 ‘≡ \$ALPHA^’ 접속 예시 화면	165
[그림 IV-21] 막스 플랑크 언어연구소 누리집	166
[그림 IV-22] 네덜란드 언어 보관소 누리집	167
[그림 IV-23] TLA의 언어 보관소	167
[그림 IV-24] 네덜란드 수어 말뭉치(Corpus NGT)	168
[그림 IV-25] 네덜란드 수어 말뭉치(Corpus NGT) 예시 화면	169

[그림 IV-26] 호주 수어 말뚱치 보관소	171
[그림 IV-27] 영국 수어 말뚱치 프로젝트 누리집의 데이터 공개 안내	172
[그림 IV-28] CAVA 영국 수어 말뚱치 보관소	172
[그림 IV-29] 미국 수어 말뚱치 보관소	174
[그림 IV-30] 한국수어 말뚱치 구축 SWOT 분석	178
[그림 IV-31] SWOT 분석에 따른 대응 전략	179
[그림 IV-32] 한국수어 말뚱치의 비전과 추진 과제	184

I . 서론



1. 연구의 필요성과 목적
2. 연구 방법

I. 서론

1. 연구의 필요성과 목적

1) 연구 필요성

○ 한국수어의 언어적 위치 확보와 한국수어 연구 기반 마련의 필요성 대두

2016년 한국수화언어법 제정으로 한국수어를 농인의 공용어로 공포하고 농인이 모든 생활영역에서 한국수어를 통하여 삶을 영위하고 필요한 정보를 제공받을 권리가 있음을 천명함으로써 한국수어는 그 언어적 위치를 확보하였다. 그러나 음성언어와는 달리 농부모 밑에 태어난 소수의 수어 사용자를 제외하고는 모어로 수어를 습득하는 경우가 거의 없다. 따라서 수어의 구조적 체계는 다양한 형태로 나타날 수 있으며, 이로 인해 수어 연구자들은 직관적 수어 체계의 기술이 더욱 어려워질 수밖에 없다. 즉, 수어는 음성언어(한국어)와는 달리 습득과 학습 환경이 열악하여 많은 변이가 나타나므로 수어 체계에 대한 실증적 연구를 위해 다양한 수어 자료들을 수집하고 기술할 필요가 있다. 따라서 수어의 구조와 사용, 수어 습득과 발달, 그리고 수어 학습에 대한 이해와 연구를 활성화시키고 발전시키기 위해서는 수어 말뭉치를 기반으로 하여 한국수어의 문법체계를 밝히는 과학적이고 체계적인 수어 연구가 이루어져야 한다.

○ 외국에서는 자국의 수어와 농문화 보존을 위한 수어 말뭉치 구축 및 이를 기반으로 한 수어 언어학 연구가 활발히 이루어지고 있음

미국, 독일, 네덜란드, 영국, 호주 등 해외 여러 나라에서는 문자가 없는 수어의 보존과 농문화의 기록, 그리고 수어 연구 기반의 제공 등을 목적으로 수어 말뭉치를 구축해 오고 있다. 수어 연구를 일찍부터 해 온 나

라들에서는 수어 말뭉치를 기반으로 하여 분류사나 일치동사 등 형태론적 주제에 대한 수어 언어학 연구들이 많이 이루어졌으며, 최근 ‘수어 연구의 이론적 이슈(Theoretical Issues in Sign Language Research, TISLR)’라는 국제 학술 대회에서는 사회 언어학, 비수지 기호, 심리 언어학, 토착 수어와 이중 언어 등에 대한 연구가 발표되는 등 다양한 주제의 수어 연구로 발전하고 있다.

○ 우리나라에서도 한국수어 말뭉치 시범 구축 사업 및 한국수어 자료 통합지원 시스템 구축 사업이 진행 중임

우리나라에서도 2015년에 한국수어 연구 활성화를 위한 과학적인 연구 기반을 제공하기 위해 ‘한국수어 말뭉치 기반 조사 및 시범 구축’ 사업을 진행하였다. 이 시범 구축 사업에서는 말뭉치 구축을 위한 지침서를 개발하고, 서울지역의 농인 30쌍을 대상으로 90시간 분량의 언어자료를 수집하여 그중에서 13시간 분량의 기본 주석을 실시하였다.

2016년에는 ‘한국수어 자료 통합지원 시스템 구축’ 사업이 진행되어 주석 지침을 수정 보완하고 수어자료 주석을 위한 시스템을 마련하였다. 또한 이 사업에서는 한국수어 말뭉치를 활용한 실증적 수어 연구 사례를 제시하기 위해 입 움직임과 관련된 주석 지침을 마련하였고, 이에 따라 한국수어에 나타난 입 움직임에 대한 연구사례를 제시하기도 하였다. 그와 더불어 해외 수어 자원과의 연계 방안을 제시하고 농사회를 대상으로 한국수어 말뭉치 구축에 대한 이해를 향상시키기 위한 워크숍도 개최하였다.

2017년에는 기 수집된 원시 말뭉치를 대상으로 250분 분량의 영상에 대한 기본 주석을 실시하였다. 또한 타입을 재정비하여 그동안 수집한 원시 말뭉치와 구축된 말뭉치를 한국수어자료 통합 지원 시스템에 연결하는 작업을 완성하였다.

2018년에는 한국수어 말뭉치의 확대를 위해 전북, 전남, 강원 경남 네 지역의 농인을 대상으로 총 14쌍의 원시 말뭉치 자료를 수집하였으며, 총

80분 분량의 기본 주석을 실시하였다.

2019년에는 균형을 갖춘 수어 말뭉치 수집을 위해 2018년 수집 지역 중 수집 분량이 가장 적었던 경남지역의 농인 총 11쌍 22명을 대상으로 원시 수어 말뭉치 자료를 추가로 확보하였다. 그리고 기본 주석이 완료된 영상 총 11개 135분 분량에 대해 기본 주석 정비, 의미 주석을 실시하였으며, 이를 토대로 수어의 도상성에 관한 사례연구를 실시하였다.

○ 한국수어 연구의 토대가 되는 원시 수어 말뭉치를 지역별 균형이 이루어지도록 해당 지역의 토박이 농인으로부터 추가 수집할 필요가 있음

한국수어 원시 말뭉치는 2015년 서울지역 농인 30쌍과 2018년 전북, 전남, 강원지역 농인 각 4쌍 및 경남지역 농인 2쌍을 대상으로 총 14쌍의 수어 자료가 수집되었다. 2018년도 수집된 2쌍에서 4쌍의 원시 자료로는 말뭉치 자료의 지역 균형성 확보 및 지역별 변이 연구 자료로 사용하기에는 한계가 있어 2019년 경남지역 농인 11쌍의 언어자료 추가 수집에 이어 2020년에는 4쌍만 수집된 전남 지역(광주 포함)의 농인을 대상으로 50시간 이상의 언어 자료를 수집하였다.

○ 기 수집된 한국수어 원시 말뭉치에 대한 기본 주석 작업 필요

한국수어 말뭉치가 한국수어 연구의 기초 자료로 사용되기 위해서는 이것에 대한 기본적인 주석 작업이 이루어져야 한다. 2019년 경남지역 11쌍의 농인으로부터 원시 수어 말뭉치를 수집한 바 있으며, 이것에 대해 글로스를 다는 기본 주석 작업이 이루어져야 지역 변이를 연구하기 위한 한국수어 연구 자료로서 활용이 가능해진다.

○ 국내외 수어말뭉치 연구 성과 및 동향을 분석하여 수어말뭉치 구축의 중장기 방향 도출

과학기술의 발달로 현대 언어학에서는 컴퓨터를 활용하여 대용량의 언어자료 처리가 가능해졌으며 이와 관련한 연구 성과들이 축적되고 있다. 한 언어에 대한 깊이 있는 분석뿐만 아니라 두 언어의 사용 양상을 분석하여 자연스러운 변환과 통번역이 가능해지고 있다. 수어 말뭉치 연구도 원시 자료 수집, 주석 달기를 완료한 분석 자료의 구축, 말뭉치를 기반으로 한 활용 연구 및 이중 언어 분석까지 확대되고 있다. 이에 국내외 말뭉치 연구 성과 및 동향을 분석하여 현재 국내 수어 말뭉치의 현 위치를 파악하고 향후 체계적으로 말뭉치를 구축하고 활용할 수 있는 방안을 도출할 필요가 있다.

2) 연구 목적

○ 본 연구의 목적은 크게 3가지이다.

첫째, 전남 지역 토박이 농인을 대상으로 한국수어 원시 말뭉치를 수집한다.

- 한국수어 말뭉치 유도 자료 수정 보완
- 전남 지역 농인 19쌍을 대상으로 원시 말뭉치 영상 촬영 및 편집
- 원시 말뭉치 수집 지침의 수정 보완 및 정교화

둘째, 기존 한국수어 말뭉치에 대한 기본 주석 및 의미 주석을 실시한다.

- 기존에 구축된 말뭉치 타입의 정비
- 주석 인력 교육 실시
- 주석 지침 수정 보완 및 정교화
- 2019년 수집된 80분 분량의 영상에 대한 기본 주석 및 의미 주석 실시

셋째, 한국수어 말뭉치 구축 중장기 계획을 제시한다.

- 국내외 수어 말뭉치 구축 및 연구 현황 분석
- 산재한 한국수어 자료 현황 파악
- 한국수어 말뭉치 구축 사업의 방향 제시

2. 연구 방법

1) 연구 내용

○ 본 연구의 내용은 크게 세 가지 영역에서 진행되었다.

한국수어 원시 말뭉치 언어 자료 수집	한국수어 말뭉치 주석		한국수어 자료 현황 및 중장기 계획
	기 구축된 말뭉치 기본 주석	기 구축된 말뭉치 의미 주석	

2) 연구 방법

(1) 한국수어 원시 말뭉치 구축

: 지역 농인이 참여하는 한국수어 원시 말뭉치 수집

- 2018년에 4쌍만을 대상으로 원시 말뭉치를 수집한 전남 지역(광주 포함)을 대상으로 하였다.
- 지역 변이를 담을 수 있는 유도 자료를 추가하였다.
- 19쌍 50시간 분량의 원시 말뭉치를 수집하였다.
- 수집한 말뭉치를 편집하여 「한국수어 자료 통합지원 시스템」에 등재한다.

(2) 한국수어 말뭉치 주석

: 기 구축된 한국수어 말뭉치에 대한 정비 및 의미 주석

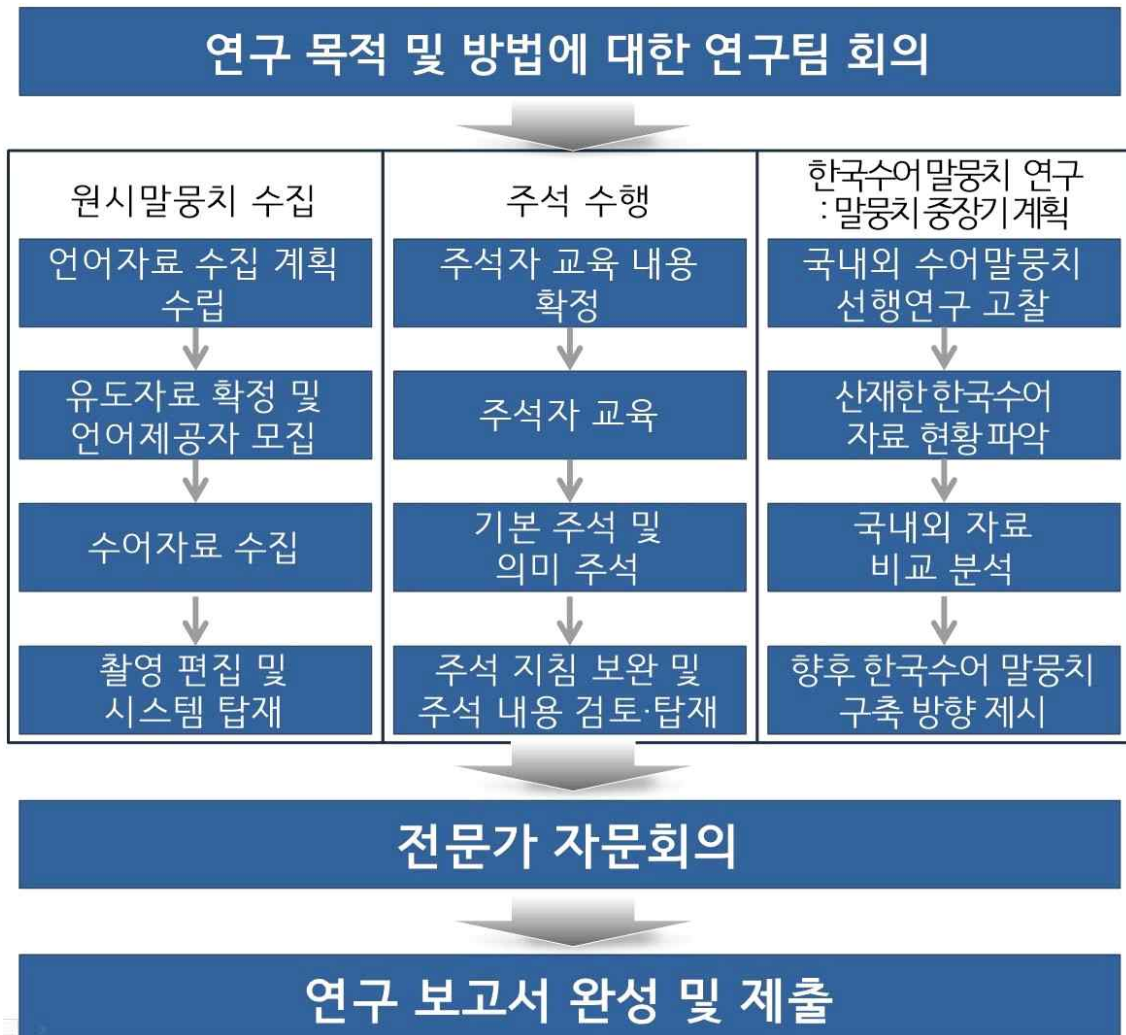
- 주석자 교육 내용을 확정하였다.
- 주석자 교육을 실시하였다.
- 타입을 정비하고 한국수어 말뭉치 지침을 수정 보완하였다.
- 경남지역 말뭉치 자료 80분 분량에 대한 기본 주석을 기록하였다.
- 경남지역 말뭉치 자료 80분 분량에 대한 의미 주석을 기록하였다.

(3) 한국수어 말뭉치 구축 중장기 계획 도출

- 국내외 수어말뭉치 선행연구를 정리하였다.
- 산재한 한국수어 자료 현황을 파악하였다.
- 향후 한국수어 말뭉치 구축 방향을 제시하였다.

3) 연구 절차

본 연구는 다음과 같은 과정에 의해 수행되었다.



[그림 I -1] 연구 절차

4) 과업별 추진 일정

내용	2020년					
	6	7	8	9	10	11
○ 연구계획 수립						
○ 지역의 한국수어 말뭉치 수집						
- 유도 자료 제작(영상 포함)						
- 언어제공자 모집 및 사회자 교육						
- 촬영 및 편집						
- 검수 및 수정						
- 시스템 탑재						
○ 기구축 수어 말뭉치 주석						
- 주석자 모집 및 기본 교육						
- 시범주석 및 검토						
- 주석 실시						
- 주석 검증 및 수정						
- 주석 통계 및 결과 분석						
- 주석 규약 수정						
○ 한국수어 말뭉치 연구						
- 국내외 선행연구 고찰						
- 산재한 수어자료 수집						
- 국내외 자료 비교 분석						
- 중장기 계획(안) 수립						
○ 과업별 최종 점검						
○ 연구 보고서 완성 및 제출						

II. 한국수어 원시 말뭉치 수집



1. 언어 자료 수집 절차
2. 유도 자료 제작
3. 언어제공자 모집 및 선정
4. 사회자 교육
5. 언어 수집 촬영
6. 영상 편집
7. 시스템 탑재

II. 한국수어 원시 말뭉치 수집

2015년 원시 말뭉치 구축은 서울 지역(경기 포함) 농인 30쌍을 대상으로 하였으며, 2018년에는 전북 4쌍, 전남(광주 포함) 4쌍, 강원 4쌍, 경남 2쌍으로 네 개 지역에서 총 14쌍의 농인을 대상으로 수어자료를 수집하였다. 언어의 빈도와 분포를 확인하고 수어 표본의 대표성을 확보하기 위해서는 한 지역에서 최소 10쌍 이상의 언어제공자로부터 수어 자료를 수집하는 것이 특정 지역 수어의 대표성과 신뢰성을 확보할 수 있다. 따라서 2019년에는 2018년에 2쌍만 수집한 경남 지역을 대상으로 11쌍의 수어 말뭉치가 추가 구축되었다. 이번 연구에서는 수어가 농학교를 거점으로 전수·발 전해 온 역사성을 근거로 2018년도 4쌍만 수집된 지역으로서 청각장애학교인 소림학교와 인화학교가 있는 전남 지역(광주 포함)의 수어 자료 수집을 다음 절차로 추진하였다.

1. 언어 자료 수집 절차

~ 6월	촬영 및 편집 업체 선정 유도 자료 수정 보완 및 제작
~ 7월 1주	언어제공자 계약서 동영상 촬영 조건에 맞는 언어제공자 모집 및 검토
7월 2주 ~ 7월 3주	언어제공자 최종 확정(19쌍, 38명) 사회자 교육 실시
7월 4주 ~ 7월 5주	촬영 사전 세팅 점검 촬영
8월 1주 ~ 8월 4주	동영상 편집
9월 1주 ~ 10월 1주	동영상 검수 및 수정

10월 2주 ~ 10월 4주	추가 촬영 및 동영상 편집
11월 1주 ~ 11월 2주	동영상 검수 및 수정
11월 3주 ~ 11월 4주	한국수어 자료 통합지원 시스템 탑재

[그림 II-1] 언어 자료 수집 절차

2. 유도 자료 제작

1) 유도 자료 제작 지침

한국수어 원시 말뭉치 수집을 위한 최초 유도 자료는 한국수어 말뭉치 구축이 처음 실시된 2015년에 제작되었다. ‘한국수어 코퍼스 구축을 위한 기초 연구(원성옥 외, 2013)’에서 신뢰도와 타당성이 검증된 항목에 대해 일차 검토를 마친 후, 홍성은과 니시오 리에(2010)가 제시한 6가지 요소를 기준으로 재논의를 통해 최종 확정하였다.

[유도 자료 6대 요소]

- ① 전체 세션 안에서의 각 과제 소요시간
- ② 총 유도 세션(elicitation session) 시간
- ③ 세션 중 휴식시간 분배의 적합성 및 언어제공자들의 스트레스 정도
- ④ 과제 배열의 적합성 및 과제 배열로 인한 과제 간의 상호작용 여부
- ⑤ 사회자와 언어제공자들 사이의 원활한 진행 여부
- ⑥ 서로 다른 학력 배경을 가진 농인들의 과제 처리 정도

유도 자료는 다음과 같은 원칙에 따라 제작되었다. 총 13가지의 과제로 확정된 유도 자료는 한글 텍스트를 최대한 배제하고 이미지 자료를 중심으로 제작한다. 이때 유도 자료에 사용할 이미지 자료의 저작권 문제는 추가적으로 확인한다. 토론 주제의 경우 다양한 주제를 추가할 수 있다. 따라서 2015년도 유도 자료를 바탕으로 언어제공자의 원활한 대화를 유도할 수 있도록 지역성, 시의성 등을 고려하여 수정·보완될 수 있으나, 구축된 한

국수어 말뭉치가 시기별, 지역별 변이형을 목적으로 연구될 수 있다는 점을 유념하여 각 과제의 목적에 부합하는 정도로 수정한다. 총 13개의 과제로 이루어진 유도 자료는 다음과 같다.

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

(1) 과제 번호 A: 수어 이름 말하기

이 과제는 언어제공자들이 자신의 수어 이름과 그 수어 이름이 어디서 유래되었는지에 대해 설명을 하는 것이다. 농문화에서 가장 대표되는 것 중 하나가 수어 이름이다. 이 과제는 언어제공자가 평소에 자주 설명하는 수어 이름을 소개함으로써 어색한 분위기를 풀고 자연스럽게 대화로 들어가기에 유용한 과제이다.



[그림 II-2] 수어 이름 말하기

(2) 과제 번호 B: 준비해 온 농담

이 과제는 사전에 언어제공자에게 연락하여 촬영 당일 다른 언어제공자에게 보여 줄 만 한 재미있는 이야기를 준비해 오도록 하여 진행하는 과제이다. 이 과제는 언어제공자들이 미리 준비해 온 이야기이므로 편안한 마음으로 자신감 있게 표현할 수 있다는 장점이 있다. 또한 농사회에서 많이 이야기되는 농담(joke)이나 유머(humor)가 나오므로 농문화의 독특한 일면을 볼 수 있다.



[그림 II-3] 준비 해 온 농담

(3) 과제 번호 C: 다시 이야기하기

다시 이야기하기는 한 명의 언어제공자가 주어진 영상 자료를 보고 그 내용을 상대 언어제공자에게 수어로 설명해 주는 과제이다. 이 과제에서는 언어제공자가 말이 없이 동작만 있는 만화영화(Tweety & Sylvester)를 보

고 그 내용을 상대방에게 이야기하도록 한다. 말로 하는 대사가 없이 행동으로만 구성된 장면이므로 분류사 수어 등 수어 본연의 특색이 드러나는 자료를 이끌어 낼 수 있다는 장점이 있다. 또한 대부분의 나라에서 유도 자료로 사용하고 있으므로 다른 나라의 말뭉치와 비교 가능한 자료이다. 이 과제는 총 3분 45초 분량의 영상으로 2015년 최초 제작 당시 6분가량의 영상을 사용하였으나, 영상이 길어 언어제공자가 지루해하는 경향이 있어 2019년 원시 말뭉치 수집부터 반복되거나 비교적 덜 중요하다고 판단되는 부분들을 편집하여 사용하였다. 처음에는 전체 영상을 보여주고, 두 번째로 3분 45초 분량의 영상을 7개 부분으로 나누어 제시하기로 하였다. 이는 언어제공자가 전체 영상을 한 번만 보고 상대방에게 설명하기에는 내용이 너무 많아 부담을 줄 수 있고, 전체 이야기에 대한 설명을 하다 보면 세부적인 내용이 누락되어 과제의 목표가 이루어지지 않기 때문이다. 또한 각각의 과제 영상 사이에 검은색 슬라이드를 삽입해서 언어제공자가 설명을 할 때는 모니터에 검은색 슬라이드만 보이도록 하였다. 이는 언어제공자가 수어를 할 때 시선의 움직임을 촬영하기 위한 것으로 모니터에 아무것도 나오지 않는 상태여야 언어제공자가 모니터를 쳐다보지 않고 상대방을 보고 영상 자료의 내용을 설명하기 때문이다. [그림 II-4] 은 과제 영상의 한 부분이다.



[그림 II-4] ‘다시 이야기하기’ 과제 제시용 ‘Tweety & Sylvester’

(4) 과제 번호 D: 수어 보고 다시 이야기하기

수어로 주어지는 영상 자료를 보고 이야기를 하는 과제로 유도 자료 중에 유일하게 수어로 제시되는 과제이다. 이 과제는 한국농아방송의 Deaf UCC ‘나라별 넥타이 고르는 법’의 영상 중에서 일부분을 잘라 사용하였다. 영상의 길이는 총 2분 21초이다. 2015년에 제작한 유도 자료의 경우 전체 영상을 한번 제시하였는데, 영상의 내용이 길어 언어제공자에게 부담이 되어 2019년에 ‘C-다시 이야기하기’ 과제와 마찬가지로 2분 21초의 전체 영상을 3개의 영상으로 잘라 제시하도록 수정하였다. [그림 II-5]는 ‘나라별 넥타이 고르는 법’ 영상 중 한 장면이다.



[그림 II-5] ‘수어 보고 다시 이야기하기’ 과제 제시용
‘나라별 넥타이 고르는 법’

(5) 과제 번호 E : 토론

이 과제는 논쟁의 여지가 있는 주제를 주고 그 주제에 대해 두 명의 언어제공자가 토론을 하게 하는 과제로, 수어 제공자가 자신이 표현하는 수어의 형식보다 토론 내용에 집중하게 하여 자연스러운 수어 산출을 유도한다. 주제를 제시할 때는 한국어 문장으로 제시하지 않고 수어 동영상을 제작하여 제시하였다. 언어제공자는 제시된 토론 주제 중에 각자 하나씩 골라 총 2개의 주제로 토론을 진행하였다. 언어제공자들의 흥미와 적극적인 참여를 유도하기 위해 다양하고 시의성이 있는 주제들로 보완하였다.

(6) 과제 번호 F: 표지판

여러 나라의 다양한 표지판을 제시하고 그 의미를 유추하게 하여 서로 이야기를 나누도록 하는 과제이다. 이 과제에서 제시하는 표지판은 다양한 유추가 가능하므로 언어제공자들의 풍부한 대화를 유도할 수 있으며, 특히 금지어 등 부정문을 유도하는 데 유용하게 많이 사용된다. [그림 II-6]은 표지판의 한 예이다.



[그림 II-6] ‘표지판’ 과제의 예

(7) 과제 번호 G: 특정 사건 묘사

이 과제는 제시된 사건들 중에 하나의 사건을 선택하여 설명하고, 그 당시에 본인이 무엇을 하고 있었는지를 이야기하는 과제이다. 언어제공자는 다양한 연령대로 구성되어 있다. 따라서 태어난 시대에 따라 특정 사건에 대해 모를 수도 있기 때문에 언어제공자들의 연령을 고려하여 1970년대부터 2010년대 사이에 일어난 역사적인 사건들을 시대별로 골고루 구성하였다. 이 과제는 사진을 보고 해당 사건을 파악할 수 있도록 사건의 대표적인 사진을 선택하여 제시하였다. 언어제공자는 사건을 묘사하며 사건이 일어난 당시 본인이 무엇을 하고 있었는지를 설명해야 한다. 농인은 정보에 접근하기 쉽지 않은 사회적 구조로 인해 어떤 사건이 발생하였을 때, 청인보다 정보 접근이 늦어지거나, 정보를 추측하게 된다. 이 과제를 통해 언어자료 수집 이외에도 농인이 어떻게 정보를 접하게 되는지 확인할 수 있다. 두 명의 언어제공자가 교대로 다음에 제시된 사건들 중 각자 하나의

사건을 선택하여 이야기를 나눈다. 언어제공자의 참여를 유도하기 위해 지역성을 고려하여 해당 지역사건을 추가할 수 있으며, 지역사건을 추가하는 경우 이전 지역사건을 삭제한다. [그림 II-7]은 특정 사건에 대한 사진을 하나의 슬라이드에 모아놓은 자료이다.

- ① 육영수 여사 피살 (1974)
- ② 삼풍백화점 붕괴 사고 (1995)
- ③ 2002 월드컵 (2002)
- ④ 세월호 침몰 (2014)
- ⑤ 남대문 화재 (2008)
- ⑥ 노무현 대통령 자살 (2009)



[그림 II-7] ‘특정 사건’ 과제의 예

(8) 과제번호 H : 주제 영역

이 과제는 사전 제작에 필요한 어휘를 수집하기에 유용한 과제로 언어 제공자가 제시된 주제에 대해 자유롭게 대화를 나누게 하였다. 제시한 주제는 날씨, 경제(은행, 증권, 노후 등), 감정, 직업, 가족과 친척, 시간(일정 등), 음식, 정치, 문화생활, 행사(잔치), 의사소통(전화, 스마트폰 등), 법, 도시와 시골, 주택과 아파트, 공공기관, 휴가(여행, 자유 시간), 패션, 종교, 외

모, 건강, 사랑(성, 결혼), 교통, 계절 등이다. 주제 영역 과제는 하나의 슬라이드에 한글로 주제가 제시되며 그와 관련된 그림 또는 사진이 함께 제시된다. 이중에서 기본적인 어휘가 많이 나올 것으로 판단되는 12개의 주제를 선정하였고, 수어 산출이 한쪽으로 치우치는 것을 방지하기 위해 1번 쌍부터 15번 쌍까지의 제공자들에게는 6개의 주제를 제시하고, 16번 쌍부터 30번 쌍까지의 언어제공자들에게는 다른 6개 주제를 제시하였다. 언어제공자들은 각각 하나씩 주제 2개를 선택하여 이에 대해 자유롭게 이야기를 나누었다. 하지만 언어제공자들의 대화가 매끄럽게 이어지지 않거나, 무슨 말을 해야 할지 모르는 경우, 사회자는 언어제공자들에게 미리 준비해 놓은 질문을 부여해서 과제가 원활히 진행될 수 있도록 하였다. 다음 [그림 II-8] 는 주제에 대한 그림 자료이다.



[그림 II-8] ‘주제 영역’ 과제의 예

(9) 과제 번호 I: 개별 수어의 다양한 표현

이 과제는 지역 방언의 표현을 유도해 내기 위한 과제이다. 가족, 미원, 색깔, 콩나물 등 지역별로 다양하게 사용되는 수어를 유도하기 위해 관련 그림을 준비하여 언어제공자에게 보여주고 수어로 표현하게 하였다. 많은 종류의 유도 자료 중에 대화가 아니라 단어 단위로 수어를 유도해 내는 과제로 다른 지역에서도 이런 유도 자료 수집을 통해 지역 방언 연구가 가능하다. 언어제공자의 출신 학교나 고향에 따라 개별적으로 다르게 사용할 것으로 예측되는 7가지 색(보라색, 분홍색, 빨강색, 회색, 밤색, 주황색, 회

색)을 보여 주었으며, 고모, 이모, 외삼촌과 같은 가족 관계 어휘는 한글로 제시하였다. 그리고 미원과 콩나물은 사진으로 제시하였다. 지역 방언을 유도하기에 적절하다고 판단되면 세부주제를 추가할 수 있다. 다음 [그림 II-9]는 ‘개별 수어의 다양한 표현’을 유도하기 위한 유도 자료이다.



[그림 II-9] ‘개별 수어의 다양한 표현’ 과제의 예

(10) 과제 번호 J: 농 사회 행사

농 사회의 주요 행사에 대한 포스터를 제시하고, 3개의 행사 중에 자신이 참석했던 행사 한 가지를 골라서 자신의 경험을 이야기하도록 하는 과제이다. 이 과제는 ‘농학교 경험 이야기하기’ 과제와 함께 가장 농인다운 대화를 유도할 수 있는 자료로 언어제공자들에게 편안하고 즐거운 분위기에서 이야기를 이끌어 낼 수 있는 장점이 있다. 지역마다 참여하는 행사가 다르므로 수어 수집을 하는 해당 지역의 행사 포스터를 제시하였다. 다음 [그림 II-10]은 서울지역의 농 사회 행사 포스터이다.



[그림 II-10] ‘농사회 행사’ 과제의 예

(11) 과제 번호 K: 그림 이야기 보고 설명하기

이 과제는 먼저 ‘홍부와 놀부’ 그림을 보고 한 명의 언어제공자가 상대방에게 ‘홍부와 놀부’ 이야기를 설명하는 과제이다. 한국 사람에게 ‘홍부와 놀부’는 매우 친숙한 이야기이므로 언어제공자에게 큰 부담을 주지 않는다. 한 슬라이드에 한 장씩 총 10장의 그림을 보여 준 다음 다시 한 슬라이드에 4장의 그림을 넣어 보여 주었다. 두 번째로 그림을 보여 줄 때, 각 슬라이드 사이에 검은색 슬라이드가 삽입되어 있는데, 언어제공자는 검은 슬라이드가 제시되어 있는 동안 상대방에게 자신이 본 그림 이야기를 설명한다. [그림 II-11]은 그림 이야기 보고 설명하기의 홍부와 놀부 그림 자료이다.



[그림 II-11] ‘그림 이야기 보고 설명하기’ 과제의 예

(12) 과제 번호 L: 옛날 수어와 현재 수어

이 과제는 옛날에는 사용했으나 현재는 사용하지 않는 수어 또는 옛날과 형태가 달라진 수어를 생각나는 대로 이야기해 보게 하는 과제이다. 이 과제를 통해 사회언어학적 변화에 대한 자료와 메타언어 자료를 수집할 수 있다.

(13) 과제 번호 M: 농학교 경험 이야기하기

2013년에 실시한 기초 연구에서는 ‘어떤 일(특정 사건)이 발생했을 때

본인이 무엇을 했는지'를 이야기하는 특정 사건 묘사(G) 과제에서 언어제공자의 경험을 이끌어 내기가 어려웠다. 그래서 본 연구에서는 농학교에서의 자신의 경험을 이야기해 보는 과제로 수정하여 제시하였다. 농학교는 농문화 생산과 전달의 핵심적인 곳으로 언어제공자들이 오랜 시간의 촬영을 통해 지쳐갈 때 마지막 유도 자료로 농학교의 경험을 이야기하는 과제를 제시함으로써 즐겁게 자신의 경험을 이야기할 수 있도록 하였다.

(14) 유도 자료 설명 동영상

확정된 유도 자료와 더불어 과제를 수행하기 위한 별도의 설명이 필요하다. 사회자의 컨디션이나 다양한 환경적 요인에 따라 유도 자료에 대한 설명이 달라질 수 있으므로 언어제공자들에게 동일한 설명을 제공할 수 있도록 별도의 영상을 제작하였다. 그리고 같은 과제를 수행하더라도 각각의 언어제공자에게 다른 역할이 부여되는데, 동영상 방식의 경우 각각의 언어제공자에게 동시에 다른 설명이 가능하다는 점에서 설명 방식으로서 유용하다. 언어제공자들의 지역 수어를 효과적으로 유도해 내기 위해 사회자는 해당 지역에 거주하는 농인으로 선정하고 유도 자료 설명 동영상을 촬영하도록 하였다. 대상 지역이 변경될 때마다 지역의 농인 사회자로 하여금 유도 자료 설명 동영상을 새롭게 제작하였다.

2) 유도 자료 내용 수정

(1) 유도 자료 내용 수정

지역을 달리하여 언어자료를 수집하는 본 연구에서는 언어의 변이형을 비교할 수 있도록 2015년, 2018년, 2019년에 사용한 유도 자료를 기반으로 아래 내용을 수정, 보완하여 재구성하였다.

총 13가지의 과제 중에서 지역의 특성 등을 반영하기 위하여 'E-토

론’, ‘G-특정 사건 묘사’, ‘H-주제 영역’, ‘I-개별 수어의 다양한 표현’, ‘J-농사회 행사’의 주제를 수정하였다.

‘E-토론’ 과제의 목적은 수어 언어제공자가 자신이 표현하는 수어의 형식보다는 치열한 토론 내용에 집중하게 하여 자연스러운 수어 산출을 유도하는 것이므로 다양한 주제를 제공하는 것이 바람직하다. 특히 토론 주제는 ‘찬반 논쟁’이 가능한 주제여야 하며, 언어제공자의 연령대나 지역 특성에 따라 선호하는 주제가 다를 수 있으므로 선택할 수 있는 주제를 다양하게 설정해놓을 필요가 있다. 5년에는 ① 1종 면허 농, ② 농학교 농인 교사, ③ 한국수어 통역사 협회 설립, ④ 메르스 등 4가지였으며, 2018년에는 ③ 한국수어 통역사 협회 설립과 ④ 메르스가 시의적절하지 않으므로 시의 적절한 주제인 ③ 수어통역사 센터 분리 ④ 폭염으로 교체한 바 있다. 2019년에는 ① 1종 면허 농, ② 농학교 농인 교사, ③ 수어통역센터 분리, ④ TV 수어통역 대 문자통역, ⑤ 남북통일, ⑥ 일회용품 사용 ⑦ 수어교원자격증 문제로 진행하였고, 본 연구에서는 이에 ⑧ 코로나 개학 연기를 추가하였다.

<표 II-1> 연도별 ‘E-토론’ 주제 비교

2020년	2015년	2018년	2019년
1. 1종 면허 농	1. 1종 면허 농	1. 1종 면허 농	1. 1종 면허 농
2. 농학교 농인 교사	2. 농학교 농인 교사	2. 농학교 농인 교사	2. 농학교 농인 교사
3. 수어통역센터 분리	3. 한국수어통역사 협회 설립	3. 수어통역센터 분리	3. 수어통역센터 분리
4. TV 수어통역 대 문자통역	4. 메르스	4. 폭염	4. TV 수어통역 대 문자통역
5. 남북통일			5. 남북통일
6. 일회용품 사용			6. 일회용품 사용
7. 수어교원 자격증 문제			7. 수어교원 자격증 문제
8. 코로나 개학 연기			

특정 사건 묘사 과제의 목적은 특정 사건 당시에 본인이 무엇을 하고 있었는지를 이야기하는 과제로 언어제공자의 연령대를 고려하여 오래된 사건이라도 기억할 만한 역사적 사건을 제시하는 것이 필요하다. G-특정 사건 묘사 과제를 수행하기 위해 지난 2015년에는 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 침몰을 제시하였다. 그러나 2018년에는 위의 사건들이 10년 이상 오래된 사건이므로 세월호 사건을 제외하고 전면 교체하여 사용한 바 있다. 2019년에는 ① 1974년 육영수 여사 피살 사건, ② 1980년 5·18 민주화 운동, ③ 1995년 삼풍백화점 붕괴, ④ 2002년 월드컵 스페인전, ⑤ 2009년 노무현 대통령 자살, ⑥ 2014년 세월호 침몰, ⑦ 2017년 행복팀 사기사건, ⑧ 2018년 남북 정상회담, ⑨ 2019년 안인득 사건으로 구성하였다. ‘2019년 안인득 사건’의 경우 경남지역 사건으로, 본 연구에서는 전남지역의 지역적 특성을 반영하기 위해 2011년 도가니 사건으로 대체하였고, 추후에도 다른 지역 사건 주제가 추가될 경우 기존의 지역 사건을 변경하기로 하였다.



[그림 II-12] 2020년도 ‘G-특정사건’ 주제별 이미지

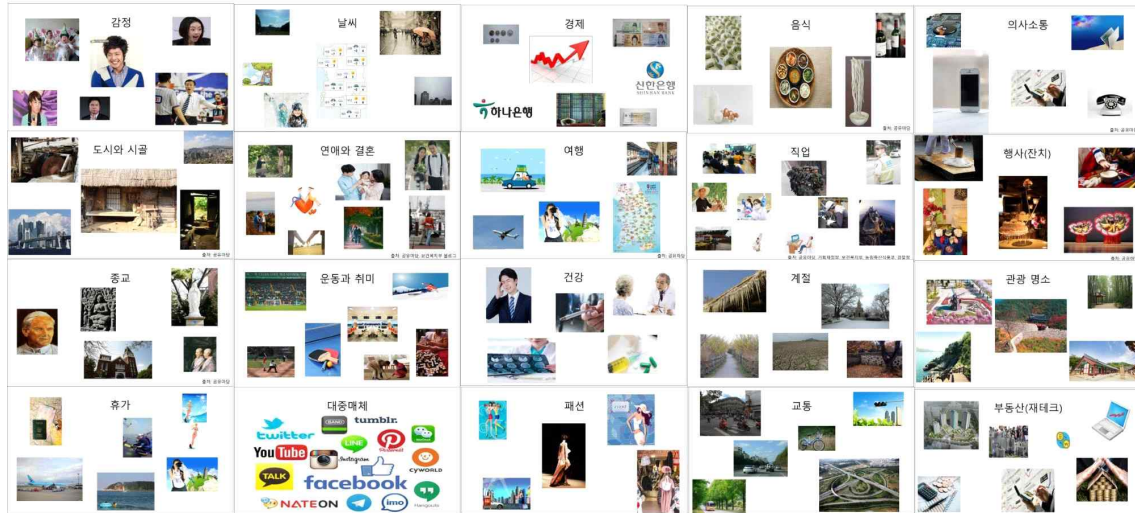
〈표 II-2〉 연도별 ‘G-특정사건’ 주제 비교

2020년	2015년	2018년	2019년
1. 1974년 육영수 여사 피살 사건	1. 1974년 육영수 여사 피살 사건		1. 1974년 육영수 여사 피살 사건
2. 1980년 5·18 민주화 운동		(전남/전북 소재에 포함됨)	2. 1980년 5·18 민주화 운동
3. 1995년 삼풍백화점 붕괴	2. 1995년 삼풍백화점 붕괴		3. 1995년 삼풍백화점 붕괴
4. 2002년 월드컵 스페인전	3. 2002년 월드컵 스페인전		4. 2002년 월드컵 스페인전
	4. 2008년 남대문 화재 사건		
5. 2009년 노무현 대통령 자살	5. 2009년 노무현 대통령 자살		5. 2009년 노무현 대통령 자살
6. 2011년 도가니 사건		1. 2011년 가슴기 살균제 사건	
7. 2014년 세월호 침몰	6. 2014년 세월호 침몰	2. 2014년 세월호 침몰	6. 2014년 세월호 침몰
8. 2017년 행복팀 사기 사건		3. 2018년 평창 동계올림픽	7. 2017년 행복팀 사기 사건
9. 2018년 남북 정상회담		4. 2018년 남북 정상회담	8. 2018년 남북 정상회담
		5. 2018년 러시아 월드컵	9. 2019년 안인득 사건
		6. 지역별 사건 전남/ 전북 5.18 민주화 운동	
		강원 고성 산불	
		경남 포항 지진	

‘H-주제 영역’의 경우 세부주제인 ‘관광명소’는 전남지역의 유명한 관광지를 중심으로 이미지를 수정하였다. 특정 주제에 집중되지 않고 다양한 수어를 수집하기 위하여 총 20개의 주제를 4개씩 5개 형태로 구분하여 촬영 일자마다 다른 주제를 제시하였다. 예를 들어 첫째 날과 둘째 날에 촬영하는 1번 쌍과 2번 쌍에게는 감정, 도시와 시골, 종교, 휴가 4가지의 주제만 제시하여 그중에서 각각 2개씩 골라서 이야기를 나누는 방식으로 진행하였다.

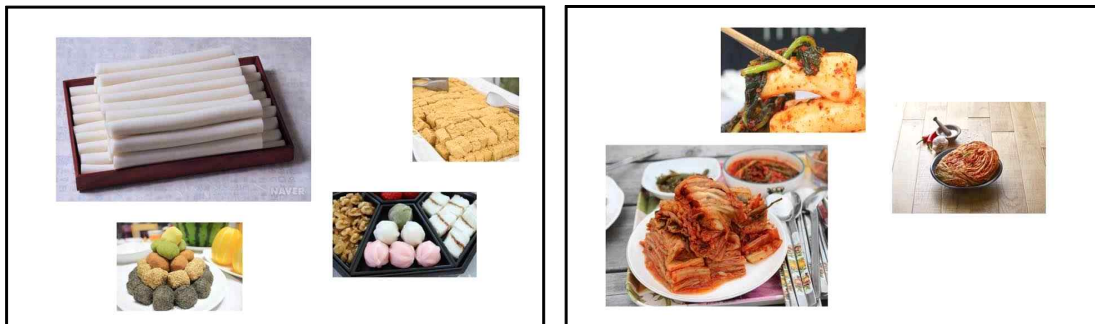
〈표 II-3〉 ‘H-주제 영역’ 과제

1번째	감정	도시와 시골	종교	휴가
2번째	날씨	연애와 결혼	운동과 취미	대중매체
3번째	경제	여행	건강	패션
4번째	음식	직업	계절	교통
5번째	의사소통	행사(잔치)	관광명소(전남)	부동산



[그림 II-13] 2020년도 ‘H-주제 영역’ 이미지

지역 방언의 표현을 유도해 내기 위한 자료인 ‘I 개별 수어의 다양한 표현’ 과제는 전남 지역에서 사용하는 개별 수어의 다양한 표현을 유도하기 위해 ‘김치’, ‘가래떡’ 자료를 추가하였다.



[그림 II-14] 전남지역에서 사용하는 개별 수어 유도 자료

J-농사회 행사는 전남 지역에서 개최된 행사 포스터로 교체하였다.

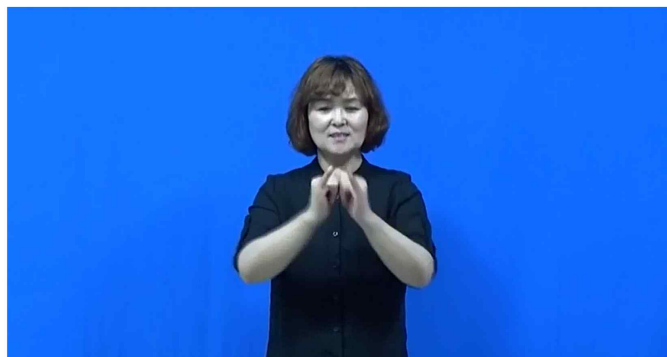


[그림 II-15] 전남지역에서 개최된 행사 포스터

(2) 유도 자료 설명 동영상 제작

전남 지역의 사회자가 2019년 유도 자료 설명 동영상을 참고하여 13개의 과제에 대한 유도 자료 설명 동영상을 촬영하였으며 연구진에서 기존의 틀과 동일한 형식으로 편집하였다. 두 명의 언어제공자가 확인하는 PPT는 관람하는 쪽과 설명 영상을 확인하는 쪽으로 구분되므로 다시 이야기하기 과제(애니메이션, 넥타이)와 흥부와 놀부 이야기 과제만 상이하고 모두 동일하게 제작되었다. 아래는 유도 자료 PPT 파일에 삽입된 유도 자료 설명 동영상의 한 장면이다.

11. 농사회 행사



[그림 II-16] 사회자의 유도 자료 설명 동영상 예시 화면

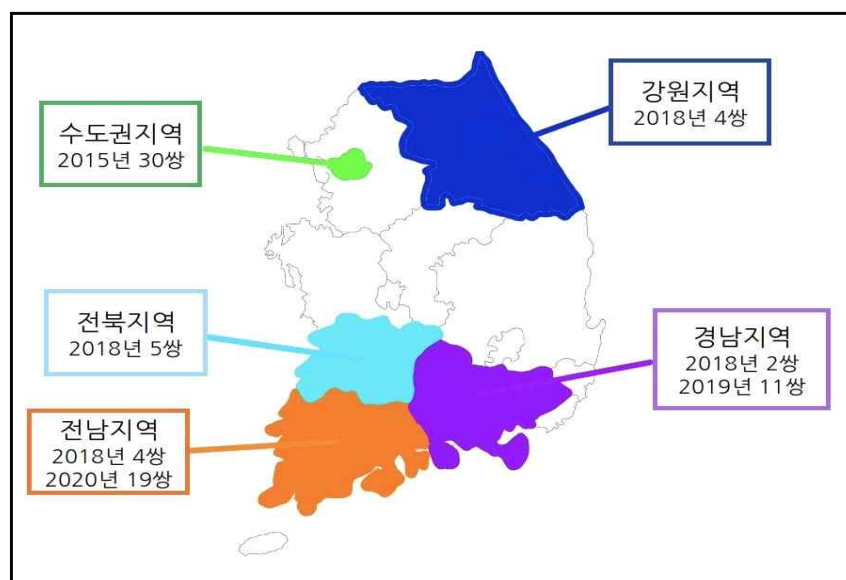
3. 언어제공자 모집 및 선정

1) 조사 지역 선정

자연수어 연구를 위한 말뭉치 연구는 실제 사용된 언어 자료를 다양하게 많이 수집하는 것이 필요하다. 특히 언어의 빈도와 분포를 확인하기 위해 다양한 지역을 대상으로 한 언어자료 수집은 필수적이다. 한 회기의 수집에서 여러 지역을 대상으로 수집하는 경우 수집된 수어 표본의 대표성이 떨어질 수 있으며 수집 작업의 효율성이 떨어진다.

실천적 측면에서 언어자료의 수집은 해당 지역에서 이루어져야 한다. 여러 지역에서 수집을 할 경우에는 여러 지역에 촬영실을 확보해야 하는 어려움과 함께 그 지역을 대표하는 유의미한 수어 언어제공자를 확보하는 데 어려움이 있다. 하지만 의미 있는 언어자료를 수집하기 위해서 해당 지역에서 최소 10쌍 이상의 토박이 수어 언어제공자로부터 원시 수어자료를 수집하는 것이 바람직하다.

이에 본 사업에서는 2018년에 4쌍을 수집한 전남 지역(광주 포함)을 대상으로 19쌍으로부터 언어자료를 추가 수집하였다. 이번 전남지역을 포함한 현재까지의 원시 말뭉치 수집 현황은 다음과 같다.



[그림 II-17] 한국수어 말뭉치 언어자료 수집 현황

2) 언어제공자 선정 조건

한국수어 말뭉치 구축을 위한 언어제공자의 적합성 확보를 위해 연구진과 전라남도 농아인협회의 회의를 통해 기본 조건 충족 여부를 점검하였다.

〈표 II-4〉 언어제공자 선정 조건

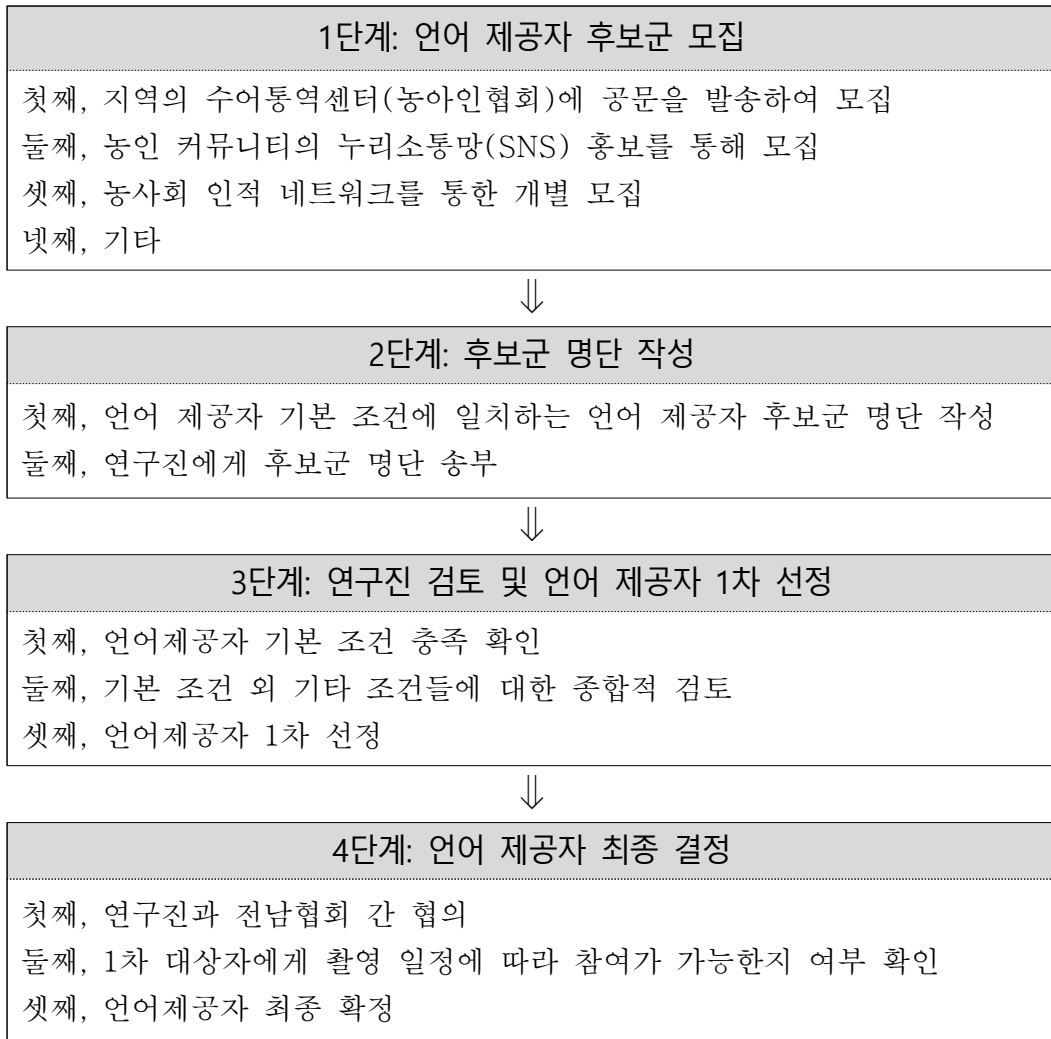
번호	조 건	가부
1	· 19세 이상 성인 농인	(예, 아니오)
2	· 한국수어를 모어로 하는 농인 - 의사소통을 위해 수어를 가장 편하고 자주 사용함	(예, 아니오)
3	· 농학교 졸업 ※ 부모가 모두 농인이면 일반학교 졸업한 농인도 가능	(예, 아니오)
4	· 현재 전남지역에서 10년 이상 거주 혹은 직장생활을 하고 있음	(예, 아니오)
5	· 주 3회 이상 농인을 만남	(예, 아니오)

이때, 언어제공자 선정의 우선 조건은 첫째, 농부모 혹은 농형제, 농배우자가 있는 농인이거나 둘째, 초등학교 이전에 수어를 습득한 농인으로 하였다.

3) 언어제공자 모집

23개 시·군 지역(광주포함)으로 이루어진 전남 지역의 언어제공자 모집은 전라남도 농아인협회의 협조로 진행되었다.

언어제공자 모집을 위해서는 후보군 모집, 후보군 명단 작성, 후보군 검토 및 언어제공자 1차 선정, 언어제공자 최종 결정의 4단계 과정으로 다음과 같이 진행하였다.



[그림 II-18] 언어제공자 모집 및 선정 절차

4) 언어제공자 기본 사항

일정 조율 등으로 최종 확정되어 본 연구에 참여한 언어제공자의 기본 사항은 다음과 같다.

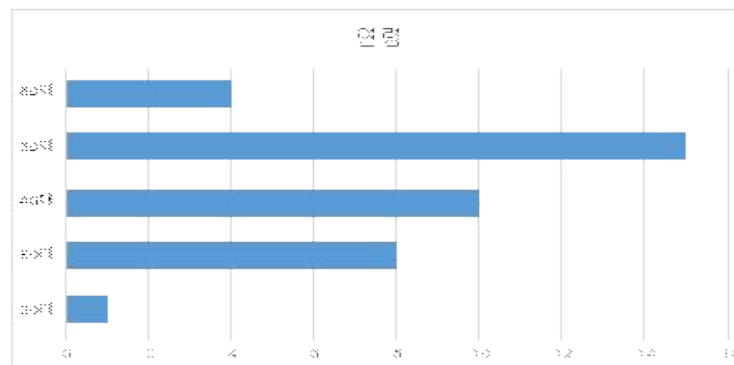
<표 II-5> 언어제공자의 기본 사항

지역	성명	성별	출생년도	지역	성명	성별	출생년도
전남	문현○	여	1965	광주	박금○	여	1988
전남	서덕○	여	1969	전남	곽을○	남	1965
전남	정준○	남	1957	전남	김경○	남	1974
전남	추영○	남	1968	전남	기정○	남	1962

광주	이원○	남	1990	전남	이명○	남	1966
광주	김신○	여	2000	전남	이○	남	1964
전남	김호○	남	1982	전남	정영○	남	1977
전남	전영○	남	1980	전남	이승○	남	1966
전남	이명○	남	1960	전남	이철○	남	1963
전남	전희○	남	1952	전남	황진○	남	1972
전남	백성○	남	1959	전남	조재○	남	1962
전남	김영○	남	1969	광주	김상○	여	1969
전남	김○	남	1981	광주	전영○	여	1972
전남	백성○	남	1973	전남	김여○	여	1972
광주	최태○	남	1990	전남	남기○	남	1966
광주	권오○	남	1990	전남	김공○	남	1988
전남	양효○	여	1968	전남	최동○	남	1987
전남	김정○	여	1976	전남	박미○	여	1977
광주	우찬○	남	1989	전남	김미○	여	1968

(가) 언어제공자 인구학적 정보

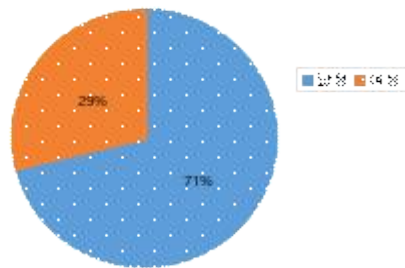
① 연령



[그림 II-19] 언어제공자의 연령 분포

언어제공자 중에는 50대가 15명으로 가장 많았고, 40대가 10명, 30대가 8명, 그리고 60대 4명, 20대 1명 순이었다.

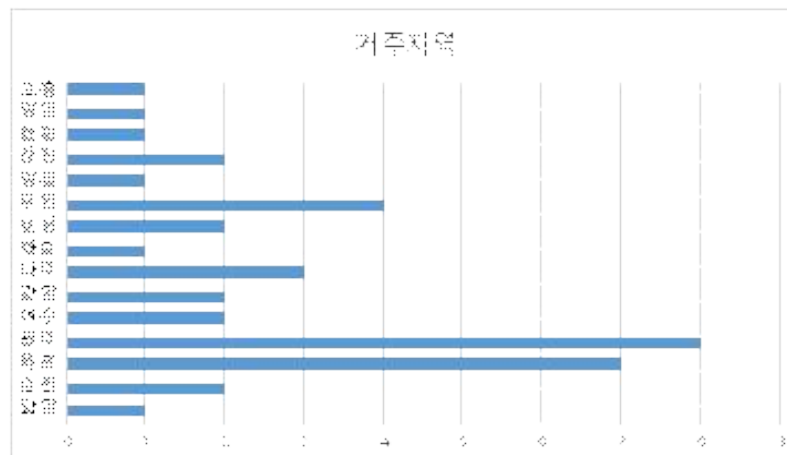
② 성별



[그림 II-20] 언어제공자의 성별

언어제공자 중에는 남성이 71%(27명), 여성이 29%(11명)였다.

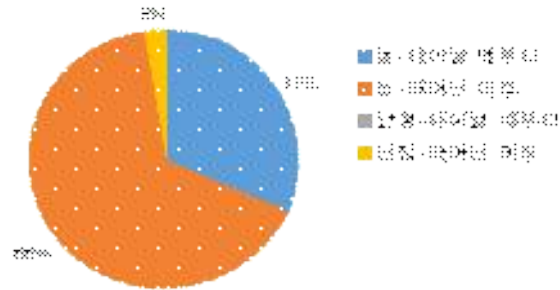
③ 거주 지역



[그림 II-21] 언어제공자의 거주 지역 현황

언어제공자의 거주 지역으로는 광주가 8명으로 가장 많았고, 목포 7명, 무안 4명, 나주 3명, 보성, 광양, 여수, 순천, 강진이 각 2명, 그리고 함평, 영광, 화순, 담양, 영암, 고흥이 각 1명이었다.

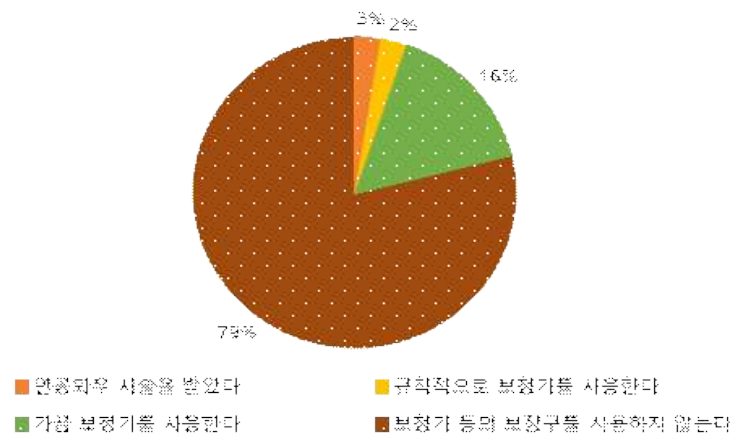
④ 청력상태/청력손실 시기



[그림 II-22] 언어제공자의 청력상태/청력손실 시기

언어제공자 중 태어난 이후 농인이 되었다는 응답이 66%(25명)로 나타났다, 태어날 때부터 농인이었다는 응답이 31%(12명)로 나타났다. 또한 태어난 이후 난청이 되었다는 응답이 3%(1명)로 나타났다.

⑤ 청각보조기기 착용 여부



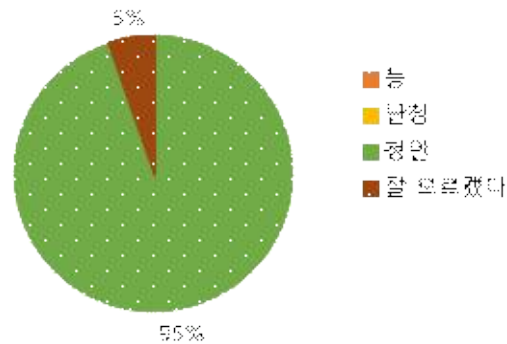
[그림 II-23] 언어제공자의 청각보조기기 착용 여부

청각보조기기 착용 여부에 관하여 ‘보청기 등의 보장구를 사용하지 않는다’ 는 답변이 79%(30명)로 가장 많았고, ‘가끔 보청기를 사용한다’

는 답변이 16%(6명), ‘인공와우 시술을 받았다’는 답변이 3%(1명), ‘규칙적으로 보청기를 사용한다’는 답변이 2%(1명)로 나타났다.

(나) 언어제공자 가족 정보

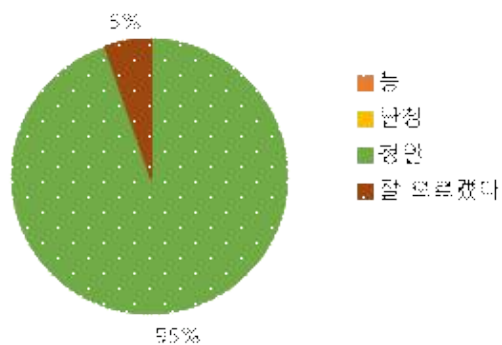
⑥ 아버지의 농인 여부



[그림 II-24] 언어제공자 아버지의 농인 여부

언어제공자의 아버지가 청인인 경우가 95%(36명)로 압도적으로 많았고, ‘잘 모르겠다’고 답변한 경우가 5%(2명)로 나타났다.

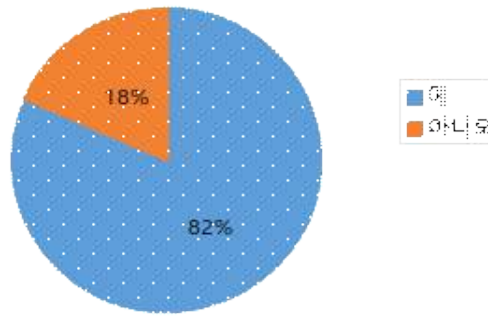
⑦ 어머니의 농인 여부



[그림 II-25] 언어제공자 어머니의 농인 여부

언어제공자의 어머니가 청인인 경우가 아버지의 경우와 마찬가지로 95%(36명)로 압도적으로 많았고, ‘잘 모르겠다’고 답변한 경우가 5%(2명)로 나타났다.

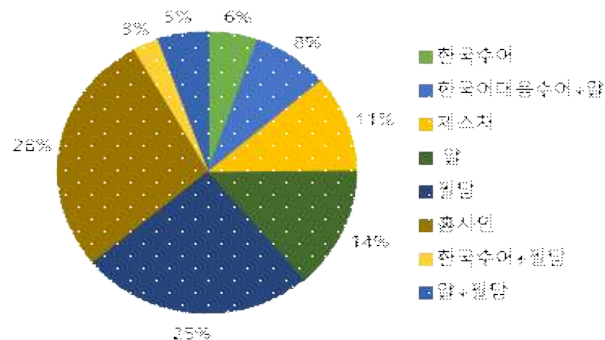
⑧ 부모와의 동거 여부(어린 시절)



[그림 II-26] 언어제공자와 부모와의 동거 여부

언어제공자 중 어린 시절 부모와 동거하였다고 답변한 경우가 82%(31명), 동거하지 않았다고 답변한 경우가 18%(7명)로 나타났다.

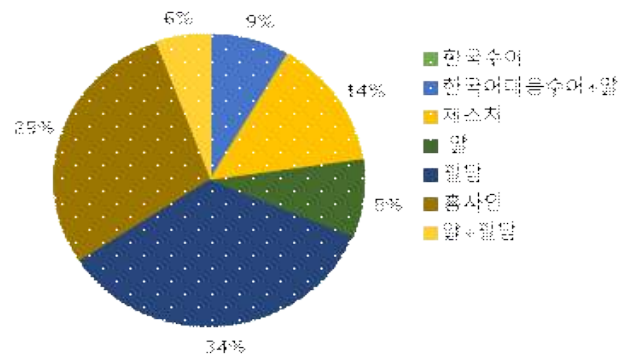
⑨ 어머니와의 주된 의사소통 방법



[그림 II-27] 언어제공자와 어머니와의 주된 의사소통 방법

언어제공자 중 어머니와의 주된 의사소통 방법으로 홈사인을 사용한다는 답변이 28%(10명), 필담을 사용한다는 답변이 25%(9명)로 나타났으며, 이어 말을 사용한다는 답변이 14%(5명), 제스처가 11%(4명), 그 외 의사소통 방법을 혼용하여 사용한다는 답변이 16%(6명), 한국수어를 사용한다는 답변이 6%(2명)으로 나타났다.

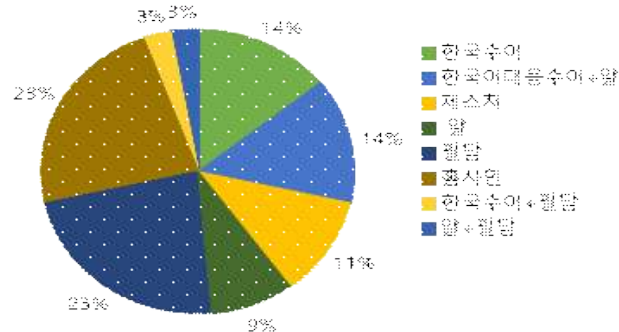
⑩ 아버지와의 주된 의사소통 방법



[그림 II-28] 언어제공자와 아버지와의 주된 의사소통 방법

언어제공자 중 아버지와의 주된 의사소통 방법으로 필담을 사용한다는 답변이 34%(12명), 홈사인을 사용한다는 답변이 29%(10명)로 나타났으며, 제스처를 사용한다는 답변이 14%(5명), 말을 사용한다는 답변이 8%(3명)로 나타났다. 그 외 의사소통 방법을 혼용하여 사용한다는 답변이 15%(5명)로 나타났다.

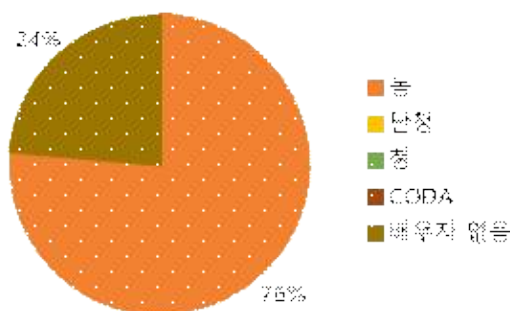
⑪ 형제 및 자매들과의 주된 의사소통 방법



[그림 II-29] 언어제공자와 형제 및 자매들과의 주된 의사소통 방법

언어제공자 중 형제 및 자매들과의 주된 의사소통 방법으로 필담, 홈사인을 사용한다는 답변이 각 23%(8명), 한국수어를 사용한다는 답변이 14%(5명)로 나타났으며, 제스처를 사용한다는 답변이 11%(4명) 말을 사용한다는 답변이 9%(3명)로 나타났다. 그 외 의사소통 방법을 혼용하여 사용한다는 답변이 20%(7명)로 나타났다.

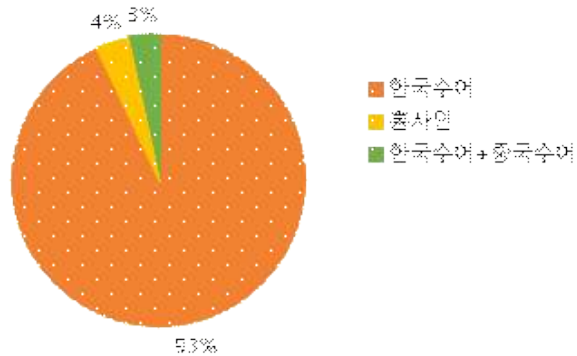
⑫ 배우자의 농인 여부



[그림 II-30] 언어제공자 배우자의 농인 여부

언어제공자의 배우자가 농인인 경우가 76%(29명), 배우자가 없는 경우가 24%(9명)로 나타났다.

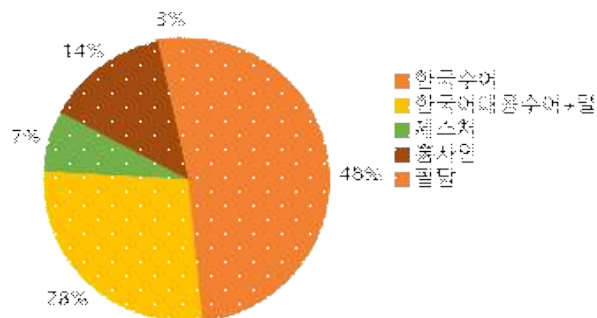
⑬ 배우자와의 주된 의사소통 방법



[그림 II-31] 언어제공자와 배우자와의 주된 의사소통 방법

언어제공자 중 배우자와의 주된 의사소통 방법으로 한국수어를 사용한다는 답변이 93%(26명)로 나타났다. 말사인을 사용한다는 답변이 4%(1명), 한국수어와 중국수어를 혼용하여 사용한다는 답변이 3%(1명)로 나타났다.

⑭ 자녀와의 주된 의사소통 방법



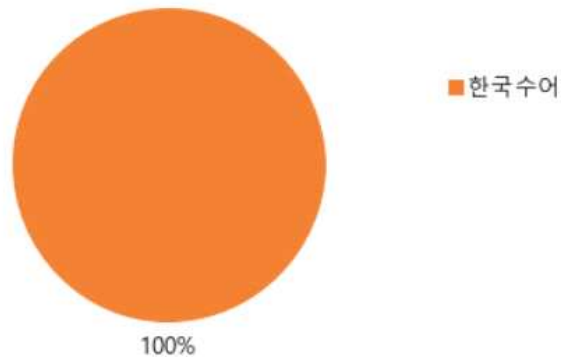
[그림 II-32] 언어제공자와 자녀와의 주된 의사소통 방법

언어제공자 중 자녀와의 주된 의사소통 방법으로 한국수어를 사용한다

는 답변이 48%(14명), 홈사인을 사용한다는 답변이 14%(4명), 한국어대응수어와 음성언어를 혼용한다는 답변이 28%(8명)로 나타났다. 또한 제스처를 사용한다는 답변이 7%(2명), 필담을 사용한다는 답변이 3%(1명)으로 나타났다.

(다) 언어제공자 언어 정보

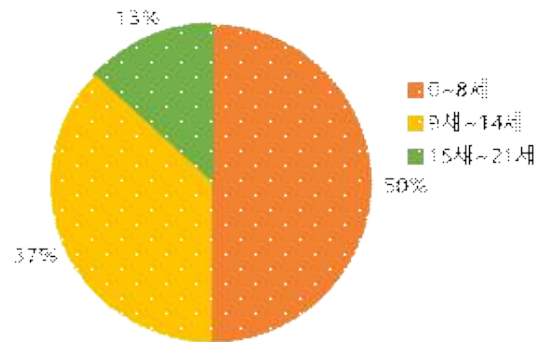
⑮ 언어 제공자(본인)의 주된 의사소통 방법



[그림 II-33] 언어제공자(본인)의 주된 의사소통 방법

언어제공자 중 자신의 주된 의사소통 방법으로 한국수어를 사용한다는 답변이 100%(38명)로 나타났다.

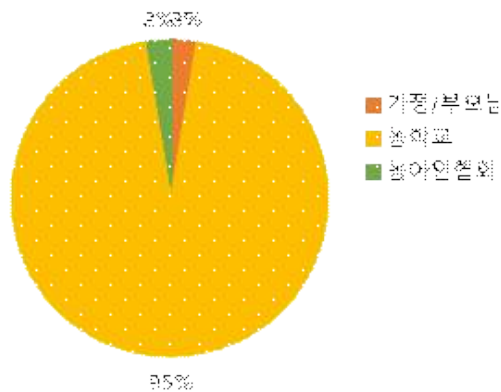
⑯ 한국수어 습득 시기



[그림 II-34] 언어제공자의 한국수어 습득 시기

언어제공자 중 한국수어 습득 시기에 대한 답변으로 0세부터 8세가 50%(19명), 9세에서 14세가 37%(14명), 15세에서 21세가 13%(5명)로 나타났다.

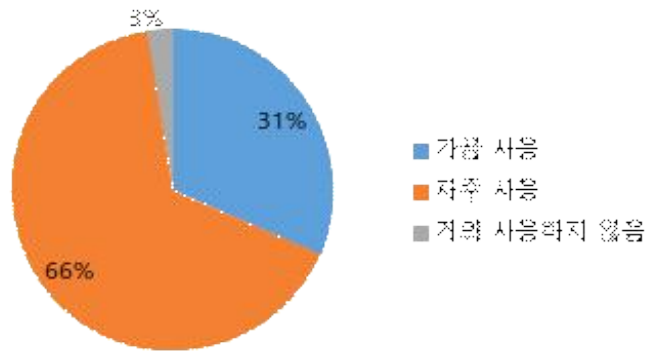
⑰ 한국수어 습득 장소



[그림 II-35] 언어제공자의 한국수어 습득 장소

언어제공자의 한국수어 습득 장소에 대한 답변으로 농학교가 95%(36명)로 압도적으로 많았고, 가정/부모님이 3%(1명), 농아인협회가 2%(1명)로 나타났다.

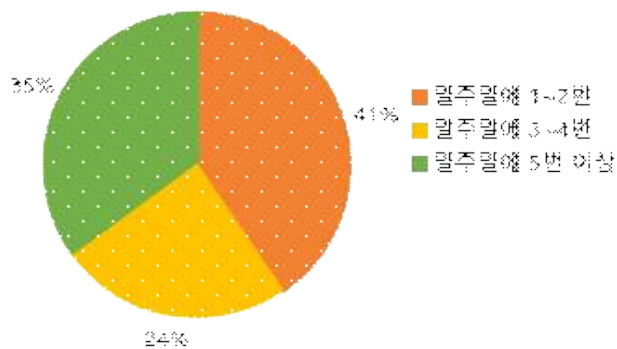
⑱ 한글 지문자(지화) 사용 유무



[그림 II-36] 언어제공자의 한글지문자(지화) 사용 유무

언어제공자의 한글 지문자 사용에 대한 답변으로 자주 사용이 66%(25명), 가끔 사용이 31%(12명). 거의 사용하지 않음이 3%(1명)으로 나타났다.

⑬ 농(난청)인과의 정기적인 만남 여부

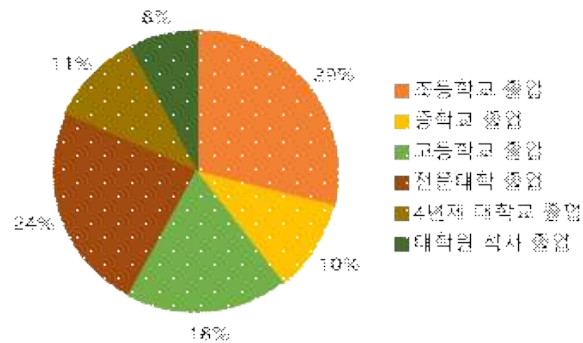


[그림 II-37] 언어제공자의 농(난청)인과의 정기적인 만남 여부

농(난청)인과의 정기적인 만남 여부에 대해 일주일에 1~2번 이상 만난다는 답변이 41%(15명), 일주일에 5번 이상 만난다는 답변이 35%(13명), 일주일에 3~4번 만난다는 답변이 24%(9명)로 나타났다.

(라) 언어제공자 학력 및 직업

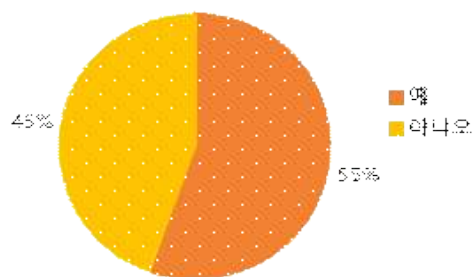
⑳ 최종 학력



[그림 II-38] 언어제공자의 최종 학력

언어제공자의 최종 학력에 대한 답변으로 초등학교 졸업이 29%(11명)로 나타났으며, 전문대학 졸업이 24%(9명), 고등학교 졸업이 18%(7명), 중학교 졸업이 11%(4명), 4년제 대학교 졸업이 10%(4명), 대학원 석사 졸업이 각각 8%(3명)로 나타났다.

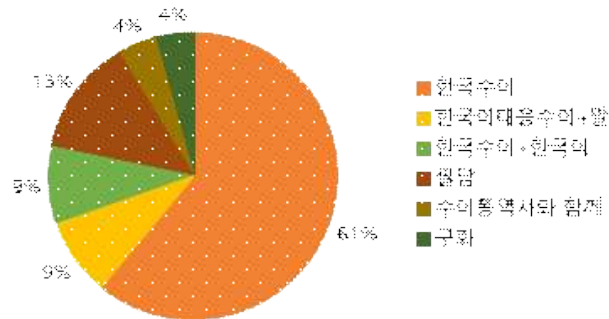
㉑ 재직 여부



[그림 II-39] 언어제공자의 재직 여부

언어제공자 중 현재 직장에 다니고 있다고 답변한 경우가 55%(21명), 현재 다니고 있지 않다고 답변한 경우가 45%(17명)로 나타났다.

② 직장에서의 주된 의사소통 방법



[그림 II-40] 언어제공자의 직장에서의 주된 의사소통 방법

직장에서의 주된 의사소통 방법으로는 한국수어가 61%(14명)로 가장 높게 나타났으며, 필답이 13%(3명), 한국어대응수어와 음성언어, 한국수어와 한국어를 혼용한다는 답변이 각각 9%(2명)로 나타났다. 그 외 수어통역사가 6%(1명), 구화를 사용한다는 답변이 5%(1명)로 나타났다.

5) 언어제공자 질문지

전남의 지역적 특성상 고령의 언어제공자가 많아 인터넷을 통한 질문지 작성과 영상으로 제공되는 질문지를 보고 한글로 작성하는 방식에 어려움이 있어 영상을 보고 언어제공자가 직접 질문지를 작성하는 방식, 농아인협회 농인 직원이 질문지의 내용을 수어로 설명하고 언어제공자가 수어로 답하는 방식 중에서 자유롭게 선택하도록 하였다(부록 4).

6) 촬영 동의

촬영 전 촬영 영상에 대한 이용 동의를 구하기 위하여 사전에 서면으로 제작된 ‘국가 언어 자원(말뭉치) 구축 및 활용 저작권 등 이용

허락 계약서(부록 3)’와 수어 영상물을 언어제공자에게 보여주고 내용에 동의할 시 서명을 하도록 하였다. 서명을 마친 언어제공자는 촬영 초반에 사회자에게 해당 내용 촬영에 수어로 동의 의사를 표시하도록 하였다.

4. 사회자 교육

1) 사회자의 역할

- 언어제공자에게 파워포인트로 작성된 유도 자료를 제시한다.
- 설명 동영상만으로 언어제공자가 과제를 제대로 이해하지 못했을 경우 과제에 대한 설명을 한다.
- 과제 소요시간을 조절한다(예, 시계 활용).
- 대화가 자연스럽게 진행될 수 있는 분위기를 조성한다.
- 언어제공자의 대화를 중재하고 촉진한다.
- 사회자 일지를 기록한다.

2) 사회자 선정 및 교육

사회자는 광주를 포함한 전남지역에 대한 언어수집을 위해 전남지역과 광주지역에서 각 1명(이00, 나00)을 선정하였다.

- 사회자 교육 일정: 총 2회 실시
 - 일시
 - 1차: 7/10(금) 9:30~17:30
 - 2차: 7/20(월) 8:30~12:00
 - 장소: 전라남도농아인협회 부설 전남수어교육원

○ 교육 내용

(1차)

- 수어 말뭉치와 유도 자료의 내용 및 목적에 대한 이해
- 언어제공자가 동영상 설명 자료를 이해하지 못할 시 필요한 보충 설명 방법
- 언어제공자가 자연스러운 분위기에서 편하게 대화할 수 있는 분위기 조성 방법
- 두 언어제공자의 수어가 균형 있게 산출될 수 있도록 조절하는 전략
- 그 외 사회자가 지켜야 할 태도
- 언어제공자와 사회자 관계 질문지, 사회자 일지 작성 방법
- 유도 자료 설명 동영상 촬영 시 고려사항 및 촬영 진행 방법



[그림 II-41] 사회자 교육 및 유도 자료 설명 동영상 촬영 연습

(2차)

- [부록 5] 언어제공자와 사회자 관계 질문지 및 [부록 6] 사회자 일지 작성 방법
- 각 과제별 적정 진행 시간 및 조절 전략

3) 유도과제 설명 동영상 제작

- 동의서 및 13개의 과제를 영상으로 촬영하였다.
- 유도 자료 설명 영상 촬영 시에는 사회자 교육 담당자의 서울지역 방언 영향을 방지하기 위하여 촬영장 접근을 금지하였다.
- 촬영된 영상을 기존 양식에 맞게 편집하였다.

5. 언어 수집 촬영

1) 촬영실 조건 및 촬영 시스템 지침

언어 수집 촬영을 위한 촬영실 조건과 촬영 시스템에 관한 지침은 다음과 같다.

- ① 촬영실의 크기는 32m² (5.3m x 6m) 이상
- ② 촬영실 창문은 빛을 차단할 수 있는 커튼이나 암막 설치
- ③ 촬영 중 청인은 촬영실에 들어갈 수 없으므로 촬영 감독과 보조 인력이 촬영을 준비하고 모니터 할 수 있는 별도의 방 확보
- ④ 촬영실은 언어제공자들이 오기 편하고 연구진들이 24시간 접근이 가능한 곳에 위치
- ⑤ 두 언어제공자를 각각 촬영할 수 있는 카메라 2대와 두 언어제공자와 사회자를 다 촬영할 수 있는 카메라 1대 등 총 3대 이상의 카메라 비치
- ⑥ 128G 이상의 카메라 메모리 용량 사전 확보
- ⑦ 카메라의 전원 연결이 중간에 끊어지지 않도록 안전장치 확보

- ⑧ 카메라는 언어제공자들이 방해를 받지 않고 대화를 할 수 있도록 최대한 눈에 띄이지 않게 설치
- ⑨ 촬영을 위한 조명, 유도 자료 제시를 위한 컴퓨터와 모니터는 다음 그림과 같이 설치

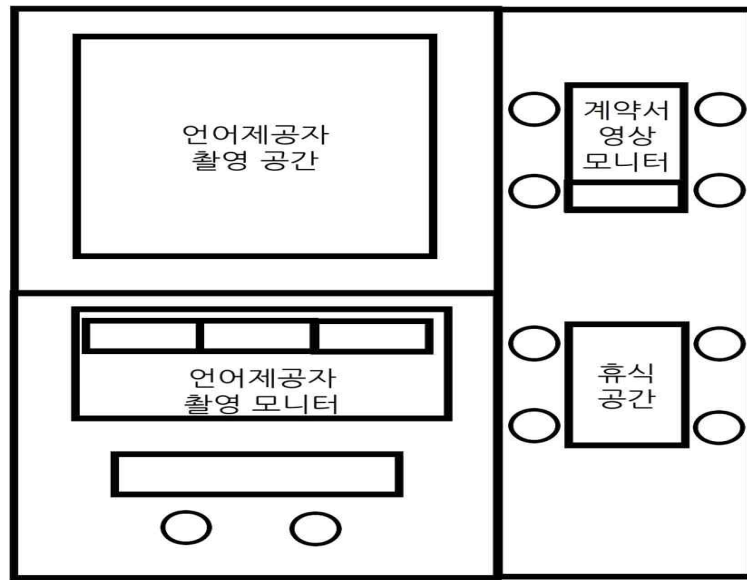


[그림 II-42] 촬영실 모습

2) 촬영장 설치

위에 제시된 지침에 따라 설치한 촬영실과 촬영 시스템에 관한 사항은 다음과 같다.

- 촬영장은 언어제공자 촬영 공간과 휴게 공간으로 구분
- 촬영장 배치도와 동일하게 촬영장 구축



[그림 II-43] 촬영장 전체 배치도

(1) 언어제공자 촬영실

- 촬영실 구비 장비는 카메라 3대, 조명 5대, 노트북 3대, 모니터 5대로 구성하였다.
- 촬영실의 모습은 다음과 같다.



[그림 II-44] 언어제공자 촬영실 배치도

(2) 휴게 공간

① 모니터 공간

- 3대의 카메라와 연결된 모니터를 설치하였다.
- 언어제공자 촬영 모습이 모니터링 되도록 하였다.



[그림 II-45] 언어제공자 모니터 공간

② 휴식 공간

- [부록 3]의 촬영 동의서, 질문지 비치해 둔 테이블과 의자가 있다.
- 계약서 및 질문지의 수어 영상 버전을 시청할 수 있도록 노트북을 비치하였다.
- 휴식 시간에 섭취할 수 있는 다과를 준비해 두었다.



[그림 II -46] 언어제공자 휴식 공간

3) 언어제공자 촬영

○ 촬영 시에는 언어제공자 및 사회자 외에는 출입을 하지 않았다.

(1) 중앙(Md)



(2) 오른쪽(Rt) 언어제공자 촬영 카메라 & 왼쪽(Lt)언어제공자



(3) 왼쪽(Lt)언어제공자 촬영 카메라 & 오른쪽(Rt) 언어제공자



4) 촬영 지침 개선 사항

촬영 시 협소한 공간으로 인해 촬영에 불편을 겪었고, 카메라가 전원에 연결되는 형태가 아닌 배터리를 사용하는 방식으로 배터리 용량 부족으로 촬영을 중단하거나 카메라 메모리가 호환되지 않아 문제가 발생하여 이에 따라 지침을 수정하였다.

수정 전	수정후
<ul style="list-style-type: none"> ① 촬영실의 크기는 32m²(5.3m x 6m) 이상 ② 촬영실 창문은 빛을 차단할 수 있는 커튼이나 암막 설치 ③ 촬영 중 청인은 촬영실에 들어갈 수 없으므로 촬영 감독과 보조 인력이 촬영을 준비하고 모니터 할 수 있는 별도의 방 확보 ④ 촬영실은 언어제공자들이 오기 편하고 연구진들이 24시간 접근이 가능한 곳에 위치 ⑤ 두 언어제공자를 각각 촬영할 수 있는 카메라 2대와 두 언어제공자와 사회자를 다 촬영할 수 있는 카메라 1대 등 총 3대 이상의 카메라 비치 ⑥ 카메라의 메모리 용량은 128G 이상 사전 확보 ⑦ 카메라의 전원 연결이 중간에 끊어지지 않도록 안전장치 확보 ⑧ 카메라는 언어제공자들이 방해받지 않고 대화를 할 수 있도록 최대한 눈에 띄이지 않게 설치 ⑨ 촬영을 위한 조명, 유도 자료 제시를 위한 컴퓨터와 모니터는 다음 그림과 같이 설치 	<ul style="list-style-type: none"> ① 촬영실의 크기는 39m²(6.3m x 6m) 이상 ② ~ ⑤ 좌동 ⑥ 카메라의 메모리 용량은 128G 이상 사전 확보, 카메라와 호환 가능 ⑦ 카메라를 전원에 연결하여 사용하며 카메라의 전원 연결이 중간에 끊어지지 않도록 안전장치 확보 ⑧ ~ ⑨ 좌동

5) 촬영 지침

언어제공자의 전남지역 토박이 수어 사용에 영향을 줄 수 있으므로 촬영 전과 촬영 중 쉬는 시간에는 가급적 언어제공자와 청인, 타 지역 농인(연구팀) 간에 대화를 나누지 않도록 하였다.

(1) 촬영 전

사회자	<ol style="list-style-type: none"> 언어제공자와 사회자 간의 관계에 관한 질문지 작성 준비물 확인 <ul style="list-style-type: none"> - 사회자 일지, 타임벨, 탁상시계, 핸드폰, 펜, 물병, 색종이
촬영자	<ol style="list-style-type: none"> 촬영실의 조명과 에어컨을 켜고 햇빛이 들어오거나 기타 문제 사항이 없는지 확인 촬영 준비 <ol style="list-style-type: none"> 1) 컴퓨터 및 모니터 정상 작동 여부 2) 카메라 정상 작동 여부 3) 언어제공자가 카메라 안에 다 들어오는지 확인(언어제공자마다 의자의 높이를 조절하거나 카메라의 각도를 조절해야 함)
협회 담당자	<ol style="list-style-type: none"> 당일 촬영할 언어제공자에게 문자나 영상으로 연락 <ul style="list-style-type: none"> - 촬영 시작 시간 안내 및 촬영 시작 1시간 전 도착 당부 - 촬영실 오기 전 통장, 신분증을 지참할 것을 안내 - 오시는 길 안내 음료와 다과 준비 언어제공자가 오면 대기실로 안내하여 관련 서류를 받음 <ol style="list-style-type: none"> 1) 국가 언어 자원(말뭉치) 구축 및 활용 저작권 등 이용 허락 계약서 2) 개인 정보 수집, 이용, 제공 및 조회 동의서 3) 자료 수집을 위한 질문지 4) 통장 사본, 신분증 사본
연구원	<ol style="list-style-type: none"> 유도 자료 정상 작동 여부 확인 언어제공자의 언어 사용 양상 파악 후 A(대화 나누는 것을 좋아하고 상황에 대한 묘사를 잘하는지), B(다른 사람의 이야기에 집중하는 것을 즐기는지)로 배치 기타 제반 사항 확인

(2) 촬영 중

사회자	<ol style="list-style-type: none"> 1. 촬영자가 준비가 완료되었는지 확인 2. 언어제공자가 준비가 완료되었는지 확인 3. 촬영자에게 촬영 준비가 완료되었음을 알림 4. 유도 자료를 제시함 5. 사회자 일지를 작성함
촬영자	<ol style="list-style-type: none"> 1. 3개 카메라의 녹화 버튼을 동시에 눌러 촬영 시작
협회 담당자	<ol style="list-style-type: none"> 1. 촬영 중에는 청인이 촬영 공간 안에 들어갈 수 없지만 촬영 공간에 들어가야 할 돌발 상황에 대처하기 위해 대기
연구원	<ol style="list-style-type: none"> 1. 촬영실 바깥에 있는 모니터로 녹화가 잘 되고 있는지 수시로 확인 2. 촬영일지 작성 3. 자료 수집을 위한 질문지가 잘 작성되었는지 확인하고 누락 된 답변 메모 4. 사회자 진행에서의 보완사항 확인

(3) 촬영 후

사회자	<ol style="list-style-type: none"> 1. 촬영자에게 촬영 종료를 알림 2. 작성한 사회자 일지를 연구원에게 전달
촬영자	<ol style="list-style-type: none"> 1. 3개의 카메라의 녹화 버튼을 눌러 촬영 종료
협회 담당자	<ol style="list-style-type: none"> 1. 언어제공자 영수증에 서명을 받음
연구원	<ol style="list-style-type: none"> 1. 자료 수집을 위한 질문지에서 누락 된 답변을 질문하여 보완 작성 2. 사회자 일지를 복사하여 촬영자에게 전달(편집 작업에 필요) 3. 사회자 진행에서의 보완사항 공유 4. 언어제공자에게 말뭉치 수집의 목적에 대해 설명

6) 촬영 시간

○ 언어제공자의 기질, 낯선 환경 등의 이유로 한 쌍의 언어제공자의

언어자료 수집을 위한 과제별 최소 촬영 시간 지침을 마련하였다.

과제명	A	B	C	D	E	F	G	H	I	J	K	L	M
최소시간(분)	5	10	10	10	25	20	30	25	10	10	10	10	15

- 촬영 시간에는 동의서부터 M. 농학교 경험 과제까지, 그리고 유도 자료를 설명하는 부분까지 모두 포함되며, 전남지역 말뭉치 수집을 위해 촬영한 시간은 <표 II-6>과 같다.

<표 II-6> 전남지역 말뭉치 전체 촬영 시간

언어 제공자	5번	6번	7번	8번	9번	10번	11번	12번	13번	14번	15번	16번	17번	18번	19번	20번	21번	22번	23번	총 계
촬영 시간	02:46:46	03:04:40	03:09:17	03:12:34	02:59:51	02:51:03	02:53:28	02:51:23	03:02:05	06:51:42	03:25:17	03:06:49	03:02:22	02:48:54	03:07:26	03:43:56	03:35:19	03:56:15	03:23:54	59:53:01

- 편집은 전체 촬영 분량 중에서 유도 자료 설명 시간, 기타 과제와 상관없는 발화 등을 제외하여 과제별로 분할한다. 편집 시간은 과제 내에서 발화한 시간만을 포함하기 때문에 전체 촬영 시간(59시간 53분 1초)과는 차이가 있다.

<표 II-7> 전남지역 말뭉치 과제별 편집 시간

과제 코드	5번 쌍	6번 쌍	7번 쌍	8번 쌍	9번 쌍	10번 쌍	11번 쌍	12번 쌍	13번 쌍	14번 쌍	15번 쌍	16번 쌍	17번 쌍	18번 쌍	19번 쌍	20번 쌍	21번 쌍	22번 쌍	23번 쌍	총 계
동의서	00:23	00:26	00:25	00:12	00:29	00:21	05:12	00:31	00:30	00:26	00:20	00:28	00:22	00:25	00:25	00:26	00:22	00:27	00:25	12:35
A	02:05	07:10	03:03	04:13	08:14	04:47	05:00	02:43	01:08	04:25	04:20	02:18	06:27	04:46	02:08	02:57	04:43	03:13	02:28	1:16:08
B	02:40	04:50	07:29	06:25	04:01	06:08	04:54	03:37	02:05	04:01	09:57	06:53	04:55	07:44	04:00	05:59	14:55	20:22	07:00	2:07:55
C	14:49	10:59	12:22	21:16	19:39	18:36	21:38	17:40	19:18	11:46	17:04	21:58	20:04	21:37	22:41	23:57	24:48	30:20	27:06	6:17:38
D	05:28	03:51	06:08	08:14	04:44	11:51	08:57	08:00	09:15	08:16	08:46	07:32	07:37	12:51	08:33	07:47	14:41	08:57	06:53	2:38:21

E	04:58	11:26	16:43	24:30	15:48	12:36	07:13	16:53	09:12	12:20	16:21	18:50	22:49	07:02	20:25	26:05	22:27	34:21	16:12	5:16:11
F	12:40	10:38	18:59	18:20	13:32	16:36	17:09	11:34	10:16	14:18	18:47	13:46	16:34	11:01	08:47	17:39	09:41	25:12	13:02	4:38:31
G	18:01	17:28	13:53	22:38	12:05	28:21	14:04	12:12	11:48	08:38	18:55	17:16	18:15	15:34	23:11	33:15	26:11	22:12	21:09	5:55:06
H	10:32	08:12	02:52	09:22	10:58	20:16	06:10	11:50	12:23	09:02	13:11	11:42	07:12	10:30	16:59	17:00	25:44	16:01	14:08	3:54:04
I	07:46	09:38	11:43	09:02	10:32	06:50	12:09	10:08	06:00	08:03	05:55	07:52	08:35	08:13	04:45	10:57	04:11	07:48	06:37	2:36:44
J	08:48	07:20	03:58	06:20	08:56	07:24	07:47	07:06	07:26	07:54	13:00	09:07	04:30	05:51	07:56	07:02	04:16	06:05	05:33	2:16:19
K	04:39	03:44	04:38	06:36	05:26	08:08	08:03	05:13	07:15	03:08	07:11	02:11	08:39	10:44	14:07	06:54	13:44	04:03	02:25	2:06:48
L	07:40	15:52	14:26	07:32	14:38	10:02	08:03	05:03	17:45	12:39	19:14	09:37	12:22	08:47	10:13	13:09	10:42	08:56	08:46	3:35:26
M	31:23	28:04	41:12	13:10	19:11	11:26	01:56	23:55	33:13	46:45	19:50	24:18	13:47	13:10	11:13	20:00	10:13	05:38	24:49	6:33:13
총계	2:11:52	2:19:38	2:37:51	2:37:50	2:28:13	2:43:32	2:08:15	2:16:25	2:22:7	2:37:14	2:51:1	2:33:8	2:33:8	2:18:5	2:33:3	3:13:7	3:06:8	3:13:5	2:36:3	49:24:59
총시간	2,964분 53초 (총 49시간 24분 59초)																			

7) 자료 수집을 위한 촬영 시 고려 사항

- 전남 지역의 경우 어업, 농업과 관련된 직업에 종사하는 언어제공자가 종종 있었으며, 그로 인해 어업, 농업과 관련된 수어가 도출될 것으로 기대하였으나 다른 직업군의 언어제공자와 쌍을 이룰 경우 해당 수어가 도출되는 경우가 극히 드물었다. 향후 가능한 같은 직업군을 가진 언어제공자를 한 쌍으로 배치하여 직업과 관련된 다양한 수어를 유도할 필요가 있다.
- 카메라, 메모리, 배터리 등의 문제로 촬영 진행에 있어 어려움이 없도록 서면 지침 전달과 더불어 연구진이 직접 진행하는 교육이 필요하다.
- 유도 자료의 경우 ‘고모’, ‘이모’ 등의 한국어 단어로 제작된 유도 자료로 제시하였는데, 이에 대해 언어제공자들이 불편해 하여 원활한 대화 유도를 방해하는 요소로 작용하였다. 음성언어 기반의

문화 특성이 짙은 자료의 경우, 한국어 단어 형태가 아닌 시각적 상징 자료(가계도)의 형태로 제공하는 것이 필요하다.

구분	세부 내용
언어제공자	<ul style="list-style-type: none"> - 같은 직업군을 가진 언어제공자를 한 쌍으로 배치하여 직업과 관련된 다양한 수어가 표현될 수 있도록 유도해야 함 - 지역의 특성을 띄고 있는 언어제공자 섭외를 위해 협회와 더불어 다른 방법이 필요
촬영 환경	<ul style="list-style-type: none"> - 카메라, 메모리 기준에 대한 지침 전달 및 실제 교육이 필요
유도 자료	<ul style="list-style-type: none"> - 고모, 이모 등의 유도 자료 제공시 한국어 단어 대신 가계도 그림을 제시하여야 함.

6. 영상 편집

1) 편집 지침

- 연구자와 편집자 간의 의사소통을 위하여 다음과 같이 편집 지침을 마련하였다.
- 파일명은 지역 코드-언어제공자 코드-촬영방향 코드의 조합으로 부여하였다. 전남지역 코드는 2018년 구축된 말뭉치와 동일하게 ‘KSLJN’으로 하였다. 언어제공자 코드는 2018년에 구축된 4쌍의 언어제공자를 1~ 4번으로 부여하고, 2020년 수집된 언어제공자들에게는 5번부터 부여하였다.

A. 파일명

- (a) 원본 영상 파일명 : [지역코드]-[언어제공자 코드]-[원본]-[촬영방향코드]
- 원본 영상이란, 동의 영상과, 과제영상, 유도 자료 설명 부분 모두를 포함하여 촬영한 전체 영상을 의미함
- (예) 전남지역 첫 번째 언어 쌍의 왼쪽(A) 언어제공자의 원본 영상 파일명
- “KSLJN5원본Lt”

(b) 동의 영상 파일명 : [지역코드]-[언어제공자 코드]-[동의서]-[촬영방향코드]

- 동의 영상이란, A 과제(이름 말하기) 시작 전 언어제공자의 촬영을 재확인하는 영상을 의미함

지역코드		언어제공자 코드 (촬영순)	촬영방향코드 (풀샷 카메라 기준)	
지역명	코드		원쪽(A)	Lt
전남	KSLJN	5 ~ 20	풀샷	Md
			오른쪽(B)	Rt

(예) 전남지역 첫 번째 언어 쌍의 오른쪽(B) 언어제공자의 동의 영상 파일명

- “KSLJN5동의서Rt”

(c) 과제별 영상 : [지역코드]-[언어제공자 코드]-[과제 코드]-[촬영방향코드]

- 과제는 총 13가지이며, 과제명 코드는 A~M

지역코드		언어제공자 코드 (촬영 순)	과제명		촬영방향코드 (풀샷 카메라 기준)	
지역명	코드		과제명	코드	원쪽(A)	Lt
전남	KSLJN	5 ~ 20	수어이름	A		
			농담	B	풀샷	Md
			수어이야기	C		
			다시이야기	D		
			토론	E	오른쪽(B)	Rt
			표지판	F		
			특정사건	G		
			주제영영	H		
			개별수어	I		
			농사회 행사	J		
			흥부와놀부	K		
			옛날수어	L		
			농학교	M		

(예시 1) 경남지역 12번쌍 E과제 영상



KSLGN12ELt

KSLGN12EMd

KSLGN12ERT

(예시 2) 전남지역 5번쌍 A과제 영상

KSLJN9ALt

KSLJN9AMd

KSLJN9ARt

* 1번쌍의 언어제공자 코드는 '5'에서 시작하므로 5번째 쌍의 언어제공자 코드는 '9'가 됨

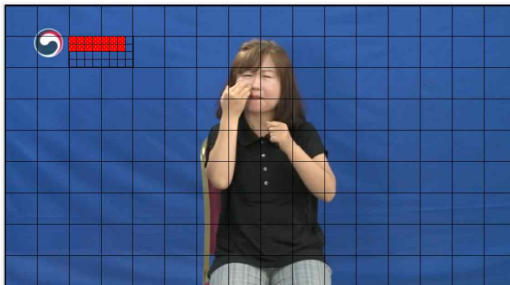
B. 파일 편집 기준

(a) 영상 화질

- 해상도: 1920*1080
- 프레임수: 30
- 화면화질: HD

(b) 국립국어원 로고

- 국어원 로고 지침서 참조
- 테두리가 있는 로고 사용, 글자크기 및 심볼과 글자간 간격 주의
- 로고 위치
: 화면을 64:36 비율로 나누었을 때, 9~15:5~6에 해당하는 지점에 글자가 들어오게 함. 아래와 같이 화면을 나누었을 때, 빨간색으로 표시한 곳에 '국립국어원' 글씨가 들어오게 함



: 단, 화면 비율에 의해 해당 지점에 로고를 삽입하였으나 로고의 위치가 적절하지 않다고 판단되는 경우 위치를 변경할 수 있음. 이것은 해당 연도에 원시 자료 촬영을 위한 카메라 세팅에 따라 위치가 변경될 수 있음을 의미함

(c) 과제별 영상 편집 시점

- 수어를 발화하기 시작하는 점과 발화가 끝난 전후에 여유가 있어야 함.
: 영상의 시작점은 유도 자료 설명 동영상 제시된 부분은 제외하되, 첫 수어 발화 시점 전 5초의 여유 필요
- : 마지막 발화가 끝난 후 5초의 여유를 두고 종료

(d) 과제별 편집파일에는 소리 삭제

(e) 원본 영상 취합 및 편집 영상 분할 방식

- 한 쌍의 언어제공사 촬영을 마친 촬영 영상은 메모리 용량, 카메라 교체 등의 이유로 상황에 따라 3개 이상의 원본 영상 파일이 있을 수 있으므로 녹화된 원본 영상들을 Lt, Md, Rt 촬영 영상의 편집 시작점을 맞추어(싱크 동기화) 하나의 원본 영상으로 제작한 후 각 과제별 편집 영상으로 분할함

C. 영상 검수 주요 내용

- (1) 누락된 영상이 있는가? (A~M 13/ Rt, Md, Lt 3/ 5~20번쌍 15)

총 영상갯수 : 과제 585개, 동의 영상 : 30개

- (2) 파일명이 제대로 되어 있는가?
- (3) 파일명과 내용이 다른 것이 있는가?
- (3) 한 개의 쌍의 영상 시간이 동일한가?
- (5) 시작점과 끝점이 동일한가? (엘란 확인)
- (6) 동영상의 시작과 끝에 여유가 있는가? (수어 발화 및 종료 전후 5초의 여유 필요. 단 여유 시간 5초 내 불필요한 영상이 있을 시 더 짧게 단축할 수 있음)
- (7) 편집 실수 여부 (앞뒤의 여유가 없는 것, bit rate 변환오류 여부)
: bit rate의 경우, 2015년 검수 내용에 있었던 것으로, [파일-속성-자세히]에서 비트속도 전송이 5000으로 된 경우 엘란에서 사진으로 불러지거나 검은 창으로 뜨는 경우가 있었다고 함. 현재 시점에서는 확인이 불가함.
- (8) 검수표 (2019년도 기준 예: 토론 및 사건의 주제는 일부 수정될 수 있음)

[illegible]

	4) 2002 월드컵											
	5) 노무현대통령											
	6) 세월호 침몰											
	7) 행복팀											
	8) 남북 정상회담											
	9) 안인득 사건											
H	주제											
	감정, 도시와시골, 종교, 휴가											
	날씨, 연애와 결혼, 운동과 취미, 대중매체											
	경제(돈), 여행, 건강, 패션											
	음식, 직업, 계절, 교통											
	의사소통, 행사(잔치), 관광명소, 부동산(재테크)											
I	개별수어											
J	농사회행사											
K	흥부와놀부											
L	옛날수어 현재수어											
M	농학교 경험											
無												

2) 영상 편집 내용 검수

- 촬영 원본 영상 및 편집본 검수는 2020년 9월 2일부터 2019년 11월 4일까지 3차례의 집합 검수 및 개별 검수로 진행되었다. 위에 제시된 영상 검수 기준을 참고하였으며, 진행 일정 및 오류 사항은 다음과 같다.

1차 검수(9/2)	- 46개 오류
오류 유형	<ul style="list-style-type: none"> • 영상의 밝기 정도가 중간에 바뀜 • 파일명 오류 • 불필요한 장면 삽입 및 영상 중간에 편집 흔적 • 화면 이상(모자이크 현상, 프레임 누락 등)

	• 영상이 없음
2차 검수(9/24)	- 11개 오류
오류 유형	<ul style="list-style-type: none"> • 파일명이 다른 복사본 삽입 • 불필요한 장면 삽입
3차 검수(11/4)	- 7개 오류
오류 유형	<ul style="list-style-type: none"> • 원본 영상 누락(편집 상 누락, 메모리 문제로 원본 영상 없음) • 파일명 오류
4차 검수(11/16)	- 0개 오류
오류 유형	<ul style="list-style-type: none"> • 오류 사항 없음 • 추가 촬영 영상에 대한 1차 검수에서도 오류 없음

- 1차 검수에서는 46개의 오류를 발견하였다. 오류로는 영상의 밝기 정도가 중간에 바뀌는 오류, 파일명의 오류, 수어 구현과 관계없는 장면 삽입 및 영상 중간에 편집된 흔적이 보이는 경우, 화면에서 모자이크 현상이 나타나거나 프레임이 누락되어 깜빡이는 경우, 영상이 없어 검은색 화면에 로고만 삽입한 경우가 있었다.
- 2차 검수에서는 11개의 오류를 발견하였으며, 오류 유형으로는 불필요한 장면 삽입, 파일명이 다른 복사본이 삽입된 경우가 있었다. 3차 검수에서는 3개의 쌍으로 이루어진 원본 파일에서 하나의 파일만 1시간가량 영상이 짧은 경우, 원본 영상이 없어 검은색 화면에 로고만 삽입한 경우, 파일명이 틀린 경우가 있어 수정 가능한 부분에 대한 수정을 요청하였고, 4차 검수에서는 오류가 발견되지 않았다.
- 4차 검수에서는 촬영 시간의 추가 확보를 위하여 2쌍의 언어제공자를 대상으로 수집한 영상에 대한 검수도 함께 이루어졌으며 추가 촬영분에 대한 편집에서는 오류가 발견되지 않았다. 결과적으로 15

쌍의 언어제공자를 대상으로 하는 연구 과업을 성취하기 위해 실시한 16쌍 분의 언어제공자의 수어 자료 수집 목표보다 초과 달성함으로써 전남 지역의 언어제공자는 과제별 영상 분량이 짧은 부분도 있었지만 여러 언어제공자의 수어 자료를 수집하였다는 점에서 유용할 것으로 사료된다.

7. 시스템 탑재

1) 언어제공자 정보 등록

- ① ‘한국수어 자료 통합지원 시스템’ 로그인 후, 첫 화면에서 ‘수어말뭉치 관리’ → ‘언어제공자 관리’ 클릭

The screenshot shows the homepage of the '한국수어 자료 통합지원 시스템'. At the top, there is a navigation bar with the system name and several menu items: '수어말뭉치 관리', '사 관리', '토론 관리', '타입 관리', '공유 게시판', '통계관리', and '오류접수'. The '수어말뭉치 관리' menu is highlighted with a red box, and a dropdown menu is visible below it, with '언어제공자 관리' also highlighted in red. The main content area is divided into several sections, each with a title and a list of items. The sections include '공지사항' (Notice), '최근 전사 목록' (Recent Transcription List), '타입 후보 요청 목록' (Type Suggestion Request List), and '최근 전사자별 전사 작업 목록' (Recent Transcription List by Transcriber). Each section has a '더보기 +' (More +) link. The '언어제공자 관리' link is the target for the next step in the process.

- ② ‘언어제공자 관리’ 접속 후 하단으로 스크롤 → ‘등록’ 클릭

한국수어 자료 통합 지원 시스템

수어 말뭉치 관리 인사 관리 토론 관리 타업 관리 공유 게시판 통계관리 오류접수





☰ > 수어 말뭉치 관리 > 언어제공자 관리

언어 제공자 목록

언어제공자 찾기 조건설정 초기화

모두 110건의 결과가 있습니다. (1/11)

엑셀 다운로드 10개씩 보기

번호	이름	원번호	지역	출생년도	성별	등록자	등록일시
110	 언어제공자116	-	경남	1959년	여성	홍성은	2019. 12. 16. 23:54:05
109	 언어제공자115	-	경남	1959년	여성	홍성은	2019. 12. 16. 23:33:03
102	 언어제공자108	-	경남	1973년	여성	홍성은	2019. 12. 12. 14:30:31
101	 언어제공자107	-	경남	1975년	남성	홍성은	2019. 12. 12. 13:51:11

<< < 1 2 3 4 5 6 7 8 9 10 > >>

등록

- ③ ‘언어제공자 등록’ 접속 후 언어제공자 정보 차례대로 입력 → 사진에서 ‘파일 선택’ 클릭 후 언어제공자 이미지 파일 업로드 → 업로드 후 ‘다음’ 클릭 → 전체 정보 입력 후 ‘저장’, ‘확인’ 클릭


한국수어 자료 통합 지원 시스템 수어 말뭉치 관리 전사 관리 토론 관리 타입 관리 공유 게시판 통계관리 오류접수

홈 > 수어 말뭉치 관리 > 언어제공자 관리

언어 제공자 등록

개인 정보 가족 정보 언어 학교 및 직업 농 사회에서의 활동

* 필수 입력사항이 있습니다.

1. 이름 *	언어제공자117		
2. 생년월일 *	1955	년 03	월 04
3. 주소 *	목포시 이포로 4번길 10-1		
사진	 <p>파일 선택 언어제공자117.png</p>		
4. 성별 *	<input checked="" type="radio"/> 남성 <input type="radio"/> 여성		
5. 현재 사는 곳 *	전남	거주 기간(에 1993 ~ 1997년)	선택하세요 ~ 선택하세요
6. 현재까지 살았던 장소 *	수어	선택하세요 ~년부터 선택하세요 ~까지 거주지 선택하세요 (지역명)	상세
7. 경력상태(손실 시기) *	<input type="radio"/> 농인 (<input type="radio"/> 태어날 때부터 <input type="radio"/> 살 이후부터) <input type="radio"/> 난청인 (<input type="radio"/> 태어날 때부터 <input type="radio"/> 살 이후부터)		
8. 청력보장구 착용유무 *	<input type="radio"/> 안공의우 시술을 받았다.(~년 이후부터 사용) <input type="radio"/> 규칙적으로 보청기를 사용한다 <input type="radio"/> 가끔 보청기를 사용한다 <input type="radio"/> 보청기 들리 보청기를 사용하지 않는다		

다음 목록

한국수어 자료 통합 지원 시스템 kshub.korean.go.kr 내용: 등록 하시겠습니까?

언어 제공자 등록

개인 정보 가족 정보 언어 학교 및 직업 농 사회에서의 활동

40. 한국수어 강사 경험	현재 한국수어를 가르치고 있다 <input type="radio"/> 예 <input type="radio"/> 아니오 과거에 한국수어를 가르쳤다 <input type="radio"/> 예 <input type="radio"/> 아니오
41. 한국수어 강사 활동 정도 [40번 '예'인 경우]	<input type="radio"/> 1년 이하 <input type="radio"/> 1~2년 <input type="radio"/> 3~4년 <input type="radio"/> 5년 이상
42. 수어 예술분야 활동에서 수어사용 여부	<input type="radio"/> 수어를 사용한다 <input type="radio"/> 수어를 사용하지 않는다
43. 우세손	글쓰기 <input type="radio"/> 오른손 사용 <input type="radio"/> 왼손 사용 <input type="radio"/> 가끔 왼손 사용 앞차림 <input type="radio"/> 오른손 사용 <input type="radio"/> 왼손 사용 <input type="radio"/> 가끔 왼손 사용 칼질하기 <input type="radio"/> 오른손 사용 <input type="radio"/> 왼손 사용 <input type="radio"/> 가끔 왼손 사용 수어할 때 <input type="radio"/> 오른손 사용 <input type="radio"/> 왼손 사용 <input type="radio"/> 가끔 왼손 사용
44. 마지막으로 하고 싶은 말	<div></div>

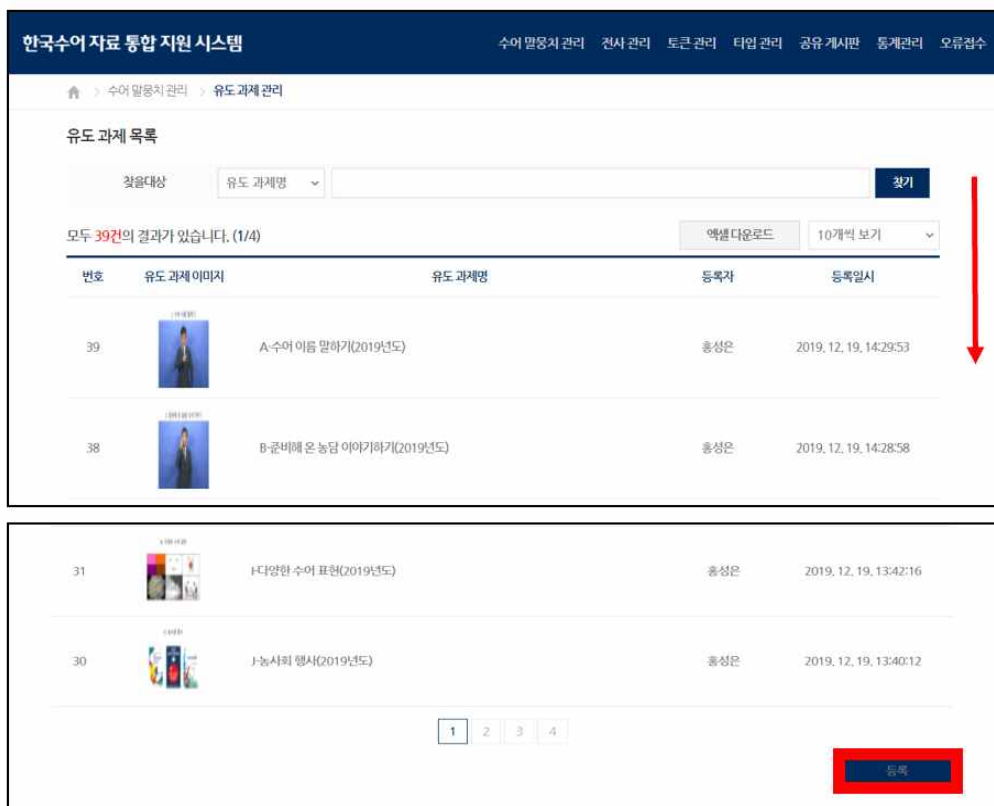
저장 확인 목록

2) 유도 과제 정보 등록

- ① ‘한국수어 자료 통합지원 시스템’ 로그인 후, 첫 화면에서 ‘수어말뭉치 관리’ → ‘유도 과제 관리’ 클릭



② ‘유도 과제 관리’ 접속 후 하단으로 스크롤 → ‘등록’ 클릭



③ ‘유도 과제 등록’ 접속 후 유도 과제 정보 차례대로 입력 → ‘유도 과제 파일 추가’ 클릭, 이미지, 영상 파일 업로드 → 업로드 후 ‘등록’, ‘확인’ 클릭

한국수어 자료 통합 지원 시스템

ksihub.korean.go.kr 내용: 등록하시겠습니까?

수어 말뭉치 관리 > 유도 과제 관리

유도 과제 등록

유도 과제명: A-수어 이름 말하기(2020년3도)

유도 과제 개요: 수어 이름 말하기

유도 과제 장르: 선택하세요.

유도 과제 파일

파일종류: 과제이미지

파일분량: 16.81KB / 10MB

파일선택: [16.81KB] KakaoTalk_20201108_21 선택 삭제

유도 과제 파일 추가

등록 목록

3) 원본 및 주석 영상 등록

(1) 원본 영상 등록

- ① ‘한국수어 자료 통합지원 시스템’ 로그인 후, 첫 화면에서 ‘수어말뭉치 관리’ → ‘원본 영상 관리’ 클릭

한국수어 자료 통합 지원 시스템

수어 말뭉치 관리

원어제공자 관리

유도 과제 관리

원본 영상 관리

원어 영상 관리

도록 찾기

공지사항

한국수어 자료 통합 지원 시스템 사용자 매뉴얼

20191217 테스트 사항입니다

통합시스템

최근 전자 목록

경남 지역 2번째 J-농사회 행사 영상

경남지역 9번째 J-농사회 행사

경남지역 11번째 J-농사회 행사

경남지역 8번째 M-농학교 경합 이야기하기

경남 지역 11번째 M-농학교 경합 이야기하기

질의 응답

인도행1

영상과 순위 오류

수화2

변경 타임 목록

수화

구월1

2020. 10. 21.

타임 후보 요청 목록

물러들디3

물러들디2

수화1

타입명 모르겠음

살습0

타임 후보 요청 목록

박정민

박정민

박정민

김수환

홍성운

최근 전자자료 전자 작업 목록

경남지역 8번째 M-농학교 경합 이야...

경남지역 11번째 J-농사회 행사

경남 지역 3번째 M-농학교 경합 이야...

경남 지역 3번째 M-농학교 경합 이야...

경남 지역 2번째 J-농사회 행사 전자...

2020. 9. 13.

2020. 9. 13.

2020. 9. 13.

2020. 9. 13.

2020. 9. 13.

② ‘원본 영상 관리’ 접속 후 하단으로 스크롤 → ‘등록’ 클릭

한국수어자료 통합 지원 시스템 수어 말뭉치 관리 전사 관리 토론 관리 타입 관리 공유 게시판 통계관리 오류접수

수어 말뭉치 관리 > 원본 영상 관리

원본 영상 정보 목록

제작일자: [] ~ []
 언어제공자: [] 초기화 조건설정
 찾을대상: 제목 [] 찾기

모두 56건의 결과가 있습니다. (1/6) 10개씩 보기

번호	사업명	원본 영상 정보 제목	저작자	등록자	등록일시
56	2019년 수어 말뭉치	경남지역 13번방 언어제공자 원본 영상	국립국어원	홍성은	2020. 2. 14. 18:01:26
55	2019년 수어 말뭉치	경남지역 12번방 언어제공자 원본 영상	국립국어원	홍성은	2020. 2. 14. 18:01:35
49	2019년 수어 말뭉치	경남지역 6번방 언어제공자 원본 영상	국립국어원	홍성은	2020. 2. 14. 17:37:07
48	2019년 수어 말뭉치	경남지역 5번방 언어제공자 원본 영상	국립국어원	홍성은	2020. 2. 14. 17:33:26
47	2019년 수어 말뭉치	경남지역 4번방 언어제공자 원본 영상	국립국어원	홍성은	2020. 2. 14. 17:29:52

1 2 3 4 5 6

등록

③ ‘원본 영상 등록’ 접속 후 원본 영상 정보 차례대로 입력 후 ‘등록’, ‘확인’ 클릭

한국수어자료 통합 지원 시스템 kshub.korean.go.kr 내용: 등록하시겠습니까?

수어 말뭉치 관리 > 원본 영상 등록

사업명: [선택하세요.]
 제목: []
 저작자: [] 제작일자: [] 촬영장소: []
 확장자: [선택하세요.] 해상도: [선택하세요.] 파일 총 용량: [] G
 영상시간: [] 시 [] 분 [] 초
 촬영일시: []

카메라: [카메라 추가]

카메라 기종	카메라명	파일명	프레임수	화면화질	연결된종위	종위	연결단언어 제공자	언어제공자	삭제
해당 정보가 없습니다.									

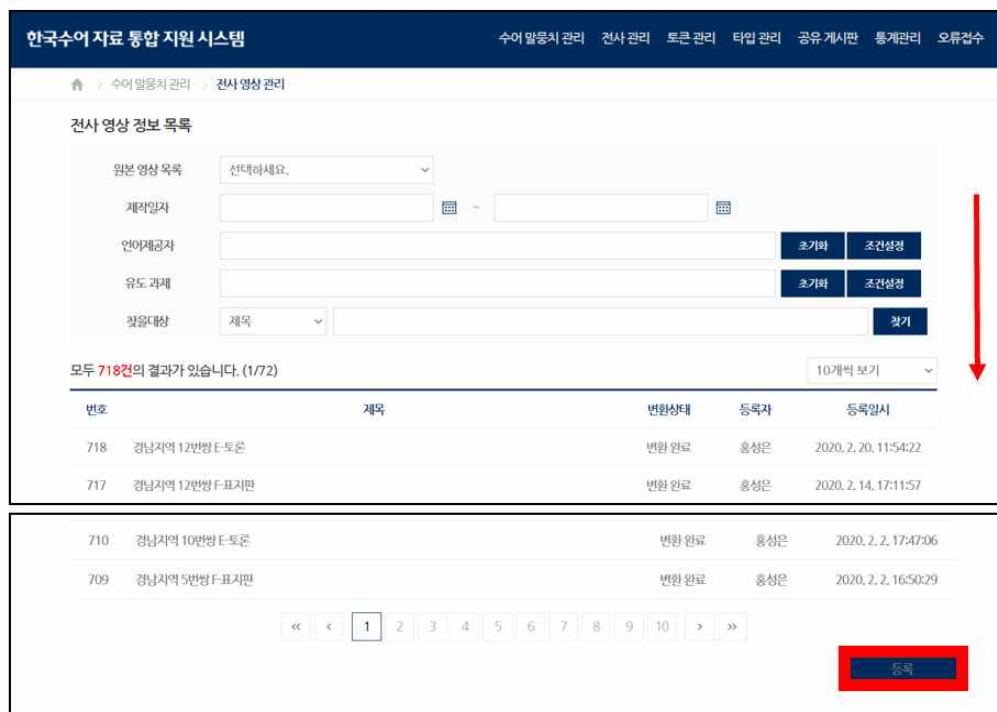
등록 목록

(2) 주석 영상 등록

- ① ‘한국수어 자료 통합지원 시스템’ 로그인 후, 첫 화면에서
‘수어말뭉치 관리’ → ‘전사 영상 관리’ 클릭



- ② ‘전사 영상 관리’ 접속 후 하단으로 스크롤 → ‘등록’ 클릭



- ③ ‘전사 영상 등록’ 접속 후 제목 입력 → 원본 영상 정보 항목에서 ‘연결’ 클릭 → 해당 원본 영상 클릭

한국수어자료 통합 지원 시스템 수어 말뭉치 관리 전사 관리 토론 관리 타입 관리 공유 게시판 통계관리 오류접수

전사 영상 등록

제목: 전남지역 12번방 E-토론

원본 영상 정보: 연결

원본 영상 정보 연결 팝업

검색대상: 제목 찾기

문두 56건의 결과가 있습니다.

번호	원본 영상 정보 제목	등록자	등록일
○ 1	경남지역 13번방 언어제공자 원본 영상	홍성은	2020. 02. 14 18: 01: 26
2	경남지역 12번방 언어제공자 원본 영상	홍성은	2020. 02. 14 18: 01: 35
3	경남지역 11번방 언어제공자 원본 영상	홍성은	2020. 05. 31 01: 24: 29

- ④ 전사 영상 붙임 항목에서 ‘선택’ 클릭 후 전사 영상 파일 업로드

한국수어자료 통합 지원 시스템 수어 말뭉치 관리 전사 관리 토론 관리 타입 관리 공유 게시판 통계관리 오류접수

전사 영상 등록

제목:

원본 영상 정보: 전남지역 12번방 언어제공자 원본 영상 연결

촬영 일자:

전사 영상 붙임

카메라명	영상품질	원본파일명	붙임파일
왼쪽 카메라	선택하세요	KSLGN12원본A	0KB / 6000MB 선택
가운데 카메라	선택하세요	KSLGN12원본Md	0KB / 6000MB 선택 삭제
오른쪽 카메라	선택하세요	KSLGN12원본R	0KB / 6000MB 선택 삭제

⑤ 유도 과제 항목에서 ‘유도 과제 추가’ 클릭

⑥ 해당 유도 과제 선택 후 ‘선택’, ‘적용’ 클릭

⑦ ‘등록’, ‘확인’ 클릭

III. 한국수어 말뭉치 주석



1. 주석 범위 및 절차
2. 주석자 및 검증자 교육
3. 기본 주석 및 의미 주석 실시
4. 기본 주석 및 의미 주석 검토
5. 타입 등록
6. 수어 말뭉치 주석 규약

Ⅲ. 한국수어 말뭉치 주석

1. 주석 범위 및 절차

2020년 연구의 주석 달기 과업은 2019년 경남지역에서 수집된 수어 말뭉치 영상 중 총 80분 분량의 영상에 대해 기본 주석 및 ‘지시’ 타입이 기록된 토큰을 제외한 기본 주석이 된 각 토큰에 대해 의미를 주석하였다.

1) 주석 범위

주석은 기본 및 의미 주석으로 하며 세부내용은 <표 Ⅲ-1>과 같다.

<표 Ⅲ-1> 기본 주석과 의미 주석의 범위

구분	세부내용
기본 주석	<ul style="list-style-type: none"> • 목적: 수집된 수어 말뭉치가 데이터(DB)로 만들어지기 위한 수어 발화의 형태 기록 • 내용: 발화된 한 개의 수어 단어의 시작과 끝을 분절한 하나의 토큰마다 글로스를 사용하여 기록 • 방법 <ul style="list-style-type: none"> ① 분절된 토큰에 입력할 글로스 확인 <ul style="list-style-type: none"> : 글로스는 한국수어 자료 통합 시스템 내 타입 목록에서 토큰에 해당하는 타입 여부 검색 : 타입 목록에 타입이 없는 경우 한국수어 자료 통합 시스템 내 ‘타입 제안’ 메뉴에 영상과 함께 새 타입 등록 : 연구진 회의를 통해 결정된 타입 결정 ② 토큰에 타입명 기록
의미 주석	<ul style="list-style-type: none"> • 목적: 토큰별 의미 기록 • 내용: 기본 주석 분절에 맞추어 [지시], [###], [%], [@] 토큰을 제외한 모든 토큰의 의미 주석 기록 • 방법 <ul style="list-style-type: none"> - 해당 토큰의 문장 맥락 안에서의 의미 기술 - 비수지 표지는 조음소로 작용하는 비수지만을 대상으로 하며,

	<p>문장 내 문법적 기능을 하는 비수지를 제외하고 의미 기술</p> <ul style="list-style-type: none"> - 대역어로 기술하되 맥락 안에서 같은 의미의 대역어가 있는 경우 함께 기술 - 해당 토큰의 의미가 동사나 문장의 형태인 경우는 술어를 기본형(~다.)으로 기술
--	---

2) 주석 진행 절차

주석 절차는 [그림 III-1]과 같다.

일시	추진 내용
6월 1일	주석자 및 검증자 선정 파일 선정 및 분배
6월 9일	주석자 온라인 교육 및 회의 1분 시범 주석(주석자)에 대한 피드백(검증자) 제공
7월 3일	분절 재교육 및 회의 주석 일치도 확인을 위한 주석자 및 검증자 논의
7월 13일	주석 진행 효율성을 위한 방안 논의
7월 25일 ~ 7월 30일	검증자, 연구진 회의 (타입후보 등록) 주석 완료
8월 1일 ~ 9월 30일	검토 실시(검증자)
9월 1일 ~ 2일	검증자, 연구진 회의 (타입후보 등록)
9월 4일 ~ 13일	주석 검토 결과 논의 (타입후보 등록)
10월 ~ 11월	주석 지침 정비 주석(기본 및 의미) 및 타입 정리

[그림 III-1] 주석 진행 절차

2. 주석자 및 검증자 교육

기본 및 의미 주석 작업자는 주석에 사용하는 엘란(ELAN) 프로그램과 주석 작업 지원을 위한 「한국수어 자료 통합지원 시스템」에 대한 이해와 사용 절차에 익숙해져야 한다. 그리고 주석은 발화되는 수어의 시작점과 끝점을 정확하게 분절하여 ID 글로스를 입력해야 하는 정교한 작업이기 때문에 주석 작업자를 대상으로 기본 교육이 필요하며, 정기적인 집합 교육을 통해 주석 일치도를 높여야 한다.

이에 본 사업에서는 원활한 주석 작업을 위해 기본 교육이 필요하지 않은 주석 유경험자를 중심으로 주석자를 모집하여 기초 교육 단계를 생략할 수 있었다. 그러나 주석 작업이 단기간 일회성으로 진행되어 왔기 때문에 오랜 시간 동안 주석 작업을 하지 않았던 주석자의 주석 숙련도 문제와 작업자 간의 일치도 문제를 고려하여 시범 주석을 계획하고 2020년 6월 1일 파일 분배가 완료된 이후 각자의 파일에 대한 1~2분가량의 표본 주석 달기를 진행하였다. 모든 주석자는 세 번 이상 주석 작업을 해 본 유경험자였지만 주석 작업 초기에 수어 분절과 글로스 확인에서 다소 어려움을 나타냈다. 이에 1분 표본 주석달기 시뮬레이션을 실시한 후에도 회의를 통해 주석자 간 일치도 확인 및 주석자가 인지하지 못하는 수어 확인 과정을 가졌다. 이 과정은 주석 작업이 일차 종료되는 시점인 7월 31일까지 온·오프라인 회의를 통해 수시로 이루어졌다. 주요 내용은 다음과 같다.

1) 주석자 교육

주석팀은 3명의 주석자로 이루어졌으며 모두 주석 유경험자로서 기초 교육은 생략하였다. 주석자들은 유경험자이지만 주석 작업이 연구가 수행될 때마다 단기적으로 이루어져 지속성이 없었기에 상당 기간의 휴식기로 인해 재적응의 시간이 필요하였다. 이에 파일 분배를 마

친 후 개별적으로 작업을 수행해 보게 한 다음 통합 시스템 및 엘란 ((EUDICO Linguistic Annotator, ELAN) 시스템 적응 차원에서의 재교육을 다음과 같이 실시하였다. 특히 주석자 간, 주석자 내, 주석자 및 검증자 간의 주석 일치도 확인은 중요하므로 1분 분량의 표본 주석에 대한 검증 작업을 실시하였으며, 분절과 타입 기록에 대한 추가 교육도 실시하였다.

○ 1차 교육

- 일시: 6월 9일(화) 19:00~21:00
- 장소: 화상 회의
- 교육내용
 - 주석의 기본 기능
 - 통합시스템 활용 및 엘란 기초 기술
 - 1분 시범 주석을 통해 본 주석 기록 방법 논의

○ 2차 교육

- 일시: 7월 3일(금) 19:00 ~ 21:00
- 장소: 한국복지대학교
- 교육내용
 - 분절 관련
 - 의미주석 교육 및 지침 안내
 - 기본주석 교육 및 검토
 - 주석자와 검증자 질의 및 논의(주석에 대한 일치도)

○ 온·오프라인 회의: 수시 실시

2) 검증자 교육

연구자 2명을 포함하여 3명으로 구성된 검증팀이 주석자의 작업 내용을 확인한 결과 주석자의 업무 시간을 지연시키는 요인이 분절과 타입명 선정에 관한 것이었음을 확인하였다. 특히 타입명 결정에서는 고정된 수어와 생산적 수어의 수형에 대한 글로스를 선택할 때 주석자가 알고 있는 명칭과 다른 글로스로 타입명이 사용되고 있을 때 동일한 수어 형태가 중복된 타입명으로 나타나기도 하였다.

이러한 점을 효율적으로 대처하기 위하여 수시로 온라인 회의를 실시하였을 뿐만 아니라 주석자의 업무 지원 및 유연한 대처가 가능하도록 주석 작업의 시작 시점부터 3명의 주석자와 3명의 검증자 간 일대일 매칭을 하여 진행하였다. 신뢰도를 높이기 위하여 3명의 검증자 간 수시 회의를 개최하여 타입명 결정 등에 대한 의사 결정을 하였다.

이러한 과정이 원활하게 수행되기 위해서는 검증자도 주석에 대한 기본 원칙을 숙지하고 있어야 하므로 주석자가 실시한 1분 주석을 검토한 결과물을 토대로 검증자 회의를 실시하였다.

3) 교육 내용

본 연구 사업에서 실시하는 주석의 범위와 주석 방법에 대한 일차 교육을 실시한 후 본격적인 주석 작업 시작에 앞서 시범적으로 1분 분량의 영상을 주석하였고 검증자가 이를 검토한 결과 주석 유경험자임에도 불구하고 주석자 간, 주석자와 검증자 간 분절 일치도가 낮게 나타나는 것으로 확인되었다. 이것은 주석자와 검증자가 분절의 기준을 해석하는 데 차이가 있었기 때문이다. 수어자가 발화한 하나의 단어를 시작되는 지점에서 끝나는 지점까지 분절한 형태를 「한국수어 자료 통합지원 시스템」의 토큰 목록에서 확인할 수 있다. 「한국수어 자료 통합지원 시스템」에서 수어 발화가 완전히 끝나지 않은 상태에서 끊어진 토큰이나 수어 형태와 관련 없는 불

필요한 연속 동작이 길게 나타나는 경우 등, 종종 잘못된 분절의 형태가 나타나 주석에서 분절의 정확도를 높이기 위한 교육을 실시하였다. 이때 분절의 기준을 설명하는 예시를 익숙하지 않은 네덜란드 수어에서 한국수어로 바꾸었다. 아래 [그림 III-2]에 제시된 분절의 예시는 영상(<https://youtu.be/NKGHB3oxny4>)으로 제작하여 [그림 III-3]의 QR코드의 형태로 주석자들에게 제공하였다.



[그림 III-2] 제스처 이동 단계의 예 ‘한국수어 [가족]’

※ 출처: De Vos, C., Torreira, F. & Levinson, S C., 2015



[그림 III-3] 분절 기준에 대한 설명 영상 QR코드

또한 주석자들이 영상을 재생하면서 작업을 할 때 분절 시점이 헛갈리는 경우, 수동으로 직접 스크롤하여 분절의 전후 시점을 살피는 것이 중요하다는 점을 반복하여 설명하였다.

3. 기본 주석 및 의미 주석 실시

1) 주석 층렬

기본 주석을 위하여 각 언어제공자(A, B)마다 우세손과 비우세손에 대한 글로스와 코멘트를 기록하기 위한 3개(우세Gloss, 비우세Gloss, 코멘트)의 층렬을 구성하였다. 의미 주석을 위한 층렬도 각 언어제공자마다 3개의 층렬(의미_우세, 의미_비우세, 의미_코멘트)을 구성하였으며, 그 내용은 아래 [그림 III-4]과 같다.



[그림 III-4] 엘란 프로그램의 기본 주석 층렬 구성

따라서 한 명의 언어제공자 수어 발화에 대한 주석 층렬은 [그림 III-5]과 같이 우세 GlossA, 의미 A, 비우세 GlossA, 비우세 의미A, 코멘트, 의미 코멘트로 총 6개의 층렬이 있으며, 주석자들은 각 토큰에 대한 층렬에 주석을 하였다.



[그림 III-5] 엘란 프로그램을 통한 주석 기록의 예

2) 기본 주석

본 연구에서는 경남지역 원시 말뭉치 자료 중 7개의 파일에 기본 주석과 의미 주석 작업을 실시하였다. 본 연구에서는 농학교 경험 과제 영상 3개(37분 32초), 농사회 행사 과제 영상 4개(42분 47초)로 영상의 총 길이는 80분 19초이다. 기본 주석을 실시한 토큰은 총 13,673개로 농학교 경험 과제 영상의 토큰이 6,150개, 농사회 행사 과제 영상의 토큰이 7,523개로 나타났다.

〈표 III-2〉 과제별 기본 주석 토큰 수

구분	농사회 행사				농학교 경험			총계
	경남 2J	경남 3J	경남 9J	경남 11J	경남 3M	경남 8M	경남 11M	
영상시간	02:41	13:24	17:22	09:20	14:16	10:10	13:06	80:19
우세A	240	816	509	857	516	505	782	4,225
비우세A	149	444	208	406	332	203	349	2,091
우세B	64	810	1474	308	936	574	797	4,963
비우세B	25	375	681	157	466	235	455	2,394
합계	478	2,445	2,872	1,728	2,250	1,517	2,383	13,673

기본 주석 층렬에는 ID 글로스 외 [%], [###], [@]와 같은 기호가 기록되는 경우가 있다. [%]는 수어가 완전히 조음되지 않았을 때 입력하여 거짓 시작임을 나타내는 표시이다. [###]는 주석자가 수어로 인지했지만 무슨 수어인지 모르는 경우로 ID 글로스 대신에 입력하였다. 이후 주석자와 검증자가 모두 모여 논의를 한 후 의문이 해결되면 ID 글로스로 바꿔 입력하도록 하였으며, 이후에도 어떤 수어인지 명확하지 않을 경우에는 [###]를 그대로 두었다. [###]는 수어가 영상화면 바깥에서 조음이 되거나 보이지 않는 경우에도 입력하였다. [@]는 수어 화자가 실수로 수어를 잘못 표현했거나 사용되지 않는 수어일 때 입력하는 표시이다. 기본 주석 기록에 나타난 [%], [###], [@]의 토큰 수는 <표 III-3>와 같으며 ‘###’, ‘@’, ‘%’의 총 토큰 수는 422개였다.

<표 III-3> 과제별 기본 주석 층렬의 [%], [###], [@] 토큰 수

구분		농사회 행사				농학교 경험			총계
		경남 2J	경남 3J	경남 9J	경남 11J	경남 3M	경남 8M	경남 11M	
영상시간		02:41	13:24	17:22	09:20	14:16	10:10	13:06	80:19
우세A	%	3	8	2	14	6	15	18	66
	#	5	21	1	14	12	15	10	78
	@	0	0	0	4	2	4	3	13
비우세 A	%	1	4	0	6	1	5	3	20
	#	0	14	1	7	3	7	1	33
	@	0	23	0	2	1	2	0	28
우세B	%	0	15	23	10	18	20	8	94
	#	1	7	3	2	15	1	2	31
	@	0	1	0	0	4	1	2	8
비우세 B	%	0	3	8	4	9	6	2	32
	#	0	8	0	0	7	0	1	16
	@	0	0	0	0	1	0	2	3
합계		10	104	38	63	79	76	52	422

기본 주석에서 나타난 타입별 토큰 수를 살펴보면 <표 III-4>와 같다. 지시 타입([지시1], [지시1_1], [지시2_1])의 빈도가 가장 높았으며 고정적 수어 타입으로는 [끝1]이 202 토큰, [같다1]이 172 토큰, [맞다1]이

162 토큰으로 많은 빈도를 나타냈으며, 상위 빈도수를 보이는 생산적 수어의 수형 타입인 [구형1]은 174 토큰 발화된 것으로 나타났다. 단, 여기서 나타나는 토큰 수는 A와 B 언어 제공자와 우세, 비우세손을 모두 합한 빈도수이다.

<표 III-4> 타입별 토큰 수

토큰명	자시1	자시1_1	자시2_1	끝1	구형1	같다1	맞다1	농1	있다1	오늘1	때문1	많다1
빈도수	807	240	210	202	174	172	162	144	141	139	121	107
토큰명	운동1	학교1	불량1	무엇1	없다1	보다1	경상 남도1	교사1	아니다1	여자1	그러나1	알다1
빈도수	102	91	90	89	86	84	68	66	65	64	58	57
토큰명	관찰 대1	들어가 다2	아니다2	교육 1	자연1	구형2	좋다1	때2	때리다2	때1	그러면1	일1
빈도수	56	56	56	54	54	53	53	52	52	49	47	47

3) 의미 주석

의미 주석은 기본 주석이 달린 토큰 중에서 [지시], [###], [%], [@] 토큰을 제외한 모든 토큰에 의미를 기록하는 작업이다. 이에 본 연구에서는 의미 주석 작업을 실시한 토큰은 총 11,766개로 농학교 경험 과제 영상의 토큰이 5,228개, 농사회 행사 과제 영상의 토큰이 6,538개로 나타났다. 과제별 의미 주석 토큰 수는 <표 III-5>와 같다.

<표 III-5> 과제별 고정된 수어의 의미 주석 토큰 수

구분	농사회 행사				농학교 경험			총계
	경남 2J	경남 3J	경남 9J	경남 11J	경남 3M	경남 8M	경남 11M	
영상시간	02:41	13:24	17:22	09:20	14:16	10:10	13:06	80:19
의미_우세A	221	696	446	681	409	409	633	3495
의미_비우세A	148	403	193	363	289	170	309	1875

의미_우세B	50	666	1241	259	774	481	700	4171
의미_비우세B	24	347	656	144	409	212	433	2225
합계	443	2112	2536	1447	1881	1272	2075	11766


* 기본 주석 중 [%], [###], [@], [지시] 토큰을 제외함

(1) 의미 주석의 범위 및 기술 방법


- [지시] 토큰을 제외한 모든 토큰에 의미 주석하므로, 기본주석 분절에 맞추어 의미 주석을 달았다.
- 해당 토큰의 문장 맥락 안에서의 의미를 기술하였다.
- 비수지 표지는 조음소로서 작용하는 비수지만을 대상으로 하고, 문장 내 문법적 기능을 하는 비수지를 제외한 의미를 기술하였다.
- 대역어 형태로 기술하되 맥락 안에서 같은 의미의 대역어가 있는 경우 함께 기술한다. 예를 들어 ‘상세하다, 자세하다’ 등 한국어 동의어가 있는 경우 2-3개 이하로 기술하였다.
- 해당 토큰의 의미가 동사나 문장의 형태인 경우는 서술어를 기본형(~다.)으로 기술하였다.

(2) 의미 주석 기록의 예

- [이상하다1]의 일반적인 의미는 이상하다, 신기하다 등으로 사용된다. 그러나 본 맥락에서는 ‘호기심이 생기다’, ‘궁금하다’의 의미로 사용되었다.

타입명	<p>이상하다1</p> 
일반적 의미	이상하다, 신기하다
맥락적 의미	수어의 형태는 [이상하다]이나 비수지 및 문맥에 따라 ‘호기심’, ‘궁금하다’ 라는 의미로 나타남

- 첫 번째 [구형4]는 납작한 원반의 공 모양의 ‘펼’을 나타내는 내는 것으로 두 공의 상태를 나타내고 있었다.

타입명	<p>구형4</p> 
의미	펼이 관문으로 밀려가다, 관문을 향해 밀려나가는 펼, 펼이 다른 펼에 맞다, 뒤에 오는 펼에 맞는 앞선 펼
설명	생산적 수어 ‘손대체 기법’

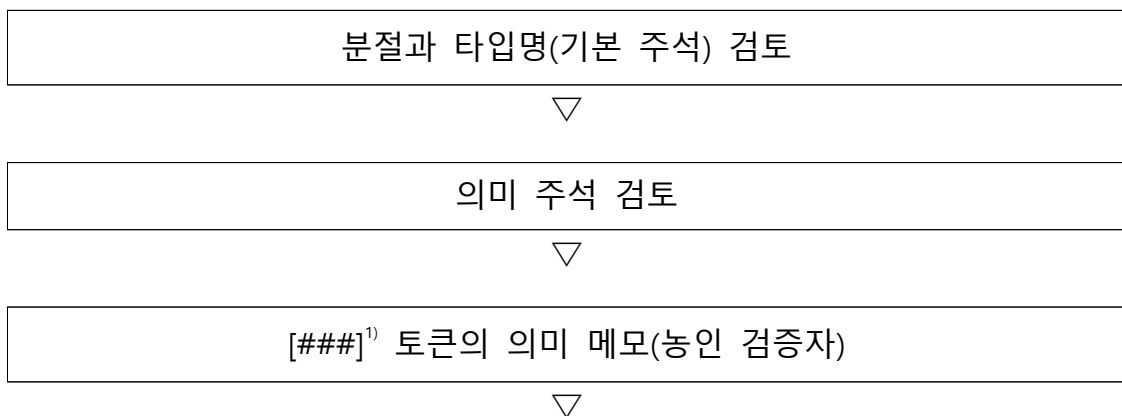
- 두 번째 [구형4]는 지팡이를 쥐고 서 있는 모습을 나타내며 ‘지팡이를 들다’를 의미하였다.

타입명	구형4 
의미	지팡이를 들다
설명	생산적 수어 ‘손이용 기법’

4. 기본 주석 및 의미 주석 검토

1) 검토 과정

검증자는 기본 주석에서 주석자가 작업한 토큰의 분절이 정확히 이루어졌는지와 더불어 [###]으로 표시되거나 누락된 토큰의 타입명을 결정하였다. 또한 의미 주석에서 해당 토큰의 의미가 잘 드러나는지에 대해 검토하였다. 검토과정은 다음과 같다.



의미를 메모한 [###] 토큰의 타입명 결정(검증자 간)



타입명 결정에 어려움이 있는 경우 연구진과 논의

[그림 III-6] 주석 검토 과정

본 연구의 주석 작업은 두 가지 방식으로 진행되었다.

첫 번째 방식은 주석자가 주석을 달 수 없었던 토큰에 ###으로 타입명을 달고 ‘주석완료’ 처리하면, 검증자의 검토과정을 거쳐 문제를 해결하였다.

두 번째 방식은 청인 주석자와 농인 검증자가 팀을 이루어 주석 작업을 하던 중 문제가 발생할 경우 주석자가 검증자에게 직접 문의하여 바로 문제를 해결하며 진행하였다.

주석 파일에 대한 주석과 검토 과정에서 진행된 위 두 가지 방식은 모두 장단점이 있었다.

첫 번째 방식의 경우, 주석을 완료하는 시간이 두 번째 방식보다 빠르며, 주석자와 검증자의 역할을 확실하게 구분하여 진행할 수 있어 시간상으로 효율적인 진행이라고 할 수 있다. 하지만 주석자가 잘못된 타입을 선택하여 주석을 진행한 경우 주석된 모든 타입을 변경하거나 재주석해야 하는 경우가 발생한다. 또한 하나의 타입을 해결하지 못하게 되어 같은 맥락에서 이루지는 다른 타입도 이해하지 못하는 경우에는 ###으로 처리하여 검증자가 많은 양의 ### 처리된 문제를 해결해야 하는 어려움이 있었다.

두 번째 방식의 경우, 주석을 완료하는 시간이 첫 번째 방식보다 더디며, 주석자와 검증자의 역할을 구분하기가 어렵다. 특히, 주석자의 질문이 바로 해결되기 어려운 경우 검증자는 문제를 해결하기 위해 다양한 방법을 모색해야 하며, 검증자가 문제 해결에 대한 책임을 지게 된다. 또한 주석자의 질문이 해당 파일의 앞뒤 맥락을 확인해야 하는 경우, 검증자는 해당

1) 주석 작업자가 모르는 수어이거나 명확하지 않은 경우 [###]로 토큰을 표시한다.

파일을 확인하는 데 많은 시간을 할애해야 한다. 하지만 첫 번째 방식보다 주석의 완성도나 신뢰도가 높다는 장점이 있다. 또한 주석 작업을 하는 데 시간이 많이 들지만 검토하는 시간은 줄어든다.

두 번째 방식을 보다 효율적으로 진행하기 위해서는 현재 종합시스템을 활용하는 방식에 변화가 필요하다. 주석 작업 중 검증자는 주석 중인 파일에 대한 접근 권한이 없기 때문에 자신의 아이디로 파일을 수정하여 반영할 수 없다. 때문에 주석자의 아이디를 이용하여 파일을 수정하거나 검증자 아이디로 수정한 파일을 이메일이나 다른 방법으로 주석자에게 전달하고 이를 주석자가 받아 다시 반영해야 하는 번거로움이 있다. 이에 청인 주석자가 ###로 처리했거나 잘못 주석한 것을 놓인 검증자가 직접 수정할 수 있도록 주석 중에도 검증자가 해당 파일에 대한 권한을 가질 수 있도록 시스템을 수정 보완해야 할 필요가 있다.

2) 분절 관련 검토 내용

검토 작업에서 분절과 관련하여 논의한 내용은 다음과 같다.

① 한손의 수어 유지, 한손의 수어가 변경되는 경우


‘수업이 있다’를 표현하기 위해 [공부1]이 우세손과 비우세손에서 함께 구현된 이후 비우세손은 [공부1]의 수어를 유지, 우세손은 [공부1]에서 [있다1]로 바뀌었다. 이에 비우세손의 [공부1]에서 어떤 시점이 분절의 끝이어야 하는지에 대해 검토하여 수정하였다.

<p>타입명</p>	<p>공부1</p> 
------------	--

주석자 기록	우세손이 사용된 [공부1]의 출발과 끝을 비우세손에도 똑같이 분절하였음.
검증자 수정	비우세손이 [공부1]을 표현한 후 약간 내린 상태라 해도 완전히 내리기까지 유지에 들어감. 그 유지에서 의미적 특성이 발견되지 않았으나, 문장의 끝과 연관될 수 있기에 유지로 인정하고 분절하였음.


② 수형과 수위가 먼저 나타난 채로 유지되는 경우

[사년전1] 수어의 수형과 수위가 명확히 나타났으나 두 손이 닿는 움직임이 나타나지 않은 채 언어제공자가 고민하는 동안 수형과 수위가 유지되었다. 이후 [사년전1] 수어에서 두 손이 닿는 움직임이 나타났는데 수형과 수위가 명확히 나타나는 시점에서 분절을 해야 하는지, 두 손이 닿는 시점에서 분절을 해야 하는지에 대해 논의하였고, 수형과 수위에서 일정 의미가 나타나므로 수형과 수위가 명확히 나타나는 시점에서 분절하였다.

타입명	사년전1 
주석자 기록	분절 지침 내용 중 준비하는 과정은 분절에 포함될 수 있으나, 두 손이 닿는 움직임을 꼭 포함해야 한다고 하여 [사년전1] 수어에서 두 손이 닿는 시점부터 분절하였음.
검증자 수정	처음 나오는 수형과 위치로 인해 [사년전1]이나 [네시간1]을 연상할 수 있는 의미가 있는 표현이기에 유지부터 분절해야 한다고 판단하였음.


③ 수어의 기본형태가 우세손과 비우세손의 움직임 시점이 다른 경우

[경험1]을 표현하기 위해서는 우세손이 목을 가리키고 양손이 전후로 두 번 스친다. 우세손이 짧게 목을 가리키는 동안에 비우세손은 아무 움직임이 없었지만 하나의 단어로 보고 우세손과 비우세손의 분절이 일치되어야 하는지, 우세손이 움직임을 시작하는 순간부터 분절되어야 하는지 논의하였고, 분절 지침에 따라 우세손과 비우세손의 수형이 시작점부터 분절하였다.

타입명	<p>경험1</p>  <p>우세 비우세</p> <p>날1 경험1 날1 경험1</p>
주석자 기록	[경험1] 수어를 우세손과 비우세손의 시작과 끝이 똑같이 분절하였음.
검증자 수정	분절 지침 그대로 비우세보다 우세손이 더 먼저 사용되기에 다르게 분절하였음.

④ 앞서 발화된 수어의 수형 그대로 제스처를 사용하는 경우

‘저기요’ 제스처의 수형이 앞서 발화된 [경상남도1]의 수형으로 사용되었다. 이에 코멘트에는 제스처 부분에만 음운변동을 분절해야 하는 건지 아니면 [경상남도1]에서 왔기에 [경상남도1]에서 출발, 제스처를 끝으로 분절해야 하는지 분절 기준에 대해 논의하였다. ‘저기요’ 제스처의 일반적인 수형과 다른 이유는 앞서 나타난 [경상남도1]에서 나타난 수형의 영향이므로 코멘트 총렬에는 [경상남도1]과 [제스처]를 함께 분절한 후 ‘제스처, 음운변동’으로 기록하였다.

타입명	경남1 제스처  <div> 우세 경상남도1 제스처 비우세 경상남도1 코멘트 음운변동 아니면 코멘트 제스처, 음운변동 </div>
주석자 기록	코멘트에는 변동이 있는 해당 제스처 분절에만 분절하였음.
검증자 수정	음운변동의 영향을 미치는 요인이 포함된 [경상남도1]도 포함해야 한다고 하여 [경상남도1]과 [제스처]를 함께 분절하였음.

⑤ 음운 변동의 형태가 고정적으로 사용되는 경우

흔히 ‘~의 날’을 나타내기 위하여 [고리1], [날1]을 사용하는데, ‘농인의 날’을 나타내는 수어에서 음운변동으로 [고리1]의 수어가 없고 형태가 변한 [날1]로 나타나 이를 음운 변동의 사용이 굳어진 것으로 보아 코멘트에서 [농1]과 [날1]을 하나로 분절하였다. 이를 음운 변동이 아닌 고정적 어휘의 형태로 볼 것인지에 대한 추후 논의가 더 필요하다.

타입명	농1 ### 날1  <div> 우세 농1 ## 날1 비우세 ## 날1 코멘트 음운변동 </div> <div> 00:00:06.719 00:00:06.588 00:00:06.649 00:00:06.710 00:00:06.771 00 농1 ## 날1 </div>
-----	--

주석자 기록	[농1]과 [날1]만 기록하였고 그 사이에는 분절하지 못하였음
검증자 수정	[고리1]와 [날1]의 상호 영향으로 음운론적 변동이 나타나 [고리1]의 수어가 나타나지 않고 약간의 변형이 있는 [날1] 수어만 나타나 [고리1]에 해당하는 분절에는 [###]로 처리하였음. 이외 다른 ‘~의 날’을 나타내는 수어에서는 [고리1]과 [날1]이 이처럼 변동이 생기지 않을 것으로 판단하였고, 농사회에서 자주 사용되는 ‘농인의 날’에만 습관적으로 사용한 결과로 음운론적 변동이 생기다가 그 형태가 굳어지는 결과가 생긴 것으로 판단하여 코멘트에는 [농1]과 [날1]에 해당하는 수어를 하나로 분절하였음.

3) 타입명 관련 검토 내용

검토 작업에서 타입명과 관련하여 논의한 내용은 다음과 같다.

① 지역적 음운변이가 나타나는 경우

현재 ‘장난1’ 타입은 양손의 1지를 펴고 머리 양 옆에서 번갈아 흔드는 형태인데 양손의 1지와 2지를 편 채 머리 양 옆에서 번갈아 흔드는 형태의 ‘장난’이 나타났다. 현재 주석 지침에는 음운 변이의 경우 기존 타입을 사용하도록 되어있는데, 이 경우에는 경남지역에서 나타나는 음운 변이로 보이므로 추후 지역적 음운변이에 대한 논의가 필요하다.

타입명	장난1 혹은 ###
-----	------------

	
주석자 기록	장난3
검증자 수정	[장난1]에 해당하는 의미를 가지며 비슷한 형태를 가지는 수어가 경남지역에서 사용되는 것으로 보임. 현재 주석 지침에 따르면 음운 변이의 경우 기존 타입명을 따르도록 되어 있어 [장난1]로 결정하였음.

② 생산적 수어의 형태로 보이는 지역적 변이가 나타나는 경우

서울 지역에서는 양손을 펴고 비우세손은 아래로, 우세손을 위로 두고 우세손을 움직여 비우세손을 쓸 듯 움직여 ‘청소를 하다’의 의미를 표현할 수 있으며, 이는 분류사의 형태로 사용하기도 한다. 그러나 경남지역에서는 고정적 어휘로 사용되기 때문에 이를 생산적 수어인 [구형1]에서 [청소하다2]로 타입명을 수정하였다.

타입명	구형1 
주석자 기록	구형1(의미층렬: 청소)
검증자 수정	[청소하다2]로 타입명을 수정하였음.

③ 지역과 관련된 수어

지역명 수어의 경우 해당 지역 출신이 아닌 주석자와 검증자 모두 수어의 의미를 알 수 없어 경남 지역 농인의 도움을 받아 타입명을 결정하였다.

타입명	마산1, 양산1, 김해1, 진주1, 사천1 등 
주석자 기록	###
검증자 수정	경남 지역 농인의 도움을 받아 [마산1], [양산1] 등으로 타입명을 결정하였음.

5. 타입 등록

주석 작업 중 주석자가 타입목록에 없는 수어의 형태가 인지되어 새로운 타입으로 판단하는 경우, 타입 후보를 [그림 III-7]와 같이 「한국수어 자료 종합지원 시스템」의 ‘타입관리’ 메뉴 내 ‘타입후보’에 후보로 등록할 수 있다.

한국수어 자료 통합 지원 시스템

수어 말뭉치 관리전사 관리토론 관리타입 관리공유 게시판통계관리오류접수

홈 > 타입 관리 > 타입 후보

타입 후보 목록

타입번호전제

처리상태전제

찾을 대상

타입 관리

타입 후보

타입 변경 목록

타입 삭제 목록

찾기

모두 779건의 결과가 있습니다. (1/78)

10개씩 보기

번호	타입 번호	타입 후보명	처리상태	등록자	등록일시
779	수지 타입 번호	대전2	등록요청	홍	2020.11.16.17:37:53
778	수지 타입 번호	말다12	등록완료	홍	2020.11.16.17:37:26
777	수지 타입 번호	월제어3	등록완료	홍	2020.11.16.17:28:40
776	수지 타입 번호	경쟁1	등록완료	홍	2020.11.16.17:28:18
775	수지 타입 번호	싸우다2	등록완료	홍	2020.11.16.17:27:48
774	수지 타입 번호	네시1	등록완료	홍	2020.9.11.16:36:48
773	수지 타입 번호	오년후1	등록완료	홍	2020.9.11.16:36:30
772	수지 타입 번호	협박1	등록완료	박	2020.9.8.21:03:12

[그림 III-7] 「한국수어 자료 통합지원 시스템」 내 ‘타입후보’ 목록

주석자가 타입 후보로 등록한 타입은 모두 155개였으나 연구팀 논의 과정에서 타입명 중복을 이유로 48개의 타입은 제외시키고, 나머지 107개를 모두 신규 타입으로 등록하였다.

등록에서 제외된 타입들 중에는 이미 등록된 타입의 변이 형태를 새로운 타입으로 인식하였거나, 동일한 수어를 타입명을 달리하여 등록 요청한 경우가 대부분이었다. 수어합성어로 인식하여 [떼어버리다1]과 [시넷물1]로 타입 후보를 등록한 경우도 있었는데, [떼어버리다1]은 [구형4A]와 [버리다1]로, [시넷물1]은 [물1]과 [강1]로 분절하여 주석하였다.

검증자 간 주석 파일 검증 과정에서 논의된 54개의 타입 중 51개의 타입이 추가로 등록되어 경상남도 지역의 말뭉치 주석을 통해 신규 등록된 타입은 모두 158개였다. 타입 목록은 [부록 7]에 제시되었다.

6. 수어 말뭉치 주석 규약

2015년 최초 한국수어 말뭉치 구축을 시작하면서 주석 과정을 정리한 ‘주석 지침’을 제작하여 일관성 있는 주석 데이터를 확보하는 데 사용해 왔다. 이러한 지침은 미리 방향을 설정해 둘 수 있는 것이 아니라 주석 데이터의 유형이 확장되어 감에 따라 새로운 규칙들이 만들어질 수 있는 구조이기 때문에 2020년도부터 그 명칭을 ‘주석 규약’으로 변경하여 사용하기로 하였다. 또한 <표 III-6>와 같이 ‘주석 지침’을 ‘주석 규약’으로 변경하여 수정·보완하는 과정에서 한국수어 말뭉치 구축 가이드를 [부록 2]에 제시하였다. 한국수어 말뭉치 구축 가이드는 원시 말뭉치 수집과 주석 데이터 구축으로 나누어 기술하였다.

<표 III-6> 한국수어 말뭉치 주석 규약의 구성

2019년 주석 지침 체계		2020년 주석 규약 체계	
영역	하위내용	영역	하위내용
I. 기본원칙		1. 개요	1.1. 주석의 기본 원칙 1.2. 번역
II. 총렬		2. 주석 총렬	2.1. 번역 총렬 2.2. 수지 글로스 총렬 2.3. 의미 총렬 2.4. 코멘트 총렬 2.5. 비수지 총렬
III. 수어 단어의 분절			
IV. 타입			
V. 글로스	1. 글로스 확정방법 2. 수와 단위 3. 명사와 동사 4. 합성어 5. 지시(가리키기) 6. 생산적 수어 7. 7. non-tokens 8. 기타	3. 분절	3.1. 분절 원칙
		4. 글로스	4.1. 고정된 수어 4.2. 생산적 수어 4.3. 결합 수어(합성어) 4.4. 기타

VI. 번역		5. 비수지 주석	5.1. 입 움직임 5.2. 기타
VII. 정밀 주석 규칙	1. 입움직임 주석의 기본 2. 총렬 3. 분절 4. 마무딩 관련 총렬의 주석 5. 마우스제스처 관련 총 렬 6. 수어의 분류 관련 총 렬 7. 기타	6. 의미 주석	6.1. 의미 주석의 기본 원칙

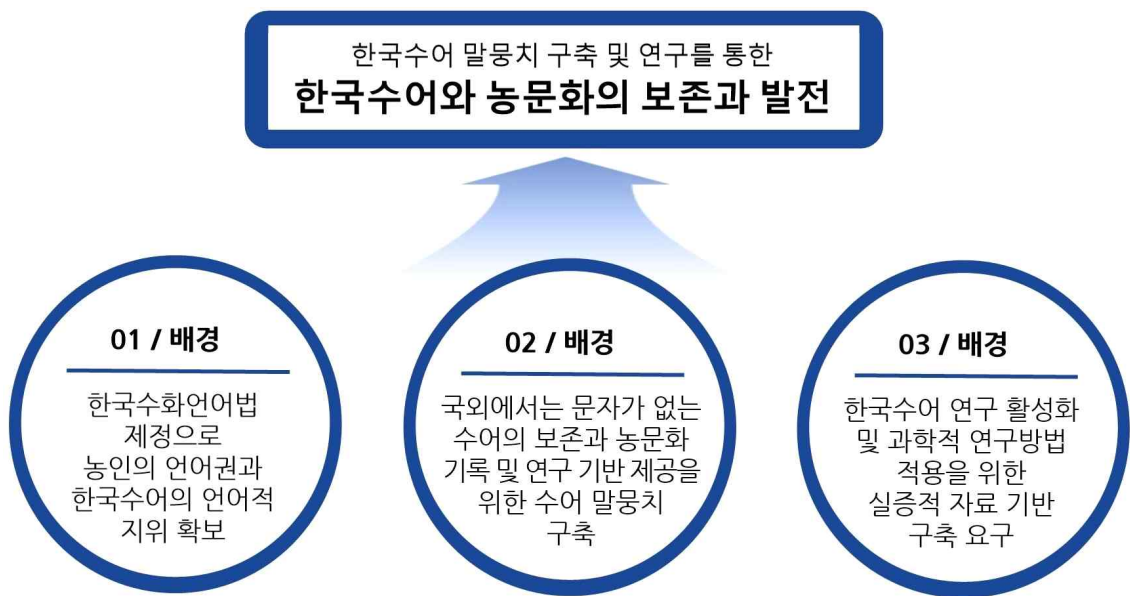
IV. 한국수어 말뭉치 구축을 위한 중장기 발전 계획



1. 한국수어 말뭉치 추진 배경 및 경과
2. 2015-2019 추진 성과
3. 한국수어 자료 실태
4. 해외 수어 말뭉치 구축 탐색
5. 해외 수어 말뭉치 공개 방법
6. 한국수어 말뭉치 구축을 위한 중장기
계획

IV. 한국수어 말뭉치 구축을 위한 중장기 발전 계획

1. 한국수어 말뭉치 추진 배경 및 경과



[그림 IV-1] 한국수어 말뭉치 추진 배경

2016년 한국수화언어법 제정으로 수어가 농인이 사용하는 공용어로 인정받으며 언어적 위치를 확보하게 되었다. 이것은 농인이 모든 생활영역에서 한국수어를 통하여 삶을 영위하고 정보를 제공받을 권리가 있음을 의미하며 수어가 언어로서 계승·보존할 가치가 있음을 의미한다.

1990년대부터 해외 여러 나라에서의 수어 말뭉치 연구는 일상에서 사용하는 언어를 분석 대상으로 하는 현대 언어학적 관점을 수용하면서 공학을 활용한 수어 자료의 기록, 저장, 코딩 및 분석이 활발하게 진행되고 있다. 국내에서는 2007년 국립국어원과 한국농아인협회가 편

찬한 『한국수어사전』이 온라인으로도 구축되어 있으나 실증적인 자료에 근거하지 못한 한계가 있었다. 문자로 표기되지 않는 한국수어의 언어적 특징을 명확하게 규명하기 위해서는 실증적이고 과학적인 연구 기반을 확충하는 것이 필요하다.

이에 2013년 9월에서 2014년 3월까지 국립국어원의 연구 용역으로 한국농아인협회에서 ‘한국수어 코퍼스 구축을 위한 기초 연구’를 수행하였다. 이 연구에서는 해외의 수어 코퍼스 연구 현황 및 사례 분석과 국내의 한국수어 사용자, 연구자, 교수 등을 대상으로 수어 말뭉치 인식 및 요구도 분석을 위해 설문을 하였다. 이 인식 및 요구도 분석의 결과는 한국수어 말뭉치 구축 연구의 타당성을 마련하는 데 기여하였으며, 이를 기초로 2015년부터 현재까지 7년에 걸쳐 매년 한국수어 말뭉치 구축 및 연구가 수행되고 있다.

2014년	2015년	2016년	2017년	2018년	2019년
한국수어 코퍼스 구축을 위한 기초 연구 <ul style="list-style-type: none"> 수어말뭉치 구축 연구 타당성마련 국외 수어코퍼스 연구현황 및 사례 분석 국내 수어 말뭉치 인식 및 요구도 분석 	한국수어 말뭉치 기반 조사 및 시범 구축사업 <ul style="list-style-type: none"> 유도자료 제작 서울·경기지역 농인 60명(30쌍) 대상 90시간 수어 영상 자료 수집 및 번역(약 47시간 12분) 주석(약 13시간 3분) 	한국수어 자료 통합 지원 시스템 개발 <ul style="list-style-type: none"> 한국수어 말뭉치의 효율적인 관리를 위한 시스템 개발·구축·정비 전문가 초청 세미나, 풀뿌리 워크숍 ‘임용직임’ 연구(‘15년 주석본의 112분 53초 분량) 	한국수어 자료 통합 지원 시스템 관리 기능 고도화 <ul style="list-style-type: none"> 토론기능, 검색기능, 주석검증기능, 관리기능 고도화 및 엘란 프로그램과의 연계 고도화 번역과주석(261분 분량) ‘수어동형이의어’ 연구 	한국수어 말뭉치 연구 및 구축 <ul style="list-style-type: none"> 지역의 말뭉치 구축 강원(8명)·경남(4명)·전남(8명)·전북(8명) 총 28명(14쌍) 41시간 12초 수어 영상 수집 기구축 데이터오류 개선 말뭉치 활용 워크숍 	한국수어 말뭉치 연구 및 구축 <ul style="list-style-type: none"> 유도자료 정비 경남지역 22명(11쌍) 36시간 분량의 수어 영상 자료 수집 의미주석 실시(121분 분량) ‘생산적수어의도상성’ 연구
예산: 백만원 사업기간: 5개월	예산: 194백만원 사업기간: 7개월	예산: 380백만원 사업기간: 8개월	예산: 200백만원 사업기간: 7개월	예산: 140백만원 사업기간: 6개월	예산: 150백만원 사업기간: 7개월

[그림 IV-2] 한국수어 말뭉치 구축 경과

2. 2015-2019 추진 성과

한국수어 말뭉치 구축 및 연구 사업은 한국수어 원시 말뭉치 수집과 수집된 말뭉치의 주석 달기와 함께 구축된 한국수어말뭉치 자료를

활용한 ‘연구’ 분야로 구성된다. 현재까지 진행된 한국수어 원시 말뭉치 수집, 원시 말뭉치 주석 달기, 구축된 자료 분석을 통한 사례연구의 추진 성과와 한계 및 시사점에 대해 각각 살펴보았다.

1) 추진 성과

(1) 한국수어 원시 말뭉치 수집

2015년 ‘한국수어 말뭉치 기반 조사 및 시범 구축 사업’에서는 농인 언어제공자의 수어 발화를 촬영한 영상 수집과 발화된 수어 영상을 주석한 자료로 구성되었다. 당시 서울·경기지역의 농인 화자를 대상으로 한국수어 원시 말뭉치가 처음 수집되었다. 이후 수어 자료 축적 및 관리를 위한 시스템을 구축하고자 주석 작업 및 사례연구만 수행하다가 2018년부터 수집을 다시 시작하였고, 2020년 현재까지 3년에 걸쳐 전북, 전남, 강원, 경남 등 4개 지역의 농인 화자를 대상으로 수어가 추가 수집되었다.

언어자료 수집을 위한 동영상 촬영은 외국의 사례를 참고하여 수어 말뭉치 촬영에 적합한 환경을 국내 실정에 맞게 구축하여 진행하였다. 촬영하는 동안 원활하고 효과적인 언어자료를 수집하기 위해 선행 연구 분석 및 2차례의 예비 시험을 통해 수정 보완된 13개의 유도자료 과제를 확정하였으며, 언어제공자는 19세 이상의 성인 농인으로 한국수어를 제1 언어로 사용하며 언어자료 수집 지역에서 10년 이상 거주하는 등의 조건에 충족하는 농인으로 선정하였다. 언어제공자는 촬영 전 언어자료 수집을 위한 질문지를 작성하고 촬영에 동의한 후 2명이 한 조가 되어 유도자료를 소재로 대화를 나누는 것을 촬영하였다. 언어제공자의 자연스러운 발화를 유도하기 위해 유도자료를 사용하지만 각 과제의 시작과 종료, 대화 촉진 등을 위해 언어제공자의 역할에 대해 충분히 숙지한 사회자가 함께 참여하였다. 이런 대화 장면은 2명의 언어제공자를 촬영하는 카메라 각 1대씩과 사회자를 정면에서 전체 화면을 담는 카메라 1대 총 3대의 카메라로 촬영하였다. 따라서 각 과제별 수집 영상 개수는 3개로, 한 쌍의 언어제공자에게 수집되는 영상은 13개 과제 총 39개 영상으로 분할, 편집되었다.

2015년에는 한국수어를 제1 언어로 사용하는 서울지역의 농인 30쌍(60명)을 대상으로 한 쌍의 언어제공자 평균 3시간 촬영하여 확보한 영상은 대략 90시간 분량이였다. 따라서 3대의 카메라로 13개의 과제를 촬영한 약 90시간 분량의 원시 말뭉치 영상은 각각 과제별 영상 파일로 분할, 편집하였다. 언어자료 수집 영상 파일은 총 1,170개이었으며 70시간 16분 분량이였다. 주석 작업을 위해 이 영상들을 엘란 프로그램에 맞는 형식으로 변환하였다.

2015년에는 수어 말뭉치 구축을 위해 활용할 수 있는 지침서와 엘란 한국어판을 개발하였다. 지침서에는 전반적인 수어 말뭉치 구축을 위한 절차와 함께 유도자료 제작 및 제시 방법, 언어제공자 모집 및 선정 방법, 언어자료 수집을 위한 동영상 촬영 및 편집 방법, 주석을 위한 구체적인 방법과 주석 규칙에 관한 내용을 담았다.

2016년에는 한국수어 말뭉치 구축 관련 지침서를 수정·보완하였다. 개정된 지침서에는 전반적인 구축 절차를 표로 제시하고, 언어자료 수집에 필요한 유도자료의 제작에서부터 언어제공자 모집 및 선정, 언어자료 수집을 위한 동영상 촬영과 언어자료의 주석에 필요한 절차 및 「한국수어 자료 통합지원 시스템」 등록 방법에 관한 내용을 제공하였다. 「한국수어 자료 통합지원 시스템」은 수어 말뭉치 영상 관리, 주석 파일 관리, 타입 관리와 같은 한국수어 말뭉치 구축에 필요한 정보 관리를 주 기능으로 하였다. 시스템 상에서 주석 관리를 하는 것은 주석 작업을 분배하고 엘란을 통해 주석한 결과물을 시스템에 탑재하여 자료를 통합 관리하는 것으로서, 이는 통합된 주석 자료를 수어 말뭉치 연구에 활용할 수 있도록 제공하기 위함이다.

2016년에 이어 2017년에도 원시 말뭉치 수집이 이루어지지 않는 않았으나, 「한국수어 자료 통합지원 시스템」이 보완됨에 따라 지침서의 관련 내용을 수정·보완하는 작업을 하였다. 주요 개정 내용은 2017년에 개발한 「한국수어 자료 통합지원 시스템」을 기반으로 수정·보완한 내용을 추가하였으며, 언어제공자에 대한 자료 수집용 질문지를 한국 사정에 맞게 수

정·보완하고, 주석 품질관리와 관련된 내용을 추가하였다.

2018년에는 한국수어 자료 확대를 위하여 2015년에 이어 두 번째로 원시 말뭉치를 수집하였다. 강원, 전남, 전북, 경남의 총 네 곳의 지역에서 총 16쌍의 언어제공자를 대상으로 촬영을 진행하였는데, 언어제공자 선정 기준에 부합하지 않는 자료를 제외하고 최종적으로는 14쌍(28명)의 언어제공자의 원시 말뭉치 자료를 구축하였다. 원시 말뭉치 자료는 지역별로 강원지역 11시간 21분 20초, 전남지역 11시간 28분 22초, 전북지역 12시간 50분 40초, 경남지역 5시간 57분 15초 분량의 촬영이 이루어져 총 41시간 12초 분량의 원시 말뭉치 자료를 수집하였다. 수집한 영상은 각 과제별 영상 파일로 분할 편집한 결과, 강원지역 4번 쌍의 J부터 M까지의 과제를 제외한 총 534개, 35시간 51분 22초 분량을 확보하였으며, 이들 영상은 모두 1차 번역까지 완료하여 「한국수어 자료 통합지원 시스템」에 탑재하였다. 또한, 향후 한국수어 말뭉치 구축의 효율성 확보를 위해 2017년 개정한 한국수어 말뭉치 구축 지침을 재정비하였다. 주요 개정 내용으로 언어 자료 수집에서 언어제공자 관리를 위한 언어제공자 고유번호 및 지역 구분과 관련된 내용을 추가하고, 언어 자료 수집 과정에서 효율적 촬영 환경 조성에 대한 내용을 수정·보완하였다.

2019년에는 한국수어 말뭉치 자료의 지역 균형성을 확보하기 위하여 2018년 수행된 원시 말뭉치 수집에서 언어제공자의 수가 가장 적었던 경남 지역을 대상으로 총 11쌍, 22명의 언어제공자로부터 총 36시간 47분 분량의 수어자료를 확보하였으며, 이를 각 과제별로 분할, 편집하고 검증하는 과정(4회)을 거쳐 완성된 영상은 총 429개, 31시간 48분 분량으로 모두 「한국수어 자료 통합지원 시스템」에 탑재하였다.

<표 IV-1> 한국수어 원시 말뭉치 수집 양적 성과

년도	지역	인원(명)		분량	
		목표	성과	촬영시간	과제별 편집 시간*
2015	서울 경기	60	60	약 90시간	70시간 16분 42초

2018	전북	8	8	41시간 12초	35시간 51분 22초
	전남	8	8		
	강원	8	8		
	경남	4	4		
2019	경남	20	22	36시간 47분	31시간 48분 08초
계	5개 지역	108	110	167시간 47분 12초	137시간 56분 12초

※ 과제별 편집 시간은 언어제공자의 자유발화를 시작 시간으로 산출한 것으로, 매 유도 과제마다 언어제공자의 자유발화를 유도하기 위하여 최초 제공되는 설명 영상 시청 및 사회자의 설명은 제외한 것으로 전체 수집한 시간보다 짧을 수 있음

<표 IV-2> 한국수어 원시 말뭉치 수집 질적 성과

연도	주요내용
2015	<ul style="list-style-type: none"> ✓ 우리나라 최초로 시도한 한국수어 말뭉치 구축 ✓ 한국수어 말뭉치 구축 지침서 개발(수어 말뭉치 구축 절차, 유도자료 제작 및 제시 방법, 언어제공자 모집 및 선정 방법, 동영상 촬영 및 편집 방법)
2016	<ul style="list-style-type: none"> ✓ 「한국수어자료통합지원시스템」 구축 ✓ 한국수어 말뭉치 구축 지침서 수정·보완
2017	<ul style="list-style-type: none"> ✓ 한국수어 말뭉치 구축 지침 수정·보완: 「한국수어자료통합지원시스템」의 수정 보완에 따른 사항
2018	<ul style="list-style-type: none"> ✓ 한국수어 말뭉치 자료의 지역 균형성 확보: 강원, 경남, 전남, 전북의 4개 지역 수어 자료를 확보함으로써 향후 지역 수어 및 변이형 연구를 위한 영상물이 확보됨 ✓ 한국수어 말뭉치 구축지침 재정비(제공자 관리를 위한 넘버링, 지역 구분 관련 내용 추가 제시, 촬영 환경 조건에 대한 지침 수정·보완)
2019	<ul style="list-style-type: none"> ✓ 지역 수어의 추가 확보: 과학적인 언어 연구를 위해 최소 10쌍 이상의 언어 자료가 확보되어야 함에 따라 경남지역 11쌍의 수어 자료를 추가 확보하여 경남지역은 총 13쌍의 수어 자료 영상 확보 ✓ 한국수어 말뭉치 수집을 위한 추가 지침 마련: 질문지 수정·보완 및 온라인 설문 구축, 지역 수어 포착을 위한 유도자료 보완, 언어 간섭 방지를 위한 지역 사회자 배치와 교육

수집된 영상의 지역 및 과제별 현황을 살펴보면 <표 IV-3>과 같다. 3년

에 걸쳐 수집한 원시 수어 말뭉치 자료는 총 5개 지역의 농인 언어제공자 110명(55쌍)으로부터 확보된 것이었다. 지역별로는 서울 70시간 16분 42초, 전북 11시간 34분 36초, 강원지역 9시간 14분 4초, 경남지역 37시간 1분 3초, 전남지역 9시간 49분 47초였다. 과제별로는 H 과제(주제 영역), E 과제(토론), G 과제(특정 사건 묘사) 순으로 23시간 8분 21초, 21시간 26분 25초, 19시간 6분 38초 순으로 가장 많은 시간의 자료를 확보하였다. F 과제(표지판)과 M 과제(농학교 이야기)도 14시간 32분 26초와 14시간 10분 8초의 분량이 확보되었다.

〈표 IV-3〉 한국수어 원시 말뭉치 지역 및 과제별 수집 현황

지역 유도과제	서울 (30쌍)	전북 (4쌍)	강원 (4쌍)	경남 (13쌍)	전남 (4쌍)	합계
A	0:58:44	0:03:28	0:04:15	0:44:57	0:04:59	1:56:23
B	2:40:26	0:32:07	0:22:33	1:04:31	0:26:08	5:05:45
C	5:22:02	0:38:59	1:01:26	2:14:09	0:52:24	10:09:00
D	2:19:05	0:03:54	0:06:40	1:11:34	0:06:44	3:47:57
E	8:52:37	2:42:54	1:19:04	6:56:57	1:34:53	21:26:25
F	5:41:16	0:47:20	1:03:03	5:46:12	1:14:35	14:32:26
G	8:33:03	1:22:47	1:06:49	6:32:01	1:31:58	19:06:38
H	12:59:02	2:13:38	1:32:10	4:54:04	1:29:27	23:08:21
I	2:22:30	0:37:47	0:30:29	1:09:41	0:24:38	5:05:05
J	5:49:44	0:37:34	0:21:19	1:49:51	0:28:27	9:06:55
K	1:51:44	0:11:50	0:16:49	0:44:56	0:16:24	3:21:43
L	4:12:11	0:34:23	0:25:17	1:14:47	0:32:48	6:59:26
M	8:34:18	1:07:55	1:04:10	2:37:23	0:46:22	14:10:08
계	390*3	52*3	48*3	169*3	52*4	698*3
	70:16:42	11:34:36	9:14:04	37:01:03	9:49:47	137:56:12

※ 본 데이터는 2015년 ~ 2019년 수집된 자료의 편집 영상의 시간을 기준으로 하였으며, 한 쌍의 언어제공자의 1개 영상의 시간을 기준으로 하였음(출처: 한국수어 통합 지원시스템)

(2) 한국수어 말뭉치 주석

2015년에 한국수어 영상 자료에 대해 국내 최초로 체계적인 수어 주석을 시도하여 총 13시간 3분 분량의 영상에 대한 주석 데이터가 완성되었다. 전체 주석 결과물의 토큰 수는 약 68,687개, 타입 수는 2,400여 개였다. 또한, 47시간 12분 길이의 한국수어 대화에 대한 한국어 번역을 제공하여 수어를 잘 모르는 사람들도 한국수어 말뭉치에 접근할 수 있도록 하였다.

2016년에는 한국수어 분석 자료를 수어사전 자원이나 해외 수어 자료와 연계하는 방안을 마련하고자 하였다. 이에 대용량 수어 분석 자료를 축적·관리하고 효과적으로 서비스하기 위한 통합지원시스템 구축 사업을 통해 한국수어 영상 주석 시스템과 한국수어 자료의 온라인 검증 관리 시스템을 개발하였다. 또한 한국수어 말뭉치 활용 및 연계 방안을 제시하기 위하여 ‘입 움직임’에 대한 사례 연구와 해외 수어 말뭉치 전문가 초청 세미나, 수어 말뭉치 홍보를 위한 농사회 대상 풀뿌리 워크숍을 수어 말뭉치 구축 연구에 포함시켜 실시하였다. 특히 ‘입 움직임’에 대한 사례 연구를 위해 비수지 중 ‘입 움직임’에 대한 주석 지침을 마련하였다. 첫째, 마우스 제스처는 타입을 추가하였으나 마우딩은 별도의 타입이 없다. 둘째, 마우스 제스처 관련 주석은 마우스 제스처 글로스와 마우스 제스처 분류(class)의 2개의 층렬로 나누어 기록하였다. 셋째, 마우딩 관련 주석은 마우딩을 기입하는 층렬 외에 스프레딩, 마우딩의 완성도, 마우딩의 분류의 4개의 층렬로 나누어 기록하였다. 넷째, 입의 움직임을 동반한 수지 기호의 분류를 기입하는 층렬과 입의 움직임과 관련된 코멘트를 기입하는 층렬을 별도로 만들어서 주석하였다. 그리하여 각 언어제공자별로 ‘입 움직임’을 주석하기 위해서 8개의 층렬을 만들어 주석하였다.

주석 지침에 대한 추가 정비도 실시하였다. ‘지시’ 중 일부 타입(‘지시 1_위치’ 또는 ‘지시 2_위치’, ‘_위치_원형’ 등)의 삭제, 손가락 사이가 붙어서 면을 나타내는 것은 ‘구형 1_면’으로, 손가락 사이가 벌어진 것은 ‘구형 1A_면’으로, 손가락 사이가 붙어서 하나의 선을 표현하는 것은 ‘구형 1_선’으로, 손가락 사이가 벌어져서 여러 개의 선을 표현하는 경우는 ‘구형 1A_선’으로 구분하는 등 생산적 수어의 수형 타입

을 추가하고 토큰으로 처리하지 않아야 할 수어(non-tokens)에 대한 3가지 정의 즉, 메타언어적 언급, 손으로는 수어를 하면서 비수지로 부정을 표현하는 수어, 대화 중 앞에서 발화한 수어를 그대로 모방 반복하는 수어는 토큰으로 처리하지 않는 것으로 추가하였다.

2017년에는 수어 말뭉치 연구를 위한 토큰 통계 기능, 대용량 토큰 정보의 빠른 검색과 다양한 방법으로 검색할 수 있는 기능, 주석 검증을 효율적으로 하기 위한 다수 검증자 기능, 「한국수어자료통합지원시스템」 관리 기능, 엘란 주석 프로그램과 타입 목록 연계를 통한 엘란 프로그램 도구 기능 등의 고도화 작업과 시스템의 안정적 운영을 위한 기능 최적화 작업 등을 수행하였다.

특히 본격적인 한국수어 말뭉치 구축 연구에 대비하여 2015년에 구축된 수어 자료를 2016년에 개발한 「한국수어자료통합지원시스템」에 탑재하면서 발생한 시스템 상의 연결 오류 문제를 해결하였다. 타입 정리 등을 하면서 발견된 분절 오류, 토큰 누락, 지문자 주석 오류, 단순한 손짓의 주석 오류 등 656개의 문제 토큰에 대한 주석 수정 작업도 실시하였다.

2015년 수집한 언어 자료(동영상) 중에서 총 261분 57초 분량의 영상에 대해 우세손과 비우세손에 대한 기본 주석 및 번역을 완료하였다. 주석의 품질 관리를 위하여 본 주석에 들어가기 전 3회의 주석 교육과 공동으로 주석 연습을 실시한 후, 주석 신뢰도를 측정하였다. 3차 신뢰도 검사에서 내적 일치도가 낮게 나타난 토큰들의 특징을 분석하여 한 차례의 추가 주석 교육을 실시한 후 본 주석을 시작하였다. 주석 결과 새로 등록된 타입 수는 1,057개, 발생 토큰 수는 23,637개이었다.

정비 과정에서 발견된 문제점을 보완할 수 있도록 수정된 주석 규칙을 제시하였다. 즉, 정비 과정에서 [지시]와 관련된 타입이 너무 세분화 되어 주석 시 많은 혼란을 가져온 사실이 발견되었다. 이에 주석의 효율성을 높이고 오류 발생 사례를 줄이기 위해 타입명을 16가지에서 8가지로 단순화시켰다. 또한, counting hand와 관련된 타입명, 신체 부위를 나타내는 타입명, 그리고 음운 변이의 형태에 대한 주석 방법을 수정하였다. 그 외 수어

합성어의 분절 방법도 정교화 하였으며, 생산적 수어에 대한 타입명도 정리하여 제시하였다.

2018년도 연구 사업에서는 2015년에 수집한 언어 자료(동영상) 중에서 6개의 영상, 총 재생 시간 82분 54초 분량의 영상을 선정하여 우세손과 비우세손에 대한 번역 및 기본 주석을 완료하였다. 주석은 ‘주석자 교육 - 주석 영상 선정 - 주석 영상 번역 - 주석 영상 구분 - 주석 점검 - 주석 결과 확립’의 순서로 이루어졌으며, 주석 결과 발생한 토큰 수는 총 13,302개이었고, 새로운 타입 201개가 생성되었다. 주석 과정에서 발견된 문제들을 바탕으로 글로스 확정 방법을 정비하고, 한국어 글로스에서 동음이의어로 나타나는 문제에 대한 해결방법을 좀 더 정밀하게 제시하였다. 그 외에 기존의 타입에 대한 정비를 실시하여, 중복 타입을 삭제하고 새로운 타입을 추가하였다.

2019년에는 그동안 문자가 없는 수어를 기록하기 위하여 하나의 수어 발화의 형태를 글로스를 빌어 기록하는 주석을 진행해 왔다. 이것은 음성 언어에서의 발화를 그대로 전사하는 것과는 달리 수어의 형태를 일관된 이름으로 기록하는 일은 주석자 간의 일치도를 확인하는 별도의 과정을 통해 정밀도를 확보해야 했기 때문이다. 기본 주석이 완료된 서울지역의 수어 자료 총 11개(135분 1초 분량)에 대해 기본 주석을 정비하고, [지시]를 제외한 토큰에 대해 의미 주석을 달았다. 그 결과 의미 주석을 기록한 고정된 수어의 토큰 수는 특정 사건 6,934개, 토론 4,723개, 농사회 행사 5,129개, 농학교 1,342개, 주제 영역 1,124개로 모두 19,252개였으며, 생산적 수어는 모두 611개였다. 생산적 수어의 각 과제별 주석 수는 특정 사건 296개, 토론 94개, 농사회 행사 138개, 농학교 42개, 주제 영역 41개로 나타났다. 대화 과제보다 이야기 과제(narration)로 분류되는 특정 사건에서 생산적 수어가 많이 나타났다.

합성어 주석은 둘 이상의 수어가 결합되어 하나의 새로운 의미를 생성하는 것으로 판단되나 음운적 변화가 없거나 실제로 하나의 단어로 합성된 상태인지 의심스러운 경우 두 개의 수어로 분리하여 각각 글로스(GLOSS)

를 부여했다. 두 단어가 결합되어 새로운 의미를 생성하나 합성어로서의 음운적 제약이 나타나지 않는 경우에는 관용적 표현이었으므로 분절을 했다. 총 8,319개의 토큰과 시스템에 등재되어 있는 2,821개의 타입을 전수 정비하였다.

〈표 IV-4〉 한국수어 말뭉치 주석의 양적 성과

년도	주석 유형	주석시간	토큰 수		타입수	번역시간
2015	기본	783분	68,687		약 2,400	2,832분 (약 47시간)
2017	기본	261분 57초	23,637		1,057	261분 57초
2018	기본	82분 54초	13,302		201 (실제 등록 143개)	82분 54초
2019	의미	135분 1초	고정된 수어	19,252	타입 정비 -새타입 28건	-
			생산적 수어	611		
			합성어	161		
계	기본	1,127분 51초 (약 18시간)	105,626		약 3,658	3,176분 51초 (약 53시간)
	의미	135분 1초 (약 2시간)	20,024			

〈표 IV-5〉 한국수어 말뭉치 주석의 질적 성과

연도	주요 성과
2015	<ul style="list-style-type: none"> ✓ 국내 최초 체계적인 수어 주석 ✓ 한국수어 대화에 대한 한국어 번역 제공 ✓ 주석 시스템 구축, 주석자 교육과 양성, 주석 규칙 규정, 엘란 한국어판 개발
2017	<ul style="list-style-type: none"> ✓ 기 구축된 말뭉치 자료를 ‘「한국수어자료통합지원시스템」’에 탑재 및 정비: 주석 파일의 오류 유형 분석, 주석 파일 등재 시 토큰 연결 실패 문제 해결, 타입 정비 ✓ 한국수어 말뭉치 구축 지침 수정 보완: 언어제공자 질문지 수정, 주석 규칙 수정, 주석 품질관리와 관련된 내용 추가 ✓ 우세손과 비우세손에 대한 기본 주석 및 번역 실시 ✓ 입 움직임 주석 실시

	<ul style="list-style-type: none"> - 주석 교육 3회 실시 → 3차례 주석 신뢰도 측정 → 추가 주석 교육 → 4차 신뢰도 검사(80% 이상 일치도) ✓ 한국수어 말뭉치 구축 계획 수립(농사회 전문가 5명, 음성언어 구축 전문가 5명 자문)
2018	<ul style="list-style-type: none"> ✓ 타입 정비 제안 <ul style="list-style-type: none"> - 글로스 확정 방법 정비(동음이의어 문제 해결방법 제시) - 중복 타입 삭제 및 새로운 타입 추가 등 ✓ 주석자 양성을 위한 주석자 교육 실시(4회)
2019	<ul style="list-style-type: none"> ✓ 주석 지침 수정 보완 및 정교화 ✓ 기 구축된 기본 주석 수정 보완(총 8,319개 토큰 정비: 분절 오류, 타입명 오류, 우세·비우세 층렬 및 언어제공자(A와B) 층렬 오류 등) ✓ 의미 주석 기록(고정된 수어, 생산적 수어, 합성어) ✓ 시스템에 등재된 총 2,821개 타입 전수 정비(타입명 중복 삭제, 타입명 수정, 타입 영상 교체, 새 타입 추가 등)

현재까지 주석된 현황을 한국수어 통합 지원시스템에 탑재된 수어 데이터를 기준으로 살펴본 결과는 <표 IV-6>과 같다. 2019년까지 서울 지역 60명의 언어제공자로부터 확보한 언어자료를 중심으로 주석이 이루어졌으며, 22번 언어쌍의 F 과제를 제외한 A 과제(이름), D 과제(넥타이), F 과제(표지판)의 주석은 모두 조금씩이라도 주석이 진행된 것으로 확인되었다. 반면, 서울지역의 B 과제(농담이야기), C 과제(애니메이션 보고 다시 이야기하기), K 과제(홍부놀부 그림 보고 이야기 설명하기) 및 2018년부터 수집된 지역의 수어 데이터에 대한 주석은 전혀 진행되지 않았다. 이는 수어 자료 수집을 위해서 ‘언어제공자 모집-촬영-편집-탑재’의 과정에 4, 5개월의 시간이 필요하며, 주석 작업을 위해 ‘작업자 간 일치도 확인-주석 실시-검증’ 등의 작업에 5개월 이상 소요되기 때문에 6, 7개월의 연구기간 내에 진행되기에는 무리가 있었을 것으로 판단된다. 2018년 4개 지역의 수어 자료 수집 당시에도 번역은 모두 완료하였으나 주석은 수집 지역과는 무관한 서울지역의 주석이 진행되었다.

<표 IV-6> 한국수어 원시 말뭉치 주석 시간 및 토큰 수 현황

지역 유도자료		서울	전북	강원	경남	전남	합계
A (이름)	번역	0:43:19	0:03:28	0:04:15	0:04:49	0:04:59	1:00:50
	주석	(8,553) 0:58:44	-	-	-	-	0:58:44
B (농담)	번역	-	0:32:07	0:22:33	0:20:15	0:26:08	1:41:03
	주석	-	-	-	-	-	-
C (애니)	번역	-	0:38:59	1:01:26	0:22:44	0:52:24	2:55:33
	주석	-	-	-	-	-	-
D (넥타이)	번역	1:31:10	0:03:54	0:06:40	0:02:01	0:06:44	1:50:29
	주석	(10,641) 2:19:05	-	-	-	-	2:19:05
E (토론)	번역	0:30:33	2:42:54	1:19:04	1:34:50	1:34:53	7:42:14
	주석	(14,443) 1:32:30	-	-	-	-	1:32:30
F (표지판)	번역	3:50:54	0:47:20	1:03:03	0:26:07	1:14:35	7:21:59
	주석	(43,724) 5:34:01	-	-	-	-	5:34:01
G (사건)	번역	1:40:45	1:22:47	1:06:49	0:31:31	1:31:58	6:13:50
	주석	(15,695) 1:43:06	-	-	-	-	1:43:06
H (주제)	번역	0:20:46	2:13:38	1:32:10	0:52:56	1:29:27	6:28:57
	주석	(12,182) 1:14:02	-	-	-	-	1:14:02
I (개별)	번역	1:29:11	0:37:47	0:30:29	0:11:59	0:24:38	3:14:04
	주석	(12,686) 1:52:17	-	-	-	-	1:52:17
J (농사회)	번역	0:21:19	0:37:34	0:21:19	0:10:03	0:28:27	1:58:42
	주석	(9,520) 1:06:51	-	-	-	-	1:49:38
K (홍부)	번역	-	0:11:50	0:16:49	0:06:18	0:16:24	0:51:21
	주석	-	-	-	-	-	-
L (옛날)	번역	1:36:44	0:34:23	0:25:17	0:10:53	0:32:48	3:20:05
	주석	(12,192) 2:00:38	-	-	-	-	2:00:38
M (농학교)	번역	0:09:16	1:07:55	1:04:10	0:18:29	0:46:22	3:26:12
	주석	(14,536) 1:52:28	-	-	-	-	2:30:00
계	번역	12:13:57	11:34:36	9:14:04	5:12:55	9:49:47	48:05:19
	주석	(154,172) 20:13:42	-	-	-	-	(154,172) 21:34:01

* 본 데이터는 2015년~2019년 주석 데이터 기준으로 하였음(출처: 한국수어 통합 지원시스템)

* 세부 현황은 [부록 8] ~ [부록 10] 참조

(3) 사례연구

사례연구는 한국수어 말뭉치를 기반으로 하여 실증적이고 과학적인 연구 결과를 제시하였다. 2016년에는 ‘한국수어에서 나타난 입 움직임’에 대한 사례연구를 실시하였다. 수어에서 ‘입의 움직임’은 음성언어에서의 조음자로서의 역할뿐만 아니라, 다양한 문법적 기능을 가지고 있다. 한국수어에서 나타난 입의 움직임을 알아보기 위해 2015년 구축된 수어 자료 중에서 우세손과 비우세손 수어에 대한 주석과 번역이 완료된 영상 총 112분 분량을 주석하고 분석하였다. 마우딩은 한글 철자법으로 주석하였으며, 마우스 제스처는 두 차례의 예비 주석을 통해 입 움직임의 기호를 확정된 후 주석 작업을 하였다.

연구 결과, 우세손 기준으로 수어의 토큰은 10,879개이었고, ‘입 움직임’이 나타난 것은 4,597개였다. 그중 마우딩이 함께 나타난 수어는 3,246개(29.8%)였으며, 마우스 제스처와 함께 나타난 수어는 1,351개(12.4%)였고, 나머지 6,282개(57.7%)의 수어는 입의 움직임 없이 수어만 나타났다. 또한 다른 나라 수어와 마찬가지로 한국수어에도 마우딩에서 스프레딩 현상이 나타났다. 즉, 전체 마우딩 중 7.7%(250개)에서 스프레딩이 나타났으며, 오른쪽과 왼쪽에서 스프레딩이 나타났다. 마우딩 완성도의 경우 대부분 한국수어에서는 모든 구어 음절을 조음하는 경우가 86.7%로 높은 비중을 차지하고 있었으며, 다음으로는 앞 음절만 조음하는 경우로 12.1%, 마지막 음절만 나타나는 경우는 1.1%로 나타났다. 마우딩의 기능을 살펴보면 단순한 마우딩이 대부분(96.2%)이었으며, 마우딩이 해당 수어에 대한 부가 정보를 제공하는 M-add가 3.0%, 해당 수어를 구체화하는 M-spec이 0.3%, 마우딩만 나타나는 Msolo가 0.5%로 나타났다.

마우스 제스처에 대한 결과 분석을 살펴보면, 가장 빈도수가 높은 마우스 제스처 타입은 입을 가장 작게 벌린 O1로 총 151개가 나타났으며, 다음으로는 입을 가장 크게 벌린 O3이 137개이었다. 다음으로는 입을 다물었다가 크게 벌리는 -O3(128개), 입이 좌우로 벌어지고 이가 살짝 보이는

H1(119개), 입을 일자로 다문 -1(77개) 순으로 나타났다. 한국수어에서의 마우스 제스처는 조음소로 작용하거나 손의 움직임을 흉내 내는 E 유형이 67.0%로 가장 많이 나타났다. 다음으로는 마우스 제스처가 수어 의미의 부가적 정보를 나타내는 A 유형이 18.9%로 나타났으며, 얼굴 표정의 일부로 나타나는 W 유형이 11.0%, 입이 입으로서의 역할을 하는 4 유형이 3.1% 순으로 나타났다.

2017년에는 한국수어에 나타난 색 이름의 다양한 형태, 지명의 다양한 형태, 이형동의어를 알아보기 위해 주석 자료와 영상을 분석하였다.

첫째, 한국수어에 나타난 색 이름의 다양한 형태를 알아보기 위해 ‘I 과제(개별 수어의 다양한 표현)’로 유도한 수어 자료 중 2015년에 주석한 52분 17초 분량의 주석 자료를 기반으로 분석하였다. 유도과제에서 제시한 색은 갈색, 분홍, 보라, 주황, 초록, 파랑, 회색으로 7가지였으며, 색과 관련하여 분석한 수어 타입은 총 26개로 [갈색], [노랑], [자주색], [파랑], [하양], [포도], [쥐], 그리고 [분홍]과 관련된 형태 4가지(변이형 포함), [보라]와 관련된 형태 4가지, [주황]과 관련된 형태 8가지, [초록]과 관련된 형태 2가지, [빨강]과 관련된 형태 2가지(변이형 포함)였다. 분석 결과 하양, 빨강, 파랑, 노랑, 회색을 나타내는 수어의 형태는 한 가지인데 비해, 혼합된 색을 나타내는 수어는 그 형태가 다양하게 나타났는데, [빨강1][파랑1][섞다1] 또는 [빨강1][약하다1]와 같이 복합어로 조어되었다. 또한 모든 색에서 공통적으로 지문자를 사용하였다.

둘째, 한국수어에 나타난 지명의 다양한 형태를 알아보기 위하여 총 34분 37초 분량의 ‘L 과제(옛날 수어와 현재 수어)’를 분석하였다. 유도과제의 특성상 해당 타입의 발생 빈도는 중요하지 않아 언어제공자 수로 분석하였는데, 분석 결과는 현재 사용되고 있는 수어 형태가 아닐 수 있다. 분석한 총 파일은 16개로 언어제공자는 32명이었고, 우세손 기준 발생 타입 수는 1,029개, 토큰 수는 8,381개였다. 분석 결과, 변이가 나타난 지명은 강원도, 남대문, 동대문, 부산, 용산, 익산(이리), 일본, 인천, 울산, 속초, 원주, 창원, 포천, 통역, 화성이었다. 9곳의 지명에 대한 22개의 타입만을 분

석한 결과 강원도는 4개의 형태(타입), 용산과 이리/익산은 3개의 형태가 나타났으며, 동대문, 부산, 일본, 울산은 2개의 형태가 나타났다.

셋째, 한국수어에 나타난 이형 동의어를 알아보기 위해 2017년 주석 완료한 총 260분 분량의 영상과 2015년 주석 완료한 영상 중 개별 수어의 ‘I 과제(개별수어의 다양한 표현 과제)’와 ‘L 과제(옛날 수어와 현재 수어)’를 제외한 146분 분량(총 406분)의 영상을 분석하였다. 분석 절차는 다음과 같다. 수어의 글로스는 한국어 어휘를 차용하여 의미를 바탕으로 정하였기 때문에 첫째, 주석 자료 중 같은 글로스를 가진 타입 중 다른 번호를 부여받은 모든 타입명을 검색하여 추출하였다. 둘째, 구체적인 대상을 나타내거나 문장 안에서 완전히 대체가 가능하다고 모어 화자가 인식한 수어만을 분석 대상으로 추출하였다. 셋째, 추출된 어휘 중 언어제공자 수가 5명 미만으로 적은 수어는 제외하였다. 이렇게 추출된 8개의 어휘, 넥타이, 샤워, 수영, 신호등, 아래, 아침, 연애, 오토바이를 나타내는 수어의 형태를 분석한 결과, 넥타이는 4개의 형태가 나타났으며, 샤워, 아침, 연애, 오토바이는 3개의 형태가, 수영, 신호등, 아래는 2개의 형태가 나타났다.

2019년에는 생산적 수어의 수형별 도상성 기법과 생산적 수어에서 나타난 수형의 의미 등 생산적 수어의 토큰에서 나타나는 도상성 기법에 대해 분석하였다. 그 결과 생산적 수어 총 611개 토큰 중 [구형1]이 264개(43.2%)로 가장 많이 나타났으며, [구형2]가 100개(16.4%)로 두 번째로 많이 나타났다. 손이용 기법은 [구형2]와 [주먹형1]에서 나타났으며, 이들 두 수형의 차이는 어떤 사물을 잡는 손의 형태나 꼭 쥐 손의 상태를 나타내고 있었다. 스케치 기법은 [구형1]과 [구형2], [육형3]에서 가장 많은 빈도수를 보였으며, 비석이나 부스, 구멍 튜브 등과 같은 입체적인 형태를 묘사하기 위해 사용하는 경우 [구형1]과 [구형2]로 나타났다. [구형1]이 스케치 기법으로 사용될 때 의미의 범위가 상당히 넓다는 것을 알 수 있었다. 측정기법에서는 사물의 크기, 넓이 및 형태를 측정하기 위해 [구형1]을 주로 사용하는 것으로 나타났다.

〈표 IV-7〉 한국수어 말뭉치 활용 사례 연구 내용

연도	주요내용
2016	‘한국수어에서 나타난 입 움직임’
2017	‘한국수어의 어휘 변이 연구’ (한국수어에 나타난 색 이름, 지명, 이형 동의어 분석)
2019	‘한국수어의 생산적 수어의 의미와 도상성 유형’

2) 한계 및 시사점

(1) 한계

한국수어 원시 말뭉치 수집을 위하여 2015년부터 2019년까지 5개의 지역에서 총 110명의 언어제공자를 대상으로 약 167시간 분량의 언어 자료를 촬영하였으며, 수집한 영상은 한 조의 언어쌍을 촬영할 때마다 3개의 영상이 확보되기 때문에 각 과제별로 분할 편집한 영상으로는 총 4,278개 총 137시간 56분 12초 분량이 2019년 말까지 수집되었으나 다음과 같은 한계를 지닌다.

첫째, 전국적 지역별 수어 특성과 변이를 확인하는 데 한계가 있다. 말뭉치 자료의 지역 균형을 확보하기 위해서 2015년 서울을 시작으로 2018년 전북, 전남, 강원, 경남 4개 지역의 원시 말뭉치를 수집하였다. 지역을 확대하였다는 것에 의의는 있으나 각 4쌍 또는 2쌍만의 자료를 확보함으로써 지역별 변이를 확인할 수 있을 만큼 충분한 분량의 언어 자료가 확보되었다고 볼 수 없었다. 이에 2019년에는 경남지역의 언어 자료 추가 수집이 있었으며, 호남과 강원, 충청, 제주 지역에 대한 추가 수집이 필요한 상황이다.

둘째, 불안정한 촬영 환경으로 지역 수어 자료 수집 과정이 원활하지 않았다. 충분한 촬영 시간 확보와 관련된 내용이 2017년에 유의점으로 제시된 이후에도 2018년과 2019년에 걸쳐 꾸준히 문제가 제기되고 있다. 촬영

환경에 대해서도 촬영실 크기, 카메라 메모리 용량, 전원 연결 문제 등 촬영 기반 확충에 대한 개선의 필요성이 계속 제기되고 있다.

셋째, 주석 전문가가 안정적으로 공급되지 않아 주석의 일관성과 신뢰성 확보에 어려움이 발생하였다. 문자가 없는 수어 자료를 주석할 때 발화된 수어에 글로스를 다는 것으로 발화된 소리를 그대로 기록하는 음성언어의 전사 기록 형태와는 다르게 주석자가 동일한 형태의 수어를 동일한 타입으로 인식하여 일관성 있게 기록하는 것은 많은 시간과 훈련을 필요로 한다.

수어 주석의 일관성 및 신뢰도를 높이기 위해서는 역량 있는 주석 인력의 확보가 무엇보다 선행되어야 한다. 또한 수어 자료에 대한 기본 주석은 수어 말뭉치 수집의 가장 기초적인 작업으로 수어 말뭉치가 확대 구축되기 위해 조속히 시행되어야 한다. 국내 최초로 한국수어 영상 자료에 대한 수어 주석을 시도한 2015년 이후 해마다 주석 작업의 중요성과 주석 인력 보충에 대한 제언을 계속해 왔는데, 국립국어원 내에 주석 전담 인력을 채용하여 지속적인 주석 작업이 이루어질 수 있도록 하는 방안이 마련되어야 할 것이다.

넷째, 「한국수어 자료 통합지원 시스템」의 안정성이 취약하며, 작업의 효율성이 떨어진다. 한국수어 말뭉치 연구 및 구축 사업은 「한국수어 자료 통합지원 시스템」으로 수집 영상 및 주석 파일을 관리하고 있다. 그래서 수어 말뭉치 수집 과업에서는 영상 납품 후 언어제공자, 과제별 편집 파일을 등록하는 일, 주석 달기 과업에서는 작업자가 시스템에서 주석 파일 다운로드와 업로드하기 과정의 수시 반복, 수정된 타입명에 대한 각 토큰별 자동 반영 등의 업무를 원활하게 지원하기 위한 시스템이지만 작업 과정에서 작동이 일시적으로 중단되어 멈춤 현상이 일어나는 문제가 반복적으로 일어나고 있다. 따라서 연구의 효율성을 위해 작업 중에 발생하는 문제를 해결하고 시스템 운영 방식을 개선하는 등 시스템 정비가 필요하다. 최근의 말뭉치 자료 구축 및 활용은 빅데이터 기반의 환경을 고려하여 개선되고 있으므로 한국수어 말뭉치도 이러한 기반 기술과의 결합이 필수적이다. 향후 연구자들과 일반인들을 위한 주석 정보를 검색, 활용할 수 있는 시스템으로 그 기능을 개선할 필요가 있다.

(2) 시사점

첫째, 지역수어의 원활한 수집을 위해 시·도 농아인협회와의 협력 관계를 구축할 필요가 있다. 지역 균형적 한국수어 말뭉치 수집과 한국수어의 지역별 변이에 대한 정보 확보를 위해서는 전국 시·도 농아인협회와 협력이 필수적이다. 지금까지 수집된 수어 말뭉치 역시 언어제공자의 모집과 선정, 촬영, 편집 등의 과정에서 지역 농아인협회의 협조와 더불어 지역 농사회의 도움을 받아 진행되었다. 앞으로도 지역 농아인협회와 긴밀한 협조가 지속적으로 필요하다.

둘째, 전문 능력을 갖추고 한국수어 말뭉치 구축 사업 이해를 하는 전문 촬영팀 확보가 이루어져야 한다. 한국수어 말뭉치 수집에서의 핵심적인 업무는 촬영과 편집이다. 현재까지 충분한 촬영 시간 확보, 적절한 촬영실 크기, 촬영에 충분한 카메라 메모리 용량, 전원 연결 등의 촬영 환경에 여러 가지 문제가 나타났다. 이는 수어 수집을 위해 지역을 옮길 때마다 그 지역의 촬영팀 수준에 따라 촬영 결과물의 수준이 달라져 품질 유지에 상당한 위험 요인으로 작용하고 있다.

아울러 충분한 촬영 시간 확보를 위해서는 사회자와 촬영 및 편집 담당자의 역할도 매우 중요하다. 따라서 수집 지역이 확정되면 해당 지역의 언어 자료 수집에 적합한 사회자가 촬영 전 충분히 연습해 볼 수 있도록 2-3차례의 교육이 시행되어야 하며, 전문 촬영팀이 이를 뒷받침 해줄 필요가 있다.

셋째, 장기적 안목에서 주석 전문 인력을 양성해야 한다. 한국수어 말뭉치 구축은 장기적인 식견으로 안정적으로 이루어져야 하는 사업이다. 해외의 사례에서 알 수 있듯이, 독일의 경우는 2009년부터 시작하여 15년 동안의 장기계획을 통해 350~400시간의 언어자료를 수집하였고, 네덜란드는 2006년부터 2년간 각 지역의 수어 자료를 수집하였으며, 그 후 2011년부터 현재까지 3명의 전일제 주석자들이 주석하고 있다. 국립국어원은 주석 전문 인력을 채용하여 지속적인 주석 작업을 해야 하고, 질 높은 주석자 확

보를 위해 수어 주석 연수회를 개최할 필요가 있다.

넷째, 축적된 수어 자료의 공개를 위한 계획 수립이 필요하다. 그동안 수집·분석된 한국수어 자료의 양이 공개를 해야 될 수준에 이르고 있다. 현재 진행되고 있는 원시 말뭉치 수집과 주석 처리가 된 자료들의 시스템 탑재, 그리고 이 자료들의 공개 범위와 방법에 대한 논의와 계획을 해야 될 시점이다. 이를 위해 해외 사례를 분석하여 그 장단점을 파악한 후, 우리나라에 적절한 공개 범위와 방법을 결정해야 한다.

다섯째, 수어 데이터베이스를 활용한 사례연구가 활발하게 이루어져야 할 것이다. 「한국수어 자료 통합지원 시스템」을 기반으로 한국수어 말뭉치 데이터베이스가 구축되면 한국수어에 대한 실증적인 연구가 가능해져 과학적이고 체계적인 연구의 틀을 갖추게 될 것이다. 아울러 증거 기반의 수어 교육을 위한 교육과정 개발, 교재 개발 등이 가능하게 될 것이다. 이런 가능성에 입각하여 수어연구와 수어교육을 위한 활발한 활동이 이루어져야 할 것이다. 아울러 한국수어 말뭉치 사업을 통해 국제 학술 활동과 한국수어 말뭉치 구축의 효율성과 그 효과의 파급 및 홍보를 위해 다른 나라들과 네트워크를 형성하고 교류할 필요도 있다.

3. 한국수어 자료 실태: 농아방송과 영상도서를 중심으로

말뭉치는 자연언어 연구를 위해 특정한 목적을 가지고 실제 사람들 사이에서 사용되는 언어의 표본을 추출한 집합체이다. 세계의 많은 음성언어들은 컴퓨터의 발달로 말뭉치 기반의 연구가 용이해졌다. 말뭉치 기반 연구는 분석의 정확성을 위해 해당 자연언어를 목적별로 분석하거나 확률상의 분포나 시계열적 접근 등을 통해 전체적으로 파악하는 것도 필요하다. 특히 언어의 빈도와 분포를 확인할 수 있는 자료로서 현대 언어학 연구에 필수적이다. 이에 음성언어 연구에서는 자연발화를 중심으로 한 말뭉치 구축뿐만 아니라 텍스트 말뭉치, 한국어 학

습자 말뭉치 등을 장르별로 구축하고 있다. 이에 반해 한국수어 말뭉치는 2015년부터 농인으로부터 수어 자료 수집 및 주석을 통해 구축되고 있으나 한정된 자원으로 수어자료를 수집하는 데 많은 제한점이 있었다.

현재 우리나라는 농인에게 수어로 정보를 제공하도록 각종 법률로 규정하고 있으며 현대 사회의 다양한 의사소통 수단, 특히 소셜 미디어(social media)의 발달로 농인의 수어 발화 자료(영상)가 꾸준히 축적되고 있다. 이처럼 많은 정보 자원으로부터 생성되는 수어가 한국수어 말뭉치 자료로 활용 가능한지를 알아보기 위하여 국내외 수어 자료 현황을 파악할 필요가 있다. 이를 위해 농인이 직접 제작·운영에 참여하고 있는 공식적인 수어 자료 중 역사와 대표성이 있는 한국농아인협회의 한국농아방송과 국립장애인도서관의 수어영상도서 자료를 중심으로 한국수어 자료 실태를 살펴보았다.

1) 해외 수어 자료 현황

독일에서는 디지털 데이터 말뭉치 기반으로 수어 텍스트 유형을 결정하는 「DGS.txt.typ」 프로젝트를 수행하기 위하여 마그데부르크(Magdeburg) 대학교의 ‘수어 디지털 도서관’을 중심으로 수어 자료를 수집하였다. 1989년부터 1996년까지 호르스트 에빙하우스(Horst Ebbinghaus)와 옌스 헤스만(Jens Hessmann)이 대학교 수화통역과의 수어 자료로 수집한 것들이 대부분이다. 그 외 수어 설교 등과 같은 자료가 필요할 때에는 새롭게 촬영하기도 하였다. 수어 자료는 대략 VHS tape 550개, DVD 20개, DV tape 13개(총 920시간) 분량이다. 이때 메타데이터에는 자료가 어디서 어떻게 수집되었는지, 언어제공자의 정보, 수어자료의 내용(주제), 영상의 길이 및 기술적 질(technical quality)과 같은 정보가 포함되었다.

구축된 자료를 살펴보면 자주 나타나는 주제가 있고 전혀 나타나지 않는 주제들이 있다. 예를 들어 농 방송국에서 방송된 아이들을 위한 수어

이야기는 많지만 의료 및 법원 주제를 가진 자료는 전혀 없어 추가 구축이 필요한 것으로 보인다. 수어자료 구축 외 목적은 등록부(register) 및 텍스트 종류 분석이다. 등록부는 상황에 맞는 언어를 말한다. 일반적으로 형식적(formal), 중립적(neutral), 비형식적(informal)인 것으로 구분이 된다. 등록부에 따라 어휘, 문법 그리고 텍스트 구조가 다를 수 있다.

텍스트 종류로는 논픽션(non-fiction)과 픽션(fiction)으로 나누어진다. 논픽션이 편지, 토론, 회의, 강연, 설교 같은 종류라면 픽션에는 소설, 동화, 전설, 시, 연극 등이 있다. 등록부와 텍스트 종류를 분석하기 위해서 부분적으로 수어 자료에 주석을 단 것도 있었다(60분 정도). 과거에는 주제를 통해서만 특정한 텍스트를 찾을 수 있었다면 현재는 텍스트 종류로 검색이 가능하다. 예를 들어 ‘역할전환(role-shift)’의 빈도수는 텍스트의 종류에 따라 다르다. 전문 강연 같은 경우는 역할전환 형상이 짧게 나타나거나 이야기 및 비공식적인 대화보다 빈도수가 낮을 것으로 예측된다.

라드바우드 대학교(Radboud University)의 크레스본(Crasborn, 2009) 박사가 주축이 된 연구팀은 네덜란드 과학 연구 기구(Netherlands Organization for Scientific Research, NWO)에 의해 마련된 기금으로 3년간 수어 언어학 말뭉치 네트워크(Sign Linguistics Corpora Network, SLCN)를 구축하여 전 세계 수어 말뭉치 및 관련 방법론에 대한 자료를 모으고 공유했다. 이 네트워크의 목표는 수어 말뭉치의 생성 및 개발에 대한 기존 지식을 정리하고, 향후 수십 년 동안 연구의 질을 높이기 위한 방법을 안내하는 데 있었다. 이를 위해 지난 2009년과 2010년에 걸쳐 ‘자료 수집’, ‘메타데이터’, ‘주석’, ‘활용’ 등 4개의 주제로 소규모 국제 워크숍을 개최하였다. 이 중 호주, 네덜란드, 영국, 독일, 자메이카, 아일랜드, 프랑스 등의 국가에서 실시된 ‘자료 수집’에 관한 내용이 발표되었다.

프랑스의 수어 말뭉치는 말뭉치 자료, 메타 자료(언어 정보), 주석 등 풍부한 자료를 가지고 있다. 그러나 산재된(Legacy) 말뭉치, 연구 중심 자료, 응용 중심(교육학) 자료 등 여전히 자료들이 이질적으로 산재해 있다. 유산 말뭉치 자료라 함은 과거 농 공동체에서 개최된 컨퍼런스나 워크숍의

비디오 영상물을 의미한다. 이들 자료에는 누리집을 통해 획득한 온라인 자료로서 말뭉치로 활용하기에 이질적이며 접근이 제한된 것도 있을 뿐만 아니라, 청각장애 아동의 말뭉치나 소수 집단의 수어 등도 있다.

연구 중심 말뭉치에는 콜린(2002)의 수어, SNCF 말뭉치(LIMSI-WebSourd, 2008-2009), CREAGEST(2007-2012), 인터뷰 ‘농인과 쓰기’ (Garcia et al., 2004), 소수 집단의 수어로서 브라질 수어(Fusellier, 2004), 모리셔스 섬 수어(Gebert, Sallandre & Fusellier, 2006), 세네갈 수어(Jirou), 시간성 말뭉치((Fusellier & Sallandre, 2005), 농아동 말뭉치(Morgenstern, 1996) 외에도 연구 목적으로 배포가 안 되는 개별적인 말뭉치로서 농 아동과 청인 아동(Blondel, Limousin, Jacob), 농학생 말뭉치(Licence Pro, DPCU 아카데미 프로그램) 등이 있다. 응용 중심 말뭉치로는 비영리 유럽 수어센터에서 관리하는 프로젝트로 약 400,000개가 넘는 수어를 수집하고 사전 형태로 문서화한 Spreadthesign(<https://www.spreadthesign.com/en.gb/search/>) 등이 있다.

자메이카 농인에게 언어를 수집하는 것은 영어, 자메이카어 크리올, 자메이카 수어와 국가 사인(Country Sign) 등 다언어 상황(multilingual situation)에서의 언어를 수집하였으며, 이러한 자료는 농인과 청인과의 일대일 대화, 농인과 청각장애인 간의 일대일 대화, 학술적인 상황과 사회적 상황에서의 집단 토론, 비공식적인 모임, 유치원에서 농인 보조인과 청인 교사 사이의 교실 상호작용, GPS 매핑, 문자 메시지 등을 통해 수집하고 있다(SLCN, 2009).

2) 국내 한국수어 자료 현황

국내 한국수어 자료는 비공식적으로 농 관련 기관을 중심으로 각 기관이 행사를 기록하기 위해 촬영한 비디오 영상 자료, 개인 연구를 위해 연구자가 개별적으로 수집한 수어 자료 등이 있다. 공식적인 수어 자료로 대표적인 것은 한국농아방송(iDBN)의 수어방송자료와 국립장애인도서관의 수어영상도서가 있다.

한국농아방송은 한국농아인협회에서 운영하고 있는 수어 방송이다. 2004년 초창기부터 한국수어 사용자인 농인 앵커가 한국수어를 방송언어로 활용함으로써 국내외 농인의 방송미디어 접근을 용이하게 하는 데 일조하고 있다. 주로 농사회 이슈, 국내외 핫이슈, 농인 관련 복지 정보 및 문화 행사 정보 등에 관해 보도하는 ‘수어뉴스’, 국제사회에서의 장애인복지 지원 현황 및 인권 선진국과 후진국에서의 농사회와 농인의 삶에 대해 영상으로 기획 제작하는 인터뷰 형식의 ‘특집방송’으로 구성되며 총 4명의 앵커가 진행하고 있다. 2004년 엄태욱, 이숙기, 양홍석 등 초창기(2004년~2009) 앵커를 거쳐 2019년부터 이은영, 지미경, 황은진, 박민호 등 4명의 앵커가 진행하고 있다. 2019년 1년간 방송된 영상은 총 60여건이었으며, 이 중 수어뉴스가 55건으로 대다수를 차지한다(<http://cafe.daum.net/deafon>).

〈표 IV-8〉 한국농아방송 2019년 유형별 영상 목록

번호	유형	영상 제목
1	뉴스 기사	- 서울시 숨은 영웅 찾기 추천 공모중
2		- 조용한 택시 감동 물결
3		- 일본 청각장애인 불편 많아
4		- 구글, 청각장애인 위한 앱 개발
5		- 제11대 임원선거 “중앙선관위” 해산으로 잠정 보류
6		- 강원도 산불 국가재난지역 선포,, 속초시지역 속초농아인교회전소
7		- 가치봄 브랜드로 새단장
8		- KBS뉴스9 수어방송 어렵다
9		- 청와대 초청 간단회서 청년문제 대변
10		- 아이돌보미
11		- 방통위 및 방송3사 인권위 진정
12		- 장애인 단체, 농인 탑승 거부 놀이공원 국가인권위 차별 진정
13		- 농인맞춤법 TV 무료 보급,,주민센터서 신청 가능
14		- 재난방송시 수어전문 인력 풀 확대

15	...	- 전국농아인대회 현황리에 개최
16		- 한국농아인협회, 정부에 서비스 지원 종합조사표
17		- 제11대 중앙회장 변승일당선! 농아인의 행복위해 앞장서길!
...		...
55	인사말	- 한국농아인협회, 업무협약 통해 일자리 창출 희망의 씨앗 심어
56		- 신년사-2019년 기해년 한국농아인협회 중앙회장 이대섭
57		- 한가위 추석 인사말-변승일 중앙회장
58	인터뷰	- 9월23일 세계 수어의 날 - “세계 수어의 날을 축하하며
59		- 변승일 회장 취임식 인터뷰
60		- 서울농학교-인터뷰

① 뉴스기사 영상 예시



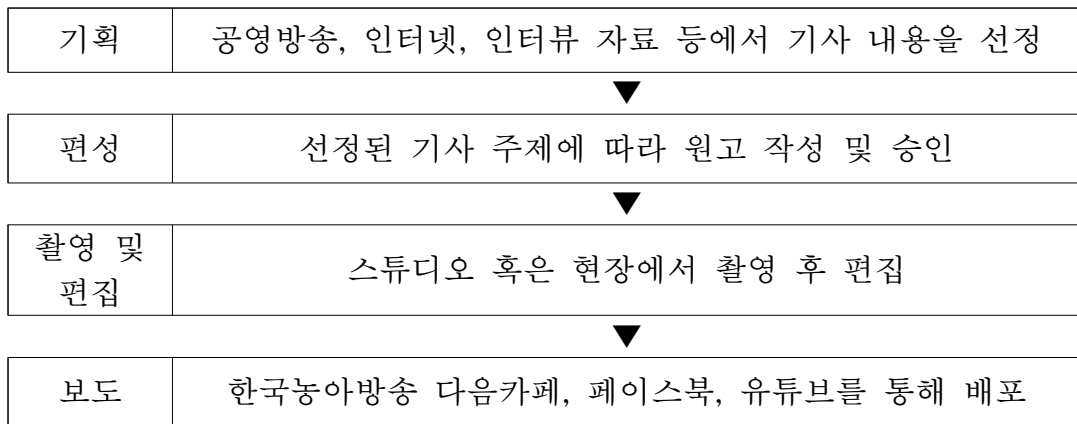
② 인사말 영상 예시



③ 인터뷰



한국농아방송은 기획 단계에서 공영방송, 인터넷, 인터뷰 자료 등에서 기사 내용을 선정한 후, 텍스트로 원고를 작성하고, 촬영 및 편집과정을 거쳐 시청자들에게 전달된다. 제작 배포 과정은 다음과 같다.



[그림 IV-3] 한국농아방송 제작 과정

국립장애인도서관의 수어 영상 도서(구. 수어영상도서자료)는 한국수어를 제1 언어로 사용하는 농인에게 한글 텍스트로 되어 있는 단행본, 연속간행물 등 활자로 된 자료에 접근할 수 있도록 한국수어로 변환하여 제작한 영상자료이다.

국립장애인도서관에서는 수어 사용 농인의 대체자료를 확충하고 평등한 도서 서비스를 제공하기 위하여 2007년부터 수어 영상 도서를 제작·보급해 오고 있다. 첫째 20편을 시작으로 점차 수량을 늘려 2018년에는 759건에 달하였고, 2019년에는 343건을 제작, 현재까지 총 3,818점에 달한다. 주제별로는 ‘총류, 철학, 종교, 사회과학, 순수과학, 기술과학, 예술, 언어, 문학, 역사’ 등이 있으며 그중에 문학이 56.5%로 가장 많다.

〈표 IV-9〉 국립장애인도서관 수어 영상 도서 주제별 제작 현황

	총류	철학	종교	사회과학	순수과학	기술과학	예술	언어	문학	역사
건수	40	236	24	637	181	233	54	11	2,153	249
퍼센트	1.0	6.2	0.6	16.7	4.7	6.1	1.4	0.3	56.4	6.5

(출처: 국립장애인도서관 누리집, 2019)

한국수어 영상도서는 크게 완역도서와 요약도서로 구분된다. 완역도서는 도서의 본문(텍스트) 전체를 한국수어로 표현한 것을 의미하며, 주로 아

동도서예 다양한 삽화 및 애니메이션 효과를 넣거나 한국수어 구연가의 분장과 의상, 소품 등을 최대한 활용하여 쉽게 내용에 접근할 수 있도록 제작한 것이다. 요약도서는 도서의 본문(텍스트) 전체를 한국수어로 표현하기 어려운 경우 도서의 내용을 논리적 전개와 지식전달의 효과성을 고려하여 함축·요약하고, 삽화 등의 효과를 삽입하여 제작한 것이다(국립장애인도서관 자료개발과, 2017). 2019년에 제작된 수어영상도서 총 343건 중 196권이 완역도서로 출간되었다.

〈표 IV-10〉 국립장애인도서관 수어 영상 도서 2019년 제작 현황

구 분		완역된 책 수	영상 건수
한국 전래 동화		132	188
세계 명작 동화		56	112
의뢰	보건	1	26
	법률	1	
기타	여행	4	17
	시	2	
합계		196 권	343 건

(출처: 국립장애인도서관 누리집, 2019)

수어영상도서로 완역된 책 196권 중 188건이 문학이고, 그중에 아동을 대상으로 하는 동화가 96%로 가장 많았다. 대표적 작품으로 ‘솔 안에 든 거인’, ‘수탉과 돼지’, ‘숙향전’, ‘신기한 나뭇잎’, ‘어우야담’, 등과 같은 한국 전래 동화가 있으며, ‘세익스피어 이야기1’, ‘솔로몬의 동굴’, ‘신드바드의 모험’, ‘장발장’, ‘정글북’, 등과 같은 세계 명작 동화가 있다.

수어영상도서는 국립장애인도서관 누리집(<http://nld.nl.go.kr/>)에 회원 가입하면 무료로 이용할 수 있다. 유튜브 국립장애인도서관 채널에서도 접근 가능하지만 국립장애인도서관 소장 자료의 일부만 이용할 수 있다.

구분		예시 자료				
한국 전래 동화	교재 표지					
	영상 캡처					
세계 명작 동화	교재 표지					
	영상 캡처					

[그림 IV-4] 국립장애인도서관 수어 영상 도서 동화 예시

수어영상도서는 도서 선정의 타당성 검토부터 도서내용에 대한 시나리오 작업과 텍스트화된 시나리오를 다시 농인 수어자가 수어로 옮기는 과정을 거쳐 제작된다. 시나리오에는 1인 서술형식, 2인 이상의 극형식과 서술형식을 조합한 형식으로 되어 있어 수어 자료로 사용하려고 하는 경우, 시나리오라는 한정된 범위 안에서 인위적인 발화 자료이기는 하나 단순 기술뿐만 아니라 상호작용의 형태의 수어를 볼 수 있다. 다만 원 도서의 내용이 수어번역본에 충실하게 반영되었는지는 여러 가지 측면에서 재고의 여지가 있다.

시나리오 단계	도서선정 / 도서내용 요약 / 한국수어 번역 및 시나리오 제작
시나리오 단계	시나리오 1, 2차 검수 및 시나리오 관리 인력 투입

시나리오 단계	한국수어 구연가 인력 관리 및 투입(사전교육 실시)
▼	
영상 제작	촬영 · 편집(자막 포함) · 제작

[그림 IV-5] 수어 영상 도서 제작 과정

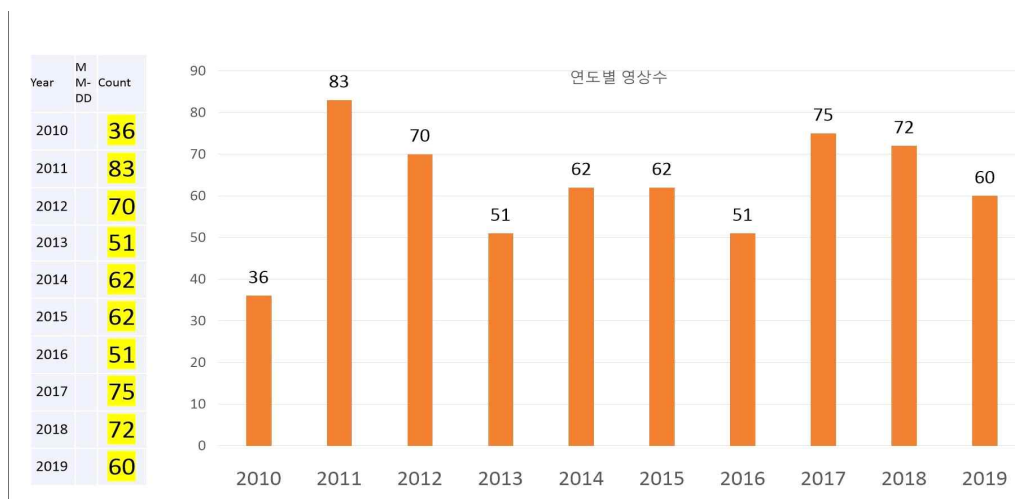
언어 연구에서 기본적인 연구방법은 자연스러운 관찰이며, 특히 다른 대화자와의 자발적으로 이루어지는 언어적 상호작용을 관찰하기 위해서 녹화는 필수적이다. 그런 면에서 한국농아방송과 수어 영상 도서는 오랫동안 축적된 자료가 많기 때문에 다량의 언어 자료를 확보한다는 차원에서는 의미 있는 일이다. 그러나 한국농아방송의 수어자료의 경우 정보 전달을 위해 앵커가 한국어 텍스트로 제작된 스크립터를 보고 촬영함으로써 번역 차원의 언어자료의 성격을 지니고 있을 뿐만 아니라, 한국어 어휘와 문법의 영향을 배제하기 어렵다. 특히 국립장애인도서관이 소장한 수어 영상 도서는 2019년 기준 3,818건으로 상당히 많은 자료에 속한다. 그러나 수어 영상 도서의 경우에도 수어 구연가의 선정 기준이 모호하며, 수어 구연가가 한국수어로 촬영하기 전에 한국어 텍스트 시나리오를 통해 번역하는 과정이 거침으로 인해 한국어 구조에서 완전히 탈피하지 못하고 있다는 사실이 다소 확인된다.

이러한 자료가 한국수어 말뭉치에서 언어학적 자료로 사용되기 위해서는 언어제공자의 초상권, 언어자료 저작권 등 연구윤리에 위배되는 요인을 제거하는 것이 필요하며, 구체적인 사례를 제시하기보다는 언어자료를 목록화 하는 것을 우선적으로 수행해야 할 것이다. 또 언어학적으로 사용되기 위해서는 기본 언어자료 외에도 언어제공 관련 메타데이터가 확보되어야 한다. 따라서 본 연구를 수행하는 동안 언어자료의 목록, 언어자료의 특성과 유형, 활용할 수 있는 연구 영역에서 확보되어야 할 메타데이터 등의 분류 기준을 마련하여 목록화 하는 작업을 수행할 필요가 있다.

3) 국내 수어 자료 실태

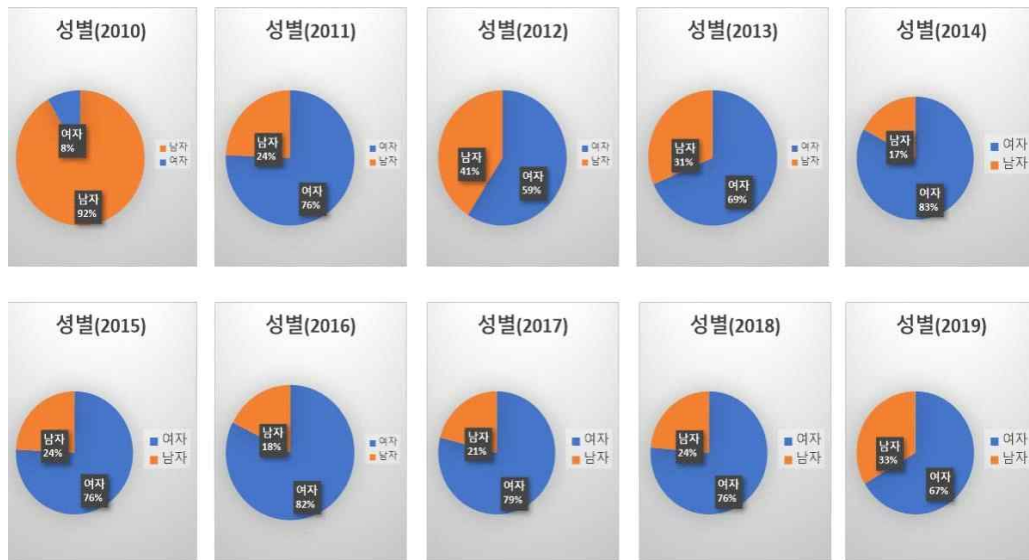
(1) 방송에서의 수어 자료 실태

방송에서의 한국수어 자료 실태 분석을 위하여 한국농아방송에서 제공되고 있는 영상을 수집하였다. 한국농아방송 다음 카페, 유튜브 등에서 제공되고 있는 영상은 2010년부터 2019년까지 총 622편이었다. 이 숫자는 한국농아방송이 제공하는 플랫폼에서 확인된 자료 중 “연결된 영상을 확인할 수 없다”라는 오류 메시지가 뜨는 것을 제외한 수치이다. 2010년도에는 현장 기록 형식의 영상물이 많이 나타나는 편이었다.



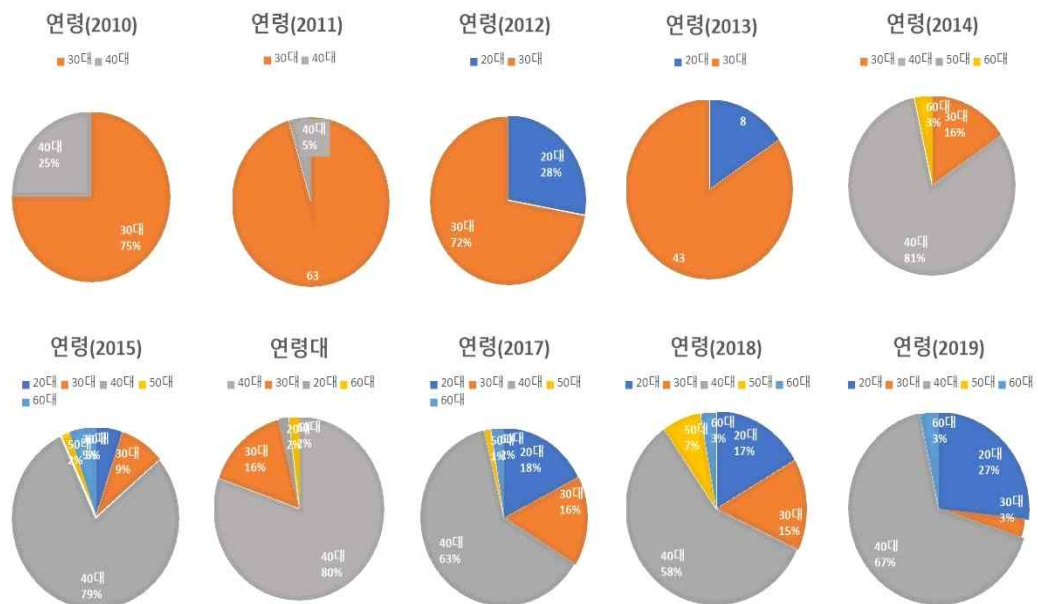
[그림 IV-6] 한국농아방송 연도별 영상 수

뉴스에 나타나는 진행자의 성별 분포는 2010년 남성이 92%로 가장 많은 분포를 보였으나 2011년부터 2019년까지는 여성이 59~83%로 매년 남성보다 많은 분포를 보이고 있다.



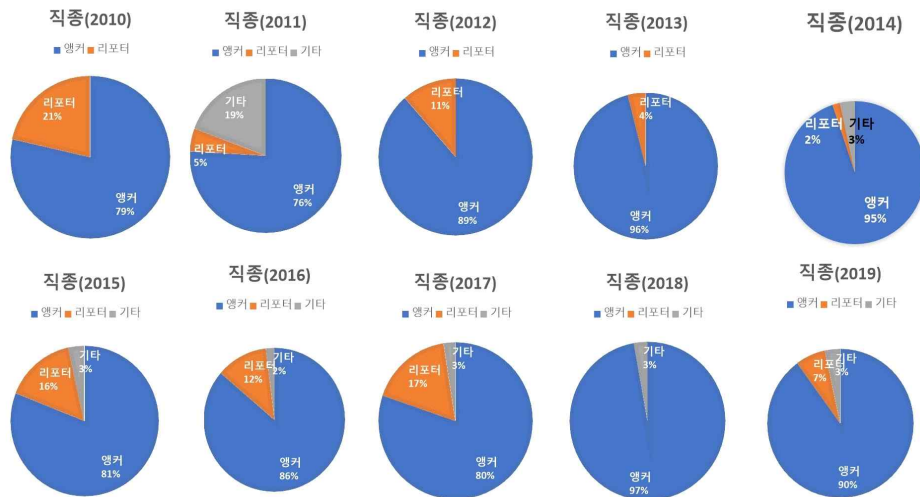
[그림 IV-7] 뉴스 진행자 성별 비율

뉴스 진행자의 연령 분포를 보면 2013년 이전에는 30대가 72~95%로 높은 분포를 보였으며 2014년 이후 40대가 58~81%의 분포로 나타났다. 특히 2017년 이후 40대가 다시 증가하는 추세를 보였다.



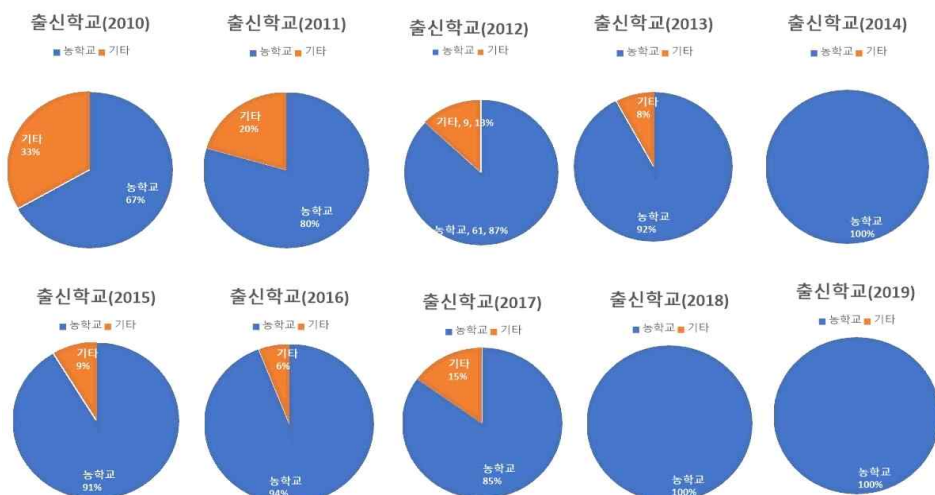
[그림 IV-8] 뉴스 진행자 연령 비율

뉴스 진행자의 역할을 살펴보면, 현재 한국농아방송은 전체 뉴스의 75% 이상이 앵커에 의해 진행되는 뉴스로 현장 중계보다는 스튜디오 내 앵커가 담당하여 진행하는 것이 더 많은 것으로 나타났다.



[그림 IV-9] 뉴스 진행자 역할 비율

뉴스 진행자의 출신학교를 살펴보면, 농학교 출신인 경우가 높게 나타났다. 2010년도에서 기타 출신이 33%로 가장 높게 나타났으나 2011년 이후 그 비율은 급격히 감소하여 이후 80~100% 내외로 농학교 출신자의 진행 비율이 높은 것으로 나타났다.



[그림 IV-10] 뉴스 진행자 출신학교 비율

뉴스 영상은 사전에 만들어진 시나리오에 의해 진행된다. 수어 뉴스의 경우에도 예외일 수 없다. 수어는 문자가 없기 때문에 한글 문장으로 제작된 시나리오를 앵커가 보면서 발화하기 때문에 일차적인 언어 표현이라기 보다는 ‘문장구역’과 같은 일종의 번역 활동의 성격을 지닌다. 그러나 현장을 기록하는 형태의 리포터 취재 뉴스에서는 취재진의 질문에 응하는 사람들이 비구조화된 형식의 수어 발화를 볼 수 있었다.

일반적으로 뉴스 앵커는 뉴스 자체를 선택하고 취재하여 자신의 관점에서 코멘트를 작성하고 해설하는 전문가이다. 우리나라의 앵커는 뉴스의 객관성 문제가 제기될 수 있어 무색무취하고 정확한 발음으로 구사하도록 하여 단순히 원고를 읽는 정도의 역할을 하는 경우도 있긴 하지만 대표적인 공식적 발화의 성격을 가진다. 이런 맥락에서 한국농아방송의 뉴스를 공식적인 한국수어 발화로 보고 그 특징을 파악하고자 하였으나 뉴스를 전달하는 과정 중 한국어 대응식 발화로 인해 분석이 용이하지 않았다. 이것은 한국어로 된 뉴스 기사를 읽고 즉시 한국수어로 바꾸어야 하는 앵커의 정보처리에서의 과부하 현상에 의한 것으로 사료된다. 이에 한국농아방송의 수어는 자연수어 발화를 수집하는 한국수어 말뭉치와는 성격이 다른 언어자료임을 알 수 있다. 다만 이들 앵커들에게서 두드러지게 발견되는 공통점은 카메라를 정면으로 응시하는 시선 처리였는데, 수어 문법 표현에서 중요한 요소인 시선 처리가 한국수어의 공식·비공식 표현의 형태 연구를 위한 자료로는 활용될 수 있을 것으로 사료된다.

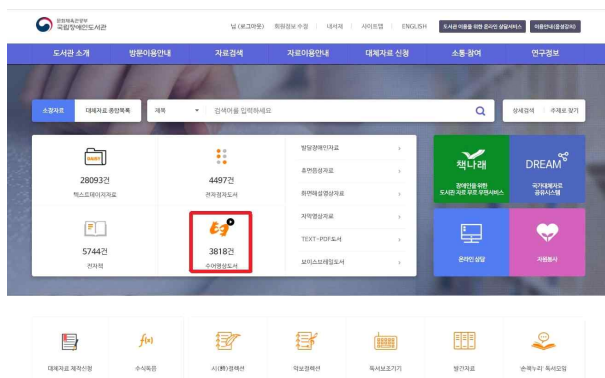
(2) 수어 영상 도서에서의 수어 자료 실태

국립장애인도서관의 자료 공개 게시판을 통해 확보된 수어 영상 도서 자료 데이터(DB) 목록에서 2007년부터 2018년도까지 제작된 도서를 확인할 수 있었다. 이 목록에는 서명, 저자, 책 발행처, 책 발행연도, 대상, 구분(완역/요약), 제작연도(수어영상 제작연도)에 대한 정보를 확인할 수 있었으나 완역 여부에 대한 정보는 2017년도 아동도서 261건

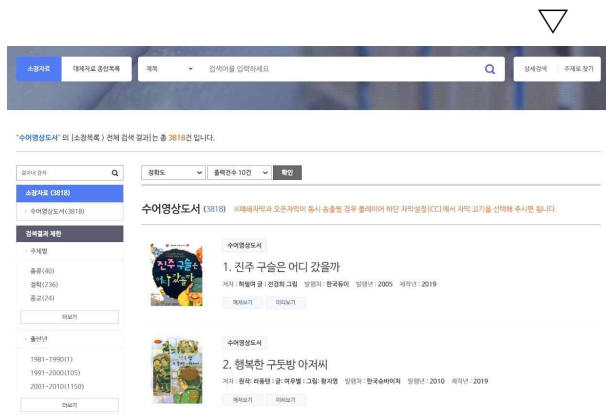
에서만 확인할 수 있었다. 수어영상의 길이는 해당 온라인 접속을 통한 상세 정보 화면에서 확인할 수 있다. 온라인 접속은 청각장애인 또는 그 가족이 회원일 때 제공받을 수 있는 서비스이다.

〈표 IV-11〉 국립장애인도서관 한국수어영상도서 자료 목록 예시

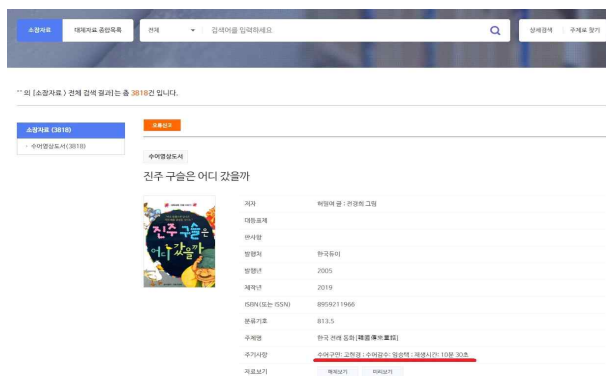
번호	제목	저자	발행처	발행년	대상	구분	제작연도
1	같이 놀자, 루이! : 모두가 행복한 진짜 통합교육 이야기	레슬리 일리 글 ; 폴리 던바 그림 ; 이승숙 옮김	한울림 스페셜	2015	All	아동 완역	2017
2	100원 부자	방미진 글 ; 박재현 그림	스콜라	2014	All	아동 완역	2017
3	4만 년 전 비밀을 품은 아이	한영미 글 ; 최희옥 그림	개암나 무	2015	All	아동 완역	2017
4	갯벌이 좋아요	유애로 글, 그림	보림	2013	All	아동 완역	2017
5	거짓말은 왜 나쁠까요?	브리지트 라베, 미셸 퓌엑 글 ; 자크 아잠 그림 ; 김이정 옮김	소금창 고	2002	All	아동 완역	2017
6	고민을 해결해 드립니다	에밀리 테이시도르 글 ; 가브리엘라 루비오 그림 ; 김영주 옮김	책속물 고기	2014	All	아동 완역	2017
7	(공기를 통해 전달되는) 소리	최원석 글 ; 지문 그림	아이앤 북	2016	All	아동 완역	2017
8	괜찮아, 선생님이 기다릴게	김영란 쓰고 그림	사계절	2014	All	아동 완역	2017
9					



국립장애인도서관 누리집
→ 로그인
→ 수어영상도서 접속

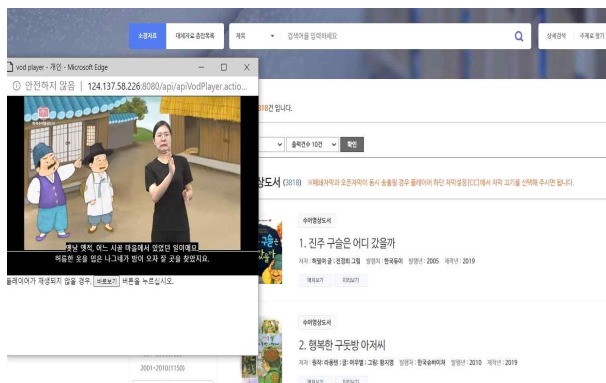


수어영상도서 목록



해당 도서 접속 후 상세 페이지

예시 도서(‘진주 구슬은 어디 갔을까’)는 총 10분 30분 길이의 수어 영상물



‘매체보기’를 접속하면 수어영상 팝업창이 나타남
수어영상 재생 시,
수어 모델의 발화, 나레이터의 음성과 자막이 동시에 송출

[그림 IV-11] 국립장애인도서관 수어 영상 도서 접속 경로

수어 영상 도서에서 구현된 수어의 특징을 확인하기 위하여 농인 연구자 2명과 청인 연구자 1명이 도서의 분류, 도서의 대상, 모델(수어 발화자), 번역 유형(요약/완역)을 중심으로 수어 영상 도서의 특징을 살펴보았다. 맨 먼저 번역물이어서 텍스트에 영향을 덜 받은 수어 영상으로 추천받은 ‘황금 물고기’를 분석하였다.

수어 영상 도서는 수어 발화자의 수어와 함께 음성으로 제공되고 있었으며, 자막은 개방형 또는 폐쇄형으로 제공 여부를 이용자의 선택에 따라 조정할 수 있었다.

<장면 1>

텍스트(음성, 자막) 옛날 어느 바닷가 마을 외딴집에 가난한 할아버지와 할머니가 살고 있었어.

수어 발화 [옛날] [바다] [과도(구형)] [지시] [동네]
[지시] [1] [집] [가난] [가난(지붕틀어짐)]
[할아버지] [할머니] [부부] [같이] [함께] [살다]

☞ (번역) : 옛날 바닷가 마을에
초라한 한 집이 있었다.
할아버지와 할머니 부부가 함께 살고 있었다.

<장면 2>

텍스트(음성, 자막) 그러던 어느 날 할아버지가 고기를 잡으러 갔다가 깜짝 놀랐지 뭐야.

수어 발화 [무엇] [날] [할아버지] [제스처(구형으로 사람배 문지르는 모양)]

(독백) { [배고프다] [물고기] [먹다] [오케이(10형)] [잡다] [오케이] }

[물고기] [1]
[바다] [걷다] [배타다] [노젓다] [배가다] [배가다(한손)]
[노젓다] [노놓다] [구형(그물 모양 펴서 잡는 모양)] [잡아
던지다] [구형(그물이 넓게 퍼지는 형태)]

(독백) { [많다] [잡다] [좋다] [기도] }

[구형(그물 모으는 형태)] [주먹형(그물 잡아당기는 모양)] [잡아당기다] [제스처(눈이 부서 눈을 잘 뜨지 못하는 형태)] [지시] [무엇]

☞ (번역) : 어느 날 할아버지는 (배고파하는 모습)

(독백) 배고프니 물고기를 먹으면 좋겠다. 물고기 한 마리
잡아야겠다.

바다로 걸어가 배를 타고 노를 저어 (바다로) 나가 노를

놓고 그물을 던졌다.
 (독백) 많이 잡히면 좋겠습니다.
 그물을 잡아당기니 (눈이 부서 눈을 잘 뜨지 못하는
 모습) 이것이 뭐지?

<장면 3>

텍스트(음성, 자막) 번쩍번쩍 빛나는 황금빛의 물고기가 그물에 걸려 나왔거든.

수어 발화 [빛나다] [노랑] [보석] [물고기] [구형(그물형태/수어 발화자가 물고기가 되어 호느적거리는 형태(역할전환))] [제스처(수어 발화자가 발화자의 위치로 돌아옴(역할 전환))] [지시]

☞ (번역) : 빛나는 황금 물고기가 그물에 걸려 있었다.

수어 영상 도서는 ‘도서 선정 → 시나리오 제작 및 연습 → 모델 촬영 → 자막 및 음성 삽입 → 완성’의 과정을 거쳐 제작하게 되는데, 이때 시나리오 과정에서 혹은 텍스트로 제작된 시나리오를 모델이 해석하는 과정에서 ‘번역’이 개입될 수 있는 구조이다. 위 수어 영상 도서에서 보듯이 어린이 대상의 동화에서 모델은 자막과 음성 텍스트보다 좀 더 상황 묘사적으로 설명하고 있음을 확인할 수 있다. 두 번째 텍스트에서 모델은 ‘고기를 잡으러 가는’ 과정을 좀 더 구체적으로 설명하면서 화자의 ‘혼잣말’ 형태와 ‘역할 전환’으로 알려진 구성된 행동이라는 수어 문법을 수행했다.

최근 다양한 양식을 연구하기 위한 말뭉치도 구축되고 있으며 가브리엘 등(Gabrielle Hodge et al., 2019)이 호주영어와 호주수어를 직접 비교하기 위하여 말뭉치를 설계한 것을 볼 때 수어 말뭉치 자료를 다양하게 수집할 필요가 있는 것으로 생각된다. 그러나 가브리엘 등의 연구가 동일한 이야기에 대해 여러 청인과 농인의 자료를 수집하는 동일 목적을 가진 말뭉치인 것과는 달리 본 수어 영상 도서는 하나의 자료를 표현하는 단일 언어자료라는 한계를 지니고 있다.

(3) 국내 수어 자료 실태 분석의 시사점

첫째, 농아방송의 수어자료와 국립장애인도서관의 수어영상도서 자료는 한국어 스크립터나 텍스트의 영향을 받아 수어가 발화되는 한계를 지니고 있다. 농아 방송과 수어영상도서를 통해 살펴본 결과, 수어는 문자가 없기 때문에 뉴스나 수어 영상 도서로 제공되기 위해서 한국어를 빌어 작성된 스크립터나 시나리오를 이용할 수밖에 없다. 수어 영상 도서 중 2008년도부터 2011년도까지 제공된 일부 영상 자료 중에는 연극의 형식을 빌어서 배역을 맡은 수어자가 내용을 충분히 숙지한 후 수어를 발화하는 영상도 있었지만 대부분의 자료에서는 한국어 텍스트의 간접 현상을 완전히 배제하기에는 한계가 있었다.

둘째, 자연수어로서의 한계로 인해 한국수어의 원형 보존이 제한적이다. 말뭉치가 다량의 실제 언어를 수집한다는 측면에서 두 영상의 수집도 의미를 갖는다고 볼 수 있겠지만 말뭉치는 언어 자료의 수집이나 구축 과정에서의 자연스러운 발화가 원래의 내용이나 형태의 누락 없이 언어적 특성이 잘 반영된 원형을 유지하여 대표성을 지닐 수 있다. 이런 측면에서 앞에서 살펴본 두 영상은 한국어를 빌어 작성된 스크립터나 시나리오를 보고 수어를 발화 함으로써 농인 앵커나 수어자의 번역 활동에 의한 수어의 형태라는 한계를 지니고 있다.

특히 수어영상도서는 시나리오를 보며 즉석에서 수어로 번역하는 활동을 하는 환경에서 제작되기 때문에 다소 한국어에 얽매일 가능성이 높은 구조여서 수어 체계에 대한 근거 자료로서 제한적이라고 할 수 있다.

4. 해외 수어 말뭉치 구축 탐색

1) 원시 수어 말뭉치 수집

(1) 호주

호주 수어 말뭉치는 트레버 존스톤(Trevor Johnston)이 호주 농인 중 수어 사용자의 감소로 인해 멸종 위기에 처한 수어를 보존하고 호주 수어 학습자들의 언어 학습뿐만 아니라 수어 연구 자료로의 활용을 목적으로 만든 것이었다. 2008년까지 구축하여 2012년부터 공개하고 있다. 호주 수어 말뭉치는 런던대학교 산하 SOAS(School of Oriental and African Studies)의 사멸 위기 언어 보관소(Endangered Languages Archive, ELAR)에 디지털 영상 아카이브(언어저장소)로 구축되어 있으며, 주석과 메타데이터 파일이 연결된 기록들로 구성되어 있다. 현재 호주 수어 말뭉치에는 2003-2005년에 걸쳐 이루어진 호주 수어 연구 프로젝트(the Sociolinguistic Variation in Auslan Project, SVIAP) ‘사회언어적 변이’의 일환으로 수집된 수어 자료도 추가되었다. 두 번에 걸쳐 수집된 자료들은 모두 농인 원어민 또는 조기 학습자/호주 수어의 준 원어민 사용자를 대상으로 녹화된 언어를 기반으로 구축된 것이다. 여기서 원어민 수어자란 부모 모두가 수어를 사용하는 농인이거나 부모 중 한쪽만 수어를 사용하는 농인, 또는 농인 형제로부터 태어날 때부터 호주 수어를 습득한 사람을 의미하며, 조기학습자 또는 수어 준원어민이란 7세 이전에 호주 수어를 습득했거나 배운 사람으로 정의하고 있다.

처음 구축된 호주 수어 말뭉치는 호주 5대 도시(시드니, 멜버른, 브리즈번, 애들레이드, 퍼스)의 참가자 100여 명으로부터 촬영되었으며 총 300시간 분량이었다. 이것은 각 참가자들에게 인터뷰, 내러티브(narrative), 설문조사에 대한 대답, 자유 대화, 그리고 그림책, 만화 영화, 호주 수어로 들려주는 녹화된 이야기 등 여러 유형의 유도자료를 사용하여 반응을 이끌어

낸 것으로 약 3시간 동안 진행된 영상물이다. 이렇게 수집된 영상물 중 사용 가능한 것들을 편집하여 언어 데이터로 제작한 영상은 약 150시간 분량이었다. 이들 영상은 주석이 용이하도록 과제별로 약 1,100개의 개별 디지털 영상 텍스트로 편집되었다. 2009년 3월을 기준으로 1,100여 개의 영상 텍스트 중 약 201개는 엘란을 사용하여 주석까지 완료하였다. 추가로 수집된 SVIAP 말뭉치는 이들 5개 도시의 211명의 참가자를 대상으로 자유 대화, 구조화된 인터뷰, 수어 어휘를 유도하는 내용 등의 유도 자료에 의해 수집된 수어 자료를 디지털 비디오 영상물로 저장하였다.

(2) 영국

영국 수어 말뭉치는 영국 전역의 수어 원어민, 준 원어민, 또는 유창한 수어 사용자로부터 수집한 자연스럽고 유도된 영국 수어를 기계가 읽을 수 있는 코퍼스로 구축하기 위해 수행된 영국 수어 말뭉치 프로젝트(British Sign Language Corpus Project, BSLCP)의 결과물이다.

이 말뭉치에서는 영국 수어의 사회언어학적 다양성과 변화에 대한 연구를 수행하고자 지역과 연령이 고려되었다. 지역별 변이 확인을 위해서 최소 30명의 표본이 요구되어 8개 지역(London, Bristol, Newcastle, Manchester, Birmingham, Belfast, Cardiff, Glasgow)을 중심으로 총 249명을 대상으로 수어 언어 자료를 수집하였다. 이들 언어제공자는 20세기 농교육에서의 언어 교육 변화 변인을 고려하여 연령대별로 구분하여 18-35세 24%(59명), 36-50세 25%(62명), 51-64세 27%(68명), 65세 이상 24%(60명)의 연령 집단별로 고르게 구성하였다. 그러나 사회적 계층이나 인종을 균형을 맞추어 다양하게 고려하지는 못하였다. 그 결과 직업이나 학력 배경으로는 대졸 또는 화이트 칼라 전문직 언어제공자(38%)가 대졸 이하의 블루칼라로서 공장이나 상업 관련 직업에 종사하는 사람들(62%)보다 적었으며, 참가자 중 20명(8%)만이 흑인 또는 아시아계였다.

영국 말뭉치는 유도자료를 사용하여 수집하였다. 이들 유도자료 중 참가자들이 촬영 세팅에 익숙해지도록 미리 준비해 온 개인적인 경험을 짧게 상대방에게 이야기하는 첫 번째 내러티브(narrative) 과제와 이어서 제시되는 30분 간의 자유 대화(30-minute conversation-free talking)는 사전에 참가자에게 인터넷 공개를 하지 않는 과제임을 알렸다. 이 과제는 지역 방언 수어 수집을 목적으로 수어 발화자가 관찰자를 의식하여 자연스럽게 본인의 수어 발화를 하지 않아 영국 수어인지 의심받을 수 있는 위험을 최소화하여 언어제공자들로부터 자연스러운 영국수어 자료를 최대한 확보하기 위함이었다. 수어 어휘 과제에서는 총 102개의 개념이 제시되었는데, 이는 높은 수준의 사회언어학적 변이형 수어가 예상되거나 이미 알려진 수어가 발화될 가능성이 있는 것들로 선정하였다.

촬영과정에서는 지역 농사회를 통해 참가자를 모집하고, 촬영과 데이터 수집이 이루어졌다. 영어의 간섭을 받거나 영국 수어와 영어의 코드 스위칭을 방지하기 위해 현장에 청인이 없도록 하였다. 또한 유도자료를 통해 언어 축진을 하는 사람은 연구원이 아닌 현장 조사원이 담당하였으며, 언어제공자가 쌍을 이루어 언어적 반응을 하는 장면을 영상에 담았다. 이때 언어제공자의 짝 구성은 나이, 친밀도, 성별을 고려했다.

촬영은 농사회 클럽이나 농 기관의 사무실처럼 친근한 공간을 활용함으로써 격식에 얽매이지 않는 다양한 수어를 사용할 수 있도록 분위기를 조성하였다.

(3) 독일

독일 수어 말뭉치는 독일 수어의 광범위한 말뭉치 구축과 말뭉치 데이터 분석을 바탕으로 독일 수어-독일어 종합사전 개발을 목표로 2009년 1월부터 2023년까지의 장기 프로젝트로 시작된 함부르크 과학 및 인문 학회(Academy of Sciences and Humanities in Hamburg)와 함부르크 대학교(Hamburg University)의 ‘독일 수어 코퍼스 프로젝트(Germa

n Sign Language Corpus project)’에 의해 수행되었다.

프로젝트 1단계에서는 독일 전역의 12개 지역에서 약 330명을 대상으로 독일 말뭉치를 수집하였다. 독일 수어 사용자의 일상 언어를 반영하기 위해 언어제공자를 지역, 성별, 연령 등 사회언어학적으로 균형 있게 수집 하였으며, 참가자는 2인 1조로 약 7시간(휴식 시간 포함) 동안 촬영하였다. 7대의 카메라로 촬영한 영상은 총 560시간 분량(750TB)이었다.

독일 수어 말뭉치도 지역 수어 변이형 도출을 위해 면접자에 의한 영향을 최대한 적게 받도록 해당 지역의 농인 현장조사원이 인터뷰를 실시하였다. 선택된 유도자료는 음운론, 형태론, 통사론, 어휘론적 현상을 포착하는 것에 목적을 둘 뿐만 아니라 필요한 경우 병렬 교차 언어 분석이 가능하도록 말뭉치 연구에서 선행되었던 과제들을 독일 수어에 맞게 수정하여 사용하였다. 즉 교차언어학 연구 목적에 적합하다고 할지라도 그림책 ‘Frog, Where Are You?’ (Mayer, 1969)이나 Pear Story(Chafe, 1980), 만화영화 ‘Tweety and Sylvester’ (Warner Brothers, 1950), 그리고 츠위저우드(Zwitsersloot)의 드로잉(2003)) 중 하나의 과제를 채택하고자 할 때에는 그 나라의 문화적 차이와 함께 일반적으로 잘 알려진 이야기라고 할지라도 농인 공동체에는 잘 알려져 있지 않은 것일 수 있다는 것을 고려하였다.

최근에는 자동 번역에 대한 연구가 활성화되면서 수어 자동 번역 및 자동 인식에 관한 관심도 높아졌다. 유럽 문화유산 온라인 기구(the European Cultural Heritage Online Organization, ECHO)에는 스웨덴 수어, 영국 수어 및 네덜란드 수어 등의 수어 말뭉치가 등록되어 있다. 수어 자동 번역 및 자동 인식에 관한 시스템 역시 말뭉치에 의존할 수밖에 없으나 ECHO에 등록된 수어 말뭉치는 한 명의 수어 발화자의 이야기로 이루어져 있어 기계가 학습하기에는 자료의 양이 너무 적거나 일반적인 것으로서 기계 자동 훈련용으로 사용하기에는 부적절하였다. 따라서 자동 수어 번역 및 인식을 위해 독일 수어로 번역된 독일 텔레

비전 일기 예보를 기반으로 하는 Phonix 코퍼스도 만들었다(B., Jan., S., Daniel., D., Philippe., Z., Morteza., & H., Zahedi., 2006).

2) 주석 달기

(1) 호주

2019년 기준으로 호주수어 말뭉치의 주석이 달린 영상은 357편이다. 주석파일은 다양한 수준의 주석을 포함하며 계속 추가되고 있다. 주석자들은 주석 작성에는 상당히 많은 시간이 소요되기 때문에 호주수어 말뭉치가 기계 판독이 가능하도록 충분히 주석을 달아서 진정한 언어 말뭉치로서의 자격을 갖추기까지는 수년이 걸릴 것으로 예상된다고 밝히고 있다.

호주 수어 말뭉치 주석 파일에는 아래와 같은 주석 유형이 있으며 보다 자세한 층위별 주석 유형은 [표 IV-5]와 같다.

- ID 글로스와 명사 및 동사 식별
- 시선 방향
- 손바닥 방향
- 구절/조항 경계
- 동사 인수의 식별
- 구성된 역할행위(constructed action)에 대한 동사 인수의 태그 지정
- 공간 변형에 대한 유무
- 자유 번역
- 직역 등

[표 IV-12] 호주 수어 말뭉치의 주석 층위 및 유형

Parent tier	Expanded name	Linguistic type
↳ Child tier		
RH ID-gloss	Gloss	BasicAnnotation
↳ RH-Mean	Meaning	BasicTag
↳ RH-GramCls	Grammatical class	GramCls
↳ RH-Transcrip	Transcription	BasicTag
↳ RH-Handsh	Handshape	BasicTag
↳ RH-Orient	Orientation	BasicTag
↳ RH-Loc	Location	BasicTag
↳ RH-Move	Movement	BasicTag
↳ RH-NonMan	Other non-manuals	BasicTag
↳ RH-OtherPhon	Other phonetic/phonological	BasicTag
↳ RH-ModOrVar	Citation modification or variation	ModOrVar
↳ RH-Freq	Lexical frequency	BasicTag
↳ RH-CAco	Co-occurrence of sign with CA	BasicTag
LH ID-gloss	Gloss	BasicAnnotation
↳ LH-Mean	Meaning	BasicTag
↳ LH-GramCls	Grammatical class	GramCls
↳ LH-Transcrip	Transcription	BasicTag
↳ LH-Handsh	Handshape	BasicTag
↳ LH-Orient	Orientation	BasicTag
↳ LH-Loc	Location	BasicTag
↳ LH-Move	Movement	BasicTag
↳ LH-NonMan	Other non-manuals	BasicTag
↳ LH-OtherPhon	Other phonetic/phonological	BasicTag
↳ LH-ModOrVar	Citation modification or variation	ModOrVar
↳ LH-Freq	Lexical frequency	BasicTag
↳ LH-CAco	Co-occurrence of sign with CA	BasicTag
ClauseLikeUnit(CLU)	Clause-like unit ('utterance/meaning unit')	BasicAnnotation
↳ RH-Arg	Argument identification	ClauseArguments
↳ RH-MacroR	Macro-role of argument	MacroRoles
↳ RH-SemR	Semantic role of argument	SemanticRoles
↳ RH-overtSUBJ?	Overt subject?	overtSUBJ?
↳ LH-Arg	Argument identification	Arguments
↳ LH-MacroR	Macro-role of argument	MacroRoles
↳ LH-SemR	Semantic role of argument	SemanticRoles
↳ LH-overtSUBJ?	Overt subject?	overtSUBJ?
CLUcomplex	CLUs overtly related to each other	BasicAnnotation
↳ OvertDependencyType	Nature of expression of dependency	BasicTag
CLUwithinCLU	Complement and embedded CLUs	BasicAnnotation
↳ OvertEmbeddedType	Nature of expression of embeddedness	BasicTag
CLUcomposite	Sentence complexity	BasicAnnotation
CLUmood	Mood	BasicAnnotation
CLUaktionsart	Event type or Aktionsart	BasicAnnotation
CLUtransitivity	Transitivity type	BasicAnnotation
LitTransl	Literal translation	BasicAnnotation
Non-manual & other		
CA	Constructed action or constructed dialogue	BasicAnnotation
↳ CA-Arg	Argument identification	ClauseArguments
↳ CA-MacroR	Macro-role of argument	MacroRoles
↳ CA-SemR	Semantic role of argument	SemanticRoles
↳ CA-overtSUBJ?	Overt subject?	overtSUBJ?
Body	Body	BasicAnnotation
Face	Global description of facial expression	BasicAnnotation
Head	Head movements	BasicAnnotation
Gaze	Direction of eye-gaze	BasicAnnotation
Eye&Brow	Eye and brow movements	BasicAnnotation
Mouthing	Mouthing (of words)	BasicAnnotation
↳ MouthingGCI	Grammatical class of mouthed English word	GramCls
MouthGestF	Mouth gestures form	BasicAnnotation
↳ MouthGestM	Mouth gestures meaning	BasicTag
FreeTransl	Free translation	BasicAnnotation
Comments	Comments	BasicAnnotation

(2) 영국

영국 수어 말뭉치 프로젝트를 추진하는 동안 주석 작업은 변이언어학(특히 음운적, 어휘적 다양성과 변화)과 어휘 빈도에 대한 연구가 함께 진행되었다.

변이언어학 연구에서는 한 손 수어의 수형과 수향의 언어학적 요소 및 사회적 요인에 따른 다양성을 조사하는 음운적 다양성과 영국 수어 어휘의 다양성을 연구하였다. 영국 수어 말뭉치 프로젝트는 언어학자들이 위기에 처한 언어 공동체의 언어 사용자와 긴밀한 협조를 통해 연구가 추진될 필요가 있음으로 인해 변이언어학에 초점을 두고 추진되었다. 즉 학교 공동체를 중심으로 하는 수어, 국가별 농인의 신분에 따른 수어, 그리고 노인들을 중심으로 알고 있는 지역 수어(방언)와 같이 소멸 위기에 놓인 영국 농사회에서 사용되는 수어를 문서화하여 보존해야 할 필요성이 대두되었기 때문이다. 그래서 주석 작업은 음운적 다양성과 어휘 빈도 연구계획과 연관되어있었다.

음운적 다양성 연구를 위해 대화 과제로 수집된 211명의 언어제공자의 6,330개 수어에 대한 주석이 이루어졌다. 모든 토큰은 간단한 코딩 시스템을 사용하여 수형과 수향에 대한 주석이 달렸고, 그중 2,110개의 토큰에는 토큰의 문법 분류와 한 손 수어인지 두 손 수어인지와 같은 추가 주석이 포함되었다.

어휘 유도 과제에서는 249명의 언어제공자를 대상으로 7,332개의 수어에 대한 주석 작업이 이루어졌으며, 이 중 약 50명의 대화 과제에서 약 25,000개 수어에 대한 주석도 완료하였다(Schembri et al, 2013).

(3) 독일

독일 말뭉치는 함부르크 대학교 수어언어학과 연구소에서 개발한 iLex 프로그램으로 주석하였다. 수집된 수어 영상 총 560시간 분량 중 65시

간에 대한 기본 주석 작업이 완료되었다. 2017년 7월 기준으로 주석된 토큰 수는 총 425,000개이다(원성옥 외, 2017).

기본 주석은 수어 발화를 세분화하고 어휘 항목을 식별하기 위한 것으로 개별 수어로의 접근을 돕는다. 글로스로 주석되며 이때 색인 수어, 알파벳 지문자 등과 같은 특별한 수어가 별도의 영역으로 묶일 수 있도록 하였으며, 생산적 수어의 경우 손 모양, 손 배열 및 시각화 기법 영역 등의 상위 집합체로 구분하여 주석을 수행할 수 있도록 하였다.

두 번째 단계로는 수어 기본형을 상세화 하는 수준에서 주석 달기를 하였다. 말뭉치 주석은 사전 제작에 필요한 어휘 분석의 요건과 밀접하게 연결된다. 따라서 해당 수어뿐만 아니라 수어의 맥락적 의미, 단어 조합 및 기타 관련된 문맥 정보가 정확하게 파악될 수 있도록 목표 토큰 외의 주변 발화도 주석할 수 있도록 하였다. 이런 측면에서 마우딩, (어휘적) 얼굴 표정, 토큰 형태의 표기 및 분류, 토큰의 문맥적 의미, 문법 범주, 공간의 장면 설정이나 수어 공간에서 특별한 장소, 물체의 위치 지정 등 장소에 대한 언급을 표기할 수 있는 공간 사용, 시각화를 위한 생산적 요소, 음운론적 변이형, 공간 방향 또는 위치 지정이나 위치 표시, 복수 표시 등과 같은 문법적 과정의 결과물인 수어의 형태를 주석하였다.

(4) 네덜란드

네덜란드 말뭉치는 MPI(Max Planck Institute for Psycholinguistics)가 심리언어학자들을 위해 개발한 ‘엘란’을 주석 도구로 사용하였다. 엘란은 원본 데이터(사운드 또는 비디오 파일)를 재생할 수 있고, 주석을 비디오의 오디오 신호와 정렬할 수 있어 온라인 주석이 가능하며, 기능이 확장되어 IMDI(ISLE Meta Data Initiative)와의 통합이 가능하도록 구성되었다.

네덜란드 수어 주석은 ECHO(European Cultural Heritage Online) 지침을 기본으로 하여 수정된 글로스 지침과 참고 리스트를 활용하여 글로스로 입력되었다. 2008년 기준 60개의 영상에 대한 주석이 완료되었으며, 12시간 분량의 수어에 대한 글로스는 총 64,000개였다. 주석 작업자는 글로스 주석을 하는 동안 수어의 수지 요소에 한정되어 글로스명을 정하는 것에 어려움을 겪었는데, 비수지 요소를 포함한 전체 구성의 의미를 충분히 살펴보지 못하여 향후 프로젝트에서는 더 많은 상세 주석을 실시할 것임을 밝혔다.

주석은 언어학 연구자들 외에도 다른 관심 분야에서도 사용될 수 있도록 작업될 수 있다. 이에 최대한 많은 자료가 많은 사람들에게 의해 이용될 수 있도록 일부 자료는 번역 주석이 제공되는데, 이는 수어통역사나 통역 전공 학생들에게 음성 지원을 통해 제공된다. 즉 그들은 먼저 특정 영상을 보고 나서 어려운 내용은 동료와 의논 후 동시통역을 한다. 이렇게 동시통역된 통역사의 음성은 미니디스크에 녹음되고, 이 오디오 파일은 WAV 파일로 처리되어 영상과 정렬하여 엘란 파일에 연결된다. 영상 재생이 가능하고 동료와 번역 내용을 의논할 수 있다는 점 때문에 네덜란드 수어 말뭉치의 음성통역 주석은 특징적이다. 이 과정의 본질은 통역과 번역을 함께 제공할 수 있다는 점이다. 음성통역 파일을 얻어내기까지 평균적으로 소요되는 시간은 영상 재생시간의 10배가 소요되었다. 이는 녹음 절차와 관련된 행정적, 기술적 업무를 모두 포함한 시간으로 수어 1시간 분량을 미니디스크에 기록하는데 10시간이 소요되었다.

3) 수어 말뭉치 활용 연구

(1) 호주

호주에서 호지(Hodge)와 가브리엘(Gabrielle) 등(2019)은 수어를 사용하는 농인과 같은 지역의 영어를 사용하는 청인에게 언어 자료를 마련해 주

기 위해 구축된 호주수어와 호주영어 말뭉치를 갖고 두 언어를 비교하는 연구를 실시하였다. 호주수어와 영어의 기본적인 어순 SVO 및 각 언어의 문법적 관계 연구(Johnston & Trevor., 2019), 호주수어의 부정 형태 중에서 머리 흔들기의 역할에 관한 연구(Johnston, Trevor., 2018), 호주수어에서의 입 움직임에 관한 연구(Johnston, Trevor & Roekel, Jane & Schembri, Adam., 2015), 그리고 호주 수어 교재에 사용된 모든 예제(예: 질문, 진술, 명령 및 요청 등)를 호주 수어 말뭉치 영상과 함께 활용한 수어교육 연구(Cresdee, Donovan & Johnston, Trevor, 2014) 등, 호주 수어 말뭉치 기반의 다양한 연구가 이루어지고 있다.

(2) 영국

영국에서는 영국수어 말뭉치가 구축되면서 2011-2015에는 수어말뭉치를 기반으로 한 영국수어사전(BSL SignBank)이 구축되었을 뿐만 아니라 일치동사 연구사업(2012-2014), 영국 통사론 연구사업(2016-2020), 농인들의 언어태도 관련 연구사업(2017-2019), 수어자동번역 연구 사업(2018-2021), 그리고 비수지 연구사업(2019-2020)이 단계적으로 추진되어 오고 있다.

영국수어의 사회언어학적 변이에 대한 연구는 오랫동안 색과 숫자 수어가 지역마다 상이하다는 연구(Brien, 1992, Deuchar, 1984, Sutton-Spence and Woll, 1999, Sutton-Spence et al., 1990) 등과 같이 어휘 수준의 연구에 집중되어 왔다. 그러나 최근 말뭉치 기반의 연구에 따르면 지역에 따른 변이로 작용하던 숫자나 색깔 관련 수어의 특징에 관한 연구는 약화되고 있으며(Stamp et al., 2012), 영국수어와 수지 영어에 대한 영국 농사회의 언어 태도(Rowley, K., Fenlon, J., & Cormier, K., 2018), 영국 농사회의 연령과 언어 변화에 따른 태도 연구(Rowley, K., Fenlon, J. & Cormier, K., 2018) 등의 주제로 확대되고 있다.

그 밖에도 영국수어 말뭉치를 기반으로 한 연구로 영국 수어에서 지시동사의 변이형, ‘1’ 손모양(손가락을 펴는 것)을 사용한 사회언어학적 변

화와 같은 영국수어 대화의 어휘 빈도(Fenlon, J., Schembri, A., Rentelis, R., & Cormier, K., 2012)나 영국수어 방언 연구(Stamp, R., Cormier, K., Evans, B., & Schembri, A., 2012), 영국수어에서 절의 위치와 지시 동사(Schembri, A., Fenlon, J., & Cormier, K., 2017), 영국 수어와 호주 수어의 두 국가 간 수어 부정어에 대한 연구(Hodge, G., Fenlon, J., Schembri, A., Johnston, T., & Cormier, K., 2018), 동사에서 나타나는 영국 수어의 변이형 연구(Fenlon, J., Schembri, A., & Cormier, K. 2018). 등이 있다.

(3) 독일

독일에서는 독일수어 말뭉치를 통해 수어의 변이형과 빈도수, 수어의 지역 분포, 수어의 공간 사용, 수어의 의미, 수어의 연어 분석 등이 이루어지고 있다. 최근 독일에서는 교육자료로 활용할 수 있는 입 제스처(mouth gestures)에 대한 실증 자료가 적어 자동 번역 시스템을 활용하여 입 제스처를 분류하는 연구(M., Brumm et al., 2020), 독일수어 말뭉치 공개 방법에 관한 연구(H., Thomas et al., 2020), 수어 의미를 중심으로 한 수어사전학에서의 수어 연어 연구(L., Gabriele & S., Marc., 2020), 독일 수어 말뭉치와 독일 수어 말뭉치 기반 사전의 관계에 관한 연구(M., Anke et al., 2020), 독일수어 문장 단위에 관한 연구(J., Elena, 2019), 독일 수어 말뭉치 기반 사전에 포함될 순차적 결합에 관한 연구(L., Gabriele et al., 2019) 등 다양한 주제로 연구들이 진행되고 있다.

(4) 네덜란드

네덜란드에서는 편 수형에서 엄지의 역할에 관한 연구(Ormel, Ellen et al., 2017), 비우세 손 생략 연구(Paligot, Aurore et al., 2016), 비우세 손의 유지 유형(Kimmelman, Vadim et al., 2016), 마우딩 연구(Bank, Ri

chard et al., 2016) 등에서 말뭉치가 활용되고 있다.

4) 해외 말뭉치 구축 탐색의 시사점

첫째, 원시 수어 말뭉치 수집의 목적이 다양하다.

호주의 경우, 농인 중 수어 사용자의 감소로 인해 소멸 위기에 처한 수어를 보존하고 호주 수어 학습자들의 언어 학습뿐만 아니라 수어 연구 자료로 활용할 목적으로 수집하고 있다. 영국의 경우에는 영국 수어 원어민의 수어를 기계로 읽을 수 있는 코퍼스를 구축할 목적으로 독일은 수어-독일어 종합사전을 개발할 목적으로 시작되었다.

둘째, 수어 말뭉치 구축 사업은 장기간 프로젝트로 수행되고 있다.

각국의 수어 말뭉치 구축 사업은 대부분 10년 이상의 기간을 두고 사업이 진행되었다. 호주는 2003년부터 연구를 진행하여 2012년에는 수어 자료를 공개하였으며, 독일은 2009년부터 2023년까지의 ‘독일 수어 코퍼스 프로젝트’라는 이름으로 독일 수어 말뭉치 구축 사업이 진행되었다.

우리나라는 1년 단위의 사업 공모를 통해 진행됨으로써 연구의 연속성이 불확실하여 장기적 계획을 수립하여 사업을 진행하지 못하는 한계를 벗어날 필요가 있다.

셋째, 지역을 대표한 언어제공자를 충분한 숫자만큼 확보하여 수어말뭉치를 수집한다.

호주는 최초 수어 말뭉치 구축 사업에서부터 호주의 5대 도시로부터 각 20명씩, 총 100명의 수어제공자를 모집하였으며, 영국은 8개 지역을 중심으로 한 지역당 최소 30명, 총 249명의 수어제공자로부터 영국수어를 수집하였고, 독일은 12개 지역의 330명 수어제공자를 대상으로 독일 수어 말뭉치를 수집하였다.

넷째, 장기간에 걸쳐 다양한 수준에서 주석 작업을 추가한다.

각국의 수어 주석 작업은 가능한 충분한 주석을 달아서 수어가 언어말

뭉치로서의 가치가 드러나도록 하고 있다. 호주는 다양한 주석 층위에서 글로스와 타입을 설정하고 있으며, 영국은 변이언어학 측면에서 음운적 다양성을 중심으로 주석 작업을 하고 있으며, 독일은 기본 주석과 상세 주석으로 나누어 주석을 함으로써 맥락적 의미, 단어 조합, 맥락적 의미와 같은 정보도 파악할 수 있게 한다.

이런 주석 작업은 단시간, 단회의 수고로 이루어지는 것이 아니라 꾸준히 하게 주석 작업을 계속 덧붙임으로써 주석이 더욱 상세화 되고, 주석의 활용성이 높아지도록 하고 있다.

다섯째, 수어 말뭉치 자료로 인해 수어의 독특한 언어 특성을 밝히는 연구가 활발하게 진행되었다.

연구 주제들은 대개 수어가 가진 독특한 언어적 특성과 관련된 것들이 많았다. 호주의 경우, 호주주어와 호주영어의 어순비교연구, 머리 흔들기의 역할에 관한 연구, 입 움직임에 관한 연구와 같이 수어의 독특한 언어 체계를 밝히는 연구들을 중심으로 연구가 진행되었다. 영국은 수어 통사론과 수어자동번역 연구를 특징적으로 진행하였으며, 독일도 이와 유사한 자동번역시스템을 이용하여 입 제스처 연구를 수행하였다. 독일은 독특하게 독일수어 말뭉치 공개 방법에 관한 연구를 수행함으로써 수어 말뭉치 구축사업의 결과가 연구자에게나 일반인에게 유익한 정보로 이용될 수 있도록 하려는 노력을 엿볼 수 있다.

〈표 IV-13〉 해외 수어 말뭉치의 수어 자료 수집 현황

	호주수어 (Auslan Corpus)	네덜란드수어 (NGT Corpus)	영국수어 (BSL Corpus)	독일수어 (DGS Corpus)
목적	<ul style="list-style-type: none"> 수어 사용자의 감소로 인해 소멸 위기에 처한 수어 보존 호주 수어 학습자를 위한 학습 자료 및 연구 자료로 활용 		<ul style="list-style-type: none"> 기계가 읽을 수 있는 영국수어 말뭉치 구축 영국 수어의 사회언어학적 다양성과 변화에 대한 연구 수행 	<ul style="list-style-type: none"> 독일 수어-독일어 종합 사전 개발 독일 수어 사용자의 일상 언어 반영을 위한 사회언어학적 균형 있는 수집
연구사업 기간	2004-2007 호주 수어 연구 프로젝트 (2003-2005) 수집 수어 자료 포함	2006-2008	2008-2014	2009-2023 (수집을 위한 촬영 기간: 2010. 01-2012. 03)
언어제공자 기준	<ul style="list-style-type: none"> 농인 원어민 수어자: 부모 모두 또는 부모 중 한쪽만 수어 사용 농인, 농인 형제를 통해 출생부터 호주 수어 습득한 사람 조기 학습자 또는 호주 수어의 준원어민 사용자: 7세 이전에 호주 수어 습득하거나 배운 사람 		<ul style="list-style-type: none"> 수어 원어민 준 원어민 유창한 수어 사용자 	
언어제공자 수	100명	100명	249명	330명
촬영 지역	5 지역 (각 지역 20명)	전국, 방언 5개로 분류	8 지역	12지역
카메라 수	2 개	4개	3개	7개
세션 시간	3시간	4시간		7시간
세션 방식	언어제공자 2명 및 사회자	언어제공자 2명 및 사회	언어제공자 2명 및 사회자	언어제공자 2명 및 사회자

		자		
유도자료, 과제	과제 10개	과제 8개	과제 4개	과제 20개
총 촬영시간	300시간	400시간		1980시간
언어 자료 시간	150시간 (영상 1,100 개)	150시간, 2,375 파일		560시간
메타데이터 형식 및 관리	IMDI	IMDI, CMDI		CMDI
홈페이지	www.auslan.org.au/about/corpus/	https://www.ru.nl/corpusngt/uk/	https://bslcorpusproject.org/	https://www.sign-lang.uni-hamburg.de/dgs-korpus

〈표 IV-14〉 해외 수어 말뭉치의 수어 자료 주석 및 활용 현황

	호주수어 (Auslan Corpus)	네덜란드수어 (NGT Corpus)	영국수어 (BSL Corpus)	독일수어 (DGS Corpus)
연구사업 기간	2004-2007 추가 주석 사업 (2008, 2012, 2020-)	2006-2008	2008-2014	2009-2023
기본주석 시간	1,100 파일 중 357개 주석됨 (2019.03)	60개 파일(12시간 분량), 주석 64,000개 수어말뭉치 중 15% 주석 (2008.01) 341개의 파일 주석(14시간 24분) 52개 파일 정밀 번역(2시간 30분) (2015년 주석 지침 기준)		기본 주석: 560시간 중 65시간, 425,000개 (2017.07) 600,000 토큰 (2020.03)

주석 프로그램	ELAN	ELAN	ELAN	iLex
주석 층위	ID 글로스, 명사, 동사 식별 시선 및 손바닥 방향 구절/ 조항 경계 동사 인수 식별 구성된 역할에 대한 동사 태그 공간 변형 유무 자유 번역 지역 등	번역 수지 글로스 수지 음운 분석(수어소) 수지 언어 분석(문법) 입 움직임 비수지 행동 정보	음운 연구용 수형, 수향 문법 분류 한손 수어와 두손 수어 등	글로스 번역(독일어 번역과 영어 번역) 마우딩, (어휘적) 얼굴 표정 문맥적 의미 문법 범주 공간 사용 생산적 요소 음운론적 변이형 등
타입목록 관리 방법 및 공개	SignBank (공개)	SignBank (공개) (http://signbank.science.ru.nl)	SignBank (공개) (http://bslsignbank.ucl.ac.uk)	연구자 포털에 공개 https://www.sign-lang.uni-hamburg.de/meinedgs/ling/start_de.html
주석지침 공유 방법	홈페이지에 탑재 Auslan Corpus Annotation Guidelines (2019년 8월)	홈페이지에 탑재 Annotation Conventions for the NGT Corpus (2015년 2월)	홈페이지에 탑재 BSL Corpus Annotation Conventions (2015년5월)	연구자포털에 탑재 Annotationskonventionen (2020년9월)
수어 형태 전사 (예: 함노시스)	엘란에서 함노시스 티어	부분적으로 엘란에서 함노시스 티어 (예: 분류사)		iLex에서 함노시스 티어
활용의 예	<ul style="list-style-type: none"> - 문법적 관계 - 부정에서의 머리 흔들기 역할 - 입 움직임 	<ul style="list-style-type: none"> - 편 수형에서 엄지의 역할에 관한 연구 - 비우세 손 생략 연구 - 비우세 손의 유지 유형 - 마우딩 연구 등 	<ul style="list-style-type: none"> - 어휘 빈도 - 방언 - 절의 위치와 지시 동사 - 영국수화와 영어의 부정 비교 등 	<ul style="list-style-type: none"> - 수어의 변이형, 빈도수, 수어의 지역 분포 - 수어의 공간 사용 - 수어의 의미 - 수어의 언어 분석

				- 자동 번역 시스템을 활용할 마우스 제스처 분류 연구 등
--	--	--	--	--

5. 해외 수어 말뭉치 공개 방법

1) 독일

함부르크 대학교, 독일 수어 및 농인 의사소통 연구소(Institute for German Sign Language and Communication of the Deaf at Hamburg University)에서 수행된 독일 수어(DGS) 말뭉치는 누리집을 통해 공개되고 있다. 누리집에서는 독일 수어 말뭉치 프로젝트는 15년 기간의 장기 프로젝트이며, 이를 통해 독일 전역 농인들의 일상 언어를 대표할 수 있도록 했으며, 향후 온라인으로 제공되는 말뭉치 기반의 ‘독일 수어-독일어 사전’이 개발될 예정임을 설명하고 있다.



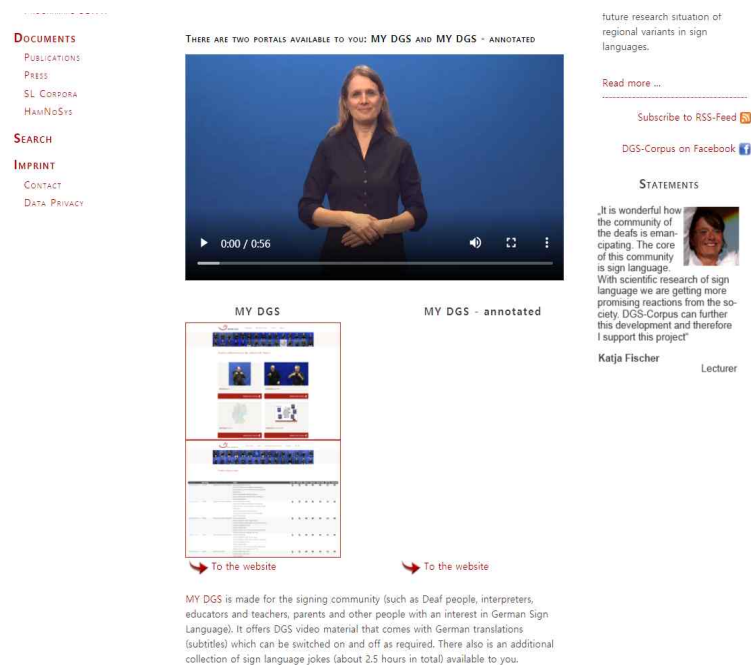
[그림 IV-12] 독일 수어 말뭉치(DGS-KORPUS) 누리집 접속 화면

※ 출처: <https://www.sign-lang.uni-hamburg.de/dgs-korpus/index.php/welcome.html>

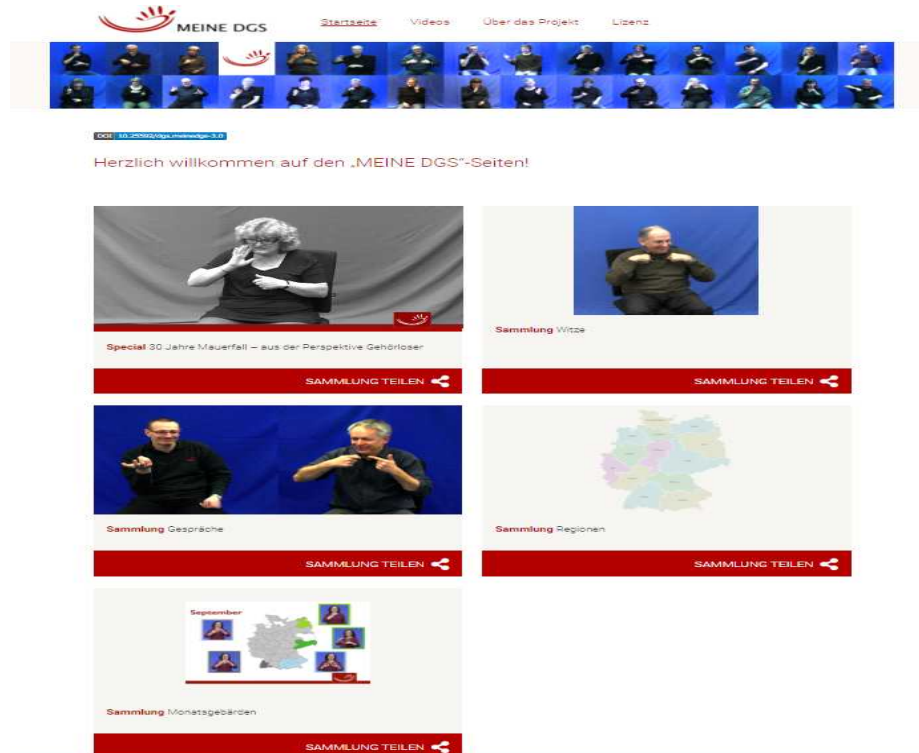
독일 수어 말뭉치 누리집의 왼쪽 메뉴에 있는 ‘공공 말뭉치(PUBLIC CORPUS)’로 접속하면 공공 독일 수어 말뭉치(PUBLIC DGS CORPUS) 자료에 대한 설명을 볼 수 있다. 공개되는 공공 독일 수어 말뭉치(PUBLIC

DGS CORPUS)는 독일 수어 말뭉치로 구축된 자료 중 약 50시간 분량의 독일 수어 영상 자료이며, 이들 영상은 여러 지역의 다양한 연령대의 농인이 독일 수어로 여러 가지 주제에 대해 이야기하는 것이다. 주로 농 공동체의 문화와 관련된 주제(47시간 이상)를 중심으로 구축된 독일 수어 말뭉치 중 약 560시간의 대화와 설명 형태의 영상 중에서 선택한 것들이다. 이 말뭉치는 [표 IV-13]와 같이 일반 공개(MY DGS)와 연구자 공개(MY DGS-annotated)로 구분하여 공개하고 있다. 원하는 방법(MY DGS 또는 MY DGS-annotated)을 선택하면 공개 화면으로 접속할 수 있다.

두 형태 모두 무료 이용이 가능하며 회원 가입 없이 이용할 수 있다.



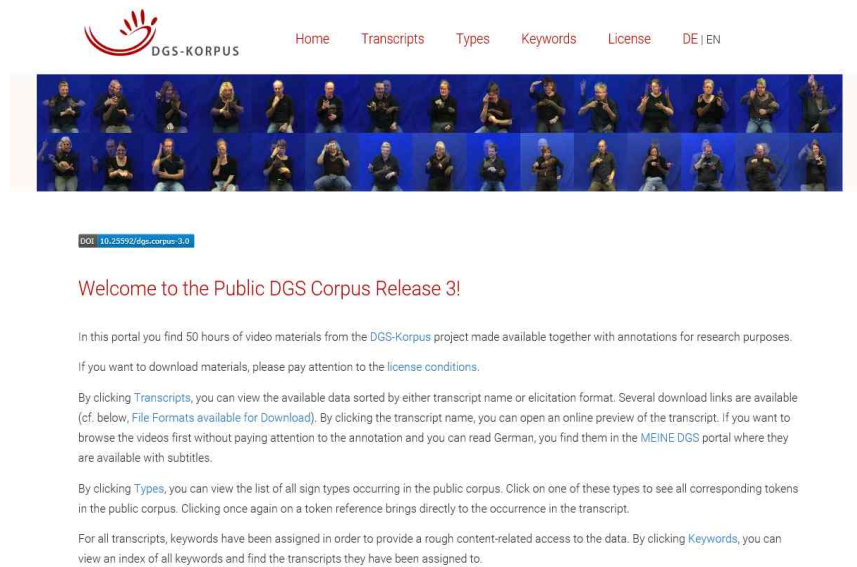
[그림 IV-13] 일반 공개(MY DGS)와 연구자 공개(MY DGS-annotated)



[그림 IV-14] 일반 공개(MY DGS) 접속 화면

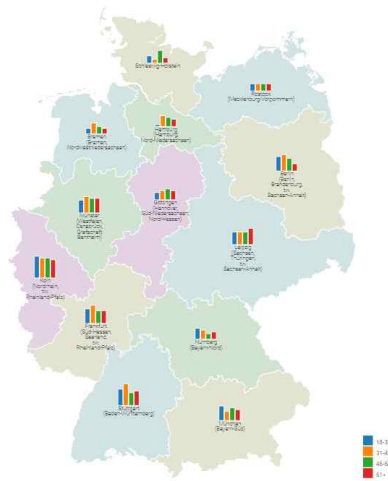
‘나의 독일 수어(MY DGS)’는 일반인 공개용이며, 수어 공동체(예: 농인, 통역사, 교육전문가 및 교사, 부모 및 기타 독일 수어에 관심 있는 사람들)를 위해 만들어졌다. 노출 여부(on/off)를 조정할 수 있는 독일어 번역(자막)과 함께 독일 수어 영상 자료가 제공된다. 2020년 9월 기준으로 ‘나의 독일 수어(MY DGS)’로 접속하면 ‘농인의 관점에서 베를린 장벽 붕괴 30년’이라는 주제의 스페셜 영상과 ‘농담 모음’, ‘대화 모음’, ‘수집 지역’, ‘월별 수어 모음(달을 세는 단위)’이라는 텍스트와 함께 대표 영상 또는 사진이 게시되어 있다. ‘베를린 장벽 붕괴 30년’, ‘농담 모음’, ‘대화 모음’을 선택하면 각 언어제공자의 영상이 게시되어 있으며, 영상의 아래에는 언어제공자의 지역, 과제 형식(토론 또는 대화), 주제에 대한 설명이 있다. ‘수집 지역’을 선택하면 독일 지도가 나타나며 독일 지도에서 하나의 지역을 선택하면 공개되는 모든 영상 중 해당 지역 언어제공자의 수어 영상을 볼 수 있다. ‘월별 수어 모음’에 접속하면 접속

하는 달(예: 9월)과 관련된 여러 수어와 독일 지도가 색깔로 대비되어 나타난다. ‘월별 수어 모음’은 1월에서 12월까지 지역별 수어 변이형을 살펴볼 수 있도록 지도와 함께 제시되어 있다. 나의 독일 수어(MY DGS)의 모든 영상은 페이스북(facebook), 트위터(twitter), 인스타그램(instagram) 등 소셜 미디어로 공유가 가능하다.



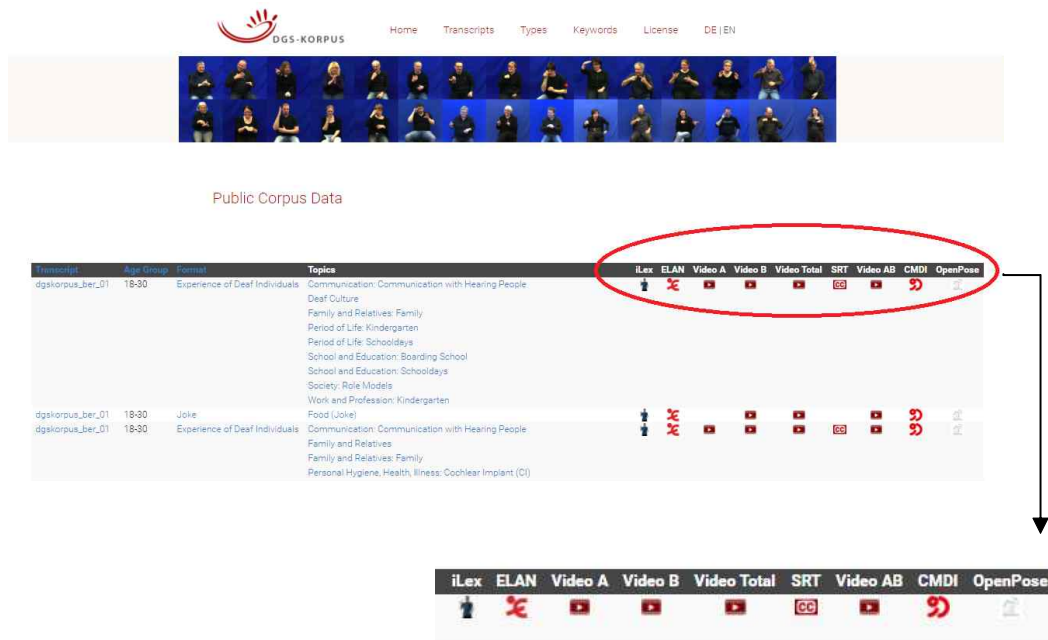
[그림 IV-15] 연구자 공개(MY DGS-annotated) 접속 화면

공공 독일 수어 말뭉치(PUBLIC DGS CORPUS) 화면에서 화살표를 선택하여 접속하면 연구자 공개용 화면(MY DGS-annotated)으로 접속할 수 있다. ‘독일 수어 말뭉치 공개 3에 온 것을 환영합니다!(Welcome to the Public DGS Corpus Release 3!)’라는 붉은 제목 아래에는 간략한 자료 수집 방법, 13개의 자료 수집 지역, 330명의 언어제공자, 주식 규칙, 인용 방법 등의 공개된 말뭉치 영상에 대한 배경 정보가 지도와 함께 자세히 설명되어 있다.



[그림 IV-16] 지역별 언어제공자 연령대별 현황

‘주석 달린 나의 독일 수어(MY DGS-annotated)’는 연구자 공개용이며, 수화언어학 분야를 연구하는 연구자들과 독일 수어, 수어 공동체, 농문화에 관심 있는 사람들이 과학적 관점에서 접근할 수 있는 자료로서 제공된다. 연구자 공개(MY DGS-annotated)는 일반인 공개(MY DGS)에서 제공되는 영상 외의 과제 영상(예: 스토리 다시 말하기, 내레이션, 교통 표지판에 대한 대화, 약속 정하기 등)을 총 100분 이상 제공하고 있다. 농담 과제 자료를 제외한 모든 동영상에는 번역, 글로스, 입 움직임 주석이 달려 있으며, 검색이 가능하고 영상과 번역이 포함된 주석 파일이 저장되어 있다.

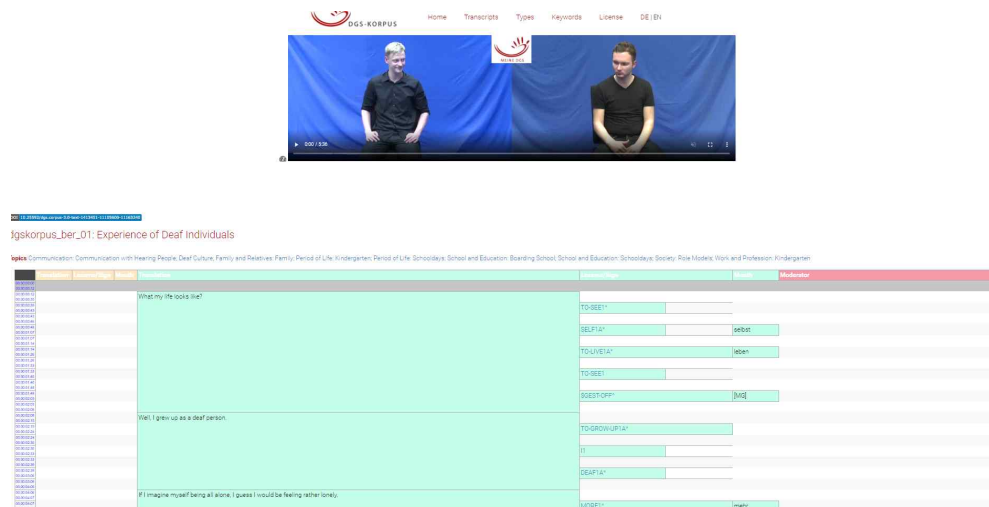


[그림 IV-17] 말뭉치 과제별 영상 및 주석 파일 저장 화면

연구자 공개(MY DGS-annotated) 누리집의 상단 메뉴 중 주석(Transcripts)을 선택하면 말뭉치 영상의 주석 파일명, 언어제공자의 연령대, 과제 영역(농담, 대화, 사건 등), 하위 주제, 언어제공자별 영상, iLex 파일과 엘란 파일 등의 순으로 정리되어 있다.

여기서 언어제공자의 수어 영상과 주석 파일을 내려받기를 해서 저장하여 사용할 수 있다. 특히 주석 프로그램은 iLex와 엘란 두 가지 형식을 모두 갖추고 있어 선택할 수 있다. 주석 프로그램을 설치한 후에 원하는 영상(언어제공자 A, 언어제공자 B, 모든 영상, 언어제공자 A와 B 영상) 파일을 내려 받을 수 있다. iLex 가져오기 파일과 엘란 파일에는 https를 통해 파일에 접근할 수 있는 URL도 제공한다. 자막(SRT) 파일은 MaxQDA와 같은 도구를 사용하여 내려 받을 수 있다. OpenPose에는 영상 자료의 주석을 사용자가 분석하여 사용할 수 있도록 제공한다. 영상이 익명화된 경우 OpenPose의 항목은 비어 두게 되며 영상 파일은 용량이 커서 압축되어 있다. CMDI에서는 참가자는 물론 각 영역에 대한 메타데이터가 포함된 파일을 내려 받을 수 있는 표시이다.

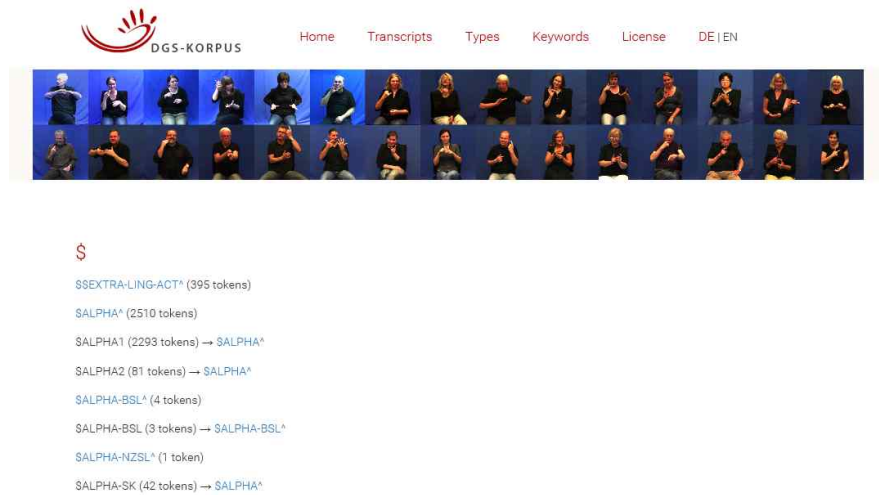
언어제공자별 영상이나 iLex 파일과 엘란 파일은 저장할 수 있게 되어 있으나 누리집 내에서도 영상과 주석 파일을 즉시 확인할 수 있도록 구성 되어 있다. 즉 목록 중 왼쪽의 주석 파일명을 선택하면 두 명의 언어제공자 영상과 영상 아래에는 주석 토큰의 시간과 주석 내용을 볼 수 있다. 영상의 왼쪽 아래에 있는 플레이 버튼을 누르면 해당 영상을 작동하며 언어 제공자의 수어 발화에 따라 아래에 주석되어 있는 내용도 이동하는 것을 확인할 수 있다.



[그림 IV-18] 파일명 ‘dgskorpus_ber_01’의 예시 화면

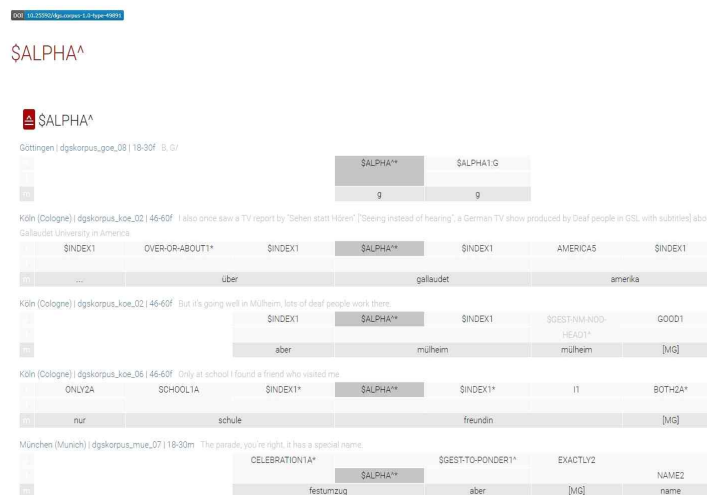
※ 출처: https://www.sign-lang.uni-hamburg.de/meinedgs/html/1413451-11105600-11163240_en.html

연구자 공개(MY DGS-annotated) 누리집의 상단 메뉴 중 타입(Types)으로 접속하면 공개된 말뭉치(public corpus)에 있는 모든 수어 타입의 목록을 볼 수 있다.



[그림 IV-19] ‘타입’ 메뉴 접속 화면

타입 중 하나를 선택하면 해당 타입이 나타나는 모든 파일의 주석 내용이 나타난다. 이 중 파일명을 선택하면 선택한 파일에서 타입이 나타나는 주석 화면이 영상과 함께 표출되는 것을 확인할 수 있다. 이것은 해당 타입이 산출되는 발화 영상과 주석 화면이 [그림 IV-20]의 형태로 제시된다.

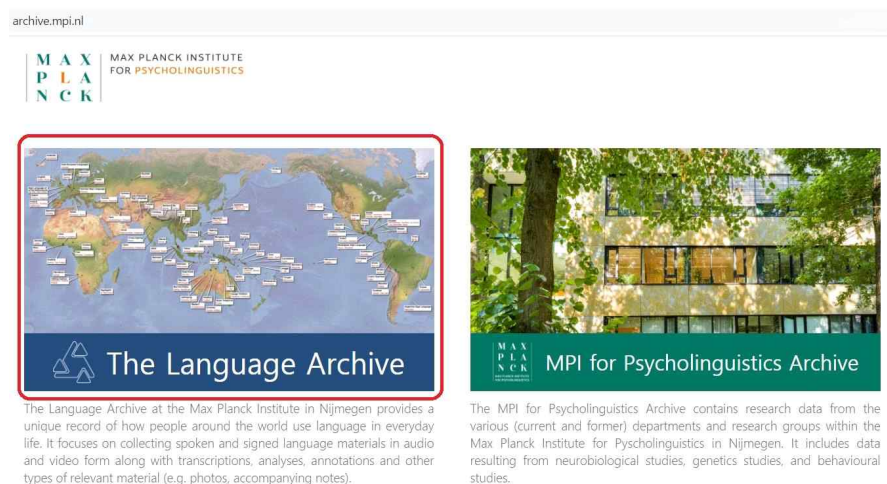


[그림 IV-20] 타입 ‘≡ \$ALPHA^’ 접속 예시 화면

주석되어 있는 모든 말뭉치 자료는 접근이 용이하도록 키워드(keyword)를 할당하였다. 상단 키워드를 선택하면 모든 키워드의 색인을 볼 수 있고 키워드에 할당된 주석을 찾을 수 있다. 입 움직임에 관한 주석을 제외하고 모든 정보는 독일어와 영어로 이용할 수 있다. 또한 주석 지침도 공유되어 있어 주석에 관한 내용을 참고할 수 있다.

2) 네덜란드

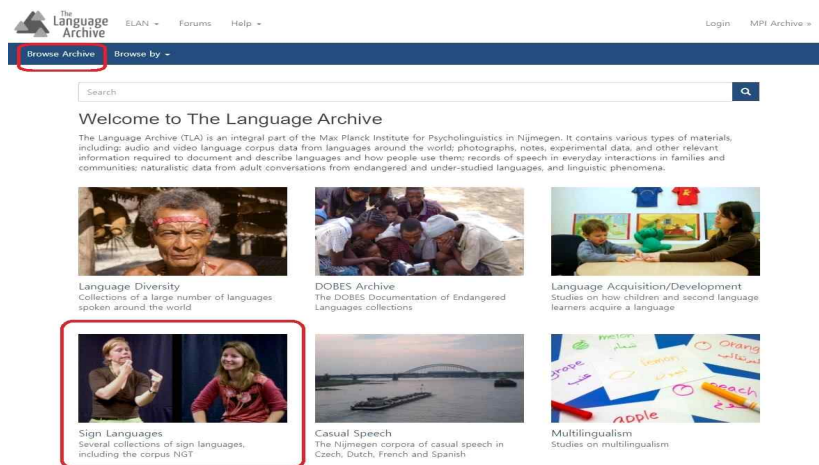
네덜란드 수어 말뭉치는 막스 플랑크 연구소(Max Planck Institute for Psycholinguistics)의 언어 보관소(TLA, The Language Archive)에서 공개하고 있다. 언어보관소에는 세계 여러 나라의 언어 데이터 및 기타 관련 정보를 보관하고 있다.



[그림 IV-21] 막스 플랑크 언어 연구소 누리집

※ 출처: <https://archive.mpi.nl/>

‘언어보관소(The Language Archive)’를 접속하면 [그림 IV-22]과 같이 언어보관소 누리집으로 연결된다.



[그림 IV-22] 네덜란드 언어보관소 누리집

※ 출처: <https://archive.mpi.nl/tla/>

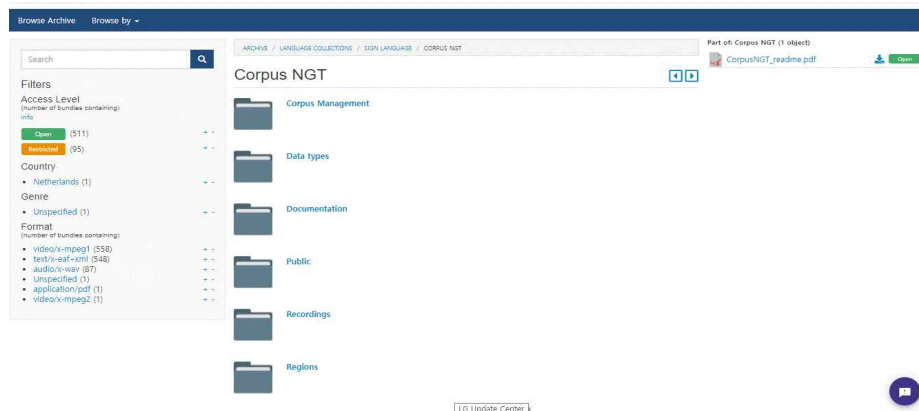
언어보관소 상단 메뉴 중 ‘보관소 검색(Browse Archive)’을 접속하면 [그림 IV-24]에서 보는 바와 같이 전 세계의 멸종 위기에 처한 100개가 넘는 언어의 시청각, 텍스트 및 기타 관련 정보를 포함하고 있는 위기의 언어 문서(DOKumentation BEdrohter Sprachen, DOBES) 보관소 그룹, 네덜란드 이중 언어 데이터가 있는 언어 모음 그룹, 막스 플랑크 심리언어 연구소에서 수집한 데이터가 보관된 MPI-PL 보관소 그룹으로 나뉘진 폴더 그림을 볼 수 있다. 네덜란드 수어 말뭉치는 이곳을 통해 공개되고 있다.



[그림 IV-23] TLA의 언어 보관소

※ 출처: https://archive.mpi.nl/tla/islandora/object/lat%3A1839_00_0000_0000_0000_0000_4

언어보관소 누리집으로 연결된 첫 페이지의 ‘수어(Sign Language)’를 접속하면 아프리카 수어, 아넘 랜드 수어, 오스트리아 수어, 네덜란드 말뭉치(Corpus NGT), Echo-수어(4개의 유럽 수어 말뭉치) 등 여러 나라의 수어 그룹이 있다. 이 중 네덜란드 말뭉치(Corpus NGT)를 접속하면 [그림 IV-24]과 같이 네덜란드 수어 말뭉치를 검색할 수 있다. 네덜란드 수어 말뭉치는 영상 파일과 엘란 주석 파일이 연결되어 있는 2,375개 파일로 구성되어 있다. 대부분의 주석 파일은 공개되어 누구든지 회원 가입 없이 데이터를 내려받을 수 있다.



[그림 IV-24] 네덜란드 수어 말뭉치(Corpus NGT)

※ 출처: https://archive.mpi.nl/tla/islandora/object/tla%3A1839_00_0000_0000_0004_DF8E_6

네덜란드 말뭉치는 ‘말뭉치 관리’, ‘사용된 유도자료 유형’, ‘문서’, ‘공개(public)’, ‘언어제공자 촬영’, ‘지역’ 그룹으로 분류되어 있다. 이 중 ‘사용된 유도자료 유형’, ‘지역’, ‘촬영’의 세 그룹에서는 제공되고 있는 정보는 메타데이터, 언어제공자당 여러 개의 동영상 파일, 주석 파일로 모두 동일하게 담고 있다. 공개 그룹에는 접근 가능한 모든 버전의 영상이 있고, 문서(documentation)에는 주석된 내용들이 담긴 주석 규약이 버전 별로 제시되어 있다.



[그림 [IV-25] 네덜란드 수어 말뭉치(Corpus NGT) 예시 화면

이러한 네덜란드 수어 말뭉치는 주로 언어학자를 위한 언어자료이지만 통역사 또는 예비 통역사, 네덜란드 말뭉치 교사 및 학습자, 청각장애인 등에게 데이터가 유용하게 사용되는 것을 목적으로 한다. 언어학자들은 문법 분석, 사회언어학 연구 및 교차 언어 비교 및 수어 및 구어 연구 등에 말뭉치를 사용할 수 있으며, 수어 교사들이 연습용 자료로도 활용할 수 있다. 통역사는 다양한 지역 수어에 익숙해질 수 있는 자료로 활용할 수 있으며, 농인 당사자는 언어 표준화에 대한 논쟁을 탐구할 수 있는 좋은 자료가 된다. 또한 청각장애 아동의 부모는 자녀와 함께 우화 이야기, 트위터, 실베스터 만화 등의 이야기를 보고 싶을 때 사용할 수 있다.

3) 호주

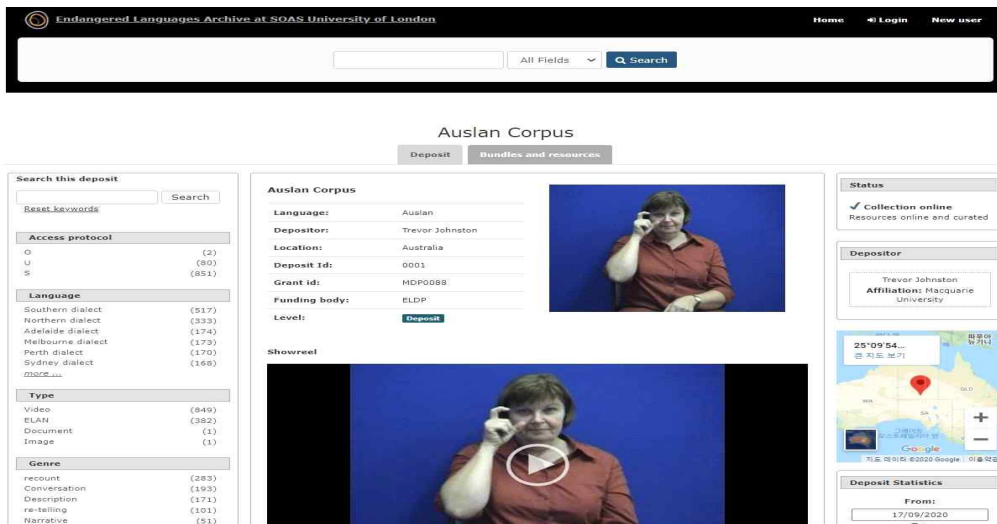
호주 수어 말뭉치는 전 세계의 멸종 위기에 처한 언어 자료를 보존하고 게시하는 디지털 저장소 ELAR(Endangered Language Archive)에 보관 및 공개되고 있다. 2005년 처음 등록된 이후 2016년까지 추가된 주석 자료가 보관되어 있으며 주석 지침은 ‘호주 수어 사인뱅크(Signbank)’에서 제공되고 있다.

이곳 호주 수어 보관소에서 공개되는 호주 수어 말뭉치는 주로 약

100명의 농인에 대한 영상과 엘란으로 제작한 주석 파일로 구성되어 있다. 영상은 인터뷰, 다시 이야기하기, 개인적인 사건 회상, 설문지 응답, 자발적인 대화 참여, 그림책 이야기 등 호주 수어를 유도할 수 있는 다양한 유도 자료를 통해 수어자의 발화를 3시간 분량의 영상으로 촬영한 후 각 주제별로 편집한 것이다. 이러한 편집 영상은 연령, 성별, 지역(방언 및 지역별) 및 텍스트 유형에 따라 분류되어 있다. 그래서 말뭉치 자료는 지역, 자료 유형(영상, 엘란 파일), 유도 과제의 유형, 주제, 참가자 정보 등에 따라 검색할 수 있다. 영상을 내려받기 위해서는 계정을 만들어야 한다.

주석 파일이 되어 있는 대표적인 영상은 ‘늑대를 올린 소년’과 ‘토끼와 거북이’ 주제의 이솝 우화에 대한 것이다. 이것은 초기 말뭉치 기반의 문법 연구에서 수어의 실증적인 자료로 활용하기 위한 것으로 영국 수어(BSL) 계통의 오래된 수어와 농사회에서 새로 생성된 수어를 비교하기 위한 기초 자료로 활용되기도 한다.

이곳 보관소에 저장된 자료에는 개인 정보 식별이 포함되어 있거나 참여자가 민감하게 여기는 주제(예: 낙태, 유전자 검사, 장애)에 대한 개인적인 의견을 공개하거나 지역 농사회의 다른 사람들에 대해 논의하는 내용이 포함되어 있다. 그래서 주석이 기재되지 않은 영상이라 할지라도 데이터의 사용 목적을 확인하고, 제3자에게 제공하지 않고 교육 목적으로만 사용하겠다는 조건에 대한 여부를 확인한 후에 사용할 수 있도록 하고 있다.



[그림 IV-26] 호주 수어 말뭉치 보관소

※ 출처: <https://elar.soas.ac.uk/Collection/MPI55247>

4) 영국

영국 수어 말뭉치는 2008년과 2011년 사이에 경제사회연구원에서 지원한 자금으로 수행된 영국수어 말뭉치 프로젝트의 일부로 수집된 것을 공개적으로 접근할 수 있도록 한 온라인 기록물로 이중으로 접근(A DUAL ACCESS CORPUS)할 수 있게 만들어졌다. 2011년 여름부터 런던 대학교의 ‘인간 의사소통 시각-영상 보관소(the University College London “Human Communication Audio-Visual Archive”, CAVA)를 통해 영국 수어 말뭉치 영역(British Sign Language Corpus Project, BSLCP)에서 영국 전역 8개 지역의 영상을 온라인에서 제공하고 있다. CAVA는 비밀번호장 동의를 한 사용자 등록을 통해 검색과 내려받기가 가능한 안전한 시스템이다. CAVA는 참가자의 메타데이터를 검색할 수 있는데 연구자, 영국수어교사, 통역자 등에게 우선적으로 제공하기 위한 것이었다. CAVA 누리집뿐만 아니라 영국수어 말뭉치 영상 자료를 사용자 편의성을 고려하여 일반 사용자들이 이용할 수 있도록 하였다. 공개 자료는 지역, 과제, 나이군 등을 선택하여 볼 수 있도록 하였다. 이것은 농사회를 대상으로 말뭉치 이용을 장려하기 위해 CAVA와 분리하여 제작된 것으로 영국수어와 영어 두 가지로

설명이 되어 있는 단순하고 이용하기 쉬운 공개 시스템이다. 이처럼 영국 수어 말뭉치에 접근하기 위해서는 두 가지 방법이 있다.



[그림 IV-27] 영국 수어 말뭉치 프로젝트 누리집의 데이터 공개 안내

※ 출처: <https://bslcorpusproject.org/data/>

CAVA

A Human Communication Audio-Visual Archive for UCL

Home / CAVA home (BSLCP)

> BSL Corpus Project Home

> CAVA home (BSLCP)

> Acknowledgements and Citation

> BSLCP Activities

> Open Access Data

> Restricted Access Data

> Search/Browse CAVA Repository

CAVA home (British Sign Language Corpus Project)

Background

The CAVA (human Communication; an Audio-Visual Archive) repository contains rights-cleared primary audio and video recordings made by UCL's human communication researchers and their associates; The British Sign Language Corpus Project (BSLCP) section of the CAVA repository contains video recordings of BSL video data from 249 deaf signers of BSL from 8 regions around the UK: London (L), Bristol (BL), Cardiff (CF), Birmingham (BM), Newcastle (N), Manchester (M), Glasgow (G) and Belfast (BF). The participants are mixed for gender, age group, age of BSL acquisition, social class and ethnicity. There are 2 levels of access. The **Open Access** data includes narratives (n) and lexical elicitation (l), and the **Restricted Access** data includes interviews (i), conversation (c) and some combined narrative followed by conversation (n-c) – for more detail see BSLCP Activities. All participants were filmed in pairs for all activities, with three cameras: one on each participant, and one on the pair. Some annotations are available as well (see below for details).



Example Video Resource filenames and what they mean:

BF9+10i = Belfast participants number 9 and 10 in the interview activity

CF17n = Cardiff participant 17 in the narrative activity

Example Annotation (.eaf) Resource filenames and what they mean:

LU12F40BHC.eaf = London participant 12, Female, age 40, Black, from Hearing family, in the Conversation activity

[그림 IV-28] CAVA 영국 수어 말뭉치 보관소

※ 출처: <https://bslcorpusproject.org/cava/>

5) 미국

미국 수어 말뭉치는 미국 수어 연구 프로젝트(ASLLRP) 자료 보관소(Database Access Interface, DAI)를 통해 공개되고 있다. 미국 수어의 농인 수어자들로부터 얻은 상당한 양의 데이터는 국립 수어 및 제스처 센터(National Center for Sign Language and Gesture Resources, NCSLGR)와 함께 수집되었다. NCSLGR에서는 수어와 구어의 제스처 구성 요소에 대한 컴퓨터 언어학 연구를 돕기 위해 실시한 프로젝트에서 얻은 여러 유형의 실험 연구 및 분석한 데이터를 제공한다. 영상 기반의 언어 자료 수집을 위해 보스턴 대학교와 럿거스(Rutgers) 대학교에 전용 시설을 설치하였는데, 각 시설에는 여러 대의 동기화된 디지털 카메라가 장착되어 대상물에 대한 다양한 관점에서의 촬영이 가능하다.

원어민 수어 사용자로부터 수집된 수어 영상 데이터는 1999년 12월부터 수집되었으며, 이 영상은 비압축 및 압축 형식으로 제공되고 있다. 영상의 언어 데이터를 분석하기 위해 주석을 지원하도록 설계된 멀티미디어 데이터 도구인 SignStream®으로 상당 부분의 영상이 주석되어 있다. SignStream®은 영상을 조작하고 특정 프레임을 세밀한 단계로 주석할 수 있으며 연구자, 교육자 및 학생에게 무상으로 배포된다. 이러한 데이터는 다른 플랫폼에서 사용할 수 있도록 텍스트 형식으로 내보낼 수도 있다.

[그림 IV-29] 미국 수어 말뭉치 보관소

※ 출처: <https://dai.cs.rutgers.edu/dai/s/dai>

6) 해외 수어 말뭉치 공개의 시사점

해외 수어 말뭉치는 구축과 공개를 함께 하고 있으며 공개되는 범위와 방식은 <표 IV-15>에서 보는 바와 같이 나라별로 상이하였다.

첫째, 공개를 위하여 별도의 페이지를 마련하고 있다. 영국과 독일은 수어말뭉치 보관소 내에서 일반인용과 연구자용 페이지를 구축하여 제공하고 있다. 일반인용은 영상만 제공되는 것이며 연구자용에서는 주석 파일도 공개된다. 호주와 네덜란드는 영국과 독일처럼 대상자별 구분이 없으며 수어 말뭉치 보관소인 누리집에 접속하면 탑재된 것 중 일부의 내용을 볼 수 있다.

둘째, 공개 영상은 수집된 전체 수어 말뭉치 중에서 선택된 일부의 과제 영상이며, 과제 영상의 선택 기준은 나라별로 상이하였다. 호주의 경우 849개의 영상이 공개되고 있으나 주석 파일은 이습 우화 과제에 대해서만 공개하고 있었다. 독일은 총 560시간의 수어말뭉치 영상 중 56시간에 한해

공개하고 있다.

셋째, 공개된 영상 및 주석 파일에 대한 다운로드가 가능하나 일부 나라에서는 회원가입을 필요로 하는 경우도 있다. 호주는 회원가입이 필수이며, 영국은 제한된 파일에 대해서는 회원가입을 요구한다. 그러나 독일과 네덜란드는 회원가입을 필요로 하지 않는다.

넷째, 공개 시기도 나라별로 다르다. 2008년부터 말뭉치 구축을 시작한 영국은 2011년, 2006년부터 말뭉치 사업이 시작된 네덜란드는 2013년, 2004년부터 말뭉치 구축이 시작된 호주는 2011년부터 수집한 수어 말뭉치를 공개하기 시작하였다. 독일은 2009년부터 장기 프로젝트로 시작한 대규모 말뭉치를 2020년부터 공개하기 시작하였다. 이는 나라별로 짧게는 3년, 길게는 11년 동안 구축된 말뭉치가 공개된 것이다.

이상의 내용들을 기반으로 우리나라 수어 말뭉치 공개를 위해 사전에 심도 있게 구체적인 검토와 논의가 되어야 할 것들은 다음과 같다.

- 수어말뭉치 탑재 장소
- 공개를 회원 가입 여부
- 공개의 내용과 범위
- 공개 시기
- 저작권 및 SNS
- 타입목록 공개 여부
- 공개 자료의 영어번역 여부
- 공개 자료 정리 방법 및 검색 기능

〈표 IV-15〉 해외 수어 말뭉치의 공개 현황

구분	호주수어 말뭉치 Auslan Corpus	네덜란드수어 말뭉치 NGT Corpus	영국수어 말뭉치 BSL Corpus	독일수어 말뭉치 DGS Corpus
탐재 장소	Endangered Languages Archive (ELAR) at SOAS University of London	The Language Archive (Max Planck Institute)	영국수어말뭉치 내 아카이브	독일수어말뭉치 내 아카이브
탐재 인터넷 주소	https://elar.soas.ac.uk/ 호주수어말뭉치와 호주수어 아카이브는 다름, 부분적으로 탐재	https://archive.mpi.nl/ 탐재된 다수 파일 공개	일반 포탈은 영상 재생만. https://bslcorpusproject.org/data/region/ 연구자 포탈 https://bslcorpusproject.org/cava/ 영상 및 전사파일 공개 (과제별 공개 및 제한된 공개) 공개: 이야기 및 개별된 수어 제한된 공개: 인터뷰 및 대화	일반 포탈: https://www.sign-lang.uni-hamburg.de/meinedgs/overview/start.html 연구자 포탈: https://www.sign-lang.uni-hamburg.de/meinedgs/ling/start_de.html
탐재 영상 시간	영상 849 개			56시간 (약 560시간 중 선택 영상)
탐재 주석 파일 개수	Aesop's fables 과제만	2,375 파일		406 파일
회원가입 필요	회원가입 필요	회원가입 필요 없음	제한된 공개 파일은 회원가입 필요	회원가입 필요 없음
영상 다운로드 가능	가능	다수 영상 다운로드 가능	공개된 파일은 가능	탐재된 영상 모두 가능
주석파일 다운로드 가능	가능	다수 주석파일 다운로드 가능	공개된 파일은 가능	탐재된 파일 모두 가능
공개 시기	2012	2013	2011	2020

6. 한국수어 말뭉치 구축을 위한 중장기 계획

1) 중장기 계획 수립 배경

한국수화언어법 제정에 따라 한국수어의 체계적 수집과 보존은 시대적 요구가 되었다. 하지만 한국수어에 대한 언어학적 체계를 수립하기에는 연구 기반이 미비한 상태에 머물고 있다. 2014년 기초 연구를 시작으로 올해까지 7년 동안 지역별 한국수어를 수집하고, 주석을 달았으며, 그 자료들을 「한국수어 자료 통합지원 시스템」에 보존 관리하고 있다. 이에 중장기 계획을 수립하여 한국수어 말뭉치 구축을 더 체계적으로 중장기적 관점에서 수행해야 할 필요성이 대두되었다.

□ 한국수화언어법 제정에 따른 한국수어 수집 및 보존 요구

- 농인 공동체의 문화유산 보존 요구
- 전국 단위의 지역 수어 변이 양상을 파악하기 위한 풍부한 원시 말뭉치 자료 확보
- 한국수어 자료를 보존하기 위한 완성도 높은 시스템 요구

□ 한국수어의 체계적 연구를 위한 주석 처리된 수어 자료 요구

- 한국수어의 구조와 특성 연구를 위한 기초 자료의 미비
- 한국수어에 대한 기본주석 및 의미주석에 대한 요구
- 전문적 주석 작업자의 안정적 공급 요구

□ 한국수어 말뭉치 자료의 보급 요구

- 7년간 5개 지역에서 수집된 수집 자료의 민간 보급 요구 증대
- 멸종 위기 언어로서 수어의 안정적 보호 및 보존
- 수어 연구자를 위한 기초 자료 제공

2) 한국수어 말뭉치 구축 SWOT 분석과 대응 전략

(1) 환경 분석

강점(S)	약점(W)
<ul style="list-style-type: none"> 가장 자연스러운 수어 자료 확보 자연언어 연구를 위한 언어 표본 데이터 확보 한국수어의 빈도, 분포 확인 가능 한국수어 연구 기반 마련 	<ul style="list-style-type: none"> 한국수어 체계에 대한 근거 미흡 말뭉치 예산 및 전문인력 투입 미비 전문 촬영팀 부재 주석 작업의 지속성 부재로 주석의 일관성 및 전문성 부재
기회(O)	위기(T)
<ul style="list-style-type: none"> 한국수화언어법 제정(2016) 코로나19 계기 공공수어 통역의 확대로 수어에 대한 관심 고조 기계번역을 위한 한국수어자료에 대한 요구 	<ul style="list-style-type: none"> 한국수어 사용인구 감소 기계번역 및 인공지능 학습을 위한 과학기술부의 수어 자료 수집 참여로 인한 자연수어 수집에 대한 필요성 상실 우려

[그림 IV-30] 한국수어 말뭉치 구축 SWOT 분석

한국수어 말뭉치 구축 사업이 갖는 강점으로는 한국수어의 지역별 수집이 가능하고, 자연스러운 상태에서의 수어를 확보할 수 있으며 한국수어의 형태 분석이 가능하고, 이렇게 수집 및 가공된 수어 자료를 전산시스템에서 관리할 수 있다는 것이다. 이는 향후 한국수어 연구를 위한 기초 자료로서 중요한 역할을 하게 될 것이다. 약점으로는 한국수어 말뭉치 구축 경험이 길지 않으며, 한국수어 체계에 대한 수어학적 근거가 미흡하고, 말뭉치 구축 예산이 한정적이어서 구축을 위한 집중적 작업이 어려우며, 한국수어 수집을 위한 전문 촬영팀이 부재하여 촬영 경험이 축적되지 않고 있고, 전문 주석인력이 안정적으로 유지되지 않아 주석의 일관성, 타당성, 신뢰성 확보에 어려움이 있다는 것을 들 수 있다.

하지만 본 사업의 기회 요소로는 2016년 한국수화언어법이 제정되어 본 사업을 추진하기 위한 법적 기반을 확보하였으며, 코로나 19로 인한 공공수어통역이 대중으로 하여금 수어에 대한 관심을 고조시켰으며, 기계번역을 위한 한국수어자료에 대한

요구가 급격하게 증가하고 있다는 것이다. 위기 요소로는 한국수어 사용인구가 점진적으로 감소하고 있으며, 기계번역을 위한 수어 수집과 한국수어 말뭉치 구축을 위한 지역별 자연수어 수집 간의 차별성에 대한 인식이 결여되어 있고 이로 인해 충분한 예산 확보가 어려워질 수 있다는 위기도 있다. 아울러 한국수어의 문법체계 수립이 아직 미진한 상태이어서 말뭉치 주석에 대한 근거가 불분명한 경우가 발생할 수도 있다.

(2) SWOT 분석에 따른 대응 전략

한국수어 말뭉치를 둘러싼 환경의 강점과 약점, 그리고 기회와 위협을 중심으로 분석된 내용은 한국수어 말뭉치의 대외적인 환경을 기반으로 교차 분석한 결과 [그림 IV-31]와 같은 전략과제가 도출될 수 있다.

강점(S)-기회(O) 전략	강점(S)-위기(T) 전략
<ul style="list-style-type: none"> ☞ 법에 기초한 장기 사업계획 수립 ☞ 지역 농아인협회와의 협력 관계 체결 ☞ 지역 농아인협회를 통한 말뭉치 인식 제고 사업 추진 ☞ 지역별 자연수어 제공자 인력풀 구축 	<ul style="list-style-type: none"> ☞ 말뭉치 공개를 통한 수어 대중성 제고 ☞ 국립국어원, 정보통신부 등 국가 기관 간의 상호협력 관계 체결 ☞ 한국수어 말뭉치 자료 공개를 통한 한국수어학 연구 활성화 ☞ 지속적인 주석 작업 체제 구축
약점(W)-기회(O) 전략	약점(W)-위기(T) 전략
<ul style="list-style-type: none"> ☞ 한국수어 말뭉치 전문 인력 양성 연수회 개최 ☞ 예산 확보를 통한 단기간 내 전국 단위 지역별 한국수어 수집 완결 ☞ 전문 주석자 양성 및 고용 	<ul style="list-style-type: none"> ☞ 전문 촬영팀 확보 ☞ 주석 데이터 일관성 및 신뢰성 확보를 위한 주석 규약 정교화 ☞ 「한국수어자료통합지원시스템」 기능 개선

[그림 IV-31] SWOT 분석에 따른 대응 전략

SWOT 분석 결과를 토대로 한국수어 말뭉치 구축 사업 발전을 위한 실행 전략들을 구체화하기 위해, 1) S-O 전략: 강점을 기회로 전환, 2) W-O 전략: 약점을 보

완하여 기회를 생성, 3) S-T전략: 강점을 통해 위협을 극복, 4) W-T 전략: 약점을 보완하며 위협을 최소화하기 위한 전략들을 구상하였다.

첫째, 강점을 기회로 만드는 S-O 전략으로 한국수어 말뭉치 구축 사업은 한국 수화언어법을 기반으로 현재 1년 단위로 수립되고 있는 계획을 장기적 관점에서 계획을 수립하여 수행할 필요가 있으며, 지역 수어 수집을 위한 언어제공자를 확보하기 위해 지역별 농아인협회와의 협력관계를 체결하여야 하고 이와 동시에 지역 농아인협회를 대상으로 수어 말뭉치에 대한 인식 제고 사업이 추진되어야 할 것이다. 아울러 지역별 자연수어 제공자 인력풀을 사전에 구축하여 언어제공자의 자격 요건을 충분히 검정함과 동시에 인력을 사전에 확보해야 할 필요가 있다.

둘째, 위협을 강점으로 극복하는 S-T 전략으로는 먼저 지금까지 축적된 한국수어 말뭉치 데이터를 공개하여 수어의 대중성을 제고해야 할 것이다. 이와 함께 국립국어원과 정보통신부 등의 국가기관 간 상호협력 체계를 구축할 필요가 있으며, 한국수어 말뭉치를 전문가에게 공개함으로써 한국수어학 연구를 활성화해야 하고 한국수어 원시 말뭉치에 대한 주석 작업을 지속적으로 할 수 있는 국립국어원 중심의 체계가 구축되어야 할 것이다.

셋째, 약점을 기회로 전환시키기 위한 W-O 전략으로서 한국수어 말뭉치 구축을 위한 전문인력 양성을 위한 연수회를 개최할 필요가 있고, 예산 확보를 통해 단기간 내 아직 수집이 되지 않은 지역의 한국수어 수집을 완결하여야 할 것이다. 동시에 전문 주석자를 양성하여 국립국어원에서 고용하여 상시 주석 작업이 이루어질 수 있도록 할 필요가 있다.

넷째, 외부의 위협과 내부 약점에 대한 영향을 최소화하기 위한 W-T 전략으로 지금까지 지역 촬영팀을 통해 1회성으로 촬영함으로써 발생했던 말뭉치 구축 사업에 대한 이해가 부족하여 영상의 질 확보의 어려움, 촬영 중단 사고 등과 같은 다양한 한계를 극복하기 위해 전문 촬영팀 확보가 요구된다. 또한 주석 데이터의 일관성 유지와 신뢰성 확보를 위해 주석 규약을 정교하게 다듬을 필요가 있으며, 무엇보다 「한국수어 자료 통합지원 시스템」의 기능을 대폭 개선하여 안정적 수어 데이터 저장과 관리가 절실하다. 나아가 해외 사례에서 본 바와 같이 가능한 빨리 수어 말뭉치 공개에 대비해야 할 것이다.

(3) 한국수어 말뭉치 구축 현황 분석

〈표 IV-16〉 한국수어 말뭉치 구축 현황

연도 주요과제	2015년 ~ 2018년	2019년 ~ 2020년
수어 자료 수집	<ul style="list-style-type: none"> - 2015년 촬영 및 편집(한 쌍 촬영 영상 3개의 합: 1,170개 영상), 일부 번역(서울 390 파일 중 110개) - 2018년 촬영 및 편집(한 쌍 촬영 영상 3개의 합: 542개) + 전체 번역(경남 26개 파일, 전북, 전남 각 52개 파일, 강원 48개 파일) 	<ul style="list-style-type: none"> - 2019년 촬영 및 편집(한쌍 촬영 영상 3개의 합: 429개), 번역 없음 - 2020년 촬영 및 편집(한쌍 촬영 영상 3개의 합: 738개), 번역 없음
	<언어제공자> 88명 <ul style="list-style-type: none"> - 서울: 30쌍(60명) - 전북, 전남, 강원: 각 4쌍 (24명) - 경남: 2쌍 (4명) 	60명 <ul style="list-style-type: none"> - 경남: 11쌍 (22명) - 전남: 19쌍 (38명) (광주 포함)
	<ul style="list-style-type: none"> - 촬영: 최소 108시간 이상 - 편집: 서울 70시간, 전북, 전남, 강원 3개지역 30.5시간, 경남 5시간 ☞ 105.5시간 - 번역: 서울 12시간, 4개지역 35시간 ☞ 47시간 	<ul style="list-style-type: none"> - 촬영: 최소 38시간 및 50시간 이상 - 편집: 경남 32시간, 전남 44시간 ☞ 76시간 - 번역: 실시되지 않음
주석 (번역 제외)	<ul style="list-style-type: none"> - 기본주석 - 연구 사례 : 입 움직임, 동형이의어, 색깔 	<ul style="list-style-type: none"> - 기본주석 + 의미주석 - 연구사례 : 생산적 수어의 도상성 -
	<ul style="list-style-type: none"> - 기본주석 서울 자료 1,213분(20시간 13분 분량) - 154,172 토큰 	<ul style="list-style-type: none"> - 의미주석 135분(2시간 15분 분량) (기존 주석 파일) - 경남 자료 기본 및 의미주석 80분 (1시간 20분)
공개	하지 않음	

국내 수어 말뭉치 수집은 2019년까지 55쌍, 110명을 대상으로 13가지 과제에 대해 한 쌍당 3시간 이상의 촬영 분량을 확보하였으며, 그 결과 13개 과제로 분할 편집된 영상은 총 2,141개 영상 파일, 137시간 56분 분량이였다. 2020년 추가로 19쌍, 38명을 대상으로 분할 편집된 738개의 영상, 43시간 분량을 추가 확보하여 총 2,879개 영상, 181시간 분량의 언어자료가 수집되었다. 이로써 국내 수

어 자료는 시간 환산을 위해 한 쌍의 언어제공자로부터 수집된 3개의 영상을 하나로 파일로 보았을 때, 2015년, 2018년, 2019년에 이어 2020년까지 5개 지역 총 148명의 언어제공자를 대상으로 945개 영상 181시간 분량의 영상이 수집되었다. 수어 자료가 대표성을 가지기 위해서는 최소 10쌍 이상의 언어제공자로부터 수어 자료가 확보되어야 하나, 2018년도 수집된 지역 중 전북과 강원도 지역의 언어 자료는 각 4쌍에 불과해 추가 수집이 필요한 상황이다. 또한 2020년 경남지역의 수어 자료에 대한 주석 작업을 진행하는 동안 기존의 주석 작업자가 수도권 지역의 인력으로서 지역 수어를 기록하는 데 제한이 있음을 파악하고 향후 수어 자료 수집에서는 수집 과정에서 지역의 수어 전문가를 통한 번역 활동이 동시에 이루어져야 할 것이다.

수어 영상을 자료화하기 위해서는 주석 달기가 필요하다. 주석 작업은 말뭉치를 활용한 연구 사례를 제시하기 위해 실시된 2016년의 ‘입 움직임’ 관련 주석 외 2015년과 2017년에는 기본 주석만 실시했다. 기본 주석은 음성언어 말뭉치에서 볼 때 발화의 형태를 그대로 기록하는 ‘전사’의 형태이지만 수어는 문자가 없기 때문에 동일한 형태를 가진 수어가 동일한 이름으로 기록되기 위해서는 모든 주석 작업자가 동일한 수어의 형태에 대해서 동일한 이름으로 인식하도록 하는 타입 영상의 확인과 입력 과정을 거쳐 작업하므로 작업에 소요되는 시간이 훨씬 많은 편이다. 이러한 까닭으로 그동안 기본 주석 작업을 중심으로 주석 달기가 진행되어 왔지만 수어 형태의 의미적 사용을 확인하기 위해 주석을 달 때 의미 주석의 필요성에 따라 2019년부터 의미 주석이 함께 진행되어 오고 있다. 2019년 수행된 주석 작업은 기본 주석이 이미 완료된 120분 분량의 수어 자료 데이터에 대해 기본 주석의 오류 수정 및 의미 주석 달기를 하는 것이었다. 그동안 서울지역을 중심으로 주석달기가 진행되었다면 2020년에는 경남지역의 수어 자료 80분을 대상으로 기본 주석 및 의미 주석을 함으로써 지역적 방언을 확인할 수 있는 계기가 마련될 수 있었다. 이러한 과정으로 진행되어 온 주석 작업은 현재까지 서울지역 수어 자료 1,213분 분량의 총 154,172개 토큰(우세손과 비우세손 포함)에 기본 주석이 완료되었으며, 이 중 120분 분량의 토큰에는 지시 타입을 제외한 토큰에 대한 의미주석도 기록되었다. 또한 경남지역 수어 자료 80분 분량의 영상에 기본 및 의미 주석이 완료되었다.

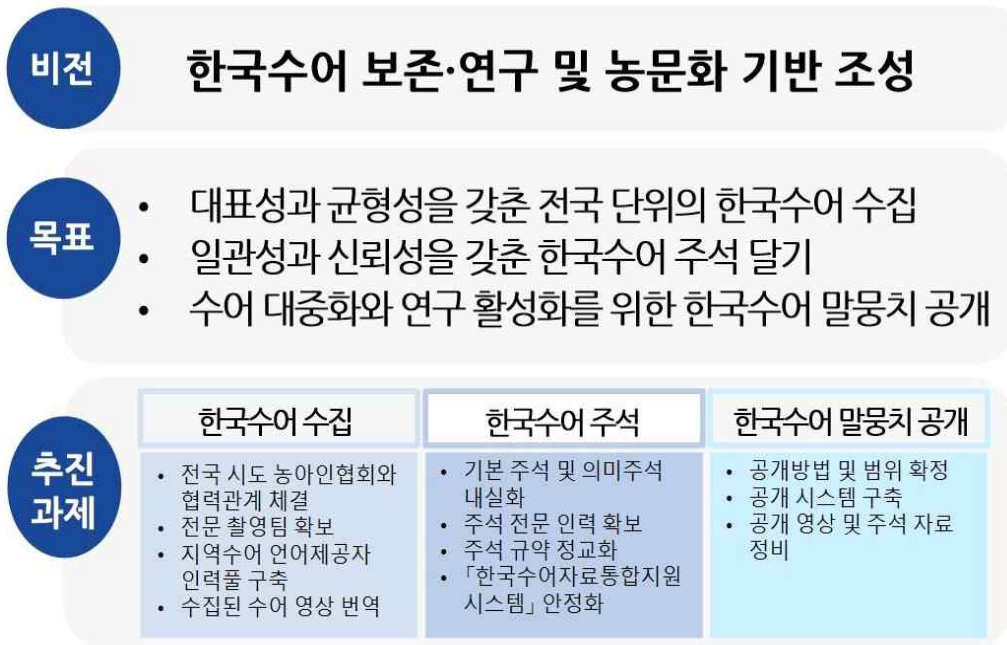
나라별로 공개하는 정보의 정도가 상이하여 일관된 기준으로 분석할 수는 없지만 <표 IV-14>에서 보는 바와 같이 10년 이상의 장기 프로젝트로 수행된 독일 자료와 비교하면 2배 이상 낮은 수행률을 보인다. 또한 수집된 영상 총 180시간의 분량 중 약 23시간에 대해서만 주석이 기록됨으로써 12.7%의 수행률에 머물고 있다.

3) 중장기 계획 수립

(1) 비전 및 추진과제 설정

한국수어 말뭉치는 단순히 실제적인 수화언어 샘플을 통해 언어를 연구하기 위한 목적으로 추진되지 않았다. 과학기술의 발달로 인공와우 사용자 및 통합교육 확대라는 사회적 변화는 수어를 모어로 사용하는 인구 축소와 한국수어의 발전과 전승을 대표하는 농학교 기반의 농사회 공동체 해체 위기를 가져왔다. 이런 맥락에서 농인 공동체의 문화를 담고 있는 한국수어의 발전과 보존을 위한 기반 마련을 위해서 한국수어 말뭉치는 지역별 균형성과 대표성을 지닐 정도의 유의미한 규모로 확보되어야 한다. 축적된 한국수어 말뭉치 자료는 수어의 언어적 특성 및 다양한 변이 관찰을 통해 수어로 하여금 언어적 지위를 확고히 할 수 있게 하고, 농인의 수어 사용에 대한 언어권을 보장하게 될 것이다.

이에 한국수어 말뭉치 구축은 현재까지 진행되어온 것과 같은 수어 자료 수집, 언어적 특성 분석을 위한 주석 달기, 수어에 대한 대중의 관심 고조와 연구자의 연구 활성화를 도모하기 위한 공개라는 3가지 핵심과제를 중심으로 추진함으로써 균형성과 대표성을 갖춘 한국수어 말뭉치로서의 가치를 보장받게 될 것이다.



[그림 IV-33] 한국수어 말뭉치의 비전과 추진 과제

(2) 과제별 추진 방향 및 성과목표

(가) 한국수어 수집

2020년 기준으로 경기지역 언어제공자 5명을 포함한 서울(60명), 광주지역 5명을 포함한 전남(46명), 경남(26명), 각 4쌍의 전북과 강원(16명)으로 총 5개 지역의 수어 자료가 수집 완료되었으며 향후 대전·충남, 충북(세종 포함), 대구·경북, 부산, 제주, 인천 8곳과 대표성을 위해 추가 수집해야 하는 전북과 강원 2곳, 소수의 인원만 확보된 경기 등 모두 11곳의 언어자료가 수집되어야 한다. 지역 농아인협회의 협조를 통해 지역의 전체 농인 수에 비례한 언어제공자를 모집하여 지역의 수어를 대표할 수 있도록 모집한다.

수집 영상의 품질을 높이기 위해서는 매년 새로운 팀으로 구성된 촬영·편집 팀에 대한 사전교육을 실시하기보다는 촬영 또는 편집 중에 발생하는 오차를 줄이고 일관성을 확보할 수 있도록 전문 촬영팀 도입을 고려해 볼 필요가 있다.

또한 지역의 수어 자료를 수집한 이듬해부터 말뭉치 주석을 시작할 때 주석 작업자가 동일 지역 출생이 아닐 시 수어의 변이 확인에 어려움이 있으므로 수집 과정에서 번역을 포함하면 주석 작성의 효율성을 높일 수 있을 것이다.

추진방향

- 전국 시·도 농아인협회와 협력을 통한 지역 균형 확보: 언어제공자 모집 환경 개선
- 촬영 및 편집의 일관성 확보를 위한 환경 개선: 전문 촬영팀 도입
- 지역 수어 언어제공자의 안정적 모집: 지역 수어 언어제공자 인력풀 구축
- 촬영된 지역 수어 영상 번역 완료: 상시 수어 번역 체제 구축

성과목표

추진 과제	성과목표	'20년 목표치	'25년 목표치	목표치 산출근거
전국 시·도 농아인 협회와 협력관계 체결	원시말뭉치 수집 지역 균형 확보 (인구 대비 모집)	5개 지역	11개 지역	<ul style="list-style-type: none"> • 대전·충남·충북, 대구·경북, 부산, 제주, 인천·경기 (2015년도 경기만 5명 수집), 전북 및 강원(추가 수집) ※ 수어 수집이 이루어지지 않은 지역 원시 말뭉치 수집
전문 촬영팀 확보	촬영조건 및 촬영 시스템 개선	촬영 211시간 편집 180시간 영상파일 2,879개 (945쌍)	촬영 350시간 편집 250시간	<ul style="list-style-type: none"> • 1개 지역 10쌍 언어제공자로부터 350시간 이상의 촬영 기준 • 매년 2개 지역의 언어 자료 수집 1개 편집 25시간 × 2개 지역 × 5년 = 250시간 ※ 동일한 조건의 촬영 및 편집 기술 확보를 위한 전문 촬영팀 구성으로 작업의 효율성 강화
지역 수어 언어제공자 인력풀 구축	지역 수어 언어제공자 모집 및 촬영된	언어제공 자 74쌍 (148명)	언어제공 자 100쌍 (200명)	<ul style="list-style-type: none"> • 충청권(33쌍) + 경상권(33쌍) + 제주(11쌍) + 강원과 전북 추가 수집(14쌍) + 인천과 경기 각 9쌍 ※ 지역 간 수어 변이 분석을 위한 최소 10쌍 이상의 언어제공자 수 확보
수어 영상 번역	촬영된 수어 영상 번역 완료	3,176분 (2019년)	19,920분	<ul style="list-style-type: none"> • 2,400분(40시간) × 5년 = 12,000분 • 미완료 영상 번역 실시: 7,920분(132시간) ※ 언어 자료 수집 지역의 인력 활용한 번역 실시

(나) 한국수어 자료 주석 달기

주석의 일관성을 유지하고, 기본주석과 의미주석의 내실화를 위해 전문 주석팀의 상시 운영이 필요하다. 상시 운영되는 주석팀에서는 동일 주석자가 지속적으로 주석 작업을 함으로써 주석의 질 개선과 신뢰성을 높일 수 있을 것이다. 현재까지 수집된 방대한 양의 수어자료는 전문 주석팀의 숙련된 주석 작업이 없는 주석이 되지 못한 채 방기 될 가능성도 있다. 4,200분(70시간) 분량의 영상이 수집된 서울의 수어자료는 현재까지 1,213분(20시간 13분)분량의 주석 달기가 완료되었는데, 이는 28.6% 수준의 주석량이다. 현재 지역별로 11쌍의 언어제공자로부터 약 2,000분 분량의 영상을 수집하고 있으며, 이 중 5~10%의 영상 자료에 대해 주석 달기를 한다면 100-200분의 분량이 된다. 2020년에 기본주석과 의미주석을 완료한 경남지역을 제외한 14개 지역에 대해 100분-200분 분량의 자료에 주석 달기를 하면 향후 총 1,400분-2,800분가량의 수어 영상 자료에 주석을 해야 한다. 매년 400분 분량의 주석달기를 한다고 가정하면 5년 후 2,000분 분량의 수어 영상 자료 주석이 완료될 것이다.

추진방향

- 기본주석 및 의미주석의 신뢰성 확보를 위한 내실화
- 주석 전문 인력 확보를 통한 주석의 일관성과 신뢰도 제고
- 주석 작업 진행에 따라 변경되는 주석 규약 내용 정비
- 「한국수어자료통합지원시스템」 안정화를 통한 수어자료 활용의 효율성 강화

성과목표

추진 과제	성과목표	'20년	'25년	목표치 산출근거
기본주석 및 의미주석 내실화	기본주석 작업 확대	1,294분 (21시간 34분)	2,000분 (33시간 20분)	• 400분 × 5년 = 2,000분 ※ 기본주석이 미완료된 영상에 대한 주석 작업 분량
	의미주석 작업 확대	135분 (2시간 15분)	2,000분 (33시간 20분)	• 400분 × 5년 = 2,000분 ※ 의미주석이 미완료된 영상에 대한 주석 작업 분량
주석 전문 인력 확보	주석전문가 양성 연수회 개최	2회 교육	4회	• 분기별 1회×4분기 = 4회 ※ 주석 전문가 저변 확대를 위한 연수회 개최 수
	주석전문가 고용	-	2명	• 주석전문가 1팀×2명 = 2명 ※ 국립국어원 전일제 주석 전문가 고용 인원
주석 규약 정교화	사례연구를 통한 주석 규약 보완	규약 정비	규약 정비	• 주석지침 규약 정비의 갯수 ※ 주석 작업 및 사례연구를 통해 확인되는 추가 내용 을 주석 규약에 보완, 정 비
「한국수어 자료통합지 원 시스템」 안정화	수어자료통합 시스템 기능 개선	154,172 토큰 (2020.9.)	200,000 토큰	• 40,000토큰 × 5년 = 200,000 토큰 ※ 「한국수어자료통합지원시스 템」 탑재 토큰 수 확대 ※ 주석 및 검증 단계 안정화 등

(다) 한국수어 말뭉치 공개

국립 기관에 의해 구축된 한국수어 말뭉치는 한국수어 연구자와 한국수어 사용자에게 공개가 이루어져야 한다. 주석이 28% 수준까지 완료된 서울지역을 시작으로 말뭉치 공개에 대한 단계적인 진행 절차가 필요하다. 이를 위해서 먼저 1차 연도에는 공개 범위, 시기, 방법을 모색하여 공개를 위한 시스템 개발을 계획하는 단계로서 일차적인 공개 범위에 해당하는 지역의 주석을 정비한다. 2차 연도에는 공개 시스템을 구축한다. 독일과 영국의 경우 공개 페이지를 달리하여 일반인용과 연구자용이 구분되어 있으며 회원 가입을 필요로 하는 것과 필요로 하지 않는 것으로 구분되어 있다. 현재 한국어 말뭉치의 경우 ‘모두의 말뭉치’를 통해 회원가입 후 필요한 말뭉치를 신청한 사람에게 공개하는 형태로 제공되고 있다. 한국수어 말뭉치 공개 과업은 한국수어 말뭉치를 어떤 형태로 공개할 것인지에 대한 계획을 수립하고 정비하여 공개 시스템에 탑재하는 것까지가 주요 과업이 될 것이다.

추진방향

- 공개 방법 및 공개 범위 확정
- 온라인 공개 시스템 구축
- 공개 영상 및 주석 자료 정비

성과목표

추진 과제	성과목표	'21년	'25년	목표치 산출근거
공개 범위 및 방법 확정	한국수어 말뭉치 공개에 따른 세부 실행계획 수립	말뭉치 공개의 필요성과 추진 전략 수립	말뭉치 공개 완료	<ul style="list-style-type: none"> • 공개 대상 영상 과제 개수 • 공개 자료의 분량 • 주석 처리된 영상 분량
온라인 공개 시스템 구축	말뭉치 보급 및 활용을 위한 공개 온라인 시스템 개발	-	수어 자료 공개 시스템 구축 및 운영	<ul style="list-style-type: none"> • 한국수어 말뭉치 온라인 공개 시스템
공개 영상 및 주석 자료 정비	말뭉치 영상 및 주석 공개 자료 완성	공개 대상 주석 정비	말뭉치 공개 영상 및 주석 자료 완비	<ul style="list-style-type: none"> • 공개한 한국수어 말뭉치 영상 개수 • 주석 자료 완비 정도

(2) 세부 추진과제별 추진 전략

과제 1

한국수어 수집

- ❖ 전국 시도 농아인협회와의 협력 관계 체결
- ❖ 전문 촬영팀 확보
- ❖ 지역 언어제공자의 인력풀 구축
- ❖ 각 지역의 원시 수어 말뭉치 영상 번역 완료

< 추진 필요성 >

- ☐ 한국수어 연구의 토대가 되는 원시 수어 말뭉치 수집에 지역별 균형이 이루어지도록 전국 각 시·도의 최소 10쌍 이상의 해당 지역의 토박이 농인으로부터 원시 자료가 요구되나, 아직 미흡한 실정임

< '20년 전국 시·도 수어 말뭉치 언어제공자 현황 >

구 분	수도권	강원도	충청도	전라북도	전라남도	경상북도	경상남도	제주도	계
언어제공자수	60	8	0	8	46	0	26	0	148
백분율 (%)	40.5	5.4	0	5.4	31.1	0	17.6	0	100

- ☐ 수집 지역 촬영 인력 확보와 촬영 시스템의 동일성 확보에 어려움 있음
- ☐ 수집과 번역을 분리하여 작업한 결과 시간이 지남에 따라 해석의 정확성 확보에 어려움 있음. 따라서 지역의 원시 말뭉치 수집과 동시에 해당 지역의 수어 사용자가 번역 작업을 완료하는 것이 효율적임.

1-1

전국 시·도 농아인협회와 협력관계 체결

기본방향

- ◆ 언어제공자 제공을 비롯한 시·도 농아인협회의 원활한 협조 기반 구축
- ◆ 지역 원시 말뭉치 수집 관련 정보 제공 등의 긴밀한 상호 협력

□ 원시 말뭉치 수집 지역 선정 시 해당 지역 농아인협회와 사전 협의

- 원시 말뭉치 수집의 전반적인 절차와 협조 사항 논의
- 언어제공자의 접근성을 고려한 촬영 장소 및 인력에 관한 논의

□ 해당 지역 한국수어 말뭉치 사업 소개 워크숍 개최

- 해당 지역 농인 대상 한국수어 말뭉치 사업의 목적과 방법, 향후 계획 홍보
- 한국수어 말뭉치 사업 소개 워크숍을 통해 원시 말뭉치 수집 작업에 농인들의 활발한 참여 효과

□ 언어제공자 모집 및 선정 과정에 홍보와 참여 협조

- 지역 여건을 고려하여 해당 지역 농아인협회와 원시 수어 말뭉치 언어제공자 모집 홍보 및 참여 요청
- 수어 말뭉치 수집 지침에 있는 언어제공자 조건에 맞도록 모집과 선정 협의

□ 해당 지역 특성에 맞는 수어 유도자료 교체 협의

- 해당 지역의 특별한 수어 변이형 및 수어 특성에 관한 정보 공유
- 해당 지역에 이슈가 되었던 특별한 사건과 농아인협회 행사에 대한 정보 공유

1-2

전문 촬영팀 확보

기본방향

- ◆ 수어 말뭉치 수집에 최적화된 촬영 환경 구축
- ◆ 지역 간 편차가 없는 촬영의 질 확보

□ 수어 말뭉치 촬영에 적합한 장소 확보

- 카메라 3대, 유도과제 제시용 모니터와 컴퓨터 등을 설치할 수 있는 12평 이상의 공간
- 언어제공자의 교통편의 접근성을 고려한 촬영 장소 지정
- 전기, 냉난방, 수도, 엘리베이터 등의 시설 설비 점검

□ 장기 전문 촬영·편집팀 운영 시스템 구축

- 지역 간 동질의 수어 영상 확보를 위해 장기 전문 촬영·편집팀 운영
- 촬영·편집 작업에 한국수어 사용 가능한 농인 채용

□ 촬영·편집 장비의 기술 발전에 맞춘 지침 수정 및 보완

- 시시각각 발전하는 최신 촬영 장비 및 편집 기술 확보
- 수어 말뭉치 영상 자료 수집에 효율적인 촬영·편집 지침의 지속적인 개선

□ 수어 말뭉치 촬영·편집 영상 검사 접수 강화

- 말뭉치 촬영 현장에서 촬영 직후 동영상 상태 검사
- 촬영 영상의 편집 및 접수 절차 세분화 및 전문화 추진

1-3

지역 수어 언어제공자 인력풀 구축

기본방향

- ◆ 지역 원시 수어 수집을 위한 최적의 언어제공자 확보
- ◆ 수어 언어제공자의 안정적 공급

□ 지역별 최소 10쌍 이상의 언어제공자 확보

- 지역별 수어 특성을 제시할 수 있는 표본의 대표성 확보를 위해 최소 10쌍 이상의 언어제공자 선정
- 지역의 농인 수에 비례하여 언어제공자 수를 조정할 수 있음

□ 언어제공자의 제반여건을 수용한 촬영 일정 조정

- 언어제공자의 조건과 더불어 연령 및 성별 균형을 고려하여 선정
- 언어제공자의 직업, 종교, 문화 활동을 고려한 탄력적인 촬영 일정으로 선정된 언어제공자의 언어 자료 확보
- 예비 언어제공자 쌍 대기

1-4

지역 원시 수어 말뚱치 영상 번역

기본방향

- ◆ 원시 말뚱치 자료의 내실화를 위한 수어 말뚱치 영상 번역 완료
- ◆ 총 19,920분의 수어 영상 번역

- ☐ 내실화 있는 번역 자료 확보로 주석 작업에서 지역 수어 인지 및 변이 확인을 위해 참고할 수 있는 자료 확보
 - 기본 주석 작업자의 주석 시 번역을 참고함으로써 주석의 오류 예방
 - 지역 특성이 반영된 수어의 정확한 해석을 보장하기 위해 해당 지역 수어 사용 인력(농인, 통역사, 교원 등) 번역 실시
- ☐ 원시 말뚱치 수집과 번역 작업으로 해당 지역 농인의 말뚱치에 대한 관심과 역할 확대
 - 말뚱치 수집 지역에서 촬영·편집 및 번역까지 완료 추진
- ☐ 많은 사람이 이용 가능한 수어 영상 자료로 확보
 - 주석 없는 말뚱치를 농인의 문화 체험 또는 청인의 통역 연습 자료 또는 교원의 예제 활용 자료 확보

과제 2

한국수어 자료 주석 달기

- ❖ 기본주석 및 의미주석 내실화
- ❖ 주석 전문 인력 확보
- ❖ 주석 규약 정교화
- ❖ 「한국수어자료통합지원시스템」 안정화

< 추진 필요성 >

□ 수집된 수어 말뭉치 영상의 기본주석 및 의미주석 미완료 자료 적체

- 2020년까지 말뭉치 수집 영상의 약 10,800 분(180시간) 중 기본주석 약 1,294 분(약 21시간 34분)(12.0%) 의미주석 약 200 분(3시간 20분)(1.9%)만이 완료된 실정임
- 수어의 특성으로 인해 말뭉치 수집에 주석이 필수로 요구됨
(수어의 화자별, 지역별 변이형과 생산적 수어 해석 및 연구에 주석 반드시 필요)

< '20년 주석 완료 현황 >

구분	기본주석	의미주석	전체
영상 분량(분)	1,294	200	10,800
완료율(%)	12.0	1.9	100.0

□ 주석 규칙 정비하여 수정·보완된 주석 지침 확정 필요

- 주석 작업에 혼란이 없도록 합리적이고 이해하기 쉬운 주석 규칙 필요
- 주석자 간의 상이한 해석으로 주석의 신뢰도 미흡

□ 전문성 있는 주석 인력 확보의 어려움으로 인한 주석의 질 저하 우려

- 주석 규칙 및 컴퓨터 주석 프로그램 교육 확대 필요
- 주석자의 잦은 교체로 인하여 주석의 일관성 확보에 어려움 있음

□ 「한국수어자료통합지원시스템」의 불안정으로 인한 주석 작업의 불편 초래

- 주석 프로그램을 이용한 검색, 탑재, 수정 등의 원활한 작업 진행에 차질 발생
- 「한국수어자료통합지원시스템」의 관리와 효율성 강화 요구됨

2-1

기본주석 및 의미주석 내실화

기본방향

- ◆ 기본주석 및 의미주석 미완료 영상에 주석 달기
- ◆ 수어 타입에 언어학 정보를 추가하여 주석 보강

□ 수어 말뭉치 영상에 기본주석 및 의미주석 달기

- 수어 말뭉치 영상 자료의 유용성을 위한 주석 달기
- 기본주석 및 의미주석의 중요성 고려한 내실화

□ 체계화된 수어 언어학 정보 주석 강화

- 수어 음운론, 형태론, 통사론, 화용론 관련 정보를 포함한 주석으로 말뭉치 기반 수어 연구 실현
- 수어 언어학의 전문성을 내포한 주석 품질 향상

2-2

전문 인력 확보

기본방향

- ◆ 연수회 개최를 통한 주석 전문가 양성
- ◆ 주석 전문 인력 채용을 통한 주석 전문가 확보

□ 주석 규칙과 컴퓨터 주석 프로그램 교육의 확대

- 주석자의 혼란 예방과 주석의 정확성을 확보하기 위한 주석 교육 강화
- 주석 소프트웨어 활용 능력 향상을 위한 교육 프로그램 개발·운영

□ 주석의 일관성 확보를 위한 전일제 전문 인력 채용

- 시간이 많이 소요되고 복잡한 주석 작업에 주석 전문 인력풀 확보
- 수어 언어학 지식과 주석 전문성을 갖춘 전일제 주석 인력 채용

2-3

주석 규약 정교화

기본방향

- ◆ 수정·보완된 주석 규칙에 따른 주석의 신뢰성 확보
- ◆ 합리적이고 이해하기 쉽도록 주석 규약 정비

□ 일반적이고 합리적인 주석 지침 확정

- 누구나 이해할 수 있고 사용하기 간편한 주석 규칙 개선
- 주석 작업에 혼란이 없도록 주석 규칙의 세분화 및 명료화 추진

□ 개선된 주석 규칙에 따른 주석 작업으로 신뢰성 확보

- 주석 규칙의 보강으로 주석자들 간의 혼선과 주석 작업의 어려움 감소
- 합리적인 주석 규칙 적용으로 주석의 신뢰도 상승

2-4

「한국수어자료통합지원시스템」 안정화

기본방향

- ◆ 「한국수어자료통합지원시스템」 기능 개선
- ◆ 「한국수어자료통합지원시스템」 활용 효율성 강화

□ 「한국수어자료통합지원시스템」의 검색, 탑재, 수정 기능 개선

- 주석자의 수어 자료 통합 시스템 사용 편이를 위한 프로그램 고도화

□ 「한국수어자료통합지원시스템」 관리 강화

- 프로그램 관리자와 긴밀한 협력으로 업무 효율화 추진
- 수어 자료 통합 지원 시스템 유지 및 보수 관리 강화

과제 3

한국수어 말뭉치 자료 공개

- ❖ 한국수어 말뭉치 공개 범위 및 방법 확정
- ❖ 한국수어 말뭉치 온라인 자료 및 공개 시스템 구축
- ❖ 공개 영상 및 주석 자료 정비

< 추진 필요성 >

- ☐ 2015년부터 추진해온 한국수어 말뭉치 구축 사업의 일부 공개 요구됨
 - 수어 말뭉치 구축의 의미에 맞도록 말뭉치 공개 필요성 공론화됨
- ☐ 수어 말뭉치 공개를 위한 공개 시스템 개발 요구됨
 - 구축된 수어 말뭉치 공개를 위한 온라인 포털 개발 시급
- ☐ 수어 사용자 및 수어를 배우는 학생, 수어 교사, 수어 연구자들이 사용 가능한 자료 확보의 어려움
 - 수어에 관심이 있는 일반인과 수어 연구자들이 활용 가능한 수어 자료 미흡한 실정임
 - 과학적인 수어 연구의 토대 마련을 위한 말뭉치 자료 공개 필요

3-1

한국수어 말뭉치 자료 공개 범위 및 방법 확정

기본방향

- ◆ 수어 말뭉치 공개 범위 확정
- ◆ 수어 말뭉치 공개 방법 수립
- ◆ 수어 말뭉치 공개 준비팀 운영

□ 말뭉치 공개 방법에 대한 전문가 집단의 자문

- 말뭉치 전문가 집단의 말뭉치 공개에 대한 의견 수렴
- 해외 수어 말뭉치 공개 사례 탐색

□ 말뭉치 공개 범위와 구성에 대한 구체적인 계획 수립

- 수어 공개 영상의 과제 및 분량, 주석의 범위 결정
- 말뭉치 공개 대상과 그에 따른 공개 자료의 구성 결정

□ 말뭉치 공개 전담 TF팀 운영

- 수어 말뭉치 공개의 치밀하고 효율적인 준비를 위한 업무 전담팀 구성
- 성공적인 수어 말뭉치 공개를 위해 TF팀 구성에 분야별 전문가 위촉

3-2

한국수어 말뭉치 자료 온라인 공개 시스템 구축

기본방향

- ◆ 전문가를 위한 한국수어 말뭉치 자료 온라인 공개 시스템 구축
- ◆ 일반인들의 접근이 용이한 수어 말뭉치 온라인 공개 시스템 개발

□ 수어 말뭉치 온라인 공개 포털 구축

- 사용이 간편하고 디자인이 복잡하지 않은 수어 말뭉치 공개 사이트 개발 논의
- 수어 자료의 검색, 영상보기, 주석 내려받기 등의 유용한 기능이 탑재된 말뭉치 시스템 구축 방안 강구
- 수어 자료 통합 지원 시스템 구축 경험을 바탕으로 시스템 개발 전문가와 긴밀한 협력 및 지원 확대

3-3

공개 영상 및 주석 자료 정비

기본방향

- ◆ 수어 말뭉치 공개 지침과 사용 방법 확정
- ◆ 수어 말뭉치 공개 범위 확정에 따른 영상 및 주석 자료 정비
- ◆ 수어 말뭉치 시범 공개를 통한 공개의 안정성 확보

□ 수어 말뭉치 공개 지침과 사용 방법 개발 및 안내

- 말뭉치 공개 지침과 포털 이용 방법 확정
- 말뭉치 구축 과정 및 말뭉치 공개 지침, 사용 방법에 대한 안내 문서 및 수어 영상 제작

□ 수어 말뭉치 공개 영상 및 주석 자료 보완 점검

- 수어 말뭉치 공개 영상 자료의 세밀한 분석 실시
- 수어 말뭉치 공개 자료의 오류를 방지하기 위한 철저한 검사 검수 시스템 마련

□ 수어 말뭉치 시범 공개를 통한 검증 실시

- 한국수어 말뭉치의 본격적인 공개에 앞서 전문가 대상 시범 공개 실시
- 한국수어 말뭉치 시범 공개의 검증을 통한 성공적인 수어 말뭉치 공개 실현

4) 연도별 추진 계획

연도별 추진 계획은 <표 IV-17> 과 같다.

5) 한국수어 말뭉치 구축 소요 예산

한국수어 말뭉치의 3가지 핵심과제를 추진하기 위해 소요될 예산은 <표 IV-18>과 같다. 사업의 효율적인 운영을 위하여 각 과제 별 별도의 운영을 고려한 사업별 운영 예산이다. 한국수어 말뭉치 수업은 전문 촬영팀에 의해 연간 2개 지역의 언어자료 수집 및 번역, 주석은 연간 400분 분량의 주석 및 검증, 공개를 위한 계획 및 실행을 위한 주석 정비 및 시스템 구축에 따른 예산이며, 공개 관련 예산은 3차 연도 이후에는 유지비용으로 산정되었다.

〈표 IV-17〉 한국수어 말뭉치 구축 연도별 추진 계획

추진과제		성과목표	2021	2022	2023	2024	2025	2026-
1. 한국수어 수집	1-1. 전국 시도 농아 인협회와 협력관 계 체결	1) 지역 균형 확보한 원시 말뭉치 수집 (연간 2곳)	⇒	⇒	⇒	⇒	⇒	
	1-2. 전문 촬영팀 확보	1) 촬영 조건 및 촬영 시스템 개선 -중앙 파견식 전문 촬영팀	⇒	⇒	⇒	⇒	⇒	
	1-3. 지역 수어 언어 제공자 인력풀 구축	1) 지역 언어제공자 모집	⇒	⇒	⇒	⇒	⇒	
	1-4. 수어 영상 번역	1) 촬영된 수어 영상 번역 완료	⇒	⇒	⇒	⇒	⇒	
2. 한국수어 자료 주석 달기	2-1. 기본주석 및 의 미주석 내실화	1) 기본 주석 작업 확대	⇒	⇒	⇒	⇒	⇒	⇒
		2) 의미 주석 작업 확대	⇒	⇒	⇒	⇒	⇒	⇒
	2-2. 주석 전문 인력 확보	1) 주석전문가 양성 연수회	⇒	⇒	⇒	⇒	⇒	
		2) 주석전문가 고용	⇒	⇒	⇒	⇒	⇒	⇒
	2-3. 주석 규약 정교화	1) 사례연구를 통한 주석 규약 보완	⇒	⇒	⇒	⇒	⇒	⇒
3. 한국수어	2-4. 한국수어자료통 합지원시스템 안 정화	1) 한국수어자료통합지원시 스템 기능 개선	⇒	⇒	⇒	⇒	⇒	⇒
	3-1. 공개 범위 및 방법 확정	1) 공개 세부 실행계획 수 립	⇒					

말뭉치 공개	3-2. 온라인 공개 시스템 구축	1) 온라인 시스템 개발		⇒ ⇒				
	3.3 공개 영상 및 주석 자료 정비	2) 영상 및 주석 공개 자료 완성	⇒	⇒ ⇒	⇒ ⇒	⇒ ⇒	⇒ ⇒	⇒ ⇒

〈표 IV-18〉 한국수어 말뭉치 구축 소요 예산

구분	사업명	세목	소요예산	산출 내역		
1차 연도 (2021)	한국수어 말뭉치 수집 (164,452,220)	인건비	35,352,220	책임연구 1인, 공동연구 2인, 연구보조 1인, 전문촬영팀 5인 (8개월)		
		운영비	129,100,000	〈전문가 활용비〉 8,800,000 × 2개 지역 17,600,000		
				〈촬영·편집비〉 38,000,000 × 2개 지역 76,000,000		
				〈번역비〉		
				수집 자료 번역주석 수당 100,000원/10분 × 2,400분 24,000,000		
				〈인쇄비〉 2,500,000원 × 1회 2,500,000		
				〈운영비〉 4,500,000원 × 2개 지역 9,000,000		
				〈전문가 활용 및 촬영·편집 1개 지역 소요 경비〉		
				* 지역당 11쌍(예비 1쌍 포함), 35시간 이상 촬영 기준		
				전문가 활용	언어제공자 수당 300,000원 × 11쌍 × 2인 6,600,000	
사회자 수당 200,000원 × 11회 2,200,000						
소계 8,800,000						
촬영· 편집	장소 임대료 500,000원 × 10일 5,000,000					
	영상 촬영 300,000원 × 10일 × 3인 9,000,000					

					영상 편집 촬영 장비(카메라, 조명, 삼각대, 컴퓨터 등) 렌더링 기타운영비 소계	300,000원 × 10일 × 2인 1,500,000원 × 10일 200,000원 × 5일 × 1인 200,000원 × 10회	6,000,000 15,000,000 1,000,000 2,000,000 38,000,000
					운영 회의비 출장 교통비 기타 운영비 소계	20,000원 × 10회 × 10인 50,000원 × 10일 × 2인 150,000원 × 10회	2,000,000 1,000,000 1,500,000 4,500,000
					인건비	50,087,630	책임연구 1인, 공동연구 3인, 연구보조 2인, 전문주석팀 (8개월)
					주석 (129,087,630) 운영비	79,000,000	<주석비> (70,000,000)
							주석 작업자 수당 120,000원/1분 × 400분 48,000,000
							주석 검증자 수당 50,000원/1분 × 400분 20,000,000
							교육비 500,000원 × 4회 2,000,000
							<운영비> (9,000,000)
							인쇄비 2,000,000원 × 1회 2,000,000
							회의비(통역비 포함) 20,000원 × 20회 × 10인 4,000,000
							기타 운영비 200,000원 × 15회 3,000,000

	한국수어 말뭉치 공개 -시스템 설계 (100,087,630)	인건비	50,087,630	책임연구 1인, 공동연구 3인, 연구보조 2인 시스템개발팀 (8개월)			
		운영비	50,000,000	<시스템 설계 의뢰>			(5,000,000)
				개발 설계비	5,000,000원 × 1회		5,000,000
				<주석 정비>			(36,000,000)
				주석 작업자 수당	60,000원/1분 × 400분		24,000,000
				주석 검증자 수당	30,000원/1분 × 400분		12,000,000
				<운영비>			(9,000,000)
				인쇄비	2,000,000원 × 1회		2,000,000
				회의비(통역비 포함)	20,000원 × 20회 × 10인		4,000,000
				기타 운영비	200,000원 × 15회		3,000,000
2차 연도 (2022)	한국수어 말뭉치 수집 (164,452,220)	인건비	35,352,220	책임연구 1인, 공동연구 2인, 연구보조 1인, 전문촬영팀 5인 (8개월)			
		운영비	129,100,000	<전문가 활용비>			8,800,000 × 2개 지역
				<촬영·편집비>			38,000,000 × 2개 지역
				<번역비>			
				수집 자료 번역주석 수당	100,000원/10분 × 2,400분		24,000,000
				<인쇄비>			2,500,000원 × 1회
				<운영비>			4,500,000원 × 2개 지역
				<전문가 활용 및 촬영·편집 1개 지역 소요 경비>			
				* 지역당 11쌍(예비 1쌍 포함), 35시간 이상 촬영 기준			
				전문가	언어제공자 수당	300,000원 × 11쌍 × 2인	6,600,000
				활용	사회자 수당	200,000원 × 11회	2,200,000

					소계		8,800,000
				촬영 . 편집	장소 임대료	500,000원 × 10일	5,000,000
					영상 촬영	300,000원 × 10일 × 3인	9,000,000
					영상 편집	300,000원 × 10일 × 2인	6,000,000
					촬영 장비(카메라, 조명, 삼각대, 컴퓨터 등)	1,500,000원 × 10일	15,000,000
					렌더링	200,000원 × 5일 × 1인	1,000,000
					기타운영비	200,000원 × 10회	2,000,000
					소계		38,000,000
				운영	회의비	20,000원 × 10회 × 10인	2,000,000
					출장 교통비	50,000원 × 10일 × 2인	1,000,000
					기타 운영비	150,000원 × 10회	1,500,000
					소계		4,500,000
	주석 (129,087,630)	인건비	50,087,630	책임연구 1인, 공동연구 3인, 연구보조 2인, 전문주석팀 (8개월)			
		운영비	79,000,000	<주석비> (70,000,000)			
				주석 작업자 수당	120,000원/1분 × 400분	48,000,000	
				주석 검증자 수당	50,000원/1분 × 400분	20,000,000	
	교육비			500,000원 × 4회	2,000,000		
					<운영비> (9,000,000)		

	한국수어 말뭉치 공개 -시스템 구축 (141,087,630)	인건비	50,087,630	인쇄비	2,000,000원 × 1회	2,000,000
				회의비(통역비 포함)	20,000원 × 20회 × 10인	4,000,000
				기타 운영비	200,000원 × 15회	3,000,000
		운영비	91,000,000	책임연구 1인, 공동연구 3인, 연구보조 2인 (8개월)		
				<시스템 구축> (30,000,000)		
				구축비	30,000,000원 × 1회	30,000,000
				(서버 및 사용자 만족 조사 포함)		
				<주석 정비> (52,000,000)		
				주석 작업자 수당	80,000원/1분 × 400분	32,000,000
				주석 검증자 수당	50,000원/1분 × 400분	20,000,000
				<운영비> (9,000,000)		
				인쇄비	2,000,000원 × 1회	2,000,000
				회의비(통역비 포함)	20,000원 × 20회 × 10인	4,000,000
				자문비	200,000원 × 3회 × 3인	1,200,000
				기타 운영비	200,000원 × 15회	3,000,000
3차 연도 이후	한국수어 말뭉치 수집 (155,452,220)	인건비	35,352,220	책임연구 1인, 공동연구 2인, 연구보조 1인, 전문촬영팀 5인 (8개월)		
		운영비	120,100,000	<전문가 활용비>	8,800,000 × 2개 지역	17,600,000
				<촬영·편집비>	38,000,000 × 2개 지역	76,000,000
				<번역비>		

				수집 자료 번역주석 수당 100,000원/10분 × 1,500분 15,000,000																																													
				<인쇄비> 2,500,000원 × 1회 2,500,000																																													
				<운영비> 4,500,000원 × 2개 지역 9,000,000																																													
				<div><전문가 활용 및 촬영·편집 1개 지역 소요 경비></div> <div>* 지역당 11쌍(예비 1쌍 포함), 35시간 이상 촬영 기준</div> <table><tr><td rowspan="3">전문가 활용</td><td>언어제공자 수당</td><td>300,000원 × 11쌍 × 2인</td><td>6,600,000</td></tr><tr><td>사회자 수당</td><td>200,000원 × 11회</td><td>2,200,000</td></tr><tr><td>소계</td><td></td><td>8,800,000</td></tr><tr><td rowspan="7">촬영· 편집</td><td>장소 임대료</td><td>500,000원 × 10일</td><td>5,000,000</td></tr><tr><td>영상 촬영</td><td>300,000원 × 10일 × 3인</td><td>9,000,000</td></tr><tr><td>영상 편집</td><td>300,000원 × 10일 × 2인</td><td>6,000,000</td></tr><tr><td>촬영 장비(카메라, 조명, 삼각대, 컴퓨터 등)</td><td>1,500,000원 × 10일</td><td>15,000,000</td></tr><tr><td>렌더링</td><td>200,000원 × 5일 × 1인</td><td>1,000,000</td></tr><tr><td>기타운영비</td><td>200,000원 × 10회</td><td>2,000,000</td></tr><tr><td>소계</td><td></td><td>38,000,000</td></tr><tr><td rowspan="4">운영</td><td>회의비</td><td>20,000원 × 10회 × 10인</td><td>2,000,000</td></tr><tr><td>출장 교통비</td><td>50,000원 × 10일 × 2인</td><td>1,000,000</td></tr><tr><td>기타 운영비</td><td>150,000원 × 10회</td><td>1,500,000</td></tr><tr><td>소계</td><td></td><td>4,500,000</td></tr></table>	전문가 활용	언어제공자 수당	300,000원 × 11쌍 × 2인	6,600,000	사회자 수당	200,000원 × 11회	2,200,000	소계		8,800,000	촬영· 편집	장소 임대료	500,000원 × 10일	5,000,000	영상 촬영	300,000원 × 10일 × 3인	9,000,000	영상 편집	300,000원 × 10일 × 2인	6,000,000	촬영 장비(카메라, 조명, 삼각대, 컴퓨터 등)	1,500,000원 × 10일	15,000,000	렌더링	200,000원 × 5일 × 1인	1,000,000	기타운영비	200,000원 × 10회	2,000,000	소계		38,000,000	운영	회의비	20,000원 × 10회 × 10인	2,000,000	출장 교통비	50,000원 × 10일 × 2인	1,000,000	기타 운영비	150,000원 × 10회	1,500,000	소계		4,500,000
전문가 활용	언어제공자 수당	300,000원 × 11쌍 × 2인	6,600,000																																														
	사회자 수당	200,000원 × 11회	2,200,000																																														
	소계		8,800,000																																														
촬영· 편집	장소 임대료	500,000원 × 10일	5,000,000																																														
	영상 촬영	300,000원 × 10일 × 3인	9,000,000																																														
	영상 편집	300,000원 × 10일 × 2인	6,000,000																																														
	촬영 장비(카메라, 조명, 삼각대, 컴퓨터 등)	1,500,000원 × 10일	15,000,000																																														
	렌더링	200,000원 × 5일 × 1인	1,000,000																																														
	기타운영비	200,000원 × 10회	2,000,000																																														
	소계		38,000,000																																														
운영	회의비	20,000원 × 10회 × 10인	2,000,000																																														
	출장 교통비	50,000원 × 10일 × 2인	1,000,000																																														
	기타 운영비	150,000원 × 10회	1,500,000																																														
	소계		4,500,000																																														

	주석 (129,087,630)	인건비	50,087,630	책임연구 1인, 공동연구 3인, 연구보조 2인 (8개월)		
		운영비	69,000,000	<주석비> (60,000,000) 주석 작업자 수당 100,000원/1분 × 400분 40,000,000 주석 검증자 수당 30,000원/1분 × 400분 20,000,000 <운영비> (9,000,000) 인쇄비 2,000,000원 × 1회 2,000,000 회의비(통역비 포함) 20,000원 × 20회 × 10인 4,000,000 기타 운영비 200,000원 × 15회 3,000,000		
	한국수어 말뭉치 공개 -시스템 유지 (6,000,000)	시스템 유지비	84,000,000	시스템 유지비 500,000원 × 12월 6,000,000 (기존 시스템 활용시 별도의 비용 발생 없음) <운영비> (78,000,000) 관리 인력 2,500,000원 × 12월 30,000,000 기타 운영비 200,000원 × 12회 48,000,000		

V. 요약 및 제언



1. 요약
2. 활용 방안 및 제언

V. 요약 및 제언

1. 요약

한국수어 말뭉치 구축 및 연구 사업은 말뭉치 수집과 주석 달기가 처음 실시된 2015년 ‘한국수어 말뭉치 기반 조사 및 시범 구축 사업’을 시작으로, 2016년 수집 영상 및 주석 파일 관리를 위한 시스템 구축, 2017년 기 구축된 수어 말뭉치 영상 및 주석 파일의 시스템 탑재 및 추가 주석이 실시되었다. 원시 말뭉치 수집이 다시 시작된 2018년에는 4개 지역 말뭉치 수집과 서울지역 수집 영상에 대한 기본 주석 및 지침과 타입 정비를, 2019년에는 2018년도 수집한 경남지역의 원시 말뭉치 추가 수집 및 기구축된 서울지역의 주석 파일의 기본 주석 정비 및 동일한 파일에 대해 의미 주석을 실시하여 왔다. 본 연구에서는 전남지역(광주 포함)의 15쌍, 50시간 이상의 추가 말뭉치 수집과 경남지역 80분 분량의 수어 자료에 대해 기본 주석 및 의미 주석을 목표로 수행하였으며, 한국수어 말뭉치에 대한 중장기 계획(안)도 수립하였다.

1) 한국수어 원시 말뭉치 구축

본 사업에서는 2015년, 2018년과 2019년에 이어 한국수어 원시 말뭉치를 구축하였다. 구축 과정은 ‘지역 선정 - 언어제공자 질문지 수정 - 유도 자료 제작 - 언어제공자 선정 - 사회자 선정 및 교육 - 원시 말뭉치 촬영 및 편집 - 「한국수어 자료 통합지원 시스템」 탑재’의 순서로 이루어졌다. 이에 따른 결과는 다음과 같다.

첫째, 지역 선정에서는 한국수어 말뭉치 자료의 지역 균형성을 확보하기 위해 2018년 수행된 원시 말뭉치 수집 지역 중 4쌍이 수집된 전남지역(광주 포함)을 선정하였다.

둘째, 언어제공자 질문지 수정에서는 언어제공자의 메타데이터를 수집하기 위한 질문지의 내용을 보다 직관적으로 파악할 수 있도록 문항의 양식을 통일하였던 2019년의 자료를 사용하였다.

셋째, 유도 자료 제작에서는 지역 수어의 포착과 더불어 풍부한 한국수어의 채집을 위해 지역 수어 모니터링을 진행하고, 2015년, 2018년, 2019년에 사용한 유도 자료를 기반으로 시의성과 지역성을 가미하여 ‘E-토론’, ‘G-특정 사건 묘사’, ‘I-개별 수어의 다양한 표현’, ‘J-농사회 행사’의 일부 주제를 수정하였다. ‘E-토론’에서는 2019년도 자료를 유지하면서 시의성을 나타내는 ‘코로나 개학 연기’를 추가하였으며, ‘G-특정 사건 묘사’에서는 지역성을 나타내는 ‘2011년 도가니 사건’을 추가하였다. ‘I-개별 수어의 다양한 표현’에서는 전남지역에서 사용하는 지역의 방언을 유도하기 위해 ‘김치’, ‘가래떡’ 자료를 추가하였으며, ‘J-농사회 행사’ 관련 전남 협회에서 개최한 행사 포스터로 교체, 제시하여 언어제공자들의 적극적인 대화 참여를 유도하였다. 또한 지역 수어를 구사하는 언어제공자들이 안정감을 느끼게 함과 동시에 서울 수어의 간섭을 배제하기 위하여 지역에서 선정된 사회자가 유도 자료 설명 동영상 촬영하도록 하였다.

넷째, 언어제공자 선정에서는 언어제공자 우선 선정 조건 등을 고려하여 최종적으로 16쌍의 언어제공자를 선정하였으나 촬영 해상도의 문제로 1쌍의 언어제공자가 추가되었으며, 최종 영상 편집 결과 촬영 시간 부족으로 2쌍의 언어제공자를 추가 촬영한 결과 최종 19쌍의 언어제공자의 수어 자료가 확보되었다.

다섯째, 사회자 선정 및 교육에서는 언어제공자들의 대화를 돕기 위한 사회자를 선정하였는데, 언어제공자의 수, 촬영 일정 등을 고려하여 2명의 사회자를 선정하여 본 사업에 대한 전체적인 내용과 더불어 사회자의 역할을 교육하였다.

여섯째, 원시 말뭉치 촬영에서는 3시간 30분 이상의 촬영 시간을 예

상하고 시작하였으나 대부분 3시간 내외의 시간으로 촬영됨은 물론 촬영된 영상 내에서도 언어 자료로 확보될 수 있는 편집 영상의 시간이 짧아 2쌍의 언어제공자의 촬영을 추가하였다. 그 결과 19쌍의 언어제공자에 대한 전체 촬영 시간은 약 59시간 53분 분량이었으며, 이를 편집하여 최종적으로 수어자료로 수집된 영상은 동의서를 촬영한 시간을 포함하여 49시간 24분 분량을 확보하였다. 본 사업 뿐 아니라 추후 이루어질 사업과의 통일성을 고려하여 편집 매뉴얼을 작성하여 공유하였고, 촬영된 수어 자료 영상의 오류를 최소화하기 위하여 3명의 연구원이 총 4차례에 걸쳐 검증과정을 거쳤다. 이러한 과정을 통해 수집된 원시 말뭉치 영상 자료를 「한국수어 자료 통합지원 시스템」에 업로드하여 원시 말뭉치를 확대·구축하였다.

2) 기본 주석 및 의미 주석

기본 주석은 주석 데이터로 만들어지지 않은 80분 분량의 기 수집 경남 지역 말뭉치 수어 자료에 대해 실시하였다. 모두 7개의 파일로 ‘과제 J(농사회 행사)’, 4개(42분 47초)와 ‘과제 M(농학교 경험)’, 3개(37분 32초)로 80분 19초 분량이었으며, 기본 주석을 실시한 결과 13,673 토큰이 나타났다. 이 중 농사회 행사 과제의 토큰이 7,523개, 농학교 경험 과제의 토큰은 6,150개였다. 기본 주석을 실시하는 과정에서 언어제공자가 수어를 완전히 조어하지 않았을 때, 수어 화자가 실수로 수어를 잘못 표현했거나 그 수어가 사용되지 않는 수어일 때, 그리고 주석자가 수어로 인지했지만 무슨 수어인지 모르거나 영상화면 바깥에서 조음되어 보이지 않을 때 각각 [%], [@], [###]의 기호로 기록하게 되는데, [%]는 212 토큰, [@]는 52 토큰, [###]는 158 토큰으로 모두 422개의 토큰이 나타났다. 이중 [###]의 빈도는 수어의 ID 글로스가 해결된 경우는 제외된 것이며 이처럼 기타 기호로 표기된 것을 제외한 실질 토큰 수는 13,250개였다.

의미 주석은 기본 주석 층렬에서 분절되어 ID 글로스로 기록된 토큰 중

에서 [%], [@], [###], [지시]를 제외한 모든 토큰에 의미를 주석하였다. 해당 토큰을 기록할 때는 문장 맥락 안에서의 의미를 중심으로 대역어 형태로 기술하였으며 이렇게 의미 주석된 토큰 수는 3개의 농학교 경험 과제에서 5,228 토큰, 농사회 행사 과제에서 6,538 토큰으로 총 11,766 토큰이었다. 고정된 수어 타입 [이상하다1]은 비수지 및 문맥에 따라 ‘호기심’, ‘궁금하다’의 의미로 쓰였으며, 일반적인 의미로 ‘괜찮다’, ‘나쁘지 않다’, ‘무방하다’ 등으로 사용되는 타입 [괜찮다1]은 ‘할 수 있다’는 의미로 쓰이는 것이 확인되었다. 생산적 수어 타입 [구형4]는 납작한 원반의 공 모양의 ‘펼’을 대체해서 나타나기도 하고, 지팡이를 잡은 손의 모양으로 나타나기도 하였다.

주석과 검증 과정에서 지속적으로 논의, 검토된 내용은 기본 주석 과정에서 정확한 수어의 형태를 확인할 수 있는 분절과 타입에 관한 것이었다. 분절에서 발생하는 오류는 ‘동일하게 시작한 두손 수어에서 한손이 한손 수어를 발화하기 위해 이동하는 동안 다른 한손이 유지되어 우세손과 비우세손의 토큰 발화시간이 서로 상이하지만 동일하게 분절하는 경우([공부1]+[지시1])’, ‘수어 발화를 시작하기 위해 움직이려는 지점부터 분절하여 수어의 시작점으로 볼 수 없는 준비 단계부터 시작점으로 분절하는 경우([사년전1])’, ‘양손 수어 타입인 [경험1]의 첫 발화가 시작점에서 한손의 움직임이 먼저 나타남으로써 양손의 분절 길이를 달리 하는 경우’ 등에서 나타나는 것으로 확인되었다.

경남지역 수어 자료의 주석 달기를 통해 새롭게 등록된 타입은 160개였다. 타입 후보로 등록된 155개의 타입 중 등록된 타입은 108개였다. 47개 타입은 중복되거나 타입명 결정 규약에 맞지 않는 것이었으며, 검증 과정에서 새롭게 확인된 타입 52개는 연구팀에서 등록하였다.

3) 한국수어 말뭉치 증장기 발전 계획

2015년 국내 최초로 구축된 한국수어 말뭉치는 한국수어의 수집과 보존,

체계적인 연구를 위한 수어 자료로서의 활용 효용성이 있는 만큼 체계적으로 운영되고 관리될 필요가 있다. 이에 그동안의 한국수어 말뭉치의 성과와 해외 말뭉치 현황을 토대로 중장기 발전 계획을 수립하였다.

한국수어 말뭉치의 성과 분석에서의 수집에서는 2019년까지 5개 지역 110명의 언어제공자를 대상으로 촬영한 약 167시간 분량의 수어자료가 각 과제별로 분할, 편집 후 총 4,278개, 137시간 56분 12초 분량의 수어 영상이 확보되어 시스템에 탑재되었다. 이들 수어 영상에 대한 주석 데이터는 총 20시간 13분 분량이며 여기서 나타난 총 토큰 수는 154,172개였다. 이러한 성과를 통해 지역별 수어 특성과 변이를 확인하기에는 한정된 지역만을 대상으로 한 점, 불안정한 촬영 환경으로 인한 영상 수집 과정이 원활하지 않는 점, 주석 전문가가 안정적으로 공급되지 않아 주석의 일관성과 신뢰성 확보에 어려움이 있는 점 등의 한계가 확인되었다.

해외 수어 말뭉치의 현황을 분석한 결과, 서로 다른 목적을 가지고 수어 말뭉치를 구축하고 있지만 장기간의 프로젝트로 수행되고 있으며, 지역을 대표할 수 있을 만큼의 충분한 수의 언어제공자로부터 수어 영상을 수집하고 있고, 지속적이고 다양한 층위의 주석 작업을 통해 수어 말뭉치를 기반으로 수어의 언어적 특성을 밝히는 다양한 연구가 진행되어 왔으며, 또한 나라별로 공개 범위와 방법은 상이하나 수어 말뭉치 구축과 함께 공개되고 있다는 점에서 우리에게 시사하는 바가 컸다.

본 연구에서는 산재한 국내 수어 자료 실태 분석을 위해 농아방송과 수어 영상 도서를 살펴보았다. 이들 영상은 문자가 없는 수어의 언어적 특징상 영상 제작을 위해 한국어를 빌어 작성된 스크립터나 시나리오를 사용하기 때문에 한국어의 영향을 받는 수어 발화라는 점에서 자연수어로서의 한계와 한국수어의 원형 보존이라는 한국수어 말뭉치의 성격과는 다소 상이하다는 점에서 한국수어 말뭉치 구축에 포함되기보다는 번역 말뭉치 또는 병렬말뭉치 형태의 별도 연구 과제로 개발되어야 할 것으로 사료된다. 실제 사용되고 있는 언어를 수집하는 한국수어 말뭉치 구축도 해결해야 할

과업들이 많으므로 한국수어 말뭉치에 구축 자체에 집중해야 할 것으로 판단된다.

한국수어 말뭉치 구축 사업은 한국수어의 지역별 다양성을 확인할 수 있을 만큼의 수집이 가능하고, 한국수어의 형태 분석을 가능하게 할 뿐만 아니라 전산시스템에서 관리할 수 있다는 강점이 있는 반면 구축 기간에 비해 주석 데이터가 미흡하여 한국수어 체계에 대한 수어학적 근거 제시에 제한점이 있으며, 한정적인 예산으로 인해 구축에 집중할 수 있는 여건을 갖추기가 어려운 점이 약점으로 확인되었다. 그러나 한국수화언어법의 제정과 코로나19로 인한 공공수어 통역의 확대로 수어에 대한 관심이 고조되었을 뿐만 아니라 기계번역을 위한 한국수어 자료에 대한 요구가 있는 만큼 안정적으로 중장기적 계획 하에 계속해서 진행되어야 한다.

한편 한국수어 사용인구의 감소, 기계번역 및 인공지능 학습을 위한 과학기술부의 수어 자료 수집 참여로 자연수어 수집에 대한 필요성에 대한 인식 상실의 우려도 있는 만큼 한국수어 말뭉치 구축의 절대적 필요성을 근거로 부단한 노력을 기울여 나가야 할 것이다.

이를 위해 향후 5년 동안 전국 단위의 수어 자료 수집과 번역, 주석되지 않은 방대한 양의 수집된 수어 영상에 대한 주석 데이터화, 그리고 대중과 연구자를 대상으로 한 공개의 핵심 과제별 추진 계획에 따라 전략적이고 집중적으로 사업을 추진해 나가야 할 것이다. 또한 5년간의 구축과 공개가 완료된 6년차부터는 수어학습자 말뭉치, 수어 영상 도서관 농아방송과 같은 자료를 통한 수어와 구어 간 비교를 위한 병렬말뭉치 등 다양한 종류의 말뭉치로 확대한 말뭉치 수집, 주석 데이터화와 기 구축된 주석 데이터의 다양한 목적별 주석을 추진해 나가야 할 것이다.

따라서 한국수어 말뭉치는 수집, 주석 데이터 구축, 공개라는 3가지 핵심과제를 수립하였다. 수집은 지역 농아인협회와의 협력 관계 체결, 전문 촬영팀 확보, 언어제공자의 안정적 공급, 영상 번역물 확보라는 4가지 추진방향을, 주석에서는 기본 주석 및 의미 주석의 내실화, 주석 전문 인력

의 확보, 주석 규약의 정교화, 「한국수어 자료 통합지원 시스템」의 안정화라는 4가지 추진방향을, 그리고 공개에서는 공개 범위 및 방법 확정, 온라인 공개 시스템 구축, 공개 영상 및 주석 자료 정비라는 3가지 추진방향에 따른 세부 성과목표를 수립하였다.

2. 활용 방안 및 제언

1) 연구의 활용 및 기대 효과

전남지역의 원시 말뭉치의 추가 수집을 통하여 지역 수어의 대표성을 갖추는 자료를 추가 확보하게 되었다. 특히 목표인 15쌍보다 많은 19쌍의 언어 자료를 수집함으로써 전남지역 수어 사용자의 다양한 수어 형태를 확인할 수 있을 것으로 기대된다.

기 수집된 경남지역의 영상에 기본 주석과 의미 주석 달기를 실시함으로써 향후 서울지역의 수어와 비교하는 연구 수행이 가능하게 되었으며, 지역 간 변이를 확인할 수 있는 계기가 마련되었다.

한국수어 말뭉치 구축 중장기 계획으로 말뭉치 구축의 방향을 모색하였으며, 이것이 연차적으로 추진된다면 향후 5년 이내 수집과 구축이 완료되고, 주석 데이터 구축은 한국수어 언어의 특성을 체계적으로 밝히는 기반이 될 수 있을 것이다. 또한 한국수어 말뭉치 공개 과업에 따라 대중과 연구자에게 공개된다면 농문화의 전승 효과 및 수어 연구 활성화의 기반이 마련될 것이다.

2) 제언

한국수어 말뭉치 구축은 장기적 관점에서 안정적으로 이루어져야 하는 사업으로 본 연구에서는 5년 이내 기본적인 수집과 주석 데이터 구축을 마

치는 계획을 제시하였다. 성공적 한국수어 말뭉치 구축 연구를 위해 다음의 것들을 제안하고자 한다.

첫째, 지역별 수어 말뭉치 수집 완료 시기를 최대한 앞당길 필요가 있다. 지역별 원시 수어 말뭉치 수집이 아직 5개 지역만 완료된 상태로 1년에 한 시·도 지역만 수집을 하면 향후 6-7년의 시간이 소요될 수 있다. 이는 시간의 흐름에 따라 기존 수어 세대가 점점 사라져 가고 있는 현실에서 한국수어의 이질적 형태가 실제로 나타나고 있다. 가급적 빠른 시간 내에 현재 최고령 세대를 이루고 있는 한국수어 모국어 화자를 통해 한국수어 원시 말뭉치를 수집해야 할 것이다.

둘째, 한국수어 말뭉치 수집의 질을 보장하기 위한 전문 촬영팀을 육성, 운영할 필요가 있다. 원시 한국수어 말뭉치의 수집의 질을 보장하기 위해서는 양질의 수어 수집 환경을 확보해야 한다. 한국수어 말뭉치 구축 사업에 대한 이해와 풍부한 수어 촬영 경험을 갖춘 인력 확보가 양질의 수어 수집 환경의 핵심이다. 모든 것을 촬영을 통해 영상으로 자료를 수집하고 보존해야 하는 수어의 특성 상 전문 촬영팀의 육성과 운영은 필수 조건으로 사료된다.

셋째, 주석의 일관성과 신뢰성을 제고하기 위한 주석 전문 인력 확보 방안이 마련되어야 한다. 주석 전문 인력 확보는 크게 연수회를 통한 주석 전문 인력 저변 확대와 국립국어원에서의 주석 전문 인력 고용과 운용의 두 가지 측면에서 접근해야 한다. 수어 영상 자료를 통해 수어 주석을 하는 행위는 매우 정교한 기술을 요함과 동시에 시간의 변화에 따라 주석을 하는 기준과 관점의 변화가 일어날 가능성이 높다. 따라서 지속적 시간 속에서 일관된 관점으로 주석을 꾸준히 진행할 때 주석의 일관성과 신뢰성을 확보할 수 있다.

넷째, 한국수어를 안정적으로 보관 관리하기 위해 「한국수어 자료 통합 지원 시스템」의 기능 개선이 신속히 이루어져야 한다. 현재의 「한국수어 자료 통합지원 시스템」은 가공된 수어 자료의 탑재, 탐색, 호출 등의 전

과정에서 동일한 문제에 대한 처리 불가 현상이 반복되고 일정한 패턴이 없어 작업을 지연시키는 등 여러 문제점을 노출시키고 있다. 시스템의 현재 기능 수준으로는 향후 축적된 수어 자료 공개를 어렵게 할 위험성이 클 것으로 판단된다.

다섯째, 축적된 한국수어 자료의 공개 계획을 체계적으로 수립하여야 한다. 지난 7년간 한국수어 말뭉치 구축 연구가 진행되어 상당량의 지역 한국수어 자료의 수집과 주석 작업이 진행되었다. 국립 기관에서 진행하여 수집된 자료와 정보를 공개하는 것 바람직하다는 것에는 이견이 없다. 하지만 공개를 위해서는 공개 시기, 공개 범위와 방법, 공개 누리집 구축, 일반인 대상의 공개와 연구 집단 대상의 공개 방법 간의 차이 등등 빠른 시간 내에 공개를 위한 체계적 계획 수립과 실제 공개가 요구된다.

여섯째, 한국수어 말뭉치 구축을 다년도 사업으로 전환해야 한다. 현재 매년 모집 공고를 내고 연구팀을 공모하여 본 사업을 진행해 오고 있다. 이런 연구 환경에서는 한국수어 말뭉치 구축 연구를 중장기 계획을 갖고 수행할 수 없다. 다년 사업으로 연구 위탁이 이루어질 때, 위탁받은 연구진은 중장기 연구계획을 수립하고 의미 있는 연구 결과물을 도출해 낼 수 있을 것이다.

참고문헌

<국내외 논저>

- 강창욱 외(2019), 『2019년 한국수어말뭉치 연구 및 구축』, 국립국어원.
- 원성욱 외(2015), 『한국수어 말뭉치 기반 조사 및 시범 구축』, 국립국어원.
- 윤병천 외(2014), 『한국수어 코퍼스 구축을 위한 기초 연구』, 국립국어원.
- 윤석민 외(2018), 『2018년 한국수어 말뭉치 연구 및 구축』, 국립국어원.
- 홍성은(2010), 「수화언어학의 유도 자료와 그 사용」. 제12회 한국수화학회 학술대회 자료집, 9-17.
- 홍성은, 니시오 리에(2010), 「독일수어 코퍼스 프로젝트에서의 유도 방법」. 제12회 한국수화학회 학술대회 자료집, 49-62.
- Cormier, K., Crasborn, O. & Bank, R.(2016), Digging into Signs: Emerging Annotation Standards for Sign Language Corpora. Paper presented at the 7th Workshop on Representation and Processing of Sign Languages: Corpus Mining [workshop part of 10th International Conference on Language Resources and Evaluation Conference], Portorož, Slovenia.
- Cormier, K., Crasborn, O. & Bank, R.(2016), The role of theory in sign language corpus annotation. Poster presented at the 12th Conference on Theoretical Issues in Sign Language Research, Melbourne, Australia.
- Cormier, K., Schembri, A.(2014), Describing sociolinguistic variation in verb directionality in British Sign Language: A corpus-based study. Final report to the Economic and Social Research Council.pdf
- Crasborn, Onno & Zwitserlood, Inge.(2008), The Corpus NGT: an online corpus for professionals and laymen. Construction and Exploitation of Sign Language Corpora. 3rd Workshop on the Representation and Processing of Sign Languages.
- Fenlon, J., Schembri, A., & Cormier, K.(2018), Modification of indicating verbs in British Sign Language: A corpus-based study. Language,

- 94(1), 84-118.
- Fenlon, J., Schembri, A., Johnston, T., & Cormier, K. (2015). Documentary and corpus approaches to sign language research. In E. Orfanidou, B. Woll, & G. Morgan (Eds.), *The Blackwell guide to research methods in sign language studies* (pp. 156-172). Oxford: Blackwell.
- Fenlon, J., Schembri, A., Rentelis, R., & Cormier, K. (2013). Variation in handshape and orientation in British Sign Language: The case of the '1' hand configuration. *Language and Communication* 33(1), 69-91. doi: 10.1016/j.langcom.2012.09.001
- Fenlon, J., Schembri, A., Rentelis, R., Vinson, D., & Cormier, K. (2014). Using conversational data to determine lexical frequency in British Sign Language: The influence of text type. *Lingua*, 143, 187-202. doi: 10.1016/j.lingua.2014.02.003
- Hanke, Thomas., Schulder, Marc., Konrad, Reiner., & Jahn, Elena. (2020). "Extending the Public DGS Corpus in Size and Depth". In Efthimiou, Eleni et al. (eds.): *Sign Language Resources in the Service of the Language Community, Technological Challenges and Application Perspectives. Proceedings of the 9th Workshop on the Representation and Processing of Sign Languages. 12th International Conference on Language Resources and Evaluation (LREC 2020)*, Marseille, France, 2020. Paris, France: European Language Resources Association (ELRA), S. 75-82.
- Hodge, G., Fenlon, J., Schembri, A., Johnston, T., & Cormier, K. (2018). Negation in two related signed languages: British Sign Language and Auslan. Paper presented at the Annual Meeting of Linguistics Association of Great Britain (LAGB), 11-14 September 2018, Sheffield, UK.
- Hodge, G., Kazuki Sekine, K., Schembri, A., & Johnston, T. (2019). Comparing signers and speakers: building a directly comparable corpus of Auslan and Australian English. *Corpora*, 14(1), 63-76.

- Johnston, T. (2016). Auslan Corpus Annotation Guidelines [November 2016 version].
- Johnston, T., & A. Schembri. (2013). 'Corpus analysis of sign languages' in C. A. Chapelle(Ed.). *The Encyclopaedia of Applied Linguistics*. UK: Blackwell Publishing Ltd. DOI: 10.1002/9781405198431.wbeal0252
- Johnston, Trevor. (2010). From archive to corpus: Transcription and annotation in the creation of signed language corpora. *International Journal of Corpus Linguistics*. 15.
- Johnston, T., Cresdee, D., Schembri, A., & Woll, B. (2015). FINISH variation and grammaticalization in a signed language: how far down this well-trodden pathway is Auslan (Australian sign language)? *Language Variation and Change*. DOI: <https://doi.org/10.1017/S0954394514000209>
- Macquarie University, Sydney. Available at: <http://www.auslan.org.au/about/annotations/>
- Rowley, K., Fenlon, J., & Cormier, K. (2018). Attitudes to age variation and language change in the British Deaf community. Paper presented at the Sign CAFÉ 1: The first international workshop on cognitive and functional explorations in sign language linguistics, Birmingham, UK, 30-31 July 2018.
- Rowley, K., Fenlon, J., & Cormier, K. (2018). Language Attitudes in the British Deaf Community: Attitudes towards British Sign Language and Sign Supported English. Paper presented at the Annual Meeting of Linguistics Association of Great Britain (LAGB), 11-14 September 2018, Sheffield, UK.
- Schembri, A., Fenlon, J., & Cormier, K. (2017). Clause position and indicating verbs in British Sign Language. Paper presented at the 2017 Annual Meeting of the Linguistic Society of America, Austin, Texas.
- Schembri, A., Fenlon, J., & Cormier, K. (2018). Variation in British Sign

Language indicating verbs. Paper presented at the New Ways of Analyzing Variation NWAV 47, New York University.

Schembri, A., Fenlon, J., Rentelis, R., Reynolds, S., & Cormier, Kearsy (2013). Building the British Sign Language Corpus. Language Documentation and Conservation 7, 136-154.

<온라인 자료>

네덜란드 언어 보관소 누리집 <https://archive.mpi.nl/tla/>

독일 수어 말뭉치 누리집 <https://www.sign-lang.uni-hamburg.de/dgs-korpus/index.php/welcome.html>

미국 수어 말뭉치 보관소 <https://bslcorpusproject.org/cava/>

영구 수어 말뭉치 프로젝트 누리집의 데이터 공개 안내

※ 출처: <https://bslcorpusproject.org/data/>

호주 말뭉치 보관소 <https://elar.soas.ac.uk/Collection/MPI55247>

CAVA 영국 수어 말뭉치 보관소 <https://bslcorpusproject.org/cava/>

CMDI <http://www.clarin.eu/content/component-metadata>

http://media.dwds.de/clarin/userguide/text/metadata_CMDI.xhtml

DGS Korpus Projekt 독일수어 코퍼스 프로젝트 www.dgs-korpus.de

EU Project Visicast 아바타 개발 프로젝트 www.visicast.co.uk

IMDI http://corpus1.mpi.nl/ds/imdi_browser

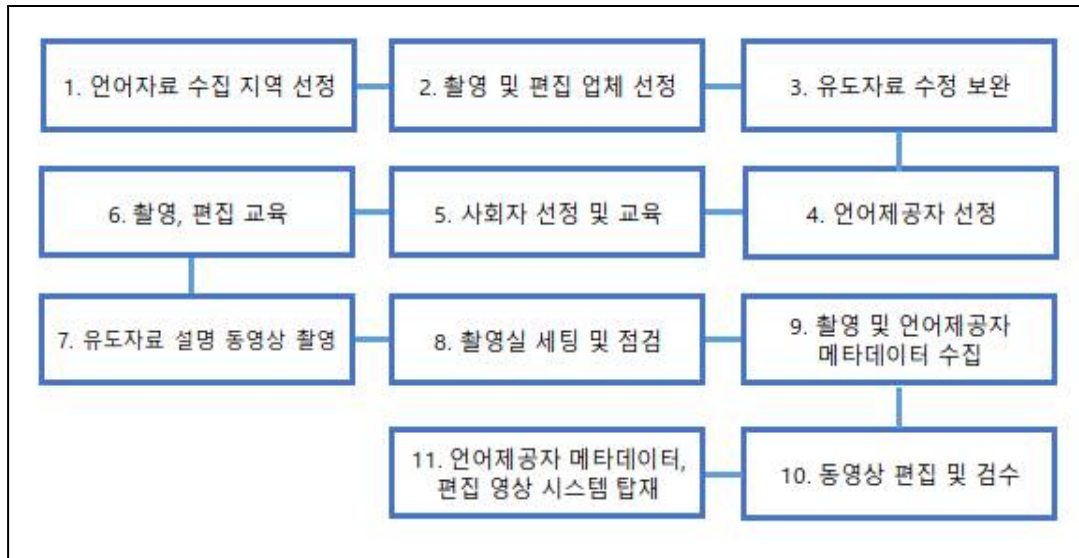
TLA의 언어 보관소 누리집 https://archive.mpi.nl/tla/islandora/object/lat%3A1839_00_0000_0000_0000_0000_4

[부 록]

[부록 1] 한국수어 말뭉치 구축 가이드

I. 원시 말뭉치 수집

1. 원시 말뭉치 수집 과정도



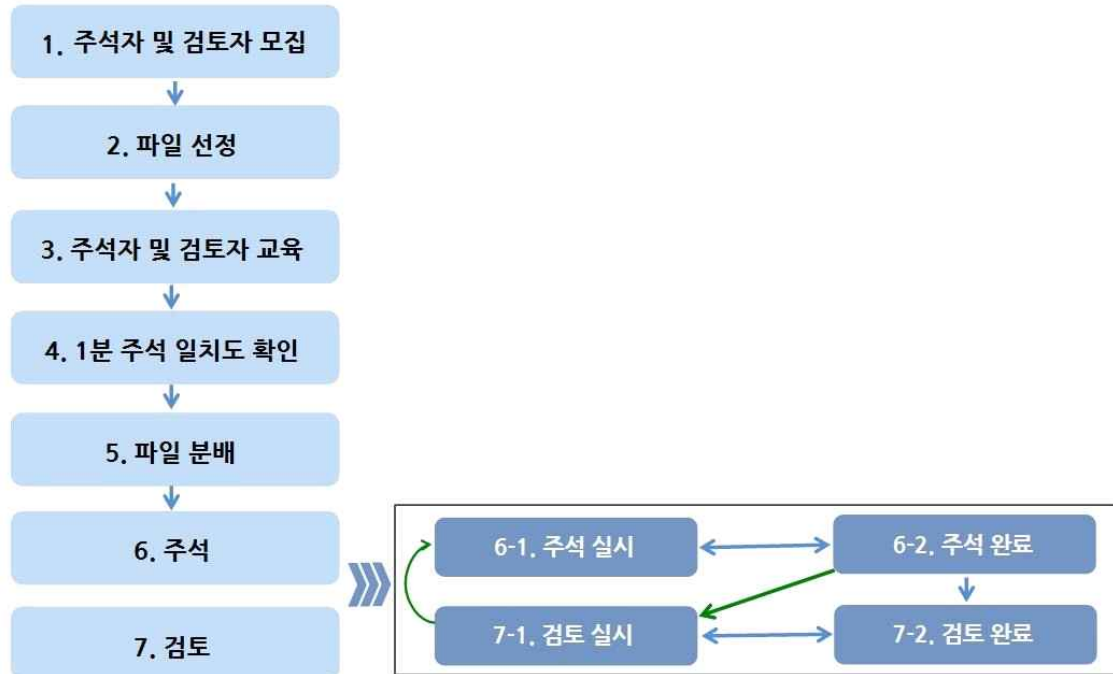
2. 과정별 주요 과업

과정 1	
언어 자료 수집 지역 선정	<ul style="list-style-type: none"> • 기존 한국수어 말뭉치 언어 자료 수집 현황 참고 • 인구 고령화, 농학교 거점 지역 등의 조건을 고려 • 수어 자료 수집 지역을 최종 선정
과정 2	
촬영 및 편집 업체 선정	<ul style="list-style-type: none"> • 수어 자료 촬영과 편집을 수행할 업체 선정
과정 3	
유도 자료 수정 보완	<ul style="list-style-type: none"> • 기존 제작된 유도 자료 중 시의성, 지역적 특성을 반영하여 주제 수정 보완
과정 4	
언어제공자 선정	<ul style="list-style-type: none"> • 수어 자료 수집 조건을 충족 언어제공자 목록 작성 • 언어제공자의 선정 조건 결정 <ol style="list-style-type: none"> ① 19세 이상 성인 농인 ② 한국수어를 모어로 하는 농인 - 의사소통을 위해 수어를 가장 편하게 자주 사용함 ③ 농학교 졸업(부모가 모두 농인인 경우 일반학교 졸업 농인도 가능)

	④ 현재 00 지역에서 10년 이상 거주 혹은 직장생활을 하고 있음 ⑤ 주 3회 이상 농인을 만남
과정 5	
사회자 선정 및 교육	<ul style="list-style-type: none"> 언어 자료 수집 지역에서 10년 이상 거주 한국수어로 편안하게 대화를 유도할 수 있는 농인 사회자 교육 실시 사회자 교육의 주요 내용 <ol style="list-style-type: none"> 수어 말뭉치와 유도 자료의 내용 및 목적에 대한 이해 언어제공자가 자연스러운 분위기에서 대화할 수 있는 분위기 조성 방법 두 언어제공자의 수어가 균형 있게 산출될 수 있도록 조절하는 전략 유도 자료 설명 동영상 촬영시 고려사항 기타 사회자가 작성하여야 할 문서 작성 방법
과정 6	
촬영, 편집 교육	<ul style="list-style-type: none"> 촬영과 편집에 관련된 지침 제작 촬영 및 편집 업체와의 미팅을 통한 지침 내용 전달 촬영 및 편집 지침 <ol style="list-style-type: none"> 촬영실 조건 및 촬영 시스템 지침 촬영 매뉴얼 편집 지침
과정 7	
유도 자료 설명 동영상 제작	<ul style="list-style-type: none"> 언어제공자의 이해를 돕기 위해 유도 자료 설명 동영상 제시 타 지역수어의 영향을 방지하기 위해 선정된 사회자가 기존 유도 자료 설명 동영상의 내용을 해당 지역 수어로 재촬영
과정 8	
촬영실 세팅 및 점검	<ul style="list-style-type: none"> 촬영실 조건 및 촬영 지침에 따라 촬영실 구축 ※ 구체적인 촬영실 구조는 보고서 참고
과정 9	
촬영 및 언어제공자 메타데이터 수집	<ul style="list-style-type: none"> 촬영 매뉴얼에 따라 촬영을 진행 자료 수집 질문지로 언어제공자 메타데이터 수집
과정10	
동영상 편집 및 검수	<ul style="list-style-type: none"> 편집 지침에 따라 영상 편집 납품된 영상 검수
과정11	
언어제공자 메타데이터 및 편집 영상 시스템 탑재	<ul style="list-style-type: none"> 언어제공자 메타데이터, 편집 영상과 기타 정보들을 「한국수어자료통합지원시스템」에 탑재

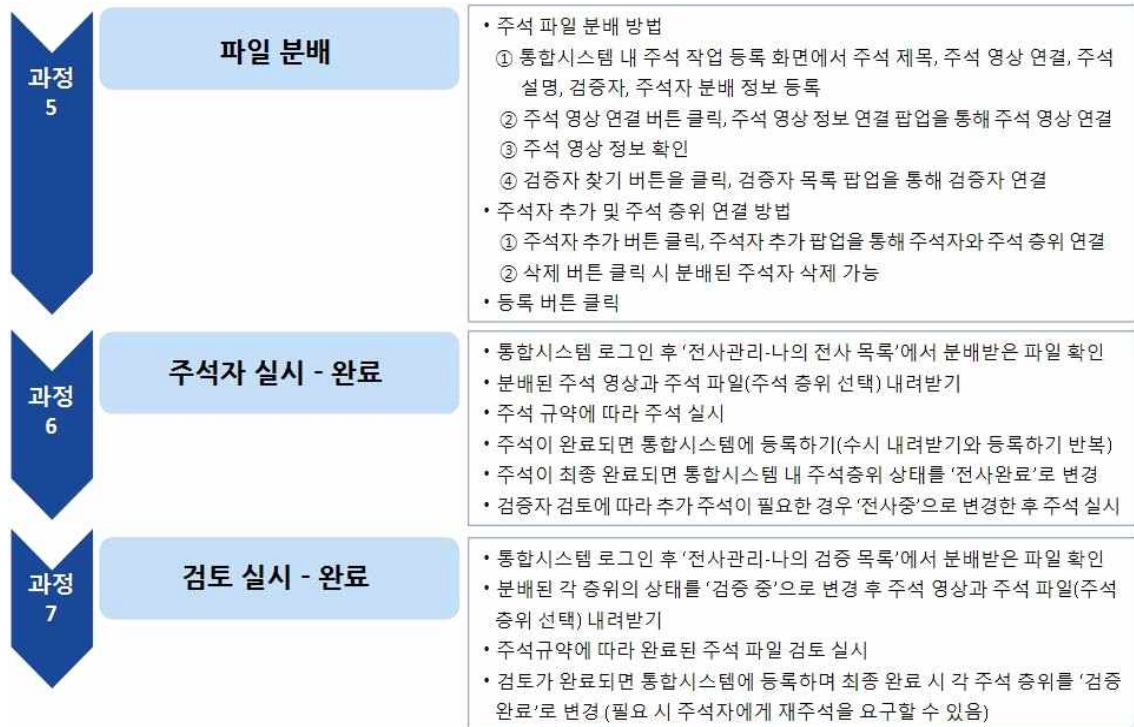
II. 말뚝치 주석 데이터 구축

1. 주석 데이터 구축 과정도



2. 과정별 주요 내용

과정 1	주석자 및 검토자 모집	<ul style="list-style-type: none"> • 농인, 수어통역사, 수어 관련 전공자 • 수어 언어학에 대한 기본적인 지식이 있는 자 우선 선발 (검토자의 경우 농인 우선 선발) • 관련 교육에 참여할 수 있는 자
과정 2	파일 선정	<ul style="list-style-type: none"> • 주석 파일 선정 • 「한국수어 자료 통합 지원 시스템」(이하 통합시스템)에서 주석 증위 등록 <ul style="list-style-type: none"> ① 부모증위, 증위타입명, 증위명, 타입 연결 여부, 타입 연결 타입 범주, 입력 토근 종류, 함께 사용되는 증위, '비고' 정보 입력 ② 증위 등록 버튼 클릭
과정 3	주석자 및 검토자 교육	<ul style="list-style-type: none"> • 수어 언어학에 대한 기본적인 지식이 있는 주석자 교육은 최소 4시간 실시 • 수어 언어학에 대한 기본적인 지식이 없는 주석자는 추가 교육 실시 • 주석자 교육 주요 내용 <ul style="list-style-type: none"> ① 말뚝치 구축에 대한 전반적인 내용 (1시간) ② 주석을 위한 기본 사항 및 원리 (1시간) ③ 주석 프로그램 ELAN (1시간) ④ 통합시스템 사용법 (1시간)
과정 4	1분 주석 일치도 확인	<ul style="list-style-type: none"> • 주석 일치도를 높이기 위한 목적 • 주석자 및 검토자 대상 1분 분량의 시범 파일 분배, 주석, 일치도 확인 • 주석 일치도가 낮은 경우 5회 이상 공동 주석 실시 • 정기적인 주석 회의 실시



[부록 2] 한국수어 말뭉치 주석 규약

1. 개요

이 규약은 2015년 국내 최초로 한국의 원시 수어 말뭉치를 수집하면서 생성된 동영상상을 주석하면서 제작된 ‘전사 지침’을 수정·보완한 것이다. 이 규약은 해마다 추가로 수행되는 주석 작업에 따라 새로운 내용이 추가되거나 수정될 수 있다.

이러한 이유로 규약의 명칭은 몇 차례 개정의 과정을 거쳤다. 한국수어 말뭉치 주석 작업이 처음 시도될 당시에는 ‘전사’, ‘전사 지침’ 등의 용어로 사용되었다. 그러나 문자가 없는 수어의 기록은 한국어 문자를 빌려와 이름(gloss)을 붙이는 방식인 것임에도 언어의 형태를 그대로 받아 적는다는 의미의 ‘전사(transcription)’로 사용되어 언어학자 간 용어의 혼동을 초래하였다. 따라서 본 규약은 2019년에 ‘전사’를 ‘주석(annotation)’으로 통일하여 사용하기로 하고 ‘전사 지침’에서 ‘주석 지침’으로 개칭하였다.

또한 2020년에는 주석에 대한 지침을 제공하기보다는 현재까지 이루어진 주석에 대한 주석 규칙 및 약속을 공유한다는 차원에서 ‘주석 규약’으로 개정하였다.

이에 본 주석 규약은 지금까지 수행된 번역 및 주석을 수행하기 위해 알아야 할 구체적인 규칙과 약속을 다루고 있다. 2015년부터 수행된 의미 있는 손 활동에 대한 기본 주석, 2016년에 실시한 비수지 중 입 움직임에 대한 주석, 그리고 2019년에 실시한 결합 수어(합성어), 생산적 수어의 도상성, 의미주석에 대한 기본 규약을 기술하였다.

향후 새로운 주석 과업에 수행되면 그에 따라 주석 규약은 수정 보완될 것이며, 기타 항목에는 앞으로 수행할 수 있는 주석 과업 내용에 대한 제목만을 제시하였다.

1.1. 주석의 기본 원칙

한국수어 말뭉치 구축을 위한 주석은 엘란(EUDICO Linguistic Annotator, ELAN) 프로그램을 사용한다. 엘란은 멀티미디어 기록물(녹음 또는 영상자료 등)에 필요한 정보를 주석하여 탐색할 수 있도록 네덜란드 막스플랑크 심리언어학 연구소에서 개발한 프로그램이다.

주석의 가장 기본은 모든 의미 있는 손 활동에 이름을 붙이는 것이다. 따라서 주석을 하고자 할 때는 가장 먼저 엘란에 수어 화자별로 양손 활동을 기록할 수 있는 두 개의 수지 층렬을 만든다.

주석을 할 때는 각 수어 단어가 시작되는 지점에서 끝나는 지점까지를 분절하여 해당 부분에 글로스를 단다. 이때, 수지 층렬을 주석 작업에서는 실제 언어자료에서 나타나는 수어 어휘(토큰)의 기본형²⁾을 타입으로 등록하며 체계적으로 관리하게 된다. 그래서 주석 작업을 수행할 때는 토큰-타입 매칭을 실행하여야 한다.

그런데 한국수어 말뭉치가 처음 구축될 당시 한국수어사전은 표제화 원칙에 따라 체계적으로 이루어진 것이 아니어서 본 주석 작업에서 사용되는 타입 목록은 한국수어사전을 참고할 수 없었다. 이에 주석 과정에서 새로운 토큰이 나타날 때마다 타입을 등록하면 타입 목록에 누적된다. 현재 타입 목록은 「한국수어 자료 통합지원 시스템」에서 관리하고 있다.

1.2. 번역

향후 한국수어 말뭉치를 공개할 경우, 수어를 모르는 사람들도 자료에 접근할 수 있도록 한국어 번역을 제공한다. 번역 목적에 따라 의역과 직역으로 제공할 수 있다. 다음은 두 번역의 특징이다.

2) 기본형(citation form)이란 농인에게 특정 수어를 물었을 때 제시하는 수어의 형태를 말하는데, 이때의 수어는 무표적인 형태이다.

① 의역(Free Translation)

- 목표어인 한국어로서 가장 자연스러운 형태로 제공한다.
- 출발어인 수어와 목표어인 한국어의 문장 길이가 일치하지 않을 수 있다.
- 수어의 지시 표현([지시])은 이해하기 쉽도록 한국어에서 가장 자연스러운 표현으로 번역한다.
- 일반적으로 수어 담화의 내용을 이해하는 데 도움이 되는 번역을 제공한다.

② 직역(Narrow Translation)

- 출발어인 수어의 절과 문장 수준으로 잘라 번역함으로 비교적 짧은 절 또는 문장으로 번역된다.
- 목표어인 한국어로 매끄럽게 번역되지 않을 수 있다.
- 수어의 지시 표현([지시])은 대명사를 사용하여 직접적으로 번역한다.
- 수어 담화 구조를 이해하는 데 도움이 되며 기계번역 알고리즘을 학습하고 테스트하는 데 유용하다.

본 말뭉치에서 채택하고 있는 번역 방식은 의역이다.

2. 주석 총렬

한국수어 말뭉치는 모두 두 명의 언어제공자가 서로 마주보며 대화하

는 형태로 구성되어 있다. 화면에 보이는 왼쪽 수어 화자는 A, 오른쪽 수어 화자는 B로 한다. 주석을 하기 위해 목적에 따라 층렬을 추가할 수 있다. 현재까지 만들어진 층렬은 다음과 같다.

2.1. 번역 층렬

지역의 원시 말뭉치를 수집하는 단계에서 번역 층렬은 향후 주석 작업자의 지역 수어 접근성을 높이기 위해 말뭉치 수집 지역의 수어 전문가(농인, 수어통역사 등)를 활용하여 기록한다. 즉 한국수어 말뭉치 수집팀은 촬영, 편집, 탑재 및 번역 작업을 수행하여야 한다.

번역 층렬은 언어제공자당 한 개의 층렬로 구성되며 「한국수어 자료 통합지원 시스템」에서 주석 파일과 영상을 내려받기하여 기록하고 저장된 파일은 다시 시스템에 등록하여야 한다.

2.2. 수지 글로스 층렬

수지 글로스 층렬은 우세손 수지 글로스 층렬과 비우세손 수지 글로스 층렬로 나눈다. 층렬의 이름은 <표1>과 같다.

수지 글로스 층렬에는 해당 수어의 글로스를 입력한다.

<표1> 한 개의 수어 영상 자료의 수지 글로스 층렬 예시

언어제공자	구분	층렬의 이름
수어 화자 A (화면에 보이는 왼쪽)	우세손	우세GlossA
	비우세손	비우세GlossA
수어 화자 B (화면에 보이는 오른쪽)	우세손	우세GlossB
	비우세손	비우세lossB

2.3. 의미 층렬

의미 층렬은 수지 글로스 층렬 바로 밑에 두고 수지 글로스 층렬과 동일하게 분절한다. 의미 층렬의 기록은 해당 수어의 맥락 속 의미를 입력하되, 한국어 표현이 다양한 경우 여러 개를 입력할 수 있다.

〈표2〉 수어 화자 A의 의미 층렬 예시

언어제공자	구분		층렬의 이름
	손	층렬	
수어 화자 A (화면에 보이는 왼쪽)	우세손	수지	우세GlossA
		의미	의미_우세A
	비우세손	수지	비우세GlossA
		의미	의미_비우세A

2.4. 코멘트 층렬

코멘트 층렬은 주석 시에 발견되는 특이사항 등을 자유롭게 텍스트로 입력할 수 있다. 단, 주석 목적에 따라 특정 내용을 기록하기로 약속하고 입력할 수 있다. 예를 들어 수지 글로스 층렬을 주석할 때 ‘음운변동’, ‘비우세손의 유지 현상’ 등이 나타나는 경우 이를 코멘트 층렬에 입력하였다.

2.5. 비수지 층렬

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

3. 분절

3.1. 분절 원칙

수어 단어는 다음과 같이 분절한다.

- 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 수향이 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.
- ⑥ 양손 수어에서 두 손의 시작과 끝이 다를 경우 별도로 수어 영상을 확인하여 정확히 분절한다.

예) 우세손과 비우세손의 구현된 [공부1]을 표현한 후, 비우세손이 약간 내린 상태로 유지하고, 우세손은 다른 단어 [가다1]할 때 그 유지된 비우세손은 잘리지 않고 완전히 내릴 때까지 기다리다가 분절한다.

- ⑦ 양손 수어의 경우 두 손의 분절이 일치하지 않을 수 있다. 즉, 비우세손이 이전의 수어를 유지하고, 우세손이 다음 수어로 바뀔 때는 따로 분절한 후, 비우세손 코멘트 층렬에 ‘유지’ 라고 입력한다.

 <p>준비하기(P) 치기(S) 유지하기(H) 내리기(R)</p> <p>제스처 이동 단계의 예 ‘한국수어 [가족]’</p> <p>※ 출처: De Vos, C.,Torreira, F. & Levinson, S C, 2015</p>	 <p>QR코드</p> <p>분절 기준에 대한 설명 영상</p>
--	--

4. 글로스

글로스는 해당 수어에 이름을 붙이는 것으로 같은 형태의 수어는 동일한 이름을 갖게 된다. 이름은 일련번호를 붙일 수도 있으나 해당 수어의 형태를 빠르게 떠올리기 위해 수어의 핵심 의미를 바탕으로 짓는다. 이러한 글로스는 수어의 의미도 번역도 아니다. 따라서 글로스를 부여하기 위해서 규칙에 대한 명확한 이해가 필요하다.

글로스를 부여할 때는 다음의 원칙에 따른다.

첫째, 글로스는 한글로 표기한다. 일반적으로 하나의 한국어 단어는 표준국어대사전에 등재된 어휘를 기본으로 한다.

둘째, 글로스는 가능한 한 해당 수어의 형태가 떠오를 수 있도록 핵심 의미를 기반으로 하되 익숙하고 알기 쉬운 표현으로 선택한다.

셋째, 동일한 수어 형태는 항상 같은 글로스를 가진다. 즉 맥락 안에서 의미가 달라진다고 하더라도 글로스는 항상 같다.

넷째, 핵심 의미가 같고 형태가 다른 수어를 구분하기 위하여 한국어 단어 뒤에 번호를 붙인 ID 글로스를 사용한다.

다섯째, 몸과 얼굴(비수지표지)은 종종 추가적인 의미 또는 전혀 다른 의미를 표현하기도 하지만, 수지 층렬 글로스 주석에서는 손의 움직임

정보만 기록한다. 예를 들어, 수어 화자가 머리를 좌우로 가로짓는 움직임과 함께 수지 기호를 표현한다면, 수지 기호에 대한 주석만 하고 부정 표현은 기록하지 않는다. 이는 글로스가 실제 번역이 아니라 단지 특정 수어에 대한 이름일 뿐이기 때문이다.

한국수어 말뭉치 구축에서는 상기 글로스 부여 원칙에 따라 다음과 같은 방법으로 글로스를 확정하였다.

- ① 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 → [꼼꼼1]

- ② 영어는 한글로 표기한다.

예) ‘엘지’ → [엘지1]

[LG1] × (잘못된 글로스 표기)

- ③ 문법적 표시는 가장 중립적인 것, 즉 무표적인 것으로 한다. 동사나 형용사의 경우 기본형으로 글로스를 붙인다.

예) 잡다, 잡는다, 잡았다, 잡겠다 → [잡다1]

예) 예쁜, 예쁘고, 예뻤다 → [예쁘다1]

- ④ 글로스는 하나의 단어로 하는 것을 원칙으로 한다. 단, 하나의 단어로 의미가 불충분한 경우 조사와 띄어쓰기 없이 붙여 쓴다.

예)



[버스가다1]

- ⑤ 수어의 형태가 다르고 의미는 같은 경우 한국어 단어 뒤에 숫자

1, 2, 3, 등을 붙인 글로스를 한다.

예)



[없다1]



[없다2]



[없다3]

⑥ 글로스로 사용한 한국어 단어가 동음이의어인 경우는 타입의 혼동을 방지하기 위해 다음과 같이 구분하여 글로스를 부여한다.

예) [예시1]

- ‘예’의 경우 ‘윗사람의 부름에 대답하거나 묻는 말에 긍정하여 대답할 때 쓰는 말’과 혼동을 방지하기 위하여 **같은 의미의 다른 어휘인 [예시1]**로 글로스를 정한다.

예) [값싸다1]

- ‘싸다’의 경우 ‘물건을 싸다’, ‘값이 싸다’ 등 다른 의미가 있기 때문에 이에 혼동을 방지하기 위하여 앞에 ‘값’을 붙여 [값싸다1]로 글로스를 정한다.

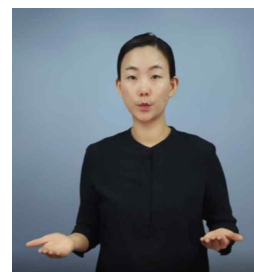
예) [일하다1]

- ‘일’의 경우 ‘자연수의 맨 처음 수’와 혼동을 방지하기 위하여 [일하다1]로 글로스를 정한다.

[예시1]



[값싸다1]



[일하다1]

⑦ 2015년 처음 주석에서는 각 토큰에 대해서 형태론적 변형과 음운

변이를 기록하여 구분하였으나, 그 후에는 상세 기록을 하지 않았다. 향후 연구 목적에 따라 형태론적인 변형과 음운 변이를 주석할 수 있다.

※ 2015년 글로스 예시

- ① [주다1]: 글로스 기본형
- ② [주다11]: [주다1]의 형태론적 변형([주다1]에서 주체와 대상이 바뀐 경우)
- ③ [주다11B]: [주다11]의 음운적 변이([주다11]에서 수형이 바뀐 경우)

4.1. 고정된 수어

사전에 등재되었거나 형태가 고정되어 어휘화된 것을 고정된 수어라고 한다. 고정된 수어의 글로스 예시는 다음과 같다.

- ① 수 또는 단위와 관련된 수어의 글로스는 다음과 같이 한다.
 - 수를 나타내는 수어는 한국어 단어로 글로스를 붙인다. 즉 숫자 ‘1’ 과 ‘2’ 는 각각 [일1]과 [이1]로 글로스를 붙인다. 숫자 ‘10’ 의 두 형태는 각각 [십1] 또는 [십2] 등으로 글로스를 붙인다.



[십1]



[십2]

- ‘첫 번째’, ‘두 번째’ 와 같은 서수는 [첫번째1], [두번째1] 등으로 글로스를 붙인다.

- ‘1년’, ‘2분’ 과 같이 숫자와 단위가 혼합된 수어의 경우는 다음과 같이 글로스를 붙인다.

예) 1년 → [일년1]

예) 2분 → [이분1]

- 알파벳 지문자로 나타내는 단위 수어는 한국어로 글로스를 붙인다.

예) ‘mm’ → [밀리미터1]

- 한 손에 나타나는 ‘counting hand(수를 세는 손)’는 펼쳐진 손가락의 수와 상관없이 모두 [counting hand]로 글로스를 붙인다. 한 손 ‘counting hand(수를 세는 손)’와 함께 나타나는 다른 손 수어는 해당 수어의 글로스를 붙인다.



[counting hand 딸]



[counting hand 남동생]

- ② 형태가 같은 수어에서 경로 수동의 여부에 따라 다음과 같이 글로스를 한다.

예) 동일한 수어 형태에서 [승용차1]이 경로 수동이 있는 경우는 [승용차가다1]로 글로스를 붙인다.



[승용차1]



[승용차가다1]

- ③ 양손을 사용하는 일치동사는 비우세손이 나타나지 않아도 되는 경우에는 다음과 같이 우세손과 비우세손 글로스를 따로 붙인다.

예) 수어 의미 ‘흔내다’의 수지 글로스 예시



우 세 손: [야단치다1]

비우세손: [사람4]

4.2. 생산적 수어

생산적 수어는 관습적인 의미를 가지지 않고 맥락에 따라 자연스럽게 생성되는 수어로 맥락을 바탕으로 이해되는 수어이다. 그러므로 한 단어로 번역 또는 한 단어로 글로스를 정하기 어렵다.

능숙한 수어 화자들 사이에서 이해가 가능한 이 수어들은 단순히 사물을 수형이나 수동으로 나타내는 것이 아니라, 생산적 수어에서 나타나는 체계적인 규칙이 적용된다.

4.2.1. 생산적 수어의 주석 원칙

생산적 수어를 주석을 할 때는 수형을 기준으로 글로스(타입명)를 정하

며 수형이 변화하는 경우 시작 수형을 기준으로 정한다. ‘생산적 수어’의 글로스는 수형의 변별자질에 따라 결정된다.

첫째, 선택된 손가락에 따라 일형, 이형, 육형, 구형 등의 지숫자를 나타내는 수형으로 글로스를 붙인다.


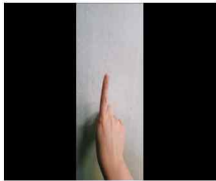
둘째, 선택된 손가락의 벌림 유무에 따라 글로스는 달라진다. 지숫자형 이름 뒤에 숫자 1, 2, 3을 붙이며 타입명을 구분한다. 이때 손가락 벌림의 정도는 변이로 보아 별도로 구분하지 않는다.

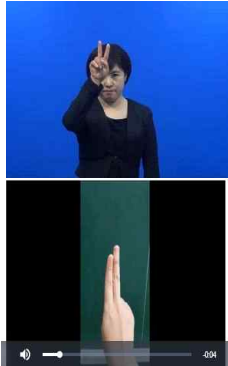




셋째, 선택된 손가락의 자세 즉, 손가락 관절의 구부림 유무에 따라 지숫자형 이름 뒤에 숫자 1, 2, 3을 붙이며 타입명을 구분한다. 이때 손가락 구부린 정도는 변이로 본다.



넷째, 선택된 손가락의 접촉 시 손의 모양이 평평하거나 원형인 경우는 다른 수형인 것으로 판단하여 타입명을 구분한다.

단, 이러한 변별자질을 고려할 때, ‘장향’, ‘손목의 구부림’까지는 고려하지 않는다.

다음은 한국수어 말뭉치 구축에서 나타난 생산적 수어의 글로스(타입)이다.

연번	타입명	그림	설명
1	일형2		<ul style="list-style-type: none"> 일형 수형의 고리 형태 ※ [고리2]와 같지만 [고리2]는 고정된 수어이고, [일형2]는 생산적 수어라서 문제 없음.
2	일형3		<ul style="list-style-type: none"> 일형 수형을 편 형태 양손 일형 수형으로 네모를 그리거나 한손 일형 수형으로 세모 혹은 원을 그리는 형태도 해당됨.

연번	타입명	그림	설명
3	이형1		<ul style="list-style-type: none"> ▪ 접촉이 없는 이형 수형 ▪ 접촉이 있는 이형 수형
4	육형1		<ul style="list-style-type: none"> ▪ 양손 육형 수형으로 초승달과 같은 모양을 표현하는 것 ▪ 육형 수형으로 네모를 표현하는 것과 양손 육형 수형을 붙여 세모 모양으로 표현하는 것이 해당됨.
5	육형2		<ul style="list-style-type: none"> ▪ 육형 수형이 벌어지고 줄어드는 것 ▪ 구부린 양손 육형 수형으로 네모를 표현하는 것 해당됨
6	육형3		<ul style="list-style-type: none"> ▪ 구부린 육형 수형
7	구형1		<ul style="list-style-type: none"> ▪ 구형 수형 ▪ 구형 수형으로 선이나 면으로 그리거나 양손으로 세모 그리거나 혹은 손가락을 벌린 구형으로 그리거나, 손가락을 벌린 구형이 선을 나타내는 것 해당됨
8	구형2		<ul style="list-style-type: none"> ▪ 구부린 구형 수형으로 그리는 것 ▪ ‘장소형’ 수형의 양손으로 공, 덩어리 등을 나타내는 그 수형 포함됨
9	구형3		<ul style="list-style-type: none"> ▪ 구형 수형을 평평하게 접촉하여 움직이는 형태

연번	타입명	그림	설명
10	구형4		<ul style="list-style-type: none"> ▪ 구형 수형을 원형으로 접촉하여 움직이는 형태 ▪ 모아진 구형 수형으로 움직이는 형태
11	십형1		<ul style="list-style-type: none"> ▪ 십형 수형으로 붓을 그리는 형태

4.2.2. 생산적 수어의 도상성

생산적 수어 어휘(productive signs)는 의미를 전달하기 위해 사물과 상황, 행위를 묘사한다. 수어 사용자는 이미지 형성 기법들을 사용해 기저 이미지를 환기시키며 맥락과 어떤 기법인지 확인하는 것은 생산적 수어 표현의 해석에 도움이 된다(Koenig, Konrad & Langer, 2008).

수어 사용자는 여러 형태적 특징 중에서 표현하고자 하는 지시체의 가장 핵심적이고 시각적인 특성을 선택하여 조어하는데, 이미지 형성 기법을 알면 생산적 수어 어휘를 적절하게 산출하는 데 도움이 될 뿐만 아니라, 생산적 수어 어휘의 의미를 이해하기 위해 기저 이미지를 인식하는 데에도 도움이 된다. 이미지 형성 기법들은 고정된 수어 어휘뿐만 아니라 생산적인 수어 어휘에도 적용된다. 2019년 한국수어 말뚝치 사례 연구에서 생산적 수어의 도상성을 분석하기 위해 활용된 이미지 형성 기법은 다음과 같다. 한국수어 말뚝치 연구에서는 독일 함부르크 대학교의 독일수어 연구소에서 제안한 이미지 형성 기법 중 손대체 기법, 손이용 기법, 스케치 기법, 측정 기법 총 4개의 기법만 활용되었다.

① 대체 기법(Substitutive technique)

- 손이 지시 대상 혹은 대상 물체의 주요 부분을 상징하는 기법
- 수어 공간 내의 수위, 수향, 수동은 대상 물체의 공간적 위치, 방향, 움직임에 상응함

예) [육형3]

[구형1]과 [주먹형1]



손대체 기법

(우) 육형3 (비) 육형3

물체를 상하좌우로 옮기는 크레인 같은
기계의 고리로 물건을 집은 형태



손대체 기법

(우)주먹형1 (비) 구형1

컨테이너, 선적물, 화물

② 손이용 기법(Manipulative technique)

- 수어 사용자의 손이 만지거나, 잡거나 혹은 대상을 사용하는 행위를 표현하는 기법

예) 손이 칼을 사용하는 모습을 보여 주거나 손으로 부채를 잡고 부채질하는 행위

- 손이 특정 위치에 가만히 있는 모습도 포함됨

예) [구형3]

[주먹형]



손이용 기법
(우) 구형3 (비) 구형3
이불을 끌어당겨 잡은 손



손이용 기법
(두손) 주먹형1
아이의 손을 잡은 손

③ 스케치 기법(Sketching technique)

- 손이 연필 혹은 붓 같은 스케치 도구이며, 수동을 통해서 사물의 모양을 그리는 기법
- 수동은 표현하고자 하는 사물의 움직임이 아니라 움직임이 없는 사물을 그리는 것임
- 사물의 모양을 수어 공간에 선으로 표현함

예) [구형2]

[십형1]



스케치 기법
(우) 구형2 (비) 구형2
구멍보트 모양을 그리는 손



스케치 기법
(우) 십형1 (비) 십형1
백화점 큰크리트 기둥 또는 원기둥이
군데군데 세워진 모양

④ 측정 기법(Measuring technique)

- 손이나 손가락이 사물의 크기나 길이를 표현하는 기법
- 손이나 손가락이 지시 대상의 크기와 그 경계선이나 테두리를 보여줌
- 측정 기법의 수형은 대부분 무표 수형(unmarked hand shapes)으로 표현 (예: 펼친 수형, 검지 수형).
- 표현하고자 하는 사물의 모양을 보여주지 않고 그 사물의 크기나 길이를 나타냄

예) [구형1]



측정 기법
(우) 구형1 (비) 구형1
맥락정보 : [버스1] [구형1] [무겁다1] [뭇1]
[구형1] 의미 : 버스(차)의 덩치가 큼

4.3. 결합 수어(합성어)

둘 이상의 수어가 순차적으로 결합된 형태의 경우, 하나의 단어인 합성어는 하나의 글로스로, 구로 나타나는 형태는 각 단어에 글로스를 붙인다. 구체적인 예시는 다음과 같다.

- ① 둘 이상의 수어가 결합된 형태로 명확하게 하나의 단어로 인지되는 경우에는 하나의 글로스로 주석한다. 예를 들어 ‘손’ + ‘돈’으로 된 수어는 두 수어를 묶어 분절한 후에 [수당1]로 글로스를 붙인다.
- ② 둘 이상의 수어가 결합된 형태에서 실제로 하나의 단어로 합성된 형태인지 의심스러운 경우는 두 개의 수어로 분리하여 각각 글로스를 붙인다. 예를 들어 ‘분홍’이라는 의미를 나타내는 수어 중 [빨강]+[약하다]로 된 수어는 [빨강1], [약하다1]로 글로스를 붙인다.
- ③ 둘 이상의 수어가 결합하여 하나의 합성어가 된 어휘는 다음과 같은 특성을 가진다.
 - 두 수어의 조합이 특정한 의미를 갖는다.
 - 조합된 두 수어 사이에 다른 수어나 신호를 삽입할 수 없다.
 - 두 수어의 결합 시에 음운론적 특성은 다음과 같다.

예) [알려주다1]

- [알다1]과 [주다1]이라는 두 수어가 결합 될 때 [알다1]의 반복 동작이 생략되고 두 수어를 연결하는 움직임(M)이 첨가되면서 수어 화자들은 하나의 단어로 인식함으로 합성어 [알려주다1]이 된다.

예) [안산1]

- [편안하다1]이라는 수어와 [산1]이라는 수어가 결합되면서

[편안하다1]에서 마지막 정지(H)가 탈락하면서 [산1]이라는 수어 사이에 연결 움직임(M)이 첨가되면서 합성어 [안산1]이 된다.

예) [만약1]

- [거짓1]과 [그러면1]이라는 수어가 결합되면서 앞 수어가 뒤 수어의 수형에 동화되어 [거짓1]의 수형이 [그러면1]의 수형과 같아져 합성어 [만약1]이 된다.

4.4. 기타

① 수어 이름

- 수어 이름의 경우 대상자의 이름 앞에 [SN]을 붙인다. 수어 이름이지만 그 사람의 한국어 이름을 모르는 경우는 [SN]으로만 글로스를 붙인다.

② 지역명

- 지역명의 경우 동일한 수어의 형태가 타입으로 존재하더라도 [지역명]으로 글로스를 붙인다.

예) ‘옥수수’와 ‘강원도’



[옥수수1]



[강원도1]

③ 외국수어/국제수화

- 수어 화자가 미국수어를 하는 경우 [미국수화]로 입력하고, 국제수어를 할 경우 [국제수화]로 입력한다. 그리고 그 수어의 의미는 코멘트 창렬에 입력한다.

④ 지시 수어

- 지시 수어는 수형에 따라 두 종류로 글로스를 정한다. ‘1지’만 편 수형은 [지시1], 손가락을 모두 편 수형은 [지시2]로 한다.
- 자세한 지시 글로스는 다음과 같다.

연번	타입명	그림	설명
1	지시1		<ul style="list-style-type: none"> ▪ 1지만 편 수형으로 지시하는 형태
2	지시1호		<ul style="list-style-type: none"> ▪ 1지만 편 수형으로 반원 또는 원을 그리는 형태 ▪ ‘반원’의 의미: 너희(의미창렬) ▪ ‘원’의 의미: 우리(의미창렬)
4	지시1_1		<ul style="list-style-type: none"> ▪ 1지만 편 수형으로 1인칭 지시
5	지시1_1 A		<ul style="list-style-type: none"> ▪ 1지와 5지를 편 수형으로 1인칭 지시
7	지시1_1 C		<ul style="list-style-type: none"> ▪ 5지만 편 수형 1인칭 지시

연번	타입명	그림	설명
11	지시2		<ul style="list-style-type: none"> 손가락을 모두 편 수형으로 지시하는 형태
12	지시2호		<ul style="list-style-type: none"> 손가락을 모두 편 수형(손바닥이 위로, 아래로)으로 반원 또는 원을 그리며 지시
13	지시2_1		<ul style="list-style-type: none"> 손가락을 모두 편 수형으로 1인칭 지시
14	지시2_1A		<ul style="list-style-type: none"> 손가락을 모두 편 수형에서 손가락 끝으로 1인칭을 지시
15	지시2_1B		<ul style="list-style-type: none"> 손가락을 모두 편 수형의 손가락을 모아 1인칭을 지시
	지시2그림		<ul style="list-style-type: none"> 이형, 구형이 형태를 그리는 것 이형과 구형이 섞여 있음

⑤ 제스처

- 다음과 같은 기준에 의해 수어와 제스처를 구분한다.

① 수어가 아니다.

- ③ 수어를 모르는 사람이 봐도 그 의미를 이해할 수 있다.
- ④ 수어를 모르는 청인도 사용하는 몸짓이다.
- ⑤ 제스처는 수지 글로스 층렬에 [제스처]로 입력하고, 코멘트 층렬에 구체적인 내용을 설명한다.

⑥ 지문자

- 지문자는 [지문자1]로 글로스로 붙인다.
- 지문자는 자음과 모음을 분리하여 코멘트 층렬에 기록한다.

예) 수지 글로스 층렬: [지문자1]

코멘트 층렬: [ㅎㅏㄱㄱ교]

⑦ 음운변동

- 수어가 앞뒤 수어의 영향을 받아 음운변동이 나타난 경우에는 수어의 원형에 해당하는 글로스를 달고 코멘트 층렬에 ‘음운변동’으로 주석을 단다.

⑧ 토큰으로 처리되지 않는 수어(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 따라 하는 것과 같은 것은 실제 수어의 사용 빈도에 포함시켜서는 안 된다. 따라서 이러한 수어는 토큰으로 처리하지 않으며 비토큰(non-token)이라 한다.
- 메타 언어적 언급은 토큰으로 처리하지 않는다.
- 손으로는 수어를 하면서 비수지로 부정을 표현하는 수어는 토큰으로 처리한다.

큰으로 처리하지 않는다. 예를 들면 어떤 ‘특정 수어’를 하면서 이렇게 수어를 하지 않는다는 의미로 머리를 흔들면 이때의 ‘특정 수어’는 토큰이 아니다.

- 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.

⑨ 그 외

- [%]: 수어가 완전히 조음되지 않은 경우 글로스 대신에 [%]로 주석을 입력하여 거짓 시작임을 나타낸다.
- [###]: 전사자가 수어로 인지했지만 무슨 수어인지 모르는 경우는 글로스 대신에 [###]로 주석을 단다. 그 후에 온라인 문의 공간에 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. 수어가 영상화면 바깥에서 조음이 된다고는 보이지 않는 경우에도 [###]로 입력한다.
- [@]: 수어 화자가 실수로 수어를 잘못 표현했고, 그 수어가 사용되지 않는 수어일 때 [@]로 입력한다.

5. 비수지 주석

한국수어 말뭉치 주석에는 ‘손 움직임’ 없이 나타나는 비수지 표지는 기록하지 않았다. 다만, 다음의 경우에만 코멘트 층렬에 입력을 하

였다. 수어 없이 긍정의 의미로 고개를 끄덕인 경우 코멘트 층렬에 ‘끄덕임’으로, 수어 없이 부정의 의미로 고개를 흔드는 경우 ‘가로저음’으로 입력하였다.

2016년 한국수어 말뭉치를 활용한 사례연구를 위해 ‘입 움직임’과 관련된 주석을 실시하였으며 이와 관련된 규약은 아래와 같다.

5.1. 입 움직임

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 음성언어로부터 온 것으로 단어 전체 또는 한 부분을 조음하는 것처럼 보이는데 이 입의 움직임은 해당 수어 단어의 앞 또는 뒤까지 지속되어 나타나기도 한다. 그러나 마우스 제스처는 음성언어의 조음과는 관련이 없는 입의 움직임이다(Crasborn, 2016).

5.1.1. 입 움직임 주석의 기본 원칙

입 움직임을 주석할 때 기본 원칙은 다음과 같다.

- ① 마우스 제스처는 타입을 추가하며, 마우딩은 별도의 타입이 없다.
- ② 마우스 제스처 관련 전사는 2개의 층렬 즉, 마우스 제스처 글로스와 마우스 제스처 분류(class) 층렬로 나눈다.
- ③ 마우딩 관련 전사는 4개의 층렬로, 마우딩을 기입하는 층렬 외에 스프레딩, 마우딩의 완성도, 마우딩의 분류 층렬로 나눈다.

5.1.2. 입 움직임 층렬

총 8개의 층렬이 2명의 언어제공자에게 각각 할당된다. 층렬의 이름은 띄어쓰기 없이 입력하며 추가된 층렬의 속성은 다음과 같다.

- ① Mouthing: 한국어 단어를 조음하는 것처럼 보이는 입의 움직임을 해당 단어의 철자법에 따라 한글로 입력하는 층렬이다.
- ② Spreading: 마우딩이 하나의 수어를 넘어서 얼마나 지속되는지를 입력하는 층렬이다.
- ③ Mouthing완성도: 마우딩으로 나타난 구어 단어의 완성도를 입력하는 층렬이다.
- ④ Mouthing분류: 마우딩을 특성에 따라 분류하고 이를 입력하는 층렬이다.
- ⑤ MouthGesture: 구어 단어를 발음하는 것 외의 모든 입의 움직임에 대한 글로스를 입력하는 층렬이다.
- ⑥ MouthGesture분류: 마우스 제스처를 특성에 따라 분류하고 이를 입력하는 층렬이다.
- ⑦ 수어분류: 마우딩 또는 마우스 제스처와 함께 나타난 수어의 품사 중에 명사와 동사를 구분하여 입력하는 층렬이다.
- ⑧ Mcoment: 입의 움직임과 관련한 주석 시에 생성되었으나 층렬에 담을 수 없는 다양한 의견 등을 자유롭게 입력하는 층렬이다.

5.1.3. 입 움직임 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 층렬은 해당 마우딩 또는 마우스 제스처의 분절 프레임과 동일하게 분절한다. 즉, 스프레딩, 마우딩 완성도, 마우딩 분류 층렬은 마우딩 층렬 분절과 마우스 제스처 분류는 마우스 제스처 층렬의 분절과 동일하게 분절한다. M코멘트 층렬 역시 해당 마우딩 또는 마우스 제스처 분절과 동일하게 분절한다.

5.1.4. 마우스 제스처 주석

(1) 마우스 제스처 기록

- ① 마우딩 이외의 모든 입의 움직임은 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.
- ③ 입의 모양은 기본 기호를 조합하여 입력한다.
 - ‘입술을 내밀고 있는데 이가 보이는 경우’ 는 [BH]
 - ‘입을 다 물고 양쪽 입꼬리가 약간 내려간 경우’ 는 [-[^]l]
 - ‘이가 보이는 상태로 왼쪽 입꼬리만 올라 간 경우’ 는 [KH]
 - ‘입이 약간 벌어지고 혀가 조금 나온 경우’ 는 [O1U1]
 - ‘혀를 여러 번 내미는 경우’ 는 [U1~]
- ④ 연속적으로 일어나는 입의 움직임은 움직임의 순서대로 기호를 조합하여 입력한다.
 - ‘입을 약간 벌리고 혀를 내밀었다가 집어넣으면서 바람을 내뿜는 경우’ 는 [O1UJ(=)]

- ‘입을 다물었다가 이가 보이게 양쪽으로 벌리면서 바람을 내뿜는 경우’ 는 [-H1(=)]

(2) 마우스 제스처 특성

- ① 마우스 제스처의 특성에 따른 분류를 입력하는 층렬이다.
- 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 움직임만으로는 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 투영(Eco)시키는 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 생성하기도 한다.
 - 4: 입이 실제 입의 역할을 하는 경우이다. 예를 들면 [먹다] [씹다] [외치다]와 같은 수어를 하면서 실제 먹거나 씹거나 외치는 입 모양을 만드는 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 수어를 하면서 만드는 입의 움직임은 얼굴 전체의 움직임과 함께 나타나므로 이 분류에 속한다.

5.1.5. 마우딩 주석

(1) 마우딩 기록

- ① 구어의 단어를 조음하는 대로 입력하되, 조음점이 보이지 않는 경우는 마우딩을 하면서 함께 나타나는 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. 예) [일본]이라는 수어를 하면서 입의 움직임은 ‘일반’으로 나타난 경우, [일반]으로 입력하고 M코멘트(Mcoment) 층렬에 [일본]이라고 입력한다.

(2) 스프레딩(Spreading) 특성

- ① 마우딩은 지속 시간이 길어 하나 이상의 수어에서 나타나기도 한다. 이러한 현상을 스프레딩이라고 하며 마우딩을 기록할

때는 이러한 스프레딩 현상을 확인하기 위하여 별도의 층렬을 만들어서 기록하여야 한다.

② 마우딩이 지속되는 시간을 분절한 후에 다음의 특성에 따라 오른쪽, 왼쪽, 양쪽, 일치, 부분 등으로 스프레딩을 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 10프레임 이상 혹은 다음 수어 지속 시간의 50%이상까지 지속될 때이다.

(3) 마우딩 완성도 특성

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

5.2. 기타 (시선, 몸, 어깨...)

향후 시선, 몸, 어깨 등 비수지 층렬에 대한 주석 작업이 이루어지면 추가되어갈 것이다.

6. 의미 주석

2015년부터 실시된 한국수어 말뭉치는 수지 글로스 층렬을 중심으로 기록되어 왔다. 글로스는 수어 어휘의 형태를 한국어를 빌려 붙여지나 한국어와 일대일 등가 관계가 성립하지는 않는다. 이에 맥락에 따라 다양한 의미로 사용되는 각 수어 어휘에 대한 글로스의 의미를 파악하기 위하여 2019년부터 수지 글로스 층렬 아래에 의미 층렬을 두어 담화 맥락 안에서 한국수어가 나타내는 다양한 의미를 기록하였다. 이렇게 기록된 의미 정보들은 수어사전 제작 등에 활용할 수 있다.

6.1. 의미 주석의 기본 원칙

첫째, 의미 주석은 [지시] 토큰을 제외한 모든 토큰에 해당 의미를 기록하는 것이며, 수지 글로스 층렬의 분절 크기에 맞추어 의미 층렬(<표2> 참조)에 주석을 한다.

둘째, 의미 주석은 해당 토큰의 문장 맥락 안에서의 의미를 기술한다. 예를 들면, 수지 글로스 층렬에 입력된 [소개1]은 맥락에 따라 ‘사회자’ 또는 ‘통역’으로 기록된다.

셋째, 비수지 표지는 조음소로 작용하는 것을 대상으로 하고 문법적 기능을 하는 비수지는 제외한 의미를 기술한다.

넷째, 대역어 형태로 기술하되 맥락 안에서 같은 의미의 대역어는 함께 기술한다.

다섯째, 해당 토큰의 의미가 동사나 문장의 형태인 경우 기본형으로 기술한다.

[부록 3] 촬영 동의서

국가 언어 자원(말뭉치) 구축 및 활용 저작권 등 이용 허락 계약서

권리자 _____(이하 “권리자”이라 함)와 이용자 국립국어원(이하 “이용자”이라 함)은 아래 저작물에 관한 저작권재산권과 권리자의 초상권, 기타 이에 부수한 권리 등(이하 ‘계약 대상 권리’라고만 한다) 일체의 권리에 대한 이용 허락과 관련하여 다음과 같이 계약을 체결한다.

다 음

제1조 (계약의 목적)

본 계약은 계약 대상 권리의 원만한 이용 허락과 관련하여 권리자와 이용자 사이의 권리관계를 정하는 것을 목적으로 한다.

제2조 (계약의 대상)

① 본 계약의 대상이 되는 저작권재산권은 다음 각 항에서 정하는 바와 같다.

1. 대상저작물은 한국수어 말뭉치로 한다(이하 ‘대상저작물’이라고만 한다)

2. 이용 대상 권리는 다음 각 목에서 정하는 일에 필요한 일체의 저작재산권과 저작인접권 중 재산권 부분을 말한다.

가. 이용자 및 이용자가 발주한 용역 사업의 수행자가 대상저작물을 일정한 형식으로 전자적 기록 매체에 담아 보존하는 일

나. 이용자 및 이용자가 발주한 용역 사업의 수행자가 한국수어 연구와 언어 정보 처리 분야에 응용하기 위해 대상저작물을 복제·변형(목차·머리말·도표·그림·각주 등의 편집 및 삭제, 언어 단위별 분리, 언어적·비언어적 정보 부착 등)하는 일

다. 이용자 및 이용자가 발주한 용역 사업의 수행자가 대상저작물 및 그 복제·변형물을 연구 및 기술 개발용으로 학계·연구기관·산업체 등이 이용할 수 있도록 제공·배포하는 일

라. 대상저작물 및 그 복제·변형물을 제공·배포받은 학계·연구기관·산업체 등이 한국수어 연구와 언어 정보 처리 분야 응용을 위하여 대상저작물 및 그 복제·변형물을 분석 및 처리하여 사용하는 것을 허락하는 일

- ② 본 계약의 대상이 되는 초상권은 다음 각 항에서 정하는 바와 같다.
가. 초상권의 대상은 제1호 가목에 기재된 저작물에 포함된 권리자의 초상 전체로 한다.
나. 이용 대상 권리는 제1호 나목에 기재된 일체의 행위에 필요한 초상권 및 퍼블리시티권으로 한다.
- ③ 기타 본 계약의 원활한 이행을 위하여 제1항 또는 제2항에서 정한 권리의 행사를 위하여 권리자는 이용자 또는 이용자가 허락한 자에 대하여 저작인격권(저작인접권에 의한 인격권 부분을 포함한다) 및 기타 법적 권리를 부당하게 행사하지 아니하기로 한다.

제3조 (이용 허락 기간)

- ① 계약 대상 권리에 대한 이용 허락 기간은 계약체결일부터 2051년 12월 31일까지로 한다.
- ② 권리자가 이용 허락 중지 의사를 밝히지 아니하면 이용 허락이 5년 단위로 자동 갱신된다.
- ③ 계약기간 중 권리자의 사망이 발생하면 특별한 사정이 없는 한 그의 적법한 상속인이 이 계약의 지위를 상속하는 것으로 한다.

제4조 (권리자의 의무)

- ① 권리자는 이용자에게 계약 대상 권리에 관하여 본 계약서 제2조에서 정한 계약 대상 권리를 제3조의 기간 동안 비독점적으로 허락한다.
- ② 권리자는 이용자에게 계약 체결일로부터 10일 이내에 계약 대상 권리의 이용을 위해 필요한 상당한 자료를 인도하여야 한다. 다만, 제2조 제1항 제1호의 가목의 저작물이 한국저작권위원회에 등록되어 있지 않은 경우 이용자가 요청하면 권리자는 이 저작물의 저작재산권을 등록한 후 위 의무를 이행한다.
- ③ 권리자는 계약 대상 권리에 제3자의 이용 허락권, 질권 등이 존재하는 경우, 이용자에게 그 사실을 사전에 알려야 한다.
- ④ 권리자는 계약 대상 권리의 저작재산권 전부 또는 일부를 제3자에게 양도하거나 이에 대하여 질권을 설정하고자 하는 경우, 사전에 이용자에게 이 사실을 통보하여야 한다.

- ⑤ 제4항의 경우 권리자는 계약기간 중 계약 대상 권리 중 일부 또는 전부를 제3자에게 양도할 경우, 이에 대하여 사전에 이용자에게 통지하여야 하여 제3자와 이용자가 계약을 체결할 수 있도록 적극 협조하여야 한다.

제5조 (이용자의 권리 및 의무)

- ① 이용자는 계약 대상 권리를 제3조의 이용 허락 기간 동안 제2조의 이용 허락을 받은 범위 내에서 비독점적으로 자유롭게 이용할 수 있다.
- ② 계약 대상 권리의 이용료는 무상으로 한다.
- ③ 이용자는 관례적으로 저작자 및 저작재산권자의 성명 등 표시를 허용하는 대상저작물을 이용하는 경우, 그 저작자 및 저작재산권자의 성명 등을 표시하여야 한다.
- ④ 이용자는 대상저작물을 이용함에 있어서 저작인격권을 침해하지 아니한다. 다만, 대상저작물의 본질적인 내용을 변경하지 않는 범위 내에서 권리자에게 그 사실을 사전에 고지한 후 사소한 수정 및 편집을 할 수 있다.

제6조 (확인 및 보증)

- ① 권리자는 이용자에게 다음 각 호의 사항을 확인하고 보증한다.
 1. 대상저작물의 저작권 이용 허락을 체결하는 데 필요한 권리 및 권한을 적법하게 보유하고 있다는 것
 2. 대상저작물의 내용이 제3자의 저작권, 상표권, 인격권을 비롯한 일체의 권리를 침해하지 아니한다는 것
 3. 대상저작물에 대하여 이용자에게 사전에 알린 제3자의 권리 외에는 이용자의 이용을 제한할 수 있는 부담이 더 이상 존재하지 아니한다는 것
- ② 이용자는 권리자에게 다음 각 호의 사항을 확인하고 보증한다.
 1. 대상저작물에 적용된 이용 허락 조건에 의해서만 대상저작물의 재이용을 허락할 것
 2. 대상저작물을 권리자 및 제3자의 인격적 권리를 침해하는 방식으로 이용하지 아니할 것

제7조 (계약내용의 변경)

본 계약 내용 중 일부를 변경할 필요가 있는 경우에는 권리자와 이용자의 서면합의에 의하여 변경할 수 있으며, 그 서면합의에서 달리 정함이 없는

한, 변경된 사항은 그 다음날부터 효력을 가진다.

제8조 (계약의 해지)

- ① 당사자는 천재지변 또는 기타 불가항력으로 계약을 유지할 수 없는 경우에 본 계약을 해지할 수 있다.
- ② 당사자는 상대방이 정당한 이유 없이 본 계약을 위반하는 경우에 4주 이상의 상당한 기간을 정하여 상대방에게 그 시정을 최고하고, 상대방이 그 기간이 지나도록 이행하지 아니하는 경우에는 계약을 해지할 수 있다. 다만, 상대방이 명백한 시정 거부 의사 표시하였거나 위반 사항의 성격상 시정이 불가능하다는 것이 명백히 인정되는 경우에는 위와 같은 최고 없이 계약을 해지할 수 있다.
- ③ 본 계약에 대한 해지권의 행사는 상대방에 대한 손해배상청구권 행사에 영향을 미치지 아니한다.

제9조 (손해배상)

당사자가 정당한 이유 없이 본 계약을 위반하는 경우, 그로 인하여 상대방에게 발생한 모든 손해를 배상할 책임이 있다. 다만, 제8조 1항의 사유로 본 계약을 이행하지 못한 경우에는 손해배상책임을 면한다.

제10조 (비용의 부담)

계약 체결에 따른 비용은 이용자가 전부 부담한다.

제11조 (분쟁해결)

- ① 본 계약에서 발생하는 모든 분쟁은 권리와 이용자가 상호 원만한 합의에 이르도록 노력하여야 하며, 분쟁이 원만히 해결되지 않는 경우에는 소제기에 앞서 한국저작권위원회에 조정을 신청할 수 있다.
- ② 제1항에 따라 해결되지 아니할 때에는 서울남부지방법원을 제1심 소송의 전속관할로 정하여 소송에 의해 해결토록 한다.

제12조 (비밀유지)

양 당사자는 본 계약의 체결 및 이행과정에서 알게 된 상대방에 관한 정보, 본 계약의 내용 및 대상저작물의 내용을, 상대방의 서면에 의한 승낙 없이 제3자에게 공개하여서는 아니 된다.

제13조 (기타부속합의)

- ① 권리자와 이용자는 본 계약의 내용을 보충하거나, 이 계약에서 정하지 아니한 사항을 규정하기 위하여 부속합의서를 작성할 수 있다.
- ② 제1항에 따른 부속 합의는 본 계약의 내용과 배치되거나 위반하지 않는 범위 내에서 유효하다.

제14조 (계약의 해석 및 보완)

본 계약서에서 명시되어 있지 아니하거나 해석상 이견이 있을 경우에는 저작권법, 민법 등을 준용하고 사회 통념과 조리에 맞게 해결한다.

제15조 (계약 효력 발생일)

본 계약의 효력은 계약 체결일로부터 발생한다.

2020년 월 일

권리자

성명: (인)

생년월일:

주소:

이용자

성명: 국립국어원장 (인)

주소: 서울특별시 강서구 금남로

[부록 4] 언어제공자 질문지

자료 수집을 위한 질문지

개인 정보

1. 이름: _____
2. 생년월일: _____
3. 주소: _____
4. 성별: ☐ 여성 ☐ 남성
5. 현재 사는 곳: _____ 거주 기간(예: 08-현재): _____
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. 청각보조기기 착용 여부

☐ 인공와우 수술을 받았다. (_____년 이후부터 사용)
☐ 규칙적으로 보청기를 사용한다.
☐ 가끔 보청기를 사용한다.
☐ 보청기 등의 보장구를 사용하지 않는다.

가족 정보

9. 아버지의 농인 여부: ☐ 농인 ☐ 난청인 ☐ 청인 ☐ 잘 모르겠다

10. 어머니의 농인 여부: ☐ 농인 ☐ 난청인 ☐ 청인 ☐ 잘 모르겠다

11. 부모와의 동거 여부(어린시절): ☐ 예 ☐ 아니오

12. 부모 외의 양육자: [11번 질문에 '아니오'라고 답한 경우]

☐ 할아버지 ☐ 할머니 ☐ 친척 ☐ 기타 _____

이들 양육자 중 남성의 농인 여부: ☐ 농인 ☐ 난청인 ☐ 청인

이들 양육자 중 여성의 농인 여부: ☐ 농인 ☐ 난청인 ☐ 청인

13. 어머니와의 주된 의사소통 방법

☐ 한국수어 ☐ 한국어대응수어+말 ☐ 제스처 ☐ 말 ☐ 필담
☐ 흠싸인 ☐ 기타 _____

14. 아버지와의 주된 의사소통 방법

☐ 한국수어 ☐ 한국어대응수어+말 ☐ 제스처 ☐ 말 ☐ 필담
☐ 흠싸인 ☐ 기타 _____

15. 형제 및 자매의 농인 여부/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 자매들과의 주된 의사소통 방법

☐ 한국수어 ☐ 한국어대응수어+말 ☐ 제스처 ☐ 말 ☐ 필담
☐ 흠싸인 ☐ 기타 _____

17. 가족 및 친척의 농인 여부

할머니	<input type="checkbox"/> 농인	<input type="checkbox"/> 난청인	<input type="checkbox"/> 청인
할아버지	<input type="checkbox"/> 농인	<input type="checkbox"/> 난청인	<input type="checkbox"/> 청인
_____	<input type="checkbox"/> 농인	<input type="checkbox"/> 난청인	<input type="checkbox"/> 청인
_____	<input type="checkbox"/> 농인	<input type="checkbox"/> 난청인	<input type="checkbox"/> 청인

18. 배우자의 농인 여부

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

19. 배우자와의 주된 의사소통 방법

☐ 한국수어 ☐ 한국어대응수어+말 ☐ 제스처 ☐ 말 ☐ 필담
☐ 홈싸인 ☐ 기타 _____

20. 자녀의 농인 여부

☐ 농인 (_____ 명)
☐ 난청인 (_____ 명)
☐ 청인 (_____ 명)
☐ 자녀 없음

21. 자녀와의 주된 의사소통 방법

☐ 한국수어 ☐ 한국어대응수어+말 ☐ 제스처 ☐ 말 ☐ 필담
☐ 홈싸인 ☐ 기타 _____

언어

22. 언어 제공자(본인)의 주된 의사소통 방법

- ☐ 한국수어 ☐ 한국어대응수어+말(음성언어) ☐ 말(음성언어)

23. 한국수어 습득 시기: _____ 살 때부터

24. 한국수어 습득 장소

- ☐ 가정/부모님
☐ 농학교
☐ 농인 교회
☐ 농아인협회
☐ 기타 _____

25. 한글지문자(지화) 사용 유무

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

26. 지역수어 사용: 지역명 _____

27. 농(난청)인과의 정기적인 만남 여부

- ☐ 일주일에 1~2번 ☐ 일주일에 3~4번 ☐ 일주일에 5번 이상

28. 본인의 한국어 능력 정도

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

29. 외국수어 사용 여부

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

학교 및 직업

30. 출신 학교

출신학교1: _____살부터 _____살까지 _____학교 _____지역

학교에서의 주된 의사소통 방법:

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

출신학교1: _____살부터 _____살까지 _____학교 _____지역

학교에서의 주된 의사소통 방법:

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

출신학교1: _____살부터 _____살까지 _____학교 _____지역

학교에서의 주된 의사소통 방법:

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

출신학교1: _____살부터 _____살까지 _____학교 _____지역

학교에서의 주된 의사소통 방법:

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

출신학교1: _____살부터 _____살까지 _____학교 _____지역

학교에서의 주된 의사소통 방법:

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

31. 학교 기숙사 입사 여부: ☐ 예 ☐ 아니오

32. 학교 기숙사 입사에 대한 추가정보 [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. 대학명/전공

대학명 _____ 전공 _____ 지역 _____

35. 직업훈련 기관명/지역

기관명 _____ 지역 _____

36. 재직 여부: ☐ 예 ☐ 아니오

37. 직장 정보 (현재 재직 중인 직장 또는 마지막 직장)

직종(직장) _____ 지역 _____
 근무 기간 _____년부터 _____년까지

직장에서의 주된 의사소통 방법:

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

38. 직장에서 작업 형태

- ☐ 팀에서 일함 ☐ 혼자 일함

39. 취미와 특별히 흥미가 있는 영역(예, 축구, 자동차, 요리, 독서, 등산 등)

농 사회에서의 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 예술분야 활동에서 수어사용 여부(예, 수어 예술분야: 연극, 시, 퍼포먼스 등)

☐ 수어를 사용한다. ☐ 수어를 사용하지 않는다.

43. 우세손

	오른손사용	왼손사용	가끔 왼손 사용
글쓰기	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
양치질	<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: 2019년 _____ 월 _____ 일

서명: _____

[부록 5] 언어제공자와 사회자 관계에 관한 질문지

언어제공자와 사회자 관계에 관한 질문지

자료수집 날짜: _____

사회자: 000

언어제공자 A: _____

언어제공자 B: _____

언어제공자 A와 언어제공자 B

- 1) ☐ 우리는 서로 매우 잘 안다.
☐ 우리는 서로 잘 안다.
☐ 우리는 서로 알지만, 자주 만나지 않는다.
☐ 우리는 과거에 만난 적이 있다.
☐ 우리는 오늘 처음 만났다.

- 2) 우리는 서로 친척 관계이거나 결혼을 통해 친척 관계이다
☐ 예 ☐ 아니오
 ('예'라고 답한 경우) 어떤 친척 관계인지(예, 조카 등)

사회자와 언어제공자 A

- 1) ☐ 우리는 서로 매우 잘 안다.
☐ 우리는 서로 잘 안다.
☐ 우리는 서로 알지만, 자주 만나지 않는다.
☐ 우리는 과거에 만난 적이 있다.
☐ 우리는 오늘 처음 만났다.

- 2) 우리는 서로 친척 관계이거나 결혼을 통해 친척 관계이다
☐ 예 ☐ 아니오

(‘예’라고 답한 경우) 어떤 친척 관계인지(예, 조카 등)

사회자와 언어제공자 B

- 1)
 - ☐ 우리는 서로 매우 잘 안다.
 - ☐ 우리는 서로 잘 안다.
 - ☐ 우리는 서로 알지만, 자주 만나지 않는다.
 - ☐ 우리는 과거에 만난 적이 있다.
 - ☐ 우리는 오늘 처음 만났다.

- 2) 우리는 서로 친척 관계이거나 결혼을 통해 친척 관계이다.
☐ 예 ☐ 아니오
 (‘예’라고 답한 경우) 어떤 친척 관계인지(예, 조카 등)

[부록 6] 사회자 일지

사회자 일지				
일 시			장 소	
참석자				
내용 및 결과	언어제공자쌍 : A: B:			
	기타 사항			
과제명	소요 시간	시작시간	종료시간	기타사항
전체				
1. 촬영 동의서	3"			
2. 수어 이름 말하기	5"			
3. 준비해 온 농담 이 이야기하기	10"			
4. 영상을 보고 수어 로 이야기하기	10"			
5. 다시 이야기하기	10"			
6. 토론	25"			A: B:
7. 표지판	20"			
8. 사건	30"			A: B:
9. 주제	25"			A: B:
10. 다양한 수어 표현	10"			
11. 농사회 행사	11"			
12. 흥부와 놀부	10"			
13. 옛날 수어와 현재 수어	10"			
14. 농학교 경험 이야 기하기	15"			

[부록 7] 후보 타입 등록 목록

연번	타입명	처리구분	설명
1	~답다1	등록완료	[~대로1]와 다른 형태이며, '특성이나 자격이 있음'의 뜻을 더하는 접미사인 수어를 새 타입으로 등록
2	가다7	등록완료	[가다1A]와 다름 / 목적지가 있는 수어를 새 타입으로 등록
3	가만있다2	등록완료	손이 아무나 대지 않다는 의미를 나타내는 새 타입 등록
4	가지다3	등록완료	[가지다2]와 의미나 사용이 달라 새 타입 등록. 빼앗다, 단순한 소유의 의미
5	감자1	등록완료	감자를 얇게 써는 수어를 새 타입으로 등록
6	개혁1	등록완료	'새롭게 뜯어고치기 위해', 비우세손이 바닥을 나타내고, 우세손은 아래에서 위로 움직여 새롭게 한다는 의미의 수어를 새 타입으로 등록
7	거리두다1	등록완료	수어자 본인과 상대 사이에 일정한 곳에 놓은 수어를 새 타입으로 등록
8	경쟁1	등록완료	시합, 대회, 경쟁 등 [사람4]과 [사람4] 겨루는 수화를 새 타입으로 등록
9	고구마1	등록완료	채소의 새 타입 등록
10	고무1	등록완료	'지우개'의 성질이 고무와 같아서 새 타입 등록
11	굴1	등록완료	과일의 새 타입 등록
12	기숙사3	등록완료	[점심1]수어와 비슷한 경남 지역의 수어를 새 타입으로 등록
13	꺼지다1	등록완료	[꺼져1]로 제안된 타입 후보. 부정적 의미가 있으며, 수어의 형태를 '꺼져'라는 글로스로 유추해내기 쉬우나 타입명 규칙에 따라 기본형으로 새 타입 등록
14	꿀보기싫다2	등록완료	입안 깊숙한 곳에 손가락을 넣어서 토하는 듯 표현한 수어를 새 타입으로 등록
15	낳다2	등록완료	일반적 표현의 [낳다1]와 달리 우세손인 아기가 직접 나오는 수어를 새 타입으로 등록
16	넘어가다1	등록완료	속아 넘어가다'와 달리 '뺨에 넘어가다'라는 의

			미가 있는 수어로 새 타입 등록
17	널다3	등록완료	널다1는 모든 손가락이 집어넣는 반면, 널다3는 엄지와 검지만 집어넣은 수어로 새 타입 등록
18	네살 많다1	등록완료	지숫자 포함한 수어 새 타입 등록
19	네시1	등록완료	시간을 나타내는 수어 새 타입 등록
20	놀리다1	등록완료	양손이 양쪽 뺨에다 얼레리 꼰레리 모양의 수어를 새 타입으로 등록
21	눈치3	등록완료	양쪽 눈동자가 재빠르게 움직이는 모양의 수어를 새 타입으로 등록
22	달리다3	등록완료	[달리다1]와는 수형과 움직임의 형태가 달라 새 타입 등록
23	닭싸움1	등록완료	한 쪽 다리만 양반다리로 한 상태의 12형이 움직이고 다른 손은 내기를 의미한 [돈1]과 함께 나타낸 수어를 새 타입으로 등록
24	대전2	등록완료	지역명은 고유의 타입명 사용. [크다+발]
25	도전1	등록완료	사람4을 도전하는 수어를 새 타입으로 등록
26	돌머리1	등록완료	[머리+돌]의 머리는 지시 형태로 나타내는데 [돌머리1]에서는 돌 모양의 주먹형으로 표현됨
27	들다3	등록완료	무언가를 들고 어깨에 대는 고정 수어를 새 타입으로 등록
28	딱맞다4	등록완료	번뜩이는 아이디어가 떠올랐을 때 사용되는 수어로 새 타입 등록
29	만취1	등록완료	[취하다1]와 달리 머리까지 짹 찬 느낌의 수어를 새 타입으로 등록
30	많다6	등록완료	거센 느낌이 날 정도로 많다로 설정하여 새 타입 등록
31	말하다4	등록완료	입말하다를 말하다 범주에서 새 타입 등록
32	멀리뛰기1	등록완료	[뛰다1]은 '높이 뛰다'이므로 '멀리뛰기1'을 새 타입으로 등록
33	모의하다1	등록완료	양손이 천 주먹형이 상호 접촉하며 움직이는 수어를 새 타입으로 등록
34	무시5	등록완료	무시하기만 하는 [무시하다1]과 달리 괴롭혀 죽이는 의미까지 있는 수어로 새 타입 등록
35	미달이1	등록완료	[문1] 수형이 같으나 움직임이 다르고, 의미가 달라 새 타입으로 등록

36	바꾸다6	등록완료	[바꾸다1]는 공간 활용을 하지 않은 일반적인 표현이며, [바꾸다6]은 너와 나의 무언가를 바꾸자는 공간일치의 수어를 새 타입으로 등록
37	벌레2	등록완료	[벌레1]는 벌레를 의미하고, [벌레2]는 지렁이, 누에를 의미함. [벌레1]가 있어 [벌레2]로 새 타입 등록
38	벌2	등록완료	빛1의 변이형, 반복이 있어 새 타입 등록
39	부리2	등록완료	[부리1]의 움직임과 달라 새 타입 등록
40	불쌍하다2	등록완료	[불쌍하다1]의 [사랑] 생략한 형태의 새 타입 등록
41	비율1	등록완료	양손이 크고 작은 정도를 나타낸 수어를 새 타입으로 등록
42	빠져나오다1	등록완료	무리에서 한사람이 빠져나오는 수어를 새 타입으로 등록 [나오다1]로 제안된 타입 후보. 탈퇴하다, 사람들 속에서 빠져나오다 등으로 사용하므로 등록 후 [빠져나오다1]로 타입명 수정
43	뻘뻘하다5	등록완료	[뻘뻘하다1]의 변이형, 움직임이 없는 수어를 새 타입으로 등록
44	뻘뻘하다6	등록완료	[뻘뻘하다5]의 변이형, [뻘뻘하다5] 보다 더한 느낌을 주는 수어를 새 타입으로 등록
45	뿔다6	등록완료	[집게형]으로 제안된 타입후보. 등록 후 [뿔다6]으로 타입명 변경
46	사탕1	등록완료	동그라미 모양 [십형1]의 뿔에 대는 수어를 새 타입으로 등록
47	선약있다1	등록완료	[약속마치다1]로 제안된 타입 후보. [약속1]+[박수형태]의 그 '마치다'는 접미사의 역할인 하나의 결합수어로 타입 등록 후 [선약있다1]로 타입명 변경
48	세월1	등록완료	[시간2]+[강1]로 볼 수 있으나 합성어로서 새 타입 등록. 분절이 어려울 정도로 두 단어가 합쳐진 경우는 하나의 타입으로 한다는 지침 추가 필요
49	셋1	등록완료	숫자의 새 타입 등록
50	숙제1	등록완료	다른 수형과 반복 횟수가 다른 [문제1]와 다른 독립적 의미를 지니는 수어이므로 새 타입 등록
51	스트라이크2	등록완료	[스트라이크1]와 다른 형태의 새 타입 등록
52	시1	등록완료	문장1의 변이형, 문장을 편집한다는 수어를 새

			타입으로 등록
53	시월1	등록완료	날짜를 나타내는 새 타입 등록
54	실1	등록완료	엄지와 검지를 대서 가늘고 길게 나타낸 수어를 새 타입으로 등록
55	실습2	등록완료	직접적으로 손대는 현상을 나타낸 수어를 새 타입으로 등록
56	십년2	등록완료	시간을 나타내는 새 타입 등록
57	싸우다2	등록완료	[싸우다1]과 달리 싸움, 전쟁, 전략적 등 범위가 큰 수어를 새 타입으로 등록
58	양궁1	등록완료	스포츠의 새 타입 등록
59	어지럽다1	등록완료	빈혈, 어지럽다고 할 때 사용된 수어를 새 타입으로 등록
60	없다12	등록완료	다 떨어져서 거의 없다는 의미의 수어를 새 타입으로 등록
61	엘지1	등록완료	알파벳 L의 검지를 이마에다 대는 수어를 새 타입으로 등록
62	연필1	등록완료	무언가를 쓰는 [쓰다1]을 그 전에 함께 실현된 혀에다 대는 표현을 새 타입으로 등록
63	열광1	등록완료	[떼쓰다]로 제안된 타입 후보. 하나로 분절되는 합성어 [생각1]+[미치도록 외치다]의 형태로 새 타입 등록
64	열쇠1	등록완료	문을 나타낸 비우세손과 열쇠를 잡고 돌리는 수형을 하나의 수어로 새 타입 등록
65	오가다3	등록완료	[사람4]가 오가는 수어를 새 타입으로 등록
66	오년후1	등록완료	시간을 나타내는 새 타입 등록
67	오르다3	등록완료	구형이 가파르게 올라가는 고정된 수어를 새 타입으로 등록
68	오르락내리락1	등록완료	[오르다3]과 [하락하다1]를 번갈아가며 표현한 하나의 단어를 새 타입으로 등록
69	오십이시간1	등록완료	시간을 나타내는 새 타입 등록
70	올림픽1	등록완료	오륜모양을 나타내기 위해 고리연결을 두세 번 한 수어를 새 타입으로 등록
71	원한1	등록완료	[한1]로 제안된 타입 후보. 한을 품다로 가슴에다 표현된 수어로 등록 후 [원한1]로 타입명 변경

			경
72	의욕상실1	등록완료	형태가 다른 [욕심1]로 생각할 수 있으나, [끈기2]가 없어진 형태의 한 단어를 새 타입으로 등록
73	이불1	등록완료	두꺼운 모양의 양손 수형이 가슴에 대는 수어를 새 타입 등록
74	이십오년1	등록완료	시간을 나타내는 새 타입 등록
75	이십육1	등록완료	숫자의 새 타입 등록
76	이십일2	등록완료	숫자의 새 타입 등록
77	인사2	등록완료	여러 사람이 인사하는 모양의 새 타입 등록
78	일등1	등록완료	숫자의 새 타입 등록
79	입댄다1	등록완료	이익 따위를 혼자 차지하거나 가로채는 관용표현인 수어를 새 타입으로 등록
80	자르다5	등록완료	칼로 자르는 모양의 고정된 수어(1형)로 새 타입 등록
81	재봉2	등록완료	무언가(옷감)를 들고 재봉질 하는 수어를 새 타입으로 등록
82	재촉하다2	등록완료	[건드리다1]로 등록된 타입 후보. '재촉하다'의 의미가 있어 타입명 [재촉하다2] 등록. 타입 목록의 [재촉하다1]는 누군가가 수어자 본인을 재촉함을 나타내는 것을 의미하는 것임
83	정신1	등록완료	생각1의 변이형, ㄹ형으로 실현된 수어를 새 타입으로 등록
84	정체성1	등록완료	자존심, 자부심, 자아를 나타낸 가슴에다 정체성의 ㄹ형이 실현된 수어를 새 타입으로 등록
85	제기차기1	등록완료	엎전을 나타낸 0형 수형+발을 나타낸 1형으로 표현된 수어 새 타입 등록
86	죽이다3	등록완료	일반적으로 죽이는 의미의 [죽이다1]와 달리 죽이다3는 특정 도구(칼)로 목에다 죽이는 수어를 새 타입으로 등록
87	중독2	등록완료	'주사맞다'같은 형태, 그러나 '중독되다' 의미가 있는 형태의 새 타입 등록
88	차근차근1	등록완료	양손이 천천히 진행해 나가는 모양의 새 타입 등록
89	청소하다2	등록완료	빗자루와 쓰레받기의 수어를 생산적 수어로 볼 수 있으나, 경남지역에서 '청소하다' 의미를 가

			진 하나의 단어로 새 타입 등록
90	촉수화1	등록완료	한손위에 다른 손이 얹는 수어를 새 타입으로 등록
91	촉하1	등록완료	의미 범위가 더 넓은 [행사2]보다는 촉하라는 기저이미지와 더 가까운 수어를 새 타입으로 등록
92	칠백1	등록완료	숫자의 새 타입 등록
93	침묵1	등록완료	T형이 입 앞에다 대서 멈추는 수어를 새 타입으로 등록
94	탁구1	등록완료	스포츠의 새 타입 등록
95	테니스1	등록완료	스포츠의 새 타입 등록
96	토너먼트1	등록완료	운동 경기에서 주로 사용된 그 수어를 새 타입으로 등록
97	토하다1	등록완료	목에서 시작한 구토 수어를 새 타입으로 등록
98	투포환1	등록완료	스포츠의 새 타입 등록
99	팽이1	등록완료	운동 도구의 새 타입 등록
100	편들다1	등록완료	[고리]와는 다른 고정된 어휘로 새 타입 등록
101	풋살1	등록완료	스포츠의 새 타입 등록
102	하락하다2	등록완료	[떨어지다2]로 제안된 타입후보. '세일, 할인' 비슷한 형태의 수어로 등록 후 [하락하다2]로 타입명 변경
103	혈안1	등록완료	양쪽 약간 닫힌 1형이 양쪽 눈에다 특정 행위를 나타낸 수어를 새 타입으로 등록
104	협력1	등록완료	배를 가리키며 모으다1를 표현한 수어를 하나의 분절으로 보고 새 타입 등록
105	협박1	등록완료	공포, 오싹하다, 아찔하다 등으로 사용된 수어를 새 타입으로 등록
106	훈련2	등록완료	[훈련1]는 [연습1+훈련2]을 하나의 타입으로 등록되어 있음. 토큰 [훈련2]에는 '연습'이 생략되어 있으므로 다른 타입으로 설정. 의미의 범위가 다름
107	휠체어3	등록완료	휠체어의 바퀴가 굴러가는 형태로 휠체어를 의미. 새 타입으로 등록
108	가위바위보1	추가논의 - 타입등록	가위이나 바위 아니면 보가 나오는 수형이 나타내는 새 타입 등록
109	간사1	추가논의	[부하1]의 움직임이 같으나 수형과 의미가 달라

		- 타입등록	새 타입 등록.
110	갑자기3	추가논의 - 타입등록	'심심하다'와 비슷한 형태를 가짐. [갑자기1, 2]와 위치와 움직임은 형태는 비슷하나 수형과 의미가 달라 새 타입 등록. '갑자기' 의미에 더하여 '충격적'이라는 의미
111	개인전1	추가논의 - 타입등록	수형과 움직임에서 [만나다1]과 비슷하나 접촉여부가 다른 [개인전1]을 새 타입 등록. 경기 시합에서 주로 사용하는 수어
112	게이트볼1	추가논의 - 타입등록	스포츠의 새 타입 등록
113	경상남도1	추가논의 - 타입등록	지역 관련 새 타입 등록
114	공짜2	추가논의 - 타입등록	대머리 치는 [공짜1]와 수어가 달라 새 타입 등록. [돈1+없다2]의 합성어 * 추후 돈이 생략되고 이마만 치는 형태의 [공짜3]이 필요
115	관절1	추가논의 - 타입등록	2개 이상의 뼈가 서로 맞닿아 연결되어 있는 곳이 다치는 수어를 새 타입 등록
116	구걸하다1	추가논의 - 타입등록	무언가를 거저 달라고 비는 수어를 새 타입으로 등록
117	구년1	추가논의 - 타입등록	시간을 나타내는 새 타입 등록
118	농사1	추가논의 - 타입등록	밭을 가는 모양의 수어를 새 타입으로 등록
119	다섯시1	추가논의 - 타입등록	시간을 나타내는 새 타입 등록
120	단체전1	추가논의 - 타입등록	움직임에서 [만나다1]과 비슷하나 수형과 접촉여부가 다른 [단체전1]을 새 타입 등록. [개인전1]의 변이형. 경기 시합에서 주로 사용하는 수어
121	단호하다1	추가논의 - 타입등록	[무조건1, 2, 3]는 무조건 연상하지만, [단호하다1]는 무조건이 있지만 '단호하다'가 더 가까워 새 타입 등록
122	달리기1	추가논의 - 타입등록	'육상'은 많은 종목의 경기를 포함하고 있고, [달리다1]과는 다른 의미로 [달리기1]로 새 타입 등록
123	당황하다3	추가논의	일반적 표현의 [당황하다]와 형태는 비슷하나

		- 타입등록	의미가 달라 새 타입 등록
124	매끄럽다1	추가논의 - 타입등록	일이나 상황이 매끄럽게 진행됨을 의미한 수어의 새 타입 등록
125	메달1	추가논의 - 타입등록	가슴에다 대는 메달을 새 타입 등록
126	며칠전1	추가논의 - 타입등록	[과거1]과 다른 형태로 새 타입 등록
127	몇2	추가논의 - 타입등록	[몇1]과 다른 형태로 새 타입 등록. 명확한 수의 개념이 나타나는 [일이삼사]로 할 것에 대해 많은 논의를 하였으나 현재 토큰에서는 '몇'의 의미로 사용되어 [몇2]로 결정
128	목발1	추가논의 - 타입등록	목발 모양이 나타낸 F형을 양손 수어의 새 타입 등록
129	밀양1	추가논의 - 타입등록	지역명 새 타입 등록
130	반갑다2	추가논의 - 타입등록	[꼬리치다1]로 제안된 타입 후보. 강아지가 꼬리를 치며 반가는 의미로 '꼬리치다'라는 타입명은 유혹하다라는 의미로 오해할 수 있음. [반갑다2]로 새 타입 등록
131	발레1	추가논의 - 타입등록	발끝으로 서서 춤(발레)을 추는 발의 형태로 새 타입 등록
132	배드민턴1	추가논의 - 타입등록	스포츠의 새 타입 등록
133	보초1	추가논의 - 타입등록	총을 나타낸 1형이 서는 수어를 새 타입으로 등록
134	볼링1	추가논의 - 타입등록	스포츠의 새 타입 등록
135	사람가다7	추가논의 - 타입등록	[사람가다4]의 변이형으로 새 타입 등록. 의미에서 '두 사람이 가다'로 주석
136	사람바꾸다4	추가논의 - 타입등록	[사람4]가 바뀌는 형태의 새 타입 등록
137	사천2	추가논의 - 타입등록	지역명 새 타입 등록. 지숫자인 [사천1]과 다른 형태의 지역명
138	삼인조1	추가논의 - 타입등록	움직임에서 [만나다1]과 비슷하나 수형과 접속 여부가 다른 [삼인조1]을 새 타입으로 등록. [개인전1]의 변이형. 경기 시합에서 주로 사용하는

			수어
139	상담1	추가논의 - 타입등록	[회의1]와 다른 형태의 새 타입 등록. 원을 크게 그리는 외부 수동이 있을 경우, [회의3]으로 등록될 수 있으나, [상담1]은 엄지만 움직이는 형태.
140	상주다1	추가논의 - 타입등록	칭찬, 보상, 축하할 때 주는 수어로 어휘화된 형태의 새 타입 등록
141	술런1	추가논의 - 타입등록	스포츠의 새 타입 등록
142	스트라이크1	추가논의 - 타입등록	한손 R 수형의 새 타입 등록 주로 볼링에서 사용하며 2번 이상 반복된 수동이 나타남
143	아니다5	추가논의 - 타입등록	[아니다2]와 비슷한 형태이나 의미가 달라 새 타입 등록
144	앞1	추가논의 - 타입등록	수어자 관점에서 보는 앞 위치로 나타내기 위해 손 바닥이 앞으로 다가는 수어를 새 타입 등록. * '다의어'에 대한 추가 지침 필요: 형태가 같고 의미(한국어로 번역할 때)가 다양할 경우 타입명은 가장 광범위하고 일반적인 의미를 가진 단어로 타입명을 결정한다. 예를 들어 [높다2]의 경우 '높다, 위, 선배 등의 의미'를 갖는 경우 [높다2]로 타입명을 정한다. [높다3]을 추가할 것을 고려-현재 해당 토큰이 없어 타입을 새로 만들 수 없으므로 추후 토큰이 있을 경우 다음 사항을 고려함. 장향(위, 아래)에 따라 의미가 달라질 수 있으므로 추후 장향이 위로 향하는 토큰이 있을 경우 타입을 만들 것인지에 대한 새로운 논의가 필요함
145	양산1	추가논의 - 타입등록	지역명의 새 타입 등록. 지역명일 경우 타입 설명에 '지역명'이라고 명시함
146	어쩌라고1	추가논의 - 타입등록	[어쩔라고1]로 제안된 타입 후보. '어찌하다' 기본형 보다는 특정 상황에서 나오는 표현으로 등록 후 표준국어맞춤법에 맞게 [어쩌라고1]로 타입명 변경
147	여섯시1	추가논의 - 타입등록	시간을 나타내는 새 타입 등록
148	열번째1	추가논의 - 타입등록	숫자의 새 타입 등록

149	영문모르다1	추가논의 - 타입등록	모든 손가락이 편 한 손의 엄지가 턱에다 두 번 접촉한 수어를 새 타입으로 등록
150	오년전1	추가논의 - 타입등록	시간을 나타내는 새 타입 등록
151	원하다4	추가논의 - 타입등록	‘귀찮다’라는 수어와 비슷한 형태이나 비수지와 형태가 조금 다름. ‘중독’의 의미도 있으나 [원 하다2]와 유사한 부분이 있으므로 [원하다4]로 새 타입 등록
152	이년전1	추가논의 - 타입등록	숫자의 새 타입 등록
153	이익2	추가논의 - 타입등록	[이익1]보다 긍정적인 의미가 있는 형태의 새 타입 등록
154	이인조1	추가논의 - 타입등록	움직임에서 [만나다1]과 비슷하나 수형과 접속 여부가 다른 [이인조1]을 새 타입 등록. [개인전 1]의 변이형. 경기 시합에서 주로 사용하는 수어
155	자료1	추가논의 - 타입등록	‘물건’이라는 수어에서 시작하여 같은 형태의 수 어로 지문자를 차용하였으나 ‘자료’라는 의미로 많이 사용되고 있으므로 [물건2]로 하지 않고 [자료1]로 새 타입 등록
156	접다2	추가논의 - 타입등록	[접다1]과 비슷한 형태이면서도 ‘어떠한 일을 접 다’, ‘그만두다’ 의미가 있는 형태의 새 타입 등록
157	코치1	추가논의 - 타입등록	[코2]+[지도하다]의 합성어를 새 타입 등록 운동 경기에서 지도자(감독, 코치)로 주로 사용 하며, 감독보다는 코치로 타입명을 설정.
158	포기3	추가논의 - 타입등록	‘항복’의 이미가 있으며 [포기2]와 유사한 부분 이 있으므로 [포기3]으로 새 타입 등록
160	혈액형	추가논의 - 기록방법	혈액형에 따르는 영어지문자의 경우 ‘영어지문 자’로 주석하고 의미에 ‘혈액형 0형’이라고 기록
161	두사람가다1	추가논의 - 기록방법	[사람가다1]의 변이형. 의미총렬에서 ‘두 사람이 가다’로 기록
162	떨어지다2	추가논의 - 기록방법	공이 떨어지는 모양의 생산적 수어. [구형4]로 주석함
163	대진2	등록요청	[토너먼트1]로 등록된 타입
164	둘러싸다2	보류	[둘러싸다1]에 해당하는 형태. [둘러싸다1]의 기

			본형에 대해 향후 한손/두손 형태 분절 필요
165	목소리1	보류	[외치다1]에 해당
166	몰려들다3	보류	[몰려들다1]에 해당. [몰려들다1]의 기본형에 대해 향후 한손/두 손 형태 분절 필요
167	가게2	등록불가	[슈퍼1]과 동일
168	가다8	등록불가	[가다6]과 동일. 두 손이 각각 개별적으로 움직여지는 [가다6]이 조직적인 움직임의 형태로 나타남
169	가로등1	등록불가	[빛1]의 반복형. [빛1]로 각각 분절하고 의미는 전체를 묶어 가로등으로 기록
170	가져오다5	등록불가	기존 타입 [가져오다3]과 중복
171	개인2	등록불가	[개인1]과 동일. [개인1]의 변이형
172	고발2	등록불가	[고발1]과 중복
173	고장3	등록불가	[고장2]와 동일. [고장2]와 수형이 같고 움직임에 의한 변이 형태
174	괴롭히다2	등록불가	[주다3]에 해당하는 수어
175	근육1	등록불가	[꼬다1]과 동일. [꼬다1]의 변이. 농인입장에서 꼬다보다는 근육이 쉽게 생각날 수 있음
176	기도3	등록불가	[기도2]와 동일. 수형의 변이
177	끈기3	등록불가	[끈기1]의 음운론적 변이
178	끓다1	등록불가	[자르다1]과 동일. '끓다'는 이미 고정된 수어로 자르다와는 다른 의미가 있음. 이미 [끓다1]을 [자르다1]로 변경하였기 때문에 자르다로 유지
179	내리다3	등록불가	[하락하다1]과 동일. 내리다, 후퇴하다, 저하의 의미
180	눅다2	등록불가	[눅다1]의 반복형. 여러 명이 누워있다는 의미
181	달리다2	등록불가	[달리다1]과 동일
182	동창2	등록불가	[동창1]과 동일
183	둘 다 책임 있다1	등록불가	[공범1]과 동일
184	떼버리다1	등록불가	향후 [구형4A]와 [버리다1]로 분절 필요
185	맞다4	등록불가	[갈다1]과 동일
186	모두2	등록불가	[지시2호]와 동일
187	버스2	등록불가	[버스1]과 동일
188	보다17	등록불가	[보다7]에 해당
189	사랑2	등록불가	[사랑1]과 동일

190	살1	등록불가	[피부2]와 동일. [고기1]과 [피부2]는 다름
191	선배1	등록불가	[높다2]에 해당
192	손씻다1	등록불가	[씻다1]과 동일
193	시냇물1	등록불가	[물1]+[강1]에 해당됨. 합성어 시냇물
194	아이고4	등록불가	[아이고2]에 해당
195	아차7	등록불가	[아하2]와 동일. 변이형
196	오르내리다1	등록불가	[오르락내리락1]에 포함
197	장난3	등록불가	[장난1]과 동일. 경남지역 수어로 보여지나 현재 변이형에 대한 타입을 구분하지 않으므로 [장난1]의 변이형으로 보고 통합함. 단, 추후 타입에 대한 변이형에 대해 재주석 또는 검토 후 지역적 특성을 구분하는 연구가 필요
198	재산1	등록불가	[가치1]과 동일
199	줍다3	등록불가	[뵈다5]에 해당
200	징그럽다1	등록불가	[꼴보기싫다2]에 해당
201	참다3	등록불가	[참다2]과 동일. 변이형
202	타입명모름	등록불가	추측/대충'의 의미로 제안된 타입 후보. [과장하다1]과 동일
203	타입명모름.	등록불가	[몰래1]과 동일
204	타입명모름...	등록불가	담화, 사연 등'의 의미로 제안된 타입 후보. [이야기하다1]과 동일
205	태도1	등록불가	[행동1]과 동일
206	태양2	등록불가	[빛1]과 동일
207	파다1	등록불가	땅을 판다', '구덩이를 만들 때 파내는 동작' 등을 나타내는 것으로 2017년에도 등록된 바 있음. 생산적 수어 [구형1]로 주석 후 의미총렬에서 '파다' 또는 '캐다'로 기록하는 것으로 함
208	퍼다1	등록불가	[열다기]에 해당
209	한곳웁기다 or같은곳웁 기다	등록불가	[웁기다1]에 해당

[부록 8] 한국수어 말뭉치 수집 시간 현황 (2015~2019)

유도 과제	서울																																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31		
A	0:01:37	0:02:01	0:01:44	0:01:13	0:01:52	0:03:13	0:01:22	0:01:56	0:01:58	0:01:00	0:02:21	0:01:41	0:02:09	0:05:18	0:01:09	0:01:19	0:01:51	0:01:33	0:00:59	0:00:47	0:00:45	0:02:43	0:02:58	0:03:52	0:00:54	0:01:13	0:01:15		0:04:41	0:02:18	0:01:02		
B	0:11:56	0:06:29	0:05:24	0:10:36	0:02:36	0:04:22	0:02:34	0:16:21	0:07:55	0:01:27	0:06:33	0:07:40	0:05:36	0:06:35	0:01:07	0:02:53	0:06:26	0:12:16	0:01:37	0:00:56	0:01:54	0:04:26	0:02:52	0:06:37	0:02:34	0:02:33	0:04:31		0:06:06	0:03:07	0:04:27		
C	0:07:26	0:13:50	0:15:09	0:07:34	0:09:13	0:12:42	0:13:19	0:13:01	0:11:58	0:10:14	0:06:37	0:12:39	0:13:27	0:15:52	0:08:19	0:11:01	0:10:30	0:08:57	0:09:57	0:07:21	0:09:28	0:08:35	0:14:11	0:12:52	0:09:10	0:01:46	0:10:57		0:15:45	0:09:26	0:10:46		
D	0:03:55	0:05:01	0:04:33	0:03:37	0:04:25	0:05:01	0:03:45	0:04:24	0:04:28	0:05:56	0:04:30	0:05:03	0:03:50	0:06:39	0:03:35	0:04:56	0:03:35	0:04:55	0:04:48	0:02:55	0:03:58	0:04:56	0:04:18	0:04:54	0:03:20	0:12:40	0:03:17		0:04:28	0:03:43	0:03:40		
E	0:17:03	0:28:27	0:22:46	0:26:20	0:17:31	0:11:53	0:11:13	0:12:17	0:28:28	0:17:11	0:23:53	0:09:17	0:03:30	0:07:07	0:16:14	0:11:28	0:17:47	0:21:11	0:13:30	0:11:21	0:12:49	0:13:35	0:18:56	0:29:14	0:13:55	0:40:02	0:18:43		0:26:35	0:16:06	0:14:15		
F	0:10:58	0:12:43	0:08:11	0:17:31	0:14:14	0:24:39	0:17:53	0:15:11	0:13:03	0:13:08	0:11:22	0:17:55	0:16:48	0:14:06	0:18:35	0:10:48	0:06:50	0:11:15	0:06:17	0:07:04	0:10:54	0:07:15	0:08:42	0:04:22	0:05:28	0:03:21	0:04:52		0:16:01	0:06:39	0:05:11		
G	0:22:37	0:26:22	0:12:43	0:17:09	0:15:17	0:19:19	0:10:32	0:11:25	0:24:55	0:18:15	0:29:11	0:11:17	0:08:28	0:13:13	0:12:30	0:18:33	0:16:37	0:20:58	0:16:24	0:10:58	0:17:50	0:18:23	0:13:33	0:23:01	0:14:57	0:16:27	0:16:44		0:16:58	0:23:32	0:14:55		
H	0:23:51	0:12:50	0:26:00	0:45:42	0:20:55	0:48:49	0:08:26	0:16:30	0:33:43	0:31:40	0:18:31	0:08:31	0:07:36	0:09:03	0:27:43	0:21:17	0:24:57	0:37:47	0:30:39	0:22:56	0:40:40	0:23:46	0:15:10	0:40:50	0:32:31	0:22:14	0:40:24		0:20:46	0:25:58	0:39:17		
I	0:02:05	0:06:51	0:06:57	0:02:48	0:06:07	0:06:46	0:03:55	0:08:36	0:03:15	0:05:54	0:11:54	0:05:26	0:05:47	0:08:24	0:04:29	0:07:18	0:02:24	0:04:17	0:03:15	0:03:22	0:03:43	0:04:34	0:03:47	0:01:44	0:04:05	0:03:45	0:02:00		0:02:05	0:03:45	0:03:12		
J	0:07:01	0:05:10	0:09:20	0:15:34	0:09:16	0:08:29	0:08:40	0:04:35	0:10:39	0:15:07	0:14:08	0:06:58	0:04:44	0:04:03	0:21:22	0:21:17	0:10:03	0:11:05	0:11:37	0:12:13	0:17:24	0:16:41	0:21:11	0:09:42	0:14:29	0:14:26	0:13:40		0:12:14	0:10:37	0:07:59		
K	0:01:52	0:03:13	0:03:46	0:03:27	0:02:39	0:08:35	0:02:59	0:03:44	0:03:06	0:02:48	0:05:32	0:03:58	0:02:30	0:02:32	0:02:21	0:03:51	0:02:53	0:11:28	0:03:18	0:01:15	0:03:32	0:02:49	0:04:05	0:05:10	0:02:45	0:05:07	0:01:58		0:04:48	0:02:31	0:03:12		
L	0:07:20	0:00:00	0:10:13	0:01:27	0:12:42	0:11:17	0:05:49	0:05:23	0:03:55	0:11:57	0:07:10	0:03:58	0:09:56	0:07:13	0:07:48	0:08:41	0:07:20	0:10:19	0:07:46	0:11:20	0:12:09	0:11:45	0:10:35	0:12:50	0:08:04	0:08:29	0:11:02		0:05:53	0:16:41	0:03:09		
M	0:09:08	0:00:00	0:20:29	0:17:27	0:19:03	0:20:46	0:04:43	0:11:20	0:09:19	0:11:39	0:07:12	0:05:37	0:07:33	0:11:13	0:22:21	0:18:27	0:28:38	0:18:52	0:27:57	0:36:38	0:17:57	0:20:16	0:30:21	0:29:16	0:16:07	0:21:26	0:16:20		0:09:16	0:18:52	0:26:05		
시간 합계	2:06:49	2:02:57	2:27:15	2:50:25	2:15:50	3:05:51	1:35:10	2:04:43	2:36:42	2:26:16	2:28:54	1:40:00	1:31:54	1:51:18	2:27:33	2:21:49	2:19:51	2:54:53	2:18:04	2:09:06	2:33:03	2:19:44	2:30:39	3:04:24	2:08:19	2:33:29	2:25:43	0:00:00	2:25:36	2:23:15	2:17:10		
유도 과제	경남													전북						강원				전남				2020년 (19쌍)					
	1	2	3	4	5	6	7	8	9	10	11	12	13	1	2	3	4	5	6	1	2	3	4	1	2	3	4						
A	0:01:03	0:03:46	0:01:59	0:03:09	0:02:34	0:08:53	0:01:17	0:06:47	0:02:22	0:02:01	0:08:09	0:01:48	0:01:09	0:00:41			0:00:58	0:01:13	0:00:36	0:01:40	0:01:09	0:00:34	0:00:52	0:01:43	0:01:59	0:00:36	0:00:41						
B	0:15:51	0:04:24	0:02:48	0:13:59	0:01:55	0:02:12	0:03:01	0:02:14	0:05:26	0:02:09	0:06:24	0:01:54	0:02:14	0:05:03			0:09:10	0:10:54	0:07:00	0:03:49	0:04:27	0:09:43	0:04:34	0:04:17	0:09:23	0:03:23	0:09:05						
C	0:10:18	0:12:26	0:06:10	0:08:24	0:07:30	0:11:58	0:11:19	0:10:29	0:09:50	0:12:05	0:12:04	0:06:43	0:14:53	0:06:26			0:12:25	0:09:31	0:10:37	0:19:32	0:16:01	0:09:28	0:16:25	0:14:21	0:17:09	0:09:05	0:11:49						
D	0:01:02	0:00:59	0:05:31	0:05:24	0:06:02	0:04:17	0:07:54	0:08:11	0:03:53	0:04:09	0:07:14	0:07:32	0:09:26	0:00:45			0:00:50	0:01:30	0:00:49	0:02:31	0:00:53	0:00:42	0:02:34	0:01:49	0:02:58	0:00:56	0:01:01						
E	0:38:28	0:56:22	0:13:30	0:28:42	0:11:39	0:12:26	0:33:36	0:26:29	0:35:36	0:39:43	0:29:57	1:04:02	0:26:27	0:26:09			0:19:11	1:25:41	0:31:53	0:26:02	0:21:21	0:15:17	0:16:24	0:19:38	0:38:35	0:16:39	0:20:01						

F	0:07:43	0:18:24	0:15:43	0:23:58	0:42:42	0:38:00	0:31:27	0:26:31	0:29:34	0:43:49	0:24:58	0:25:29	0:17:54	0:12:52			0:15:34	0:10:12	0:08:42	0:12:28	0:09:29	0:27:46	0:13:20	0:19:05	0:20:49	0:14:03	0:20:38				
G	0:18:54	0:12:37	0:25:01	0:28:02	0:32:11	0:37:05	0:33:07	0:23:33	0:26:33	0:59:46	0:24:58	0:45:08	0:25:06	0:22:41			0:26:02	0:11:47	0:22:17	0:18:03	0:11:52	0:20:40	0:16:14	0:15:33	0:27:34	0:27:51	0:21:00				
H	0:35:26	0:17:30	0:19:03	0:13:30	0:12:38	0:37:33	0:22:16	0:15:54	0:30:12	0:15:51	0:27:06	0:14:32	0:32:33	0:50:22			0:40:38	0:15:57	0:26:41	0:14:22	0:11:28	0:32:05	0:34:15	0:05:16	0:20:22	0:40:21	0:23:28				
I	0:06:11	0:05:48	0:04:48	0:03:08	0:10:33	0:03:18	0:05:11	0:05:43	0:04:02	0:05:44	0:03:32	0:03:34	0:08:09	0:18:10			0:03:42	0:10:37	0:05:18	0:11:12	0:05:45	0:08:02	0:05:30	0:07:11	0:05:12	0:02:55	0:09:20				
J	0:07:22	0:02:41	0:13:24	0:06:09	0:07:58	0:10:56	0:07:07	0:07:24	0:17:22	0:07:52	0:09:20	0:64:24	0:12:16	0:02:39			0:02:36	0:27:28	0:04:51	0:04:24	0:08:36	0:08:19		0:05:49	0:04:59	0:07:50	0:09:49				
K	0:02:35	0:03:43	0:01:37	0:03:03	0:03:01	0:03:00	0:04:06	0:04:18	0:03:28	0:05:38	0:02:39	0:02:17	0:05:31	0:01:15			0:02:26	0:04:28	0:03:41	0:05:58	0:05:50	0:05:01		0:01:52	0:05:55	0:04:27	0:04:10				
L	0:06:01	0:04:52	0:05:04	0:04:55	0:04:45	0:02:14	0:05:13	0:04:06	0:16:52	0:03:18	0:08:38	0:02:40	0:06:09	0:08:47			0:04:56	0:09:31	0:11:09	0:09:39	0:07:01	0:08:37		0:16:06	0:05:49	0:01:18	0:09:35				
M	0:06:01	0:12:28	0:14:16	0:02:05	0:19:29	0:04:31	0:15:33	0:10:10	0:09:34	0:12:10	0:13:06	0:18:55	0:19:05	0:07:10			0:21:21	0:12:27	0:26:57	0:16:43	0:30:27	0:17:00		0:18:10	0:13:59	0:06:02	0:08:11				
시간 합계	2:36:55	2:36:00	2:08:54	2:24:28	2:42:57	2:56:23	3:01:07	2:31:49	3:14:44	3:34:15	2:58:05	3:14:34	3:00:52	2:43:00	0:00:00	0:00:00	2:39:49	3:31:16	2:40:31	2:26:23	2:14:19	2:43:14	1:50:08	2:10:50	2:54:43	2:15:26	2:28:48				

서울지역 70시간 16분 42초
 전북지역 11시간 34분 36초
 강원지역 9시간 14분 04초
 경남지역 37시간 01분 03초
 전남지역 9시간 49분 47초

[부록 9] 한국수어 말뭉치 기본(의미) 주석 현황 (2015~2019)

[illegible]

[부록 10] 한국수어 말뭉치 번역 시간 현황 (2015~2019)

유도 과제	서울																														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
A	0:01:37	0:02:01	0:01:44	0:01:13	0:01:52		0:01:22	0:01:56	0:01:58	0:01:00	0:02:21	0:01:41	0:02:09		0:01:09	0:01:19	0:01:51	0:01:33	0:00:59		0:00:45	0:02:43	0:02:58		0:00:54		0:01:15		0:04:41	0:02:18	
B																															
C																															
D	0:03:55		0:04:33	0:03:37	0:04:25			0:04:24	0:04:28	0:05:56	0:04:30	0:05:03	0:03:50		0:03:35	0:04:56	0:03:35	0:04:55	0:04:48		0:03:58	0:04:56	0:04:18				0:03:17		0:04:28	0:03:43	
E	0:17:03																		0:13:30												
F	0:10:58		0:08:11	0:17:31			0:17:53	0:15:11	0:13:03	0:13:08	0:11:22	0:17:55	0:16:48		0:18:35		0:06:50	0:11:15	0:06:17		0:10:54		0:08:42		0:05:28		0:04:52		0:16:01		
G	0:22:37						0:10:32	0:11:25		0:18:15								0:20:58											0:16:58		
H																													0:20:46		
I	0:02:05		0:06:57	0:02:48			0:03:55	0:08:36	0:03:15	0:05:54	0:11:54	0:05:26	0:05:47		0:04:29		0:02:24	0:04:17	0:03:15		0:03:43	0:04:34	0:03:47		0:04:05		0:02:00				
J																			0:11:37					0:09:42							
K																															
L	0:07:20			0:01:27			0:05:49	0:05:23	0:03:55	0:11:57	0:07:10	0:03:58	0:09:56		0:07:48		0:07:20		0:07:46								0:11:02		0:05:53		
M																													0:09:16		
시간 합계	1:05:35	0:02:01	0:21:25	0:26:36	0:06:17	0:00:00	0:39:31	0:46:55	0:26:39	0:56:10	0:37:17	0:34:03	0:38:30	0:00:00	0:35:36	0:06:15	0:22:00	0:42:58	0:48:12	0:00:00	0:19:20	0:12:13	0:19:45	0:09:42	0:10:27	0:00:00	0:22:26	0:00:00	1:18:03	0:06:01	0:00:00
유도 과제	경남													전북						강원				전남				2020년 (19쌍)			
	1	2	3	4	5	6	7	8	9	10	11	12	13	1	2	3	4	5	6	1	2	3	4	1	2	3	4				
A	0:01:03	0:03:46												0:00:41				0:00:58	0:01:13	0:00:36	0:01:40	0:01:09	0:00:34	0:00:52	0:01:43	0:01:59	0:00:36	0:00:41			
B	0:15:51	0:04:24												0:05:03				0:09:10	0:10:54	0:07:00	0:03:49	0:04:27	0:09:43	0:04:34	0:04:17	0:09:23	0:03:23	0:09:05			
C	0:10:18	0:12:26												0:06:26				0:12:25	0:09:31	0:10:37	0:19:32	0:16:01	0:09:28	0:16:25	0:14:21	0:17:09	0:09:05	0:11:49			
D	0:01:02	0:00:59												0:00:45				0:00:50	0:01:30	0:00:49	0:02:31	0:00:53	0:00:42	0:02:34	0:01:49	0:02:58	0:00:56	0:01:01			
E	0:38:28	0:56:22												0:26:09				0:19:11	1:25:41	0:31:53	0:26:02	0:21:21	0:15:17	0:16:24	0:19:38	0:38:35	0:16:39	0:20:01			
F	0:07:43	0:18:24												0:12:52				0:15:34	0:10:12	0:08:42	0:12:28	0:09:29	0:27:46	0:13:20	0:19:05	0:20:49	0:14:03	0:20:38			

G	0:18:54	0:12:37												0:22:41			0:26:02	0:11:47	0:22:17	0:18:03	0:11:52	0:20:40	0:16:14	0:15:33	0:27:34	0:27:51	0:21:00				
H	0:35:26	0:17:30												0:50:22			0:40:38	0:15:57	0:26:41	0:14:22	0:11:28	0:32:05	0:34:15	0:05:16	0:20:22	0:40:21	0:23:28				
I	0:06:11	0:05:48												0:18:10			0:03:42	0:10:37	0:05:18	0:11:12	0:05:45	0:08:02	0:05:30	0:07:11	0:05:12	0:02:55	0:09:20				
J	0:07:22	0:02:41												0:02:39			0:02:36	0:27:28	0:04:51	0:04:24	0:08:36	0:08:19		0:05:49	0:04:59	0:07:50	0:09:49				
K	0:02:35	0:03:43												0:01:15			0:02:26	0:04:28	0:03:41	0:05:58	0:05:50	0:05:01		0:01:52	0:05:55	0:04:27	0:04:10				
L	0:06:01	0:04:52												0:08:47			0:04:56	0:09:31	0:11:09	0:09:39	0:07:01	0:08:37		0:16:06	0:05:49	0:01:18	0:09:35				
M	0:06:01	0:12:28												0:07:10			0:21:21	0:12:27	0:26:57	0:16:43	0:30:27	0:17:00		0:18:10	0:13:59	0:06:02	0:08:11				
시간 합계	2:36:55	2:36:00	0:00:00	0:00:00	0:00:00	0:00:00	0:00:00	0:00:00	0:00:00	0:00:00	0:00:00	0:00:00	0:00:00	2:43:00	0:00:00	0:00:00	2:39:49	3:31:16	2:40:31	2:26:23	2:14:19	2:43:14	1:50:08	2:10:50	2:54:43	2:15:26	2:28:48				

연구 책임자 강창욱

공동 연구원 원성옥, 홍성은, 변강석, 이은영, 송미연, 남민우

연구 보조원 명혜진, 홍장미, 이은주

담당 연구원 남미정(학예연구사)
조주연(연구원)

발행인: 소강춘

발행처: 국립국어원

서울시 강서구 금남화로 154

전화 02-2669-9695, 전송 02-2669-9737

인쇄일: 2020년 11월 27일

발행일: 2020년 11월 27일

인 쇠: 우리동네출력소

※ 이 책은 국립국어원의 용역비로 수행한 ‘2020년 한국수어 말뭉치 구축 및 연구’ 사업의 결과물을 발간한 것입니다.