

국립국어원 2017-01-19

| |
|----------------------|
| 발간등록번호 |
| 11-1371028-000670-01 |

한국수어 말뭉치 연구 및 구축

제 출 문

국립국어원장 귀하

국립국어원과 체결한 연구용역 계약에 따라 ‘한국수어 말뭉치 연구 및 구축’에 관한 연구 최종보고서를 작성하여 제출합니다.

■ 사업기간: 2017년 5월 23일 ~ 2017년 12월 15일

2017년 12월 15일

연구 책임자: 원성옥(한국복지대학교)

연구 기관 : 한국복지대학교 산학협력단

연구 책임자 : 원성옥

공동 연구원 : 허일, 홍성은, 변강석, 남기현, 이은영

연구 보조원 : 박정민, 김다해

한국수어 말뭉치 연구 및 구축

본 연구의 목적은 2015년 ‘한국수어 말뭉치 기반 조사 및 시범 구축’에서 수집한 언어 자료에 대한 기본 전사와 더불어 향후 본격적으로 한국수어 말뭉치 구축 사업을 실시하기 전에 기 전사 자료를 2016년에 개발한 한국수어 자료 통합 시스템 내에서 정비하고, 이에 따라 보다 체계적이고 효율적인 한국수어 말뭉치 구축을 위한 방안을 제시하는 것이다. 연구의 결과를 요약하면 다음과 같다.

첫째, 2015년 ‘한국수어 자료 통합 지원 시스템’이 없는 상태에서 전사한 말뭉치 자료를 2016년 구축된 시스템에 탑재하고 이를 정비하였다. 이 과정에서 일부 자료가 시스템과 연결이 되지 않는 오류가 발생하였는데 이것을 해결하는 과정에서 오류 유형을 분석하였다. 즉, 연결하는 과정에서 발생한 오류의 유형을 분석하여 총 13,340개 토큰의 연결 실패 문제를 해결하였으며, 추가 타입의 등록(한글지문자, 영어 지문자, 얼굴 이름 등), 2개의 타입을 1개의 타입으로 통일(39건), 타입명 수정(14건), 새로운 타입 등록(36건) 등 타입을 정비하였다. 또한 시스템에서의 연결 실패 문제를 해결하고 타입 정리 등을 하면서 발견된 분절 오류, 토큰 누락, 지문자 전사 오류, 단순한 손짓의 전사 오류 등 656개의 문제 토큰에 대한 재전사를 실시하였다.

둘째, 2015년 수집한 언어 자료(동영상) 중에서 총 261분 57초 분량의 영상에 대해 우세손과 비우세손에 대한 기본 전사 및 번역을 완료하였다. 전사의 품질 관리를 위하여 본 전사에 들어가기 전 3회의 전사 교육과 공동 전사를 실시한 후, 전사 신뢰도를 측정하였다. 3차 신뢰도 검사 결과를 분석하여 일치도가 낮게 나타난 토큰들의 특징을 분석하여 한 차례의 추가 전사 교육을 실시한 후 본 전사를 시작하였다. 우세손과 비우세손에 대한 기본 전사 및 번역을 완료한 파일은 토론 과제(과제 번호 E) 4개, 주제 영역 과제(과제 번호 H) 3개, 농사회 행사(과제 번호 J) 4개, 농학교 경험 이야기(과제 번호 M) 4개로 총 15개이다. 전사 결과 새로 등록된 타입 수는 1,057개이고 발생 토큰 수는 23,637개이다.

셋째, 말뭉치를 기반으로 한 실증적인 연구 사례를 제시하였다. 연구

는 기 전사된 한국수어 말뭉치에서 연구 목적에 따라 언어 자료를 선택하여 실시하였다. 그 하나는 한국수어에 나타난 색 이름의 다양한 형태를 알아보기 위해 ‘개별 수어의 다양한 표현(과제 번호 I)’ 과제로 유도한 언어 자료 중 2015년에 전사한 52분 17초 분량의 전사 자료를 기반으로 분석하였다. 한국수어에서 나타난 지명의 다양한 형태를 알아보기 위해서는 총 34분 37초 분량의 ‘옛날 수어와 현재 수어’ 과제 영상을 분석하였다. 이 두 연구에서는 유도 과제의 특성 상 해당 타입의 발생 빈도는 중요하지 않아 언어 제공자 수로 분석하였으며, 분석한 결과는 현재 사용되고 있는 수어 형태가 아닐 수 있다. 마지막으로 한국수어에 나타난 이형동의어를 알아보기 위해 2017년 전사 완료한 총 250분 분량의 영상과 2015년 전사 완료한 총 146분 분량의 영상 중 개별 수어의 ‘다양한 표현 과제(과제 번호 I)’와 ‘옛날 수어와 현재 수어 과제(과제 번호 L)’을 제외한 영상을 대상으로 분석하였다. 분석한 어휘는 넥타이, 샤워, 수영, 신호등, 아래, 아침, 연애, 오토바이를 나타내는 다른 형태의 수어이다. 이들 연구에서 분석 결과는 한정된 과제와 언어 제공자를 대상으로 하였기 때문에 이 데이터의 결과를 일반화하는 데에는 한계가 있지만 한국수어 말뭉치를 기반으로 하는 실증적이고 과학적인 연구의 사례를 제시하는 데에 그 의의가 있다.

넷째, 한국수어 말뭉치 구축 계획을 수립하기 위하여 농사회 전문가 5명과 음성 언어 구축 전문가 5명의 자문 의견을 받았다. 자문 의견과 해외 수어 연구 현황에 관한 워크숍 결과를 토대로 하여 향후 말뭉치 구축 방안을 제시하였다.

다섯째, 2015년 개발하고 2016년 개정한 말뭉치 구축 지침을 한국수어자료 통합 지원 시스템의 수정 보완 사항에 따라 수정 보완하여 제시하였다. 주요 개정 내용은 2017년 개발한 한국수어 자료 통합 지원 시스템을 기반으로 수정 보완한 내용을 추가하였으며, 언어 제공자에 대한 자료 수집을 위한 질문지를 한국 사정에 맞게 수정 보완하여 제시하고, 전사 품질관리와 관련된 내용을 추가하였다. 또한 정비 과정에서 발견된 문제점을 보완하기 위해 [지시], 생산적 수어 그리고 counting hand와 관련된 타입명 등 전사 규칙을 수정하였다.

주요어: 한국수어, 한국수어 말뭉치, 전사, 한국수어 연구

<Abstract>

Research and building of the KSL Corpus.

The Korean Sign Language (KSL) Corpus was build in 2015 and in 2016 the "Integrative System for Korean Sign Language Data" was developed. The "Integrative System for KSL Data" is a system, which administrates the language data and the annotation files. It also regulates the consistency of the token-type matching process within the annotation.

The aim of this research is to improve the annotation of the KSL Corpus by using the "Integrative System for Korean Sign Language Data". Furthermore this research will provide strategies in order to continue to build a corpus in a systematic and effective way. The results of this research are as follows.

First, the corpus data which was build in 2015 without any kind of support system was systemized by the "Integrative System for KSL Data". After feeding the system with the corpus data there were disconnection errors because some tokens didn't match with the types. These 13,340 disconnection errors were analyzed and connected manually. In this process we established new kind of types (Korean finger alphabet, english finger alphabet, sign names etc.), unified two types into one (39 samples), edited type names (14 samples) and entered new types (36 samples). We also fixed segmentation error, added omitted annotation tags, corrected wrong finger alphabet entries and further simple annotation errors, when they appeared during the manual connection process. There were 656 token in a whole which had to be re-annotated because of the mentioned reasons.

Second, 261Min and 57 seconds of the language data which was collected in 2015 has been translated into Korean and

annotated regarding the dominant and non-dominant hand. In order to provide quality assurance the annotators had to complete a three-day-instruction, a joint annotation training and their annotations were measured. The third reliability test was analyzed in detail and the results lead to an additional training session before the annotators started their official annotation. The translated and annotated files regarding the dominant and non-dominant hand are 4 "discussion tasks" (task E), 3 "subject area tasks" (task M), 4 "Deaf event tasks" (task J) and 4 "Deaf experience tasks" (task M). That means 15 new files have been annotated. The new annotated data contains 23,637 tokens and 1,057 types.

Third, we were able to do examples of research based on the corpus data. We studied the various forms of signs which occurred to express colours in KSL. In this case we based our findings on the 2015 annotated data (52 min 17 sec) of the task "isolated signs" (task I). For analyzing various forms of signs which represent different places in Korea such as cities and provinces we used annotated data (34 min and 27 sec) of the task "new vs old signs". In both cases it is rather interesting to see how many informants used a certain sign than to study the frequency of the corresponding types and in the latter case the result might display signs which are not used nowadays. We also studied synonym signs by using 250 min of data which was annotated 2017 and 146 min of data which was annotated 2015. We excluded the task "isolated signs" and "old vs new signs" and the synonyms we analyzed contain the meaning of neck tie, shower, swim, traffic light, below, morning, dating and motorbike. Our research findings are restricted because the number of tasks and informants are limited. That means further research is needed in order to generalize our results, but our results are of value

because they represent the first research based on data of the KSL Corpus.

Fourth, we questioned 5 deaf person and 5 hearing experts on spoken corpora in order to collect opinions on further corpus building. Based on these ideas as well as the content of our workshop "Current corpus based research in International Sign Language Linguistics" we developed strategies to continue building the KSL Corpus in the future.

Fifth, we revised the KSL Corpus Guidelines which were written in 2015 and adjusted in 2016 by adding the new functions of the "Integrative System for Korean Sign Language Data" and by adapting the questionnaire for the deaf informants to the korean circumstances. Besides we changed some annotation guidelines regarding the index signs, productive signs and counting hand signs.

Keywords: KSL, KSL Corpus, annotation, KSL research

【목 차】

| | |
|--------------------------------------|----|
| I. 서론 | 1 |
| 1. 연구 목적 및 필요성 | 1 |
| 2. 연구 방법 및 범위 | 2 |
| 2.1 기구축 한국수어 말뭉치 검토 및 문제점 분석 | 2 |
| 2.2 기구축 한국수어 말뭉치 전사 및 분석 연구 수행 | 3 |
| 2.3 향후 한국수어 말뭉치 구축 방향 및 방법론 마련 | 3 |
| 2.4 한국수어 말뭉치 구축 지침서 개정 | 3 |
| II. 선행 연구 고찰 | 5 |
| 1. 독일의 말뭉치 기반 수어 연구 현황 | 5 |
| 1.1 독일수어 말뭉치 연구 사업 소개 | 5 |
| 1.2 독일수어 말뭉치 연구 주제 | 5 |
| 1.3 독일수어 전자사전 | 6 |
| 1.4 변이형 연구 | 7 |
| 1.5 독일수어 말뭉치 분석 도구 | 11 |
| 1.6 독일수어 말뭉치와 수어 사전 편찬 | 12 |
| 2. 수어의 어휘 체계 | 14 |
| 2.1 고정된 어휘와 생산적 어휘 | 15 |
| 2.2 변이형과 수정 | 16 |
| 3. 수어 전사에서의 분절 | 23 |
| 3.1 음운론적 관점 | 23 |
| 3.2 음성학적 관점 | 25 |
| 3.3 수어 단어의 실증적인 길이에 대한 증거 | 26 |
| 4. 수어 연구를 위한 엘란(Elan)의 활용 | 28 |
| 4.1 층렬별 통계 값 보기 | 28 |
| 4.2 단일 전사 파일(eaf 파일) 검색 | 29 |
| 4.3 검색 및 변경 | 29 |
| 4.4 다수의 파일 층렬별 빈도 분석 | 30 |
| 4.5 다수의 전사 파일 검색하기 | 30 |
| 4.6 다중 파일 검색 및 변경 | 31 |

| | |
|--|----|
| 4.7 다수 전사 파일 다중 검색 | 32 |
| 4.8 다중 파일 검색 예 | 35 |
| Ⅲ. 기구축 한국수어 말뭉치 정비 | 37 |
| 1. 기구축 한국수어 말뭉치 연결 실패 문제 해결 | 37 |
| 1.1 추가 타입 등록 | 38 |
| 1.2 타입 정비 | 40 |
| 2. 기구축 한국수어 말뭉치의 전사 규칙 정비 | 42 |
| 2.1 [지시] | 42 |
| 2.2 Counting hand | 43 |
| 2.3 신체를 지시하는 경우 | 44 |
| 2.4 타입명이 음성 언어로 ‘구’의 형태일 때 | 45 |
| 2.5 등록된 수어와 모든 형태가 같으나 수위만 다를 경우 | 45 |
| 2.6 생산적 수어 | 46 |
| 2.7 [육형1] / [육형2] / [육형3] | 47 |
| 2.8 [육형], [십형]의 변이형 | 47 |
| 3. 정비 결과 | 48 |
| 3.1 분절 오류 | 48 |
| 3.2 토큰 누락 | 48 |
| 3.3 토큰 삭제 | 48 |
| 3.4 지문자 오류 | 48 |
| Ⅳ. 전사 및 사례 연구 | 49 |
| 1. 전사 방법 및 절차 | 49 |
| 1.1 전사를 위한 기본 사항 | 49 |
| 2. 전사자 교육 및 공동 전사 | 51 |
| 2.1. 전사자 교육 | 51 |
| 2.2 공동 전사 | 53 |
| 3. 전사 품질 관리 | 54 |
| 3.1 전사 신뢰도 검사 결과 | 55 |
| 3.2 3차 신뢰도 검사 개요 및 분석 | 56 |

| | |
|------------------------------------|-----|
| 3.3 4차 신뢰도 검사 개요 및 분석 | 67 |
| 4. 전사 결과 | 71 |
| 5. 한국수어의 어휘 변이 연구 | 73 |
| 5.1 한국수어에 나타난 색 이름 | 73 |
| 5.2 한국수어에 나타난 지명 | 84 |
| 5.3 한국수어에 나타난 이형 동의어 분석 | 89 |
| V. 한국수어 말뭉치 구축 계획 | 95 |
| 1. 한국수어 말뭉치 구축을 위한 자문 | 95 |
| 1.1 자문 방법 | 95 |
| 1.2 자문을 위한 질문 | 95 |
| 1.3 자문 의견 결과 | 96 |
| 2. 한국수어 말뭉치 구축 방안 | 109 |
| 2.1 언어 자료 수집 방법 | 111 |
| 2.2 언어 자료 수집을 위한 지역 수와 선정 기준 | 111 |
| 2.3 언어 자료 수집을 위한 주체 | 111 |
| 2.4 전사 방안 | 111 |
| 2.5 농인 수어 연구자 양성 | 114 |
| 2.6 한국수어 말뭉치 기반 수어 연구 | 114 |
| 2.7 한국수어 말뭉치 구축 자료 공개 | 114 |
| 2.8 한국수어 말뭉치 구축 자료 접근성 | 116 |
| 2.9 한국수어 말뭉치 구축 로드맵 | 116 |
| 3. 워크숍 개최 | 119 |
| VI. 한국수어 말뭉치 구축 지침 | 121 |
| 1. 2017년 개정 방향 | 121 |
| 2. 언어 수집 과정에서의 수정 보완 | 121 |
| 2.1 촬영 시간 | 121 |
| 2.2 두 언어 제공자 간의 관계 | 122 |
| 2.3 사회자의 역할 | 122 |
| 2.4 촬영 환경 | 123 |

| | |
|--|-----|
| 2.5 촬영 동영상의 편집 및 관리 | 123 |
| 3. 전사 지침에서의 수정 보완 | 124 |
| VII. 결론 | 127 |
| 1. 연구 결과 요약 | 127 |
| 1.1 기구축 한국수어 말뭉치의 정비 | 127 |
| 1.2 전사 | 128 |
| 1.3 한국수어 말뭉치 기반 연구 사례 | 129 |
| 1.4 한국수어 말뭉치 구축 계획 | 130 |
| 1.5 한국수어 말뭉치 구축 지침 개정 | 133 |
| 2. 활용 방안 및 제언 | 133 |
| 2.1 연구의 활용 및 기대 효과 | 133 |
| 2.2 제언 | 135 |
| VIII. 참고문헌 | 137 |
| <부록 1> 과제별 연결 실패 건수 | 141 |
| <부록 2> 타입 삭제, 타입 수정, 새로운 타입 등록 목록 | 147 |
| <부록 3> ELAN(EUDICO Linguistic Annotator)과 수어 검색·분석 | 158 |
| <부록 4> 2017년 개정 한국수어 말뭉치 구축 지침 | 198 |
| <부록 5> 언어 제공자 모집 공문 | 250 |
| <부록 6> 2017년 개정 질문지 | 251 |
| <부록 7> 영상 촬영, 배포, 판권 소유 및 자료 활용 동의서 | 258 |
| <부록 8> 개인 정보 수집 · 이용 · 제공 및 조회 동의서 | 259 |
| <부록 9> 언어 제공자와 사회자 관계에 관한 질문지 | 260 |
| <부록 10> 촬영 일지 | 261 |

【표 목차】

| | |
|--|----|
| <표 1> 고정된 어휘와 생산적 어휘를 구분하는 기준 | 16 |
| <표 2> 연결 실패 이유와 해결 방법 | 37 |
| <표 3> 기존 [지시] 타입 목록 | 42 |
| <표 4> 지시의 8가지 종류 | 43 |
| <표 5> 기존 Counting hand에 해당되는 타입 목록 | 43 |
| <표 6> 기존 [구형]에 해당되는 타입 목록 | 46 |
| <표 7> 기존 [육형]에 해당되는 타입 목록 | 47 |
| <표 8> 추진 내용 및 방법 | 51 |
| <표 9> 전사 교육 내용 | 52 |
| <표 10> 추진 내용 및 방법 | 53 |
| <표 11> 공동 전사 추진 내용 | 53 |
| <표 12> 전사자 신뢰도 검사 결과 | 56 |
| <표 13> 각 과제별 세부 정보 | 72 |
| <표 14> 갈색을 나타내는 수어 타입과 언어 제공자 수 | 75 |
| <표 15> 분홍을 나타내는 수어 타입과 언어 제공자 수 | 76 |
| <표 16> 보라를 나타내는 수어 타입과 언어 제공자 수 | 78 |
| <표 17> 주황을 나타내는 수어 타입과 언어 제공자 수 | 80 |
| <표 18> 초록을 나타내는 수어 타입과 언어 제공자 수 | 82 |
| <표 19> 파랑을 나타내는 수어 타입과 언어 제공자 수 | 82 |
| <표 20> 회색을 나타내는 수어 타입과 언어 제공자 수 | 83 |
| <표 21> 하양을 나타내는 수어 타입과 언어 제공자 수 | 83 |
| <표 22> 빨강을 나타내는 수어 타입과 언어 제공자 수 | 83 |
| <표 23> 노랑을 나타내는 수어 타입과 언어 제공자 수 | 84 |
| <표 24> 2개 수어로 색을 표현한 언어 제공자 수 | 84 |
| <표 25> 강원도 | 85 |
| <표 26> 남대문 | 86 |
| <표 27> 동대문 | 86 |
| <표 28> 부산 | 87 |
| <표 29> 용산 | 87 |

| | |
|-----------------------------------|-----|
| <표 30> 이리 / 익산 | 88 |
| <표 31> 인천 | 88 |
| <표 32> 울산 | 88 |
| <표 33> 일본 | 89 |
| <표 34> 넥타이 | 91 |
| <표 35> 샤워 | 91 |
| <표 36> 수영하다 | 92 |
| <표 37> 신호등 | 92 |
| <표 38> 아래 | 93 |
| <표 39> 아침 | 93 |
| <표 40> 오토바이 | 94 |
| <표 41> 두 자문 집단에게 의뢰한 자문 항목 | 96 |
| <표 42> 한국수어 말뭉치 연차별 추진 내용 | 117 |
| <표 43> 생산적 수어의 글로스 | 236 |
| <표 44> 입의 움직임을 나타내는 기본적인 기호 | 245 |

【그림 목차】

| | |
|---|----|
| <그림 1> 독일수어 전자사전 | 6 |
| <그림 2> 독일수어 전자사전 검색 방법 연구 | 7 |
| <그림 3> 변이형 연구 | 8 |
| <그림 4> [원하다] | 9 |
| <그림 5> [여자] | 9 |
| <그림 6> [또는] | 10 |
| <그림 7> [지금] | 10 |
| <그림 8> [이사하다] | 11 |
| <그림 9> [이사하다]의 나이별 빈도 | 11 |
| <그림 10> 변이형 연구 분석 도구 | 12 |
| <그림 11> 수어 말뭉치에서 전사를 통해 정리 되는 수어 체계 | 14 |
| <그림 12> 수어 사전에서의 수어 체계 | 14 |
| <그림 13> 핀란드수어 [BLACK] | 24 |

| | |
|---|----|
| <그림 14> 핀란드수어 독립된 단어 [TO-SEND] 산출 표상 | 27 |
| <그림 15> 주석 통계 | 29 |
| <그림 16> 정규 표현식을 선택하여 [지시]를 검색할 경우 | 29 |
| <그림 17> 검색된 [지시]의 검색 범위를 좁혔을 경우 | 29 |
| <그림 18> [지시]와 [구형]을 함께 사용한 주석을 검색한 예 | 29 |
| <그림 19> 오타 [아니다11]을 검색하여 [아니다1]로 수정한 예와 그 결과 | 30 |
| <그림 20> 다수의 파일 층렬별 빈도 분석 | 30 |
| <그림 21> 다수의 전사 파일 검색 | 31 |
| <그림 22> 다수의 전사 파일 검색 결과 | 31 |
| <그림 23> 다중 파일 검색 및 변경 | 32 |
| <그림 24> Substring Search | 33 |
| <그림 25> Single Layer Search | 33 |
| <그림 26> 다중 파일 검색 예 | 35 |
| <그림 27> 비우세 [counting hand]와 우세 [지시1_비1] | 44 |
| <그림 28> [눈1] | 45 |
| <그림 29> [미원2] | 46 |
| <그림 30> 엘란(Elan) 전사 예 | 50 |
| <그림 31> 전사 교육 장면 | 52 |
| <그림 32> 공동 전사 및 재교육 모습 | 54 |
| <그림 33> 일치도가 낮은 토큰- [구형1] | 57 |
| <그림 34> 일치도가 낮은 토큰- [풍물1] | 58 |
| <그림 35> 일치도가 낮은 토큰- [일1] | 59 |
| <그림 36> 일치도가 낮은 토큰- [들어가다2] | 59 |
| <그림 37> 일치도가 낮은 토큰- [촬영11] | 60 |
| <그림 38> 일치도가 낮은 토큰- [희생1] | 61 |
| <그림 39> 일치도가 낮은 토큰- [지시1_비1] | 61 |
| <그림 40> 일치도가 낮은 토큰- [지시1_1] | 62 |
| <그림 41> 일치도가 낮은 토큰- [지시2_1] [지시2_1] / [이십살1] | 63 |
| <그림 42> 일치도가 낮은 토큰- [지시1_1] | 64 |

| | |
|---|-----|
| <그림 43> 일치도가 낮은 토큰- [농1] | 64 |
| <그림 44> 일치도가 낮은 토큰- [아래1] | 65 |
| <그림 45> 일치도가 낮은 토큰- [지시2호] | 65 |
| <그림 46> 일치도가 낮은 토큰- [이전1] | 66 |
| <그림 47> 일치도가 낮은 토큰- [지시2_1] | 67 |
| <그림 48> 일치도가 낮은 토큰- [대부분1] / [무엇1] | 68 |
| <그림 49> 일치도가 낮은 토큰- [지시1_1] | 69 |
| <그림 50> 일치도가 낮은 토큰- [없다11] | 69 |
| <그림 51> 일치도가 낮은 토큰- [제스처] | 70 |
| <그림 52> 일치도가 낮은 토큰- [완전1] | 71 |
| <그림 53> 갈색: [밤1]+[약하다1] / [갈색1]+[약하다1] - 50대 | 75 |
| <그림 54> 분홍: [빨강1]+[약하다1]-70대 | 76 |
| <그림 55> 분홍: [빨강1]+[약하다1] - 60대 | 76 |
| <그림 56> 분홍: [빨강1]+[약하다1]-50대 | 77 |
| <그림 57> 분홍: [빨강1]+[퍼지다1]-60대 [빨강1]+비수지-30대 | 77 |
| <그림 58> 보라: [포도1]+[색1] - 50대 | 79 |
| <그림 59> 보라: [빨강1]+[파랑1]+[중1] [빨강1]+[파랑1]+[섞다1] -70대 | 79 |
| <그림 60> 주황: [밤1]+[노랑1] / [수화1]+[없다1] - 60대 | 81 |
| <그림 61> 주황: [ㅂ ㅓ ㄴ]/ [노랑1]+[빨강1] [여자1]+[예쁘다]+[웃을 입다] - 60대 | 81 |
| <그림 62> 주황: [하양1]+[밤1]/ [밤1]+[빨강1] - 70대 | 81 |
| <그림 63> 워크숍 모습 | 119 |
| <그림 64> 유도 자료 등록 | 206 |
| <그림 65> 언어 제공자 등록 창 | 209 |
| <그림 66> 촬영실 전체 배치도 | 210 |
| <그림 67> 촬영실 모습 | 211 |
| <그림 68> 원본 영상 정보 등록 | 213 |
| <그림 69> 전사 영상 등록 | 216 |
| <그림 70> 엘란(Elan) 전사 예 | 217 |

| | |
|-------------------------------|-----|
| <그림 71> 기본적인 총렬 | 220 |
| <그림 72> 의미 총렬 | 221 |
| <그림 73> [소개] | 222 |
| <그림 74> [먹다1] | 222 |
| <그림 75> [먹다2] | 222 |
| <그림 76> [먹다3] | 222 |
| <그림 77> [궁궐] | 223 |
| <그림 78> [중국] | 223 |
| <그림 79> [공] | 224 |
| <그림 80> [큰 공] | 224 |
| <그림 81> [시골] | 224 |
| <그림 82> [시골] | 225 |
| <그림 83> [없다1] | 225 |
| <그림 84> [없다2] | 225 |
| <그림 85> [없다3] | 225 |
| <그림 86> ID 글로스의 구조 | 226 |
| <그림 87> [꼼꼼하다] | 227 |
| <그림 88> [버스] | 227 |
| <그림 89> [버스가다] | 227 |
| <그림 90> [없다1] | 228 |
| <그림 91> [없다2] | 228 |
| <그림 92> [없다3] | 228 |
| <그림 93> [물건싸다] | 228 |
| <그림 94> [값싸다] | 229 |
| <그림 95> [십1], [십2] | 229 |
| <그림 96> [일년1] | 229 |
| <그림 97> [이분1] | 230 |
| <그림 98> [counting hand] | 230 |
| <그림 99> [버스1], [버스가다1] | 231 |
| <그림 100> [야단치다] | 231 |
| <그림 101> [수당1] | 232 |

| | |
|------------------------------|-----|
| <그림 102> [빨강1], [약하다1] | 232 |
| <그림 103> [오다], [가다] | 233 |
| <그림 104> [오가다] | 233 |
| <그림 105> [편안], [산] | 233 |
| <그림 106> [안산] | 234 |
| <그림 107> [거짓], [-면] | 234 |
| <그림 108> [만약] | 234 |

I. 서론

1. 연구 목적 및 필요성

한국수어 연구의 활성화와 과학적인 연구 기반 제공을 위해 국립국어원에서는 2015년 ‘한국수어 말뭉치 기반 조사 및 시범 구축’ 사업을 진행하였다. 이 시범 구축 사업에서는 한국수어 말뭉치 구축을 위한 지침서를 개발하고, 서울 지역의 농인 30쌍을 대상으로 90시간 분량의 언어 자료를 수집하여 그 중에서 13시간 분량의 기본 전사를 실시하였다. 그러나 이 시범 구축 사업에서는 문자 체계가 없는 수어의 특성으로 인한 전사의 어려움이 있었고, 또한 전사를 위한 시스템의 부족으로 전사자 간의 신뢰도를 확보하기 어려웠다.

문자가 없는 수어의 기록은 음성 언어의 기록과는 다르기 때문에 한국수어 말뭉치 구축 과정에서 전사의 일치도를 확보하는 데에 큰 어려움이 있었다. 각 나라에서는 수어를 전사하기 위해서 자국의 문자 체계를 사용하여 모든 의미 있는 손 활동에 이름, 즉 글로스(gloss)를 붙여 기록한다. 이때 사용하는 글로스는 해당 수어를 쉽게 떠올릴 수 있도록 해당 수어의 의미와 관련이 있는 이름을 사용한다. 하지만 그 글로스가 의미 그 자체는 아니기 때문에 대역어와 혼동을 가져오기도 한다. 따라서 전사 과정에서는 수어 타입과 글로스를 확인하고 전사자 간의 전사 일치도를 확보하기 위해 주기적으로 전사 회의를 실시해야 하는 과정 등 많은 시간과 노력을 필요로 한다.

2015년 시범 구축한 한국수어 말뭉치 구축에서도 다른 나라 수어 말뭉치와 마찬가지로 한국어의 문자인 한글을 사용하여 한국어 단어의 의미와 관련된 글로스를 통한 전사를 하였다. 전사는 타입명의 결정 및 전사의 일치도 확보를 위해 일주일에 한 번씩 전사 회의를 하면서 진행하였다. 그러나 시스템도 없는 상황에서 다수의 전사자들이 한국어 글로스를 사용한 전사는 타입명의 불일치 및 중복 등 오류가 많이 발견될 수밖에 없었다. 이에 2016년에는 ‘한국수어 자료 통합 지원 시스템 구축’ 사업이 진행되어 수어 자료 전사를 위한 시스템을 마련하

고 그에 따른 전사 지침을 수정 보완하였다.

이에 본격적으로 한국수어 말뭉치 구축 사업을 진행하기 전, 기존에 전사된 자료를 2016년에 개발한 한국수어 자료 통합 시스템 상에 등록하는 과정에서 발생한 타입과 토큰의 연결 문제 등 시스템 내에서의 한국수어 전사 자료를 정비할 필요가 있다. 또한 향후 본격적인 한국수어 말뭉치 구축을 위해 기 수집된 한국수어 자료를 2016년에 만든 시스템에서 전사를 실시해 봄으로써 시스템 내에서의 한국수어 자료 구축 및 전사 지침을 개선·보완하고, 향후 보다 체계적이고 효율적인 한국수어 말뭉치 구축을 위한 방안을 제시할 필요가 있다.

따라서 본 연구에서는 다음의 연구 목적을 설정하였다.

첫째, 2015-2016년에 전사 완료한 한국수어 말뭉치 중 한국수어 자료 통합 지원 시스템 내에서의 연결 오류를 파악하여 수정하고 정비한다.

둘째, 2015년에 기구축한 한국수어 말뭉치를 검토하여 문제점을 파악하고 그 해결 방안을 제시하고 구축 지침서를 개선·보완하여 제공한다.

셋째, 기구축 한국수어 말뭉치의 미 전사된 언어 자료 중에서 수지를 대상으로 250분 분량의 전사를 실시하고 전사 결과를 분석하고, 한국수어 말뭉치 기반 수어 연구 사례를 제공한다.

넷째, 향후 한국수어 말뭉치 구축 방향 및 방법론을 제시하고 한국수어 말뭉치 구축의 로드맵을 제시한다.

2. 연구 방법 및 범위

2.1 기구축 한국수어 말뭉치 검토 및 문제점 분석

- 기구축 한국수어 말뭉치(2015년)에 대한 문제점 검토

- 기구축 한국수어 말뭉치 시스템에서의 연결 오류 유형 분석
- 기구축 한국수어 말뭉치 시스템에서의 연결 문제 해결 방안 모색
- 2015-2016년 지침에 따른 한국수어 말뭉치 정비

2.2 기구축 한국수어 말뭉치 전사 및 분석 연구 수행

- 전사자 교육 내용 확정
- 전사자 교육 실시
- 기구축 한국수어 말뭉치에서 수지 대상 전사 실시(250분 이상)
- 전사된 한국수어 말뭉치 기반 사례 연구

2.3 향후 한국수어 말뭉치 구축 방향 및 방법론 마련

- 국외 수어 말뭉치 구축의 최신 동향 조사 및 선행 연구 분석
- 향후 한국수어 말뭉치 구축 계획 제시
- 사업 내용 관련 전문가 자문회의 및 워크숍 개최

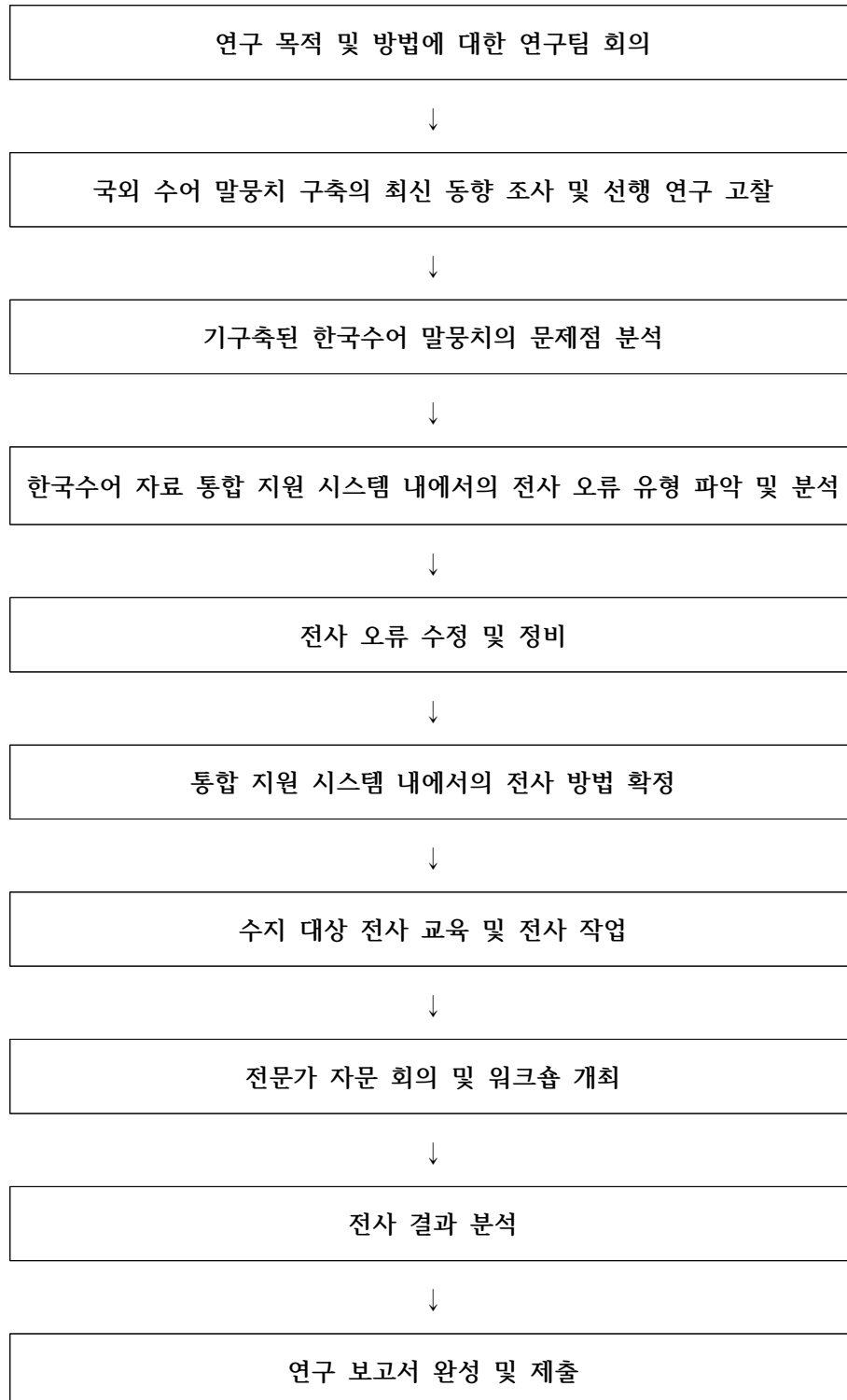
2.4 한국수어 말뭉치 구축 지침서 개정

- 기구축 한국수어 말뭉치에서 나타난 수정 보완 사항 파악
- 한국수어 자료 통합 시스템 내에서의 한국수어 말뭉치 구축 지침 보완 및 개선

- 2015년 기구축한 한국수어 말뭉치의 검토 및 문제점 분석은 2016년에 구축한 한국수어 자료 통합 지원 시스템 내에서 실시한다.
- 수지를 대상으로 하는 250분에 대한 전사는 ‘한국수어 말뭉치 기반 조사 및 시범 구축’에서 수집한 언어 자료 중 미 전사된 자료를 대상으로 한다.
- 한국수어 말뭉치 기반 수어 연구 사례는 2015년 ‘한국수어 말뭉치 기반 조사 및 시범 구축’에서 기 전사하고 2017년에 수정·보완한 전사한 자료와 2017년에 추가로 250분 전사한 자료 중 연

구 목적에 따라 대상을 선택하여 분석한다.

연구는 다음의 절차에 따라 실시한다.



II. 선행 연구 고찰

1. 독일의 수어 말뭉치 기반 수어 연구 현황

1.1 독일수어 말뭉치 연구 사업 소개

독일수어 말뭉치 연구 사업은 2009년 시작했으며 2023년까지 수행할 예정이다. 사업 지원 기관은 Akademie der Wissenschaften Hamburg 이며 연구 기관은 함부르크 대학교 수화언어학과 연구소다. 독일수어 말뭉치 연구 사업은 두 목적을 가지고 있다. 독일수어 말뭉치를 구축하고 수어 말뭉치 기반 독일수어 전자사전 편찬이다.

독일수어 말뭉치를 구축하기 위해 전국 12곳에서 총 330명의 농인 언어 제공자를 모집하였다. 165쌍에게 다양한 유도 자료를 제시하여 언어 데이터를 수집하였으며 7대의 카메라로 촬영했다. 촬영한 수어 자료는 총 560시간 분량이다(대략 750TB). 현재 수어 영상 65시간의 기본 전사를 완료하였고 전사된 토큰 수는 425,000개이다(2017.7.18. 기준). 전사는 함부르크 대학교, 수화언어학과 연구소에서 개발한 iLex 프로그램으로 한다. iLex는 수어 전사 프로그램이자 데이터베이스이다.

독일수어 말뭉치 연구 프로젝트는 아래와 같은 연구 제목들을 다루고 있다.

1.2 독일수어 말뭉치 연구 주제

- 수어의 변이형과 빈도수
한 수어의 여러 개의 변이형이 존재 할 경우 사용 빈도가 가장 높은 변이형이 어떤 것인지 토큰 수로 분석한다.
- 수어의 지역 분포
존재하는 변이형이 어떤 지역에서 사용되는지 살펴본다.
- 수어의 공간 사용
수어의 공간적 사용이 어떤 형태로 표현되는지 살펴본다.
- 수어의 의미



어느 맥락에서 어떤 의미로 수어가 사용 되는지 분석한다.

- 수어의 연어

주로 어떤 수어 단어들이 함께 나타나는지 분석한다.

1.3 독일수어 전자사전

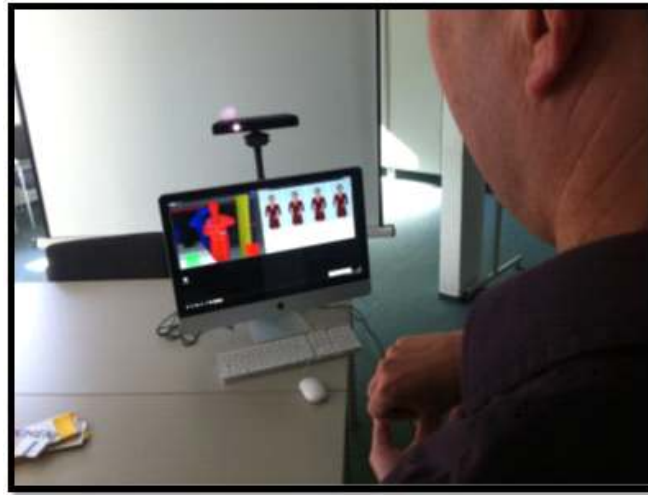
위 연구 주제의 결과는 바로 사전 편찬에 활용할 계획으로 사전에 수어 말뭉치에서 뽑아낸 용례를 활용할 예정이다. 아직 사전 편찬은 진행되지 않았지만 아래와 같은 표제항이 하나의 예이다.

| | | | |
|--|---|---|---|
| <p>ENDE</p>  <p>Variante 1 Variante 2</p>  | <p><u>Anmerkung zur Form:</u> Kann auch mit Wiederholung ausgeführt werden.</p> <p><u>Raumnutzung:</u> Gebärde kann – je nach Bezug – an verschiedenen Orten im Gebärdenraum ausgeführt werden.</p> <p><u>Verbreitung:</u> deutschlandweit</p> <p><u>formgleich/-ähnlich:</u> Variante 2: → STEMPEL (Variante 1), → ARBEITEN2</p> | <p><u>Bedeutungen:</u></p> <ol style="list-style-type: none"> 1. [vollständig/bis zum Abschluss bearbeiten]: erledigen, schaffen, beenden, fertig machen (z.B. Aufgabe, Arbeit) 2. (mit Erfolg) abgeschlossen, beendet (z.B. Schule, Ausbildung) 3. [Signal zur zeitlichen Gliederung]: vorbei, zu Ende, abgeschlossen; <p><u>typische Konstruktion:</u> [Ereignis/Tätigkeit] ENDE DANN ...</p> <ol style="list-style-type: none"> 4. nur das, mehr nicht, das war's, das ist alles; <p><u>typische Position:</u> [letzte Gebärde in der Äußerung]</p> <ol style="list-style-type: none"> 5. ... | <p><u>Beispiele:</u></p>  |
|--|---|---|---|

<그림 1> 독일수어 전자사전

독일수어 전자사전은 표제어 6,000개를 목적으로 한다. 일반적으로 수화언어학자들은 수어의 특성을 감안해서 사전 표제어 5,000-6,000개가 적합하다고 본다.

독일수어 말뭉치 연구 사업에서는 수어 전자사전의 검색을 여러 방법으로 접근하려고 한다. 글자로 검색하는 방법, 수어의 형태로 검색하는 방법과 더불어 카메라 인식을 통해 사전 사용자가 검색하고자 하는 수어를 카메라 앞에서 보여주면 해당 표제어를 찾는 방법도 연구하고 있다.



<그림 2> 독일수어전자사전 검색 방법 연구

1.4 독일수어 변이형 연구

같은 의미를 갖고 있지만 수어 형태가 살짝 다른 경우가 있다. 그것을 변이형이라고 한다. 수어는 글이 없기 때문에 음성 언어와 달리 변이형이 많다. 존재하는 변이형들이 지역과 관계가 있는지 조사하기 위해서 독일수어 말뭉치 연구팀은 여러 개의 수어를 통계 내어 지도로 표현했다. 또한 변이형 연구에 언어 제공자의 나이도 감안하였다.

Neue Suche Suche über Gebärdenform ▾

Die bisher eingestellten Suchparameter [\[Hilfe\]](#)

Bitte wählen Sie die gewünschte Anzahl der Hände: [\[Hilfe\]](#)

Bitte wählen Sie die gewünschte Handform: [\[Hilfe\]](#)

Bitte wählen Sie die gewünschte Handstellung: [\[Hilfe\]](#)

Bitte wählen Sie die gewünschte Lokation: [\[Hilfe\]](#)

Bitte wählen Sie die gewünschte Bewegung: [\[Hilfe\]](#)

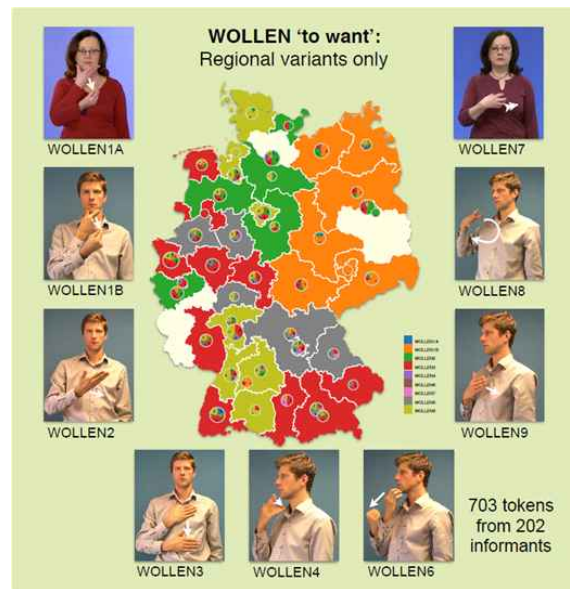
führen zu 1310 (von 1310) Treffern:

<그림 3> 변이형 연구

1.4.1 독일수어 [원하다]¹⁾

지도를 통해 알 수 있듯이 [원하다] 수어는 기본형, 보편적으로 사용되는 수어가 없다. 전국에서 사용되는 [원하다] 수어는 9개이다.

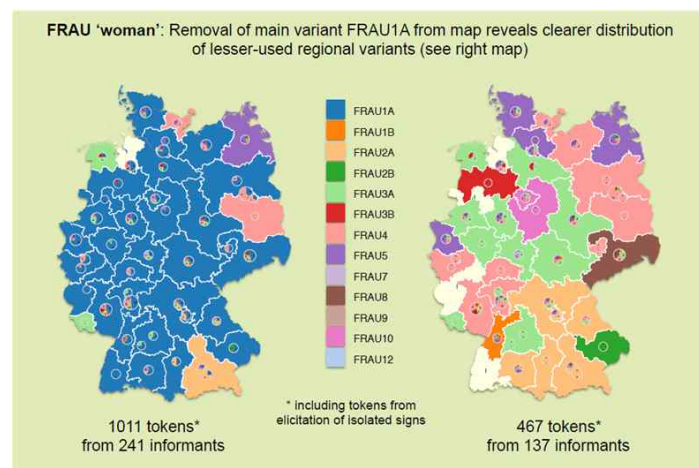
1) []는 수어 단어, 글로스를 의미한다. 알파벳을 사용하는 나라에서는 이를 표현하기 위해 보통 대문자로 사용하나 한글을 사용하는 우리나라 수어 연구자들은 []을 사용하여 수어 단어를 나타낸다. 본 보고서에서는 수어 단어, 타입을 포함한 글로스를 []로 표기하였다.



<그림 4> [원하다]

1.4.2 독일수어 [여자]

첫 지도를 보면 [여자] 수어는 변이형이 존재하지 않고 거의 전국적으로 똑같은 수어가 사용되는 것으로 보인다. 그러나 가장 높은 빈도수를 가진 [여자] 수어를 제외하면 다양한 변이형이 보인다.

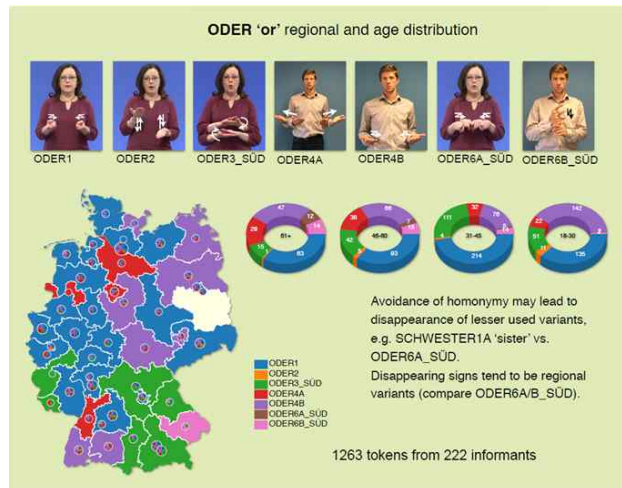


<그림 5> [여자]

1.4.3 독일수어 [또는]

[또는] 수어의 변이형은 7개인 것으로 나타났다. 7개의 변이형을 나이별로 분석하면 언어 제공자들이 젊어질수록 [여동생] 수어와 같은

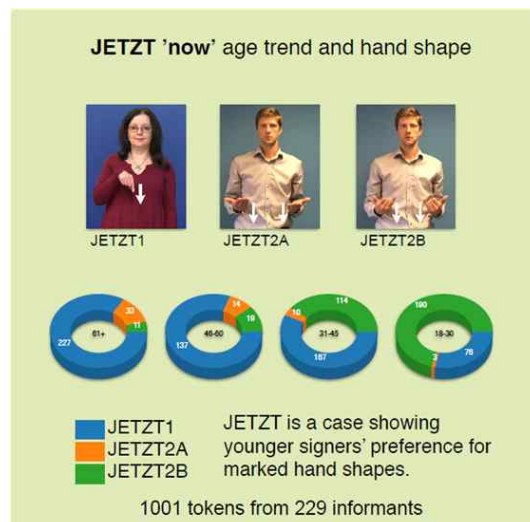
형태를 가진 수어는 사용하지 않는 것을 알 수 있다. 변이형 중 다른 수어와의 동음이의어(Homonym)가 있을 경우 그 변이형은 없어질 확률이 제일 크다.



<그림 6> [또는]

1.4.4 독일수어 [지금]

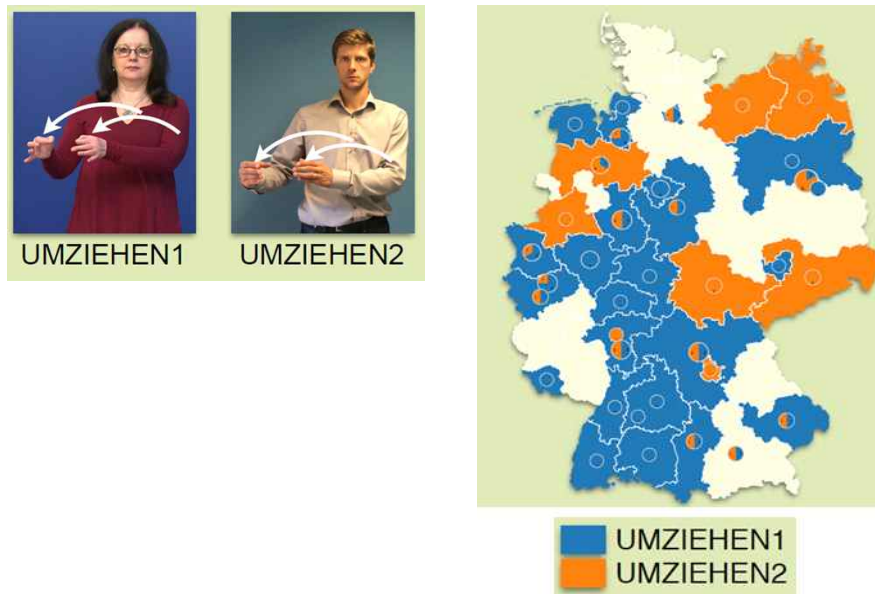
[지금] 수어는 변이형이 3개로 나타난다. 나이별로 분석하면 젊은 언어 제공자는 단순한 수형을 가진 변이형보다 좀 세련된 수형을 가진 수어를 선호한다는 것을 알 수 있다.



<그림 7> [지금]

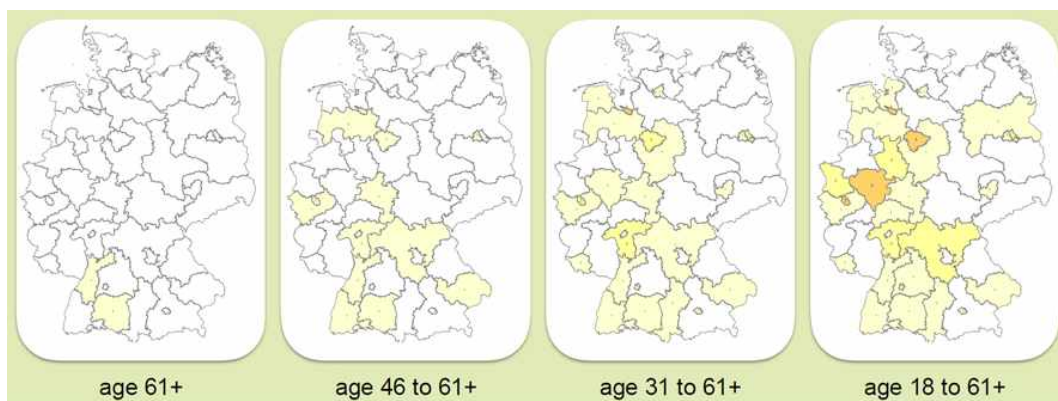
1.4.5 독일수어 [이사하다]

[이사하다] 수어는 2개의 변이형이 있다.



<그림 8> [이사하다]

지도로 보면 두 변이형은 그저 방언 같지만 지도를 언어 제공자의 나이별로 살펴보면 첫 번째 변이형은 주로 젊은 언어 제공자들이 사용하고 두 번째 변이형은 없어지는 추세라는 것도 알 수 있다.

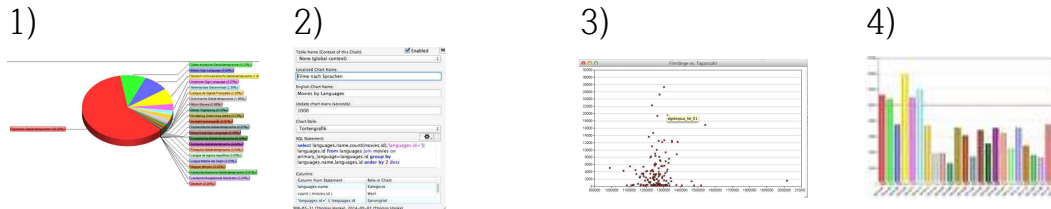


<그림 9> [이사하다]의 나이별 빈도

1.5 독일수어 말뭉치 분석 도구

독일수어 말뭉치 연구진은 1.4에서 기술된 변이형 연구를 진행하기 위해서 아래와 같은 분석 도구들을 개발하였다.

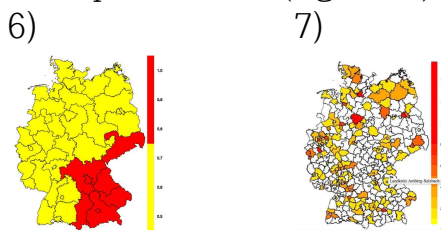
* Business Charts in iLex (fig.1-4)



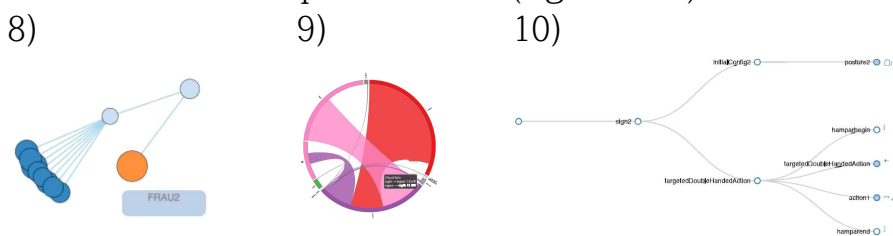
* Graphs, Nets and Lattices in iLex (fig. 5)



* Maps in iLex (figs.6-7)



* Interactive Graphs in iLex (figs. 8-10)



<그림 10> 변이형 연구 분석 도구

1.6 독일수어 말뭉치와 수어 사전 편찬

독일수어 말뭉치 연구진은 독일수어 말뭉치 기반 수어 사전 편찬을 준비하면서 독일수어 말뭉치에서 사용하는 전사 방법들이 수어 사전을 위한 전사 방법과 다르다는 것을 알게 되었다. 독일수어 말뭉치에서 사용하는 전사 방법과 수어 사전을 위한 전사 방법의 차이는 아래와 같다.

1.6.1 독일수어 말뭉치 전사

- 토큰 타입 매칭을 위해 타입 목록이 필요하다.
- 수어 구분을 큼직하게 하여 우선 수어의 형태를 보고, 의미를 그 다음으로 본다.

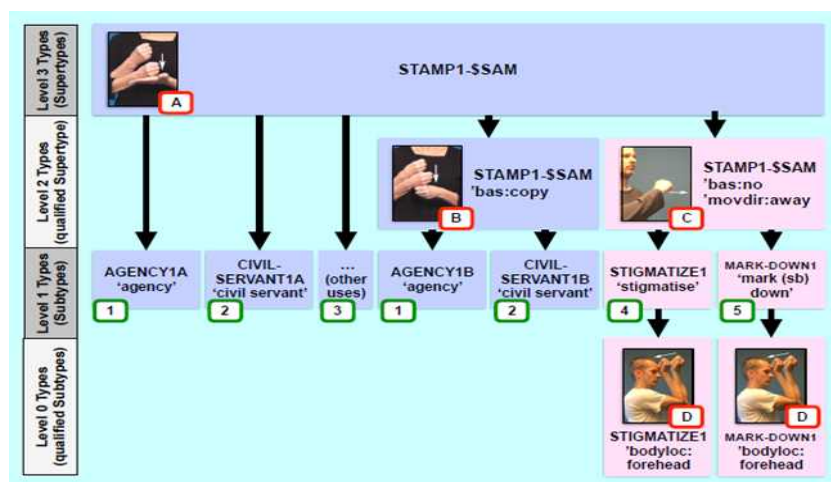
- 독일수어 말뭉치 전사에서는 변이형, 어형, 형태론적 수정(modification)²⁾ 등은 하나로 묶어 전사할 수도 있다.
- 독일수어 말뭉치 전사는 결과보다 과정이다.
- 전사를 할수록 더 많은 정보를 얻게 됨으로 과거 적용된 전사 규칙이 변경되어야 한다.
- 독일수어 말뭉치 전사를 완료하는 것은 거의 불가능하다고 봐야 한다. 연구 주제에 따라 새로운 질문이 생기고 그에 따른 추가 전사가 필요하다.

1.6.2 수어 사전 편찬

- 수어 사전 형식으로 수어를 기술하기 위해서는 꾸준한 독일수어 말뭉치 전사가 필요하다.
- 우선적으로 수어의 의미와 수어의 사용을 본다.
- 사전 규칙을 따라 표제어를 작성해야 한다.
- 수어 사전이라는 것은 한 순간을 기준으로 전사한 것을 확정해야 한다.
- 모든 표제어는 같은 기준으로 작성 되어야 한다.

독일수어 말뭉치에서 전사를 통해 정리 되는 수어 체계는 수어 사전에서 보는 수어 체계와 다르다.

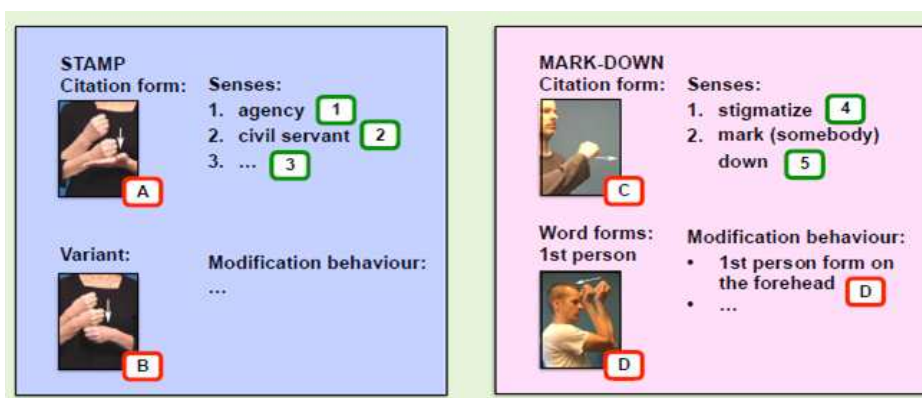
2) Langer, G./Köig, S./Konrad, R.(2007)에 의하면 수어를 하나의 개별적인 수어로 보지 않고 여러 개를 비교하여 보면 수어들끼리 서로 관계가 있는 것을 알 수 있다. 하나의 기저 이미지 안에 여러 수어가 있고, 그 수어들끼리는 서로 연관된다. 도상성이 있는 어휘화된 수어들은 대개 형태론적인 수정이 가능하다. 형태론적인 수정이란 기저 이미지의 한 수어의 형태에서 의미를 확장하거나, 의미를 더욱 자세하게 하기 위해 수어의 형태를 수정하는 것이다. 의미를 확장하거나 더 자세하게 표현하기 위해서 기저 이미지에 정보를 추가한다. 연관된 수어들은 맥락에 맞추어 수정되는 경우가 많으며 어떤 경우에는 맥락에 맞춰서 나타나는 수어의 다양한 형태들이 의미가 추가되어 형태론적인 수정형으로 사용될 때도 있고 특별한 의미로 어휘화 되어 사용되기도 한다.



수어말뭉치

<그림 11> 독일수어 말뭉치에서 전사를 통해 정리 되는 수어 체계

독일수어 말뭉치에서는 형태를 중심으로 전사하기 때문에 수어의 형태가 비슷하거나 기저 이미지가 같으면 한 수어 가족이라고 생각한다. 그러나 똑같은 두 수어를 사전 편찬의 시각으로 보면 두 수어는 각각의 표제어로 정리 된다.



수어사전

<그림 12> 수어 사전에서의 수어 체계

2. 수어의 어휘 체계

수어의 어휘 체계를 검토하기 위해 비언어적인 시각적-제스처 행위와 언어적인 시각적-제스처 행위를 구별하는 것이 필요하다. 비언어적인 시각적-제스처 행위에는 몸짓, 관습화된 제스처, 마임 등이 속한다. 언어적인 시각적-제스처 행위에는 수어의 생산적 어휘(productive signs)가 속한다. 생산적 어휘는 농인들이 일관성 있게 다시 산출할 수

있고 결과적으로 특정한 합의된 값이 수형, 수향, 수위, 수동에 부여될 수 있는 비교적 안정적이고 구별이 가능한 시각적-제스처 행위이다 (Johnston & Schembri, 1999: 115). 생산적 어휘의 구성소들이 도상적일지라도 의미와 형태가 관습적으로 연결되어 비언어적인 시각적-제스처와 구별된다. 제스처와 생산적 어휘는 관습적 기준에 따라 구분되고 다시 생산적 어휘는 어휘소와 구별된다. 생산적 어휘는 해당 언어의 가능하거나 혹은 실제적인 언어적 단위들로 생성된 의미를 가진다면 어휘소는 해당 언어의 분리된 단위들에 관습적으로 주어진 의미를 부여한다. 어휘소는 확인과 반복이 가능한 인용형을 가진다. 합성적 의미보다 더 특징적이거나 예측이 가능하지 않고 상대적으로 맥락에 덜 의존적인 의미를 가진다. 수어에서 어휘소는 확립된 어휘(established signs) 혹은 고정된 어휘(frozen signs)로 달리 표현할 수 있다. 어휘소는 수어 사전에 표제어로 들어갈 후보인데 반해, 생산적 어휘는 문법 혹은 사전의 문법 부분에서 취급되어야 한다(같은 책: 136-7쪽).

한국수어 말뭉치의 전사에서 생산적 어휘와 고정된 어휘의 구분은 매우 중요한 과정이다. 둘 중 어느 것에 속하나에 따라 글로스가 달리 붙여진다. 또한 한국수어 말뭉치의 전사 과정에서 변이형(variation)과 형태론적 수정(modification)의 구분도 중요하다. 그러므로 고정된 어휘와 생산적 어휘의 구분과 변이형과 수정의 차이를 차례대로 살펴보고 하겠다.

2.1 고정된 어휘와 생산적 어휘

수어에서 고정된 어휘와 생산적 어휘를 구별하는 음운적, 형태 통사적, 의미적 기준과 그 밖의 기준은 <표 1>과 같다(같은 책: 136쪽).

<표 1> 고정된 어휘와 생산적 어휘를 구분하는 기준

| 기준 | 고정된 어휘 | 생산적 어휘 |
|-----------------|--|--|
| 음운적 기준 | 생산적 어휘보다 비교적 작은 범 위의 수형, 수향, 수위, 수동을 사용하는 인용형을 가짐 | 비교적 넓은 범위의 수형, 수향, 수위, 수동을 사용 |
| | 수형, 수향, 수위, 수동, 비수지 자질에서 이형태 변이형을 야기 | 수형, 수향, 수위, 수동, 비수지 자질에서 작은 변이형은 의미의 변화를 야기 |
| | 우세 조건과 대칭 조건과 같은 특정한 음운적 제약을 준수 | 어휘소에서 발견되는 우세 조건과 대칭 조건과 같은 모든 음운적 제 약을 준수하지 않을 수 있음 |
| 형태 통사적 기준 | 단일 형태소일 수 있음: 수어 단 어의 의미는 비교적 자의적이거 나 단어의 합성적 의미와 관련하 여/보다 더 특정적임 | 다형태소적임: 수어 단어의 의미 는 의미에 기여하는 특정한 수형, 수향, 수위, 수동과 함께 합성적 임 |
| | 의미는 맥락을 떠나서 비교적 분 명함 | 의미는 비교적 맥락-의존적임 |
| | 모든 품사로 분류됨: 명사, 동사, 의문사, 수식어, 연결어, 감탄사 | 맥락 속에서 명사로 행동할지라도 수어 발화에서 서술어 역할을 수 행할 가능성이 높음 |
| 의미적 기준 | 수어 단어의 형태와 의미의 관계 가 방언과 다른 수어들에서 다양 함 | 형태와 의미의 관계가 방언과 다 른 수어들에서 비교적 덜 다양함 |
| | 어휘소를 사용하는 동안 수어 발 화자는 상대방과 눈 맞춤의 가능 성이 더 높음 | 생산적 어휘를 사용하는 동안 수 어자는 상대방과 눈 맞춤의 가능 성이 낮고 손을 볼 가능성이 더 높음 |
| | 어휘소를 사용하는 동안 수어자 는 음성 언어 어휘 항목의 입 패 턴을 동반할 가능성이 높음 | 생산적 어휘를 사용하는 동안 수 어자는 음성 언어의 어휘 항목의 입 패턴을 동반할 가능성이 낮음 |
| 다른 기준 | | |
| | | |

2.2 변이형과 형태론적 수정

여러 나라 수어에서 사회 언어적 변이형이 관찰되었다. 변이형이란 의미에서 변화가 최소한이라면 표준 형태로부터 한두 가지 측면이 다른 어휘소의 대안적 형태를 말한다. 이 변이형은 나이와 출신 지역, 민족, 그리고 사회·경제적 지위와 같은 사회적 요인, 동화와 축소와 같은

음운 과정에 의한 요인 그리고 문법화와 같은 언어적 요인에 의해 발생한다. 수어에서 사회 언어적 변이형은 세 가지 유형으로 나눌 수 있다(Schembri & Johnston, 2012).³⁾

2.2.1 음운적 변이형

음운적 변이형(phonological variation)은 “동일한” 수어 어휘를 다른 형태로 사용하는 것으로, 달리 표현하면 다른 수형 혹은 수동을 사용하는 것이다(Johnston, 2003: 439-443). 또한 한손 단어가 두 손으로 표현될 때가 있고, 두 손 단어가 한손으로 표현될 때가 있다.

미국수어, 호주수어, 뉴질랜드수어에서 위치에서의 변이형과 미국수어에서 수형에서의 변이형 그리고 한손 단어와 두 손 단어에서의 변이형을 차례대로 살펴보도록 하겠다.






1) 위치 변이형

Woodward, Erting & Oliver(1976)는 지역과 민족에 따른 변이형을 연구하였다. 미국수어에서 RABBIT, COLOUR의 인용형은 얼굴 혹은 손에서 산출될 수 있는데 애틀랜타 지역의 수어 사용자들은 뉴올리언스 지역의 수어 사용자들보다 이 어휘들을 얼굴이 아닌 손에서 산출하였고 흑인 수어 사용자들도 손에서 산출하였다. 또한 미국수어 DEAF는 언어적 요인과 지역적 요인 그리고 화용적 요인에서 연구하였다. DEAF의 인용형은 뺨에서 턱으로 이동하는 동작이다(a). 변이형은 두 가지로 나타난다. 첫째, 인용형의 동작 순서가 턱에서 뺨으로 이동한다(b). 둘째, 이동이 없고 턱만 접촉한다(c). 변이형이 발생하는 언어적 요인으로는 합성어를 들 수 있다. DEAF-WORLD, DEAF-CULTURE와 같은 합성어에서는 b와 c로 나타났으며 DEAF가 서술어일 때는 a로 나타났다. 지역적인 요인을 살펴보면 캔자스, 미주리, 버지니아 지역에서는 b와 c를 사용하였고 보스턴 지역의 수어 사용자들은 a를 사용하였다. 특히 보스턴 지역의 나이든 수어 사용자들은 a를 더 많이 사용했고 젊은 수어 사용자들은 b와 c를 사용했다. b와 c는 내레이션에서 a는 대화에서 더 많이 사용하였다. 위치에서의 변이형에 대한 강력한

3) 세 가지 변이형에 대한 내용은 Schembri와 Johnston(2012)을 요약·정리하였다.

요인은 문법적 기능이다. 명사, 동사, 형용사는 인용형으로 나타나는데 반해 (예, FATHER, UNDERSTANDING, DIZZY), 전치사(예, FOR)와 의문사(예, WHY)는 더 낮은 위치에서 산출되었다.

2) 수형 변이형

Bayley 외(2000)는 수어 어휘 5,356개에서 수형 변이형에 대한 강력한 요인은 문법적 기능이지만 사회적 요인 또한 중요함을 제시하였다. WH-의문사의 경우는  변이형, 대명사(INDEX1)의 경우는  변이형이 많이 나타났다. 지역적으로는 캘리포니아, 캔자스/미주리, 루지애나, 매사추세츠 지역에서는 인용형을 선호했지만 메릴랜드와 버지니아 그리고 워싱턴 지역에서는 인용형을 선호하지 않았다. 젊은 수어 사용자들은 나이든 수어 사용자들보다  과  수형을 더 많이 사용했고 비-원어민 수어 사용자는  수형을 선호했다.

3) 한손 단어와 두 손 단어 변이형

미국수어에서 CAT, COW, GLASSES는 한손 혹은 두 손으로 산출할 수 있다(Woodward & DeSantis, 1977). 백인 수어 사용자들은 한손 변이형을 더 많이 사용했다. 남쪽 지역의 수어 사용자들은 다른 지역의 수어 사용자들에 비해 두 손 변이형을 더 많이 사용했다. 나이가 많은 수어 사용자들은 젊은 수어 사용자들보다 두 손 변이형을 많이 사용했다.

Deuchar(1981)은 영국수어에서 두 손 단어에서 비우세손의 탈락(weak drop)이 가능함을 제시하였다. 대칭적인 두 손 단어에서 비우세손 탈락은 빈번하였다. 또한 담화별로도 차이가 있었다. 201개 토큰 중에서 6%만이 형식적인 상황에서 비우세손 탈락이 있었지만 비공식적인 상황에서는 50%가 비우세손 탈락이 있었다.

뉴질랜드수어에서는 젊은 수어 사용자들은 나이든 수어 사용자들보다 비우세손 탈락을 선호했다. 많은 한손 단어들이 과거에는 두 손 단어로 사용되었다(McKee et al., 2011)

2.2.2 어휘적 변이형

어휘적 변이형(lexical variation)은 ‘동일한 의미를 위해 수어 화자들이 완전히 다른 단어를 사용하는 것’이다(Johnston, 2003: 439-443). 어휘적 변이형은 종종 수어 사용자들의 연령, 성, 지역, 민족, 교육적 배경과 체계적인 연관을 맺고 있다.

1) 지역

지역적 변이형이 생긴 이유는 과거에 농인들이 여행의 기회가 적었다는 여건상의 요인뿐만 아니라 수어를 위한 집중적인 훈련 프로그램의 부재로 농학교별로 수어가 다를 수 있었기 때문이다. 그러나 미국의 경우에는 표준화된 핵심 어휘부가 풍부하게 확립되었는데 이는 모든 농교사들이 코네티컷 하트퍼드에 소재한 농학교에서 교육을 받았기 때문이다.

지역적 변이형은 특정 의미적 범주에 집중되어 보인다. 호주수어에서 북쪽과 남쪽의 변이형을 살펴보면 주목할 만한 것은 이 두 지역의 변이형은 수, 색채, 그리고 다른 개념들을 위해 전통적으로 사용된 어휘들이라는 점이다(Johnston, 1998). 특정 의미 영역인 색채 어휘(colour signs)를 보면 어휘 항목의 핵심 부류는 이 두 지역에서 실제로 상이하다.

2) 연령

농인들은 가족들이 종종 청인이고 수어 습득 연령이 매우 늦으며 세대 간의 전달에 문제가 있기 때문에 세대 간의 수어의 차이가 존재한다.

영국수어의 경우 어린 수어 사용자들은 나이 든 수어 사용자들보다 더 적은 어휘적 변이형을 인식했고 어린 수어 사용자들은 더 많은 구화(mouthing)를 사용했다.

호주수어의 경우에는 51세 이상의 수어 사용자들이 50세 미만의 수어 사용자들보다 더 빈번하게 지문자를 사용했고 71세 이상은 특히 그러했다(Schembri & Johnston, 2007).

Padden과 Gunsauls(2003)는 연령과 사회 계층이 지문자와 보통명

사 사용에 영향을 준다는 것을 보고했다. 나이가 많은 노동자인 수어 사용자들은 보통 명사를 더 많이 지문자로 표현했다. 대학 교육을 받은 농인 원어민 수어 사용자들은 더 많이 지문자를 사용했다. 이와 관련하여 Lucas 외, 2001)은 중간 계층의 수어 사용자들이 노동자인 수어 사용자들보다 지문자를 더 많이 사용한다고 보고했다.

3) 성

호주수어에서 여성과 남성이 소수의 어휘적 변이형을 다르게 사용할 수 있다는 일화적인 보고가 있으며 어휘적 변이형에 관한 상당히 중요한 연구는 아일랜드수어에서 이루어졌다(Le Master et al., 1991; Leeson & Grehan, 2004). 한 세기가 넘게 아일랜드 농공동체는 더블린에 있는 남학생과 여학생을 위한 기숙제 농학교(농 여학생을 위한 세인트 메리 학교와 농 남학생을 위한 세인트 조셉 학교)에서 수어사용의 다른 전통으로 인해 다른 어휘들을 유지하게 되었다. Le Master와 Dywer(1991)는 153개의 자극 세트를 사용하여 항목 중 106개가 다르다는 것을 확인했고 그 중 63%는 어떤 방식으로든 연관되어 있음을 보고했다.

4) 민족과 종교

아프리카계 미국인이 구사하는 미국수어에서 변이형이 존재함이 확인되었다(Aramburo, 1989; Lucas et al., 2001). 앞서 살펴본 아일랜드수어에서 성별의 차이가 있었던 것처럼 어휘적 변이형의 출현은 미국 농교육의 역사적 맥락을 반영한다. 19세기와 20세기 분리교육 기간 동안에 남부 일부 주에 아프리카계 미국 농아동들을 위한 학교가 설립되었다. Aramburo(1989)에 따르면, 아프리카계 미국 수어 사용자들은 FLIRT, SCHOOL, BOSS와 같은 독특한 어휘적 변이형을 가짐을 보여주었다.

보다 최근의 흑인들의 미국수어 변이형에 대한 연구는(McCaskill et al., 2011) 특정한 어휘적 변이형의 사용뿐만 아니라 수많은 다른 차이들을 확인할 수 있음을 제시한다. 즉 백인 수어 사용자들과 비교할 때 흑인 수어 사용자들은 두 손 변이형 사용이 많으며 KNOW를 포함하는

어휘 부류에서 어휘들을 더 낮게 산출하였으며 훨씬 더 많은 반복을 사용하였다.

2.2.3 문법적 변이형(grammatical variation)

미국수어와 브라질수어에서 형태 통사적 변이형 연구는 거의 없었다. 성, 나이, 사회 계층, 지역에 따라 수어 사용자들 사이에서 일관된 차이를 보여주는 실증적인 연구도 아직 없다. 아일랜드수어에서 동시적인 구문과 화제 표지에서 일부 성의 차이가 보고되었지만 데이터 세트는 작았으며 결정적이지는 않다(Leeson & Grehan, 2004). 대조적으로 원어민과 비원어민 수어 사용자의 문법 기술의 차이는 문헌에서 여러 번 보고되었다(Boudreault & Mayberry, 2006).

McKee와 동료들은 미국수어에서는 65%, 뉴질랜드수어에서는 절의 절반, 호주수어에서는 2/3에서 주어가 생략됨을 발견했다. 이 수어들에서 주어를 생략하는 이유는 앞 절에서 동일한 지시체를 확인하는 주어를 사용했거나 혹은 역할 전환을 사용했을 경우이며 호주수어와 미국수어의 경우에는 영어가 수어의 절에 영향을 주었기 때문이다(영어는 pro-drop 언어가 아님).

2.2.4 형태론적 수정

수어의 어휘소는 내부적으로 혹은 외부적으로 수정된다(Johnston & Schembri, 1999: 143). 어휘소의 수정 형태(modified forms)는 변이형과 구별될 필요가 있다. 어휘소의 변이형 발음과 달리 수정은 체계적인 방식으로 의미의 차이를 전달하기 위해 일어난다. 수어에서 수정은 인용형에는 없는 어휘소의 윤색이다. 어휘소 인용형의 내부적인 수정은 방향, 표정, 그리고 수형의 변화를 포함하더라도 주로 공간과 움직임 포함한다. 호주수어의 경우 OPERATE와 CAESARIAN은 의미의 특정성(specificity) 원리에 따라 별개의 어휘소로 구분한다. CAESARIAN은 맥락을 떠나서 ‘위를 수술하다’가 아니라 ‘복부 절개를 통해 여성의 자궁에서 아기를 들어 올리는 수술’을 의미한다. 그러나 OPERATE-LOC는 특정적으로 ‘개심 수술’, ‘심장 이식’, ‘관상동맥 바이패스’를 말하는 것이 아니므로 의미의 특정성 원리에서 어휘소가 아

나라 OPERATE의 수정으로 간주한다. OPERATE-LOC의 산출 위치는 의미에 예측할 수 있는 어떤 것도 부가하지 않기 때문이다.

다음으로 움직임 포함의 경우를 보면 호주수어에서 명사와 동사 사이의 대립이 오로지 움직임의 특성에 기초하여 구별되는 때에도 그 패턴은 약해 보인다. 명사 형태는 제한되고 혹은 반복된 움직임으로 실현되고 동사 형태는 단일하고 연속된 움직임으로 실현된다. 많은 어휘소들이 이 패턴을 따르는 듯 보일지라도 명사와 동사 의미를 모두 가진 다른 많은 어휘소들은 각 의미를 위해 어떤 체계적인 수정을 보여주지 않는다. 어휘소는 명사 혹은 동사 의미를 강조하기 위해 수정된다(같은 책: 149쪽). 호주수어의 SIT와 CHAIR은 때로는 움직임의 수정으로 구별될지라도 하나의 동일한 어휘소의 수정으로 간주되며 사전에서 독립된 어휘소로 들어가지 않는다.

2.2.5 정리

한국수어에서 한국수어 말뭉치를 바탕으로 변이형 연구를 한다면 변이형 중에 어떤 변이형이 가장 빈번하게 나타나는지 또는 어떤 변이형을 인용형으로 보는 것이 좋을지를 실증적인 자료를 바탕으로 하여 연구할 수 있다.

음운적 변이형과 관련하여 특정한 수어 어휘가 복합어 혹은 단일어로 사용되었을 때 변이형이 나타날 수 있으며 문장 속에서 해당 수어 어휘의 앞뒤 어휘들도 영향을 주는 것을 관찰할 수 있다.

어휘적 변이형과 관련하여 한국수어 말뭉치에서 의미는 같으나 형태가 달라 두 개의 글로스로 전사된 것들이 있다면 언어 제공자들이 어떤 변이형을 선호하는지 관찰할 수 있다. 예를 들어 한국수어 말뭉치에서 나타난 ‘강원도1~3’의 변이형은 언어 제공자의 나이, 농학교, 그리고 고향에 따라 관찰할 수 있다.

문법적 변이형과 관련하여 어떤 수어 표현은 대화 속에서 사용되는 형태와 이야기를 할 때 사용되는 형태가 다를 수 있다. 이러한 문법적 변이형을 연구할 수 있다.

변이형 연구는 현재 쓰이는 다양한 수어 사용을 기록하고 사회 언어적인 요소들의(나이, 지역, 성별 등등) 영향을 제시할 수 있다. 많은

변이형 중에서 어떤 변이형들은 통시적인 변화의 모습을 보여주는 경우도 있다. 그러므로 변이형 연구를 통해서 앞으로 한국수어가 어떠한 모습으로 변화할지를 예측하는데 시사점을 제공할 것이다.

3. 수어 전사에서의 분절

한국수어 말뭉치의 전사에서 분절에 관하여 수화언어학자들이 가지고 있는 선형적인 단어(linear sign)의 길이에 대한 주류의 관점과 최근의 경향성을 살펴보도록 하겠다.⁴⁾ 수어 단어 구조의 가장 영향력이 있는 음운론적 모델을 살펴본 다음에 음성학적 관점을 제시하도록 하겠다.

3.1 음운론적 관점

과거 몇 십 년 동안 선형적인 단어의 길이에 대한 통념이 정확하지 않을 수 있다는 문제 제기를 뒷받침하는 증거들이 축적되었다. 선형적인 단어의 경계는 실제로 우리가 알고 있는 것보다 더 긴 단위라는 것이다. 이에 대한 주된 증거는 수어 단어 산출(sign articulation)과 수어 단어 수용(sign recognition)에서 온다. 이 두 증거의 상세한 내용은 음운론적 관점과 음성학적 관점 다음에 제시하기로 하겠다.

수어의 흐름(sign stream) 속에서 선형적인 단어의 시작점과 끝점이 무엇인지는 수어 연구와 음운론적 연구의 시작 단계에서 중시되지 못했다. 그러나 암묵적인 합의는 선형적인 단어는 손이 두 공간적인 지점 사이를 움직이는 동안의 움직임 기간이고, 일부 단어는 정지 기간이라는 점이다.

3.1.1 동시적인 단위

Stokoe와 그의 추종자들은 수어 단어를 구조적으로 동시적인 단위로 정의하였다. 수어 단어는 시간적으로 구성된 단위가 아니라는 그의

4) 분절에 대한 내용은 Jantunen(2015)을 요약·정리하였다.

확신과 함께 수어 단어의 실증적인 범위에 대한 어떤 질문도 구조주의 언어학의 방법론적 전통에서는 중시되지 못했다. 왜냐하면 언어의 추상적인 에믹 구성소를 확인하는데 주된 관심을 두었던 전통 속에서 형태소가 어디에서 시작되는지와 같은 질문을 할 이유가 없었기 때문이다. 전통적인 방법론은 형태소의 내부 구성소만을 기술하기 위해 고안되었다. Stokoe와 그의 추종자들은 수어 단어가 선형적인 단위임을 고려하지 않았지만 수어 단어의 길이를 수동의 길이와 같다고 간주했다고 보는 것이 타당할 것이다. ‘위’, ‘오른쪽’, ‘앞’과 같은 방향적인 명칭으로 혹은 손목, 손가락 관절로 산출된 더 작은 수동의 경우에는 ‘바깥으로 돌리다’, ‘펴다’ 등으로 분류하였다.

수어 단어에서 대부분의 구조적 자질들은 수어 단어가 수용되기 전에 이미 나타난다. 예를 들어 핀란드수어 BLACK의 경우 Stokoe 모델에서 보면 신체를 향한 움직임을 가진다. 이 단어는 수어자를 향해 움직이기 시작하는 지점에서 시작한다. 단어의 끝은 손이 움직임을 멈추는 위치이다. 이 단어의 위치는 이마이고 화살표는 움직임의 방향을 기술한다. BLACK과 다른 단일어들에서 손을 시작 위치에 올린 다음에 움직임을 일어나고 끝 위치에서 내린다. 그러나 손의 시작과 끝 움직임은 Stokoe와 수어 음운론의 현대적인 연구에서 수어 단어의 일부로 취급되지 않았으며 전이 움직임(transitional movement) 혹은 전이(transitions)로 분류하고 언어적 분석에서 제외하였다.



<그림 13> 핀란드수어 [BLACK]

3.2.2 연속적인 단위

수어 단어 경계의 특성에 대한 보다 확실한 인식은 Liddell과 Stokoe의 제안을 반대하는 아이디어가 도입된 이후부터 중시되었다.

수어 단어도 연속적인 구조라는 것이다. Liddell의 MH 모델에서 수어 단어의 길이에 대한 문제는 움직임(M)과 정지(H) 분절의 개념으로 대답을 주었다. 움직임(M)은 조음자(전형적으로 손)가 움직이고 있는 시간으로 정의되고, 정지(H)는 조음자의 모든 자질들이 변하지 않는 때의 움직임이다. MH 모델에서 수어 단어는 분절들로 구성되며 전형적으로 수어 단어는 M, H, MH, HMM의 연속으로 기술된다.

3.2 음성학적 관점

Liddell(1984)은 시간적으로 구성된 단위로서의 수어 단어 연구를 시작하였다. 그러나 수어 단어의 경계를 정하기 위한 실제적인 기준들이 분명하지 않다. 음성학적 모델의 맥락에서 이 기준은 MH 모델이 출판된 지 25년이 지난 아주 최근에서야 Liddell과 Johnson(2011)에 의해 검토되었다. 그것은 MH 모델(Liddell, 1984)과 정지-움직임 모델(Liddell & Johnson, 1989)에 기초한 것이고 미국수어 CHICAGO의 경계를 결정하는 방식-결과적으로 길이-에 관한 분명한 입장을 밝혔다. 수어 단어의 시작을 결정하는 기준은 다음과 같다(Liddell & Johnson, 2011).

첫째, 수어 단어들을 비교하고 이에 기초하여 같은 시작되는 위치를 결정한다. 맥락과 수어의 연속 속에서 손이 다른 방향에서 시작 위치로 향하는 기간을 전이로 분류한다.

둘째, 수어 단어의 시작을 손의 움직임의 방향이 변하는 위치와 연결한다.

셋째, 원어민 수어 사용자들의 직관을 참고한다. 원어민 수어 사용자들은 두 움직임의 집합점을 가진 수어 단어의 대략적인 시작을 가늠할 수 있다.

그러나 Liddell과 Johnson은 수어 단어의 끝점을 확인하기 위한 과정에 대해서는 어떠한 실제적인 기술도 제공하지 않았다. Liddell과 Johnson(2011)이 기술한 기준은 수어 말뭉치 언어학에서 최근에 나타난 수어 단어의 음성학적 식별 기준과 매우 흡사하다. 수어 말뭉치의 몇 가지 지침들은 다소 다르지만 공통점이 있다. 지침들은 우세 손의

양극단의 연속적인 움직임은 수어 단어의 식별을 위한 시작점으로 한다. 이는 네덜란드 수어 말뭉치의 전사 규칙에서 발췌한 다음 글에 나타나 있다(Crasborn & Zwitserlood, 2008: 6).

수어 단어의 시작: 첫 번째 프레임에서 손은 단어의 시작 위치에서 마지막 위치로 이동하기 위해 움직이기 시작한다. 혹은 (손이 공간에서 움직이지 않을 경우에) 첫 번째 프레임에서 수형이 변하기 시작한다. 혹은 (손이 공간에서 움직이지 않고 수형도 변하지 않는 경우에) 첫 번째 프레임에서 수향이 변하기 시작한다.

수어 단어의 끝: 첫 번째 프레임에서 수어 단어가 끝난 이후에 수형이 변하기 시작하고 혹은 첫 번째 프레임에서 손이 단어의 마지막 위치로부터 움직이기 시작한다.

말뭉치 전사 규칙에 포함된 수어 단어 식별 기준은 수어 단어의 시작과 끝을 결정하기 위해 이용할 수 있는 가장 정확한 음성학적 기준이다. 이러한 기준의 출현은 수어 연구에서 직관에 의존하던 연구에서 데이터의 보다 광범위한 수집과 이용으로의 전환을 반영한다.

3.3 수어 단어의 실증적인 길이에 대한 증거

수어 단어의 길이에 대한 주류적인 관점은 수어 단어의 움직임을 기술하기 위해 수행된 매우 초기의 작업으로부터 기초한다. 수어 단어의 시작점과 끝점은 주류의 관점이 인정하는 것보다 더 긴 단위이다. 이를 위한 증거를 수어 단어의 산출과 수용에서 찾을 수 있다.

3.1.1 수어 단어 산출(sign articulation)

수형이나 비수지 요소들은 수어 단어를 산출하려고 손을 움직이기 전부터 나타난다. 예를 들면 핀란드수어 TO-SEND를 식별하기 위한 주류의 기준, 즉 Johnson과 Liddell 그리고 말뭉치 전사 지침을 따르면, 10-17번(8컷) 프레임이 수어 단어의 길이이다. 그러나 10번 프레임도 첫 번째는 아니다. 수어 단어를 구성하는 수형과 수향의 조음 자질이 처음 나온 것이 아니기 때문이다. 19번 프레임도 이 자질을 감지할 수 있는 마지막 프레임이 아니다. TO-SEND의 처음 수형은 5번 프

레이스에서 이미 완전하게 형성되었고 마지막 수형의 형상도 20번 프레임에서 여전히 감지할 수 있다. 수향은 이미 17번 프레임에서 도달했고 19번 프레임에서 여전히 명백하게 감지할 수 있다. 결과적으로 단어를 구성하는 조음 자질의 처음과 마지막 발생을 고려하여 TO-SEND를 분절한다면 주류 관점이 취하는 것보다 대략 2배가 될 것이다(그림 14 참조).



<그림 14> 핀란드수어 독립된 단어 [TO-SEND] 산출 표상

이에 덧붙여 Vitaniemi 외(2014)에 따르면 자동적인 수어 단어 식별을 위한 최근 컴퓨터를 활용한 연구에서는 따르면 핀란드수어 단어들에서 첫 번째 조음 자질은 주류 관점이 취하는 것보다 평균 3.9프레임 더 앞에서 측정되었고 마지막 자질은 평균 1.9프레임 뒤로 측정되었다.

3.1.2 수어 단어 수용(sign recognition)

수어 단어는 많은 경우에 산출이 시작되었다고 간주되기 이전에 이미 상대에게 수용된다. 수어 단어가 독립적으로 그리고 맥락 속에서 음성학적 형태가 나타나기 전에 이미 인식된다는 연구에는 Grosjean(1981)가 있다. Grosjean은 비디오상으로 개별 단어들의 시작 위치를 인위적으로 바꿈으로써(예를 들어 실험에 참여한 사람들이 자극 단어를 점진적으로 더 많이 볼 수 있도록 조절함) 단어를 인식할 수 있었을 때는 단어의 형태가 대략 47% 나타났을 때임을 관찰하였다.

후속 연구로 Clark과 Grosjean(1982)은 맥락과 함께 구나 문장에서 수어 단어의 인식이 Grosjean(1981)에서 발견한 것보다 더 일찍 일어남을 제시하였다. 수어의 연속적인 흐름 속에서 수어 단어를 성공적으로 인식하기 위해서는 수어 단어가 37% 정도 나타났을 때 가능했다는 것이다. 이 연구에서 원어민 수어 사용자는 자극 단어의 시작을 앞의 단어가 끝났다고 판단한 순간과 연결하였다. 실제로 많은 경우에서 수어 단어의 인식 혹은 수용할 수 있는 지점은 단어 자체의 기간이 아니라 그보다 앞선 전이 움직임의 기간에서이다.

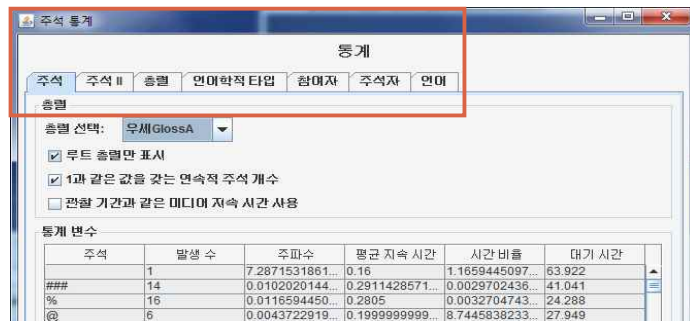
4. 수어 연구를 위한 엘란(Elan)의 활용

기존에 전사 된 한국수어 말뭉치는 엘란(Elan) 프로그램을 활용하여 검색하여 분석을 하거나 통계를 낼 수 있다. 검색 및 분석, 통계는 필요에 따라 단일 전사 파일(eaf파일)에서 이뤄지거나 다수 전사 파일(eaf파일)을 대상으로 검색하여 분석하거나 통계를 낼 수 있다.

엘란(Elan)을 활용한 검색 및 분석, 통계 방법은 아래와 같다.

4.1 층렬별 통계값 보기

엘란(Elan) 메뉴 ‘보기’의 ‘주석 통계’를 통해 주석, 층렬, 언어학적 타입, 참여자, 주석자, 언어 항목을 살펴볼 수 있다.

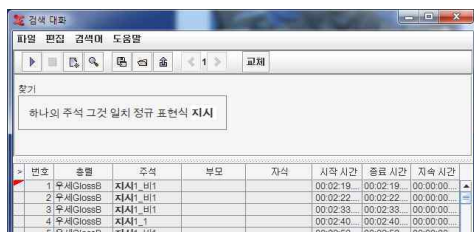


<그림 15> 주석통계

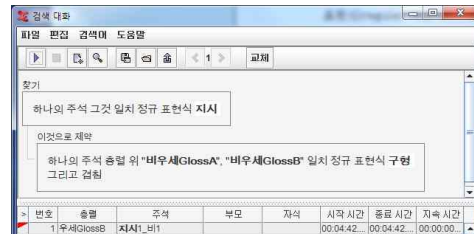
4.2 단일 전사 파일(eaf 파일) 검색

엘란(Elan) 메뉴 중 ‘검색’의 ‘찾기’에서 찾고자 하는 층위를 선택하여 검색하고자 하는 주석을 입력한다. 보다 세부적인 내용을 추가하고자 한다면 ‘제약 추가(Add constraint)’를 통해 검색 조건을 추가하거나 ‘정규 표현식(regular expression)’을 선택하여 검색의 범위를 좁히거나 넓힐 수 있다.

예) [지시]와 [구형]이 함께 사용된 주석을 검색



<그림 16> 정규 표현식을 선택하여 [지시]를 검색할 경우



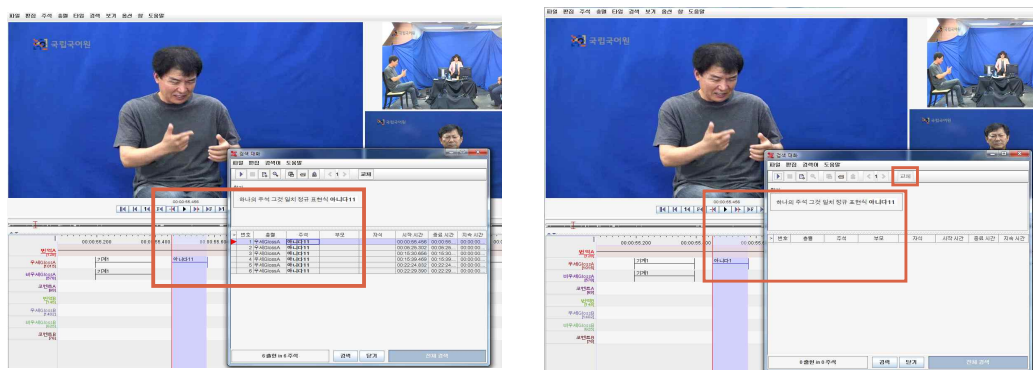
<그림 17> 검색된 [지시]의 검색 범위를 좁혔을 경우



<그림 18> [지시]와 [구형]을 함께 사용한 주석을 검색한 예

4.3 검색 및 변경

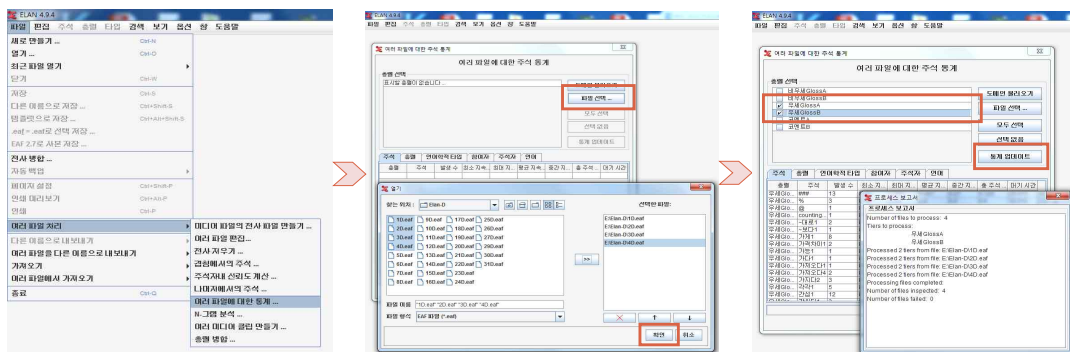
수정하고자 하는 주석을 검색하여 원하는 주석명으로 변경할 수 있다. 먼저 ‘검색’의 ‘찾기’에서 수정하고자 하는 주석을 검색한다. ‘검색 대화’팝업창의 ‘교체’를 클릭하여 원하는 주석명을 입력한다. 입력 후 수정한 내용을 주석 모드에서 확인할 수 있다. 주석을 검색하여 수정한 결과는 <그림 19>와 같다.



<그림 19> [아니다11]을 [아니다1]로 수정한 결과

4.4 다수의 파일 층별별 빈도 분석

다수의 파일에 대해 층별별 주석 빈도를 분석할 경우, 파일 -> 여러 파일 처리 -> 여러 파일에 대한 통계를 클릭한다. 도메인이나 다수의 파일에서 빈도를 분석하고자 하는 층별을 선택한 후 ‘통계 업데이트’를 클릭하면 선택한 여러 파일에서 원하는 층별의 주석 빈도 통계를 구할 수 있다.



<그림 20> 다수의 파일 층별별 빈도분석

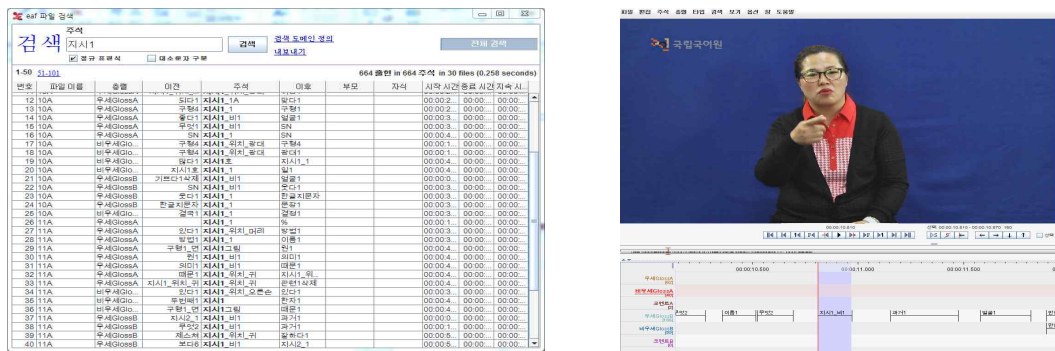
4.5 다수의 전사 파일 검색하기

다수의 전사 파일을 검색할 경우, ‘검색’의 ‘여러eaf검색’에서 도메인을 선택한 후 찾고자 하는 주석명을 입력한다.



<그림 21> 다수의 전사파일 검색

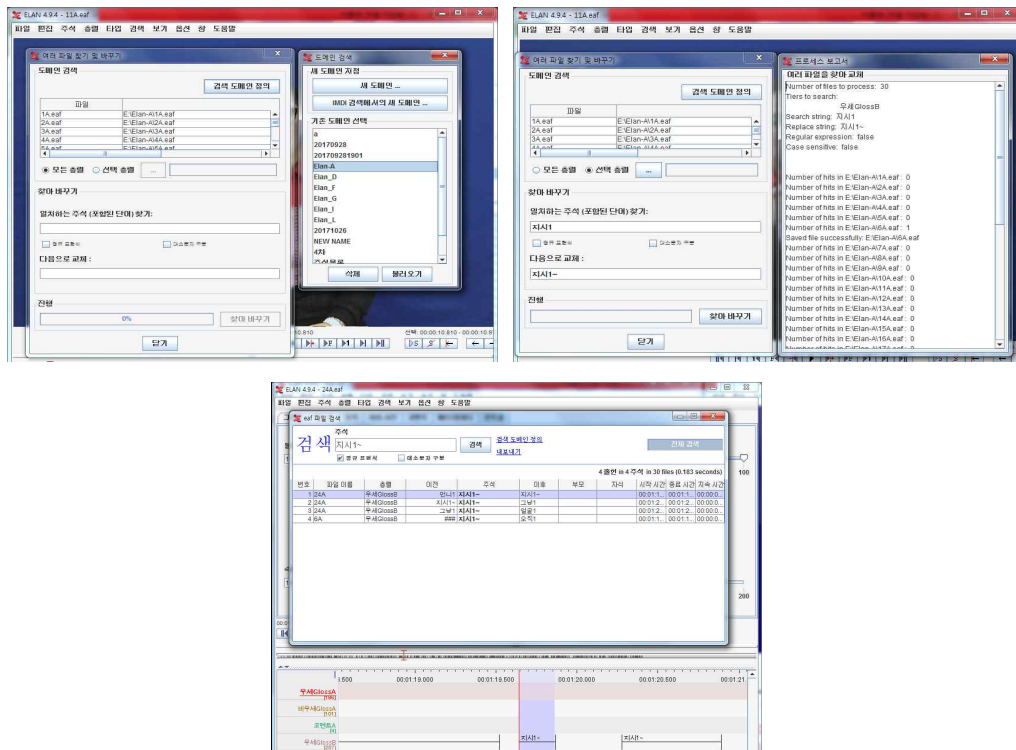
결과 값은 1)번호 2)파일 이름 3)층렬 4)이전 5)주석 6)이후 7)부모 8)자식 9)시작 시간 10)종료 시간 11)지속 시간으로 나타난다. 필요에 따라 주석을 클릭하면 해당 전사 파일이 실행되고 선택한 주석을 주석 모드에서 확인할 수 있다.



<그림 22> 다수의 전사 파일 검색 결과

4.6 다중 파일 검색 및 변경

여러 파일을 대상으로 원하는 주석명을 변경하고자 할 때는 ‘검색’의 ‘여러 파일 찾기 및 바꾸기’를 클릭한다. 이 때 바꾸고자 하는 파일을 한 폴더로 묶어 도메인으로 설정해야 도메인에 속한 파일의 주석명을 변경할 수 있다.



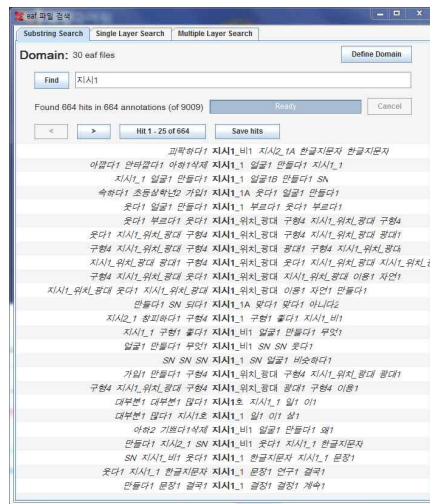
<그림 23> 다중 파일 검색 및 변경

4.7 다수 전사 파일 다중 검색

여러 전사 파일을 대상으로 다중 검색을 하고자 할 때는 ‘검색’의 ‘여러eaf 구조적 검색’을 클릭한다. ‘여러 eaf 구조적 검색’에서는 Substring Search, Single Layer Search, Multiple Layer Search 세 가지 검색이 가능하다.

4.7.1 Substring Search

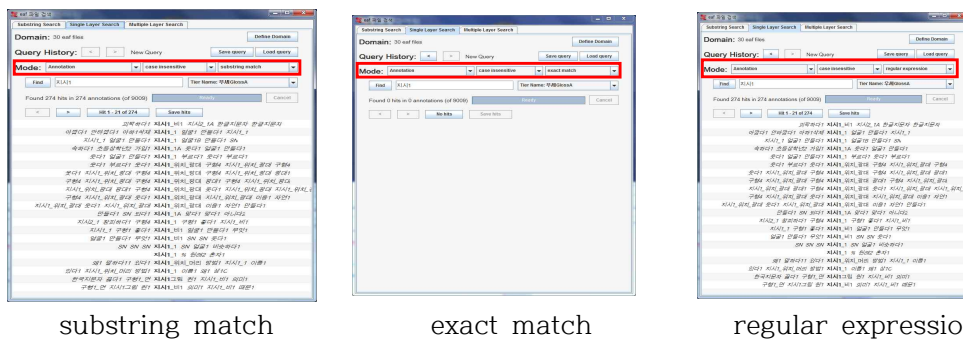
Substring Search에서는 도메인에 해당되는 전사 파일의 모든 층렬을 검색한다. 주석명 [지시1]로 검색한 결과는 아래와 같다.



<그림 24> Substring Search

4.7.2 Single Layer Search

Single Layer Search에서는 원하는 층렬을 선택하여 검색할 수 있다. 도메인에 해당되는 전사 파일의 우세GlossA를 대상으로 주석명을 [지시1]로 검색하였을 때 선택한 Mode에 따라 아래와 같은 결과 값을 얻을 수 있다.



substring match

exact match

regular expression

<그림 25> Single Layer Search

4.7.3 Multiple Layer Search

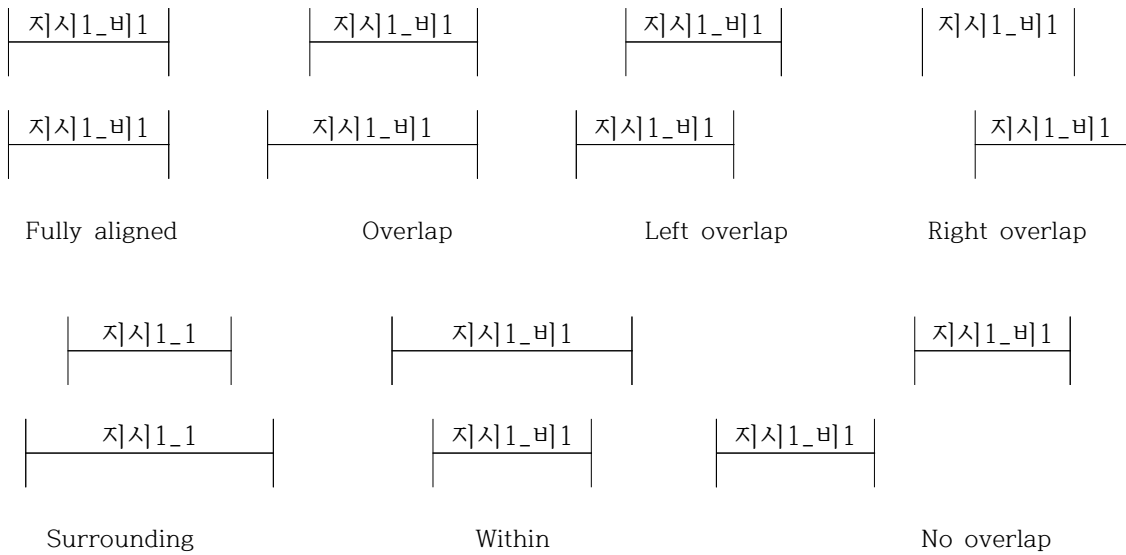
Multiple Layer Search에서는 도메인에 해당되는 전사 파일의 모든 층렬을 검색한다. Substring Search와 다른 점은 검색 조건을 보다 상세히 구조화하여 검색의 범위를 좁힘으로 찾고자 하는 결과를 도출하는데 있다.

Multiple Layer Search검색 시 구간에 대한 조건 제한은 아래와 같다.

- * = N annotations: 같은 층렬의 두 주석 간의 주석 수가 N수와 같을 경우
- * > N annotations: 같은 층렬의 두 주석 간의 주석 수가 N수 보다 많을 경우
- * < N annotations: 같은 층렬의 두 주석 간의 주석 수가 N수 보다 적을 경우
- * = X milliseconds: 같은 층렬의 A 주석에서 B 주석이 발생하기까지 걸리는 시간이 X milliseconds와 같을 경우
- * > X milliseconds: 같은 층렬의 A 주석에서 B 주석이 발생하기까지 걸리는 시간이 X milliseconds보다 길 경우
- * < X milliseconds: 같은 층렬의 A 주석에서 B 주석이 발생하기까지 걸리는 시간이 X milliseconds보다 짧을 경우

Multiple Layer Search 검색 시 겹치는 주석을 구하는 조건 제한은 아래와 같다.

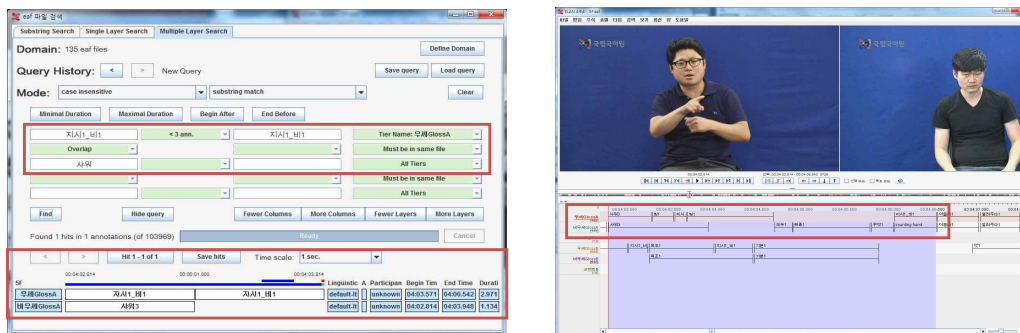
- * Fully aligned: 다른 층렬의 두 주석이 시작 시간과 종료 시간이 동일한 경우
- * Overlap: 다른 층렬의 두 주석이 겹친 경우이나 완전히 일치된 경우와 다르며, 왼쪽 중첩, 오른쪽 중첩, 주변 및 내부에 포함된 경우
- * Left overlap: 다른 층렬의 겹치는 두 주석이 왼쪽 구간일 경우
- * Right overlap: 다른 층렬의 겹치는 두 주석이 오른쪽 구간일 경우
- * Surrounding: 위 층렬의 주석 시간이 아래 층렬의 주석시간 보다 짧을 경우
- * Within: 위 층렬 주석 시간이 아래 층렬의 주석 시간 보다 길 경우
- * No overlap: 다른 층렬의 두 주석이 어느 쪽으로나 겹치지 않는 경우



4.8 다중 파일 검색 예

여러 파일을 대상으로 검색 조건을 구조화하여 검색한 예는 아래와 같다.

구간은 [지시1_비1](A)과 [지시1_비1](B)사이 주석 값의 수가 3개보다 적은 우세GlossA 층렬이다. 이 구간에 [지시1_비1](A)에 겹치는 [샤워]를 모든 층렬에서 검색한 경우이다. 그 결과 총 1개가 검색되었다.



<그림 26> 다중 파일 검색 예

검색 결과 5F 전사파일의 우세GlossA 층렬에서 [지시1_비1](A), [빛1], [지시1_비1](B)의 구간에 [지시1_비1](A)에 겹치는 [샤워3]이 검색되었다.

Ⅲ. 기구축 한국수어 말뭉치 정비

1. 기구축 한국수어 말뭉치 연결 실패 문제 해결

2015년 구축한 한국수어 말뭉치는 시스템이 없는 가운데 전사를 실시하였다. 이에 시스템 없이 전사한 내용을 시스템에 연결하는 과정에서 발생한 연결 실패 문제를 해결하고 기구축 한국수어 말뭉치를 정비하였다.

2015년 기구축한 한국수어 말뭉치를 2016년 구축한 통합 지원 시스템에 연결하면서 나타난 오류의 유형은 다음과 같았으며, 유형별 해결 방법을 통해 연결 실패 문제를 해결하였다. 연결 실패 이유는 다음과 같다.

<표 2> 연결실패 이유와 해결 방법

| 번호 | 연결실패 이유 | 예시 | 해결 방법 |
|----|---|------------------------|------------------------------|
| 1 | 토큰 글로스 오타 | 어런니1, 남짜2 | 오타 수정 |
| 2 | 한글 지문자를 낱자로 전사 | ㅎ ㄴ ㅁ | 한글 지문자 타입 추가 |
| 3 | 영어 지문자를 낱자로 전사 | CHICAGO | 영어 지문자 타입 추가 |
| 4 | 수어이름 전사 | SN홍길동 -> SN언어 제공자14 | SN |
| 5 | 전사자가 언어 제공자의 수어를 모를 경우 | ### | ### 타입 연결 |
| 6 | 언어 제공자가 수어를 하다가 멈춘 경우 | % | % 타입 연결 |
| 7 | 언어 제공자가 수어를 잘못된 경우 | @ | @ 타입 연결 |
| 8 | 입 움직임이 손으로 가려져 안 보인 경우 | @@ | @@ 타입 연결 |
| 9 | 변이형(Variation)과 형태론적 수정(Modification)의 연결 | 도와주다11, 시골2a | 변이형과 형태론적 수정을 구분하지 않고 타입에 연결 |

총 전사 파일에서 연결 실패로 나타난 토큰은 총 13,340개이다. 전사 파일별 연결 실패 토큰 수는 <붙임1>에 있다.

2015년 기 전사된 한국수어 말뭉치를 정비하기 위한 연결 실패 문제는 분석을 통해 타입을 추가하여 연결 실패 문제를 해결하였다. 또한 그 해결 과정에서 나타난 새로운 오류 유형에 대해서는 타입 정비 등을 통해 일부 정비하고 전사 규칙을 수정 보완하였다. 그 내용을 정리하면 다음과 같다.

1.1 추가 타입 등록

1.1.1 영어 지문자

1) 연결 실패 내용

- 영어 지문자는 알파벳 대문자로 전사되었다.(예:LIM)
- 각각의 알파벳 타입이 존재하지 않아 연결 실패가 나타났다.

2) 해결 내용

- 알파벳 대문자로 전사된 영어 지문자를 [영어지문자]라는 타입을 만들어 연결하였다.
- 현재는 LIM같은 지문자를 시스템에서 인식하여 자동적으로 [영어지문자] 타입에 연결되고 있다.

1.1.2 한글 지문자

1) 연결 실패 내용

- 한글 지문자는 각 낱자를 분리하여 전사되었다.
(예: ㅁㅏ ㄹㅁㅏㅇㅈㅣ)
- 한글 자모음에 대한 각각의 타입이 존재하지 않음으로 연결 실패가 나타났다.

2) 해결 내용

- 각 한글 자음과 모음으로 전사된 한국 지문자를 [한글지문자]라는 타입을 만들어 연결하였다.
- 현재는 한글 지문자를 시스템에서 인식하여 자동적으로 [한글

지문자] 타입에 연결되고 있다.

1.1.3 얼굴이름

1) 연결 실패 내용

- 각각의 얼굴 이름에 대한 타입이 존재하지 않았다.

2) 해결 내용

- [SN] 타입명을 추가하고 SN_000이라는 모든 토큰을 연결하였다.

1.1.4 그 외

1) 전사자가 알지 못하는 수어

(1) 연결 실패 내용

- 전사자가 알지 못하는 수어의 경우 연결 실패로 나타났다.

(2) 해결 내용

- 전사자가 알지 못하는 수어를 매칭하는 [###] 타입을 추가 등록하여 연결하였다.

2) 언어 제공자가 수어를 하다 멈춘 경우

(1) 연결 실패 내용

- 언어 제공자가 수어를 하다가 멈춘 경우 연결 실패가 나타났다.

(2) 해결 내용

- 언어 제공자가 수어를 하다가 멈춘 경우를 매칭하는 % 타입을 추가 등록하여 연결하였다.

3) 언어 제공자가 수어를 잘못된 경우

(1) 연결 실패 내용

- 언어 제공자가 수어를 잘못된 경우 @ 타입이 없어 연결 실패가 나타났다.

(2) 해결 내용

- @ 타입을 추가 등록하여 연결하였다.

1.2 타입 정비

1.2.1 2개의 타입을 1개의 타입으로 통일 (총 39건)

1) [파헤치다] / [간섭]

(1) 논의 내용

- [파헤치다], [간섭]이 같은 수어로 보여 구별에 어려움이 있었다.

(2) 해결 내용

- 의미는 다르나 기저 이미지는 같아 [간섭1]로 통일하였다.

2) [합하다] / [모으다]

(1) 논의 내용

- [합하다], [모으다]의 수어가 비슷해 타입 매칭이 모호하였다.

(2) 해결 내용

- 기저이미지가 같기 때문에 하나의 타입으로 보았다.
- [합하다1]을 삭제하고 [모으다1]로 통일하였다.

3) [점1]

(1) 논의 내용

- 점이 하나인 경우('.')와 여러 개('...') 인 경우가 있었다.
- [점1]로 보고 3개로 분절하여 입력할 것인지 혹은 형태론적 수정(modification)으로 보고 [점1]로 전사 할 것인지에 대한 여부를 논의하였다.

(2) 해결 내용

- 형태론적 수정(modification)으로 보아 [점1]로 전사하였다.

1.2.2 타입명 수정 (총 14건)

1) [배아프다1]

(1) 논의 내용 - [배아프다1] 타입명 적절성 여부를 논의하였다.

(2) 해결 내용

- 같은 수어이나 맥락에 따라 아까울 때, 배 아플 때, 질투가 나

서 배가 아플 때 등 다양한 의미로 사용된다.

- 타입명을 [배아프다1]로 할 때 의미의 폭이 좁아져 타입을 매칭 하는데 장애가 될 수 있다.
- 여러 맥락에서 타입을 매칭할 수 있도록 타입명을 [배아프다1]을 [몸부림1]로 수정하였다.

2) [건배1] / [술잔이 오간다]

(1) 논의 내용 - 수어가 비슷하여 타입을 매칭하기 모호하였다.

(2) 해결 내용

- 기저이미지가 같으나 의미가 다르다.
- 두 수어의 차이점은 사람 수와 연관이 있다. (두 사람이 술잔을 부딪친다. 여러 사람이 술잔을 나눈다.)
- [술2] 타입을 등록하여 타입 영상에 두 수어를 함께 제시하였다.

1.2.3 새로운 타입 등록 (36개)

1) 논의 내용

- 2015년 타입 목록을 정리하였다. 정리 과정에서 삭제된 타입이 시스템에 등록이 되지 않아 연결 실패가 발생하였다.

2) 해결 내용

- 새로운 타입이 필요한 경우 시스템에 등록한다.
- 새로운 타입은 후보로 등록하고 검토 후 타입 목록에 추가하였다.

1.2.4 변이형 토큰의 매칭

1) 논의 내용

- 변이형은 2015년 초기 전사를 하다가 중도에 전사하지 않기로 결정하여 변이형 토큰을 연결할 타입이 존재하지 않았다.
- 그러나 기 전사된 변이형 토큰은 그대로 두기로 하여 타입이 존재하는 변이형 토큰이 발생한다. 이에 새로운 타입명(변이형에 대한)을 추가 또는 기본형에 연결할지에 대한 논의하였다.

2) 해결 내용

- 타입명이 존재하지 않는 변이형 토큰은 기본형 타입에 수동으로 연결하였다. 예) 도와주다1b를 [도와주다1] 타입에 연결

2. 기구축 한국수어 말뭉치의 전사 규칙 정비

기구축 한국수어 말뭉치의 정비 과정에서 다음과 같은 타입명에 대한 전사 규칙을 수정 보완할 필요가 발견되었다. 그 내용을 요약하면 다음과 같다.

2.1 [지시]

1) 논의 내용

- [지시] 타입이 세분화 되어 있어 전사 시 혼란이 생길 수 있다.
- 타입명이 단순한 것이 효율적이며 전사의 오류 발생을 줄인다.

<표 3> 기존 [지시] 타입 목록

| 기존 [지시] 타입 목록 | | | |
|---------------|--------------|----|------------|
| 1 | [지시1_위치RA] | 9 | [지시2_위치RT] |
| 2 | [지시2_위치LT] | 10 | [지시1_위치LT] |
| 3 | [지시2_위치MT] | 11 | [지시1_위치M] |
| 4 | [지시1_위치_오른손] | 12 | [지시1_위치MA] |
| 5 | [지시1_위치L] | 13 | [지시1_위치MB] |
| 6 | [지시1_위치LA] | 14 | [지시1_위치MD] |
| 7 | [지시2_위치_손] | 15 | [지시1_위치MT] |
| 8 | [지시2_위치M] | 16 | [지시1_위치R] |

2) 결정 내용

- [지시]를 편 손가락의 수에 따라 [지시1]과 [지시2]로 나눈다.
- [지시1]을 가리키는 대상(일인칭과 그 외)에 따라 [지시1_1]과 [지시1-비1]로 나눈다.
- [지시]로 그리는 호, 원은 모두 [지시1호], [지시2호]로 하였다.

- 예) ‘우리’의 경우 [지시2호])
- 지시로 그리는 그림은 [지시1그림], [지시2그림]으로 하였다.
- 예) 서울 지역 30번 쌍 F-표지판 전사 (00:05:00.840 - 00:05:01.090)
- [지시1], [지시2]의 변이형을 타입 매칭할 시 [지시1], [지시2] 중 두드러진 것에 매칭하였다.
- 예) U형 (두 손가락을 편 수형)으로 사선으로 긋는 동작
(서울 지역 8번 쌍 F-표지판 전사 (00:01:45.567 - 00:01:45.815) =>
[지시2그림]로 전사

<표 4> [지시]의 8가지 종류

| | | | | |
|-----|------------|-------------|-----------|------------|
| 지시1 | 1) [지시1_1] | 2) [지시1_비1] | 3) [지시1호] | 4) [지시1그림] |
| 지시2 | 1) [지시2_1] | 2) [지시2_비1] | 3) [지시2호] | 4) [지시2그림] |

2.2 Counting hand

- 1) 논의 내용
- Counting hand에 해당되는 다양한 수어가 있으나 현 타입 목록으로 한계가 있었다.
 - 이를 분류하는 것은 연구 목적에 따라 달라질 수 있다.
 - 기초 연구에 맞게 Counting hand 타입을 정리할 필요가 있었다.

<표 5> 기존 [Counting hand]에 해당되는 타입 목록

| 기존 Counting hand에 해당되는 타입 목록 | | | |
|------------------------------|---------------------|---|--------------------|
| 1 | [counting hand 남동생] | 2 | [counting hand 남자] |
| 3 | [counting hand 여자] | 4 | [counting hand 딸] |
| 5 | [counting hand 아들] | 6 | [counting hand 형] |

- 2) 결정 내용
- [counting hand] 타입을 추가 하였다.
 - 전사 방법 변경하였다.
 - 기존 전사 방법은 아래와 같다.

우세: [counting hand 남동생]

비우세: [counting hand 남동생]

- 변경 된 전사 방법은 아래와 같다.

우세손: 주 의미를 나타냄/ 비우세손: [counting hand]

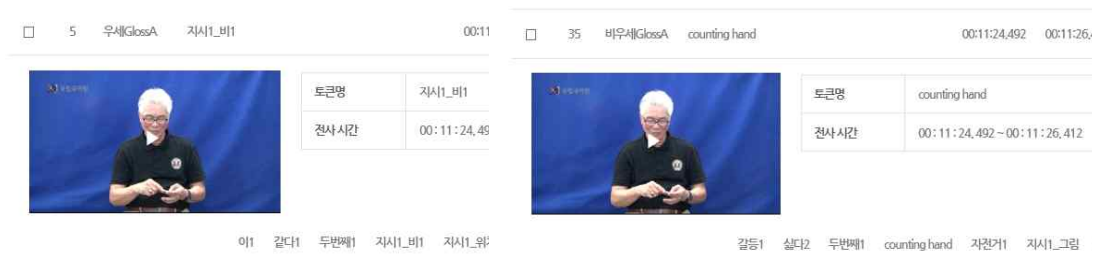
예1) [둘째 아들]의 전사 방법

우세손: [사람] 또는 [지시1_비1] 등

비우세: [counting hand]

예2) 서울 지역 8번 쌍 F-표지판 전사 (00:11:24.492 - 0:11:26.412)

우세: [지시1_비1] / 비우세: [counting hand]



<그림 27> 비우세 [counting hand]와 우세 [지시1_비1]

2.3 신체를 지시하는 경우

1) 논의 내용

- [지시1_위치_눈] 타입이 존재한다.
- ‘눈’과 ‘눈을 지시’하는 타입 구분이 모호하다.
- 타입명을 단순하게 하는 것이 효율적이며 전사의 오류 발생이 적다.
- 신체부위에 대해 말할 때 사용되는 지시는 [다리1] 혹은 [지시1_비1] 중 어느 것으로 매칭해야 하는지 논의하였다.

2) 해결 내용

- [눈1]과 같이 타입을 등록한다면, 다리2로 해야 하지만 다리를 지시하는 것은 고정된 수어가 아니므로 지시로 분류하였다.
- 다리가 아플 때 정확히 아픈 곳을 지칭해서 설명하기 때문에 지시로 연결였다.
- 타입명을 단순화하여 정리하였다.
- 신체를 지칭하는 경우는 [눈1], [코1]로 매칭하기로 하였다.

예) [눈1] : 서울 지역 20번 쌍 M - 농학교 경험 이야기
(00:01:36.196-00:01:36.356)

1
우세GlossA
눈1
00:01:36.196
00:01:36.356
연결 완료
경미회
2017. 9. 1. 15:39:28
닫기



| | |
|-------|---------------------------|
| 토큰명 | 눈1 |
| 전사 시간 | 00:01:36.196~00:01:36.356 |

~대로1 독수리1 새1 눈1 ~대로1 지시1 폭발1

<그림 28> [눈1]

- 눈 주위, 코 주위를 나타내기 위해 원을 그리는 경우에는 [눈원1], [코원1]로 매칭하기로 하였다.

2.4 타입명이 음성 언어로 ‘구’의 형태일 때

- 1) 논의 내용
 - 수어가 ‘구’형태일 때 타입명 규칙과 분절 여부에 대해 논의하였다.
- 2) 결정 내용
 - 음성 언어로 의미를 표현하면 ‘구’의 형태여도 수어 기준으로 보았을 때 하나의 수어로 여겨질 경우는 분절하지 않고 하나의 수어로 타입을 등록하였다.
 - 주로 두 수어가 한 수어로 나타남으로 분절하기에 어려움이 있었다. 예) [마음에안들다1]

2.5 등록된 수어와 모든 형태가 같으나 수위만 다를 경우

- 1) 논의 및 결정 내용
 - 토큰이 등록된 수어와 수위만 다른 경우는 새로운 타입으로 등록하지 않고 변이형(Variation)으로 보도록 하였다.
- 예) [미원2] 서울 지역 12번 쌍 I-개별 수어의 다양한 표현 전사
(00:04:44.935-00:04:45.314): 수위는 ‘불’이나 영상에서의 수위 ‘턱’

1
우세(GlossA)
미원2*
00:04:44,935
00:04:45,314
연결 완료
서영재
2017.4.11. 14:27:40
닫기



| | |
|-------|-----------------------------|
| 토론명 | 미원2* |
| 전사 시간 | 00:04:44,935 ~ 00:04:45,314 |

가지다2
미원2*
지시1_위치M
미원2*
지시1_위치M
오해1
지시1_위치M

<그림 29> [미원2]

2.6 생산적 수어

2.6.1 [구형_선], [구형_면]

1) 논의 내용 - [구형_선], [구형_면]의 구분이 모호하다.

<표 6> 기존 [구형]에 해당되는 타입 목록

| 기존 [구형]에 해당되는 타입 목록 | | | |
|---------------------|---------|---|---------|
| 1 | [구형1_면] | 2 | [구형1_선] |
| 3 | [구형2] | 4 | [구형3] |
| 5 | [구형4] | | |

- [구형1]의 필요 여부에 따른 타입 추가가 필요하다.
 - 손바닥에 놓인 무언가를 가리킬 때 등 ‘선’, ‘면’, ‘그림’이외 [구형1]이 사용되는 경우가 있다.
예) 넥타이 주제 중, 넥타이를 보여주는 맥락에서 비우세 [구형1], 우세 [지시1_비1]이 나온다. 이 때, 비우세는 [구형1], [구형1_면] 중 무엇으로 해야 하는지 모호하다.
 - 현재 구형1_면, 구형1_선만 있는데, 구형1과 구형1_면은 기법이 다르기 때문에 구분 필요하다. (구형1_면은 SASS가 아니며 생산적 수어에서 구형1은 개체를 나타내기 때문에 구분.)
- 2) 결정 내용
- [구형]을 [구형_선](구형으로 선을 그릴 때)과 [구형_면](손바닥 면의 움직임을 통해 면을 표현할 때)로 구분하였다.
 - [구형1] 타입을 추가 하였다.

2.7 [육형1] / [육형2] / [육형3]

1) 논의 내용

- [육형1], [육형2], [육형3]과 관련 된 연결 실패 유형이 있다.
- [육형1_세모], [육형1_네모], [육형2_네모] 타입의 필요 유무를 논의 하였다.

<표 7> 기존 [육형]에 해당되는 타입 목록

| 번호 | 기존 [육형]에 해당되는 타입 목록 | | |
|----|---------------------|---|----------|
| 1 | [육형1] | 2 | [육형1_세모] |
| 3 | [육형2] | 4 | [육형1_네모] |
| 5 | [육형3] | 6 | [육형2_네모] |

- 육형 규칙 재검토하였다.
- 재검토 후 타입 유지하고 [육형1_세모], [육형1_네모], [육형2_네모]을 하위타입으로 정리하였다.

2.8 [육형], [십형]의 변이형

- 1) 연결실패 내용 - [육형], [십형]의 변이형 타입 매칭이 필요하였다.
- 2) 논의 내용
 - [육형], [십형]의 변이형이 유사하여 구분이 모호하다.
 - [육형], [십형]의 변이형의 정의가 필요하다.
 - [육형1] / [육형1_세모] / [육형1_네모] / [육형2_네모]
 - [육형]의 변이형은 엄지와 검지가 맞닿지 않고 나머지 손가락이 퍼졌거나 접힌 수형이다. 예) ‘지역’을 뜻하는 수어.
 - [십형]은 엄지와 검지가 맞닿아 있다.
 - [십형]의 변이형은 엄지와 검지가 맞닿아 있으며 나머지 손가락이 접혔을 경우 십형의 변이형이다.

3. 정비 결과

한국수어 자료 통합 지원 시스템에서의 연결 오류 및 타입 정리 등을 통해 기 전사된 한국수어 말뭉치를 정비하였으나 시스템 내에서 해결할 수 없는 문제가 발생하였다. 엘란(Elan)파일을 다운 받아서 재전사를 하고 다시 시스템에 올리는 과정에서 또 다른 연결 실패가 발생하는 등의 문제가 나타났다. 그리하여 연구진이 파악하여 정비한 토큰은 656개이며, 그 유형을 분류하면 다음과 같다.

3.1 분절 오류

- 양손을 전사해야 하나 한 손만 전사 된 경우이다.
- 한 수어인데 두 토큰으로 전사한 경우이다.
(예: [안산1]을 [안전1A], [산1]으로 전사하거나 [사과1]을 [빨강1] [사과1]로 전사한 경우)
- 두 수어인데 전사자가 한 토큰으로 전사한 경우이다.
(예: [중국2] [나라1]을 [중국2]로 전사하거나 [눈1]+[싸다1]를 관용표현 전사규칙을 미숙지하여 [눈싸다1]로 전사한 경우)

3.2 토큰 누락

영상에 수어가 나왔으나 수어가 빨라 전사를 하지 못한 경우이다.

3.3 토큰 삭제

수어가 없는데 전사자가 그냥 손짓 같은 것을 토큰으로 전사했을 경우
우에 토큰을 삭제하였다.

3.4 지문자 오류

지문자에서 오류가 발생했을 경우이다.

(예: ㄴ ㅋ ㆁ ㅌ ㅍ ㅇ ㅣ 를 ‘ㄴ’, ‘ㅋ’, ‘ㆁ’, ‘ㅌ’, ‘ㅍ’, ‘ㅇ’, ‘ㅣ’로 전사)

IV. 전사 및 사례 연구

1. 전사 방법 및 절차

2015년 기구축한 한국수어 말뭉치 전사의 수정 보완 작업을 통해 보다 체계화되고 정교화 된 전사 지침을 마련하고, 그에 따라 비우세손과 우세손, 코멘트 그리고 번역 등 4개의 층렬을 250분 이상 전사한다. 이 때 전사는 한국수어 자료 통합 지원 시스템에 회원 가입을 한 후, 로그인을 하여 시스템 내에서 영상 자료와 엘란(Elan) 프로그램을 다운 받아 실시한다.

1.1 전사를 위한 기본 사항

1.1.1 기본 사항

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

1.1.2 엘란(Elan)의 특성 및 예시

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



<그림 30> 엘란(Elan) 전사 예

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

2. 전사자 교육 및 공동 전사

2.1. 전사자 교육

전사를 위하여 페이스 북 등을 통해 총 13명의 전사자를 모집하고 전사 교육을 실시하였다. 농인이 5명, 청인이 8명의 전사자를 모집하였다. 전사 교육 목표는 첫째로 한국수어 말뭉치 연구를 위한 한국수어 말뭉치 기본 지식 습득이며, 둘째 전사자 간의 일치도 확보를 위한 엘란(Elan) 기능 및 시스템 사용 방법 습득이다. 전사 교육의 내용 및 방법은 <표 8>과 같다.

<표 8> 추진 내용 및 방법

| | |
|----------|---|
| 추진 내용 | <ul style="list-style-type: none"> □ 일시: 2017. 6. 21.(수) ~ 2017. 6. 23(금), 10:00~16:00, 총 3회(18시간) □ 장소: 한국복지대학교 입학관 508호(수화통역과) □ 대상: 농인 전사자 5명, 청인 전사자 8명 □ 강사: 홍성은(한국수어 말뭉치 및 과제 설명, 타입·토큰·글로스, 시스템 사용법) 이은영(전사 지침, 전사 규칙, 엘란(Elan) 사용법) 김다해(엘란(Elan) 소개, 엘란(Elan) 설치, 엘란(Elan) 사용법) |
| 추진 방법 | <ul style="list-style-type: none"> □ 전체 전사자를 대상으로 교육에 적합한 날짜를 조사함 □ 농인과 청인의 의사소통 및 교육 방법 특성을 고려하여 교육 내용에 따라 농·청인 분반 및 공통 교육으로 실시 함 <ul style="list-style-type: none"> - 21일: 농인 대상 교육 - 농인 강사가 주요 교육 내용을 전담 - 22일: 농·청인 공통 교육 - 청인과 농인 강사가 역할을 분담하여 교육. 수화통역사를 두 명 배치. - 23일: 청인 대상 교육 - 청인 강사가 주요 교육 내용을 전담. □ 교육 중 질의응답을 자유롭게 하여 전사에 대한 궁금증을 즉시 해소시키고, 보다 정확하게 교육 내용을 전달코자 함 |

전사 교육 내용은 <표 9>와 같다.

<표 9> 전사 교육 내용

| 일자 | 시간 | 강사 | 교육 내용 | 비고 |
|--------------|-------|------------|--|--------------|
| 6.21. (수) | 10-11 | 홍성은 이은영 | 인사. 사업 소개. 한국수어 말뭉치 설명, 과제 설명 | 수화통역 농인강사 |
| | 11-14 | 이은영 | 엘란(Elan) 소개, 엘란(Elan) 설치, 엘란(Elan) 사용법 | 농인강사 |
| | 14-16 | 이은영 | 시스템 사용법 | 농인강사 |
| 6.22. (목) | 10-12 | 홍성은 | 타입, 토큰, 글로스 | 수화통역 |
| | 13-14 | 홍성은 | 타입, 토큰, 글로스 | 수화통역 |
| | 14-16 | 이은영 | 전사 지침, 전사 규칙 | 음성통역 |
| 6.23. (금) | 10-12 | 김다해 | 엘란(Elan) 소개, 엘란(Elan) 설치, 엘란(Elan) 사용법 | 청인강사 |
| | 13-14 | 홍성은 | 사업 소개. 한국수어 말뭉치 설명, 과제 설명 | 청인강사 |
| | 14-16 | 홍성은 | 시스템 사용법 | 청인강사 |

전사 교육 장면은 <그림31>과 같다.



<그림 31> 전사 교육 모습

전사 교육에는 농인 5명, 청인 8명, 뜻풀이 수어 사전 연구자 2명 총 15명이 참여하였다. 전사 신청자 중 93%가 전사 교육에 참여하였으며 결석자는 추후 교육하기로 하였다.

2.2 공동 전사

전사자 교육을 통해 습득한 엘란(Elan) 및 시스템 사용에 대한 실습과 개인 전사들 간의 전사 신뢰도 및 일치도 확보를 위해 다음과 같이 공동 전사를 실시하였다. 공동 전사의 내용 및 방법은 <표10>과 같다.

<표 10> 추진 내용 및 방법

| | |
|-------|--|
| 추진 내용 | <ul style="list-style-type: none"> □ 일시 : 2017. 6. 26.(월) ~ 2017. 6. 28(수), 09:00~18:00, 총 3회(24시간) □ 장소 : 한국복지대학교 인학관 211호(공통강의실) □ 대상 : 농인 전사자 5명, 청인 전사자 8명 □ 담당자 : 홍성은, 남기현, 이은영, 박정민, 김다해 |
| 추진 방법 | <ul style="list-style-type: none"> □ 전체 전사자를 대상으로 공동 전사 기간 동안 개인별 총 6시간을 선택하여 참여 □ 공동 전사 기간 동안 연구자 1인, 보조 연구원 1인 이상이 돌아가며 공동 전사 운영을 담당 □ 공동 전사 중 질의응답을 자유롭게 하여 전사에 대한 궁금증을 즉시 해소하고, 전사자 간 일치도를 높임 □ 공동 전사 기간 중 실시한 전사 결과의 일치도를 분석하여 공동 전사 및 재교육 실시 |

공동 전사의 추진 내용은 <표11>과 같다.

<표 11> 공동 전사 추진 내용

| 일자 | 시간 | 담당자 | 참석자 |
|---------------|-------------|-----------------|---|
| 6. 26. (월) | 09:00~16:00 | 홍성은, 이은영 김다해 | 이○○, 정○○, 김○○, 김○○, 이○○, 정○○ |
| 6. 27. (화) | 09:00~12:00 | 홍성은, 남기현 | 이○○, 이○○, 유○○, 임○○, 김○○ |
| | 13:00~16:00 | | 이○○, 이○○ |
| 6. 28. (수) | 09:00~12:00 | 홍성은, 이은영 | 송○○, 이○○, 유○○, 임○○, 김○○ |
| | 13:00~16:00 | | 이○○ |
| | | | 재교육 및 2차 공동 전사 - 김○○, 김○○, 이○○, 정○○, 송○○ |

공동 전사 장면은 <그림 32>와 같다.



<그림 32> 공동 전사 및 재교육 모습

농인 5명, 청인 8명, 뜻풀이 수어 사전 연구자 1명 총 14명이 참여하였다. 전사 신청자 중 87.5%가 공동 전사에 참여하였으며 결석자는 추후 전사 교육과 함께 전사 방법에 대한 상세 교육을 하였다.

3. 전사 품질 관리

전사의 품질을 관리하기 위해 다음과 같은 방법을 시행하였다.

- 모든 전사자는 일주일에 1번 실시하는 전사 회의에 의무적으로 참석한다. 전사 회의에서는 시스템 안에서 해결하지 못한 문제점에 대해 논의하고 해결 방법을 찾는다.
- 연구진 2명을 검토자로 지정하여 모든 전사 과정을 검토하게 한다. 전사자가 전사 작업을 마치고 시스템에 올리면(시스템에서는 ‘전사 완료’로 표시) 전사 파일이 검토자에게 분배 된다(시스템에서는 ‘검토 중’으로 표시). 검토자가 검토를 마치면 전사 파일은 시스템에 검토 완료 되어 등록된다. 즉, 모든 전사는 전사 중 ⇨ 전사 완료 ⇨ 검토 중 ⇨ 검토 완료를 거쳐 시스템에 등록된다.
- 새로운 타입을 등록할 때는 전사자 회의와 연구진의 검토를 거쳐 등록한다. 시스템에서는 새로운 타입이 후보로 표시되었다가 전사 회의

를 통해 새로운 타입으로 등록된다.

- 같은 영상 자료에 대해 1분간 전사하여 전사자 간 신뢰도를 확보한 후 본 전사를 실시한다.

3.1 전사 신뢰도 검사 결과

전사 품질 관리를 위해 전사 교육을 마치고 공동 전사 후에 3차례의 전사 신뢰도를 측정하였다. 1차 검사는 서울 지역 24번 쌍 M-농학교 경험이야기 과제(1분 3초), 2차 검사는 서울 지역 14번 쌍 M-농학교 경험이야기 과제(44초), 3차 검사는 서울 지역 31번 쌍 M-농학교 경험 이야기 과제(1분 16초), 4차 검사는 서울 지역 14번 쌍 M-농학교 경험이야기 전사로 실시하였다.

신뢰도 검사는 엘란(Elan) 프로그램에 있는 전사 신뢰도 측정 기능(Caculate Inter Annotator Reliability)을 사용하여 실시하였다. 먼저 전사 일치도를 확인하기 위해 전사 파일을 농인인 공동 연구원이 전사하고 2명의 공동 연구원이 검토하여 기준(정답)을 정하고 각 전사자들의 전사와 비교하여 그 신뢰도를 측정하였다.

엘란(Elan) 프로그램의 신뢰도 측정 방법은 음성 언어 전사의 신뢰도 측정을 위해 개발된 것이므로 분절 일치도가 51%이상인 경우에 한해 각각의 토큰에 대한 일치도를 계산한다. 따라서 수어 전사에서는 시각 언어의 특성상 분절의 일치를 확보하기 어려워 높은 신뢰도를 얻을 수 없었다. 이에 2차 신뢰도 검사 후에는 분절에 대한 재교육을 함께 실시하고 3차 신뢰도를 측정 하였다. 이후 3개월 (7월~9월)간의 전사를 진행하였으며 4차 신뢰도를 측정하였다. 4차에 걸친 신뢰도 검사 결과는 다음과 같다.

<표 12> 전사자 신뢰도 검사 결과

| 번호 | 전사자 | 1차 | 2차 | 3차 | 4차 |
|----|-----|-------|-------|-------|-------|
| 1 | 김○○ | 중도 포기 | 69.7 | 70.76 | 71.08 |
| 2 | 송○○ | 미참여 | 43.08 | 69.59 | 69.88 |
| 3 | 정○○ | 7.69 | 29.03 | 73.09 | 미참여 |
| 4 | 김○○ | 43.90 | 45.83 | 60.23 | 미참여 |
| 5 | 이○○ | 56.04 | 56.92 | 79.53 | 84.34 |
| 6 | 이○○ | 56.96 | 66.04 | 67.25 | 79.51 |
| 7 | 유○○ | 66.67 | 66.67 | 74.85 | 92.77 |
| 8 | 임○○ | 64.15 | 67.74 | 69.00 | 89.16 |
| 9 | 김○○ | 62.96 | 73.53 | 74.85 | 84.34 |
| 10 | 이○○ | 51.28 | 70.97 | 76.02 | 95.18 |
| 11 | 정○○ | 62.86 | 74.6 | 중도 포기 | 89.16 |
| 12 | 이○○ | 미참여 | | | 73.49 |
| 13 | 이○○ | 미참여 | | | 83.13 |
| 평균 | | 47.55 | 59.81 | 71.51 | 82.91 |

3차 신뢰도 측정 후 신뢰도 검사에서 일치도가 낮게 나타난 토큰의 특징을 분석하여 그 결과를 토대로 한 차례의 전사 교육을 추가로 실시한 후 본 전사에 들어갔다. 3개월의 전사 기간을 거친 후 4차 신뢰도 측정 결과 80%이상의 일치도를 보였다. 3차와 4차 신뢰도 검사에 대한 분석 결과는 다음과 같다.

3.2 3차 신뢰도 검사 개요 및 분석

3차 신뢰도 검사는 서울 지역 31번 쌍 M-농학교 경험 이야기를 1분 16초로 편집하여 전사자 10명을 대상으로 실시하였다. 동일 영상을 연구자가 전사한 후 연구자와 연구 보조원이 검토한 전사 결과를 검토

기준으로 정하였다. 3차 신뢰도 검사에서 실시한 전사 결과, 발생 토큰 수는 180개, 발생 타입 수는 77개로 검토에 해당 되는 타입은 % 타입을 제외한 76개였다. 토큰 수로는 176개가 검토 해당 사항이었다. 전사 결과 [지시1_1]이 12회, [지시1_비1]이 10회, [교사1]이 10회, [있다1]이 9회, [지시2_1]이 9회로 발생하였다.

기준 전사 결과와 전사자 전사 결과를 비교하여 일치도가 낮은 토큰을 분석하였으며 그 내용은 아래와 같다.

3.2.1 일치도가 낮은 토큰 분석

1) [구형1]

- 토큰 맥락: [행사1]+[지시1_비1]+[구형1]¹+[지문자:부스]+[구형1]²
+[삼성1]+[학교1]+[있다1]

- 번역

: 수어 문화제라.... ‘부스’를 하나 삼성학교가 맡았었어요.

- 전사자 간 발생 토큰

(1) @ / % / 구형1_선 / ### / 구형1 / 구형1_면

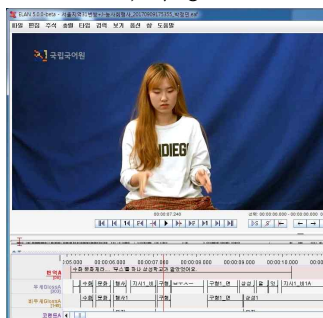
(2) 구형1_면 / 구형1 / 빌딩3/ 박물관1 / ###

- 일치도가 낮은 이유

(1) 생산적 수어와 고정된 수어의 구분이 모호하다.

(2) [구형1]과 [구형1_면]의 구분 규칙이 불명확하다.

1) 구형1



2) 구형1

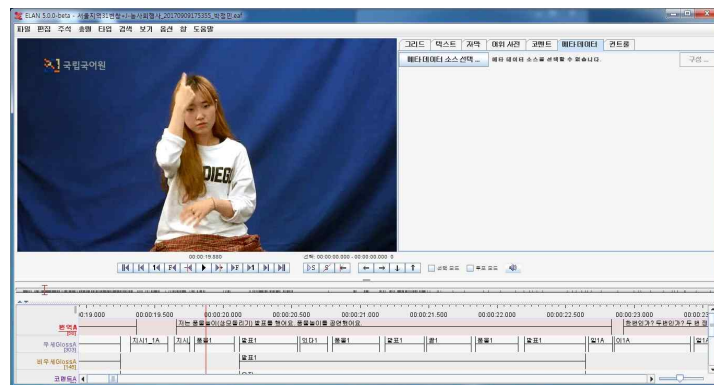


<그림 33> 일치도가 낮은 토큰- [구형1]

2) [풍물1]

- 토큰 맥락: [지시1_1]+[지시1_비1]+[풍물1]¹+[발표1]
+[있다1]+[풍물1]²+[발표1]+[끝1]

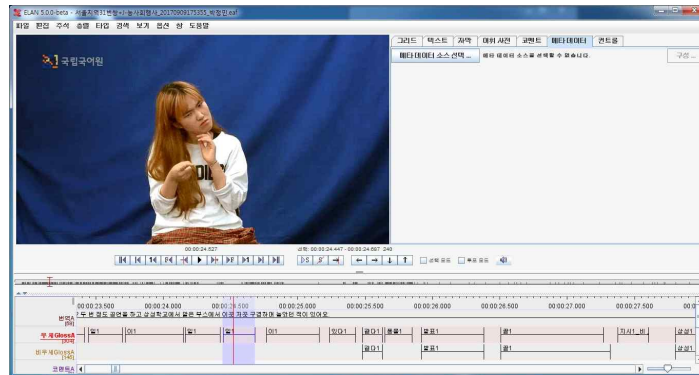
- 번역: 저는 풍물놀이(상모돌리기) 발표를 했어요. 풍물놀이 공연을 했어요.
- 전사자 간 발생 토큰
 - (1) 원1 / 풍물1 / 돌리다1 / ###
 - (2) 원1 / 풍물1 / 지시1_비1 / 다음1 / ###
- 일치도가 낮은 이유
 - (1) 생산적 수어와 고정된 수어의 구분이 모호하다.
 - (2) 의미 맥락을 파악하지 못하여 타입 구분에 오류가 생겼다.



<그림 34> 일치도가 낮은 토큰- [풍물1]

3) [일1]

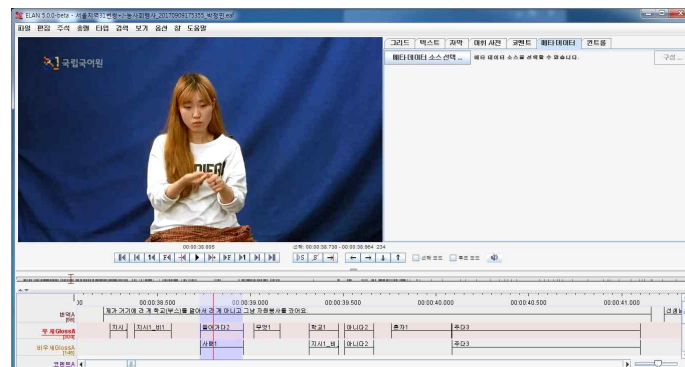
- 토큰 맥락: [풍물1]+[발표1]+[일1]+[이1]+[일1]+[이1]+[일1]
+[일1]+[이1]+[있다1]+[같다1]
- 번역: 풍물 발표를 한번인가? 두 번인가? 두 번 정도 공연을 했던 것 같아요.
- 전사자 간 발생 토큰: 지시1_비1 / 일형3 / 미분절 (3명)
- 일치도가 낮은 이유
 - (1) 의미 맥락을 파악하는 데 오류가 발생하였다.
 - (2) 앞 토큰과 동일하여 반복 수어로 판단하였다.
 - (3) 미분절하였다.
 - (4) [일1]을 두 번 하는 과정에서 [주먹형1]을 포함하였다.



<그림 35> 일치도가 낮은 토큰- [일1]

4) [들어가다2]

- 토큰 맥락: [지시1_1]+[지시1_비1]+[들어가다2]+[무엇1]
+[학교1]+[아니다2]
- 번역: 내가 거기에 간 건 학교를 말아서 간 게 아니고
- 전사자 간 발생 토큰: 혼자1(5명) / 지시1_1 / ###
- 일치도가 낮은 이유
 - (1) 오류 발생 수어로 인해 의미 파악이 모호하다.
 - (2) 다양한 각도에서 수어를 보아야 명확한 의미 파악이 가능하다.



<그림 36> 일치도가 낮은 토큰- [들어가다2]

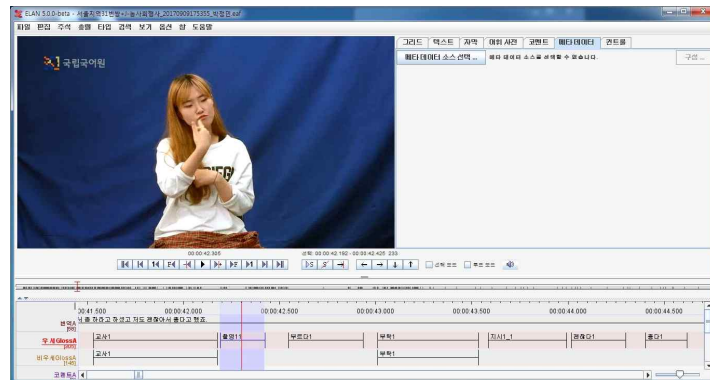
5) [촬영11]

- 토큰 맥락: [교사1]+[촬영11]+[부르다1]+[부탁1]+[지1_1]
+[괜찮다1]+[좋다1]
- 번역: 선생님이 영상 통화로 부르셔서 부탁하셨습니다. ‘나는 괜찮다 좋다’라고…….
- 전사자 간 발생 토큰

: 촬영1 / ### / 보다17/ 물어보다1 / 누구1 / 지시1_1

- 일치도가 낮은 이유

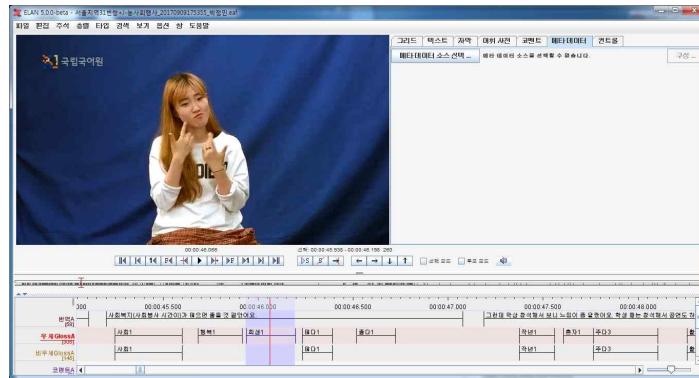
- (1) 의미 맥락과 함께 오류 발생 수어를 파악하기 모호하다.
- (2) 여러 각도에서 수어를 보아야 명확한 의미 파악이 가능하다.
- (3) 빠르게 지나가는 수어를 인지하지 못하였다.



<그림 37> 일치도가 낮은 토큰- [촬영11]

6) [희생1]

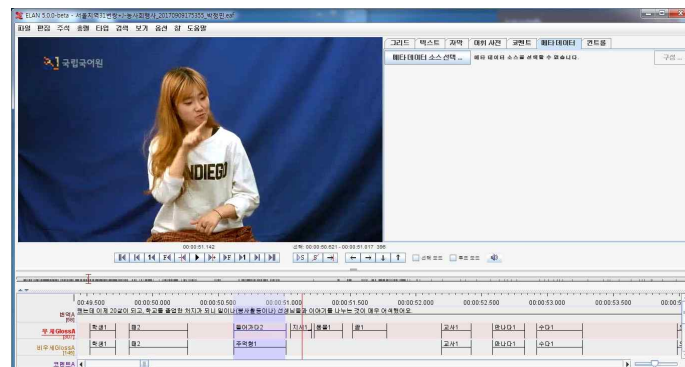
- 토큰 맥락: [사회1]+[행복1]+[희생1]+[많다1]+[좋다1]
- 번역: 사회복지의 봉사활동이 많으면 좋아요.
- 전사자 간 발생 토큰
: 제스처 / 주다1(3명) / 미분절(5명, [많다1])
- 일치도가 낮은 이유
 - (1) 수어를 파악하기 모호하다. (미분절, 주다1).
 - (2) 뒤의 수어 [많다1]과 연결되어 [희생1]을 알기 어렵다.
 - (3) 전사자가 토큰 목록을 미숙지하였다. (주다1).
 - (4) [주다3]의 수어와 유사하나 전사자 3명이 [주다1]로 입력하였다.
 - (5) 이는 타입명에 붙은 숫자는 빈도에 따른 순차가 아닌, 수어 발견 순차임을 미숙지하였기 때문으로 보인다.
 - (6) 그로 인해 [주다3]을 [주다1]로 전사한 것으로 보인다.
 - (7) [희생1]과 [주다3]의 수어 형태가 비슷함으로 의미 맥락에서 [희생1]과 [주다3]을 구분할 필요가 있다.



<그림 38> 일치도가 낮은 토큰- [희생1]

7) [지시1_비1]

- 토큰 맥락: [학생1]+[때1]+[들어가다2]+[지시1_비1]
+[풍물1]+[끝1]
- 번역: 학생 때 그 풍물놀이에 참여했어요.
- 전사자 간 발생 토큰: [무엇1] / ### / [풍물1](5명)
- 일치도가 낮은 이유
 - (1) 수어를 파악하기 모호하다.
 - (2) 뒤에 따라오는 [풍물1]과 연결되고, 수형이 같다.

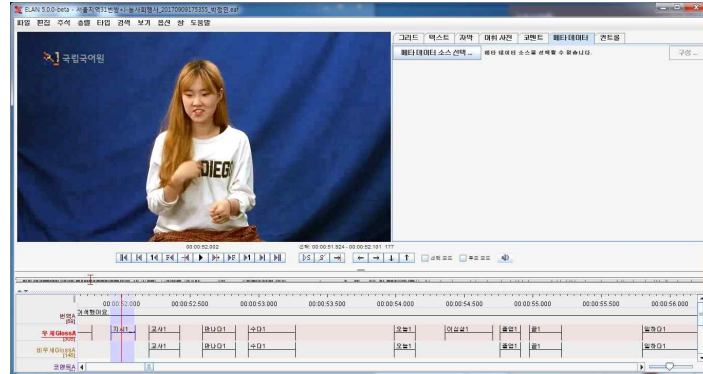


<그림 39> 일치도가 낮은 토큰- [지시1_비1]

8) [지시1_1]

- 토큰 맥락: [지시1_1]+[교사1]+[만나다1]+[수다1]
- 번역: 선생님과 만나 이야기를 나눴다.
- 전사자 간 발생 토큰
 - : [지시2_1] / ### / 미분절(5명,[교사1])
- 일치도가 낮은 이유

- (1) 수어를 파악하기 모호하다.
- (2) [지시1_1]의 수형이 뒤 [교사1]의 수형으로 변이되었다.
- (3) [교사1]을 하기 위한 연결된 수동으로 볼 수 있다.



<그림 40> 일치도가 낮은 토큰- [지시1_1]

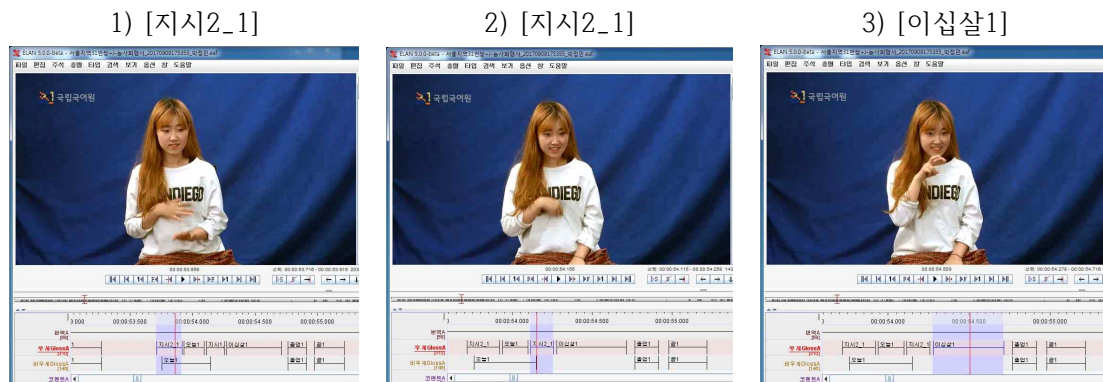
9) [지시2_1] / [이십살1]

- 토큰 맥락
 - : [지시2_1]¹⁾+[오늘1]+[지시2_1]²⁾+[이십살1]³⁾
+[졸업1]+[끝1]
- 번역: 나는 이제 스무살이라 졸업을 했어요.
- 전사자 간 발생 토큰
 - (1) [지시1_1] / [가다2] / 미분절(4명,[오늘1])
 - (2) [혼자1] / [솔직하다1] / ### / 미분절(6명, 뒷수어)
 - (3) [이십오1] / [이십1](4명) / [노래1] / [쥐1] / 미분절
- 일치도가 낮은 이유
 - (1) [지시2_1]
 - ① 수어를 파악하기 모호하다.
 - ② [지시2_1]의 수형이 뒤 [오늘1]의 수형과 같다.
 - ③ [오늘1]을 하기 위한 연결된 수동으로 볼 수 있다.
 - ④ 우세 [지시2_1]과 비우세 [오늘1]이 함께 나타남으로 [오늘1]로 볼 수 있다.
 - ⑤ 주어로서의 ‘나’를 의미하는 것으로 보았다.
 - (2) [지시2_1]
 - ① 수어를 파악하기 모호하다.
 - ② [이십살1]을 하기 위한 연결된 수동으로 볼 수 있다.

수어가 명확하지 않아 [혼자1], [솔직하다1]등의 다른 의미로 해석 될 수 있다.

(3) [이십살1]

- ① 전사자가 [이십살1]과 [이십1] 두 타입이 있음을 알지 못하였다.
- ② 타입을 미숙지 한 것으로 보인다.
- ③ 의미 맥락을 파악하는 데 오류가 발생하였다.
- ④ 언어 제공자의 수어가 불확실하다.
- ⑤ 나이를 나타내는 수동이 아니다.

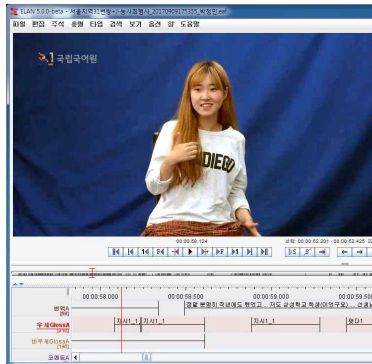


<그림 41> 일치도가 낮은 토큰- [지시2_1] / [지시2_1] / [이십살1]

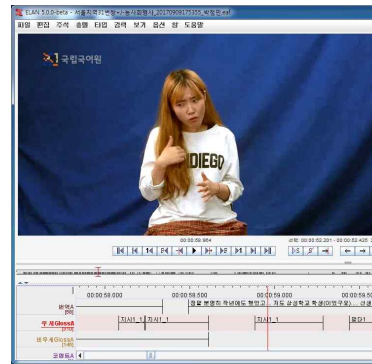
10) [지시1_1]

- 토큰 맥락: [지시1_1]¹⁾+ [지시1_1]+[지시1_1]²⁾+ [맞다1]
+ [확실하다1]+[지시2_1]+[작년1]+[끝1]
- 번역: 나.. 나? 내가... 맞아요. 확실히 저는 작년에 했어요.
- 전사자 간 발생 토큰
 - 1) [지시2_1] / 미분절(3명, [지시1_1])
 - 2) [지시2_1] / 미분절(6명, [지시1_1])
- 일치도가 낮은 이유
 - (1) 언어 제공자가 [지시1_1]과 [지시2_1]의 수형이 모호하여 수어를 파악하기 어렵다.
 - (2) 비수지를 파악하지 못하여 의미 맥락을 파악하는 데 오류가 발생한 것으로 보인다.
 - (3) 이에 반복 수어로 인지하여 분절하지 않았다.

1) [지시1_1]



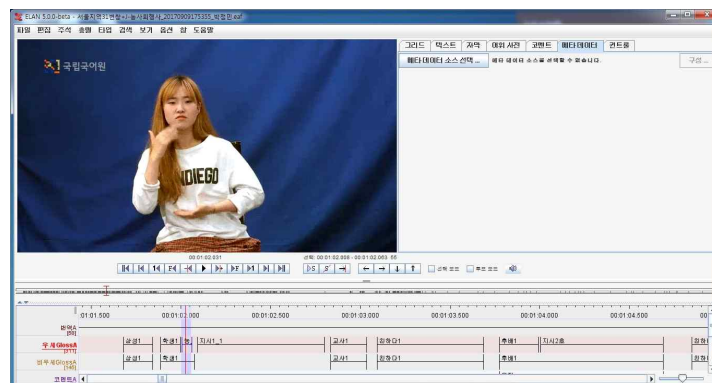
2) [지시1_1]



<그림 42> 일치도가 낮은 토큰- [지시1_1]

11) [농1]

- 토큰 맥락: [지시1_1]+[삼성1]+[학생1]+[농1]+[지시1_1]
+[교사1]+[친하다1]
- 번역: 삼성농학교 학생일 때 저는 선생님과 친했어요.
- 전사자 간 발생 토큰: 미분절 (9명 [학생1])
- 일치도가 낮은 이유
 - (1) 앞 [학생1]과 뒤 [지시1_1] 사이의 찢나에 발생하였다.
 - (2) 이에 수어를 파악하기 어렵다.

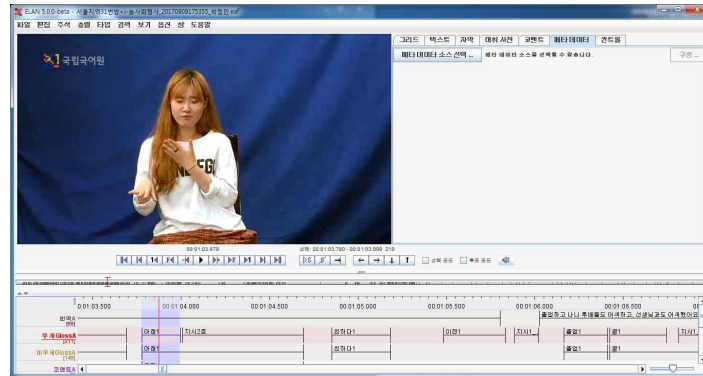


<그림 43> 일치도가 낮은 토큰- [농1]

12) [아래1]

- 토큰 맥락: [아래1]+[지시2호]+[친하다1]
- 번역: 후배들과 친하다
- 전사자 간 발생 토큰: [후배1](8명) / [이후1]

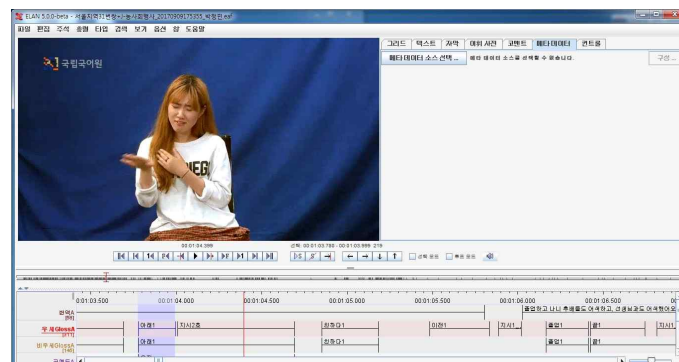
- 일치도가 낮은 이유
 - (1) 타입 목록에 오류가 있었다.
 - (2) [아래1], [후배1]의 타입 정리가 필요하다.



<그림 44> 일치도가 낮은 토큰- [아래1]

13) [지시2호]

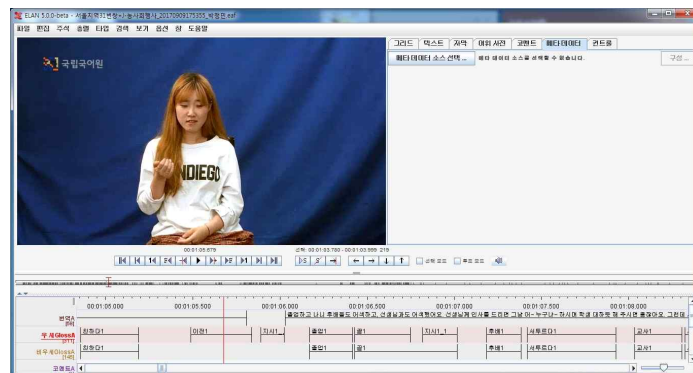
- 토큰 맥락: [아래1]+[지시2호]+[친하다1]
- 번역: 후배들과 친하다
- 전사자 간 발생 토큰: [우리1] / ### / [지시2] / 미분절
- 일치도가 낮은 이유
 - (1) 의미 맥락을 파악하는 데 오류가 발생하였다.
 - (2) 수형의 의미는 ‘우리’이나 비수지를 포함한다.
 - (3) ‘우리’, ‘여기 있는 것 모두’, ‘다’을 [지시1호] 또는 [지시2호]임에 대한 교육이 필요하다.
 - (4) ‘우리’, ‘여기 있는 것 모두’, ‘다’에 해당되는 수어를 고정된 수어로 인식함에 따라 [지시1호], [지시2호]로 생각하지 못하였다.



<그림 45> 일치도가 낮은 토큰- [지시2호]

14) [이전1]

- 토큰 맥락: [친하다1]+[이전1]+[지시2_1]+[졸업1]+[끝1]\
+[지시2_1]+[아래1]+[서투르다1]
- 번역
: 친했어요. 전에도 졸업하고 나니 후배들도 어색하고...
- 전사자 간 발생 토큰: % / 미분절(7명)
- 일치도가 낮은 이유
 - (1) 수어를 파악하기 모호하다.
 - (2) 앞 [친하다1]과 뒤 [지시2_1] 사이의 연결 과정으로 볼 수 있다.
 - (3) 손의 움직임으로 인식하였으나 의미가 없거나 모르는 수어 또는 하려다가 그만둔 수어로 인식할 수 있다.



<그림 46> 일치도가 낮은 토큰- [이전1]

15) 기타

- 반복하는 수어의 분절이 일치 하지 않는 경우이다.
: [교사1] / [서투르다1] / [당황1] / [다르다1]

3.2.2 모두 일치하는 토큰

수어의 형태가 앞뒤 수어와 확연히 달라 앞뒤 수어에 영향을 받지 않는 경우 토큰의 일치도가 높았다. 또한 [수어1], [문화1], [삼성1], [있다1], [학생1], [발표1], [끝1], [놀다1], [지시1_1], [졸업1], [이후1], [되돌리다1], [괜찮다1], [좋다1], [행복1], [느끼다1], [일하다1], [만나다1], [맞다1], [다르다1]과 같이 수어의 형태가 고정되어 있고, 의미가 뚜렷한 토큰의 경우 전사자의 전사 내용이 모두 일치하

였다.

3.3 4차 신뢰도 검사 개요 및 분석

4차 신뢰도 검사는 본 전사를 시작하고 3개월 후에 실시하였다. 서울 지역 14번 쌍 M-농학교 경험 이야기를 56초로 편집하여 전사자 11명을 대상으로 실시하였다. 동일 영상을 이은영 연구자가 전사한 후 홍성은, 박정민이 검토를 마친 전사 결과를 검토 기준으로 정하였다. 4차 신뢰도 검사에서 실시한 전사 결과, 발생 토큰 수는 97개, 발생 타입 수는 50개로 검토에 해당 되는 타입은 @와 % 타입을 제외한 48개였다. 토큰 수로는 83개가 검토 해당 사항이었다. 전사 결과 다수 발생한 타입은 [없다11](9회), [학교1](9회), [지시2_1](6회), [형1](6회)이었다..

기준 전사 결과와 전사자 전사 결과를 비교하여 일치도가 낮은 토큰을 분석하였으며 그 내용은 아래와 같다.

3.3.1 일치도가 낮은 토큰 분석

1) [지시2_1]

- 토큰 맥락

: 지시2_1]¹⁾[학교1]+[들어가다1]+[아니다1]+[그러면1]

- 번역 : 내가 학교에 들어가지 않았다면

- 전사자 간 발생 토큰 : 지시1_1 / 지시1_비1

- 일치도가 낮은 이유

: [지시2_1]과 [지시1_1]의 형태를 구별하기 어렵다.

- 개선 방법 : 다양한 각도에서의 촬영이 필요하다.



<그림 47> 일치도가 낮은 토큰- [지시2_1]

2) [대부분1] / [무엇1]

- 토큰 맥락
: [대부분1]¹⁾+ [무엇1]²⁾+ [없다11]+ [대화1]+ [안통하다1]
- 번역: 대부분의 무학력자들은 (어떠냐면) 대화할 때 소통이 잘 되지 않거든요.
- 전사자 간 발생 토큰
1) 깨달다2(6명) / 훌륭하다2 / 갑자기2 / ###
2) 지시1 / 토큰 없음(5명)
- 일치도가 낮은 이유
: [대부분1]과 [깨달다2]의 형태와 비슷하다.
: 언어 제공자의 습관, 편의 등에 의해 발화의 위치가 눈 쪽으로 이동된다.
: 이로 인해 [대부분1]을 [깨달다2]로 전사한 전사자가 많다.
: 이전 수어 [대부분1]과 연결되어 나타남으로 [무엇1]을 못 볼 가능성이 있다.
: [무엇1]의 형태가 [대부분1]의 형태와 비슷하여 [대부분1]의 수어를 거두는 과정으로 볼 수 있다.
- 개선 방법
: 전사를 시작하기 전 내용을 확인하여 영상의 내용과 언어 제공자의 언어 습관을 파악할 필요가 있다.
: 주석 입력 전 영상을 다양한 속도로 확인할 필요가 있다.

1) [대부분1]



2) [무엇1]



<그림 48> 일치도가 낮은 토큰- [대부분1] / [무엇1]

3) [지시1_1]

- 토큰 맥락
: [지시1_1]¹⁾+ [학교1]+ [글1]+ [깨달다2]+ [크다1]+ [제일1]

+ [중요1]

- 번역
: 나는 학교에서 글을 배우는 게 무척 중요하다고 생각해요.
- 전사자 간 발생 토큰: 지시1_비1 / 지시2_1(4명)
- 일치도가 낮은 이유
: [지시2_1]과 [지시1_1]의 형태를 구별하기 어렵다.
- 개선 방법 : 다양한 각도에서의 촬영이 필요하다.



<그림 49> 일치도가 낮은 토큰- [지시1_1]

4) [없다11]

- 토큰 맥락: [학교1]+[없다1]+[그러면1]+[없다11]+[많다1]
+ [지시1_비1]+[없다11]+[시골1]+[없다11]¹⁾
+ [없다11]²⁾+ [없다11]³⁾+ [지시1호]+[많다1]
- 번역: 학교가 없으면 배울 수 없는 농인이 많죠. 저기 시골에
가 보면 배움을 받지 못한 농인들이 여기 저기 많이 있
어요.
- 전사자 간 발생 토큰: 토큰 없음
- 일치도가 낮은 이유: 3회 반복되는 [없음11]을 한번 분절하였
다.
- 개선 방법: 반복 수어에 대한 전사 규칙을 보다 세밀하게 정하
고 이를 교육한다.

1) [없다11]



2) [없다11]



3) [없다11]



<그림 50> 일치도가 낮은 토큰- [없다11]

5) [제스처]

- 토큰 맥락: [형1]+[지시2_1]+[제스처]+[지시2_1]
+[제스처]¹⁾+ [제스처]²⁾
- 번역: 오빠가 “내가 최고지?”라고 물어서 “오빠가 최고야”라고 했어요.
- 전사자 간 발생 토큰: 토큰 없었다.
- 일치도가 낮은 이유
 - : 2회 반복되는 [제스처]를 단순 반복으로 보았다.
 - : 이에 다른 의미로 사용된 [제스처]를 한번으로 분절하였다.
 - : [제스처]¹⁾에 연결되는 비수지와 [제스처]²⁾에 연결되는 비수지를 인지하지 못하였다.
- 개선 방법: 반복 수어에 대한 전사 규칙을 보다 세밀하게 정하고 이를 교육하는 것이 필요하다.
 - : 수어를 보다 면밀히 읽을 수 있는 전사자 양성이 필요하다.

1) [제스처]



2) [제스처]



<그림 51> 일치도가 낮은 토큰- [제스처]

6) [완전1]

- 토큰 맥락: [지시2_1]+[학교1]+[다행1]+[형1]+[왜1]
+[시골1]+[없다11]+[완전1]¹⁾
- 번역: 나는 학교를 다녀 다행이에요. 오빠가 왜 보냈냐면 시골에는 학교를 다니지 못한 농인이 심각할 정도로 많거든요
- 전사자 간 발생 토큰: 모두1 / 많다1 / ### (3명))
- 일치도가 낮은 이유

- : 수어의 형태가 변형되어 사용되었다.
- : 수어 하나에 비수지가 다양하게 포함되어 의미를 이룬다.
- : 이로 인해 타입을 연상하는 데에 방해를 받은 것으로 보인다.

- 개선 방법

- : 다양한 요소를 가진 수어 변이를 타입과 연결할 수 있는 경험 축적이 필요하다.
- : [완전1]과 같은 타입을 수집하여 전사 교육 시 활용한다.



<그림 52> 일치도가 낮은 토큰- [완전1]

3.3.2 모두 일치하는 토큰

3차 일치도 검사 때와 같이 수어의 형태가 앞뒤 수어와 확연히 달라 앞뒤 수어에 영향을 받지 않을 경우 토큰의 일치도가 높았다. 또한 [맞다1], [학교1], [아니다1], [그러면1], [대화1], [안통하다1], [낮다1], [글1], [제일1], [중요1], [없다1], [많다1], [지시1_비1], [시골1], [졸업1], [남동생1], [데려오다1], [고맙다1], [다행1] 과 같이 수어의 형태가 고정되어 있고, 의미가 뚜렷할 경우 전사자 간 일치도가 높았다.

4. 전사 결과

총 261분 57초 분량의 영상에 대해 우세손과 비우세손에 대한 기본 전사 및 번역을 완료하였다. 기본 전사 및 번역을 완료한 파일은 토론 과제(과제 번호 E) 4개, 주제 영역 과제(과제 번호 H) 3개, 농사회 행사(과제 번호 J) 4개, 농학교 경험 이야기(과제 번호 M) 4개로 총 15개이며, 발생 토큰 수는 23,637개이다. 토큰 수는 우세손을 기준으로 산출하였으며, 전사 및 번역을 완료한 각 과제별 세부 정보는 다

음의 <표 13>과 같다.

<표 13> 각 과제별 세부 정보

| 연번 | 과제명 | 영상 시간 | 토큰 수 | 새 타입 수 |
|----|--------------------------|----------|--------|--------|
| 1 | 서울 지역 31번 쌍 J-농사회 행사 | 00:07:59 | 615 | 39 |
| 2 | 서울 지역 14번 쌍 J-농사회 행사 | 00:04:03 | 295 | 35 |
| 3 | 서울 지역 20번 쌍 J- 농사회 행사 | 00:12:13 | 1081 | 61 |
| 4 | 서울 지역 31번 쌍 M-농학교 경험 이야기 | 00:26:05 | 2326 | 80 |
| 5 | 서울 지역 31번 쌍 E-토론 | 00:14:15 | 801 | 12 |
| 6 | 서울 지역 24번 쌍 M-농학교 경험 이야기 | 00:29:16 | 3134 | 85 |
| 7 | 서울 지역 24번 쌍 E-토론 | 00:29:14 | 3507 | 59 |
| 8 | 서울 지역 20번 쌍 M-농학교 경험 이야기 | 00:33:38 | 2032 | 124 |
| 9 | 서울 지역 20번 쌍 H-주제 영역 | 00:22:56 | 2417 | 205 |
| 10 | 서울 지역 20번 쌍 E-토론 | 00:11:21 | 1083 | 87 |
| 11 | 서울 지역 16번 쌍 J-농사회 행사 | 00:21:17 | 1798 | 46 |
| 12 | 서울 지역 16번 쌍 H-주제 영역 | 00:21:17 | 2173 | 117 |
| 13 | 서울 지역 14번 쌍 M-농학교 경험 이야기 | 00:11:13 | 1019 | 72 |
| 14 | 서울 지역 14번 쌍 H-주제 영역 | 00:09:03 | 842 | 51 |
| 15 | 서울 지역 14번 쌍 E-토론 | 00:07:07 | 504 | 2 |
| | 총 15개 파일 | 261분57초 | 23,627 | 1,075 |

5. 한국수어의 어휘 변이 연구

5.1 한국수어에 나타난 색 이름

5.1.1 분석 대상 및 방법

한국수어에서 나타난 색 이름의 다양한 형태(타입)를 알아보기 위하여 분석한 자료는 ‘개별 수어의 다양한 표현(과제 번호 1)’ 과제로 유도한 영상이다. 사용된 유도 과제는 다양한 형태의 수어가 존재할 것이라고 생각되는 7가지의 색깔과 미원, 콩나물은 시각 자료로, 고모, 이모, 삼촌과 같은 가족 관계를 나타내는 명칭은 한글로 제시하여 수어로 어떻게 표현하는지 질문하였다.

분석에 사용한 전사 자료는 2015년에 전사 완료한 52분 17초 분량의 영상에 대해 우세손과 비우세손의 움직임을 기본 전사한 파일이다. 유도 과제의 특성상 타입의 발생 빈도는 의미가 없으므로 언어 제공자 수로 분석하였다. 이에 본 사례 연구에서 제시하는 분석 결과는 한정된 과제와 언어 제공자에게서 추출한 데이터를 사용하였으므로 결과의 해석을 제한적으로 일반화하여야 한다. 분석한 총 파일은 25개로 언어 제공자는 총 50명이며, 우세손 기준 발생 타입 수는 754개, 토큰 수는 10,082개였다.

5.1.2 분석 결과

유도 과제에서 제시한 색은 갈색, 분홍, 보라, 주황, 초록, 파랑, 회색으로 7가지이다. 이 과제를 통해 분석한 색 관련 수어 타입은 총 27개로 갈색, 노랑, 분홍(4), 보라(4), 빨강(2), 자주색, 주황(8), 초록(2), 파랑, 하양과 포도, 쥐이다(팔호 안은 타입 개수임). 분석 결과 하양, 빨강, 파랑, 노랑회색을 나타내는 수어의 형태는 1가지인데 비해, 혼합된 색을 나타내는 수어는 그 형태가 다양하게 나타나거나 [빨강1]+[파랑1]+[섞다1] 또는 [빨강1]+[약하다1]와 같이 복합어로 표현하였다. 또한 모든 색에서 공통적으로 지문자를 사용하였다.

연령대별로 보면, 모든 연령대에서 발생하는 타입은 [갈색1], [노랑1], [분홍1], [빨강1], [초록1A], [파랑1], [하양1], [포도1], [쥐1]

이었으며, 60대 이상에서 발생하지 않는 타입은 [분홍1A·B·C], [보라 1,1A·B], [자주색1], [주황1~8], [초록2]였으며, 이들 언어 제공자들은 분홍, 주황, 보라를 복합어로 표현하거나, 여성의 옷 색으로 표현한 언어 제공자도 있었다. 영국수어 연구 결과, (Woll, 1983) Bristol과 Cardiff에서는 젊은 농인들은 나이든 농인이 사용하는 갈색, 초록색, 보라색, 노란색을 나타내는 수어를 사용하지 않았으며, 같은 수어로 여러 색을 표현하면서 마우딩으로 구분한다고 하였다. 이는 다양한 색을 나타내는 수어 형태를 나이든 농인보다 젊은 농인이 많이 사용하는 우리나라 결과와 다르다. 유도 과제에서 제시한 색깔별로 나타나는 수어 형태를 요약하면 다음과 같다.

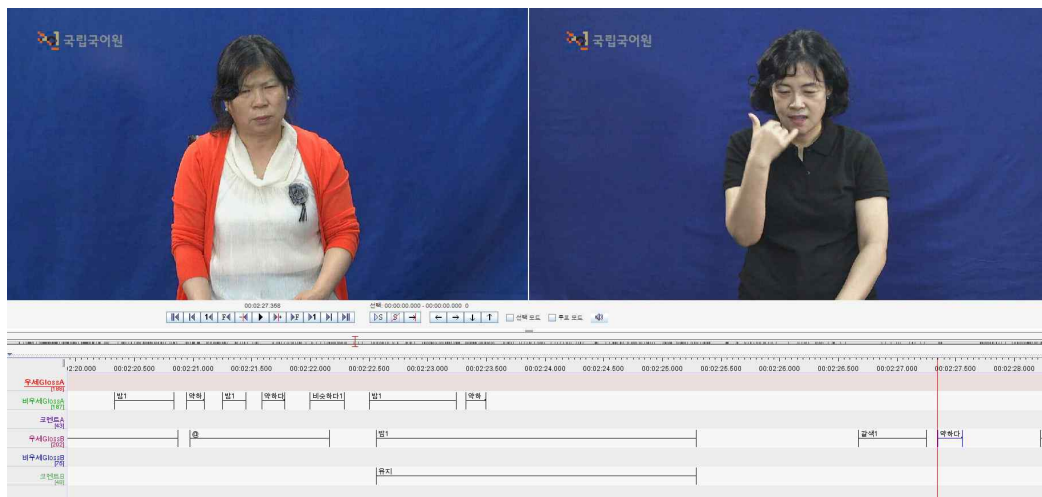
1) 갈색

갈색은 2가지 형태의 수어로 나타났다. 46명의 언어 제공자가 [갈색1]로, 9명의 언어 제공자가 [밤1]로 표현하였다. 그 외 지문자로 [ㅎㅏㅓ ㅓㅓ ㅓㅓ ㅓ] 표시하거나, 두 개 이상의 수어를 조합해서 표현하기도 하였다. 즉, [밤1]+[약하다1]로 하거나, [빨강1]+[노랑1]+[섞다1] / [밤1]+[노랑1]+[섞다1]처럼 표현하기도 하였다. 또한 한 언어 제공자는 [갈대]라는 사물을 나타내는 수어로 대체하여 갈색을 표현하기도 하였다. 연령대별로 살펴보면, 20대, 30대, 40대, 60대의 언어 제공자는 모두 [갈색1]로 표현하였으며, 그 중에서는 [밤1]을 함께 사용하는 언어 제공자도 있었다. 그에 비해 50대 언어 제공자 9명 중에는 7명은 [갈색1]로, 2명은 [밤1]로 표현하였고, 70대 언어 제공자는 2명은 [갈색1], 또 다른 2명은 [밤1]로 표현하였다. 2가지 타입에 대한 연령별 언어 제공자 수는 <표 14>에 요약하였으며, 다양한 형태로 나타내는 예는 <그림 53>에 제시하였다.

<표 14> 갈색을 나타내는 수어 타입과 언어 제공자 수

| 언어 제공자 타입 | | 20대 (6명) | 30대 (15명) | 40대 (14명) | 50대 (9명) | 60대 (2명) | 70대 (4명) | 계 (50) |
|--------------|---|-------------|--------------|--------------|-------------|-------------|-------------|-----------|
| 갈색1 |  | 6 | 15 | 14 | 7 | 2 | 2 | 46 |
| 밤1 |  | 1 | 1 | 2 | 2 | 1 | 2 | 9 |

()속은 분석 대상 과제에서의 언어 제공자 수



<그림 53> 갈색: [밤1]+[약하다1] / [갈색1]+[약하다1]- 50대

2) 분홍

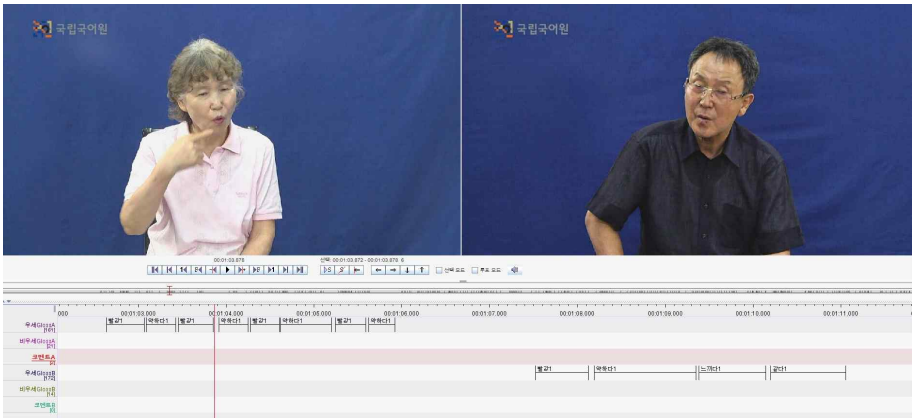
분홍을 나타내는 수어는 42명의 언어 제공자가 [분홍1]로, 5명 언어 제공자는 [분홍1A, B, C]와 같은 음운 변이의 형태로 표현하였으므로 분홍을 나타내는 수어는 1개의 타입이 나타났다고 볼 수 있다. 연령별로 살펴보면, 20대에서 50대까지의 언어 제공자는 대부분 [분홍1] 또는 그의 음운 변이 형태인 1개의 수어 단어로 표현한 것에 비해, 60대 이상의 언어 제공자 중 2명은 [분홍1]을 사용하면서 동시에 그리고 나머지 4명은 [빨강1]과 함께 [약하다], 또는 비수지 등으로 분홍색을 표현하였다. 그 외 70대의 언어 제공자 중에는 [빨강1]이라는 수어로 표현한 사람도 있었다. 즉, 분홍을 [빨강1]+[약하다1] 또는 [약하다1]+[빨강1] 등으로 표현하였다. 이렇게 볼 때 분홍색을 나타내는 수어의 형태는 1가지이지만 2개의 어휘를 사용한 복합어 형태로 표현하기도 하였다.

각 형태를 표현한 연령별 언어 제공자 수는 <표 15>에 요약하였으며, 다양한 형태로 나타내는 예는 <그림 54>에서 <그림 57>까지 제시하였다.

<표 15> 분홍을 나타내는 수어 타입과 언어 제공자 수

| 언어 제공자 | | 20대 (6명) | 30대 (15명) | 40대 (14명) | 50대 (9명) | 60대 (2명) | 70대 (4명) | 계 |
|--------|--|-------------|--------------|--------------|-------------|-------------|-------------|----|
| 타입 | | | | | | | | |
| 분홍1 |  | 4 | 14 | 14 | 8 | 1 | 1 | 42 |
| 분홍1A |  | - | 1 | - | - | - | - | 1 |
| 분홍1B |  | - | 1 | - | - | - | - | 1 |
| 분홍1C |  | 1 | 1 | 1 | - | - | - | 3 |

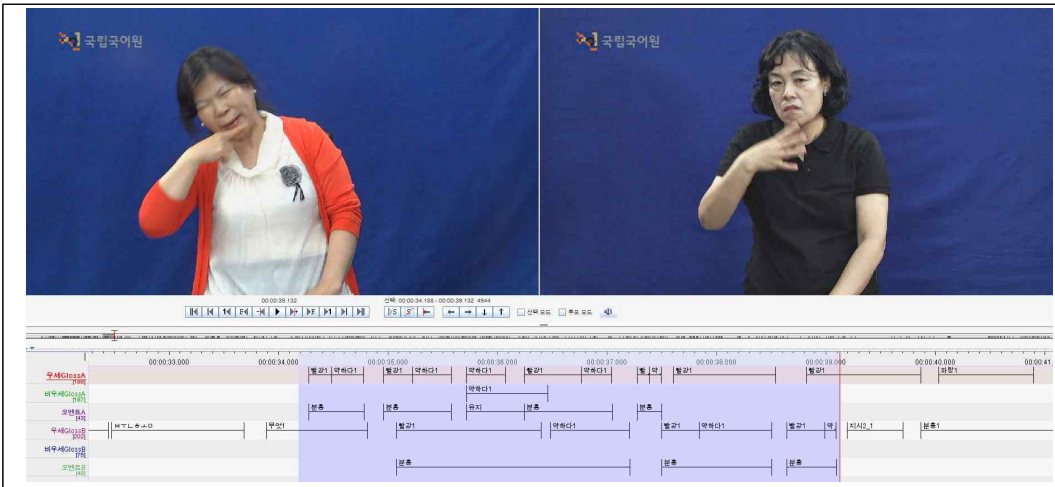
()속은 분석 대상 과제에서의 언어 제공자 수



<그림 54> 분홍: [빨강1]+[약하다1]-70대



<그림 55> 분홍: [빨강1]+[약하다1] - 60대



B: “분홍(지화)”을 어떻게 해요?
A: [빨강1]+[약하다1] 라고 해요.
B: 아... 그렇게 해요?
A: [약하다1]는 거죠. 빠~~강도 아니고
B: 나는 이렇게 해요. (수어를 보여주며)

이 담화에서 [빨강1]+[약하다1]은 수어 간 간격이 거의 없어 한 수어로 보일 정도로 [빨강1]+[약하다1]이 붙어 있음.

<그림 56> 분홍: [빨강1]+[약하다1]-50대



<그림 57> 분홍: [빨강1]+[퍼지다1]-60대 / [빨강1]+비수지-30대

3) 보라


보라는 4가지 형태의 수어로 나타났다. 총 50명의 언어 제공자 중 42명의 언어 제공자가 [보라1] 또는 [보라1A, B]와 같은 음운 변이의 형태로 표현하였다. 그리고 8명의 언어 제공자는 [포도1], 4명의 언어 제공자는 [보라2]라는 표현을 사용하기도 하였다. 그리고 70대 언어 제공자 1명은 [자주색1]이라는 수어를 하면서 마우딩으로 ‘자주색’이

라고 한 후에 지문자로 자주라고 표현하였다. 그 외에도 지문자로 보라라고 쓰거나, [포도1]과 [색]이라는 수어로 표현하기도 하였다. 또한 60대 이상의 언어 제공자들은 [빨강1]+[파랑1]+[섞다1] 또는 [빨강1]+[파랑1]+[중1] 등으로 표현하였다. 이렇게 볼 때 보라색을 나타내는 수어 형태는 4가지이며, 2개 이상의 어휘가 연결된 복합어로 나타내기도 하였다. 4가지 타입을 언급한 언어 제공자 수는 <표 16>에 요약하였으며, 다양한 형태로 나타내는 예는 <그림 58>, <그림 59>에서 제시하였다.

<표 16> 보라를 나타내는 수어 타입과 언어 제공자 수

| 언어 제공자 타입 | | 20대 (6명) | 30대 (15명) | 40대 (14명) | 50대 (9명) | 60대 (2명) | 70대 (4명) | 계 (50명) |
|--------------|---|-------------|--------------|--------------|-------------|-------------|-------------|------------|
| 보라1 |  | 5 | 13 | 13 | 8 | - | - | 39 |
| 보라1A |  | - | 1 | - | - | - | - | 1 |
| 보라1B |  | 1 | 1 | - | - | - | - | 2 |
| 보라2 |  | 2 | 1 | - | 1 | - | - | 4 |
| 자주색1 |  | - | - | - | - | - | 1 | 1 |
| 포도1 |  | 1 | 2 | 1 | 2 | 1 | 1 | 8 |

()속은 언어 제공자 수



A: 이건 수어가 없는데... 수어가 없죠?

B: [보라1]라고 하죠.

A: 아... [보라1]이라고 해요? 처음 보네요. 아~ 그렇게 하는구나.

B: “보라색 그거 가지고 와” 할 때 [보라1]을 써요.

A: 지문자 ‘ㅂ’, ‘ㄹ’을 하는 거군요. 아...

B: [포도1]+[색1]? [포도1]+[색1]이라고도 하죠.

A: 맞아요. [포도1]+[색1]라고 하죠.

<그림 58> 보라: [포도1]+[색1] - 50대



<그림 59> 보라: [빨강1]+[파랑1]+[중1] / [빨강1]+[파랑1]+[섞다1] -70대

4) 주황

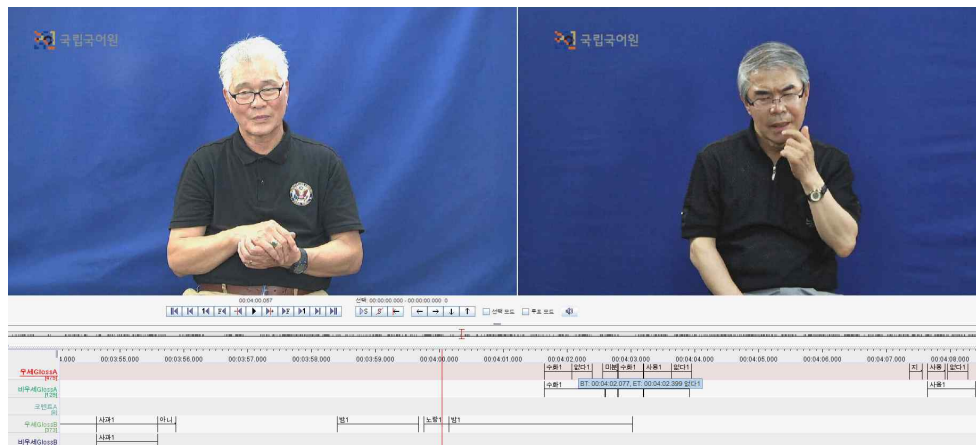
주황은 9가지 형태의 수어로 가장 많은 형태가 나타났다. 30명의 언어 제공자가 [주황4]로, 11명 언어 제공자가 [주황1]로, 각각 6명이 [주황5]와 [주황6]으로 표현하였으며, 각각 4명의 언어 제공자가 [주황3]과 [감1], 2명이 [주황2], 그리고 각각 1명의 언어 제공자가 [주황7]과 [주황8]로 표현하였다. 연령별로 살펴보았을 때, 60대 이상의 언어 제공자 모두는 단일 수어 형태로 주황을 표현하지 않았으며, [밤

1]+[노랑1], [노랑1]+[빨강1], [하양1]+[밤1], [밤1]+[빨강1]과 같은 복합어 형태로 주황을 표현하였다. 그 외 [버찌니], [수화1]+[없다1]라는 대답이 나오기도 하였다. 뿐만 아니라 [사과1]과 같은 사물로 대체하고자 하는 형태도 나타났다. 9가지 타입을 언급한 언어 제공자 수는 <표 17>에 요약하였으며, 다양한 형태로 나타내는 예는 <그림 60>, <그림 61>, <그림 62>에 제시하였다.

<표 17> 주황을 나타내는 수어 타입과 언어 제공자 수

| 언어 제공자 | | 20대 (6명) | 30대 (15명) | 40대 (14명) | 50대 (9명) | 60대 (2명) | 70대 (4명) | 계 |
|--------|---|-------------|--------------|--------------|-------------|-------------|-------------|----|
| 타입 | | | | | | | | |
| 주황1 |  | 1 | 2 | 5 | 3 | - | - | 11 |
| 주황2 |  | - | 1 | 1 | - | - | - | 2 |
| 주황3 |  | - | 2 | 2 | - | - | - | 4 |
| 주황4 |  | 5 | 12 | 12 | 1 | - | - | 30 |
| 주황5 |  | 1 | 2 | 2 | 1 | - | - | 6 |
| 주황6 |  | 1 | 3 | - | 2 | - | - | 6 |
| 주황7 |  | - | - | 1 | - | - | - | 1 |
| 주황8 |  | - | - | 1 | - | - | - | 1 |
| 감1 |  | - | 1 | - | 2 | - | - | 4 |

()속은 언어 제공자 수



<그림 60> 주황: [밤1]+[노랑1] / [수화1]+[없다1] - 60대



<그림 61> 주황: [바 1 1] / [노랑1]+[빨강1] / [여자1]+[예쁘다]+[옷을 입다] - 60대



<그림 62> 주황: [하양1]+[밤1] / [밤1]+[빨강1] - 70대

5) 초록

초록은 2가지 형태의 수어로 나타났다. 44명의 언어 제공자가 [초록 1A]로, 3명의 언어 제공자가 [초록2]로 표현하였다. 그 외 60대 이상의 언어 제공자들은 [풀1]이라는 사물로 대체하는 형태와 청, 또는 초록이라고 지문자로 나타내기도 하였다. 2가지 타입을 언급한 언어 제공자 수는 <표 18>에 요약하였다.

<표 18> 초록을 나타내는 수어 타입과 언어 제공자 수

| 언어 제공자 | | 20대 (6명) | 30대 (15명) | 40대 (14명) | 50대 (9명) | 60대 (2명) | 70대 (4명) | 계 |
|--------|--|-------------|--------------|--------------|-------------|-------------|-------------|----|
| 타입 | | | | | | | | |
| 초록1A |  | 6 | 15 | 11 | 9 | 1 | 2 | 44 |
| 초록2 |  | 2 | 1 | - | - | - | - | 3 |

()속은 언어 제공자 수

6) 파랑

파랑은 1가지 형태의 수어로 나타났다. 50명의 언어 제공자 모두가 [파랑1]로 표현하였다. 그 외에 지문자를 사용하거나 주변 사물로 대체하는 형태로 나타나기는 하였지만 단일 형태가 나타났다. 1가지 타입을 언급한 언어 제공자 수는 <표 19>에 요약하였다.

<표 19> 파랑을 나타내는 수어 타입과 언어 제공자 수


| 언어 제공자 | | 20대 (6명) | 30대 (15명) | 40대 (14명) | 50대 (9명) | 60대 (2명) | 70대 (4명) | 계 |
|--------|---|-------------|--------------|--------------|-------------|-------------|-------------|----|
| 타입 | | | | | | | | |
| 파랑1 |  | 6 | 15 | 14 | 9 | 2 | 4 | 50 |

()속은 언어 제공자 수

7) 회색

회색은 1가지 형태의 수어로 나타났는데, 50명의 언어 제공자 중 47명의 언어 제공자가 [쥐1]로 표현하였으며, 지문자로 표현하기도 하였다. 1가지 타입을 언급한 언어 제공자 수는 <표 20>에 요약하였다.

<표 20> 회색을 나타내는 수어 타입과 언어 제공자 수


| 언어 제공자 | | 20대 (6명) | 30대 (15명) | 40대 (14명) | 50대 (9명) | 60대 (2명) | 70대 (4명) | 계 |
|--------|---|-------------|--------------|--------------|-------------|-------------|-------------|----|
| 타입 | | | | | | | | |
| 쥐1 |  | 6 | 14 | 14 | 8 | 2 | 3 | 47 |

()속은 언어 제공자 수

8) 유도 과제 이외 색

유도 과제에서 제시하지 않은 색 중에 [하양1], [빨강1], [노랑1]이라는 수어가 나타났다. 이는 분홍, 주황, 갈색, 보라와 같은 색을 표현하기 위해 나타났다. 또한 [하양1]은 ‘미원’을, [노랑1]은 콩나물을 유도하는 과정에서도 나타나서 발생 빈도가 높았다. 예를 들면 분홍은 [빨강1]+[하양1]+[섞다1A] (언어 제공자 1명)로, 주황은 [빨강1]+[노랑1]+[모으다1] 또는 [빨강1]+[노랑1]로, 갈색은 [빨강1]+[노랑1]+[섞다1A]로 표현하였다. 하양, 빨강, 노랑을 나타내는 수어 타입과 언어 제공자 수는 <표 21>, <표 22> <표 23>에 요약하였다.

<표 21> 하양을 나타내는 수어 타입과 언어 제공자 수

| 언어 제공자 | | 20대 (6명) | 30대 (15명) | 40대 (14명) | 50대 (9명) | 60대 (2명) | 70대 (4명) | 계 |
|--------|---|-------------|--------------|--------------|-------------|-------------|-------------|----|
| 타입 | | | | | | | | |
| 하양1 |  | - | 7 | 4 | 6 | 2 | 4 | 23 |


()속은 언어 제공자 수

<표 22> 빨강을 나타내는 수어 타입과 언어 제공자 수

| 언어 제공자 | | 20대 (6명) | 30대 (15명) | 40대 (14명) | 50대 (9명) | 60대 (2명) | 70대 (4명) | 계 |
|--------|---|-------------|--------------|--------------|-------------|-------------|-------------|----|
| 타입 | | | | | | | | |
| 빨강1 |  | 4 | 8 | 5 | 7 | 2 | 4 | 30 |
| 빨강1A |  | - | - | 1 | - | - | - | 1 |

()속은 언어 제공자 수

<표 23> 노랑을 나타내는 수어 타입과 언어 제공자 수

| 언어 제공자 | | 20대 (6명) | 30대 (15명) | 40대 (14명) | 50대 (9명) | 60대 (2명) | 70대 (4명) | 계 |
|--------|---|-------------|--------------|--------------|-------------|-------------|-------------|----|
| 타입 | | | | | | | | |
| 노랑1 |  | - | 6 | 4 | 7 | 2 | 4 | 23 |

()속은 언어 제공자 수

다음은 색 이름을 2개의 수어로 표현하는 언어 제공자 수를 알아보았다. 주로 50대 이상에서 2개의 수어를 사용하여 색을 표현하는 경향이 있었다. 2개의 수어로 표현하는 언어 제공자 수를 <표 24>에 요약하였다.

<표 24> 2개 수어로 색을 표현한 언어 제공자 수

| 언어 제공자 | 20대 (6명) | 30대 (15명) | 40대 (14명) | 50대 (9명) | 60대 (2명) | 70대 (4명) | 계 |
|--------|-------------|--------------|--------------|-------------|-------------|-------------|----|
| 형태 | | | | | | | |
| 빨강+하양 | - | - | - | 1 | - | 1 | 2 |
| 빨강+약하다 | 1 | 3 | 2 | 4 | 1 | 4 | 15 |
| 빨강+노랑 | - | 1 | 3 | 3 | 1 | 1 | 9 |
| 빨강+파랑 | - | - | - | 1 | - | 2 | 3 |
| 노랑+밤 | - | - | - | 2 | 1 | 1 | 4 |
| 갈색+약하다 | - | - | - | 1 | - | - | 1 |
| 밤+약하다 | - | - | - | - | - | 1 | 1 |
| 밤+노랑 | - | - | - | 2 | - | 1 | 3 |

5.2 한국수어에 나타난 지명

5.2.1 분석 대상 및 방법

한국수어에서 나타난 지명의 다양한 형태를 알아보기 위하여 2015년과 2017년에 전사한 모든 파일에서 지명에 해당하는 타입을 추출하였다. 추출 결과 변이가 나타나는 지명은 모두 ‘옛날 수어와 현재 수어(과제 번호 L)’에서 나타났다. 따라서 지명의 변이 형태를 분석하기 위해 총 34분 37초 분량의 ‘옛날 수어와 현재 수어’ 과제 영상만을 대상으로 분석하였다. 유도 과제의 특성상 해당 타입의 발생 빈도는 중요하지 않아 언어 제공자 수로 분석하였으며, 분석한 결과를 제한적

으로 일반화하여야 한다. 또한 본 사례 연구는 지금 수어의 형태가 옛날 수어와 형태가 다른 것에 대해 이야기해 달라는 과제에서 추출한 데이터이므로 현재 사용되고 있는 수어 형태가 아닐 수 있다. 분석한 총 파일은 16개로 언어 제공자 수는 32명이고, 우세손 기준 발생 타입 수는 1,029개, 토큰 수는 8,381개였다.

5.2.2 분석 결과

변이가 나타난 지명은 강원도, 남대문, 동대문, 부산, 용산, 익산(이리), 인천, 울산, 속초, 원주, 창원, 포천, 통영, 화성이 있었고 국가명으로는 일본이 있었다. 그러나 본 연구에서는 1명의 언어 제공자만이 언급한 속초(2), 원주(3), 창원(2), 포천(2), 통영(2), 화성(2)은 분석에서 제외하고(괄호 속은 나타난 타입의 개수임) 8곳의 지명과 1곳의 국가명, 22개의 타입만을 분석하였다. 각 지명에 대한 수어 형태는 다음과 같다.



강원도는 4개의 타입이 나타났는데, 18명의 언어 제공자가 [강원도 2]를 언급하였으며, 15명의 언어 제공자가 [강원도1]을 언급하였다. [강원도3]은 20대 언어 제공자를 제외하고 1명씩 언급하였으며, [강원도4]는 30대 언어 제공자 1명만 언급하였다. 4가지 수어 형태를 언급한 언어 제공자 수는 <표 25>에 요약하였다.

<표 25> 강원도

| 언어 제공자 | | 20대 | 30대 | 40대 | 50대 | 계 |
|--------|---|--------|---------|--------|--------|--------------|
| 타입 | | | | | | |
| 강원도1 |  | 2(13%) | 5(33%) | 6(40%) | 2(13%) | 15 (100%) |
| 강원도2 |  | 2(11%) | 6(33%) | 8(44%) | 2(11%) | 18 (100%) |
| 강원도3 |  | - | 1(33%) | 1(33%) | 1(33%) | 3 (100%) |
| 강원도4 |  | - | 1(100%) | - | - | 1 (100%) |


남대문은 2개의 타입이 나타났는데, 8명의 언어 제공자가 [남대문1]을 언급하였으며, 6명의 언어 제공자가 [남대문2]를 언급하였다. 2가지 수어 형태를 언급한 언어 제공자 수는 <표 26>에 요약하였다.

<표 26> 남대문

| 언어 제공자 | | 20대 | 30대 | 40대 | 70대 | 계 |
|--------|---|--------|--------|--------|--------|-------------|
| 타입 | | | | | | |
| 남대문1 |  | 1(13%) | 2(25%) | 3(38%) | 2(25%) | 8 (100%) |
| 남대문2 |  | 1(17%) | 1(17%) | 2(33%) | 2(33%) | 6 (100%) |


동대문은 2개의 타입이 나타났는데, 6명의 언어 제공자가 [동대문1]을 언급하였으며, 2명의 언어 제공자만이 [동대문2]를 언급하였다. 2가지 수어 형태를 언급한 언어 제공자 수는 <표 27>에 요약하였다.

<표 27> 동대문

| 언어 제공자 | | 30대 | 40대 | 70대 | 계 |
|--------|---|--------|--------|--------|-------------|
| 타입 | | | | | |
| 동대문1 |  | 1(17%) | 3(50%) | 2(33%) | 6 (100%) |
| 동대문2 |  | 1(50%) | 1(50%) | - | 2 (100%) |

부산은 2개의 타입이 나타났는데, 5명의 언어 제공자가 [부산1]을 언급하였으며, 1명의 60대 언어 제공자만이 [부산2]를 언급하였다. 2가지 수어 형태를 언급한 언어 제공자 수는 <표 28>에 요약하였다.

<표 28> 부산

| 언어 제공자 | | 20대 | 40대 | 60대 | 계 |
|--------|---|--------|--------|---------|-------------|
| 타입 | | | | | |
| 부산1 |  | 1(20%) | 3(60%) | 1(20%) | 5 (100%) |
| 부산2 |  | - | - | 1(100%) | 1 (100%) |




용산은 3개의 타입이 나타났는데, 3명의 언어 제공자가 [용산3]을 언급하였으며, 2명의 언어 제공자가 [용산2]를 언급하였다. [용산1]은 20대 언어 제공자 2명만이 언급하였다. 3가지 수어 형태를 언급한 언어 제공자 수는 <표 29>에 요약하였다.

<표 29> 용산

| 언어 제공자 | | 20대 | 40대 | 70대 | 계 |
|--------|---|---------|--------|--------|-------------|
| 타입 | | | | | |
| 용산1 |  | 2(100%) | - | - | 2 (100%) |
| 용산2 |  | 1(50%) | 1(50%) | - | 2 (100%) |
| 용산3 |  | 1(33%) | 1(33%) | 1(33%) | 3 (100%) |



이리의 경우는 지명이 이리에서 익산으로 바뀌었기 때문에 이리와 익산으로 글로스를 한 토큰으로 함께 검색하였다. 이리/익산은 3개의 타입이 나타났는데, 4명의 언어 제공자가 [이리1]을 언급하였으며, 4명의 언어 제공자가 [이리2]를 언급하였다. 또한 [익산1]은 3명의 언어 제공자가 언급하였다. 3가지 수어 형태를 언급한 언어 제공자 수는 <표 30>에 요약하였다.

<표 30> 이리 / 익산

| 언어 제공자 | | 30대 | 40대 | 계 |
|--------|---|--------|--------|-------------|
| 타입 | | | | |
| 이리1 |  | 1(25%) | 3(75%) | 4 (100%) |
| 이리2 |  | 1(25%) | 3(75%) | 4 (100%) |
| 익산1 |  | 2(67%) | 1(33%) | 3 (100%) |



인천은 2개의 타입이 나타났는데, 9명의 언어 제공자가 [인천1]을 언급하였으며, 11명의 언어 제공자가 [인천2]를 언급하였다. 2가지 수어 형태를 언급한 언어 제공자 수는 <표 31>에 요약하였다.

<표 31> 인천

| 언어 제공자 | | 30대 | 40대 | 60대 | 계 |
|--------|---|--------|--------|--------|--------------|
| 타입 | | | | | |
| 인천1 |  | 3(33%) | 4(44%) | 2(22%) | 9 (100%) |
| 인천2 |  | 4(36%) | 5(45%) | 2(18%) | 11 (100%) |


울산은 2개의 타입이 나타났는데, 8명의 언어 제공자가 [울산1]을 언급하였으며, 8명의 언어 제공자가 [울산2]를 언급하였다. 2가지 수어 형태를 언급한 언어 제공자 수는 <표 32>에 요약하였다.

<표 32> 울산

| 언어 제공자 | | 20대 | 30대 | 40대 | 60대 | 계 |
|--------|---|--------|--------|--------|--------|-------------|
| 타입 | | | | | | |
| 울산1 |  | 2(25%) | 3(38%) | 3(38%) | - | 8 (100%) |
| 울산2 |  | 1(13%) | 3(38%) | 3(38%) | 1(13%) | 8 (100%) |

일본은 2개의 타입이 나타났는데, 4명의 언어 제공자가 [일본1]을 언급하였으며, 5명의 언어 제공자가 [일본2]를 언급하였다. 2가지 수어 형태를 언급한 언어 제공자 수는 <표 33>에 요약하였다.

<표 33> 일본

| 언어 제공자 | | 30대 | 40대 | 50대 | 60대 | 계 |
|--------|---|--------|--------|--------|--------|-------------|
| 타입 | | | | | | |
| 일본1 |  | 1(25%) | 2(50%) | - | 1(25%) | 4 (100%) |
| 일본2 |  | 3(60%) | - | 1(20%) | 1(20%) | 5 (100%) |

5.3 한국수어에 나타난 이형 동의어 분석

5.3.1 분석 대상 및 방법

한국수어에 나타난 이형 동의어를 알아보기 위해 2017년 전사 완료한 총 260분 분량의 영상과 2015년 전사 완료한 영상 중 개별 수어의 ‘다양한 표현 과제(과제 번호 I)’와 ‘옛날 수어와 현재 수어 과제(과제 번호 L)’을 제외한 146분 분량의 영상을 대상(총 406분)으로 분석하였다.

이형 동의어 분석을 위해 수어의 글로스를 한국어 어휘를 차용하여 의미 바탕으로 정하였기 때문에 전사 자료 중 먼저 번호만 다른 같은 타입명을 검색하여 추출하였다. 다음으로 추출한 수어 어휘에서 수어 모어 화자인 공동 연구원의 언어적 직관을 바탕으로 완전한 동의어로 보기 어려우며 맥락이나 연어 관계를 통해 유의어의 수준을 더 깊이 분석해야 할 필요가 있는 수어 어휘는 제외하였다. 그리하여 구체적인 대상을 나타내거나 문장 안에서 완전히 대체가 가능하다고 모어 화자가 인식한 수어만을 분석 대상으로 추출하였다. 그런 다음 다시 추출된 어휘 중 언어 제공자 수가 5명 미만으로 적은 수어를 제외하였다. 이에 따라 분석한 어휘는 샤워와 관련된 어휘 등 8개와 관련된 이형

동의어이다.





5.3.2 분석 결과

넥타이, 샤워, 수영, 신호등, 아래, 아침, 오토바이를 나타내는 수어의 형태를 분석한 결과, 넥타이는 4개의 형태가 나타났으며, 샤워, 아침, 오토바이는 3개의 형태가, 수영, 신호등, 아래는 2개의 형태가 나타났다. 각 타입을 사용한 언어 제공자 수와 토큰 수를 다음 <표 34>에서 <표 40>로 요약하였다.

넥타이와 관련된 수어 229개의 토큰 중 [넥타이1]은 15명의 언어 제공자에 의해 83개의 토큰이 발생하였는데, 주로 30대와 40대 언어 제공자에게서 나타났으며 60대 이상의 언어 제공자에게서는 나타나지 않았다. [넥타이2]는 11명의 언어 제공자에 의해 31개의 토큰이 발생하였는데, 주로 40대와 50대 언어 제공자에게서 나타났으며 70대 언어 제공자 중에서도 2명이 사용하였다. [넥타이3]은 6명의 언어 제공자에 의해 12개의 토큰이 발생하였는데, 주로 30대에서 나타났다. 또한 [넥타이4]는 23명의 언어 제공자에 의해 103개의 토큰으로 가장 많이 발생하였는데, 30대 언어 제공자에게서 가장 많이 나타났으며 모든 연령대에서 골고루 사용하는 것으로 나타났다.

또한 20대, 30대, 40대 언어 제공자들은 주로 [넥타이1]과 [넥타이4]를 사용하였으나 50대 언어 제공자는 [넥타이2]와 [넥타이4]를 많이 사용하였다. 특히 60대에서는 [넥타이1]과 [넥타이2]가 나타나지 않았으며, 70대 언어 제공자에게서는 [넥타이1]과 [넥타이3]이 나타나지 않았다. 4가지 타입에 대한 언어 제공자수와 발생 토큰 수는 <표 34>에 요약하였다.


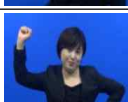

<표 34> 넥타이

| 언어 제공자 연령 타입 | | 20대 | 30대 | 40대 | 50대 | 60대 | 70대 | 계 |
|-----------------|---|--------------|--------------|--------------|--------------|-------------|-------------|-----------------|
| 넥타이1 |  | 2(13%) 17 | 6(40%) 32 | 5(33%) 24 | 2(13%) 10 | - | - | 15(100%) 83 |
| 넥타이2 |  | 1(9%) 1 | 1(9%) 4 | 3(27%) 9 | 4(36%) 12 | - | 2(18%) 5 | 11(100%) 31 |
| 넥타이3 |  | 1(17%) 3 | 2(33%) 6 | 1(17%) 1 | 1(17%) 1 | 1(17%) 1 | - | 6(100%) 12 |
| 넥타이4 |  | 3(13%) 16 | 7(30%) 37 | 5(22%) 11 | 5(22%) 23 | 2(9%) 15 | 1(14%) 1 | 23(100%) 103 |

이탤릭체의 숫자는 발생 토큰 빈도임

샤워와 관련된 수어는 총 408개의 토큰이 발생하였는데, 그 중에서 [샤워1]은 31명의 언어 제공자가 사용하였으며 발생 토큰은 116개이며, 주로 50대에서 많이 나타났다. [샤워2]는 28명의 언어 제공자에 의해 109개의 토큰이 나타났으며, 30대에서 가장 많이 나타났다. [샤워3]은 43명의 언어 제공자에 의해 183개의 토큰이 발생하여 가장 많이 나타났으며 30대, 40대 50대에서 골고루 나타났다. 또한 연령대별로 보았을 때, 20대 언어 제공자는 [샤워2]와 [샤워3]을 주로 사용하였으며, 30대와 40대 언어 제공자는 3가지 타입을 골고루 사용하였다. 50대 이상 언어 제공자는 [샤워1]과 [샤워3]을 주로 사용하는 것으로 나타났다. 3가지 타입에 대한 언어 제공자 수와 발생 토큰 수는 <표 35>에 요약하였다.



<표 35> 샤워

| 언어 제공자 연령 타입 | | 20대 | 30대 | 40대 | 50대 | 60대 | 70대 | 계 |
|-----------------|---|--------------|---------------|---------------|---------------|-------------|-------------|-----------------|
| 샤워1 |  | 2(6%) 4 | 7(22%) 26 | 6(19%) 19 | 11(35%) 35 | 3(9%) 9 | 2(6%) 6 | 31(100%) 116 |
| 샤워2 |  | 5(17%) 18 | 11(39%) 48 | 6(21%) 25 | 5(17%) 17 | 1(3%) 3 | - | 28(100%) 109 |
| 샤워3 |  | 4(9%) 7 | 12(28%) 34 | 11(26%) 54 | 10(23%) 48 | 4(9%) 23 | 2(4%) 17 | 43(100%) 183 |

이탤릭체의 숫자는 발생 토큰 빈도임

수영과 관련된 수어는 9명의 언어 제공자에 의해 총 13개의 토큰이 발생하였는데, 그 중에서 [수영하다1]은 4명의 언어 제공자에 의해 4개의 토큰이 나타났으며, [수영하다2]는 5명의 언어 제공자에 의해 9개의 토큰이 나타났다. 연령에 따른 차이는 거의 보이지 않았다. 2가지 타입에 대한 언어 제공자 수와 발생 토큰 수는 <표 36>에 요약하였다.

<표 36> 수영하다

| 언어 제공자 연령 타입 | | 30대 | 40대 | 50대 | 60대 | 70대 | 계 |
|-----------------|---|--------------------|--------------------|--------------------|--------------------|--------------------|---------------------|
| 수영하다1 |  | 1(25%) <i>1</i> | - - | 1(25%) <i>1</i> | 1(25%) <i>1</i> | 1(25%) <i>1</i> | 4(100%) <i>4</i> |
| 수영하다2 |  | 2(40%) <i>5</i> | 1(20%) <i>2</i> | 1(20%) <i>1</i> | 1(20%) <i>1</i> | - - | 5(100%) <i>9</i> |

이탤릭체의 숫자는 발생 토큰 빈도임

신호등과 관련된 수어는 10명의 언어 제공자에 의해 총 22개의 토큰이 발생하였다. [신호등1]은 5명의 언어 제공자에 의해 8개의 토큰이 나타났고, 40대 언어 사용자가 많이 사용하였다. [신호등3]은 5명의 언어 제공자에 의해 13개의 토큰이 나타났으며, 연령에 따른 차이는 거의 보이지 않았다. 두 가지 타입에 대한 언어 제공자 수와 발생 토큰 수는 <표 37>에 요약하였다.

표 37> 신호등

| 언어 제공자 연령 타입 | | 30대 | 40대 | 50대 | 60대 | 계 |
|-----------------|---|--------------------|--------------------|--------------------|--------------------|----------------------|
| 신호등1 |  | 1(20%) <i>1</i> | 3(60%) <i>4</i> | 1(20%) <i>3</i> | - - | 5(100%) <i>8</i> |
| 신호등3 |  | 1(20%) <i>4</i> | 1(20%) <i>1</i> | 2(40%) <i>5</i> | 1(20%) <i>3</i> | 5(100%) <i>13</i> |

이탤릭체의 숫자는 발생 토큰 빈도임

아래와 관련된 수어는 17명의 언어 제공자에 의해 총 36개의 토큰이 발생하였다. [아래1]은 10명의 언어 제공자에 의해 15개의 토큰이 나타났으며, 50대 언어 사용자가 가장 많이 사용하였다. [아래2]는 7

명의 언어 제공자에 의해 21개의 토큰이 나타났는데, 연령에 따른 차이는 거의 보이지 않았다. 2가지 타입에 대한 언어 제공자 수와 발생 토큰 수는 <표 38>에 요약하였다.

<표 38> 아래

| 언어 제공자 연령 타입 | | 20대 | 30대 | 40대 | 50대 | 계 |
|-----------------|---|--------------------|--------------------|---------------------|--------------------|-----------------------|
| 아래1 |  | 1(10%) <i>2</i> | 3(30%) <i>5</i> | 1(10%) <i>1</i> | 5(50%) <i>7</i> | 10(100%) <i>15</i> |
| 아래2 |  | 2(28%) <i>5</i> | 2(28%) <i>2</i> | 2(28%) <i>10</i> | 1(14%) <i>3</i> | 7(100%) <i>21</i> |

이탤릭체의 숫자는 발생 토큰 빈도임

아침과 관련된 수어는 18명의 언어 제공자에 의해 총 36개의 토큰이 발생하였다. [아침1]은 13명의 언어 제공자에 의해 31개의 토큰이 나타났는데, 50대 언어 사용자가 가장 많이 사용하였다. [아침2]는 3명의 언어 제공자에 의해 3개의 토큰이 나타났는데, 20대와 30대에서는 나타나지 않았으며 40대 이후의 언어 제공자에 의해 1번씩 발생하였다. [아침3]은 40대와 50대 언어 제공자 1명씩에 의해 2개의 토큰이 나타났다. 즉, [아침1]은 20대에서부터 60대까지 모든 연령대의 언어 제공자들에게서 나타나는데 비해, [아침2]와 [아침3]은 20대와 30대 언어 제공자들에게서는 나타나지 않았다. 3가지 타입에 대한 언어 제공자 수와 발생 토큰 수는 <표 39>에 요약하였다.




<표 39> 아침

| 언어 제공자 연령 타입 | | 20대 | 30대 | 40대 | 50대 | 60대 | 계 |
|-----------------|---|--------------------|---------------------|--------------------|--------------------|--------------------|-----------------------|
| 아침1 |  | 2(15%) <i>4</i> | 5(38%) <i>13</i> | 2(15%) <i>6</i> | 3(23%) <i>7</i> | 1(7%) <i>1</i> | 13(100%) <i>31</i> |
| 아침2 |  | - - | - - | 1(33%) <i>1</i> | 1(33%) <i>1</i> | 1(33%) <i>1</i> | 3(100%) <i>3</i> |
| 아침3 |  | - - | - - | 1(50%) <i>1</i> | 1(50%) <i>1</i> | - - | 2(100%) <i>2</i> |

이탤릭체의 숫자는 발생 토큰 빈도임

오토바이와 관련된 수어는 50명의 언어 제공자에 의해 총 176개의 토큰이 발생하였다. [오토바이1]은 33명의 언어 제공자에 의해 139개의 토큰이 나타났는데, 20대에서 70대까지 모든 연령대의 언어 사용자에게서 나타났다. [오토바이2]는 15명의 언어 제공자에 의해 34개의 토큰이 나타났는데, 60대와 70대에서는 나타나지 않았다. [오토바이4]는 30대와 50대 언어 제공자 1명씩에 의해 3개의 토큰이 나타났다. 즉, [오토바이1]은 20대에서부터 70대까지 연령대의 언어 제공자들에게서 나타나는데 비해, [오토바이2]와 [오토바이4]는 60대와 70대 언어 제공자들에게서는 나타나지 않았다. 3가지 타입에 대한 언어 제공자 수와 발생 토큰 수는 <표 40>에 요약하였다.

<표 40> 오토바이

| 언어 제공자 연령 타입 | | 20대 | 30대 | 40대 | 50대 | 60대 | 70대 | 계 |
|-----------------|---|--------------------|----------------------|---------------------|---------------------|---------------------|-------------------|------------------------|
| 오토 바이1 |  | 4(12%) <i>8</i> | 13(39%) <i>60</i> | 8(24%) <i>37</i> | 3(9%) <i>15</i> | 4(12%) <i>13</i> | 1(3%) <i>6</i> | 33(100%) <i>139</i> |
| 오토 바이2 |  | 3(20%) <i>4</i> | 3(20%) <i>7</i> | 4(26%) <i>10</i> | 5(33%) <i>13</i> | - | - | 15(100%) <i>34</i> |
| 오토 바이4 |  | - | 1(33%) <i>1</i> | - | 1(66%) <i>2</i> | - | - | 2(100%) <i>3</i> |

이탤릭체의 숫자는 발생 토큰 빈도임

V. 한국수어 말뭉치 구축 계획

1. 한국수어 말뭉치 구축을 위한 자문

향후 한국수어 말뭉치 구축 계획 수립을 위해 농사회 전문가 자문 회의 내용과 음성 언어 말뭉치 전문가들의 서면을 통한 자문 의견을 요약 정리한 결과이다.

1.1 자문 방법

전문가 집단 자문은 2회에 걸쳐 실시하였다. 1회에는 농사회 전문가 5명에게 관련 질문을 통하여 의견을 수집하였고, 2회에는 음성 언어 말뭉치 구축 전문가 5명에게 전자 메일을 통해 자문 의견을 수집하였다.

1.2 자문을 위한 질문

농사회 전문가 자문 회의와 음성 언어 말뭉치 전문가 서면 자문을 받기 위해 두 집단 공통적인 질문지와 집단에 맞춘 상이한 질문지로 구성하였다.

언어 자료 수집과 전사 방법, 한국수어 말뭉치 공개 시기와 방법에 대한 질문은 두 전문가 집단 모두에게 동일하였다. 농사회 전문가 집단에게는 한국수어 말뭉치 구축에서 농사회와의 연계 및 협력 방안과 홍보 방법에 대한 자문을 구하였고, 음성 언어 말뭉치 전문가 집단에게는 한국수어 말뭉치 시스템 개발과 향후 한국수어 말뭉치 기반 연구 방향에 대한 자문을 의뢰하였다.

<표 41> 두 자문 집단에게 의뢰한 자문 항목

| 농사회 전문가 집단 | 음성 언어 말뭉치 전문가 집단 |
|---------------------------------------|-----------------------------|
| 언어 자료 수집 방법 | |
| 언어 자료 전사 방법 | |
| 한국수어 말뭉치 공개 방법 | |
| 한국수어 말뭉치에 대한 홍보 방법 | 한국수어 말뭉치 시스템 개발 |
| 한국수어 말뭉치 구축 과정에서의 농사회와의 연계 및 협력 방안 | 한국수어 말뭉치의 활용도를 높이기 위한 방안 |

1.3 자문 의견 결과

농사회 전문가의 의견을 먼저 제시한 후 이어서 음성 언어 말뭉치 전문가의 의견을 제시하였다.

1.3.1 언어 자료 수집 방법

언어 자료 수집 방법, 언어 자료 수집을 위한 지역 선정 기준, 언어 자료 수집을 위한 횟수, 언어 자료 수집 주체로 세분하여 자문 의견을 요청하였다.

1) 언어 자료의 수집 방법

농사회 전문가 집단에서는 언어 자료 수집 방법에 대한 특별한 의견을 제시하지 않았으나 음성 언어 말뭉치 전문가 집단에서는 자연스러운 상황 속에서의 수집의 필요성을 제안하였다. 2015년 구축된 한국수어 말뭉치는 13개의 과제로 구성된 유도 자료를 통해 수집되었다. 유도 자료를 통한 자료 수집이 여러 가지 상황이나 영역, 사용역을 가정한 발화 자료를 얻고자 한 점을 높이 평가하였으나 이에 추가하여 자연스러운 상황과 대화 상황의 언어 자료 수집을 제안하였다.

- 자연스러운 환경 속에서의 담화, 대화 등 준비되지 않고 유도되지 않은 수어 자료는 수어의 5가지 수어소 외에도 수어 동작의 크기 등 다른 차원의 요소가 덧붙여진 전사가 가능할 것이다.
- 대화에서는 상대방의 말에 호응하거나 반응하는 다양한 청자 반응 (backchannels)이 나타나며 청자 반응 내에도 여러 가지 형태의 어휘 항목들이 있다. 손이나 몸짓으로 나타나기도 하는 청자 반응이 수어와 구분되면서도 어떤 방식으로 실현되는가는 흥미로운 부분이기에 좀 더 많은 대화 과제가 필요해 보인다. 또한 토론과 같은 공적 말하기 외에 일상 대화를 관찰할 수 있는 과제도 추가되면 좋겠다. 특히 일상 대화에서는 축약 등 언어적 변이 현상이 나타나기 쉬운데 몸짓이나 손짓 등 표현 수단이 한정된 수어에서 이 현상이 어떻게 실현되는가는 상당히 중요한 지점일 수 있을 것 같다.

2) 언어 자료 수집을 위한 지역 선정 기준

농사회 전문가 집단은 언어 자료 수집을 위한 지역 선정 기준에 대해 일치된 의견을 제시하였다. 전국 17개 지역에 소재한 한국농아인 협회를 중심으로 17개 시·도 지역을 6개의 도별로 묶어서 언어 자료를 수집할 것을 제안하였다. 6개의 지역 구분은 강원도, 경기도, 충청도, 전라도, 경상도, 제주도였다.

음성 언어 말뭉치 전문가 집단은 한국수어 말뭉치 구축 목적과 사회학적 요인 변수를 고려하여 언어 자료 수집 지역을 선정해야 한다고 제안하였다.

- 한국수어 말뭉치가 "표준"수어의 쓰임을 반영해야 한다면, 전국 규모로 지역/성별/연령별 수어 사용자에 대한 통계 자료를 확보하여 그 비율만큼의 언어 제공자를 선정하는 것이 타당할 것이며, 만약 특정 지역의 수어 연구를 위한 것이라면, 언어 제공자 선정 시 해당 지역성을 감안하는 것이 타당할 것이다.
- 여러 연령, 성별, 계층 등 사회적 변수에 따른 농인 인구 분포를 일

차적으로 고려할 뿐만 아니라 행정 구역과 농아인협회 지역 등이 함께 고려되어야 할 것이다. 특히 전국 농학교가 있는 지역은 교육 현장이나 교육 수준 등 언어 사용과 학습, 사회 활동의 또 다른 거점이 될 수 있으므로 이 기준 역시 중요하게 고려되어야 할 것이다.

위 의견과 달리 지역별, 대상별로 언어 수집을 확대하지 말고 표준이 되는 수어들을 정비하는 것의 필요성을 제안한 의견도 있었다.

- 수어에 표준화된 형태가 없고 지역별 개인별 변이형이 존재한다면, 그것을 일일이 조사하여 정리하는 것보다는 표준이 되는 수어의 표지들을 정비하여 농인 간의 소통에 도움을 주는 쪽으로 목표를 한정하는 게 좋을 것 같다. 따라서 지역별, 대상별 확대는 바람직하지 않다.

3) 언어 자료 수집을 위한 지역 선정 순서

언어 자료 수집을 위한 지역 선정 순서에 대한 농사회 전문가 집단의 의견은 다양하였다. 언어 자료를 수집할 수 있는 지역, 농학교가 설립된 지역, 서울부터 시작하여 농인 인구가 많은 지역, 대도시로부터 시작하여 중소 도시 그리고 제주도 지역까지 언어 자료를 수집할 것을 제안하였다.

4) 언어 자료 수집을 위한 지역의 수

농사회 전문가 집단은 일정 기간 안에 언어 자료를 수집하는 지역의 수에 대해 구체적인 횟수를 제안하였다. 언어 자료를 충실하게 그리고 집중적으로 수집하기 위해 1년에 1~2지역(1명), 2~3지역(3명)이 적당하다는 의견이 있었으며 인력이 감당할 수 있는지 여부에 따라 조절할 수 있다고 하였다. 음성 언어 말뭉치 전문가들의 의견에서도 1~2개 지역 혹은 2~3개 지역을 선정하여 시범적으로 언어 자료를 수집할 것을 제안하였다. 그러나 언어 자료 수집을 위한 지역의 수를 임의적으로 결정하지 말고 더 큰 틀에서 한국수어 말뭉치의 구축 목적과 범위를 정해 연차별로 필요 예산과 인원에 따라 결정하여야 한다는 의견이

있었다.

- 1년 동안 언어 자료 수집 지역은 여러 가지 사회학적 변수, 요인을 고려해 가장 다양한 언어 사용과 많은 사용 인구를 가진 지역 등 보편적 언어 사용 양상을 파악할 수 있는 지역 설정이 필요해 보인다. 지나치게 세분화하거나 한 곳에 편중하기보다 보편성과 대표성을 띠는 2~3개 지역 정도를 시범적으로 선정해 자료 수집을 할 필요가 있을 것이다.
- 수집 지역의 수는 투입 예산과 작업 가능한 인력의 수에 따라 결정되므로, 총 사업 기간, 연차별 투입 예산 규모, 수집 참여 가능 인력 등에 대한 자료 제시가 먼저 필요하다. 만약 전국에서 사용되고 있는 수어에 대한 조사를 목적으로 한다면, 전 지역을 수어 채집 지역으로 정의하고 연차별 예산을 고려하여 채집 기간 및 지역의 수를 결정하는 것이 바람직할 것이다.

5) 언어 자료 수집 주체

언어 자료 수집의 주체에 대한 질문에서는 두 집단에서 다소 이견을 보여주었다. 농사회 전문가 집단은 한국농아인협회를 주체로 하여 언어 자료를 수집하면서 연구 기관인 대학과의 협조의 필요성을 제안하였다.

- 언어 자료의 수집은 대상자 선정 등의 과정에서 농사회와의 협력이 중요하다. 농아인협회를 중심으로 하는 것이 바람직하다.

음성 언어 말뭉치 전문가 집단은 국립국어원이 주체가 되어야 한다는 의견이 더 많았다.

- 언어 자료의 수집 주체는 국립국어원이 주체가 되어야 할 것이다. 전문적 지식이나 이론, 관점 등은 대학의 연구 기반을 활용해야 하고, 이 과정에서 자문 등 여러 분야에서 도움을 얻거나 때로는 실질적 작업을 함께 수행할 수도 있을 것이다. 농아인협회나 시·도 협회

등은 농아인들의 실제 언어생활, 사회생활을 고려한 자문, 현황 파악과 사업 추진을 위한 진행, 언어 사용 문제에 대한 다양한 조언 등을 해 줄 수 있어 또 한 축이 된다. 그러나 이 여러 기관의 협조를 구하면서 여러 업무를 조율해 언어 자료를 수집하는 사업을 지속할 수 있는 힘, 경제적 지원과 계획 등은 국립국어원이 주체가 되어 할 수밖에 없는 일이다.

- 수어 자료 수집의 주체는 한국농아인협회와 같이 해당 분야의 전문성을 갖춘 단체가 되는 것이 바람직할 것이다. 반면 수어도 언어의 한 갈래이므로, 한국수어 말뭉치의 구축과 보존 그리고 활용 확산을 위한 정부 정책 수립 및 예산 확보의 역할은 현재와 같이 국립국어원에서 전담을 하되 그 세부 계획은 전문 단체와의 협의를 통해 결정하는 것이 바람직하다.

1.3.2 언어 자료 전사 방법

전사의 질과 효율을 높이기 위한 질문이다. 농사회 전문가 집단은 전사는 전사의 전문성을 위해 용역 방식보다는 국립국어원에서 전사 전문 인력을 채용하는 방식이 좋다는 의견(3명)과 두 방식 모두 무방하다는 의견(2명)이 있었다.

- 전사는 한국수어 말뭉치 구축 사업에 가장 중요한 작업이므로 전문적인 지식을 갖추고 훈련을 받은 사람 즉 전문가가 지속적으로 해야 한다고 생각한다. 그러므로 용역 방식과 국어원에서 전사자 채용 방식 중 어느 쪽이 더 전문적인가에 따라 결정하면 좋겠다.
- 전사는 유경험자가 지속적으로 맡는 것이 바람직하므로 국립국어원에서 전사자를 정식 직원으로 채용함으로 안정적으로 장기간 근무하면서 전사 업무를 수행하도록 하는 것이 바람직하다.

음성 언어 말뭉치 전문가 집단은 전사 작업과 말뭉치의 완성도와 충실성, 수준 등은 전사자의 작업에 좌우하므로 국립국어원에서 전사자를 채용하여 안정적인 전사 전문가를 확보할 뿐만 아니라 체계적인 관

리가 필요하다는 의견을 제시하였다.

- 전사자는 국립국어원에서 일괄 고용하여, 체계적인 관리를 할 필요가 있다. 자칫하면 애써 구축한 한국수어 말뭉치 영상이 참고 할 동영상 비디오에 그칠 가능성이 높다.
- 기존과 같은 용역 방식은 사업의 기획 방향이나 작업의 내용을 잘 이해한 전사자를 직접 구할 수 있다는 점에서는 초기 단계에서는 불가피하고 또 효율적인 측면도 있다. 그러나 사업의 지속성과 일관성 등을 생각하면 국립국어원에서 이들을 지속적으로 관리할 수 있는 방식을 생각해야 한다고 본다. 이들을 중심으로 국립국어원에서 전사자를 채용하거나 그들과 연계할 수 있는 채용 방식이 필요하다. 개별 연구기관이나 연구자에 종속된 채용 방식의 불안정성을 극복하기 위해 초기에 관계 맺은 연구자 풀, 작업자 풀 안에서 국어원이 전사자를 채용하거나 더 직접적인 연계를 맺을 수 있는 고용 방식을 고민하여 운영하는 것이 좋을 것 같다.

1.3.3 한국수어 말뭉치 공개 방법

한국수어 말뭉치 공개 방법은 공개 대상과 시기, 공개 방법 나누어 볼 수 있다. 농사회 전문가 집단에서는 농인과 청인이 모두 자유롭게 이용할 수 있도록 제공하자는 의견(2명)과 연구자 중심으로 제공하자는 의견(3명)이 있었다.

- 공개는 오직 수어 연구자들에 한해 공개하는 것이 좋다. 수어연구자의 소속과 한국수어 말뭉치 자료 이용 목적 등을 명시한 소정의 자료 이용 청원서를 작성하도록 해서 하자가 없는 한 공개를 허용하도록 하는 것이 바람직하다. 자료 이용 시 반드시 출처를 명시하도록 조건을 달게 하여 수어 말뭉치 자료의 가치를 극대화하는 것이 좋겠다.
- 국어원 홈페이지에 별도로 한국수어를 개설하여 공개하도록 하되,

많은 농인들은 물론 청인들도 자유롭게 이용할 수 있도록 한다.

음성 언어 말뭉치 전문가 집단의 의견은 농사회 전문가 집단과 차이가 있었다. 농사회 전문가 집단은 연구자들에게 우선 공개하자는 의견이 더 많았다면 음성 언어 말뭉치 전문가 집단은 농인과 청인, 연구자와 일반인 모두에게 공개되어야 한다는 의견이 많았다(4명 중 3명).

- 한국수어 말뭉치의 공개 범위를 차등하여야 할 이유가 명확하지 않다. 농인이든 연구자든 모든 사람이 전체 한국수어 말뭉치를 자유롭게 접근할 수 있도록 하는 것이 바람직하다.
- 최종적으로는 연구자와 농인, 청인 등 모든 사용자들에게 공개되어야 할 것이다. 그러나 시범 운영 단계나 한국수어 말뭉치 구축 보완 단계에서는 이용자에 따라 공개 수준의 차이를 두어야 할 것이다.

음성 언어 말뭉치 전문가 집단은 공개 범위에 대해서는 제한적인 공개(1명)와 무제한적인 공개(3명)가 가능하다는 의견을 제시하였다. 제한 없이 공개해야 한다는 의견이 다수였지만 수어 말뭉치의 특성과 시범 단계 등 실행상의 이유로 제한적으로 공개할 수 있지만 기본적으로는 누구나가 이용할 수 있어야 한다는 점은 동일하였다.

- 수어 전문가 등록, 수어 관련 대학원생, 수화통역사, 수어교원 등에게 영상과 전사 자료를 공개하고, 농인과 청인 수어 화자들에게는 영상을 공개하되, 이후에 요청이 있다면 전사 자료를 공개할 수 있다고 생각한다.
- 자료는 동의서를 받았다면 제한하지 말고 원 자료까지 모두 공개하는 게 가장 바람직하다.
- 시범 운영 단계나 한국수어 말뭉치 구축 보완 단계에서는 이용자에

따라 공개 수준의 차이를 두어야 할 것이다.

- 다른 데이터 공개의 경우와 마찬가지로, 특정 정치적 혹은 종교적 신념에 배치가 되는 한국수어 말뭉치나 개인 정보 등이 포함된 한국수어 말뭉치들은 일반 공개가 바람직하지 않으므로 공개를 제한해야 할 수도 있다.
- 한국수어 말뭉치의 경우, 그 활용에 구축 중인 시스템이 반드시 필요하다면, 시스템이 어느 정도 안정화된 이후에 한국수어 말뭉치를 공개하는 것이 타당할 것이다. 불안정한 시스템은 자칫 한국수어 말뭉치 자체의 품질 및 신뢰도를 낮추는 결과를 가져올 수 있기 때문이다.

한국수어 말뭉치 공개 시기에 대해서도 두 집단에서 의견이 차이가 있었다. 농사회 전문가 집단은 사업 주관 기관과 연구 기관에서 어느 정도 양과 질적 측면에서 구축이 되었다고 확신했을 때 공개할 수 있다고 답변한 것에 반해 음성 언어 말뭉치 전문가 집단은 구축 초기부터 공개하는 것이 바람직하다는 의견이었다. 한국수어 말뭉치 구축 초기부터 공개하는 것은 사용자들로부터 피드백을 받아 완성도 높은 한국수어 말뭉치를 구축할 수 있다는 이점이 있음을 알 수 있다.

- 모든 데이터가 그렇듯이, 한국수어 말뭉치 역시 구축 초기부터 공개를 시도하는 것이 바람직하다. 물론 보다 완벽하고 보다 풍부한 상태의 한국수어 말뭉치를 공개하고자 하여 공개시기를 늦추는 경우도 있으나, 초기 공개를 통해 사용자들로부터 많은 피드백을 받고 그것을 이후 구축 과정에 반영한다면 구축된 한국수어 말뭉치의 활용 가능성은 더욱 높아질 것이다.
- 한국수어 말뭉치는 연구진들이 생각하는 균형적 한국수어 말뭉치의 일차 작업이 완료된 후 구축 지침과 함께 공개하는 것이 적절하다고 생각한다. 한국수어 말뭉치 공개는 구축 지침과 함께 이루어져

야 의견과 보완 방법 등 다양한 의견을 받을 수 있을 것이다.

마지막으로, 음성 언어 말뭉치 전문가 집단은 웹사이트를 통해 자유롭게 접근하도록 하는 방법에 의견이 일치하였다.

1.3.4 한국수어 말뭉치 시스템 개발

한국수어 말뭉치 시스템 개발에 대한 음성 언어 말뭉치 전문가 집단의 종합적인 의견은 다음과 같다.

첫째, 주요 사용자들의 요구 사항을 수집·분석하여 그 결과를 반영하여 연차별로 시스템의 기능을 개발하는 것이 타당하다.

- 전사자 간의 신뢰도 외에, 시스템 구축자와의 소통도 매우 필요하다. 특히 시스템의 수정 요구와 이에 대한 반영 여부가 공개적으로 공유될 수 있는 장치 개발이 반드시 필요하다. 사용자를 도입하여, 이에 따른 문제점을 미리 파악해서 이를 보완하는 게 시급하다. 특히 한국수어 말뭉치의 시스템을 활용한 집단에 대한 예비 사용자 평가를 도입할 필요가 있다. 그에 따라 시스템의 구성과 수정이 결정될 수 있을 것이다.

둘째, 새로운 전사 체계를 개발하기보다는 현재 음성 언어 말뭉치와 해외 수어 말뭉치 구축에서 사용하고 있는 전사 도구를 참고하여 한국수어를 위한 전사 도구를 개발하는 것이 필요하다.

- 말뭉치 전사 도구인 엘란(Elan)을 시스템에 탑재할 수도 있지만 필수적인 것은 아니다. 엘란(Elan)으로 분석한 결과를 일관된 기호로 옮기는 작업에 더 중요하기 때문이다. 특히 마지막 단계에서 필요한 정보가 정해진다면 엘란(Elan)이 도구로 합당한지, 잉여적 정보를 제공하는지의 파악이 가능해질 것이다. 이를 위해 동영상 분석이 가능한 몇 개의 프로그램을 모두 사용해 보고 장단점을 정리하는 게 바람직하다.

- 엘란(Elan) 프로그램은 언어 자료 특히 다양한 경로로 이루어지는 의사소통의 총체인 대화를 기록하고 한국수어 말뭉치로 구축하기에 좋은 점들을 가지고 있지만 연구의 목적에 따라 그것을 변용하거나 사용 방법을 익히고, 또 자료를 저장하거나 공개하는 데에 여러 가지 문제를 갖고 있기도 하다. 따라서 한국수어 말뭉치의 특징에 맞게 프로그램을 이용한 방법과 전사의 도구, 지침 등은 향후의 한국수어 말뭉치 구축, 유사한 다른 말뭉치 구축에도 도움이 될 것이기에 공개된다면 좋을 것 같다.

셋째, 한국수어 말뭉치 시스템의 목적에 맞게 개발하는 것이 필요하다.

- 시스템에 탑재하기에 더 적절한 전사 도구나 통합 전사 기능을 갖춘 시스템을 개발한다면 더 바람직한 일이겠지만 그것을 언어와 관련된 여러 분야에서 함께 사용하기에 쉽지는 않다. 어휘 자료를 구축하는 것과 대화 자료를 구축하는 것, 문장 단위를 구축하는 것, 제스처 자료를 구축하는 것은 전사의 가장 기초적인 단위와 수준에서부터 달라지는 일이기 때문이다. 따라서 이와 같은 도구와 시스템을 개발할 때 중요한 것은 각각의 말뭉치에 맞게 프로그램을 만들고 그 방법을 공개하는 것, 또는 다양한 용도를 고려해 공유 가능성이 높은 시스템을 만드는 것이다. 따라서 공개할 시스템 역시 특정한 영역에 제한을 두기보다 이와 같은 가능성을 열어둔 기초적 시스템, 다양한 사용자가 접근하기 쉬운 시스템이어야 할 것 같다.

1.3.5 이용자의 편의를 높이기 위해 필요한 검색 방법

현재 한국수어 말뭉치 구축 지원 통합 시스템에서는 타입명과 토큰명으로 검색이 가능한데 이용자들의 용이한 검색을 위한 방법이 무엇인지 의견을 요청하였다. 이에 대하여 먼저는 한국수어 말뭉치와 언어학적 단위에 대한 이해가 없는 일반 이용자와 한국수어를 잘 모르는 이용자들이 쉽게 검색할 수 있는 방법을 제안하였다.

- 수어를 아는 이들이 특정한 항목을 찾기 위해 검색하는 방법으로서

는 곤란할 것이며 오히려 어휘 항목에 대한 언어학적 지식, 단위에 대한 인식이 없이도 다양한 방식으로 검색할 수 있어야 한국수어 말뭉치에 대한 접근을 높일 수 있을 것이다.

다음으로 의미 중심의 검색 방법을 제안하였다. 타입명과 토큰명은 수어 표현의 의미가 아닌 명칭이므로 검색을 용이하게 하기 위해 수어의 의미와 동작으로 찾아가는 방법이 가능해야 함을 제안하였다.

- 타입 토큰은 [형식] 중심의 검색이다. 형식은 해당 용어를 알아야 검색할 수 있는 방법이니, 해당 언어 형식(혹은 용어 언어 형태)을 검색하기 위한 길잡이 해설(의미 중심으로 ‘00’을 표현하고 싶을 때)같이 검색어에 잡히게 하는 게 좋다. 이는 전사를 하는 과정에서 나오는 다양한 전사의 용어들을 역으로 활용할 수 있을 것이다. 혹은 다양한 지역 차이, 방언의 차이 등도 함께 고려해서 검색의 용이성을 높여야 한다. 혹은 행동 양상(몸짓)을 검색할 수 있게 해야 한다(00 동작).

추가적으로 본 사업에서 사용한 유도 자료의 특성을 활용한 검색 방법의 모색의 필요성과 공개하기 전에 수어 전문가들을 대상으로 시범 운영하여 검색 시 나타나는 문제점 등을 검토하는 것을 제안하였다.

- 여러 형태와 변이형을 최대한 다양하게 검색할 수 있어야 할 것이며 검색을 위한 카테고리, 기준도 공개하여 그것을 통한 검색 가능성도 생각해 볼 수 있다. 또한 구축된 자료의 속성들을 생각한다면 유도 과제의 종류나 자료의 장르, 범주 등을 고려해 검색에 활용할 수 있는 방법도 필요할 것 같다.
- 공개하기 전에 수어 전문가를 대상으로 시범 운영을 실시한 후 이 용자 결과 분석해서 어떤 방법이 나올지 탐색이 필요하다고 생각된다.

1.3.6 한국수어 말뭉치의 활용도를 높이기 위한 방법

한국수어 학습자를 위한 수어 학습과 교육자료, 수어 연구 자료로 써 기여하기 위해 한국수어 말뭉치 구축에서 고려해야 할 점, 추가해야 할 기능, 교사 및 학습자를 위한 웹 인터페이스 개발에 제안할 점, 한국수어 사전 편찬을 위해 고려할 점 등 한국수어 말뭉치의 활용도를 높이기 위한 방법에 관한 의견을 요청하였다.

첫째, 한국수어 사전과 연계하는 방안이다. 한국수어 말뭉치를 활용하여 한국수어 사전에 누락된 비수지 기호 정보를 보완할 수 있으며 향후 수어 사전의 필수 어휘를 선정하고 수집하는데 활용할 수 있다고 제안하였다.

- 현재 한국수어 사전에 그림 정보는 비수지 기호가 거의 없고, 영상 정보는 비수지 기호를 볼 수 있으나, 정보로 제공되지 않기 때문에, 비수지 기호 정보를 보완한 한국수어 사전 편찬에 활용되었으면 한다.
- 궁극적으로 수어의 표준화와 수어 사전을 구축하고자 한다면, 사전에 담을 필수 어휘를 선정해야 하며, 품사별 균형을 정리해야 한다. 특히 주요 용언과 필수적 명사들을 정비하여, 이를 목표로 삼아야 할 것이다. 동영상 자료의 분석은 과제 중심의 귀납적 결과물이므로, 목표로 하는 수어 용어가 나타나지 않을 수도 있기 때문이다. 따라서 기존 사전의 필수 어휘들을 점검하고 이 중 수어에서 반드시 필요로 하는 어휘들을 선정하고 이를 수집할 방법을 연역적으로 모색해야 할 것 같다. 향후 오픈 사전 형식으로 지역별 변이형 등에 대해 추가 수집 제시하는 것도 방법이다.

둘째, 수어 학습에 대한 의견이다. 이는 수어 사전 편찬과도 긴밀한 관련이 있음을 언급하였다.

- 수어 학습과 교육은 수어 사용자와 주변인들, 일반인들 모두를 대상으로 한 것이므로 기초 어휘를 충실히 구축하되 다양한 용법과 변이형에 대한 정보를 함께 저장하고 또 추후에 공개하거나 제시할 방법

을 미리 생각해 자료 구축 작업을 진행하고 시스템도 마련해야 할 것이다. 이는 한국수어 사전 편찬과도 연관된 부분으로 사전의 편찬 항목을 어떻게 구조화하는가는 실제 그 자료들을 구축하고 여러 정보에 대한 주석을 다는 작업과 직결되는 것이다. 따라서 기존의 사전 편찬 항목과 편찬 시스템을 참고해 수어 편찬에 필요한 항목과 주석 시스템, 제시 방법을 미리 고려해 자료 구축의 시스템을 마련하는 것, 보완하는 것이 가장 효율적이고 필요한 일일 것이다. 교사 및 학습자를 위한 웹 인터페이스 개발을 위해서는 실제 사용자들에 대한 수요 조사를 이 단계에서 폭넓게 미리 실시하는 것을 제안한다.

1.3.7 한국수어 말뭉치 홍보 방법

한국수어 말뭉치 홍보 방법에 대한 농사회 전문가 집단의 의견은 농인들이 많이 이용하고 있는 소셜 네트워크 서비스와 수어 관련 학과가 있는 학교, 농인 관련 단체들에게 안내 제공 그리고 자체 홈페이지 개설 등의 홍보 방법을 제안하였다.

1.3.8 한국수어 말뭉치 구축 과정에서 농사회와의 연계 및 협력 방안

농사회 전문가 집단에게만 질문한 내용이다. 한국수어 말뭉치 구축 과정에서 농사회와의 연계 및 협력 방안은 다양한 의견들이 제시되었다.

첫째, 각 지역 수어의 전사 과정에서는 해당 지역의 언어 제공자들 혹은 그 지역 수어에 능숙한 농인과 청인들과의 협력이 필요하다.

둘째, 농사회에서 언어학적 이해를 갖춘 농인들을 중심으로 한 연구 인력의 구성이 필요하다.

셋째, 전국 지역을 대표하는 지식층 농인들로 구성된 모니터링단을 구성하여 그들의 피드백을 받는 것이 필요하다.

넷째, 농인 수어 연구자 양성이 필요하다.

다섯째, 전사 과정에서 수어를 보고 제대로 정리할 수 있는 농인 전사자를 선정하는 것이 필수적이며 최종 점검 과정에서는 보다 능력을 갖춘 농인이 정리를 담당하는 것이 필요하다.

1.3.9 기타

향후 한국수어 말뭉치 구축을 위한 자유로운 의견을 요청한 것에 대한 음성 언어 말뭉치 전문가 집단의 의견이다.

첫째, 한국수어 말뭉치 구축과 활용의 목적을 최우선적으로 정의하여 그에 따른 적절한 방법론을 정의하는 것이 필요하다.

둘째, 한국수어 자료에 축약형, 방언형과 한국수어 사전에 등재된 수어를 함께 제시하는 것이 필요하다.

셋째, 한국수어 말뭉치 구축은 농인들의 소통을 위한 의미 있는 연구일 뿐만 아니라 말뭉치 언어학 측면에서는 수어 영상을 분석하는 최신의 말뭉치 연구의 산물이 되므로 학계와 농사회가 체계적인 수집과 분석을 위해 노력할 때 의미 있는 결과물의 산출이 가능하다.

2. 한국수어 말뭉치 구축 방안

2013년부터 2016년까지 수행된 일련의 한국수어 말뭉치 구축 사업을 통해 한국수어 말뭉치 구축을 위한 방법론과 지침서 개발 등 기초적인 체계를 마련해왔다. 그러나 2015년에 서울 지역의 언어 자료를 수집한 이후 전국적인 언어 자료 수집이 이루어지지 않아 지역별 수어의 특징뿐만 아니라 보편적인 언어 사용 양상을 전체적으로 파악할 수 있는 자료수집이 실시되지 못하고 있다. 본 연구에서는 지난 2년간의 한국수어 말뭉치 구축 및 전사 경험과 한국수어 연구자 및 농인 당사자, 그리고 음성 언어 말뭉치 전문가로부터 얻은 자문 내용을 토대로 하여 향후 한국수어 말뭉치 구축 방안을 제안하고자 한다. 모든 자료의 수집은 그 목적에 따라 자료의 특징과 규모, 그리고 수집 방법이 달라지므로 한국수어 자료의 수집 역시 그 목적에 따라 어떤 자료들을 어떤 방법으로 수집할 것인가가 결정된다. 따라서 앞으로의 언어 자료 수집 과정에서는 한국수어 말뭉치 구축의 목적을 보다 명확히 하고 그 목적에 따라 자료 수집 방법을 수정·보완할 필요가 있다. 한국수어 말뭉치의 구축 목적에 따른 다양한 구축 방법이 있을 수 있으나 본 연구에서는 지역별 한국수어 수집을 목표로 하여 한국수어 말뭉치 구축을 위해 고려할 사항을 다음과 같이 제안한다.

2.1 언어 자료 수집 방법

2.1.1 목표 어휘 및 문법 항목의 명시

유도 자료를 통해 언어 자료를 수집할 경우 각 과제별로 예상되는 목표 어휘와 문법 항목을 기술한다. 일상적인 주제를 담은 유도 자료는 한국수어의 필수적인 기본 어휘 목록을 정비하는데 도움이 되며 토론처럼 특정 주제를 담은 유도 과제는 격식적인 표현을 수집하는데 도움이 된다. 다양한 유도 과제별로 이끌어낼 수 있는 예상 어휘와 문법을 정리하여 자료 수집의 효율성을 높인다.

2.1.2 연역적 방식의 언어 자료 수집

한국수어 관련 사전 편찬을 위해 언어 자료를 수집한다면 사전에 들어갈 필수 어휘를 품사별 균형성을 고려하여 선정해야 하며 해당 목표 어휘들을 연역적으로 수집하는 방법 또한 모색할 수 있다.

2.1.3 자료 수집 방법의 다양화

현재까지 구축한 한국수어 말뭉치 자료는 유도 과제를 사용하여 이끌어냈다. 향후 지역별 수어를 수집하더라도 기존 자료와의 연관성 및 지속성을 유지하는 것이 중요하므로 현재까지 사용한 유도 자료를 가지고 언어 자료를 수집하는 것이 바람직하다. 그러나 자연스러운 환경 속에서 나타난 담화, 대화 등의 자료나 일상 언어와 통속적 언어 등 사적인 발화 자료를 수집할 필요가 있으므로 자료 수집 방법을 다양화하고 유도 자료를 통한 수집 방법과의 적절한 균형이 이루어지도록 자료 수집 방법을 계획하는 것이 필요하다.

2.1.4 언어 자료 자원의 확대

기존에 있는 다양한 형태의 한국수어 교재 및 동영상 자료 등을 수집할 필요가 있다. 이는 앞에서 언급한 자료 수집 방법의 다양화에 기여할 수 있는 방법이라 하겠다. 농인들의 회의 장면, 대화 장면 등의 영상 자료를 수집할 수 있으나 개인 정보와 농인의 얼굴이 나타나므로 초상권을 고려하여 수집에서 세심한 주의가 필요하다.

2.2 언어 자료 수집을 위한 지역 수와 선정 기준

현재까지 수집된 언어 자료는 서울 지역 수어에 한정되어 있어 지역 내 수어의 특징 파악과 지역 간 수어 비교가 불가능하다. 향후 한국수어 말뭉치 구축에서 지역별 수어 수집을 목표로 한다면 해당 지역성을 고려하여 지역을 선정할 필요가 있다. 전국 17개 시·도 지역마다 한국농아인협회가 있으므로 언어 제공자 선정과 언어 자료 수집을 위해 각 시·도 농아인협회와 협력하는 방안이 있다. 지역 수어를 수집하기 위해서는 해당 지역의 농사회 및 농인 관련 단체의 협조가 반드시 필요하다. 이를 위해 언어 자료 수집 이전에 해당 지역에서 농인 중심의 워크숍을 개최하는 등 지역 홍보를 진행한다면 한국수어 말뭉치의 언어 자료 수집이 더욱 수월할 것이다. 언어 자료를 수집하기 위한 지역의 수는 예산과 참여 인력에 따라 달라지겠지만 수어의 경우는 동영상 촬영과 편집 그리고 전사가 필요하므로 1년에 1~2지역이 적절할 것으로 보인다. 또한 언어 자료 수집 때는 촬영한 영상의 자동 백업 장치의 설정 등 촬영 및 편집 전문가의 참여가 필요하다.

2.3 언어 자료 수집을 위한 주체

한국수어 말뭉치 구축은 언어 자료의 수집과 전사 그리고 관리와 공개 등 어느 한 기관이 수행하기에는 매우 복잡하고 전문적인 과업이다. 따라서 농사회의 적극적인 참여가 요구되며 전문적인 지식이나 이론 등 대학의 연구 기반을 활용할 수 있어야 한다. 그러나 국민의 언어생활에 대한 기록과 자료 수집이라는 큰 틀에서 볼 때 한국수어 말뭉치 구축은 국가가 주체가 되어야 하며 사업을 지속적으로 수행해가는 것이 가장 중요하다.

2.4 전사 방안

2.4.1 전문성을 갖춘 전사자 인력 확보

한국수어 말뭉치의 전사의 질은 전사자에 따라 크게 좌우된다. 전

사자는 수어에 대한 메타 언어적인 시각을 갖고 수어 자료를 분석할 수 있어야 하므로 단기간에 전문성을 갖추 수 없다. 2016년 한국수어 말뭉치 사업에서 시범 구축의 경험이 있는 연구자와 전사자가 참여하기는 하였지만 새로운 전사자를 투입할 수밖에 없는 상황에서 많은 시간을 교육에 투자하므로 그 효율성이 떨어질 수밖에 없었다. 해외의 사례에서도 보면, 네덜란드의 경우는 전일제 전사자 3명을 고용하여 전사를 하고 있다. 또한 독일에서는 전일제 전사자와 전사 교육자 외 수어 언어학 및 통역학을 전공하는 학부생들이 공동 전사를 하고 있다. 이에 향후 한국수어 말뭉치 구축 사업에서는 전사자를 채용하여 전사를 실시할 필요가 있다.

또한 전사자의 양적 확대와 전사 능력의 질적 향상을 위해 전사자 양성 교육 실시도 한 방법이 될 수 있다. 한국수어 관련 학과와 연계하여 전사 관련 연구 실시, 매년 일정 기간 동안 한국수어 전문가들을 대상으로 한 전사자 양성 교육 실시, 전사자 수준별 교육 내용 다양화 등 전사자의 전문성 확보를 위한 다각도의 방안이 모색되어야 하겠다.

2.4.2 전사 품질 관리

전사자 간 교차 검토와 정기적인 전사 회의를 통한 의견 조정 과정을 시행하여 전사자 간 혹은 전사자 내 전사의 일치도와 정확성 등 전사의 질과 효율을 높이도록 한다.

2.4.3 전사 도우미 어플리케이션 개발

향후 한국수어 말뭉치 구축의 효율성을 높이기 위한 방안으로 ‘전사 도우미 어플리케이션’의 개발을 제안한다. 이 어플리케이션은 연구자나 전사자가 수어 전사를 하다가 궁금한 내용이 있으면 정리를 해서 어플리케이션을 통해 농민들에게 물어보는 것이다. 그러면 회원 가입을 한 농민들이 질문에 대한 대답을 해 주고, 대답에 참여한 농민 회원들은 점수를 얻게 되고 이 점수에 따라 상품 등을 지급하는 형태이다. 독일수어 말뭉치에서 개발한 어플리케이션을 참고할 수 있다.

2.4.4 자동 전사 방안에 대한 연구

최근의 말뭉치 자료 구축 및 활용은 빅데이터 기반의 환경을 고려하여 개선되고 있으므로 한국수어 말뭉치도 이러한 기반 기술과의 결합이 필수적이다. 특히 영상 인식 기술의 놀라운 발전 성과를 도입하여 자동 전사와 데이터 구축을 하는 것이 지금까지의 말뭉치 구축 작업의 한계를 극복하고 다음 단계로 발전하는 데 필요하므로 이에 대한 연구가 추진되어야 한다.

2.4.5 의미 전사 확보

전사 양을 확보하는 것도 중요하지만 기본 전사와 번역 외에 의미 전사를 하는 것이 중요하다. 2017년 새로 등록한 타입들은 타입 수준에서 의미를 기록했으나 토큰 의미에 대한 기록은 전혀 없다. 이에 추후 전사 진행 시 의미 층렬을 작성하여 기본 의미에서 벗어나는 경우 전사자가 기본 의미를 상세히 기록하여 이것을 시스템에서 관리 할 수 있도록 하는 것이 필요하다. 시스템을 통해 의미를 관리할 경우 다양한 의미 검색이 가능하여 한국수어 말뭉치 활용이 확대될 것이다.

2.4.6 수어 형태에 대한 전사

기 전사에서는 글로스만 전사되어 있어 타입명을 관리하고 이를 검색하는 것이 가능하다. 그러나 타입명을 모를 경우 해당 타입을 찾을 수가 없다. 이에 모든 수어의 형태를 수어 문자 형태로 전사하는 것이 필요하다. 그러나 그 규모가 방대함으로 타입 목록의 수어 형태 만이라도 HamNoSys 또는 SignWriting으로 전사하는 방안이 필요하겠다. 수어 형태를 어떤 방식으로 전사 할 것인지에 대한 모색이 필요하나 수형 사전에서 사용되었던 수형 사진을 함께 사용하는 방법도 바람직하다.

2.4.7 농인들이 편리하게 사용할 수 있는 시스템

현재 한국수어 자료 통합 지원 시스템은 수어 사용자보다 음성 언어 사용자를 위한 시스템이라 할 수 있다. 앞으로 있을 각 지역 전사 계획을 고려한다면 청인 전사자보다 농인 전사자가 많이 필요할 것이다. 현재 한국의 전사 환경은 외국과 달리 전사자들이 함께 전사를 할

수 있는 공간이 없고 개인의 공간에서 전사를 해야 한다. 그렇다면 서로 수어로 소통하고 회의를 하며, 동시에 영상 자료를 보며 공유하는 시스템이 갖추어져야 한다. 예를 들어 영상 쪽지 기능, 영상 회의 기능 등이 필요하다. 농인의 시각적인 필요에 따라 청인의 기준을 벗어나 보다 효과적이고 혁신적인 접근 방법을 모색하는 것이 바람직하다.

2.5 농인 수어연구자 양성

모든 수어 연구에 농인들의 참여가 중요하지만 한국수어 말뭉치는 특히 농인의 참여 없이는 불가능한 분야이다. 아무리 수어를 잘 아는 통역사라 할지라도 농인들만의 대화를 전사할 때는 한계를 느낄 것이고 지역 수어의 전사를 시작하면 더욱 그럴 것이다. 따라서 전사 과정에 참여할 모국어 농인 전사자뿐만 아니라 전사 후 분석까지 수행할 수 있는 농인 연구자를 양성해야 한다. 현재 국내에서 농인 연구자를 양성할 수 있는 여건이 부족한데, 적어도 한국수어 말뭉치 연구에서 농인 연구자와 농인 전사자를 양성할 수 있는 방안 마련이 시급하다.

2.6 한국수어 말뭉치 기반 수어 연구

한국수어 말뭉치 구축과 전사 작업의 규모가 워낙 커서 연구진은 한국수어 말뭉치의 전사와 관리를 하다 보면 연구 주제에 따라 연구에 집중하기가 쉽지 않다. 그러나 한국수어 말뭉치 기반 연구는 과거에 전사자료 없이 실시했던 연구를 확인할 수 있는 기회이기 때문에 매우 중요하다. 더 나아가 한국수어 말뭉치를 기반으로 하여 새로운 주제에 대한 연구를 수행하는 것이 바람직하다.

2.7 한국수어 말뭉치 구축 자료 공개

2.7.1 한국수어 말뭉치 구축 공개 대상과 범위

한국수어 말뭉치의 시스템 안정화 등 실행상의 이유로 일부 연구자들에게 제한적으로 공개할 수 있지만 최종적으로는 모든 이용자들에게 공개되어야 한다. 한국수어 말뭉치를 공개할 때 일반 이용자에게는

한국수어 영상을 제공하며 연구자들에게는 전사 파일까지 공개하는 등 이용자들에 따라 다르게 적용할 수 있다.

2.7.2 한국수어 말뭉치 구축 공개 시기

한국수어 말뭉치는 영상 자료 중심이기 때문에 일반 음성 언어 말뭉치보다 고려할 사항이 더 많으므로 성급한 공개보다는 잘 정리된 시범 자료를 공개하는 방법을 제안한다. 또한 한국수어 말뭉치 자료의 공개를 통해 언어 및 수어 연구자들로부터 보완을 위한 다양한 의견이나 새로운 방향 설정에 대한 피드백을 받아 한국수어 말뭉치의 완성도와 활용도를 높일 수 있어야 하므로 피드백을 제공해 줄 연구자들이 활용할 수 있을 정도의 충분한 영상과 전사 자료 등이 구축된 후에 공개되어야 한다는 점도 공개 시기를 결정할 때 고려해야 할 사항이다.

2.7.3 한국수어 말뭉치 구축 공개 방식

한국수어 말뭉치 구축 자료는 모든 이용자가 쉽게 접근할 수 있도록 웹 사이트를 통해 제공되어야 한다. 2016년 한국수어 말뭉치 등 한국수어 관련 자료를 한 곳에 모아서 관리할 수 있는 한국수어 자료 통합 지원 시스템을 마련하였다. 이용자들이 개인 컴퓨터에서 직접 자료를 내려 받는 방식은 자료가 쉽게 유출될 수 있는 위험성이 있다. 한국수어 말뭉치의 언어 제공자들은 한국수어 말뭉치 사업팀에 영상 자료에 대한 권리를 위임하였으며 이것이 한국수어 말뭉치 구축과 학술적 목적을 위해 사용되는 것에 동의하였다. 그러므로 한국수어 말뭉치 사업은 이 영상들이 유출되어 위의 목적 외로 사용되는 것을 막기 위해 주의를 기울여야 한다. 따라서 언어 자료에 대한 보안성을 높이기 위해 영상 파일의 다운로드를 허용한다면 워터마크(watermark)와 같은 장치를 사용하여 본래의 용도 외로 불법적으로 사용하지 못하도록 방안을 마련해야 한다. 또한 한국수어 말뭉치 자료는 언어 제공자들의 얼굴이 공개되기 때문에 유도 과정에서 나누었던 개인적인 대화 내용이 여과 없이 공개되는 것은 곤란하다. 따라서 한국수어 말뭉치 자료를 어떤 방법으로 공개할지 추후 세심한 검토 과정을 거친 후에 공개해야 한다.

2.8 한국수어 말뭉치 구축 자료 접근성

한국수어 말뭉치는 한국수어 관련 전공자나 연구자 외에 한국수어를 모르는 일반인들이나 다른 학문 분야의 전문가들에게도 열려 있어야 한다. 그러므로 한국수어를 모르는 이용자들이 접근할 수 있도록 수어 영상에 대한 한국어 번역을 제공할 뿐만 아니라 타입명과 토큰명과 같은 형식 중심의 검색 방법 외에 의미 중심의 검색 방법을 마련하는 것이 필요하다. ID 글로스는 토큰의 의미가 아니기 때문에 한국수어를 아는 사람조차 타입명과 토큰명으로는 쉽게 찾기 어려울 때가 있으므로 의미 정보로 원하는 토큰들을 검색하는 방법은 유용한 검색 방법이 될 것으로 보인다. 이를 위해 전사할 때 의미를 더 상세하게 기록하는 방법이 있을 것이며 의미 전사를 관리할 수 있는 시스템의 보완도 필요하겠다.

2.9 한국수어 말뭉치 구축 로드맵

한국수어 말뭉치 구축을 위한 연차별 계획은 10년 사업으로 실행하되 3단계로 나누어 추진하는 방안을 제안한다. 각 연차별 추진 내용은 <표 42>과 같다.

<표 42> 한국수어 말뭉치 연차별 추진 내용

| 구분 | 기존 완료 (2015-2017년) | 1단계 (2018-2020년) | 2단계 (2021-2023년) | 3단계 (2024-2027년) |
|----------------|---|---|---|--|
| 언어 자료 수집 지역 | • 서울 지역(언어 제공자 30쌍, 90시간 수집) | | • 전국 8도와 특별자치도 중 3개 지역 선정 | • 전국 8도와 특별자치도 중 3개 지역 선정 |
| 언어 자료 수집 방법 | • 유도 자료 작성 | • 자료 수집 방법 검토 및 보완 | • 자료 수집 방법 검토 및 보완 | • 자료 수집 방법 검토 및 보완 |
| 전사 분량 | • 기본 전사: 13시간 + 250분 • 정밀 전사: 112분 (비수지 기호) | • 서울 지역 수어 전사: 250분 • 지역 수어 전사: 250분 • 정밀 전사: 100분 • 의미 전사(엘란 파일에 의미 층렬 추가) • 타입의 수형 전사 | • 서울 지역 수어 전사: 250분 • 지역 수어 전사: 250분 • 정밀 전사: 100분 • 의미 전사(엘란 파일에 의미 층렬 추가) • 타입의 수형 전사 | • 서울 지역 수어 전사: 250분 • 지역수어 전사: 250분 • 정밀 전사: 100분 • 의미 전사(엘란 파일에 의미 층렬 추가) • 타입의 수형 전사 |
| 한국어 번역 | • 2015년 13시간 전사분 중 일부 완료 • 2017년 250분 전사분 완료 | • 기본 전사와 정밀 전사분 한국어 번역: 600분 | • 기본 전사와 정밀 전사분 한국어 번역: 600분 | • 기본 전사와 정밀 전사분 한국어 번역: 600분 |
| 한국수어 말뭉치 연구 방안 | • 한국수어 말뭉치 구축 지침 및 전사 지침 작성 • 비수지 연구 사례 | • 한국수어 말뭉치 공개 방법 마련 - 국외 선행 연구물 검토 - 한국수어 말뭉치 공개에 대한 모델 3개 선정 | • 지역 수어 비교 및 관리 | • 한국수어와 한국어 비교 연구 |

| 구분 | 기존 완료 (2015-2017년) | 1단계 (2018-2020년) | 2단계 (2021-2023년) | 3단계 (2024-2027년) |
|-------------------|---|--|---|---|
| 한국수어 말뭉치 홍보 방법 | <ul style="list-style-type: none"> • 농인을 대상으로 한국수어 말뭉치 워크숍 개최 • 농인 풀뿌리 워크숍 개최 | <ul style="list-style-type: none"> • 해당 지역 한국수어 말뭉치 사업 소개 워크숍 개최 | <ul style="list-style-type: none"> • 해당 지역 한국수어 말뭉치 사업 소개 워크숍 개최 | <ul style="list-style-type: none"> • 해당 지역 한국수어 말뭉치 사업 소개 워크숍 개최 |
| 한국수어 말뭉치 공개 방법 | | <ul style="list-style-type: none"> • 시범 공개 <ul style="list-style-type: none"> - 3과제 수어 영상 공개 - 1과제 전사 자료 공개 - 전사 지침 공개 • 한국수어 말뭉치 이용자 의견 수렴 및 보완 | <ul style="list-style-type: none"> • 공개 <ul style="list-style-type: none"> - 5과제 수어 영상 공개 - 1과제 전사 자료 공개 - 전사 지침 공개 • 한국수어 말뭉치 이용자 의견 수렴 및 보완 | <ul style="list-style-type: none"> • 공개 <ul style="list-style-type: none"> - 5과제 수어 영상 공개 - 1과제 전사 자료 공개 - 전사 지침 공개 • 한국수어 말뭉치 이용자 의견 수렴 및 보완 |
| 시스템 | <ul style="list-style-type: none"> • 전사 파일 관리 • 타입 목록 • 메타 데이터 관리 • 토큰 검색 • 단순 통계 기능 | <ul style="list-style-type: none"> • 한국수어 말뭉치 시스템 고도화 • 농인에게 적합한 접근 방법 마련(영상채팅 등) • 검색 강화 • 통계 기능 강화 | <ul style="list-style-type: none"> • 전사 도우미 어플리케이션 개발 | |
| 한국수어 말뭉치 활용 방안 | | <ul style="list-style-type: none"> • 한국수어 사전과의 연계 방안 <ul style="list-style-type: none"> - 기초 어휘 선정 - 지역적 변이형 - 용례 추출 - 수어 의미론 연구 | <ul style="list-style-type: none"> • 한국수어 교육 및 수화 통역과의 연계 방안 <ul style="list-style-type: none"> - 농문화 수집 - 비수지 기호 수집 - 수화통역사를 위한 영상 / 번역 공개 자료 연구 | <ul style="list-style-type: none"> • 한국수어와 한국어 비교 연구 방안 <ul style="list-style-type: none"> - 통사론 연구 - 제스처 연구 • 지역적 변이형 비교 연구 |

3. 워크숍 개최

한국수어 말뭉치 연구 및 구축을 위한 워크숍은 총 2회 진행하였다. 1차 워크숍은 2017년 8월 24일 개최하였다. 홍성은, 허일이 독일을 중심으로 한 해외 수어 말뭉치 기반 수어 연구 현황과 엘란(Elan)프로그램을 사용한 수어 연구 방법 소개를 주제로 진행하였다. 2차 워크숍은 2017년 9월 22일에 개최하였으며 네덜란드를 중심으로 한 해외 수어 말뭉치 기반 수어 연구 현황에 대해 변강석을 통해 들을 수 있었다. 워크숍 참석자는 1차에는 연구 참여자 7명, 전사자 7명, 국립국어원 3명, 희망자 3명 총 20명이었고 2차에는 연구자 7명, 국립국어원 6명, 희망자 5명 총 18명이었다.

워크숍 장면은 <그림 63>과 같다.



<그림 63> 워크숍 모습

VI. 한국수어 말뭉치 구축 지침

1. 2017년 개정 방향

해외 사례에서도 알 수 있듯이 한국수어 말뭉치 구축은 짧은 시간에 이루어질 수가 없다. 그러나 한국수어 말뭉치 기반 조사 및 시범 구축(원성옥, 2015) 사업은 7개월이라는 짧은 기간 동안 언어 자료의 수집에서부터 전사 환경에 이르기까지 연구 기반이 전무한 상태에서 이루어지므로 많은 시행착오와 어려움이 있었다. 이에 2015년 한국수어 말뭉치 시범 구축에서 제시한 문제점 및 어려움을 먼저 분석하고, 이를 해결하는 방향으로 언어 자료 수집 과정에서부터 전사에 이르기까지 전반적인 내용을 수정 보완하였다. 또한 2017년 개발한 한국수어 자료 통합 시스템을 기반으로 하여 지침서를 수정하여 향후 보다 효율적이고 체계적인 한국수어 말뭉치 구축에 활용될 수 있도록 하였다. 2017년 개정 한국수어 말뭉치 구축 지침은 <부록 4>에 제시한다.

2. 언어 수집 과정에서의 수정 보완

2.1. 촬영 시간

2015년 한국수어 말뭉치 구축 과정에서는 연구 일정이 촉박하여 18일 동안의 짧은 기간 동안 30쌍의 언어 제공자 모두를 촬영하였다. 따라서 언어 제공자와 일정 조절이 쉽지 않았으며, 하루에 2쌍에서 3쌍의 언어 제공자를 촬영함으로 인해 사회자의 피로가 누적되는 것을 비롯하여 촬영 일정을 정하는 데에도 많은 어려움이 있었다.

또한 언어 제공자가 약속 시간을 착각하거나 지각하는 경우가 있었고, 연이은 촬영으로 인한 촬영 준비의 지연, 설문지를 미리 작성해 오지 않은 경우 설문지 작성으로 인한 지연 등으로 촬영 시간이 계획보다 늦어지는 경우가 많았다. 이로 인해 촬영 일정에

여러 가지 어려움이 발생했다. 이에 2017년 지침에서는 언어 제공자 선정 및 모집에서 촬영 시간 및 준비물에 대한 자세한 안내와 점검 및 충분한 촬영 시간을 확보하는 내용을 유의점으로 추가하였다.

2.2 두 언어 제공자 간의 관계 및 메타 데이터 수집

A 언어 제공자와 B 언어 제공자가 어색한 사이라면 대화에 어려움이 있었다. 특히 서로 아는 사이지만, 불편한 관계인 사람도 있었다. 또한 두 언어 제공자 사이의 나이 차이가 많은 경우는 젊은 언어 제공자가 나이 많은 언어제공자와 대화를 시작하는데 불편해하고 어려워하는 경향이 있었다. 처음 보는 사람인 경우 1시간 정도가 지나서야 자연스럽게 대화를 이어갈 수 있었다. 그러므로 다음에는 추천된 A가 직접 B를 추천하게 하는 방식으로 이러한 상황을 방지할 수 있을 것이다. 그리고 A와 B 언어 제공자가 서로 농인 가족이 있고 가정 환경이 비슷하면 서로 충분히 이해하고 인식 차이를 거의 느끼지 않기 때문에 대화가 편안하게 진행되는 경향이 있었다. 이에 2017년 지침에서는 언어 제공자 선정 및 모집의 유의점에 언어 제공자 선정 방식에 대한 내용을 추가하였다.

또한 기존의 자료 수집을 위한 질문지의 형식은 너무 복잡하고 한국의 상황과 맞지 않는 부분이 있어 한국의 상황에 적절하게 질문의 내용을 단순화하는 등으로 수정하였다.

2.3. 사회자의 역할

짧은 촬영 일정으로 인해 하루에 3번 촬영을 하는 경우도 있었으며 이에 따라 사회자에게 신체적 그리고 정신적으로 상당한 무리가 있었다. 또한 사회자가 촬영하는 경우 두 언어 제공자에게 각각 유도 자료를 제시하므로 두 대의 컴퓨터를 따로 클릭하며 조절해야 하는 번거로움도 있었다. 결국 언어 제공자를 유도하는 것

과 컴퓨터 2대에 대한 조작을 동시에 해야 하는 어려움이 있었다. 이에 2017년 지침에서는 언어 자료 수집의 유의점에 사회자의 어려움을 반영한 촬영실 조건 및 촬영 시스템에 대한 내용을 추가하였다.

2.4. 촬영 환경

대부분 언어 제공자들은 촬영실 접근에 어려움이 없었으나, 나아가 드신 언어 제공자의 경우 촬영실을 찾는 데 약간의 어려움이 있었다. 따라서 촬영실 위치에 대한 더 자세한 안내가 미리 제공되어야 할 필요가 있다. 또한 3시간 가까이 촬영을 하기 때문에 의자도 불편하고 모니터를 보는 것을 불편해하기도 하였다. 따라서 더욱 안락한 환경을 조성할 필요가 있으며, 촬영 모니터실과 언어 제공자 휴게 공간을 분리하여 언어 제공자들이 촬영 모니터를 볼 수 없게 해야 한다.

카메라의 메모리(SM과 CF)칩 용량은 64GB 이상으로 하는 것이 안전하다. 또한 카메라는 충전 배터리를 사용하는 것은 중간에 전기가 나갈 위험이 있어 사용하지 않는 것이 좋다. 따라서 직접 코드에 연결하여 사용하여야 하는데, 전기선이 보이지 않아야 하고, 다니면서 건드리지 않도록 유의하여야 한다. 무엇보다 카메라가 작동하고 있는지를 수시로 점검할 필요가 있다. 이에 2017년 지침에서는 촬영 환경과 관련한 내용들을 수정하면서 언어 자료 수집의 유의점에 촬영실 조건 및 촬영 시스템에 대한 내용을 추가하였다.

2.5. 촬영 동영상의 편집 및 관리

한국수어 말뭉치 구축을 위해서는 언어 자료 수집을 위한 동영상 촬영과 편집 그리고 수집된 자료를 전사하는 작업이 그 핵심이다. 따라서 언어 자료 수집을 위한 동영상 촬영 시 자동 백업 장치의 설정 등 촬영 및 편집 전문가의 참여가 필요하다. 더불어 전

산 언어학이나 전산을 전공한 전문가의 참여가 있어야 한다. 또한 촬영이 끝나고 나면 바로 그날 촬영한 동영상에 문제가 있는지에 대한 점검이 필요하다. 따라서 촬영을 마친 후에는 반드시 3대의 카메라로 촬영한 영상에 문제가 있는지를 모두 확인한 후에 폴더 별로 구분하여 저장하고, 촬영한 동영상의 오류가 발견되면 재촬영 일정을 잡는 등 신속하게 대처하여야 한다. 이에 2017년 지침에서는 촬영 환경과 관련한 내용들을 수정하면서 언어 자료 수집의 유의점에 편집 및 관리에 대한 내용을 추가하였다.

3. 전사 지침에서의 수정 보완

기구축 한국수어 말뭉치를 한국수어 자료 통합 지원 시스템 내에 정비하는 과정에서 나타난 문제점과 본 연구 사업에서 전사하는 과정에서 발견한 문제점 등을 고려하여 다음과 같은 내용을 2017년 개정 지침에 수정 반영하였다. 첫째, 한국수어 말뭉치 구축에서 전사는 매우 시간이 오래 걸릴 뿐 아니라 문자가 없는 언어이므로 전사 시 그 일치도를 확보하는 데에 어려움이 있는 작업이다. 따라서 전사자 품질 관리를 위한 내용을 추가하였다.

둘째, [지시]와 관련된 타입이 세분화되어 있어 전사 시 혼란을 초래하였다. 따라서 타입을 16개에서 8개로 줄이고 타입명을 단순화하였다.⁵⁾

셋째, Counting hand에 해당하는 타입명을 정하는 규칙이 비효율적이어서 기본 전사의 목적에 맞게 전사 방법을 변경하여 제시하였다.

넷째, 신체를 지시하는 경우, 신체 부위명과 신체를 지시하는 타입의 구분이 모호하여 전사 방법과 타입명 결정 방법을 수정하였다.

다섯째, 기본 전사의 목적에 맞게 생산적 수어에 대한 타입 목록을 재분류하고 단순화하였다.

여섯째, 기본 전사에서는 해당 수어가 등록된 타입과 수위나 수

5) Ⅲ. 기구축 한국수어 말뭉치의 전사 규칙 정비 참조

형 등 하나의 수화소가 다른 경우 이를 새로운 타입으로 등록하지 않고 변이형으로 보도록 수정하였다. 예를 들어 [지시1_비1]에서 기본형은 검지만 편 수형이나 엄지까지 편 경우 새로운 타입을 등록하지 않았다. [있다]의 경우 [있다] 이전의 수어 위치에 따라 수위가 변화될 수 있으나 그 의미에는 변화가 없으므로 이를 변이형으로 보아 새로운 타입으로 등록하지 않았다.

일곱째, 타입명이 한국어로 ‘구’의 형태라 하더라도 수어 기준으로 하나의 수어로 여겨질 경우, 수어를 분절하지 않고 하나의 타입으로 등록하였다. (예: [마음에안들다1])

VII. 결론

1. 연구 결과 요약

2015년 서울 지역을 중심으로 한국수어 말뭉치 시범 구축 사업이 이루어졌으며, 그 결과를 토대로 효율적인 한국수어 말뭉치 구축을 위해 한국수어 자료 통합 지원 시스템이 2016년 구축되었다. 이에 본 연구에서는 2015년 시범 구축 사업에서 전사한 자료들을 지원 시스템에 연결하는 과정에서 발생하는 문제들을 해결하고, 시스템을 활용한 기본 전사 작업을 실시하였다. 또한 한국수어 말뭉치를 기반으로 하는 한국수어 연구 사례와 향후 한국수어 말뭉치 구축 계획을 제시하고 구축 지침을 수정 보완하였다. 연구의 결과를 요약하면 다음과 같다.

1.1 기구축 한국수어 말뭉치의 정비

첫째, 2015년 시스템 없이 전사한 한국수어 말뭉치를 2016년에 구축한 시스템에 연결하는 과정에서 발생한 오류의 유형을 분석하여 총 13,340개 토큰의 연결 실패 문제를 해결하였다.

둘째, 연결 문제 해결 과정에서 나타난 오류 문제를 해결하기 위해 한글 지문자, 영어 지문자, 얼굴 이름에 대한 추가 타입을 등록하였다.

셋째, 2개의 타입을 1개의 타입으로 통일(39건), 타입명 수정(14건), 새로운 타입 등록(36건) 등 타입을 정비하였으며, 2015년 전사 초기에 전사된 변이형 토큰 중 타입명이 존재하지 않는 경우에는 기본형 타입에 연결하였다.

넷째, 시스템에서의 연결 실패 문제를 해결하고 타입 정리 등을 하면서 발견된 분절 오류, 토큰 누락, 지문자 전사 오류, 단순한 손짓의 전사 오류 등 656개의 문제 토큰에 대한 재전사를 실시하였다.

1.2 전사

2015년 수집한 언어 자료(동영상) 중에서 총 261분 57초 분량의 영상에 대해 우세손과 비우세손에 대한 기본 전사 및 번역을 완료하였다. 전사 절차 및 결과는 다음과 같다.

첫째, 페이스 북 등을 통해 농인 5명 청인 8명 총 13명의 전사자를 모집하여 전사 교육을 실시하였다. 전사 교육은 6월 21일부터 23일까지 총 3회 18시간 실시하였으며, 한국수어 말뭉치에 대한 기본 지식과 ELAN 및 시스템 사용법 등을 교육하였다.

둘째, 전사 교육 후 6월 26일부터 28일까지 총 3회에 걸쳐 공동 전사를 실시하여 개인별 총 6시간을 기본으로 공동 전사에 참여하도록 하였다.

셋째, 전사의 품질 관리를 위하여 본 전사를 시작하기 전 3차례에 걸쳐 전사 신뢰도를 측정하였다. 신뢰도 검사 결과는 1차 47.55%, 2차 59.81%, 3차 71.51%이었다. 3차 신뢰도 검사 결과를 분석하여 일치도가 낮게 나타난 토큰들의 특징을 분석하여 한 차례의 추가 전사 교육을 실시한 후 본 전사를 시작하였다. 3개월 동안 전사를 실시한 후 다시 4차 신뢰도 검사를 실시한 결과 82.91%의 일치도를 확보하였다.

넷째, 우세손과 비우세손에 대한 기본 전사 및 번역을 완료한 파일은 토론 과제(과제 번호 E) 4개, 주제 영역 과제(과제 번호 H) 3개, 농사회 행사(과제 번호 J) 4개, 농학교 경험 이야기(과제 번호 M) 4개로 총 15개이다. 전사 결과 사용한 타입 수는 1,057개이고 발생 토큰 수는 23,637개, 새롭게 등록한 타입은 295개이다.

다섯째, 전사 결과 가장 많이 나타난 타입은 [지시1_비1]과 [지시1_1]로 1,436개와 666개의 토큰이 나타났으며, 다음으로는 [같다1](347개), [맞다1](336개), [무엇1](328개), [지시2_1](328개) [있다1](281개) 순으로 나타났다.

1.3 한국수어 말뭉치 기반 연구 사례

한국수어 말뭉치를 기반으로 한 사례 연구는 이형 동의어를 알아보기 위해 3가지로 나누어 실시하였다.

첫째, 한국수어에 나타난 색 이름의 다양한 형태를 알아보기 위해 ‘개별 수어의 다양한 표현(과제 번호 1)’ 과제로 유도한 언어 자료 중 2015년에 전사한 52분 17초 분량의 전사 자료를 기반으로 분석하였다. 유도 과제에서 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 1가지인데 비해, 혼합된 색을 나타내는 수어는 그 형태가 다양하게 나타났으며 [빨강1]+[파랑1]+[섞다1] 또는 [빨강1]+[약하다1]와 같이 복합어로 표현하였다. 또한 모든 색에서 공통적으로 지문자를 사용하였다.

연령대별로 보면, 모든 연령대에서 발생하는 타입은 [갈색1], [노랑1], [분홍1], [빨강1], [초록1A], [파랑1], [하양1], [포도1], [쥐1]이었으며, 60대 이상에서 발생하지 않는 타입은 [분홍1A·B·C], [보라1,1A·B], [자주색1], [주황1~8], [초록2]였으며, 이들 언어 제공자들은 분홍, 주황, 보라를 복합어로 표현하거나, 여성의 옷 색으로 표현한 언어 제공자도 있었다.

둘째, 한국수어에서 나타난 지명의 다양한 형태를 알아보기 위하여 총 34분 37초 분량의 ‘옛날 수어와 현재 수어’ 과제 영상을 분석하였다. 유도 과제의 특성 상 해당 타입의 발생 빈도는 중요하지 않아 언어 제공자 수로 분석하였으며, 분석한 결과는 현재 사용되고 있는 수어 형태가 아닐 수 있다. 분석한 총 파일은 16개로 언어 제공자는 32명이고, 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. 그러나 본 연구에서는 1명의 언어 제공자만이 언급한 속초(2), 원주(3), 창원(2), 포천(2), 통영(2), 화성(2)은 분석에서 제외하고(괄호 속은 나타난 타입의 개수임) 9곳의 지명 22개의 타입만을 분석하였다. 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 전사 완료한 총 260분 분량의 영상과 2015년 전사 완료한 영상 중 개별 수어의 ‘다양한 표현 과제(과제 번호 I)’와 ‘옛날 수어와 현재 수어 과제(과제번호 L)’을 제외한 146분 분량(총406분)의 영상으로 대상으로 분석하였다. 분석 절차는 다음과 같다. 이형 동의어 분석을 위해 수어의 글로스를 한국어 어휘를 차용하여 의미 바탕으로 정하였기 때문에 전사 자료 중 먼저 번호만 다른 같은 타입명을 검색하여 추출하였다. 다음으로 수어 모어 화자인 공동 연구원의 언어적 직관을 바탕으로 완전한 동의어로 보기 어려우며 맥락이나 언어 관계를 통해 유의어의 수준을 더 깊이 분석해야 할 필요가 있는 수어 어휘는 제외하였다. 그리하여 구체적인 대상을 나타내거나 문장 안에서 완전히 대체가 가능하다고 모어 화자가 인식한 수어만을 분석 대상으로 추출하였다. 그런 다음 다시 추출된 어휘 중 언어 제공자 수가 5명 미만으로 적은 수어를 제외하였다. 이에 따라 분석한 어휘는 샤워와 관련된 어휘 등 8개와 관련된 이형 동의어이다. 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.

1.4 한국수어 말뭉치 구축 계획

한국수어 말뭉치 구축 계획을 수립하기 위하여 농사회 전문가 5명과 음성 언어구축 전문가 5명의 자문 의견을 받았다. 자문 의견

과 해외 수어 연구 현황에 관한 워크숍 결과를 토대로 하여 향후 한국수어 말뭉치 구축 방안을 제시하였다.

첫째, 언어 자료 수집 방법은 목표 어휘 및 문법 항목을 명시하여 다양한 유도 과제별로 이끌어낼 수 있는 예상 어휘와 문법을 정리하여 자료 수집의 효율성을 높인다. 또한 자료 수집 방법을 다양화하고 해당 목표 어휘의 수집을 위한 연역적 방식의 언어 자료 수집 방법을 모색한다.

둘째, 기존에 있는 다양한 형태의 수어 교재 및 동영상 자료 등을 수집 하는 등 언어 자료 자원을 확대할 필요가 있다.

셋째, 언어 자료 수집을 위한 지역 수와 선정 기준은 전국 17개 시·도 지역마다 한국농아인협회가 있으므로 언어 제공자 선정과 언어 자료 수집을 위해 각 시·도 농아인협회와 협력하여, 언어 자료 수집 이전에 해당 지역에서 농인 중심의 워크숍을 개최하는 등 지역 홍보를 통해 한국수어 말뭉치의 언어 자료 수집의 수월성과 효율성을 꾀한다. 또한 언어 자료를 수집하기 위한 지역의 수는 예산과 참여 인력에 따라 달라지겠으나 수어의 경우는 동영상 촬영과 편집 그리고 전사가 필요하므로 1년에 1~2지역이 적정할 것으로 보인다. 또한 언어 자료 수집 때는 촬영한 영상의 자동 백업 장치의 설정 등 촬영 및 편집 전문가의 참여가 필요하다.

넷째, 언어 자료 수집을 위한 주체는 국민의 언어생활에 대한 기록과 자료 수집이라는 큰 틀에서 볼 때 국가가 주체가 되어야 하며 전문가와 농사회의 적극적인 협조를 통해 사업을 지속적으로 수행해가는 것이 가장 중요하다.

다섯째, 한국수어 말뭉치의 전사의 질은 전사자에 따라 크게 좌우된다. 전사자는 수어에 대한 메타 언어적인 시각을 갖고 수어 자료를 분석할 수 있어야 하므로 단기간에 전문성을 갖추 수 없다. 이에 향후 한국수어 말뭉치 구축 사업에서는 전사만을 담당할 전일제 전사자를 채용하여 전사를 실시할 필요가 있다. 또한 전사자의 양적 확대와 전사 능력의 질적 향상을 위해 전사자 양성 교육 실시도 한 방법이 될 수 있다.

여섯째, 전사의 품질 관리를 위해 전사자간 교차 검토와 정기적

인 전사 회의를 통한 의견 조정 과정을 시행하여 전사자간 혹은 전사자내 전사의 일치도와 정확성 등 전사의 질과 효율을 높이도록 한다.

일곱째, 향후 한국수어 말뭉치 구축의 효율성을 높이기 위한 방안으로 ‘전사 도우미 어플리케이션’의 개발을 제안한다. 이 어플리케이션은 연구자나 전사자가 수어 전사를 하다가 궁금한 내용이 있으면 정리를 해서 어플리케이션을 통해 농인들에게 물어보는 것이다. 그러면 회원 가입을 한 농인들이 대답을 해 주고, 대답에 참여한 농인 회원들은 점수를 얻게 되고 이 점수에 따라 상품 등을 지급하는 형태이다.

여덟째, 최근의 말뭉치 자료 구축 및 활용은 빅데이터 기반의 환경을 고려하여 개선되고 있으므로 한국수어 말뭉치도 이러한 기반 기술과의 결합이 필수적이다. 특히 영상 인식 기술의 놀라운 발전 성과를 도입하여 자동 전사와 데이터 구축을 하는 것이 지난한 한국수어 말뭉치 구축 작업의 한계를 극복하고 다음 단계로 발전하는 데 꼭 필요하다.

아홉째, 한국수어 말뭉치의 공개 대상과 범위는 농인과 청인, 연구자와 일반인 등 모든 이용자들에게 공개되어야 한다. 시스템의 안정화 등의 실행상의 이유로 일부 연구자들에게 제한적으로 공개할 수 있지만 최종적으로는 모든 이용자들에게 공개되어야 한다. 그러나 공개 시기는 한국수어 말뭉치는 영상 자료 중심이기 때문에 일반 말뭉치보다 고려할 사항이 더 많으므로 성급한 공개보다는 잘 정리된 시범 자료를 공개하는 방법을 제안한다.

마지막으로 한국수어 말뭉치는 한국수어 관련 전공자나 연구자 외에 한국수어를 모르는 일반인들이나 다른 학문 분야의 전문가들에게도 열려 있어야 하므로, 수어 영상에 대하여 한국어 번역을 제공해야 하며, 궁극적으로는 의미 중심의 검색 방법이나 행동 양상과 같은 것으로도 검색할 수 있도록 해야 한다.

1.5 한국수어 말뭉치 구축 지침 개정

2015년 개발하고 2016년 개정한 말뭉치 구축 지침을 한국수어 자료 통합 지원 시스템의 수정 보완 사항에 따라 수정하였으며, 향후 한국수어 말뭉치 구축의 효율성 확보를 위해 보다 정교화하였다. 주요 개정 내용은 다음과 같다.

첫째 2015년 개발하고 2016년에 개정한 지침서를 토대로 하였으며, 2017년 개발한 한국수어 자료 통합 지원 시스템을 기반으로 수정 보완한 내용을 추가하고 정교화하여 지침서로 개정하였다.

둘째, 언어 제공자에 대한 자료 수집을 위한 질문지를 한국 사정에 맞게 수정 보완하여 제시하였으며, 전사 품질 관리와 관련된 내용을 추가하였다.

셋째, 정비 과정에서 발견된 문제점을 보완할 수 있도록 수정된 전사 규칙을 제시하였다. 즉, 정비 과정에서 [지시]와 관련된 타입이 너무 세분화되어 전사 시 많은 혼란을 가져온 것이 발견되었다. 이에 전사의 효율과 오류 발생을 줄이기 위해 타입명을 16가지에서 8가지로 단순화시켰다. 또한 counting hand와 관련된 타입명과 신체 부위를 나타내는 경우의 타입명, 그리고 음운 변이의 형태에 대한 전사 방법을 수정하였다. 그 외 수어 합성어의 분절 방법에 대한 내용도 정교화하였으며, 생산적 수어에 대한 타입명도 정리하여 제시하였다.

2. 활용 방안 및 제언

2.1 연구의 활용 및 기대 효과

본 연구의 활용 방안 및 기대 효과를 다음의 다섯 가지로 요약한다.

첫째, 본 연구에서는 기구축된 한국수어 말뭉치를 한국수어 자료 통합 지원 시스템 안에 통합·구축하면서 기 전사된 내용을 정비하였다. 더불어 시스템을 통한 전사 작업을 실시하였다. 이러한 과정

에서 발생한 문제점을 해결하는 방안을 마련하고 이를 토대로 한국수어 말뭉치 구축 지침서를 수정 보완하였다. 이 개정된 한국수어 말뭉치 구축 지침서는 향후 보다 체계적이고 효율적인 한국수어 말뭉치 구축을 위한 지침을 제공할 것이다.

둘째, 아직은 그 양이 제한적이기는 하지만 2015년과 2016년에 이어 본 연구에서 전사한 약 256분 분량의 기본 전사와 번역은 향후 한국수어-한국어 사전 편찬이나 수어 연구에 필요한 실증적인 기초 자료들을 제공할 것이다. 또한 올해 전사한 한국수어 말뭉치에는 다양한 연령대의 언어 제공자들이 농학교 경험과 농사회 행사에 대한 이야기를 나누고 다양한 주제 영역에 대한 토론을 하는 언어 자료들이 포함되어 있다. 이는 농사회와 농문화에 대한 매우 값진 자료를 제공할 것이다.

셋째, 본 연구에서는 비록 제한적인 한국수어 말뭉치이지만 한국수어 말뭉치를 기반으로 한 수어 연구 사례를 제시하였는데, 이는 향후 한국수어 말뭉치 기반의 보다 과학적인 수어 연구 활성화에 기여할 것이다. 지금까지의 한국 수어 연구가 연구자의 경험이나 농인 화자의 언어적 직관에 의해 이루어져 온 것에 비해 본 연구에서 제시한 연구의 사례는 수어 연구의 한계를 극복하고 한 차원 더 높은 수준으로 발전할 수 있는 가능성을 보여 준 것이다. 특히 연구 목적에 따라 한국수어 말뭉치를 기반으로 어떻게 연구할 것인가에 대한 사례를 제시하므로 한국수어 말뭉치를 활용한 실증적이고 과학적인 한국수어 연구 방법론을 제공하였다.

넷째, 이렇게 한국수어 말뭉치 기반의 연구들은 향후 수어 교원 및 수화통역사 양성에 필요한 수어 교재 뿐 아니라, 농아동 및 학생을 위한 교재 개발에 필요한 실증적인 자료를 제공하는 등 질 높은 수어 교재 개발에 기여할 것이다. 예를 들어, 한국수어 말뭉치를 기반으로 한 이형동의어는 수어의 다양한 어휘 지도에 필요한 교재 개발 시에 활용될 수 있다.

다섯째, 음성 언어의 전사와는 달리 문자 체계가 없는 수어의 전사는 외국에서도 전사의 어려움이 많다. 이에 본 연구에서는 수어 전사의 품질 관리를 위해 전사 교육, 공동 전사 후 3차례의 전사

일치도 검사를 하고 그 결과 분석을 통한 1차례의 교육을 더 실시한 후에 본 전사를 실시하였다. 또한 전사 과정에서도 주 1회의 전사회의를 실시하였으며, 전사 진행 3개월째에 다시 한 번 전사 일치도 검사를 실시하므로 전사 신뢰도 확보하기 위한 노력을 하였다. 특히 본 연구에서는 전일제로 일하는 전문 전사자가 없는 우리나라 상황에서 전사의 신뢰도 확보를 위한 방법을 제시하므로 향후 보다 신뢰할 수 있는 전사 결과물을 도출할 수 있는 방안을 제공하였다.

여섯째, 본 연구에서는 농사회 전문가와 음성 언어 말뭉치 전문가의 자문 의견을 받고, 해외 수어 연구 현황에 관한 워크숍을 실시한 후 그 결과를 토대로 향후 한국수어 말뭉치 구축 방안을 제시하였다. 이는 장기적인 관점에서의 한국수어 말뭉치 구축을 위한 로드맵을 제공할 것이다.

2.2 제언

한국수어의 사용은 초창기에는 필요에 의해, 그 후 권리의 차원에서 이루어졌다. 그러나 이제는 우리나라의 수어의 사용도 한국수화언어법의 제정 및 시행과 더불어 언어 자원의 확보라는 차원으로 한걸음 진보하였다. 특히 국립국어원의 특수언어진흥과가 만들어지고 전문 인력을 확보하여 한국수어 자원의 보존과 발전을 위한 국가적 차원의 정책과 연구가 이루어지고 있는 것은 매우 환영할 일이다. 이에 본 연구 결과를 바탕으로 하여 한국수어 발전의 가장 기본적인 토대가 되는 한국수어 말뭉치 구축을 위해 다음과 같이 제언하고자 한다.

첫째, 2015년 서울 지역 농인 60명을 대상으로 90시간 분량의 언어 자료를 수집한 이후, 부분적인 전사와 함께 한국수어 자료 통합 지원 시스템의 구축 등이 이루어졌으나, 더 이상의 언어 자료 수집이 이루어지지 않았다. 따라서 한국수어의 발전과 보존을 위해 더 늦기 전에 서둘러 전국 규모의 언어 자료 수집이 필요하다.

둘째, 언어 수집과 더불어 전사 작업은 한국수어 말뭉치 구축에

서 매우 중요하다. 따라서 수집한 언어 자료를 지속적으로 전사하고 번역할 필요가 있다. 그러나 현재와 같이 언어학적 배경이 부족한 다수의 전사자를 매번 교육시켜 전사에 투입하는 것은 전사 품질을 보증하는 데에 어려움을 초래할 수밖에 없다. 따라서 국립국어원 특수언어진흥과에서 전사할 수 있는 인력을 확보하여 지속적인 전사 작업을 해야 할 것이다.

셋째, 한국수어 말뭉치 구축은 한국수어 사용자인 농인과 농사회의 적극적인 참여 및 협조가 필요하다. 따라서 언어 자료 수집에서도 각 지역의 농아인협회가 주체가 되어야 하며, 향후 한국수어 말뭉치 구축의 효율성을 높이기 위해 ‘전사 도우미 어플리케이션’을 개발을 제안한다. 이 어플리케이션은 한국수어 말뭉치 홍보 뿐 아니라, 스마트폰을 사용하는 농인들의 한국수어 말뭉치에 대한 관심 및 참여를 끌어낼 수 있을 것이다.

VIII. 참고문헌

- Aramburo, A. (1989). Sociolinguistic Aspects of the Black Deaf Community. In: Lucas, C.(ed.), The Sociolinguistics of the Deaf Community. San Diego, CA: Academic Press, 103-121.
- Bayley, R. / Lucas, C. / Rose, M. (2000). Variation in American Sign Language: The Case of deaf. In: Journal of Sociolinguistics, 4, 81-107.
- Blanck, D. / Bleicken, J. / Hanke, T. / Hanß, A. / Hofmann, I. / Jeziorski, O. / König, L. / König, S. / Konrad, R. / Langer, G. / Rathmann, C. / Salden, U. / Wagner, S. (2016). „DGS Korpus - Aktueller Projektstand 2016“. Poster presented at the „Akademientag“ on the 18. 05.2016 in Hamburg.
- Boudreault, P. / Mayberry, R. (2006). Grammatical Processing in American Sign Language: Age of First Language Acquisition Effects in Relation to Syntactic Structure. In: Language and Cognitive Processes 21, 608-635.
- Clark, L., E. / Grosjean, F. (1982). Sign recognition processes in American Sign Language: The effect of context. Language and Speech 25(4), 325-340.
- Deuchar, M. (1981). Variation in British Sign Language. In: Woll, B. / Kyle, J. / Deuchar, M. (eds.), Perspectives on British Sign Language and Deafness. London: Croom Helm, 109-119.
- Fenlon, J. / Cormier, K. / Schembri, A. (2015). Building BSL SignBank: The Lemma Dilemma Revisited. International Journal of Lexicography, 28(2), 169-206.
- Grosjean, F. (1981). Sign and word recognition: A first comparison. Sign Language Studies, 32, 98-125.

- Hanke, T. (2016). "Towards a Visual Sign Language Corpus Linguistics". In: Efthimiou, E. et al. (Eds.): Workshop Proceedings. 7th Workshop on the Representation and Processing of Sign Languages: Corpus Mining. Language Resources and Evaluation Conference (LREC), Portorož, Slovenia, 28 May 2016. ELRA. S. 89-92.
- Hanke, T. / Konrad, R. / Langer, G. / Müller, A. / Wähl, S. (2017). "Detecting Regional and Age Variation in a Growing Corpus of DGS". Poster presented at the Workshop "Corpus-based Approaches to Sign Language Linguistics: Into the Second Decade". 24.07.2017.
- Jantunen, T. (2015). How long is the sign? *Linguistics*, 53(1), 93-124.
- Johnston, T. (2003). Language Standardization and Signed Language Dictionaries, *Sign Language Studies*, 3(4), 341-468.
- Johnston, T. / Schembri, A. (1999). On Defining Lexeme in a Signed Language. *Sign Language and Linguistics*, 2(2), 115-185.
- Johnston, T. / Schembri, A. (2007). *Australian Sign Language (Auslan): An Introduction to Sign Language Linguistics*. Cambridge: Cambridge University Press.
- Langer, G./Köig, S./Konrad, R.(2007). 'Hintergrundinformationen' in Konrad, R./König, S./Schwarz, A./Hanke, T./Prillwitz, S.(Eds). *Fachgebardenlexikon Gesundheit und Pflege*. Seedorf: Signum.
- Langer, G. / Troelsgård, T. / Kristoffersen, J. / Konrad, R. / Hanke, T. / König, S. (2016). "Designing a Lexical Database for a Combined Use of Corpus Annotation

- and Dictionary Editing". In: Efthimiou, E. et al. (Eds.): Workshop Proceedings. 7th Workshop on the Representation and Processing of Sign Languages: Corpus Mining. Language Resources and Evaluation Conference (LREC), Portorož, Slovenia, 28 May 2016. ELRA, 143-152.
- Le Master, B. / Dwyer, J. (1991). Knowing and Using Female and Male Signs in Dublin. In: Sign Language Studies 73, 361-369.
- Leeson, L. /Grehan, C. (2004). To the Lexicon and Beyond: The Effect of Gender on Variation in Irish Sign Language. In: Van Herreweghe, M. / Vermeerbergen, M. (eds.), To the Lexicon and Beyond: Sociolinguistics in European Deaf Communities. Washington, DC: Gallaudet University Press, 39-73.
- Lucas, C. (1995). Sociolinguistic Variation in ASL: The Case of DEAF. In: Lucas, C. (Ed.), Sociolinguistics in Deaf Communities. Washington. DC: Gallaudet University Press, 325.
- Lucas, C. / Bayley, R. / Reed, R. / Wulf, A. (2001). Lexical Variation in African American and White Signing. In: American Speech 76, 339-360.
- Lucas, C. / Bayley, R. /Valli, C. (2001). Sociolinguistic Variation in American Sign Language. Washington, DC: Gallaudet University Press.
- McCaskill, C. / Lucas, C. /Bayley, R. / Hill, J. (2011). The Hidden Treasure of Black ASL: Its History and Structure. Washington, DC: Gallaudet University Press.
- McKee, D. / McKee, R. /Major, G. (2011). Numeral Variation in New Zealand Sign Language. In: Sign Language Studies 12, 72-97.

- Padden, C. / Gunsauls, D. (2003). How the Alphabet Came to Be Used in a Sign Language. In: Sign Language Studies 4, 10-33.
- Schembri, A. / Johnston, T. (2012). Sociolinguistic Aspects of Variation and Change. In: Pfau, R. / Steinbach, M. / Woll, B. (Eds), Sign Language: An International Handbook. Berlin: Walter de Gruyter, 788-816.
- Schembri, A. / Johnston, T. (2013). Sociolinguistic Variation and Change in Sign Languages. In: Bayley, R. / Cameron, R. / Lucas, C. (Eds.), Oxford Handbook of Sociolinguistics (pp. 503-524). Oxford: Oxford University Press.
- Woll, B. (1983). Historical Change in British Sign Language. Unpublished Report, Deaf Studies Unit, University of Bristol
- Woodward, J. / DeSantis, S. (1977). Two to One It Happens: Dynamic Phonology in Two Sign Languages. In: Sign Language Studies, 17, 329-346.
- Woodward, J. / Erting, C. / Oliver, S. (1976). Facing and Hand(l)ing Variation in American Sign Language. In: Sign Language Studies, 10, 43-52.

<부록 1> 과제별 연결 실패 건수

| 번호 | 과제명 | 파일 번호 | 파일명 | 발생 건수 |
|----|------------------------|----------|-------------------------------|----------|
| 1 | 수어 이름 말하기 (918건) | 1 | 서울 지역 1번 쌍 A- 수어 이름 말하기 전사 영상 | 20 |
| | | 2 | 서울 지역 2번 쌍 A-수어 이름 말하기 전사 영상 | 14 |
| | | 3 | 서울 지역 3번 쌍 A-수어 이름 말하기 전사 영상 | 29 |
| | | 4 | 서울 지역 4번 쌍 A-수어 이름 말하기 전사 영상 | 12 |
| | | 5 | 서울 지역 5번 쌍 A-수어 이름 말하기 전사 영상 | 23 |
| | | 6 | 서울 지역 6번 쌍 A-수어 이름 말하기 전사 영상 | 43 |
| | | 7 | 서울 지역 7번 쌍 A-수어 이름 말하기 전사 영상 | 14 |
| | | 8 | 서울 지역 8번 쌍 A-수어 이름 말하기 전사 영상 | 55 |
| | | 9 | 서울 지역 9번 쌍 A-수어 이름 말하기 전사 영상 | 43 |
| | | 10 | 서울 지역 10번 쌍 A-수어 이름 말하기 전사 영상 | 43 |
| | | 11 | 서울 지역 11번 쌍 A-수어 이름 말하기 전사 영상 | 32 |
| | | 12 | 서울 지역 12번 쌍 A-수어 이름 말하기 전사 영상 | 49 |
| | | 13 | 서울 지역 13번 쌍 A-수어 이름 말하기 전사 영상 | 19 |
| | | 14 | 서울 지역 14번 쌍 A-수어 이름 말하기 전사 영상 | 61 |
| | | 15 | 서울 지역 15번 쌍 A-수어 이름 말하기 전사 영상 | 20 |
| | | 16 | 서울 지역 16번 쌍 A-수어 이름 말하기 전사 영상 | 5 |
| | | 17 | 서울 지역 17번 쌍 A-수어 이름 말하기 전사 영상 | 42 |
| | | 18 | 서울 지역 18번 쌍 A-수어 이름 말하기 전사 영상 | 40 |
| | | 19 | 서울 지역 19번 쌍 A-수어 이름 말하기 전사 영상 | 21 |
| | | 20 | 서울 지역 20번 쌍 A-수어 이름 말하기 전사 영상 | 4 |
| | | 21 | 서울 지역 21번 쌍 A-수어 이름 말하기 전사 영상 | 5 |
| | | 22 | 서울 지역 22번 쌍 A-수어 이름 말하기 전사 영상 | 73 |
| | | 23 | 서울 지역 23번 쌍 A-수어 이름 말하기 전사 영상 | 58 |

| 번호 | 과제명 | 파일 번호 | 파일명 | 발생 건수 |
|----|---|----------|---|----------|
| 1 | 수어 이름 말하기 (918건) | 24 | 서울 지역 24번 쌍 A-수어 이름 말하기 전사 영상 | 72 |
| | | 25 | 서울 지역 25번 쌍 A-수어 이름 말하기 전사 영상 | 11 |
| | | 26 | 서울 지역 26번 쌍 A-수어 이름 말하기 전사 영상 | 10 |
| | | 27 | 서울 지역 27번 쌍 A-수어 이름 말하기 전사 영상 | 15 |
| | | 28 | 서울 지역 29번 쌍 A-수어 이름 말하기 전사 영상 | 50 |
| | | 29 | 서울 지역 30번 쌍 A-수어 이름 말하기 전사 영상 | 26 |
| | | 30 | 서울 지역 31번 쌍 A-수어 이름 말하기 전사 영상 | 9 |
| 2 | 수어를 보고 수어로 다시 이야기 하기 (넥타이) (1458건) | 1 | 서울 지역 1번 쌍 D -수어를 보고 수어로 다시 이야기하기 (넥타이) 전사 영상 | 100 |
| | | 2 | 서울 지역 3번 쌍 D -수어를 보고 수어로 다시 이야기하기 (넥타이) 전사 영상 | 43 |
| | | 3 | 서울 지역 4번 쌍 D -수어를 보고 수어로 다시 이야기하기 (넥타이) 전사 영상 | 63 |
| | | 4 | 서울 지역 5번 쌍 D -수어를 보고 수어로 다시 이야기하기 (넥타이) 전사 영상 | 46 |
| | | 5 | 서울 지역 6번 쌍 D -수어를 보고 수어로 다시 이야기하기 (넥타이) 전사 영상 | 47 |
| | | 6 | 서울 지역 9번 쌍 D -수어를 보고 수어로 다시 이야기하기 (넥타이) 전사 영상 | 75 |
| | | 7 | 서울 지역 10번 쌍 D -수어를 보고 수어로 다시 이야기하기 (넥타이) 전사 영상 | 102 |
| | | 8 | 서울 지역 11번 쌍 D -수어를 보고 수어로 다시 이야기하기 (넥타이) 전사 영상 | 41 |
| | | 9 | 서울 지역 12번 쌍 D -수어를 보고 수어로 다시 이야기하기 (넥타이) 전사 영상 | 78 |
| | | 10 | 서울 지역 13번 쌍 D -수어를 보고 수어로 다시 이야기하기 (넥타이) 전사 영상 | 45 |
| | | 11 | 서울 지역 14번 쌍 D -수어를 보고 수어로 다시 이야기하기 (넥타이) 전사 영상 | 94 |
| | | 12 | 서울 지역 15번 쌍 D -수어를 보고 수어로 다시 이야기하기 (넥타이) 전사 영상 | 87 |
| | | 13 | 서울 지역 16번 쌍 D -수어를 보고 수어로 다시 이야기하기 (넥타이) 전사 영상 | 22 |

| 번호 | 과제명 | 파일 번호 | 파일명 | 발생 건수 |
|----|---|----------|---|----------|
| 2 | 수어를 보고 수어로 다시 이야기 하기 (넥타이) (1458건) | 14 | 서울 지역 17번 쌍 D -수어를 보고 수어로 다시 이야기하기 (넥타이) 전사 영상 | 25 |
| | | 15 | 서울 지역 18번 쌍 D -수어를 보고 수어로 다시 이야기하기 (넥타이) 전사 영상 | 25 |
| | | 16 | 서울 지역 19번 쌍 D -수어를 보고 수어로 다시 이야기하기 (넥타이) 전사 영상 | 44 |
| | | 17 | 서울 지역 20번 쌍 D -수어를 보고 수어로 다시 이야기하기 (넥타이) 전사 영상 | 28 |
| | | 18 | 서울 지역 21번 쌍 D -수어를 보고 수어로 다시 이야기하기 (넥타이) 전사 영상 | 36 |
| | | 19 | 서울 지역 22번 쌍 D -수어를 보고 수어로 다시 이야기하기 (넥타이) 전사 영상 | 66 |
| | | 20 | 서울 지역 23번 쌍 D -수어를 보고 수어로 다시 이야기하기 (넥타이) 전사 영상 | 127 |
| | | 21 | 서울 지역 24번 쌍 D -수어를 보고 수어로 다시 이야기하기 (넥타이) 전사 영상 | 53 |
| | | 22 | 서울 지역 26번 쌍 D -수어를 보고 수어로 다시 이야기하기 (넥타이) 전사 영상 | 67 |
| | | 23 | 서울 지역 27번 쌍 D -수어를 보고 수어로 다시 이야기하기 (넥타이) 전사 영상 | 73 |
| | | 24 | 서울 지역 29번 쌍 D -수어를 보고 수어로 다시 이야기하기 (넥타이) 전사 영상 | 34 |
| | | 25 | 서울 지역 30번 쌍 D -수어를 보고 수어로 다시 이야기하기 (넥타이) 전사 영상 | 15 |
| | | 26 | 서울 지역 31번 쌍 D -수어를 보고 수어로 다시 이야기하기 (넥타이) 전사 영상 | 22 |
| 3 | 표지판 (5763건) | 1 | 서울 지역 3번 쌍 F-표지판 전사 영상 | 242 |
| | | 2 | 서울 지역 4번 쌍 F-표지판 전사 영상 | 320 |
| | | 3 | 서울 지역 5번 쌍 F-표지판 전사 영상 | 616 |
| | | 4 | 서울 지역 6번 쌍 F-표지판 전사 영상 | 563 |
| | | 5 | 서울 지역 7번 쌍 F-표지판 전사 영상 | 356 |
| | | 6 | 서울 지역 9번 쌍 F-표지판 전사 영상 | 246 |
| | | 7 | 서울 지역 11번 쌍 F-표지판 전사 영상 | 203 |

| 번호 | 과제명 | 파일 번호 | 파일명 | 발생 건수 |
|----|-----------------------------------|----------|----------------------------------|----------|
| 3 | 표지판 (5763건) | 8 | 서울 지역 12번 쌍 F-표지판 전사 영상 | 803 |
| | | 9 | 서울 지역 13번 쌍 F-표지판 전사 영상 | 255 |
| | | 10 | 서울 지역 15번 쌍 F-표지판 전사 영상 | 392 |
| | | 11 | 서울 지역 16번 쌍 F-표지판 전사 영상 | 166 |
| | | 12 | 서울 지역 18번 쌍 F-표지판 전사 영상 | 323 |
| | | 13 | 서울 지역 19번 쌍 F-표지판 전사 영상 | 303 |
| | | 14 | 서울 지역 20번 쌍 F-표지판 전사 영상 | 164 |
| | | 15 | 서울 지역 21번 쌍 F-표지판 전사 영상 | 52 |
| | | 16 | 서울 지역 23번 쌍 F-표지판 전사 영상 | 130 |
| | | 17 | 서울 지역 24번 쌍 F-표지판 전사 영상 | 75 |
| | | 18 | 서울 지역 25번 쌍 F-표지판 전사 영상 | 61 |
| | | 19 | 서울 지역 26번 쌍 F-표지판 전사 영상 | 109 |
| | | 20 | 서울 지역 27번 쌍 F-표지판 전사 영상 | 83 |
| | | 21 | 서울 지역 30번 쌍 F-표지판 전사 영상 | 223 |
| | | 22 | 서울 지역 31번 쌍 F-표지판 전사 영상 | 78 |
| 4 | 특정 사건 묘사 (1010건) | 1 | 서울 지역 1번 쌍 G-특정 사건 묘사 전사 영상 | 363 |
| | | 2 | 서울 지역 6번 쌍 G-특정 사건 묘사 영상 | 404 |
| | | 3 | 서울 지역 7번 쌍 G-특정 사건 묘사 전사 영상 | 198 |
| | | 4 | 서울 지역 8번 쌍 G-특정 사건 묘사 전사 영상 | 45 |
| 5 | 개별 수어의 다양한 표현 (2532건) | 1 | 서울 지역 1번 쌍 I-개별 수어의 다양한 표현 전사 영상 | 25 |
| | | 2 | 서울 지역 3번 쌍 I-개별 수어의 다양한 표현 전사 영상 | 179 |
| | | 3 | 서울 지역 4번 쌍 I-개별 수어의 다양한 표현 전사 영상 | 26 |
| | | 4 | 서울 지역 7번 쌍 I-개별 수어의 다양한 표현 전사 영상 | 93 |

| 번호 | 과제명 | 파일 번호 | 파일명 | 발생 건수 |
|----|-----------------------------------|----------|-----------------------------------|----------|
| 5 | 개별 수어의 다양한 표현 (2532건) | 5 | 서울 지역 8번 쌍 I-개별 수어의 다양한 표현 전사 영상 | 183 |
| | | 6 | 서울 지역 9번 쌍 I-개별 수어의 다양한 표현 전사 영상 | 21 |
| | | 7 | 서울 지역 10번 쌍 I-개별 수어의 다양한 표현 전사 영상 | 128 |
| | | 8 | 서울 지역 11번 쌍 I-개별 수어의 다양한 표현 전사 영상 | 289 |
| | | 9 | 서울 지역 12번 쌍 I-개별 수어의 다양한 표현 전사 영상 | 160 |
| | | 10 | 서울 지역 13번 쌍 I-개별 수어의 다양한 표현 전사 영상 | 55 |
| | | 11 | 서울 지역 15번 쌍 I-개별 수어의 다양한 표현 전사 영상 | 188 |
| | | 12 | 서울 지역 16번 쌍 I-개별 수어의 다양한 표현 전사 영상 | 106 |
| | | 13 | 서울 지역 17번 쌍 I-개별 수어의 다양한 표현 전사 영상 | 26 |
| | | 14 | 서울 지역 18번 쌍 I-개별 수어의 다양한 표현 전사 영상 | 88 |
| | | 15 | 서울 지역 19번 쌍 I-개별 수어의 다양한 표현 전사 영상 | 25 |
| | | 16 | 서울 지역 20번 쌍 I-개별 수어의 다양한 표현 전사 영상 | 66 |
| | | 17 | 서울 지역 21번 쌍 I-개별 수어의 다양한 표현 전사 영상 | 76 |
| | | 18 | 서울 지역 22번 쌍 I-개별 수어의 다양한 표현 전사 영상 | 164 |
| | | 19 | 서울 지역 23번 쌍 I-개별 수어의 다양한 표현 전사 영상 | 215 |
| | | 20 | 서울 지역 25번 쌍 I-개별 수어의 다양한 표현 전사 영상 | 134 |
| | | 21 | 서울 지역 26번 쌍 I-개별 수어의 다양한 표현 전사 영상 | 87 |
| | | 22 | 서울 지역 27번 쌍 I-개별 수어의 다양한 표현 전사 영상 | 54 |
| | | 23 | 서울 지역 30번 쌍 I-개별 수어의 다양한 표현 전사 영상 | 73 |
| | | 24 | 서울 지역 31번 쌍 I-개별 수어의 다양한 표현 전사 영상 | 71 |
| 6 | 옛날 수어와 현재 수어 (1659건) | 1 | 서울 지역 1번 쌍 L-옛날 수어와 현재 수어 전사 영상 | 139 |
| | | 2 | 서울 지역 4번 쌍 L-옛날 수어와 현재 수어 전사 영상 | 10 |
| | | 3 | 서울 지역 7번 쌍 L-옛날 수어와 현재 수어 전사 영상 | 53 |
| | | 4 | 서울 지역 8번 쌍 L-옛날 수어와 현재 수어 전사 영상 | 87 |

| 번호 | 과제명 | 파일 번호 | 파일명 | 발생 건수 |
|----|-------------------------------|----------|----------------------------------|----------|
| 6 | 옛날 수어와 현재 수어 (1659건) | 5 | 서울 지역 9번 쌍 L-옛날 수어와 현재 수어 전사 영상 | 28 |
| | | 6 | 서울 지역 10번 쌍 L-옛날 수어와 현재 수어 전사 영상 | 24 |
| | | 7 | 서울 지역 11번 쌍 L-옛날 수어와 현재 수어 전사 영상 | 106 |
| | | 8 | 서울 지역 12번 쌍 L-옛날 수어와 현재 수어 전사 영상 | 96 |
| | | 9 | 서울 지역 13번 쌍 L-옛날 수어와 현재 수어 전사 영상 | 58 |
| | | 10 | 서울 지역 14번 쌍 L-옛날 수어와 현재 수어 전사 영상 | 71 |
| | | 11 | 서울 지역 15번 쌍 L-옛날 수어와 현재 수어 전사 영상 | 52 |
| | | 12 | 서울 지역 17번 쌍 L-옛날 수어와 현재 수어 전사 영상 | 45 |
| | | 13 | 서울 지역 19번 쌍 L-옛날 수어와 현재 수어 전사 영상 | 52 |
| | | 14 | 서울 지역 27번 쌍 L-옛날 수어와 현재 수어 전사 영상 | 733 |
| | | 15 | 서울 지역 29번 쌍 L-옛날 수어와 현재 수어 전사 영상 | 74 |
| | | 16 | 서울 지역 30번 쌍 L-옛날 수어와 현재 수어 전사 영상 | 31 |

<부록 2> 타입 삭제, 타입 수정, 새로운 타입 등록 목록

1. 삭제한 타입 목록

| 번호 | 타입명 | 결정 사항 |
|----|-----------------|------------------|
| 1 | [데려오다1]/[이끌다1] | [이끌다1] 삭제 |
| 2 | [어긋나다1]/[안통하다1] | [어긋나다1] 삭제 |
| 3 | [오다1]/[오다2] | [오다2] 삭제 |
| 4 | [때리다1]/[때리다2] | [때리다1]/[때리다2] 삭제 |
| 5 | [질문1]/[물어보다1] | [물어보다1] 삭제 |
| 6 | [때2]/[때1A] | [때1A] 삭제 |
| 7 | [세계1]/[나라1] | [세계1] 삭제 |
| 8 | [경고1]/[선언1] | [경고1] 삭제 |
| 9 | [사용1]/[낭비1] | [낭비1] 삭제 |
| 10 | [기쁘다1]/[재미1] | [기쁘다1] 삭제 |
| 11 | [연결1]/[관련1] | [관련1] 삭제 |
| 12 | [고치다1]/[바꾸다1] | [고치다1] 삭제 |
| 13 | [이야기하다1]/[뉴스1] | [뉴스1] 삭제 |
| 14 | [몰래1]/[몰래2] | [몰래2] 삭제 |
| 15 | [흡수1] / [빨아먹다1] | [흡수1] 삭제 |
| 16 | [후지다1]/[나쁘다2] | [후지다1] 삭제 |
| 17 | [바람1]/[가을1] | [가을1] 삭제 |
| 18 | [답다1]/[여름1] | [여름1] 삭제 |
| 19 | [편안하다1]/[안전1] | [안전1] 삭제 |
| 20 | [자꾸1]/[또1] | [자꾸1] 삭제 |
| 21 | [돌보이다1]/[독특1] | [돌보이다1] 삭제 |
| 22 | [넘어가다1]/[패스1] | [패스1] 삭제 |

| 번호 | 타입명 | 결정 사항 |
|----|----------------|----------------------|
| 23 | [약하다1]/[피곤하다1] | [피곤하다1] 삭제 |
| 24 | [보다6]/[보다61] | [보다61] 삭제 |
| 25 | [원인1]/[동기1] | [동기1] 삭제 |
| 26 | [몇1]/[얼마1] | [얼마1] 삭제 |
| 27 | [아하1]/[아하5] | [아하1] 삭제 |
| 28 | [비우다1]/[없다10] | [비우다1] 삭제 |
| 29 | [바닥1]/[평평하다1] | [평평하다1] 삭제 |
| 30 | [가만있다1]/[두다1A] | [가만있다1] 삭제 |
| 31 | [숫자1]/[번호1] | [번호1] 삭제 |
| 32 | [얼마1]/[몇1] | [얼마1], [얼마1A] 삭제 |
| 33 | [모으다1]/[합하다1] | [합하다1] 삭제 |
| 34 | [하다1_] | [하다1_] 삭제 |
| 35 | [여러명1] | [여러명1] 삭제 |
| 36 | [진짜1] | [진짜1] 삭제 |
| 37 | [소극1] | [소극1] 삭제 |
| 38 | [사백만원1] | [사백만원1] 삭제 |
| 39 | [보다17]/[보다7] | [보다17] 삭제 |
| 40 | [길1]/[도로1] | [도로1] 삭제 |
| 41 | [비키다1]/[제쳐두다1] | 제스처-> 타입 삭제 |
| 42 | [세번째2] | Counting hand-> 타입삭제 |
| 43 | [몰염치1]/[뻔뻔하다1] | [몰염치1],[몰염치2] 삭제 |
| 44 | [정도1]/[적당하다1] | [정도1] 삭제 |

2. 수정한 타입 목록

| 번호 | 타입명 | 결정 사항 |
|----|----------------|-------------------------------------|
| 1 | [탈출1] | [도망2] 수정 |
| 2 | [운명1] | [우연1] 수정 |
| 3 | [부르다3] | [보내다1] 수정 |
| 4 | [문자1] | [문장1] 수정 |
| 5 | [가져가다1] | [가져오다1] 수정 |
| 6 | [배수자1] | [백수1] 수정 |
| 7 | [한글1] | [글1] |
| 8 | [관련1A] | [연결1A] 수정 |
| 9 | [삼성농학교1]/[삼성1] | [삼성1] 수정 |
| 10 | [모임1]/[단체1] | [모임1] 삭제 [단체1] 영상-> [모임1] 수정 |
| 11 | [자르다3]/[끊다1] | [자르다3] 삭제 [끊다1] ->[자르다1] 수정 |
| 12 | [비키다1]/[제쳐두다1] | [제스처]로 병합 |
| 13 | [세번째2] | ‘counting hand’로 병합 |
| 14 | [몰염치] | [몰염치1]/[몰염치2] -> 뻔뻔하다2, 뻔뻔하다3 수정 |
| 15 | [기숙사1] | [잠1] 수정 |
| 16 | [기숙사2] | [기숙사1] 수정 |

3. 새롭게 등록한 타입 목록

| 번호 | 타입명 | 번호 | 타입명 |
|----|---------------|----|---------|
| 1 | 제스처 | 21 | 기숙사1 |
| 2 | 농2 | 22 | 십칠일1 |
| 3 | 섞다1 | 23 | 지시2그림 |
| 4 | 대충1 | 24 | 구형1A_면 |
| 5 | 생존1 | 25 | 구형1A_선 |
| 6 | 넘다1 | 26 | 마음대로하다1 |
| 7 | 공범1 | 27 | 이형1A |
| 8 | 높다2 | 28 | 앰블런스1 |
| 9 | 원2 | 29 | 초록2 |
| 10 | 새롭다1 | 30 | 지시2호 |
| 11 | 따로2 | 31 | 이발2 |
| 12 | 고등2 | 32 | 뜻밖에1 |
| 13 | counting hand | 33 | 고추1 |
| 14 | 낱두다1 | 34 | 미원18 |
| 15 | 뿌리다6 | 35 | 망신2 |
| 16 | 주먹형1 | 36 | 팽권2 |
| 17 | 쌓다2 | 37 | 없다11 |
| 18 | 일형3 | 38 | 중양1 |
| 19 | 구형1 | 39 | 백년1 |
| 20 | 포악하다1 | 40 | 갔다오다1 |

| 번호 | 타입명 | 번호 | 타입명 |
|----|-----------|----|-------|
| 41 | 오래3 | 61 | 백이십1 |
| 42 | 농아인협회중앙회1 | 62 | 기타등등1 |
| 43 | 보청기1 | 63 | 진해1 |
| 44 | 합덕1 | 64 | 한번에1 |
| 45 | 머리카락1 | 65 | 열살1 |
| 46 | 종합1 | 66 | 수화3 |
| 47 | 구형4A | 67 | 통하다1 |
| 48 | 쌀1 | 68 | 꼭소2 |
| 49 | 월요일1 | 69 | 하나님1 |
| 50 | 눈원1 | 70 | 아이고2 |
| 51 | 시작2 | 71 | 감동1 |
| 52 | 미국수화 | 72 | 명예1 |
| 53 | 국제수화 | 73 | 수출1 |
| 54 | 일본수화 | 74 | 반전1 |
| 55 | 사계절1 | 75 | 키1 |
| 56 | 예3 | 76 | 몸부림1 |
| 57 | 독특2 | 77 | 상상하다1 |
| 58 | 삼성1 | 78 | 성경책1 |
| 59 | 독수리1 | 79 | 시끄럽다1 |
| 60 | 가난2 | 80 | 목마르다2 |

| 번호 | 타입명 | 번호 | 타입명 |
|-----|--------|-----|---------|
| 81 | 일곱시1 | 101 | 풍물1 |
| 82 | 안산1 | 102 | 둘중하나1 |
| 83 | 이천십사1 | 103 | 이천1 |
| 84 | 양심1 | 104 | 팔천1 |
| 85 | SN_김현기 | 105 | 팔만1 |
| 86 | 술2 | 106 | 십만1 |
| 87 | 이갈다1 | 107 | 육만1 |
| 88 | 벼락1 | 108 | 머리원1 |
| 89 | 때리다1 | 109 | 마음에안들다1 |
| 90 | 때리다2 | 110 | 간판1 |
| 91 | 보리1 | 111 | 바둑1 |
| 92 | 미분1 | 112 | 다섯번째1 |
| 93 | 근육1 | 113 | 삼십오년후1 |
| 94 | 목메달다1 | 114 | 목사1 |
| 95 | 선물1 | 115 | 실패2 |
| 96 | SN_김영삼 | 116 | 결석1 |
| 97 | 몇개월1 | 117 | 불교1 |
| 98 | 그림자1 | 118 | 많다5 |
| 99 | 이십살1 | 119 | SN_이재란 |
| 100 | 몇년전1 | 120 | 오층1 |

| 번호 | 타입명 | 번호 | 타입명 |
|-----|-------|-----|-------|
| 121 | 해고1 | 141 | 노아1 |
| 122 | 재량1 | 142 | 십사만1 |
| 123 | 권위1 | 143 | 이스라엘1 |
| 124 | 잔업1 | 144 | 사천1 |
| 125 | 열한시1 | 145 | 저축1 |
| 126 | 열시1 | 146 | 유혹1 |
| 127 | 성추행1 | 147 | 로마1 |
| 128 | 삼십오년1 | 148 | 천만1 |
| 129 | 당황2 | 149 | 폭포1 |
| 130 | 싫다4 | 150 | 한시1 |
| 131 | 하늘2 | 151 | 불가능4 |
| 132 | 잊다2 | 152 | 반복1 |
| 133 | 동사무소1 | 153 | 팔월1 |
| 134 | 사라지다3 | 154 | 사십년1 |
| 135 | 무례1 | 155 | 아이디어2 |
| 136 | 과거2 | 156 | 추첨1 |
| 137 | 매주1 | 157 | 삼십만1 |
| 138 | 아시아2 | 158 | 신분증1 |
| 139 | 회개1 | 159 | 뺨다1 |
| 140 | 무시4 | 160 | 대들다1 |

| 번호 | 타입명 | 번호 | 타입명 |
|-----|-------|-----|-------|
| 161 | 제정1 | 181 | 배탈1 |
| 162 | 봉고차1 | 182 | 경기도1 |
| 163 | 유도1 | 183 | 이천삼백1 |
| 164 | 날씬하다1 | 184 | 유혹2 |
| 165 | 시키다2 | 185 | 의심3 |
| 166 | 몇년후1 | 186 | 우물1 |
| 167 | 천안1 | 187 | 평택1 |
| 168 | 의정부1 | 188 | 소심1 |
| 169 | 십오일1 | 189 | 기숙사2 |
| 170 | 청음회관1 | 190 | 목메달다2 |
| 171 | 충청도1 | 191 | 되돌리다2 |
| 172 | 칠십일1 | 192 | 일요일1 |
| 173 | 할수없다1 | 193 | 우편1 |
| 174 | 화내다2 | 194 | 성냥1 |
| 175 | 새벽1 | 195 | 스캐너1 |
| 176 | 십일2 | 196 | 포토샵1 |
| 177 | 십일일1 | 197 | 끔찍하다1 |
| 178 | 돈내기1 | 198 | 팔십1 |
| 179 | 백2 | 199 | 칠십이1 |
| 180 | 수박1 | 200 | 부부싸움1 |

| 번호 | 타입명 | 번호 | 타입명 |
|-----|--------|-----|--------|
| 201 | 경상북도1 | 221 | 이천일1 |
| 202 | 늦었다2 | 222 | 구십칠1 |
| 203 | 달1 | 223 | 구십사1 |
| 204 | 미남1 | 224 | 힘2 |
| 205 | 벌다4 | 225 | 빌리다2 |
| 206 | 애화1 | 226 | 다행2 |
| 207 | 반2 | 227 | 말레이시아1 |
| 208 | 칠년1 | 228 | 태국1 |
| 209 | 갈로뎃대학1 | 229 | 배부르다1 |
| 210 | 미국2 | 230 | 대충보다1 |
| 211 | 계속2 | 231 | 앞서다1 |
| 212 | 성폭행1 | 232 | 비자1 |
| 213 | 이천사1 | 233 | 썩다1 |
| 214 | 용인1 | 234 | 별1 |
| 215 | 이천오1 | 235 | 달러1 |
| 216 | 이천육1 | 236 | 돈주다2 |
| 217 | 강남1 | 237 | 듣다3 |
| 218 | 인공와우1 | 238 | 여덟번째1 |
| 219 | 출마1 | 239 | 치매1 |
| 220 | 삼월1 | 240 | 푸켓1 |

| 번호 | 타입명 | 번호 | 타입명 |
|-----|--------|-----|-------|
| 241 | 합의2 | 261 | 두살많다1 |
| 242 | 카드1 | 262 | 두살적다1 |
| 243 | 칠십살1 | 263 | 한살적다1 |
| 244 | 무당1 | 264 | 월제어2 |
| 245 | 입다물다2 | 265 | 카드2 |
| 246 | 모기1 | 266 | 칠만1 |
| 247 | 버리다2 | 267 | 오십만원1 |
| 248 | 끈기2 | 268 | 백년2 |
| 249 | 전기흐르다1 | 269 | 영세1 |
| 250 | 삼분1 | 270 | 일년후1 |
| 251 | 주무르다1 | 271 | 경기도2 |
| 252 | 딱맞다3 | 272 | 이년후1 |
| 253 | 순서1 | 273 | 연맹1 |
| 254 | 해화1 | 274 | 데려오다3 |
| 255 | 여우1 | 275 | 임신1 |
| 256 | 쌍둥이1 | 276 | 회복1 |
| 257 | 푸켓2 | 277 | 그립다1 |
| 258 | 동굴1 | 278 | 스토커1 |
| 259 | 충청남도1 | 279 | 밀가루1 |
| 260 | 마포1 | 280 | 신부1 |

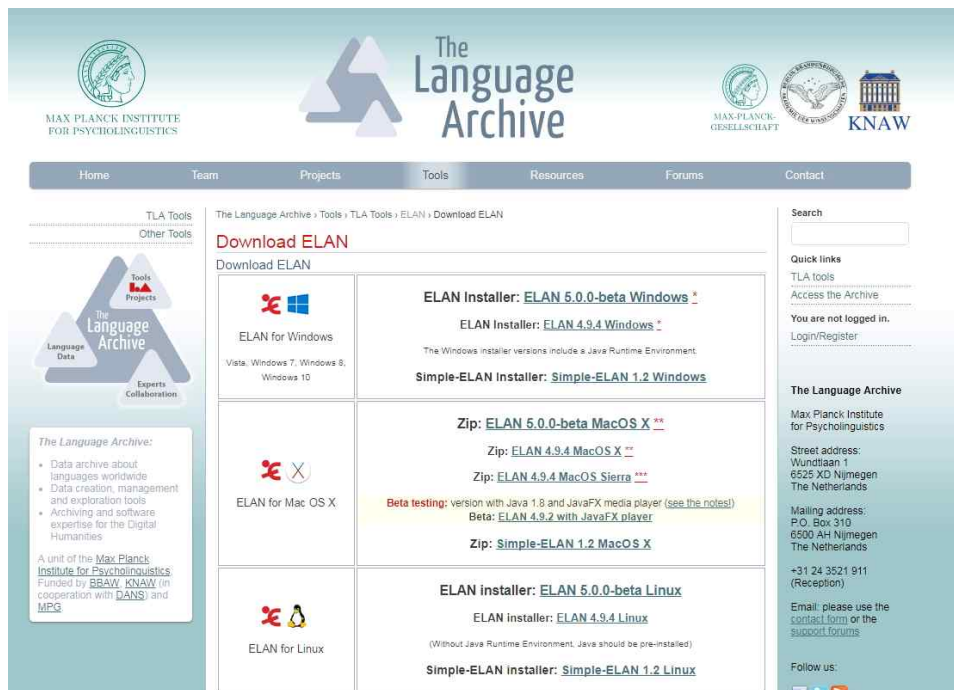
부록 2. 타입 삭제, 타입 수정, 새로운 타입 등록 목록

| 번호 | 타입명 | 번호 | 타입명 |
|-----|-------|-----|-------|
| 281 | 정도1 | 289 | 이십칠1 |
| 282 | 칩십오1 | 290 | 통장정리1 |
| 283 | 육십칠1 | 291 | 주사1 |
| 284 | 의지1 | 292 | 당진1 |
| 285 | 안통하다2 | 293 | 영상통화1 |
| 286 | 명동1 | 294 | 추수1 |
| 287 | 이백만1 | 295 | 찢다1 |
| 288 | 예산1 | | |

ELAN(EUDICO Linguistic Annotator)과 수어 검색·분석

© ELAN 설치(최신버전 5.0.0-beta)

<https://tla.mpi.nl/tools/tla-tools/elan/download/>



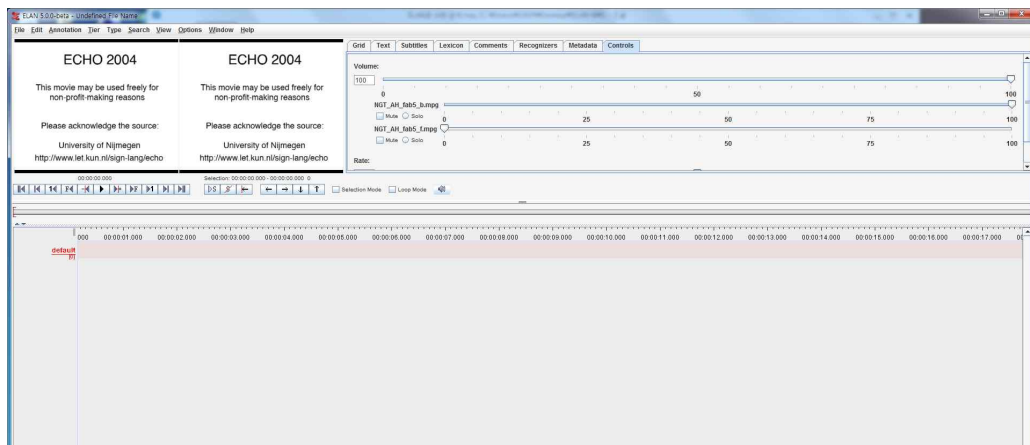
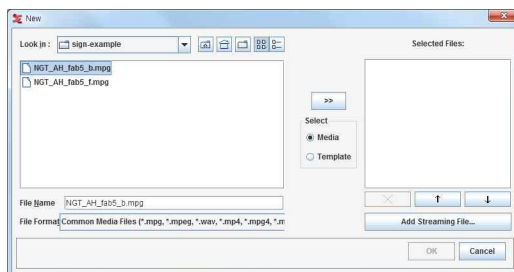
© 시스템 요구 사항

Requirements:

- Java 1.5 or higher (included in the standalone versions for Windows and Linux)
- Mac OS 10.4 or higher with the proper Java version
- Windows 98, ME, 2000, XP, Vista or 7 (32 & 64 bit)
- At least 256 Mb RAM, 512Mb or more preferred
- 1.5 Ghz or faster processor (CPU)

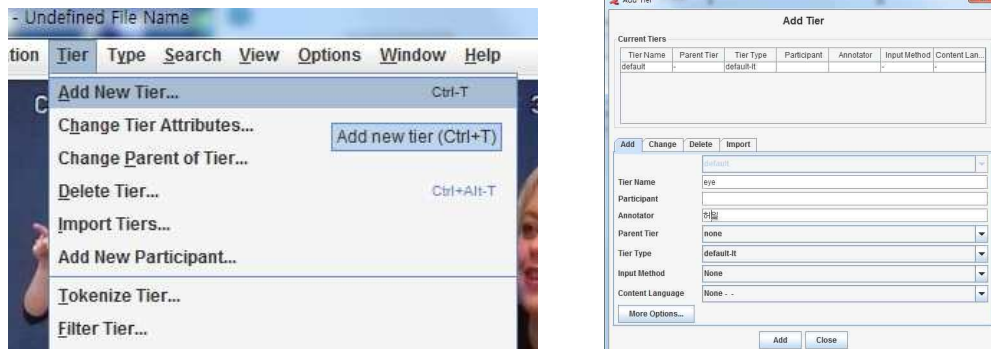
◎ 시작하기

ELAN 메뉴 File -> New



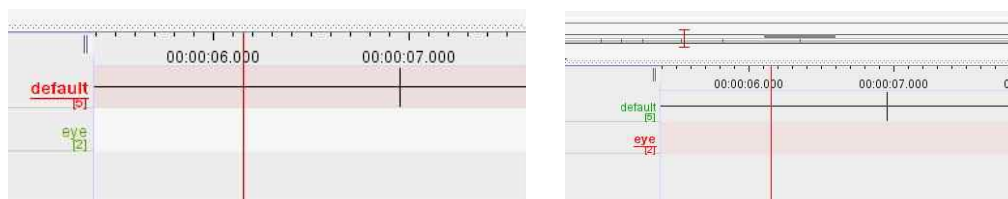
◎ 새로운 Tier의 추가

Tier -> Add New Tier



◎ 입력 Tier의 활성화

좌측 Tier명을 더블클릭하면 입력 티어가 활성화(분홍색)된다.



◎ 단축키 모음

- 파일

File options

| | |
|-------------------------|-----------------------------|
| CTRL+S | Saves the current project |
| CTRL+SHIFT+S | Save as ... |
| CTRL+SHIFT+ALT+S | Save as template |
| CTRL+W | Close the current window |
| CTRL+Q | Exit the application |
| CTRL+O | Open a document |
| CTRL+N | Create a new document |
| CTRL+P | Prints the current document |
| CTRL+SHIFT+P | Page Setup |
| CTRL+ALT+P | Print Preview |
| SHIFT+UP | Activate previous window |
| SHIFT+DOWN | Activate next window |

- 영상 재생

Playing a document & making selections

| | |
|-------------------------|---|
| CTRL+SPACE | Starts or stops the Media Player |
| CTRL+B | Go to the beginning of the video/audio fragment |
| CTRL+PAGE UP | Go to the previous scroll view |
| SHIFT+LEFT | Go back one second |
| CTRL+LEFT | Go back one frame |
| CTRL+SHIFT+LEFT | Go back one "pixel" on the annotation density viewer |
| CTRL+SHIFT+RIGHT | Go to the next "pixel" on the annotation density viewer |
| CTRL+RIGHT | Go to the next frame |
| SHIFT+RIGHT | Go to the next second |
| CTRL+PAGE DOWN | Go to the next scroll view |
| CTRL+E | Go to the end of the media fragment |
| CTRL+ALT+R | Switch to predefined playback rate |
| CTRL+SHIFT+R | Switch to predefined volume |
| CTRL+G | Go to time |

- 전사

| | |
|-----------------------|---|
| ALT+N | Opens the Inline Edit box for inserting a new annotation (only available if a time interval or annotation is selected) |
| ALT+M | Opens the Inline Edit box for modifying an annotation (only available if an annotation is selected) |
| ALT+SHIFT+M | Opens the Local Data Category Selection for associating the annotation with a data category of the ISO Data Category Registry. |
| ALT+Delete | Deletes the value of the selected annotation |
| ALT+D | Deletes an annotation (only available if an annotation is selected) |
| CTRL+SHIFT+N | New annotation before |
| CTRL+SHIFT+N | New annotation after |
| CTRL+SHIFT+T | Add new tier type |
| CTRL+ALT+L | Edit linked file |
| CTRL+SHIFT+C | Edit controlled vocabulary |
| ALT+LEFT | Go to previous Annotation |
| ALT+RIGHT | Go to next Annotation |
| ALT+UP | Go annotation upward |
| ALT+DOWN | Go annotation downward |
| CTRL+T | Add new tier |
| CTRL+ALT+T | Delete tier |
| CTRL+UP | Set previous tier active |
| CTRL+DOWN | Set next tier active |
| CTRL+ALT+SHIFT | Show empty slots in multitier viewer |
| CTRL+ENTER | Saves the changes made to the Inline Edit box or the Edit annotation box |
| ESC | (from within the Inline Edit box or the Edit annotation box): exits the box without saving the changes |
| ESC | (in all other contexts): deselects a selection |
| SHIFT+ENTER | Allows you to switch between the Inline Edit box and the Edit annotation box |
| ALT+click | Allows you to multiple annotations on multiple tiers. |
| SHIFT+E | Opens the Inline edit box for the active annotation |

- 복사 및 붙이기

Copy and paste

| | |
|-------------------------|-----------------------------|
| CTRL+C | Copy Annotation |
| CTRL+ALT+C | Copy Annotation Group |
| CTRL+D | Duplicate Annotation |
| CTRL+V | Paste Annotation |
| CTRL+ALT+V | Paste Annotation Group |
| CTRL+SHIFT+V | Paste Annotation Here |
| CTRL+ALT+SHIFT+V | Paste Annotation Group Here |
| CTRL+ALT+G | Copy the current time |

- 선택 작업

Working with selections

| | |
|--------------------------------|--|
| SHIFT+SPACE | Play selection |
| CTRL+SHIFT+SPACE | Play around selection |
| ALT+SHIFT+C | Clear selection |
| CTRL+SHIFT+Z | Clear selection and cancel selection mode |
| CTRL+/- or CTRL+SHIFT+K | Move the crosshair to the right or left of the selection |
| ALT+CTRL+/- | Move the crosshair to the center of the selection |
| CTRL+K | Toggle selection mode |
| CTRL+L | Toggle loop mode |
| SHIFT+A | Put the selection in the center of the Timeline Viewer |

- 검색

Searching

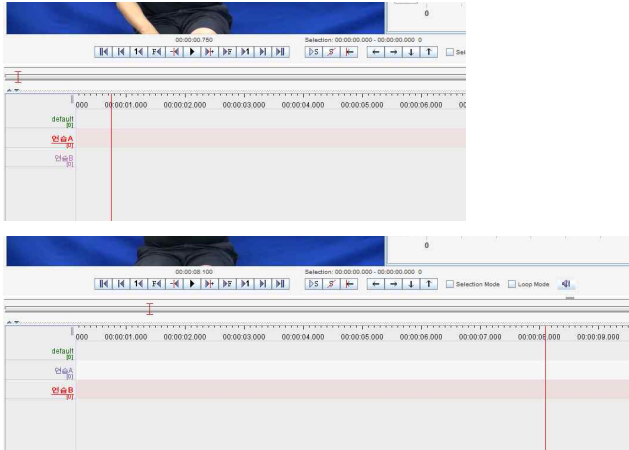
| | |
|---------------------|------------------------------|
| CTRL+F | Find |
| CTRL+SHIFT+F | Search in multiple eaf files |

General

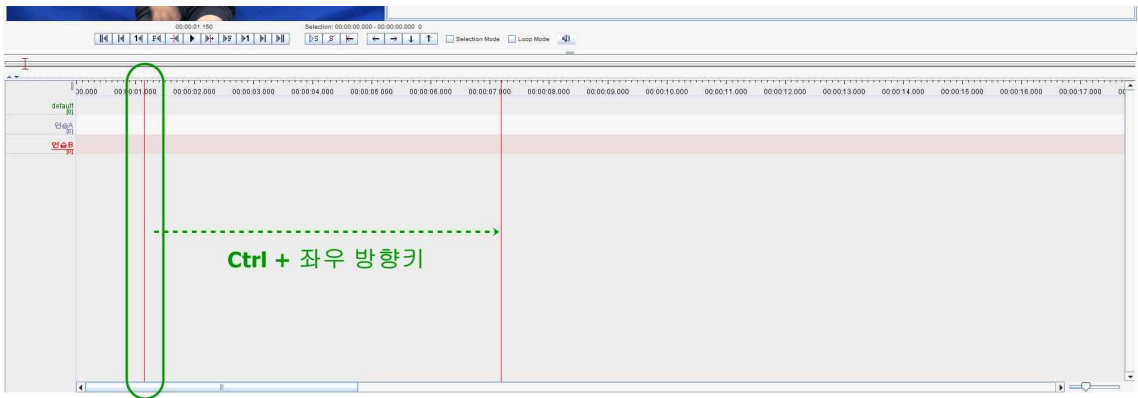
| | |
|---------------|------|
| CTRL+Z | Undo |
| CTRL+Y | Redo |

◎ 영상 보고 전사하기

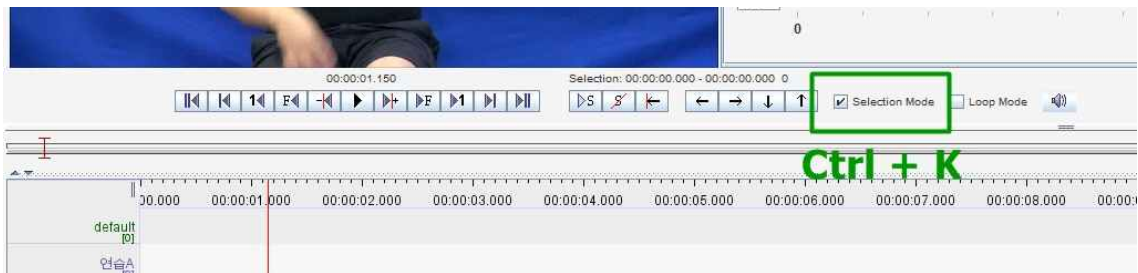
1. Ctrl+ ⌵ : 전사할 층렬 선택하기



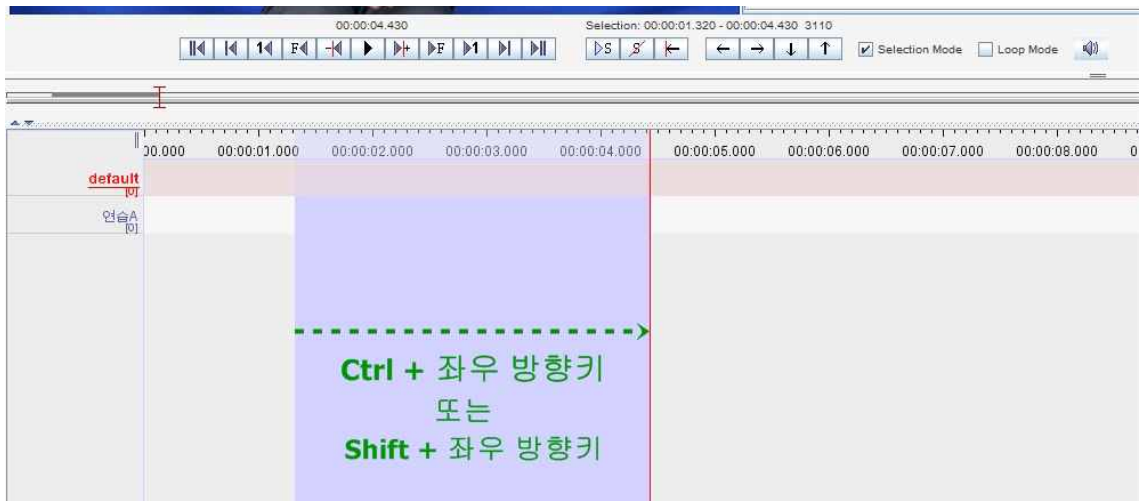
2. Ctrl+ ⇐ : 전사할 구간으로 이동하기



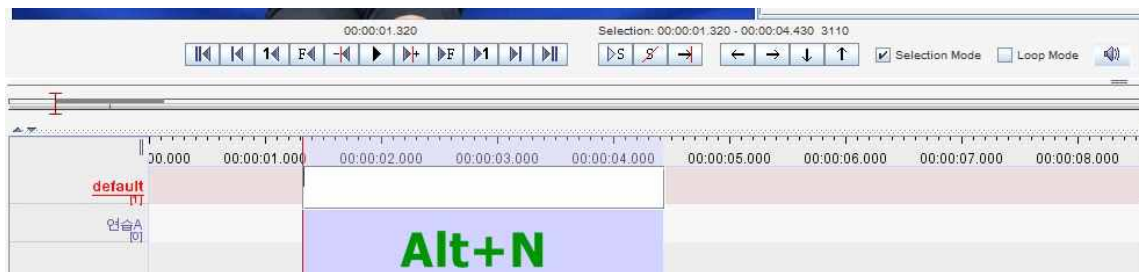
3. Ctrl + K : 분절 시작하기



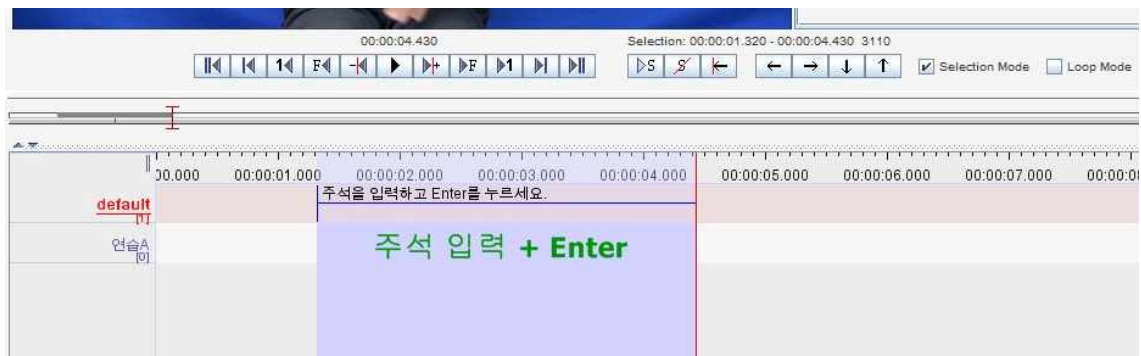
4. Ctrl+ ⇐ 또는 Shift + ⇐ : 분절하고 싶은 구간 선택하기



5. Alt + N : 주석 공간 만들기



6. 주석 입력 + Enter : 주석 달기



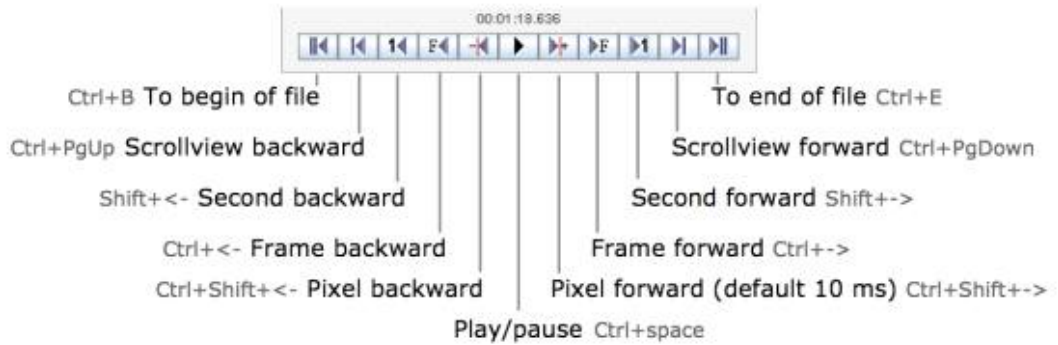
7. Alt+Shift+C : 선택 구간 해지 하기



8. 분절 수정 : Alt+분절 시작 또는 끝에서 마우스 조정

9. 분절 구간 이동 하기 : Alt+ ⇐

◎ 영상 재생



맨 앞으로: Ctrl+B

맨 뒤로: Ctrl+E

Scrollview 앞으로(한 페이지 앞): Ctrl+Page Up

Scrollview 뒤로(한 페이지 뒤): Ctrl+Page Down

1초 앞으로: Shift+Left

1초 뒤로: Shift+Right

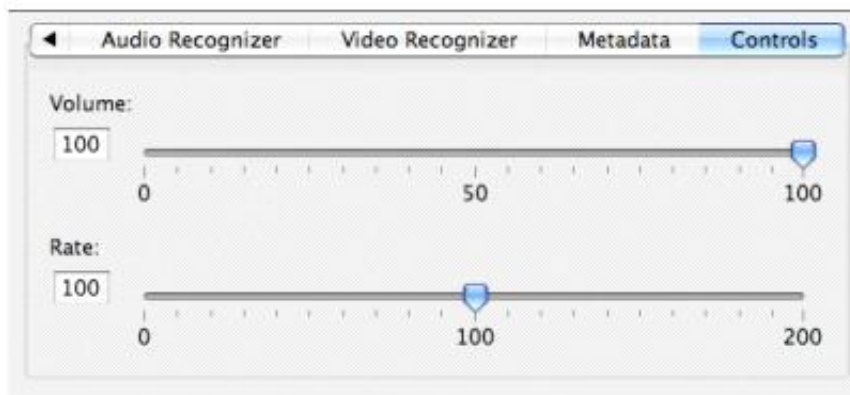
1프레임 앞으로(40 ms, 0.04초): Ctrl+Left

1프레임 뒤로(40 ms, 0.04초): Ctrl+Right

1픽셀 앞으로(10 ms, 0.01초): Ctrl+Shift+Left

1픽셀 뒤로(10 ms, 0.01초): Ctrl+Shift+Right

◎ 볼륨과 속도의 조절

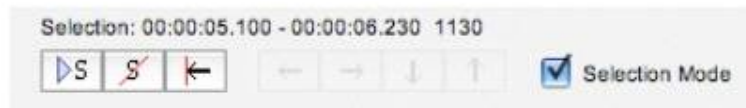


◎ 선택된 분절 조정

선택된 분절만 재생: Shift+Space

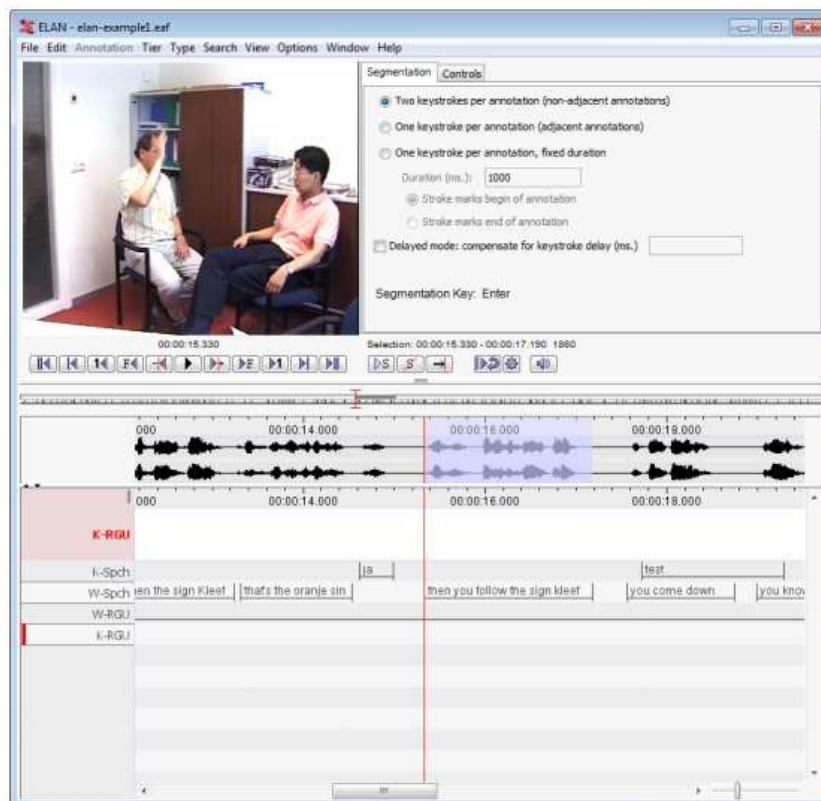
선택 취소: Alt+Shift+C

좌우 경계선 바꾸기: Ctrl+/



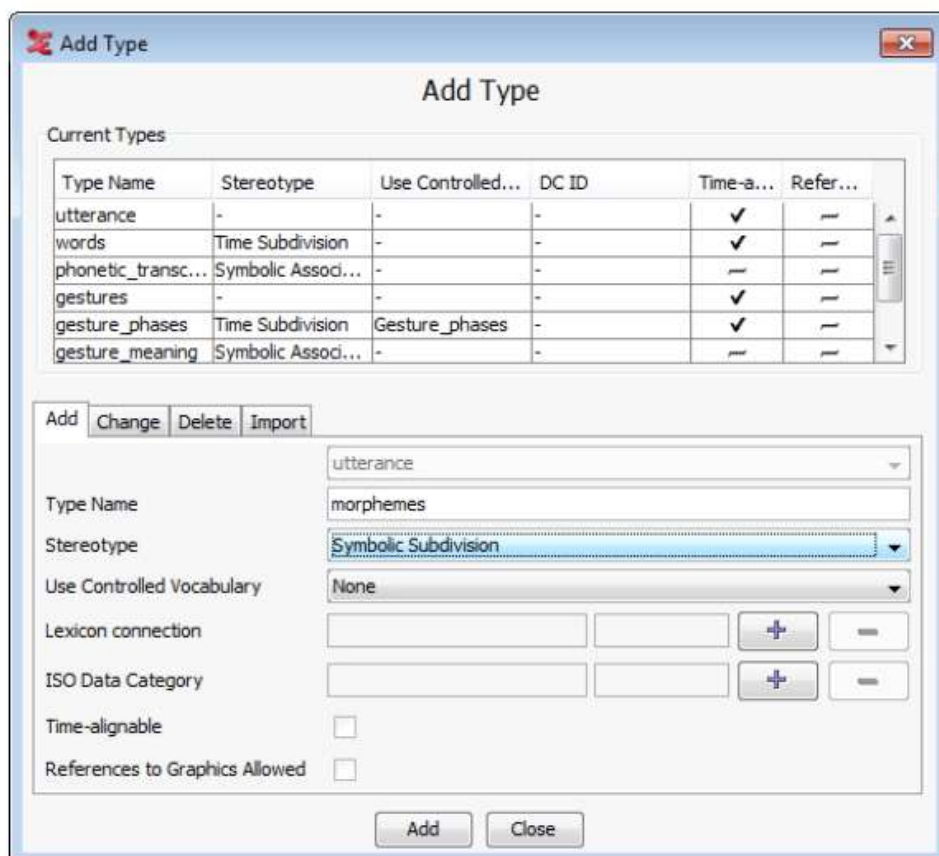
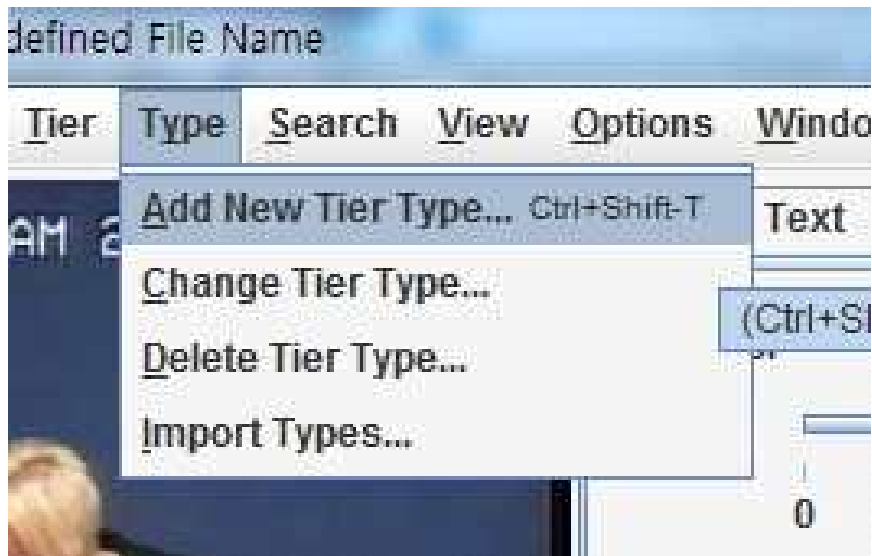
◎ 분절 모드

영상을 재생하거나 멈추면서 Enter를 누르면 분절의 시작과 끝이 지정되어 빠르게 분절 작업을 진행할 수 있다.



◎ 티어 유형의 정의

Type -> Add New Tier Type



- Stereotype

2. The predefined Stereotypes are:



"Time Subdivision"

- Time aligned subdivision of parent interval, no gaps

"Symbolic Subdivision"

- Ordered, symbolic subdivision of parent interval

"Included In"

- Time aligned subdivision of parent interval, gaps are allowed

"Symbolic Association"

- One-to-one association with the parent annotation

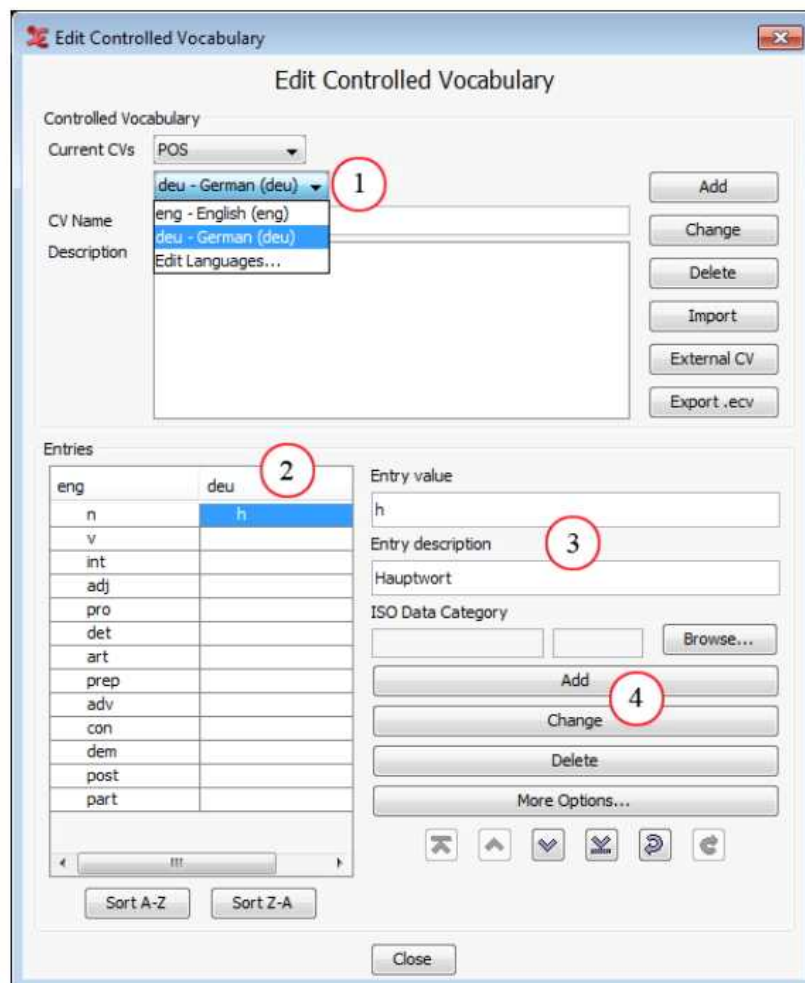
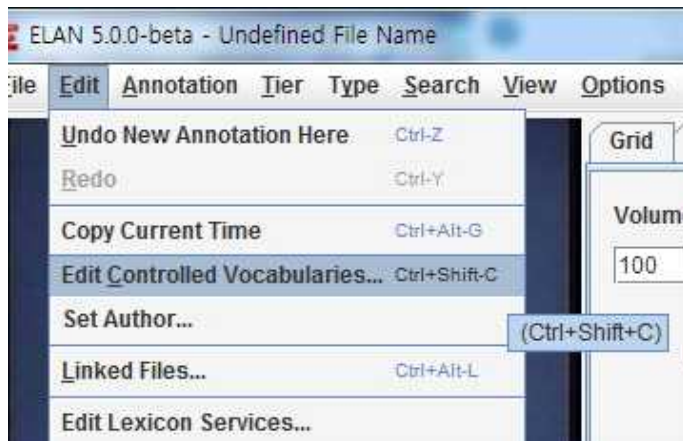
| | |
|---|---|
| <ul style="list-style-type: none"> • Time Subdivision | <p>independent tier. The annotations cannot overlap.</p> <p>The annotation on the parent tier can be sub-divided into smaller units, which, in turn, can be linked to time intervals. Note that there are no time gaps allowed, i.e., the smaller units have to immediately follow each other.</p> <p>E.g., an utterance transcribed on a parent tier can be sub-divided into words – each of which is then linked to its corresponding time interval. Annotations on such tiers are time-alignable. They differ from annotations on independent tiers in that they are assigned to an interval that is contained within the interval of their parent annotation.</p> |
| <ul style="list-style-type: none"> • Symbolic Subdivision | <p>Similar to Time Subdivision, except that the smaller units cannot be linked to a time interval.</p> <p>E.g., a word on a parent tier can be sub-divided into individual morphemes (which are not linked to a time interval).</p> |
| <ul style="list-style-type: none"> • Included In | |

Tiers, tier types, controlled vocabularies and templates.

| | |
|---|--|
| <ul style="list-style-type: none"> • Symbolic Association | <p>All annotations fall within the borders of the parent tier. However, there can be gaps between the child annotations.</p> <p>E.g., a sentence with a silence can be split up into words while the silence corresponds to a gap in the child annotations (i.e. the separate words).</p> <p>The annotation on the parent tier cannot be sub-divided further, i.e., there is a one-to-one correspondence between the parent annotation and its referring annotation.</p> <p>E.g., one sentence on a parent tier has exactly one free translation. Or one word has exactly one gloss.</p> |
|---|--|

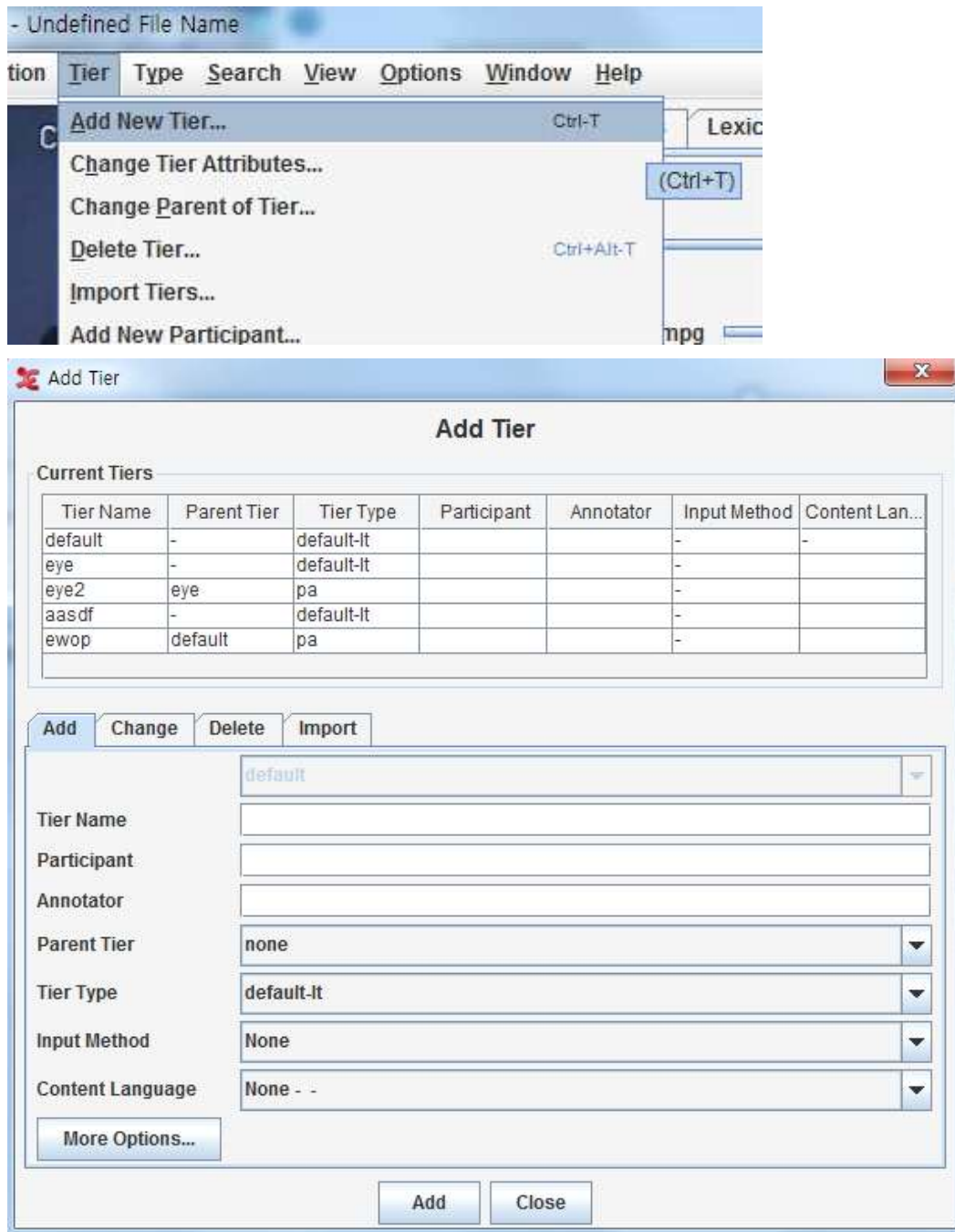
- Controlled Vocabularies

Edit -> Edit controlled vocabularies



◎ 티어 추가하기

Tier -> Add New Tier



◎ 티어 간 위계 설정(부모 티어 - 자식 티어)

Table 2.1. Nested tier dependencies

| tier: | type: | hierarchical relation: |
|---------------------------------|-------------|------------------------|
| ref (referent) | independent | parent of tx and ft |
| tx (text) | referring | parent of mb |
| mb (morpheme break) | referring | parent of gl and ps |
| gl (gloss), ps (part of speech) | referring | - |

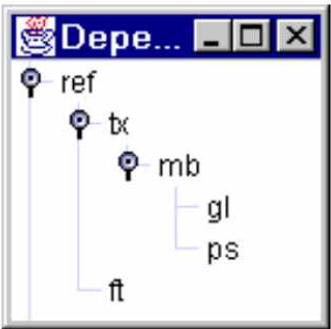


Figure 2.2. Tier dependencies

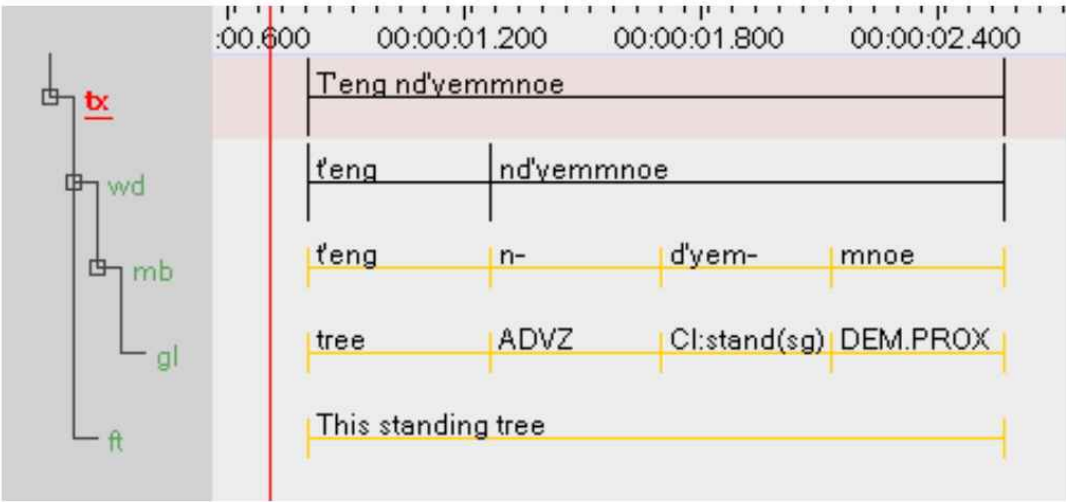
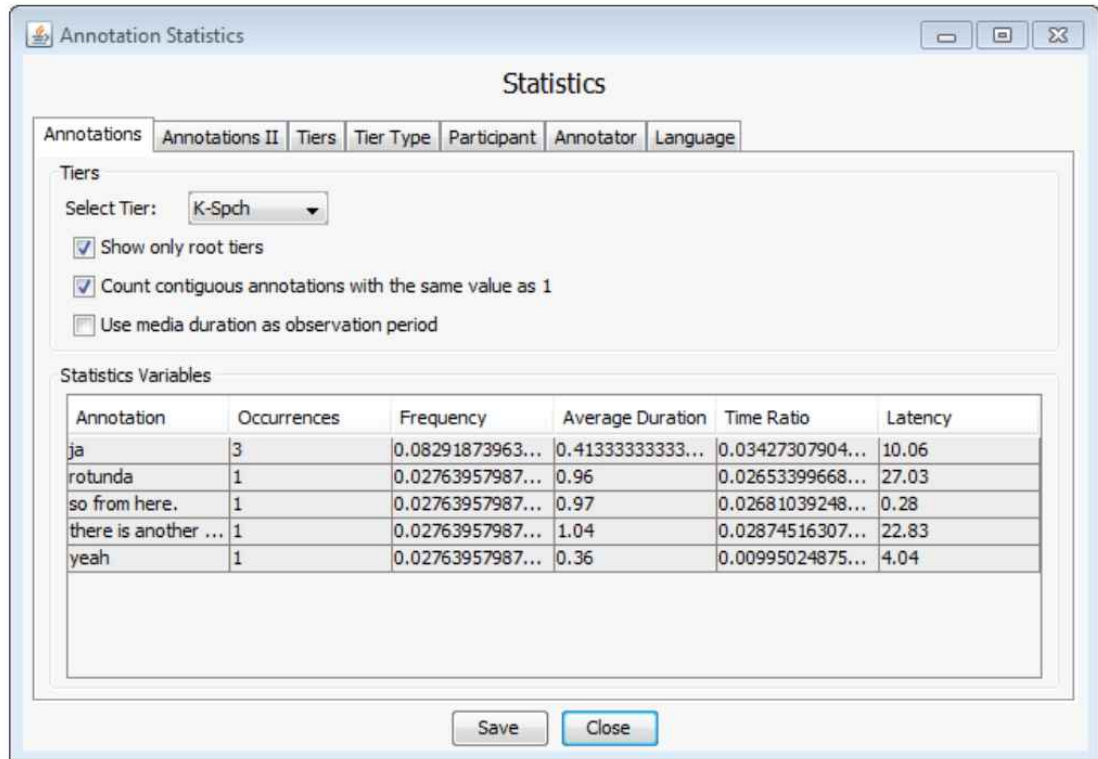


Figure 2.3. Tier dependencies in the timeline viewer

◎ 티어별 통계값 보기

View > Annotations Statistics



◎ 단일 전사 파일(eaf 파일) 검색

1. Click on Search menu.
2. Go to Find (And Replace)... (alternatively you can press CTRL+F)

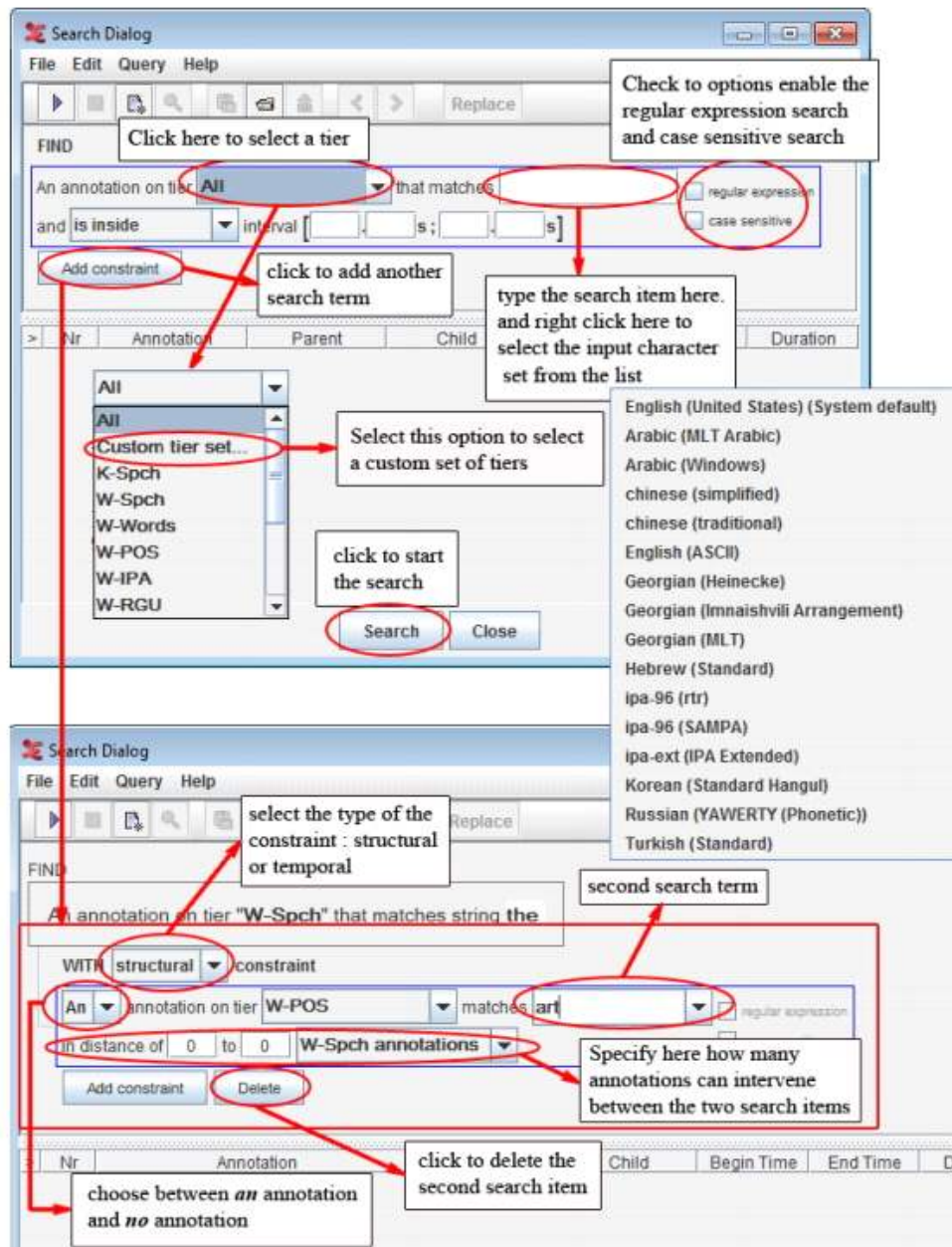


Figure 4.1. Search dialog

- 검색 조건

- *is inside*: the annotation is completely contained within a given interval



- *overlaps*: at least a part of the annotation is contained within the given interval



- *overlaps only begin time of*: the annotation only has its end part in common with the given interval



- *overlaps only end time of*: the annotation only has its begin part in common with the given interval



- *is within ... around*: the annotation is contained in an interval around either the begin time or the end time



- *is within ... around begin time of*: the annotation is contained in an interval around the begin time



- *is within ... around end time of*: the annotation is contained in an interval around the begin time



- 검색 결과

the search item

the value of each matching annotation with its value of parent and child annotations.

All child annotations are listed and it concatenates all the annotations with linguistic type 'time subdivision'

| Nr. | Annotation | Parent | Child | Begin Time | End Time | Duration |
|-----|------------------------------------|-------------------|---|--------------|-------------|--------------|
| 1 | so you go out of the | the matching text | you go out of the Institute to the Saint Anna | 00:00:00.780 | 00:00:04.00 | 00:00:03.220 |
| 2 | and then you go the | the matching text | and then you go the other Saint Anna Straat to | 00:00:04.600 | 00:00:10.00 | 00:00:05.400 |
| 3 | and you follow then the sign Kieef | the matching text | (and you follow then the sign Kieef) (and you follow then the sign Kieef) | 00:00:11.150 | 00:00:13.00 | 00:00:01.850 |
| 4 | that's the orange single | the matching text | that's the Orange Single (barte ba orange single) | 00:00:13.310 | 00:00:14.00 | 00:00:00.690 |
| 5 | then you follow the sign kieef | the matching text | (then you follow the sign kieef) (sen ju: bla:u 8) | 00:00:15.330 | 00:00:17.00 | 00:00:01.670 |
| 6 | you come down to the | the matching text | (you come down to the) (ju: kum daun tu be) | 00:00:21.860 | 00:00:22.00 | 00:00:00.140 |
| 7 | and then you follow the sign kieef | the matching text | (and then you follow the sign kieef) (and den) | 00:00:29.370 | 00:00:31.00 | 00:00:01.630 |
| 8 | and then you go the this way | the matching text | (and then you go the this way) (and den ju: go) | 00:00:32.379 | 00:00:34.00 | 00:00:01.621 |

the number of matches found

11 occurrences in 8 annotations

begin, end time and duration of each annotation

Search Close Search complete

Figure 4.5. Search results

◎ 검색 조건의 저장

1. Specify a query in the search dialog
2. Either choose Query > Save or click on the save icon in the toolbar.
3. Enter a filename for the active query.
4. Choose save.

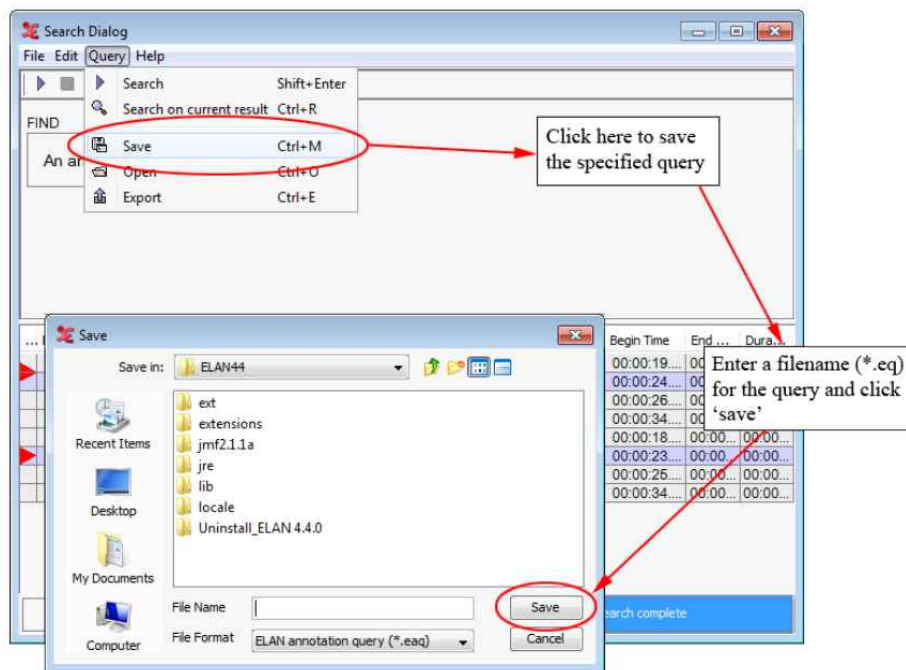


Figure 4.8. Saving a query

◎ 저장한 검색 조건 불러오기

1. Either choose Query > Open or click on the Open icon in the toolbar.
2. Select an existing query from the file dialog.
3. Choose Open

◎ 검색 결과의 저장

1. In the query dialog, select Query > Export matches
2. Specify a file name
3. Click on save

| | | | | | | | |
|---|--------------|---|--------------|--------------------------------------|--------------|--------------|-----|
| the end time of the annotation (both in hh:mm:ss.mmm format and seconds) | | | | | | | |
| W-Words | 00:00:02.190 | 2.19 | 00:00:02.570 | 2.57 | 00:00:00.380 | 0.38 | the |
| W-Words | 00:00:03.200 | 3.2 | 00:00:03.340 | 3.34 | 00:00:00.140 | 0.14 | the |
| W-Words | 00:00:05.500 | 5.5 | 00:00:05.700 | 5.7 | 00:00:00.200 | 0.2 | the |
| Tier name | | the begin time of the annotation (both in hh:mm:ss.mmm format and seconds) | | the duration of 1 annotation unit | | query result | |

Figure 4.9. Exported search results

◎ 검색 및 변경

Edit > Replace

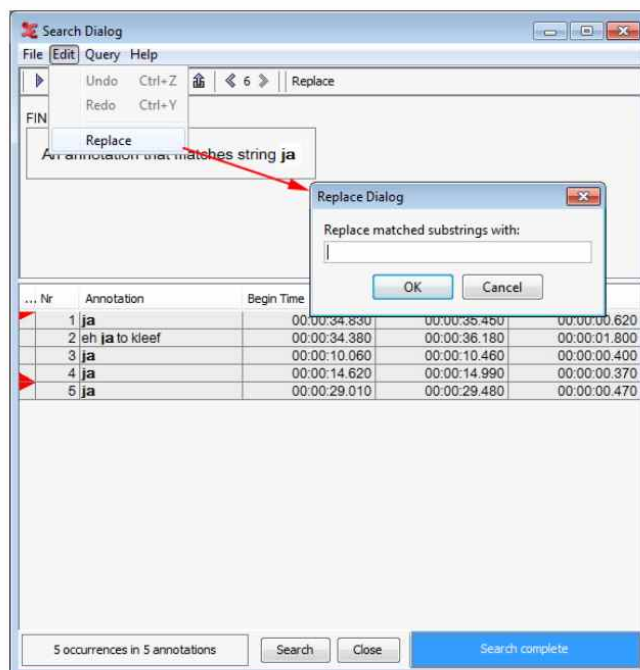


Figure 4.11. Search and replace

◎ 검색결과를 클릭하면 주석 모드에서 확인 가능

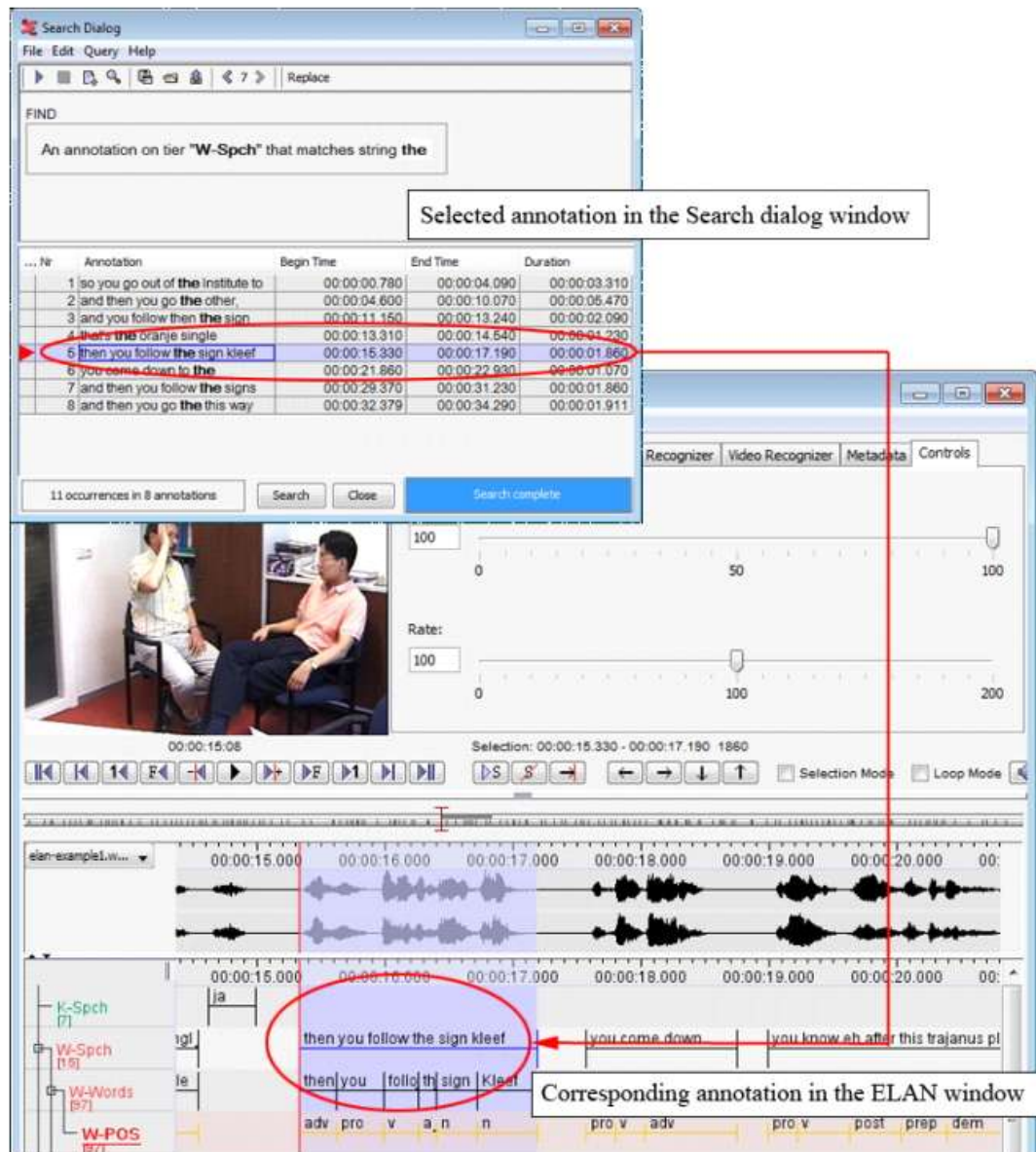


Figure 4.12. From search results to annotations

◎ 단일 파일 검색 예

Annotation Statistics

Statistics

Annotations Annotations II Tiers Tier Type Participant Annotator Language

Tiers

Select Tier: **우새GlossA**

☒ Show only root tiers

☒ Count contiguous annotations with the same value as 1

☐ Use media duration as observation period

Statistics Variables

| Annotation | Occurrences | Frequency | Average Duration | Time Ratio | Latency |
|------------|-------------|-----------------|------------------|-----------------|---------|
| 지시1_위치M | 33 | 0.0263452650... | 0.3270303030... | 0.0086157000... | 46.862 |
| 문1 | 27 | 0.0215552168... | 0.3456666666... | 0.0074509199... | 67.838 |
| 종다1 | 24 | 0.0191601927... | 0.4030833333... | 0.0077231543... | 64.377 |
| 지시1_1 | 21 | 0.0167651686... | 0.2943809523... | 0.0049353463... | 0.502 |
| 지시1_위치R | 20 | 0.0159668273... | 0.3197 | 0.0051045946... | 408.999 |
| 무엇1 | 18 | 0.0143701445... | 0.2585555555... | 0.0037154807... | 19.574 |
| 한국1 | 18 | 0.0143701445... | 0.4333888888... | 0.0062278609... | 306.064 |

Save Close

Annotation Statistics

Statistics

Annotations Annotations II Tiers Tier Type Participant Annotator Language

Tiers

Select Tier: **눈매리방향A**

☒ Show only root tiers

☒ Count contiguous annotations with the same value as 1

☐ Use media duration as observation period

Statistics Variables

| Annotation | Occurrences | Frequency | Average Duration | Time Ratio | Latency |
|------------|-------------|------------------|------------------|------------------|---------|
| 뒤로 | 3 | 0.00239502409... | 0.4696666666... | 0.00112486298... | 105.054 |
| 앞뒤로 | 10 | 0.00798341365... | 0.3276 | 0.00261536631... | 639.862 |
| 앞으로 | 26 | 0.02075687551... | 0.38273076923... | 0.00794429493... | 34.177 |
| 오른쪽 | 5 | 0.00399170682... | 0.4535999999... | 0.00181063821... | 655.91 |
| 왼쪽 | 14 | 0.01117677912... | 0.32821428571... | 0.00366837857... | 435.569 |
| 좌우로 | 3 | 0.00239502409... | 0.2776666666... | 6.65018357859... | 466.707 |
| 종합 | 6 | 0.00479004819... | 0.6046666666... | 0.00289638247... | 31.88 |

Save Close

Annotation Statistics

Statistics

Annotations Annotations II Tiers Tier Type Participant Annotator Language

Tiers

Select Tier: **눈썹A**

☒ Show only root tiers

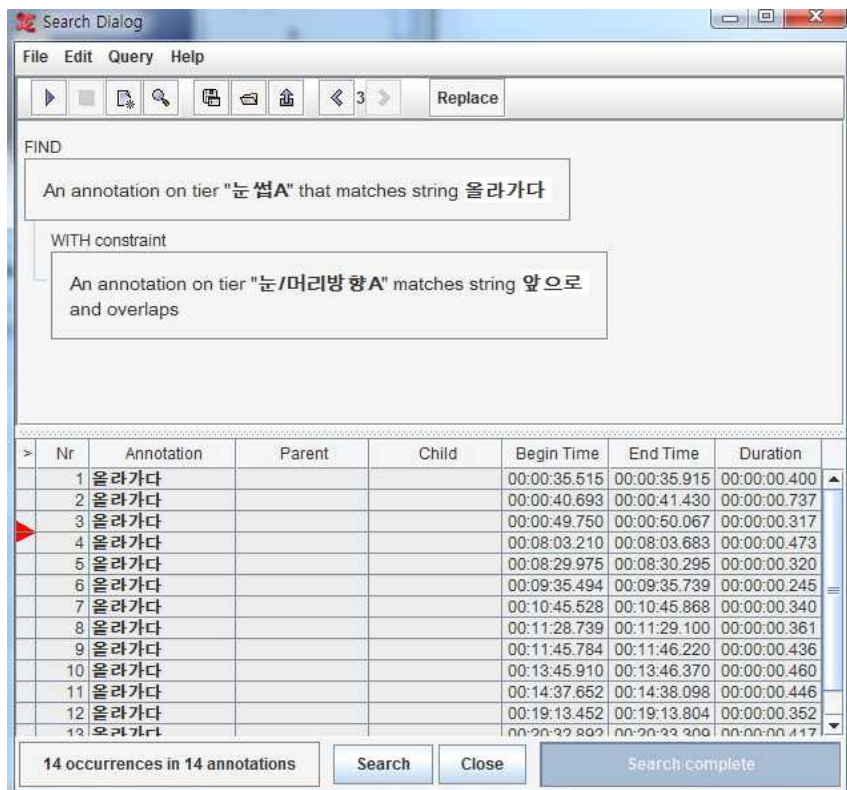
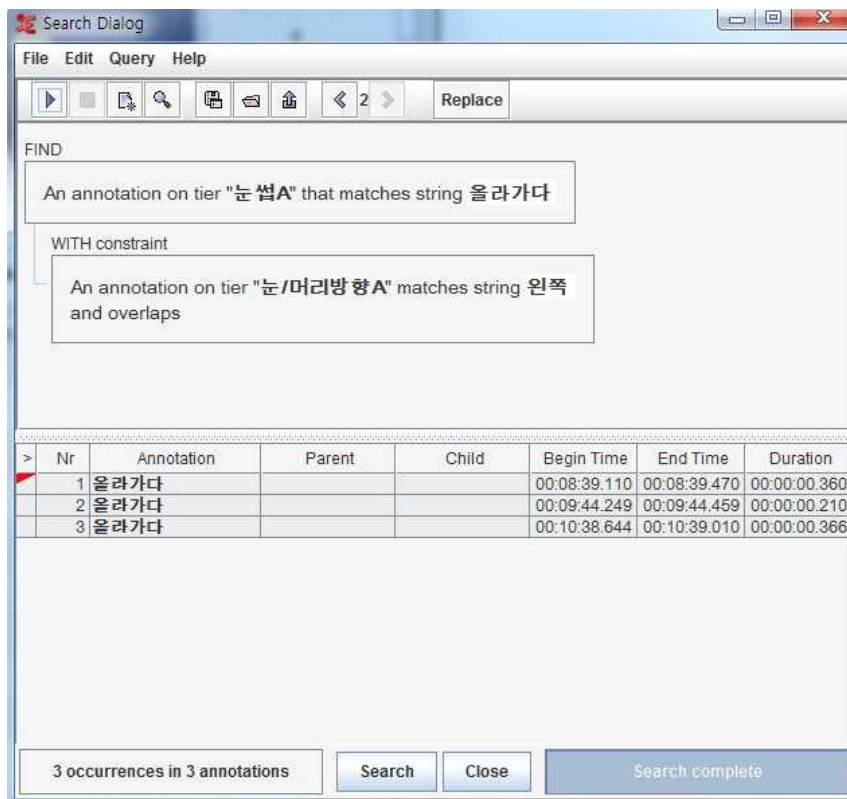
☒ Count contiguous annotations with the same value as 1

☐ Use media duration as observation period

Statistics Variables

| Annotation | Occurrences | Frequency | Average Duration | Time Ratio | Latency |
|------------|-------------|------------------|------------------|------------------|---------|
| 내려가다 | 1 | 7.98341365978... | 0.375 | 2.99378012241... | 511.887 |
| 올라가다 | 33 | 0.02634526507... | 0.39169696969... | 0.01031936049... | 31.88 |
| 종합 | 13 | 0.01037843775... | 0.49330769230... | 0.00511976318... | 88.033 |
| 저푸리다 | 28 | 0.02235355824... | 0.34903571428... | 0.00780219016... | 65.522 |

Save Close



- 검색 조건 예: Temporal distance

Table 4.1. Annotation example

| tier | annotations | | | | | | | |
|---------------------------|----------------------------|------|-----|-------|-----|-----|---------|-----|
| st (sentence) | He sees trees and flowers. | | | | | | | |
| tx (word) | he | sees | | trees | | and | flowers | |
| mb (morpheme break) | ... | see | -s | tree | -s | ... | flower | -s |
| ps (part of speech) | ... | V | SUF | N | SUF | ... | N | SUF |

FIND

An annotation on tier **tx** that matches **\w** ☒ regular expression
 and **is inside** interval [] s; [] s ☐ case sensitive

tx matches \w (any non-empty annotation)

Add constraint

WITH constraint

An annotation on tier **ps** matches regular expression **^N\$**
 in distance of 0 to 0 tx annotations

ps matches N

WITH constraint

An annotation on tier **mb** matches string **-s**
 in distance of -X to +X ref annotations

mb matches -s

tx 티어에서 trees, flowers 선택됨, sees 선택 안 됨

Search Dialog

File Edit Query Help

Replace

An annotation on tier "K-Spch"

WITH constraint

No annotation on tier "W-Spch" that overlaps

WITH constraint

An annotation on tier "W-Spch" that is within 2000000 ms around begin time

| Nr | Annotation | Parent | Child | Begin Time | End Time | Duration |
|----|------------|--------|-------|--------------|--------------|--------------|
| 1 | ja | | | 00:00:14.620 | 00:00:14.990 | 00:00:00.370 |
| 2 | rotunda | | | 00:00:27.030 | 00:00:27.990 | 00:00:00.960 |

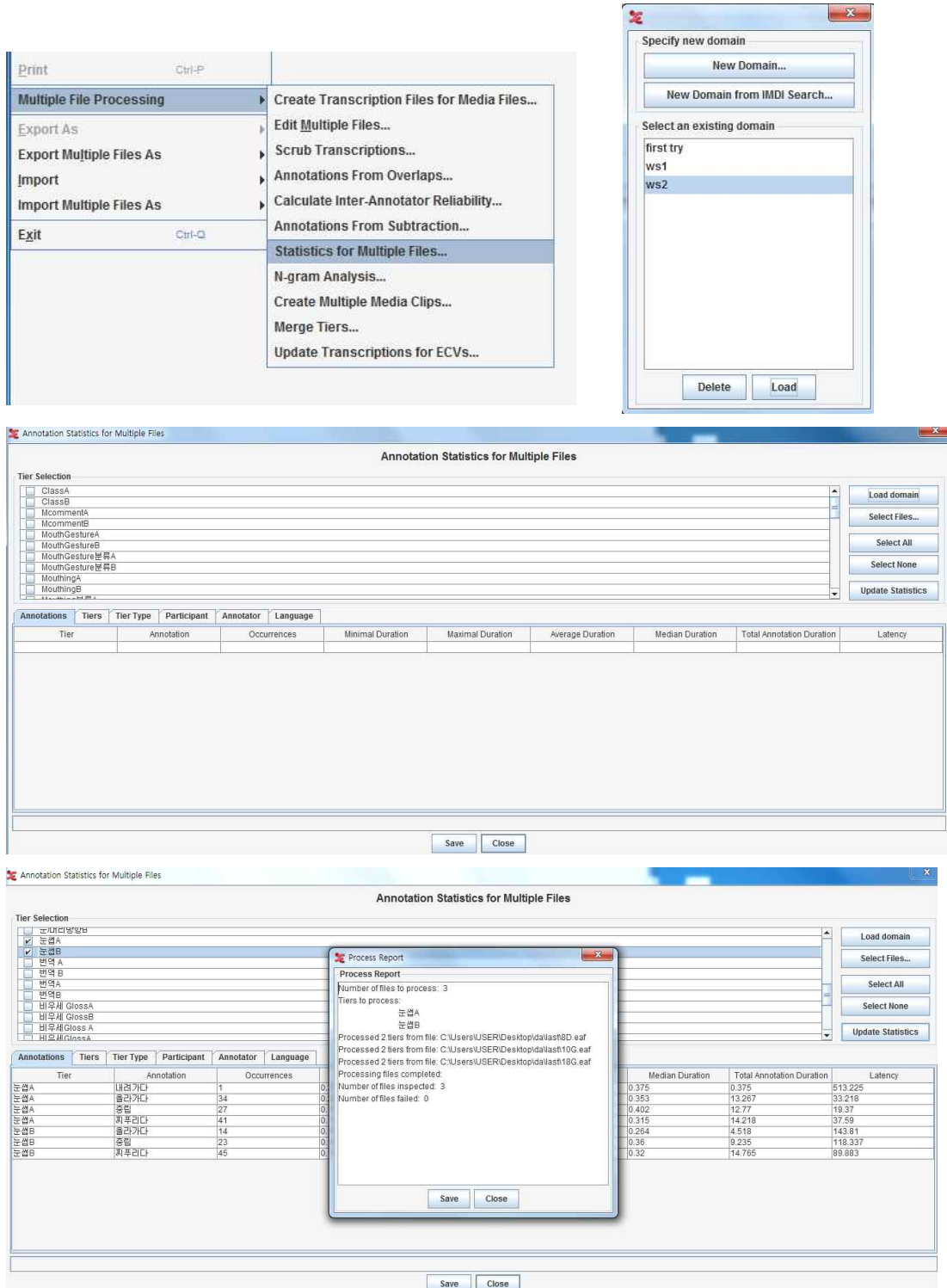
2 occurrences in 2 annotations Search Close Search complete

Timeline view showing K-Spch and W-Spch annotations with a red circle highlighting the search results.

Figure 4.4. Advanced searching example

◎ 다수의 파일 티어별 빈도분석

File -> Multiple File Processing -> Statistics for Multiple Files...



◎ 다수의 전자 파일 검색하기

Search > Search multiple eaf...



Figure 4.13. Searching through multiple annotation files

◎ 다중 파일 검색 결과

| Nr | File | Tier | Before | Annotation | After | Begin Time | End Time | Duration |
|----|---------------|---------------|---------------------|---------------|----------------------|--------------|--------------|--------------|
| 13 | elan-example1 | W-Words | of | the | town | 00:00:08.130 | 00:00:08.260 | 00:00:00.130 |
| 14 | elan-example1 | W-Words | then | the | sign | 00:00:12.200 | 00:00:12.340 | 00:00:00.140 |
| 15 | elan-example1 | W-Words | 's | the | Oranje | 00:00:13.694 | 00:00:13.836 | 00:00:00.142 |
| 16 | elan-example1 | W-Words | follow | the | sign | 00:00:16.263 | 00:00:16.404 | 00:00:00.141 |
| 17 | elan-example1 | W-Words | to | the | Rhine | 00:00:22.635 | 00:00:22.930 | 00:00:00.295 |
| 18 | elan-example1 | W-Words | follow | the | signs | 00:00:30.305 | 00:00:30.465 | 00:00:00.160 |
| 19 | elan-example1 | W-Words | go | the | this | 00:00:33.220 | 00:00:33.620 | 00:00:00.400 |
| 20 | elan-example1 | W-RGMe | | Going out of | Going out of M... | 00:00:01.600 | 00:00:02.240 | 00:00:00.640 |
| 21 | elan-example1 | W-RGMe | A roundabout | Going to the | Going along Or... | 00:00:12.120 | 00:00:12.680 | 00:00:00.560 |
| 22 | elan-example1 | W-RGMe | Going along Or... | Going to the | Going down | 00:00:17.000 | 00:00:17.200 | 00:00:00.200 |
| 23 | pear story | Clause Tra... | and then you s... | and so he | and he starts pi... | 00:00:10.440 | 00:00:14.040 | 00:00:03.600 |
| 24 | pear story | Clause Tra... | and so he clim... | and he starts | and he puts the... | 00:00:14.050 | 00:00:17.070 | 00:00:03.020 |
| 25 | pear story | Clause Tra... | and he starts pi... | and he puts | OK | 00:00:17.283 | 00:00:19.423 | 00:00:02.140 |
| 26 | pear story | Clause Tra... | and so then he ... | and then he | and starts puttin... | 00:00:23.413 | 00:00:25.173 | 00:00:01.760 |
| 27 | pear story | Clause Tra... | and then he cli... | and starts | and then you h... | 00:00:25.183 | 00:00:28.553 | 00:00:03.370 |
| 28 | pear story | Clause Tra... | and then he go... | and the man | and starts picki... | 00:00:38.943 | 00:00:40.663 | 00:00:01.720 |
| 29 | pear story | Clause Tra... | and the man go... | and starts | | 00:00:40.673 | 00:00:43.843 | 00:00:03.170 |
| 30 | pear story | morph | with | the | ladder | 00:00:13.476 | 00:00:13.752 | 00:00:00.276 |
| 31 | pear story | morph | off | the | tree | 00:00:16.312 | 00:00:16.689 | 00:00:00.377 |
| 32 | pear story | morph | puts | the | pears | 00:00:18.084 | 00:00:18.351 | 00:00:00.267 |
| 33 | pear story | morph | down | the | ladder | 00:00:24.668 | 00:00:24.919 | 00:00:00.251 |
| 34 | pear story | morph | all | the | pears | 00:00:26.531 | 00:00:26.868 | 00:00:00.337 |
| 35 | pear story | morph | and | the | man | 00:00:39.158 | 00:00:39.373 | 00:00:00.215 |

Figure 4.14. Searching through multiple annotation files: result

◎ 다중 파일 검색 및 변경

Search > Find & Replace in Multiple Files...

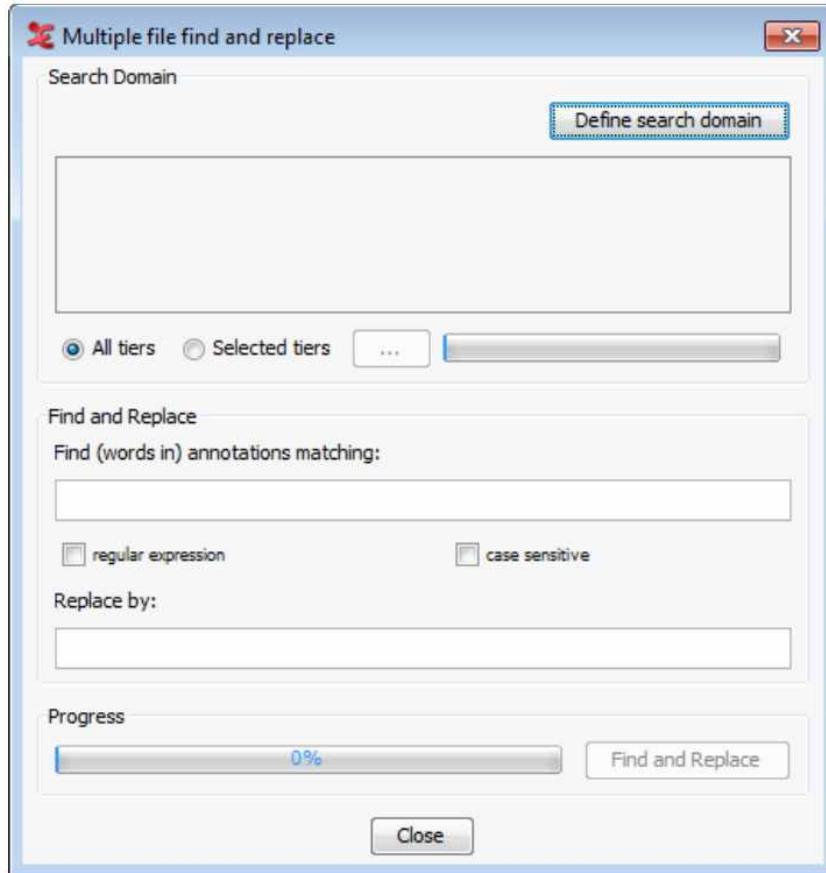


Figure 4.15. Find & replace in multiple files dialog

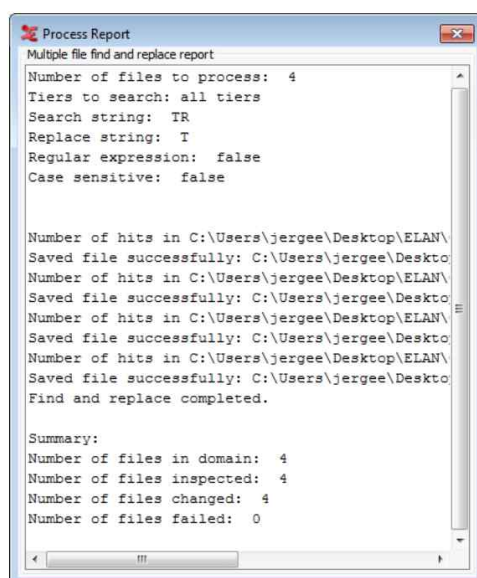


Figure 4.17. Process report dialog

◎ 다수 전사 파일 다중 검색

Search > Structured search multiple eaf....

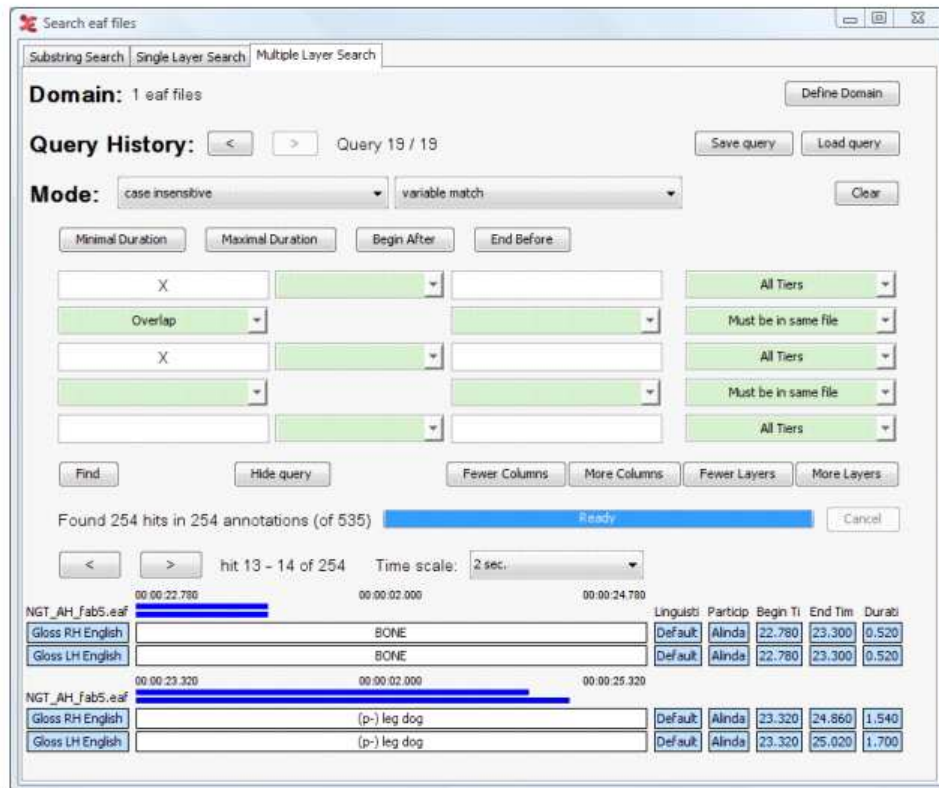


Figure 4.23. Variable Match

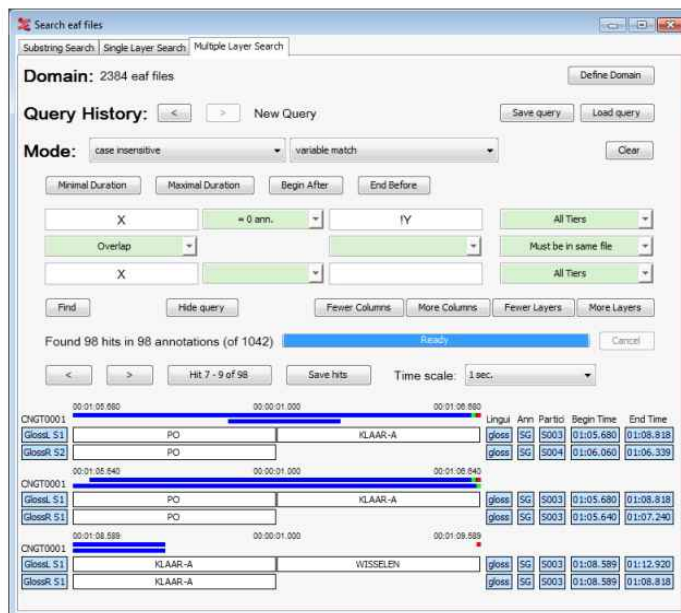


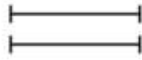
Figure 4.24. Multiple Variable Match

- 검색 조건 제한

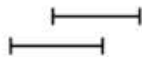
- **= N annotations:** between the annotations containing the two search strings, there must be exactly N annotations.
- **> N annotations:** between the annotations containing the two search strings, there must be more than N annotations.
- **< N annotations:** between the annotations containing the two search strings, there must be less than N annotations.
- **= X milliseconds:** between the annotations containing the two search strings, there must be exactly X milliseconds.
- **> X milliseconds:** between the annotations containing the two search strings, there must be more than X milliseconds.
- **< X milliseconds:** between the annotations containing the two search strings, there must be less than X milliseconds.
- **No constraints:** there are no constraints.
- **Clear:** clear the current constraint.

- 검색 조건2

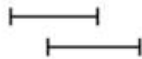
- **Fully aligned:** the begin time and end time of both annotations are the same:



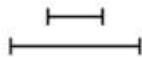
- **Overlap:** part of both annotations overlap. This includes the other options Fully aligned, Left overlap, Right overlap, Surrounding and Within.
- **Left overlap:** the begin time and end time of the annotation matching the lower search string lie *before* the begin time and end time of the annotation matching the upper search string:



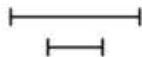
- **Right overlap:** the begin time and end time of the annotation matching the lower search string lie *after* the begin time and end time of the annotation matching the upper search string:



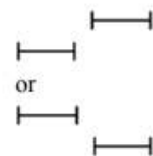
- **Surrounding:** the begin time of the annotation matching the lower search string lies before the begintime of the annotation matching the upper search string and end time of the annotation matching the lower search string lies after the end time of the annotation matching the upper search string:



- **Within:** the begin time of the annotation matching the lower search string lies after the begintime of the annotation matching the upper search string and end time of the annotation matching the lower search string lies before the end time of the annotation matching the upper search string:



- **No overlap:** the begin time of the annotation matching a search string lies after the end time of the annotation matching the other search string:



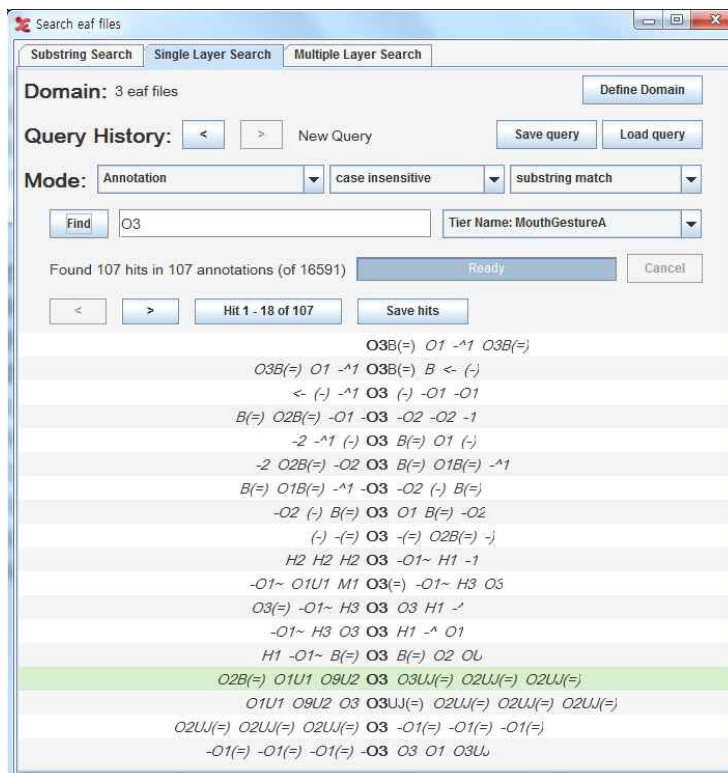
- 검색 조건3

- **begin time - begin time = X milliseconds**: the begin time of the annotations matching the upper search string must lie exactly X milliseconds before the begin time of the annotation matching the lower search string.
- **begin time - begin time < X milliseconds**: the begin time of the annotations matching the upper search string must lie less than X milliseconds before the begin time of the annotation matching the lower search string.
- **begin time - begin time > X milliseconds**: the begin time of the annotations matching the upper search string must lie more than X milliseconds before the begin time of the annotation matching the lower search string.
- **begin time - end time = X milliseconds**: the begin time of the annotations matching the upper search string must lie exactly X milliseconds before the end time of the annotation matching the lower search string.
- **begin time - end time < X milliseconds**: the begin time of the annotations matching the upper search string must lie less than X milliseconds before the end time of the annotation matching the lower search string.
- **begin time - end time > X milliseconds**: the begin time of the annotations matching the upper search string must lie more than X milliseconds before the end time of the annotation matching the lower search string.
- **end time - begin time = X milliseconds**: the end time of the annotations matching the upper search string must lie exactly X milliseconds before the begin time of the annotation matching the lower search string.
- **end time - begin time < X milliseconds**: the end time of the annotations matching the upper search string must lie less than X milliseconds before the begin time of the annotation matching the lower search string.
- **end time - begin time > X milliseconds**: the end time of the annotations matching the upper search string must lie more than X milliseconds before the begin time of the annotation matching the lower search string.
- **end time - end time = X milliseconds**: the end time of the annotations matching the upper search string must lie exactly X milliseconds before the end time of the annotation matching the lower search string.
- **end time - end time < X milliseconds**: the end time of the annotations matching the upper search string must lie less than X milliseconds before the end time of the annotation matching the lower search string.
- **end time - end time > X milliseconds**: the end time of the annotations matching the upper search string must lie more than X milliseconds before the end time of the annotation matching the lower search string.
- **No constraint**: there are no constraints.
- **Clear**: clear the current constraint.

◎ 다중 파일 검색 예: 마우스 제스처(O1 O2 O3)

| 연번 | 입모양 | 기호 | 사진 |
|----|---------------------|----|---|
| 1 | 입 모양에 관계없이 가장 작게 벌림 | O1 |  |
| 2 | 입 모양에 관계없이 중간정도로 벌림 | O2 |  |
| 3 | 입 모양에 관계없이 가장 크게 벌림 | O3 |  |

| 연번 | 기호 | 설명 | 빈도 | 사진 |
|----|-----|-------------------------------|-----|--|
| 1 | O1 | 입 모양에 관계없이 가장 작게 벌림 | 151 |  |
| 2 | O3 | 입 모양에 관계없이 가장 크게 벌림 | 137 |  |
| 3 | -O3 | 입을 다물었다가 열면서 입을 벌림 ex) 가능 '파' | 128 |  |
| 4 | H1 | 입이 좌우로 벌어지고 이가 살짝 보임 | 119 |  |
| 5 | -1 | 입을 다물 | 77 |  |



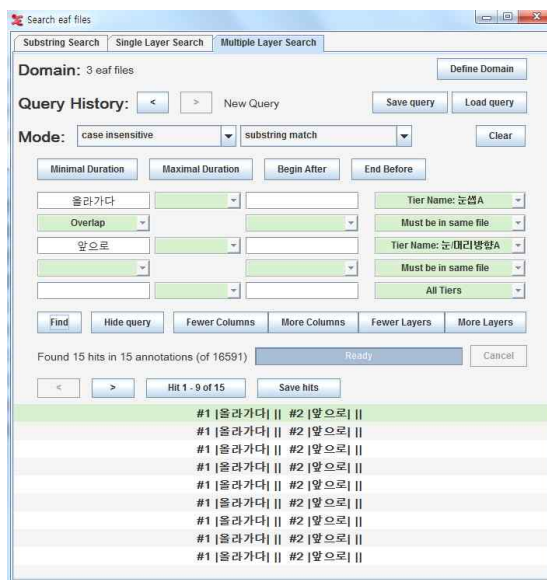
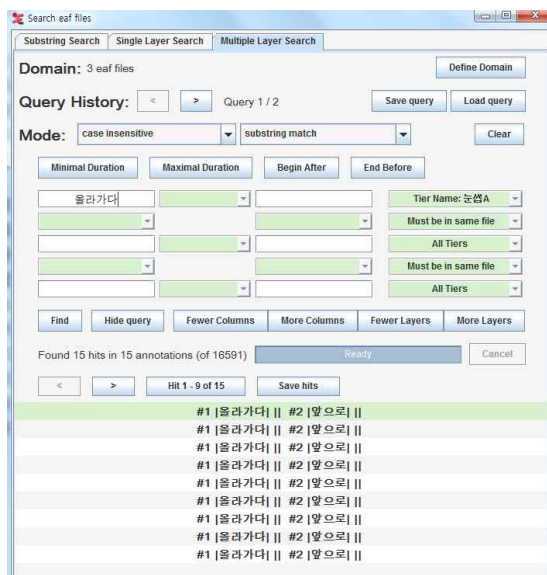
◎ 다중 파일 검색 예: 눈썹이 올라가고, 머리를 앞으로 내민 경우

(11) [한국수어] 농학교 교사^a 농학생^b 한국수어^c a가르치다^b

[한국어] 농학교 교사는 농학생에게 한국수어를 가르친다.

(12) [한국수어] 농학교 교사^a 농학생^b 한국수어^c b가르치다^a

[한국어] 농학교 학생은 교사에게 한국수어를 가르친다.



© Regular Expression 검색

Table A.1. Characters

| | |
|---------------------|---|
| <code>x</code> | The character <code>x</code> |
| <code>\\</code> | The backslash character |
| <code>\On</code> | The character with octal value <code>On</code> ($0 \leq n \leq 7$) |
| <code>\Onn</code> | The character with octal value <code>Onn</code> ($0 \leq n \leq 7$) |
| <code>\Omnn</code> | The character with octal value <code>Omnn</code> ($0 \leq m \leq 3, 0 \leq n \leq 7$) |
| <code>\xhh</code> | The character with hexadecimal value <code>0xhh</code> |
| <code>\uhhhh</code> | The character with hexadecimal value <code>0xhhhh</code> |
| <code>\t</code> | The tab character (<code>'\u0009'</code>) |
| <code>\n</code> | The newline (line feed) character (<code>'\u000A'</code>) |
| <code>\r</code> | The carriage-return character (<code>'\u000D'</code>) |
| <code>\f</code> | The form-feed character (<code>'\u000C'</code>) |
| <code>\a</code> | The alert (bell) character (<code>'\u0007'</code>) |
| <code>\e</code> | The escape character (<code>'\u001B'</code>) |
| <code>\cx</code> | The control character corresponding to <code>x</code> |

Table A.2. Character classes

| | |
|------------------------------------|---|
| <code>[abc]</code> | <code>a</code> , <code>b</code> , or <code>c</code> (simple class) |
| <code>[^abc]</code> | Any character except <code>a</code> , <code>b</code> , or <code>c</code> (negation) |
| <code>[a-zA-Z]</code> | <code>a</code> through <code>z</code> or <code>A</code> through <code>Z</code> , inclusive (range) |
| <code>[a-d[m-p]]</code> | <code>a</code> through <code>d</code> , or <code>m</code> through <code>p</code> : <code>[a-dm-p]</code> (union) |
| <code>[a-z&&[def]]</code> | <code>d</code> , <code>e</code> , or <code>f</code> (intersection) |
| <code>[a-z&&[^bc]]</code> | <code>a</code> through <code>z</code> , except for <code>b</code> and <code>c</code> : <code>[ad-z]</code> (subtraction) |
| <code>[a-z&&[^m-p]]</code> | <code>a</code> through <code>z</code> , and not <code>m</code> through <code>p</code> : <code>[a-lq-z]</code> (subtraction) |

Table A.3. Predefined character classes

ce: <http://docs.oracle.com/javase/7/docs/api/java/util/regex/Pattern.html>

374

REGULAR
EXPRESSION SEARCH

| | |
|----|---|
| . | Any character (may or may not match line terminators [http://docs.oracle.com/javase/7/docs/api/java/util/regex/Pattern.html#lt]) |
| \d | A digit: [0-9] |
| \D | A non-digit: [^0-9] |
| \s | A whitespace character: [\t\n\x0B\f\r] |
| \S | A non-whitespace character: [^\s] |
| \w | A word character: [a-zA-Z_0-9] |
| \W | A non-word character: [^\w] |

Table A.4. POSIX character classes (US-ASCII only)

| | |
|------------|---|
| \p{Lower} | A lower-case alphabetic character: [a-z] |
| \p{Upper} | An upper-case alphabetic character: [A-Z] |
| \p{ASCII} | All ASCII: [\x00-\x7F] |
| \p{Alpha} | An alphabetic character: [\p{Lower}\p{Upper}] |
| \p{Digit} | A decimal digit: [0-9] |
| \p{Alnum} | An alphanumeric character: [\p{Alpha}\p{Digit}] |
| \p{Punct} | Punctuation: One of !"#\$%&'()*+,-./:;<=>?@[\]^_`{ }~ |
| \p{Graph} | A visible character: [\p{Alnum}\p{Punct}] |
| \p{Print} | A printable character: [\p{Graph}] |
| \p{Blank} | A space or a tab: [\t] |
| \p{Cntrl} | A control character: [\x00-\x1F\x7F] |
| \p{XDigit} | A hexadecimal digit: [0-9a-fA-F] |
| \p{Space} | A whitespace character: [\t\n\x0B\f\r] |

Table A.5. Classes for Unicode blocks and categories

| | |
|--------------------|---|
| \p{InGreek} | A character in the Greek block (simple block [http://docs.oracle.com/javase/7/docs/api/java/util/regex/Pattern.html#abc]) |
| \p{Lu} | An uppercase letter (simple category [http://docs.oracle.com/javase/7/docs/api/java/util/regex/Pattern.html#abc]) |
| \p{Sc} | A currency symbol |
| \P{InGreek} | Any character except one in the Greek block (negation) |
| [\p{L}&&[^\p{Lu}]] | Any letter except an uppercase letter (subtraction) |

Table A.6. Boundary matchers

| | |
|----|-------------------------|
| ^ | The beginning of a line |
| \$ | The end of a line |
| \b | A word boundary |
| \B | A non-word boundary |

375

REGULAR
EXPRESSION SEARCH

| | |
|----|---|
| \A | The beginning of the input |
| \G | The end of the previous match |
| \Z | The end of the input but for the final terminator [http://docs.oracle.com/javase/7/docs/api/java/util/regex/Pattern.html#lt], if any |
| \z | The end of the input |

Table A.7. Greedy quantifiers

| | |
|--------|---|
| X? | X, once or not at all |
| X* | X, zero or more times |
| X+ | X, one or more times |
| X{n} | X, exactly n times |
| X{n, } | X, at least n times |
| X{n,m} | X, at least n but not more than m times |

Table A.8. Reluctant quantifiers

| | |
|----------|---|
| X?? | X, once or not at all |
| X*? | X, zero or more times |
| X+? | X, one or more times |
| X{n}? | X, exactly n times |
| X{n, }? | X, at least n times |
| X{n, m}? | X, at least n but not more than m times |

Table A.9. Possessive quantifiers

| | |
|----------|---|
| X?+ | X, once or not at all |
| X*+ | X, zero or more times |
| X++ | X, one or more times |
| X{n}+ | X, exactly n times |
| X{n, }+ | X, at least n times |
| X{n, m}+ | X, at least n but not more than m times |

Table A.10. Logical operators

| | |
|-----|---|
| XY | X followed by Y |
| X Y | Either X or Y |
| (X) | X, as a capturing group [http://docs.oracle.com/javase/7/docs/api/java/util/regex/Pattern.html#cg] |

Table A.11. Back references

| | |
|----|--|
| \n | Whatever the n th capturing group [http://docs.oracle.com/javase/7/docs/api/java/util/regex/Pattern.html#cg] matched |
|----|--|

Table A.12. Quotation

376

REGULAR
EXPRESSION SEARCH

| | |
|----|---|
| \ | Nothing, but quotes the following character |
| \Q | Nothing, but quotes all characters until \E |
| \E | Nothing, but ends quoting started by \Q |

Table A.13. Special constructs (non-capturing)

| | |
|------------------|---|
| (?:X) | X, as a non-capturing group |
| (?idsux-idsux) | Nothing, but turns match flags on - off |
| (?idsux-idsux:X) | X, as a non-capturing group [http://docs.oracle.com/javase/7/docs/api/java/util/regex/Pattern.html#cg] with the given flags on - off |
| (?=X) | X, via zero-width positive lookahead |
| (?!X) | X, via zero-width negative lookahead |
| (?<=X) | X, via zero-width positive lookbehind |
| (?<!X) | X, via zero-width negative lookbehind |
| (?>X) | X, as an independent, non-capturing group |

2017년 개정 한국수어 말뭉치 구축 지침

1. 2017년 개정 한국수어 말뭉치 구축 지침

1.1 개요

본 지침서는 2017년 개발한 한국수어 자료 통합 지원 시스템을 기반으로 한 한국수어 말뭉치를 구축을 위한 지침서이다. 본 지침서는 2015년 한국수어 말뭉치 기반 조사 및 시범 구축 사업에서 서울 지역의 농인 60명을 대상으로 언어 자료를 수집하고 수집한 언어 자료를 대상으로 번역 및 기본 전사를 실시한 결과와 2016년 비수지 중 마우스 제스처와 마우딩 전사를 실시한 결과를 토대로 제시한 지침서를 수정·보완하였다.

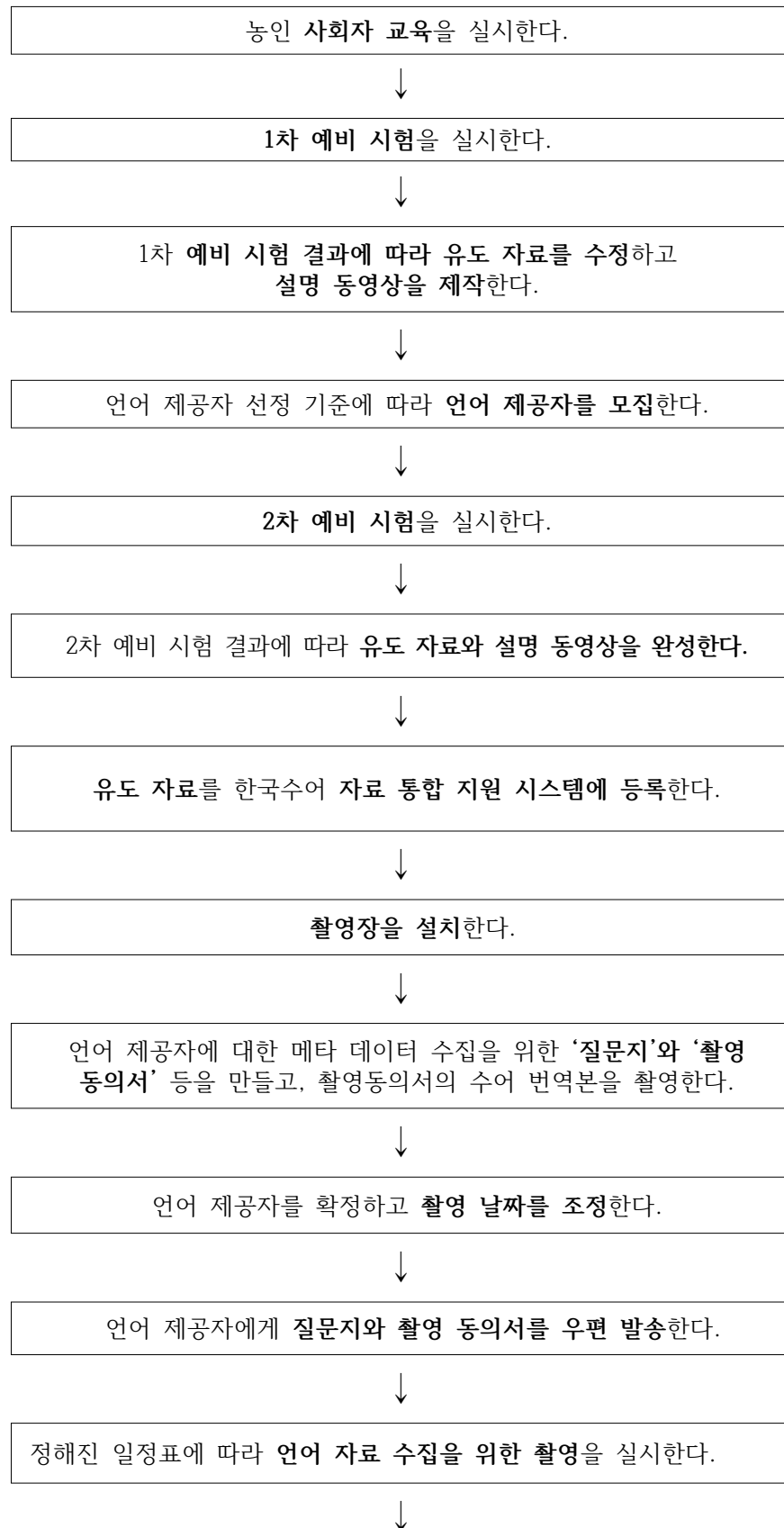
본 지침서에서는 전반적인 구축 절차를 표로 제시한 후, 언어 자료 수집에 필요한 유도 자료의 제작에서부터 언어 제공자 모집 및 선정, 언어 자료 수집을 위한 동영상 촬영과 언어 자료의 전사에 필요한 절차 및 한국수어 자료 통합 지원 시스템에 관한 등록 방법을 제공한다. 한국수어 자료 통합 지원 시스템은 수어 말뭉치 관리, 전사 템플릿 관리, 타입 관리를 통해 전사에 필요한 정보를 관리한다. 또한 전사 관리를 통해 전사 작업을 분배하고 엘란(Elan)을 통해 전사한 결과물을 시스템에 구축하여 자료를 통합한다. 통합한 전사 자료는 한국수어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

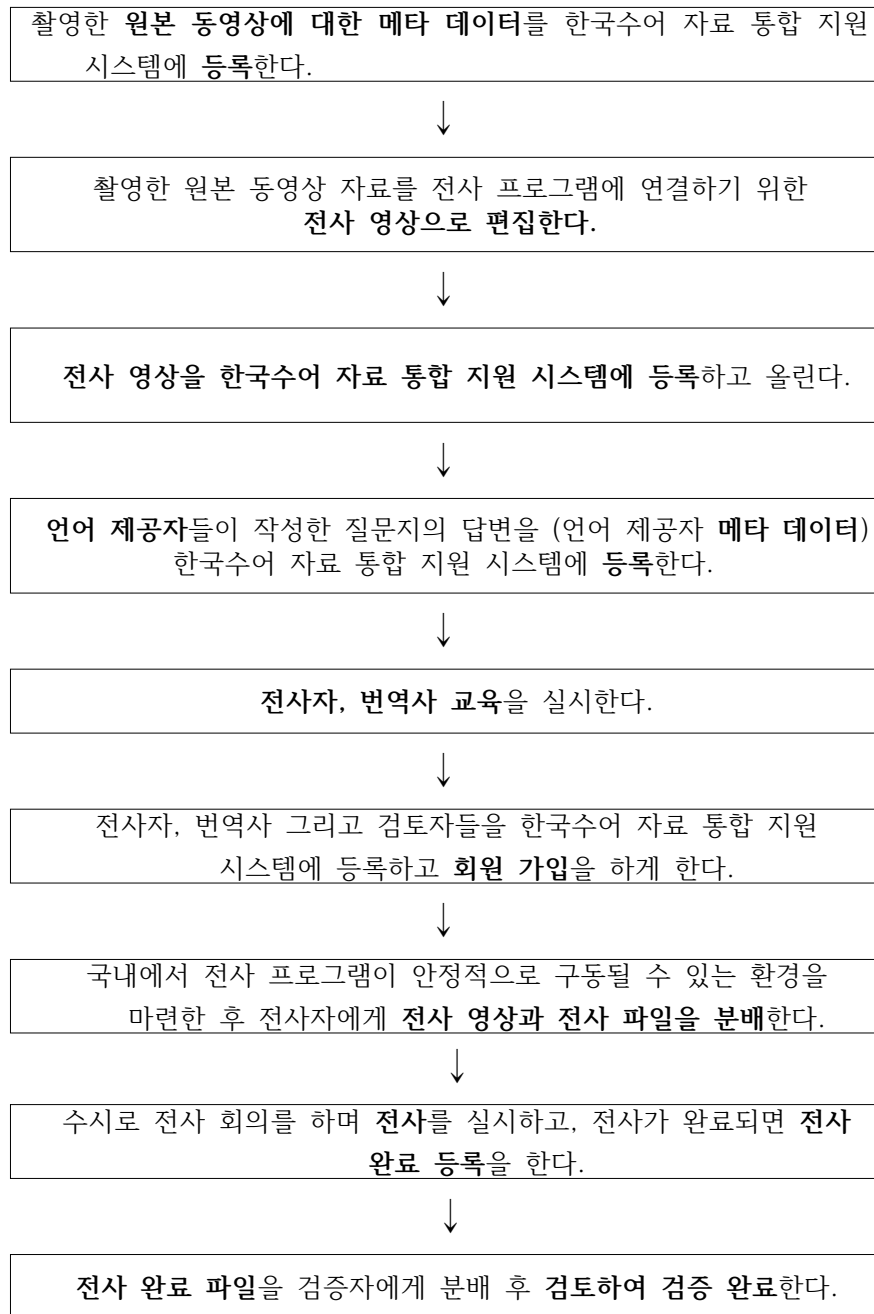
1.2 구축 절차

한국수어 말뭉치를 구축하기 위한 절차는 다음과 같다.

연구 목적에 따라 유도 자료를 선택하여 제작한다.







1.3 한국수어 말뭉치 수집 지침

1.3.1. 유도 자료

1) 유도 자료의 목적 및 필요성

수어를 연구하거나 한국수어 말뭉치를 구축하기 위해서는 자연스러운 상황에서 다양한 수어 어휘와 표현 등을 수집하여야 한

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

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

2) 유도 자료의 종류

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

(1) 언어 입력: 문자 언어와 수어

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

어진 수어 데이터는 자연스러운 데이터가 아니고, 맥락을 고려하지 않은 단순한 번역일 수 있다는 단점이 있다.

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

(2) 그림, 사진, 만화

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

있다는 단점이 있다.

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

③ 만화: 딱딱하지 않은 비공식적인 분위기를 만들어 낼 수 있기 때문에, 가장 자연스럽고 편안하게 데이터를 얻어낼 수 있다. 유명한 만화 인물을 보여 줄 경우 언어 제공자들이 바로 알아보는 효과가 있을 수 있다. 하지만, 만화 속에서만 사용하는 규칙이나 표현법이 있기 때문에 만화에 익숙하지 않은 사람들의 경우에는 내용을 이해하는 데 어려움이 따를 수 있다. 또한 유명한 만화는 대부분 저작권 문제가 있기 때문에 사용 시 꼭 유의해야 한다.

(3) 동영상과 애니메이션

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

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

(4) 주제, 동화와 우화

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

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

(5) 놀이

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

3) 유도 자료 제시 방법 및 등록

- (1) 유도 자료의 제시는 모든 언어 제공자에게 같은 방법으로 이루어져야 한다. 따라서 수어로 유도 자료를 설명하는 수어 동영상 제작한다.
- (2) 완성된 설명 동영상 파일을 파워포인트로 제작된 유도 자료에 삽입한다. 이때 동영상 파일의 용량을 줄이도록 한다. 또한 윈도우 PC로 유도 자료를 제시할 경우 유도 자료를 제시할 PC에서 유도 자료 제작 작업을 해야 한다. 그렇지 않으면 유도 자료를 옮길 때마다 영상과 파워포인트를 다시 연결해야 하는 문제가 생긴다.
- (3) 두 대의 컴퓨터에 각각의 언어 제공자에게 제시할 유도 자료를 불러낸다.
- (4) 사회자가 순서대로 각각의 언어 제공자에게 유도 자료를 제시한다.
- (5) 한 언어 제공자에게 유도 자료를 제시할 때 다른 언어 제공자에게는 아무런 그림이 없는 검은 화면을 제공한다.
- (6) 언어 제공자가 유도 자료의 내용을 잘 이해하지 못할 때는 사회자가 보충 설명을 한다.
- (7) 언어 제공자들의 대화가 단절될 때 사회자는 대화를 이어갈 수 있도록 질문을 제공한다.
- (8) 유도 자료로 제작하여 사용한 특정 유도 과제는 한국수어 자료 통합 지원 시스템에 등록한다.

- (9) 유도 자료 등록 창은 <그림 64>와 같으며, 등록은 다음과 같이 진행한다.

유도 과제 등록

1 유도 과제명

유도 과제 개요

2 유도 과제 파일

유도 과제 파일 추가

| 파일종류 | 파일붙임 | 파일설명 | 삭제 |
|--------|-------------|------|----|
| 수어설명영상 | 0KB / 100MB | | 삭제 |
| 과제영상 | 0KB / 100MB | | 삭제 |
| 과제사진 | 0KB / 10MB | | 삭제 |

3 등록 목록

<그림 64> 유도 자료 등록

- 가. 유도 자료 등록 화면으로 유도 자료명, 유도 자료 개요 등록 양식을 제공한다.
- 나. 유도 자료 파일 등록 영역으로 유도 자료 파일 추가 버튼 클릭 시 유도 자료 파일 붙임 양식이 생성된다.
- 파일 종류는 수어 설명 영상, 과제 영상, 과제 사진 중 선택한다.
 - 파일 붙임은 선택 버튼을 클릭하여 파일을 첨부하고, 삭제가 필요할 경우에는 삭제 버튼을 클릭하여 첨부한 파일을 삭제한다.
 - 파일 설명은 유도 자료 파일에 대한 파일 설명을 등록한다.
 - 삭제 버튼 클릭 시 해당 유도 자료 파일 양식을 삭제한다.

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

1.3.2. 언어 제공자 모집 및 선정

1) 언어 제공자 모집

- ① 언어 제공자는 농사회의 중심인 각 지역의 농아인협회를

통해 모집한다. 각 지역의 농아인협회를 통해 모집할 경우, 한국농아인협회(중앙회)를 통해 각 지역의 농아인협회에 언어 제공자 모집 협조 공문을 발송하면 효과적이다(부록 5 참조).

- ② 농인 사회의 페이스북과 같은 누리소통망(SNS)을 이용하여 언어 제공자 모집 광고를 한다.
- ③ 각 지역의 농인 사회의 리더를 통해 개별적 모집도 동시에 한다.
- ④ 언어 제공자의 기본 조건은 다음과 같다.
 - 19세 이상 농인
 - 가장 편하고 많이 사용하는 의사소통 방법이 한국수어임
 - 현재 언어 자료 수집 지역에 10년 이상 거주하고 있음
 - 일주일에 3번 이상 농인을 만남
 - 농학교를 졸업함

2) 언어 제공자 선정

- ① 모집된 언어 제공자는 연구진 회의를 통해 기본 조건을 충족했는지를 점검한다.
- ② 기본 조건이 충족된 언어 제공자를 대상으로 우선 다음과 같은 선발 조건에 따라 1차 선정을 한다.
 - 농부모, 농형제, 또는 농배우자가 있음
 - 초등학교 이전에 한국수어를 습득하였음
- ③ 1차 선정된 언어 제공자를 대상으로 나이와 성별을 고려하여 최종 언어 제공자를 확정한다.
- ④ 확정된 언어 제공자들과 촬영 날짜를 정한다.

3) 언어 제공자에 대한 자료 수집 및 촬영 동의서

- ① 선정된 언어 제공자의 메타 데이터를 수집하기 위한 질문지를 제작한다. 질문지는 개인 정보, 가족 정보, 사용 언어에 대한 정보, 학교 및 직업에 대한 정보, 농인 공동체 안에서

의 활동 등에 관한 내용으로 구성되어 있다. 질문지를 한국 실정에 맞게 수정하였으며 2017년 개정한 질문지를 <부록 6>에 제시한다.

- ② 언어 제공자의 초상권과 저작권 문제 해결을 위해 동의서를 제작한다. “영상 촬영, 배포, 판권 소유 및 자료 활용 동의서”(부록 7)와 “개인 정보 수집 · 이용 · 제공 및 조회 동의서”(부록 8)을 참조한다.
- ③ “영상 촬영, 배포, 판권 소유 및 자료 활용 동의서”는 언어 제공자의 편의를 위해 수어로 번역한 동영상 CD도 함께 제작한다.
- ④ 자료 수집을 위한 질문지, 영상 촬영, 배포, 판권 소유 및 자료 활용 동의서, 수어 번역본 CD, 개인 정보 수집·이용·제공 및 조회 동의서를 미리 언어 제공자에게 우편 발송하여 각자 집에서 작성해 줄 것을 부탁한다.
- ⑤ 언어 제공자가 작성해 온 질문지와 동의서는 촬영날 수거한다.
- ⑥ 촬영 당일은 언어 제공자와 사회자 관계에 대한 질문지(부록 9 참조)를 작성한다.

4) 언어 제공자 자료(메타 데이터 등록)

- ① 언어 제공자가 작성한 질문지 내용(언어 제공자 메타 데이터) 및 언어 제공자 사진을 한국수어 자료 통합 지원 시스템에 등록한다.
- ② 언어 제공자 등록 창은 <그림 65>과 같으며, 등록 방법은 다음과 같다.

1 언어 제공자 등록

개인 정보 가족 정보 언어 학교 및 직업 농인 공동체 안에서의 활동(예, 농인 클럽 내에서의)

2

12. 아버지 ☐ 농인 ☐ 난청인 ☐ 청인 ☐ CODA(농부모 청인자녀)

13. 어머니 ☐ 농인 ☐ 난청인 ☐ 청인 ☐ CODA(농부모 청인자녀)

14. 나는 부모님과 함께 3개월부터 7개월까지 살았다 ☐ 예 ☐ 아니오

15. 나는 다른 양육자의 돌봄을 받았다 ☐ 예 ☐ 아니오

16. 어머니와의 주된 의사소통 방법 ☐ 한국수어(KSL) ☐ 유사수어체계(Signed Korean) ☐ 몸짓이나 제스처 ☐ 말(음성 언어)

17. 아버지와의 주된 의사소통 방법 ☐ 한국수어(KSL) ☐ 유사수어체계(Signed Korean) ☐ 몸짓이나 제스처 ☐ 말(음성 언어)

18. 형제 자매 ☐ 농인 ☐ 난청인 ☐ 청인 ☐ 없음 ☐ 나보다 나이가 많다 ☐ 나보다 나이가 적다 ☐ 삭제

19. 형제 및 자매들과의 주된 의사소통 방법 ☐ 한국수어(KSL) ☐ 유사수어체계(Signed Korean) ☐ 몸짓이나 제스처 ☐ 말(음성 언어)

20. 가족이나 친척들 중 농인이나 난청인이 있습니까? ☐ 예 ☐ 아니오 ☐ 삭제

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

22. 자녀수 ☐ 농인 ☐ 난청인 ☐ 청인 ☐ CODA(농부모 청인자녀) ☐ 배우자 없음

23. 배우자와의 주된 의사소통 방법 ☐ 한국수어(KSL) ☐ 유사수어체계(Signed Korean) ☐ 몸짓이나 제스처 ☐ 말(음성 언어)

24. 자녀와의 주된 의사소통 방법 ☐ 한국수어(KSL) ☐ 유사수어체계(Signed Korean) ☐ 몸짓이나 제스처 ☐ 말(음성 언어)

3

이전 다음 목록

<그림 65> 언어 제공자 등록 창

- 가. 언어 제공자 등록 화면으로 언어 제공자 등록 항목이 많아 등록 분류별 탭으로 구분하여 언어 제공자 정보를 등록할 수 있도록 구성하였다. 해당 탭 클릭 시 해당 분류 등록 화면으로 이동한다.
- 나. 가족 정보 탭 클릭 시 나오는 화면으로 언어 제공자 가족에 대한 정보를 등록하는 화면으로 ‘18. 형제 자매’처럼 여러 개 정보를 입력하는 정보는 추가, 삭제 버튼을 두어 여러 개 정보를 추가, 삭제 할 수 있다.
- 다. 언어 제공자 버튼 영역으로 이전 버튼 클릭 시 가족 정보 이전 탭인 개인 정보로 이동한다. 다음 버튼 클릭 시에는 가족 정보 다음 탭은 언어 탭으로 이동한다. 목록 버튼 클릭 시에는 언어 제공자 목록 화면으로 이동한다.

5) 유의점

- ① 언어 제공자와 촬영 일자를 조절할 때는 충분한 시간을 확보하여야 하며, 하루에 1팀을 촬영하는 것을 권장한다.
- ② 약속된 촬영 날 하루 전에는 약속 시간을 다시 한 번 문자를 통해 알리고, 미리 제공한 질문지와 동의서를 당일 날

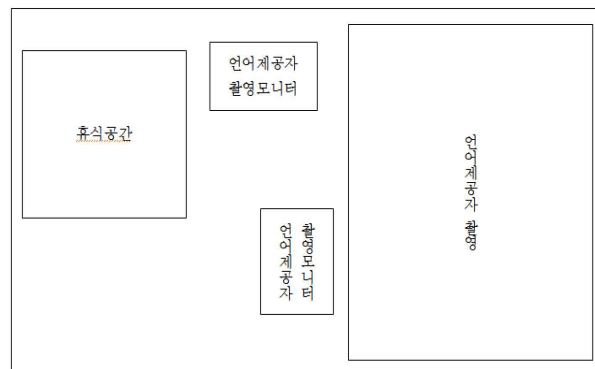
꼭 가져오도록 확인 점검한다.

- ③ A와 B 언어 제공자 쌍은 서로 친숙한 관계에 있는 사람으로 할 것을 권장한다.
- ④ 재촬영 또는 섭외한 언어 제공자의 돌발 상황을 대비한 언어 제공자도 미리 확보할 것을 권장한다.

1.3.3. 언어 자료 수집

1) 촬영실 조건 및 촬영 시스템

- ① 촬영실의 크기는 32m²(5.3m x 6m) 이상이어야 한다.
- ② 촬영실 창문은 빛을 차단할 수 있는 커튼이나 암막이 필요하다.
- ③ 촬영 중 청인은 촬영실에 들어갈 수 없으므로 촬영 감독과 보조 인력이 촬영을 준비하고 모니터할 수 있는 별도의 공간이 필요하다. 전체 촬영실의 배치는 <그림 66>을 참조한다.



<그림 66> 촬영실 전체 배치도

- ④ 언어 제공자들이 오기 편하고 연구진들이 24시간 접근이 가능한 곳이어야 한다.
- ⑤ 두 언어 제공자를 각각 촬영할 수 있는 카메라 2대와 두 언어 제공자와 사회자를 다 촬영할 수 있는 카메라 1대 등 총 3대 이상의 카메라가 필요하다.
- ⑥ 카메라의 메모리 용량은 64GB 이상으로 하는 것이 안전하다.
- ⑦ 카메라의 전원 연결이 중간에 끊어지지 않도록 유의한다.

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



<그림 67> 촬영실 모습

2) 농인 사회자의 역할 및 교육

- ① 언어 자료 수집을 위한 촬영 시에는 최대한 언어 자료를 풍부하게 유도해 내기 위해 농인 사회자가 필요하다.
- ② 사회자의 역할은 다음과 같다.
- 언어 제공자에게 유도 자료 파워포인트 자료를 제시한다.
 - 설명 동영상만으로 언어 제공자가 과제를 제대로 이해하지 못했을 경우 과제에 대한 설명을 한다.
 - 과제 소요 시간을 조절한다.
 - 대화가 자연스럽게 진행될 수 있는 분위기를 조성한다.
 - 언어 제공자의 대화를 중재하고 촉진한다.
 - 사회자 일지를 기록한다.
- ③ 사회자를 위한 교육 내용은 다음과 같다.
- 한국수어 말뭉치와 유도 자료의 내용 및 목적에 대한 이해
 - 언어 제공자가 동영상 설명 자료를 이해하지 못할 시 필요한 보충 설명 방법
 - 언어 제공자가 자연스러운 분위기에서 편하게 대화할 수 있는 분위기 조성 방법
 - 언어 제공자의 대화가 단절되었을 때 필요한 대화 촉진 전략
 - 두 언어 제공자의 수어가 균형 있게 산출될 수 있도록 조절하는 전략

- 각 과제의 소요시간을 균형 있게 조절하는 전략
- 그 외 사회자가 지켜야 할 태도

3) 동영상 촬영

- ① 언어 자료 수집을 위한 동영상 촬영의 진행을 위해서는 농인 사회자, 촬영 감독, 농인 사회자를 보조하는 보조원 등 최소 3인이 필요하다. 각각의 역할은 다음과 같다.
 - 농인 사회자는 유도 자료의 제시 등 언어 제공자들이 원활하게 수어 대화를 나눌 수 있도록 진행을 맡는다.
 - 촬영 감독은 카메라와 조명의 상태를 점검하고 각 언어 제공자의 수어가 잘 보일 수 있도록 촬영을 진행한다. 쉬는 시간에 메모리 카드를 교체하고 촬영 파일을 즉시 저장한다.
 - 보조원은 원활한 촬영의 진행을 위해 언어 제공자의 안내, 질문지와 동의서 등 필요 서류의 회수, 녹화가 잘 되는지의 확인, 촬영이 끝난 동영상 자료의 관리 등 보조적인 업무를 맡는다.
- ② 동영상 촬영에 들어가기 전에 질문지와 동의서를 받는다.
- ③ 원활한 언어 자료 수집을 위해 동영상 촬영에 들어가기 전에 두 언어 제공자와 사회자 간의 친밀감(rapport) 형성을 위해 간단한 다과를 하며 시간을 갖는 것이 좋다.
- ④ 다과 시간을 갖는 동안 보조원과 촬영 감독은 카메라와 유도 자료 제시를 위한 컴퓨터 등을 점검한다.
- ⑤ 촬영이 시작되면 촬영이 끝날 때까지 촬영실 안에 청인 관계자는 접근하지 않는다.
- ⑥ 사회자는 과제의 시작을 알리기 위해 각 과제를 시작할 때마다 벨을 누른다.
- ⑦ 하루에 1~ 2쌍의 언어 제공자를 촬영한다.
- ⑧ 촬영이 끝난 후 촬영된 영상에 문제가 없는지를 점검하고, 3대의 카메라로 촬영한 영상을 폴더별로 구분하여 저장함과 동시에 백업 영상을 만든다.

- ⑨ 촬영 일지를 기록한다.(부록 10 참조)
- ⑩ 검수를 마친 동영상들은 한국수어 자료 통합 지원 시스템에 등록한다.
- ⑪ 원본 영상 등록 창은 <그림 68>과 같고, 등록 방법은 다음과 같다.

원본 영상 등록

1

사업명: 2015년 한국수어 시범 말뭉치

제목: 11번째 원본 영상

저작자: 국립국어원 제작일자: 20150701 촬영장소: 스튜디오

확장자: MTS 해상도: 1280X720 파일 총 용량: 10 G

영상시간: 01 시 12 분 50 초 수어장르: 토론 유도 자료 종류: 이미지

2

카메라

카메라 추가

| 카메라 기종 | 카메라명 | 파일명 | 프레임수 | 화면화질 | 연결된층위 | 층위 | 삭제 |
|--------|------|-----|------|------|-------|----|----|
| | | | 25 | HD | | | |

연결 삭제

3

언어제공자

언어제공자 추가

| 이름 | 성별 | 출생년도 | 지역 | 삭제 |
|-----|----|-------|----|----|
| 최봉규 | 남성 | 1989년 | 서울 | |

삭제

4

등록 목록

<그림 68> 원본 영상 정보 등록

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

라. 원본 영상 정보 버튼 영역으로 등록, 목록 버튼을 제공한다.

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

4) 유의점

- ① 촬영 시 언어 제공자와 사회자가 앉을 의자는 촬영 시간 3시간 이상 사용을 해야 하므로 편안함과 동시에 수화를 하는 데 방해가 되지 않는 것으로 준비한다.
- ② 촬영에 들어가기 전 최대한 카메라의 메모리 용량은 촬영 중간에 끊어지지 않도록 넉넉한지를 확인해야 하며, 3대의 카메라에 똑같은 용량의 메모리 카드 3개를 사용해서 같은 시점에서 메모리 카드를 대체하도록 한다.
- ③ 언어 제공자에게 제공하는 모니터 2개는 똑같은 모델로 사용하는 게 중요하며 유도 자료 중 색깔을 정확히 봐야 하는 과제가 있다면 색깔이 변함없이 보이는지 모니터에서 직접 확인한다.
- ④ 사회자가 언어 제공자에게 유도 자료를 제시할 때, 언어 제공자의 모니터 2개를 동시에 조절할 수 있도록 설정을 따로 한다.
- ⑤ 사회자는 언어 제공자들이 서로 하고자 하는 말을 자유롭게 할 수 있도록 배려한다.
- ⑥ 촬영 감독은 카메라 등 촬영 환경에서 문제가 생길 때 즉시 확인하여 조치해야 하므로 농인이 할 것을 권장한다.
- ⑦ 촬영한 동영상은 촬영을 마친 후 바로 문제가 있는 지를 확인하여 오류가 발견될 시 재촬영 일정을 잡는 등 신속하게 대처할 것을 권장한다.
- ⑧ 촬영한 동영상의 편집 및 관리를 위해서 구축 팀에 수어가 가능한 촬영 및 편집 전문가를 참여시킬 것을 권장한다.

1.4 자료 처리 지침

1.4.1. 동영상 편집 절차 및 방법

영상 편집은 수어를 아는 편집자한테 맡긴다.

- ① 맥 PC에서 편집 가능한 형식으로 1차 변환을 한다.
- ② 3대의 카메라에서 촬영한 영상을 동기화한다.
- ③ 과제별로 동영상을 분절한다.
- ④ 각 동영상을 eaf 파일에서 불러내기에 최적화된 형식의 파일로 2차 변환을 한다.
현재 엘란(Elan) 프로그램 버전에서는 MPEG1(1280×720 비트율3000) 파일 형식이 가장 적합하다.
- ⑤ 최종 변환을 마친 동영상 파일을 eaf 파일과 묶어 전사를 위한 파일을 제작한다.
- ⑥ 엘란(Elan) 프로그램이 오류 없이 가장 잘 구동되는 환경은 다음과 같다.

Windows:

- JDS (Java Direct Show) —Java — Microsoft Media Foundation (.mp4, .m4a, .m4v(win &only), .wmv, .wma, .asf)
- DirectX/DirectShow (Windows Media Player) - QuickTime - JMF (Java Media Framework)

- ⑦ 편집을 마친 동영상 파일명은 언어 제공자 쌍-과제-촬영 카메라 순으로 나타낸다.

1.4.2 편집 동영상 검수 및 등록

편집이 완료된 동영상은 다음의 절차에 따라 검수 확인 절차를 거친다.

- ① 과제별로 오른쪽, 왼쪽, 중앙 카메라 촬영 파일이 모두 있는지 확인한다.
- ② 같은 과제의 영상에서 오른쪽, 왼쪽, 중앙 카메라로 찍은 3

- 개의 동영상의 시간이 동일한지 확인한다.
- ③ 엘란(Elan) 프로그램인 eaf 파일에서 오른쪽, 왼쪽, 중앙 3개 동영상 파일을 불러와 다음의 사항을 확인한다.
 - 앞뒤에 여유가 있는가?
 - 파일명과 영상의 내용이 맞는가?
 - ④ 문제가 있는 영상을 발견하면 기록하여 편집자와 함께 수정하도록 한다.
 - ⑤ 편집 검수 완료 동영상은 전사 영상이 되며, 이를 한국수어 자료 통합 지원 시스템에 등록한다.
 - ⑥ 전사 영상은 <그림 69>과 같이 시스템에 등록한다(자세한 내용은 ‘한국수어 자료 통합 지원 시스템 사용 설명서’ 참고)

전사 영상 등록

1

2

3

4

5

| | | | |
|----------|---------------------------|---------|-----------------|
| 제목 | | | |
| 원본 영상 정보 | 5번방 원본 영상 연결 | | |
| 전사 영상 불임 | 카메라명 | 영상품질 | 원본파일명 |
| | 원쪽 | 선택하세요 ▼ | KSLSL5XLT.MTS |
| | 오른쪽 | 선택하세요 ▼ | KSLSL5XRT.MTS |
| 유도 과제 | 유도 과제명 | | 등록자 |
| | 과제 번호 L'애플 수어와 현재 수어 | | 시스템 관리자 |
| | 등록일 | | 삭제 |
| | 2016. 11. 29. | | 삭제 |

등록
목록

<그림 69> 전사 영상 등록

1.5. 전사 지침

1.5.1. 개요

1) 기본 사항

- ① 전사를 위해 엘란(Elan) 프로그램을 사용한다.
- ② 전사를 위한 전사자 교육을 실시한다.
- ③ 전사의 일치도를 높이기 위해 공동 전사를 원칙으로 한다.
공동 전사가 어려운 경우 최소한 처음 5회 이상은 공동 전

사를 실시한다.

- ④ 전사의 일치도를 높이기 위해 주 1회 이상의 전사회의를 실시한다.
- ⑤ 전사 토큰 타입 매칭(token-type matching) 일치도를 높이기 위해 한국수어 자료 통합 지원 시스템을 사용한다.
- ① 언어 자료 영상에서 2명의 언어 제공자가 각각 찍힌 영상(full shot)의 왼편 언어 제공자는 A, 오른편 언어 제공자는 B이다.
- ② eaf 파일(엘란(Elan) 프로그램)에 전사를 마친 파일명은 한국수어 자료 통합 지원 시스템에서 내려 받은 파일명과 동일해야 하므로 절대 수정하지 않도록 한다.

2) 엘란(Elan)의 특성 및 예시

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



<그림 70> 엘란(Elan) 전사 예

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

1.5.2. 전사자 품질 관리

전사자들은 전사 경험이 없으며 외국과 달리 같은 장소에서 전사를 할 수 있는 상황이 아니므로 품질 관리에 특별한 신경을 써야한다.

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

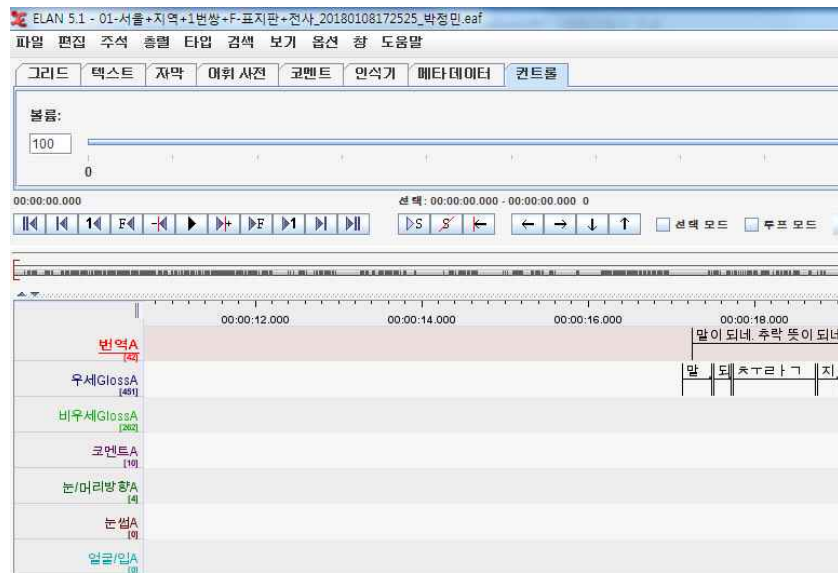
1.5.3. 기본 전사 규칙

1) 기본 원칙

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

2) 층렬

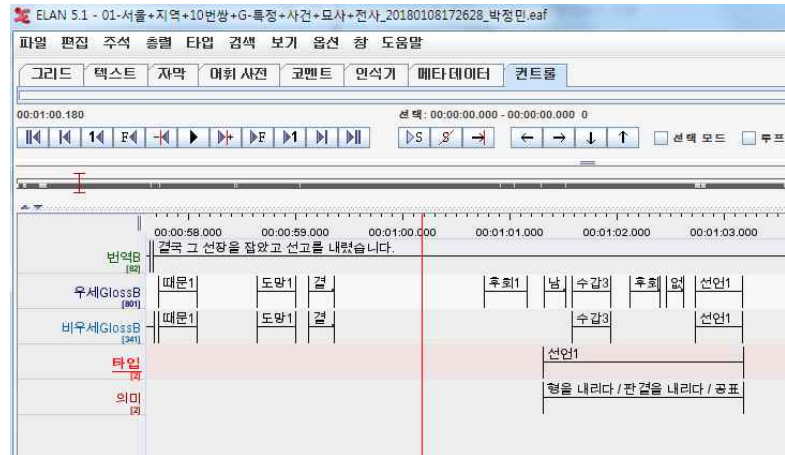
- ① 수어 데이터를 체계적으로 전사할 때는 전사 정보를 여러 층렬로 나누어 입력한다.
- ② 기본적인 층렬은 다음과 같다.
 - 번역 층렬
 - 수지 글로스 층렬
 - 비수지 층렬
 - 코멘트 층렬



<그림 71> 기본적인 층렬

- ③ 수지 글로스 층렬은 우세손과 비우세손으로 나눈다.
- ④ 비수지 층렬은 연구 목적에 따라 다양하게 나눌 수 있지만 전사의 편리를 위해서 눈과 머리 방향, 눈썹, 얼굴과 입 등으로 나누는 것이 좋다. 눈과 머리의 방향은 주로 함께 움직이며 문법적인 기능을 나타내므로 하나의 층렬을 할당하고, 눈썹은 의문문, 화제화 등 문장의 형태를 나타내므로 또 다른 층렬을 할당한다.
- ⑤ 수어 화자들은 수지 없이도 비수지 기호를 사용하여 의미 전달을 한다. 따라서 비수지 기호 층렬을 따로 입력하지 않고 수지 층렬만을 입력할 때에도 다음의 경우는 입력을 해야 한다.
 - 수어 없이 긍정의 의미로 고개를 끄덕인 경우 코멘트 층렬에 ‘끄덕임’으로 입력한다.
 - 수어 없이 부정의 의미로 고개를 흔든 경우 ‘가로저음’으로 입력한다.
- ⑥ 수지 글로스 층렬 아래에 의미 층렬을 둘 수 있다. 즉, <그림72>와 같이 타입 층렬과 의미 층렬을 만들고 타입 층렬에는 맥락에 따라 다양한 의미로 사용되는 해당 수어의 글로스를 입력하고, 의미 층렬에는 맥락에서 결정된 의미를 입력한다. 이렇게 기록된 의미 정보들은 수어 사전 제작 시에 활

용할 수 있다.



<그림 72> 의미 층렬

3) 수어 단어의 분절

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

4) 타입

- ① 전사 시 글로스를 붙인 개별적인 수어 어휘 하나하나를 토큰이라 하고, 이러한 토큰들 중에서 같은 형태의 토큰을 묶은 것을 타입이라 하는데, 이는 실제적이고 개별적인 것이 아니라 추상적인 개념이다.
- ② 어휘는 형태(form)와 의미(meaning)로 이루어지는데, 도상성이 강한 수어 어휘에서는 기저 이미지(underlying image)도

고려되어야 한다. 즉, 같은 형태, 같은 의미 또는 같은 그림일 경우 같은 타입으로 분류(예: [소개 <그림 73>])하였다. 형태는 같지만 기저 이미지가 다른 경우⁶⁾, 또는 핵심 의미는 같지만 기저 이미지가 다른 경우는 다른 타입(예: [먹다1 <그림 74>], [먹다2 <그림 75>], [먹다3 <그림 76>])으로 분류한다.



<그림 73> [소개]



<그림 74> [먹다1]



<그림 75> [먹다2]



<그림 76> [먹다3]

③ 기저 이미지(underlying image)란 수어 화자들의 머릿속에

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

그려지는 공통적인 이미지로 도상성이 강한 수어 어휘 기저에 자리하고 있는 이미지이다. 그 이미지는 수어의 의미 한 부분을 직접적으로 표현(예: [궁궐 <그림 77>])하기도 하고, 수어 의미와 간접적으로 연관(예: [중국 <그림 78>])되기도 한다 (Langer, G.König, S.Konrad, R.2007). 즉, [궁궐]이란 수어는 궁궐 지붕의 모양을 직접적으로 표현하고 있으며, [중국]이라는 수어는 중국을 대표할 수 있는 전통적인 옷 모양을 나타낸 것이다.



<그림 77> [궁궐]



<그림 78> [중국]

- ④ 수어 어휘를 나타내는 타입은 기본형, 변이, 형태론적인 수정으로 나눌 수 있다.
- 기본형은 관련 수어들 중에서 제일 단순하고 가장 흔히 사용되는 수어이며 맥락 없이도 독립적으로 사용될 수 있는 형태이다 (Mandel, 1997).
 - 형태론적인 수정은 수어의 형태를 확장하거나 정보를 추가하기 위해 수어의 한 형태가 바뀐 것으로, 도상성이 강한 수어에서 나타난다. 이는 하나의 기저 이미지 형태에 의미를 확장하거나 상세화하기 위하여 형태가 수정되는 경우를 말한다. 예) [공 <그림 79>], [큰 공 <그림 80>]



<그림 79> [공]



<그림 80> [큰 공]

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



<그림 81> [시골]



<그림 82> [시골]

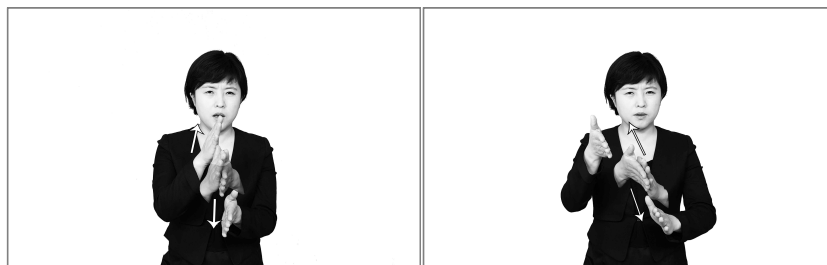
- ⑤ 같은 글로스를 사용하는 다른 타입의 수어일 경우는 ID글로스 바로 뒤에 숫자 1, 2, 3, 등으로 나타낸다. ID 글로스 뒤에 하나의 숫자만 붙어있는 경우 해당 타입의 기본형을 의미한다. 예) 없다1 <그림 83>, 없다2 <그림 84>, 없다3 <그림 85>



<그림 83> [없다1]



<그림 84> [없다2]

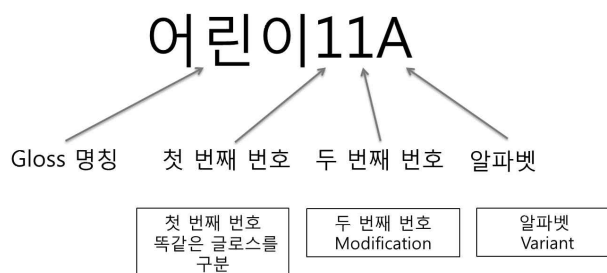


<그림 85> [없다3]

- ⑥ 형태론적인 수정은 해당 타입의 ID 글로스 다음에 1, 2, 3 등으로 나타낸다. 변이형은 해당 타입의 ID 글로스 다음에

알파벳 A, B, C 등으로 나타낸다.

- ⑦ 연구 목적에 따라 형태론적인 수정과 변이를 다 나타내지 않고 기본형으로 전사할 수 있다.
- ⑧ ID 글로스의 구조는 <그림 86>와 같다.



<그림 86> ID 글로스의 구조

- ⑨ 한국어로는 두 개의 단어이지만 수어로는 하나의 수어인 경우는 가장 중립적이고 가장 쉽게 떠올릴 수 있는 한국어 단어로 글로스를 부여한다. 예) 도움을 주다, 도움을 받다의 경우 글로스를 [돕다1]로, 주다, 받다는 [주다]로 글로스를 정한다.

5) 글로스(ID glosses)

(1)글로스(ID glosses) 확정 방법

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



<그림 87> [꼼꼼하다]

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

예) 잡다, 잡는다, 잡았다, 잡겠다 → [잡다]

예) 예쁜, 예쁘고, 예뻐다 → [예쁘다]

- ④ 글로스는 하나의 단어로 하는 것을 원칙으로 한다. 단 하나의 단어로 불충분한 경우 [버스가다]처럼 띄어쓰기 없이 붙여 쓴다. <그림 88>는 버스, <그림 89>는 버스가다로 글로스한 수어이다.



<그림 88> [버스]



<그림 89> [버스가다]

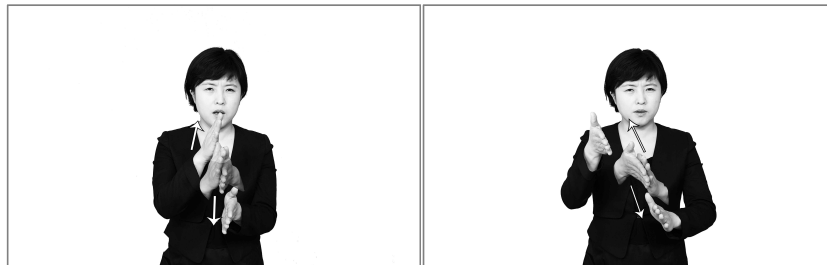
- ⑤ 수어 동의어(같은 의미이지만 형태가 다른 수어)로 글로스가 같은 경우는 글로스 뒤에 1, 2와 같은 숫자를 표시하여 [없다1 <그림 90>], [없다2 <그림 91>], [없다3 <그림 92>]과 같이 글로스를 붙인다.



<그림 90> [없다1]



<그림 91> [없다2]



<그림 92> [없다3]

⑥ 글로스로 사용한 한국어 단어가 동음이의어인 경우는 그 수어가 쉽게 유추될 수 있도록 글로스에 의미를 추가한다. 예를 들어 글로스가 ‘물건을 안에 넣고 보이지 않게 썬워 가리거나 돌려 말다.(표준국어 대사전 싸다01)’인 경우 [물건싸다 <그림 93>]로, ‘물건 값이나 사람 또는 물건을 쓰는 데 드는 비용이 보통보다 낮다(표준국어 대사전, 싸다05).’인 경우는 [값싸다 <그림 94>]로 글로스를 붙인다.



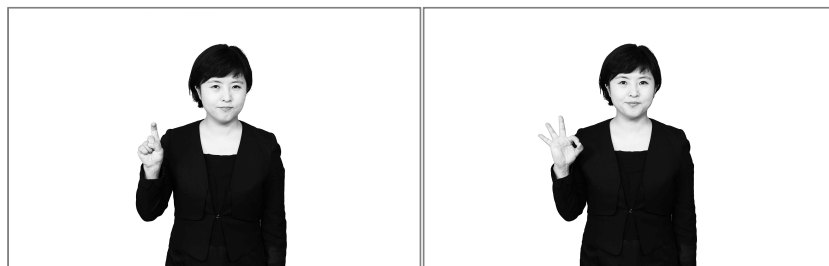
<그림 93> [물건싸다]



<그림 94> 값싸다

(2) 수와 단위

- ① 수를 나타내는 수어는 한국어 단어로 전사한다. 즉 1은 [일1]로, 2는 [이1], 10은 <그림 95>과 같이 [십1] 또는 [십2] 등으로 글로스를 붙인다.



<그림 95> [십1], [십2]

- ② 첫 번째, 두 번째와 같은 서수는 [첫번째1], [두번째1] 등으로 글로스를 붙인다.
 ③ 1년, 2분과 같이 수가 수어에 혼입되는 경우는 다음과 같이 글로스를 붙인다.

1년 → [일년1] <그림 96>

2분 → [이분1] <그림 97>



<그림 96> [일년1]



<그림 97> 이분1

- ④ 비우세손에 나타나는 'counting hand(손으로 세기)'는 펼쳐진 손가락의 수와 상관없이 모두 [counting hand <그림 102>]로 전사하고, 비우세손의 'counting hand'와 함께 나타나는 우세손의 수어는 해당 수어로 글로스를 붙인다.



<그림 98> [counting hand]

- ⑤ 알파벳 지문자로 나타내는 단위 수어는 한국어로 글로스를 붙인다.
'mm' → [밀리미터1]

(3) 명사와 동사

- ① 명사와 동사의 형태가 같은 수어는 이동이 없는 경우는 명사로 글로스하고, 이동이 있는 경우는 동사로 글로스를 붙인다. 예를 들어 <그림 103>처럼 동일한 수어 형태에서 움직임이 없으면 [버스1]로, 움직임이 있는 경우는 [버스가다1]로 글로스를 붙인다.



<그림 99> [버스1], [버스가다1]

- ② 일치동사 중 <그림 104>의 ‘야단치다’처럼 비우세손이 나타나지 않아도 되는 경우에는 다음과 같이 우세손과 비우세손 글로스를 따로 붙인다.

예) 우세 글로스: [야단치다1]

비우세 글로스: [사람4]



<그림 100> [야단치다]

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

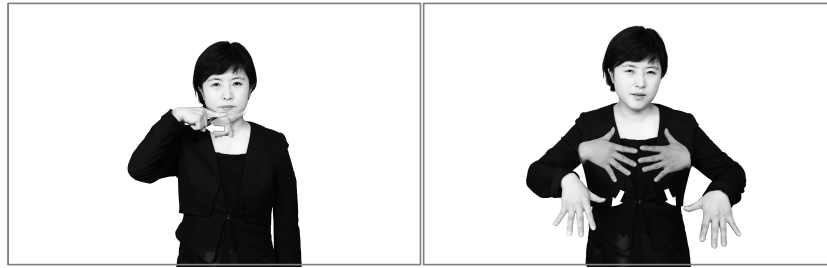
(4) 합성어

- ① 둘 이상의 수어가 결합된 형태이나 명확하게 하나의 단어로 인지되는 경우에는 하나의 글로스로 주석한다. 예를 들어 <그림 105>의 ‘손+돈’으로 된 수어는 두 수어를 묶어 분절 한 후에 [수당1]로 글로스를 붙인다.



<그림 101> [수당1]

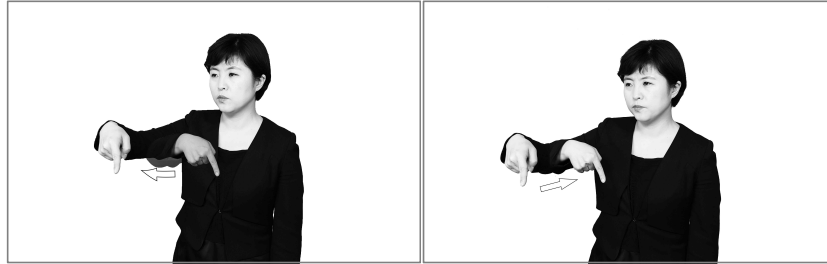
- ② 둘 이상의 수어가 결합된 형태에서 실제로 하나의 단어로 합성된 상태인지 의심스러운 경우는 두 개의 수어로 분리하여 각각 글로스를 붙인다. 예를 들어 <그림 106>의 분홍이라는 수어는 [빨강1], [약하다1]로 글로스를 붙인다.



<그림 102> [빨강1], [약하다1]

- ③ 둘 이상의 수어의 연쇄가 고정되어 하나의 합성어라고 인정하는 기준은 다음과 같다.
- 두 수어의 조합이 특정한 의미를 갖는다.
 - 조합된 수어의 의미가 예측되지 않을 수 있다.
 - 조합된 두 수어 사이에 다른 수어나 신호를 삽입할 수 없다.
 - 두 수어의 결합 시에 나타난다고 증명된 음운론적 특성이 적용된다.
- ④ Valli & Lucas(2000)가 제시한 합성어에서 나타나는 특성은 다음과 같다.
- 첫 번째 수어의 마지막과 다음 수어의 처음 사이에 움직임(M)이 들어간다. 그런데 여기서 주의할 점은 앞 수어의 마지막과 뒤 수어의 처음 사이에 첨가된 움직임(M)은 다음의 움직임에 동화가 된다는 것이다.

예를 들어 우리 자료의 [오가다 <그림 104>]라는 수어에서도 [오다]와 [가다 <그림 103>]라는 두 수어가 결합될 때 두 수어를 연결하는 움직임(M)이 첨가되면서 수어 화자들은 하나의 단어로 인식하였다.

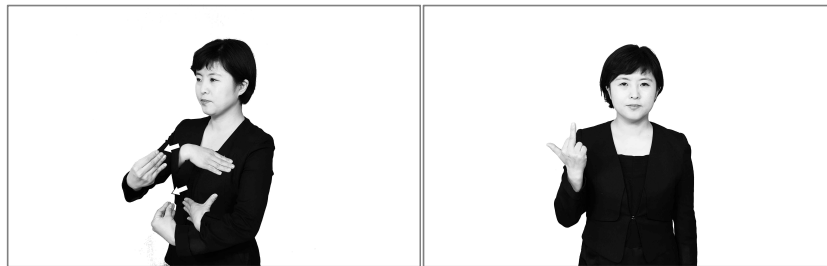


<그림 103> [오다], [가다]



<그림 104> [오가다]

- 두 단어가 결합할 때 움직임 사이의 접촉이 없는 정지(H)동작이 생략된다. 예를 들어 [안산]이라는 지명의 수어는 [편안]이라는 수어와 [산]이라는 수어가 결합하여 이루어진 복합어인데, [편안]에서 마지막 정지 동작(H)이 탈락하면서 [산]이라는 수어의 첫 움직임(M)과 [편안]라는 수어의 마지막 사이에 다른 움직임(M)이 첨가되었다(<그림 105>, <그림 106> 참조).

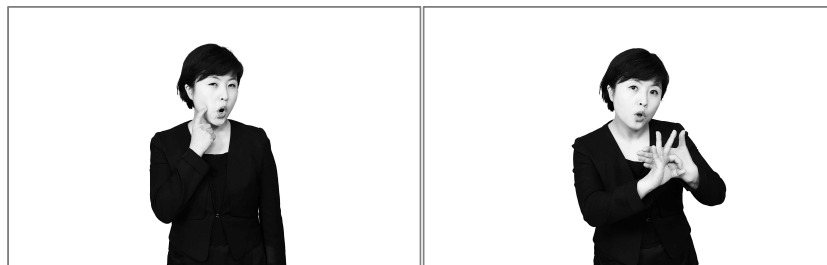


<그림 105> [편안], [산]

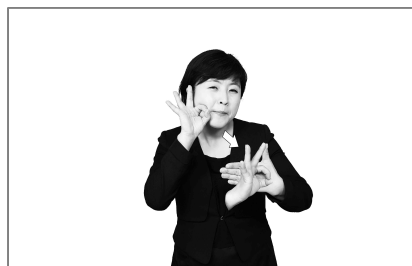


<그림 106> [안산]

- 동화 현상이 나타난다. 즉, 어떤 한 분절이 가까이에 있는 다른 분절의 특성을 갖게 되는 현상이 나타난다. 예를 들어 [만약]이라는 수어는 [거짓]+[면]이라는 수어로 이루어진 합성어인데, [거짓]이라는 수어를 할 때 뒤의 [-면]이라는 수어의 수형에 동화되어 1만 편 수형이 1지와 5지로 동그라미를 만들고, 2, 3, 4지를 편 [-면]이라는 수어의 수형과 같아지므로 역 행동화 현상이 일어난다. (<그림 107>, <그림 108> 참조)



<그림 107> [거짓], [-면]



<그림 108> [만약]

(5) 지시(가리키기)

- ① 검지만 편 수형(1형)인 경우 [지시1]로 글로스를 붙인다.
- ② 손가락을 다 편 수형(9형)인 경우 [지시2]로 글로스를 붙인다.
- ③ 자신을 지칭하는 경우는 수형에 따라 [지시1] 또는 [지시2]



















뒤에 [_1], 타인을 지칭하는 경우는 [_비1]를 붙여 글로스를 붙인다. 수어에서는 2인칭과 3인칭을 나타낼 때 손의 지시에서는 차이가 없고 눈 응시(eye gaze)와 맥락으로 나타내므로, 손의 글로스에서는 2인칭과 3인칭을 구분하여 글로스를 붙이지 않고 한 타입으로 보아 [_비1]로 글로스를 붙인다.

- ④ [지시1] 또한 [지시2]이 움직임이 있을 경우 글로스에 [호]를 붙인다. 움직임은 주로 반원으로 나타난다.
- ⑤ [지시1] 또한 [지시2]가 선을 따라 그리는 경우 아니면 하나의 모양/형태를 손끝으로 표현하는 경우 글로스에 [그림]을 붙인다.
- ⑥ 지시는 총 8개의 타입으로 나뉜다:
 [지시1_1], [지시1_비1], [지시1호], [지시1그림]
 [지시2_1], [지시2_비1], [지시2호], [지시2그림]
- ⑦ 한국수어에서 신체 부위를 의미하는 수어는 그 부위를 지시하며 나타난다. 따라서 수어 화자가 팔을 가리키며 ‘팔’이라는 수어를 사용할 경우 글로스 규칙을 적용하여 [팔1]로 글로스를 붙인다. 신체부위를 가리키는 수어를 [지시1_비1] 아니면 [지시2_비1]로 전사하지 않는다.

6) 생산적 수어(productive sign)

생산적 수어는 관습적인 의미를 가지지 않고 맥락 속에서 자연적으로 만들어진 수어로 맥락 안에서 이해되는 어휘이다. 그러므로 한 단어로 번역 또는 한 단어로 글로스를 정하기 어렵다. 그러나 이 수어들은 능숙한 수어 화자들 사이에서 이해가 가능하다. 그리고 이 수어들은 단순히 사물을 수형이나 수동으로 나타내는 것이 아니라, 생산적 수어에서 나타나는 체계적인 규칙이 적용된다. 이에 전사 시에는 수형에 기준을 두어 글로스를 정한다. 글로스의 예는 <표 43>와 같다.

<표 43> 생산적 수어의 글로스

| | 타입명 | 수형 | 설명 | 예시 |
|----|------|---|--|---|
| 1 | 일형1 |  | 1지만 편 형태로 얇고 긴 물체 또는 그림을 나타냄 |  |
| 2 | 육형1 |  | 1지와 5지의 손가락이 퍼진 상태로 하나의 형태를 표현 |  |
| 3 | 육형2 |  | 1지와 5지의 손가락이 퍼져 있으나 마디가 구부러진 상태로 수어를 시작하는 것 |  |
| 4 | 구형1 |  | 1, 2, 3, 4, 5지를 모두 펴고 손가락 사이의 벌어짐이 없는 손으로 납작하고 평평한 면 아니면 구획을 나누어 줄 때 선의 의미를 나타내는 것 |  |
| 5 | 구형1A |  | 1,2,3,4,5지를 모두 펴고 손가락 사이가 벌어진 손으로 납작하고 평평하며 아주 넓은 면 아니면 여러 개의 선을 나타내는 것 |  |
| 6 | 구형2 |  | 1, 2, 3, 4, 5지를 모두 편 손이지만 마디가 구부러진 형태 |  |
| 11 | 구형3 |  | 5지와 1, 2, 3, 4지 손가락의 접촉이 있으며 손가락의 마디가 구부러지지 않는 것 |  |
| 12 | 구형4 |  | 5지와 나머지 손가락 끝의 접촉이 있고 '0'형으로 보이는 것 |  |
| 13 | 십형1 |  | 1지와 5지의 손가락 끝이 동그랗게 붙어 있으며 나머지 세 손가락은 모두 펴져 있는 것 |  |

7) 토큰으로 처리하지 않아야 할 수어들(Non-tokens)

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

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

8) 기타

- ① 지문자에 글로스를 붙일 때는 자음과 모음을 분리하여 글로스를 붙인다. 예) ‘학교’ → [ㅎㅏ ㄱㅑ ㄱㅑ ㄴㅑ]
- ② 이름 수어인 경우는 대상자의 성과 이름 앞에 [SN]을 붙인다. 수어 이름이지만 그 사람의 한국어 이름을 모르는 경우는 [SN]으로만 글로스를 붙인다.
- ③ 수어가 앞뒤 수어의 영향을 받아 음운 변동이 나타난 경우에는 수어의 원형에 해당하는 글로스를 달고 코멘트 층렬에 ‘음운변동’으로 주석을 단다.
- ④ 수어가 완전히 조음되지 않은 경우는 글로스 대신에 [%]로 주석을 입력하여 거짓 시작임을 나타낸다.
- ⑤ 전사자가 수어로 인지했지만 무슨 수어인지 모르는 경우는 글로스 대신에 [###]로 주석을 단다. 그 후에 온라인 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.

- ⑥ 수어가 영상 화면 바깥에서 조음이 된다는지 하여 보이지 않는 경우에도 [###]로 입력한다.
- ⑦ 수어 화자가 실수로 수어를 잘못 표현했고, 그 수어가 사용되지 않는 수어일 때 [@]로 입력한다.
- ⑧ 코멘트 층렬에는 언어학적 유형 등 자유로운 텍스트를 입력할 수 있다. 그러나 코멘트 층렬에 입력 시에 너무 길게 입력을 하는 것은 좋지 않다.
- ⑨ 수어 화자가 미국수어를 하는 경우 '미국수화'로 입력하고, 국제수어를 할 경우 [국제수화]로 입력한다. 그리고 그 수어의 의미는 코멘트 층렬에 입력한다.

⑩ 제스처

가. 다음과 같은 기준에 의해 수어와 제스처를 구분한다.

- 수어가 아니다.
- 수어를 모르는 사람이 봐도 그 의미를 이해할 수 있다.
- 수어를 모르는 청인도 사용하는 몸짓이다.

나. 제스처는 코멘트 층렬에 [제스처]로 입력한다.

1.5.4. 번역

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

① 자유 번역(Free Translation)

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

② 정밀 번역(Narrow Translation)

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

1.5.5. 정밀 전사 규칙

2015년에 진행된 한국수어 말뭉치 기반 조사 및 시범 구축 사업의 지침서에서는 수지 글로스에 대한 세부 전사 규칙만을 제시하였다. 그러나 본 연구에서는 한국수어 말뭉치 기반 수어 연구 사례를 제시하기 위해 한국수어에서의 입의 역할 및 기능에 대한 분석을 하기로 하였다. 따라서 비수지 기호 중 입 움직임에 대한 연구를 위한 세부 전사 지침을 마련하였다.

1) 입 움직임 전사의 기본 원칙

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

언어로부터 온 것이 아니며, 마우딩은 음성 언어로부터 왔거나 빌려온 것이다(Crasborn, 2016).

입 움직임 전사의 기본 원칙은 다음과 같다.

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

2) 층렬

각 언어 제공자에 대한 기본 층렬인 ‘번역’, ‘우세 글로스’, ‘눈/머리 방향’, ‘눈썹’, ‘얼굴/입’, ‘코멘트(comment)’ 층렬 다음에 입 움직임과 관련한 8개의 층렬을 추가한다. 추가한 층렬은 마우딩(Mouthing) 관련 층렬로 ‘마우딩’, ‘스프레딩(Spreading)’, ‘마우딩 완성도’, ‘마우딩 분류’가 있으며, 마우스 제스처(Mouth Gesture) 관련 층렬로 ‘마우스 제스처’, ‘마우스 제스처 분류’ 층렬이 있다. 그 외에 ‘수어 분류’ 층렬과 ‘M 코멘트’ 층렬 등 8개의 층렬이 2명의 언어 제공자에게 각각 할당되어 총 16개의 층렬을

추가한다. 층렬의 이름은 띄어쓰기 없이 입력하며 추가된 층렬의 속성은 다음과 같다.

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

3) 분절

- ① 수어가 시작되고 끝나는 지점과 상관없이 마우딩과 마우스 제스처가 시작하고 끝나는 지점을 분절의 기준으로 삼는다.
- ② 마우딩 또는 마우스 제스처의 마지막 모양이 의미 없이 오랜 시간 유지되는 경우는 분절에 포함하지 않는다.
- ③ 마우딩 또는 마우스 제스처 관련 속성의 층렬은 해당 마우딩 또는 마우스 제스처의 분절 프레임과 동일하게 분절한다.

다. 즉, 스프레딩, 마우딩 완성도, 마우딩 분류 층렬은 마우딩 층렬 분절과 마우스 제스처 분류는 마우스 제스처 층렬의 분절과 동일하게 분절하다. M코멘트 층렬 역시 해당 마우딩 또는 마우스 제스처 분절과 동일하게 분절한다.

4) 마우딩 관련 층렬의 전사

(1) 마우딩

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

(2) 스프레딩(Spreading)

- ① 마우딩은 그 지속 시간이 길어 하나 이상의 수어에서 나타나기도 한다. 이러한 현상을 스프레딩이라고 하는데, 이 층렬은 이러한 스프레딩 현상을 확인하기 위한 층렬이다.
 - ② 마우딩이 지속되는 시간을 분절한 후에 다음의 특성에 따라 오른쪽, 왼쪽, 양쪽, 일치, 부분 등으로 입력한다.
- 오른쪽: 마우딩의 시작점이 의미와 관련된 수어의 시작점과 같으나, 아무 관련이 없는 다음 수어까지 지속되는 경우이다. 즉, 마우딩의 지속 시간이 해당 수어의 지속 시간보다 길어 다음 수어까지 덮는다. 예를 들어 [학교]라는 수어를 하면서 마우딩

으로 ‘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

- ① 마우딩은 해당 구어 단어의 일부분만 나타나기도 한다. 이 층렬은 마우딩이 해당 단어의 모든 음절이 나타났는지 부분적으로 나타났는지에 대한 완성도를 확인하는 층렬이다.
- ② 마우딩 완성도는 다음의 특성에 따라 완, 앞, 뒤, 앞뒤 등으

로 입력한다.

- 완: 구어 단어의 모든 음절을 조음한 경우이다. 예를 들어 ‘어린이’ ‘조심’ 등 모든 음절이 나타난 경우 [완]으로 입력한다.
- 앞: 구어 단어의 첫 음절만 조음한 경우이다. 예를 들어 ‘조심’에서 ‘조’ 음절만 나타난 경우 [앞]으로 입력한다.
- 뒤: 구어 단어의 마지막 음절만 조음한 경우이다. 예를 들어 ‘조심’에서 ‘심’ 음절만 나타난 경우 [뒤]로 입력한다.
- 앞뒤: 구어 단어의 첫 음절과 마지막 음절만 조음한 경우이다. 예를 들어 ‘자동차’에서 ‘동’ 음절은 탈락되고 ‘자차’라는 음절만 나타난 경우 [앞뒤]로 입력한다.

(4) 마우딩 분류

① 마우딩의 기능 및 특성에 따른 분류를 입력하는 층렬이다.

② 마우딩의 분류는 다음의 특성에 따라 M, M-add, M-spec로 입력한다.


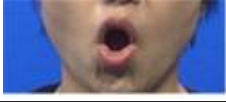
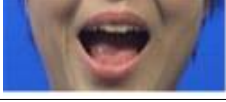






- M: 일반적인 마우딩으로 [학교]라는 수어를 하면서 ‘학교’라는 마우딩이 나타나는 경우이다.
- M-add: 수어에 해당하는 구어 단어의 마우딩이 나타나는 것이 아니라 다른 마우딩이 나타나면서 그 마우딩이 부가 정보를 제공해 주는 경우이다. 예를 들어 [학교]라는 수어를 하면서 마우딩은 ‘맞다’라고 하므로 “학교가 맞다.”라는 정보를 제공한다.
- M-spec: 수어와 함께 나타나는 마우딩이 해당 수어의 의미를 구체화 시키는 경우이다. 예를 들어 [새]라는 수어를 하면서 마우딩은 ‘비둘기’라고 하므로 [새]라는 수어의 의미를 구체화 시킨다.
- M-solo: 수어 없이 마우딩만 나타나는 경우이다.

5) 마우스 제스처 관련 층렬

(1) 마우스 제스처

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

<표 44> 입의 움직임을 나타내는 기본적인 기호

| 연번 | 입 모양 | 기호 | 사진 |
|----|----------------------|----|---|
| 1 | 입 모양에 관계없이 가장 작게 벌림 | O1 |  |
| 2 | 입 모양에 관계없이 중간 정도로 벌림 | O2 |  |
| 3 | 입 모양에 관계없이 가장 크게 벌림 | O3 |  |
| 4 | 입을 모으면서 입술을 내뺌 | B |  |
| 5 | 입이 좌우로 벌어지고 이가 보임 | H1 |  |
| 6 | 입이 좌우로 크게 벌어지고 이가 보임 | H2 |  |
| 7 | 입을 다물 | -1 |  |
| 8 | 입을 더 세게 다물 | -2 |  |
| 9 | 양쪽 입 꼬리가 약간 내려감 | ^1 |  |
| 10 | 양쪽 입 꼬리가 많이 내려감 | ^2 |  |
| 11 | 왼쪽 입 꼬리만 올라감 | < |  |

| 연번 | 입 모양 | 기호 | 사진 |
|----|------------------|-----|---|
| 12 | 오른쪽 입 꼬리만 올라감 | > |  |
| 13 | 왼쪽 볼에 바람을 넣음 | (- |  |
| 14 | 오른쪽 볼에 바람을 넣음 | -) |  |
| 15 | 양 볼에 바람을 넣음 | (-) |  |
| 16 | 양 볼에 바람을 뺌 |)-(|  |
| 17 | 볼 안에서 혀로 볼을 밀어냄 | o> |  |
| 18 | 아랫니만 보임 | M |  |
| 19 | 윗니만 보임 | W |  |
| 20 | 지속 또는 반복적으로 움직임 | ~ | |
| 21 | 양악을 엇갈리게 움직임 | Z |  |
| 22 | 양 볼의 바람을 내뿜음 | (=) |  |
| 23 | 입 안에서 혀를 약간 다물 | C |  |
| 24 | 내민 혀를 입 안으로 집어넣음 | UJ | |

- ③ 입의 모양은 기본 기호를 조합하여 입력한다. 예를 들어 ‘입술을 내밀고 있는데 이가 보이는 경우’는 [BH]로, ‘입을 다물고 양쪽 입꼬리가 약간 내려간 경우’는 [-^1]로, ‘이가 보

이는 상태로 왼쪽 입꼬리만 올라 간 경우’는 [<H]로, ‘입이 약간 벌어지고 혀가 조금 나온 경우’는 [O1U1]로 ‘혀를 여러 번 내미는 경우’는 [U1~]로 입력한다.

- ④ 연속적으로 일어나는 입의 움직임은 움직임의 순서대로 기호를 조합하여 입력한다. 예를 들어 ‘입을 약간 벌리고 혀를 내밀었다가 집어 넣으면서 바람을 내뿜는 경우’는 [O1U](=)]로, ‘입을 다물었다가 이가 보이게 양쪽으로 벌리면서 바람을 내뿜는 경우’는 [-H1(=)]로 입력한다.

(2) 마우스 제스처 분류

- ① 마우스 제스처의 특성에 따른 분류를 입력하는 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 복사(Eco)를 하는 경우도 이 분류에 속한다. 이 분류에 속하는 입의 움직임은 [가능하다]라는 수어에서처럼 손의 움직임은 없이 ‘입을 닫았다가 열면서 바람을 순식간에 내는(파)는 입의 움직임만으로도 의미를 나타내기도 한다.
 - 4: 입이 실제 입의 역할을 하는 경우이다. 예를 들면 [먹다]

[씹다] [외치다]와 같은 수어를 하면서 실제 먹거나 씹거나 외치는 입의 모양을 하는 것이다.

- W: 입의 모양이 독립적으로 나타나는 것이 아니라, 전체 얼굴 표정(Whole-face)의 한 부분으로 나타나는 경우이다. 이 분류의 입 움직임은 주로 감정을 표현한다. 예를 들면 [징그럽다]라는 하는 수어를 하면서 나타나는 입의 움직임은 얼굴 전체의 움직임과 함께 나타나므로 이 분류에 속한다.

6) 수어의 분류(Class) 관련 층렬

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

입 모양 또는 입 제스처와 함께 나타난 수어의 품사 중에 명사와 동사를 구분하여 입력하며 동사는 단순 동사, 일치 동사, 공간 동사로 분류하여 주석한다.

7) 기타


- ① 수어와 함께 나타나는 마우스 제스처 중에 독립적으로 사용될 수 있으며 명확한 의미를 가진 것들이 있다. 이러한 마우스 제스처가 수어 없이 독립적으로 나타난 경우 ‘M 코멘트’층렬에 [!!!]를 입력한다

예를 들어 [가능하다]라는 수어의 경우 입술의 긴장을 유지한 채로 입을 닫았다가 여는 마우스 제스처가 동반되지만, 이 마우스 제스처가 수어 없이 홀로 나타났을 때 ‘M 코멘트’층렬에 [!!!]를 입력한다.

- ② 입의 움직임이 있으나 입의 일부 또는 전체가 가려져 판단하기 어려운 경우는 ‘M 코멘트’ 층렬에 [@@]를 입력한다.

- ③ 화자가 어떤 의도를 가지고 입을 움직였다가 주저하거나 도로 입을 다물기도 한다. 이때는 해당 티어에 주석 대신에 [%]를 입력하여 거짓 시작임을 나타낸다.

<부록 5> 언어 제공자 모집 공문

| | | | |
|---|-----------------------------------|--------|--|
|  사단법인 한국농아인협회 | <h2 style="margin: 0;">업무협조전</h2> | 담당부서 | |
| | | 담당자 | |
| | | E-mail | |
| 주소 : 우)133-834 서울특별시 성동구 아차산로 7길 21(성수2가3동) 보라빌딩 4층 전화 : 02)461-3562 / 팩스 : 02)461-2651-2 | | | |

시행일자 : 2015. 5. 28.

수 신 서울농아인협회장

제 목 언어제공자(농인) 추천 의뢰 협조 요청

-
1. 귀 회의 투명한 발전을 기원합니다.
 2. 국립국어원에서 각종 수어 사전 편찬과 수어 연구의 기반을 조성하기 위한 목적으로 「한국어 말동치(코퍼스) 조사 및 시범 구축」 연구사업을 진행하고 있습니다.
 3. 이와 관련, 언어제공자(농인) 추천을 의뢰하니 산하 지부에도 시달하여 불입의 추천기준에 적합한 언어제공자(농인)를 추천받을 수 있도록 협조하여 주시기 바랍니다.

- 가. 제목: 언어제공자(농인) 추천 의뢰
 나. 추천 인원: 각 지부별 4명(서울 거주 10년 이상인 농인)
 다. 추천기간: 2015. 5. 28(목) ~ 6. 5(금)
 라. 요청 사항
 - 추천 조건에 해당하는 농인 추천 의뢰
 - 회신방법: e-mail(ley-512@hanmail.net)
 - 회신기한: 2015. 6. 5(금)까지
 마. 문의: 이 (농인)010- , 이 (청인)010-

- 붙임 1. 언어제공자 추천 조건 1부.
 2. 언어제공자 추천 양식 1부.
 3. 개인정보동의서 1부. (메일발송)끝.

사단법인 한국농아인협회[직인생략]

<부록 6> 2017 자료 수집을 위한 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)
8. 청력상태/청력손실 시기
☐ 농 (☐ 태어날 때부터 ☐ ____살 이후부터)
☐ 난청 (☐ 태어날 때부터 ☐ ____살 이후부터)
9. 청력보장구 착용 여부
☐ 인공와우 시술을 받았다.
☐ 규칙적으로 보청기를 사용한다.
☐ 가끔 보청기를 사용한다.
☐ 보청기등의 보장구를 사용하지 않는다.

가족 정보

10. 아버지의 청력상태: ☐ 농인 ☐ 청인

11. 어머니의 청력상태: ☐ 농인 ☐ 청인

12. 부모와의 동거 유무: ☐ 예 ☐ 아니오

13. 부모 외의 양육자: [12번 질문에 '아니오'라고 답한 경우]

☐ 할아버지 ☐ 할머니 ☐ 친척 ☐ 기타 _____

14. 어머니와의 주된 의사소통 방법

☐ 한국수어 ☐ 한국어대응수어+말 ☐ 제스처 ☐ 말 ☐ 필담
☐ 흠싸인 ☐ 기타 _____

15. 아버지와의 주된 의사소통 방법

☐ 한국수어 ☐ 한국어대응수어+말 ☐ 제스처 ☐ 말 ☐ 필담
☐ 흠싸인 ☐ 기타 _____

16. 형제 및 자매의 청력상태/나이

- ① ☐ 농인 ☐ 난청인 ☐ 청인
- ② ☐ 농인 ☐ 난청인 ☐ 청인
- ③ ☐ 농인 ☐ 난청인 ☐ 청인
- ④ ☐ 농인 ☐ 난청인 ☐ 청인
- ⑥ ☐ 형제 및 자매 없음

17. 형제 및 자매와의 주된 의사소통 방법

☐ 한국수어 ☐ 한국어대응수어+말 ☐ 제스처 ☐ 말 ☐ 필담
☐ 흠싸인 ☐ 기타 _____

18. 가족 및 친척의 청력상태

| | | |
|-------|-----------------------------|------------------------------|
| 할머니 | <input type="checkbox"/> 농인 | <input type="checkbox"/> 난청인 |
| 할아버지 | <input type="checkbox"/> 농인 | <input type="checkbox"/> 난청인 |
| _____ | <input type="checkbox"/> 농인 | <input type="checkbox"/> 난청인 |
| _____ | <input type="checkbox"/> 농인 | <input type="checkbox"/> 난청인 |

19. 배우자의 청력상태

☐ 농인 ☐ 난청인 ☐ 청인 ☐ 배우자 없음

20. 배우자와의 주된 의사소통 방법

☐ 한국수어 ☐ 한국어대응수어+말 ☐ 제스처 ☐ 말 ☐ 필담
☐ 흠싸인 ☐ 기타 _____

21. 자녀의 청력상태

☐ 농인 (----- 명)
☐ 난청인 (----- 명)
☐ 청인 (----- 명)
☐ 자녀 없음

22. 자녀와의 주된 의사소통 방법

☐ 한국수어 ☐ 한국어대응수어+말 ☐ 제스처 ☐ 말 ☐ 필담
☐ 흠싸인 ☐ 기타 _____

언어

23. 언어 제공자(본인)의 주된 의사소통 방법

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

24. 한국수어 습득 시기: ____살 때부터

25. 한국수어 습득 장소

- ☐ 농학교
☐ 농인 교회
☐ 농아인협회
☐ 기타 _____

26. 한글지문자(지화) 사용 유무

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

27. 지역수어 사용: 지역명 _____

28. 농(난청)인과의 정기적인 만남 유무: ☐ 예 ☐ 아니오

29. 본인의 한국어 능력 정도

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

30. 외국수어 사용 여부

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

학교 및 직업

31. 출신 학교

출신학교1:_____ (제일 오래 다녔던 학교 이름을 적어주세요)

주된 의사소통 방법:

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

출신학교2:_____ (학교 이름을 적어주세요)

주된 의사소통 방법:

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

출신학교3:_____ (학교 이름을 적어주세요)

주된 의사소통 방법:

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

32. 학교 기숙사 입사 유무: ☐ 예 ☐ 아니오

33. 학교 기숙사 입사에 대한 추가정보

[33번에서 '예'라고 답한 경우]

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

34. 최종 학력: ☐ 초등학교 졸업 ☐ 중학교 졸업
 ☐ 고등학교 졸업 ☐ 전공과 졸업
 ☐ 전문대학 졸업 ☐ 4년제 대학교 졸업
 ☐ 대학원 석사 졸업 ☐ 대학원 박사 졸업
 ☐ 기타 _____

35. 직업훈련 기관명/지역

기관명 _____ 지역 _____

36. 대학명/전공

대학명 _____ 전공 _____

37. 직장에서 작업 형태

☐ 팀에서 일함 ☐ 혼자 일함

38. 직장 퇴직 여부: ☐ 예 ☐ 아니오

[38번에서 '예'라고 답한 경우] 마지막 직장에 관한 다음 질문에 답하십시오

39. 마지막 직장 정보

나는 _____년부터 _____로써(직종) 일했다.

40. 취미와 특별히 흥미가 있는 영역(예, 축구, 자동차, 요리, 독서, 등산 등)

농 사회에서의 활동

41. 한국수어 강사 경험

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

42. 한국수어 강사 활동 정도

[43번에서 ‘예’라고 답한 경우] ☐ 자주 ☐ 가끔 ☐ 드물게

43. 수어 예술분야 활동에서 수어사용 여부(예, 수어 예술분야: 연극, 시, 퍼포먼스 등)

☐ 수어를 사용한다 ☐ 수어를 사용하지 않는다

44. 우세손

| | 오른손 사용 | 왼손 사용 | 어느 때는 왼손 사용 |
|-------|--------------------------|--------------------------|--------------------------|
| 쓰기 | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 양치질하기 | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 칼질하기 | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 수어할 때 | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

47. 마지막으로 하고 싶은 말

날짜와 서명 _____

<부록 7> 영상 촬영, 배포, 판권 소유 및 자료 활용 동의서

영상 촬영, 배포, 판권 소유 및 자료 활용 동의서

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

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

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

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

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

자료제공자(을)

성명: _____

주소: _____ 우편번호: _____

전화: _____ Email: _____

서명: _____ 날짜: 2015 / ____ / ____

(위 내용을 확인하고 동의합니다.)

국립국어원 한국수어 말뭉치 시범 구축 사업팀 귀중

<부록 9> 언어 제공자와 사회자 관계에 관한 질문지

언어제공자와 사회자 관계에 관한 질문지

자료수집 날짜: _____
 사회자: _____
 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 ☐ 아니요
 ('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 ☐ 아니요
 ('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 ☐ 아니요
 ('예'라고 답한 경우) 어떤 친척 관계인지(예, 조카 등) _____

<부록 10> 촬영 일지

| 촬영 일지 | |
|---------------|--|
| 일시 | 2015. 6. 29. 8:15 - 18:00 |
| 장소 | 청담동 스튜디오 |
| 참석자 | 이OO, 신OO, 배OO [언어 제공자: 박OO, 목OO] |
| 내용 및 결과 | <ul style="list-style-type: none"> ● 주요업무: 언어 제공자 2팀 촬영 - 오전 촬영 시간: 10:00-12:50 - 오전 휴식 시간: 10분 - 언어 제공자 한 명이 늦게 와서 촬영 시작이 지연되었다. 언어 제공자들이 '자료 수집을 위한 질문지' 및 '촬영동의서'를 미리 받아보지 못해서 촬영실에서 작성하느라 시간이 지체되었다. - 오후 촬영 시간: 13:45-16:40 - 오후 휴식시간: 7분 - 촬영을 마친 후 언어 제공자들이 못다 한 이야기가 있다고 하여 추가로 10분 정도 촬영하였다. ● 촬영 장비 설치시 언어 제공자의 머리 윗부분의 공간을 여유 있게 하였으나 대화 중 손을 번쩍 드는 부분이 몇 번 나와 손이 카메라의 화면 밖으로 나갔다. ● 언어 제공자가 질문지를 미리 작성해오지 않을 경우 현장에서 다시 질문지를 설명하고 작성해야 하니 여기에 소요되는 시간이 많아진다. ● 언어 제공자가 작성해 온 질문지를 보니 잘못 작성한 부분이 있다. 작성해 온 질문지를 받아, 다시 한 번 확인해야 한다. |

연구 책임자 원성옥

공동 연구원 남기현, 변강석, 이은영, 허일, 홍성은

연구 보조원 김다해, 박정민

담당 연구원 이수미(학예연구사)
김민정(연구원)

발행인: 송철의

발행처: 국립국어원

서울시 강서구 금낭화로 154

전화 02-2669-9775, 전송 02-2669-9737

인쇄일: 2017년 12월 15일

발행일: 2017년 12월 15일

인 쇄:

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