

국립국어원 2021-01-33

발간등록번호

11-1371028-000867-01

2021년 말뭉치 함의 분석 및 연구

연구책임자
송상현



제 출 문

국립국어원장 귀하

국립국어원과 체결한 연구용역 계약에 따라 '2021년 말뭉치 함의 분석 및 연구 보고서'를 작성하여 제출합니다.

■ 사업 기간: 2021년 5월 ~ 2021년 12월

2021년 12월 17일

연구책임자: 송상헌(고려대학교)

연구 기관 고려대학교 산학협력단
 성신여자대학교 연구산학협력단
 (주)나라지식정보

연구책임자 송상헌
공동연구원 강아름, 김경민, 김태우, 이도길, 장하연, 최윤지, 김일환, 박진
 호, 정연주, 정슬아, 조경찬, 유현조, 조정아

연구 기관: 고려대학교 산학협력단 · 성신여자대학교 연구산학협력단 · (주)나라지식정보

연구진	
책임 연구원	송상헌(고려대학교)
공동 연구원	강아름(충남대학교)
	김경민(고려대학교)
	김태우(부산대학교)
	이도길(고려대학교)
	장하연(부산외국어대학교)
	최윤지(인하대학교)
	김일환(성신여자대학교)
	박진호(서울대학교)
	정연주(홍익대학교)
	정슬아(성신여자대학교)
	조경찬(나라지식정보(주))
	유현조(나라지식정보(주))
조정아(나라지식정보(주))	
연구 보조원	김다미(서울대학교)
	김성태(고려대학교)
	김푸른솔(서울대학교)
	노강산(고려대학교)
	백인영(서울대학교)

	신운섭(고려대학교)
	왕규현(고려대학교)
	이강혁(서울대학교)
	이규민(고려대학교)
	이지민(고려대학교)
	정우현(서울대학교)
	최진(서울대학교)
	최지선(성신여자대학교)
보조원	황동진(고려대학교)
	홍윤서(성신여자대학교)
	조새하(성신여자대학교)
	이수정(성신여자대학교)
	서지수(성신여자대학교)
	박소영(성신여자대학교)

<국문 초록>

2021년 말뭉치 함의 분석 및 연구

본 사업은 한국어 화자의 직관을 수집하여 정보 산업에 활용 가능한 데이터로 변환함으로써, 국가 주도 구축 언어자원의 품질 향상을 도모함에 목적이 있다. 본 사업의 대상이 되는 언어자원은 확신성 담화이다. 예를 들어, 영희가 “철수는 밥을 먹었는지 모른다.”고 말한다면, 영희는 철수가 밥을 먹었다는 사실을 함의하는지 또는 화자(영희)가 이를 확신하는지 문제가 된다. 전통적인 언어 이론에서는 화자의 (비)확신성이 모절의 술어에 의해서 결정된다고 보았으나, 최근의 말뭉치와 실험 연구에서는 이분법적 구별이 어려운 척도 양상으로 화자의 확신성이 나타나는 것으로 보고되고 있다. 따라서 본 사업에서는 한국어 화자의 확신성을 연구하고자 국어 말뭉치에서 대상 담화를 추출하고 리커트 7점 척도로 평가하였다.

과업의 효율적인 수행을 위하여, 본 사업팀은 여러 공정을 병렬적으로 수행하였다. 관련 기초 연구 분석 및 예비 실험을 수행한 후, ‘대상 담화의 추출’, ‘확신성 언어실험’, ‘언어정보 부착’을 병렬적으로 진행하였다. 대상 담화의 추출에서는 내포절과 보문소를 특징으로 하는 한국어 문형을 수집하였다. 수집 목표는 장르별로 문어·신문 1,000건, 구어 500건이며, 한국어 화자의 직관을 파악하기에 적합한 문형을 수집하였다.

2020년 수행한 ‘말뭉치 함의 분석 및 연구’ 사업과는 달리, 2021년 사업에서는 대상 담화의 적합성을 판별할 (유사)보문소 목록을 198개로 확장하였으며, 함의 취소 운용소는 (필수 포함 요소가 아니라) 대상 담화에 포함되거나 포함되지 않을 수 있는 요소로 반영하였다. 또한 2020년 사업과의 차이점은 대상 담화를 추출하는 방법을 달리하여, 대상 담화의 장르 균형성과 한국어 보문소 목록을 참조하였다는 점이다. 이를 기준으로

165,225건의 한국어 확신성 담화 사례를 국립국어원 말뭉치에서 기계적으로 추출하고, 일부를 작업자가 주석처리하였다.

확신성 실험은 수집된 문형에 대한 한국어 화자의 직관을 계량화하는 과업이다. 실험을 수행하기 위하여 수집된 확신성 담화에서 1,806개 문항을 선별하고 실험 문항으로 설계하였다. 또한 관련 실험 설계 지침을 정비하여 향후 사업 수행을 위한 토대를 마련하였다. 확신성 실험의 절차는 다음과 같았다. 60명의 한국어 화자를 모집하여 세 그룹으로 나누고, 담화 사례별로 20명 내외의 한국어 화자가 7점 리커트 척도로 확신성을 평가하도록 하였다.

화자의 확신성에 기여하는 요소를 살펴보고자 1,806개 문항에 대하여 언어정보 부착을 수행하였다. 한국어 문형의 구조에 따라 내포절의 인칭, 시제소, 보문소, 확신성 유형 정보를 포함하여 다양한 언어정보를 온라인 워크벤치에서 수동으로 작업하였다. 공동연구원으로 구성된 상위작업자의 검수로 최종적으로 확보된 1,562개의 담화 사례에 대하여 확신성 실험 결과 및 언어정보 주석을 부착하였다. 또한, 말뭉치를 재가공하여 한국어 인공지능 평가에 활용할 수 있는 방법론을 제시하고, 그 결과를 공개하였다.

본 사업의 신뢰성을 담보하기 위하여 국어 및 컴퓨터공학 전문가 자문을 받아 국어정보 자원으로서의 성격을 명확히 하였다. 특히, 과업수행내용을 토대로 관련 지침을 정비하고, 적합한 한국어 문형 수집 절차에 관한 방법론을 정립함으로써 향후 고품질 국어 말뭉치 자원의 재생산과 확장에 기여하였다.

주요어: 함의 분석, 확신성 담화 수집, 확신성 언어실험, 언어정보 부착

차례

제 1장 사업 개요

1. 사업의 목적	2
2. 사업 수행 범위	3
3. 사업 수행 일정	4

제 2 장 대상 담화 추출 및 주석

1. 확신성 담화의 대상과 범위	7
2. 확신성 담화의 추출	10
3. 확신성 담화의 주석 작업	19
4. 확신성 담화의 구축 지침 보완 방안	25

제 3 장 확신성 언어 실험

1. 확신성 실험의 대상과 범위	35
2. 확신성 실험의 수행 지침 및 사례	36
3. 확신성 실험의 결과	47

제 4 장 확신성 언어 실험 결과의 국어학적 분석

1. 확신성 언어 실험 결과의 활용과 범위	59
2. 확신성 언어정보 부착	60
3. 확신성 언어 실험 결과의 인공지능 평가 활용 방안	69

제 5 장 결론 및 제언

1. 결론	84
2. 제언	86

붙임 1. 추론_확신성 분석 말뭉치 2021 구축 지침서

1. 개요	90
2. 작업 절차	96
3. 주석 지침	99
4. 실험 지침	128
<표 1> (유사) 보문소 목록	135
<표 2> 모문 술어 목록	136
참고 문헌	138
Abstract	140

제 1 장

사업 개요

1. 사업의 목적

본 사업은 화자의 발화 속에 담긴 내포 명제의 사실성과 국어의 구조가 상호작용하는 양상을 파악하고자 한다. 이를 위하여 한국어 화자의 확신성 직관을 리커트 척도를 사용한 언어 실험으로 수집하였다. 또한 나아가 본 사업은 실험 수행 결과물을 언어자원으로 구축하여 관련 산업의 요구를 충족시키고자 한다. 특히, 그동안 산업계에서는 언어자원의 활용성을 증대하기 위하여 인공지능이 자연어의 복잡한 구조를 제대로 이해하는지를 평가하는 구체적인 방법론 개발의 필요성을 꾸준히 제기해 왔다. 이에 따라 본 사업은 언어자원을 재가공하여 한국어 인공지능을 평가하는 새로운 방법론과 지침을 제안하였다.

국립국어원에서는 온라인으로 ‘모두의 말뭉치’를 공개하여 언어자원의 이용가능성을 확장하고 있다. 이러한 기초의 연장에서, 본 사업은 기존 문어·구어 말뭉치에서 함의 분석의 대상이 되는 ‘확신성 담화’를 추출하여, 정량화된 실험적 방법론으로 국어 이해의 양상을 계량화하는 것을 주요 사업 내용으로 하였다. 확신성 담화는 생성된 내포 명제의 사실성에 대한 화자의 직관을 반영하고 있다. 예를 들어, ‘철수가 영희를 역 앞에서 보았다고 말했다’는 문장은 ‘철수가 영희를 역 앞에서 보았다’는 내포 명제를 함의(entailment)한다.

한편 본 사업의 필요성은 한국어 발화가 내포하는 명제의 사실성을 이해하는 한국어 화자가 단순히 참·거짓이 아닌 복잡한 화용론적 맥락에서 국어 발화를 해석하기 때문일 수 있다는 학술적 연구와 관련이 있다. 예를 들어, ‘철수가 영희를 역 앞에서 보았는지 모르겠다고 말했다’는 문장이 동일한 내포 명제를 함의하는지 또는 내포 명제의 사실성을 취소(cancellation)하는지가 화자(청자)의 직관에 따라 매우 다양한 양상으로 나타날 수 있다. 이러한 미묘한 차이를 포착하기 위해서는 실험 방법론에 입각한 계량적인 함의 분석이 필요한 것이다.

또한, 4차 산업혁명의 시대에 국어 말뭉치는 인공지능 구축·평가 데이

터로도 활용될 필요가 있다. 문제는 인공지능이 한국어 화자의 직관을 반영하는 함의 관계를 얼마나 정확하게 예측할 수 있는가이다. 이를 해결하기 위해 본 사업에서는 정교한 실험적 환경에서 수집한 한국어 화자의 언어이해 양상을 재가공하여, 인공지능의 자연어이해(Natural Language Understanding) 수준을 평가하는 방법을 제안하였다. 다시 말해 본 사업의 의의는 기업들이 개발한 대규모 인공지능의 고난도 성능을 평가할 최적화된 방법론을 제시하였다는 점에서 찾을 수 있다.

요약하면, 본 사업은 산업적·학술적 요구에 따라 ‘2021년 말뭉치 함의 분석 및 연구’에서 다음과 같은 사항을 제안하였다. 국립국어원에서 구축한 말뭉치에서 함의 분석의 대상이 되는 담화를 추출하고, 관련 주석 지침을 정비하였다. 또한 추출된 담화를 정량적 실험에 적합한 형태로 변형하고, 한국어 화자의 확산성 판단을 수집하는 실험을 설계하고 진행하였다. 마지막으로 실험의 결과를 재가공하여 인공지능 평가에 적합한 형태로 변형하여 실제 한국어 인공지능 평가 결과를 제시하였다.

2. 사업 수행 범위

본 사업은 국립국어원 신문기사, 문어, (준)구어 말뭉치를 대상으로 확산성 담화를 수집하여 1,500건의 담화로 구성된 확산성 언어자원을 구축하였다. 본 사업은 각 공정 단계에서 예비적 공정과 본 공정을 이원화하여 공정의 완성도를 높였다. 본 사업은 장르별 균형성과 오류율을 고려하여, 1,805건의 담화 사례를 수집하였다. 수집한 사례의 범위는 문어 960건, 신문 303건, 구어 421건, 대화 121건이다(제 2 장). 또한 상위 검수자가 추가적인 선별 절차를 거쳐 최종적으로 1,562건의 확산성 담화 사례를 정리하였다. 한국어의 확산성과 밀접한 관련을 맺고 있는 한국어 함의취소운용소(‘-르지 모르다’ 등)와 보문소(‘ㄴ_믿음’ 등)의 결합이 존재하는 문장이 사업의 대상으로 포함되었다. 또한 장르별로 대상 담화의 성질

이 다르므로 해당 정보를 반영하여 대상 담화를 추출하였다.

향후 관련 국어 말뭉치 사업의 연속성을 담보하고 효율적인 개발을 위해, 담화 수집 대상 및 방법에 대한 상세한 기술과 제외 대상에 대한 지침을 마련하였다. 문어와 신문은 앞뒤 3문장 문맥을 함께 수집, 구어는 10문장 문맥을 수집하는 지침을 제안하였으며, 내포 명제를 기술하는 지침을 구체적인 사례와 함께 예시를 다수 제안하였다. 또한 시제, 복원, 인칭대명사, 양태를 기준으로 내포 명제의 언어적 구조를 파악할 수 있는 주석 작업 지침을 제안하였다. 구체적인 지침과 관련 사례는 별도의 파일로 정리하여 용이하게 참조할 수 있도록 하였다.

언어 실험은 60명의 한국어 화자를 모집하여 3개월에 걸쳐 1,806건의 담화를 대상으로 진행되었다(제 3 장). 실험언어학 분야에서 다수의 논문을 출판한 경험이 있는 박사급 공동연구원이 실험의 구성과 문항의 구조를 설계하였다. 60명의 한국어 화자는 대상 담화와 실험의 구성에 대한 교육을 받았으며, 신뢰할 수 없는 응답 패턴을 보이는 참여자는 실험 결과에서 제외되었다. 최종 실험결과에는 급내상관계수를 포함한 통계분석이 적용되었다.

언어정보 부착은 내포절의 시제와 인칭, 함의취소운용소를 대상으로 한다. 관련 사례와 함께 실제 주석 정보 부착의 예시를 제안하였다(제 4 장). 실험 결과와 부착된 언어정보를 종합하여 인공지능 평가를 진행하였다. 인공지능 평가는 한국어 언어 모형인 KrBERT, KoBERT 등을 대상으로 하였다. 함의 분석에서 나타난 한국어 화자의 확신성 판단 점수를 재가공하여, 인공지능의 추론 능력 수준을 살펴보았다.

3. 사업 수행 일정

본 사업은 2021년 5월 17일 착수하여 2021년 12월 17일까지 약 7개월간 수행되었다. 매달 온라인 회의와 함께 월간 보고서를 작성하였으며,

구체적인 진행은 다음의 <표 1>에 제시하였다. 변동사항에 대한 보고 및 공정 수행 결과에 대한 정보 교환은 발주기관과 수시로 진행하였다.

<표 1. 사업 공정 수행 경과>

과업 구분	5월	6월	7월	8월	9월	10월	11월
대상 담화 추출	■	■					
대상 담화 주석		■	■	■	■		
확신성 언어실험				■	■	■	■
언어정보 부착						■	■
인공지능 평가							■

제 2 장

대상 담화 추출 및 주석

1. 확신성 담화의 대상과 범위

확신성 담화는 어떤 대화 또는 발화가 함의하고 있는 내용의 사실성이 국어의 문장구조와 맥락에 의하여 결정되는 사례를 대상으로 하였다. 화자가 함의하고 있는 내용의 사실성은 진실이거나 거짓, 또는 모호할 수 있으며, 이러한 의도는 다양한 한국어 술어에 의해 표현된다. 이 때문에 화자의 의도와 발화의 술어가 어떠한 관계를 맺고 있는지가 중요하다. 또한 화용적 맥락과 밀접한 관계를 맺고 있는 까닭에 담화의 내포명제의 사실성에 대한 화자의 확신성은 달라질 수 있다. 따라서 본 사업은 언어 구조적 관계를 실험으로 계량화하고, 언어자원으로 재가공함으로써 국어 연구 및 인공지능 산업에 활용할 수 있는 담화를 수집하였다. 구체적인 예는 아래 <표 2>와 같다.

<표 2. 대상 담화 사례>

대상 담화와 실험 구조	추론 관계
<p>뉴욕의 겨울은 눈이 많이 오고 매우 춥다. 플로리다에서 온 존은 어제 눈이 온 것을 알지 못한다. 그는 눈을 거의 경험하지 못했다.</p> <p>[질문] 다음 문장에 대해 필자는 어느 정도로 확신한다고 생각하십니까?</p> <p>“어제 (뉴욕에) 눈이 많이 왔다”</p>	<p>함의</p>
<p>철수는 영희를 좋아하지만 아직 고백하지 않았다. 철수는 영희가 자신의 사랑을 받아줄 것으로 기대한다. 모두 철수가 언제 고백할지 궁금해한다.</p>	<p>중립</p>

<p>[질문] 다음 문장에 대해 필자는 어느 정도로 확신한다고 생각하십니까?</p> <p>“영희가 철수의 사랑을 받아줄 것이다”</p>	
<p>철수는 바나나를 구황작물이라고 생각해. 바나나가 땅에서 자라는 줄 알지? 절대 사실을 알려주지 않을거야.</p> <p>[질문] 다음 문장에 대해 필자는 어느 정도로 확신한다고 생각하십니까?</p> <p>“바나나가 땅에서 자란다”</p>	<p>모순</p>

또한 본 사업은 영어에 대한 선행 연구인 de Marneffe et al.(2019)을 참조하였다. 관련 연구에서 (비)사실성에 대한 명제의 사실성을 함의/중립/모순의 삼분적 구조로 구분하였다. 이러한 삼분적 구조는 내포 명제의 사실성의 이론적 양상에 관한 것으로, 실제 한국어 화자가 평가하는 내포 명제(가설)의 사실성은 **확신성 정도(degree of commitment)**에 따라 차이를 보인다. 언어학에서는 이를 명확한 범주적 경계가 없는 경사적 속성 (gradient nature)의 문제로 다루고 있다.

이러한 양상은 평가 척도를 어떻게 설정하느냐에 따라 달리 평가될 수 있다. 예를 들어, 5점 또는 7점 리커트 척도에 따라 확신성의 크기가 달라질 수 있다. 이에 관해서는 제 3 장에서 상술한다. 따라서 잠정적으로 본 사업에서 함의 관계는 내포절에 대한 화자의 확신성이 참으로 확신되는 경우라고 볼 수 있다. 중립 관계는 내포절에 대한 화자의 확신성이 참인지 거짓인지 모르는 불확신성으로 나타나는 경우이다. 모순은 내포절에 대한 화자의 확신성이 거짓으로 확신되는 경우이다.

수집한 사례의 범위는 문어 960건, 신문 303건, 구어 421건, 대화 121

건으로 도합 1,806건이다. 각 사례별로 확산성 정도를 계량화하기 위한 언어 실험을 수행하였고, 관련 언어정보를 부착하였다. 본 사업은 해당 사례를 추출하기 위하여 국립국어원에서 구축한 말뭉치로 범위를 한정하였다. 이는 기존 국어 정보화 사업의 연속성을 담보하고 확장성을 꾀하기 위함이다. 적합한 문형과 사례를 추출하기 위하여 기계적 자동화 추출 기법과 수동적 분류 기법을 상호보완적으로 사용하였다. 기계적 추출 단계에서는 한글 자모 분리와 형태소 분석을 적용하여 기존 국어원 구문분석 말뭉치와 세종 형태분석 구어 말뭉치를 활용하였다.

자동 추출 알고리즘의 국어적 기준을 설정하기 위하여 198개의 (유사)보문소 목록을 작성하여 활용하였다(그림 1 참조). 넓은 범위의 보문소 목록을 작성한 목적은 한국어 문형의 분포적 균형성을 확보하기 위한 것이다. 예를 들어, 장르에 따라 보문소 ‘_다는 생각’이 ‘_라는 소식’보다 빈번하게 출현하거나 그렇지 않을 수 있다. 이때, 특정 문형에 편향되는 문제를 해소하기 위하여 장르를 다양화하거나 담화 추출의 대상을 확장하는 방법을 선택할 수 있었다. 이러한 접근은 컴퓨터 도구를 활용한 자동 방법에 기초하였다.

이와 달리, 수동 분류는 박사급 공동연구원의 자문을 받아 작업 규칙을 정하고 검수를 받았다. 수동 분류의 기준은 내포 명제의 명확성, 언어구조적 일관성 등을 고려하였다. 해당 분류 기준에 적합하지 않은 문형은 공정의 다음 단계인 확산성 언어 실험과 언어정보 부착 및 인공지능 시범 평가에서 제외되었다.

또한 본 사업은 지난 ‘20년 말뭉치 함의 분석 및 연구’ 사업의 지침을 보완하는 것을 과업의 범위로 하였다. 이러한 지침 보완은 대상 담화 자동 수집 및 수동 분류의 절차를 보강하여 공정의 효율성과 정확성을 높이는 것을 주요 목적으로 한다. 해당 지침의 내용은 기존 지침의 미비점을 추가로 설정하고 관련 국어 예시와 해결 방안을 함께 기술하는 형태로 작성되었다.

2. 확신성 담화의 추출

2.1. 예비 추출

확신성 담화의 기계적 추출은 예비 추출 단계와 본 추출 단계로 나뉜다. 예비 추출은 다음 단계의 공정에 적합한 한국어 문형을 추출하기 위한 자동 추출 및 수동 분류의 기준을 확정하는 것을 목적으로 한다. 이러한 샘플링 단계는 별도의 기준을 두지 않는 것과 별도의 기준을 두어 (1) 작업자 층위 (2) 국어 보문소 유형별로 공정 적합성에 어떠한 영향을 미치는지 판별하는 데 효과적이다.

본 추출은 국어의 구어, 신문, 문어 장르와 대상 담화의 함의적 관계에 영향을 미치는 내포 명제를 기준으로 반복적으로 추출하였다. 이러한 반복적 추출은 이전 단계에서 불균형하였던 데이터 분포를 해소하고, 보다 다양한 한국어 문형을 추출할 수 있다는 장점이 있다. 특히, 데이터 구축의 관점에서 볼 때, 무작위적 추출과 일회적 추출은 데이터 편향성을 증폭시키는 경향이 있다는 점에서 본 사업의 공정은 중요하다. 구체적인 예는 아래 <표 3>과 같다.

<표 3. 확신성 담화의 추출 단계>

예비추출 샘플링	본 추출	
	장르 추출	내포명제 추출
precommit0 (10,041) - 별도의 기준을 두지 않음	commit0 (478) - 보문소 “줄” 추출	commit4 (803) - 실험 240 문항 추출
precommit1 (10,041) - 작업자 층위별 적합성을 기준으로 함.	commit1 (51,882) - 문어 장르 1차 추출	commit5 (1,114) - 실험 문항 작업 및 지침
precommit2 (10,000개) - 보문소 유형별 추출	commit2 (10,000) - 문어 장르 2차 추출	commit6 (811) - “모순” 문항 1차 추출
	commit3 (10,000) - 신문 장르 추출	commit7 (945) - “모순” 문항 2차 추출

1차 예비 추출 단계(precommit0-1)에서는 무작위로 10,041건의 담화 구조(문어 1,837/신문 1,886/구어 6,318)를 추출하였다. precommit0은 완전 무작위로 추출한 것인 반면, precommit1은 대학원 석박사과정에 재학 중인 보조연구원과 박사급 공동연구원의 작업자 층위를 분리하여 추출하였다. 이는 작업자 층위 간에 적합한 대상 담화를 판별하는 경향성이 존재하는지 파악하기 위함이다. 이를 토대로 일일 워크숍 또는 작업자 교육을 실시하였다.

1차 사전 작업에서는 무작위로 담화 구조를 선별하였기에 문어와 신문의 경우 내포문의 비율이 10% 미만에 그쳤다. 내포문이 포함되는 문형에서만 술어와 결합하여 화자의 확신성이 달라지는 형태를 발견할 수 있기 때문에, 무작위 선별은 대상 담화 수집에 매우 부적합한 방법임을 알 수 있다. 또한, 구어는 문장의 형태가 불완전한 구어적 생략이 빈번하였다. 2차 사전 작업에서는 이를 보완하기 위하여 보문소를 기준으로 확장된 담화 구조를 추출하였다.

2차 예비 추출 단계(precommit2)에서는 198개의 보문소를 기준으로 확장된 담화 구조(문어 2,500/신문 2,500/구어 5,000)를 추출하였다. 보문소는 내포절이 전체 문장(모문)의 보문이 되도록 만드는 요소이다. 따라서 보문소를 기준으로 추출하면 내포절이 포함된 문형이 정확하게 추출될 가능성이 높아진다.

<그림 1. 보문소 목록>

ㄴ_가능성	ㄴ_믿음	ㄴ_지_	다는_뜻	ㄹ_답	라"고_
ㄴ_가정	ㄴ_반성	ㄴ_질문	다는_말	ㄹ_뜻	라_
ㄴ_개념	ㄴ_분위기	ㄴ_처지	다는_비판	ㄹ_마음	라고_
ㄴ_거_	ㄴ_비판	ㄴ_태도	다는_사실	ㄹ_말	라고는_
ㄴ_거라고_	ㄴ_사건	ㄴ_해석	다는_생각	ㄹ_맘	라고도_
ㄴ_걸_	ㄴ_사례	ㄴ_형상	다는_소리	ㄹ_문제	라는_개념
ㄴ_것_	ㄴ_사실	ㄴ_형태	다는_소문	ㄹ_비판	라는_거
ㄴ_것도_	ㄴ_상태	ㄴ_희망	다는_소식	ㄹ_사건	라는_걸_
ㄴ_것만_	ㄴ_상황	ㄴ가_	다는_얘기	ㄹ_사실	라는_것
ㄴ_것으로	ㄴ_생각	ㄴ가는_	다는_의견	ㄹ_상황	라는_게_
ㄴ_것은_	ㄴ_설명	ㄴ가를_	다는_의미	ㄹ_생각	라는_뜻
ㄴ_것을_	ㄴ_소리	ㄴ지_	다는_의지	ㄹ_설명	라는_말
ㄴ_것이_	ㄴ_소문	ㄴ지가_	다는_의혹	ㄹ_소리	라는_비판
ㄴ_것이라고_	ㄴ_소식	ㄴ지는_	다는_이야기	ㄹ_얘기	라는_사실
ㄴ_게_	ㄴ_신념	ㄴ지도_	다는_일	ㄹ_예상	라는_생각
ㄴ_견해	ㄴ_양상	ㄴ지를_	다는_입장	ㄹ_예정	라는_얘기
ㄴ_결과	ㄴ_얘기	냐_	다는_주장	ㄹ_예측	라는_의미
ㄴ_경우	ㄴ_어려움	냐를_	다를_	ㄹ_의도	라를_
ㄴ_경향	ㄴ_예상	냐"고	ㄹ_가능성	ㄹ_의미	ㄹ도_
ㄴ_경험	ㄴ_예측	냐_	ㄹ_가정	ㄹ_의심	ㄹ만_
ㄴ_교훈	ㄴ_욕구	냐가_	ㄹ_개념	ㄹ_의지	ㄹ으로_
ㄴ_권리	ㄴ_욕망	냐고_	ㄹ_거_	ㄹ_이야기	ㄹ은_
ㄴ_기능	ㄴ_의견	냐는_	ㄹ_거라고_	ㄹ_일	ㄹ을_
ㄴ_기대	ㄴ_의도	느냐_	ㄹ_걸_	ㄹ_입장	ㄹ이_
ㄴ_기분	ㄴ_의문	니_	ㄹ_것_	ㄹ_자세	마_
ㄴ_기쁨	ㄴ_의미	니가_	ㄹ_것도_	ㄹ_전제	으면_
ㄴ_기억	ㄴ_의사	니는_	ㄹ_것만_	ㄹ_주장	자_
ㄴ_기회	ㄴ_의심	니도_	ㄹ_것으로	ㄹ_주제	자고_
ㄴ_논리	ㄴ_의지	니를_	ㄹ_것은_	ㄹ_줄_	
ㄴ_뉴스	ㄴ_의혹	다"고_	ㄹ_것을_	ㄹ_줄은_	
ㄴ_느낌	ㄴ_이야기	다_	ㄹ_것이_	ㄹ_줄을_	
ㄴ_답	ㄴ_일	다"고_	ㄹ_것이라고_	ㄹ_해석	
ㄴ_뜻	ㄴ_입장	다고_	ㄹ_게_	ㄹ_확률	
ㄴ_마음	ㄴ_자세	다고는_	ㄹ_경우	ㄹ_희망	

또한 내포절이 함의하는 사실에 대한 화자의 확신성이 문제가 되기 때문에, 이와 밀접한 관련을 맺는 국어적 요소들이 존재한다. 이러한 기능을 하는 요소를 함의 취소 운용소라고 한다. 함의 취소 운용소는 화자의 확신성 판단에 영향을 미치는 핵심적 요소로, 내포절이 함의하는 명제에

대한 진리치를 바꿀 수 있다. 이를 유형화하면 크게 부정/의문/조건/양태의 함의 취소 운용소가 있다.

<그림 2. 함의 취소 운용소의 예>

부정	_____	못, 안/아니, 없, -지 말/아니하/않-, -지(는) 못하-
의문	_____	-까, -는가, -는데-, -나, -냐, -잖아, -지 등(의문형 종결어미)
조건	_____	-(으)면(야), -(니)다/라면, 어/아도
양태	인식 양태	-(-으)ㄴ가 싶-, -(-으)ㄴ/-(-으)ㄹ 것 같-, -(-으)ㄹ 것이-, -(-으)ㄴ/-(-으)ㄹ 듯하 듯하-, -(-으)ㄹ 수 있/없-, -(-으)ㄹ 텐데, -(-으)ㄹ지 모르-, -겠(추측)-, -(-으)ㄴ 모양이-, -나 보-, -어/아 보이-
	비인식 양태	의무: -(-으)ㄹ 필요가 있-, -(-으)면 되-, -어야 되-, -어야 하- 능력(가능/불가능): -(-으)ㄹ 만하-, -(-으)ㄹ 수 있/없-, -(-으)ㄹ 줄 알/모르- 의지: -겠-, -고 싶-, -고자 하-, -(-으)려 하-

그러나 공정 수행에서 함의 취소 운용소를 쉽게 발견하기 어려웠다는 점을 파악했다. 이러한 점은 자동 공정으로 확신성 담화를 모두 추출하기는 어렵다는 점을 시사하며, 후속 공정에서 수동 분류를 거치는 작업이 필요함을 보여준다.

예비추출의 성과는 다음과 같다. (1) 작업자 층위간 경향성 파악 (2) 유사 보문소의 목록 작성으로 내포문을 포함한 문형 추출의 정확성 제고 (3) 화자의 확신성 판단에 핵심적인 함의 취소 운용소의 파악이다.

2.2 본 추출

본 추출의 주요 목적은 후속 공정의 대상이 되는 대상 담화를 자동적으로 수집하여 수동적으로 선별하는 것이다. 앞선 예비 추출에서 고려한 198개 유사 보문소가 하나 이상 반드시 포함되는지 자동수집 및 수동선별 단계에서 검수하였다. 보문소는 화자가 내포절이 생성하는 가설의 사실성에 대한 태도를 나타내는 언어적 구조와 관련이 있으므로 반드시 포함되어야 한다. 이와 달리 함의 취소 운용소는 화자의 확신성 정도와 관련 있는 수의적인 요소로, 포함되거나 포함되지 않을 수 있다.

예를 들어, ‘철수가 밥을 먹었는지 모르겠다’는 인식 양태 함의 취소 운용소 ‘-는지 모르다’가 내포 명제에 대하여 화자가 확신하고 있지 못함을 보여준다. 반대로, 함의 취소 운용소가 없더라도 ‘철수가 밥을 먹었다고 생각한다’는 문장에서 화자의 확신성 태도가 드러난다. 이때, 보문소 ‘_다고 생각’은 생략될 수 없다.

이러한 양상에 기초하여, 자동 수집 단계에서 유사 보문소 목록과 함의 취소 운용소를 모두 포함하는 자동 검색 알고리즘을 작성하여 대상 담화를 추출하였지만, 수동 선별 단계에서는 함의 취소 운용소가 아닌 보문소도 다수 포함되도록 허용하였다. 단, 보문소는 반드시 포함되도록 선별하였다.

본 추출 단계에서는 후속 공정의 목적에 따라 장르별 추출(commit0~3)과 내포명제 추출(commit4~7)로 구성된다. 확신성 담화를 추출할 때, 무작위로 추출하지 않고 국어학적 분석에 의거하여 보문소 목록과 함의 취소 운용소 목록을 작성하여 타당하고 일관된 기준에 따라 추출하였다는 점이 본 사업 공정의 특징이다. 이러한 방법론은 향후 관련 말뭉치를 확장하고 후속 사업을 기획함에 있어서 효율성과 정확성을 높일 수 있다는 점에서 효과적이다.

2.2.1. 장르별 추출 (commit0-3)

장르별 추출은 특정 문형 패턴의 빈번한 출현으로 인한 데이터 불균형 해소의 관점에서 4단계로 접근하였다. 이 외에 유사 보문소 또는 함의 취소 운용소와 관련 있는 특정 문형 패턴이 별도로 관리할 필요가 있다고 판단되면, commit(number) 파일로 관리하였다. 예를 들어, 보문소 ‘-줄’은 화자의 확신성과 관련하여 다른 보문소 또는 함의 취소 운용소와 구별되므로 별도로 관리하였다. 이는 장르별로 특정 보문소가 매우 빈번하게 출현하는 반면, 나머지 보문소는 상대적으로 빈도가 낮다는 점에 착안하였다.

‘commit0 (478개 사례)’는 보문소 ‘-줄’이 매우 특이한 문형 패턴을 형성하므로 샘플 데이터에서 별도로 작업할 필요가 있었다.

‘commit1 (51,882개 사례)’은 “문어 장르”에서 추출하였다. 문어 장르는 책, 잡지, 보고서 등으로 구성되었으나 세부 장르별 불균형이 매우 심하였다. 따라서 5만 건 이상의 사례를 추출하여 빈도가 적어 주변화된 세부 장르를 고르게 추출하였다.

‘commit2 (10,000개 사례)’는 문어 장르 불균형성을 해소하기 위하여 분포를 다양하게 하였다. 또한 다른 장르에 대하여 데이터 균형을 확보하기 위하여 표준적 크기인 10,000개로 줄였다.

‘commit3 (10,000개 사례)’는 신문 장르에서 추출하였다. 그러나 신문 장르의 특징인 빈번한 직접 인용으로 인하여 보문소 “다고”의 비율이 매우 높았다. 따라서 실험과 인공지능 평가에서 특정 문형 패턴에 과적합될 수 있다는 문제가 존재하였다.

2.2.2. 내포명제 추출 (commit4-7)

commit4-7의 4단계는 함의/중립/모순의 라벨 부착과 균형 분포를 고려하여 내포 명제를 추출하는 것을 목적으로 하였다. 이는 언어 실험과 인공지능 평가에서 특정 응답 패턴이 반복되어 실험 참여자 또는 인공지능 휴리스틱(heuristic)에 의존하는 부정적인 패턴을 발생시킬 수 있기 때문이었다.

‘commit4 (803개 사례)’는 추출-주석-실험-평가의 전체 한국어 함의 분석 공정을 수행하기 위하여 구성되었다. 확산성 실험의 수행을 위한 240개의 적합한 문형을 잠정적으로 추출하였다.

‘commit5 (1,114개 사례)’는 commit4와 동일한 목적에서 수행되었으며 commit5_101-105로 회차별 실험에 적합하게 다시 세분화하였다. 또한 향후 확산성 실험의 수행과 인공지능 자연어 이해 평가를 위한 문형 추출을 고려하여 별도의 사례별 지침 작성을 위한 기초 자료로 활용하였다.

‘commit6 (811개 사례)’는 데이터 불균형을 보완하기 위해 추출된 파일이다. 확산성 판단의 패턴 중 내포 명제가 모순 또는 거짓인 담화 사례가 8% 내외로 매우 낮은 빈도로 발견되었다. 이러한 불균형을 해소하기 위해 내포 명제가 모순 또는 거짓인 대상 담화를 추가로 수집하였다.

‘commit7 (945개 사례)’는 내포 명제가 모순 또는 거짓인 담화 사례의 비율을 20% 이상으로 보충하기 위해 추가로 수집되었다.

위의 대상담화 수집 절차를 거쳐 향후 실험 공정 및 언어정보 부착, 그리고 인공지능 평가의 기초 자료로 사용할 대상 담화를 국립국어원 말뭉치에서 추출하였다.

<표 4. 확신성 담화의 추출 단계>

예비추출	본 추출	
	장르 추출	내포명제 추출
precommit0 (10,041) - 별도의 기준을 두지 않음	commit0 (478) - 보문소 “줄” 추출	commit4 (803) - 실험 240 문항 추출
precommit1 (10,041) - 작업자 층위별 적합성을 기준으로 함.	commit1 (51,882) - 문어 장르 1차 추출	commit5 (1,114) - 실험 문항 작업 및 지침
precommit2 (10,000개) - 보문소 유형별 추출	commit2 (10,000) - 문어 장르 2차 추출	commit6 (811) - “모순” 문항 1차 추출
	commit3 (10,000) - 신문 장르 추출	commit7 (945) - “모순” 문항 2차 추출
	commit8 (55,948) - 구어 장르 추출	commit9 (500) - 구어 장르 1차 추출
	commit14 (502) - 문어 장르 추가 추출	commit10 (157) - 구어 장르 “모순” 추출
	commit15 (503) - 문어 장르 추가 추출	commit11 (500) - 구어 장르 3차 추출
		commit12 (500) - 구어 장르 4차 추출
		commit13 (500) - 구어 장르 5차 추출

2.2.3. 구어장르 추출 (commit8-13)

구어 장르의 경우, 문어 및 신문 장르와 다른 특성으로 인해 별도의 방

식을 적용하여 추출을 진행하였다. 구어는 문어에 비해 짧은 길이로 구성되고, 발화자가 교체되는 구어의 특성이 존재한다. 이러한 상황에서 보문소가 등장하는 담화의 문맥을 정확히 파악하기 위해, 대상 담화 이전 10개의 문장과 이후 10개의 문장을 수집하였다.

‘commit8 (55,948개 사례)’는 국립국어원 구어 말뭉치를 대상으로 보문소가 등장하는 문장을 추출하였다. 구어 장르에서 확신성 담화를 추출하는 문제에 대해 다각적으로 탐지하기 위해 5만 건 이상의 사례를 추출하여 확신성 담화의 유형을 살폈다.

‘commit9 (500개 사례)’는 구어 장르에서 이전 문맥의 파악 및 생략된 문맥의 복원 문제에 중점을 두고 수집하였다.

‘commit10 (157개 사례)’는 구어 장르의 모순 유형을 중심으로 수집하여 그 특성을 살폈다.

‘commit11 (500개 사례)’는 구어 장르의 보문소 유형 및 분포에 대해 파악하기 위해 수집하였다.

‘commit12 (500개 사례)’는 수집한 문장에서 나타나는 정치적, 사회적 문제에 대한 스크리닝을 중심으로 수집하였다.

‘commit13 (500개 사례)’는 앞선 수집에서 등장한 문제를 다각적으로 살펴 실험이 가능한 문형 수집을 위한 목적으로 수집하였다.

2.2.4. 문어 추가추출 (commit14-15)

실험 및 정제 과정에서 문어 장르에 대해 정제한 결과, 특정 보문소별로 충분한 수가 확보되지 않아, 빈도가 적게 등장한 보문소를 대상으로 기반으로 문어 추가추출을 진행하였다. ‘commit14 (502개 사례)’와 ‘commit15 (503개 사례)’ 모두 문어 장르의 보문소 편향을 완화하기 위해 동일한 방식으로 수집하였다.

3. 확산성 담화의 주석 작업

3.1. 주석 작업자 교육

작업자 교육은 확산성 언어 정보 부착에 대한 기본적 이론을 익히고 작업 지침을 숙지하는 것뿐만 아니라, 작업자가 워크벤치를 이용하여 직접 정보 부착 작업을 해 볼 수 있도록 국어학 전문가와 함께 실습하였다. 교육은 총 두 번의 워크숍을 통해 진행되었다. 작업자가 본격적인 작업에 들어가기에 앞서 국어학 전문가와 함께 실시한 Intensive 1-day Workshop에서는 확산성 언어 정보에 해당되는 시제, 인칭, 함의최소운용소에 대한 기본적 이론을 중심으로 교육이 이루어졌다. 전문가와 함께 실제 워크벤치에서 다양한 예문을 이용하여 정보를 부착하며 작업에 대한 이해도를 높였다.

또한, 함의 분석 사업 준비 단계 또는 화자의 확산성 담화를 추출하는 예비 샘플링 절차에 참여하였던 경험을 보완하였다. 이는 사업단 내부 워크숍을 수 회 개최하는 방식으로 이루어졌다. 워크숍의 목적은 작업자의 직관 차이로 인해 발생할 수 있는 오류를 최소한으로 하기 위함이다. 이를 위해 의견 공유 및 논의 시간을 갖고 전문가의 판단하에 정보부착 지침을 마련하였다. 예시로, 내포절과 모절의 시제와 시제소 정보를 부착할 때, 시제와 시제소 내포명제를 기준으로 정보를 부착할 것인지, 혹은 대상 문장 그대로의 형태를 기준으로 할 것인지에 대하여 작업자들의 기준이 상이하였기 때문에 상위 작업자(박사급 공동연구원)의 지침을 숙지하고 이행하였다. 이러한 워크숍 참여를 토대로 박사급 공동연구원과의 충분한 논의 후 시제는 내포명제를 기준으로, 시제소는 형태를 기준으로 정보를 부착한다는 기준을 세울 수 있었다. 이를 바탕으로 작업 결과물 중간 점검 및 품질 관리를 위한 'Interim Check Workshop'을 실시했다.

박사급 공동연구원이 지금까지의 작업 결과물을 살펴보고 작업자에게 개선해야 할 부분을 설명하였다. 이후에는 각 작업자가 언어 정보 부착 작업을 진행하며 메모한 의문점 및 특이사항을 바탕으로 전문가에게 질문하고 전문가는 이에 대해 답변을 하며 작업자들이 이후 언어정보 부착 작업을 진행할 때 숙지할 지침을 다시 한번 강화하였다.

3.2. 온라인 주석 작업 환경

주석 작업은 온라인 환경에서 제공하는 워크벤치에서 작업 공정이 진행되었다. 아래의 그림은 온라인 워크벤치의 구성을 보여준다.

<그림 3. 온라인 워크벤치 (담화 사례 주석)>

-중3 여학생, 「학교 선생님의 말 중 무시당하거나 기분 나쁜 말」 중에서 우리집도 철거가 될 것 같다. 그럼 우리는 어디에서 살까 하고 생각하면 무섭다.

▲ Show/Hide the Previous Context

[WCRW1900001362.1] 802/945 (문어)

우리 집이 철거가 안 되면 좋겠다.

우리 집이 철거가 안 되면/EC./SY; 좋겠다.

▼ Show/Hide the Next Context

동네가 없어질 것 같다. 사람이 없는 동네는 싫다. 철거가 안 되면 정말 좋겠다.

내포 구문 해당 여부:
 해당함

보문소:
 면

보문술어:
 좋다

내포명제:
 우리 집이 철거가 안 될 것이다

관계:
 함의 중립 모순 해당없음

기타:
 의견 불일치
 확인 요망

내포 구문 해당 여부에 대하여 ‘해당 없음’ 버튼을 추가하여, 추출된 담화 사례가 확신성 판단에 적합한지 먼저 판단할 수 있도록 하였다. ‘해당함’의 버튼을 누르면 보문 술어, 내포 명제, 확신성 관계 등 추가적인 주석 정보를 선택할 수 있도록 하였다.

마지막으로 가장 마지막 항목에서 ‘다음 문장’을 선택하거나, 가장 첫 항목에서 ‘이전 문장’을 선택하는 경우, 그리고 등록되지 않은 index로 이동하는 경우 빈 페이지로 이동하며, 그 이후 오류가 발생하게 되는데, 이러한 경우에 예외 처리를 하였다.

3.3. 주석 결과물

아래의 도표는 온라인 워크벤치에서 작업된 결과물을 도표로 나타낸 것이다. 온라인 워크벤치에서 내포 구문 해당 여부가 ‘해당함’으로 체크된 사례만 추가적인 담화 사례 주석을 진행하였다. 좌측의 열은 대상 담화의 메타 정보를 나타내고, 우측의 열은 실제 작업 결과물의 내용이다. 모든 작업 결과는 .csv 파일 형태로 저장되고 추후 .json 파일 등으로 변환될 수 있다.

<표 4. 담화 사례 주석의 항목과 내용>

idx	NLRW2000000013.334
genre	문어
prev	아파트에 사는 광주와 전남 서민들이 다른 지역에 비해 높은 월세 부담에 시달리고 있는 것으로 나타났다. 광주·전남지역 아파트 전월세 전환율이 전국에서 가장 높은 수준을 기록했기 때문이다. 특히 전남지역의 아파트 전월세 전환율은 전국 최고를 기록했다.
current	한국감정원이 2018년 11월 신고 기준 실거래 정보를 활용해 전월세 전환율을 산정한 결과, 전국 주택종합 전월세 전환율은 6.1%로 전달에 비해 소폭 하락했다고 11일 밝혔다.
next	유형별로는 아파트 4.7%, 연립다세대주택 5.7%, 단독주택 7.5% 순이었고 지역별로는 수도권 5.7%, 지방 7.2%로 조사됐다. 전월세

	전환율은 전세금을 월세로 전환할 때 적용되는 비율로, 이 비율이 높으면 전세에 비해 상대적으로 월세 부담이 높다는 의미이며 낮으면 반대이다. 다만, 전월세 전환율 상승(하락 또는 보합)이 절대적인 월세부담 증가(감소 또는 보합)를 의미하는 것은 아니다.
pred	밝히다
comp	다고_
context+target	특히 전남지역의 아파트 전월세 전환율은 전국 최고를 기록했다. 한국감정원이 2018년 11월 신고 기준 실거래 정보를 활용해 전월세 전환율을 산정한 결과, 전국 주택종합 전월세 전환율은 6.1%로 전달에 비해 소폭 하락했다고 11일 밝혔다. 유형별로는 아파트 4.7%, 연립다세대주택 5.7%, 단독주택 7.5% 순이었고 지역별로는 수도권 5.7%, 지방 7.2%로 조사됐다.
prop	전국 주택종합 전월세 전환율은 2018년 11월에 6.1%로 전달에 비해 소폭 하락했다
relation	함의

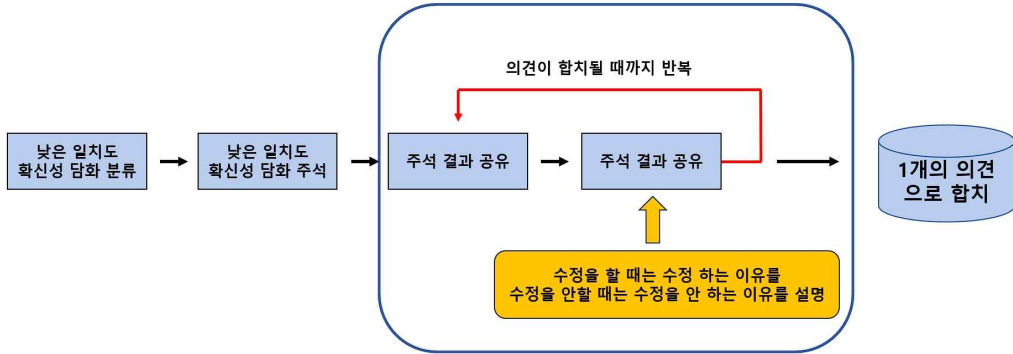
담화 구조에 대한 정보는 (1) 고유번호(idx) (2) 장르 (genre) (3) 선행 문맥 (prev) (4) 대상 문장(current) (5) 후행 문맥(next) (5) 모절 술어 (pred) (6) 보문소 (comp) (7) 선후행 맥락을 연결한 대상 문장 (context+target) (8) 내포 명제(prop) (9) 확신성 관계(relation)를 포함하고 있다.

3.4. 주석 검증 절차

담화 사례 주석 작업의 정확성과 신뢰성을 확보하기 위하여 1차 주석 작업 이후에 작업자간 교차 검증 절차를 수행하였다. 교차 검증은 이전 작업자의 작업물을 전수로 재검수하는 방식으로 진행되었다. 검증 절차는 크게 세 단계로 나뉜다. 먼저 A작업자가 1차 작업을 수행한다. 1차 작업 중에 발생하는 의문 및 추가적인 검증이 필요할 것으로 판단되는 사례에 '확인 요망' 또는 '민감 사안'과 같은 별도의 표시를 한 후 그 이유를 작성하였다.

<그림 4. 주식 검증 절차 >

델파이 기법



1차 작업을 마치면, B작업자가 동일한 자료를 대상으로 교차 검증을 수행하였다. 교차 검증 시, B작업자는 A작업자와 의견이 불일치하거나 여전히 추가적인 검증이 더 필요할 것으로 판단되는 사례에 1차 작업과 동일하게 별도의 표시를 한 후 그 이유를 작성하였다. 이러한 교차검증 과정을 체계화하여 매 워크벤치마다 이전 작업의 교차검증 작업자와는 다른 작업자를 배정하였다.

이후 두 작업자가 검증한 자료를 상위 층위의 검수자(박사급 공동연구원)가 마지막으로 검수하는 절차를 거치며 이전의 작업자 간의 의견에 차이가 있는 부분 또는 추가 검증이 필요한 부분에 대해 따로 판단하여 주식 작업의 결과를 선택하였다. 즉, 이와 같은 세 단계를 과정을 통해 복수의 작업자가 동일한 자료에 대한 분석을 수행하여 그 결과를 통합하였다.

<표 5. 담화 사례 주석 검증 예시>

대상 문장	1차 작업자 메모	검증 내용
<p>특히 이웃까지 다 들릴 큰 소리로 그들 형제를 방 안에 몰아 놓고 뒤따라 들어올 때는 금세 매가 그의 등허리에 떨어지르 /VV+ETM;_/SY; 줄 /NNG;_/SY;알았다.</p>	<p>매를 맞을 줄 알았는데 뒷 문맥에서 매를 맞지 않았음을 확인할 수 있어서 관계를 모순으로 하였습니다만 다른 분의 의견도 필요할 듯 하여 메모 남깁니다.</p>	<p>* 검수 의견: 저 문장만 봐서는 중립이 맞는 것 같지만 뒷 문맥을 참고한다면 모순으로 보는 것이 맞는 것 같습니다.</p>
<p>합산규제 연장 여부에 따라 인수합병(&) 판도가 달라질 것으로 예상되는 가운데 합산규제가 자칫 유튜브나 넷플릭스 등과의 경쟁을 제한하는 것 아니냐/EC:는 /IX;_/SY;지적이 나온다.</p>	<p>내포명제와 관계에 확신이 없어 메모 남깁니다.</p>	<p>* 검수 의견: ~한 것이 아니냐 의 경우는 아니다 로 내포명제를 설정할 수 없다고 생각함. 제한하다 로 끝나야지 더 자연스러운 것 같습니다.</p>

4. 확산성 담화의 구축 지침 보완 방안

4.1. 지침 신규대조표

본 사업은 ‘2020년 함의 분석 및 연구’와 유사한 성격을 지니고 있으나, 기존의 지침 내용을 변경하고 새롭게 규정하고 있다. <표 6.>에서는 주요 변경점 및 신규 내용이 요약되어 있다.¹⁾

<표 6. 20년과 21년 지침 신규대조표>

	20년 지침	21년 지침
분석 대상 구조와 탐색	(1) 시범 탐색 단계에서 ‘국립국어원 신문 말뭉치’ 6만 문장을 사용한다. (2) 대상 문장은 I) 내포절, ii) 보문소, iii) 모문 술어, iv) 함의취소운용소를 포함한다.	(1) (변동) 시범 탐색과 본 과업에서 동일하게 장르의 균형을 고려하여 문어, 신문, 구어 장르에서 추출한다. (2) (변동) 모절에 함의취소운용소는 포함되지 않을 수 있다.
분석 대상의 수	본 연구 과제의 목표는 내포절의 (비)사실성 함의에 있어서 화자의 확산성에 영향을 주는 요인을 분석하고 <u>1,800개</u> 의 문장으로 구성된 “한국어 확산성 말뭉치”를 구축하는데 있다.	(변동) 본 연구 과제의 목표는 기존 사업을 보완하고, <u>1,500개</u> 이상의 담화에 화자의 확산성 실험 결과를 포함여 말뭉치를 새롭게 구축하는데 있다.
	(1) 보문소: 본 연구 과제에서는 문헌을 기반으로 (비)사실성 구조를 구성하는 보문소의 종류와 표기 형태를 다음 <u>6가지</u> 로 한정	(1) (변동) 보문소: 1차 연구보다 <u>다양한 유형(총 198개)의 보문소</u> 를 선정하여 다양한 한국어의 언어 사용을 반영한다. 보문소와

1) 보다 자세한 지침과 사례는 붙임 1.에서 부연하였다.

<p>대상 담화의 보문소와 모문 술어 구조</p>	<p>하여 사용한다.</p> <p>-음/ㅁ, -기, 것을, 것으로, -다고/라고, 줄</p> <p>(2) 모문 술어: 사실성 여부를 기준으로 다음 3가지로 나누어 쓴다.</p> <p>① 사실성 술어</p> <p>기억하다, 깨닫다, 발견하다, 밝히다, 이해하다, 잇다, 지적하다, 확인하다, 후회하다</p> <p>② 비(非)사실성 술어</p> <p>받아들이다, 보다, 보이다, 생각하다, 알다, 여기다, 인정하다, 주장하다, 확신하다</p> <p>③ 반(反)사실성 술어</p> <p>기대하다, 당부하다, 믿다, 바라다, 예상하다, 요구하다, 원하다, 착각하다, 희망하다</p>	<p>술어의 유형을 가급적 제한하지 않고, 내포절에 대한 화자의 확신성을 담은 용례를 최대한 다양하게 포착한다.</p> <p>(2) (변동) 1차 연구보다 넓은 맥락을 고려하여 내포문 명제의 화자 확신성을 추론한 정보를 부착함으로써 문장을 넘어서는 맥락이 내포문의 명제(가설)에 대한 확신성에 영향을 미치는 경우까지 반영한다.</p>
<p>대상 담화의 맥락 구조</p>	<p>(1) 대상 담화는 선행 문장과 대상 문장으로 구성되어 있다. 선행 문장은 확신성을 판단하는데 영향을 미치는 경우에 한하여 제시한다. 최대 2개의 선행 문장을 제시할 수 있다.</p> <p>(2) 내포 명제는 대상 문장을 단문으로 바꾼 것을 말한다.</p>	<p>(1-1) (변동) 대상 담화는 대상 문장과 선행 문장, 그리고 후행 문장 또한 포함한다.</p> <p>(1-2) (변동) 선행 문장과 후행 문장의 길이를 장르별로 달리하여 제시한다. 문어는 3문장까지 제시하고, 구어는 5~10문장까지 제시한다.</p> <p>(2) 변동 없음</p>
<p>대상 문장 선정 원칙</p>	<p>(1) 모절에 함의 최소 운용소를 포함하는 내포문일 것</p>	<p>(1) (변동) 모절에 함의최소운용소는 포함되지 않을 수 있다.</p>

	<p>(2) 술어가 인용문, 관형절을 형성하고 있어 함의 관계가 파악에 문제가 있는 문장은 제외한다.</p> <p>(3) 난해하거나 부자연스러운 문장은 제외한다.</p>	<p>(2-1) (신규) 부가절, 의사관계절, 의문사절, 낯선 옛글 또한 제외한다. 또한 장르별로 판단을 달리할 수 있다.</p> <p>(2-2) 정보구조는 선행행 문장(문어 3문장, 구어 5~10 문장)을 참조하여 보강할 수 있으면 제외하지 않고 사용한다.</p> <p>(3) 변동 없음</p>
<p>내포 명제 기술 원칙: 오류, 복원, 인칭대명사, 대용어</p>	<p>(1) 주어, 술어 및 문장 성분은 그대로 기술하되, 맞춤법과 기호 사용 오류는 수정한다.</p> <p>(2) 내포 명제의 주어가 없는 경우, 문맥상 예측할 수 있는 선행 주어를 복원한다. 논항이 아닌 부가어는 제거할 수 있다.</p> <p>(3) 1·2인칭 대명사 ‘나’ 또는 ‘너’는 나(필자), 나(화자), 너(청자), 나(사람이름), 너(사람이름) 등으로 기술한다. 또는 대명사가 생략된 사례에 한정하여 직접적으로 ‘청자’를 복원할 수 있다. 다만, ‘우리’는 괄호 속에 병기하지 않을 수 있다. 또한 구어 장르와 달리 문어 장르는 2인칭 대명사가 독자를 가리키므로 당신(청자 또는 독자)로 기술하지 않는다.</p>	<p>(1) 대상 문장은 맞춤법과 기호 사용 오류 또한 그대로 두는 것을 원칙으로 한다. 대상 문장에서 생성된 내포 명제는 확실성 평가에 사용되므로 오류를 수정한다.</p> <p>(2) (변동) 선행 주어와 선행 목적어를 복원한다. 논항이 아닌 부사어나 수식어 또한 복원한다. 주격 표지 ‘은/는’ ‘이/가’는 복원 과정에서 가장 자연스러운 것으로 한다.</p> <p>(3) (변동) 재귀사 ‘자기, 자신’은 선행어에서 복원하되, ‘나, 너, 우리, 그’는 복원하지 않는 것을 원칙으로 한다. 단, 확실성을 파악하기 어려운 경우 복원할 수 있다. ‘청자’와 ‘독자’는 모두 괄호 속에 병기하지 않는다. 의인화된 문장은 제외한다.</p>
<p>설문 조사 또는 언어 실험</p>	<p>(1) 7점 척도를 사용한다 (-3 ~ +3).</p> <p>(2) 시험 공정은 말뭉치에서 추출한 100문항과 인위적으로 생성한 344문항을 대상으로 24명의 참여자가 설문 조사를 하였다. 설문 시간은 세트당 30분으로 한다.</p> <p>(3) 응답 세트는 본 문항 20개, 통제 문항</p>	<p>(1) (변동) 리커트 7점 척도를 사용한다(1 ~ 7). 1은 확실성이 매우 낮은 것을 가리키며, 7은 확실성이 매우 높은 것을 가리킨다.</p> <p>(2) (변동) 기관윤리위원회의 심의 대상에 해당하므로 IRB 승인을 받았다. 참여자의 적격성 또는 “성실도”를 확인하기 위하여</p>

	<p>3개로 구성되어 있다. 문어, 신문, 구어 장르를 포괄하여 총 90세트이다. 각 세트별로 8명에서 20명 내외의 응답을 수집하였고, 총 응답자 수는 844명이다. 다만, 108명의 응답은 제외되어 최종적으로 736명의 응답을 유효응답으로 하였다.</p> <p>(4) 기초통계분석(평균, 표준편차 등)외에 크리펜도르프 알파(Krippendorff's alpha)를 사용하여 응답자 신뢰성 분석을 한다.</p>	<p><u>119명의 대상자에게 48개의 필러(filler) 문항을 풀게 하였다.</u> 최종적으로 60명이 선정되었다. 피로효과(fatigue effect)를 고려하여 <u>세트당 20분의 설문 시간을 두었다.</u></p> <p>(3) (변동) 응답 세트는 40~50개 문항으로 구성되어 있다. <u>응답자는 최종적으로 선정된 60명이며, 예비 설문을 풀고, 설문에 대한 해설과 사례를 교육받았다. 응답자는 무작위로 20명으로 구성된 3개 그룹으로 나뉘어, 응답 세트를 풀도록 하였다.</u> 응답자의 평균 응답 시간은 15분 내외로 측정되었다.</p> <p>(4) (변동) 기초통계분석(평균, 표준편차 등)외에 <u>급내상관계수(Intraclass correlation coefficient)</u>를 사용하여 응답자 신뢰성 분석을 한다.</p>
<p>언어 정보 주석 지침: 시제, 인칭, 함의취소 운용소</p>	<p>(1) 모절 술어와 내포절 술어의 형태(morph)를 기준으로 시제 정보를 부착한다. 이형태(allomorph)를 구별하지 않는다. 의문형 종결어미는 시제소를 주석하지 않는다.</p> <p>(2) 모절과 내포절의 인칭이 일치하는지 Y/N으로 주석한다. 주어가 표면에 드러나지 않는 일반칭은 1인칭과 동일한 [1] 값을, 필자/화자가 배제되는 경우는 3인칭과 동일한 [3]값을 부착한다.</p> <p>(3) 모절의 함의취소운용소 정보, 즉 ‘부정, 의문, 조건, 양태’ 정보를 부착한다. 함의취소운용소정보가 중첩되는 경우, (침표)를 사용한다 예: [부정, 의문].</p>	<p>(1) 작년도 지침을 준용하여 변동은 없으나, <u>온라인 워크벤치에서 시제 목록을 보여주고 작업자가 적합한 시제를 선택하도록 한다.</u></p> <p>(2) (변동) 주어가 표면에 드러나지 않으면 <u>복원하여 인칭을 주석하도록 하고, 복원할 수 없으면 ‘알 수 없음’으로 둔다.</u> 마찬가지로 인칭 목록을 먼저 제시하여 작업 오류를 방지한다.</p> <p>(3) (변동) 모절의 함의취소운용소는 수의적인 정보이므로, <u>함의취소운용소가 있으면 그 종류를 표기하고, 없으면 ‘없음’으로 둔다.</u> 함의취소운용소 정보가 중첩되면 +(덧셈)를 사용한다 예: [부정 + 의문].</p>

4.2. 내포 명제 기술 원칙 (붙임 참조)

확신성 담화의 구축 지침은 ‘2020년 함의 분석 및 연구’ 사업에서 제안된 구축 지침의 미비점을 보완하고, 보다 세밀하게 작성하였다. 지침 제안 사항은 내포 명제 기술 원칙과 관련하여 ‘오류’, ‘복원’, ‘인칭대명사’를 중심으로 유형을 정리하고 기술하였다. 주요 용어 설명과 자세한 기술 원칙, 그리고 구체적인 사례는 붙임에 정리하였다.

4.3. 주요 내포 명제 기술 사례 (붙임 참조)

아래 사례는 내포 명제를 기술하는 방법론 중 (1) 오류, (2) 복원, (3) 인칭대명사와 관련된 기술 지침과 사례(문어 장르)들을 일부 소개하였다. 오류는 맞춤법 오류(한국(O), 현국(X)), 잘못된 띄어쓰기(한국(O), 한 국(X)), 인코딩 오류(한국(O), □난국(X)) 등을 말한다. 복원은 대상 문장에 포함된 대명사(이 책, 그 물건, 저것 등)가 가리키는 것을 선후행 문장에서 가져와 바꿔쓰는 것을 말한다. 인칭대명사(그, 그녀 등) 또한 선후행 문장에서 해당 대명사가 가리키는 인물을 말한다.

4.3.1. 오류

◆ 맞춤법 오류, 띄어쓰기, 기호 사용 오류 등은 모델의 견고성 측면에서 오타를 임의로 수정하거나 배제하지 않는 것을 제안한다. 자연어 텍스트 환경에서도 오타가 존재하기 때문이다.

<오류 사례 1: 맞춤법 오류>

<대상 문장>

하지만 A조합원장은 **부속토**를 나눠주면서 전체 조합원들을 대상으로 공지하지 않아 특

정 조합원에게 특혜를 줬다 의혹을 하고 있다.

<내포 명제>

부속토를 나눠주면서 특정 조합원에게 특혜를 줬다.

<오류 사례 2: 띄어쓰기>

<대상 문장>

일본도쿄는 국내 유명 최고경영인들이 자료수집이나 휴가차 방문하는 도시이다. 특히 삼성그룹의 이견희 회장은 매년 시간만 나면 일 본으로 새로운 사업을 구상하러 간다고 한다.

<내포 명제>

삼성그룹의 이견희 회장은 매년 시간만 나면 일 본으로 새로운 사업을 구상하러 간다.

◆ 단, 아래의 사례처럼 실험 문항에 포함될 수 있는 내포 명제에 오타가 발생하는 경우는 오타를 수정할 것을 제안한다.

<오류 사례 3: 맞춤법 오류>

<대상 문장>

100여편의 작품 중 다수의 우수작품들이 독특한 심상을 가지고 있어 우열을 가리기 어 렵다고 전하면서 최종적으로 김용옥 수필가 겸 시인 작품‘지혜의 발견’수필을 선정했다.

<내포 명제>

100여편의 작품 중 다수의 우수작품들이 독특한 심상을 가지고 있어 우열을 가리기 어 렵다.

4.3.2. 복원

◆ 내포 명제는 다음과 같은 방식으로 주어와 술어 및 다른 문장 성분을 가급적 그대로

기술하는 것을 원칙으로 할 것을 제안한다.

<복원 사례 1>

<대상 문장> 성조와 발음이 아무리 중요하다고 해도 ‘언어는 소통을 위한 도구’임을 기억하면 문제는 의외로 쉽게 풀린다
<내포 명제> 언어는 소통을 위한 도구이다.

<복원 사례 2>

<대상 문장> 연초에는 지금과 정반대의 상황이 벌어졌던 것을 기억할 것이다.
<내포 명제> 연초에는 지금과 정반대의 상황이 벌어졌다.

◆ 대상 문장의 내포절에 주어 없이 명제 내용을 파악하기 어려운 경우에는 문맥상 예측할 수 있는 선행 주어를 복원하여 제시한다.

<복원 사례 3: 내포절의 주어 없음>

<대상 문장2> 참다 못해 노조를 만들었지만, 한 달 만에 직장을 잃었다"고 했다.
<내포 명제> 노동자들이 참다 못해 노조를 만들었지만, 한 달 만에 직장을 잃었다

2) <선행 문장> 해고를 당한 노동자들은 "아사히글라스에서 일하는 9년 동안 최저임금 수준만 받으면서 기계처럼 일했다

4.3.2. 인칭 대명사

◆ 내포절에 제시된 인칭 대명사의 선행어를 대상 담화에서 찾을 수 있는 경우에는 선행어를 복원한다.

<인칭대명사 사례 1: 선행어 복원>

<대상 담화> 아무리 머리가 비상하고 탐험정신이 강한 인간들 이라고 해도 자기들의 마음속에 행복이 숨겨져 있다는 것을 깨닫기는 어려울 것이오.
<내포 명제> 인간들의 마음속에 행복이 숨겨져 있다.

<인칭대명사 사례 2: 선행어 복원>

<대상 문장 ³⁾ > 이제 그는 병호의 요구를 피할 수 없음을 깨달은 것 같았다.
<내포 명제> 한봉주는 병호의 요구를 피할 수 없다.

◆ ‘자기’, ‘자신’ 등의 인칭 대명사의 선행어는 대상 담화에서 찾을 수 있는 경우에는 선행어를 복원한다.

<대상 문장 ⁴⁾ > 얼른 자신의 추태를 추스르고 지금까지 그 <재미있던 일>을 생각하고 있던 것처럼 얼
--

3) <선행 문장> 이 말에 **한봉주는** 한참동안 입을 다물었다.

버무리면서, 이번에는 정말로 <재미있던 일>을 생각하기 시작했다.

<내포 명제>

미레야는 지금까지 그 <재미있던 일>을 생각하고 있었다.

위의 사례들은 향후 말뭉치 구축 과정에서 발생하는 오류를 일관되게 처리하고, 구축되는 말뭉치의 품질을 향상시킬 수 있다. 또한 대규모 말뭉치 구축에 필요한 작업자 교육 자료로 활용 가능하다. 보다 구체적인 사례와 자세한 지침은 붙임에 기술하였다.

4) <선행 문장> 「미레야 씨?」 의아해하는 유트의 목소리에, 미레야는 혼자 실실 웃고 있던 것을 깨닫고 얼굴을 붉혔다. 「에, 재미있던 일이라……

제 3 장

확신성 언어 실험

1. 확산성 실험의 대상과 범위

사업의 범위에는 일반 언어 사용자 대상 실험을 통해 내포문을 기반으로 생성한 가설에 대하여 화자가 확산하는 정도를 점수화하는 것이 포함된다. 또한 일반 언어 사용자 8명 이상, 평가 척도 7단계 이상으로 평가하여야 한다. 그리고 관련 실험 결과 및 분석 방법론을 구체적으로 제시하는 것이 포함된다.

이에 따라 본 사업은 수집된 대상 담화 1,650건(과업 요구량 대비 10% 추가 실험)을 실험에 적합한 형태의 가설로 변형하여 일반 언어 사용자의 확산성 판단을 점수화하였다. 또한 일반 언어 사용자 60명(미응답자 2명 제외)을 모집하여, 7단계 리커트 척도로 평가하였다. 60명은 한국어가 모국어인 피험자이다. 피험자를 20명 내외의 3개 그룹으로 나누어, 각 그룹은 실험을 병진적으로 실시하였다. 그리고 관련 실험은 실험통사론을 전공한 박사급 공동연구원이 설계하여 예비 실험과 본 실험을 진행하였다.

관련 선행 연구인 Marneffe et al. 2019에 따르면 영어의 전제투사 원리가 모질의 술어적 선택과 함께 화용론적 상호작용에 의하여 복잡한 양상이 나타난다고 밝혔다. 이 때문에 영어의 사례에서는, 화자의 확산성 점수 또는 판단은 양극단에 치우치지 않고, 점진적으로 증가하거나 감소하는 형태를 보였다. 구체적으로는 내포 명제(가설)가 진리값 또는 사실성이 거짓이거나 참이라도 화자의 확산성은 술어의 선택에 따라 다소 모호한 태도로 나타날 수 있다. 예를 들어, “I know that John liked Mary”와 “I suspected that John liked Mary”은 내포 명제 “John liked Mary”의 사실성에 대한 다른 태도를 취하고 있다. 또한 맥락에 영향을 주는 통사적 요소들, 인칭, 시제 등이 이러한 태도에 영향을 주는 것으로 나타났다. 특히, Marneffe et al. 2019는 영어 사용자의 확산성을 대규모 언어 실험을 통해 포착하였다.

본 사업은 이에 착안하여 한국어의 문장 구조와 사용자의 특성에 적합한 확신성 실험을 설계하고 대규모 실험을 수행하였다. 또한 관련 문항 설계 지침과 실험 수행 지침을 구체적으로 사례와 함께 정리하였다. 이와 관련하여 한국어의 확신성 판단이 어떠한 양상을 보이는지가 본 사업의 범위에 포함된다. 따라서 관련 통계분석을 진행하였다. 특히, 엄격한 통계분석 방법은 급내상관계수를 활용하여 한국어 언어 사용자별, 그리고 문항별 응답 패턴의 신뢰성을 분석하였다.

2. 확신성 실험의 수행 지침 및 사례

2.1. 예비 실험 수행 절차

본 실험을 수행하기 이전에 확신성 판단 실험 설계를 수행하고 공정의 효율성을 도모하기 위해 예비 실험을 수행하였다. 예비 실험의 목적은 60명 이상의 확신성 판단 실험 참여자의 사전교육을 수행하고, 사전에 미응답자 또는 실험에 소극적인 참여자를 제외하기 위함이다. 본 사업은 9개의 예비 샘플 문항을 설계하여 60명의 실험 참여자 인원을 확정하고 2명의 참여희망자는 본 실험에서 제외하였다.

예비 샘플 문항은 각 함의 관계가 동일한 비율로 함의/중립/모순 각 세 문항이 선택되었다. 이는 후속 공정에서 내포 명제의 사실성에 대한 화자의 확신성이 어떠한 패턴으로 나타나는지 파악하기 위함이다. 이후 별도의 해설을 피실험자에게 제공하였다. 아래는 9개의 예비 실험 문항 중, 각각 ‘함의’ 관계와 ‘모순’ 관계에 대한 문항 예시들이다.

< 문항 예시 1 >

[문맥] 귀사는 자율규제단체를 통한 개인정보보호 자율규제 정책이 행정안전부 고시 「 개인정보보호 자율규제단체 지정 등에 관한 규정」을 통해 시행되고 있음을 알고 있습니다.

니까?

[질문] 다음 문장에 대해 필자는 어느 정도로 확신한다고 생각하십니까?

자율규제단체를 통한 개인정보보호 자율규제 정책이 행정안전부 고시 「개인정보보호 자율규제단체 지정 등에 관한 규정」을 통해 시행되고 있다.

[라벨] 함의

[예상 점수] 5-7

[해설]

문항 1의 필자는 주어진 문장 “자율규제단체를 통한…”을 알고 있는지 질문하고 있습니다. 여기서 필자는 법률 규정이 실제 시행되고 있고, 이러한 사실을 알고 있는지를 묻는 의도를 가지고 있습니다. 따라서 필자는 주어진 문장이 사실임을 확신한다고 볼 수 있습니다.

[구성의도]

본 문항의 모절의 진실적(veridical) 응답(responsive)술어 알다(know)는 함의취소용소 “-ㄴ니까?”와 결합하였다. 또한 (비)사실성(veridicality) 술어와 연관성이 높은 보문소 “-음/ㄴ”과 결합하고 있다. 이 때, 화자(i.e., the holder of an attitude)가 “-음/ㄴ”이 이끄는 내포절의 명제에 대하여 진실한 답(true answer)을 알고 있다고 볼 근거가 문맥에 있다. 단, 내포절의 주어와 모절의 주어 “귀사”는 일치하지 않는다.

본 문항은 행정청의 시행령에 관한 이해당사자의 인식 수준을 조사하기 위한 설문조사에서 추출되었다. 따라서 필자의 질문은 위의 「개인정보보호 자율규제단체 지정 등에 관한 규정」이 적법하게 고시되어 구속력이 있음을 전제한다. 그러나, 응답자는 (1) 필자(내포절의 주어, 행안부)가 누구인가를 유추하는 과정 또는 (2) 의문문의 통사적 형태가 확신성을 저해한다고 응답자가 추론할 수 있다는 점에서 확신성 판단이 유보될 수 있다. 그러나 내포절에 대한 필자의 태도는 모호하거나 거짓임을 전제하지 않으므로, 1-4점을 부여한 응답자는 실험문항을 구성의도와 다르게 해석한 것이다.

<문항 예시 2>

<p>[문맥] 7월이 되니 날이 급하게 더워졌다. 습기로 가만히만 있어도 온몸이 축축했다. 여름에는 꽃이 별로 없을 줄 알았다. 집집마다 담장 아래 접시꽃이 한창이다.</p> <p>[질문] 다음 문장에 대해 필자는 어느 정도로 확신한다고 생각하십니까?</p> <p>여름에는 꽃이 별로 없다</p> <p>[라벨] 모순 [예상 점수] 1-3</p>
--

<p>[해설]</p> <p>필자는 여름에 꽃이 피지 않는다고 생각했지만 실제로 7월의 무더위를 겪어보니 집집마다 꽃이 피어있는 경험을 했습니다. 때문에 주어진 “여름에는 꽃이 별로 없다”가 거짓이라고 확신하고 있다고 볼 수 있습니다. 이 때, 근거는 주어진 문장 뒤에 이어지는 “집집마다 담장 아래 접시꽃이 한창이다”이다 입니다.</p>
<p>[구성의도]</p> <p>모절의 진실적 응답술어 “-알다”와 의문 보문소 “-줄”이 결합하였다. 화자는 내포절의 명제가 거짓임을 알고 있으므로 사실성이 있다.</p> <p>일간지의 수필 기고면에서 변형없이 그대로 추출하였다.</p>

위의 사례를 포함한 9개 문항에 대하여 62명의 실험 참여 희망자가 응답하였다. 사후 기술통계 분석을 실시한 결과, 아래의 도표처럼 피실험자의 평균 확신성 점수와 문항 구성 의도가 일치하는 것으로 나타났다.

<표 7. 예비 실험 기술 통계 분석>

	문항 1	문항 2	문항 3	문항 4	문항 5	문항 6	문항 7	문항 8	문항 9
평균	6.34	5.88	6.18	3.49	3.57	3.15	2.4	2.17	2.48
표준편차	1.06	1.17	1.20	1.22	1.11	1.26	1.74	1.71	1.01

95% 신뢰구간	[6.07, 6.59]	[5.59, 6.16]	[5.89, 6.47]	[3.11, 3.78]	[3.30, 3.84]	[2.85, 3.45]	[1.99, 2.83]	[1.75, 2.58]	[2.23, 2.71]
예상점수	5 - 7	5 - 7	5 - 7	3 - 5	3 - 5	3 - 5	1 - 3	1 - 3	1 - 3

예비실험에서 평균은 실제 응답자 62명의 평균 확신성 점수를 말한다. 예상 점수는 연구원이 사전 실험 설계 단계에 함의 관계에 대한 확신성 점수를 추정한 값을 말한다. 모든 문항에 대하여 연구원이 추정한 점수와 95% 신뢰구간이 상당 부분 일치하므로, 연구원의 문항 설계가 응답자의 확신성 점수 패턴에 적합하였음을 알 수 있다.

또한 예비 실험의 결과를 토대로 적정 문항 응답 시간을 도출할 수 있었다. 적정 문항 응답 시간은 1회의 실험에 적정한 개수의 문항을 제시하는 기준이 된다. 함의 분석 응답의 정확도가 매우 높은 응답자 5인을 선별하여 별도의 문항으로 정확한 응답시간을 재측정하였다. 선별된 5인의 응답자가 별도의 42개 문항에 대하여 응답한 시간을 측정하여 평균 소요 시간을 적정 문항 응답 시간으로 정하였다.

- 응답자 1. 820초 (13분 32초)
- 응답자 2. 600초 (10분 07초)
- 응답자 3. 900초 (15분 00초)
- 응답자 4. 930초 (15분 32초)
- 응답자 5. 1020초 (17분 00초)
- 평균. 854초 (문항당 20.33초)

문항 당 평균 20.33초 소요되는 것으로 파악되었다. 또한 응답자의 문항이해도가 점진적으로 높아지므로, 추후 본 실험에서는 확신성 실험 1회에 50개 문항을 실험하는 것으로 정하였다.

이러한 예비 실험 결과를 토대로, 본 실험에서 사용될 문항 구성에 대

한 설계 실험 수행 절차를 계획하였다.

2.2. 본 실험

2.2.1 문항 설계 지침 및 주요 사례 (붙임 참조)

본 실험에 사용된 문항에 대한 설계 지침 및 구체적인 사례는 별도 첨부 파일에 정리하였다. 다양한 한국어 문형의 처리에 관한 지침을 정비하였다. 작업자의 착오 또는 실험 참여자의 오해에서 발생하는 여러 오류 유형을 정리하고, 아래와 같은 39개 지침을 첨부하였다. 여기서는 문항 설계에 관한 개략적인 내용과 주요 사례를 소개한다. 보다 구체적인 사례는 붙임에 자세히 기술되어 있다.

2.2.1.1. 한국어의 정보 구조

지침 (1) 내포 명제로 자연스러운 한국어 문장 기술이 힘든 경우, 앞뒤 문맥을 고려할 수 있다. 단, 앞문맥을 우선적으로 고려할 것을 제안한다. 그러나 아래의 예시처럼 뒷문맥이 지나치게 긴 경우, 뒷 문맥을 고려하지 않는다.

<대상 담화⁵⁾>

포격은 잠시도 쉬지 않고 밤새도록 계속되었다. 그리고 그는 인천이라는 항구 도시가 이제는 지구상에서 영원히 말살되었을 것으로 생각했다.

<내포 명제>

인천이라는 항구 도시가 이제는 지구상에서 영원히 말살되었을 것이다

지침 (2) 시제정보를 내포절에 기술하여 정보구조적으로 완성시키는 것을 제안한다.

<대상 담화>
나는 스스로가 경멸스러워질 지경이었다. 그때까지 여자와 잠을 자본 일이 없었다는 사실은 또 얼마나 수치스레 여겨졌던가.
<내포 명제>
그때까지 나는 여자와 잠을 자본 일이 없었다

지침 (3) 구조적으로 내포명제가 명확하지 않더라도 하더라도, 의미적으로 내포명제를 보강할수 있으면 사용할 것을 제안한다.

<대상 담화6>
일반적으로 민법 제4편 친족법과 제5편 상속법을 가족법이라 한다.
<내포 명제>
"민법" 제5편 상속법은 가족법이다

2.2.1.2. 한국어의 통사 구조

지침 (4) 보문절이 아닌 부가절(adjunct)은 사용하지 않는 것을 제안한다. 따라서 실험에서 제외한다.

<대상 담화7>
예준 역시 래프팅을 한 번도 해본 적이 없던 터라 ^{부가절} 흔쾌히 ‘오케이 사인’을 보냈다.

- 5) <선행/후행 문장> ... 한데 다음날 아침 상륙용 주정으로 도착해서 바라본 인천은, 그토록 치열했던 함포의 포격에도 불구하고 그의 예상과는 달리 상당 부분이 파괴를 모면한 채 반 이상이 건재했다.
- 6) <선행/후행 문장> 정치·경제 용어사전 가족법 家族法 친족적 공동생활, 신분의 승계(承繼), 신분에 기인하는 재산의 승계를 규율하는 실체법.

<내포 명제>

예준은 래프팅을 한 번도 해본 적이 없었다 (제외)

지침 (5) 의사관계절(pseudo relative clause)는 실험에서 제외한다.

<대상 담화⁸⁾>

그가 몸을 움직일 때마다 온몸의 뼈가 무너져내리는 것 같은 소리의 의사관계절가 났다.

<내포 명제>

온몸의 뼈가 무너져내리다. (제외)

지침 (6) 내포 명제에 의문사가 포함되는 경우는 제외할 것을 제안한다.

<대상 담화⁹⁾>

상처 난 마음을 어떻게 의문사 다독거려야 하는지 그는 알 수 없었다

<내포 명제>

상처 난 마음을 어떻게 다독거려야 했다 (제외)

2.2.1.3. 한국어의 의미 구조

지침 (7) 내포절 안에 있는 양태 표지는 그대로 유지하는 것을 원칙으로 할 것을 제안한다.

- 7) <선행/후행 문장> 콘서트 관람을 마치고 서연이 불쑥 래프팅 제안을 했는데, 수철이 맞장구를 쳤다.
- 8) <선행/후행 문장> 그녀의 머리는 헝클어져 있었고 드문드문 새로 자라난 흰머리가 보였다. 희는 이불을 개어 인방에 가져다 놓고 나와 씩크대에 가래침을 뺐었다.
- 9) <선행/후행 문장> 울고 있을지도 모른다. 다케유키는 마음이 아팠다. 하지만 도와줄 방법이 없었다.

<대상 담화>

실제로 일반 구원파 신도들은 현재 유병언 씨에 장례식에 국과수에서 시신을 넘겨주면 그것이 유병언 회장일 것이라고 믿습니까?

<내포 명제>

국과수에서 넘겨주는 시신은 유병언 회장일 것이다. [인식-추측]

<대상 담화>

상대방의 입장에서 **생각해야 한다**는 것을 이해했다면 고민하거나 망설이지 말고 즉시 실행에 옮기면 된다.

<내포 명제>

상대방의 입장에서 **생각해야 한다**. [비인식-의무]

<대상 담화>

나는 윤리학이 어떤 쟁점에 대해 대답해 줄 수 있다고 믿지 않는다.

<내포 명제>

윤리학은 어떤 쟁점에 대해 대답해 줄 수 있다. [비인식-능력(가능)]

지침 (8) 다만, 다음의 경우에는 양태 표지를 제거하고 기술할 것을 제안한다.

<대상 담화>

의사 선생님은 제가 목이 악화됐을 거라고 예상하신 것 같습니다.

<내포 명제>

내(필자)는 목이 악화되었다.

위의 사례에서는, 주어가 1인칭인 명제가 내포절에 있으며, 이를 그대로 내포 명제에 추측의 인식 양태 표지를 사용하여 기술하면 어색하다.

위의 지침과 사례들은 구체적인 지침의 일부이다. 보다 자세한 사항은

별도로 첨부된 지침 파일에서 종합적으로 관리되었다.

2.2.2. 실험 수행 지침 및 절차 (붙임 참조)

2.2.2.1 IRB 승인

본 사업의 언어 실험 수행에 앞서 고려대학교 생명윤리위원회의 심의를 거쳐 IRB 승인을 받았다(연구과제명: 말뭉치 함의 분석을 위한 온라인 언어 실험 / 승인번호: KUIRB-2021-0192-02).

2.2.2.2 피실험자 모집

예비 실험에서 문항들을 가장 많이 맞힌 연구대상자를 “성실도”가 높은 참여자로 판단하였고, 이 중 본 사업의 언어실험에 참여하기를 원하는 인원을 모집하였다. “성실도”가 높은 사람들 중 본 실험 참가 희망자는 총 63명이었으나, 이 중 참가 희망을 묻는 사전 설문 조사에 배치된 문항 적응용 샘플 문항에서 한 번호로만 찍는 등의 아웃라이어(outlier)로 판단되는 인원 3명을 제외하였고, 이에 따라 총 60명의 피실험자를 모집하게 되었다.

2.2.2.3 실험 수행

다음은 전자우편 안내문에 제시된 문항 설명문(문어/구어)이다.

<그림 5. 전자우편 상의 문항 설명문(문어/구어)>

주어진 [문장]에 대해서 [보기]로 주어진 내용의 말하는 이(화자) 혹은 글쓴이(필자)가 어느 정도 확신하는 것 같으냐를 1~7의 정도로 평가해 주시는 방식입니다.

[보기] 철수는 내일 날씨가 흐릴 것으로 예상한다.

[문장] 내일 날씨가 흐릴 것이다

위 [보기]에서 "철수"는 주어입니다. "철수"가 판단한 것과 위 [보기]의 필자 및 화자의 견해가 반드시 같은 것은 아닙니다. 위 [문장]에 대한 필자 혹은 화자의 판단을 기준으로 평가해 주시면 됩니다.

또한 실험을 진행하시는 여러분들의 판단을 묻는 것이 아니니 이점도 참고 부탁드립니다.

언어 실험은 코로나 상황으로 인하여 비대면으로 실험을 진행하였다. 온라인 언어 실험 플랫폼은 피실험자의 편의와 친숙도를 고려하여 구글 서베이(Google survey)를 활용하였다. 먼저 문항 풀이 이전에 연구참여 설명서 및 연구참여동의서 등을 배치하였으며, 이후 실험에 대한 안내문을 배치하였다. 구어/문어 및 대화문 실험 안내문의 예시는 아래와 같다.

<그림 6. 대화문 문항>

1/40

[대화]

상만: 어, 그래. 은호원 씨는 마케팅팀으로 가면 되겠다.

호원: 아... 네...

우진: 안되겠는데요?

강호: 망했다.

기택: 왜? 왜?

강호: 치킨치킨..

기택: 렉!

우진: 은호원 씨는 영업팀에서 쓰시죠.

우진: 마케팅이 잘 참는 걸로 되는 것도 아니고, 갑질이 소원인데, 그럴 일도 아니구요. 게다가 무책임하기까지 하던데.

상만: 거 참. 서 부장.... 영업팀엔 그래도 남직원이 낫지!

우진: 성차별적인 위험한 발언이십니까?

[질문] 다음 [문장]의 내용에 대해 위 대화의 우진은 어느 정도로 확신한다고 생각하십니까?

[문장] 마케팅은 잘 참는 걸로 된다

위 [대화]를 읽고 [질문]에 답해주세요. (#401001) *

	1	2	3	4	5	6	7	
거짓이라고 확신한다	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	참이라고 확신한다

뒤로 다음 양식 지우기

실험은 리커트 7점 척도로 진행하였으며, [보기](혹은 [대화]) 문장에 대한

[질문]의 답으로 거짓으로 확신할수록 낮은 점수를, 참이라고 확신할수록 높은 점수를 부여하도록 하였다.

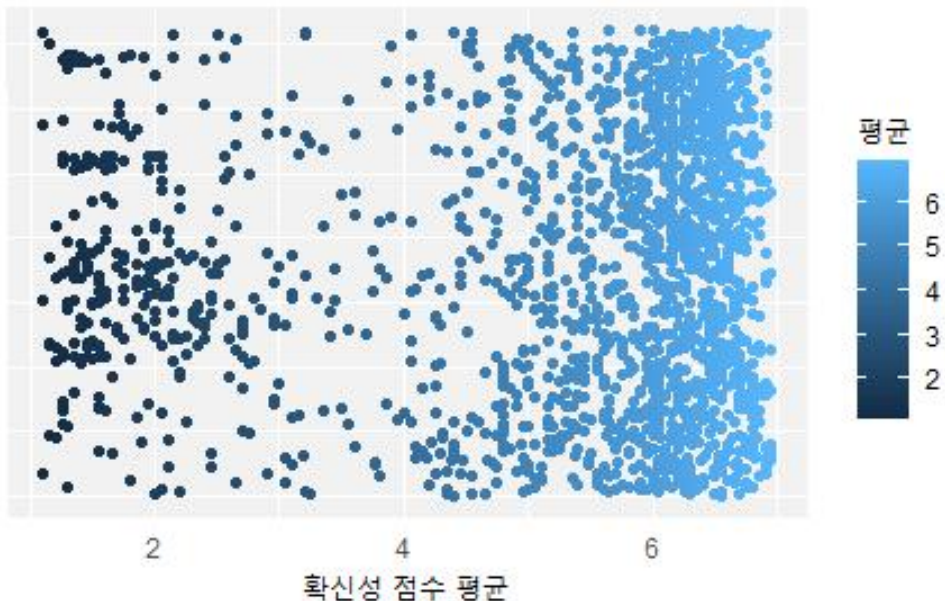
실험 답례비는 1회당 문화상품권 10,000원권이다(개인 전자우편 계정으로 문화상품권의 PIN번호를 발송하였다).

3. 확신성 실험의 결과

3.1. 응답자 점수 평균 분포

총 1,805 문항에 대하여 확신성 실험이 수행되었다. 최초 과제 제안서는 1,500 문항을 대상으로 하였으며, 기타 사유로 인한 문항 손실과 실험 실패로 인한 오류율을 고려하여 목표량의 20%를 추가로 수행하였다. 전체 문항에 대한 응답 분포는 아래 그림과 같다. 가로축은 1,805 문항에 대하여 20명 내외의 참여자가 응답한 점수이다.

<그림 7 확신성 실험의 문항별 응답점수 평균 분포>



위의 문항 응답 자료의 특징적인 부분은 평균점수 분포가 매우 편향되어 있다는 점이다. 특히, 중립과 함의의 경계에 있다고 볼 수 있는 4~6점 구간에 분포하는 문항이 많다는 점은 화자의 확신성이 중립과 함의의 사이에서 모호하게 나타날 수 있다는 점을 시사한다.

<그림 8. 확신성 실험 문항 분석과 함의 관계 라벨 부착>

IDX	TARGET	PROMPT	MEAN	SD	95% CI lo	95% CI up	label
105025	최근 출간된 '나는 습관을 조금 바꾸기로 했다'는 이런 삶의 태도에 대한 그의 고찰을 '습관'이라는 주제를 실천 강령으로 삼았다. 작가는 프리랜서로 전향하며 습관도 관리와 노력이 필요를 통감했다 한다.	습관도 관리와 노력이 필요하다	5.4	1.1877	4.879468	5.920532	entailment
105026	혹시 이런 식으로 오만무례한 자세를 가진 목사가 있었다면 그들에게야말로 '그들의 민중교회'는 바야흐로 끝장이 나고 만 길세. 소련이 무너졌기 때문에 민중 교회가 허덕이고 운동권이 무기력해졌다는 소리는 참으로 간사한 기회주의자들의 허튼 수작일세.	소련이 무너졌기 때문에 민중 교회가 허덕이고 운동권이 무기력해졌다	1.5	0.9459	1.085441	1.914559	contradiction
105029	이는 소모적인 이데올로기 대립이라는 불길에, 한국 사회의 '신용 자본'을 태워먹는 지도층 전반에 대한 불신의 불길에 끼이는 휘발유가 아닐 수 없다. 그렇다고 전향서를 내야 한다는 것이 아니다.	전향서를 내야 한다	1.35	0.5871	1.092692	1.607308	contradiction
105032	내려가기 위해 지스가 카인 쪽으로 향했다. 안전하게 카인의 무릎을 거쳐 내려갈 생각이었다.	지스는 카인의 무릎을 거쳐 내려갈 것이다	5.45	1.0501	4.989773	5.910227	entailment
105033	그가 이번에 문학이론서인 이 책을 집필한 것은 기독교문학에 대한 사명 때문이라고 고백하고 있다. 그러나 사명이 있다고 누구나 이 일을 할 수 있는 것은 아니다.	사명이 있다면 누구나 문학이론서를 집필하는 일을 할 수 있다	1.25	0.5501	1.008908	1.491092	contradiction

실험 결과인 리커트 척도를 토대로 함의 관계 정보를 부착하였다. 구체적으로는, 각 문항별 응답자의 확신성 점수의 평균을 구하고 95% 신뢰구간을 도출하여 이를 고려하였다. 이때, 예비 실험과 동일하게 1~3점의 구간을 모순으로 하였고, 3~5점 구간을 중립으로, 5~7점 구간을 함의로 하였다. 예를 들어, 어떤 문항의 신뢰구간의 길이가 [5.5, 6.5]라면, 신뢰 구간이 5~7점 사이에 모두 포함되므로 함의 관계 정보가 부착되었다.

반대로, 어떤 문항의 신뢰구간이 [1.5, 3.0]이라면 모순 관계의 구간에 포함되므로 함의 관계 라벨은 “모순”으로 부착되었다. 이를 엑셀 수식을 이용하여 자동적으로 개별 문항에 따라 부착하였고, 그 비율은 아래 표와 같다. <표 8>은 실제 확신성 실험에서 사용된 전체 문항 수의 함의 관계가 라벨 분포를 보여준다. 1,805건의 문항에 대하여 20명 내외의 한국어 화자가 응답하였다. <표 9>는 전체 실험 문항 1,805건에서 내부 검토를

거쳐 243건을 제외한 최종 산출물의 함의 관계 라벨 분포이다.

<표 8. 전체 확산성 실험 문항 수>

<표 9. 최종 산출물 수>

내포 명제 추론 관계	문항 수	비율	내포 명제 추론 관계	문항 수	비율
함의	1,314	72.80	함의	1,129	72.28
중립	208	11.52	중립	172	11.01
모순	282	15.62	모순	260	16.65
전체	1,805	100 (%)	전체	1,562	100 (%)

<표 8>과 <표 9>에서 나타난 바와 같이 중립/모순에 비하여 함의 관계에 문항 분포가 상당히 편중되어 있다. 특히, 함의에 가까운 중립이나, 반대로 중립에 가까운 함의로 화자의 확산성이 나타나는 경우가 별도의 분류 없이 포함되었기 때문일 수 있다. 그러나 별도의 분류 범주를 설정하더라도, 연속형 변수(continuous variable)을 범주적 변수(categorical variable)로 바꾸는 과정에서 상당한 정보의 손실이 발생하며, 자의적(ad-hoc)이라는 약점을 충분히 방어하기 어렵다. 따라서 문항을 별도로 분류하여 목적에 맞게 사용하는 작업은 사용자의 의도에 따라 바뀔 수 있다.

3.2. 확산성 실험 결과의 통계적 분석

3.2.1. 급내상관계수를 활용한 응답자 신뢰도 분석

급내상관계수(Intra Class Correlation; ICC)를 이용한 확산성 언어 실

험에 참여한 응답자의 신뢰성을 평가를 하였다. 사회과학에서 흔히 사용되는 급내상관계수는 평가자들의 응답에 대한 신뢰도를 측정하는데 사용하는 통계적 도구이다. 본 사업에는 기존 연구에 따라 ICC1k 또는 ICC2k 중 하나가 확신성 실험과 관련하여 가장 신뢰할 만한 통계값이라고 제안한다 (Yum et al., 2021). 해당 연구는 의학 전문가 10인의 언어적 직관을 수집하기 위하여 실험언어학의 방법론을 사용한 것으로, ICC를 이용하여 의학 전문가 직관의 신뢰성을 통계분석하였다.

ICC는 무작위로 응답자 그룹을 형성하여 응답자 간의 응답패턴을 분석하는 방법이다. 이를 통하여 응답자의 응답 패턴이 일관적인지 확인할 수 있다. 각 평가별로 총 6가지 ICC값과 F값, 자유도값, p-value등이 나타난다. ICC1 (단일 절대 평가)은 응답을 무작위로 하나 선정하여 각 문항마다 매번 다른 응답으로 평가하는 단방향 ANOVA 모델이다. 응답과 절대 합치도간의 편차에 민감하다. ICC2 (단일 무작위 평가)는 응답 번호 (1-7)의 절대 합치도(absolute agreement)를 평가하였다. 각 문항별로 k개의 응답을 무작위로 추출한다. ICC3 (단일 고정 평가)는 더 큰 응답의 모수를 위한 일반화는 진행하지 않았다. 각 문항별로 고정된 k개의 응답을 사용한다. ICC2와 ICC3는 평가자 간의 평균 편차를 없애는 역할을 하지만 응답별로 평가자 간의 상호관계에 민감하다. ICC2와 ICC3의 차이는 랜덤한 추출에 얼마나 민감한가를 보여준다.

ICC1k (평균 절대 평가), ICC2k (평균 무작위 평가), ICC3k (평균 고정 평가)는 각각 k명의 응답자를 반영한 결과이다. 본 과제에서는 급내상관계수를 이용하여 응답 인원들의 내포 명제에 대한 확신성이 비슷한 패턴을 보이는지 또는 상이한 패턴을 보이는지 파악하였다. 이때, 리커트 7점 척도를 직접 사용하지 않고, 함의(5-7점) - 중립(3-5점) - 모순(1-3점)으로 변환하였다. 이는 함의 관계 라벨의 분포와 응답자의 신뢰성을 파악하기 위함이다. 또한, 각 척도의 1-7점 점수에는 중요성을 부여하기 어렵지만, 이를 함의 관계별로 다시 그룹화하여 통계분석하면 확신성의 양상을 분석하는데 용이하다. 구체적으로는, 함의와 중립 사이에서 모호

한 응답이 상당수 분포하였는데, 모호한 응답이 많을수록 ICC값이 떨어진다.

급내상관계수 측정은 총 3가지 방법으로 진행하였다. 첫 번째는 문항별 급내상관계수 종합 측정 (문항-인원-평가)이다. 1,805건의 실험 문항은 20명 내외의 응답이 있었고, 응답의 분포가 균일할수록 ICC값이 높다. 반대로, 20명의 응답자가 하나의 문항에 대하여 제각각의 점수를 부여하면 ICC값이 떨어진다. 두 번째는, 함의 관계별 급내상관계수 종합 측정 (함의-인원-평가)이다. 함의관계라벨을 그룹화하여 급내상관계수를 측정하는 것이다. 예를 들어, 1,805건의 실험 문항에서 1,314건이 ‘함의’에 해당하는 문항이었는데, 이 문항들이 하나의 그룹을 형성한다. 이때, 그룹 내에서 ICC값을 측정하는 것이다. 따라서 ‘함의’ 문항 그룹에 비하여 ‘모순’ 문항 그룹의 응답이 제각각일 경우, ICC값이 상대적으로 떨어진다.

세 번째는, 함의별 급내상관계수 조합 측정 (함의-인원-평가)이다. 이 분석은 문항 그룹 내부의 응답 간 신뢰성 분석이 아닌, 문항 그룹과 또 다른 문항 그룹을 대치시켜 신뢰성을 분석하는 것이다. 따라서 문항 함의별 급내상관계수 조합의 경우, 3개의 함의 태그를 기준으로 2개의 조합을 만들어 급내상관계수를 구하였다 (함의-중립, 함의-모순, 중립-모순).

아래 <표 10>~<표 14>는 전체 확산성 실험 문항 1,805건을 대상으로 ICC 신뢰성 분석하여 얻은 통계 결과이다. 일반적으로 상관관계가 0.40 미만이면 신뢰성이 매우 낮은 것으로 평가된다. 1.00에 가까울수록 신뢰도가 높다. 또한 평가방법에는 ‘단일’ 평가가 있고 ‘평균’ 평가가 있다. 단일 평가는 하나의 응답자를 가정하며, 평균 평가는 하나 이상의 응답자들의 응답 평균을 가정한다. 따라서 응답자의 수가 많을수록 단일 평가와 평균 평가의 신뢰도 계수(또는 상관관계)가 크게 차이날 수 있다. 아래 <표 10>~<표 14>에서 나타나듯이, 평균 평가의 신뢰성은 문항별, 라벨별, 함의별로 매우 높게 나타나며(0.86 ~ 0.97), 따라서 확산성 실험의 응답은 신뢰할 만하다.

<표 10. 전체 데이터셋 기준 문항별 급내상관계수 종합>

유형	평가 방법	상관관계	F 통계량	자유도	95% 신뢰구간
ICC 1	단일 절대 평가	0.715	43.83	1804	[0.7 0.73]
ICC 2	단일 무작위 평가	0.715	44.11	1804	[0.7 0.73]
ICC 3	단일 고정 평가	0.717	44.11	1804	[0.7 0.73]
ICC 1k	평균 절대 평가	0.977	43.83	1804	[0.98 0.98]
ICC 2k	평균 무작위 평가	0.977	44.11	1804	[0.98 0.98]
ICC 3k	평균 고정 평가	0.977	44.11	1804	[0.98 0.98]

<표 11. 전체 데이터셋 기준 함의별 급내상관계수 종합>

유형	평가 방법	상관관계	F 통계량	자유도	95% 신뢰구간
ICC 1	단일 절대 평가	0.662	34.43	1804	[0.65 0.68]
ICC 2	단일 무작위 평가	0.662	34.57	1804	[0.65 0.68]
ICC 3	단일 고정 평가	0.663	34.57	1804	[0.65 0.68]
ICC 1k	평균 절대 평가	0.970	34.43	1804	[0.97 0.97]
ICC 2k	평균 무작위 평가	0.970	34.57	1804	[0.97 0.97]
ICC 3k	평균 고정 평가	0.971	34.57	1804	[0.97 0.97]

<표 12. 전체 데이터셋 기준 함의 라벨별 급내상관계수 (모순-중립)>

유형	평가 방법	상관관계	F 통계량	자유도	95% 신뢰구간
ICC 1	단일 절대 평가	0.410	12.83	201	[0.36 0.47]
ICC 2	단일 무작위 평가	0.410	13.00	201	[0.36 0.47]
ICC 3	단일 고정 평가	0.413	13.00	201	[0.36 0.47]
ICC 1k	평균 절대 평가	0.922	12.83	201	[0.91 0.94]
ICC 2k	평균 무작위 평가	0.922	13.00	201	[0.91 0.94]
ICC 3k	평균 고정 평가	0.923	13.00	201	[0.91 0.94]

<표 13. 전체 데이터셋 기준 함의 라벨별 급내상관계수 (모순-함의)>

유형	평가 방법	상관관계	F 통계량	자유도	95% 신뢰구간
ICC 1	단일 절대 평가	0.966	493.092	221	[0.96 0.97]
ICC 2	단일 무작위 평가	0.966	493.44	221	[0.96 0.97]
ICC 3	단일 고정 평가	0.966	493.44	221	[0.96 0.97]
ICC 1k	평균 절대 평가	0.997	493.092	221	[1. 1.]
ICC 2k	평균 무작위 평가	0.997	493.44	221	[1. 1.]
ICC 3k	평균 고정 평가	0.997	493.44	221	[1. 1.]

<표 14. 전체 데이터셋 기준 함의 라벨별 급내상관계수 (중립-함의)>

유형	평가 방법	상관관계	F 통계량	자유도	95% 신뢰구간
ICC 1	단일 절대 평가	0.262	7.03	1281	[0.24 0.28]
ICC 2	단일 무작위 평가	0.262	7.09	1281	[0.24 0.28]
ICC 3	단일 고정 평가	0.264	7.09	1281	[0.25 0.28]
ICC 1k	평균 절대 평가	0.857	7.03	1281	[0.85 0.87]
ICC 2k	평균 무작위 평가	0.858	7.09	1281	[0.85 0.87]
ICC 3k	평균 고정 평가	0.859	7.09	1281	[0.85 0.87]

아래 <표 15> ~ <표 19>는 데이터 필터링을 거친 1,562개 문항에 대하여 동일한 방법으로 신뢰성 분석을 하였다 (단, 문항 오류 또는 문항의 적절성으로 인하여 5개 문항이 제외되기 이전의 1,567개 문항에 대하여 신뢰성 분석을 진행하였다). 마찬가지로, 신뢰성 분석의 방법에 상관없이 응답자의 신뢰성은 매우 높게 나타난다.

<표 15. 데이터 필터링 후 문항별 급내상관계수 종합>

유형	평가 방법	상관관계	F 통계량	자유도	95% 신뢰구간
ICC 1	단일 절대 평가	0.723	45.53	1566	[0.71 0.74]
ICC 2	단일 무작위 평가	0.723	45.79	1566	[0.71 0.74]
ICC 3	단일 고정 평가	0.724	45.79	1566	[0.71 0.74]
ICC 1k	평균 절대 평가	0.978	45.53	1566	[0.98 0.98]
ICC 2k	평균 무작위 평가	0.978	45.79	1566	[0.98 0.98]
ICC 3k	평균 고정 평가	0.978	45.79	1566	[0.98 0.98]

<표 16. 데이터 필터링 후 함의별 급내상관계수 종합>

유형	평가 방법	상관관계	F 통계량	자유도	95% 신뢰구간
ICC 1	단일 절대 평가	0.671	35.80	1566	[0.66 0.69]
ICC 2	단일 무작위 평가	0.671	35.92	1566	[0.66 0.69]
ICC 3	단일 고정 평가	0.672	35.92	1566	[0.66 0.69]
ICC 1k	평균 절대 평가	0.972	35.80	1566	[0.97 0.97]
ICC 2k	평균 무작위 평가	0.972	35.92	1566	[0.97 0.97]
ICC 3k	평균 고정 평가	0.972	35.92	1566	[0.97 0.97]

<표 17. 데이터 필터링 후 함의 라벨별 급내상관계수 (모순-중립)>

유형	평가 방법	상관관계	F 통계량	자유도	95% 신뢰구간
ICC 1	단일 절대 평가	0.406	12.64	184	[0.35 0.46]
ICC 2	단일 무작위 평가	0.406	12.87	184	[0.35 0.47]
ICC 3	단일 고정 평가	0.411	12.87	184	[0.36 0.47]
ICC 1k	평균 절대 평가	0.920	12.64	184	[0.9 0.94]
ICC 2k	평균 무작위 평가	0.921	12.87	184	[0.9 0.94]
ICC 3k	평균 고정 평가	0.922	12.87	184	[0.91 0.94]

<표 18. 데이터 필터링 후 함의 라벨별 급내상관계수 (모순-함의)>

유형	평가 방법	상관관계	F 통계량	자유도	95% 신뢰구간
ICC 1	단일 절대 평가	0.967	503.78	190	[0.96 0.97]
ICC 2	단일 무작위 평가	0.967	504.50	190	[0.96 0.97]
ICC 3	단일 고정 평가	0.967	504.50	190	[0.96 0.97]
ICC 1k	평균 절대 평가	0.998	503.78	190	[1. 1.]
ICC 2k	평균 무작위 평가	0.998	504.50	190	[1. 1.]
ICC 3k	평균 고정 평가	0.998	504.50	190	[1. 1.]

<표 19. 데이터 필터링 후 함의 라벨별 급내상관계수 (중립-함의)>

유형	평가 방법	상관관계	F 통계량	자유도	95% 신뢰구간
ICC 1	단일 절대 평가	0.253	6.76	1107	[0.23 0.27]
ICC 2	단일 무작위 평가	0.253	6.81	1107	[0.24 0.27]
ICC 3	단일 고정 평가	0.254	6.81	1107	[0.24 0.27]
ICC 1k	평균 절대 평가	0.852	6.76	1107	[0.84 0.86]
ICC 2k	평균 무작위 평가	0.852	6.81	1107	[0.84 0.86]
ICC 3k	평균 고정 평가	0.853	6.81	1107	[0.84 0.87]

결과를 종합하여 보면 응답자들은 모순-함의 라벨을 결정하는 데 매우 높은 일치도를 보였다(급내상관계수: 0.977). 즉, 모순과 함의를 구분하는 데는 극히 일부를 제외한 인원이 서로 일치하는 언어적 직관을 보여주었다고 판단할 수 있다. 반면 중립-함의 라벨을 결정하는 데는 상대적으로 낮은 일치도를 보였으며 특히 단일 응답자를 기준으로 평가한 ICC1, ICC2, ICC3에서 매우 낮은 점수를 보였다(0.262, 0.262, 0.264). 이는 특정 문장쌍이 중립인지 함의인지 판단하는 데 있어 개개인의 가치관이 다르다는 것을 의미한다.

이와 달리, 중립-함의와 상대적으로 모순-중립의 경우에는 비교적 높

은 일치도를 보였다. 이는 응답자들이 문장을 함의라고 판단할 때보다 모순이라고 판단할 때 특정 어구 등의 언어적 표현에 민감하게 반응한 결과라고 해석할 수 있다.

3.2.2. 아웃라이어 응답자 분석

아웃라이어(outlier) 응답자는 다른 응답자와 달리 통계적으로 확연히 구별되는 응답패턴을 가진 응답자를 말한다.

<그림 10. 개인별 제외 후 ICC 측정>

excluded	max ICC	average ICC	ICC1	ICC2	ICC3	ICC1k	ICC2k	ICC3k
#15	0.976641658	0.849486537	0.722003	0.722032	0.723237	0.976501	0.976504	0.976642
#16	0.97659439	0.849269057	0.721625	0.721653	0.722823	0.976458	0.976461	0.976594
#04	0.977238358	0.852484031	0.727447	0.727471	0.728508	0.977119	0.977122	0.977238
#10	0.976788711	0.850208001	0.7233	0.723329	0.72453	0.976649	0.976652	0.976789
#12	0.976829304	0.850519363	0.723964	0.723985	0.724887	0.976724	0.976727	0.976829
#03	0.976879427	0.850648707	0.724087	0.724116	0.725329	0.976738	0.976742	0.976879
#09	0.976770339	0.850114085	0.723128	0.723157	0.724368	0.976629	0.976632	0.97677
#08	0.976937379	0.850931162	0.724593	0.724621	0.725841	0.976796	0.976799	0.976937
#17	0.976415372	0.848379589	0.720014	0.720043	0.721257	0.976273	0.976276	0.976415
#18	0.976506122	0.848820889	0.720805	0.720834	0.72205	0.976364	0.976367	0.976506
#19	0.976626997	0.849428809	0.721913	0.721941	0.723109	0.97649	0.976494	0.976627
#07	0.976992246	0.851213973	0.725113	0.725141	0.726325	0.976855	0.976858	0.976992
#13	0.976819493	0.850351144	0.72355	0.723579	0.724801	0.976677	0.976681	0.976819
#01	0.977512448	0.853970669	0.730238	0.730254	0.730952	0.977433	0.977434	0.977512
#06	0.977007433	0.851296634	0.725268	0.725296	0.72646	0.976873	0.976876	0.977007
#05	0.976846651	0.850501255	0.723835	0.723863	0.72504	0.97671	0.976713	0.976847
#14	0.976700296	0.849802642	0.722598	0.722625	0.723752	0.976569	0.976572	0.9767

때문에, 아웃라이어 응답자가 부여한 화자의 확신성 점수는 신뢰하기 어렵다. 따라서 응답자들이 다양한 문항에 골고루 응답을 표시하였는지 판단할 필요성이 있었다. 또한, 특정 인원의 응답이 다른 사람들의 직관과 크게 벗어나는가 그렇지 않은가를 판단하는 작업 또한 중요하다.

각 인원을 제외 후 최대 급내상관계수와 평균 급내상관계수를 측정하여 비교해 보았을 때 차이가 많이 나는 인원의 경우 응답 기준이 일반적이지 않다고 판단할 수 있었다. 평가 결과 각 인원을 제외한 이후임에도 전체적인 급내상관계수의 차이가 없는 것으로 보아 특이사항은 없음을 확인하였다.

<그림 11. 인원별 응답 분포 비교>

peep	1	2	3	4	5	6	7	diversity
#01	131	82	50	126	165	326	687	7
#03	146	84	50	125	213	400	549	7
#04	128	86	62	200	236	404	451	7
#05	142	97	47	179	198	372	532	7
#06	113	107	48	189	231	404	475	7
#07	130	88	51	164	253	396	485	7
#08	128	98	42	153	231	403	512	7
#09	142	75	61	157	202	352	578	7
#10	143	88	59	174	193	359	551	7
#12	125	104	41	140	176	293	688	7
#13	136	128	32	124	168	424	555	7
#14	147	97	41	168	218	397	499	7
#15	124	109	50	135	249	389	511	7
#16	152	106	39	153	207	365	545	7
#17	164	92	46	129	168	349	619	7
#18	140	97	41	146	215	379	549	7
#19	122	120	42	125	173	424	561	7

각 인원별 1에서 7까지의 선택지별 응답 빈도를 측정하여 얼마나 다양한 응답을 하였는가 측정하여 3개 이하의 선택지를 고른 인원의 경우 설문에 불성실하게 답변한 것으로 간주하였다. 모든 인원이 7개의 선택지를 모두 사용하였으므로 특이사항은 없는 것으로 확인되었다.

제 4 장

확신성 언어 실험 결과의 국어학적 분석

1. 확신성 언어 실험 결과의 활용과 범위

확신성 언어자원의 구축과 활용의 대상에는 언어정보를 부착하는 공정과 인공지능 평가가 포함된다. 언어정보 부착의 범위는 수집된 대상 담화 사례 1,500건에 대하여 말뭉치의 형식으로 시제, 인칭 등의 언어정보를 부착하는 것이다. 작업에 대한 온라인 교육은 단기간의 집중 교육과 작업의 완성도에 따라 수시로 이루어졌다(Intensive 1-day Workshop 또는 Interim Check Workshop). 또한 작업자(석박사과정 보조연구원)의 결과물에 대하여 상위 층위의 작업자(박사급 공동연구원) 별도의 검증절차를 수행하여 오류율을 낮추고 작업 품질을 향상시켰다.

언어정보 부착은 한국어 화자의 확신성을 계량적 척도에서 측정하는 방법과 달리, 함의 관계에 영향을 줄 수 있는 통사/의미 이론적 구성물을 보완적으로 태깅하는 것을 포함한다. 여기에는 전통적인 함의 취소 운용소는 포함되거나 포함되지 않을 수 있으나, 시제와 인칭 정보, 주어 일치 여부, (유사)보문소 등 한국어 언어 이론에서 제안되어왔던 구성물에 대한 태깅을 수행하였다. 이러한 언어 정보는 함의 관계에 대한 구체적인 관점을 제시할 뿐만 아니라, 제 3 장에서 수행되었던 한국어 언어 사용자의 확신성 양상과 비교할 수 있는 토대를 제공한다. 이는 기존의 한국어 연구를 풍부하게 할 수 있다는 점에서 의의가 있다.

인공지능 평가는 확신성 언어실험의 통계적 결과를 종합분석하여 자연어 추론 벤치마크로 재가공하는 공정이 포함되었다. 자연어 추론 벤치마크는 인공지능의 한국어 능력을 평가하는 일종의 테스트 도구로 기능한다. 인공지능 평가에서는 내포절의 사실성에 대한 일반 언어 사용자의 확신성 판단과 한국어 인공지능 모델(KorBERT, KoBERT 등)이 유사한 패턴을 보이는지 상관관계를 분석하였다. 상관관계는 일반적으로 Pearson's r 척도로 측정한다.

2. 언어 정보 부착 수행

2.1. 확산성 언어정보 부착의 작업 지침 (붙임 참조)

해당 내용은 붙임에 작업 지침과 함께 구체적인 사례를 종합하여 정리하였다.

2.2. 작업자 교육

작업자 교육은 확산성 언어정보 부착에 대한 기본적 이론을 익히고 작업 지침을 숙지하는 것뿐만 아니라 작업자가 워크벤치를 이용하여 직접 정보 부착 작업을 해 볼 수 있도록 국어학 전문가와 함께 실습하였다. 교육은 총 두 번의 워크숍을 통해 진행되었다. 작업자가 본격적인 작업에 들어가기에 앞서 국어학 전문가와 함께 실시한 Intensive 1-day Workshop에서는 확산성 언어 정보에 해당되는 시제, 인칭, 함의 취소 운용소에 대한 기본적 이론을 중심으로 교육이 이루어졌다. 전문가와 함께 실제 워크벤치에서 다양한 예문을 이용하여 정보를 부착하며 작업에 대한 이해도를 높였다.

전문가와의 충분한 논의 후 시제는 내포명제를 기준으로, 시제소는 형태를 기준으로 정보를 부착한다는 기준을 세울 수 있었다. 이를 바탕으로 작업자는 각자 정보부착 작업을 진행하였고 한 달 뒤에 다시 모여 작업 결과물 중간 점검 및 품질 관리를 위한 Interim Check Workshop을 실시했다. 전문가는 지금까지의 작업 결과물을 살펴보고 작업자에게 개선해야 할 부분을 설명하였다. 이후에는 각 작업자가 언어정보 부착 작업을

진행하며 메모한 의문점 및 특이사항을 바탕으로 전문가에게 질문하고 전문가는 이에 대해 답변을 하며 작업자들이 이후 언어정보 부착 작업을 진행할 때 숙지할 지침을 다시 한번 강화하였다.

2.3. 온라인 워크벤치

원활한 언어정보 부착 작업을 위해 확신성 담화의 주석 작업에 활용된 워크 벤치를 기능적으로 확장하였다. php 및 MySQL 소프트웨어를 사용하여 온라인 워크벤치를 준비하였다. 해당 워크벤치의 구성은 아래 사진과 같다. 앞선 워크벤치와 마찬가지로 선행, 후행 맥락은 버튼을 클릭하였을 때만 등장하도록 하였으며, 맥락 정보와 내포명제 정보는 더 이상 수정하지 못하되 표시만 되도록 구상하였다.

<그림 12. 언어정보 부착 워크벤치>

▲ Show/Hide the Previous Context

[WBRW1900013174.1] 6/189 [202006] (문어)

"무염식의 자연식 요법으로 모든 질병을 치유할 수 있다고 말할 수는 없습니다.

▼ Show/Hide the Next Context

본문소: <input type="text" value="다고"/>	국어원 본문소: 다고
본문술어: <input type="text" value="말하다"/>	국어원 본문술어: 말하다

Context + Target
 그때까지만 해도 거슨은 커피관장을 시도해보지 않았기 때문에 그에 대해서는 말하지 않았다. 당시에는 의사들 각자의 견해와 기회에 따라 식염수관장이나 온수관장 등이 이루어지고 있었다. "무염식의 자연식 요법으로 모든 질병을 치유할 수 있다고 말할 수는 없습니다. 그러나 대부분의 질병 치료에 큰 도움을 준다고 말할 수는 있을 것 같습니다."

내포명제:
 무염식의 자연식 요법으로 모든 질병을 치유할 수 있다

내포절: 모절과의 주어 일치 여부: <input type="radio"/> 일치 <input checked="" type="radio"/> 불일치 시제: <input type="radio"/> 과거 <input checked="" type="radio"/> 현재 <input type="radio"/> 미래 시제소: <input type="radio"/> 은/는 <input type="radio"/> 았/았 <input type="radio"/> (으)르_것/겠 <input checked="" type="radio"/> 없음 <input type="checkbox"/> 일반적인 사실 기술 인칭: <input type="radio"/> 1st <input type="radio"/> 2nd <input checked="" type="radio"/> 3rd <input type="radio"/> 알 수 없음	모절: 시제: <input type="radio"/> 과거 <input checked="" type="radio"/> 현재 <input type="radio"/> 미래 시제소: <input type="radio"/> 은/는 <input type="radio"/> 았/았 <input type="radio"/> (으)르_것/겠 <input checked="" type="radio"/> 없음 인칭: <input type="radio"/> 1st <input type="radio"/> 2nd <input type="radio"/> 3rd <input checked="" type="radio"/> 알 수 없음
---	---

ECO:
 표지:
 양태 부정 의문 조건 그 외 없음

기타:
 민감 사안
 확인 요망

국어원 비교:
 context+target 수정
 국어원에 대한 답변:

메모:

해당 워크벤치는 내포절과 모절에 대한 시제, 시제소, 인칭 정보 및 함의 취소운용소(ECO) 정보 등을 태깅할 수 있도록 설계되었다. 이에 더해, 국립국어원에서 제시한 수정 사항을 내부 정보와 더불어 표기하여 내부 정보를 수정하고, 국립국어원에서 남긴 의견에 효율적이며 체계적으로 답변

할 수 있도록 하였다.

이때, 아래 그림과 같이 보문소와 보문술어의 경우 국립국어원 제안 사항을 항상 표시하되, 내부 정보와 다를 경우에는 UI상 색깔을 다르게 표시하여 작업자가 이를 반영할 수 있도록 하였다. 맥락정보와 내포명제의 경우, 워크벤치 화면이 너무 길어지는 경우를 방지하기 위하여 내부 맥락 및 내포명제가 국어원에서 제시한 맥락 및 내포명제와 차이가 있을 경우에만 표시되도록 하였다.

또한 해당 워크벤치에서는 미처 거르지 못한 정치적, 성적 내용 등의 민감할 수 있는 사항을 거르기 위하여 ‘민감 사항’이라는 체크박스를 제공하였으며, 이전 워크벤치와 마찬가지로 ‘확인 요망’ 체크박스를 통해 작업자가 재검토를 요청할 수 있도록 하였다.

<그림 13. 국립국어원 제안 내용과 내부 정보가 다른 예시>

보문소:	국어원 보문소:
<input type="text" value="_것_"/>	<input type="text" value="나_것_"/>
보문술어:	국어원 보문술어:
<input type="text" value="이다"/>	<input type="text" value="이다"/>

Context + Target
 6월 시에 따르면 올해 노인일자리 및 사회활동 지원사업에 총 558억3천만원을 투입해 2만363명의 어르신들에게 일자리를 제공한다. 이는 지난해(411억5천500만원, 1만5천899명)에 비해 예산은 146억7천500만원(35.6% 증)이, 참여자 수는 4천464명(28.1%)이 늘어난 것이다.

내포명제:
 지난해에 비해 예산은 146억7천500만원(35.6% 증)이, 참여자 수는 4천464명(28.1%)이 늘어났다

국어원 내포명제:
 지난해에 비해 노인일자리 사업 예산은 146억7천500만원이, 참여자 수는 4천464명이 늘어났다

위의 작업 예시는 언어정보 부착에서 보문소와 보문술어에 대한 작업 과정이다. 확신성 담화의 핵심적인 정보인 보문소와 보문술어는 국립국어원 제안 사항을 표시하여 참조할 수 있도록 하였다. 본 사업팀의 작업자들이 도출한 견해와 국립국어원의 제안 사항이 상이할 경우 인터페이스의 색을 달리하였다. 위의 작업 절차는 발주기관인 국립국어원과 본 사업팀의 의사소통을 원활히 하기 위함이다.

또한 언어정보 부착은 3개 층위에서 종합적으로 작업 검수를 하였다. 낮은 층위에서는 기초작업자 또는 보조연구원이 언어정보를 부착하였고,

높은 층위에서는 의미론을 전공한 박사급 공동연구원이 검수하고, 공동연구원이 보완하였다.

2.4. 작업 결과물

내포절과 모절의 시제정보와 주어의 인칭, 내포절과 모절의 주어 일치 여부에 대해 처리하였다. 모절과 내포절 모두 과거, 현재, 미래를 확인하여 해당 정보를 부착하였다. 내포절의 시제소는 내포절 명제를 따로 추출한 것이 아닌 문장 상에서 나타난 시제소에 대해 처리하였다. 이때 시제소는 현재를 나타내는 은/는, 과거를 나타내는 었/았, 미래를 나타내는 (으)ㄴ-것/겠으로 작성하였다. 그 중 문장에서 시제소가 뚜렷하게 나타나지 않는 경우 ‘없음’으로 설정하였다.

<표 20. 언어 정보 부착의 기준>

내포절					모절				
시제	과거		현재	미래	시제	과거	현재	미래	
시제소	은/는	였/았	(으)ㄴ- -것/ -겠	없음	시제소	은/는	였/았	(으)ㄴ- -것/ -겠	없음
인칭	1인칭	2인칭	3인칭	알 수 없음	인칭	1인칭	2인칭	3인칭	알 수 없음
모절과의 주어 일치 여부	일치		불일치		인칭	1인칭	2인칭	3인칭	알 수 없음

주어의 인칭은 문장과 context에서 주어가 명확하게 드러나지 않은 경우에는 ‘알 수 없음’으로 처리하였고 그 외 1인칭, 2인칭, 3인칭임을 부

착하였다.

내포절의 주어는 모절의 주어와 일치하는 경우 일치, 그렇지 않은 경우는 불일치로 작성하였다. 아래의 도표는 시제 및 주어 및 함의 취소 운용소에 대한 언어 정보를 부착한 결과물을 도표로 나타낸 것이다.

담화 구조에 대한 정보는 (1) 고유번호(idx) (2) 장르 (genre) (3) 선행 문맥 (prev) (4) 대상 문장(current) (5) 후행 문맥(next) (5) 모절 술어 (pred) (6) 보문소 (comp) (7) 선후행 맥락을 연결한 대상 문장 (context+target) (8) 내포 명제(prop) (9) 확신성 관계(relation)를 포함하고 있다.

<표 21. 언어정보 부착의 예시>

idx	SBRW1900010062.1
genre	구어
prev	P4: 실제로 한 대학에서 실험을 했는데 P2: 예. P4: 같은 염도 하지만 매운맛의 정도만 다르게 했었는데 P1: 예. P4: 이~ 실험자의 대부분이 더 매운 음식도 싱겁다 이렇게 느꼈었다고 합니다. P2: 오~ P4: 그 이유를 보니까 P2: 예. P4: 매운맛의 정도가 올라갈수록 짠맛을 정상적으로 느끼지 못했다는 건데 P1,P2: 음~
current	P4: 결과적으로는 데 많은 양의 소금이 들어가야 짠맛을 느낄 수 있었다라는 이야기가 되겠죠.
next	P1: 그랬다. 예. P2: 예. P1: 짜여니 님이 지금 스트레스도 풀리고 건강에도 좋은데 실제로 그런 건가요 느낌인 건가요라고 질문을 주셨어요. P6: {laughing} P5: 그~ 스트레스와 관련해서는 실제로 매운 것을 먹으면 스트레스에 더 취약하다는 쪽이 정설입니다. P1: 예. P5: 매운 걸 먹게 되면 통증을 느끼잖아요. P2: 음~ P5: 아리니까 통증을 느끼게 되면 다른 스트레스 일시적으로 잊어버릴 수가 있거든요. P1: 예.
pred	되다
comp	라는_이야기
context+target	P4: 실제로 한 대학에서 실험을 했는데 P2: 예. P4: 같은 염도 하지만 매운맛의 정도만 다르게 했었는데 P1: 예. P4: 이~ 실험자의

	대부분이 더 매운 음식도 싱겁다 이렇게 느꼈었다고 합니다. P2: 오~ P4: 그 이유를 보니까 P2: 예. P4: 매운맛의 정도가 올라갈수록 짠맛을 정상적으로 느끼지 못했다는 건데 P1,P2: 음~ P4: 결과적으로는 데 많은 양의 소금이 들어가야 짠맛을 느낄 수 있었다라는 이야기가 되겠죠. P1: 그랬다. 예. P2: 예. P1: 쭈여니 님이 지금 스트레스도 풀리고 건강에도 좋은데 실제로 그런 건가요 느낌인 건가요라고 질문을 주셨어요.
prop	매운맛의 정도가 올라갈수록 많은 양의 소금이 들어가야 실험자의 대부분이 짠맛을 느낄 수 있었다
내포절-모절 주어 일치 여부	불일치
내포절 시제	과거
내포절 시제소	있/았
내포절 주어 인칭	3rd
모절 시제	현재
모절 시제소	없음
모절 주어 인칭	3인칭
함의 취소 운용소 표지	-겠-
함의 취소 운용소	양태

워크벤치의 언어정보 부착 결과물을 기준으로 하여 작업자들이 제작한 prop과 국어원의 prop이 다를 경우에는 서로 다른 두 개의 prop 중 어 떠한 것을 고를지에 대한 소견을 정리하여 기술하였다. 이 외에도 국어원 비고를 통해 개별 결과물의 불완전성에 대해 지적이 있는 경우, 재검수 후 검수자의 판단에 따라 상황에 적절한 답변을 기술하였다. 또한 국어원 으로부터 정치/사회적으로 민감하거나 논란의 소지가 있는 이슈이기에 제외 요청이 있는 경우에는 재검수 후 제외에 대한 동의 의사를 밝힌 후 제거하였다.

국어원과의 소통을 더욱 원활히 하기 위해 워크벤치에 국어원에 대한 답변란을 추가하여 개별 결과물에 대한 수정 및 제거 여부를 명확히 밝

히고자 하였다. 아래의 두 사례는 각각 국어원과 prop에 있어서 견해가 다른 경우와 prop에 대한 국어원의 질의가 있는 경우에 답변을 기술한 것이다. 첫 번째 사례에서는 작업자들이 선정한 prop을 유지하겠다는 견해를 밝히고 있으며 두 번째 사례에서는 시제 서술에 대한 국어원의 질의에 대해 답변을 하며 근거를 밝히고 있음을 확인할 수 있다.

<표 22. 언어 정보 부착의 예시>

idx	NLRW2000000005.437
genre	신문
prev	울산시에 따르면 울산 지역 악취 민원은 2016년 739건, 2017년 637건, 2018년 735건에 달한다. 빛공해의 경우 2016년 62건에 불과하던 민원이 2018년에는 88건으로 45%가량 급증했다. 민원이 가장 많이 들어오는 소음·진동 관련 민원은 2016년 3007건, 2017년 4344건, 2018년 2773건으로 집계됐다.
current	특히 대규모 아파트 공사와 원룸 공사가 많았던 2017년 울주군과 북구에서 소음·진동 민원이 최대 3배 가까이 증가한 것으로 확인됐다.
next	이중 울산시에 환경분쟁조정 신청이 접수된 분쟁은 2016년 3건, 2017년 8건, 2018년 4건으로, 2016년 지반침하 분쟁 조정신청 1건을 제외하곤 나머지 14건이 전부 소음과 관련된 민원이다. 환경부와 울산시에서는 환경분쟁 민원의 원인을 규명하고 신속하게 해결하기 위해 환경분쟁조정위원회를 구성해 운영하고 있다. 환경분쟁조정위원회는 민사소송에 비해 상대적으로 적은 비용으로 피해 구제를 받을 수 있다는 장점이 있으나, 주민들이 이용하기엔 절차가 복잡하다.
pred	확인되다
comp	ㄴ_것으로_
context+target	민원이 가장 많이 들어오는 소음·진동 관련 민원은 2016년 3007건, 2017년 4344건, 2018년 2773건으로 집계됐다. 특히 대규모 아파트 공사와 원룸 공사가 많았던 2017년 울주군과 북구에서 소음·진동 민원이 최대 3배 가까이 증가한 것으로 확인됐다.
prop	2017년 울주군과 북구에서 소음·진동 민원이 최대 3배 가까이 증가하였다
국어원 prop	2017년 울주군과 북구에서 소음·진동 민원이 최대 3배 가까이 증가했다
내포절-모절 주어 일치 여부	불일치
내포절 시제	과거
내포절 시제소	없음

내포절 주어 인칭	3rd
모절 시제	과거
모절 시제소	였/았
모절 주어 인칭	3rd
함의 취소 운용소 표지	없음
함의 취소 운용소	없음
기타 (민감 사안/확인 요망)	
국어원에 대한 답변	하였다/했다는 직관의 차이인 것 같으며 큰 의미가 없으므로 내포 명제를 유지하는 것이 좋을 것으로 사료됩니다.

2.5. 검증 절차 수행

언어정보 부착 작업의 정확성과 신뢰성을 확보하기 위하여 1차 부착 작업 이후에 작업자간 교차 검증 절차를 수행하였다. 예를 들어, 연구보조원 A가 모절의 시제소 ‘-였/았’이 없다고 판단하여 언어 정보가 누락 되었다면, 연구보조원 B가 검수하여 누락된 정보를 보완하였다. 교차 검증은 1차 작업자의 작업 결과를 전수로 재검수하는 방식으로 진행되었다. 이와 함께, 교차검증 작업자는 1차 작업자와 의견이 불일치하거나 여전히 추가적인 검증이 더 필요할 것으로 판단되는 사례에 ‘확인 요망’ 또는 ‘민감 사안’과 같은 별도의 표시를 한 후 그 이유를 작성하였다.

또한 1차 작업자가 작업 중 재확인이 필요하다고 판단하여 남긴 메모에 대해 교차검증 작업자는 이를 참고하여 작업물을 수정하거나, 수정을 하지 않을 시에도 그 이유를 메모에 남겨 주요 이슈를 공유할 수 있도록 하였다. 아래 <표 22> 모절의 시제소(-였/았)와 인칭(1인칭, 3인칭)에 관한 1차 작업 및 2차 검증 내용 예시이다. 실제 작업은 온라인 워크벤치를 경유하여 이루어졌다.

<표 23. 언어 정보 부착 작업 예시>

대상 문장	1차 작업 내용	2차 검증 내용
당연히 대궐 안에 그녀가 없음을 알면서도 그녀의 존재감이 너무도 또렷해 자신의 모든 감각을 의심했다.	모절 시제소 (없음), 모절 인칭 (1인칭)	주석 이슈: 모절 인칭을 <u>1인칭에서 3인칭으로 수정할 것.</u>
혼사를 꺼내놓고 이처럼 쉽게 물러선다는 건 이미 만반의 준비를 해놓았다는 뜻인지도 몰랐다.	모절 시제소 (없음) 모절 인칭 (불명확, 논의 필요)	주석 이슈: 모절 시제소를 <u>없음에서 있음으로 변경할 것.</u> ('았') 검토 의견: 공동연구원에 문의

또한 명확한 언어 정보 부착이 어려울 경우, 보조연구원 보다 상위의 공동연구원에 검토의견을 문의하도록 하였다. 이와 함께, 교차검증 과정을 체계화하여 워크벤치에 주석 작업이 보완되지 않은 새로운 원시 데이터를 업로드를 하면 이전 단계의 교차검증 작업자와는 다른 작업자를 배정하였다.

3. 확산성 언어 실험 결과의 인공지능 평가 활용 방안

본 사업은 합의 분석과 관련된 확산성 언어자원을 재가공하여 인공지능 평가에 활용하는 구체적인 방법론을 모색하였다. 이를 토대로 대표적인 한국어 인공지능 모델의 평가결과를 도출하였다. 먼저 확산성 언어자원을 인공지능 평가 도구로 활용하기 위해서는 데이터의 불균형을 해소할 필요가 있었다. 이는 인공지능이 통계적 정보에 굉장히 민감하기 때문이다. 또한 인공지능은 판별의 대상이 되는 문장의 길이(토큰 수)에 매우 민감하므로, 언어적 정보가 아닌 텍스트 구조적 특징이 인공지능의 성능

에 문제가 될 수 있다. 이러한 점은 오류 분석에서 다루었다.

또한 확신성 언어자원을 이용한 인공지능 평가에는 회귀 과제가 아닌 분류 과제를 인공지능 평가 과제로 제안하였다. 이는 가장 일반적인 인공지능 평가 방법 중 하나인 GLUE 벤치마크 또는 자연어추론(Natural language inference; NLI) 과제가 회귀과제가 아닌 분류과제로 인공지능을 평가하고 있기 때문이다. 본 사업에서는 이를 준용하여 인공지능 평가를 시행하였다.

마지막으로 언어자원을 재가공하는 공정에서 함의/중립/모순의 비율을 1:1:1에 가깝게 조정하려고 노력하였다. 이는 정답 비율이 편향되어 있으면, 인공지능 평가에 착오적 결과를 초래할 수 있기 때문이다. 따라서 인공지능에게 제시하는 선택지가 되는 함의/중립/모순의 정답 비율이 매우 중요하며, 비율이 균형적 분포를 이루고 있지 않으면 평가가 왜곡될 우려가 있다. 그러나 국어 말뭉치에서 관찰되는 일반적인 언어 사용에서는 내포 명제의 사실성 또는 확신성이 모순 관계인 사례의 비율이 4%에 불과하였다. 본 사업에서는 이 비율을 제 2 장에서 상술하는 것처럼 보완하려고 하였으나, 공정의 어려움으로 이를 크게 보완하지는 못하였다. 따라서 향후 과제에서는 다른 자료(국어 연구 논문 등)에서 내포 명제의 사실성이 모순 관계인 담화 구조를 추가적으로 선별하여 300개 사례 이상이 포함되도록 할 필요가 있다.

3.1. 자연어추론 벤치마크와 인공주석물

자연어추론 벤치마크는 인공지능의 언어이해 수준 또는 언어추론 능력을 객관적으로 평가하기 위한 데이터셋을 말한다. 최근에는 인공지능의 취약점이 추론능력으로 지적되고 있어 평가 데이터셋의 중요성을 더하고 있다. 일반적으로, 논리적 관계가 존재하거나 존재하지 않는 서로 다른

두 문장을 주고, 그 관계를 인공지능에게 묻는 것이다.

문장 1	철수는 영희에게 비둘기를 보았냐고 물었다.
문장 2	영희는 비둘기를 보았다.
선택지	함의/중립/모순
정답	중립 관계

위의 예시에서 문장 1과 문장 2는 ‘영희’ ‘비둘기’ ‘보다’와 같은 동일한 단어로 구성되었지만, 문장 1이 사실이라고 할 때, 문장 2가 반드시 참이라고 할 수 없으므로 논리적 관계는 중립이다. 만약 인공지능이 단순히 어휘가 유사하다고 서로 다른 두 문장이 논리적으로도 유사하다고 판별한다면, 인공지능의 언어추론 능력은 부족하다고 볼 수 있다.

인공지능 모델에 따라 차이가 있지만 자연어추론 벤치마크에서 인공지능은 80 ~ 90% 내외의 정답률을 보인다. 그러나 최근 컴퓨터과학 연구에서는 인공주석물(annotation artifact)에 의하여 정답률이 부풀려지고 있다는 지적이 있었다. 이는 평가 데이터셋이 엄격하게 논리관계를 묻지 못하고, 인공지능이 논리관계가 아닌 패턴을 포착하여 잘못된 방법 또는 휴리스틱으로 높은 정답률을 보인다는 것이다.

해외 선행 연구에서는 이러한 평가의 취약성을 해소할 방법으로 함의 분석 말뭉치 또는 CommitmentBank (de Marneffe et al. 2019)를 재가공하여 활용하는 것이 제안되었다. 이는 복잡한 언어구조에 대한 이해 없이는 정확한 함의 분석이 어렵다는 점에 착안하여, 인공지능이 잘못된 휴리스틱으로 정답률을 부풀리는 것을 최소화할 수 있기 때문이다.

3.2. 평가 방법

자연어추론 벤치마크와 인공주석물에 대한 문제의식에서 출발하여, 최신 한국어 인공지능 모델의 언어추론 능력을 평가하기 위해 본 사업의

결과물을 활용할 것을 제안한다. 또한 한국어의 특성에 적합한 형태로 확신성 판단 실험 결과와 언어정보 주석을 재가공할 수 있는 구체적인 방법론을 제시하였다.

본 사업의 결과물은 자연어추론 벤치마크와 달리 서로 다른 문장의 관계가 인위적이지 않다. 이는 국립국어원 말뭉치에서 추출된 일반 언어 사용자의 자연스러운 발화를 대상으로 하기 때문이다. 또한 대상 담화에서 생성된 가설이 원래 문장의 내포절이기 때문이다.

문장 1	식약청에서 허가한 치료약을 사용해 개발도중 독성문제로 인해 실패할 확률이 거의 없다는 장점이 있다.
문장 2	식약청에서 허가한 치료약을 사용하면 개발도중 독성문제로 인해 실패한다
선택지	함의/중립/모순
정답	모순 관계

위의 예시에서 문장 2는 문장 1의 내포절에서 구성된 내포 명제로, 모문의 술어인 '-ㄴ 할 확률이 거의 없다는 장점이 있다'와 자연스럽게 호응한다.

그러나 모문의 술어가 형태적으로 '없다' '못하다' '아니다'의 부정어 표현인 경우, 내포 명제의 사실성을 반드시 모순으로 확신하는 것은 아니다. 예를 들어, '우리 집이 허물어지지 않았으면 좋겠어'는 부정어 표현 '않았으면'이 있지만 내포 명제 '우리 집이 허물어진다'는 사실을 함의하지 않는다. 이러한 평가 데이터를 보완하는 것은 중요한데, 이는 인공지능이 단순히 부정어의 형태적인 정보만 학습하여 논리적 관계에 대한 이해 없이 정답을 맞힐 수 없도록 하기 때문이다. 따라서 인공주석물 문제를 해소하기 위한 중요한 평가자료이다.

3.3. 평가 결과

본 연구의 인공지능 평가 방법과 절차, 그리고 결과는 다음과 같았다. 학습 데이터가 인공지능 언어모델에 얼마나 잘 활용될 수 있는지 판단하기 위해 한국어를 지원하는 대표적인 BERT 모델인 M-BERT, KoBERT, KR-BERT를 이용하였다. 전체 데이터 가운데 70%(1,256쌍)는 훈련 데이터로, 30%(313쌍)는 검증 데이터로 활용하였다. 정답 라벨과 예측 라벨 간에 학습에 따른 선형적인 상관관계의 강도를 파악하기 위해 상관계수를 측정하였다. 예측 정확도를 측정하여 모델의 실질적인 활용도를 측정하였다.

<표 24. Google Multilingual BERT (M-BERT)>

학습횟수	손실값	학습률	피어슨 상관계수	스피어맨 상관계수	정확도
1	3.42	1.60E-05	0.04	0.04	0.39
2	3.20	1.87E-05	0.08	0.02	0.44
3	3.20	1.69E-05	0.08	0.02	0.45
4	3.18	1.51E-05	0.07	0.01	0.45
5	3.17	1.33E-05	0.08	0.05	0.45
6	3.16	1.15E-05	0.08	0.01	0.45

<표 25. M-BERT 검증 데이터셋 결과>

학습횟수	손실값	학습률	피어슨 상관계수	스피어맨 상관계수	정확도
1	3.42	-	0.07	0	0.05
2	3.20	-	0.17	0.10	0.48
3	3.20	-	0.17	0.10	0.48

4	3.18	-	0.17	0.10	0.48
5	3.17	-	0.17	0.10	0.48
6	3.16	-	0.17	0.10	0.48

<표 26. KoBERT>

학습횟수	손실값	학습률	피어슨 상관계수	스피어맨 상관계수	정확도
1	1.72	2.00E-05	0	0	0.37
2	1.59	1.78E-05	0.11	0.06	0.45
3	1.58	1.55E-05	0.11	0.06	0.45
4	1.58	1.33E-05	0.11	0.06	0.45
5	1.58	1.11E-05	0.11	0.06	0.45
6	1.57	1.11E-05	0.11	0	0.06

<표 27. KoBERT 검증 데이터셋 결과>

학습횟수	손실값	학습률	피어슨 상관계수	스피어맨 상관계수	정확도
1	1.72	-	0.10	0.06	0.46
2	1.59	-	0.10	0.06	0.46
3	1.58	-	0.10	0.06	0.46
4	1.58	-	0.10	0.06	0.46
5	1.58	-	0.10	0.06	0.46
6	1.58	-	0.10	0.06	0.46

<표 28. KR-BERT >

학습횟수	손실값	학습률	피어슨 상관계수	스피어맨 상관계수	정확도
1	1.77	5.34E-6	0.00	0.00	0.34

2	1.54	1.07E-05	0.09	0.04	0.46
3	1.39	1.60E-05	0.24	0.19	0.49
4	0.99	1.98E-05	0.69	0.66	0.66
5	0.46	1.93E-05	0.88	0.88	0.86
6	0.21	1.87E-05	0.96	0.96	0.95

<표 29. KR-BERT 검증 데이터셋 결과 >

학습횟수	손실값	학습률	피어슨 상관계수	스피어맨 상관계수	정확도
1	1.61	-	0.167	0.10	0.45
2	1.59	-	0.167	0.10	0.45
3	1.55	-	0.367	0.32	0.43
4	1.52	-	0.482	0.40	0.47
5	1.66	-	0.546	0.48	0.44
6	1.92	-	0.539	0.47	0.46

모델마다 학습의 효율성에 큰 차이를 보였다. M-BERT, KoBERT와 달리 소규모 언어모델을 지향한 KR-BERT에서 학습 효율을 보여주었다. 훈련 성과를 보았을 때 지속적으로 성능 향상이 나타나는 것으로 보아 올바르게 학습이 이루어지는 것으로 관찰할 수 있었다. 검증 성과를 보아 4 에포크에서 최적의 성능을 나타냈으며 이후로 훈련 데이터에 과적합되어 나머지 새로운 테스트 데이터에 대한 일반화에 실패할 수 있었다.

최적의 결과를 보인 에포크 4에서는 피어슨 점수가 0.5에 근접하여 어느 정도 상관관계를 보인 것이라고 볼 수 있으나, 이는 해석의 여지에 따라 다르며, 일반적으로 0.6 이상을 유의미한 상관관계로 볼 수 있다는 관점에 따라 인공지능 모델의 확신성 이해 능력이 충분하지 않은 것으로 볼 수 있었다. 다만, 정확도 또한 0에서 6까지 라벨을 예측하는 7분지 예측 태스크라는 점을 감안할 때, 매우 어려운 분류과제의 성격에 비하여

일정 부분 유의미한 성과를 보인 것으로 판단되었다.

3.4. 인공지능 모델 오류 분석

본 사업팀에서는 인공지능 모델이 내포 명제의 확산성과 관련하여 한국어 사용자의 직관과 상이한 패턴을 보이는 사례를 대상으로 오류 분석을 실시하였다. 이는 인공지능과 인간의 언어 능력이 어떤 차이를 보이는지 살펴볼 수 있는 중요한 사례이다.

분석 결과를 제시하면, 한국어 인공지능(KR-BERT)은 모순적 추론 관계를 가진 내포 가설을 잘 이해하지 못하는 것으로 나타났다. 한국어의 자연스러운 발화와 한국어 화자의 직관을 토대로 인공지능을 평가하였다는 점에서 인공지능에 취약하지 않은 평가이다. 그러나 아래 사례에서 관찰되듯이, 일부 오류는 다음의 요소들 중 하나 이상이 관여하였을 수 있다. (1) 전제 문장이 매우 길거나(60단어 이상) 복잡하여 충분한 교육을 받은 성인도 어려워할 문장이며, (2) 구어적 특성이 많은(어떻게 -> 어뜨케) 전제 문장은 문어체 텍스트로만 학습한 인공지능에게는 알 수 없는 어휘가 많은 텍스트(Out of Vocabulary (OOV) 문제)로 처리되었을 수 있으며, (3) 내포 명제를 생성하는 과정에서 휴리스틱을 의도적으로 배제하였다는 점에서 오류가 발생하였을 수 있다.

<그림 14. BERT를 이용한 예측 결과 분석 접근법>

함의 예측 결과		중립 예측 결과		모순 예측 결과	
모델이 옳게 예측한 문장	모델이 틀리게 예측한 문장	모델이 옳게 예측한 문장	모델이 틀리게 예측한 문장	모델이 옳게 예측한 문장	모델이 틀리게 예측한 문장
예측 결과를 기반으로 한 오류 분석					

위의 분류체계에 따라 각 함의 관계별로 인공지능 모델이 확신성 평균 점수를 정답으로 할 때, 예측이 맞거나 틀리는 경우를 분류하여 사례를 수집하였다.

<표 30. 오류 분석의 예시 (함의)>

확신성 함의			
전제 문장	내포 명제	정답	예측
<p>지난 십오일 경남 김해의 한 이십 대 여성이 사십 대 남성을 살해한 끔찍한 사건이 벌어졌습니다.처참했던 사건 현장 살해 직후 여자는 바로 경찰에 자수를 했다고 하는데요. 그녀는 왜 이런 범행을 저지른 것 일까요? 피의자 어머니의 미용실에서 처음 만난 두 사람 이후 가끔 연락을 주고 받았다고 하는데요. 거 부에도 불구하고 사건 당일 남자는 여자의 집까지 찾아와 실랑이를 벌 였다고 합니다.</p>	<p>김해의 한 이십 대 여 성과 사십 대 남성은 피의자 여성 어머니의 미용실에서 처음 만난 이후 가끔 연락을 주고 받았다</p>	5	5
<p>경직된 분위기를 부드럽게 바꿀 수 있는 가장 좋은 방법이 바로 웃음을 이용하는 것이다. '웃음은 만국 공통어'라는 말이 있을 정도로, 웃 음에는 사람들의 마음을 열어 주는 힘이 있다. 대화는 통하지 않아도 함께 웃을 수 있다면 서로의 마음</p>	<p>웃음은 만국 공통어이다</p>	5	5

<p>은 이미 받은 열려 있는 것이나 다름없다.</p>			
<p>분리대와 경계블럭 등 안전시설물이 설치돼 있었지만 버스 정류장을 제외한 자전거 전용도로의 일부 구간이 군데 군데 끊겨 있었다. 일부 자전거 운전자들은 자전거 전용도로가 끊기자 인도·자전거 겸용도로로 이동했지만 상당수는 자전거 전용도로가 끊겨있다는 것을 모른채 차도를 그대로 내달렸다. 그때마다 뒤따르던 차량들은 차도를 달리는 자전거 운전자들을 향해 경적을 울리거나 아슬아슬하게 스쳐 지나갔다.</p>	<p>자전거 전용도로가 끊겨있다</p>	<p>6</p>	<p>5</p>
<p>정말 그이들은 '혜성처럼 갑자기' 등장한 것일까. 그이들의 이력을 조금이라도 아는 이라면 결코 '갑자기'라는 단어를 쓸 수 없으리라. 나는 김준호와 손심심이 거둔 우리 소리의 충격이 우리들 고정관념의 빈틈을 찌르는 변법자강이라고 단언한다.</p>	<p>김준호와 손심심이 거둔 우리 소리의 충격이 우리들 고정관념의 빈틈을 찌르는 변법자강이다</p>	<p>6</p>	<p>5</p>

위와 같은 분석체계에 따라 인공지능 모델의 함의판단에 대하여 오류 분석을 진행하였다. 인공지능 모델이 문장의 함의 판단에 대해서는 어느 정도 정확한 모습을 보이나 그 정도에 대해서는 파악이 어려운 것으로 보였다.

<표 31. 오류 분석의 예시 (중립)>

확신성 중립			
전제 문장	내포 명제	정답	예측
누가 보아도 치욕스러운 결혼이었다. 그렇다 해도 평범하고 행복한 여인으로 살 수 있기를 바라는 마음은 죄가 아닐 것이다. “오라버니, 아무 걱정 마세요.” 오히려 덕혜가 영친왕을 위로했다.	덕혜는 평범하고 행복한 여인으로 살 수 있을 것이다	2	2
센터에는 총 5명의 운영 인력을 투입된다. 동구는 센터 개소를 통해 지역청년의 능동적 사회참여기회를 보장하고, 자립기반을 형성해 권익 증진에 기여할 것으로 기대하고 있다.	센터 개소를 통해 지역청년의 능동적 사회참여기회를 보장하고, 자립기반을 형성해 권익 증진에 기여할 것이다.	4	4
추모객들의 눈물 속에서 '천개의 바람이 되어'를 부르고 내려온 합창단은 무대 위에서 끝끝내 참아냈던 눈물을 터뜨리기도 했다. 친구의 품에 안겨 서럽게 울던 정명여자중 2학년 박주경(17·여)양은 "세월호 선체가 있는 목포신항에 와서 노래를 부르려니 사고 당시가 생각나서 157눈물이 멈추질 않는다"며 "안전을 더 소중하게 생각하는 사회가 됐으면 좋겠다"는 바람을 전했다. 추모객들은 맹골수도 깊은 바다에	안전을 더 소중하게 생각하는 사회가 될 것이다	3	4

서 부식된 세월호 선체 앞에서도 슬픔을 감추지 못했다.			
Comenius 같은 교육사상가는 인류의 이상주의에 호소하여 그들로 하여금 제도적인 생활보다도 이성을 존중하게 했다. Comenius는 엄격한 계급적인 차별이 심한 사회에서는 이것을 없애기 위해서 교육을 널리 보급해야 할 것이라고 생각했다. 그러나 Machiavelli는 교육은 전쟁의 씨를 뿌리고 집권자의 권역을 높이는 데 이용되어야 할 것이라고 생각했다.	심한 계급적인 차별을 없애기 위해서 교육을 널리 보급해야 한다	3	5

또한 인공지능이 중립의 범주에 들어서는 판단은 대부분 옳게 하였으나 그 정도에 대한 판단을 옳게 하지 못한 것으로 보였다. 이때, 중간에 섞인 ‘Comenius’와 같은 영어가 모델의 판단에 혼란을 준 것으로 보인다.

<표 32. 오류 분석의 예시 (모순)>

확신성 모순			
전제 문장	내포 명제	정답	예측
‘그럼 저 여자는 제2 황비인가? 근데 서로 꽤 친한가 보네.’ 그들의 대화를 들으면서 지스는 보통 귀족이나 왕족들은 가족 간에 사이가	보통 귀족이나 왕족들은 가족 간에 사이가 안 좋다	1	1

<p>안 좋다고 생각했던 편견을 바꿨다. 카인이야 원체 성격이 그러니까 넘어가더라도 황후와 다른 황비들 간의 사이는 좋아 보였다.</p>			
<p>화석이 간직한 비밀은 또 없을까요? 생물들이 나타난 차례를 알 수 있어요 지구에는 수많은 생물이 살아가고 있어요. 하지만 모든 생물이 한꺼번에 생겨난 것은 아니에요.</p>	<p>모든 생물이 한꺼번에 생겨났다</p>	<p>0</p>	<p>0</p>
<p>여기에 고기만 있었더라면 어르게 보면 조금 더 물릴 수가 있고 조금 더 찢찢할 수 있는 그런 느낌들이 취나물이 향기와 그 맛으로 함께 포장을 또 해주니까 먹는 데 있어서 불편함이 없는데요. 네. 음~ 언제 씹었냐는 듯이 입안이 개운해지는 거 같은데요</p>	<p>취나물은 고기와 함께 먹을 때 불편하다</p>	<p>0</p>	<p>5</p>
<p>희재: 옛날 생각나네. 꼭 춘천 가는 열차 탄 기분이야. 희자: 대체 어딜 가는데 그래? 혜신: (미소로) 아주 괜찮은 맛집이 있대요. 희자: 지금 그 맛집 찾아가는 거야? 혜신: ... 네. 희재: 야, 경치 좋다. 누님, 바깥 구경 좀 하세요. 희자: 우리 인생이 이 기차 같으면 얼마나 좋겠냐? 볼 거다 보면서 여유부릴 수 있다면 말이다. 희재: 그러게요. 요</p>	<p>우리 인생이 기차 같다</p>	<p>1</p>	<p>4</p>

즘 사람들은 다들 빨리 가려고만 해요.			
--------------------------	--	--	--

모순의 경우 절대적인 학습 데이터의 부족으로 비교적 정확한 예측이 어려운 양상을 보여준다. 이는 향후 과제에서 모순 함의 관계 데이터를 추가로 수집할 필요가 있음을 시사한다. 담화문의 경우나 전제 문장이 지나치게 긴 경우 예측이 어려워지는 모습을 보여주었다.

제 5 장

결론 및 제언

1. 결론

본 사업의 대상과 범위는 한국어 확산성 담화를 수집하고, 관련 확산성 언어 실험과 1,500건의 언어정보 부착 공정을 수행하는 것이었다. 본 사업이 달성한 성과는 아래 표에 요약되어 있다.

<표 33. 사업 요약>

	확신성 담화 추출 자료 수	필터링 이전 실험 문항 수	필터링 이후 실험 문항 수	언어정보 부착 수
문어/신문	106,115	1,264	1,021	1,021
구어	59,110	421	420	420
대화	-	121	121	121
총합	165,225	1,806	1,562	1,562

단위(수)는 어절 수 또는 문장 수가 아닌 개별 확산성 담화의 사례를 단위로 하였다. 확산성 담화 추출 자료 단계에서는 ‘대화’가 ‘구어’에 포함되었다. 필터링은 윤리적으로 민감한 사례이거나, 실명이 언급되어 논란의 여지가 있을 수 있는 실험 문항을 제외한 것이다. 본 사업에서는 필터링된 실험 문항을 일률적으로 제거하지 않고, 별도로 마킹하여 관리하였다. 언어정보 부착은 확산성 실험 종료 후, 보문소, 시제, 인칭 등의 정보를 부착한 것이다.

위 사업 수행을 통해 본 사업팀이 달성한 성과는 다음과 같다. 첫째,

고품질의 한국어 확신성 언어 자원을 구축하였다. 기존 국어 말뭉치를 토대로 하여 사업의 연속성을 확보하면서 1,500건 이상의 정제된 확신성 담화를 구축하였다. 또한, 60명 이상의 한국어 화자가 대규모 언어 실험에 참여함으로써 신뢰할만한 실험 결과를 도출하였다. 한편 다양한 컴퓨터 도구를 활용하여 공정을 자동화함으로써 향후 추가적인 말뭉치 구축에 필요한 공정 구축의 토대를 마련하였다. 이처럼 언어학계와 전산학계에 기여할 수 있는 실용적 성과를 거두었다는 점에서 본 사업은 소기의 목적을 달성하였다고 말할 수 있다.

둘째, 박사급 공동연구원과 다수의 (보조)연구원의 면밀한 검토를 거쳐 확신성 말뭉치 구축 지침을 정비하였다. 이는 향후 관련 사업의 효율성 제고에 도움을 줄 것이다. 말뭉치 구축에 참여하는 작업자들의 일관성 있는 작업을 위해 구체적인 지침 마련은 매우 중요하다. 본 사업은 담화 추출, 실험 수행, 언어정보 부착의 공정별로 체계화된 지침을 마련하였고, 구체적인 사례와 함께 별도의 지침 가이드라인을 정비하였다. 확신성 말뭉치 구축 사업이 단발성이 아닌 연속적, 장기적으로 발전할 수 있는 토대가 마련된 것이다.

셋째, 한국어의 함의 관계와 화자의 확신성 이해에 관한 대규모 언어 연구 수행 결과, 기존의 관련 영어 연구에서 나타난 확신성 패턴과 유사한 패턴을 한국어에서도 발견하였다. 다만 연구원과 자문단의 종합적인 견해에 따르면, 영어와 유사하게 내포 가설의 함의 관계와 한국어 화자의 확신성은 반드시 일치하는 것은 아니었다. 한국어 화자의 직관에 따라 이론적 함의 관계를 이해하는 양상은 다소 차이가 있었으며, 이는 화용론적 맥락 요소 등이 관여하였기 때문이라고 볼 수 있다.

넷째, 확신성 국어 자원을 활용하여 인공지능 평가를 수행함으로써 4차 산업 혁명과 정보 산업의 발전에 직접적으로 기여하는 실용적인 성과를 거두었다. 최근의 대규모 인공지능 개발에서는 자연어이해 수준을 점검할 수 있는 벤치마크의 필요성이 더욱 강조되고 있는 상황이지만 구조화된 자연어이해 평가가 부족하다는 문제가 있었다. 본 사업이 확신성 국어 자

원을 재가공하여 수행한 공정은, 한국어 인공지능 평가의 준거기준을 보강하여 제시했다는 점에서 가치가 있다.

종합하면, 본 사업은 확신성 담화 주석 정보 부착과 관련 대규모 언어 실험 수행을 통해 한국어 화자의 국어 사용 양상을 정량적으로 기술하였다. 특히 한국어 연구의 측면에서 대규모 언어 실험을 계획하고 성공적으로 수행하였다는 점에서 주목할 만하며, 이를 활용하여 인공지능 평가의 한 방향을 구체적으로 제시했다는 점에서 의의가 있다.

2. 제언

본 사업은 고품질의 확신성 언어자원을 구축하였다는 점에서 사업의 목적을 달성하였다고 평가할 수 있다. 그러나 사업 수행의 과정에서 발견된 두 가지 문제점은 한계로 남아 있다.

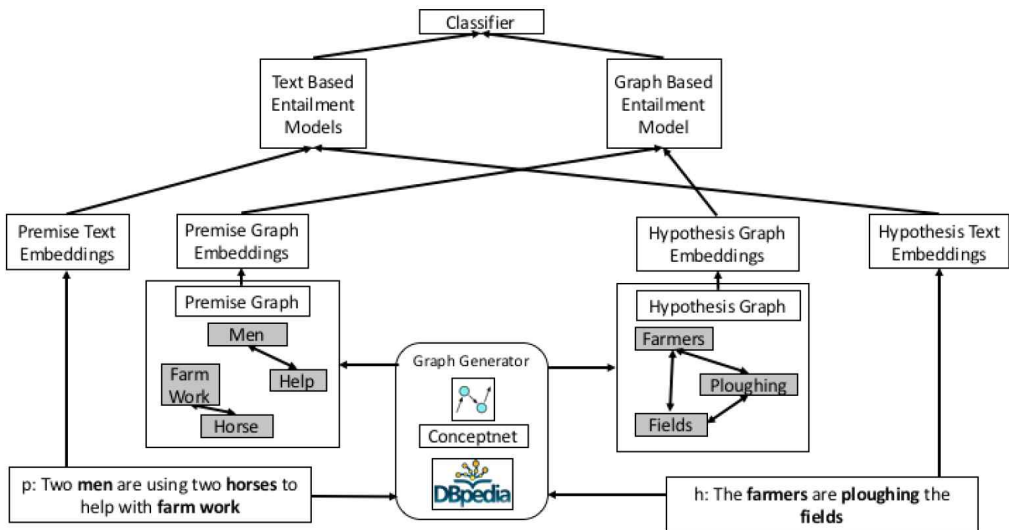
첫째, 구축된 한국어 확신성 담화의 함의 관계가 상당히 편향되어 있다. 이는 실제 언어 사용에서 내포 명제의 사실성과 이에 대한 확신성 정도가 고르게 분포하는 것이 아니라, 편향된 패턴을 보임을 의미할 수 있다. 이러한 언어 사용 패턴을 재가공하여 7분지 분류 문제로 만들 때, 데이터 편향의 문제가 발생할 수 있다. 예를 들어, 전체 분류 문제에 대하여 문제의 이해와 상관없이 하나의 응답만 답변하여도 데이터 편향에 따라 정답률이 매우 높게 나올 수 있기 때문이다. 이러한 문제는 인공 주석물 문제와 연관이 있으며, 본 사업은 이러한 문제를 충분히 해소하지 못하였다. 따라서, 향후 함의 분석 과제에서는 문항 응답의 평가가 고르게 분포할 수 있도록 모순/중립 함의관계와 관련된 추가적인 실험과 문항 추출 작업이 필요할 것이다.

둘째, 확신성 언어 실험 수행 과정에서 문항 응답이 유실되거나 일부 실험 참여자가 응답을 완료하지 않는 등의 문제가 있었다. 하나의 문항에 20명의 실험 참여자가 응답을 하는 것이 원칙이었으나, 실험 수행에 사

용된 구글 서버가 전산적 오류로 인하여 응답을 누락시켜 19명의 응답만 기록된 문항이 일부 존재하였다. 또한 동시 접속한 인원이 중복 응답하여 21명의 응답을 기록한 문항도 존재하였다. 이는 대면 오프라인 실험을 진행할 수 없는 환경에서 불가피하게 발생한 문제로, 향후 온라인 환경에서 관련 언어 실험을 수행할 경우 사전에 이러한 전산 오류 문제를 예방해야 할 것이다.

마지막으로 본 사업팀은 향후 연구과제로 인공지능 평가를 포함한 산업적 활용방안과 국어 말뭉치 연구자원으로서의 활용방안을 제안한다. 인공지능 산업의 성장으로 언어모델의 크기가 기하급수적으로 증가하게 되면서 자연어 이해의 중요성 역시 폭발적으로 증가하고 있다. 자연어 이해의 평가 방법론으로 함의 분석과 확신성 말뭉치가 적극 활용된다면 한국어 언어 모델의 성능 평가에 기여할 수 있다. 또한 지식 그래프(knowledge graph)를 활용하여 모델의 학습이 추가적으로 이루어졌을 때, 함의 분석과 같은 언어적으로 복잡한 현상을 이해하는 성능이 향상되는지 살펴볼 필요가 있다.

<그림 15. ConseqNet 시스템의 아키텍처 (Wang et al. 2019)>



Wang et al. (2019)에 따르면, 단순히 텍스트 정보를 분류하는 것을 넘어서 구조화된 정보를 이용한 지식 추론을 통해 상당한 수준의 성능 향상을 이루어낼 수 있다. 지식 추론이란 전제와 가설이 되는 텍스트로부터 상식 그래프(common sense knowledge graph)를 생성하고, 이를 임베딩하여 두 문장 각각의 그래프 기반 함의 모델을 만들어 분류하는 방식이다. 단순히 텍스트 정보뿐만 아니라 상식 그래프 또한 추가적으로 이용하기 때문에 모델의 성능 향상이 가능한 것이다.

상식 그래프를 이용한 자연어 이해 과제에서도 본 사업에서 구축한 텍스트 데이터가 유용하게 활용될 수 있어 향후 보완적인 시도가 필요하다. 구체적으로는, 향후 과업에서는 상식 그래프(knowledge graph)를 학습한 한국어 인공지능이 그렇지 못한 한국어 인공지능보다 더 뛰어난 성능을 보이는지 평가할 수 있다. 이 때, 본 사업이 구축한 확신성 언어자원을 활용하면, 한국어 문장구조에서 자연스럽게 나타나는 복잡한 추론 구조를 인공지능에 물어봄으로써 상식 그래프의 효과성을 입증할 수 있다. 본 사업의 확신성 언어자원은 부자연스럽게 생성된 인공주석물이 아닌 말뭉치의 자연스러운 발화와 한국어 화자의 직관을 계량화한 언어자원이므로 더욱 유용할 것이다.

붙임 1.

**추론_확신성 분석 말뭉치
2021 구축 지침서**

1. 개요

1.1. 본 과제의 목적

어떤 발화에서, 화자(필자)가 전달하는 메시지를 이해하려고 할 때, 우리는 화자가 확신하고 있는(또는 화자가 확신하고 있다고 알고자 하는) 내용이 무엇인지 알아야 한다. 언중간의 일반적인 화행과 관습적인 문법 형식에 따라 화자의 확신(speaker's commitment)은 사건의 내용을 객관적으로 진술하는 평서문의 형태로 서술될 때 가장 명확하게 판단할 수 있다(남기심·고영근 2014). 하지만 실제 언어 사용에서 화자의 생각은 서법(mood)의 범주 내에서 의문문, 명령문, 청유문 등으로 다양하게 표현될 수 있으며(임동훈 2008, 2011), 청자(독자)는 평서문으로 실현되지 않은 자리에서 명제를 인식해 낸 후, 인식한 명제에 대한 화자의 확신성(참/거짓의 정도)을 추론한다.

본 사업의 목적은 다양한 문법 범주의 내포문에서 인식되는 명제에 대한 화자의 확신성 정보를 부착한 한국어 말뭉치를 구축하는 것으로, 내포문에서 추출한 명제에 대해 화자가 확신하는 정도를 모문 및 주변 맥락과 관련지어 분석한다. 전통적인 의미 이론에서는 내포 명제의 (비)사실적 함의는 모절(matrix clause)의 술어에 의해 결정되는 것으로 알려져 왔다. 이 술어들은 '사실성 술어(factive predicate)'와 '비사실성 술어(non-factive predicate)'의 이분법 체계로 분류되었다(Kiparsky & Kiparsky 1971; Karttunen 1971, 2013). 예를 들어 (1)에서 know와 같은 동사는 대표적인 사실성 술어인데, 아래와 같이 화자가 내포절에 가지는 참의 진리값을 전제한다. 반면 (2)에서 비사실성 술어인 believe는 내포절의 참의 진리값을 전제하지 않는다.

- (1) \neg . Mary says: John knows that it is snowing. \Rightarrow It is snowing.
 \perp . know(p) \Rightarrow p
- (2) \neg . Mary says: John believes that it is snowing. \nRightarrow It is snowing.
 \perp . believe(p) \nRightarrow p

범언어적으로 이러한 (비)사실성 전제는 여러 언어에서 발견되며(Iatridou 2000; Horn 2014), 한국어에서도 내포절의 (비)사실성에 영향을 미치는 보문소와 보

문 술어에 대한 많은 연구가 진행되었다(강범모 1983; 이정민 2018, 2020; Han 1996, 2006; Lee 2019a, 2019b).

하지만 최근 코퍼스과 실험을 통한 연구에서 (비)사실성 술어의 구분은 전통적인 이분법으로 이루어지는 것이 아니라 척도 양상을 나타낸다고 주장되었다(Marneffe et al. 2019). 이에 따라 확신성 말뭉치에 기반한 화자의 확신성 실험을 통해 45개의 영어 (비)사실성 술어들의 문맥을 함의(entailment), 중립(neutral), 모순(contradiction)의 삼분적 구조로 구분하였다.

국립국어원에서 2020년에 수행한 ‘말뭉치 함의 분석 및 연구’(국립국어원, 2020) 사업(이하 1차 연구)은 Marneffe et al.(2019)의 연구에 따라, 함의 취소 운용소(entailment canceling operator)를 포함한 7개의 보문소와 27개의 술어로 이루어진 담화 구조를 분석하여 어떠한 문맥에서 주어진 언어 표현에 대한 화자의 확신성이 영향을 받는지를 분석하였다.

1차 연구의 결과물은 내포문의 사실성 연구를 수행할 때 유용한 자료이며, 인공지능 언어모델이 내포문과 모문 사이의 함의 관계를 추론할 수 있는지 평가하는 용도로 활용할 수 있다. 그러나 1차 연구에는 다음과 같은 한계가 있다. 우선, 제한된 수의 술어, 보문소, 함의 취소 운용소로 구성된 문맥만을 대상으로 하였기 때문에 내포문과 모문간 함의 관계에 대한 다양한 사례를 반영하지 못한다. 따라서 언어 모델이 이 결과물에 기반한 추론 과제를 모두 해결하였다고 하더라도, 내포문과 모문 간의 함의 관계를 잘 추론한다는 근거로는 충분하지 않다. 다음으로, 실제 언어 사용에서 화자의 확신성은 문장 이상의 범위를 고려하여야 파악할 수 있는 경우들도 있으나, 1차 연구는 내포문과 모문 간의 관계만을 고려하는 경향이 있었다.

따라서 ‘2021년 말뭉치 함의 분석 및 연구’는 1차 연구를 다음과 같은 방향으로 확장한다. 첫째, 1차 연구보다 다양한 유형(총 198개)의 보문소를 선정하여 다양한 한국어의 언어 사용을 반영한다. 보문소와 술어의 유형을 가급적 제한하지 않고, 내포절에 대한 화자의 확신성을 담은 용례를 최대한 다양하게 포착한다. 둘째, 본 과제에서는 함의취소운용소를 포함하지 않는 문맥도 자료에 포함한다. 1차 연구는 보문소, 술어, 함의취소운용소를 모두 포함한 문맥만을 추출하였다. 이는 특정 연구를 위해서는 바람직한 수집 방법일 수 있으나, 내포문에서 인식되는 명제의 확신성을 인공지능이 전반적으로 추론할 수 있는지 평가하

기에는 제한적이다. 또한 함의취소운용소를 포함하지 않는 문맥까지 자료에 포함하면 일반적으로 알려진 양태(modal) 표지 외에 화자의 확신성에 영향을 미치는 요소가 있는 경우 이를 자연스럽게 반영할 수 있다. 셋째, 본 연구는 1차 연구보다 넓은 맥락을 고려하여 내포문 명제의 화자 확신성을 추론한 정보를 부착함으로써 문장을 넘어서는 맥락이 내포문의 명제(가설)에 대한 확신성에 영향을 미치는 경우까지 반영한다.

1.2. 주요 용어

담화 구조 정보를 위한 주요 용어는 다음과 같다.

가. 대상 담화: 내포 명제에 대한 화자의 확신성을 파악하기 위한 연구 대상 문장과 선행 문장과 후행 문장을 포함한 전체 문장들.

(3) 이것은 무식한 계층에 속한 남녀노소들이 부른 노래로 선비들이 무가치한 것으로 취급했기 때문에 관심 밖에 있었던 것이다. 대단히 유감스러운 일이 아닐 수 없다. 이 책의 특징은 두 가지를 들 수 있다. 첫째, 오늘날 돌이켜보면 향요는 매우 소중한 노래임을 알 수 있다. 우리 민족문화의 편모라도 얻어보려고 한다면 우리 기층의 어른들과 어린이들이 불렀던 노래를 무시하고는 안 되기 때문이다. 이 노래는 아주 오래된 것으로 석대(昔代)의 사조(思潮), 언어학, 문학발달사 등을 살피는 데 있어서도 유용하기 때문에 그 가치를 돈으로 환산할 수 없다. 둘째, 이 노래는 가치가 높은 것으로 생각했기 때문에 남한과 북한 지역을 포함하여 전국 방방곡곡에 산재해 있는 자료를 수집하려고 노력했다.

나. 대상 문장(current): 확신성 판단의 대상이 되는 문장. 내포절과 보문소, 모절 술어로 구성된다. 아래는 (4)의 대상 담화에서 대상 문장을 추출한 예시이다.

(4) 첫째, 오늘날 돌이켜보면 향요는 매우 소중한 노래임을 알 수 있다.

다. 선행 문장(prev): 대상 문장 앞에 출현하는 문장. 문어는 3개, 구어는 5~10개로 제시된다. 아래는 (5)의 대상 담화에서 선행 문장을 추출한 예시이다.

(5) 이것은 무식한 계층에 속한 남녀노소들이 부른 노래로 선비들이 무가치한 것으로 취급했기 때문에 관심 밖에 있었던 것이다. 대단히 유감스러운 일이 아닐 수 없다. 이 책의 특징은 두 가지를 들 수 있다.

라. 후행 문장(next): 문장 뒤에 출현하는 문장. 문어는 3개, 구어는 5~10개로 제시된다. 아래는 (6)의 대상 담화에서 후행 문장을 추출한 예시이다.

(6) 우리 민족문화의 편모라도 얻어보려고 한다면 우리 기층의 어른들과 어린이들이 불렀던 노래를 무시하고는 안 되기 때문이다. 이 노래는 아주 오래된 것으로 석대(昔代)의 사조(思潮), 언어학, 문학발달사 등을 살피는 데 있어서도 유용하기 때문에 그 가치를 돈으로 환산할 수 없다. 둘째, 이 노래는 가치가 높은 것으로 생각했기 때문에 남한과 북한 지역을 포함하여 전국 방방곡곡에 산재해 있는 자료를 수집하려고 노력했다.

마. 실험 문항(context+target): 대상 문장과 선(후)행 문장의 결합. 아래는 (7)의 대상 담화에서 실험 문항을 추출한 예시이다.

(7) 이 책의 특징은 두 가지를 들 수 있다. 첫째, 오늘날 돌이켜보면 향요는 매우 소중한 노래임을 알 수 있다. 우리 민족문화의 편모라도 얻어보려고 한다면 우리 기층의 어른들과 어린이들이 불렀던 노래를 무시하고는 안 되기 때문이다.

바. 내포 명제(prop/prompt): 화자의 확신성에 대한 판단의 대상이 되는 명제. 대상 문장에 보문으로 쓰인 내포절을 하나의 명제 형식(단문)으로 바꾼 문장을 말한다. 아래는 (8)의 대상 문장에서 내포 명제를 추출한 예시이다.

(8) 향요는 매우 소중한 노래이다

사. 보문소(comp): 내포절이 전체 문장(모문)의 보문이 되도록 만드는 요소. 본 과제에서는 사전 작업을 기반으로 함의/중립/모순 구조를 구성하는 보문소와 유사보문소의 종류와 표기 형태를 198가지로 한정하여 사용한다(구체적인 용례는 부록을 참고). 아래는 (9)의 대상 문장에서 보문소를 추출한 예시이다.

(9) 보문소: ‘마을’

아. **모절 술어(pred)**: 모절의 술어. 내포명제에 대한 화자의 확신성을 판단하는 데 있어서 보문소와 더불어 중요한 요소로 작용한다. 아래는 (10)의 대상 문장에서 모문 술어의 예시이다.

(10) 술어 ‘알(다)’

- ㄱ. 대상 문장: 첫째, 오늘날 돌이켜보면 향요는 매우 소중한 노래임을 알 수 있다.
- ㄴ. 내포 명제: 향요는 매우 소중한 노래이다.

자. **함의 취소 운용소(entailment cancelling operator: ECO)**: 내포 명제에 대한 함의 투사를 증명할 수 있는 언어적 기제이다. 본 연구 과제에서 대상으로 삼은 함의 취소 운용소는 ‘부정, 의문, 조건, 양태’ 네 가지로 유형화하였다.

[그림 1] 함의 취소 운용소의 예

부정	—————	못, 안/아니, 없, -지 말/아니하/않-, -지(는) 못하-
의문	—————	-까, -ㄴ가, -ㄴ데-, -나, -냐, -잖아, -지 등(의문형 종결어미)
조건	—————	-(으)면(야), -(ㄴ)다/라면, 어/아도
양태	인식 양태	-(-으)ㄴ가 싶-, -(-으)ㄴ/(-으)ㄴ 것 같-, -(-으)ㄴ 것 이-, -(-으)ㄴ/(-으)ㄴ 듯하 듯하-, -(-으)ㄴ 수 있/없-, -(-으)ㄴ 텐데, -(-으)ㄴ지 모르-, -겠(추측)-, -(-으)ㄴ 모양이-, -나 보-, -어/아 보이-
	비인식 양태	의무: -(-으)ㄴ 필요가 있-, -(-으)면 되-, -어야 되-, -어야 하- 능력(가능/불가능): -(-으)ㄴ 만하-, -(-으)ㄴ 수 있/없-, -(-으)ㄴ 줄 알/모르- 의지: -겠-, -고 싶-, -고자 하-, -(-으)려 하-

함의 취소 운용소의 구체적인 예는 아래와 같다.

(11) 부정

- ㄱ. 대상 문장: 그는 자신의 부모나 친구, 심지어는 아내까지도 배역을 수행하는 배우라

는 사실을 전혀 눈치채지 못한 채 30여 년을 살아온 것이다.

ㄴ. 내포 명제: 트루먼쇼 주인공의 부모나 친구, 심지어는 아내까지도 배역을 수행하는 배우였다.

ㄷ. 모문 술어: 눈치채다

ㄹ. ECO: -지 못하다

(12) 의문

ㄱ. 대상 문장: 우리 인생이 기차 같으면 얼마나 좋겠냐?

ㄴ. 내포 명제: 우리 인생이 기차 같다.

ㄷ. 보문소: 으면

ㄹ. 모문 술어: 좋다

ㅁ. ECO: 냐

(13) 조건

ㄱ. 대상 문장: 집에서 가족들이 걱정하며 기다릴 것을 생각하면 단 하루도 마음이 편할 리 없습니다.

ㄴ. 내포 명제: 집에서 가족들이 걱정하며 기다릴 것이다.

ㄷ. 생각하다

ㄹ. ECO: -면

(14) 양태(인식)

ㄱ. 대상 문장: 권력은 무조건 나쁘다는 생각은 좀 편협한 것 같아.

ㄴ. 내포 명제: 권력은 무조건 나쁘다.

ㄷ. ECO: -(으)ㄴ 것 같-

(15) 양태(비인식/의무)

ㄱ. 대상 문장: 복합기, 프린터, 복사기, 스캐너 등 사무실 필수 사무기기도 잘 돌아가고 있는지 살펴봐야 한다.

ㄴ. 내포 명제: 복합기, 프린터, 복사기, 스캐너 등 사무실 필수 사무기기도 잘 돌아가고 있다.

ㄷ. ECO: -어야 하-

(16) 양태(비인식/능력)

ㄱ. 대상 문장: 데이브가 창문으로 빠져나와 기관차 위로 올라갔을 때 그는 자신이 다치지 않은 것은 물론 침대 칸도 전복되지 않았고, 단 한 명의 승객도 상처 입지 않은 것을 알 수 있었습니다.

ㄴ. 내포 명제: 단 한 명의 승객도 상처 입지 않았다.

ㄷ. ECO: -(으)ㄴ 수 있-

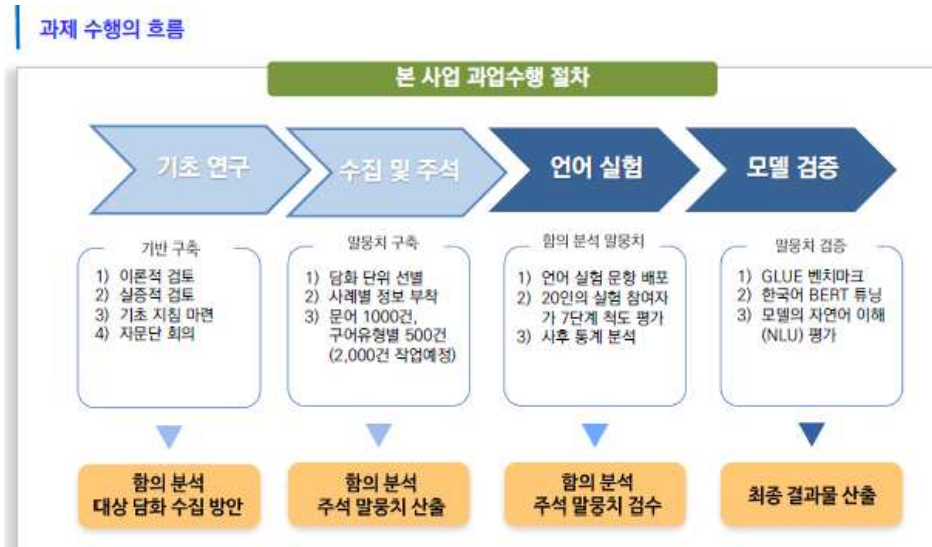
(17) 양태(비인식/의지)

- ㄱ. 대상 문장: 굳이 구별해야 한다면 여행에 대한 즉흥 이야기라고 말하고 싶다.
- ㄴ. 내포 명제: 그것은 여행에 대한 즉흥 이야기이다.
- ㄷ. ECO: -고 싶-

2. 작업 절차

2.1 구축 절차

[그림 2] 과업수행 절차



2.2 탐색 대상 및 단계

■ 자료 추출 개요

- 대상이 되는 내포문이 포함된 문장 추출을 위해, 보문소를 기준으로 문장 추출을 진행하였다.
- 추출은 문어, 신문, 구어를 장르에 따라 균등한 개수를 추출하였다(문어 및 신문 1,000개, 구어 500개).
- 문어와 신문은 앞뒤 3문장 문맥을 함께 수집하고 구어는 10문장 문맥을 수집하

였다.

■ 자료 추출 방법

- 국립국어원 구문분석 말뭉치, 세종 형태분석 구어 말뭉치를 바탕으로 보문소를 선정하였다.
- 선행 연구를 통해 유사 보문소 목록을 선정하였다((유사)보문소 전체 목록은 부록 참조).
- 정확한 추출을 위해, 자모 분리와 형태소 분석을 적용하여 추출하였다.
- 편향된 분포를 피하기 위해 랜덤하게 추출하였다.

■ 자료 추출 결과

보문소 목록, 추출 개수, 장르를 나누어서 예비 추출 3회, 본 추출 8회 진행하였다.

- 예비 추출: 작업별 방향과 기계적 추출의 3단계로 진행
 - Precommit1(10,410개): 주석작업자의 작업 방향의 검토를 위해 별도의 자료 추출 기준을 세우지 않고 작업하였다.
 - Precommit2(10,041개): 데이터의 구성은 이전과 같으나, 확산성 실험을 운용하는 연구원과 연구보조원 관점에서, 그리고 모델을 이용한 자연어 이해 평가를 진행하는 보조연구원의 시각에서 데이터의 적합성을 검수하였다.
 - Precommit3(10,000개): 이전 두 단계의 의견을 종합하여 ‘보문소’를 기준으로 다시 추출한 것으로, 데이터의 배치와 구성을 달리하였다.
- 본 추출
주석 작업 및 실험과 병진적으로 진행하며, 각 단계의 문제점을 보완하여 진행하였다.

Commit1-3: 장르에서 특정 문형의 불균형 확인, 보문소 기준 추출이 적합함을 확인하였다.

- Commit1(51,882개): 문어 장르에서 추출하였다.
- Commit2(10,000개): “문어” 장르의 불균형을 해소하기 위해 분포를 다양하게 하여 추출하였다.
- Commit3(10,000개): “신문” 장르에서 추출. “-다고” 문형 빈출하였다.

Commit4-7: 추출 대상을 다양화, 함의/중립/모순 균형을 목적으로 추출하였다.

- Commit4(803개): 문어 장르에서 추출하였다.
- Commit5(1,114개): commit4와 동일한 목적에서 수행되었으며, 101-105의 회차별 실험에 적합하게 세분화하였다.
- Commit6(811개): 모순 관계의 충원을 위해 별도 추출하였다.
- Commit7(945개): Commit6을 보완하기 위해 가정 데이터를 찾아서 별도로 추출하였다.

■ 수정 및 보완

- 추출 작업과 주석 및 실험 작업을 병진적으로 진행함으로써, 주석/실험 작업에서 보고되는 문제점을 추출 작업으로부터 보완 가능하였다.

[그림 3] 작업자 워크벤치

The screenshot shows a web-based interface for text analysis. At the top, there's a header with a toggle for 'Show/Hide the Previous Context' and a text ID: [WARW190000006] 101/469 [101119]. (문어). Below this is a text snippet: '그러면서 다 읽고 나면 인생은 허무하다고 떠올려댄다.' A second toggle for 'Show/Hide the Next Context' is visible. The main area contains several sections: '보문소:' and '국어원 보문소:' with input fields containing '다고'; '보문술어:' and '국어원 보문술어:' with input fields containing '떠들어대다' and '떠들어댄다→떠들어대다'; 'Context + Target' with a description: '그러나 모두 하루 만에 백만장자가 되고 하루 만에 일거지가 되려고 안달한다. 그러면서 다 읽고 나면 인생은 허무하다고 떠올려댄다.'; '내포명제:' with the text '인생은 허무하다'; '내포절:' with two columns of radio button options for '모절' and '모절:' including choices like '과거', '현재', '미래', '은/는', '었/았', '으_것/겠', '없음', and '일방적인 사실 기술'; 'ECO:' with a '표지:' field and radio button options for '양태', '부정', '의문', '조건', '그 외', and '없음'; '기타:' with checkboxes for '민감 사안' and '확인 요망'; '국어원에 대한 답변:' with a text input field; and '메모:' with another text input field. At the bottom, there are buttons for '저장', '이전 문장', '다음 문장', and a '로 이동' dropdown menu.

■ 민감 사항

- 성(gender), 정치, 종교와 관련한 민감 사례를 담았거나, 실명이 언급되어 논란의 여지가 있을 수 있는 담화들은 분석에서 제외하였다.

3. 주석 지침

3.1 문어 지침

내포 명제 기술 시 다음의 원칙들을 따른다.

1) 오류

■ 맞춤법 오류, 띄어쓰기, 기호 사용 오류 등은 임의로 수정하거나 배제하지 않는다. 자연어 텍스트 환경에서도 오타가 존재하기 때문이다.

(18) 실험 문항: 하지만 A조합원장은 부속토를 나눠주면서 전체 조합원들을 대상으로 공지하지 않아 특정 조합원에게 특혜를 줬다 의혹을 하고 있다.

내포 명제: 부속토를 나눠주면서 특정 조합원에게 특혜를 줬다.

(19) 실험 문항: 일본 도쿄는 국내 유명 최고경영인들이 자료수집이나 휴가차 방문하는 도시이다. 특히 삼성그룹의 이건희 회장은 매년 시간만 나면 일본으로 새로운 사업을 구상하러 간다고 한다.

내포 명제: 삼성그룹의 이건희 회장은 매년 시간만 나면 일본으로 새로운 사업을 구상하러 간다.

■ 단, 아래의 사례처럼 실험 문항에 포함될 수 있는 내포 명제에 오타가 발생하는 경우는 오타를 수정한다.

(20) 실험 문항: 100여편의 작품 중 다수의 우수작품들이 독특한 심상을 가지고 있어 우열을 가리기 어렸다고 전하면서 최종적으로 김용옥 수필가 겸 시인 작품 ‘지혜의 발견’수필을 선정했다.

내포 명제: 100여편의 작품 중 다수의 우수작품들이 독특한 심상을 가지고 있어 우열을 가리기 어렵다.

2) 복원

■ 내포 명제는 다음과 같은 방식으로 주어와 술어 및 다른 문장 성분을 가급적 그대로 기술하는 것을 원칙으로 한다.

(21) ㄱ. 실험 문항: 성조와 발음이 아무리 중요하다고 해도 ‘언어는 소통을 위한 도구’임을 기억하면 문제는 의외로 쉽게 풀린다.

내포 명제: 언어는 소통을 위한 도구이다.

ㄴ. 실험 문항: 연초에는 지금과 정반대의 상황이 벌어졌던 것을 기억할 것이다.

내포 명제: 연초에는 지금과 정반대의 상황이 벌어졌다.

■ 대상 문장의 내포절에 주어와 술어 명제 내용을 파악하기 어려운 경우에는 문맥상 예측할 수 있는 선행 주어를 복원하여 제시한다.

(22) ㄱ. 실험 문항: 해고를 당한 노동자들은 "아사히글라스에서 일하는 9년 동안 최저임금 수준만 받으면서 기계처럼 일했다. 참다 못해 노조를 만들었지만, 한 달 만에 직장을 잃었다"고 했다.

내포 명제: 노동자들이 참다 못해 노조를 만들었지만, 한 달 만에 직장을 잃었다

ㄴ. 실험 문항: 이날 박종진은 수능을 마친 큰 딸 박인 양으로 소개하며 "내 진짜 여자친구"라고 말해 눈길을 끌었다. 그는 "딸이 19살까지는 부드러웠는데 고3이라 공부하면서 좀 까칠해졌다"고 덧붙였다.

내포 명제: 박종진의 딸이 고3이라 공부하면서 좀 까칠해졌다

■ 대상 문장의 내포절과 선행 문장에 주어와 술어 명제 내용을 파악하기 어려운 경우에는 문맥상 예측할 수 있는 후행 주어를 복원하여 제시한다. 내포 명제는 문맥에 따라 주어와 함께 ‘은/는’ 이나 ‘이/가’를 복원하여 제시한다. 조사를 복원할 때에는 해당 문장만으로 의미를 전달하는 것이 가장 자연스럽게 자주 쓰일 법한 것으로 복원하도록 한다.

(23) ㄱ. 실험 문항: 그는 "선발에 대한 욕심은 있긴 하다. 그러나 팀이 필요한 게 먼저다. 팀에 도움이 돼야 선수의 가치가 있다"며 "써주는 대로 열심히 할 생각이다"고 했다. 박준표가 올해 마운드에서 제 몫을 해주길 바라는 팬들의 기대에는 자신감으로 응답했다

내포명제: 박준표는 팀이 써주는 대로 열심히 할 것이다

- ㄴ. 실험 문항: 국립국제교육원에서 다도(茶道)를 배우고, 거문고·가야금과 같은 전통악기를 보는 등 새로운 문화를 배우면서 한국에 매료됐다. 그는 "2주간 한국을 다니면서 언젠가 조국에 정착해 살아야겠다고 생각했다"며 "혼자 우즈베키스탄과 한국을 오가다가 아내, 자녀와 함께 지난해 8월 정착했다"고 했다. 차 지부장은 고려인들의 잃어버린 문화와 언어를 되찾아 주는 것이 중요하다고 강조했다.

내포 명제: 차 지부장은 혼자 우즈베키스탄과 한국을 오가다가 아내, 자녀와 함께 지난해 8월 정착했다

■ 대상 문장의 내포절에 목적어가 없어 명제 내용을 파악하기 어려운 경우에는 문맥상 예측할 수 있는 선행 목적어를 복원하여 제시한다.

- (24) 실험 문항: 박 국장님, 이번 총회 결과에 대해 국내의 분위기가 너무 안 좋습니다. 언론과 국회에서 정부를 비난하는 목소리가 엄청납니다. 대통령까지 책임져야 한다는 주장도 있습니다.

내포 명제: 대통령까지 "총회결과에 대해" 책임져야한다.

■ 보다 명확한 내포 명제 구성을 위해 부사어나 수식어 등의 보조적인 정보를 복원하여 제시한다.

- (25) ㄱ. 실험 문항: 이제는 신용 우량자가 대우받는 시대가 됩니다. 개인의 신용관리가 매우 중요하다고 볼 수 있습니다. 체계적인 자금계획과 재테크 플랜을 통해 철저하게 개인의 신용관리를 하면서, 풍성해진 재테크 상품들을 통해 재산을 증식해 나가야 합니다.

내포 명제: 신용 우량자가 대우받는 시대에는 개인의 신용관리가 매우 중요하다.

- ㄴ. 실험 문항: 광주 광덕중·고등학교 오래시간 불러온 교가 교체작업에 나섰다. 이는 최근 친일 음악가 김성태에 의해 교가가 작곡됐다는 사실을 알았기 때문이다. 충격을 받은 신홍수 광주 광덕중·고등학교 이사장이 전면에 나서서 교가 교체를 표명하면서 작업 속도도 빨라지고 있다.

내포 명제: 친일 음악가 김성태에 의해 광주 광덕중·고등학교 교가가 작곡됐다.

㉔. 실험 문항: 도대체 어떤 맛인 지 양배추즙을 시음해 보고 싶다 했더니 한 여
직원이 급히 건네준다. 마셔보니 정말 첨가물 한방울 없는 원물 그대로다.
'보약이려니' 하고 마시는 중장년층이 주고객이 아닐까 싶었지만 의외로 최대
고객은 젊은 여성층이다. 다이어트, 변비 예방에 즉효약이라는 입소문 때문이
라 한다.

내포 명제: 중장년층이 양배추즙의 주고객이다.

㉕. 실험 문항: 이해관계를 떠나 내가 만난 사람의 행복을 진지하게 빌어주는 게
중요해요. 그런 선의가 출발점이죠. 인내심을 갖고 선의를 쌓아야죠. 선의를
가진 사람이 많다는 생각이 들도록요. 선의는 인프라이입니다. 쌓아가야죠.

내포 명제: 내가 만난 사람의 행복을 진지하게 빌어주는 선의를 가진 사람이
많다

3) 인칭 대명사

■ 내포절에 제시된 인칭 대명사의 선행어를 대상 담화에서 찾을 수 있는 경우에는 선
행어를 복원한다.

(26) ㉖. 실험 문항: 한번은 주인 여자가 혹시 무슨 일이 생긴 건 아닌가, 걱정이 되어 방문을
열어 본 적도 있었다. 그 이후로도 주인 여자는 가끔씩 뒤란으로 돌아와 인기척을 냈
는데, 그것은 그에게 사람 소리를 내고 살라는 힌트 같은 것이었다. 그녀는 자기 집
에서 끔찍한 일이 일어나는 사태를 몹시 두려워했음이 분명하다. 그는 거의 아무것도
하지 않고 지냈다.

내포 명제: 주인 여자는 자기 집에서 끔찍한 일이 일어나는 사태를 몹시 두려워했다.

㉗. 실험 문항: 무실[水谷]이란 곳으로 장가를 들어 무실 아재라 불리우는 사람인데, 우리
보다는 십년 이상 연상이고 항렬도 높았다. 따라서 우리와 쉽게 어울릴 수 있는 사람
이 아니었으나, 우연히 우리 자리에 끼여들어 객원을 자처하더니 나중에는 좌장(座
長)격이 되어 흉허물없이 지냈다. 하지만 어떻게 보면 그는 처음부터 우리와 함께할
수 있는 사람이었는지도 모르겠다. 무실 아재가 가까운 도시의 농업고등학교를 졸업
하고 고향으로 돌아온 것은 그로부터 이십 년 전이었다.

내포 명제: 무실 아재는 처음부터 우리와 함께할 수 있는 사람이었다.

■ '자기', '자신' 등의 인칭 대명사의 선행어는 대상 담화에서 찾을 수 있는 경우에는
선행어를 복원한다.

- (27) 실험 문항: 「에, 재미있던 일이라……」 얼른 자신의 추태를 추스르고 지금까지 그 <재미 있던 일>을 생각하고 있던 것처럼 얼버무리면서, 이번에는 정말로 <재미있던 일>을 생각해 보기 시작했다. 그리고 미래야는 오래지 않아 그것을 찾아낼 수 있었다.
 내포 명제: 미래야는 지금까지 그 <재미있던 일>을 생각하고 있었다.

■ ‘우리’나 ‘나’ ‘너’ ‘그’는 대상 담화 내에서 복원되지 않고 확신성 판단에 무리가 없다면 그대로 사용한다.

- (28) 가. 실험 문항: 피곤에 찌들어 전철을 타는 이들의 얼굴에는 하루를 시작하는 기쁨도 생기도 찾아볼 수 없습니다. 하지만 바쁜 일정에서 벗어나 조금만 더 자신에 대해 깊이 생각해 보면, 분명히 우리에게는 원대한 꿈이 있었다는 걸 기억하게 됩니다.
 내포 명제: 우리에게는 원대한 꿈이 있었다.
- 나. 실험 문항: 그때 나에겐 오직 준희뿐이었습니다. 친구도 별로 없었어요. 준희는 거의 나를 파먹고 살았던 것 같습니다. 배가 고파 보였습니다. 외로워 보였습니다.
 내포 명제: 준희는 거의 나를 파먹고 살았다
- 다. 실험 문항: 기쁠 때나 슬플 때나 안타까울 때나 괴로울 때나 나와 매표소가 늘 열려 있기 때문인가 봐요. 난 정말 웃기도 잘하고 울기도 잘해요. 남자답지 못하다고 편지를 주셔도 할 수 없어요.
 내포 명제: 난 남자답지 못하다.

■ 문어와 신문 말뭉치에서 선정한 담화의 경우, 2인칭 대명사 ‘당신’ 등이 지시하는 대상은 대부분 글을 읽는 독자들이다. 따라서 괄호 속에 ‘청자’ 등을 병기하지 않는 것을 원칙으로 한다.

- (29) 가. 실험 문항: 바람이 차다. 그러나 당신은 추운 줄도 모른다. 옥상의 한가운데 서서 두 팔을 짝 벌리고 달빛을 한껏 받아들인다.
 내포 명제: 당신은 춥다.
- 나. 실험 문항: 스스로에게 진지하게 물어봐야 한다. 지금 당신의 나이가 몇 살이든, 처한 환경이 어떠한든 중요한 것은 지금도 늦지 않았다는 사실이다. 늦었다고 생각할 때가 가장 빠르다고 하지 않던가.
 내포 명제: 당신은 지금도 늦지 않았다.

■ 의인화된 문장의 경우, 그 자체로 자연스러운 한국어 문장인지 검토한다.

그렇지 않은 경우 제외한다.

(30) ㄱ. 실험 문항: 무슨 일이든 뜻대로 할 수 있단다. 이번엔 내가 너를 도와주마. 심해어는 몸이 두둥실 떠오르는 것을 느꼈어요.

내포 명제: 심해어는 몸이 두둥실 떠올랐다. (제외)

ㄴ. 실험 문항: “니야옹!”(잘 먹었습니다!) 접시를 깨끗하게 비운 지스는 작은 목소리로 가릉거처음에는 이런 예의를 차린 적이 별로 없었지만, 문득 생각해 보니 고양이치고 너무 호사한다는 생각에 양심이 찢려 언제부턴가 자신도 모르게 하게 된 인사였다.

내포 명제: 지스는 고양이치고 너무 호사한다. (제외)

4) 대용어

■ 대동사 표현 ‘그럴’, ‘그러한’ 또한 복원한다.

(31) 실험 문항: 그리고 형은 어땠는지 모르지만, 나는 그 과제가 곤혹스러워지기 시작했다. 할 수만 있다면 어머니의 진술을 그만 멈추게 하고 싶었다. 그러나 “그럴” 권리가 나에게 없다는 사실을 나는 이해했다.

내포 명제: 어머니의 진술을 멈추게 할 권리가 나에게 없다.

■ ‘둘은’, ‘두권’, ‘양자가’ ‘서로’ 등 대용어 표현은 선후행 문맥을 통해 복원한다.

(32) ㄱ. 실험 문항: 신학과 정신분석의 상호 침투는 침범이 아니다. 학문적이고 임상적인 대화이다. 서로 경계하면 서로 적이 될 수밖에 없다. 방어 본능을 해제하면, 둘은 상호 고유한 영역이 있으면서도 별개가 아님을 발견할 수 있다.

내포 명제: 방어 본능을 해제하면, 신학과 정신분석은 별개가 아니다.

ㄴ. 실험 문항: 내가 배울 점이 더 많을 것 같다. 서로 잘하면 시너지 효과를 내지 않을까 싶다. 그 선수에게 배울 점이 있으면 빨리 배우고 빨리 적응할 수 있게 도울 생각이다.

내포 명제: 나와 그 선수가 서로 잘하면 시너지 효과를 낼 것이다.

■ ‘이’, ‘이런’과 ‘이러한’은 선행 문맥에서 가장 가까운 지시적 표현으로 복원한다.

(33) ㄱ. 실험 문항: 유 열사의 생가는 용두리 지령부락 입구에 복원·조성돼 있다. 그런데 이 복원된 생가는 실제 유관순 열사의 생가터가 아니라고 마을주민들은 기억하고 있다.

내포 명제: 용두리 지령부락 입구에 복원된 생가는 실제 유관순 열사의 생가터가 아니다.

ㄴ. 실험 문항: 양이원영 사무처장은 “현재 석탄발전소 가동률이 80%, LNG발전소 가동률이 45%인데 둘의 가동률을 바꾸면 미세먼지가 크게 줄어들 수 있다. 이 경우 가구당 평균 2000~3000원의 전기요금이 오를 것”이라고 전했다.

내포 명제: 석탄발전소와 LNG발전소의 가동률을 바꾸면 가구당 평균 2000~3000원의 전기요금이 오를 것이다.

ㄷ. 실험 문항: 고(故) 김수환 추기경의 담당기자로서 가장 가까운 곳에서 김 추기경이 뿜어내는 경이로운 인간적 향기를 체험한 저는 프란치스코 교황의 여러 면모를 알게 되면서 당시와 비슷한 감동을 느꼈습니다. 알면 알수록 정말 이런 분과 동시대인으로 같은 대지를 딛고 살아가고 있다는 사실에 감사하게 되었습니다.

내포 명제: 김 추기경과 동시대인으로 같은 대지를 딛고 살아가고 있다.

ㄹ. 실험 문항: 평면의 종이 위에서 곧 튀어나올 듯 입체적인 표현 기법도 신기하다. 사물을 채우고 있는 색깔은 볼수록 감동을 준다. 달, 석탑 등 실물에 가까운 색깔은 물감으로 이러한 표현이 가능할까 싶을 정도다. “이름있는 작가라고 해서 작품이 모두 좋은 것만은 아닌것 같아요. 오히려 이름이 알려진 작가의 작품일수록 냉철한 시선을 가져야하죠.”

내포 명제: 달, 석탑 등 실물에 가까운 색깔도 물감으로 입체적인 표현이 가능하다.

■ 가까운 선행 맥락에 의거하여 내포 명제를 관습적인 상황(habitual situation)에 맞게 ‘재편’할 수 있다.

(34) 대상 문장: 문간에서 대기하던 왕의 검식관(檢食官)이 먼저 한 입 먹고 잠시 기다려서 독이 없음을 증명한 다음에야 왕 앞에 전달되었다. 왕은 보통 오후 2시에 정찬을 먹었는데 이는 공식적인 공개행사로 많은 사람들이 몰려들어서 왕이 식사하는 것을 관람하였다. 루이 14세는 경이로운 대식가였다.

내포 명제: 왕이 보통 오후 2시에 식사한다.

5) 정보구조

■ 구조적으로 내포명제가 명확하지 않더라도, 선후행 문맥을 통해 의미적으로 내포명제를 보강할 수 있으면 사용한다.

(35) ㄱ. 실험 문항: ‘간헐적 단식 2부-단식 모방 다이어트의 비밀’ 편. 굶어야 단식이 라는 편견을 깬 단식법이 등장했다. 먹는데도 공복 효과를 낸다고 한다. 미국 대학 발터 롱고 박사가 장수와 질병 예방법으로 단식을 추천했고, 안전하게 단식하는 방법으로 단식 모방 다이어트(FMD) 식단을 개발했다.

내포 명제: 간헐적 단식은 먹는데도 공복 효과를 낸다.

ㄴ. 실험 문항: 원가드 3는 충격감지 강도를 고객들의 기호에 맞게 설정할 수 있다는 게 특징이다. 가가호호 주거환경에 최적화시켜 기능을 사용할 수 있다는 것이다. 또 원가드 클라우드 서버를 통해 보다 안정적인 서비스 제공이 가능해졌다

내포 명제: 원가드 3는 가가호호 주거환경에 최적화시켜 기능을 사용할 수 있다.

ㄷ. 실험 문항: 광주 동구 동명동 아이플렉스 206호 그의 사무실을 찾아갔다. 풍기는 분위기는 마치 뷰티·미용 전문가 아닐까 싶었는데 얘기를 나눠보니 IT 분야의 해박한 지식을 갖추고 있었다. ◆불편하지만 돈을 쓸수밖에 없는 아이템없을까 "이 분야 창업을 고민하면서 있으면 좋은 거 말고, 불편하지만 돈을 쓸 수밖에 없는 아이템을 만들어보자." 엄 대표는 런치팩 사업을 구상하면서 세운 목표다.

내포 명제: 엄 대표는 뷰티·미용 전문가이다

■ 내포절의 명제가 복잡할 경우 단순화할 수 있도록 한다.

(36) ㄱ. 실험 문항: 따라서 부부갈등이 심각하지 않은 부부의 경우에도 부부관계를 증진시킴으로써 결혼 만족도를 높이고자 하는 욕구가 높아지고 있다(임선영, 2006). 만족스러운 결혼생활을 예측할 수 있는 객관적인 지표(指標)는 찾기 힘들지만 여러 가지 특성들이 관련되어 있는 것은 분명하다(권명득 외, 1992).

내포 명제: 결혼 만족도에는 여러 가지 특성들이 관련되어 있다.

ㄴ. 실험 문항: 이런 '간판 위주'의 교육이 갖는 폐해를 통계청의 경제 활동 인구

조사 결과는 흥미롭게 보여준다. 조사 응답자 전체의 40.5%가 전공과 불일치하는 직업을 가지고 있다고 응답했으며, 16.8%가 상당히 불일치하는 직업을 가지고 있다고 대답했다. 전공과 매우 일치하는 직업을 가졌다고 답한 사람은 겨우 16.9%에 불과했다.

내포 명제: 조사 응답자의 16.8%는 전공과 상당히 불일치하는 직업을 가지고 있다.

■ 두 개의 내포 명제가 있는 경우 첫 번째 명제를 대상 내포 명제로 삼는다.

(37) 실험 문항: 정신이란 말을 통해서 볼 때 신을 정신적인 현상으로 인식하고 있음을 알겠다. 결국 인간은 죽은 사람이 꿈에 보이기 때문에 사후세계가 있다고 믿게 되었을 것이며, 내세관이 싹튼 것도 꿈이 출발점이며 종교의 출발점이라고도 보여진다.

내포 명제: 내세관이 싹튼 것은 꿈이 출발점이다

■ 어순은 정보적 구조에 따라 작업자의 판단에 따라 재배치 가능하다.

(38) 실험 문항: 또 남북 간에 문화·역사 분야의 교류 사업을 우선적으로 추진하고, 남포·해주 등과 지방정부 간 협력 사업도 추진하겠다. 시민들이 평화를 통해 삶이 실질적으로 개선되는 것을 체감하실 수 있도록 '평화지수'도 적극적으로 개발·활용할 계획이다.

내포 명제: 평화를 통해 시민들의 삶이 실질적으로 개선될 것이다.

■ 지시 대상의 정보는 문맥을 통해 최대한 명세화(specific) 하도록 한다.

(39) 실험 문항: 범인 A씨는 해가 바뀐 이날 오전 0시 10분, 도쿄 시부야구 다케시타 거리에서 차량으로 행인들이 움집한 곳에 돌진했다. 그는 연말연시를 맞아 차량 통행이 금지됐던 메이지진구 인근의 도로에서 행인 8명을 들이받았고, 살인미수 혐의로 즉각 체포됐다. A씨는 체포 직후 자신의 행동을 '테러'라며 규정했고, "사형에 대한 보복"이라고 옴진리교 사형수들의 형 집행에 불만이 있음을 내비쳤다. 앞서 일본 정부는 지난해 7월 도쿄 지하철역 사린가스 테러사건 등과 관련해 교주 아사하라 쇼코 등 옴진리교 관계자 13명의 사형을 집행한 바 있다.

내포 명제: 범인 A씨는 옴진리교 사형수들의 형 집행에 불만이 있었다.

6) 관계절

■ 보문절이 아닌 부가절(adjunct)은 사용하지 말아야 한다. 따라서 실험에서 제외한다.

(40) 실험 문항: 콘서트 관람을 마치고 서연이 불쑥 래프팅 제안을 했는데, 수철이 맞장구를 쳤다.

대상 문장: 예준 역시 래프팅을 한 번도 해본 적이 없던 터라 흔쾌히 '오케이 사인'을 보냈다.

내포 명제: 예준은 래프팅을 한 번도 해본 적이 없었다. (제외)

■ 의사관계절(pseudo relative clause)는 실험에서 제외한다.

(41) 실험 문항: 그녀의 머리는 헝클어져 있었고 드문드문 새로 자라난 흰머리가 보였다. 희는 이불을 개어 안방에 가져다 놓고 나와 씩크대에 가래침을 뱉었다.

대상 문장: 그가 몸을 움직일 때마다 온몸의 뼈가 무너져내리는 것 같은 소리 pseudoRC가 났다.

내포 명제: 온몸의 뼈가 무너져내린다. (제외)

7) 의문사

■ wh-question은 제외한다.

(42) 실험 문항: 울고 있을지도 모른다. 다케유키는 마음이 아팠다. 하지만 도와줄 방법이 없었다.

대상 문장: 상처 난 마음을 어떻게 다독겨려야 하는지 그는 알 수 없었다.

내포 명제: 상처 난 마음을 어떻게 다독겨려야 했다. (제외)

■ 영어와 달리 한국어의 wh-word와 indefiniteness 사이에 혼동이 오는 경우는 단순화하는 것이 아닌 제외한다.

(43) 실험 문항: 여기서 우리는 다시 '사형취상'의 방법을 적용해야 한다. 무엇이 동작이고, 그 동작에서 아들이 아버지보다 나은 어떤 행동을 하였는가를 확인해야 하

는 것이다.

내포 명제: "아들이 아버지보다 나은어떤" 행동을 하였다" (제외)

■ wh-word와 indefinites 사이를 판별할 수 있는 경우 indefinites인 경우 내포 명제의 대상으로 삼는다.

(44) 실험 문항: 노란색 번호판을 단 영업용 트럭이었다. 비록 하잘것없는 운전 기술이었지만, 아버지가 무엇이랄도 '기술'을 가지고 있다는 사실은 자식들에게 이를 테면 기적과도 같은 놀라운 일이었다.

내포 명제: 아버지가 "무엇이랄도" 기술을 가지고 있었다.

8) 양태

■ 내포절 안에 있는 양태 표지는 그대로 유지하는 것을 원칙으로 한다.

(45) ㄱ. 실험 문항: 부상에 시달리며 별다른 활약을 펼치지 못했던 양학선은 수원시와의 재계약 이후 인천 국가대표 선수촌에 들어가 컨디션 회복에 집중하고 있다. 양학선이 올 시즌을 발판으로 삼아 수원시의 숙원인 올림픽 첫 메달 획득을 이뤄낼 수 있을지 주목된다.

내포 명제: 양학선이 수원시의 숙원인 올림픽 첫 메달 획득을 이뤄낼 수 있을 것이다 [비인식-능력(가능)]

ㄴ. 실험 문항: 실제 마늘에 관한 자료를 찾고 정리하는 과정에서 우리 주변에 늘 함께 하는 것이 소중한 것이라는 것을 느꼈고 아주 가까운데 보물섬이 있다는 것이 더욱 놀라왔다. 조물주가 마늘 속에 모든 비밀을 감추어 둔 것 같은 느낌을 받았다. 인간을 사랑하는 조물주가 정성을 다하여 준비한 식물중의 하나처럼 생각된다.

내포 명제: 조물주가 마늘 속에 모든 비밀을 감추어 둔 것 같다. [인식/추측]

ㄷ. 실험 문항: 특히, 노인과 청년층이 쉽게 사먹을 수 있는 음식 값이 가파르게 올라 서민들의 가계경제에 적지 않은 부담을 안겼다는 분석이다. 문제는 이미 음식 값이 오를 만큼 오른 상황에서 최저임금 인상 등을 빌미로 외식업계가 추가 가격인상에 나설 수 있다는 것이다. 6일 호남지방통계청에 따르면 지난 한해 광주지역 소비자물가지수는 전년보다 1.2% 상승했다.

내포 명제: 외식업계가 추가 가격인상에 나설 수 있다 [비인식/능력(가능)]

ㄹ. 실험 문항: 강호축에 4차산업혁명 과학기술을 도입해 경박단소(輕薄短小)형의

첨단산업을 유치하고 백두대간 국민쉼터를 조성하는 내 용의 강호축 종합개발계획을 4차 국가균형발전 5개년 계획에 반영하기 위함이다. 국가 균형발전과 대한민국 미래 100년 먹거리 창출을 위해 충청~강원~호남을 잇는 강호축을 꼭 개발해야 한다는 것이 충북의 입장이다. 충남도는 내년 국비 6조 3863억원을 확보하며 처음으로 정부 예산 6조원 시대를 열었다.

내포 명제: 국가 균형발전과 대한민국 미래 100년 먹거리 창출을 위해 충청~강원~호남을 잇는 강호축을 꼭 개발해야 한다 [비인식/의무]

■ ‘(불)가능하다’ 등의 모문 술어에 따라 내포절의 양태가 결정될 수 있다.

- (46) 실험 문항: 진로 설정과 탐색을 위해서는 다양한 직간접 경험이 필요하다. 하지만 우리가 모든 것을 직접 경험하기는 불가능하다. 그래서 간접경험을 많이 해야 하는데, 간접경험의 대표적인 방법이 바로 '독서'다.
- 내포 명제: 우리는 모든 것을 직접 경험할 수 있다

■ 양태를 나타내는 우연적 구성은 보문소와 모문 술어로 분리될 수 있다.

- (47) ㄱ. 실험 문항: 동주는 눈을 뜨고 벌떡 일어났다. 나무 침대 아래 용접봉과 용접기의 선들이 어지럽게 널려 있었다. 작업실에서 자동차 엔진을 용접하다가 잠이 들었던 모양이었다. 동주는 침대에서 일어나 창가로 다가갔다.
- 내포 명제: 동주는 작업실에서 자동차 엔진을 용접하다가 잠이 들었다
- 보문소: ㄴ 모양
- 모문 술어: 이다
- ㄴ. 실험 문항: 수안은 반 년 남짓 그 놀이에 매달려 있더니 계절이 바뀔 무렵 노트를 집어던지곤 방바닥에 드러누웠습니다. 그 시절 그 아이는, 어쩌면 세상에 없는 언어로 얘기하고 싶었는지도 모릅니다.
- 보문소: ㄴ지도
- 모문 술어: 모르다
- ㄷ. 실험 문항: 장사에 도움 되니까 쌍수 들어 환영하기는 했다만 집권당사 건물 올릴 때는 찝찝했던 게 사실이다. 정치하는 것들 가까이 오면 재수 옴팡이라든디. 중얼거렸던 게 나름대로 선견지명이었나 보다.
- 보문소: 나
- 모문 술어: 보다

■ 문맥이나 내포명제가 과도하게 어려워져서 작업자 스스로도 모호할 경우, 실험 문항에서 제외한다. 예컨대, 다음의 예시에서 인식양태와 비인식(의무) 양태가 상충한다.

(48) 실험 문항: 먹고 죽자고 자신만만하게 외치며 시작한 술자리는 의외로 빨리, 그리고 조용히 끝났다.

대상 문장: 아무래도 재경을 배려해주어야 할 것 같다는 의견일치 때문이었다.

내포 명제: 재경을 배려해 주어야 한다. (제외)

■ 양태를 수반한 내포절 명제의 경우, 데이터에 대한 판단이 엇갈리는 경우 실험에서는 제외한다.

(49) 실험 문항: 그러나 사단 작전과의 지시는 천만 뜻밖의 것이었다. 어제까지도 아군에게 절대적으로 필요했던 그 고지가 오늘은 작전 지역에서 제외되었으니 고지에서 전병력을 철수시키라는 것이다.

내포 명제: 고지에서 전병력을 철수시킨다. (제외)

■ ‘필자/화자’가 전달한 내용을 가능하면 유지할 수 있도록 내포절에 쓰인 보조사 및 부사류를 변경하지 않는다.

(50) ㄱ. 실험 문항: 내가 미쳤지, 왜 이런 일을 저지르는지 알 수가 없어. 다시는 폭력을 휘두르지 않을 테니까 용서해 주구려. 그녀는 때때로 아순시온으로 돌아가고 싶은 마음에 사로잡히기도 했다.

내포 명제: 그녀는 때때로 아순시온으로 돌아가고 싶었다.

ㄴ. 실험 문항: 지금은 연탄을 사용하는 곳이 별로 없지만, 예전에는 추운 겨울날 까만 연탄에 불을 지펴 방을 따뜻하게 하기도 하고, 그 불로 맛있는 음식을 만들어 먹기도 했습니다. 그런데 연탄은 그런 떨감의 역할만 했던 것은 아니었습니다.

내포 명제: 연탄은 떨감의 역할만 했었다.

■ 다만, 내포 명제의 주어나 목적어에 해당하는 경우 보조사의 형태를 수정할 수 있다.

- (51) 실험 문항: 아이들의 신체가 체계적이고 골고루 발달해야 공부도 열심히 할 수 있다. 여러분도 잘 알다시피, 건강만큼 좋은 선물이 있을까 싶다. 어린이들에게 컴퓨터도 중요하고 영어도 중요하지만 더욱 중요한 것은 자신의 신체를 잘 돌보고 어떻게 엑스사이즈를 해야 하는지 그래서 자신의 몸을 사랑하게 만드는 일을 유아기 때부터 심어 줘야 한다는 것이다.
내포 명제: 건강은 좋은 선물이다

■ 다만, 내포절에서 부가적인 정보만을 제공하는 단순 수식 성분은 내포 명제 기술 시 제거할 수 있다.

- (52) ㄱ. 실험 문항: 지금은 연탄을 사용하는 곳이 별로 없지만, 예전에는 추운 겨울날까만 연탄에 불을 지펴 방을 따뜻하게 하기도 하고, 그 불로 맛있는 음식을 만들어 먹기도 했습니다. 그런데 연탄은 그런 역할만 했던 것은 아니었습니다.
내포 명제: 연탄은 역할만 했었다
- ㄴ. 실험 문항: 둘째, 부자들은 부자들의 행동과 부자들의 모임에서 관심을 떼지 않는다. 돈이 있는 곳에서 돈에 대한 정보가 나오는 걸 알고 있기 때문이다. 오죽하면 부자들이 사는 동네에서 살면 재테크의 반은 성공했다는 얘기가 나오겠는가.
내포 명제: 부자들이 사는 동네에서 살면 재테크의 반은 성공한다.
- ㄷ. 실험 문항: 또 낮은 전압에서도 효율이 80%이상 유지되며, 충전이 가능한 전압으로 키워 변환시켜 주는 회로 기술도 접목시켰다. 이와같은 기술은 기존에도 있었지만 상용화 수준의 에너지 출력의 성과는 이번이 처음이라고 연구진은 설명했다. 연구진은 패치 착용시 불편함을 제거하고 착용시 심미감, 움직이는 상황에서의 특성 고려해 상용화가 이뤄지도록 연구가 더 필요하다고 말했다.
내포 명제: 상용화 수준의 에너지 출력의 성과는 이번이 처음이다

■ 모문 술어가 ‘아니다’, ‘없다’, ‘않다’ ‘와 같은 부정어인 경우는 ‘이다’, ‘있다’의 부정형이 아닌 부정어 형태 그대로 보문 술어로 간주한다.

- (53) ㄱ. 실험 문항: 유 열사의 생가는 용두리 지령부락 입구에 복원·조성돼 있다. 그런데 이 복원된 생가는 실제 유관순 열사의 생가터가 아니라고 마을주민들은 기억하고 있다.
내포 명제: 용두리 지령부락 입구에 복원된 생가는 실제 유관순 열사의 생가터가 아니다.

ㄴ. 실험 문항: 2연승 거두며 공동2위 올랐지만 김민욱-랜드리 발목 다쳐 교체 프로농구 부산 kt가 2018년 마지막 날과 새해 첫날에 걸쳐 치러진 ‘농구영신’ 매치를 승리로 장식했다. 센터 김민욱의 부상 악재만 없었다면 더욱 기분 좋은 출발이 될 뻔했다. kt는 지난달 31일 밤 11시 창원체육관에서 열린 2018-2019 프로농구 창원 LG전에서 79-70으로 이겼다.

내포 명제: 센터 김민욱의 부상 악재는 없었다.

ㄷ. 실험 문항: 광릉 숲 경계로부터 20미터 이상 떨어져 있으면 매립장 건설이 가능하다는 회신도 환경부로부터 받았습니다. 양남시는, 다이옥신이나 중금속은 비산재에 포함돼 있고, 바닥재는 그렇게 환경적으로 위험하지 않은 것으로 보고 있습니다. 이번 에 건설되는 매립장에는 바닥재만 매립하는 데다가 갖가지 안전 조치를 취하기 때문에 환경적으로 염려할 이유가 전혀 없습니다.

내포 명제: 바닥재는 다이옥신이나 중금속처럼 환경적으로 위험하지 않다.

■ ‘않/안-’ 등 부정어를 포함한 내포 명제 형태 역시 그대로 유지한다.

(54) ㄱ. 실험 문항: 김 위원장의 지난 1일 신년사에 대한 첫 반응으로, 2차 북-미 정상회담에 대한 의지를 재확인한 것이다. 트럼프 대통령은 1일(현지시각) 저녁 6시10분께 트위터에 “김정인이 북한은 핵무기를 만들지도 시험하지도 않으며 전파하지도 않을 것이라고 말했다. 또한 그는 트럼프 대통령을 언제라도 만날 준비가 돼 있다고 말했다”는 미 공영방송 <피비에스>(PBS) 기사의 한줄을 올렸다.

내포 명제: 북한은 핵무기를 만들지도 시험하지도 않으며 전파하지도 않을 것이다.

ㄴ. 실험 문항: 그러나 이 책을 읽은 독자들은 좋은 서비스를 제공하는 습관을 들이면 성격도 긍정적인 방향으로 바뀐다는 사실을 알게 될 것이다. 요즘 이구동성으로 경기가 안 좋다고 한다. 백화점이나 호텔의 매상도 예년과 다르다고 한다. 그러나 이런 불황에도 호황인 업체들이 있다.

내포 명제: 요즘 경기가 안 좋다.

9) 토픽 마커(topic marker)

■ 내포절 주어가 문맥상 ‘주제/화제’ 표지 ‘은/는’인데 ‘이/가’로 중화되어 있는 경우, ‘은/는’으로 바꾸어 표시하는 것을 원칙으로 한다. 전체 문장을 읽고 내포절의 확실성을 판단하는 것이 본 연구의 실험 조건임을 고려할 때, 주어가 대부분 ‘주제/화제’인 경우가 많으므로 ‘은/는’을 붙이는 것이 자연스럽게 때문이다.

(55) ㄱ. 실험 문항: 신고를 받고 출동한 지구대 경찰은 10대 후반에서 20대 초반 남녀가 함께 생활하는 원룸에서 온몸이 상처투성이인 20대 지적장애인 A(23)씨를 발견했다. 경찰은 주변 사람들의 눈치를 심하게 보던 A씨가 지적장애인이라는 것을 금방 알아챘다. 가족의 품에 다시 안긴 피해자 A씨의 입에서는 충격적인 사건의 전말이 쏟아져 나왔다.

내포 명제: 주변 사람들의 눈치를 심하게 보던 A씨는 지적장애인이었다.

ㄴ. 실험 문항: 『동의보감』, 『향약구급방』, 『규화보전』 등과 같은 옛 의학서에는 수없이 많은 금기 식품들이 지정되어 있는데, 대개는 과학적으로 근거가 없는 것이다. 하지만 당시의 사람들은 그것을 믿고 있었기 때문에 먹고 싶어도 먹지 못하는 안타까움을 많이 겪어야 했다. 임산부가 달걀을 먹으면 안된다는 것도 그 많은 금기사항들 가운데 하나였다. 임산부나 수유부에게는 특별히 많은 음식을 금기하여 조심시켰는데, 대개 아이가 잘못될 수 있다는 위협을 바탕으로 깔고 있다.

내포 명제: 임산부는 달걀을 먹으면 안된다.

10) 시제

■ 내포 명제는 내포절 술어의 시제를 그대로 쓰는 것을 원칙으로 한다.

(56) ㄱ. 실험 문항: 이란의 행보를 한국도 주목할 수밖에 없다. 큰 이변이 없는 한 이란은 이라크를 잡고 선두로 토너먼트 라운드에 진출할 것으로 보인다.

내포 명제: 큰 이변이 없는 한 이란은 이라크를 잡고 선두로 토너먼트 라운드에 진출할 것이다

ㄴ. 실험 문항: 행정기관 전산망에는 쌍둥이 모자가 영암 모 아파트에 거주한다고 나왔으나, 현장 조사를 해보니 애초부터 이 곳에 살지 않았다는 주변인들의 진술을 얻는 데 그친 것이다. 전남도교육청은 쌍둥이 친모가 아들들을 데리고 출국했을 가능성을 고려해 출입국사무소에 문의했으나 출국기록은 없었다. 지난 17일 교육당국 의뢰로 경찰 수사가 개시되면서 상황은 예사롭지 않게 흘러갔다.

내포 명제: 쌍둥이 친모가 아들들을 데리고 출국했다

■ 모문의 술어는 과거 시제로, 내포절의 술어는 현재 시제로 표시되는 경우 중, 내포절의 시제 해석이 모문에 의존적인 경우에는 내포절에 과거 시제를 복원하여 표시한다 (박진호 2011).

(57) ㄱ. 실험 문항: 그렇게 말하려던 것이 아니에요. 내가 어리석었어요. 내가 그의 의도를 간파했어야 했어요. 나를 '기억의 집'으로 불러 이 노트를 넘겨줄 때, 나는 그가 무엇인가를 모의하고 있다는 사실을 감지했어야 했어요.

내포 명제: 그가 무엇인가를 모의하고 있었다. (*그는 무엇인가를 모의하고 있다.)

ㄴ. 대상 문장: 그래서 우리의 질문은 언제나 그 내용과 대상으로 향한다. 왜 우리는 의도의 주체에 대해서는 질문할 줄을 몰랐을까. 누가, 또는 무엇이 그것을 의도하는가를 질문하지 않은 것은, 혹시 우리가 너무 개와 뉴스 가치에 대한 격언을 잘 이해하고 있기 때문이 아닐까

내포 명제: 우리는 의도의 주체에 대해서 질문했다. (*우리는 의도의 주체에 대해서 질문한다.)

■ 다만, 내포절의 명제가 초시간적(즉, 총칭) 명제인 경우에는 모문의 시제를 따르지 않고 현재 시제로 기술한다(전영철 2003).

(58) ㄱ. 실험 문항: '그렇긴 하네.' 안의 설명을 듣고 나자 마나연초를 펴서 차원력을 회복하는 건 분명 한계가 있다는 걸 깨달을 수가 있었다. 더욱이 마나연초는 공짜로 구할 수 있는 물건도 아니었다.

내포 명제: 마나연초를 펴서 차원력을 회복하는 건 분명 한계가 있다. (총칭)

ㄴ. 실험 문항: 새로운 시대에 발맞춰 혁신을 시도하는데 규제가 발목을 잡고 있다는 지적이다. 규제 해소는 정부가 앞으로 중점 추진할 대목이라고 입을 모았다.

내포 명제: 규제 해소는 정부가 앞으로 중점 추진할 대목이다. (총칭)

■ '기대하다' '바라다' '빌다' '희망하다' 와 같은 바램을 나타내는 모문 술어의 경우 내포절은 미래 시제를 복원하여 표시한다.

(59) ㄱ. 실험 문항: 책의 부피를 조금이라도 줄이기 위해 참고문헌은 초판에서처럼 각 주로 대체하였고 '찾아보기' 부분도 인명의 경우는 외국인도 '가나다'순으로 배열한 초판을 그대로 두었고 사항의 경우에 개정판 작업의 일환으로 일부만 고쳤다. 아무쪼록 개정판이 초판에 비해 더 나은 책으로 독자들에게 다가갈 기를 기대한다.

내포 명제: 개정판이 초판에 비해 더 나은 책으로 독자들에게 다가갈 것이다.

ㄴ. 실험 문항: 그런 차에 세상사에 대한 그의 명쾌한 법률 이야기가 한 권의 책으로 우

리 곁에 다가온다니 더없이 기쁘고 또 고맙다. 아무쪼록, 이 한 권의 책이 만나는 사람 모두에게 터널 끝의 밝은 빛이 되기를 바란다.

내포 명제: 이 한 권의 책이 만나는 사람 모두에게 터널 끝의 밝은 빛이 될 것이다.

㉔. 실험 문항: 그래서 감정을 압축하는 중이야. 무작정 기다리기만 하는 게 능사일지 모르겠지만…… 아무튼 작업 잘 되길 빌게.

㉕. 실험 문항: 나는 이 책이 이땅의 여성 현실을 이해하는 데 도움이 되었으면 합니다. 그리고 특히 젊은 여성들이 읽어서 여성의 인간화를 막는 구조적인 모순을 물리치는 데 힘을 얻게 되기를 희망합니다.

내포 명제: 젊은 여성들이 이 책을 읽어서 여성의 인간화를 막는 구조적인 모순을 물리치는 데 힘을 얻게 될 것이다.

■ 보문소가 ‘㉔ 가능성’, ‘㉕ 경우’일 경우 내포 명제는 미래 시제로 복원한다.

(60) ㉔. 실험 문항: 무역분쟁 장기화와 세계 철강 수요의 절반가량을 차지하는 건설경기 둔화는 철강 업황 회복을 늦출 위험 요인으로 꼽힌다. 다만 전문가들은 국내 철강업체의 기초 체력이 좋아져 예전처럼 실적이 곤두박질칠 가능성은 거의 없다고 입을 모은다. 금융정보업체 에프앤가이드에 따르면 포스코의 올해 예상 영업이익은 5조2247억원이다.

내포 명제: 국내 철강업체의 실적이 곤두박질칠 것이다

㉕. 실험 문항: 총 120조원을 2028년까지 단계적으로 투입한다. 업계에서는 SK하이닉스가 클러스터 조성 비용의 상당 부분을 부담할 것으로 예상하고 있다. 반도체클러스터 구미유치가 성사될 경우 침체된 구미경제는 획기적 전환점을 맞게 된다.

내포 명제: 반도체클러스터 구미유치가 성사될 것이다

■ 문맥적 정보에서 ‘머지않아’, ‘곧’과 같은 미래 시제를 나타내는 부사어가 있는 경우 내포절을 미래 시제로 복원한다.

(61) 실험 문항: 용역계약 민간 청소업체 두 곳에서 물 청소차 6대와 진공 흡입차 13대를 동원해 폭 12미터 이상 도로에 매일 한 차례씩 물청소를 한다니. 이제 한국의 부자들이 머지않아 미국의 부자들을 따라하지 않을까 모르겠다. 미국 부자의 대명사인 '하이퍼리치'(hyper rich)의 생활처럼 주말에 전용기로 날아다니고, 저녁 식사에 500달러짜리 포도주를 마시는 것 말이다.

내포 명제: 한국의 부자들이 머지않아 미국의 부자들을 따라할 것이다

11) 기타

■ 낫선 옛글은 제외한다.

(62) 실험 문항: 청하웁건대, 전하께서는 하루빨리 여총지화를 깨달으시어 궁인 장씨를 꺾박으로 내치심이 옳은 줄로 아옵니다.

내포 명제: 전하께서는 하루빨리 여총지화를 깨달으시어 궁인 장씨를 꺾박으로 내치심이 옳다. (제외)

3.2 구어 지침

내포 명제 기술 시 다음의 원칙들을 따른다.

1) 대상 문장 형성

■ 대상 문장은 턴테이킹이 이루어지더라도 선후 문장을 통해 완전한 문장 구조를 형성한다.

(63) 대상 담화:

P3: 어쨌든 돼지 책은

P2: 네.

P3: 어~ 이렇게 엄마가 하는 일이 꼭 엄마가 해야 하는 일이 아님을

P1: 음~

P2: 음~

P3: 보여주고 있는

P1: 네.

P3: 동화책이었습니다.

P1: 네.

P2: 네.

P1: 자 다음 책은

P3: 네. 다음 책은 그림보 출판사에서 나온 서선연 글 오승민 그림에 호랑이를 탄 엄마입니다.

P1: 네.

대상 문장: 어쨌든 돼지 책은 어~ 이렇게 엄마가 하는 일이 꼭 엄마가 해야 하는 일이
아님을 보여주고 있는 동화책이었습니다.

내포 명제: 엄마가 하는 일이 꼭 엄마가 해야 하는 일은 아니다.

2) 복원

■ 대상 문장의 내포절과 선행 문장에 주어가 없어 명제 내용을 파악하기 어려운 경우에는 문맥상 예측할 수 있는 선행행 주어를 복원하여 제시한다.

(64) ㄱ. 실험 문항:

이거를 갖다가 그~ 한국 당 선관위가 뭐 전당 대회 그~ 날짜를 변경하는 것이 뭐 가능한지 뭐 장소라든지 비용 문제 분명히 고려해야 되잖아요. 그니까 이런 걸 지금 검토한 것으로 전해지고 있고 이게 잠시 후 두 시에 관련해서 회의를 진행할 입장입니다.네. 이제 얼마 안 남았잖아요.

내포 명제: 한국 당 선관위가 두 시에 전당 대회에 관련해서 회의를 진행할 것이다

ㄴ. 실험 문항:

P3: 우리나라가 전체 살인 중에서 이 존속 살해가 오 프로가 된다는 거예요.

P1: 아하

P3: 그니까 살인 사건 백 건 중에서 다섯 건은 존속 살인 사건이라는 거죠.

P1: 예.

P3: 이 비율은 이제 얼마나 높은가 하면은 미국 같은 경우에는 존속 살인 비율이 이 프로 영국 같은 경우 일 점 오 프로

P1: 아하

P3: 이런 주요 국가들에 비해서 서너 배 높게 나타나는 것으로 밝혀졌습니다.

P1: 네. 심은숙 선생님

P2: 예.

P1: 자~ 이렇게 인제 가족 간에 불화가 끔찍한 결과에 이르기기도 하는데

P2: 예.

P1: 대개 유산 상속 문제일 수도 있고

P2: 네.

P1: 부모를 누가 모시느냐 이 다툼일 때도 있고 그래요.

대상 문장: 이런 주요 국가들에 비해서 서너 배 높게 나타나는 것으로 밝혀졌습니다.

내포 명제: 미국과 영국에 비해서 우리나라는 존속살인 비율이 서너 배 높게 나타난다

■ 대상 문장의 내포절과 선행 문장에 목적어가 없어 명제 내용을 파악하기 어려운 경우에는 문맥상 예측할 수 있는 선행 목적어를 복원하여 제시한다. 이 때 문맥을 통해 의미적으로 내포 명제를 보강할 수 있으면 사용한다.

(65) ㄱ. 실험 문항:

지금 보면 재미있는 것이 육십 대 이상에서는 오십이 점 삼 퍼센트가 어~ 무임 승차 상향 조정에 찬성한다. 그러니까 한 칠십 세 정도로 해서 상향해도 좋겠다는 의견이 이렇게 많이 있는 거 같고요. 그리고 또 재미있는 게 젊은 세대 이십 대가 삼십팔 퍼센트가 무임 승차 폐지해라 폐지해야 된다 이런 주장이 나왔어요.

내포 명제: 무임승차 연령을 칠십 세 정도로 상향해도 좋겠다

ㄴ. 대상 담화:

P2: 그런데 지금 보세요. 이번 지방 선거 어~ -피- 투표 한 장에 투입되는 비용이 에~ 만 원을

P1: 아~

P2: 조금 넘는다고 집계가 돼 있는데

대상 문장: 조금 넘는다고 집계가 돼 있는데

내포 명제: 이번 지방 선거 투표 한 장에 투입되는 비용이 만 원을 조금 넘는다.

■ 보다 명확한 내포 명제 구성을 위해 문맥을 통해 부사어나 수식어 등의 보조적인 정보를 복원하여 제시한다.

(66) 실험 문항:

P1: 그거 인제 지구상에 하나밖에 없는 단어 아닙니까? 뽀로로.

P2: 예. 그렇습니다

P1: 예.

P2: 그~ 어~ 우리나라를 비롯해서 여러 나라에 이미 저작권 등록. 상표 등록이 되어 있기 때문에

P1: 네.

P2: 유사한 이름으로 나오기는 좀 어려울 겁니다.

P1: 예 근데 어ㄹ까? 지 지 으~ 지으시게 된 겁니까.

대상 문장: 유사한 이름으로 나오기는 좀 어려울 겁니다.

내포 명제: 뿌로로와 유사한 이름의 상표가 나올 것이다.

3) 인칭 대명사

■ ‘너’, ‘우리/저희’, ‘자기 자신’ 등은 선행 문맥에서 가장 가까운 지시적 표현으로 복원한다.

(67) 가. 실험 문항:

유라: 죽일 듯이 사랑하거나, 죽일 듯이 미워하거나, 내가 아는 건 두 가지밖에 없으니까

정금: ...

유라: 죽지도 않으니까 나는.. 약을 먹어도, 술을 먹어도..

정금: 내 이혼에 대해서.. 지금 와서 생각해보면 (목소리가 갈라져서 헛기침 한번 하고).. 애들 아파하고 나 그렇게 이혼한 거에 대해서, 전적으로 너에게 책임 있다고 생각하지 않는다..

정금: 내 남편이었던 사람.. 작은 유혹에 무너져버린 우리의 결혼생활.. 지금 생각해 보면 너 아니어도 우리 지금쯤 헤어져 있었을 거야 아마.. 그러니까 이제 그 생각으로부터는 최소한 자유로와져라..

대상 문장: 내 이혼에 대해서.. 지금 와서 생각해 보면 (목소리가 갈라져서 헛기침 한번 하고).. 애들 아파하고 나 그렇게 이혼한 거에 대해서, 전적으로 너에게 책임 있다고 생각하지 않는다.

내포 명제: 내가 남편하고 이혼한 것에 대해서 전적으로 유라에게 책임 있다.

나. 실험 문항:

사실은 몇 년 전에 어느 농민께서 저한테 하신 말씀입니다. 이런 얘기를 하더라고요. 저희가 정부 앞에 가서 정말 돌을 던지고 데모하는 것은 무조건 지원해달라는 돈을 달라는게 아닙니다. 우리가 진정으로 바라는 것은 희망입니다.

내포 명제: 농민들은 돈을 달라고 정부 앞에 가서 돌을 던지고 데모한다.

다. 실험 문항:

네. 소크라테스는요. 자기 자신을 스스로 아테네의 등어로 자처했다 그러죠. 등어가 흠스 플라이 이래가주구요. 좀 낫설죠?

내포 명제: 소크라테스는 스스로 아테네의 등어로 자처했다

■ 다만 인칭 대명사가 대상 담화 내에서 복원되지 않고 확신성 판단에 무리가 없다면 그대로 사용하며 지시 대상을 병기하지 않는다.

(68) 실험 문항:

은영: 나 열 안나요, 그냥 나쁜 꿈을 꿔어요.
형우:
은영: 너무 아픈 꿈인데..... 깨워줘서 고마워요.
형우: 고맙습니다.
은영: 형우 씨 이제 나한테 화 풀렸나 보다.. 그쵸?
형우: 형우 씨 화 안 났습니다.
내포 명제: 형우 씨가 나한테 이제 화가 풀렸다

■ 이때, 낮춤말 ‘저’가 쓰인 경우에는 ‘나’로 바꾸어 기술한다.

(69) 실험 문항:

P7: 노력하자.
P6: 언행일치
P5: 음~
P7: 예.
P5: 음~
P7: 예. 생각하는 대로 살자.
P7: 근까 옳은 것을 생각하고 그렇게 살려고 노력을 해야겠다는 생각을 요즘 많이 하고 있어서 저두 지나처럼 살려고 많이 노력하고 있습니다.
P6: 음~ 제가 알기로는 작가님이 최소주의의 삶을 좀 살고 계시다고 들었어요.
P6: 어~ 쯤 맛있는 거 많이 먹구 이러면
P7: 음.
P6: 다음에 또 먹고 싶으니까
P7: 네.
내포 명제: 나는 옳은 것을 생각하고 그렇게 살려고 노력을 할 것이다.

4) 대용어/지시사

■ ‘그’, ‘그렇게’, ‘하다’와 같은 대용적 표현과 ‘여기’, ‘거기’, ‘이때’와 같은 지시사는

선후행 문맥을 통해서 복원한다.

(70) ㄱ. 실험 문항:

P2: 요즘에 나오는 퍼레이드 보편은 어떠세요? 좀

P4: 어~ 일단 김정은 정권에 와서 -달라졌- 크게 달라졌다고 봐요.

P1: 음

P4: 형식이라든가 규모

P1: 예.

P4: 예. 상당수 많이 커졌죠.

P1: 음

P4: 그까 그만큼 비용도 많이 증가됐다는 걸 볼 수 있는 거 같습니다. 제가 볼 때

P1: 그렇죠.

P5: 네.

P1: 그래서요.

내포 명제: 요즘 김정은 정권의 퍼레이드는 형식이나 규모 만큼 비용도 많이 증가됐다

ㄴ. 실험 문항:

P1: 뭐 아무래도 이런 문제가 있다 싶으면은

P3: 네.

P1: 어 반드시 인제 이런 상담센터를 통해서 치료를 받아야겠죠?

P3: 네네.

P1: 그러니까 뭐 집에서 어떻게 대충 어 말 훈련시켜서 된다 이걸 아닌 거 같아요.

P3: 어 맞아요. 저희는 아무래도 아동의 언어발달에 어 집중적으로 저희가 지원을 하고 있기 때문에

P1: 예.

P3: 간혹 그렇게 생각하시는 분들도 계시기는 하지만

P1: 네.

P3: 어 저희 전문가의 도움이 필요하다고 생각이 됩니다.

내포 명제: 간혹 집에서 대충 말을 훈련시키면 아동 언어발달 문제를 해결할 수 있다고 생각하시는 분들도 계신다

ㄷ. 실험 문항:

P1: 야 여기서 오늘 쿡 포인트가 될 수도 있을 것 같은데요.

P2: 네.

P1: 어 꿀뚜기 요리하실 때 연유를 넣으면은

P2: 네.

P1: 비린내를 또 말끔히 잡을 수 있다는 걸 또 우리가 알 수 있겠네요.

대상 문장: 야 여기서 오늘 쿡 포인트가 될 수도 있을 것 같은데요.

내포 명제: 연유가 오늘 쿡 포인트가 될 수도 있다.

ㄹ. 실험 문항:

P1: 사실 우리가 제주도에 놀러 간다라고 하면

P2: 네.

P1: 공항에 내려서

P2: 네.

P1: 이렇게 조금만 가면 그 유명한 그 갈치 집 있잖아요.

P2: 네네네네.

P1: 거기에 항상 들려서 먹고 출발하는 경험들이 좀 많았거든요.

P2: 오~ 네. 그러셨구나 그래서 만약에 펜션이나 이런 데서 요리가 가능을 하시면

내포 명제: 나는 제주도에 놀러갔을 때 항상 유명한 갈치 집에 들려서 먹고 출발했다

ㅁ. 실험 문항:

그만큼 어 어목한 시기에 삼 월 일 일에 이십 년 삼 월 일 일에 학교 뒤편 기숙사 뒤편에서 독립만세를 외쳤습니다. 이때 어 삼십여 명 정도가 했다고 해요. 그런데 그 중 여섯 명만 확인돼 있는 거예요.

내포 명제: 이십 년 삼 월 일 일에 학교 뒤편 기숙사 뒤편에서 삼십여 명 정도가 독립만세를 외쳤다

5) 정보구조

■ 구조적으로 내포 명제가 명확하지 않다고 하더라도, 선후행 문맥을 통해 의미적으로 내포 명제를 보강할 수 있으면 사용한다.

(71) ㄱ. 실험 문항:

P2: 메이저 리그는 추 출전을 못 하겠다는 거고 그러다 보니까 올림픽에서 야구 빠져라.

P1: 음

P2: 빼 빠린거예요.

P1: 그렇죠.

P2: 공가 이제 메이저리그는 그래 그럼 우리끼리 대회를 만들겠다.

P1: 예.

P2: 이 피파 월드컵처럼 그게 더블유비씨예요.

P1: 더블유비씨

P2: 근데 그 메이저리그가 너무 황포를 부리다 보니까 이제 이게 일본이 화가 난 거예요.

P1: 요번에 그 프리미어 트웰브

P2: 예.

P1: 대회는 일본이 주도한 그런 대회죠?

내포 명제: 메이저리그가 더블유비씨 대회를 만들겠다고 황포를 부렸다.

ㄴ. 실험 문항:

P2: 그러기도 하고. 요즘은 그~ 강남에 있는 젊은이들이 구호선을 타고 그~ 인천공항에 와서 많은 시간을 보내고 인천공항에 예 같이 볼거리. 먹을거리.

P1: 아 비행기 타러 오는 건 아닌데?

P2: 아. 뭐.

P1: 예.

P2: 문화예술의 그~ 향기가 난단 어~ 그거 말씀대로

P1: 음.

P2: 어~ 와서 즐기는 사람들입니다.

내포 명제: 강남에 있는 젊은이들이 구호선을 타고 인천공항에 비행기 타러 온다.

■ 대상 문장에서 중복되는 부분은 삭제한다.

(72) 실험 문항:

P1: 앞선 브리핑 내용 듣고 오겠습니다.

P1: 자 이 제주도에선 시신을 유기하지 않았다 이런 발표 내용인데 이 내용이 사실이 아닌 것으로 드러났나 봐요.

P2: 네. 경찰이요. 이 고유정이 제주도 내에는 시신을 유기하지 않았다 라고 방금 보신 것처럼 이야기를 했는데

P1: 예.

내포 명제: 고유정이 제주도에 시신을 유기하지 않았다는 내용은 사실이 아니다.

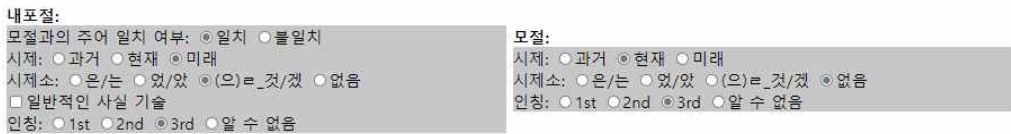
3.3. 언어 정보 부착(주석) 지침

대상 문장(target sentence)의 언어 정보를 분석하여 주석할 때에는 다음의 원칙들을 따른다.

3.3.1. 시제

■ 워크벤치에서 체크 박스로 내포절과 모절의 시제 표지 목록을 표시한다. 이를 통해 간편하고 오류가 줄어든 태깅 방식을 지향한다.

[그림 4] 작업자 워크벤치(시제/인칭 부분)



■ 모절 술어와 내포절 술어의 형태(形態, morph)를 기준으로 시제 정보를 부착한다.¹⁰⁾ 시제 정보는 내포절 술어에서 모절 술어의 순서로 제시한다. 이때 이형태(異形態, allomorph)를 구분하여 주석하지 않으며 시제소가 드러나지 않는 경우에는 [없음] 값을 부착한다. 아래와 같이 대표형으로 주석한다.

- (73) ㄱ. 그분은 언젠가 모든 것이 끝나게 되어 있다는 종말이 더 강조되어야 한다고 생각했
을 거야. [현재, 은/는, 과거, 었/았]¹¹⁾
 ㄴ. 중국인은 아빠 같은 거렁뱅이에게만 진정한 친구의 손을 내민다는 것을 알아야 할
것 같구나. [현재, 은/는, 현재, 없음]
 ㄷ. 그~ 더 늘어날 것이라고 저는 어~ 단언할 수 있습니다.[미래, (으)ㄹ 것, 현재, 없음]

■ 모절의 함의 취소 운용소가 '의문'인 경우, 의문형 종결어미(-ㄴ가, -는가, -느냐 등)의 일부분 (ㄴ, 는, 느 등)을 시제소로 주석하지 않는다.

- (74) 내가 부부에 대해, 남녀에 대해 무엇인가 제대로 알고 있기는 한 것인가?
 [현재, 없음, 현재, 없음]

10) 이에 따라 모절의 시제는 주로 발화시나 사건시가 기준이 되는 절대 시제의 값이, 내포절의 시제는 주로 모절의 사건시에 의존하여 상대적으로 결정되는 상대 시제의 값이 부착된다.

11) 괄호 속의 시제 정보는 [내포절 시제, 내포절 시제소, 모절 시제, 모절 시제소] 순이다.

3.3.2. 인칭

■ 워크벤치에서 체크 박스로 내포절과 모절의 인칭 목록을 표시한다. 이를 통해 간편하고 오류가 줄어든 태깅 방식을 지향한다(그림 4 참조).

■ 모절 주어와 내포절 주어의 인칭 정보를 1, 2, 3으로 구분하여 부착한다. 모절 주어와 내포절 주어가 일치하는 경우에는 [일치] 값을, 일치하지 않는 경우에는 [불일치] 값을 부착한다.

- 1인칭을 부착하는 경우: 주어가 필자/화자로 해석이 되는 경우이다.
- 2인칭을 부착하는 경우: 청자로 해석이 되는 경우이다.
- 3인칭을 부착하는 경우: 필자/화자가 술어의 주체로서 분명하게 배제되는 경우이다.

(75) ㄱ. 나를 '기억의 집'으로 불러 이 노트를 넘겨줄 때, 나는 그가 무엇인가를 모의하고 있다는 사실을 감지했어야 했어요. [1, 3, 불일치]

내포절 주어: 그 - 3인칭

모절 주어: 나 - 1인칭

ㄴ. 경찰이 그걸 보고 당신이 빠리에 있는 줄로 생각할지도 모르거든요. [2, 3, 불일치]

내포절 주어: 당신 - 2인칭

모절 주어: 경찰 - 3인칭

■ 주어가 표면에 드러나지 않는 경우 작업자가 주어를 복원하여 처리한다.

(76) ㄱ. 사람들은 아주 작은 일에도 감동을 받고 좋아한다는 것을 알아야만 한다 [3, 1, 불일치]

내포절 주어: 사람들 - 3인칭

모절 주어: 우리(미명시) - 1인칭

ㄴ. 지금 당신의 나이가 몇 살이든, 처한 환경이 어떠한든 중요한 것은 지금도 늦지 않았다는 사실이다. [2, 3, 불일치]

내포절 주어: 당신(미명시) - 2인칭

모절 주어: 중요한 것 - 3인칭

■ 미명시된 주어를 복원할 수 없는 경우 '알 수 없음'으로 처리한다.

(77) ㄱ. 누군가 이걸 열어서 먹는다는 사실을 전제하지 않는다면야 모든 피스타치오가 다 평화겠쇼.

내포절 주어: 누군가 - 3인칭

모절 주어: 알 수 없음('전제하다'의 주체가 누구인지 해석 불가)

ㄴ. 그래야 잠재적 범죄자들이나 또는 범죄를 저지르려고 마음을 갖고 있는 사람들에게 범죄를 하면 언젠가는 잡히게 된다는 걸 명확하게 보여주지 않으면

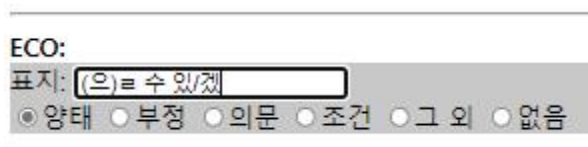
내포절 주어: 잠재적 범죄자들이나 또는 범죄를 저지르려고 마음을 갖고 있는 사람들 - 3인칭

모절 주어: 알 수 없음('보여주다'의 주체가 누구인지 해석 불가)

3.3.3. 함의 취소 운용소(ECO)

■ 워크벤치에서 체크 박스로 함의 취소 운용소가 있으면 그 종류를 표시한다. 없으면 '없음'에 표시한다.

[그림 5] 작업자 워크벤치(ECO 부분)



■ 모절의 함의 취소 운용소 정보, 즉 '부정, 의문, 조건, 양태' 정보를 부착한다. [양태]의 경우 '양태 분류, 양태 표지'를 추가로 부착한다. '양태 분류'는 인식 양태에 대해서는 [인식] 값을, 비인식 양태에 대해서는 [가능, 불가능, 의무, 의지] 등 2차 분류 값을 부착한다.

(78) ㄱ. 중국인은 아빠 같은 거렁뱅이에게만 진정한 친구의 손을 내민다는 것을 알아야 할 것 같구나. [(으)ㄹ 것 같-, 양태/인식]

ㄴ. 그만큼 합격자의 -교- 그 학생의 교과 성적이 학생의 발목을 붙잡을 수 있는 상황임을 인제 좀 분명하게 인식하셔야 되구요 [-어야 되-, 양태/비인식/의무]

ㄷ. 돈 이전에 사람이 있다는 사실 유념한다면 이런 비극은 없지 않을까요. [-다면, 조건]

ㄹ. 왜 우리는 의도의 주체에 대해서는 질문할 줄을 몰랐을까. [-ㄹ까, 의문]

ㅁ. 죠다쉬는 트럼프 한 장이 남아 있다는 것을 눈치채지 못했다. [-지 못하, 부정]

■ 모절에 함의 취소 운용소가 둘 이상인 경우에는 ‘양태/부정/의문/조건’의 순서로 + 를 사용하여 주석한다.

- (79) ㄱ. 소득 1만 달러 시대를 사는 한국 시민들의 주요 관심사가 건강·환경과 관련된 삶의 질 향상에 있다는 점을 감안한다면, 생명의료윤리에 대한 이 같은 높은 관심은 자연스러운 것이라 할 수 있겠습니까. [-ㄹ 수 있-/-겠-, 양태/인식+양태/인식]
- ㄴ. 과연 스프링클러 설치가 재건축을 해야 될 만한 사유인지 한번 따져봐야 될 거 같애요.[-어야 되-/- ㄹ 것 같-, 양태/비인식/의무+양태/인식]
- ㄷ. 그러나 그들이 어찌 그윽하면서 넓은 경내가 그 사이에 있는 줄을 알겠는가? [-겠-/-ㄴ가, 양태/인식+ 의문]

4. 실험 지침

4.1. 실험 수행 지침 및 절차

4.1.1. IRB 승인

사람을 포함하는 언어 실험 수행은 사전에 생명윤리위원회의 심의를 거쳐 IRB 승인을 받을 필요가 있다.

4.1.2. 피실험자 모집

피실험자는 한국어를 모국어로 사용하는 19세 이상의 사람들로 한정하였다. 피실험자 모집은 본 사업의 수행에 앞서 진행된 별건의 사전 언어 실험에 참여한 적이 있는 사람들 119명 가운데 “성실도”가 높은 인원 60명을 선정함으로써 이뤄졌다. “성실도”를 평가한 기준은 다음과 같다.

먼저, 사전 언어 실험 설문지의 문항 중 명백히 비문법법적인 문장과 명백히 문법적인 문장들로 구성된 48개의 필러(filler) 문항들을 배치하였다(7점 척도의 질문에서 비문법적인 문장은 1~3점으로, 문법적인 문장은 4~7점으로 선택했을 시 맞춘 문장으로 간주한다).

이 문항들을 가장 많이 맞춘 연구대상자를 “성실도”가 높은 참여자로 판단하였고, 이 중 본 사업의 언어실험에 참여하기를 원하는 인원을 모집하였다. “성실도”가 높은 사람들 중 본 실험 참가 희망자는 총 63명이었으나, 이 중 참가 희망을 묻는 사전 설문 조사에 배치된 문항 적응용 샘플 문항에서 한 번호로만 찍는 등의 아웃라이어(outlier)로 판단되는 인원 3명을 제외하였고, 이에 따라 총 60명의 피실험자를 모집하게 되었다.

4.1.3. 실험 수행

피실험자에게 피로 효과(fatigue effect)가 나타나지 않도록 20분 이내에 풀 수 있게 설계하였으며, 이에 따라 1개 세트 당 40~50문항 사이로 구성하였다. 60명을 매 실험마다 무작위로 구성한 3개 집단(20명/20명/20명)으로 나누었고, 1개 집단 당 1개의 실험 세트에 대한 실험을 진행하였다. 각 집단은 실험마다 무작위로 섞여서 이루어졌기 때문에 집단별 실험 구성원이 고정됨으로써 생길 수 있는 편향을 방지하였다.

본 사업의 언어 실험은 일반적인 수용성 판단 언어 실험에 비해 문항의 이해 및 풀이가 까다롭기 때문에, 피실험자가 본 언어 실험의 문항 풀이 방법에 대해 충분히 이해하도록 하기 위하여 각별히 주의하였다. 이를 위해, 전자우편으로 언어 설문 조사 온라인 링크를 보낼 때, 실험 문항을 푸는 데 알아야 할 안내 사항을 적었고, 아울러 피실험자들이 참가 희망 설문 조사에서 풀었던 문항 적응용 샘플 문항에 대한 해설도 함께 제시하였다. 다음은 전자우편 안내문에 제시된 문항 설명문(문어/구어)이다.

주어진 [문장]에 대해서 [보기]로 주어진 내용의 말하는 이(화자) 혹은 글쓴이(필자)가 어느 정도 확신하는 것 같으나를 1~7의 정도로 평가해 주시는 방식입니다.

[보기] 철수는 내일 날씨가 흐릴 것으로 예상한다.

[문장] 내일 날씨가 흐릴 것이다

위 [보기]에서 "철수"는 주어입니다. "철수"가 판단한 것과 위 [보기]의 필자 및 화자의 견해가 반드시 같은 것은 아닙니다. 위 [문장]에 대한 필자 혹은 화자의 판단을 기준으로 평가해 주시면 됩니다.

또한 실험을 진행하시는 여러분들의 판단을 묻는 것이 아니니 이점도 참고 부탁드립니다.

아래는 마찬가지로 언어 실험 안내를 보낼 때 포함한 샘플 문항과 해설의 예시이다.

<p>[문맥] 귀사는 자율규제단체를 통한 개인정보보호 자율규제 정책이 행정안전부 고시 「개인정보보호 자율규제단체 지정 등에 관한 규정」을 통해 시행되고 있음을 알고 있습니까?</p>
<p>[질문] 다음 문장에 대해 필자는 어느 정도로 확신한다고 생각하십니까?</p> <p><i>자율규제단체를 통한 개인정보보호 자율규제 정책이 행정안전부 고시 「개인정보보호 자율규제단체 지정 등에 관한 규정」을 통해 시행되고 있다.</i></p>
<p>[해설]</p> <p>문항 1의 필자는 주어진 문장 "<u>자율규제단체를 통한...</u>"을 알고 있는지 질문하고 있습니다. 여기서 필자는 법률 규정이 실제 시행되고 있고, 이러한 사실을 알고 있는지 묻는 의도를 가지고 있습니다. 따라서 필자는 주어진 문장이 사실임을 확신한다고 볼 수 있습니다.</p>

언어 실험은 코로나 상황으로 인하여 비대면으로 실험을 진행하였다. 온라인 언어 실험 플랫폼은 피실험자의 편의와 친숙도를 고려하여 구글서베이(Google survey)를 활용하였다. 먼저 문항 풀이 이전에 연구참여설명서 및 연구참여동의서 등을 배치하였으며, 이후 실험에 대한 안내문을 배치하였다. 구어/문어 및 대화문 실험 안내문의 예시는 아래와 같다.

말뭉치 함의 분석을 위한 온라인 언어 실험 (202)

oneiric66@korea.ac.kr (공유되지 않음) 계정 전환

아래 설명문을 읽으시고 '다음' 버튼을 누르시면 실험이 시작됩니다.

[문항 안내]

- (1) 주어진 문장을 읽고 [보기]의 필자(혹은 화자)가 해당 문장을 어느 정도로 확신한다고 생각하는지 물어보는 문항입니다.
- (2) 1점에서 7점 사이를 선택하는 방식으로, 필자가 거짓이라고 매우 높게 확신한다고 생각한다면 1점, 참이라고 매우 높게 확신한다고 생각한다면 7점을 선택합니다.
- (3) 가급적 1점에서 7점 사이의 점수에서 다양한 숫자를 골고루 선택해 주시기 바랍니다.
- (4) 너무 깊이 고민하지 않고 직관적으로 드는 생각을 바로 바로 선택해 주시기 바랍니다. 너무 오랫동안 판단하시는 것이 오히려 좋지 않습니다.

뒤로 다음 양식 지우기

말뭉치 함의 분석을 위한 온라인 언어 실험 안내(401)

ku.expsyn@gmail.com (공유되지 않음) 계정 전환

아래 설명문을 읽으시고 '다음' 버튼을 누르시면 실험이 시작됩니다.

[문항 안내]

- (1) 화자들간의 대화를 읽고 특정 화자가 [질문] 아래 제시되는 [문장]의 내용과 관련해 어느 정도로 확신한다고 생각하는지 물어보는 문항입니다.
- (2) 1점에서 7점 사이를 선택하는 방식으로, 특정 화자가 [문장]의 내용이 거짓이라고 매우 높게 확신한다고 생각한다면 1점, 참이라고 매우 높게 확신한다고 생각한다면 7점을 선택합니다.
- (3) 너무 깊이 고민하지 않고 직관적으로 드는 생각을 바로 바로 선택해 주시기 바랍니다. 너무 오랫동안 판단 하시는 것이 오히려 좋지 않습니다.

실제 실험 환경의 예시는 다음과 같다. 내포문을 시각적으로 강조하는 등의 편집은 실험의 결과를 왜곡할 수 있기 때문에 행하지 않았다.

1/47

[보기] 롯데건설이 '롯데캐슬 3.0'이라는 디자인 모델을 선보였고 대우건설은 상반기 중으로 새로운 '푸르지오'를 내놓게 된다. LH(한국토지주택공사)는 상반기 중으로 아파트 브랜드를 선보이겠다고 이미 공언한 바다. 이번에 새 브랜드가 나오게 되면서 LH는 15년동안 다섯 번째의 브랜드가 나오게 된다.

[질문] 다음 문장에 대해 위 보기의 필자(화자)는 어느 정도로 확신한다고 생각하십니까?

[문장] LH는 상반기 중으로 새 아파트 브랜드를 선보일 것이다

위 [보기]를 읽고 [질문]에 답해주세요. (#202001) *

1 2 3 4 5 6 7

거짓이라고 확신한다

참이라고 확신한다

뒤로

다음

양식 지우기

[대화]

상만: 어, 그래. 은호원 씨는 마케팅팀으로 가면 되겠다.

호원: 아... 네...

우진: 안되겠는데요?

강호: 망했다.

기택: 왜? 왜?

강호: 치킨치킨..

기택: 헉!

우진: 은호원 씨는 영업팀에서 쓰시죠.

우진: 마케팅이 잘 참는 걸로 되는 것도 아니고, 갑질이 소원인데, 그럴 일도 아니구요. 게다가 무책임하기까지 하던데.

상만: 거 참. 서 부장.... 영업팀엔 그래도 남직원이 낫지!

우진: 성차별적인 위험한 발언이십니까?

[질문] 다음 [문장]의 내용에 대해 위 대화의 우진은 어느 정도로 확실하다고 생각하십니까?

[문장] 마케팅은 잘 참는 걸로 된다

위 [대화]를 읽고 [질문]에 답해주세요. (#401001) *

1 2 3 4 5 6 7

거짓이라고 확신한다

참이라고 확신한다

뒤로

다음

양식 지우기

실험은 리커트 7점 척도로 진행하였으며, [보기](혹은 [대화]) 문장에 대한 [질문]의 답으로 거짓으로 확신할수록 낮은 점수를, 참이라고 확신할수록 높은 점수를 부여하도록 하였다. 실험 답례비는 1회당 문화상품권 10,000원권이다(개인 전자우편 계정으로 문화

상품권의 PIN번호를 발송하였다).

4.1.4. 응답자 신뢰도 분석

본 실험은 1,805건의 담화에서 생성된 내포 명제에 대하여 각각 20명 내외의 확산성 평가 결과를 수집하였다. 추가적인 평가기준을 고려하여 최종적으로는 1,562건의 담화 사례를 제출하였다. 본 실험에서는 설문조사 또는 심리학에서 주로 사용되는 크리펜도르프의 알파(Krippendorff's alpha)가 아닌 급내상관계수(Intraclass Correlation Coefficient)를 응답자 신뢰도 분석에 사용하였다. 급내상관계수는 의학통계에 주로 사용되는, 척도로 임의의 집단 내부의 구성원들이 서로 얼마나 닮아있는지를 평가한다. 때문에, 본 실험에서는 평가 기준별로 여러 집단을 설정하여 급내상관계수를 계산하였다. 아래의 급내상관계수 도표는 1,805건의 문항에 대하여 20명의 응답자를 각각의 집단으로 설정하고 계산한 것이다.

유형	평가 방법	상관관계	F 통계량	자유도	95% 신뢰구간
ICC 1	단일 절대 평가	0.715	43.83	1804	[0.7 0.73]
ICC 2	단일 무작위 평가	0.715	44.11	1804	[0.7 0.73]
ICC 3	단일 고정 평가	0.717	44.11	1804	[0.7 0.73]
ICC 1k	평균 절대 평가	0.977	43.83	1804	[0.98 0.98]
ICC 2k	평균 무작위 평가	0.977	44.11	1804	[0.98 0.98]
ICC 3k	평균 고정 평가	0.977	44.11	1804	[0.98 0.98]

일반적으로, 집단 내의 평균을 구하여 급내상관계수를 계산하는 ‘평균’ 절대/무작위/고정 평가를 준용하고, ‘단일’ 절대/무작위/고정 평가는 참조 자료로 활용한다. 평균 평가 모두가 0.977로 나타나므로 응답자의 신뢰도가 매우 높은 것으로 나타난다. 일반적으로 0.8이 넘으면 매우 신뢰할만한 응답으로 이해할 수 있다.

또한, 본 실험에서는 전체 문항 1,805건의 라벨 분포를 잠정적으로 산출하였다. 내포 명제의 함의 관계를 자연어추론 평가세트로 가공한다고 가정하였을 때, 아래와 같이 라

벨 분포가 나타날 수 있다. 아래의 예시는 리커트 척도 점수가 [5-7] 구간에 있는 내포 명제는 ‘함의’로 보고, [3-5] 구간에 있는 내포 명제는 ‘중립’으로 보았다. 나머지는 ‘모순’으로 보았다.

내포 명제 함의 관계	문항 수	비율
함의	1,314	72.80
중립	208	11.52
모순	282	15.62
전체	1,805	100 (%)

아래의 급내상관계수 도표는 내포 명제의 함의 관계에 따라 집단을 형성하여 평가한 것이다. 구체적으로는, 내포 명제가 ‘중립’인 담화와 ‘함의’인 담화를 집단별로 비교한 것이다. 평균 평가를 기준으로 급내상관계수는 0.857 - 0.859로 나타났다. 이는 내포 명제의 함의 관계를 판단하는 응답자가 집단별로 매우 닮아있다는 점을 보여준다.

유형	평가 방법	상관관계	F 통계량	자유도	95% 신뢰구간
ICC 1	단일 절대 평가	0.262	7.03	1281	[0.24 0.28]
ICC 2	단일 무작위 평가	0.262	7.09	1281	[0.24 0.28]
ICC 3	단일 고정 평가	0.264	7.09	1281	[0.25 0.28]
ICC 1k	평균 절대 평가	0.857	7.03	1281	[0.85 0.87]
ICC 2k	평균 무작위 평가	0.858	7.09	1281	[0.85 0.87]
ICC 3k	평균 고정 평가	0.859	7.09	1281	[0.85 0.87]

<표 1> (유사) 보문소 목록

ㄴ_가능성	ㄴ_믿음	ㄴ_지_	다는_뜻	ㄹ_답	라"고_
ㄴ_가정	ㄴ_반성	ㄴ_질문	다는_말	ㄹ_뜻	라_
ㄴ_개념	ㄴ_분위기	ㄴ_처지	다는_비판	ㄹ_마음	라고_
ㄴ_거_	ㄴ_비판	ㄴ_태도	다는_사실	ㄹ_말	라고는_
ㄴ_거라고_	ㄴ_사건	ㄴ_해석	다는_생각	ㄹ_맘	라고도_
ㄴ_걸_	ㄴ_사례	ㄴ_형상	다는_소리	ㄹ_문제	라는_개념
ㄴ_것_	ㄴ_사실	ㄴ_형태	다는_소문	ㄹ_비판	라는_거
ㄴ_것도_	ㄴ_상태	ㄴ_희망	다는_소식	ㄹ_사건	라는_걸_
ㄴ_것만_	ㄴ_상황	ㄴ_가_	다는_얘기	ㄹ_사실	라는_것
ㄴ_것으로	ㄴ_생각	ㄴ_가는_	다는_의견	ㄹ_상황	라는_게_
ㄴ_것은_	ㄴ_설명	ㄴ_가를_	다는_의미	ㄹ_생각	라는_뜻
ㄴ_것을_	ㄴ_소리	ㄴ_지_	다는_의지	ㄹ_설명	라는_말
ㄴ_것이_	ㄴ_소문	ㄴ_지가_	다는_의혹	ㄹ_소리	라는_비판
ㄴ_것이라고_	ㄴ_소식	ㄴ_지는_	다는_이야기	ㄹ_얘기	라는_사실
ㄴ_게_	ㄴ_신념	ㄴ_지도_	다는_일	ㄹ_예상	라는_생각
ㄴ_견해	ㄴ_양상	ㄴ_지를_	다는_입장	ㄹ_예정	라는_얘기
ㄴ_결과	ㄴ_얘기	나_	다는_주장	ㄹ_예측	라는_의미
ㄴ_경우	ㄴ_어려움	나를_	다를_	ㄹ_의도	라를_
ㄴ_경향	ㄴ_예상	냐"고	ㄹ_가능성	ㄹ_의미	ㄹ_도_
ㄴ_경험	ㄴ_예측	냐_	ㄹ_가정	ㄹ_의심	ㄹ_만_
ㄴ_교훈	ㄴ_욕구	냐가_	ㄹ_개념	ㄹ_의지	ㄹ_으로_
ㄴ_권리	ㄴ_욕망	냐고_	ㄹ_거_	ㄹ_이야기	ㄹ_은_
ㄴ_기능	ㄴ_의견	냐는_	ㄹ_거라고_	ㄹ_일	ㄹ_을_
ㄴ_기대	ㄴ_의도	느냐_	ㄹ_걸_	ㄹ_입장	ㄹ_이_
ㄴ_기분	ㄴ_의문	니_	ㄹ_것_	ㄹ_자세	마_
ㄴ_기쁨	ㄴ_의미	니가_	ㄹ_것도_	ㄹ_전제	으면_
ㄴ_기억	ㄴ_의사	니는_	ㄹ_것만_	ㄹ_주장	자_
ㄴ_기회	ㄴ_의심	니도_	ㄹ_것으로	ㄹ_주제	자고_
ㄴ_논리	ㄴ_의지	니를_	ㄹ_것은_	ㄹ_줄_	
ㄴ_뉴스	ㄴ_의혹	다"고_	ㄹ_것을_	ㄹ_줄은_	
ㄴ_느낌	ㄴ_이야기	다_	ㄹ_것이_	ㄹ_줄을_	
ㄴ_답	ㄴ_일	다"고_	ㄹ_것이라고_	ㄹ_해석	
ㄴ_뜻	ㄴ_입장	다고_	ㄹ_게_	ㄹ_확률	
ㄴ_마음	ㄴ_자세	다고는_	ㄹ_경우	ㄹ_희망	

<표 2> 모문 술어 목록 (최종 결과물)

가정하다	나타나다	말해주다	상실하다	오해하다	조성되다
간절하다	나타내다	망설이다	생각(이)들다	옳다	좋다
감사하다	남다	맞다	생각되다	외치다	주다
감지하다	낫다	명백하다	생각하다	요구하다	주목되다
강조하다	낫다	모르다	생생하다	우기다	주목하다
강하다	낫아지다	모으다	서러워하다	우세하다	주장하다
갓다	내다	목격하다	설교하다	원하다	중요하다
같다	내다보다	물론이다	설명하다	위장하다	즐겁다
가우똥하다	내비치다	미루다	속상하다	유념하다	증대되다
거부한다	놀라다	믿다	속상해하다	유명하다	증명되다
걱정되다	놀랍다	바라다	숙여 보다	유명해지다	증명하다
걱정하다	높다	반증하다	쉽다	유추하다	지적하다
검토하다	놓치다	받다	스쳐지나가다	유행하다	지켜보다
겪다	눈치채다	발견하다	시사하다	응시하다	직감하다
결의하다	느껴지다	발생되다	신다	의미하다	짐작하다
경고하다	느끼다	발표되다	실감하다	이다	집계되다
경험하다	다르다	발표하다	실망하다	이야기하다	집중하다
고려하다	다짐하다	밝혀내다	싫다	이용하다	착각하다
고백하다	단언하다	밝혀주다	싫다	이해하다	체감하다
고소하다	단정하다	밝혀지다	쓰다	인식시키다	추궁하다
고요하다	달구다	밝히다	아니다	인식하다	칭하다
고집하다	담다	방지하다	아뢰다	인정하다	커지다
공개하다	답하다	배우다	안내하다	인지하다	크다
공언하다	대답하다	보게 되다	알게 하다	없다	터득하다
과시하다	덧붙이다	보내주다	알다	입장이다	통하다
관람하다	돌다	보다	알려주다	있다	틀림없다
광고하다	돌아보다	보여주다	알려지다	잇다	파다하다
교란시키다	동일하다	보여지다	알리다	잊어버리다	파악되다
구분하다	되다	보유하다	알아내다	자각하다	파악하다

굳히다	두렵다	보이다	알아두다	자백하다	판단하다
귀뜸해주다	뒷받침하다	부르다	알아주다	자부하다	편협하다
규명하다	드러나다	부르짖다	알아차리다	자아내다	편협하다
규정하다	드물다	부연하다	알아채다	자위하다	평가받다
그러다	듣다	부인하다	암시하다	잡다	표현하다
그럴싸하다	들다	부탁드리다	얘기하다	재고하다	피력하다
그렇다	들리다	분명하다	어려워지다	적히다	필요하다
기다리다	들어간다	분석되다	어렵다	전망되다	하다
기대받다	들어있다	분석하다	어이없다	전망하다	함축하다
기대하다	들통나다	불가능하다	언급되다	전제하다	해내다
기록하다	따지다	불편해지다	얼버무리다	전하다	해석하다
기억되다	떠들어대다	비치다	없다	전해지다	확신하다
기억하다	떠오르다	빌다	없어지다	절감하다	확인되다
기원하다	뚜렷하다	빠지다	여기다	점쳐지다	확인하다
깨달다	뜻하다	사로잡히다	여쭙보다	정의하다	횡횡하다
꾸미다	막다	살펴보다	연장하다	제기되다	희망하다
끌다	많다	살피다	열리다	제시하다	희박하다
나다	많아지다	상기시키다	염려스럽다	제언하다	
나무라다	말씀드리다	상기하다	예상되다	조명하다	
나오다	말하다	상상하다	예상하다	조사되다	

참고 문헌

- 강범모 (1983), 한국어 보문명사 구문의 의미 특성, *어학연구* 19-1, 서울대 어학연구월, 53-73.
- 남기심·고영근 (2001), 『표준국어문법론』, 박이정.
- 박진호 (2011), 시제, 상, 양태, *국어학* 60, 국어학회, 289-322.
- 이정민 (2018), 「‘알다’라는 말의 분석」, 『의미·통사 구조와 인지』, 한국문화사, 34-39.
- 이정민 (2020), 「태도 동사 ‘알다’의 사실성 교체 -한국어와 그 밖의 알타이 제어의 특성-」, 『국어학』 96, 국어학회, 3-58.
- 임동훈 (2008), 한국어의 서법과 양태 체계, *한국어 의미학* 26, 211-249.
- 임동훈 (2011), 한국어의 문장 유형과 용법, *국어학* 60, 323-359.
- 전영철 (2003), 한국어의 총칭표현들, *한국언어정보학회 프로시딩즈*, 13-24.
- de Marneffe, M.-C., Simons, M. & Tonhauser, J. (2019), The CommitmentBank: Investigating projection in naturally occurring discourse. *Proceedings of Sinn und Bedeutung* 23, 107-124.
- Han, C-H. (1996), Comparing English and Korean Counterfactuals: the Role of Verbal Morphology and Lexical Aspect in Counterfactual Interpretation, *The proceedings of ESCOL '96*, 1-15.
- Han, C-H. (2006), Variation in Form-Meaning Mapping between Korean and English Counterfactuals, *Journal of East Asian Linguistics* 15, 167-193.
- Horn, L. (2014), The cloud of unknowing. In J. Hoeksema & D. Gilbers (eds.). *Black Book, A Festschrift in Honor of Frans Zwarts*, 178-196.
- Iatridou, S. (2000), The Grammatical Ingredients of Counterfactuality, *Linguistic Inquiry* 31-2, 231-270.

- Jiang, N., & de Marneffe, M. C. (2019). Evaluating BERT for natural language inference: A case study on the CommitmentBank. *In Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP)*, 6086-6091.
- Karttunen, L. (1971), Implicative verbs. *Language*, 47(2), 340-358.
- Karttunen, L. (2013), You will be lucky to break even. In Tracy Holloway King & Valeria de Paiva (eds.). *Quirky Case to Representing Space: Papers in Honor of Annie Zaenen*, 167-180. Stanford, CA: CSLI publications.
- Kiparsky K. & P. Kiparsky(1971), Fact In D. Steinberg and L. Jakobovits(eds.) *Semantics: An Interdisciplinary Reader in Philosophy Linguistics and Psychology*, Cambridge University Press, 345-69.
- Lee, C. (2019a), Factivity alternation of attitude 'know' in Korean, Mongolian, Uyghur, Manchu, Azeri, etc. and content clausal nominals. *Journal of Cognitive Science*, 20(4), 449-508.
- Lee, C. (2019b), Content nominals with and without factive presupposition, *MS. SNU*.
- Wang, X., Kapanipathi, P., Musa, R., Yu, M., Talamadupula, K., Abdelaziz, I., Chang, M., Fokoue, A., Makni, B., Mattei, N. & Witbrock, M. (2019, July). Improving natural language inference using external knowledge in the science questions domain. *In Proceedings of the AAAI Conference on Artificial Intelligence*, 33, 1, 7208-7215.
- Yum, Y., Lee, J. M., Jang, M. J., Kim, Y., Kim, J. H., Kim, S., ... & Joo, H. J. (2021). A Word Pair Dataset for Semantic Similarity and Relatedness in Korean Medical Vocabulary: Reference Development and Validation. *JMIR Medical Informatics*, 9(6), e29667.

<Abstract>

2021 Research and Analysis of Korean CommitmentBank

The present research project aimed to build a high-quality corpus of Korean speaker's attitude or (non-)commitment about the (non)factivity of given events, e.g., *John eats lobster*. For example, a Korean speaker may show firm commitment about the declarative sentence *chelswuka papul mekessta*, but the same speaker may show uncertainty about the negative sentence *chelswuka papul mekessnunci molukeyssta*. Although the categorical distinction of speaker (non-)commitment appears to be straightforward, in reality, speaker (non-)commitment forms a gradient pattern due to the linguistic complexity of comprehending the (non)factivity.

We extended the previous works on the speaker (non-)commitment by conducting large-scale linguistic judgments on Korean discourses. In addition, the current project also aimed to build annotation guidelines for tagging the linguistic factors contributing the speaker (non-)commitment about the (non)factivity (i.e., truthfulness of the utterance). We first automatically gathered 165,225 sentences or discourses containing Korean (pseudo-)complementizers. These unique lexical items combine the embedded clause and the main predicate, such that resultant constructions show the speaker (non-)commitment about the (non-)factivity. To extract the discourses, we used Korean Language Institute Parsed Corpora and Sejong Morphological Analysis Corpora. Then, (assistant)researchers manually inspected the subset of 165,225 sentences using online workbench. Importantly, our project workers excluded the irrelevant discourses and left sentences that

potentially featured the linguistic complexity of Korean speaker (non-)commitment.

These items were examined by a group of linguists, generating more than 1,805 instances of speaker commitment (960 instances of written language, 303 instances of newspapers, 421 instances of spoken language, 121 instances of conversations). To gather the relevant linguistic intuition about the discourses, we conducted 7-point-Likert scale linguistic judgment experiments recruiting 60 Korean native speakers. These participants were equally divided into three groups, and each group assessed sets of 40 instances of discourses. Note that responses of one or two participants are missing due to the invalid answers and system errors. Importantly, we specified the step-by-step protocols for the replication of our language experiment and the guidelines for constructing the sentence stimuli.

At the last stage of the construction of Korean CommitmentBank, linguistic information like person, tense, and complementizer in the sentence stimuli is tagged by a group of linguists. The final version of our Korean CommitmentBank consists of 1,562 instances of Korean speaker (non-)commitment. This corpus shows how linguistic components of discourses contribute the comprehension of (non)factivity of events. Notably, annotation guidelines are specified with concrete examples for the future extension of our project.

As a benchmark computational analysis, we recasted Korean CommitmentBank for natural language understanding. The analysis tests whether language AI captures the Korean speaker (non-)commitment. This is motivated by the recent attempts to test the robustness of generalizable AI. The result shows that language AI called KR-BERT captures the linguistic intuition of Korean language speakers, but there are still room for improvements. We expect that future research may use our Korean commitment corpus for the

validation of model's robustness and generalization capacity.

Keywords: Korean speaker commitment, corpus annotation,
Large-scale veridicality judgement, Natural language understanding

<기획·연구>

국립국어원 이승재(언어정보과장)

국립국어원 김소희(언어정보과 학예연구사)

국립국어원 서근화(언어정보과 연구원)

<연구 참여자>

연구 책임자 송상헌

공동 연구원 강아름, 김경민, 김태우, 이도길, 장하연,

최윤지, 김일환, 박진호, 정연주, 정슬아,

조경찬, 유현조, 조정아

발행인: 국립국어원장

발행처: 국립국어원

서울시 강서구 금남화로 154

전화 02-2669-9775, 전송 02-2669-9727

인쇄일: 2021년 12월 17일

발행일: 2021년 12월 17일

인 쇄: 현대문화사 02-922-3376

※ 이 책은 국립국어원의 용역비로 수행한 ‘2021년 말뭉치 함의 분석 및 연구’ 사업의
결과물을 발간한 것입니다.