



저작자표시-비영리-동일조건변경허락 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.
- 이차적 저작물을 작성할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



동일조건변경허락. 귀하가 이 저작물을 개작, 변형 또는 가공했을 경우에는, 이 저작물과 동일한 이용허락조건하에서만 배포할 수 있습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

교육학박사학위논문

한국어 학습자의
학술적 데이터 설명문 쓰기 연구

2023년 8월

서울대학교 대학원
국어교육과 한국어 교육전공
이 지 연

한국어 학습자의 학술적 데이터 설명문 쓰기 연구

지도교수 김 호 정

이 논문을 교육학 박사학위 논문으로 제출함.

2023년 5월

서울대학교 대학원
국어교육과 한국어 교육전공
이 지 연

이지연의 교육학 박사학위 논문을 인준함.

2023년 7월

위 원 장 _____ (인)

부위원장 _____ (인)

위 원 _____ (인)

위 원 _____ (인)

위 원 _____ (인)

국문초록

데이터 설명문은 객관성을 드러내고자 하는 학술적 텍스트에서 필수적으로 사용된다. 수치를 나타내고 개념을 연결하는 시각적 데이터가 언어적 기호와 융합하여 필자의 주장을 객관적으로 뒷받침하는 중심적인 역할을 하기 때문이다. 따라서 한국어로 학술적 글쓰기를 본격적으로 수행하는 외국인 대학생에게 데이터 설명문의 교수-학습은 매우 중요하다. 그러나 데이터 설명문은 이제까지 학술적 텍스트 작성과 관련해 연구가 활발하지 않았다. 그 이유는 데이터 설명문이 상위 장르에 내포되어 그에 따라 양식성이 변화하는 비고정적인 속성을 갖고 있기 때문이다.

이에 본고에서는 이러한 가변적인 장르 내포적 텍스트 탐구의 출발점으로 학부생 외국인 학습자가 작성하는 학술적 데이터 설명문이라는 사용역을 설정하였다. 또한 이 사용역 내에서 나타나는 학술적 데이터 설명문의 담화 구조와 언어 사용을 알아보려고 하였다. 본고에서는 이러한 분석을 바탕으로 데이터 설명문의 교수-학습을 위해 상위 장르의 맥락을 중심으로 학습자에게 적절한 비계와 어휘문법적 자원을 제공할 수 있는 자료 기반 학습(data-driven learning)의 교수-학습 내용을 구안할 필요가 있다고 제안하였다.

본고에서는 데이터 설명문의 수집을 위해 한국어로 데이터 설명문 작성 요구가 높은 경제학 및 경영학 전공 대학생을 대상으로 표와 그래프로 구성된 '전공 수업 보고서' 쓰기 과제를 제시하였다. 분석 결과는 외국인 대학생을 기준으로 언어권별, 한국어 숙달도별로 나누어 살펴보았으며, 한국인 대학생들과의 공통점과 차이점도 함께 제시하였다.

수집한 데이터 설명문은 담화 구조와 언어 사용을 기준으로 분석하였다. 담화 구조의 경우 데이터 설명문의 순서와 더불어 메타 기능적, 인지적 의미를 함께 고려하여 이동 마디를 구축하였다. 구체적으로는 데이터 기술의 방식에 체계 기능적 다중 양식 담화 분석을 통한 메타 기능적 의미를 부착하고, CEFR, IELTS, TOPIK의 데이터 설명문 평가 기준에서

확인할 수 있는 인지적 능력을 활용해 이동 마디를 설정하였다. 그 결과 학습자의 데이터 설명문 이동 마디는 크게 ‘A. 주제 및 방향 요약, B. 수치/결과 나열, C. 해석 및 결론’으로 도출되었다. 언어 사용의 경우는 이렇게 설정한 이동 마디별 말뭉치의 3-gram을 통하여 학습자 집단별 공통점과 차이점을 알아보았다.

담화 구조의 집단별 비교는 언어권별(중국어:베트남어) 및 한국어 숙달도별(중급:고급)로 ‘상위 이동 마디 순서 유형 비교’ 및 ‘이동 마디 단계 빈도의 집단별 평균 차이 검증(t-test)’을 통해 수행하였다. 상위 이동 마디 순서 유형은 어떠한 형태로 이동 마디를 배열하여 데이터 설명문을 구성하는지 알아보려는 것이고, 이동 마디 단계 빈도는 각 이동 마디 단계가 지닌 메타 기능적 속성과 인지적 부하를 학습자가 집단별로 얼마나 인식하고 사용하는지 알아보려는 것이다.

이동 마디 순서 유형은 중국어권 학습자가 AB와 같은 C 누락형을 유의하게 많이 사용하였으며, 베트남어권 학습자는 ABABC와 같은 ABC 유형의 변이형을 다수 사용하였음이 밝혀졌다. 또한 중급 학습자는 AB, 고급 학습자는 ABC 유형에 가장 높은 선호도를 나타냈다. 이동 마디 빈도는 A 이동 마디의 경우, 베트남어권 학습자가 ‘(소)제목을 활용한 주제 제시(A1)’, ‘연구의 주제와 내용 및 방법 설명(A2)’을 유의하게 많이 사용하였고, 고급 학습자의 빈도 평균도 A2에서 중급 학습자에 비해 유의하게 높게 나타났다. B 이동 마디에서는 베트남어 학습자에 비해 중국어권 학습자의 ‘시각적 데이터 순서대로 나열(B1-2)’과 ‘제일 큰/작은 수치에 주목(B2-4)’ 이동 마디 사용 빈도가 유의하게 높았다. 반면 베트남어권 학습자는 중국어권 학습자에 비해 ‘수치적 비교(B2-2)’, ‘시각적 데이터 언급 표지 사용(B3)’, ‘수치의 의미 제시(B4)’를 유의하게 사용하는 것으로 나타났다. 고급 학습자는 B3 이동 마디가, 중급 학습자는 ‘연대기적 나열(B1-1)’ 이동 마디가 상대 집단에 비해 유의하게 많이 나타났다. ‘주어진 데이터/상식적 지식에 근거한 해석(C1)’에서는 베트남어권 학습자, 고급 학습자들이 각각 중국어권 학습자, 중급 학습자에 비해 유의하게 빈도가 높았다. 다만 ‘데이터의 비판점과 개선 방향(C2)’은 네 집단

모두 저조하게 사용하였고, 사용 빈도 차이 또한 유의하지 않았다. ‘문제 해결 방안/ 후속 연구/ 미래 방향 제시(C3)’에서는 고급 학습자의 선택 빈도가 유의하게 높았다.

언어 사용 양상은 각 이동 마디별 말뭉치를 형태 태깅 기준 3-gram으로 분석하였다. 그 결과를 바탕으로 2회 이상 중복된 어휘문법적 요소를 추출하였고, 이에 따라 각 이동 마디별 어휘문법 목록을 작성하였다. 전반적으로 중급 및 중국어권 학습자의 어휘문법 사용이 단조롭거나 다양하지 못하며, 수가 적은 것으로 나타났다. 학습자 숙달도별로 보았을 때, 중급 학습자는 학술적 장르의 특성 파악이 어려워 ‘-버니다’체 등의 장르에 맞지 않는 문체를 생성하거나, 고급 학습자에 비해 제한적인 어휘 선택이 나타나는 것을 확인할 수 있었다. 학습자 언어권별 전이도 몇 가지 확인할 수 있지만, 이를 교정할 수 있는 한국어 상위 장르가 명시적으로 제시되지 못해 모어에서 전이된 어휘를 고수하고 있는 것으로 보인다.

학습자들이 내포 장르로서 데이터 설명문을 구성하기 위해서는 상위 장르에 따른 담화 구조를 인식하고 맥락적인 언어 사용의 발달을 꾀해야 한다. 이에 본고에서는 특정 상위 장르에 내포된 데이터 설명문의 담화 구조와 언어 사용을 학습자가 파악하여 변용할 수 있는 자료 기반 학습(data-driven learning, DDL)을 제안하였다.

이상에서 살펴본 바와 같이 본고는 이제까지 연구되지 않았던 학술적 데이터 설명문의 구성을 학습자 작문을 통해 파악하고, 그 특성을 토대로 다른 상위 장르에서 데이터 설명문을 작성할 수 있도록 교수-학습의 방향을 제안하였다는 데 의의가 있다.

주요어 : 데이터 설명문, 장르 내포, 체계 기능적 다중 양식 담화 분석, 이동 마디, n-gram, 자료 기반 학습

학 번 : 2016-30523

목차

| | |
|--|----|
| I. 서론 | 1 |
| 1. 연구의 목적과 필요성 | 1 |
| 2. 선행 연구 | 7 |
| 2.1. 장르 연구의 대상 변화 | 7 |
| 2.2. 학술적 텍스트에서의 데이터 설명문 연구 | 11 |
| 3. 연구 방법과 연구 대상 및 자료 | 16 |
| 3.1. 연구 방법 | 16 |
| (1) 담화 구조 분석 | 16 |
| (2) 언어 사용 분석 | 21 |
| 3.2. 연구 대상 및 자료 | 27 |
| | |
| II. 학술적 데이터 설명문 분석을 위한 이론적 고찰 | 33 |
| 1. 학술적 데이터 설명문의 개념과 범위 | 33 |
| 2. 체계 기능적 속성에 따른 이동 마디 구성 | 35 |
| 2.1. 학술적 데이터 설명문의 체계 기능적 다중 양식 담화 분석 .. | 35 |
| (1) 학술적 데이터 설명문의 다중 양식적 속성 | 35 |
| (2) 학술적 데이터 설명문의 메타 기능적 의미 | 40 |
| 2.2. 기존 데이터 설명문 이동 마디에서의 메타 기능적 의미 | 50 |
| 3. 정보 매개적 속성에 따른 인지적 및 언어적 구성 | 61 |

| | |
|--|-----|
| 3.1. 학술적 데이터 설명문의 매개적 속성 | 61 |
| 3.2. 매개적 속성으로서의 인지적 및 언어적 구성 요소 | 63 |
| 4. 학술적 데이터 설명문의 분석 프레임 | 70 |
| 4.1. 학술적 데이터 설명문의 구조 | 70 |
| 4.2. 학술적 데이터 설명문의 어휘문법 | 78 |
| | |
| Ⅲ. 한국어 학습자의 학술적 데이터 설명문 양상 분석 | 81 |
| 1. 학습자의 학술적 데이터 설명문의 담화 구성 | 81 |
| 1.1. 이동 마디 순서 유형과 빈도 | 81 |
| (1) 외국인 대학생과 한국인 대학生の 비교 | 82 |
| (2) 학습자 변인별 이동 마디 순서 유형과 특성 | 91 |
| 1) 학습자 언어권별 이동 마디 순서 유형과 특성 | 91 |
| 2) 한국어 숙달도별 이동 마디 순서 유형과 특성 | 103 |
| 1.2. 이동 마디 구성 요소별 특징 | 109 |
| (1) 이동 마디 A: 주제 및 방향 요약 | 110 |
| 1) 외국인/한국인 대학生の 이동 마디 단계 분포 및 빈도 | 110 |
| 2) 학습자 언어권별 이동 마디 단계 분포 및 빈도 | 113 |
| 3) 학습자 숙달도별 이동 마디 단계 분포 및 빈도 | 115 |
| (2) 이동 마디 B: 수치/결과 나열 | 116 |
| 1) 외국인/한국인 대학生の 이동 마디 단계 분포 및 빈도 | 116 |
| 2) 학습자 언어권별 이동 마디 및 단계 분포 및 빈도 | 123 |
| 3) 학습자 숙달도별 이동 마디 및 단계 분포 및 빈도 | 127 |

| | |
|--|-----|
| (3) 이동 마디 C: 해석 및 결론 | 133 |
| 1) 외국인/한국인 대학생의 이동 마디 단계 분포 및 빈도 | 133 |
| 2) 학습자 언어권별 이동 마디 및 단계 분포 및 빈도 | 135 |
| 3) 학습자 숙달도별 이동 마디 및 단계 분포 및 빈도 | 137 |
| 2. 학습자 학술적 데이터 설명문의 어휘문법 사용 | 140 |
| 2.1. 주제 및 방향 요약(A) | 140 |
| (1) (소)제목을 활용한 주제 제시(A1) | 140 |
| (2) 연구 주제와 내용 및 방법 설명(A2) | 147 |
| (3) 데이터 현상에 대한 배경 설명(A3) | 151 |
| 2.2. 수치/결과 나열(B) | 156 |
| (1) 구체적인 정보 추출(B1) | 156 |
| (2) 의도적 그룹화(B2) | 163 |
| (3) 시각적 데이터 언급 표지 사용(B3) | 177 |
| (4) 수치의 의미 제시(B4) | 182 |
| 2.3. 해석 및 결론(C) | 188 |
| (1) 주어진 데이터/상식적 지식에 근거한 해석(C1) | 188 |
| (2) 데이터의 비판점과 개선 방향(C2) | 193 |
| (3) 문제 해결 방안/후속 연구/미래 방향 제시(C3) | 195 |

IV. 학습자 변인별 학술적 데이터 설명문 양상에 대한 논의와 교육적 제언 199

| | |
|--|-----|
| 1. 학술적 데이터 설명문 담화 구성에서의 학습자 변인별 특성 | 199 |
| 1.1. 학습자 변인별 이동 마디 순서 유형의 특성 | 199 |

| | |
|---|-----|
| 1.2. 학습자 변인별 이동 마디 사용 빈도의 특성 | 208 |
| (1) 주제 및 방향 요약에서 언어권별 차이의 해석 | 208 |
| (2) 수치/결과 나열에서의 인지적 차이 확인 | 210 |
| (3) 해석 및 결론에서의 숙달도 요인 부각 | 215 |
| 2. 학술적 데이터 설명문 언어 사용에서의 학습자 변인별 특성 | 216 |
| 2.1. 상위 장르의 어휘문법 레퍼토리 | 216 |
| 2.2. 기학습된 어휘문법의 전이와 강화 | 218 |
| 2.3. 어휘 선택에서의 다양성과 차이 | 220 |
| 3. 학습자의 학술적 데이터 설명문 양상에 따른 교수-학습 제언 | 222 |
| 3.1. 학습자의 학술적 데이터 설명문 양상에 따른 쓰기 교수-학습의 필요성 | 222 |
| 3.2. 학술적 데이터 설명문 교수-학습을 위한 원리와 그 적용 | 225 |
| (1) 데이터 설명문의 장르 내포적 성격에 따른 교육 원리 | 225 |
| (2) 자료 기반 학습의 방법과 데이터 설명문에의 적용 | 227 |
| | |
| V. 결론 | 232 |
| | |
| 참 고 문 헌 | 240 |
| | |
| [부록] | 250 |
| | |
| Abstract | 253 |

표 목 차

| | |
|--|----|
| <표 1> 연구 절차와 내용 | 6 |
| <표 2> Nordrum & Eriksson(2015)이 변용한 Biber et al.(2007)의 하향식 이동 마디 분석 일부 | 16 |
| <표 3> 학습자 데이터 설명문 분석의 예 | 19 |
| <표 4> Biber et al.(2007)의 하향식 이동 마디 분석 중 미시 분석에의 이용 단계 | 21 |
| <표 5> n-gram 추출 비교 | 25 |
| <표 6> 한국인 대학생 참여자 정보 | 29 |
| <표 7> 외국인 대학생 참여자 정보 | 29 |
| <표 8> 경제/경영 전공 대학생 작문 말뭉치 자료 표시 | 30 |
| <표 9> 참여자 말뭉치 정보 | 31 |
| <표 10> 학술적 텍스트에서 시각적 데이터의 체계 기능적 의미 | 47 |
| <표 11> Swales & Feak(1994)의 데이터 설명문 이동 마디 | 51 |
| <표 12> 데이터 설명문의 텍스트적 메타 기능 중 이동 마디(Guinda, 2011:130) | 53 |
| <표 13> Nordrum & Eriksson(2015:70)의 데이터 설명문 이동 마디와 단계 | 55 |
| <표 14> 경제학 및 경영학 학술지의 데이터 설명문 이동 마디(이지연, 2023:162-163) | 59 |
| <표 15> CEFR(2020) ‘데이터 설명하기’가 포함하는 인지적/언어적 능력 | 64 |
| <표 16> IELTS ‘Writing task 1’이 제시하는 인지적/언어적 능력 | 67 |
| <표 17> TOPIK 53-54번 평가 기준 | 68 |

| | |
|---|-----|
| <표 18> 학습자 데이터 설명문의 이동 마디와 표지 부착 | 73 |
| <표 19> 이동 마디 단계를 적용한 분석의 예 | 77 |
| <표 20> 외국인 대학생과 한국인 대학생의 이동 마디 순서 유형 비교 | 82 |
| <표 21> 외국인/한국인 대학생의 52개 이동 마디 순서 유형 중 사용 개수와 비율 | 84 |
| <표 22> 외국인/한국인 대학생이 선택한 상위 5개 이동 마디 순서 유형 | 84 |
| <표 23> ABCBC 유형 데이터 설명문의 예 | 86 |
| <표 24> ABABC 유형 데이터 설명문의 예 | 87 |
| <표 25> ABCBCBC 유형 데이터 설명문의 예 | 88 |
| <표 26> 외국인/한국인 대학생의 ‘A/B-시작’ 이동 마디 선호도 | 90 |
| <표 27> 언어권별 이동 마디 순서 빈도 | 92 |
| <표 28> 중국어권/베트남어권 학습자의 32개 이동 마디 순서 유형 중 사용 개수와 비율 | 93 |
| <표 29> 중국어권/베트남어권 학습자가 선택한 상위 4~5개 이동 마디 순서 유형 | 94 |
| <표 30> 중국어권/베트남어권 학습자의 이동 마디 순서 유형 비교 .. | 95 |
| <표 31> AB 유형 데이터 설명문의 예 | 95 |
| <표 32> ABABAB 유형 데이터 설명문의 예 | 96 |
| <표 33> B 유형 데이터 설명문의 예 | 98 |
| <표 34> ABABABC 유형 데이터 설명문의 예 | 100 |
| <표 35> ABABCBC 유형 데이터 설명문의 예 | 101 |
| <표 36> 한국어 숙달도별 이동 마디 순서 빈도 | 103 |
| <표 37> 중급/고급 학습자의 34개 이동 마디 순서 유형 중 사용 개수와 | |

| | |
|--|-----|
| 비율 | 104 |
| <표 38> 중급/고급 학습자가 선택한 상위 3~4개 이동 마디 순서 유형 | 105 |
| <표 39> 외국인/한국인 대학생의 A 이동 마디 분포 | 110 |
| <표 40> 한국인/외국인 대학생의 A 이동 마디 빈도 평균 t-test 결과 · | 111 |
| <표 41> 중국어권/베트남어권 학습자의 A 이동 마디 분포 | 113 |
| <표 42> 중국어권/베트남어권 학습자의 A 이동 마디 빈도 평균 t-test 결과 | 114 |
| <표 43> 중급/고급 학습자의 A 이동 마디 분포 | 115 |
| <표 44> 중급/고급 학습자의 A 이동 마디 빈도 평균 t-test 결과 | 115 |
| <표 45> 외국인/한국인 대학생의 B 이동 마디 분포 | 117 |
| <표 46> 외국인/한국인 대학생의 B 이동 마디 선택 빈도 평균 t-test 결과 | 119 |
| <표 47> 중국어권/베트남어권 학습자의 B 이동 마디 분포 | 123 |
| <표 48> 중국어권/베트남어권 학습자의 B 이동 마디 선택 빈도 평균 t-test 결과 | 125 |
| <표 49> 중급/고급 학습자의 B 이동 마디 분포 | 127 |
| <표 50> 중급/고급 학습자의 B 이동 마디 선택 빈도 평균 t-test 결과 · | 128 |
| <표 51> 중급 학습자 B2-계열 사용 양상 | 129 |
| <표 52> 외국인/한국인 대학생의 C 이동 마디 분포 | 133 |
| <표 53> 한국인/외국인 대학생의 C 이동 마디 선택 빈도 평균 t-test 결과 | 134 |
| <표 54> 중국어권/베트남어권 학습자의 C 이동 마디 분포 | 135 |

| | |
|---|-----|
| <표 55> 중국어권/베트남어권 학습자의 C 이동 마디 선택 빈도 평균 t-test 결과 | 136 |
| <표 56> 중급/고급 학습자의 C 이동 마디 분포 | 137 |
| <표 57> 중국어권/베트남어권 학습자의 C 이동 마디 선택 빈도 평균 t-test 결과 | 137 |
| <표 58> A1 언어권별 어휘문법 목록 | 143 |
| <표 59> A1 숙달도별 어휘문법 목록 | 143 |
| <표 60> 표절 의심 작문과 그 대상 | 146 |
| <표 61> A2 언어권별 어휘문법 목록 | 148 |
| <표 62> A2 숙달도별 어휘문법 목록 | 149 |
| <표 63> A3 언어권별 어휘문법 목록 | 153 |
| <표 64> A3 숙달도별 어휘문법 목록 | 153 |
| <표 65> B1 언어권별 어휘문법 목록 | 158 |
| <표 66> B1 숙달도별 어휘문법 목록 | 160 |
| <표 67> B2 언어권별 어휘문법 목록 | 165 |
| <표 68> B2 숙달도별 어휘문법 목록 | 169 |
| <표 69> B3 언어권별 어휘문법 목록 | 179 |
| <표 70> B3 숙달도별 어휘문법 목록 | 180 |
| <표 71> B4 언어권별 어휘문법 목록 | 183 |
| <표 72> B4 숙달도별 어휘문법 목록 | 185 |
| <표 73> C1 언어권별 어휘문법 목록 | 190 |
| <표 74> C1 숙달도별 어휘문법 목록 | 191 |
| <표 75> C2 언어권별 어휘문법 목록 | 193 |

| | |
|------------------------------|-----|
| <표 76> C2 숙달도별 어휘문법 목록 | 193 |
| <표 77> C3 언어권별 어휘문법 목록 | 196 |
| <표 78> C3 숙달도별 어휘문법 목록 | 196 |

그림 목차

| | |
|---|-----|
| <그림 1> UAM corpus tool 6.2e를 통한 분석의 예 | 18 |
| <그림 2> 학술적 데이터 설명문 연구의 비언어-언어 스펙트럼상 위치 · | 38 |
| <그림 3> 학술적 데이터 설명문 구조 형성 원리 | 72 |
| <그림 4> 학술적 데이터 설명문의 언어 요소 | 79 |
| <그림 5> TOPIKⅡ 제60회 53번 문제 | 107 |
| <그림 6> 세부 학습자군별 이동 마디 순서 빈도 | 202 |
| <그림 7> 2022학년도 성균관대학교 수시모집 논술우수전형 논술 시험 기출 문제(성균관대학교 입학처, 2023.6.20. 검색) | 206 |
| <그림 8> 세부 학습자군별 A 이동 마디 빈도 | 209 |
| <그림 9> 세부 학습자군별 B1 이동 마디 빈도 | 211 |
| <그림 10> 세부 학습자군별 B2 이동 마디 빈도 | 212 |
| <그림 11> 세부 학습자군별 B3, B4 이동 마디 빈도 | 214 |
| <그림 12> 세부 학습자군별 C 이동 마디 빈도 | 216 |
| <그림 13> 국립국어원 한국어 학습자 말뭉치에서 나타난 ‘데이터’ 어휘 콘코던스 | 228 |

I. 서론

1. 연구의 목적과 필요성

한국어 학습자는 한국어 숙달도가 높아지며 점차 직업, 학문 등의 특수 목적 한국어 학습에 접근한다. 데이터 설명문은 학습자가 직업 분야로 나아갈 때도, 학문 분야로 진출할 때도, 또는 특수 목적이 아닌 일반적으로 한국어를 사용하여 뉴스나 신문을 볼 때도 자주 마주하게 되는 유형의 텍스트이다. 또한 데이터 설명문은 데이터 리터러시의 관점에서 데이터를 소통하는 능력을 대변한다고 볼 수 있다. 송유경(2021)은 데이터 리터러시를 “단순히 데이터를 분석하는 것뿐만 아니라 데이터를 가지고 효과적인 의사소통을 하는 능력을 포함”한다고 하였다. 이러한 관점에서 데이터 설명문 역시 데이터를 활용한 의사소통의 한 방식이라 할 수 있다.¹⁾

그러나 이같이 데이터 활용과 소통이 중요해지고 있음에도 데이터 설명문의 교육적 의의는 한국어 교육 내에서 적극적으로 고려되지 않고 있다. 시각적 데이터와 그를 설명하는 언어적 서술로 구성된 ‘데이터 설명문’은 그동안 언어교육사에서 관심을 끄는 주요 연구 분야가 아니었다. 전통적인 장르로 구분되지 않아 그 특수성을 파악하기가 어려웠을 뿐더러, 문어 텍스트뿐만이 아닌 시각적 수치 데이터를 포함하는 형식적 이질성 역시 연구자들이 접근하기 어려운 분야였으리라 생각된다. 그러나 데이터 설명문은 학술적 텍스트 내에서도 그가 갖는 고유한 형식이 있는 바, 외국인 대학생²⁾들이 그것을 인식하여 상황맥락에 적합하게 구성해

1) 본고에서 사용하는 ‘데이터(data)’는 특정한 정보를 내포하는 수치들의 모임을 뜻하며, 이들이 표, 그래프 등으로 표현된 것을 ‘시각적 데이터’로 부른다. 이는 뒤에 등장할 ‘자료 기반 학습(data-driven learning)’과의 용어 사용에서 차이를 두고자 함인데, 자료 기반 학습에서의 data는 말뭉치를 뜻하고, 관련 교육학적 연구에서 이를 이미 ‘자료’로 번역하여 사용하고 있기 때문이다.

2) 본고에서는 제2언어로서 한국어 학술적 데이터 설명문 교육이 필요한 외국인 대학생을 ‘학습자’로 칭한다. 다만 데이터 설명문 분석을 위해 비교 집단으로 모집한 ‘한국인 대학생’과 비교 및 대조가 필요할 때 ‘학습자’와 동일한 대상을 ‘외국인 대학생’으로도 칭

내는 것은 어려운 일임을 알 수 있다.

또한 외국인 학습자가 작성 시 어려움을 겪는다는 것 외에도 한국어 교육이 데이터 설명문에 주목해야 하는 이유는 다음과 같이 세 가지로 정리할 수 있다.

첫째, 대학에서 요구하는 학술적 텍스트는 매개적이며 담화 통합적 성격을 갖는다. 즉, 학술적 텍스트는 다른 텍스트의 글 또는, 표, 그림, 수식, 그래프 등의 자료를 인용하거나 재진술하여 자신의 주장을 강화하는 특성이 있다. 서로 다른 두 기호로 동일한 의미와 내용을 담도록 전달하는 것은 매개(mediation, CEFR, 2020)라는 개념으로 설명될 수 있다. 학술 환경에서의 매개는 기존 한국어 교육 연구들이 언어 기호로만 구성된 텍스트의 재진술 및 인용을 ‘담화 통합’이라는 주제로 다룬다(이선영, 2017; 송은정, 2018; 김현민, 2020). 그러나 시각적 기호로 표상된 매체가 필자의 학술적 텍스트에 새로이 서술되는 것에 대한 본격적인 연구는 한국어 교육에서 찾아보기 어렵다. 실제 학술적 텍스트에서의 담화 통합은 시각적 기호의 통합이 빈번함에도 그러하다.

둘째, 데이터 설명문은 언어적 자료와 시각적 데이터가 함께 어우러져 의미를 전달하는 ‘다중 양식 조화(multimodal ensemble)’를 추구한다. 교과서, 수업 과제 보고서, 실험 보고서, 학술 논문, 학위 논문과 같이 학술적 환경에서 사용되는 다양한 텍스트에서는 그래프, 표, 그림 등의 시각적 데이터를 풍부하게 사용하는 경향이 있다. 이들을 글과 어떻게 어울리게 하여 배치할 것인가, 어떻게 글로 표현할 것인가는 학술적 의사소통에 지대한 영향을 끼친다. 이에 대해 Nordrum & Eriksson(2015:61)에서는 “효과적인 다중 양식 조화는 작문에서 시각적 데이터가 얼마나 많이 언급되어야 하고 변환되어야 하는지를 저자가 잘 알고 선택하는 것에 달려 있다”고 하였다. 이러한 다중 양식 조화는 목표 언어의 비모어 필자는 물론이고 모어 필자라 하더라도 미숙한 필자는 이를 인식하고 적용하기 어렵다. 그럼에도 학습자 변인에 따른 다중 양식 조화 양상에 대해서는 아직 연구되지 않았다.

한다.

셋째, 데이터 설명문은 TOEFL, IELTS, TOPIK 등 언어 시험에서 학술성을 평가하는 요소로서 자주 제시된다. 이는 데이터 설명문이 분명히 학술적 텍스트 구성의 하나로 중요한 역할을 수행함을 뜻한다. 다만 이것이 단일 장르가 아닌 다양한 학술적 상위 장르에 내포된 미니 장르(mini-genre, Lemke, 1998:95) 중 하나이기 때문에 이를 단독으로 떼어 담화 관습을 장르 분석적으로 파악하였을 때 그 결과는 상위 장르에 따라 매우 다양하게 나타난다.³⁾ 그간의 학술적 텍스트 연구가 전문가의 모범 작문을 상정하고 그와 학습 필자 작문을 비교하는 장르 중심적 방식으로 주로 이루어졌던 것을 생각하면 데이터 설명문의 이러한 가변적 속성은 연구가 어려운 것으로 기피되었을 법하다. 그러나 이로 인해 학습자들이 텍스트 구성에 어려움을 겪게 된다면 이에 대한 연구 또한 필수적으로 이루어져야 할 것이다.

특히 본고에서는 학습자의 데이터 설명문 구성 방식을 활용하여 다양한 상위 장르에서 데이터 설명문을 구성할 수 있도록 방향을 제시하고자 한다. ‘학위 논문’, ‘실험 보고서’와 같은 상위 장르는 학습자가 이미 고정된 양식을 장르 중심으로 학습하여 동일한 장르를 재생산할 수 있다는 장점이 있다. 그러나 이러한 기존의 장르 중심 방식으로 데이터 설명문을 교수-학습하려면 대단히 많은 상위 장르의 데이터 설명문을 모두 추출하고 분석해야 하는 번거로움이 있다.

이러한 점들을 고려하여 본고에서는 대학생 학습자가 학술적 장에서 데이터 설명문을 어떻게 구성하는지를 분석하여 파악하고, 이를 원형으로 삼아 다양한 상위 장르에서 데이터 설명문을 변용할 수 있도록 하는 자료 기반 학습(data-driven learning)이 유용하다고 보았다. 즉, 학습자가 스스로 다양한 상위 장르에 따른 차이를 변별하고 자신의 데이터 설명문에 반영할 수 있는 능력을 기르는 것이 궁극적으로 내포 장르의 확장에 필요한 교육 방법이라 보았다.

위에서 언급한 바와 같이 학술적 데이터 설명문은 상위 장르와 분과의

3) 예를 들어, II장에서 확인할 수 있는 Swales & Feak(1994)의 학술적 데이터 설명문의 보편적 이동 마디는 Nordrum & Eriksson(2015)에서 다룬 응용 화학 분야 논문의 데이터 설명문 이동 마디와 매우 다르다.

영향을 많이 받으므로 이를 통제하여 학습자의 전공 요인은 ‘경제학 및 경영학’으로, 학위 과정은 ‘학부’로 한정한다. 경제학 및 경영학으로 학문 분과를 제한한 까닭은 해당 분과가 데이터 설명문이 다수 사용되는 경학문(soft science) 분야이기 때문이다. 이공계로 불리는 중학문(hard science) 분과는 본디 학문 목적 영어 분야 데이터 설명문 연구의 주 대상이었다. 그러나 한국의 중학문 분과에서 지배적으로 사용되는 언어는 영어이기 때문에 이공계열 텍스트를 한국어로 쓰고자 하는 외국인 학습자의 수요가 상대적으로 매우 적다(이슬비, 2016:35). 따라서 외국인 대학생이 한국어로 데이터 설명문을 쓰고자 하는 요구는 경학문에 집중하는 경향이 있다. 본고에서는 특히 그중에서도 수치적, 시각적 데이터를 다수 사용하는 경제학 및 경영학 전공 외국인 대학생들에게 데이터 설명문 교육이 유용할 것이라 판단하였다. 이와 같은 이유로 본고에서는 경제학 및 경영학으로 전공 요인을 통제하고, 학습자들의 다른 변인인 모어, 한국어 숙달도에 따라 학술적 데이터 설명문의 양상이 어떠한 차이를 보이는지 확인할 것이다.

구체적으로, 본고에서 학습자들에게 작성하도록 요청한 학술적 데이터 설명문은 수업 시간에 제출하는 과제로 상정하고, 표 1개와 그래프 2개로 구성하였다([부록] 참고). 이들을 주제에 따라 자유롭게 배치하여 하나의 완결된 글을 작성하는 것이다. 이러한 과제는 학부 과정에 재학 중인 학습자들에게 요구되는 가장 현실적인 학술적 데이터 설명문이라는 데서 학습자 텍스트의 실제성을 확보할 수 있다.

이렇게 수집한 학술적 데이터 설명문에서는 담화 구성을 이동 마디로 두고, 언어 사용을 이동 마디 내 어휘문법으로 두어 분석하도록 한다. 다음 장에서 자세히 설명할 것이지만, 상위 장르에 속하는 텍스트의 이동 마디는 ‘이 다음에 어떠한 내용이 나온다’와 같은 순서로 충분히 제시할 수 있다. 그러나 상위 장르에 내포되어 나타나는 데이터 설명문은 이러한 텍스트 진행 순서보다는 데이터 서술 양상을 상세하게 포착하여 이동 마디를 나눌 필요가 있다. 따라서 이동 마디를 구성할 때 그 순서를 데이터 설명문의 체계 기능 다중 양식 담화 분석을 위한 메타 기능 요소인

‘관념적’, ‘대인적’, ‘텍스트적’ 기능의 측면에서 분석하여 이동 마디로 제시하였다. 이는 학습자가 시각적 데이터의 어떠한 부분을 선택하는지, 그 선택한 부분의 표현 방법과 정도는 어떠한지를 알아보기 위한 것이다. 또한 n-gram 분석을 통해 이동 마디별 어휘문법적 사용 양상을 확인하였다.

본고에서는 이러한 방식으로 데이터 설명문의 구성 요소를 분석하여 학습자의 언어권별, 한국어 숙달도별 변인에 따라 그 특성을 알아내고자 한다. 이렇게 알아낸 학습자별 변인에 따른 특성을 바탕으로 한 교육 방법으로 자료 기반 학습(data-driven learning)의 유용성을 제시한다. 이상을 정리하여 연구 문제를 설정하면 다음과 같다.

1. 한국어 학습자의 학술적 데이터 설명문은 학습자 변인에 따라 어떠한 구조 및 언어적 특징을 나타내는가?

1.1. 한국어 학습자의 학술적 데이터 설명문에서 담화 구성은 학습자 언어권 및 숙달도 변인에 따라 어떠한 공통점 및 차이점을 나타내는가?

1.2. 한국어 학습자의 학술적 데이터 설명문에서 언어 사용은 학습자 언어권 및 숙달도 변인에 따라 어떠한 공통점 및 차이점을 나타내는가?

2. 한국어 학습자의 학술적 데이터 설명문 분석 결과를 토대로 도출할 수 있는 학습자 변인별 교육적 시사점은 무엇인가?

이를 위해 각 장별로 연구 절차와 방법을 정리하면 다음과 같다. 우선 I 장에서는 데이터 설명문 교육의 연구사적 위치를 조망한다. 이는 학술적 장르 쓰기에 대한 선행 연구와 데이터 설명문의 연구사를 확인하여 본고의 연구사적 위치를 제시하고자 하기 위함이다. 그다음으로는 연구 방향에 맞는 연구 방법을 상술하도록 한다. 특히 연구 참여자인 대학생들이 작성한 데이터 설명문의 분석을 위한 연구 방법을 제시한다.

II 장에서는 문헌 연구를 통해 데이터 설명문이라는 텍스트 유형에 대해 정의하고 그 특징을 다중 양식, 이동 마디, 매개(mediation)의 측면에서

알아보도록 한다. 그 과정에서 학술적 데이터 설명문 양상을 구체적으로 파악할 기준을 체계 기능적 다중 양식 이론 및 정보 매개적 인지 능력 관점에서 재편한 이동 마디와 그 단계에 해당하는 어휘문법으로 세분화한다.

Ⅲ장에서는 언어권별로는 중국어권 학습자와 베트남어권 학습자, 숙달도별로는 중급 학습자와 고급 학습자로 학습자군을 나누어 외국인 대학생의 데이터 설명문에 나타나는 이동 마디 및 어휘문법의 특성을 알아본다. 구체적으로는 이동 마디 순서 유형, 이동 마디 단계 선택 빈도, 이동 마디별 어휘문법의 차이를 알아볼 것이다. 이 과정에서 한국인 대학생의 데이터 설명문도 비교를 위해 함께 분석한다. 이를 위해 모든 집단 대학생의 데이터 설명문을 말뭉치로 만들어 이동 마디를 코딩한다. 그 다음으로 이동 마디별 품사 태깅을 통해 n-gram을 추출한다.

Ⅳ장에서는 앞에서 분석한 학습자 데이터 설명문의 양상을 중급 중국어권, 고급 중국어권, 중급 베트남어권, 고급 베트남어권 학습자군으로 세분화하여 정리하고 해석한다. 이들 분석의 결과는 데이터 설명문 교육이 다양한 상위 장르의 데이터 설명문으로 확장과 변용을 견인해야 한다는 점에서 자료 기반 학습의 활용에 대한 시사점을 제시한다.

V장은 연구의 의의와 후속 연구를 통한 발전 방향을 제시한다. 이를 정리한 전체 논문의 구성과 연구의 내용은 다음과 같다.

<표 1> 연구 절차와 내용

| 장 | 연구 절차 | 연구 내용 및 방법 |
|------|--------------------------------|--|
| I 장 | 연구 목적과 연구 문제 연구 방법 및 계획 | 문헌 연구 |
| II 장 | 데이터 설명문 양상 분석을 위한 이론적 탐색 | 문헌 연구 1) 체계 기능적 다중 양식 담화 분석 2) 정보 매개의 인지적 접근 |

| | | |
|------|-------------------------------------|---|
| III장 | 학습자 데이터 설명문의 양상 | <p style="text-align: center;">실험 연구 및 양적 분석</p> <p>1) 학습자 변인별 이동 마디 순서 및 하위 이동 마디 선호 빈도 분석</p> <p>2) 학습자 변인별 이동 마디에 속한 n-gram 분석</p> |
| IV장 | 학습자 데이터 설명문 해석과 교육적 적용에 대한 제언 | <p style="text-align: center;">실험 결과 해석 및 교육 방안에 대한 시사점</p> <p>1) 학습자 변인별 데이터 설명문 양상 해석</p> <p>2) 자료 기반 교육 활용에 대한 시사점 도출</p> |
| V장 | 결론 | |

2. 선행 연구

2.1. 장르 연구의 대상 변화

텍스트 분석에서 장르에 대한 연구는 일반적으로 신수사학과, 체계 기능 언어학, 특수 목적 언어 교육의 세 가지로 나뉜다. 신수사학과의 장르는 특정 사회 집단의 관심과 가치를 제도적, 역사적 맥락에서 통합하고 개인을 위한 특정 사회적 역할과 필자-독자 간의 관계를 강화하고자 한다. 즉, 전형화된 수사적 행위가 어떠한 집단에게 인정받으며, 그는 누구와 무엇에 이득이 되는지, 이면적 맥락에는 무엇이 있는지 찾아내고자 한다(Hyland, 2004/이수미·이소연 역, 2019:74-75). 체계 기능 언어학 역시 신수사학과에서와 마찬가지로 장르를 사회적 가치의 재생산 관점에서 바라본다. 체계 기능 언어학에서는 장르를 사회적 목적을 이루기 위해 진화되어 온 수사적 구조로 파악하고 있다(Hyland, 2004/이수미·이소연 역, 2019:58). 따라서 체계 기능 언어학 기반의 작문 교육에서는 아직 장르의 의도를 모르는 학생들에게 이를 파악할 수 있는 문식성을 길러주기

위해 장르의 유형과 그에 수반하는 언어적 자질을 명시적으로 전달한다. 이들 두 견해에서 장르는 텍스트로 표상되는 사회적 목적에 관심을 두고 있다. 이들 입장에서, 장르가 갖는 수사적 고정성은 특정 사회적 힘을 재생산하기 때문에 이에 숨겨진 의도와 의미를 인식하고 파악하는 것이 중요하다.

이와는 달리 특수 목적 언어 교육의 경우에는 보다 세밀하게 한정된 맥락 속에 존재하는 텍스트를 장르로 선정하여 다룬다(Bawarshi & Reiff, 2010/정희모 외 역, 2015:82-83). 그 목적이 특정 맥락과 관습에 고정된 수사를 새로운 사회에 진입한 초심자들이 익혀 의사소통을 할 수 있도록 하는 데 있기 때문이다. 따라서 특수 목적 언어 교육은 텍스트를 이루는 구조와 언어에의 사고보다는 그 양식의 훈련에 집중한다. 제2언어/외국어 학습자를 대상으로 한 장르 연구가 주로 특수 목적 언어 교육에서 이루어진 것은 사회적으로 더 깊은 사고를 위해서는 그 사고의 발전을 위한 이해와 표현이 제2언어/외국어 자원을 우선적으로 요하기 때문이다.

따라서 특수 목적 언어 교육 패러다임하의 제2언어 작문 교육 연구는 주로 모어 전문 필자가 생성한 전범 텍스트와 제2언어 학습 필자가 작성한 텍스트를 비교하여 모어 전문 필자의 그것을 따라갈 수 있는 교수-학습적 논의로 이루어져 있다.⁴⁾ 따라서 특수 목적 언어 교육의 장르 분류는 학문 목적 맥락이라면 교재, 학위 논문, 학술지 논문 초록, 보고서 등으로 실제 사용하는 텍스트의 이름인 경우가 많다.⁵⁾ 즉, 특수 목적 언어 교육의 대상이 되는 텍스트는 전문 필자가 작성한 완성된 '순수한' 양식을 지닌 텍스트가 대부분이다.

4) 이 세 분야에서 다루는 장르에 대한 더 자세한 비교는 Hyland(2004/이수미·이소연 역, 2019), Bawarshi & Reiff(2010/정희모 외 역, 2015), 유민애(2017:72)등을 참고하라.

5) 이는 앞에 다룬 두 이론과는 차이가 나는 지점이다. 신수사학에서의 장르는 언어 참여자의 사회적 위치 및 관계에 따라 나뉜다. 예를 들어, 교실에서 교수자의 장르 집합은 강의계획서, 과제, 수업계획, 공지사항 등이나 학습자의 장르 집합은 필기, 교사에게 보내는 이메일, 시험 답안 등으로 제시된다(Bawarshi & Reiff, 2010/정희모 외 역, 2015:144). 체계 기능 언어학에서의 장르는 서사, 경험적 이야기, 논쟁, 설득과 같은 더 기초적인 분류를 제시한다. 즉, 구체적인 텍스트의 이름을 특정하지 않는다. 텍스트의 양식성, 구조 등이 갖는 공통적 '사회적 힘과 의미'를 찾아내는 것이 목적이기 때문이다.

그러나 이러한 순수한 장르의 개념은 장르의 혼합이 일어난다 하더라도 이를 배제하며 형성된 경우가 많다. Bhatia(1997:191)가 언급한 바대로 장르의 내포나 혼합을 배제하는 것은 장르가 갖는 특성을 본문에 한정하기 때문에 특수 목적 언어 교육에서 교육적 결과를 만들어내는 데 유용하다. 그러나 실제 텍스트에서는 장르가 혼합되고(genre-mixing), 한 장르가 다른 장르에 포함되는 장르 내포(genre-embedding)가 빈번하게 일어나기 때문에(Bhatia, 1997:191) 순수한 장르의 양식성만을 대상으로 텍스트를 연구하는 것은 실제성이 떨어진다.

그럼에도 그간 학문 목적 한국어의 장르 분석에서는 이러한 순수성에 천착하여 텍스트 본문을 형성하는 거시 구조 또는 미시 구조를 분석하는데 초점을 맞추었다. 한국어 학위 논문의 서론이 Swales(1990)의 서론 이동 마디 분석 모형인 CARS와 다름을 밝힌 박은선(2006), 학위 논문의 '선행 연구' 부분의 이동 마디를 다루는 김민혜(2014), 박수연(2015), 학위 논문의 결과와 결론을 분석한 박지희(2016), 조성해(2022) 등에서는 학술적 텍스트 본문의 거시 구조를 이동 마디로 제시하고 있다.

또한 담화 표지, 정형 표현 등의 미시적 언어 사용을 토대로 학술적 텍스트의 특성을 확인하려는 연구들도 있다. 홍혜란 외(2020)에서는 학술 논문, 전공 개론서, 신문, 일반 교양 도서 네 장르의 말뭉치를 수집하여 장르별 메타 담화 표지를 제시하고자 하였다. 그 과정에서 말뭉치에는 초록, 참고문헌, 도표와 그래프를 삭제하였다고 하였다. 위의 조성해(2022)에서는 또한 학위 논문 결론의 응집성 표지를 확인하였다. 윤새롬(2019)에서는 학술 논문 100편의 서론에서 예문, 각주, 페이지 번호를 제외한 나머지를 대상으로 하여 이동 마디별 다단어 단위(multi-word unit)⁶⁾를 추출하였다. 최지희(2022)에서도 인문, 사회, 예술, 자연과학, 공학 다섯 분야 학술 논문에서 메타 정보, 주석, 도표, 그림, 참고 문헌, 초록을 삭제한 본문 내용만을 대상으로 n-gram 분석을 통한 학술적 정형 표현을 제시하고 있다.

6) 윤새롬(2019)에서는 다단어 단위를 모든 종류의 단어 연속체로 정의하고 있다. 특히 '반드시 -아/아야 하다'와 같은 불연속적인 구성도 다단어 단위에 포함하고 있다.

이들 학술적 장르에 대한 연구는 모두 본문에 해당하는 텍스트 구조와 언어 사용을 대상으로 하며, 실제적으로 학술적 텍스트를 구성하는 다양한 요소에 상대적으로 덜 주목하는 경향이 있다고 할 수 있다. 그러나 본고에서는 학술적 텍스트를 구성하는 다양한 요소 역시 학술적 텍스트를 작성해야 하는 학습자에게 학습되어야 할 중요한 요소라 본다. 따라서 이동 마디와 언어 사용에 대한 연구 역시 본문 이외의 요소에도 적용되어야 한다고 보았다.

학술적 텍스트의 이러한 덜 주목되는 다양한 요소를 Lemke(1998:95)에서는 미니 장르(mini-genre)라 칭한 바 있다. 이는 본문에 덧붙여 텍스트를 구성하는 도표, 표, 캡션(caption), 참고문헌, 주석, 감사의 말 등이 속한다. 한국어 교육에서도 이러한 학술적 텍스트에서의 미니 장르에 대한 인식이 증대되고 있다. 학술적 논문의 각주를 연구 대상으로 삼고 있는 박나리(2013:229)에서는 미니 장르를 ‘하위텍스트(sub-text)’⁷⁾로 명명하고, 이들을 학술적 텍스트를 총체적으로 규명하기 위해 연구되어야 할 것으로 언급하고 있다. 이와 더불어 이윤진(2016)에서는 외국인 유학생의 학술적 텍스트의 제목 작성에 대해 초점을 맞추고, 제목의 종결 형식, 부적절한 표현 사용, 정보량과 핵심어 선택 적정성, 부호 사용 문제의 쟁점을 제시한 바 있다. 그러나 이들 미니 장르 중 하나인 ‘도표, 표’와 관련된 논의는 본고의 선제적 논의인 이지연(2023)에서 다룬 것이 처음이다.

도표와 표를 언어 기호로 변환하고 언어 기호 사이에 배치하는 행위를 인식하는 것에 대해 O'Halloran, Lemke 등은 ‘텍스트에 대한 다중 양식적 접근’으로 설명하였다. 시각적 데이터 및 데이터를 언급하고 그를 분석하는 언어적 기호는 그 자체로서 특수한 용법을 지닌다. Lemke(1998:109)는 과학 논문을 구성하는 요소들이 의미를 형성하는 데는 언어적 의미 요소와 더불어 시각 제시 요소, 수학적 요소, 기술적-실행적(감각적) 요소 등의 관습적 배치 역시 유용함을 주장하였다. 시각적 데이터를 사용하는 학술적 텍스트라면 과학 분과가 아니라 하더라도 이

7) 이에는 목차, 초록, 각주, 참고문헌 등이 속한다.

러한 시각적 데이터의 관습적 배치가 갖는 효과성을 인지하고 있을 것이다. 따라서 본고의 ‘데이터 설명문’이 주목하는 데이터의 시각적 제시는 학술적 장르의 효과를 만들어내는 내포 장르의 역할로 볼 수 있다. 즉, 내포 장르 연구는 이제까지 연구해 왔던 다른 상위 장르들과 같이 학술적 장르 연구의 하나이며, 학문 목적 언어 교육의 범위 내에 있는 요소라 할 수 있다.⁸⁾ 따라서 이전 본문 중심 장르 연구들의 방법론을 따르되, 텍스트의 범위를 데이터 설명문에 국한하여 이동 마디와 n-gram을 통한 정형 표현을 파악할 예정이다.

2.2. 학술적 텍스트에서의 데이터 설명문 연구

데이터 설명문이 갖는 특수한 언어 사용에도 불구하고 이제까지 연구가 많이 되지 않은 이유는 이들이 다양한 상위 장르에 내포되어 그 장르에 맞도록 변화하는 특성이 ‘순수한 장르’로 간주되지 않았기 때문이다. 학술적 데이터 설명문의 구조를 연구한 Guinda(2011:117)의 언급처럼 데이터 설명문은 “단독으로는 잘 나타나지 않으며 실험 보고서, 교재, 연구 논문 등의 거대 장르 내 한 부분으로 나타나는 경우가 많다.” 그럼에도 불구하고 데이터 설명문은 사람들이 명확히 인식할 수 있는 특성을 갖고 있고, 연구 논문, 교과서, 시험, 강의, 수업 토론, 발표, 기업 회의, 주식 거래 상황, 의학적 진단, 정치적 논의 등 어디에나 존재하기에 연구의 필요성이 있다(Guinda, 2011:116).

이제까지 발표된 몇 안 되는 학술적 데이터 설명문 연구는 데이터 설명문을 ‘정보 전달(information transfer, Widdowson, 1973)’, ‘그래프 묘사/쓰기(Graph description/writing, Valentina, 2016; Yang, 2012)’, ‘언어 평가 과제 이름(Valentina, 2016; 이명화, 2017; 최혜민, 2017; 김금숙, 2019; 이호인, 2020)⁹⁾, ‘구어/문어로 데이터 설명하기 활동(유럽 공통 참조 기준, Common European Framework of Reference, 2020)’, ‘데이터 설명문

8) 다만 내포 장르라는 점 때문에 상위 장르 내의 특수한 사용역이라 볼 수 있을 것이다.

9) 예를 들어, IELTS의 ‘Academic writing task 1’으로 불리거나 ‘한국어 능력시험Ⅱ 쓰기 53번 문제’로 제시된다.

(data commentary, Swales & Feak, 1994; Guinda, 2011; Nordrum & Eriksson, 2015; 이지연, 2023)', '시각적 설득(visual persuasion, Miller, 1998)' 등 고정되지 않은 여러 이름으로 부른다. 이들 연구를 크게 둘로 나누면 언어 과제 및 평가 연구와 학술적 텍스트에서의 시각성 연구 맥락으로 나눌 수 있다.

데이터 설명문은 Widdowson(1973)이 시각적 데이터에 처음 주목하여 '정보 전달(information transfer)'이라는 개념을 도입한 이래 언어 교육 내에서 인식되기 시작하였다. 정보 전달은 주로 시각적으로 제시되는 수식, 다이어그램, 그래프, 표 등의 비언어적 자료를 언어적으로 제시하는 과제 또는 활동으로 설명된다(Swales & Feak, 1995:64-66). Widdowson(1979:73)은 "비언어적 표현 정보인 다이어그램, 표, 그래프 등을 기반으로 언어적 양식을 연습한다면, 비언어적 장치는 언어적 설명을 촉진하는 데 쓰일 수 있다"고 하였다. 이에 따라 이후 영어 교육 분야에서는 비언어적 표현 정보를 이용한 언어 교육 과제 및 평가 연구가 진행되었다(Bartolic, 1981; Johns & Davies, 1983).

이러한 흐름에 따라 언어 평가 영역에서도 시각적 데이터가 등장하는 언어 평가 항목에 대한 연구가 진행되었다. 우선 그래프 쓰기(graph writing) 시험을 보는 학습자에 대한 연구는 Yang(2012;2016)과 Valentina(2016)가 있다. Yang(2012)에서는 그래프 쓰기 과제 중 학습자의 전략과 과제 해결 중의 어려움을 연구하였다. 이 연구에서는 학습자들이 그래프 쓰기 중 그래프 이해, 그래프 해석, 그래프 번역 전략을 사용함을 밝혀내었다. 또한 그래프 쓰기 중 어려움으로는 어휘적 지식과 그 사용, 그래프 친숙성, 주제 지식, 시험 요령 전략 등이 있음도 확인하였다. Yang(2016)은 이를 발전시켜 학습자의 그래프 친숙성, 영어 쓰기 능력, 주제 지식이 그래프 쓰기 시험 능력과 어떠한 관련이 있는지를 고찰하였다. 그 결과 그래프 친숙성을 제외한 나머지 두 변수는 테스트 수행에 긍정적 영향을 끼침을 발견하였다. Valentina(2016)는 그래프 쓰기인 IELTS의 Academic writing 1 과제를 성공적으로 수행하는 데 필요한 네 가지 요소를 실험 연구를 통해 밝혀내었다. 말하기/듣기/읽기/쓰기

기능의 연합 및 전이, 비판적 사고와 쓰기, 적합한 문체의 사용, 그래픽 문식성이 그것이다. 이들 연구는 시각적 데이터를 제시하는 과제 역시 학습자의 일반적 쓰기 능력이 주요함을 확인하는 것으로 결론을 내리고 있다.

한국어 교육에서도 이와 같이 언어 평가에서의 데이터 설명문 사용을 다룬 연구가 있다. 이명화(2017)에서는 한국어 학습자의 등급에 따른 한국어 능력시험Ⅱ의 53번 그래프 설명문 오류를 정리한 바 있다. 그러나 이 연구에서는 조사 사용 오류가 많다는 장르 일반적 결론을 도출하며 그래프 쓰기의 특수성을 부각하지 못하고 있다. 최혜민(2017)에서는 한국어 능력시험Ⅱ 53번 그래프 쓰기 문항의 평가 기준이 54번 자기 주장 글쓰기 유형과 동일한 것을 비판하며, ‘자료 해석 행위’가 그래프 쓰기 수행을 평가하는 기준으로 반영되어야 함을 주장하였다. 이는 데이터 설명문 쓰기가 다른 쓰기 유형과 차이가 있음을 인식한 연구라 할 수 있다. 또한 한강우(2022)에서도 53번 문항 답안의 구조와 평가가 54번 주장하는 글쓰기와는 달리 두괄식으로 제시되는 것이 바람직하다고 주장한다. 김금숙(2019)에서는 한국어 능력시험 53번 문항을 ‘담화 양식, 주제와 지시문, 쓰기 동기화 자료, 평가 기준’의 요소로 나누어 과제 개선 방향을 제시하였다. 이 논문에서는 시각적 데이터를 단순히 언어 평가를 위한 ‘쓰기 동기화 자료’로 인식하고, 평가에 적합한 글의 구성과 형태를 고민하고 있다. 이효인(2019)은 한국어 능력시험Ⅱ 53번 그래프 쓰기의 텍스트 구조를 나열, 대응, 인과가 결합한 것으로 보고, 이들의 구조적 표지를 교수하여야 함을 주장한다. 이 논문은 데이터 설명문 자체의 구조를 인식하고 그를 구체적으로 드러낼 수 있는 표지를 제시했다는 데서 의의가 있다. 다만 시험용 데이터 설명문을 작성하기 위한 교육 방법이 실제 학술적 데이터 설명문 작성에도 적용될 수 있을지를 고민하게 한다. 유럽 공통 참조 기준(이하 CEFR, 2020)에서는 매개(mediation) 활동의 하나로 ‘구어/문어로 데이터 설명하기’를 들고 있다. 여기에서는 평가 등급에 따라 달성해야 하는 인지적, 언어적 수준이 달라짐을 들고 있다. 특히 등급이 낮을 때는 일반적 분야의 일기도, 약도 등의 시각적 데이터

를 해석하여 표현할 것을 요구하고, 등급이 올라갈수록 특수한 분야의 전문적인 데이터 해석을 요구한다. 언어적 수준이 전문 분야의 인지적 지식과 함께 발전해 감을 시사한다는 점에서 CEFR의 평가 기술은 의의가 있다고 하겠다.

이들 평가 관련 연구에서 시각적 데이터에 대한 연구가 따로 존재하는 것은 시각적 데이터 사용이 학술적으로 중요한 의미를 지니고 있다는 방증이라 할 수 있다. 다만 평가 관련 연구에서의 시각적 데이터는 언어 사용을 촉진할 수 있는 도구로서 작용하며, 실제 목적의 데이터 설명문의 구성과 언어 사용이 어떠해야 한다는 데까지는 논의하지 않고 있다.

이와 달리 시각적 데이터를 단순히 언어적 구조와 형태 연습을 위한 도구가 아닌 구분되는 하나의 텍스트로 인식하고 접근한 것은 Swales & Feak(1994)이 처음이다. Swales & Feak(1994)은 대학원생을 위한 학술적 글쓰기 교재에 ‘데이터 설명문(data commentary)’을 하나의 장으로 구분해 그 이동 마디(move) 구성과 어휘문법적 특징을 제시하였다. 이에 더해 Swales & Feak(1995)에서는 네 명의 비영어권 대학원생을 대상으로 사례 연구를 진행해 바람직한 데이터 설명문의 구조에 대해서도 논한다. 즉, 데이터 설명문이라는 텍스트의 성격이 비언어적인 자료를 사용하기에 전형적으로 선택적이고 해석적인 코멘터리의 형태를 취함을 언급한다(Swales & Feak, 1995:67). Guinda(2011)에서도 이와 같은 정의를 받아들여 마드리드 대학 항공우주공학 전공 대학생들의 영어 데이터 설명문을 분석한 바 있다. 이 논문에서의 데이터 설명문 과제는 한 문단의 데이터 설명문을 쓰는 짧은 것이기 때문에 언어 사용에 주목하였다. 특히 주제의 변환을 알리는 구조를 you, we와 같은 메타 담화(meta-discourse)로 파악하고, dramatic/drastic과 같은 추세 변화의 어휘가 주로 무엇과 결합하는지를 다루었다.

데이터 설명문의 언어적 형식과 언어 교육적 접근에서 더 나아가 데이터 설명문이 갖는 시각적 효과가 텍스트와 어우러져 텍스트의 의미를 구성하는 데 어떠한 역할을 하는지에 주목하는 연구도 있다. 이는 주로 학술적 발표와 문서에 나타나는 다중 양식성(multimodality)에 주목하여 체

계 기능적 관점에서 연구되기 시작하였다. Miller(1998)에서는 과학 학술지 논문에서 시각적 데이터가 신뢰성과 설득력을 담보한다고 하였다. 여기에서는 메시지와 시각적 데이터의 관계를 알아보기 위해 체계 기능 언어학에서 사용하는 세 가지 메타 기능을 사용하여 학술지 및 유명 보도매체의 시각적 데이터와 그에 부연된 언어적 사용을 비교 분석하고자 하였다. Paltridge(2013:352-353)에서는 도표, 차트, 다이어그램과 같은 시각화 요소가 학문적 장르에서 어떻게 발전되어 왔는지 설명하고 있다. 이러한 다중 양식적 분석은 체계 기능 연구와 결합하여 더욱 상세한 분석의 틀을 제시하고 있다. Guo(2004)는 대학 생물학 교재의 관습적 담화를 파악하기 위해 O'Halloran(2008a)이 제시한 체계 기능-다중 양식 담화 분석(Systemic functional-multimodal discourse analysis, SF-MDA)의 분석 방법을 사용한다. Nordrum & Erikson(2015)은 응용 화학 분야 연구 논문과 학위 논문의 데이터 설명문 분석을 위해 Biber et al.(2007)의 이동 마디(move) 분석을 적용하며 체계 기능적 이론을 적용했음을 제시한다. 말뭉치를 활용해 언어 기호의 구조적, 어휘문법적 접근을 중요시하는 기존의 이동 마디 분석과는 달리, 이 연구에서는 시각적 데이터의 위치와 개수, 텍스트에서 언급한 이미지의 종류 등을 함께 고려함으로써 다중 양식적 이동 마디 분석을 수행하였다. 이를 따라 이지연(2023)에서도 체계 기능적 다중 양식 담화 분석을 이동 마디 구성에 적용하여 경제학 및 경영학 학술지의 구조를 살펴본 바 있다.

이를 종합해 보면 학문 목적 언어 교육에서 데이터 설명문에 대한 접근은 시각적 데이터의 언어적 구조 및 의미를 해설하는 방법과 효과를 연구하는 것임을 알 수 있다. 그 과정에서 대다수의 연구가 체계 기능 이론의 메타 기능(metafunction)을 차용하는 것도 특징이라 볼 수 있다.

3. 연구 방법과 연구 대상 및 자료

3.1. 연구 방법

(1) 담화 구조 분석

본고에서는 학습자의 데이터 설명문 분석을 위해 이동 마디 중심의 담화 구조 분석(거시 분석)과 해당 이동 마디별 언어 분석(미시 분석)을 수행한다. 우선 담화 구조 분석에서는 Biber et al.(2007)의 하향식 이동 마디 분석과 근거 이론의 방법 중 일부인 반복적 비교 분석을 아울러 사용하도록 한다.

Nordrum & Eriksson(2015)은 학술지 데이터 설명문 분석을 위해 Biber et al.(2007:13)의 하향식 이동 마디 분석 방법을 일부 변형하여 제시하고 있다. Biber et al.(2007)의 하향식 이동 마디 분석 방법은 원래 한 상위 장르 전체의 이동 마디를 추출하기 위해 고안되었기에 그 단계가 일곱 가지에 해당하는데, Nordrum & Eriksson(2015)에서는 전체 학술지 논문 중 데이터 설명문만을 집중하여 다루므로 이 단계를 세 가지로 축소하여 다음과 같이 차용하였다. 그리고 데이터 설명문의 추출이라는 단계를 하나 더 추가하였다.¹⁰⁾

<표 2> Nordrum & Eriksson(2015)이 변용한 Biber et al(2007)의 하향식 이동 마디 분석 일부

| | 단계 | 이 접근에서 얻게 되는 것 |
|---|-------------------------|---|
| X | 데이터 설명문 추출하기 | 석사 논문/학술 논문의 결론 및 논의 부분에서 데이터 설명문이 어디에서 시작되고 끝나는지 결정한다. |
| 1 | 의사소통적/기능적 분류(수사적 이동 마디) | 데이터 설명문의 지엽적 맥락에서 각 텍스트 분절의 수사적 기능을 결정한다. |

10) 본고에서 나머지 분석 단계를 고려하지 않은 까닭은 데이터 설명문이 장르 그 자체가 아니라 학술 논문 장르 내 내포된 장르이기 때문에 그러하다. 그러나 본고에서는 4단계 '각 단위의 언어적 분석'과 5단계 '담화 분류의 언어적 기술'은 작문을 실제 수행하는 학습자에게 중요 정보를 제공한다고 보고 미시 분석에서 다루도록 한다.

| | | |
|---|--------|---------------------------|
| 2 | 세부 분절화 | 전체 텍스트를 이동 마디로 세분화한다. |
| 3 | 유형화 | 모든 이동 마디를 그 유형에 따라 유형화한다. |

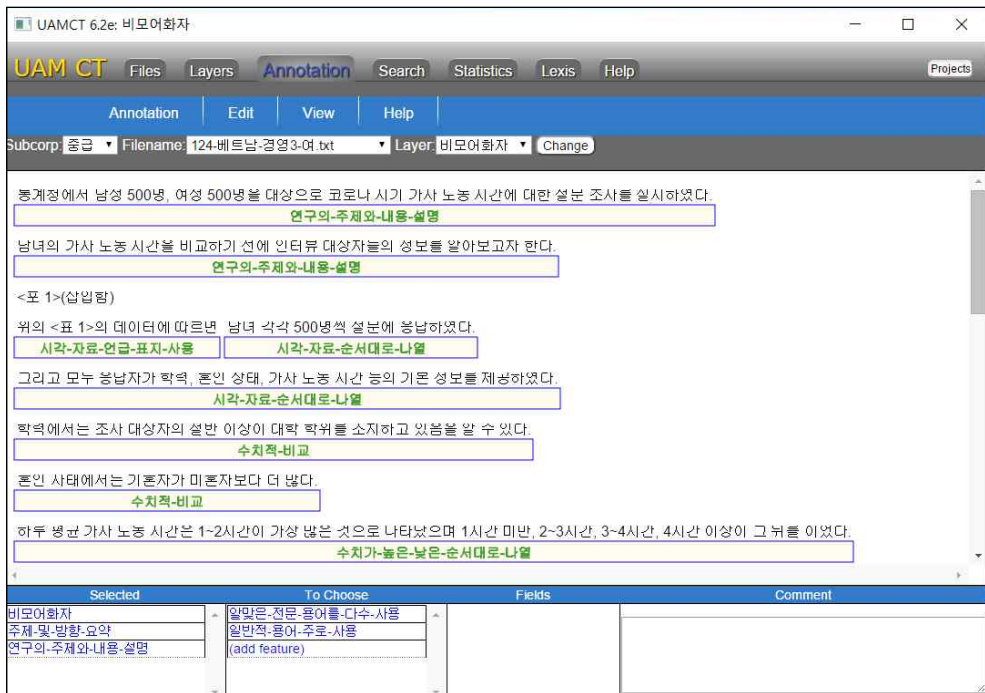
데이터 설명문의 이동 마디에 대해서는 선행 연구들에서도 제시한 바 있다. 그러나 대상 텍스트가 명확하지 않거나(Swales & Feak, 1994), 전문 필자의 학회지 논문이어서(Nordrum & Eriksson, 2015) 본고에서 수집한 학습자 데이터 설명문과 차이가 있는 경우가 있다. 또한 학습자가 수업 중에 작성한 데이터 설명문이어서 본고와 조건이 모두 유사하지만, 이동 마디의 구조를 온전히 사용하기에는 전공 차이와 다루는 텍스트 차이로 인해 보완이 필요한 것도 있었다(Guinda, 2011). 따라서 본고에서는 수집한 학습자 데이터 설명문의 구조에 집중하여 위 1,2,3단계 분석을 적용하고자 하였다.

다만 본고에서는 학습자에게 데이터 설명문에 국한된 과제를 받았기 때문에 ‘데이터 설명문 추출하기’ 단계는 사용하지 않고, 그다음으로 제시한 1,2,3단계를 수행하였다. 그러나 ‘의사소통적/기능적 분류’에서 수사적 기능을 결정하는 것이 보통 언어적 기호에 국한되기 때문에 시각적 데이터의 활용을 풍부하게 파악하기 어렵다고 보았다. 또한 ‘세부 분절화’와 ‘유형화’ 역시 의사소통적/기능적 분류의 한계 때문에 데이터 설명문을 구성하는 학습자들의 인지적 양상을 포착하기 어려웠다. 이에 세분화와 유형화의 방법과 기준을 학습자 데이터 설명문 자체에서 찾기로 하였다. 이를 위해 본고에서는 하향식 이동 마디 분석 방법에 반복적 비교 분석(constant comparative method)을 적용하였다.

반복적 비교 분석은 근거 이론(ground theory)의 방법론 중 일부를 차용해 다양한 질적 연구에 사용하는 것으로, 정보가 수집됨에 따라 이를 비교하고 대조하여 하나의 큰 현상으로 설명하는 것을 뜻한다(DePoy & Gitlin, 2016:318). 반복적 비교 분석은 자료를 읽으면서 자료에 이름을 붙이고 분류하는 ‘개방 코딩’으로 시작한다. 다음으로 진행되는 ‘범주화’ 단계에서는 코드들 간의 비교와 대조를 통해 코드들이 속하는 범주를 형성한다. 이 범주들의 이름을 명명할 때 선행 연구에서의 이름을 차용할

수도 있다. 마지막으로 ‘범주 확인’은 범주에 맞게 코딩한 자료들이 잘 들어가는지 확인하는 단계이다(유기웅 외, 2019:324).

본고에서는 한국인 대학생 및 외국인 대학생의 데이터 설명문을 말뭉치로 구축하여 UAM Corpus TOOL(Version 6.2e)을 이용한 코딩을 실시하였다. 실제 코딩의 예는 다음 <그림 1>과 같다.



<그림 1> UAM corpus tool 6.2e를 통한 분석의 예

우선 한국인 대학생 자료 10편, 외국인 대학생 자료 20편(중급 10편/고급 10편)을 줄 코딩하여 이름을 붙였다. 그 과정에서 텍스트 순서에 따라 범주화하여 이동 마디를 형성하였다. 이동 마디에 포함되지 않는 부분과 주어진 이동 마디로 묶은 문장 및 구 중 이동 마디가 표상하고자 하는 의미와 맞지 않는 부분은 따로 표시하였다. 다음으로는 앞에서 배제된 문장 및 구를 포함하여 제시할 수 있는 포괄적 이동 마디를 구축하기 위해 원자료를 반복적으로 확인하였다. 이미 분류한 이동 마디의 이

름과 위치가 데이터 설명문 구조에 적합한지 판단하며, 그를 통해 이동 마디의 특성을 잘 표현할 수 있는 이름으로 이동 마디 이름을 수정하였다. 이 과정에서는 선행 연구에서 제시된 이동 마디 이름과 속성을 참고하였으며, 선행 연구들이 내포하는 체계 기능적 다중 양식 담화 분석의 메타 기능적 의미와 인지적 특성 역시 고려하여 더욱 세밀한 범주를 구성하였다. 마지막으로 수집한 자료 모두를 앞에서 개발한 이동 마디로 분석하며, 이에 포함되지 않거나 새로운 범주를 만들 수 있는 코드를 모아 지속적으로 범주의 구성과 위계를 변경하며 설정하였다. 이를 통해 이동 마디와 하위 이동 마디 또는 단계를 상세화하였다. 위 <그림 1>을 124번 학습자 작문을 통해 구체적으로 분석한 예는 다음과 같다.¹¹⁾

<이동 마디A: 주제 및 방향 요약>

[A2: 연구의 주제와 내용 및 방법 설명]

<이동 마디B: 수치/결과 나열>

[B1-2: 시각적 데이터 순서대로 나열]

[B2-1: 수치가 높은/낮은 순서대로 나열]

[B2-2: 수치적 비교]

[B2-3: 특성에 따른 그룹화]

[B2-4: 제일 큰/작은 수치에 주목]

[B2-5: 특이성/중요 지점 지적]

[B3: 시각적 데이터 언급 표지 사용]

[B4: 수치의 의미 제시]

<이동 마디C: 해석 및 결론>

[C1: 주어진 데이터/상식적 지식에 근거한 해석]

11) <그림 1>과 달리 이 분석 예시에서는 이동 마디 이름을 대체하는 표지를 부착한다. 이 표지는 각 이동 마디 이름이 너무 길어 분석에 사용하기 어렵기 때문에 대체적 기호로 제시한 것이다. 이는 II장 3절에서 자세히 언급할 것이다. 본문에서는 우선 124번 학습자의 작문에 나타난 이동 마디만을 굵은 글씨로 제시하였다.

<표 3> 학습자 데이터 설명문 분석의 예

(A2) 통계청에서 남성 500명, 여성 500명을 대상으로 코로나 시기가 가사 노동 시간에 대한 설문 조사를 실시하였다. 남녀의 가사 노동 시간을 비교하기 전에 인터뷰 대상자들의 정보를 알아보려고 한다.

<표 1>(삽입함)

(B3) 위의 <표 1>의 데이터에 따르면 (B1-2) 남녀 각각 500명씩 설문문에 응답하였다. (B1-2) 그리고 모두 응답자가 학력, 혼인 상태, 가사 노동 시간 등의 기본 정보를 제공하였다. (B2-2) 학력에서는 조사 대상자의 절반 이상이 대학 학위를 소지하고 있음을 알 수 있다. (B2-2) 혼인 상태에서는 기혼자가 미혼자보다 더 많다. (B2-1) 하루 평균 가사 노동 시간은 1~2시간이 가장 많은 것으로 나타났으며 1시간 미만, 2~3시간, 3~4시간, 4시간 이상이 그 뒤를 이었다.

(A2) 다음으로 맞벌이 상태별 하루 가사 노동 시간을 나타내고 있다. (B3) 다음과 같다.

<그림 1>(삽입함)

(B3) <그림 1>를 살펴보면 (B2-5) 맞벌이 또는 외벌이의 경우도 여성이 남성보다 가사 노동을 했다. (B1-2) 먼저 맞벌이의 경우 남성의 가사 노동 시간이 가장 적으로 조사되었고 (25시간) 여성의 가사 노동 시간이 약 140시간이었다. (B1-2) 남편이 외벌이를 하는 경우 (B2-4) 여자들이 가사 노동을 하는 시간은 가장 많은 것으로 보인다 (약 245시간). (B1-2) 또한 아내가 외벌이를 하는 경우에서도 (B2-5) 여자들이 남자보다 집안일에 더 많은 시간을 보냈다.

(A2) 마지막으로 남성과 여성의 하루 평균 가사 노동 시간 비교를 하고자 한다.

<그림 2> (삽입함)

(B2-3) 남성은 2019년 48명에서 2020년 61명으로 증가해서 2021년에는 78명에 이르렀다. (B2-3) 반면 여성은 2019년 174명에서 2020년 187명으로 증가했다. (B1-2) 그러나 2021년에는 180명으로 감소했

다. (B2-2)2021년에는 여성의 가사 시간이 줄어들었지만 여성의 가사 시간이 여전히 남성에 비해 두 배 많았다.

(B4)위의 데이터에 의하면 코로나 시대에도 여성이 남성보다 더 많이 집안일을 하고 있다는 결론을 내릴 수 있다. (C1)코로나 기간 동안 남성은 집에 있는 시간이 길어지고 하루 평균 가사 노동 시간이 증가한 반면 여성의 가사 노동 시간은 감소했다.

[124-중급-베트남어 경영3-여]

이렇게 수집된 학습자 데이터 설명문 이동 마디의 순서 유형, 단계별 빈도의 차이는 학습자 변인별로 제시될 예정이다. 이동 마디 순서는 학습자 작문의 구성 순서를 파악하고 어떠한 반복이 일어나는지 파악할 수 있기에 이동 마디 연구에서 다수 분석된다(박은선, 2006; 유민애, 2017; 김진웅, 2019). 다만 이동 마디 단계의 수가 다양하고 그 조합이 대단히 많기 때문에 이들을 모두 나열하여 비교하는 것은 한계가 있다. 따라서 본고에서는 하위 단계까지는 다루지 않고 상위 단위인 이동 마디로만 이를 확인하여 학습자 변인별로 비교하도록 한다.

또한 이동 마디 순서에서 다루지 못한 이동 마디 단계는 별도로 얼마나 많이 선택했는지 빈도를 추출할 것이다. 이를 바탕으로 학습자 변인별 이동 마디 단계의 선택 빈도 차이를 t-test를 통해 알아볼 것이다. 이는 이동 마디의 출현 빈도가 이동 마디의 필수성을 확인하는 분석 방법으로 자주 사용됨을 활용한 것이다.

(2) 언어 사용 분석

이동 마디의 범주화 작업 이후에는 각 하위 이동 마디와 단계(step)에서의 언어적 분석을 실시한다. 이는 위 <표 2>의 Biber et al.(2007:13)에서 제시한 네 번째, 다섯 번째 단계의 미시 분석에 해당한다.

<표 4> Biber et al.(2007)의 하향식 이동 마디 분석 중 미시 분석에의 이용 단계

| | 단계 | 이 접근에서 얻게 되는 것 |
|---|---------------|---|
| 4 | 각 단위의 언어적 분석 | 각 담화 단위의 어휘/문법적 특징을 분석한다. |
| 5 | 담화 분류의 언어적 기술 | 말뭉치의 특정 기능 유형 담화 단위에서 각 기능 분류의 전형적인 언어적 특징을 기술한다. |

이 절차를 수행하기 위해 본고에서도 Biber et al.(2007)에서 제시한 바와 같이 양적 분석인 말뭉치 분석을 수행한다. 이는 말뭉치의 언어적 특성을 계량적으로 파악할 수 있는 방법이며, 이를 통해 거시적으로 분석한 이동 마디와 단계에 해당하는 어휘문법적 특징을 밝혀낼 수 있다.

일반적으로 텍스트의 특수성을 드러내는 어휘문법적 특징은 여러 방법으로 밝혀낼 수 있는데, 말뭉치 언어학에서 주로 사용하는 방식은 핵심어(keyword), 연어(collocation), n-gram 등이 있다. 본고에서는 핵심어나 연어와 같이 내용어를 기반으로 한 어휘문법적 특징보다는 기능적이고 문법적 구성을 포함하는 단위를 밝혀내고자 한다. 학습자의 데이터 설명문에서 발견된 고빈도 내용어가 과제로 제시한 표와 그래프의 제목과 범례인 경우가 대부분이었기 때문이다. 따라서 본고에서는 n-gram을 기반으로 형태적 고정성을 지닌 연쇄적 표현에 초점을 맞출 것이다. 이러한 고빈도 n-gram에 기반한 연쇄적 고정 표현은 ‘어휘 꾸러미(lexical bundle)’, ‘정형 표현’, ‘다단어 단위(multi-word unit)’ 또는 ‘문형’이라 불린다(Biber et al, 2004; 최준 외, 2010; 남길임, 2013; 허희정, 2016; 윤새롬, 2019). 본고에서는 추출한 n-gram 연쇄에 이러한 이름을 붙이지 않는다. 다만 ‘어휘문법적 특성’이라는 범주 내에서 내용어 단독으로만 존재하는 구성과 내용어+기능어 결합, 기능어+기능어 결합으로 이루어진 구성을 모두 밝혀 적을 것이다.¹²⁾ 또한 ‘기능어+내용어’, 또는 ‘내용어+기능어’ 구성에서 어미나 조사의 이형태는 앞말에 자연스럽게 이어지는 것

12) 여기에서의 내용어는 과제에 제시된 표와 그래프 제목 및 범례를 제외한 것들이 속한다. 예를 들면, ‘증가하다, 나타나다’와 같은 것들이 있다. 또한 ‘내용어+기능어’ 구성은 ‘-에 대해-’, ‘-를 실시하-’와 같이 나타난다. ‘기능어+기능어’ 구성은 ‘-이었다’, ‘-있지만’ 등이 있다.

이나 대표형 하나만을 기재하며, 내용어가 유사하고 조사와 어미 구성이 같은 어휘문법의 경우에는 기능어는 그대로 둔 채 내용어만을 ‘/’를 통해 제시한다. 예를 들면, ‘-에 대한/관한’과 같은 구성이 그러하다.

n-gram 추출은 장르적 범주가 크고 반복의 빈도가 높은 연구에서 주로 사용되므로, 그보다 작은 이동 마디 단위에서 고정성을 알아보려는 연구를 수행하도록 한다. 본고에서는 윤새롬(2019)의 방식을 따라 이동 마디와 단계를 먼저 나누고 그 안에서 자주 사용되는 어휘문법을 추출할 것이다. 일반적으로 한국어의 n-gram 연구에서는 영어에서의 말뭉치 분석 단위와 같이 어절 단위의 분석을 행하거나, 조사 및 어미를 포함하여 형태소 단위를 기준으로 삼아 분석을 행하는 두 가지 방식을 확인할 수 있다. 전자의 경우 직관적이고 편리한 분석이 가능한 반면 한국어의 교착어적 성격을 나타내는 기능어들을 확인할 수 없다는 단점이 있다. 반면 후자는 예를 들어, ‘-ㄴ 경우에’와 같이 3-gram 추출을 했을 때 관형형 어미까지 밝혀낼 수 있으므로 기능어의 확인이 가능하다.

본고에서는 한국어의 교착어적 특성을 볼 때, 문장 구조와 고정성에서는 어미와 조사 등의 기능어가 중요하기 때문에 연속되는 형태소의 구성인 n-gram을 설정하는 것이 바람직하다고 본다. 따라서 본고에서는 최준외(2010) 및 최지희(2022)에서 분석한 것처럼 n-gram을 형태소 단위로 태깅하여 추출할 것이다. 한국어의 형태소 교착 및 어미의 사용을 더 상세히 태깅하는 kkma 태그¹³⁾를 부착하여 이동 마디별 품사 태깅을 수행하고, 태깅이 잘못된 것들은 모두 수작업을 통해 바로잡았다. n-gram의 추출은 Microsoft Excel의 textjoin함수를 활용한다.

n-gram으로 정형 표현을 추출할 때 Biber et al.(2004)와 같은 경우 영어에서는 어절 단위의 4-gram을 유효한 것으로 설정하는 데 반해, 최준외(2010) 및 남길임(2013), 최지희(2022)는 한국어의 어미가 갖는 기능을

13) 서울대학교 IDS 연구실에서 개발한 형태소 분석기이다. 이는 국립국어원에서 제공하는 세종 품사 태그보다 어미를 세분화하여 형태소 교착을 보다 더 자세히 포착할 수 있다. 또한 작문에서 나타나는 오폭기의 품사 역시 잘 유추하여 부착하는 경향이 있다. 이는 현재 kkma.snu.ac.kr 홈페이지에서 간단히 품사 태깅이 가능하며, python으로도 실행 가능한 라이브러리를 제공한다.

극대화하여 형태소 단위의 5-gram으로 기준을 둘 것을 제안하였다. 그러나 n 이 커질수록 빈도가 작아지기 때문에 이동 마디 단위라는 아주 작은 말뭉치를 상정하는 본고에서는 적어도 어간과 어미의 결합을 확인할 수 있으며, 조사와 어미를 구분할 수 있는 3-gram을 기준으로 두었다. 다만 이동 마디별 말뭉치가 각 작문에서 추출한 분절된 구 또는 문장들의 모임이기 때문에 앞뒤 문맥을 다른 말뭉치들보다 판별하기 어렵다는 단점이 있다. 이에 n 의 수치를 3과 4 둘로 두고 결합 양상을 확인하여 빈도를 셀 때 참고하도록 한다. 다만 이동 마디의 성질에 따라 n 의 기준이 달라질 수 있는데, 예를 들어, ‘(소)제목을 활용한 주제 제시(A1)’와 같은 이동 마디에서는 말뭉치 거의 대부분을 ‘코로나 시기 성별 가사 노동 시간 연구’라는 문구가 차지한다. 따라서 A1에서는 n -gram에 따른 구성을 제시하지 않고 해당 이동 마디 단계에 나타나는 제목을 중복되는 것을 정리하여 모두 기재할 예정이다. 나머지 소제목 역시 ‘서론’, ‘본론’, ‘결론’과 같은 명사 또는 ‘설문 응답자 정보’와 같은 명사군으로 이루어져 있어 따로 분석이 필요하지는 않다고 보았다.

이러한 분석이 끝난 후 3-gram 목록에서 동일 어휘문법이 몇 번 나타나야 목록에 수록할 것인지를 결정해야 한다. 연구자는 자료의 크기에 따라 빈도와 범위의 기준을 자의적으로 조정해야 한다(Byrd et al. 2010). 본 연구에서 빈도 2개 이상이라는 대단히 보수적인 기준을 잡은 이유는 두 가지가 있다.

첫째, 어휘문법을 추출할 텍스트 길이가 매우 짧다. 어휘문법 추출 텍스트의 범위는 하나의 이동 마디 내 단계이므로 중급 A1과 같이 200자인 것도 있고 중급 B4처럼 6,000자인 것도 있다. 이 때문에 한 이동 마디 단계에 어절 수가 적을 경우에는 2개마저도 반복되지 않는 경우가 빈번하다는 점을 고려하였다.

둘째, 또한 본고에서는 태깅한 단위가 계속 반복되는 일반적인 n -gram이 아닌 태깅한 단위를 절대적으로 n 개로 끊어서 제시하는 n -gram이기 때문에 동일한 단위가 반복되는 경우가 적다. 아래 <표 5>에서 볼 수 있듯, 일반적인 n -gram을 추출하게 되면 하나의 단위를 지

속적으로 반복하기 때문에 실제로는 동일한 단위가 과도하게 추출된다.

<표 5> n-gram 추출 비교

| B1-1 일반 n-gram 추출 | B1-1 본고 3-gram | B1-1 본고 4-gram |
|-------------------|----------------|----------------|
| (전략) | | |
| 187 분 2021 | 2019 년 부터 | 2019 년 에 는 |
| 187 분 으로 | 2019 년 에 | 2020 년 61 분 |
| 187 분 을 | 2020 년 까지 | 2021 년 에 는 |
| 2019 년 48 | 2020 년 남자 | 2921 년 78 분 |
| 2019 년 부터 | 61 분 2921 | 48 분 2020 년 |
| 2019 년 에 | 61 분 이 | 78 분 이 다 |
| 2019 년 에 | 까지 7 계단 | 78 분 이 라고 |
| 2019 년 에 | ↳ 증가세 를 | 가사 노동 시간 은 |
| 2020 년 까지 | ↳ 다 반면 시간 | 가사 노동 시간 은 |
| 2020 년 187 | 남 자 하루 | 경우 는 2019 년 |
| 2020 년 61 | 년 78 분 | 까지 7 계단 하락 |
| 2020 년 남자 | 년 78 분 | 년 남자 61 분 |
| 2020 년 에 | 노동 시간 은 | 노동 시간 을 비교 |
| 2021 년 78 | 다 2019 년 | 다 |
| 2021 년 까지 | 다 남성 쪽 | 돌파 하 었 지만 |
| 2021 년 에 | 다가 2021 년 | 등 꾸준 하 ↳ |
| (중략) | | |
| 가사 노동 시간 | 분 이 다 | 의 하루 평균 가사 |
| 가사 노동 시간 | 상승 하 었 | 이 1 시간 을 |
| (중략) | | |
| 년 남자 61 | 평균 가사 노동 | |
| 년 187 분 | 평균 가사 노동 | |
| 년 48 분 | 평균 가사 시간 | |
| 년 61 분 | 하 면 여자 | |
| 년 까지 174 | 하 었 다 | |
| 년 까지 7 | 하락 하 었 | |

이에 본고에서는 3-gram을 기준으로 4-gram까지 함께 활용하여 특정 표현이 2개 이상으로 출현하는지를 보려 한다. 예를 들어, 위 표에서 보면, 맨 왼쪽의 형태 반복적 3-gram은 ‘(수치)년에’가 다수 출현하지만 4-gram으로 옮겨가 확인해 보면 실제로는 ‘(수치)년에는’과 같은 구성을 담지 못해 생긴 것임을 알 수 있다. 따라서 실제 이상의 빈도가 야기할 문제를 고려하여 본고에서는 2개 이상의 실제 나타나는 어휘문법에 집중

한다.¹⁴⁾

어휘문법의 양상이 다양해 3-gram 길이에서 공통된 형태를 추출하기 어려운 경우에는 공통적으로 발생하는 최소의 구성이라도 찾아내 제시하도록 한다. 예를 들어 ‘-는 것’과 같은 경우는 앞뒤로 다양한 내용어와 기능어가 함께하므로 하나의 고정된 형태를 제시하기 어렵다. 따라서 자주 등장하는 결합적 구성이 활용이나 어근/어간의 다양성 때문에 고정적 형태를 제시하지 못한다면 앞뒤로 결합 여부를 ‘-’로 제시한 후 최소 고정 형태만을 제시하도록 한다. 여기서 내용어+기능어, 또는 기능어+기능어 결합형이 아닌 조사나 어미를 단독으로 반복 사용하는 것은 추출에 넣지 않는다. 예를 들어, 주격조사 ‘-이’ 또는 관형형 어미 ‘-ㄴ’은 아무리 많이 등장한다 할지라도 그것의 단독 등장은 어휘문법으로 처리하지 않는다.

이에 더해 과제로 제시한 표와 그래프에서 나타나는 내용어는 배제하도록 한다. 이미 주어진 시각적 데이터에 나타나는 어휘를 그대로 보고 사용하는 것은 불필요한 반복을 만든다. 위의 <표 5> 가운데에 기재한 본고의 3-gram 분석에서도 전체 200자가량의 텍스트에 ‘평균 가사 N’이 세 번 이상 출현한다. 따라서 이들 내용어는 추출 대상에서 제외한다. 그러나 나머지 내용어에 대해서는 활용형까지 완전히 같지는 않더라도, 어근이나 어간이 동일한 내용어가 2회 이상 등장한다면 앞뒤에 교체 정보를 밝히는 ‘-’를 삽입하여 추출하도록 한다. 예를 들어, 베트남어권 학습자의 B1-1에서는 ‘증가하였다, 증가하는 경향을 보였다’와 같은 구성을 발견하였기 때문에 ‘증가하-’를 추출하였다.

또한 베트남어권 학습자 A2 말뭉치에서 추출한 ‘통하여 하루’와 같은

14) 또한 이들의 빈도와 범위(소속 데이터 설명문의 편수)를 따로 기록하지는 않는다. 이제까지 최대로 나타난 동일 구성이 18개(‘-르 수 있’, 중국어권 B4)이기 때문에 개수가 매우 적어 빈도 표기의 의미가 없다고 보았다. 또한 범위 역시 말뭉치 수집 범위가 텍스트 하나가 아닌 구나 문장 단위여서 1번 데이터 설명문에 속한 문장이나 구가 말뭉치 내에서 2번 데이터 설명문의 구나 문장과 앞뒤로 나열해 있다. 이러한 경우 어휘문법이 추출되면 어느 것이 어휘문법이 소속된 텍스트인지 구분하기 어렵다. 하나의 텍스트 단위라면 이러한 경우가 적으나 본고의 말뭉치는 각 텍스트에서 몇 개의 문장과 구를 뽑아 각 이동 마디 단계를 구성하므로 단절 부분이 대단히 많다.

의미 추출이 불가능한 구, 절 및 ‘다 이 그래프’와 같이 문장의 경계가 뚜렷한 연속체들도 제외하도록 한다. 다른 어휘문법적 요소 추출에 관련 항목이 들어가 있는 경우에는 문장이나 절의 의미를 추출할 수 있지만, 이렇듯 분절된 단위로는 어휘문법 구조를 알기 어렵기 때문이다.

말뭉치의 규모에 대해서 권혁승 외(2018:153)에서는 특수 목적 언어 교육을 위한 말뭉치는 이상적 규모는 없으며 적절한 규모는 구체적 연구 목적과 필요에 따라 다르다고 하였다. 특히 교육적 목적으로 미니 코퍼스나 하부 코퍼스와 같은 작은 규모의 코퍼스를 구축하여 사용하는 것도 유용한 것으로 보고 있다(권혁승 외, 2018:153). 일반적으로 특수 목적 영어에서는 전문 영역 어휘 선정에서 백만 어절을 기준으로 두고, 이를 대표성과 특이성을 파악할 수 있는 규모로 삼는다. 그러나 영어를 기반으로 한 특수 목적 소규모 말뭉치 구축에 대해서는 김동영(2021)에서 Nelson(2010)과 Koester(2010)를 인용하여 한 장르당 2만 어절로 텍스트 수집을 하는 것이 바람직하다고 제시하고 있다. 응용 화학 석사 학위 논문과 학회지 논문의 데이터 설명문 말뭉치를 연구한 Nordrum & Eriksson(2015)에서도 전자는 22,818어절, 후자는 15,817어절로 소규모 말뭉치 어절 기준을 따르려 노력한 것을 볼 수 있다.

이에 따라 본고에서도 n-gram 분석을 위해 수집하는 자료의 어절 수를 2만 어절가량으로 정한다. 이는 특히 본 연구처럼 완결된 텍스트 전체를 대상으로 하지 않아 적은 분량의 텍스트만을 대상으로 하는 데 적합하다 할 수 있다.

3.2. 연구 대상 및 자료

본 연구는 Google docs를 이용한 비대면 온라인 설문조사 방식으로 참여자의 참여 동의 여부와 간단한 인구학적 정보 및 데이터 설명문 작문을 수집하였다. 또한 대학 재학생의 참여를 위해 대학 온라인 커뮤니티, 유학생/외국인 온라인 커뮤니티 및 온라인 채팅방 등에 연구 참여 링크를 게시하였다. 이를 통한 학습자 자료 수집 기간은 2022년 6월 3일부터

2022년 12월 31일까지였다.

이 연구 참여자들은 한국 대학에서 경영 또는 경제 관련 전공을 하고 있어야 하고, 18세 이상의 성인이어야 한다. 텍스트의 우수성 여부를 평가하려는 것이 목적이 아닌 학술적 데이터 설명문 작성에 있어 대학생이 어떠한 특성을 보이는지 확인하는 것이 목적이기 때문에 재학 중인 대학의 이름이나 참여자 이름 등의 불필요한 개인정보는 수집하지 않았다. 따라서 인구학적 정보는 성별, 연령¹⁵⁾, 전공(경제학 계열/경영학 계열), 학년¹⁶⁾만을 수집하며 외국인 대학생의 경우 한국어 능력(중급/고급)과 모어 정보를 추가적으로 수집하였다. 외국인 대학생의 국적은 고려 변인이 아니었으나, 한국 대학 유학생의 국적 비율이 중국과 베트남 순으로 높아서 그러한지 중국과 베트남 학생의 참여도가 높아 실험 진행 중 언어권 학생들을 집중적으로 모집하여 모어별 차이도 추가적으로 알아보도록 하였다.¹⁷⁾

연구 참여 대상은 한국인 대학생 50명, 한국어 능력이 중급 이상인 외국인 대학생 100명(중급 50명/고급 50명)으로 정하였으나 실제 모집 완료 인원은 한국인 대학생 50명, 외국인 대학생 각 70명(중급), 55명(고급)이다.¹⁸⁾ 구체적인 참여자 정보는 한국인 대학생의 경우는 다음 <표

15) 연령의 경우 한국인 대학생은 20대 46명, 30대 3명, 40대 1명이었고, 외국인 대학생은 20대 126명, 30대 1명으로 20대가 압도적으로 많아 연령이 큰 변수가 되지 않는 것으로 판단하였다.

16) 학년을 변인으로 두고 분석을 하려 하였을 때는 전문 용어 및 개념 사용과 같은 전공 지식이 개입하는 양상을 보여 전공 교육이 아닌 언어 교육에서 다루기에는 적절하지 않다고 보았다. 다만 학년의 개입 양상이 글쓰기 구조와 어휘에서 나타나는 것이 사실이므로 추후 연구자가 다룰 수 있는 전공 영역에서 학습자 학년 변인을 다시 확인할 필요는 있다.

17) 교육안전정보국 교육통계과(2021:25)에서는 출신 국가별 학위과정 외국인 유학생을 중국(49.8%), 베트남(20.8%), 몽골(4.1%), 일본(1.7%), 미국(1.1%), 기타(22.5%)순으로 제시하고 있다.

18) 불성실한 응답, 분량에 맞지 않는 텍스트, 표절 등을 고려하여 한국인 100명, 외국인 중급 100, 외국이 고급 100명까지 모집할 수 있도록 IRB 승인을 득하였다(서울대학교 생명윤리위원회 심의 승인: IRB No. 2206_001-010). 또한 각 집단별로 2만 어절에 근접하도록 과제 분량을 1000자(약 400어절)*50명으로 계산하였으나, 참여자 모집의 어려움으로 인해 600자(약 250어절)로 분량이 수정되어 인원수의 변동이 있었다. 한국인 대학생 50명이 작성한 작문의 말뭉치 규모와 외국인 대학생의 말뭉치 규모를 맞추기 위해 자료 수집 기한 내에 최대한 모집한 외국인 대학생의 수가 중급 70명, 고급 55명이다.

6>, 외국인 대학생의 경우 <표 7>과 같다.

<표 6> 한국인 대학생 참여자 정보

단위: 명

| 전공 | 성별 | | 계 |
|--------|----|----|----|
| | 남 | 여 | |
| 경영학 계열 | 9 | 20 | 29 |
| 경제학 계열 | 7 | 14 | 21 |
| | | | 50 |

<표 7> 외국인 대학생 참여자 정보

단위: 명

| 전공 | 성별 | | 모어 | 한국어 능력 | | 계 | |
|--------|----|----|----------|----------|----|-----|----|
| | 남 | 여 | | 중급 | 고급 | | |
| 경영학 계열 | 10 | 71 | 중국어: 33 | 몽골어: 0 | 45 | 36 | 81 |
| | | | 베트남어: 38 | 영어: 2 | | | |
| | | | 말레이어: 3 | 포르투갈어: 1 | | | |
| | | | 터키어: 1 | 미얀마어: 0 | | | |
| | | | 러시아어: 2 | 스페인어: 1 | | | |
| 경제학 계열 | 7 | 37 | 중국어: 26 | 몽골어: 2 | 25 | 19 | 44 |
| | | | 베트남어: 12 | 영어: 0 | | | |
| | | | 말레이어: 0 | 포르투갈어: 1 | | | |
| | | | 터키어: 2 | 미얀마어: 1 | | | |
| | | | 러시아어: 0 | 스페인어: 0 | | | |
| | | | | | | 125 | |

한국인 대학생의 답안은 총 52편을 수집하였으나 불성실 답안 1편과 중복 답안 1편을 제외하고 50편을 사용하였다. 외국인 대학생 답안의 경우 총 모집 편수는 131편이나, 150자 미만 작성으로 분량이 심각하게 미달하거나 중도 포기하여 문장을 제대로 끝맺지 않은 답안 2개 및 동일한 답안 4개를 제외하고 125편을 사용한다.¹⁹⁾ 또한 자료 구분은 모두 기호

19) 이동 마디 코딩 과정에서 맨 처음과 맨 마지막 문단을 다르게 하고 중간 내용이 동일

로 하였다. 기호 표기는 다음과 같다.

<표 8> 경제/경영 전공 대학생 작문 말뭉치 자료 표시

| 구분 | 인원 | 기호 |
|---------|--------------------|--|
| 한국인 대학생 | 50명 | 수집 연번-한국어-전공(학년)-성별 예) [1-한-경영4-여] |
| 외국인 대학생 | 중급: 70명 고급: 55명 | 수집 연번-한국어 능력 수준-모어-전공(학년)-성별 예) [1-중급-중국-경영2-남] |

참여자들에게는 표 하나, 그래프 두 개를 제공하고 이를 사용하여 수업 보고서를 쓰는 것처럼 데이터 설명문을 작성해 줄 것을 요청하였다.²⁰⁾ 학부 학생으로서 참여자들이 속해 있는 학술적 상황에서 가장 빈번하게 요구되는 학술적 텍스트는 수업 보고서이다. 따라서 수업 보고서에서의 데이터 설명문 쓰기 과제는 실제성을 확보할 수 있다고 보았다. 참여자에게는 600자 이상(약 150어절)을 써 줄 것을 요청하였다.²¹⁾

한국인 대학생의 작문에서 수집된 총 글자 수는 55,312자이며 어절 수로는 17,004어절이다. 평균 텍스트 길이는 1106.24자, 340.08어절이다. 문장당 어절 수는 한 문장의 길이가 얼마나 긴지를 확인하는 지표이다. 한국인 대학생의 경우 문장당 16.79어절이 도출되었다. 또한 이들 텍스트에서 추출한 TTR(Type-Token ratio)²²⁾은 0.18이다.

한 편([중급-34-중국-경제3-여]와 [중급-39-중국-경영3-여])과 아예 텍스트 전체가 모두 똑같은 편([중급-117-베트남어-경영3-여]와 [중급-122-베트남어-경영2-남] 및 [중급-118-베트남어-경영3-여]와 [중급-126-베트남어-경영2-여])이 발견되었다. 이는 동일인이 작성하였거나 동일 글을 다른 두 사람이 쓴 것으로 간주해 39와 122, 126을 삭제하였다. 그러나 이미 UAM Tool에 이 이후에 해당하는 작문의 연번을 저장한 다음이기 때문에 번호 수정이 불가하여 총 사용 편수는 125명이나 학습자 자료 번호는 128번까지 있음을 밝힌다.

20) [부록]을 참고하여라.

21) 초기 IRB 심의에서는 한국인/외국인 대학생 모두에게 1,000자 쓰기를 요구하였으나, 외국인 대학생의 경우 1,000자의 글쓰기 자체가 부담이 되어 그러한지 참여율이 매우 저조하였다. 그래프 한 개를 제시하는 TOPIK 그래프 쓰기가 2~300자를 기준으로 하기에 본고에서도 이와 같은 조건에 맞추어 표와 그래프 총 3개에 대한 쓰기를 600자로 조정하여 수정 심의를 받았다.

22) Type은 전체 텍스트에서 유일하게 나타나는 형태를 뜻하고, Token은 전체 형태의 개수를 나타낸다. 타입/토큰 비율이 높을수록 다양한 단어를 사용했다는 의미가 된다. 본고에서는 타입과 토큰을 Sketch Engine을 통해 파악하였다.

외국인 대학생들의 작문에서 추출한 총 글자 수는 93,179자이며, 어절 수는 28,927어절이다. 급수별로 나누어 제시하면, 중급 46,232자, 14,793어절, 고급 46,947자, 14,134어절이다. 평균 텍스트 길이는 중급 660.45자, 201.91어절 고급 853.58자, 268.96어절이며, 문장당 어절 수는 중급 12.22어절/문장, 고급 14.35어절/문장이다. TTR은 중급 0.12, 고급 0.13이어서 0.18인 한국인 대학생에 비하면 다양한 어휘를 사용하는 정도에 차이가 있었다. 이를 정리하면 다음 <표 9>와 같다.

<표 9> 참여자 말뭉치 정보

| | 한국인 대학생 말뭉치 | 외국인 대학생 말뭉치 | |
|---------------------------|-------------------|-------------|------------|
| | | 중급 | 고급 |
| 수집 편 수 | 50 | 70 | 55 |
| 글자 수 | 55,312자 | 46,232자 | 46,947자 |
| | | 93,179자 | |
| 어절 수 | 17,004어절 | 14,793어절 | 14,134어절 |
| | | 28,927어절 | |
| 평균 텍스트 길이 | 1106.24자 | 660.45자 | 853.58자 |
| | 340.08어절 | 201.91어절 | 268.96어절 |
| 문장당 어절 수 | 16.79어절/문장 | 12.22어절/문장 | 14.35어절/문장 |
| TTR (Type/Token ratio) | 0.18 | 0.12 | 0.13 |

한국인 대학생에 비하면 외국인 대학생들의 전체 글자 및 어절 수가 부족하나, 숙달도에 따른 텍스트 길이와 한국어 모어성을 고려했을 때 수용 가능한 정도로 생각된다.²³⁾ 또한 평균 텍스트 길이와 문장당 어절 수는 한국인 대학생>외국인 대학생(고급)>외국인 대학생(중급) 참여자 순으로 긴 것으로 나타났다. 이는 참여자가 가용할 수 있는 언어 레퍼토리의 차이로 생각된다.

데이터 설명문 과제는 저학년 학생들을 고려하여 일반적인 대학생이 흥미를 갖고 이해할 수 있을 만하며, 경제학 및 경영학에서 관심을 갖고

23) Nordrum & Eriksson(2015)에서도 비교/참조 말뭉치가 7,000어절가량 차이난다.

다루는 여러 현상 중 하나를 선정하고자 하였다. 또한 주어진 데이터를 통해 학습자가 가치적인 판단을 하고 이를 합리화할 수 있도록 논쟁적인 주제를 제시하고자 하였다.²⁴⁾ 이에 따라 본고에서 참여자에게 제시한 과제의 주제는 최근 전세계적으로 가장 주목받았던 현상인 ‘코로나’, 참여자 모두가 소속되어 있는 ‘성별’, 누구나 일상적으로 경험해 봤음직한 ‘가사’와 경제학 및 경영학의 주요 연구 분야인 ‘노동’을 고려해 ‘코로나 시기 성별 가사 노동 시간 연구’로 정하였다.

과제에서 제시한 시각적 데이터 유형은 일반적으로 학술적 데이터 설명문에서 자주 사용되는 것으로, 수치적 정보를 정리하여 전달하는 표, 각기 다른 변수들의 수치를 비교하여 나타내는 막대 그래프, 시계열적 수치를 나타내는 선 그래프로 구성되어 있다.²⁵⁾ 과제에 사용된 수치는 통계청에서 제시하는 성별 가사 노동 관련 자료를 활용하였다. 다만 전반적인 추세는 동일하게 유지하되 구체적인 수치와 연도는 조금씩 바꾸어서 ‘코로나 시기’의 그래프로 변형하였다. 특히 과제의 <표 1>의 경우에는 가사 노동 시간을 제외한 성별, 학력 등의 수치는 연구자가 임의로 정하였다.

본고에서는 이들 시각적 데이터를 학습자들이 어떻게 언어화하는지 알아보하고자 하였다. 구체적으로, 수치를 어느 정도로 해석하고 그 의미를 어떻게 선택해 전달하는지, 이러한 선택은 어떠한 구조로 나타나는지, 그 구조별 표현의 방식은 어떠한지 확인하려 하였다. 이렇게 구안한 데이터 설명문 과제는 [부록]에서 확인할 수 있다.

24) 그러나 이 과정에서 객관적 판단이 아닌 가치적 판단이 개입될 수 있는 국가, 민족, 종교와 관련된 내용은 최대한 피하려 하였다. 그러나 논쟁적이지 않은 일반적 수치는 학술적 데이터 설명문보다는 기업 보고서나 신문 등에서 더 자주 사용되므로 필자가 데이터 수치에 대한 자신의 입장을 표명할 수 있는 주제로 선정하려 노력하였다. 참고로 Swales & Feak(1994)에서는 연구 참여자에게 ‘아동의 성별에 따른 공격성 차이’에 대한 표를 제시하고 데이터 설명문을 작성할 것을 요청한 바 있다.

25) 이외에 상자 그림이나 줄기-잎 그림처럼 통계적 처리 이전에 데이터의 경향성을 알아 보려는 목적으로 다소 전문적인 영역에 사용되는 시각적 데이터 및 히트맵과 원 그래프처럼 면적과 색 등으로 왜곡된 정보를 제시할 수 있는 가능성이 있는 시각적 데이터는 제외하였다.

Ⅱ. 학술적 데이터 설명문 분석을 위한 이론적 고찰

1. 학술적 데이터 설명문의 개념과 범위

Miller(1998:30)에서는 학술지에 투고된 학술 논문은 주장과 설득을 목적으로 하며, 실험실에서 바로 전달된 것 같은 시각적 데이터는 이 목적을 효과적으로 달성할 수 있게 한다고 하였다. 특정 연구 분야에 숙련된 전문 독자는 표와 그래프, 그림 자료 등의 시각적 데이터 확인만으로도, 또는 이들 시각적 데이터에 주석으로 부연된 짤막한 정보 표시 정도로 실험 결과의 의미를 파악하고 텍스트에서 저자가 주장하는 것과 다른 사실을 발견할 수 있다. 그러나 시각적 데이터에 대한 상세한 설명이 필요한 독자도 있을 것이기 때문에 저자의 실험 해석에 대한 텍스트가 제공될 수 있다(Miller, 1998:33). Miller(1998)가 다루는 이 시각적 데이터와 그에 대한 부연 텍스트는 본고에서 다루는 학술적 데이터 설명문의 구성과 목적에 정확하게 부합한다. 그러나 Miller(1998)는 이에 대해 단순한 ‘시각적 설득(visual persuasion)’ 및 ‘시각적 데이터(visuals)’에 부연하는 ‘주석(gloss)’ 또는 이를 적극적으로 설명하는 ‘텍스트’라고만 표현하고 있어 시각적 데이터가 담고 있는 데이터를 언어적 기호로 전달하는 것을 대상화한 용어까지는 제시하지 못하였다.

본고에서 사용하는 ‘데이터 설명문’이라는 용어는 Swales & Feak(1994)의 data commentary를 번역한 것이다.²⁶⁾ 이 용어를 처음 사용한

26) ‘commentary’라는 단어는 어떠한 대상, 사건에 대해 의견이나 설명을 표현하는 것(Britannica dictionary, 2023.6.7. 검색)으로서 해설, 논평 등으로 번역된다. 보통 종교 서적이거나 고전에 주석을 다는 것으로 그 장르를 삼는 경우도 있으나, 본고에서처럼 시각적 데이터에 대한 변환과 해석으로서 commentary라는 용어를 쓴 것은 Swales & Feak(1994)가 처음이다. 그러나 이 commentary를 별개의 장르에 포함된 것으로 보고 번역한 역어는 없는 실정이며, ‘해설문, 논평문’은 각각 문학 작품에 대한 해설을 담은 글, 특정한 시사 문제나 언론 보도와 관련된 글로 수치적 자료를 글로 표현하는 총체적 의미로서의 데이터 코멘터리와 매우 다르다. 본고에서는 데이터 코멘터리가 시각적 데이터를 언어로 전달하고 설명하는 것을 목적으로 함에 주목하여 ‘데이터 설명문’으로

Swales & Feak(1994:77)는 데이터 설명문을 ‘데이터를 논의하기 위해 표, 그래프, 수치 또는 다른 종류의 비언어적 수단을 주 텍스트와 융합하여 또는 부록으로 첨부하여 제시하는 쓰기의 하위 과업’으로 정의하고 있다(이지연, 2022:244 재인용). 즉, Swales & Feak(1994, 1995)의 연구에서는 ‘데이터 설명문’을 그래프, 표 등의 수치 정보를 담고 있는 비언어적 데이터를 언어적으로 제시하는 것으로 보고 있는 것이다. 이들은 시각적 데이터가 보조적 역할이 아닌 실제 학술적 텍스트에서 갖는 속성을 파악하고자 노력하였다. 이는 비언어적 기호에서 언어적 기호로 그 내용을 전부 옮기는 것이 아니라, 비언어적 기호에서 정보를 선택하여 해석하는 것을 데이터 설명문의 속성으로 보는 것이다. 따라서 Swales & Feak(1994, 1995)에서는 데이터 설명문을 나름의 의미를 표상하는 ‘시각적 데이터와 그를 부연하는 언어 기호’로 간주한다. 즉, 이들 범위를 선정할 때에도 시각적 데이터가 가시적으로 제시되고, 언어 기호가 그에 부연되는 구조를 반드시 포함한다.²⁷⁾

이와 마찬가지로, Nordrum & Eriksson(2015:59)에서도 데이터 설명문을 ‘시각적 데이터에 대한 언어적 해설’로 제시하고 있다. 이 연구에서는 응용 화학 학술지 논문과 같은 분야 석사 학위 논문에서 결과(result) 또는 결과와 논의(result and discussion) 장만을 추출해 데이터 설명문이라 조작적으로 정의하였다. 시각적 데이터와 그 부연이라 하더라도, ‘연구 방법’ 장의 시각적 데이터는 ‘결과’를 나타내는 장의 시각적 데이터와는 다른 속성을 가질 것이라 판단하였기 때문이다(Nordrum & Eriksson, 2015:67). 따라서 학술적 데이터 설명문은 단순히 아무런 시각적 데이터에 부연하는 언어적 양식이 아닌, 특정한 연구 결과를 수치적으로 분석하여 논의하는 성격을 갖고 있는 것이라 할 수 있다.

번역하였다.

27) 예를 들어, Swales & Feak(1994:80)에서는 특정 시각적 데이터에 부연된 학술적 영어 데이터 설명문은 다음과 같이 항상 ‘위치(location)’와 ‘요약(summary)’으로 시작한다고 하였다.

| 위치 | 요약 |
|---------------|--|
| Table 5 shows | the most common modes of computer infection for U.S. businesses. |

따라서 표, 도표, 그림이 묘사하는 데이터가 시각적 기호에서 언어 기호로 변환된다는 점과, 이 과정에서 데이터의 ‘선택적’인 변환이 일어난다는 점, 학문 분과에 대한 저자와 독자의 지식이 이 선택에 대한 해석에 주도권을 행사할 수 있다는 점이 데이터 설명문의 속성이라 볼 수 있다. 즉, 데이터 설명문의 속성은 다음과 같이 정리될 수 있다. 첫째, 다양한 양식이 한 텍스트에 나타나는 다중 양식적 속성이 있다. 둘째, 저자-독자 관계의 주도권을 논하는 체계 기능 이론의 대인적 메타 기능 양상을 보인다. 셋째, 저자와 독자가 데이터를 해석하는 데 개입하는 관념적 메타 기능 및 CEFR(2020)에서 데이터 설명문의 숙달도에 영향을 끼치는 요인으로서 지적하는 인지적 차이가 존재한다. 이들은 서로가 혼재된 복잡한 양상을 띠고 있으며, 이에 따라 체계 기능 이론 및 언어 평가에서 전제하는 인지적 숙달도와 관계가 있다. 이를 통해 데이터 설명문은 다중 양식, 체계 기능적 메타 기능, 언어 평가에서의 인지적 측면을 모두 살펴보아야 적절하게 이해가 된다고 볼 수 있다.

정리하자면, 학술적 영역에서의 데이터 설명문은 ‘특정 연구 결과를 포함하는 시각적 데이터를 해석하여 언어적으로 제시하는 텍스트’로 정의할 수 있다. 그리고 시각적 데이터에서 중요 정보를 추출하고 제시하는 방식은 저자의 의도와 목적에 영향을 받는다. 또한 그 데이터 설명문을 작성하거나 읽는 저자 및 독자의 학문 분과적 지식 수준은 데이터 해석에 영향을 끼친다. 이러한 속성을 분석하기 위한 이론에는 다중 양식 담화 분석, 체계 기능적 메타 기능, 언어 평가의 인지적 부분이 있다. 이는 다음 절에서 더 자세히 확인해 볼 것이다.

2. 체계 기능적 속성에 따른 이동 마디 구성

2.1. 학술적 데이터 설명문의 체계 기능적 다중 양식 담화 분석

(1) 학술적 데이터 설명문의 다중 양식적 속성

일반적으로 텍스트, 문(文)은 ‘문자’로 구축된 하나의 완결된 내용을 뜻한다. 그러나 기존에 문자로만 표현할 수 있었던 내용들이 기술의 발전으로 문자 이외의 여러 방식으로 전달되고 있다. 이는 전통적이고 관습적인 양식을 고수하는 학술 논문에서도 자주 관찰되고 있다.

특히 이공계라 불리는 중학문(hard science) 분야에서의 학술 출판은 종이뿐만이 아닌 온라인 공간에서도 이루어지는바, 기존의 정지된 그림과 표로만 제시할 수 있었던 다양한 실험 내용과 결과가 동영상으로도 삽입이 가능해졌고, 2차원 종이로 출판할지라도 문자로만 제시되었던 논문 초록을 ‘그림 초록(graphical abstract)’이라는 새로운 방식으로 요구하기도 한다(Elsevier, 2023.7.13. 검색). 이미 학술 출판 분야에서는 언어적 기호뿐만이 아닌 다양한 양식의 기호들이 함께 사용되어 학술적, 장르적 의미를 구성하는 다중 양식성(multimodality)이 표면화되기 시작한 것이다. 경학문(soft science) 분야의 학술 출판에서는 아직까지는 2차원의 종이 또는 텍스트 유형 파일이 주로 사용되는 매체이기 때문에 학술적 내용을 전달할 때 언어적 기호와 더불어 정적인 시각적 요소, 즉, 표, 그래프, 다이어그램, 수식 등이 주로 사용됨을 볼 수 있다. 동작성은 떨어지지만, 경학문에서도 역시 학술적 내용의 의미 전달과 이해에 시각적 데이터가 중요한 자원으로서 기능함을 볼 수 있는 것이다.

따라서 학술적 텍스트에서 그림, 그래프 등의 시각적 데이터가 언어적 기호인 글과 어떠한 관계를 이루고 있는지, 그 특성은 무엇인지, 어떻게 구현되고 있는지, 궁극적으로 이들이 어떻게 의미와 장르적 관습을 형성하는지에 주목할 필요가 있다. 본고에서 데이터 설명문을 다루는 것도 학술 장르에서 시각적 양식과 문자라는 양식이 학술적인 텍스트를 조직하는 데 어떠한 역할을 하는지 살펴보려는 다중 양식적 연구의 흐름 가운데 자리한다. 이러한 다중 양식적 연구는 시각적 이미지도 문자 못지않게 의미를 전달하는 기호적 양식임을 주장하는 사회기호학적 의사소통 이론에서 출발하였다. 이에 대해 Fei(2004:220)에서는 “의미가 문자 단독으로만 구성되지 않으며, 사회는 기호적 자원을 혼합하고 공동 배치하여

의미를 만든다”라고 하였다.

이러한 관점을 학술적 텍스트에 적용한 연구들은 다음과 같다. 우선 기존 장르 이론에서 중요하게 다루지 않았던 시각적이고 위상적 특성들을 찾아내 연구한 Lemke(1998), O'Halloran(2008a)에서는 텍스트에서 미니 장르(mini-genre)라 명명한 시각적 요소와 그의 텍스트 내 위치에 집중하였다. Lemke(1998)는 이들 중 시각적 데이터인 도표와 표는 현상의 지속적 변화와 시공간적 관계를 서술하는 데 필요한 요소라 하였으며, Nordrum & Eriksson(2015:61)에서는 효과적인 다중 양식 조화(multimodal ensemble)를 위해 작문에서 얼마나 시각적 요소가 언급되어야 하고 변환되어야 하는가를 저자가 잘 알고 선택해야 한다고 하였다. 즉, 한 장르를 관습적으로 구성하기 위해 시각적 요소의 배열과 그를 글로 해석하는 정도가 의도적으로 선택된다는 것이다. 이에 더해 Roth(2013:162)는 다양한 양식의 시각물(graphic) 제시를 읽는 것과 주장하는 것 및 사용 가능한 원본 데이터 간의 관계 이해를 융합하여야 데이터 설명문을 쓸 때 정보를 취사선택할 수 있다고 하였다. Guo(2004)는 생물학 교재의 분석에서 통계적 그래프가 두 변인 간의 수치적 관계를 시각적으로 다룬다는 데서 두 개체를 유형적으로 묶어 지향하는 역할을 한다고 하였다. 또한 이 논문에서 그래프는 데이터에 따라 신뢰 여부를 알 수 있지만, 그 모습 자체로 신뢰성을 나타내는 힘을 가지며, ‘여기를 보라’는 무표적 명령을 지니게 된다고 하였다(Guo, 2004:203).

위 연구들에서는 학술적 텍스트 내에서 시각적 데이터의 배치와 해석에 관련된 분과 지식의 정도, 저자-독자 관계에 대한 속성과 의미를 주장하고 있다. 이를 뒷받침하기 위해 이들 연구에서는 학술적 텍스트 구성에서 나타나는 언어적 기호-시각적 요소의 융합과 그 의미를 구체적으로 분석하기 위해 체계 기능 언어학의 메타 기능적 원리를 적용한 다중 양식 담화 분석(Multimodal discourse analysis: MDA 또는 MMDA)을 사용한다.

‘체계 기능적 다중 양식 담화 분석(Systemic functional-multimodal discourse analysis, SF-MDA)’은 Halliday(1978, 2004)의 체계 기능 이론

에서 주장하는 관념적 기능(ideational function), 대인적 기능(interpersonal function), 텍스트적 기능(textual function)의 세 메타 기능을 분석의 기준으로 삼고, 이들을 문자 기호를 넘어선 다른 기호들의 분석에까지 확장한 것이다. 여기서 다중 양식은 문자, 소리, 시각적 이미지, 수학적 기호, 조각, 건축, 제스처 등 다양한 양식의 기호들이 얽혀 의미를 형성함을 뜻한다(O'Halloran, 2008a:444). 이러한 접근은 언어와 시각적 요소가 혼합된 텍스트를 분석할 때 기호적 자원이 관념적 의미를 제공하는 동시에 사회적 관계를 발생시키며, 이들은 텍스트적 메타 기능인 담화 구성을 통해 파악할 수 있다는 주장을 기반으로 한다(Lemke, 1998:91; O'Halloran, 2008a:444). 따라서 체계 기능적 다중 양식 담화 분석의 분석 범위는 매우 다양하다. 이 중 본고의 위치는 다음 그림과 같이 '언어적 자료 중심'과 매우 가깝다.



<그림 2> 학술적 데이터 설명문 연구의 비언어-언어 스펙트럼상 위치

그렇다면 언어적 자료 중심의 체계 기능 언어학을 적용한 기존 학술 장르 텍스트 분석 연구를 차용할 수 없는 것인가 확인해 보아야 한다. 장르를 구분하는 이론 틀로서 체계 기능 언어학은 필자/화자의 언어적 선택과 그 표현의 기저에 어떠한 사회적/기호적 함의가 있는지 알아내는 연구들의 기반이 되었다(이슬비, 2016; 신범숙, 2020a; 소지영, 2020). 그러나 본고에서는 기존 연구에서의 접근과는 다르게 개별적 언어 분석에 체계 기능 언어학을 사용하지 않고, 이동 마디 구축에 필요한 이론 틀로서 체계 기능 이론의 메타 기능만을 사용하도록 한다.

이러한 이유로는 첫째, 본고의 대상이 되는 데이터 설명문은 문자 기호만을 분석해서는 그 구성 체계와 의미를 제대로 파악할 수 없기 때문이다. 체계 기능 언어학적 이론에 기대어 언어적 측면만을 분석하게 된다.

면 데이터 설명문의 언어 기호는 시각적 데이터의 반복 그 이상의 의미는 없다. 예를 들어, 체계 기능 언어학에서 관념적 기능을 언어적으로 구현하는 방식은 동성(transitivity) 체계에 따라 ‘과정(process, 서술어), 참여자(participant, 서술어의 필수적 보충어), 배경(circumstance, 주로 부사어로 구현)’과 같은 구체적인 언어 단위를 언급하는 것이다(이관규 외, 2021:23). 그러나 본고에서 분석하는 데이터 설명문에서는 표 1개와 그래프 2개를 설명할 것을 요청한다. 따라서 이 데이터 설명문의 참여자(participant, 필수적 보충어)는 표와 그래프의 변인에 해당하기 때문에 동성을 판단하기 위한 참여자가 대단히 많으며 이들이 두 번 이하로 출현하는 경우가 많아 분석이 용이하지 않다. 이에 수반되는 과정(process, 서술어) 또한 동일한 어휘와 표현이 반복되어 분석이 풍부하지 않다는 단점이 있다. O’Halloran(2008b:106)에서는 특히

$$f'(x) = \lim_{h \rightarrow 0} \frac{\sqrt{x+h-1} - \sqrt{x-1}}{h}$$

와 같은 수식에 대한 체계 기능적 다중

양식 담화 분석에서는 체계 기능 언어학적 분석과는 달리 x, h, 1과 같은 요소를 ‘주요 참여자(key participant)’로 두고, 극한(limitation) 풀이 과정을 ‘전개적 과정(operative process)’으로 따로 명명해 언어에서 서술어로 대치할 수 있는 ‘과정(process)’ 개념이 수식에서는 매우 다르게 적용됨을 보여주었다.

이에 더해 대인적 기능에서 가치 판단을 나타내는 태도는 언어로 나타나기보다는 시각적 데이터의 존재 자체로 설명되는 측면도 있다. 학술적 텍스트에 나타나는 수치 정보를 담은 시각적 데이터는 그 사실성 여부와는 관계없이 독자와의 대인적 관계에서 신뢰성을 제시하며, 해당 시각적 데이터를 보라는 텍스트 필자의 명령적 어조 역시 담고 있다(Guo, 2004). 그러나 이러한 작용은 모두 언어 기호로 표현되지 않는다.

둘째, 본고에서 다루고자 하는 외국인 대학생의 데이터 설명문은 한국어 모어로 사용하지 않는 필자에 의해 작성되었다. 즉, 외국인 대학생은 한국어 자원 및 한국 사회(또는 집단)에 대한 사회적 기호 해석 자원이 부족하다. 그 때문에 이들을 대상으로 외면적으로 나타나는 텍스트의 구조를 통해 잠재적인 의도와 힘을 파악하고 이를 변환하거나 재생산할

수 있는 사회적 함의를 깨치는 것을 교육적 목표로 잡기에는 조금 더 시간이 필요하다. 한국어가 모어가 아닌 필자들이 더 초점을 두어야 할 부분은 텍스트의 관습성이 명시적으로 어떻게 표현되어 있는지 학습을 통해 언어 자원을 축적하는 것이라 할 수 있다. 따라서 언어 및 기호적 선택에 선행하는 잠재적 의미를 인식하고자 하는 교육적 접근은 한국어 비모어 필자/독자들에게는 한국어 모어 필자/독자와는 다른 방식으로 적용되어야 한다.

따라서 본고에서는 체계 기능 이론의 메타 기능적 요소를 주요한 분석의 틀로 사용하나, 그 개념적인 면에서 시각성을 포함한 텍스트 분석에 필요한 바를 선택한다. 또한 텍스트의 언어적인 부분에 국한하여 논의하는 체계 기능 언어학의 분석 기준은 시각적 요소 및 장르 내포적 요소를 적극적으로 드러낼 수 있을 때만 부차적으로 사용한다. 그러나 체계 기능적 다중 양식 담화 분석에서 사용하는 메타 기능이 어떠한 의미로 해석되는지 그 범위를 알아보기 위해 다음 항에서는 스펙트럼의 다양한 위치에 존재한 체계 기능적 다중 양식 담화 분석 연구도 고찰한다.

(2) 학술적 데이터 설명문의 메타 기능적 의미

본고에서는 체계 기능 이론에서의 메타 기능적 의미가 시각적 데이터와 함께 등장하는 텍스트에서는 어떻게 적용되는지 살펴볼 것이다. 특히 시각적 데이터의 메타 기능적 해석은 구체적인 측면에서 학자들마다 조금씩 견해가 다르다. 따라서 <그림 2>의 스펙트럼 전체에서 시각적 데이터가 갖는 메타 기능 의미를 정리해 볼 것이다. 이를 위해 체계 기능적 다중 양식 담화 분석 연구인 Miller(1998), Lemke(1998), O'Halloran(1996), Kress & Leeuwen(2006)을 통해 학술적 텍스트 내에서 시각적 데이터와 언어적 기호가 함께 나타나는 현상을 어떠한 메타 기능적 의미를 부착하여 분석하는지 확인할 것이다.

Miller(1998)는 과학 학술지 논문을 대상으로, Lemke(1998)의 경우 과학적 학술적 텍스트를 대상으로 하였으며, O'Halloran(1996)의 경우 수학적

학술적 텍스트 및 수학 교재를, Kress & Leeuwen(2006)은 표와 다이어그램에 있어서는 과학적 학술적 텍스트에서의 시각적 데이터 특성을 논하였다. 이러한 학술적인 시각적 데이터의 특성은 경제/경영 학술적 텍스트에서도 유사하게 나타난다. 구분관(2014:51)에서는 경제학 학술 논문 텍스트를 분석하며 그 특성을 자연과학에서처럼 수리적 모형화를 추구하는 것이라 하였다. 특히 경제/경영 학술적 텍스트에서 발견되는 데이터 설명문의 시각적 데이터들은 변수들의 수치적 관계에 중점을 둔 그래프 및 통계 결과를 나타낸 수치 표, 모형을 설명하기 위한 수식 등이 대부분으로, 과학적 학술적 텍스트에서 다루는 시각적 데이터와 크게 다르지 않다. 따라서 경제/경영 학술적 텍스트의 데이터 설명문에 해당하는 부분은 경학문 계열 중에서도 중학문과 유사한 시각적 데이터를 사용한다 할 수 있다. 이에 본고의 분석에서도 중학문을 대상으로 한 위 선행 연구들의 ‘그래프, 도표, 표’에 대한 메타 기능적 해석을 차용할 수 있다고 보았다.

체계 기능적 다중 양식 담화 분석은 체계 기능 언어학의 언어적 분석과는 달리 시각적 요소의 분석을 위해 추상적 수준에서 조정된 접근을 요한다. Eggins(2004/김서형 외 역, 2021:49-50)에서는 할리데이의 관념적 메타 기능의 의미를 언어학적 수준에서 ‘경험을 나타내는 방법에 대한 의미’라 하였고, 대인적 메타 기능의 의미를 ‘다른 사람과의 역할 관계에 대한 의미와 서로에 대한 태도에 관한 의미’로, 텍스트적 메타 기능의 의미를 ‘말하고 있는 것들이 어떻게 서로 연결되어 있고, 어떻게 이전에 말한 것, 우리 주변의 맥락과 관련되어 있는지에 대한 의미’로 정의하였다.²⁸⁾ 이들 메타 기능은 하나 이상이 함께 출현하는 경우가 있다(이관규

28) 체계 기능적 다중 양식 담화 분석 연구자들은 할리데이의 메타 기능을 조작적으로 정의하여 자신의 연구에 사용하였다. Miller(1998)와 Kress & Leeuwen(2006)은 할리데이의 용어 그대로 ‘관념적(ideational), 대인적(interpersonal), 텍스트적(textual)’이라는 용어를 사용하였고, 많은 부분 체계 기능 언어학의 용어와 개념을 사용해 텍스트를 분석하였지만, Lemke(1998)에서는 동일한 대상을 ‘제시적(presentational), 지향적(orientational), 구성적(organizational) 기능’이라는 용어로 제시하며, O’Halloran(1996)에서는 ‘표현적 기능(representation function)’, ‘양식적 기능(modal function)’, ‘조직적 기능(compositional function)’으로 조정하여 제시한다.

외, 2021:49).

구체적으로, 관념적 메타 기능은 경험적/논리적 두 개의 요소로 나뉜다. 이들은 각각 실세계에서의 경험과 실세계에서의 논리적 관계를 뜻한다 (O'Halloran, 1996:45). 따라서 시각적 데이터에 등장하는 변수들의 논리적 비교, 올바른 시각적 데이터의 해석을 돕기 위한 장의 표시, 시각적 데이터가 내포하는 분과적 지식, 시각적 데이터의 관습적 양식과 같은 학술적 실재를 이해하기 위한 지식으로 이루어진다. Miller(1998:36)에서는 과학 학술지에서 시각적 데이터를 사용하는 것이 글로 표현하기 어려운 관념적 메타 기능, 즉, 논리적인 요소를 나타내는 것이라 본다. 따라서 과학 학술지의 데이터 설명문은 원인과 결과를 의미하는 변수들의 논리적 비교 관계를 강조한다고 하였다.

한편 Lemke(1998)는 과학 학술적 텍스트에서 언어적 기호와 비언어적 기호의 '배치'를 중심으로 메타 기능적 의미를 해석하였다. Lemke(1998:92-94)에서는 언어가 주가 되는 과학 학술적 텍스트는 시각, 제스처, 그래픽, 수식 등 다양한 자원 체계를 배치하며 이들이 상호작용하여 멀티미디어 장르를 구성한다고 하였다. 이때 위에서 언급한 세 메타 기능이 각 기호적 양상의 기능적 자원을 활용하여 의미를 형성한다고 보았다. 구체적으로, 관념적 측면에서 과학 학술적 텍스트의 시각적 데이터는 텍스트의 다른 요소와의 관계를 통해 대상, 행위, 관계의 구성 등을 보여줄 수 있다. 즉, 어휘적 이름과 도표적 이미지는 함께 작용하여 그 장면의 전형적 스크립트 측면을 보여주는 표시적 기능을 한다.

학술적 데이터 설명문 내의 수학적 시각적 데이터에 집중한 O'Halloran(1996:161)의 연구는 수학에서의 체계 기능적 공간 체계를 정리하며, 표현적 기능 관점(관념적)에서 그래프와 다이어그램의 기능은 기하학적 배치와 같은 사실적 인식, 벤다이어그램/데이터 그래프와 같은 수학적 상징이 나타내는 현실에 대한 인식, 곡선과 직선으로 나타나는 지속적인 공변인에 대한 위상적 양식, 수식과 그래프의 변환 및 관계, 변인 패턴 간 비교 등이 있다고 하였다. 독자들은 위와 같은 그래프, 다이어그램의 기능적 배치와 관습성을 경험적 의미로 소화하게 된다

(O'Halloran, 1996:169). 예를 들어, 그래프는 변인들끼리의 패턴을 글이 표현할 수 없는 방식의 추세로 나타낸다. 이들의 인식은 그래프 변인 간 패턴에 대한 배경지식, 즉, 눈금이 어떠한 배열인가, 축의 의미는 무엇인가, 그래프의 교점은 무엇을 나타내는가와 같은 경험적 의미에 근거해 이루어진다. 여기에서 전문가 독자와 새로운 독자의 경험적 차이가 드러나게 된다. 이러한 양식성에 대한 경험적, 논리적 지식은 관념적 요소로 치환될 수 있다.

Kress & Leeuwen(2006)에서는 특히 시각적 데이터 자체의 속성과 그 구성 및 배치에 집중하여 이들의 관념적, 대인적, 텍스트적 기능을 제시하였다. 특히 이들은 그래프, 도표, 표에서 '발생한 사건은 어떠한 시간적 단계를 거치는가(특히 왼쪽에서 오른쪽으로의 이동)', '부분 수치들이 전체 수치들에 대해 가리키는 총체적 의미는 무엇인가(원 그래프의 부분적 구성과 그 합)', '상황이 진보/퇴보했는가(선 그래프의 증감 표시)', '수량의 차이가 나타내는 바는 무엇인가(참여자 속성의 다층적 비교)' 등의 질문을 통해 관념적 측면을 파악하고자 하였다(Kress & Leeuwen, 2006:59).

학술적 데이터 설명문에서 대인적 메타 기능은 전문적 지식을 두고 저자와 독자가 어떻게 주도권을 갖느냐에 집중한다. 많은 연구들에서 주로 저자의 태도와 주장에 따른 특정 부분의 강조 서술 및 시각적 데이터 배치를 통한 시선 집중 명령이 저자의 우세함을 나타낸다고 설명한다. 예를 들어, Miller(1998:32)는 과학 학술지의 시각적 데이터에서는 저자의 주장과 관점을 강조하고 나머지 발견이나 견해를 주석으로 처리하거나 아예 무시하는 것으로 축소한다고 하였다. 그러나 이러한 행위는 독자의 전문적 수준에 따라 달라질 수 있다. 본문의 텍스트들은 저자의 의도에 따라 독자들을 인도하기 때문에 숙련되지 않은 독자는 저자의 해설에 의존할 수밖에 없다. 그러나 시각적 데이터의 아래 또는 다른 페이지에서 주석으로 제시되는 간략한 언어적 설명은 전문 독자가 시각적 데이터를 스스로 판단할 수 있도록 하는 것이라 하였다.

Lemke(1998)에서 다루는 대인적 메타 기능은 시각적 요소가 촉진하는

평가적 지향을 알아볼 수 있는데, 저자/독자가 자신의 배경에 따라 장면
에 대해 지니는 긍정적, 부정적 지향을 의미한다. 텍스트 생산자와 해석
자는 이미지가 제시된 장면에 대해 ‘태도(stance)’를 구성한다. 이를 통해
해석자는 생산자가 어떻게 전달하려는 바를 확신하는지를 판단하며, 전
달하려는 바에 대해 호/불호, 평범/비범, 필요/의무와 같은 평가적 지향
을 갖게 된다.

O’Halloran(1996)에서 제시하는 양식적 기능(대인적)에서는 텍스트와 독
자의 관계에서 의무의 양상이 거의 없다. 다만 존재한다면, 수학적 시각
적 데이터들은 대인적으로는 주로 무표적 명령형인 ‘여기를 보라’ 또는
직선과 관련되어 ‘여기까지만 해석하라’ 정도의 한정성만이 존재한다. 이
는 또한 주목의 요소를 강조하는 역할도 도맡고 있는데, 독자는 곡선,
축, 변인의 속성, 방향과 같은 그래프의 특정한 특징에 직접적으로 시선
을 주며, 여기에는 라벨/제목과 같은 텍스트, 폰트 사이즈, 색, 음영, 선
모양, 화살표, 크기 등 명시적으로 표시된 특징이 부여된다. 이러한 시선
사로잡기는 각 요소의 상호작용이 두드러진 것(prominence)이어야 한다.
글을 읽는 ‘독자’가 아닌 그 내용을 사용하는 ‘사용자’는 독자와는 다른
띄엄띄엄 읽는 경로를 보여준다(Kress & Leeuwen, 2006). 따라서 이들
의 시선을 사로잡는 두드러진 것은 관념적 의미를 표상하는 시각적 데이
터일 수밖에 없다. 수학적 텍스트에서 시각적 요소의 관습적 균일화는
신뢰성과 관계가 있다. 그래프 크기나 범주의 차이를 조정하여 두드러진
특징을 가리는 것은 관습적 균일화에 맞지 않아 그래프가 나타내는 신뢰
성을 하락시킨다(O’Halloran, 1996:161-168).

Kress & Leeuwen(2006:114-116)에서는 시각적 데이터를 분석할 때
대인적 관점에서 저자와 독자 간 상호작용을 고려해야 한다고 하였다.²⁹⁾
저자와 독자는 직접적 관계가 없기 때문에 서로의 상을 추정하여 고정된

29) 해당 연구에서는 시각적 데이터 내 참여자(participants, 그래프의 변인에 해당) 간 관
계 및 시각적 데이터 내 참여자와 생산자, 해석자의 관계도 다루고 있지만 전자는 주로
관념적 의미에서 설명되고, 후자는 다이어그램이나 그래프보다는 인물화와 같은 그림의
시선을 대상으로 설명하므로 본고의 논의에 적합하지 않아 다루지 않는다. 더 자세한
내용은 Kress & Leeuwen(2006)을 참고하라.

다. 따라서 저자가 자신의 시각적 데이터를 더 수용 가능하게 만들고 싶다면 사회적 기관의 맥락에서 엄격하게 정의된 관습과 가치, 믿음을 따라 그 안에서 작품이 생성되고 순환되게 해야 한다. 다만 시각적 데이터는 독자가 메시지를 읽기만 해야 하는 것에 반해 생산자는 메시지를 더 유창하고 빈번하게 생성하고 독자의 피드백을 반영해 읽고 쓰기를 모두 할 수 있으므로 이는 저자 우위의 특수한 행위라고도 평가된다.

또한 시각적 의사소통에서 중요한 양식성(modality)은 메시지의 신뢰성과 관련이 있다. 이 신뢰성은 절대적 진실보다는 생산자와 해석자가 공유하는 상대적 진실과 관련이 있기 때문에 언어학에서는 정도의 형용사 또는 부사(영어의 경우 조동사)로 표현된다. 다만 장(field)에 따라 표현이 달리 수용되는 경우가 종종 있다. 예를 들어, 과학적-기술적 이미지의 사실성과 신뢰성은 ‘이것을 사용할 수 있는가’, ‘이것에서부터 실제 차원을 측정할 수 있는가’, ‘우리가 이것으로부터 어떻게 실험을 설정할 것인가’를 찾아낼 수 있는가’ 등의 질문에 기반을 두고 있다(Kress & Leeuwen, 2006:154-173). 특히 Kress & Leeuwen(2006:165)에서는 “사회문화적 엘리트에 의해 사용되는 추상적 코딩은 이미지가 총체성에 대해 개별성을 낮추고 필수적 자질에 대한 구체성을 낮출수록 올라간다. 이러한 텍스트를 생산하고 읽을 수 있는 능력은 사회적 구별짓기의 표지로서 이 코딩 방향에 근거한다”고 하였다. 즉, 학술적 텍스트의 추상적이고 단순한 시각적 데이터 사용은 학문적 훈련을 받은 저자와 독자의 관계를 상징한다.

마지막으로 텍스트적 메타 기능은 다른 메타 기능들에 있는 메시지를 조직하여 단일하게 전체를 구성하는 역할을 한다(Miller, 1998:32). 시각적 데이터에 대한 텍스트적 메타 기능은 주로 위치와 경계, 색 등이 주요 요소이다. 본고의 학습자 데이터 설명문에서는 학습자들이 직접 그래프, 표, 도표 등을 만들거나 제시한 것이 아니기 때문에 이러한 양상은 많이 나타나지는 않았다. 그러나 시각적 데이터를 언급하거나 제목 등을 사용하여 특정 내용이 들어가는 위치를 분절하는 것, 또는 특정 요소를 현저하게 많이 사용하여 독자의 시선을 끄는 것 등은 살펴볼 수 있다고

본다.

구체적으로 Miller(1998:32-40)에서는 텍스트적 메타 기능에 속하는 요소로 테마(thema)와 레마(rhema), 구정보(given)와 신정보(new) 등을 제시한다. 시각적 데이터에서 테마는 주제가 아닌 범위, 또는 초점의 변화를 지적한다. 예를 들어, ‘<그림 2>에서는…’과 같은 어구가 테마에 속하며, 논의의 시작점이 그림 1에서 그림 2로 완전히 옮겨갔음을 뜻한다. 또한 새로운 정보를 담고 있는 시각적 데이터는 항상 구정보의 오른쪽에 위치해야 한다.

이와는 달리 Lemke(1998:93-94)에서 다루는 시각적 데이터의 텍스트적 기능은 텍스트 구성을 시각적 요소와 텍스트의 배치, 색, 글자 크기 등의 표현 양식으로 추구하는 것으로 볼 수 있다. 예를 들어, 시각적 데이터를 텍스트 캡션과 병치할 때 이들의 크기가 상대적 중요성을 결정할 수 있는 조직적 기능을 한다. O’Halloran(1996:161-180) 역시 Lemke(1998)와 같이 시각적 데이터의 배치와 조직에 초점을 맞춘다. O’Halloran(1996)의 조직적 기능(텍스트적)에서는 ‘제시되는 것’과 이것들이 ‘얼마나/어떻게 독자들에게 전달되는가’의 측면에서 의미를 구성한다. 시각적 데이터들의 배치에서 나타나는 수평선, 수직선, 대각선의 의미를 파악하기 위해서는 그 배경까지 보아야 하는 게슈탈트 이론을 끌어올 수 있다. 이러한 시각적 요소의 위치, 2차원/3차원에서의 원근감, 직선과 곡선의 사용, 이들 요소의 과정에 이름을 붙이기 위한 상호 연결, 본문과의 연결을 담당하는 응집성, 과정을 나타내는 다이어그램 종류 등이 시각적 요소를 포함한 텍스트 조직에 영향을 미친다. 텍스트 조직에서 본문의 언어적 텍스트와 시각적 배치는 제목, 캡션과 같이 특수한 언어 텍스트와만 어울려 공간적으로 분리된 위치에 존재한다. 다만 본문의 언어 텍스트에서는 이들 시각적 배치의 참조를 명시적으로 제시하고, 이를 숫자와 선 같이 언어와 상이한 기호로 반복하는 역할을 한다.

Kress & Leeuwen(2006)에서 시각적 데이터의 텍스트적 접근에는 세 가지 체계가 협응하여 관념적, 대인적 의미를 구성한다. 첫째, 정보 가치는 이미지의 자리를 통해 부여된다. 메시지의 출발은 주로 왼쪽, 위, 중

심이며, 메시지에서 드러내고자 하는 바는 오른쪽, 아래, 가장자리에 나타난다. 따라서 이미 주어진 구정보는 전자에, 새로 나타나는 신정보는 후자에 배치되는 경향이 있다. 즉, ‘모두가 이미 알고 있는 것’, ‘상식적이고 보편적인 내용’, ‘이상적인 가치’는 구정보로서 왼쪽, 위, 중심에 자리하고, ‘문제시되는 것’, ‘논쟁할 만한 것’, ‘이슈에 대한 정보’, ‘실제적인 것’은 신정보로서 오른쪽, 아래, 가장자리에 나타난다. 둘째, 현저성 요소는 해석자의 시선을 끄는 것과 관련이 있다. 전경과 배경, 상대적 사이즈, 톤/색의 대비, 날카로움의 차이 등과 같은 요소 배치의 구현과 관계가 있다. 현저성은 요소들 중 가장 중요한 위계를 형성하며 이들 중 하나를 선택하는 것이 더 중요하고 집중을 요함을 뜻한다. 셋째, 경계를 설정하는 프레임(framing)이다. 프레임 내 요소의 개별성이 더 강할수록 분리된 단위로 제시되고, 분리성이 없는 것은 단체적 정체성을 강조한다. 이들 프레임의 연결 시 더 강조되어 시선을 끄는 현저한 요소에서 덜한 요소로 옮겨간다. 텍스트가 논문과 같이 소제목, 강조 도구(기울임꼴, 굵은 글씨, 밑줄), 전형적 요소의 줄 또는 특정 현상의 특징, 표, 다이어그램 등을 더 많이 사용할수록 텍스트를 순차적으로 읽기보다는 비선형적으로 ‘사용하는’ 읽기 과정이 나타난다. 이러한 텍스트 구성은 독자에게 그 순서나 연결을 맡긴다는 점에서 상호작용적이다(Kress & Leeuwen, 2006:175-208). 다만 이들의 분석은 일반적인 이미지-텍스트 배열에서 유효하므로 시각적 배치의 기준과 방법이 정해져 있는 학술적 텍스트에는 적합하지 않은 기준일 수 있다. 다만 표의 음영이라든지 시각적 데이터의 자리를 여백 등으로 배치하는 것, 이들을 응집하기 위한 고정적 표현이 있는 점 등은 이와 연결하여 사용할 만하다.

언어 기호 외 다른 기호가 함께 나타나는 학술적 텍스트를 분석하는 기준으로서 체계 기능적 다중 양식 담화 분석에 대한 위 내용을 정리하면 다음 <표 10>과 같다.

<표 10> 학술적 텍스트에서 시각적 데이터의 체계 기능적 의미

| | 관념적 | 대인적 | 텍스트적 |
|-----------------------|--|---|--|
| Lemke(1998) | ·시각적 데이터와 그를 나타내는 언어적 이름이 장면의 전형적 스크립트 형성 | ·생산자/해석자가 자신의 배경에 따라 장면에서 지니는 긍정/부정 지향 표현 | ·시각적 데이터와 텍스트 병치 시 배치, 색, 글자 크기 등을 통해 중요성 등 표현 |
| O'Halloran(1996) | ·그래프, 다이어그램이 나타내는 분과 학문적(수학적) 의미 ·전문 독자의 경험적 지식이 의미 이해에 중요 요소 | ·독자의 주목 요소 강조 ·두드러짐을 표시하는 관습적 방식에 기반하여 독자의 신뢰성 추구 | ·본문의 언어 텍스트와 구분된 시각적 요소의 위치 ·본문과 시각적 데이터의 연결을 위한 제목, 캡션, 참조 ·기호를 달리한 내용의 반복 |
| Kress & Leeuwen(2006) | ·그래프, 다이어그램이 나타내는 기호적 의미(요소 간 거리, 방향, 위치, 시간, 포함 여부, 수량, 수치 등) | ·해석자보다 생산자 우위의 텍스트 및 시각적 데이터 ·내용에 신뢰성을 나타내는 양식성은 집단이 공유하는 특성이며 교육받은 정도에 따라 다름 ·학술적 텍스트의 경우 추상적이고 단순한 시각적 데이터 사용 | ·배치의 위치에 따른 정보의 가치 차이 ·시각적으로 현저하게 나타나는 요소가 다른 요소에 비해 더 중요 ·독자의 비선형적 읽기와 경계를 위한 시각적 데이터를 현저성 요소로 제공 |

종합하자면, 위 학자들이 제시한 학술적 텍스트에서 데이터 설명문 구성 요소의 메타 기능은 다음과 같이 정리할 수 있다. 첫째, 관념적 메타 기능은 경험적인 것과 논리적으로 나뉘는데, 과학과 수학의 성격을 지닌

학술적 데이터를 시공간적 관계 형태의 시각 이미지로 배열하는 것은 주로 논리적 의미를 발생시키는 것으로 보인다(O'Halloran, 2008b:134). 따라서 과학적, 수학적 성격의 시각적 데이터에서는 세상에 대한 경험과 논리를 그래프, 다이어그램과 그를 이해할 수 있는 논리적 관습성으로 표상한다고 볼 수 있다. 그래프와 다이어그램, 표, 수식을 구성하는 요소들과 그들이 결합했을 때 나타내는 의미를 이해할 수 있는 것이 과학과 수학 영역에서는 관념적 메타 기능이라 할 수 있을 것이다. 그러나 이들이 다른 분과까지 확장하게 된다면, 그래프와 다이어그램, 표가 지니는 의미 역시도 해당 학문의 이론에 따라 달리 정의될 수 있다. 예를 들어, 연도별 성별 가사 노동량을 나타내는 선 그래프의 남녀 직선 차이가 오른쪽으로 갈수록 줄어드는 것은 남녀의 가사 노동 시간 격차가 줄어들어 성역할의 사회적 고정성이 탈피되고 있는 것으로 해석할 수 있다. 이렇듯 그래프와 다이어그램, 표, 수식에 투영되는 전문적 분과 지식은 관념적 메타 기능을 이루는 것이라 볼 수 있다.

둘째, 대인적 메타 기능은 표상된 시각적 데이터의 은유적, 관습적 의미를 표현하고 이해할 수 있는 필자/독자인지, 그들의 시선을 끌기 위한 주목 요소를 어떻게 구현하는지, 이것이 형성하는 필자/독자의 태도는 어떠한지를 중시한다. 이는 텍스트가 전달하는 메시지의 '신뢰성'과 관계가 있다. 학술적 텍스트에서, 본문 사이에 그래프를 삽입하는 관습적 시각 제시 형태를 고수한다면, 설사 그 내용이 사실이 아닐지라도 첫 인상에서는 신뢰성 있는 실험 결과로 인식하게 된다(Lemke, 1998:107). 또한 텍스트 필자는 이미지와 텍스트를 조합하여 독자에게 자신의 관점을 관철할 수 있는 위치에 있게 된다. 특히 텍스트 필자는 색, 크기, 이미지화 등의 요소를 사용하여 중요성을 강조할 수 있고, 직접적 참조 명령을 할 수도 있다. 그러나 이러한 방식은 예상 독자에 대한 필자의 실제 태도에 따라 위압적일 수도, 설득적일 수도 있다. 예를 들어, 예상 독자가 글을 평가하는 동료 평가자이거나 교수자라면 필자의 태도는 강요적이기보다는 설득적일 것이다.³⁰⁾ 따라서 대인적 메타 기능은 관습적 기호 사용과

30) 이러한 필자 태도의 언어적 구현은 본고의 분석 방향과는 달라 상세 기준으로 다루지

배열을 통해 독자에게 신뢰를 줄 수 있는 행위라 정리할 수 있다.

셋째, 텍스트적 메타 기능은 위에 다룬 관념적, 대인적 기능을 보강하는 역할을 하며, 일관성과 응집성을 가지고 있어야 한다. 특히 이들은 텍스트 장르의 관습성을 따라야 하는데, 언어적 주장 제시를 따라가기 위해 이는 매우 중요하다(Lemke, 1998:101). 예를 들어, 시각적 데이터는 언어적 본문과는 구분된 자리를 마련해야 한다든지, '<표 3>을 참조하라'와 같은 관습적 방식으로 시각적 데이터와 언어적 본문을 연결하는 응집성을 가진다든지 하는 것이 그것이다.

2.2. 기존 데이터 설명문 이동 마디에서의 메타 기능적 의미

체계 기능적 다중 양식 담화 분석은 시각적 양식을 포함한 다중 양식이 지니는 의미와 효과, 조화, 배열 등에 집중하여 메타 기능적 의미를 추출하는 데 사용된다. 그러나 이러한 양상은 연구자의 역량에 따라 텍스트에서 부분적으로 추출할 수밖에 없다. 이는 체계 기능 언어학에서도 마찬가지인데, 언어 기호로 이루어진 텍스트의 일부에서 나타나는 메타 기능적 의미를 찾아내기 때문에 전체 텍스트에서의 메타 기능 양상 파악이 어렵다.

본고에서 다루는 데이터 설명문은 다중 양식 연구 중에서도 언어적인 면에 기울어져 있지만, 시각적 기호와 언어적 기호의 배열이라는 구조가 메타 기능적 의미를 지니기 때문에 텍스트 전체 구조를 조망하는 작업을 간과할 수 없다. 즉, 학술적 데이터 설명문의 다중 양식성은 배열과 같은 구조에 기인하기 때문에 메타 기능적 의미를 전체 구조에 적용해 분석해야 함을 뜻한다. 이에 따라 본고에서는 시각적 데이터와 언어적 텍스트를 구조적으로 포괄하고, 이에 메타 기능적 의미를 부착하기 위해 거시적 구조로서 이동 마디를 제시한다. 특히 앞에서 언급한 바와 같이 장르

는 않는다. 학술적 텍스트의 필자 태도에 대해서는 이슬비(2016)에서 상세히 다루고 있으므로 이를 참조하라.

내포적 성격을 갖는 데이터 설명문은 단독으로 떼어서는 보편적인 이동 마디 순서로 표현하기 어렵다. 따라서 데이터 설명문의 속성을 반영할 수 있도록 이동 마디에 체계 기능 이론의 메타 기능적 의미가 추가적으로 설정되어야 한다.

그렇다면 우선 데이터 설명문의 이동 마디에 대한 연구는 어떻게 이루어졌는지부터 살펴볼 필요가 있다. Swales & Feak(1994)에서는 영미권 대학에 재학 중인 영어 비모어 학습자를 위해 일반적인 학술적 데이터 설명문의 이동 마디와 그에 수반하는 특징을 다음 <표 11>과 같이 제시하였다.

<표 11> Swales & Feak(1994)의 데이터 설명문 이동 마디

| 이동 마디 | |
|--|---------------------------|
| 요소를 배치하고 제언을 요약한다 (Location elements and summaries) | |
| 제언을 강조한다 (highlight statements) | 데이터의 추세를 짚어 내거나 정규성을 확인한다 |
| | 덜 중요한 것과 더 중요한 것을 구분한다 |
| 설명문을 완료한다 (concluding a commentary) | 알맞은 강도로 자신의 주장을 제시한다 |
| | 앞에 다룬 내용의 설명/영향력(필수) |
| | 예측 못한 결론/불만족스러운 데이터(선택) |
| | 가능한 후속 연구/전망(선택) |

이들은 체계 기능적 메타 기능의 의미를 적용하여 이론화를 하지는 않았지만 각 이동 마디에서는 메타 기능적 의미가 나타나고 있다. ‘요소를 배치하고 제언을 요약한다’는 텍스트를 시작하는 관습적 구조라 할 수 있다. Swales & Feak(1994:80)에서는 영어로 된 데이터 설명문의 시작은 언급하려는 시각적 데이터의 위치와 그 요약 두 요소로 이루어지며, 시각적 데이터를 언급한 후 내용을 서술하는 구조로 되어 있다고 하였다. 예를 들어, “표 5는 미국 기업에서 가장 흔한 컴퓨터 바이러스 감염의 양상을 보여준다(Swales & Feak, 1994:80)”와 같이 ‘표 5’라는 위치와 그 내용에 대한 요약이 문장을 구성하고 있다. 이 이동 마디에서는 내용

적 부분을 알고 제시하는 관념적 메타 기능과 시각적 데이터가 있는 위치를 강조하고 분리하는 텍스트적 메타 기능을 제시한다.

‘제언을 강조한다’는 데이터 설명문 구성에서 가장 중요한 위치를 차지하고 있으며, 필자의 지력을 보여줄 수 있는 기회가 된다(Swales & Feak, 1994:85). 데이터를 올바르게 해석할 수 있는가, 그중에서 무엇이 중요한가, 모든 정보들 중 얼마만큼의 정보를 제시할 것인가, 주장하는 것과 어떠한 관련성을 갖는가는 필자와 독자의 학문적이고 분과적인 지식이 결정하는 관념적 성격이 크다. 또한 대인적 측면에서는 데이터를 해석할 때 그 확신성에 따라 독자에게 제시하는 강도를 확률성, 거리, 일반화 등의 기술을 사용하여 조정함을 설명하기도 한다(Swales & Feak, 1994:86-88). 이 과정에서 하지 말아야 할 것이 세 가지 있는데 이는 다음과 같다.

- 단어로 자세한 사항들을 단순히 반복하지 마라
- 모든 정보를 다 다루려 시도하지 마라
- 합리적이거나 방어할 수 있는 것 이상을 주장하지 마라

(Swales & Feak, 1994:86, 연구자 번역)

이들 제한 사항은 이른바 양의 격률(Grice, 1989:26)을 지키라는 것이다. 데이터 설명문 작성에서 이를 수행하기 위한 방법은 데이터 제시 시 단순 수치 나열보다는 비교의 방법을 사용한다거나, 일반적인 경우와는 다르게 주요한 요소를 다른 것보다 먼저 나오게 하는 경우를 들 수 있다(Swales & Feak, 1994: 90-93). 시각적 데이터에서 현저하거나 강조된 정보를 우선적으로 언급하는 이와 같은 행위는 텍스트적 메타 기능과 관련이 있다. 이와 같이 시각적 데이터를 학술적 지식을 지니고 분석/이해하는 것, 장르 관습에 따라 데이터를 반복하거나 제시하는 순서를 고려하는 것, 독자의 권력 및 데이터의 확실성을 고려하여 주장의 강도를 조절하는 것을 볼 때 이 이동 마디는 관념적, 텍스트적, 대인적인 메타 기능을 모두 반영한다.

‘설명문을 완료한다’는 데이터 설명문을 마무리하는 단계에서 영향력,

문제, 예외 등을 논의한다. 이 이동 마디는 설명/영향력을 항상 제시하며, 만약 필요하다면 예측하지 못했던 결론이나 불만족스러운 데이터를 그 다음에 삽입한다. 마지막으로 적절하다면 가능한 후속 연구 또는 전망을 제시한다. 즉, 뒤의 두 부분은 선택적이나 설명/영향력을 제시하는 부분은 필수적인 것이다. 이 과정에서 데이터의 신뢰성과 확실성은 대인적 메타 기능의 하나인 강화와 약화의 언어적 기법을 사용해 표현한다. 데이터의 의미나 영향력을 논할 때 관념성이 드러난다거나, 불만족스러운 데이터나 예측하지 못한 결론에 대해 완화적 표현을 사용하는 대인적 특성이 나타나는 점, 글의 완료라는 장르 관습적 텍스트성이 확인된다는 점에서 이 이동 마디도 체계 기능적 속성을 보인다. 이러한 전체적 구조는 전형적인 학술적 데이터 설명문의 이동 마디와 그 특징을 나타낸다고 할 수 있다.

다음으로, Guinda(2011)에서는 Hyland(2015)의 메타 담화 이론을 활용하여 영어가 모어가 아닌 대학생들의 데이터 설명문을 관념적, 대인적, 텍스트적 메타 기능으로 구분하였다. 그러나 본고의 의도와 달리 이 논문에서는 관념적, 대인적, 텍스트적 언어 표현에 중점을 두고, 이동 마디는 텍스트적 기능에서만 다루었다.³¹⁾ 다만 본고에서는 텍스트적 기능으로 분류된 이러한 이동 마디 자체에서도 메타 기능적 속성을 확인할 수 있다고 보았다. 즉, Guinda(2011)가 제시한 이동 마디는 텍스트적 메타 기능에 속하나, 본고에서는 이동 마디를 메타 기능보다 상위에 놓고 데이터 설명문을 분석하므로 이들 이동 마디에 다른 메타 기능들을 드러낼 수 있는 요소들을 추가하도록 한다. 여기에서는 데이터 설명문에서 가능한 이동 마디를 네 개로 분류하여 제시하는데, 이는 다음 <표 12>와 같다.

31) Guinda(2011:129-130)에서 제시하는 체계 기능적 속성은 주로 언어적 표현에서 드러난다. 예를 들어, 관념적 속성은 주목하는 것이 명사적인 상태/조건이나 과정적인 동사이거나 하는 것에서 찾을 수 있고, 대인적 기능은 독자를 인도하는 표현(접속 표지, 주제 전환자, 어구 해설, 내부 지시, 독자 참여 표지)이 무엇이 있는가 하는 것이 있다. 텍스트성은 위에서 언급한 바와 같이 이동 마디가 있으며 내용 전개에서 원인과 결과 표지, 시간과 단계 순서 제시, 비교와 대조 표지 등을 사용하는 것 또한 있다.

<표 12> 데이터 설명문의 텍스트적 메타 기능 중 이동 마디(Guinda, 2011:130)

| 이동 마디 | |
|--------------------|-----------|
| 파노라마적 목적의 시각적 제시 | |
| 데이터의 전반적인 틀 평가 | |
| 데이터의 자세한 평가 | 그룹화 |
| | 그룹 특성 검토 |
| | 그룹 내 비교 |
| 더 넓은 맥락에서 영향력을 논의함 | 특수한 사례 연구 |

첫 번째로 다루는 ‘파노라마적(한 눈에 볼 수 있는) 목적의 시각적 제시’는 위의 Swales & Feak(1994)의 ‘요소를 배치하고 제언을 요약한다’보다 상위적이고 포괄적인 텍스트의 시작을 전제한다. “이 그래프는 그 시기의 매출 상승을 보여준다(Guinda, 2011:130)”와 같은 문장은 그래프의 위치를 지적하고 내용을 요약함과 동시에 해당 문단을 시작하는 역할을 한다. 즉, 장면의 전형적 속성을 전체적으로 조망하며 틀을 짓는다는 점에서 관념성을 확인할 수 있고, 그래프를 언급하며 시각적 데이터와의 응집성 또는 위치의 구분성을 드러내는 텍스트적 기능을 알 수 있다.

둘째, ‘데이터의 전반적인 틀 평가’는 “해당 연도 동안 판매가 감소 추세였다(Guinda, 2011:130)”와 같이 데이터가 가리키는 현상을 전체적으로 평가하는 부분을 뜻한다. 이 평가는 중립적일 수도 있고, 필자의 태도가 반영될 수도 있다. 독자나 장르가 요구하는 태도를 담화 표지를 통해 표현하기 때문에 이는 대인적일 수 있다. 이에 더해 전문 용어 사용 정도와 명료성을 조절하는 방식, 분과 학문적 지식 인용 등으로 관념성을 드러낼 수 있다.

셋째, ‘데이터의 자세한 평가’는 데이터의 의미에 대해 여러 측면에서 전달하려는 노력 중 하나이다. 산만하게 제시되어 있는 데이터를 동질적인 데이터별 또는 변인별로 그룹화하여 제시하거나, 그 그룹의 특성을 하나씩 언급하는 것이 이에 속한다. 이에 더해 나눈 그룹의 특징을 알아보기 위해 소속 변인 간 비교를 하거나, 이상치와 같은 특수한 데이터만을 짚어 내어 해석하기도 한다. 이는 단순 나열과 비교하여 데이터 특성

을 그룹화, 특성 검토, 비교, 사례만 뽑는 등의 선택은 장르 관습성으로서의 텍스트적 메타 기능과 데이터의 인지적 이해라는 관념적 메타 기능을 동시에 나타낸다.

넷째, ‘더 넓은 맥락에서 영향력을 논의함’은 후속 연구나 다른 해석의 가능성을 언급하거나 데이터의 비관점 등으로 주로 제시된다. 이 이동 마디는 데이터가 갖는 의미를 더 넓은 맥락에 적용하여 해석할 수 있는 지식적 측면을 요구하기 때문에 관념성을 드러낸다고 할 수 있다. 또한 독자를 의식한 주장의 강약 조절 등의 대인적 메타 기능도 찾을 수 있다.

Swales & Feak(1994) 및 Guinda(2011)의 이동 마디는 텍스트의 시작과 끝을 나타내는 특징이 확실한 것과, 데이터를 언급하고 사용하는 방식을 다양한 전략을 차용하여 제시한다는 데서 공통점이 있다. 다만 Swales & Feak(1994)에서 ‘제언을 강조한다’로 간단히 제시한 것을 Guinda(2011)에서는 더 명료하고 자세하게 구분하여 제시하였다는 것과 이동 마디 내에서 표현되는 내용과 방식에 메타 기능적 의미를 부여하였다는 것에서는 차이가 있다.

데이터 설명문 단독으로만 조직된 텍스트들은 위의 두 이동 마디를 거의 따르는 것처럼 보인다. 그러나 Nordrum & Eriksson(2015)에서는 응용 화학 분야 석사 논문 및 학회지 논문에서 데이터 설명문이 속하는 결론 및 논의 부분만을 추출하여 분석하였기 때문에 일반적 데이터 설명문을 다룬 Swales & Feak(1994) 및 Guinda(2011)의 이동 마디와는 일부 차이가 있다. 이는 순수한 장르로서의 데이터 설명문과 내포된 장르로서의 데이터 설명문의 구성 차이로도 볼 수 있다.

<표 13> Nordrum & Eriksson(2015:70)의 데이터 설명문 이동 마디와 단계

| 이동 마디 | 단계 |
|-------|----------------------|
| 배경 정보 | 절차-방법 |
| | 분과 학문 지식 |
| | 시각적 데이터 제시 선택에 대한 설명 |

| | | |
|---------------|--------------------------|---------------------------|
| 시각적 데이터 제시 | 언급 내용 있음 | 요약-언급 |
| | | 결과-언급 시각적 데이터에 대한 설명 및 |
| | 언급 내용 없음 | 해석 언급 명시적 언급 후에 |
| | | 명시적 언급 없음 |
| 결과에 대한 언급 | 결론/시사점의 구체화 | |
| | 선행 연구/연구 질문/다른 데이터와의 관련성 | |
| | 해석 | |
| | 해석의 구체적 내용 | |
| | (재)가설 | |
| | 후속 연구 | |
| 다른 장들과의 교차 참조 | | |

구체적으로, <표 13>에서 제시하는 이동 마디는 크게 배경 정보, 시각적 데이터 제시, 결과에 대한 언급으로 나눌 수 있다. 우선 ‘배경 정보’는 ‘요소를 배치하고 제언을 요약한다’보다 더 단도직입적으로 글을 시작하고, 단계는 더 세분화되었다. Guinda(2011)에서 제시하였던 ‘파노라마적 목적의 시각적 제시’는 드물게 등장한다. 즉, 응용 화학 학술지의 경우에는 데이터 설명문 시작에서 존재해야 했던 ‘어떤 내용이고 무엇을 할 것’이라는 전개가 서론이라는 선행 장을 통해 별도로 제시되기 때문에 데이터 분석이 추가 되는 ‘결과와 분석’ 장에서는 이들을 언급하지 않는다. 다만 보다 전문적인 방식으로 데이터 설명문을 시작하고 있다. 예를 들어, 어떠한 방식으로 제시하고자 하는 결과가 도출되었는지 상세히 설명하는 ‘절차-방법’, 그를 선택하게 된 또는 해석하게 된 학문적 배경을 근거로 제시하는 ‘분과 학문 지식’, 연구 방법 및 절차를 도표로 제시하거나 분과 학문적 지식을 인용하며 그래프와 표를 배치할 때 나타나는 ‘시각적 데이터 제시 선택에 대한 설명’이 단계(step)로서 세분화된다.

둘째로, ‘시각적 데이터 제시’ 이동 마디는 전체 텍스트 내에서 문단 등으로 차지하는 자리가 있는 것은 아니다. 시각적 데이터 제시와 그에 대한 논평은 사실 그 앞 이동 마디인 ‘배경 정보’나 뒤 이동 마디인 ‘결과에 대한 언급’에서 나타난다. 다만 이 두 이동 마디에서 공통적으로 나타

나는 ‘시각성’에 대한 언급이 데이터 설명문에서 반복적이고 중요한 요소로 작용하기 때문에 이를 단독으로 분화하여 두 이동 마디 사이에 배치한 것으로 보인다. 이는 앞의 두 이동 마디에서는 볼 수 없는 것이나 시각적 데이터를 포함한 체계 기능적 담화 분석에서는 고려해볼 수도 있는 구성이라 여겨진다. 이와 같이 텍스트의 특정 위치에 등장하지 않지만 텍스트 전반적으로 나타나며 연구자가 집중하고자 하는 요소 역시 이동 마디로 나타낼 수 있음을 확인할 수 있다.

‘결과에 대한 언급’은 Swales & Feak(1994)가 제시하는 ‘제언을 강조한다’와 ‘설명문을 완료한다’를 혼합한 부분이고, Guinda(2011)의 나머지 세 이동 마디인 ‘데이터의 전반적인 틀 평가’, ‘데이터의 자세한 평가’, ‘더 넓은 맥락에서 영향력을 논의함’에 해당하는 단계를 충실하고 다양하게 수록하였다. 이는 데이터의 수치가 나타내는 결과가 무엇이고 이를 얼마나 어떻게 표현해야 하는지를 고민한다는 데서 데이터 설명문의 정수라 할 수 있다. 다만 추후 본고의 분석에서도 나타나지만 한국어의 논문에서는 후속 연구에 대한 내용은 ‘결과 및 분석’이 아니라 후행하는 ‘결론’장에서 주로 분석되는 경향이 있어 나타나지 않는 등의 차이가 있기도 하고, 데이터의 객관적 정보 제시와 해석이 구분되는 경우 등도 있다.

관념적 메타 기능과 대인적 메타 기능은 Nordrum & Eriksson(2015)의 이동 마디 모형에서 ‘배경 정보’와 ‘결과에 대한 언급’ 중 일부로 구현된다. 관념적 메타 기능은 연구의 분과 지식에 대해 논하는 부류에 속하고, 대인적 메타 기능은 생산자의 주장 및 관련 연구에 대한 태도, 독자 설득/우위 및 신뢰성 획득을 위한 관습적 장치 등으로 표현된다. 위의 이동 마디로 따지자면, 관념적 메타 기능은 ‘절차-방법’, ‘분과 학문 지식’, ‘(재)가설’, ‘결론/시사점의 구체화’, ‘해석’, ‘해석의 구체적 내용’ 등의 학문적 지식을 나타내는 단계로 표상되고, 대인적 메타 기능은 ‘시각적 데이터 제시 선택에 대한 설명’, ‘선행 연구/연구 질문/다른 데이터와의 관련성’, ‘다른 장들과의 교차 참조’ 등 학술지의 관습적 구조를 통한 신뢰성 구축과 같은 방식으로 제시된다. 다만 위의 Swales & Feak(1994)의 이동 마디에서도 그러했던 것처럼 각 메타 기능만을 단독으로 나타내는

이동 마디는 존재하지 않는다. 언급한 메타 기능을 지지하기 위한 텍스트적 메타 기능이 각 요소에 함께 존재하기 때문이다. 이는 하나의 절, 문장, 문단에 관념적, 대인적, 텍스트적 기능이 동시에 내재한다는 이관규 외(2021:29)에서도 확인할 수 있다.

그러나 지배적인 메타 기능의 경향성은 존재하는데, 특히 Nordrum & Eriksson(2015)의 두 번째 이동 마디인 ‘시각적 데이터 제시’는 데이터 설명문의 메타 기능적 이해에서 텍스트적인 것으로 지적할 만하다. 이 이동 마디는 관습적이고 고정적인 표현을 통해 시각적 데이터와 언어적 본문을 연결한다. 따라서 위에서 언급했던 것처럼 하나의 이동 마디를 형성하는 대신 ‘배경 정보’와 ‘결과에 대한 언급’ 두 이동 마디에 걸쳐서 제시되는 경향이 있다. 즉, ‘시각적 데이터 제시’는 하나의 이동 마디를 이루지 못하고 다른 이동 마디에 기생하는 특성을 보인다. 따라서 본고에서는 ‘시각적 데이터 제시’에 해당하는 이동 마디가 따로 존재한다기보다는 모든 이동 마디에 분산된 것으로 보았다. 즉, 시각적 데이터와 관련된 이동 마디는 관념적/대인적 메타 기능에 속하는 이동 마디와 동시에 출현할 수도, 단독으로 출현할 수도 있는 것으로 본다. 또한 이는 주로 텍스트와 시각적 데이터의 일관성과 응집성 및 경계, 중요성 측면에서 단계로 구현될 것이다. 메타 기능이 상위 이동 마디에 동일하게 적용되지 않거나, 하위 이동 마디들이 각 상위 이동 마디들을 이동하며 출현하는 점, 특정 상위 이동 마디는 그 순서와 위치가 텍스트에 명시적으로 정해지지 않았다는 점에서 이는 일반적으로 익숙한 형태의 이동 마디는 아니다. Nordrum & Eriksson(2015:67) 역시 자신들이 제시한 이 이동 마디를 보완이 더 필요한 틀이라 평가하였다.

따라서 데이터 설명문의 이동 마디를 구축할 때는 데이터 제시와 평가에 관련해서 개별 텍스트의 특성을 수용할 수 있어야 한다. 또한 텍스트에서 데이터를 언급하는 장면을 보다 면밀히 파악하여 이동 마디를 구축해야 한다. 이러한 연구들의 논의를 토대로 이지연(2023)에서는 한국어 경제학 및 경영학 학술지에 나타나는 데이터 설명문의 이동 마디를 분석한 바 있다. 이 연구에서는 Nordrum & Eriksson(2015)에서와 같이 학술

지 내 데이터 설명문을 결과와 논의 장에 국한하여 추출하였고, 이들의 이동 마디를 바탕으로 반복적 내용 분석을 통해 새로운 이동 마디를 구성하였다. 이는 이지연(2023)의 분석 대상 데이터 설명문이 Nordrum & Eriksson(2015)의 대상 데이터 설명문과는 달리 한국어로 구성되었으며, 학문 분과도 경제학으로 다르기 때문에 이를 반영한 것이다. 그 결과 한국어 경제학 및 경영학 데이터 설명문 이동 마디는 ‘주제 및 방향 요약’, ‘수치/결과 나열’, ‘해석 및 결론’으로 나타나며, 그를 구성하는 하위 이동 마디에 관념적, 대인적, 텍스트적 의미를 부여하였다.

<표 14> 경제학 및 경영학 학술지의 데이터 설명문 이동 마디(이지연, 2023:162-163)

| 이동 마디 | 단계 | | 메타 기능 | |
|----------------------|--------------------------|--|---|--------------|
| 주제 및 방향 요약 | 소제목으로 주제 제시 | | 대인적 | |
| | 연구의 주제와 내용 및 방법 설명 | 주어에 3인칭 대명사 사용 알맞은 전문 용어를 다수 사용 | | |
| | 데이터 현상에 대한 배경 설명 | 기타 데이터 통해 배경 해석 | 관념적 | |
| | | 현상에 대한 기존 이론 소개 | | |
| 연구 목표 또는 방법 제시 | 용어의 정의 | | | |
| | 연구 수행의 이론적 근거 | | | |
| | 연구 방법과 도구의 검증 | | | |
| 수치/결과 나열 | 시각적 데이터 언급 | 구체적인 정보 추출 | 연대기적 나열 시각적 데이터 순서대로 나열 | 관념적/ 텍스트적 |
| | | 의도적 그룹화 | 수치가 높은/낮은 순서대로 나열 특이성/중요 지점 지적 | 대인적/ 텍스트적 |

| | | | | |
|------------|----------------------|--------------|---|---------|
| 해석 및 결론 | | | 제일 큰/작은 수치에 주목 특성에 따른 그룹화 수치적 비교 | |
| | | | 시각자료 언급 표지 사용 | |
| | | 수치의 의미 제시 | 추론하기 확신하기 | 대인적 |
| | 연구 결과/수치 의미 해석 | | 전문 분과 지식/선행 연구/다른 데이터 인용 해석 통계적/수리적 | 관념적/대인적 |
| | | | 검증 결과 데이터의 비판점과 개선 방향 | 관념적 |
| | | | 문제 해결 방안/후속 연구/미래 방향 제시 | 관념적/대인적 |

여기서는 관념성을 ‘다루고 있는 내용과 그 환경에 대한 지식적 이해’, ‘분과 전문 이론과 지식’으로 보고 연구의 내용을 제시하거나 해석하는 이동 마디에 주로 부착한다. 또한 대인성은 독자의 신뢰를 깨지 않도록 관습적으로 구성된 것이나, 필자가 독자에게 주로 제시하고 싶은 데이터를 강조하는 기술 방식이 이에 속한다. 텍스트성은 시각적 데이터와 언어적 자료를 잇는 특수한 언급이나 순서를 인식하게 하는 언급으로 제시된다.

이들 선행 연구에서의 이동 마디와 메타 기능 연합은 연구자에 따라 해석과 분류가 조금씩 달라질 수 있으나 텍스트 전체 구조를 조망하며 다

중 기호를 해석하려는 시도는 본고에서도 동일하게 적용된다. 따라서 한국어 학습자의 데이터 설명문 분석에서는 이들을 어떻게 해석하고 변환할지 마지막 절에서 고찰해 보도록 하겠다.

3. 정보 매개적 속성에 따른 인지적 및 언어적 구성

3.1. 학술적 데이터 설명문의 매개적 속성

학술적 데이터 설명문의 작성에는 관념적 메타 기능에 속하는 학문 분과적 지식 외에도 시각적 데이터에서 어떤 데이터를 얼마나 선택해 어떻게 제시하는지에 대한 인지적 능력도 필요하다. 따라서 학습자의 데이터 설명문에서 이러한 인지적/언어적 능력의 양상을 파악하기 위해 이에 대한 구체적인 내용을 알아볼 필요가 있다.

Swales & Feak(1995:75-76)에서는 대학생들이 자신의 학술적 글을 쓸 때 비판적 사고를 바탕으로 어느 부분의 데이터를 선택하고 강조해야 하는지가 중요하다고 보았다. 따라서 시각적 데이터를 있는 그대로 묘사하는 ‘정보 전달(information transfer)’보다는 해석, 비판, 기대를 사용역에 맞는 어휘와 문법을 사용하여 제시하는 ‘데이터 설명문(data commentary)’ 작성에 초점을 두어야 한다고 주장하였다. 따라서 학술적 데이터 설명문은 데이터가 갖는 정보를 저자가 목적과 의미를 투영하여 독자가 기대하는 양식으로 다시 제시하는 텍스트라 할 수 있다. 이렇듯 한 양식에서 정보를 선택적으로 골라 장에 맞게 다른 양식으로 전달하는 과정은 유럽공통참조기준(Common European Framework of Reference, CEFR)에서도 ‘매개(mediation)’라는 용어를 통해 주목하고 있다.

CEFR(2020:33)에서는 의사소통적 언어 활동과 전략이 말하기, 듣기, 읽기 쓰기의 전통적인 4기능으로는 목표와 거시 기능을 표상하기 어렵다고 판단하고 이를 생산, 수용, 상호작용, 매개의 네 요소로 개편하였다. 생산, 수용, 상호작용에 비해 생소한 매개는 전달자가 정보를 수용하여 이

해하고 그를 상호작용에서의 독자 판단을 통해 그에게 맞는 방식으로 생산하는 행위를 일컫는다(김호정 외, 2021).³²⁾ 또한 CEFR(2020)에서 주목하는 것은 의미를 담고 표현하는 개체가 문자에 국한되지 않는다는 점이다. 따라서 문자, 수어, 시각적 데이터 등 모두가 의미를 매개하는 기호가 될 수 있다. 이는 데이터 설명문을 구성하는 통계 결과표, 그래프, 다이어그램 등을 언어 기호로 표현하는 것도 포함한다고 볼 수 있다.

따라서 본고에서는 매개의 개념이 학술적 데이터 설명문 작성에서 필요한 인지적 능력을 표상하고 있다고 보았다. 학술적 데이터 설명문 작성은 작성자가 시각적 데이터의 내용과 의미를 전문 독자에게 전달하기 위해 고민하는 과제이기 때문이다. 즉, 저자가 통계 결과표, 그래프, 다이어그램과 같은 전문적 시각 양식이 내포하는 정보를 자신의 주장에 맞게 더 잘 이해할 수 있도록 장애 맞는 언어 양식으로 독자에게 전달하는 것을 데이터 설명문의 매개적 특성이라 본다.

이러한 데이터 설명문의 매개적 특성은 ‘구어/문어로 데이터 설명하기’라는 항목이 매개의 하위 범주에 목록으로 명시적으로 제시된 데서도 확인할 수 있다(CEFR, 2020:90). 본고에서 다루는 데이터 설명문은 이 중 텍스트 매개의 ‘구어/문어로 데이터 설명하기’에 해당된다.

‘데이터 설명하기’³³⁾는 “도표(그래프, 다이어그램 등)에 나타난 언어 텍스트 정보를 변환하는 것(CEFR, 2020:96)”을 일컬으며, 구체적으로는 다음과 같은 행위를 포함한다.

- 익숙한 주제에 대한 그림 자료(예: 순서도, 일기도)를 묘사한다.
- 추세를 그래프로 나타낸다.
- 막대그래프에 대해 해설한다(comment).
- 그림으로 제시된 경험적 데이터의 현저한 지점을 선택하고 해석한다.

32) 즉, 내용의 의미를 이해할 수 있게 하는 모든 전달 전략이 이에 속하는데, 구체적으로는 ‘이전 지식과 연계하기, 언어를 받아들여 적용하기, 복잡한 정보를 나누기, 난해한 텍스트를 자세히 서술하기, 텍스트를 간소화하기’ 등이 있다(CEFR, 2020:35).

33) ‘구어/문어로 데이터 설명하기’ 범주에서 본고는 ‘문어로 데이터 설명하기’만을 다룰 예정이기 때문에 CEFR에서 제시하는 ‘데이터 설명하기’ 범주 설명과 기술문도 문어에 해당하는 것만을 선택하여 사용하도록 한다.

여기에서는 순서도를 묘사하는 단순한 정보 전달도 있지만, 두 번째, 세 번째와 네 번째 같은 학술적 장에서의 학술적 의도를 지닌 행위는 Swales & Feak(1995)이 강조했던 것과 같이 정보를 선택하고 해석해야 한다는 점을 명확히 하고 있다. 또한 Guinda(2011:118)에서는 시각적 데이터를 맞닥뜨렸을 때 언어 사용자에게 요구되는 기능들이 “CEFR(2001)에서 명확히 밝힌 바대로 일반적 능력에 해당하며, 학습할 수 있는 능력(Savoir-apprendre) 또는 자율적 학습의 일부분을 형성하는 학습과 체험적 능력의 융합으로 볼 수 있다”고 하였다. 즉, 학술적 데이터 설명문은 전문 분과의 지식을 학습하는 과정에서 인지적/언어적 발달을 동시에 도모할 수 있는 학술적 실제성을 띤 과업이라 할 수 있다.

3.2. 매개적 속성으로서의 인지적 및 언어적 구성 요소

앞에서 제시한 데이터 설명문의 매개적 속성은 데이터를 취사선택하는 인지적 행위와 이를 다양한 상황의 독자에게 이해시키려는 언어적 표현으로 이루어진다. 이러한 능력은 숙달도와 관련이 있기 때문에 주로 평가 영역에서 그 구체적인 내용을 다루고 있다. 본고에서는 데이터 설명문 작성 시 선언적 지식과 방법의 측면에서 학습자의 인지적 능력, 목표어 언어학적 측면에서 학습자의 언어 능력이 중요하다는 것을 인식하고 이를 학습자 데이터 설명문 분석에서 활용할 예정이다. 이에 따라 데이터 설명문의 인지적/언어적 능력에 대해 평가에서의 논의를 다수 활용할 것이다. 다만 이 연구는 쓰기 평가를 논하는 것이 아니므로 평가 기준을 참고할 뿐, 학습자의 숙달도를 평가하지는 않을 것이다. 즉, 평가 영역에서 논하는 데이터 설명문의 인지적/언어적 요소들을 확인하여, 이들이 어떻게 실제 학습자 데이터 설명문에 나타나는지 확인할 것이다. 구체적으로는 인지적/언어적 평가 기준을 학습자 데이터 설명문의 이동 마디와 언어 사용 분석에서 적용할 예정이다. 따라서 본고에서는 CEFR(2020)이 제시하는 등급 기준을 달성한 상태보다는, 그 등급에 접근하는 과정적

측면에서의 학습자 데이터 설명문에 더 집중하려 한다.

이를 위해 평가 영역에서의 언어적/인지적 요소를 확인해 보도록 한다. 우선 CEFR(2020)의 텍스트 매개 활동인 ‘구어/문어로 데이터 설명하기’의 등급 평가 기준을 인지적 능력과 언어적 능력으로 나누어 어려운 정도를 다음 <표 15>와 같이 위에서부터 아래로 나열하였다.

<표 15> CEFR(2020) ‘데이터 설명하기’가 포함하는 인지적/언어적 능력

| 인지적 능력 | 언어적 능력 |
|---|--|
| <ul style="list-style-type: none"> ·개념적으로 복잡한 주제 이해 ·주어진 데이터의 개념적 복잡성 이해 ·주어진 데이터가 속한 분야의 전문성 이해 ·주어진 데이터에서 중요한 정보 파악(현재성, 중요 지점 등) ·주어진 데이터에서 구체적인 정보 추출(전반적 추세, 주요 사항 등) | <ul style="list-style-type: none"> ·명확하고 신뢰성 있게 표현하는 문어 능력 ·데이터 해석을 문어로 제시할 수 있는 능력 ·사전이나 참조 자료의 도움을 받아 문어를 구성할 수 있는 능력 ·문장을 작성할 수 있는 능력 |

CEFR(2020)에서는 등급에 따라 인지적 어려움과 언어적 어려움이 동시에 증가하도록 평가 기술을 조정하고 있다. 학습자의 제2언어 능력이 특정 등급에 도달한 상태에서는 그 등급에 맞도록 인지적 어려움과 언어적 어려움이 해소된 상태일 것이다. 다만 학습자들이 데이터 설명문을 작성하는 ‘과정’에서는 인지적 어려움과 언어적 어려움의 ‘관계’가 결과적 상태에서만큼 정적이지 않을 것이라 본다.

일반적인 정보 처리 모형에서 다루는 작업 기억(working memory)의 한정성에 미루어 짐작해 볼 때 제2언어 학습의 과정에서는 Skehan(1998)의 한정적 수용 가설과 관련된 양상이 있음을 간과할 수 없다. Skehan(1998)의 한정적 수용 능력 가설(limited capacity hypothesis)에서는 학습자의 작업 기억 용량에는 한계가 있어 과제의 인지적 부담이 증가하면 제2언어 산출 능력이 저하된다고 보고 있다(이선진 외, 2023:136 재인용). 이 한정적 수용 능력 가설의 기반은 교육학의

정보 처리 모형이다. 이 모형에서는 인간은 작업 기억이 유한하여 한꺼번에 여러 정보를 처리할 수 없다고 본다. 다만 모어 사용이나 자전거타기처럼 오랜 기간 훈련하여 자동화된 정보는 장기 기억으로 넘어갔기 때문에 작업 기억에 영향을 주지 않는다(이지연, 2014:24-29). 이에 따라 한국인 대학생이 한국어로 데이터 설명문을 쓸 때는 한국어 사용에 대한 작업 기억의 부담이 떨어져, 데이터의 수치 비교와 변환 등을 할 수 있는 인지적 활동이 가능한 용량이 남아 있는 반면, 외국인 대학생은 한국어를 사용하는 것 자체가 작업 기억에 부담을 주기 때문에 수치 계산과 같은 복잡한 비교에 할당할 용량이 부족하다는 것을 알 수 있다.³⁴⁾ 본고에서는 이러한 정보 처리적 관점에서 데이터 설명문의 이해와 표현 방식에 대한 인지적 부담의 정도를 추론해 보도록 하겠다.

CEFR(2020)의 평가 기술문에서는 ‘데이터 설명하기’가 포함하는 인지적 능력에는 개념적 복잡성, 자료가 속한 분야의 전문성, 자료에서의 중요 정보 파악과 추출이 포함된다. 이를 학습자의 학술적 데이터 설명문 분석에서 다루는 인지적 능력에 적용해 보자면, 학습자가 데이터에 내재된 복잡한 개념을 얼마나 잘 이해하고 있는가, 그 개념을 파악하기 위해 필요한 전문 분과적 지식은 얼마나 갖추었는가, 해당 데이터에서 어떤 정보가 어떤 목적과 장에서 유용한지 판단할 수 있는가 등이 있다. 이는 주로 데이터의 수치를 언어적으로 변환하는 이동 마디에서 주로 사용될 수 있는데, 구체적인 내용은 추후 분석 프레임에서 다시 다루도록 한다.

등급에 따른 목록별로 살펴보면, ‘주어진 데이터에서 구체적인 정보 추출’과 같이 있는 그대로의 정보를 훑어내어 작성하는 것은 인지적 능력이 가장 적게 요구된다. 시각적 데이터를 직접 눈으로 보고 있기 때문에 내용이 고정된 채 그 기호만 변화시키는 것은 큰 인지적 노력이 필요하지 않은 것으로 보인다. 이러한 단순한 기호 간 전이는 기호 그 자체에 대한 미숙성이 클 때 나타나는 경향이 있다. 즉, 시각적 데이터를 제2언어로 변환하는 과정에서 제2언어에 미숙한 저자는 데이터 처리에 인지적

34) 제2언어와 정보 처리에 대한 더 자세한 논의는 Skehan(1998), Robinson(2001)을 참조하여라.

노력을 더 많이 하기보다는 제2언어로 문어를 구성하는 데 더 많은 노력을 쏟을 것이기 때문이다(이선진 외, 2023:159). 그다음으로 인지적 능력을 조금 더 요하는 작업은 ‘주어진 데이터에서 중요한 정보 파악’이다. 모든 정보를 추출하여 언어적 변환을 하는 것에서 더 나아가 데이터의 특성에 조금 더 집중해야 하기 때문에 데이터에 대한 인지적 노력이 더 필요하다. ‘주어진 데이터가 속한 분야의 전문성 이해’는 데이터가 표상하는 전문 분과적 의미에 대한 지식이 필요함을 강조한다. 즉, 제2언어 숙달도가 뛰어나다 할지라도 특정 분과에 대한 이해가 있지 않으면 해당 데이터를 이해하기 어렵기 때문에 이는 인지적 노력이 조금 더 필요하다고 할 수 있다. 여기서 더 나아가 ‘주어진 데이터의 개념적 복잡성 이해’의 경우에는 분과 지식을 통해 시각적 데이터가 나타내는 어려운 개념을 이해해야 한다. 이보다 더 인지적 노력이 필요한 정도는 ‘개념적으로 복잡한 주제 이해’이다. 주제 자체가 분과 학문에서 천착해야 하는 어려운 개념의 풀이와 관계가 있다면, 다른 요소에 할당할 수 있는 작업 기억의 용량이 부족할 것이라 생각된다.

한편 언어적 능력은 관습성과 관련이 있는 ‘명확하고 신뢰성 있게 표현’, 데이터 해석의 언어적 사용에 집중한 ‘데이터 해석을 문어로 제시’, 문어 구성 중 어휘문법적 레퍼토리가 부족한 것에 집중한 ‘사전이나 참조 자료의 도움을 받아 문어를 구성’, 제2언어/외국어로 ‘문장을 작성’과 같은 결과로 제시된다. 앞에서 언급했듯 본고에서는 이러한 위계를 통해 학습자의 능력을 평가할 의도는 없으므로 이러한 작문의 결과는 학습자의 이동 마디별 말뭉치 분석에서 나타나는 어휘문법적 사용을 확인한 후 텍스트의 관습성 부족이나 데이터 해석 방법 숙지 부족, 어휘문법적 레퍼토리 부족, 제2언어 문장 구성 능력 부족 등으로 학습자 데이터 설명문 해석에서 활용하도록 한다.

또한 CEFR 말고도 다른 쓰기 평가 기준 관련 선행 연구들에서는 데이터 설명문의 구성 지식과 언어를 어떻게 처리하는지도 확인해 보도록 하겠다. 데이터 설명문이라는 텍스트를 따로 지정하는 쓰기 평가에는 IELTS와 TOPIK이 있다. IELTS는 학문 목적 학습자를 대상으로 하는

‘학술적(academic)’ 시험과 일반 목적 학습자를 대상으로 한 ‘일반적(general)’ 시험이 있다.³⁵⁾ 이 학술적 시험의 첫 번째 작문 과제가 테이 터 설명문 작성이다. 이 평가의 기준은 ‘과제 달성’, ‘일관성과 응집성’, ‘어휘 자원’, ‘문법적 범위와 정확성’으로 나뉘며, 숙달도는 Band 0부터 Band 9까지 10단계에 걸쳐 판단된다. 여기서의 인지적 요소는 ‘과제 달성’에 해당하며 나머지는 언어적 요소로 판단할 수 있다. Writing task 1의 등급 기술서를 바탕으로 인지적 요소와 언어적 요소를 나누어 정리해 보면 다음 <표 16>과 같다(IELTS 공식 홈페이지, 2023.4.19. 검색).

<표 16> IELTS ‘Writing task 1’이 제시하는 인지적/언어적 능력

| 인지적 능력 | 언어적 능력 |
|---|--|
| <ul style="list-style-type: none"> ·과제 분량 준수 ·과제의 이해 정도 ·과제의 주요 요소/중요 항목을 전체성과 관련성에 따라 정확하고 적절하게 다루는 정도 ·정보의 적절한 선택과 전체적 제시 ·그래프의 주요 추세, 차이, 단계에 대한 전체적 제시 | <ul style="list-style-type: none"> ·아이디어의 논리적 구조화 ·정보와 주제의 일관성 있는 순서와 진행 ·응집성 장치의 적절하고 정확하며 자연스러운 사용 ·단락 나누기의 자연스러움 ·어휘의 범위와 그 특성에 따른 자연스럽고 정교한 사용 ·문체와 언어에 대한 인식과 자연스러운 사용 ·구두법, 철자법, 조어법의 오류 정도 ·주제와 과제에 맞는 어휘를 정확하고 적절하게 사용 |

35) 따라서 band의 특정 부분에 일반적/학술적 영역에만 적용되는 평가 기준을 따로 표시해 놓는다. 예)band 6의 ‘과제 달성’에서는 학술과 일반을 아래와 같이 나누어 놓았다.

| | |
|--------|---|
| band 6 | <ul style="list-style-type: none"> ·과제의 요구 사항 제시 ·(학술) 적절하게 선택된 정보를 개관하여 제시 ·(일반) 일반적으로 명확한 목적을 제시: 어조에서 부적절한 것이 발견될 수 있음 ·주요 요소, 중요 요소를 적절하게 제시하나 세부적으로는 관계없거나, 부적절하거나, 불명확함 |
|--------|---|

| | |
|--|--|
| | ·문장의 구성에서 복문의 정확하고 적절한 사용 ·문법 오류의 지배적인 정도 |
|--|--|

IELTS는 인지적 능력에 평가의 조건을 세밀하게 제시해 놓았다. 과제의 분량에 대한 것이 그러한데 본고에서는 그다음부터 제시되는 내용에 더 주목하도록 한다. 이해 정도와 주요 요소를 적절하게 다루는 것, 정보의 선택과 제시, 그래프의 추세, 차이, 단계 제시 등이 학습자 데이터 설명문의 ‘수치/결과 나열’ 이동 마디에서 구체적으로 확인될 수 있다고 보기 때문이다. 언어 사용은 CEFR에서와 마찬가지로 이동 마디별 n-gram 분석 결과에서 나타난 양상을 해석하는 데 사용한다.

TOPIK의 경우에는 데이터 설명문에 국한된 평가 양식이 없다. 다만 ‘내용 및 과제 수행’, ‘글의 전개 구조’, ‘언어 사용’의 세 평가 기준으로 글쓰기 전반을 평가한다. ‘내용 및 과제 수행’은 과제 이행과 주제 관련 내용 구성의 다양성을 평가한다. ‘글의 전개 구조’는 글의 구성의 명확성과 논리성, 단락 구성, 논리 전개를 나타내는 담화 표지 사용, 조직적 연결을 주로 보고 있다. ‘언어 사용’은 문법과 어휘의 다양하고 풍부한 정도, 그의 적절한 선택과 정확한 사용, 목적과 기능에 따라 격식에 맞는 글쓰기, 맞춤법의 정확한 사용 등을 들고 있다. 이를 하위 요소까지 정리하면 다음과 같다(TOPIK 공식 홈페이지, 2023.4.29. 검색).

<표 17> TOPIK 53-54번 평가 기준

| | |
|------------|--------------------------------|
| 내용 및 과제 수행 | 주어진 과제를 충실히 수행하였는가? |
| | 주제에 관련된 내용으로 구성하였는가? |
| | 주어진 내용을 풍부하고 다양하게 표현하였는가? |
| 글의 전개 구조 | 글의 구성이 명확하고 논리적인가? |
| | 글의 내용에 따라 단락 구성이 잘 이루어졌는가? |
| | 논리 전개에 도움이 되는 담화 표지를 적절하게 사용하여 |

| | |
|-------|---|
| | 조직적으로 연결하였는가? |
| 언어 사용 | 문법과 어휘를 다양하고 풍부하게 사용하며 적절한 문법과 어휘를 선택하여 사용하였는가? |
| | 문법, 어휘, 맞춤법 등의 사용이 정확한가? |
| | 글의 목적과 기능에 따라 격식에 맞게 글을 썼는가? |

TOPIK에서는 인지적 능력을 과제 수행의 정도와 내용 지식, 담화 지식의 세 가지로 나누는 것으로 보인다. 다른 평가 기준들처럼 도표의 이해와 활용에 대한 인지적 접근은 평가에 가지적으로 드러나기보다는 ‘내용 및 과제 수행’에서 내용 지식과 관련되어 제시된다. 다만 최혜민(2017)에서는 TOPIK 53번 데이터 글쓰기 과제에서는 ‘자료 해석 행위’가 그래프 쓰기 수행을 평가하는 기준으로 반영되어야 함을 주장하였다. 이는 시각적 데이터 해석이 데이터 설명문의 인지적 능력에 해당함을 연구자들도 지적하는 것이라 할 수 있다.

따라서 이들 평가 기준에서 다루는 데이터 설명문의 인지적 측면은 시각적 데이터가 각각 또는 함께 연합하여 나타내는 의미를 이해하여 표현하는 것, 시각적 데이터 표현이 단순한 수치 나열에 그치는 것이 아닌 추세와 그에 대한 해설을 덧붙이며, 제일 높거나, 크거나, 낮거나, 작은 등 현저히 눈에 띄는 지점을 우선 언급하는 것 등이라 할 수 있다. 이러한 방식으로 데이터 설명문에서 수치를 제시하는 방식은 Swales & Feak(1995)의 연구와 Guinda(2011)의 이동 마디에서도 언급된 바 있다. 이들을 정리하자면, 1. 주제와 과제 수행에 관련된 내용을 개념적으로 온전히 이해하는 것, 2. 자료에서 적절하고 정확한 정보를 선택하는 것, 3. 정보를 적절한 양식과 방식, 순서로 서술하는 것으로 볼 수 있다. 또한 이들 평가 기준을 인지적 노력의 측면에서 정리해 볼 때 그 정도는 데이터의 단순 나열일 경우가 가장 적고, 그다음이 데이터에서의 적절한 정보 추출, 전문적 지식을 개입한 데이터의 이해, 데이터가 나타내는 개념과 주제 이해 순으로 인지적 노력이 다수 필요함을 알 수 있다.

한편 언어적 능력은 응집성 및 텍스트성과 어휘문법 및 구두점, 맞춤법 사용으로 나눌 수 있다. 응집성 및 텍스트성은 글의 일관성 있는 순서와

진행, 단락의 자연스러운 구분, 일관성과 논리적 전개를 나타내는 표지 사용, 조직적 연결을 들 수 있다. 어휘문법적 부분은 언어학적 단위별로 구분하자면 표기(구두점, 맞춤법) 사용, 어휘의 풍부함/정확함/적절함, 주제/내용/장에 맞는 어휘/문법/문체 사용, 문법 오류가 내용 이해를 결정 짓는 정도, 복문의 적절성과 정확성을 들 수 있다. 이들 언어적 능력은 평가에서 정오 또는 적절성을 판단하는 기준으로 사용하기 때문에 본고에서는 이들 기준을 평가적으로 사용하기보다는 양상 파악의 참고 기준 정도로 삼는다.

4. 학술적 데이터 설명문의 분석 프레임

4.1. 학술적 데이터 설명문의 구조

이제까지의 논의를 살펴본바, 데이터 설명문을 이루는 요소는 텍스트 구조적, 미시 언어적으로 재편하여 제시할 수 있다. 이에 따라 학습자의 학술적 데이터 설명문 분석을 위한 프레임으로서 텍스트의 구조에 해당 하는 것을 재배열할 것이다.

우선 고려해야 할 것은 실제적인 텍스트 진행 순서이다. Swales(1990)가 제안한 이동 마디 분석에서는 본래 이동 마디 구축을 평가자 2인 이상이 협의하여 진행한다. 그러나 본고의 이동 마디 구축은 장르 일반적이고 언어 중심의 순서를 알아내려는 것보다는 데이터 설명문의 다중 양식적 서술의 의미를 이동 마디 순서에 반영하려는 것이다. 이에 더해 데이터 설명문의 정보 매개적 특성을 이루는 인지적 사항들이 어떻게 반영되는지도 파악하려는 것이 목적이다. 따라서 모형이 이미 구축된 상태에서 텍스트의 분석의 타당성을 피하려기보다는 새로운 모형을 구성하기 위해 실제 자료와 이론은 어떻게 교환되어야 하는가, 선행 연구에서 개발된 이동 마디와는 무엇이 다른가를 학습자의 데이터 설명문과 선행 이론의 탐구를 통해 밝혀내야 한다. 이에 따라 본고에서는 자료와 코드의 반복적 분석을 수행하는 반복적 내용 분석을 따라 학술적 데이터 설명문

의 이동 마디를 설정하려 한다고 언급한 바 있다.

이동 마디는 장르적 텍스트 분석 시 텍스트의 구조와 순서를 파악하는 방법이다. 앞선 연구에서 볼 수 있듯 데이터 설명문의 이동 마디는 체계 기능 이론의 메타 기능적 의미를 부착할 수 있다. 즉, 시각적 데이터와 언어적 자료가 함께 배치되고, 상호 변환되며, 그 과정에서의 선택과 강조, 제시 기술을 사용하기 때문에 이동 마디 구성 역시도 언어적 기호와 더불어 메타 기능적 의미를 지니고 있다고 할 수 있다. 이러한 과정에서 CEFR, IELTS, TOPIK이 언급하는 인지적 능력을 추가적으로 부착한다. 이들의 평가 기술에서는 중요 정보의 포착과 그를 상세히 서술하는 방식의 차이가 등급을 나누는 인지적 능력의 차이로 작용함을 보여준다. 데이터 설명문에서 수치 서술을 분석할 때 위치적인 이동 마디로만 표현한다면 단순히 ‘수치/결과 나열’ 정도로 제시할 수 있을 것이다. 그러나 데이터의 선택과 강조, 제시 기술 등은 텍스트 내의 등장 순서나 메타 기능으로는 설명할 수 없는 인지적 속성도 갖고 있다. 따라서 수치 정보의 가치에 대한 인식과 그 표현은 ‘수치가 높은/낮은 순서대로 나열’ 또는 ‘제일 큰/작은 수치에 주목’과 같은 단계에서 구체적으로 제시할 수 있다.

이러한 인지적 능력이 중요한 까닭은 인간의 언어 사용과 관련하여 정보 처리 모형에서 제시하는 작업 기억의 유한함 때문이다. 작업 기억에서 처리할 수 있는 정보는 5 ± 2 로 추정된다(문은식 외, 2007). 그 이상의 정보가 투입되면 망각이 되는데, 이를 해결하기 위해서는 다양한 전략을 통해 정보를 자동화해야 한다(DeKeyser, 2001). 제2언어/외국어를 사용할 때 이들이 자동화되었다면, 즉, 제2언어/외국어 숙달도가 높다면 작업 기억에서 상대적으로 인지적 부하를 덜 받기 때문에 다른 인지적 작업들을 하기에 용이해진다. Skehan(1998)에서는 이에 언어 교육에서의 과제 배열이 학습자의 작업 기억 용량에 영향을 주는 인지적인 면을 반드시 고려해야 한다고 하였다(이지연, 2014:28-29 재인용). 따라서 인지적 부하가 높은 데이터 설명문의 이동 마디는 한국어가 자동화된 정도에 따라 달리 선택될 수도 있을 것이라는 추론을 할 수 있다.

| | |
|---|--------------------------|
| 이동 마디 | |
| 체계 기능적 다중 양식 담화 분석 기반의 메타 기능 | 정보 매개의 인지적 능력 |

<그림 3> 학술적 데이터 설명문 구조 형성 원리

이에 본고에서는 체계 기능적 메타 기능과 정보 매개의 인지적 요인을 고려하여 체계 기능적 의미를 부착한 이동 마디를 위 <그림 3>과 같은 원리에 의해 구축하도록 한다. 그 결과 구축된 학습자의 학술적 데이터 설명문 이동 마디는 아래의 <표 18>과 같다.³⁶⁾

36) 다음 장에서는 이들 이동 마디를 토대로 학습자 변인별 담화 구조의 차이를 분석할 것이다. 그 과정에서 이동 마디 이름을 반복하거나 나열하는 것이 번거로우므로 아래 <표 18>에서 확인할 수 있는 바처럼 A,B,C와 같은 표지를 부착하도록 한다.

<표 18> 학습자 데이터 설명문의 이동 마디와 표지 부착

| 이동 마디 | 단계 | 메타 기능 | 인지적 능력 | |
|------------------|-----------------------------|-------------------------|--|---|
| A. 주제 및 방향 요약 | A1. (소)제목을 활용한 주제 제시 | 대인적/텍스트적 | 적절한 양식과 방식, 순서로 서술 주제와 과제 수행에 관련된 내용을 개념적으로 온전히 이해 | |
| | A2. 연구의 주제와 내용 및 방법 설명 | 관념적/텍스트적 | | |
| | A3. 데이터 현상에 대한 배경 설명 | 관념적 | | |
| B. 수치/결과 나열 | B1. 구체적인 정보 추출 | B1-1. 연대기적 나열 | 관념적/텍스트적 | 적절한 양식과 방식, 순서로 서술 |
| | | B1-2. 시각적 데이터 순서대로 나열 | 관념적/텍스트적 | |
| | B2. 의도적 그룹화 | B2-1. 수치가 높은/낮은 순서대로 나열 | 대인적/텍스트적 | 자료에서 적절하고 정확한 정보를 선택하는 것 (인지적 노력 요구 정도가 다름) |
| | | B2-2. 수치적 비교 | 대인적/텍스트적 | |
| | | B2-3. 특성에 따른 그룹화 | 대인적/텍스트적 | |
| | | B2-4. 제일 큰/작은 수치에 주목 | 대인적/텍스트적 | |
| | | B2-5. 특이성/중요 지점 지적 | 대인적/텍스트적 | |
| | B3. 시각적 데이터 언급 표지 사용 | 대인적/텍스트적 | 적절한 양식과 방식, 순서로 서술 | |
| B4. 수치의 의미 제시 | 대인적 | | | |
| C. 해석 및 결론 | C1. 주어진 데이터/상식적 지식에 근거한 해석 | 관념적/대인적 | 주제와 과제 수행에 관련된 내용을 개념적으로 온전히 이해 | |
| | C2. 데이터의 비판점과 개선 방향 | 관념적/대인적 | | |
| | C3. 문제 해결 방안/후속 연구/미래 방향 제시 | 관념적/대인적 | | |

이들 이동 마디는 학습자의 데이터 설명문에서 항상 모두 출현하는 것은 아니며, 단계 역시 순차적이거나 필수적으로 선택되는 것은 아니다. 학습자에 따라 특정 이동 마디를 사용하지 않는 경우도 있고, 단계의 경우에는 보통 한 이동 마디에 한두 가지만 나타나는 경우가 대부분이다. 이들의 메타 기능적 의미 및 인지적 능력의 적용을 구체적으로 살펴보면 다음과 같다.

우선 ‘A. 주제 및 방향 요약’의 첫 단계는 ‘A1. (소)제목을 활용한 주제 제시’로 독자에게 ‘여기를 보라’는 명시적 지시를 통해 저자 우위의 대인적 메타 기능을 보여준다(O’Halloran, 1996:168). 이와 더불어 독자의 시선을 사로잡고, 내용을 구분하는 텍스트적 메타 기능을 나타낸다(Miller, 1998:32). 이러한 텍스트적 메타 기능은 전문 독자들이 필요한 부분만을 발췌해서 읽는 이른바 ‘비선형적 읽기’를 안내하는 역할을 하기도 한다. 이러한 판단은 (소)제목 아래에 관련된 정보가 있을 것이라는 고정된 양식성에 대한 독자의 신뢰에 근거한다.

‘A2. 연구의 주제와 내용 및 방법 설명’은 데이터 설명문에 제시하려는 주제와 내용이 무엇인지, 제공된 표와 그래프가 무엇을 위해, 누구로부터 어떻게 수집한 자료에 의한 것인지를 저자가 제시하는 부분이다. 본고의 시각적 데이터는 연구자가 학습자들에게 제시했기 때문에 학습자가 직접 만들어낸 시각적 데이터가 아니지만, 학습자 데이터 설명문에서는 표에 나온 대상들에게 ‘설문 조사’한 것을 분석/해석/설명한다는 언급이 다수 발견되었다. 내용과 방법에 대한 구체적인 설명이 관념적 메타 기능에 속한다면, 텍스트적 메타 기능은 Crismore et al. (1993)이 제시한 주제 부각어(topicalizers)로 수행된다. 정혜승(2012:470)에서는 이 주제 부각어를 ‘다들 화제와 목표를 지시’한다고 하였는데, A2는 다음에 올 내용을 예고하고 직접 그 내용으로 진입하게 한다는 점에서 주제 부각어와 같은 자질을 갖고 있다 할 수 있다.

‘A3. 데이터 현상에 대한 배경 설명’은 과제인 ‘코로나 시기 성별 가사노동’과 관련하여 저자들이 갖고 있는 지식적 측면을 확인할 수 있는 부분이다. 이들 단계에서는 주제와 관련된 내용을 개념적으로 이해하고 관

런 사항을 수록한다.

‘B1-1. 연대기적 나열’ 및 ‘B1-2. 시각적 데이터 순서대로 나열’은 시각적 데이터에 나타난 데이터를 전체적으로 제시하는 방법 중 하나이다. 이 단계에 속하는 문장 및 구는 제시된 데이터의 위계와 그룹에 대해 고려하지 않고 단순한 나열을 하기 때문에 인지적 난도가 높지 않다는 데 그 특징이 있다. 또한 이 이동 마디 단계는 데이터의 내용과 추세를 전체적으로 제시하여 논의의 전형적인 장을 만들어 주는 관념적 속성이 보인다. 이지연(2023)에서 나타난 경제학 및 경영학 학술지에서는 이렇게 전체적으로 데이터의 내용을 나열하여 논의의 초점과 전반적인 현상에 집중하게 한 후 특이점이 나타나는 데이터를 선택하는 경향을 보였다. 학습자 데이터 설명문에서도 이와 동일하게 나타날지는 분석을 통해 밝힐 것이다.

‘B2-1. 수치가 높은/낮은 순서대로 나열’은 전반적으로 모든 데이터를 나열한다는 측면에서 ‘B1. 구체적인 정보 추출’ 단계에 속한 하위 단계들과 결이 유사하나, 인지적 측면에서 그리고 텍스트적 메타 기능 측면에서 현저하게 높은 중요한 정보를 인식하고 그 순서에 따라 배열한다는 점에서 B1과 차이가 있다고 보았다. 이는 또한 독자에게 제일 많거나 높거나 적은 수치에 집중하게 하는 저자의 의도가 개입한다는 점에서 대인적 메타 기능도 확인할 수 있다.

이렇듯 저자의 의도가 개입한 선택적인 데이터 나열이라는 측면에서 공통점을 띠기 때문에 ‘B2-2. 수치적 비교, B2-3. 특성에 따른 그룹화, B2-4. 특이성/중요 지점 지적’은 대인적 특성을 갖고 있다. 이들 단계에서는 무엇이 중요하고 적절한 정보인지를 저자의 관점에서 선택한다. 그리고 이들을 언어 기호로 언급하며 시각적 데이터와 언어 자료의 응집성을 나타내는 역할을 한다. 또한 주어진 자료에서 중요 정보를 파악하고 언급할지 여부를 선택하는 ‘판단’은 CEFR(2020)과 IELTS의 인지적 능력에 속하기 때문에 인지적 요소가 포함된 이동 마디 단계라 볼 수 있다. 구체적으로, B2-2의 경우는 수치가 무엇에 비해 얼마나 적고 많은지를 계산해야 하기 때문에 인지적 노력이 다른 이동 마디에 비해 더 소요

되는 편이다. 그러나 그 다음에 나타나는 B2-3의 경우, 그룹화를 하는 기준은 주어진 시각적 데이터의 범례에 해당하는 ‘남자, 여자’, ‘맞벌이 외벌이’, ‘기혼, 미혼’ 등이므로 새로운 그룹을 직접 생성하지 않는 한 인지적 노력은 상대적으로 적게 필요하다. B2-4의 경우 두드러진 부분을 인식하고 제시하는 데서 어느 정도 인지적 노력이 필요하기는 하나, 그에 대한 해석이 후행하거나 염두에 두어지지 않는다면 이 역시도 B2-2 보다는 인지적 노력이 적게 든다고 할 수 있다.

‘B3. 시각적 데이터 언급 표지 사용’은 Miller(1998:36)에서 해석한 바와 같이 저자가 명시적으로 독자에게 시각적 데이터와의 사이에 관계를 만들어 주는 것으로 볼 수 있다. 이는 또한 적절한 양식으로 “시각적 데이터와 언어적 자료로 시선을 오가며 그 순서를 인식하게 하는(이지연, 2023:166)” 텍스트적인 응집성을 확인할 수 있도록 해 준다. 또한 독자에게 어디를 볼지 정해줌으로써 그에게 영향력을 행사할 수 있는 대인적 메타 기능도 수행한다.

‘B4. 수치의 의미 제시’는 ‘C1 주어진 데이터/상식적 지식에 근거한 해석 이동 마디’의 ‘해석 및 결론’과 유사한 듯 보이나, 해당 절 앞에 제시된 수치에 국한된 분석과 추정을 제시한다는 점에서 차이가 있다. 그 의미 제시의 근거는 저자의 데이터 선택에 의한 판단인 경우가 많다. 그를 독자들에게 설명하거나 설득하기 위해 다시 한 번 정리해 제시하는 것이 B4에 해당한다. 따라서 이는 매우 대인적인 단계라 할 수 있다.

‘C1. 주어진 데이터/상식적 지식에 근거한 해석, C2. 데이터의 비판점과 개선 방향, C3. 문제 해결 방안/후속 연구/미래 방향 제시’는 앞에서 제시했던 모든 데이터의 결과를 종합하는 것이다. 이는 “저자의 의도를 실현하는 부분에서 타당성과 논리성을 이끌어내기 위한 요소로 사용되므로 논리적 메타 기능도 갖고 있다고 할 수 있다(이지연, 2023:167).” 다만 C2의 경우에는 학습자 데이터 설명문에서 자주 사용되지 않는 경향을 보였다. Swales & Feak(1994:76)에서 데이터 설명문이 해석, 비판, 텍스트적 기대를 사전에 심어 주어야 한다는 주장을 했기 때문에 핵심적이라 생각되지만, 학술지에서는 이 이동 마디가 데이터 설명문으로 정의된 범

위 안에 존재하기보다는 전략적으로 결론 부분에 따로 등장하는 경향을 보인다. 이상의 이동 마디 단계를 적용하여 분석한 예는 다음과 같다.

<표 19> 이동 마디 단계를 적용한 분석의 예

(A2)코로나 시기 성별 가사 노동 시간 연구에 대한 설문 조사를 실시하였다. (B3)<표1>에 따르면 (B1-2)설문 조사를 참여한 대상자는 남500명과 여 500명 중에서 학력이 대졸의 사람이 가장 많았으며 620명이었다. 그 다음으로 고졸, 대학원졸, 종졸의 순으로 이루어졌다. 하루 평균 가사 노동 시간에 대하여 1시간에서 2시간을 일하는 답변자가 가장 많았고 1시간 미만, 2시간이상 등의 답변이 그 뒤를 이었다. (B2-4)혼인 상태에 대해 대부분의 설문 참여자는 기혼 상태이다.

(B3)맞벌이 상태별 하루 가사 노동 시간에 대한 조사 결과에 따르면 (B4)여성이 일하는 것과 상관 없이 모두 가사 노동 시간이 남성보다 훨씬 더 많다는 것을 알 수 있다. (B1-2)특히, 맞벌이의 가정에서도 여성이 가사 노동 시간은 평균 129분이며 남성은 27분의 큰 차이가 있었다. 남성이 일하는 외벌이 가정에서는 여성의 노동 시간은 훨씬 더 많고 240분의 큰 수차에 도달했다. (B4)아내가 일하는 외벌이 가정에서 여전히 여성의 노동 시간이 더 많다는 결과를 보였다.

(B3)<그림2>를 살펴보면 (B2-3)여성의 하루 평균 가사 노동 시간은 2020년의 187분에서 2021년의 180분으로 내리지만 큰 변화가 없었다. 반면에 남성의 하루 평균 가사 노동 시간은 2019년부터 2021년까지 48분에서 78분으로 큰 폭으로 올랐다. (B4)여전히 여성의 반도 안되는 수치였지만 남성이 가정에서 일하려고 노력하고 있는 것을 알 수 있다. (C1)이는 코로나때문에 재택근무 하는 남성 직장인이 많아지면서 가사 노동 시간이 증가하는 것과도 관련이 있습니다. 그래도 여성은 일하는지와 상관없이 가정에서 가사를 많이 하는 편이다. 직장 일을 하면서 가사 일도 해야 한다는 것은 여성들에게 큰 부담이다. 잘못된 계산식은 오랫동안 여성이 가사를 해야 한다는 사회 고정관념 때문이라고 생각한다. 여성이 가사를 안 하면 오히려 비판을 받을 수도 있다.

(C3)우리 사회는 여성의 부담을 조금이라도 줄이기 위해 여자가 집에서 일해야 하는 사회 고정관념을 바꿔야 한다. 부부가 서로 도와주는 가정 분위기를 만들기 위해 같이 노력해야 한다.

4.2. 학술적 데이터 설명문의 어휘문법

본고에서는 학술적 글쓰기에 필요한 데이터 설명문의 특성을 알아보기 위해 학술적 텍스트의 말뭉치 언어학적 접근에 주목한다. 즉, 장르를 밝히기 위한 이동 마디 분석과 대규모 언어 처리를 중심으로 하는 말뭉치 분석을 따로 보는 것이 아닌, 이들 모두를 함께 이용하는 것이다.

이들을 따로 볼 경우 Swales(1990), Bunton(2002), 박은선(2006)으로 이어지는 이동 마디 연구 흐름에서는 학술적 텍스트의 거시적인 CARS 모형만을 추출할 수 있다. 이는 학습자들이 학술적 텍스트의 거시적인 순서만을 확인하게 할 뿐 실제적으로 그를 구성하는 어휘문법적 미시 체계에 대해서는 말해주지 못한다. 반면 Biber(2006), 남길임(2013), 남길임 외(2014), 허희정(2016) 등은 말뭉치를 이용한 특정 텍스트 분야 전체의 정형 표현 추출이 곧 장르성을 드러낸다고 보아 각 학술적 텍스트 내의 구성과는 별개로 빈도, 범위 및 정형 표현을 추출한다. 이들 연구는 학술적 텍스트의 이동 마디와 단계별로 정형 표현이 달라질 수 있다는 세밀한 접근은 하지 못한다는 단점이 있다.

이러한 면을 보완하기 위해 본고에서는 학술지 논문의 서론을 수집한 백만 단어 말뭉치에서 어휘 꾸러미(lexical bundle)를 찾아내어 각 이동 마디와 연결한 Cortes(2015)의 연구 및 한국어 교육학 학위 논문 서론의 각 이동 마디에서 고정적이고 보편적으로 사용되는 다단어 단위를 추출한 윤새롬(2019)의 접근과 같이 이동 마디별로 고정 표현을 추출하는 방식을 사용하기로 한다. 즉, 학술적 텍스트 전체의 특성이 아닌, 데이터 설명문에 해당하는 부분의 각 이동 마디별 정형 표현을 알아보는 것이다. 이는 각 이동 마디가 갖는 의미를 어휘문법적으로 표상하는 역할을 맡는다.

또한 매개적 접근에서도 인지적 능력을 나타내는 기준뿐 아니라 언어적 능력에 대한 판단 요소를 구조와 관련이 있는 응집성 측면과 어휘문법과

관련이 있는 미시적 측면으로 제시하고 있다. 특히 미시적 측면에서는 학습자 오류와 같은 부분을 중점적으로 다루고 있는데, 이는 데이터 설명문의 어휘문법적 사용에서 아직 무엇이 정확하고 적절한지에 대한 데이터 수집이 다수 이루어지지 않았기 때문에 본고에서는 제외하겠다고 위에서 언급한 바 있다. 본고는 학습자의 쓰기 숙달도가 아닌 이들이 작성한 데이터 설명문의 특성에 대해 확인하고자 하므로 n-gram 목록에서 나타나는 일반적인 특성에 초점을 두어 탐구하도록 한다. 특히 본고에서 사용되는 이동 마디는 메타 기능적 의미를 지닌 것이 대부분이다. 따라서 이에 속한 어휘문법에서 이러한 메타 기능적 의미를 발견할 수 있는 구성이 있다면 이를 파악하도록 한다.

정리하자면 학습자 데이터 설명문 분석을 위한 어휘문법 요소는 이동 마디별 n-gram 분석을 통한 내용어, 내용어+기능어, 기능어+기능어 구성이다. 이들의 해석에는 응집성과 체계 기능적 메타 기능이 개입한다. 이는 아래 <그림 4>와 같이 표현될 수 있다.

| n-gram을 통해 추출한 내용어와 기능어 | |
|------------------------------------|-------------------------|
| 매개와 평가 관련 논의에서의 응집성 | 이동 마디별 메타 기능 |

<그림 4> 학술적 데이터 설명문의 언어 요소

보다 정밀한 논의를 위해 본고에서는 n-gram 추출 시 A1을 제외한 나머지 이동 마디에서 제시한 과제의 제목, 표 및 그래프의 제목과 변수 이름은 제외하도록 한다. 내용어의 확인이 중요하기는 하나, 시각적 데이터에서 제시한 어휘를 그대로 사용하는 것은 학습자가 생각해 낸 내용어

가 아니기 때문에 의미가 없는 것으로 생각된다. 예를 들어, ‘가사 노동’은 4-gram 구성에서 ‘가사 노동 시간과’, ‘가사 노동 시간의’, ‘가사 노동 시간이’, ‘가사 노동 연구에’, ‘가사 노동의 무게’와 같이 다수 나타나나 이 어휘 자체가 반복되는 것은 학습자의 언어 능력과 큰 관련이 없다고 보았다. 또한 위에서 본 것처럼 내용어뿐만 아니라 내용어+기능어, 기능어+내용어, 기능어+기능어 구성도 정형 표현에 해당하므로 함께 확인하도록 한다.

n-gram은 단위가 높아질수록 반복되는 어휘문법의 개수가 줄어든다. 한 이동 마디의 말뭉치 규모가 매우 작기 때문에 본고에서는 일반적으로 내용어와 기능어의 가장 작은 단위로 처리하는 3-gram을 기준으로 하고, 4-gram을 참고하여 확장된 형태를 확인한다.

Ⅲ. 한국어 학습자의 학술적 데이터 설명문 양상 분석

본 장에서는 한국인 대학생과 외국인 대학생의 차이가 외국인 대학생의 어떤 변인에 따라 나타는지 확인하는 것을 목표로 둔다. 따라서 앞으로 본 장에서 제시할 이동 마디 순서 유형, 이동 마디 단계의 빈도 분석, 어휘문법적 목록 분석의 순서를 다음과 같이 제시한다. 첫째, 한국인 대학생과 외국인 대학생의 차이를 우선 확인할 것이다. 둘째, 외국인 대학생의 언어권 및 한국어 숙달도에 따른 차이를 파악할 것이다.

한국 내 유학생 국적 비율은 중국어권 1위, 베트남어권 2위로 연구 참여자 모집 시에도 이 두 언어권 학습자의 참여율이 높았다. 이에 언어권은 연구 참여자의 모어 중 50명 이상 군집이 확인되는 중국어권(59명)과 베트남어권(50명)으로 나누어 두 집단만을 살펴본다.³⁷⁾ 숙달도 기준에서는 인원을 50명 이상 모집하였기 때문에 중급 70명, 고급 55명을 대상으로 분석한다.

1. 학습자의 학술적 데이터 설명문의 담화 구성

1.1. 이동 마디 순서 유형과 빈도

이동 마디 순서 유형은 대체적으로 상위 5개에 집중적으로 나타나는 경향이 있다. 따라서 전체 순서 유형과 대비하여 상위 5개 순서 유형이 얼마나 더 많이 나타나는지를 백분율로 확인한다. 이는 비교하고자 하는 각 학습자군별 상위 5개 순서 유형이 각기 다르기 때문이기도 하고, 나타내는 수치 중 가장 높은 것이 ABC 유형에서 외국인 학습자군의 30개 텍스트이므로 이들을 통계적으로 비교하기에는 수치가 너무 적어 유의한 결과가 나오지 않기 때문이다.

37) 나머지 언어권별 참여자 수가 각 1~3명이어서 통계적으로 의미 있는 양상의 확인이 어렵다고 판단하였다.

(1) 외국인 대학생과 한국인 대학生の 비교

본 절에서는 이동 마디 순서는 이동 마디 대분류인 ‘A. 주제 및 방향 요약’, ‘B. 수치/결과 나열’, ‘C. 해석 및 결론’으로만 제시하도록 한다. 또한 같은 이동 마디 내에 속하는 단계가 앞뒤로 존재하여도 이는 하나의 상위 이동 마디로 간주한다. 예를 들면, 위의 <표 19>와 같이 ‘(A2)-(B3)-(B1-2)-(B2-4)-(B3)-(B4)-(B1-2)-(B4)-(B3)-(B2-3)-(B4)-(C1)-(C3)’ 이동 마디 구성을 보이는 데이터 설명문이 있을 때, 그 이동 마디 순서는 ABC로 둔다. 다만 ‘(A1)-(B2-1)-(C1)-(A1)-(B1-2)-(C3)’과 같이 상위 이동 마디가 구분되는 형태라면 ABCABC로 표현한다. 이렇게 축소된 형태로 이동 마디 순서를 제시하는 까닭은 이동 마디의 하위 단계가 16개에 달하여 이들 부류를 모두 나열하면 이동 마디 순서의 경향성을 파악하기 어렵기 때문이다.³⁸⁾ 따라서 다음 절에서 하위 이동 마디의 빈도 차이와 그 해석을 구체적으로 다루도록 하고 여기에서는 이동 마디 순서에 집중하도록 한다. 이러한 기준에 맞추어 외국인 대학생과 한국인 대학生の 이동 마디 순서를 비교한 것은 아래 <표 20>과 같다.

<표 20> 외국인 대학생과 한국인 대학生の 이동 마디 순서 유형 비교

| 연 번 | 이동 마디 순서 유형 | 외 국 인 | 한 국 인 | 연 번 | 이동 마디 순서 유형 | 외 국 인 | 한 국 인 |
|--------|-------------|-------------|-------------|--------|-------------|-------------|-------------|
| 1 | AB | 25 | 1 | 27 | ABCABCB | 0 | 1 |
| 2 | ABAB | 4 | 0 | 28 | ABCABCBCBAB | 0 | 1 |
| 3 | ABABAB | 4 | 0 | 29 | ABCAC | 0 | 1 |
| 4 | ABABABAB | 1 | 0 | 30 | ABCB | 2 | 0 |
| 5 | ABABABABABC | 3 | 1 | 31 | ABCBABC | 0 | 1 |

38) 유민애(2017)의 이동 마디 연구에서는 11개 이동 마디가 완전히 일치하여 3회 이상 나타나는 구조는 총 3개이고, 이 구조에 속하는 외국인 학습자 텍스트는 137편 중 9편밖에 없음을 밝힌 바 있다. 또한 본고의 <표 20>에서 볼 수 있듯, 상위 이동 마디의 배열만으로도 대단히 많은 유형이 존재함을 알 수 있다.

| | | | | | | | |
|---------|----------------|----|----|----|------------------|-----|----|
| 6 | ABABABCAC | 1 | 0 | 32 | ABCBABCBCBABC | 1 | 0 |
| 7 | ABABABABC | 1 | 0 | 33 | ABCBC | 8 | 4 |
| 8 | ABABABABCB | 1 | 0 | 34 | ABCBCABC | 2 | 0 |
| 9 | ABABABAC | 1 | 1 | 35 | ABCBCBAC | 0 | 1 |
| 10 | ABABABACBC | 1 | 0 | 36 | ABCBCBC | 1 | 3 |
| 11 | ABABABC | 5 | 2 | 37 | ABCBCBCBCBC | 0 | 2 |
| 12 | ABABABCABC | 1 | 0 | 38 | ABCBCBCBCABCBCAC | 0 | 1 |
| 13 | ABABABCBAABCBC | 1 | 0 | 39 | ABCBCBCBCBCBCAC | 1 | 0 |
| 14 | ABABABCBC | 1 | 1 | 40 | ACBABABABC | 1 | 0 |
| 15 | ABABABCBCBCBC | 0 | 1 | 41 | B | 4 | 0 |
| 16 | ABABACBC | 0 | 1 | 42 | BAB | 3 | 0 |
| 17 | ABABC | 6 | 5 | 43 | BABAB | 1 | 0 |
| 18 | ABABCABC | 0 | 2 | 44 | BABABABC | 1 | 0 |
| 19 | ABABCBAABCABC | 0 | 1 | 45 | BABABABCBC | 1 | 0 |
| 20 | ABABCBCA | 0 | 1 | 46 | BABCB | 1 | 0 |
| 21 | ABABCBC | 7 | 1 | 47 | BACBABA | 1 | 0 |
| 22 | ABABCBCB | 1 | 0 | 48 | BC | 0 | 1 |
| 23 | ABABCBCBC | 0 | 2 | 49 | BCB | 0 | 1 |
| 24 | ABABCBCBCBC | 1 | 0 | 50 | BCBC | 1 | 1 |
| 25 | ABC | 30 | 10 | 51 | BCBCBC | 0 | 1 |
| 26 | ABCABABCABC | 0 | 1 | 52 | BCBAB | 1 | 0 |
| 합계: 175 | | | | | | 125 | 50 |

<표 20>에서 한국인 및 외국인 대학생 모두를 대상으로 확인한 전체 이동 마디 순서 유형은 총 52개이며, 외국인 대학생은 34개, 한국인 대학생은 27개를 선택하였다. 이는 외국인 대학생이 전체 이동 마디 순서 유형 중 65%를, 한국인 대학생이 51%를 사용함을 알 수 있다. 따라서 외국인 대학생들은 한국인 대학생들에 비해 더 다양한 이동 마디 순서 유형을 사용하는 것을 확인할 수 있다.

<표 21> 외국인/한국인 대학生の 52개 이동 마디 순서 유형 중 사용 개수와 비율

| | 사용 유형 개수 | 전체 유형 중 사용 유형 비율 |
|---------|----------|------------------|
| 외국인 대학생 | 34 | 65% |
| 한국인 대학생 | 27 | 51% |

전체 52개 이동 마디 순서 유형 중 31개가 한 편의 데이터 설명문에만 해당된다. 다만 1개씩 나타나는 유형들은 두 집단이 서로 겹치지 않아 외국인 대학생만이 사용하는 이동 마디 유형과 한국인 대학생만이 사용하는 이동 마디 유형이 존재함을 알 수 있다. 즉, 한국인 대학생과 외국인 대학생이 공유하는 이동 마디 순서 유형은 52개 중 10개에 불과하다. 이러한 이동 마디 순서 유형의 다양성 또는 과편성은 텍스트의 장르성이 명확하지 않거나 장르성을 잘 모르는 필자들이 텍스트를 구성했을 때 나타난다. 유민애(2017:94)에서는 외국인 학습자들의 텍스트에 나타나는 이러한 과편적 경향성을 두고 한국인에 비해 일관성 없이 여러 이동 마디 순서 유형을 사용하여 이동 마디의 의미적 기능을 약화한다고 말한 바 있다. 본고의 데이터 설명문은 미니 장르로서 고정적인 이동 마디 순서가 상위 장르에 따라 변화하기 때문에 한국인 대학생과 외국인 대학생 모두가 대단히 다양한 순서 유형을 보이는 것으로 생각된다.

그럼에도 그 속에서 학습자들은 특정 이동 마디 순서 유형을 집중적으로 선호하는 것을 볼 수 있다. 이동 마디 순서 유형을 빈도가 높은 순으로 상위 5개를 선정하고, 이들이 전체 이동 마디 순서 유형 선택에서 차지하는 비율을 알아보면 다음과 같다.

<표 22> 외국인/한국인 대학생이 선택한 상위 5개 이동 마디 순서 유형

| 외국인 대학생 | | | | 한국인 대학생 | | | |
|---------|-------------|----|-------|---------|-------------|----|-------|
| | 이동 마디 순서 유형 | 편수 | % | | 이동 마디 순서 유형 | 편수 | % |
| 1 | ABC | 30 | 24.0% | 1 | ABC | 10 | 20.0% |
| 2 | AB | 25 | 20.0% | 2 | ABABC | 5 | 10.0% |

| | | | | | | | |
|---|---------|----|-------|---|---------|----|------|
| 3 | ABCBC | 8 | 6.4% | 3 | ABCBC | 4 | 8.0% |
| 4 | ABABC | 6 | 4.8% | 4 | ABCBCBC | 3 | 6.0% |
| 5 | ABABCBC | 5 | 4.0% | | | 22 | 44% |
| | | 74 | 59.2% | | | | |

외국인 대학생의 경우에는 상위 5개 유형이 과반인 전체 이동 마디 순서 유형의 59.2%를 차지하여 나머지 47개보다 더 대표성이 있음을 보여주고 있다. 한국인 대학생의 경우에는 5위에 해당하는 이동 마디 순서 유형 4개가 모두 동일하게 2편으로 나타나 4위까지만 수집하였다. 상위 4개 유형에 속하는 데이터 설명문이 전체의 44%를 차지하고 있고, 만약 5위에 해당하는 4개 유형까지 포함한다면 60%를 차지해 이 역시도 특정 이동 마디 순서 유형에 쏠림 현상이 있음을 확인할 수 있다.

ABC 유형은 앞 장의 <표 19>에서 볼 수 있듯, 일반적인 텍스트 구조인 서론-본론-결론 구조와 유사해서인지 외국인/한국인 대학생 모두가 제일 선호하는 이동 마디 순서 유형이다. 이 유형에서는 연구의 방향이나 주제, 설문조사라는 연구 방법에 대해 간략히 소개하고 주어진 모든 시각적 데이터를 한 번에 제시한다. 이 과정에서 각 시각적 데이터를 따로 언급하여 문단을 나누거나 시각적 데이터의 데이터를 합쳐서 일괄적으로 또는 관련성에 따라 결합하여 제시하는 경우도 있다. 그리고 마지막으로 이들 시각적 데이터가 나타내는 현상에 대하여 자신의 해석이나 분석을 덧붙인다. 그러나 외국인 대학생이 두 번째로 선호하는 이동 마디 유형인 AB 유형은 한국인 대학생 집단에서는 한 편밖에 발견할 수 없었다. 이는 이후 외국인 대학생 변인별 분석에서 다시 알아보도록 하겠다.

ABCBC, ABABC 유형은 외국인/한국인 대학생 모두 선호도가 중간 정도에 속한다. 또한 세부 구성에 특징이 있지만 전반적으로 ABC 유형의 흐름과 유사하다. 각 이동 마디 순서의 세부 특징은 실험 과제가 응답자 정보에 대한 표 1개, 성별 가사 노동 관련 그래프 2개를 다루기 때문에 이들을 저자 기준에 따라 두 개로 분리해 서술한 구성이라 할 수 있다. 따라서 표와 그래프 2개를 각각의 분석 대상 데이터로 간주하고 데이터

설명문을 작성한다면 자주 사용되는 이동 마디 순서 유형이라 할 수 있다.

예를 들어, ABCBC 유형은 다음 <표 23>과 같이 A에 어떤 연구를 하려는지 언급하고, B에서 표와 첫 번째 그래프의 정보 추출에 집중한 뒤 C에 해당하는 관련 해석을 덧붙인다. 그리고 그다음 그래프에서 동일한 과정을 반복한다.

<표 23> ABCBC 유형 데이터 설명문의 예

(A2)남, 여 같은 수의 표본을 조사한 것이다. (B2-1)학력은 대학생이 가장 많고 고졸 및 대학원졸이 그 뒤를 잇는다. (B2-4)하루 평균 가사노동 시간은 1시간~3시간이 가장 많다. (B2-1)표본에 속하는 대상의 다수 즉 약 75퍼센트는 기혼자이므로 부부 각 일방의 가사노동 시간을 추측할 수 있으나, 독신자(비혼, 미혼, 기혼, 사별) 비율이 25퍼센트 가까이 되어 (B4)자료 분석에 교란요인이 될 수는 있을 것이다. 데이터에서 가장 먼저 눈에 띄는 점은 (1)어떤 형태의 가족이든 간에 여성이 가사노동에 투입하는 시간이 더 길다는 것이다. 남성 외벌이 가정에 있어서 여성이 가사노동에 투입하는 시간이 길다는 것은 충분히 납득가능하다. (C1)하지만 맞벌이의 경우 그리고 심지어 여성 외벌이 가정의 경우에도 여성이 가사노동에 투입하는 시간이 길다는 것은 전통적 성역할을 고수하려는 사회 분위기 외에는 다른 이유를 생각하기 어렵다. (2)다음으로 눈에 띄는 점은, (B3)그림 제2에 의할 때 (B2-3)남성의 가사노동 시간이 2019년 이후 현재에 이르기까지 지속적으로 상승(일 평균 48분에서 일 평균 78분으로) 해왔음에도 불구하고, 여성의 가사노동 시간은 거의 일정하게(최단 일 평균 174분에서 최장 일 평균 178분까지) 유지되고 있다는 점이다. (B4)이는 다시 말해 남성이 가사노동에 참여하는지 여부 그리고 얼마나 오랜 시간을 투입하는지 여부와 무관하게 여성의 가사노동 투입시간은 거의 변화가 없었다는 것을 의미한다. (C1)아마 그 이유는 남성의 가사노동이 그 가족에게 제공하는 효용이 크지 않아(설것이를 해도 그다지 깨끗하게 하지 않음) 그 가정에 속한 여성이 종래와 별반 다를 바 없을 정도의 노력을 투입해야만 가사노동의 실효성이 보장되기 때문일 것이다. 또 다른 원인으로는, 가사노동의 결과 수준에 대한 기대치가 남성의 경우보다 여성의 경우에 높기 때문(예를 들면 남성이 기대하는 설것이 수준보다 여성

이 기대하는 설것이 결과물의 청결도 수준이 높아서)일 수도 있다고 추측된다.

[9-한-경영4-남]

<표 24>는 첫 번째 A에 응답 정보와 배경을 서술하고 뒤를 잇는 B에서 이를 해석한 뒤, 그다음 A에서 그래프 2개를 순차적으로 함께 제시한다. 그리고 뒤따르는 B에서 이들 그래프의 의미를 혼합하여 해석한다. 그리고 C에서 마지막으로 향후 현상의 개선 방향이나 자신의 해석을 덧붙인다.

<표 24> ABABC 유형 데이터 설명문의 예

(A2)코로나 시기 성별 가사 노동 시간을 알아보기 위해 1000명 중에 남성 500명과 여성 500명, 비혼인 280명과 기혼 720명을 대상으로 하루 평균 가사 노동 시간에 대한 설문조사를 실시했다. (B3)설문 조사 응답자의 기본 정보는 <표 1>와 같다. (B2-1)설문 응답자 중에 대학 졸업자는 620명으로 가장 많으며 고졸 학교 졸업자는 260명, 대학원 졸업자는 110명, 마지막으로 중학교 졸업자는 10명 순이었다. 응답자의 하루 평균 가사 노동 시간은 1시간부터 2시간까지는 402명으로 가장 많으며, 1시간 미만은 280명, 2시간부터 3시간까지는 227명, 3시간부터 4시간까지는 74명, 그리고 4시간 이상은 17명 순이었다. (B4)이를 통해 대부분의 응답자는 고등학교를 졸업했으며 하루 평균 집안일을 하는 시간은 최대 3시간이라는 것을 알 수 있다. (A2)먼저, 남성과 여성의 하루 평균 가사 노동 시간을 비교하기 위해 2019년부터 2021년까지 3년동안 조사하였는데, (B2-3)여성의 가사노동 시간은 2019년에 174분에서 2020년에 187분으로 살짝 높아지고 마지막으로 2021년에 180분으로 내리는 반면 남성의 경우는 시간이 지나갈 수록 48분으로부터 61분, 그리고 78분으로 지속적으로 높아졌다. (B4)이처럼 코로나 시기에 여성의 가사 노동 시간이 큰 변화를 보여주지 않는 것과 달리 남성의 경우는 그 시간이 점점 늘어갔다는 것을 볼 수 있다. (B2-2)그러나 그 사실이 있음에도 불구하고 여성의 가사 노동 시간은 여전히 남자보다 3배나 많다는 것을 볼 수 있다. (B3)다음으로 맞벌이 상태별 하루 집안일을 하는 시간을 조사하였으며 결과는 <그림 1>와 같다. (B2-3)맞벌이, 남편인 가장, 아내인 가장의 가정을 순으로 남성의 가사 노동 시간은 27분, 31분, 66분으로 나타

났으며 여성의 경우는 129분, 240분, 106분이었다. **(B4)**따라서 한 가정에 돈을 버는 사람은 누구든 집안일을 더 많이 하는 사람은 항상 여성인 것을 알 수 있다. 특히 맞벌이 가정의 남편과 외벌이(남편) 가정의 남편의 가사 노동 시간은 큰 차이가 없으나 여성은 2배나 더 많이 한다는 것을 볼 수 있다. **(C1)**따라서 누구나 집에서 지내는 시기인 코로나 시기에 남자의 가사 노동 시간이 늘어남에도 불구하고, 여성은 남성과 마찬가지로 경제 활동에 참여함에도 불구하고 집안일은 여전히 여성의 부담이라는 현실을 알게 되었다.

[86-고급-베트남어-경영1-여]

나머지 이동 마디 순서 유형 중 하나인 ABABCBC는 외국인 대학생의 상위 이동 마디 순서에서 나타난다. 이 역시도 AB 유형과 마찬가지로 다음 절 외국인 대학생의 이동 마디 순서 분석에서 다시 다루기로 한다. 마지막으로 한국인 대학생의 이동 마디 순서에서 주로 나타나는 ABCBCBC는 표와 그래프를 하나씩 나열할 때 주로 나타나는 이동 마디 순서 유형이다. 이에 대한 예는 다음 <표 25>에서 볼 수 있다.

<표 25> ABCBCBC 유형 데이터 설명문의 예

(B3)코로나19 바이러스는 2019년 12월부터 전 세계를 강타하였다. 강한 전파력으로 인해 대면으로 진행되던 모든 활동들은 비대면으로 전환되기 시작하였고, 마스크를 착용하지 않으면 실내 생활을 아예 하지 못할 정도로 거리두기 수칙이 강화되었다. 밖에서 생활하는 시간보다 집 안에서 생활하는 시간이 늘어감에 따라, 남성과 여성 모두의 하루 평균 가사 노동 시간이 증가하기 시작하였다.

<그림 2> (삽입함)

(B3)위의 <그림 2>은 코로나가 심한 영향을 끼치기 전인 2019년부터, 코로나가 초절정기로 사회 활동이 거의 중단되었던 2020년, 그리고 코로나의 영향력이 2020년과 비슷한 2021년, 세 해의 남성과 여성 하루 평균 가사 노동 시간을 정리하였다. 이 표에 의하면, **(B2-3)**남 여의 하루 평균 가사 노동 시간의 합은 2019년 222분, 2020년 248분, 2021년 258분으로 지속적으로 증가하였다. **(B4)**이는 곧 집 안에서 머무르는 시간이 압도적으로 길어짐에 따라 야기된 결과이다.

(B3)또한 <그림 2>에서 알 수 있듯, **(B4)**남성과 여성의 하루 평균 가

사 노동 시간에 차이는 2019년 이래로 점점 줄어들고 있다. **(B1-1)**2019년, 여성과 남성의 하루 평균 가사 노동 시간은 각각 174분과 48분으로, **(B2-2)**여성이 남성보다 가사 노동 시간이 3.625배 높은 반면, **(B1-1)**2021년에는 그 차이가 줄어 여성과 남성 각각 180분과 78분 **(B2-2)**(약 2.3배)를 기록하였다. **(B4)**물론 아직까지 여성의 하루 평균 가사 노동 시간이 남성의 하루 평균 가사 노동 시간의 2배 이상 정도 높은 것이 현실이지만, 이 차이는 점차 줄어들고 있음을 알 수 있다. 남성의 하루 평균 가사 노동 시간 증가율이 여성의 하루 평균 가사 노동 시간 증가율 보다 크다.

<그림 1.> (삽입함)

(B3)<그림 1>은 2022년, 남성과 여성의 하루 가사 노동 시간은 직장 과 연계하여 분석하고 있다. **(B2-3)**맞벌이 가정의 경우 남성은 하루에 27분, 여성은 하루에 129분 정도를 가사 노동을 하는 것으로 드러났다. **(B2-2)**이는 약 4.78배 차이로, **(B4)**남성과 여성의 가사 노동 시간 차이가 극심하다는 것을 알 수 있다. **(B2-3)**이 차이는 남편이 외벌이일 경우 더욱 증폭되어 **(B2-2)**가사노동 차이가 약 7.74배 까지 증가한다는 것을 알 수 있다. **(B2-3)**아내가 외벌이하는 경우, 남성과 여성의 가사노동 시간 차이가 줄어들어, 여성이 106분, 남성이 66분으로 그 차이가 약 **(B2-2)**1.61배 정도 하는 것으로 드러났다. **(B4)**물론 아내가 외벌이하는 경우 차이가 줄어들지만, **(C1)**그럼에도 불구하고 여성이 남성의 하루 가사 노동 시간보다 더 높다는 것으로 보아 가사 노동은 여성이라는 유교적 관점이 한국 사회에 남아있는 것으로 추정해볼 수 있다.

<표 1> (삽입함)

(B3)<표 1>은 <그림 1>과 <그림 2>의 설문 응답자의 기본 정보를 나타내고 있다.

<표 1>에서 **(B4)**하루 평균 가사 노동 시간이 4시간 이상이라고 응답한 사람은 17명으로, 외벌이 남편이 있는 가정의 경우 여성이 평균 240분, 약 4시간 정도 일한다는 것을 고려해보았을 때, 기혼자 720명 중 남편 외벌이 가정은 약 17명이 될 것으로 추정된다. 이는 곧 남편 외벌이의 가정은 그 수가 매우 소수임을 알 수 있다.

(C1)코로나19 상황동안 남성 외벌이 가정이 줄어들면서, 맞벌이 혹은 아내 외벌이 가정이 증가하였고 그런 결과로 인해 남편의 가사 노동 시간이 가파르게 증가하는 반면, 여성의 하루 평균 가사 노동 시간 증가율은 그렇게 높지 않다는 것을 유추해볼 수 있다.

(B3) <표 1>을 분석해보면, <그림 1>과 <그림 2>의 데이터는 남, 여 각각 500명씩 설문한 결과이다. (B2-1) 또한 설문 대상 1000명은 비혼은 280명, 기혼은 720명으로 구성되었고 1000명의 응답자 중 중졸이하는 1%, 고졸은 25%, 대졸 및 대학원졸이 73%로 드러났다. (C1) 2020년 기준 대한민국 국민 학력별 인구분포가 중학교 이하 11%, 고등학교 39%, 고등교육 이상이 51%인 것을 고려해보았을 때, (C2) 본 설문은 대한민국의 평균 가사 노동 시간을 정확하게 반영한다기보다는 고학력자의 비중이 더 많이 반영된 지표라고 할 수 있다는 점에서 한계가 있다.

[35-한-경영4-여]

위에서 살펴본 각 집단별 상위 5개 유형을 포함하여 전체적으로 확인할 수 있는 또 다른 경향성으로는, ‘A- 시작 유형’이 ‘B-시작 유형’보다 많다는 것이다. 위에서 본 바와 같이 상위 5개 유형은 모두 ‘A-시작 유형’이다. 전체적으로 보아도 A로 시작하는 이동 마디 순서 유형은 40개인데 반해 B로 시작하는 이동 마디 순서 유형의 개수는 12개로 매우 적으며 그를 채택한 데이터 설명문 편수 또한 ‘A-시작 유형’ 157편, ‘A-시작 유형’ 18편으로 차이가 큼을 알 수 있다. 구체적으로 외국인 대학생은 125편 중 111편, 한국인 대학생은 50편 중 46편의 데이터 설명문이 ‘A-시작 유형’이다. 대다수의 대학생이 전공 과제에서의 데이터 설명문 시작을 A ‘주제 및 방향 요약’으로 수행하는 경우가 많으며, 본론인 B ‘수치/결과 나열’부터 시작하는 것은 드물다고 볼 수 있다. 한국인 대학생은 ‘A-시작 유형’ 이동 마디 순서 선택 비율이 92%에 달해 88.8%인 외국인 대학생에 비해 이 유형의 이동 마디 순서를 더 선호하는 것으로도 볼 수 있다.

<표 26> 외국인/한국인 대학생의 ‘A/B-시작’ 이동 마디 선호도

| | 외국인 대학생 | 한국인 대학생 |
|-----------------------|------------|---------|
| A-시작 유형 편수 (선택 비율) | 111(88.8%) | 46(92%) |
| B-시작 유형 편수 (선택 비율) | 14(11.2%) | 4(8%) |

위의 결과를 정리하자면, 한 이동 마디 순서에 한두 데이터 설명문만 존재하는 경우가 매우 많아서 각 집단 의 거의 반수에 해당하는 대학생들이 이동 마디 순서 유형 선택에 파편화된 경향을 보인다. 이는 수집한 데이터 설명문에 장르성이 고정되어 나타나지 않거나, 대학생 저자들이 데이터 설명문 장르성에 대한 합의가 이루어지지 못한 결과일 수 있다. 또한 이런 파편화된 정도는 한국인 대학생보다 외국인 대학생에게서 그 비율이 더 높게 나타난다.

다만 대학생들이 선호하는 이동 마디 순서 유형은 있다. 선택 빈도가 높은 이동 마디 순서 유형 상위 5개를 뽑아 본 결과, 외국인 대학생 집단에서는 이들이 과반을, 한국인 대학생 집단에서는 이것이 반수 가까이 차지하였다. 이는 대학생들의 데이터 설명문 구조 선택이 특정 유형에 편중되어 있음을 뜻한다. 특히 ABC 유형은 두 집단 모두에서 1위를 차지하였고, ABCBC 및 ABABC도 두 집단 모두에서 중간 수준을 나타내었다. 다만 선호하는 이동 마디 순서 유형은 집단마다 조금씩 차이가 있는 것으로 나타났다. 상위 5개 이동 마디 순서 유형 ‘A-시작 유형’이기는 하지만, 한국인 대학생 집단에서는 ‘AB’가 나타나지 않는 반면, 외국인 대학생 집단에서는 2위를 차지하였다. 또한 가장 순위가 낮은 이동 마디 순서 유형은 각각 상대 집단의 상위 5개에 들지 못하였다. 그렇다면 한국인 대학생들과 차이나는 외국인 대학생들의 이동 마디 순서 유형은 언어권과 숙달도 중 어떠한 변인에 의한 것인지 더욱 자세히 알아보도록 하겠다.

(2) 학습자 변인별 이동 마디 순서 유형과 특성

1) 학습자 언어권별 이동 마디 순서 유형과 특성

우선 외국인 대학생의 언어권별 특성은 어떠한지 알아보도록 하겠다. 위에서 언급한 바와 같이 한국인 대학생과의 수를 맞추기 위해 한 언어권에 50명이 넘는 집단만을 선택하도록 한다.³⁹⁾ 그 결과 중국어권 학습

자 59명과 베트남어권 학습자 50명이 언어권별 두 집단으로 지정되었다. 이들 학습자군이 선택한 이동 마디 순서 빈도는 다음 <표 27>과 같다.

<표 27> 언어권별 이동 마디 순서 빈도

| 연 번 | 이동 마디 순서 유형 | 중 | 배 | 연 번 | 이동 마디 순서 유형 | 중 | 배 |
|--------|---------------|----|---|--------|------------------|---|---|
| 1 | AB | 19 | 2 | 27 | ABCACB | 0 | 0 |
| 2 | ABAB | 2 | 1 | 28 | ABCABCBCBAB | 0 | 0 |
| 3 | ABABAB | 3 | 1 | 29 | ABCAC | 0 | 0 |
| 4 | ABABABAB | 0 | 1 | 30 | ABCB | 1 | 1 |
| 5 | ABABABABABC | 0 | 4 | 31 | ABCBABC | 0 | 0 |
| 6 | ABABABCAC | 0 | 1 | 32 | ABCBABCBCBABC | 0 | 1 |
| 7 | ABABABABC | 0 | 1 | 33 | ABCBC | 4 | 4 |
| 8 | ABABABABCB | 0 | 1 | 34 | ABCBCABC | 1 | 1 |
| 9 | ABABABAC | 0 | 1 | 35 | ABCBCBAC | 0 | 0 |
| 10 | ABABABACBC | 0 | 1 | 36 | ABCBCBC | 0 | 0 |
| 11 | ABABABC | 0 | 5 | 37 | ABCBCBCBCBC | 0 | 0 |
| 12 | ABABABCABC | 0 | 1 | 38 | ABCBCBCBCABCBCAC | 0 | 0 |
| 13 | ABABABCBCBCBC | 0 | 1 | 39 | ABCBCBCBCBCBCAC | 1 | 0 |
| 14 | ABABABCBC | 0 | 1 | 40 | ACBABABABC | 0 | 1 |
| 15 | ABABABCBCBCBC | 0 | 0 | 41 | B | 3 | 0 |
| 16 | ABABACBC | 0 | 0 | 42 | BAB | 1 | 0 |
| 17 | ABABC | 1 | 5 | 43 | BABAB | 0 | 0 |
| 18 | ABABCABC | 0 | 0 | 44 | BABABABC | 1 | 0 |
| 19 | ABABCABCABC | 0 | 0 | 45 | BABABABCBC | 0 | 1 |
| 20 | ABABCBCA | 0 | 0 | 46 | BABCB | 0 | 1 |
| 21 | ABABCBC | 2 | 5 | 47 | BACBABA | 1 | 0 |
| 22 | ABABCBCB | 1 | 0 | 48 | BC | 0 | 0 |
| 23 | ABABCBCBC | 0 | 0 | 49 | BCB | 0 | 0 |

39) 나머지 언어권별 집단은 말레이어 3명, 터키어 3명, 러시아어 2명, 영어 2명, 몽골어 2명, 포르투갈어 2명, 미얀마어 1명, 스페인어 1명으로 그 수가 너무 적어 언어권별로는 다루지 않았다.

| | | | | | | | |
|----|-------------|----|---|----|--------|----|----|
| 24 | ABABCBCBCBC | 1 | 0 | 50 | BCBC | 0 | 0 |
| 25 | ABC | 16 | 8 | 51 | BCBAB | 0 | 0 |
| 26 | ABCABABCABC | 0 | 0 | 52 | BCBCBC | 1 | 0 |
| 계 | | | | | | 59 | 50 |

전체 이동 마디 순서 유형에서 이들 두 언어권 학습자가 모두 선택하지 않은 20개 유형을 제거하면 총 52개 유형 중 32개 유형만이 남는다. 중국어권 학생들은 이 32개 유형 중 17개(53%)만을 사용하고, 베트남어권 학생들은 32개 중 24개(75%)만을 사용한다. 이를 볼 때 베트남어권 학습자들이 중국어권 학습자들보다 이동 마디 순서를 더 다양하게 선택함을 알 수 있다. 따라서 베트남어권 학습자들은 데이터 설명문의 구조를 비교적 덜 고정적으로 인식함을 알 수 있다. 한편, 두 집단 모두 한국인 대학생들이 사용한 28개 유형보다는 절대적인 개수가 적었다. 이는 언어권 별 학습자들이 사용하는 이동 마디 순서 유형의 개수의 절대적 다양성은 상대적으로 크지 않음을 나타낸다. 이는 앞서 이야기했던 바와 같이 데이터 설명문 과제 자체의 학술적 전형성이 고착화된 것이 아니고, 한국인 대학생들이라 하더라도 이러한 고정성을 인식하는 것은 아니라는 사실을 재차 확인하게 해 준다. 따라서 이동 마디 순서 유형 선정 분포가 오히려 물리는 경향이 있음을 알 수 있다.

<표 28> 중국어권/베트남어권 학습자의 32개 이동 마디 순서 유형 중 사용 개수와 비율

| | 사용 유형 개수 | 전체 유형 중 사용 유형 비율 |
|-------|----------|------------------|
| 중국어권 | 17 | 53% |
| 베트남어권 | 24 | 75% |

그렇다면 이들 중 중국어권 학습자와 베트남어권 학습자가 다수 선택하는 이동 마디 순서 유형을 상위 5개까지 추출해 비교해 보도록 하겠다. 베트남어권 학습자들이 다섯 번째로 선호하는 이동 마디 순서 유형은 두 개여서 4위까지만 기재하였다. 그 결과 중국어권 학습자의 이동 마디 순서 유형 상위 5개는 전체 선택의 76%를 차지할 정도로 압도적인

쏟림 현상을 보여주고 있다. 또한 베트남어권 학습자의 경우에도 상위 4 위까지의 이동 마디 순서 유형 선택 비율이 46%, 5위에 해당하는 두 유형까지 포함했을 때는 62%로 과반인 것을 볼 수 있다. 외국인/한국인 대학생의 경우에서와 마찬가지로 이들 상위 유형 5개가 전체 32개 유형에서 사용상 반수 이상인 것을 알 수 있다. 다만 중국어권 학습자의 작문은 ABC와 AB 유형에 과반이 몰린 반면, 베트남어 학습자는 ‘A-()-C’ 내에서 변이형을 사용하는 것으로 보인다.

<표 29> 중국어권/베트남어권 학습자가 선택한 상위 45개 이동 마디 순서 유형

| 중국어권 학습자 | | | | 베트남어권 학습자 | | | |
|----------|-------------|----|-------|-----------|-------------|----|-----|
| | 이동 마디 순서 유형 | 편수 | % | | 이동 마디 순서 유형 | 편수 | % |
| 1 | AB | 19 | 32.2% | 1 | ABC | 8 | 16% |
| 2 | ABC | 16 | 27.1% | 2 | ABABABC | 5 | 10% |
| 3 | ABCBC | 4 | 6.7% | 3 | ABABC | 5 | 10% |
| 4 | ABABAB | 3 | 5% | 4 | ABABCBC | 5 | 10% |
| 5 | B | 3 | 5% | | | 23 | 46% |
| | | 45 | 76% | | | | |

중국어권 학습자와 베트남어권 학습자가 공통적으로 다수 선택하는 이동 마디 순서 유형은 ABC 유형이다. 이는 한국인 대학생들 역시도 가장 선호하는 유형이어서 한국인 대학생과 외국인 대학생 모두에게 보편적으로 여겨지는 구조라 생각할 수 있다. ABCBC와 ABABC 유형의 경우에도 한국인 대학생과 외국인 대학생 전체 집단에서 상위 2~4위를 했던 이동 마디이고 ABC와 공통적으로 ‘A-()-C’ 형태인 유형이라 할 수 있다.

이를 제외한 이동 마디 순서 유형에서는 중국어권 학습자와 베트남어권 학습자의 차이가 드러난다. 우선 중국어권 학습자의 경우에는 AB, ABABAB, B 유형과 같이 ‘해석 및 결론’ 이동 마디를 기술하지 않는 ‘C-누락형’ 이동 마디 유형 선택이 눈에 띈다. 특히 마지막 B 유형은 A와 C 이동 마디를 모두 기술하지 않고 있다. 따라서 중국어권 학습자와 베트남어권 학습자들의 이동 마디 순서 유형을 비교하면 ‘B-시작 유형’

과 맨 마지막의 ‘C-누락 유형’에 차이가 있음을 알 수 있다.

<표 30> 중국어권/베트남어권 학습자의 이동 마디 순서 유형 비교

| | 중국어권 학습자 | 베트남어권 학습자 |
|-----------------------|---------------|------------|
| B-시작 유형 편수 (선택 비율) | 7 (11.8%) | 2 (4%) |
| C-누락 유형 편수 (선택 비율) | 30 (50.8%) | 8 (16%) |

‘B-시작 유형’의 경우 중국어권 학습자에게서는 11.8%, 베트남어권 학습자에게서는 2%밖에 발견되지 않아 전반적으로는 두 집단에서 모두 낮게 나타났다. 다만 중국어권 학습자가 베트남어권 학습자에 비해 이 유형을 비교적 많이 선택했음을 알 수 있다. 그러나 ‘C-누락 유형’의 경우 중국어권 학습자 과반이 선택했음을 알 수 있다. 그렇다면 이러한 누락 유형이 나타내는 양상이 학습자 모어의 영향 때문인지 학습자가 작성한 이동 마디 유형의 예를 보며 확인해 보도록 하겠다.

<표 31> AB 유형 데이터 설명문의 예

(A2) 통계청에서 2022년에 남녀 각 500명을 대상으로 설문 조사를 실시하였다. 설문 응답자의 경우는 다음과 같다. (B1-2) 학력을 보면 중졸 이하, 고졸, 대졸, 다학원졸은 각각 10명, 260명, 620명, 110명이고 하루 평균 가사 노동 시간은 1시간 미만이 280명, (B2-4) 시간 ~ 2시간이라고 대답하는 사람은 402명으로 가장 높은 수치로 나타났다. (B1-2) 2시간~ 3시간이 227명이고 3시간~ 4시간이 74명이었다. (B2-5) 하루 평균 노동 시간 4시간 이상의 경우는 17명밖에 안 된다. (B1-2) 혼인 상태는 비혼인이 280명으로 나타났으며 기혼인이 720명을 있었다.

(B3) 남녀 기혼 상태의 맞벌이 상태별 하루 가사 노동 시간은 다음 <그림 1>과 같다. (B2-3) 맞벌이 가정에는 여성이 129분이 나타났는데 남성이 27분만 있었다. 외벌이 가정에는 남편 일하는 경우에 남성은 일을 해야 하기 때문에 31분밖에 안 됐는데 여성이 240분이 있었다. 아내가 일하는 경우에는 여성은 일하는데 하루 평균 가사 노동 시간이 106분이고 남성의 66분보다 더 많은 것으로 나타났다. (B4) 이 그래프를 통해 남성

이 일을 하든 안 하든, 여성이 남성보다 가사 노동 시간이 더 많은 것을 볼 수 있다. (B2-3)남성과 여성의 하루 평균 가사 노동 시간 비교에 보면 여성은 2019년의 174분, 2020년의 187분, 2021년의 180명으로 높은 데이터로 나타났다. (B2-2)남성은 2019년의 48명, 2020년의 61명, 2021년의 78명으로 꾸준히 증가하고 있는데 여성보다 두배나 적었다. (B4)비록 남성이 가사 노동을 하는데 평균 가사 노동 시간은 여성에 비해 훨씬 적다.

[85-고급-중국어-경영1-여]

우선 AB 유형은 주어진 표의 정보를 모두 기술한 후 그래프 두 개 사이의 구분 없이 데이터를 모두 기술하거나, 연결 등을 전혀 신경 쓰지 않고 표1, 그림1, 그림2의 내용만을 각각 한 문단으로 나타내는 경우가 많다. 즉, 데이터 내용만을 나열하는 구조로 되어 있다. 또한 대부분의 경우 AB 유형은 글의 길이가 다른 유형에 비해 짧은 경우가 많다. 한국어 숙달도가 고급임에도 텍스트의 길이를 길게 구성하지 못한 점 등을 볼 때 이것이 학습자의 모어인 중국어의 영향인지는 단정하기 어려울 것으로 보인다. 우서양(2016)에서는 중국인 대학생 및 대학원생을 대상으로 한국어 설명문 쓰기 연구를 한 바 있는데, 이때 텍스트 구조를 ‘처음-본문’으로 선택한 학습자들이 50%가 넘었다고 한 바 있다. 그러나 설명문 작성을 위해 제시한 유도 자료가 TOPIK 53번 문제라는 점에서 이러한 특성은 중국어권의 특성이라기보다는 TOPIK 53번 구조의 영향이라 볼 수 있다. 특히 생략되는 C유형이 데이터의 ‘해석’에 관련되는 부분이기 때문에 시각적 데이터 세 개가 나타내고자 하는 바를 찾아내고 이를 논리적으로 연결하거나 배경 지식을 사용해 해석하는 행위가 저학년의 학습자에게는 어려울 것이라 추정할 수도 있다. 본고에서는 학년별 비교는 다루지 않기 때문에 C 이동 마디 유형과 관련해서는 뒤의 숙달도와 관련이 있는지 다시 한 번 확인해 보도록 하겠다.

<표 32> ABABAB 유형 데이터 설명문의 예

(A1)코로나 시기 성별 가사 노동 시간 연구

<표 1> **(A2)**남녀 각 500명을 대상으로 설문 응답자 기본 정보에 대해 조사를 실시하였다. **(B1-2)**먼저 학력은 중졸 이하는 10명을 있다. 고졸은 260명을 있다. 대학원졸은 110명을 있다. 그 다음에는 하루 평균 가사 노동 시간은 1시간 미만은 280명을 있다. 1시간~2시간은 402명을 있다. 2시간~3시간은 227명을 있다. 3시간~4시간은 74명을 있다. 4시간 이상은 17명을 있다. 마지막은 혼인 상태은 비혼(미혼, 사별, 이후 등)중에 280 명을 있다. 기혼(사실혼 포함)중에 720명을 있다.

(B4)이상의 설문 조사 결과를 통해 1시간~2시간 동안 가장 많다.

<그림 1> **(A2)**남성과 여성을 대상으로 맞벌이 상태별 하루 가사 노동 시간에 대해 조사를 실시하였다. **(B2-3)**먼저 남성은 맞벌이를 27분으로 나타냈다. 여성은 맞벌이를 129분으로 나타냈다. 그 다음에는 남성은 외벌이(남편)을 31분으로 나타냈다. 여성은 외벌이(남편)을 240분으로 나타냈다. 마지막은 남성은 외벌이(아내) 를 66분으로 나타냈다. 여성은 외벌이(아내)를 106분으로 나타냈다.

(B4)이상의 설문 조사 결과를 통해 여성의 항상 노동 시간 가장 많다 .

<그림 2> **(A2)**남성과 여성을 대상으로 하루 평균 가사 노동 시간 비교에 대해 조사를 실시하였다. **(B2-3)**먼저 여성의 하루 평균 가사 노동 시간 살펴보다 . **(B1-1)**2019년은 하루 평균 가사 노동 시간이 174분으로 나타냈다. 그 다음에는 2020년은 187분으로 나타냈다. 마지막은 180분으로 나타냈다. 그리고 남성의 하루 평균 가사 노동 시간은 2019년을 48분으로 나타냈다. 그 다음에는 2020년은61분으로 나타냈다. 마지막은 2021년은 78분으로 나타냈다.

(B4)이상의 설문 조사 결과를 통해 하루 평균 노동시간 남성보다 여성이 더 길다.

[48-중급-중국어-경영1-여]

중국어권 학습자의 또 다른 특징적인 이동 마디 순서 유형 사용으로는 ABABAB 유형이 있다. 이 유형의 데이터 설명문 구조는 각 표와 그래프에 따라 ‘언급-데이터 나열’하는 구조를 반복하고 있다. 표나 그래프에서 특징적이거나 주제적이라 생각하는 수치를 마지막에 뽑아 제시하는

것으로써 논평이나 해석을 같음한다. 각 표와 그래프의 내용을 충실히 나열하기는 하였으나 이들을 짜임새 있게 연결하거나 하지는 않는 점으로 보아 이 유형을 선택하는 학습자들 역시도 데이터 설명문을 구성할 때 어떠한 구조가 적합한지에 대해 알지 못하거나 고민하고 있음을 볼 수 있다.

<표 33> B 유형 데이터 설명문의 예

<표 1 삽입>

(B3)위의<표 1 >은 2022년에 통계청에서 남성과 여성이 각 500명을 대상으로 설문 응답자에 대한 설문 응답자 기본 정보이다. (B1-2)설문 정보에 따르면 학력에 대해 중학교 및 이하를 졸업한 남녀가 10명이고 고등학교를 졸업한 260명이 있으며 , 대학교를 졸업한 남녀가 620명으로 가장 많고 대학원을 졸업한 남녀가 110명이다. (B1-2)한편으로 하루평균 가사 노동 시간의 경우는 1시간 미만의 남녀가 280명, 1시간부터 2시간까지 가사를 하는 사람은 402명, 2시간부터 3시간까지 하는 사람이 227명, 3시간부터 4시간까지 74명이고 4시간 이상의 남 녀 단 17명이 있다. (B1-2)그리고 혼인 상태에 대해 비혼 남녀가 280명이 있고 기혼 남녀가 720명이 있는 정보이다. (B4)즉, 요즘 은 대다수 사람은 대학에서 졸업한 경우가 많다. 항상 가사 노동 시간 은 1시부터 2시까지 이다. 그리고 기혼한 사람이 많다는 것을 알 수 있다.

<그림1 삽입>

(B3)<그림 1 >는 2022년에 통계청에서 맞벌이 상태별 하루 가사 노동 시간에 대한 설문 조사이다. (B2-3)이 조사 결과에 따르면 맞벌이 의 경우 남성은 하루에 27분 정도 가사 노동을 했지만 여성은 129 분 정도 했다. 남편 외벌이의 경우 남성은 31분만 가사 노동을 했지만 여성의 가사 노동 시간이 240분으로 조사되었다. 반면에, 아내 외벌이의 경우 남성은 66분이고 여성은 106분이다. (B4)즉, 요즘 은 맞벌이든지 외벌이든지 여성은 남성보다 하루에 가사 노동 시간이 더 긴 다는 것이다.

<그림2 삽입>

(B3)<그림 2>는 2022년에 통계청에서 남성과 여성의 하루 평균 가사 노동 시간에 대한 설문 조사이다. (B1-1)조사 결과에 따르면 남성은 2019년의 48분에서 2020년의 61분, 2021년의 78분으로 꾸준히 증가한 것으로 나타났다. 여성은 2019년의 174분에서 2020년의 187분으로 증가했지만 2020년의 187분에서 2021년의 180분으로 감소한 것으로 나타났다. (B4)즉, 남성의 하루 평균 가사 노동 시간이 꾸준히 증가했지만 여성의 하루 평균 가사 노동 시간이 더 긴 다는 것이다.

[54-중급-중국어-경영2-여]

마지막으로 B유형은 제시된 표와 그래프에 시선을 이동하며 이들 시각적 데이터의 데이터의 수치를 나열하고 그 수치의 의미를 제시하는 구조로 이루어져 있다. 이 역시도 표와 그래프마다 문단을 분리하고 있으며, 각 문단은 별개의 데이터 설명문으로 구성된다. 즉, 하나의 글로 통합되지 못하고 있음을 보여준다. 내용을 하나로 잇지 못하는 분절적인 글 구성은 구조에 대한 몰이해는 확인할 수 있으나, 이것이 중국어 글쓰기 구조의 영향인지는 알기 어렵다.

오히려 유민애(2017:166)에서는 중국에서 모든 국민이 지속적으로 평가받는 ‘술직(述職)’이라는 텍스트의 구조가 ‘자기 결심, 희망’으로 끝맺는 C3과 같은 구조를 지니고 있다고 하였기 때문에 ‘C3. 미래 방향 제시’는 중국어권 학습자들이 주로 생성할 법한 텍스트 마무리 단계라 볼 수 있다. 따라서 중국어권 학습자들이 나타내는 이러한 데이터 설명문 이동 마디 순서 유형은 모어의 전이보다는 다른 변인의 영향일 가능성이 있다고 본다.

한편 베트남어권 학습자가 선택한 상위 5개 이동 마디 순서 유형은 모두 A로 시작해 C로 끝나는 유형이다. 그 사이에 어떤 이동 마디가 반복되든 기본적인 ABC 유형을 지키고 있는 것이다. ‘ABC’ 및 ‘ABABC’ 이동 마디 유형은 앞의 한국인 대학생과의 비교에서 이미 예시한 바 있기 때문에 아래에서는 ABABABC와 ABABCBC 유형의 예를 살펴보고 하겠다.

<표 34> ABABABC 유형 데이터 설명문의 예

(A1)코로나 시기 성별 가사 노동 시간 연구

(A2)가 연구는 코로나 시기 성별 가사 노동 시간을 설문하며 분석한다. 이 연구는 맞벌이 상태별 하루 가사 노동 시간, 남성과 여성의 하루 평균 가사 노동 시간 비교, 그리고 설문 응답자 기본 정보를 언급한다.

(A1)1. 맞벌이 상태별 하루 가사 노동 시간:

그림1

(B3)그래프를 살펴보면 **(B4)**맞벌이하는데도 여성들이 늘 남성들보다 하루 가사 노동 시간을 더 많이 한다는 것을 알 수 있다.**(B2-3)**여성이 집안일을 하는 시간은 남편이 외부출근자일 때 가장 많으며, 부부가 함께 일할 때, 아내가 외부 출근할 때 순으로 줄어든다.

(A1)2. 남성과 여성의 하루 평균 가사 노동 시간 비교:

그림2

(B3)그림 2를 통해 **(B4)**여성의 평균 가사 노동 시간은 여전히 남성보다 많다는 것을 알 수 있다. **(B2-5)**2019년 여성의 가사노동 시간은 남성의 3배 이상이다. **(B3)**그렇지만2019년부터 2021년까지 여성에 비해 남성 평균 가사 노동 시간은 꾸준하고 더 많이 늘어난 것으로 나타났다. **(B2-2)**이에 따라 2021년까지 남성과 여성의 평균 가사 노동 시간에 대해 격차는 약 2배 수준으로 줄었다.

(A1)3. 설문 응답자 기본 정보

표1

(A2)남성과 여성의 근무시간을 더 잘 이해하기 위해, 이 연구는 남자 500명과 여자 500명을 포함한 1,000명을 대상으로 설문조사를 실시하였다.**(B1-2)**이 조사 설문의 내용은 학력, 하루 평균 가사 노동 시간과 혼인 상태로 나뉠 수 있다. 실현한 조사 설문 1000명 중에서 비혼한 280명과 기혼한 720명이 있다 . **(B2-4)**대졸인 응답자 수가 620명으로 1위에 차지한다. 중졸 이하인 응답자가 10명으로 가장 낮게 나타난다. **(B2-1)**응답한 하루 평균 가사 노동 시간에 대해 1시간 미만이 280명, 1~2시간이 402명, 2~3시간이 227명, 3~4시간이 74명, 4시간 이상이 17명을 차지하고 있다.

(B4)이 연구를 통해 우리는 코로나 시기에 여성이 남성보다도 더 많은

노동 시간을 보내고 있다는 결론을 내릴 수 있다. 그렇지만, 남성의 평균 노동 시간에 대해 점점 늘었고 남성과 여성 사이에 격차를 감소된다. **(C1)**이는 사회에서 남성과 여성의 평등권이 더욱 두드러진다는 것을 증명한다.

[119-중급-베트남어-경제3-여]

ABABABC 유형은 앞의 중국어권 학습자가 주로 선택하는 ABABAB 유형과 크게 다르지 않은 구성을 보인다. 표와 그래프 각각을 분절하여 제시하는 것은 같지만, 소제목으로 구획을 지어 각 시각적 데이터의 데이터 분석이 하나의 글로 기능함을 보여준다. 또한 마지막에 한 문장이거나 수치의 의미가 아닌 그를 통해 해석한 전체 현상을 설명한다. 물론 모든 ABABABC 유형이 이와 같이 소제목으로 구분되는 것은 아니나 ‘이번에는, 마지막으로([82-고급-베트남어-경영2-여])’와 같이 각 시각적 데이터의 데이터 분석 내용을 연결하는 기능을 하는 장치를 삽입하고 있다는 데서 유사하게 분리의 기능을 수행하는 것을 볼 수 있다.

<표 35> ABABCBC 유형 데이터 설명문의 예

(A2)2022년 통계청 자료를 바탕으로 코로나 기간 동안 성별에 따른 가사노동 시간을 조사한 결과이다. 첫 번째는 다음과 같은 표에 따라 남성과 여성을 포함한 조사 참가자의 정보 데이터를 분석하는 것이다.

표1

(B3)위의 표에 따르면 **(B1-2)**설문조사 참여자는 1000명으로 남자 500명, 여자 500명으로 중학교와 중졸 이하, 고졸, 대졸, 대학원졸의 4개 학력에 속한다. **(B2-1)**하루에 가사 노동 시간을 조사한 결과 1시간에서 2시간 사이의 집안일을 하는 사람이 402명으로 가장 많은 비중을 차지했고 1시간 미만 집안일을 하는 사람이 280명 그 뒤를 이었다. 그런 다음 2시 ~ 3시, 3시 ~ 4시, 4시 위가 마지막 자리를 차지한다. **(B2-2)**기혼자가 미혼자보다 340명 더 많다.

(A2)둘째, 다음 그림을 바탕으로 남녀의 가사 노동 시간을 혼인상태에 따라 비교한다.

그림1

(B3)그림1을 보면 **(B2-3)**맞벌이 그룹의 경우에 여성의 가사 노동 시간

이 남성보다 여성이 129분, 남성이 27분으로 훨씬 더 많은 것을 알 수 있다. (B4)이것은 남편과 아내가 함께 돈을 벌기 위해 일하기 때문에 사회의 불평등을 보여주지만 여성이 집에 돌아와도 여전히 더 많은 가사일을 해야 한다. (B2-3)외벌이(남편) 그룹에서는 남편만 돈 벌기 위해 일할 때 여자는 풀타임으로 집안일을 해야 한다는 것을 알 수 있다. 여성이 집안일을 해야 하는 시간은 240분, 남성은 31분이다. 마지막으로 외벌이(아내) 그룹에서 즉 돈을 벌기 위해 일을 하고 남자보다 40분 더 가사 노동을 해야 하는 한국 사회의 현실을 본다. (C1)나는 모든 사회에서 항상 집안일에 특별한 가족을 돌보는 부담을 가지고 있기 때문에 여성에게 불공평하다고 생각한다.

그림2

(B3)그림 2에 따르면 좋은 신호를 볼 수 있다. (B1-1)2019년부터 2021년까지 3년 동안 남성이 집안일을 하는 시간이 늘었다. (B2-3)여성의 가사 노동 시간은 시간이 지남에 따라 꾸준히 증가했지만 남성의 가사 노동 시간은 개선되었다. (B1-1)구체적으로 2020년에는 남녀의 가사 노동 시간이 13분 증가했고 2021년에는 여성이 늘어나고 남성이 감소하게 된다. (B4)그러나 전체 시간을 보면 더 고르게 분포되어 있다. (C1)이는 한국 사회가 점차 변화하고 페미니즘이 인식되고 있음을 보여준다.

[117-중급-베트남어-경영3-여]

마지막으로 ABABCBC 유형은 첫째와 둘째 시각적 데이터에서는 ‘첫 번째, 둘째’와 같이 연결의 응집성을 보여주는 장치를 사용하여 텍스트의 단일성을 보여주는 것 같으나 갑자기 두 번째 자료를 논하는 문단 끝에서 자신의 생각과 해석을 삽입함으로써 세 번째 자료와의 단절이 나타났다. 그리고 세 번째 자료의 데이터 나열을 마친 후 마지막으로 해석을 한 문장으로 제시한다. 이 유형의 다른 텍스트들도 이와 마찬가지로 시각적 데이터 두 개를 논한 뒤 해석을 부연하고, 마지막 시각적 데이터의 수치를 나열한 후 데이터가 나타내는 현상에 대해 해석하거나 논평한다. 이는 표의 경우에는 해석할 만한 의미가 없다고 생각해 그래프들에 대해서만 해석을 붙였으나 그림 1과 그림 2 사이의 응집성이 떨어지는 것을 볼 수 있다.

이들 베트남어권 학습자가 자주 사용하는 이동 마디 순서 유형은 ‘A-()-C’로 구성된 중국어권 학습자의 이동 마디 순서 유형과 크게 다른 특성을 확인할 수는 없었다. 위의 ABABCBC 유형에서 첫 번째 등장하는 C의 위치가 적절하지 않다고 여길 수는 있다. 그러나 이 역시도 데이터의 적절한 나열과 해석의 위치에 대한 미숙성으로는 볼 수 있을지 몰라도 베트남어 구조의 특징으로는 볼 수 없을 듯하다. 정리하자면 두 언어권의 이동 마디 순서 유형의 차이는 ‘C-누락형’에서 분명히 존재하지만, 이것은 언어권의 영향이라기보다는 텍스트 구조 인식의 미숙성 때문으로 보이는 경우가 다수였다.

2) 한국어 숙달도별 이동 마디 순서 유형과 특성

한국어 숙달도별 이동 마디 순서 유형에서는 학습자 언어권을 통제하지 않고 모든 외국인 학습자들의 한국어 숙달도에 따라 어떤 이동 마디 순서 유형을 선택했는지 알아보도록 한다. 따라서 위의 언어권별 숙달도와 합계가 같지 않다. 학습자의 언어 숙달도별 이동 마디 순서 유형 조사에서는 아래 <표 36>과 같이 중급 학습자와 고급 학습자 사이의 유형별 빈도 차이를 알아보도록 한다.

<표 36> 한국어 숙달도별 이동 마디 순서 빈도

| 연번 | 이동 마디 순서 유형 | 중급 | 고급 | 연번 | 이동 마디 순서 유형 | 중급 | 고급 |
|----|--------------|----|----|----|---------------|----|----|
| 1 | AB | 19 | 6 | 27 | ABCABCBCB | 0 | 0 |
| 2 | ABAB | 3 | 1 | 28 | ABCABCBCBABC | 0 | 0 |
| 3 | ABABAB | 3 | 1 | 29 | ABCAC | 0 | 0 |
| 4 | ABABABAB | 0 | 1 | 30 | ABCB | 1 | 1 |
| 5 | ABABABABABC | 1 | 3 | 31 | ABCBABC | 0 | 0 |
| 6 | ABABABCAC | 1 | 0 | 32 | ABCBABCBCBABC | 0 | 1 |
| 7 | ABABABABC | 0 | 1 | 33 | ABCBC | 3 | 5 |
| 8 | ABABABABCBCB | 0 | 1 | 34 | ABCBCABC | 1 | 1 |

| | | | | | | | |
|----|---------------|----|----|----|------------------|----|----|
| 9 | ABABABAC | 0 | 1 | 35 | ABCBCBAC | 0 | 0 |
| 10 | ABABABACBC | 0 | 1 | 36 | ABCBCBC | 1 | 0 |
| 11 | ABABABC | 3 | 2 | 37 | ABCBCBCBCBC | 0 | 0 |
| 12 | ABABABCABC | 1 | 0 | 38 | ABCBCBCBCABCBCAC | 0 | 0 |
| 13 | ABABABCBCBC | 0 | 1 | 39 | ABCBCBCBCBCBCAC | 1 | 0 |
| 14 | ABABABCBC | 0 | 1 | 40 | ACBABABABC | 0 | 1 |
| 15 | ABABABCBCBCBC | 0 | 0 | 41 | B | 3 | 1 |
| 16 | ABABACBC | 0 | 0 | 42 | BAB | 2 | 1 |
| 17 | ABABC | 1 | 5 | 43 | BABAB | 0 | 0 |
| 18 | ABABCABC | 0 | 0 | 44 | BABABABC | 1 | 0 |
| 19 | ABABCBCABCABC | 0 | 0 | 45 | BABABABCBC | 0 | 1 |
| 20 | ABABCBCA | 0 | 0 | 46 | BABCB | 1 | 0 |
| 21 | ABABCBC | 4 | 3 | 47 | BACBABA | 1 | 0 |
| 22 | ABABCBCB | 0 | 1 | 48 | BC | 0 | 0 |
| 23 | ABABCBCBC | 0 | 0 | 49 | BCB | 0 | 0 |
| 24 | ABABCBCBCBC | 0 | 1 | 50 | BCBC | 1 | 0 |
| 25 | ABC | 18 | 12 | 51 | BCBCBC | 0 | 0 |
| 26 | ABCABABCABC | 0 | 0 | 52 | BCBAB | 0 | 1 |
| 계 | | | | | | 70 | 55 |

여기에서는 두 집단 학습자 모두가 선택하지 않은 이동 마디 순서 유형이 18개 있어 52개 유형 중 34개만이 사용된다. 아래 <표 37>에서 볼 수 있듯, 중급 학습자들은 이 34개 이동 마디 순서 유형 중 21개(전체 유형의 61%)만을 사용하였고, 고급 학습자들의 경우에는 26개(전체 유형의 76%)만을 사용하였다. 고급 학습자들의 이동 마디 순서 유형 사용 개수가 중급 학습자들보다 높은 것을 보아 고급 학습자가 데이터 설명문의 구조를 더 자유롭게 인식하고 있음을 알 수 있다.

<표 37> 중급/고급 학습자의 34개 이동 마디 순서 유형 중 사용 개수와 비율

| | 사용 유형 개수 | 전체 유형 중 사용 유형 비율 |
|----|----------|------------------|
| 중급 | 21 | 61% |
| 고급 | 26 | 76% |

이번에도 이들 34개 유형 중 상위 5개에 해당하는 이동 마디 유형을 추출하려 하였으나, 중급의 경우 4위부터는 이동 마디 유형의 선택 빈도가 3회로 동일한 것이 5편이나 있어 3위까지만 제시하고, 고급의 경우에도 동일한 이유로 4위까지만 제시하도록 한다. 그 결과 중급 학습자의 경우는 상위 3개 이동 마디 순서 유형이 전체 선택의 58.5%를 차지할 정도로 선호 유형의 편중이 컸으며, 고급 학습자의 경우도 상위 4개 유형이 전체 선택 유형의 50.9%를 차지하여 나머지 22개 이동 마디 순서 유형 보다 압도적으로 선호됨을 알 수 있다.

<표 38> 중급/고급 학습자가 선택한 상위 3~4개 이동 마디 순서 유형

| 중급 학습자 | | | | 고급 학습자 | | | |
|--------|-------------|----|-------|--------|-------------|----|-------|
| | 이동 마디 순서 유형 | 편수 | % | | 이동 마디 순서 유형 | 편수 | % |
| 1 | AB | 19 | 27.1% | 1 | ABC | 12 | 21.8% |
| 2 | ABC | 18 | 25.7% | 2 | AB | 6 | 10.9% |
| 3 | ABABCBC | 4 | 5.7% | 3 | ABABC | 5 | 9.1% |
| | | 41 | 58.5% | 4 | ABCBC | 5 | 9.1% |
| | | | | | | 28 | 50.9% |

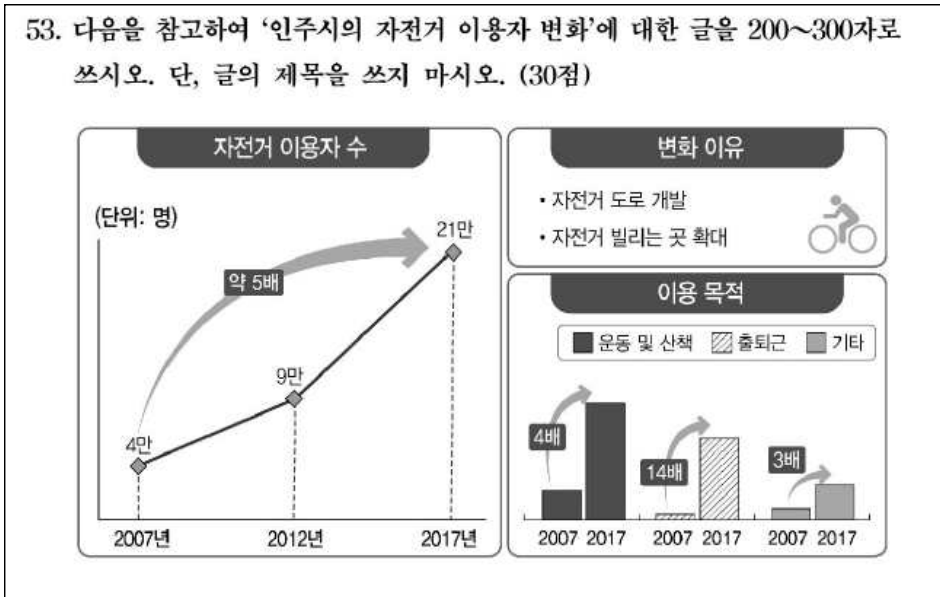
이번에도 ABC 유형은 중급에서는 2위, 고급에서는 1위로 모든 집단에서 상위 선택 유형임을 알 수 있다. 따라서 ABC 유형은 모든 집단에서 선호되는 보편적인 이동 마디 순서 유형이라 할 수 있다.

AB 유형은 중급에서는 1위, 고급에서는 2위로 이 역시도 외국인 대학생들에게는 선호되는 구조라 할 수 있다. 언어권별로는 중국어권에서 다수 나타나는 경향을 보였으나 분석의 예에서 확인할 수 있듯, 이는 모어

의 구조에서 전이된 것이라기보다는 글의 짜임과 흐름이라는 보편적인 글쓰기 구조에 대한 훈련이 되지 못한 것으로 보인다. 그리고 이러한 보편적 글쓰기를 저해하는 다른 요인이 개입한 것으로 보인다. 이를 좀 더 구체적으로 보자면, 두 가지 분석을 할 수 있다. 첫째는 학습자가 표와 그래프 정보를 어떻게 적절하게 제시하고 엮어야 하는지에 대한 적절한 판단을 하지 못하여 그러할 수도 있다는 것이다. AB 유형을 선택한 외국인 학습자 25명 중 14명이 1학년, 3명이 2학년이라는 점에서 일반적인 학술적 글쓰기 훈련을 받지 못해 서론-본론-결론의 구조가 익숙하지 못했음을 추론할 수 있다. 게다가 자료 수집 기간이 6월부터여서 연구 참여자들이 전기 입학한 1학년, 또는 후기 입학한 2학년인 경우에는 한국 대학 경험이 1년 미만으로, 일반적인 한국어의 학술적 글쓰기 구조를 훈련하기 어려웠을 시점이기도 하다. 또한 AB 유형을 선택한 중급 학습자가 19명, 고급 학습자가 6명이라는 점에서 학술적 글쓰기 구조를 훈련하지 못하는 데에는 중급이라는 한국어 숙달도가 영향을 미쳤을 가능성도 있다.

둘째는 연구 참여자들이 TOPIK 시험 53번 답안 작성 유형의 영향을 받아 그럴 수도 있다는 추정을 하게 한다. 한국의 대학에 입학하려는 인문사회계 외국인 유학생들에게는 일반적으로 TOPIK 3급 이상의 점수가 요구되기 때문에(교육부, 2019), TOPIK II 시험 유형에 대한 노출과 학습이 반드시 있을 것으로 생각된다. 이 TOPIK II 시험에는 다음과 같은 데이터 설명문 유형의 글쓰기 평가 문제가 있다(토픽 공식 홈페이지, TOPIK 제60회, 2018. 10. 시행).

53. 다음을 참고하여 ‘인주시의 자전거 이용자 변화’에 대한 글을 200~300자로 쓰시오. 단, 글의 제목을 쓰지 마시오. (30점)



<그림 5> TOPIK II 제60회 53번 문제

위 53번 문제의 지시문에서 알 수 있듯, ‘글의 제목을 쓰지 않고’, ‘200~300자의 짧은 분량’ 내에 데이터의 ‘내용’을 모두 전달해야 한다. 따라서 다른 부차적인, 그러나 일반적인 학술적 텍스트였다라면 중요했을 독자 견인 표현이나 시각적 데이터 언급 표현은 요구되지 않으며, 그래프 두 개를 문단까지 나누어 가며 표현하기에는 어려운 조건이기도 하다. 이에 TOPIK 홈페이지에서 제공하는 제60회 ‘정답 및 채점 기준표’의 모범 답안 구조는 다음과 같이 AB 구조이다.

인주시의 자전거 이용자 변화를 살펴보면, 자전거 이용자 수는 2007년 4만 명에서 2012년 9만 명, 2017년에는 21만 명으로 지난 10년 동안 약 5배가 증가하였다. 특히 2012년부터 2017년까지 자전거 이용자 수가 급격히 증가한 것으로 나타났다. 이러한 변화의 이유로는 자전거 도로의 개발과 자전거를 빌리는 곳이 확대된 것을 들 수 있다. 자전거 이용의 목적을 보면 10년 동안 운동 및 산책은 4배, 출퇴근은 14배, 기타는 3배 늘어난 것으로 나타났는데 이용 목적 중 출퇴근이 가장 높은 증가율을 보였다.

따라서 한국어 숙달도가 낮은 외국인 대학생들은 데이터 설명문이 장르 고정성과 구조가 확고하지 않기 때문에 이를 작성할 시 구조의 정형성을 판단하기 어렵고, 한국어 학습 중 접했던 데이터 설명문에 가장 유사한 장르가 TOPIKⅡ 53번 문제였기 때문에 이를 일반적인 데이터 설명문의 구조로 인식하고 있을 가능성이 높다.

그 외 ABABCBC, ABABC, ABCBC와 같은 ‘A-()-C’ 이동 마디 순서 유형은 앞에서 예시한 바와 같이 ABC 유형을 원형으로 세 시각적 데이터 중 어떤 것들을 유사한 것으로 간주하고 묶어 제시하는가의 차이로 볼 수 있다. 그 과정에서 수치의 해석(C)이 어떤 묶음에 해당하는가에 따라 세부 이동 마디 순서가 달라지는 것으로 확인된다. 이러한 구성은 한국인 대학생, 중국어권 및 베트남어권 학습자, 중급 및 고급 학습자 모두에게서 나타나는 일상적인 이동 마디 순서라 할 수 있다.

정리하자면, ABC 유형이 한국인을 비롯한 모든 학습자 집단에서 데이터 설명문을 나타내는 가장 선호도 높은 이동 마디 순서임을 알 수 있다. 또한 ‘A-()-C’ 구조는 기본적으로 ABC 유형의 확장형이라 세 개의 시각적 데이터를 어떻게 묶어 제시하는지에 따라 차이가 생기는 것을 알 수 있다. 이러한 ABC 확장 유형은 집단에 관계없이 보편적으로 나타나는 것을 확인하였다. 이는 시각적 데이터를 이용하여 설명하는 데이터 설명문의 구조에는 시각적 데이터 제시를 해야 한다는 제약이 있어 ABC류 확장 유형을 언어권이나 숙달도에 관계없이 학습자들이 보편적으로 사용함을 알 수 있다.

C가 누락된 이동 마디 순서 유형은 중국어권 학습자 집단과 중급 학습자 집단에서 높은 빈도로 선택된다. 다만 중국어의 보편적인 글쓰기 구조가 C 이동 마디를 요구한다는 점에서 이는 중국어 글쓰기 구조의 영향은 아닐 것이라 보았다. 이에 학습자들이 작성한 C-누락형 데이터 설명문을 분석한 결과 TOPIKⅡ 53번 문항의 모범 답안과 구조가 같아 학습자들의 데이터 설명문 구조가 TOPIK의 영향을 받았을 가능성이 있

다. 특히 저학년 학습자 및 한국어 숙달도가 중급인 학습자들에게 주로 C-누락형이 나타나 한국어의 데이터 설명문 구조에 대한 확신이 없는 학습자들이 이를 주로 선택함을 알 수 있다. 즉, 전공 과제로서 작성해야 하는 한국어 데이터 설명문 구조에 대한 지식이 없을 때, 이미 그 구조를 경험해 본 텍스트 중 가장 데이터 설명문에 가까운 유형인 TOPIKⅡ 53번 구조를 인출했을 가능성이 있다.

1.2. 이동 마디 구성 요소별 특징

이제까지 다룬 이동 마디 순서 유형은 한 편의 데이터 설명문 전체의 흐름을 보기 위해 이동 마디의 하위 단계를 제외하고 상위 이동 마디만을 대상으로 하였다. 그러나 이동 마디 세부 순서를 구성하는 단계를 얼마나 사용했는지도 학습자의 상세한 데이터 설명문 구조를 파악하기 위해서 필요하다. 이에 본 절에서는 각 이동 마디를 구성하는 하위 단계까지 확장하여 이들의 개수가 학습자 변인별로 어떠한 공통점과 차이점이 있는지 알아보며, 학습자 작문에서 이들 특성의 의미가 무엇인지 해석해보고자 한다.

우선 학습자군별 데이터 설명문에서 이동 마디 단계의 빈도를 세기 전에 이동 마디 단계가 몇 편의 학습자 작문에 포함되었는지 분포를 확인해 보아야 한다. 한 편의 학습자 데이터 설명문에서 과도하게 많은 특정 이동 마디 단계가 출현하는지를 알아보아야 하기 위한 것도 있고, 고정적으로 사용되는 이동 마디 단계가 있는지 알아볼 수 있는 방법이기도 하기 때문이다. Kanoksilapatham(2005:272)에서는 전문 필자가 작성한 생명화학 학회지 논문의 각 장에서 나타나는 특정 이동 마디 단계의 빈도가 60%가 넘는다면 관습적인 것으로, 그렇지 않으면 선택적인 것으로 본다. 예를 들어, 10개 학술 논문을 대상으로 서론에 A1 단계가 있는 논문의 개수가 6개 이상이라면 이 A1 단계는 관습적이다. 이동 마디 단계의 본고에서도 이와 같은 Kanoksilapatham(2005)의 연구 방법을 따라 학습자 데이터 설명문에서 이동 마디 단계의 분포를 계산하도록 한다.

다만 본고는 전문 필자가 작성한 텍스트를 대상으로 하는 것이 아니기 때문에 ‘관습적’ 선택이라는 개념보다는 학습자 집단에서 선호하는 ‘고정적’ 이동 마디 단계로 파악하도록 한다.

이렇듯 이동 마디 단계의 분포를 확인한 후에는 학습자군별로 수집한 각 단계의 빈도수 평균을 비교한다. 본고의 데이터 설명문은 Swales(1990)나 Kanoksilapatham(2005)에서 다루는 학술 논문처럼 장이 구분되어 한 장이 하나의 이동 마디 순서 유형으로만 고정돼 있는 것이 아니고, 한 데이터 설명문 내에서 이동 마디와 단계가 다회 반복되어 그 빈도를 알아볼 필요도 있기 때문이다. 이들 대상은 앞 장에서와 마찬가지로 한국인 대학생 50편, 중국어권 학습자 59편, 베트남어권 학습자 50편, 중급 학습자 70편, 고급 학습자 55편에서 추출한 이동 마디 단계 수를 대상으로 한다. 이렇게 수집된 이동 마디 단계의 집단 간 빈도 차이는 Microsoft Excel을 통해 빈도의 평균을 t-test로 검정하였다.

(1) 이동 마디 A: 주제 및 방향 요약

1) 외국인/한국인 대학생의 이동 마디 단계 분포 및 빈도

본고에서는 우선 Kanoksilapatham(2005)의 방법을 따라 ‘A. 주제 및 방향 요약’ 이동 마디 단계가 외국인 대학생과 한국인 대학생의 각 데이터 설명문에 나타나는 정도가 어떠한지 그 범위를 알아볼 것이다. 외국인/한국인 대학생 데이터 설명문 전체에서 이동 마디 단계가 나타나는 편수를 세어 전체 편수 대비 비율을 알아본다. 데이터 설명문 한 편에 해당 이동 마디 단계가 하나라도 출현하면 1편으로 친다. 또한 이동 마디 단계는 선택을 안 할 수도 있고, 중복으로 출현할 수도 있으므로 각 이동 마디 단계가 나타나는 데이터 설명문 편수의 합이 전체 데이터 설명문 편수와 같지는 않다.

<표 39> 외국인/한국인 대학생의 A 이동 마디 분포

| 이동 마디와 단계 | | 외국인 | 한국인 |
|---------------|--------------------|-------------------|----------------|
| A | 주제 및 방향 요약 | | |
| A1 | (소)제목을 활용한 주제 제시 | 28(22%) | 16(32%) |
| A2 | 연구의 주제와 내용 및 방법 설명 | 104(83.2%) | 42(84%) |
| A3 | 데이터 현상에 대한 배경 설명 | 18(14.4%) | 24(48%) |
| 전체 데이터 설명문 편수 | | 125 | 50 |

<표 39>와 같이 A 이동 마디 단계 분포 확인 결과 외국인과 한국인 대학생들의 데이터 설명문 전체 편수 중 60%이상 분포하는 이동 마디 단계는 A2 하나만 존재하였다. 이는 A2 단계가 대부분의 경우 본격적인 글의 시작을 나타내는 단계이기 때문이다. 따라서 데이터 설명문의 A2 이동 마디 단계는 두 집단에서 고정적으로 사용하는 것을 알 수 있다. 다만 A1과 A3의 경우에는 외국인 대학생이든 한국인 대학생이든 수업 과제로 제출해야 하는 데이터 설명문 A 이동 마디 단계의 고정적 구조에 대한 인식이 집단 내에서 합의된 것이 아님을 보여준다. 이 역시 데이터 설명문의 상위 장르에 따른 비고정적 속성을 반영한다고 볼 수 있다.

그러나 이러한 분포만으로는 집단 간 선호되는 이동 마디 단계를 확인하기 어려워 이동 마디 단계의 전체 개수를 세어 집단 간 빈도의 평균 차이가 유의한지 알아보도록 하겠다. 이 과정에서는 대학생들의 데이터 설명문에서 추출한 이동 마디 단계를 학습자 변인별로 나누어 학습자들이 선택한 이동 마디 단계의 빈도 차이를 알아보도록 하겠다. 이를 위해 Microsoft Excel을 통해 t-test를 실시한다. 다음 <표 40>은 A 이동 마디에 속하는 단계들의 선택 빈도 평균을 t-test를 통해 차이가 유의한지 검정한 것이다.

<표 40> 한국인/외국인 대학생의 A 이동 마디 빈도 평균 t-test 결과

| | 빈도 | | 평균 | | 표준편차 | | t | p |
|----|-----|-----|-------------|------|------|------|-------|------|
| | 한국인 | 외국인 | 한국인 | 외국인 | 한국인 | 외국인 | | |
| A1 | 38 | 52 | 0.76 | 0.41 | 1.64 | 1.05 | 1.64 | 0.1 |
| A2 | 99 | 238 | 1.98 | 1.90 | 1.83 | 1.53 | 0.28 | 0.78 |
| A3 | 62 | 32 | 1.24 | 0.25 | 1.67 | 0.70 | 5.46* | 0.00 |

*p<0.05

A 이동 마디에 속하는 단계들의 한국인/외국인 대학생 집단 간 차이는 A3을 제외하고는 유의하지 않았다. 외국인 대학생과 한국인 대학생의 A1 빈도수 평균은 한국인 대학생 0.76, 외국인 대학생이 0.41로 한국인 대학생이 더 높은 편이다. 그러나 이들 두 집단의 평균 차이는 95%의 유의 기준에서 유의하지 않았다. 이는 텍스트 전체를 아우르는 하나의 제목뿐 아니라 ‘1.서론’처럼 소제목의 개수도 모두 센 것이어서 외국인 대학생들의 수가 더 많은 만큼 평균의 차이가 얼마 없었을 것으로 생각된다.

‘연구의 주제와 내용 및 방법 설명(A2)’은 텍스트 도입으로서의 역할을 하기 때문에 두 집단에서 모두 자주 사용한다. 또한 이들 이동 마디 단계는 과제로 제시한 세 시각적 데이터를 따로 분리해서 하나씩 제시하는 구성이라면 각 시각적 데이터 도입에서 반복해서 사용하기 때문에 두 집단 모두에서 선택 빈도가 비교적 높은 편이다. 이렇듯 두 집단 모두 다수 사용하기 때문에 A2의 선택 빈도는 한국인 대학생과 외국인 대학생의 차이가 유의하지 않다. 달리 말하면, 이 단계는 두 집단에서 보편적으로 다수 사용된다고 볼 수 있다.

‘데이터 현상에 대한 배경 설명(A3)’은 한국인 대학생의 선택 빈도 평균이 외국인 대학생에 비해 그 차이가 유의하게 높았다. 한국인 대학생의 경우에는 아래 ㄱ)과 같이 데이터 설명문 도입에 데이터가 나타내는 현상을 배경으로 제시하는 것을 볼 수 있다. 이들이 기술하는 데이터 현상은 주로 코로나 시기의 특성과 전통적인 가사 노동의 여성 편향성이

있다.

ㄱ) 코로나19 바이러스는 2019년 12월부터 전 세계를 강타하였다. 강한 전파와 강력으로 인해 대면으로 진행되던 모든 활동들은 비대면으로 전환되기 시작하였고, 마스크를 착용하지 않으면 실내 생활을 아예 하지 못할 정도로 거리두기 수칙이 강화되었다. 밖에서 생활하는 시간보다 집 안에서 생활하는 시간이 늘어감에 따라, 남성과 여성 모두의 하루 평균 가사 노동 시간이 증가하기 시작하였다. ... (후략)

[35-한-경영4-여]

이는 데이터 설명문의 구성 목적을 주어진 데이터의 수치 전달에 두는 것과 그 데이터가 의미하는 현상이 무엇인지에 대해 논하려는 데에 두는 것 사이의 차이로 볼 수 있다. 이러한 차이는 학술적 데이터 설명문에 대한 경험의 차이에서 기반한 것이라 본다.

2) 학습자 언어권별 이동 마디 단계 분포 및 빈도

이번에는 외국인 대학생 중 중국어권 및 베트남어권 학습자들이 어느 A 이동 마디 단계를 고정적으로 사용하고 있는지를 그 분포를 통해 알아보도록 하겠다.

<표 41> 중국어권/베트남어권 학습자의 A 이동 마디 분포

| | 이동 마디와 단계 | 중국어권 | 베트남어권 |
|----|--------------------|------------------|----------------|
| A | 주제 및 방향 요약 | | |
| A1 | (소)제목에 활용한 주제 제시 | 8(13.6%) | 18(36%) |
| A2 | 연구의 주제와 내용 및 방법 설명 | 54(91.5%) | 46(92%) |
| A3 | 데이터 현상에 대한 배경 설명 | 9(15.2%) | 7(14%) |
| | 전체 데이터 설명문 편수 | 59 | 50 |

중국어권 학습자와 베트남어권 학습자의 A1, A3 단계는 모두 그 분포가 60%를 넘지 못해 고정적으로 사용되는 것은 없는 것으로 나타났다.

그러나 두 학습자군 모두 A2를 90%이상 사용하여 언어에 관계없이 학습자들이 데이터 설명문 작성 시 고정적으로 사용하는 단계를 A2로 볼 수 있다는 것이다.

그렇다면 중국어권 학습자와 베트남어권 학습자의 각 이동 마디 단계의 빈도수는 어떠한 차이가 있는지 알아보기 위해 t-test를 시행하여 언어권별로 평균 차이를 검증하였다. 그 결과는 아래 <표 42>와 같다.

<표 42> 중국어권/베트남어권 학습자의 A 이동 마디 빈도 평균 t-test 결과

| | 빈도 | | 평균 | | 표준편차 | | t | p |
|----|-----|------|------|-------------|------|------|--------|------|
| | 중국어 | 베트남어 | 중국어 | 베트남어 | 중국어 | 베트남어 | | |
| A1 | 8 | 42 | 0.14 | 0.84 | 0.34 | 1.51 | -3.19* | 0.00 |
| A2 | 79 | 142 | 1.34 | 2.84 | 0.84 | 1.84 | -5.57* | 0.04 |
| A3 | 15 | 13 | 0.25 | 0.26 | 0.65 | 0.77 | -0.04 | 0.97 |

*p<0.05

제목 사용에 있어서는 중국어권보다 베트남어권 학생들의 사용 빈도 평균이 유의하게 높았다. 여기서의 ‘A1. (소)제목을 활용한 주제 제시’는 데이터 설명문 한 편에 나타난 모든 소제목을 코딩하여 그 수를 셸 것이다. 중국어권 학습자가 베트남어권 학습자보다 9명 더 많이 모집되었음을 볼 때 베트남어권 학생들은 소제목을 활용하여 데이터 설명문의 주요 분석 단위를 분절하는 것을 알 수 있다. 반면 중국어권 학생들은 단계가 8회밖에 나타나지 않아 데이터 설명문에 제목을 사용하는 경우가 거의 없는 것으로 보인다.

‘A2. 연구의 주제와 내용 및 방법 설명’ 역시 베트남어권 학습자들이 중국어권 학습자들보다 유의하게 많이 사용함을 알 수 있다. 이는 ‘(소)제목을 활용한 주제 제시’를 다수 사용하는 것과도 연관이 있는데, (소)제목을 활용한 그래프나 표를 구분해서 하나씩 다루다보면, 대체적으로 제목 밑에 그래프나 표에 어떠한 내용이고 어떻게 연구되었다는 언급이 수반되기 때문이다. 따라서 베트남어 학습자의 이와 같은 두 이동 마디 단계의 사용은 관련이 있는 것처럼 보인다.

‘A3. 데이터 현상에 대한 배경 설명’의 선택 빈도는 두 언어권 학습자의

차이가 유의하지 않았으며, 다른 단계보다 비교적 덜 사용함을 알 수 있다. 이러한 경향성은 두 집단이 언어에 관계없이 공유하는 것을 볼 수 있다. 즉, 외국인 대학생들에게서 A3 단계 사용이 저조한 이유는 모어의 영향이 아니라 볼 수 있다.

3) 학습자 숙달도별 이동 마디 단계 분포 및 빈도

이번에는 외국인 학습자의 한국어 숙달도별 집단에서 이동 마디 단계가 고정적으로 나타나는지 알아보도록 하겠다.

<표 43> 중급/고급 학습자의 A 이동 마디 분포

| 이동 마디와 단계 | | 중급 | 고급 |
|---------------|--------------------|------------------|------------------|
| A | 주제 및 방향 요약 | | |
| A1 | (소)제목에 활용한 주제 제시 | 13(18.5%) | 15(27.3%) |
| A2 | 연구의 주제와 내용 및 방법 설명 | 62(88.5%) | 51(92.7%) |
| A3 | 데이터 현상에 대한 배경 설명 | 12(17.1%) | 5(9%) |
| 전체 데이터 설명문 편수 | | 70 | 55 |

중급/고급 학습자군 모두에서는 A2가 각각 88.5%, 92.7%로 분포 비율이 60%가 넘어 고정적으로 사용하는 단계라 할 수 있다. 학습자 숙달도별 이동 마디 단계 선택 빈도 및 이들의 빈도수 평균 차이가 유의한지 알아보려 t-test를 수행한 결과는 아래 <표 44>와 같다.

<표 44> 중급/고급 학습자의 A 이동 마디 빈도 평균 t-test 결과

| | 빈도 | | 평균 | | 표준편차 | | t | p |
|----|-----|-----|------|-------------|------|------|--------|------|
| | 중급 | 고급 | 중급 | 고급 | 중급 | 고급 | | |
| A1 | 24 | 28 | 0.34 | 0.51 | 0.98 | 1.13 | -0.87 | 0.38 |
| A2 | 115 | 123 | 1.64 | 2.24 | 1.44 | 1.56 | -2.18* | 0.03 |
| A3 | 21 | 11 | 0.30 | 0.20 | 0.72 | 0.67 | 0.78 | 0.43 |

*p<0.05

그 결과 중급과 고급 학습자의 빈도 차이는 A2에서만 유의하게 나타났다. 즉, 고급 학습자의 빈도수 평균이 중급 학습자에 대해 유의하게 높았다. 이는 A2를 중급 학습자보다는 고급 학습자가 다수 사용한다는 뜻인데, 이는 'B-유형' 시작 구조가 주로 중급에서 나타남과 관련이 있어 보인다.

다만 A1과 A3는 빈도와 평균 차이가 크지 않고, 통계적으로도 유의하지 않기 때문에 두 집단에서 모두 평균적으로 유사한 빈도로 잘 사용하지 않는 것으로 보인다. 따라서 숙달도에 관계없이 A1과 A3을 적게 사용하는 경향은 외국인들이 공통으로 경험했던 학술적 데이터 설명문인 TOPIK 영향으로 보인다. 이를 전체적으로 정리하면 다음과 같다.

첫째, 한국인 대학생들의 A3 단계 사용 빈도가 외국인 대학생들보다 유의하게 높았다. 사용 분포 역시 한국인 대학생이 48% 사용하여 약 반수에 해당하는 학생들이 사용한 것을 알 수 있다. 다만 그 분포가 60% 미만이기 때문에 한국어 학습자들에게 고정적으로 나타나는 단계는 아니라 할 수 있다.

둘째, 언어권별 학습자의 A 이동 마디 단계 사용 빈도 차이는 A1과 A2에서 베트남어권 학습자가 유의하게 크다. 즉, A1과 A2를 베트남어권 학습자가 더 많이 사용한다. 사용 분포로 보았을 때 A1은 분포가 60% 미만이어서 고정적으로 나타나는 단계는 아니고, 일부 학습자들만이 2~3개씩 사용하는 것으로 보인다. A2는 중국어권과 베트남어권 학습자 모두가 90% 이상 사용하여 고정적으로 나타나는 단계라 할 수 있다.

셋째, A2는 중급과 고급 학습자 60%이상의 분포를 보여 두 집단에서 모두 고정적으로 나타난다. 그러나 고급 학습자의 A2 단계 사용 빈도가 중급 학습자보다 유의하게 높았다.

(2) 이동 마디 B: 수치/결과 나열

1) 외국인/한국인 대학생의 이동 마디 단계 분포 및 빈도

데이터 설명문에서 가장 중요한 위치를 차지하는 ‘수치/결과 나열’ 이동 마디는 크게 ‘B1. 구체적인 정보 추출’과 ‘B2. 의도적 그룹화’ 두 부분으로 나뉜다. 시각적 데이터를 언어적 기호로 단순히 변환하는 일반적인 나열보다는 의도에 따른 비교가 인지적으로 더 큰 노력을 요하는 것임은 자명하다. 또한 시각적 데이터와 언어적 기호의 응집성 있는 배치를 위해 시각적 데이터를 언급하여 데이터 설명문을 구성하는 B3 역시 독자를 견인하는 대인적 의미의 담화 구성 지식을 학습자가 알고 활용할 수 있는지 확인할 수 있는 단계라 생각된다. 마지막으로 B4는 제시된 수치가 어떤 의미인지 언급하는 부분이다. 이는 다음 이동 마디에 나올 저자의 전체 데이터 해석을 이어 주는 단계라 할 수 있다. 외국인 대학생과 한국인 대학생이 이러한 단계 중 무엇을 고정적으로 사용하는지 아래 <표 45>에서 분포를 확인해 보도록 하겠다.

<표 45> 외국인/한국인 대학生の B 이동 마디 분포

| | 이동 마디와 단계 | 외국인 | 한국인 |
|------|-------------------|-------------------|----------------|
| B | 수치/결과 나열 | | |
| B1 | 구체적인 정보 추출 | | |
| B1-1 | 연대기적 나열 | 17(13.6%) | 11(22%) |
| B1-2 | 시각적 데이터 순서대로 나열 | 89(71%) | 17(34%) |
| B2 | 의도적 그룹화 | | |
| B2-1 | 수치가 높은/낮은 순서대로 나열 | 72(57.6%) | 18(36%) |
| B2-2 | 수치적 비교 | 60(48%) | 33(66%) |
| B2-3 | 특성에 따른 그룹화 | 109(87.2%) | 37(74%) |
| B2-4 | 제일 큰/작은 수치에 주목 | 50(40%) | 25(50%) |
| B2-5 | 특이성/중요 지점 지적 | 30(24%) | 5(10%) |
| B3 | 시각적 데이터 언급 표지 사용 | 98(78.4%) | 35(70%) |
| B4 | 수치의 의미 제시 | 117(93.6%) | 48(99%) |
| | 전체 데이터 설명문 편수 | 125 | 50 |

외국인 대학생 집단에서는 B1-2, B2-3, B3, B4가, 한국인 대학생 집단에서는 B2-2, B2-3, B3, B4의 분포가 60%를 넘어 고정적으로 사용되는 것을 볼 수 있다. B2-3, B3, B4 이동 마디는 두 집단 모두에서 60%가 넘기 때문에 데이터 설명문 이동 마디에서 외국인과 한국인을 가리지 않고 고정적으로 사용함을 알 수 있다.

B2-3에 해당하는 문장은 시각적 데이터 내 변인이 그룹별로 나뉘어 있어 이들에서 나타나는 수치를 서술하는 경우가 많기 때문에 두 집단 모두에서 다수 나타난 것으로 보인다. 예를 들어, ‘맞벌이 가정, 남성이 외벌이인 가정, 여성이 외벌이인 가정’과 같은 그룹화를 통해 해당 그룹의 남녀 가사 노동 시간을 비교하는 문장은 시각적 데이터에서 이들 그룹을 가시적으로 확인할 수 있기 때문에 이들 그룹의 이름을 적절히 제시하는 언어적인 노력에만 집중한다면 충분히 이 이동 마디 단계를 구성할 수 있다.

B3은 Swales & Feak(1994:80)에서 데이터 설명문의 구성 요소인 ‘위치’에 해당하기 때문에 당연히 언급되어야 한다고 보았다. 실제로 많은 데이터 설명문에서 설명을 위해 대상을 지정하는 장치로서 다음과 같이 이를 활용하고 있다.

(ㄱ) <그림1>을 보면, 맞벌이와 외벌이 가구 모두에서 여성이 남성에 비해 훨씬 많은 시간을 가사 노동에 쓰고 있음을 알 수 있다.

[34-한-경제4-여]

(ㄴ) <그림2>과 같은 남성과 여성의 하루 평균 가사 노동 시간을 비교하는 결과 따르면 여성은 남성보다 가사 노동 시간이 훨씬 더 많은 것으로 나타났다.

[123-고급-중국어-경영2-여]

B4는 외국인 데이터 설명문에서는 80.8%가, 한국인 데이터 설명문에서는 99%가 선택해 데이터 설명문의 마무리가 단순히 수치를 제시하는 것에서 끝나는 것이 아니라 그에 수치에 대한 의미를 정리해 재서술하는 과정을 거침을 알 수 있다. 외국인 학습자들이 C 이동 마디 단계를 자주

생략하는 것을 생각해 보았을 때, 상식이나 일반적 지식 등을 활용해 더 확장하는 것까지는 어렵게 생각하나, 주어진 수치의 의미를 숫자를 빼고 언어화하는 것은 대다수의 학습자들이 채택하는 고정성이 있음을 확인할 수 있다. 이러한 고정성은 어휘문법 사용에서도 드러나는 것처럼 보이는 데, 이는 (ㄷ)과 (ㄹ)의 예를 볼 때 더욱 확실히 알 수 있다.⁴⁰⁾

(ㄷ) 이는 맞벌이 유무에 상관없이 여성이 남성에 비해 상대적으로 더 많은 시간을 하루 가사 노동 시간에 할애함을 알 수 있다.

[24-한-경제4-여]

(ㄹ) 남자의 하루 평균 가사 노동 시간은 증가하는 추세를 보이고 여자의 시간은 감소하는 추이를 보이나 여자가 투입하는 시간은 여전히 남자보다 더 많다는 것을 알 수 있다.

[2-고급-베트남-경제4-여]

외국인 대학생과 한국인 대학생의 차이는 B1-2와 B2-2 사용에서 나타난다. 외국인 대학생들은 시각적 데이터가 제시하고 있는 수치를 그대로 왼쪽에서 오른쪽, 또는 위에서 아래로 제시하는 B1-2 단계의 사용이 89%로 대단히 많은 데 비해 한국인 대학생은 동일 단계 분포 비율이 34%에 불과하다. 한편 B2-2 단계 사용의 고정성이 나타나는 한국인 대학생 집단의 데이터 서술 방식은 주로 수치들을 비교해 그룹 간 어떠한 특성이 있는지를 비교하는 것을 고정적으로 제시한다. 즉, 외국인 대학생들은 시각적 데이터를 기준을 세워 그에 따라 배열하기보다는 제시된 그대로를 서술하는 경향이 크고, 한국인 대학생들은 단순한 시각적 데이터의 반복으로서의 언어적 서술을 피하고 주어진 데이터 간의 관계를 파악해 제시하려는 경향이 있는 것이다. 그렇다면 두 집단에서의 이동 마디 단계 사용 빈도 평균의 차이는 어떠한지 알아보기로 하겠다.

40) 더 자세한 언어적 분석은 다음 절에서 제시하도록 한다.

<표 46> 외국인/한국인 대학생의 B 이동 마디 선택 빈도 평균 t-test 결과

| | 빈도 | | 평균 | | 표준편차 | | t | p |
|------|-----|-----|-------------|-------------|------|------|--------|------|
| | 외국인 | 한국인 | 외국인 | 한국인 | 외국인 | 한국인 | | |
| B1_1 | 35 | 17 | 0.28 | 0.34 | 0.85 | 0.82 | 0.42 | 0.67 |
| B1_2 | 277 | 41 | 2.22 | 0.82 | 2.20 | 1.38 | -5.00* | 0.00 |
| B2_1 | 205 | 31 | 1.64 | 0.62 | 2.20 | 1.04 | -3.13* | 0.00 |
| B2_2 | 149 | 95 | 1.19 | 1.9 | 1.49 | 2.07 | 2.18* | 0.03 |
| B2_3 | 279 | 87 | 2.23 | 1.74 | 1.50 | 1.55 | -1.93 | 0.06 |
| B2_4 | 83 | 40 | 0.66 | 0.8 | 0.96 | 0.94 | 0.85 | 0.40 |
| B2_5 | 37 | 6 | 0.30 | 0.12 | 0.58 | 0.32 | -2.52* | 0.01 |
| B3 | 291 | 146 | 2.33 | 2.92 | 1.82 | 1.98 | 1.89 | 0.06 |
| B4 | 357 | 196 | 2.86 | 3.92 | 1.68 | 1.97 | 3.58* | 0.00 |

*p<0.05

<표 46>에서는 B1-2, B2-1, B2-2, B2-5, B4의 차이가 유의하게 확인되었다. 반면 B1-1, B2-3, B2-4, B3은 그 차이가 유의하지 않아 두 집단이 유사하게 선택하는 것을 알 수 있다. B1-2, B2-1, B2-5는 외국인 대학생의 선택 빈도 평균이 유의하게 높았다. B2-2와 B4는 한국인 대학생에게서 선택 빈도 평균이 유의하게 높았다.

시각적 데이터의 정보를 자세히 나열하는 B1-1과 B1-2는 시각적 데이터를 그대로 언어 기호로 변환한 것이기 때문에 인지적, 언어적 난도가 높지 않다. 다만 B1-1은 두 집단의 차이가 유의하지도 않고, 빈도수가 높지도 않아 두 집단 모두에서 잘 사용하지 않는 것으로 보인다. 과제에 제시한 그림2가 시계열적 데이터이기는 하나, 남성과 여성으로 그룹화해서 추이를 살피는 것이 대부분이고, 연도를 주어로 삼아 추세를 확인하는 것은 선호하지 않는 것으로 생각된다. 따라서 B1-1은 두 집단 모두에서 별로 사용하지 않는 경향성을 나타낸다. 반면 'B1-2. 시각적 데이터 순서대로 나열'은 외국인 대학생의 사용이 유의하게 많았다.

B2단계는 데이터를 수치적으로 계산하고, 필자의 의도를 개입하여 독자와의 거리를 설정해야 하기 때문에 인지적 노력이 비교적 많이 든다. 또

한 다양한 서술 방식의 적용이 요구되기 때문에 언어적 레퍼토리 역시도 B1보다는 많이 필요하다. ‘B2-1. 수치가 높은/낮은 순서대로 나열’과 ‘B2-5. 특이성/중요 지점 지적’은 외국인 대학생이 유의하게 더 많은 빈도로 사용하였다. 이는 B2-1이 B1-2에 비해 인지적 부담이 조금은 있지만 시각적 데이터를 확인하며 그 수치의 많고 적음 순서로 배열하는 것이라 다른 단계에 비해 부담이 덜해 외국인 대학생의 선호도가 높은 것으로 생각된다. ‘B2-5. 특이성/중요 지점 지적’은 주어진 시각적 데이터에서 두드러지는 부분이나 이상치에 가까운 부분을 짚어내는 것이다. 이는 자료에 대한 인지적 노력이 수반되어야 해 외국인 학습자들에게 부담이 될 것으로 생각하였으나 이에 대한 선택 빈도가 오히려 한국인 대학생에 비해 유의하게 높다. 한국인 대학생의 경우에는 이 이동 마디 선택이 6회로 대단히 적은데, 이러한 선택 뒤에는 그에 대한 분석이 반드시 뒤따르기 때문에 그 부담을 피하기 위해 아예 선택을 하지 않는 것으로 생각된다. 그에 반해 외국인 대학생은 두드러지는 중요 지점을 지적하는 경우는 많으나 단순한 언급으로 끝나는 경우가 많고, 이에 대해 분석을 추가하는 경우는 많지 않다. 이는 두드러진 부분에 대한 지적이 그 해석을 동반해야 하는 것을 외국인 대학생들이 인지하지 못했을 수도 있고, 두드러진 부분의 해석까지 수행하기에 인지적이든 언어적이든 자원이 부족했을 수도 있어 그러한 것으로 생각된다. 이러한 경향은 아래 (ㄱ)과 (ㄴ)을 통해 확인할 수 있다. (ㄱ)은 한국인 대학생이 B2-5를 언급하고 그에 대한 분석이 후행하는 것을 보여준다.

(ㄱ) (B2-5)외벌이를 하는 경우를 보면 특이한 점이 있는데, 외벌이를 하는 성별이 하루 가사 노동 시간 분배에 큰 영향을 끼치지 못한다는 것이다. (B2-2)남편이 외벌이를 하는 경우에는 남성 31분, 여성 240분으로 여성이 남성보다 8배 많은 시간을 가사 노동에 투입하는 것을 알 수 있다. 이에 비해 아내가 외벌이 하는 경우에는 남편이 외벌이 하는 경우보다 남성의 하루 가사 노동 시간이 두배 늘었지만 여성의 하루 가사 노동 시간보다는 적은 것을 확인 할 수 있다. (B4)두 상황 모두 아내가 남편보다 가사 분담을 더 많이 하고 있는 것이다. 상황이 반대인데 결과가 반대로 나오지 않는 점은 기본적으로 여성이 남성보다 가사 노동을 많이 분담하고 있다는 것으로 이해할

수 있으며, 다음 자료 또한 그 사실을 뒷 받침 해준다.

[30-한-경제4-여]

반면 (ㄷ)의 경우 외국인 대학생은 그래프에서의 특이한 추세를 확인하기는 했지만 이것이 어떠한 의미인지 해석하는 데까지는 나아가지 못하고, 전체적인 의미만을 후술하는데 그친다. 즉, 2021년의 여성 가사 노동 시간 감소세에 대한 해석을 덧붙이는 것이 아니라 이에 대해 단순 언급만 한다. 그다음 C1에서는 전체 데이터 설명문에 대한 해석만을 제시하고 있는 것이다.

(ㄷ) (B3)<그 림2>에 따르면 (B2-3)남성은 노동 시간이 48분에서 78분까지 지속적으로 증가 한 것으로 나타났다. 여성은 2019년~2020년 174분에서 187분으로 늘어간 것으로 드러났다. **(B2-5)2021년에는 180분으로 감소한 것으로 보였다.**

<그 림2>(삽입함)

(C1)조사를 통해 맞벌이든 외벌이(남편)든 외벌이(아내)든 여성이 남성보다 가사노동에 더 많은 시간을 알 수 있다.

[115-고급-중국어-경영1-여]

한편, 'B2-2. 수치적 비교'와 'B4. 수치의 의미 제시'는 한국인 대학생들의 선택 빈도가 95% 수준에서 유의하게 높았다. B2-2의 경우는 기준 집단과 비교 집단 사이의 수치를 산술적으로 계산해야 하고 그에 대한 관계성 파악에 인지적 노력이 드는 것과, 비교에 사용되는 어휘 사용 미숙성에 대한 부담 때문에 외국인 대학생들은 쉽게 선택하지 못한 것으로 생각된다. 반면 한국인 대학생들은 어휘 사용에서 가용할 수 있는 인지적 용량이 있기 때문에 수치 비교와 같은 행위를 하기에 상대적으로 용이하다고 볼 수 있다. 따라서 한국인 대학생들은 시각적 데이터를 그대로 나열하는 B1-2 마디를 피하고 주로 B2-2와 같이 데이터를 한 번 기준에 따라 묶어 비교한 뒤 제시하는 경향이 있다. 이는 한국인 대학생

데이터 설명문의 66%가 채택할 정도로 선호도가 높다.

B4의 경우 주로 앞뒤에 제시한 데이터의 의미를 정리하여 제시하는 역할을 수행한다. 한국인 대학생들은 데이터를 시각적으로 주어진 대로 나열하기보다는 그룹화, 수치적 비교 등의 방식으로 데이터를 재구성해 제시하는 것을 선호하는데, 그 과정에서 데이터가 궁극적으로 나타내는 의미가 무엇인지를 부연하는 편이다. 즉, 한국인 대학생들의 데이터 설명문의 기술은 데이터 제시로 끝나는 것이 아닌, 그것이 나타내는 의미를 명확히 하고자 하는 경향이 있다.

2) 학습자 언어권별 이동 마디 및 단계 분포 및 빈도

이번에는 위에서 살펴본 외국인 학습자들의 B 이동 마디 순서 선택 경향이 언어권에도 적용되는지 확인해 보도록 하겠다. 아래의 <표 47>은 중국어권 학습자와 베트남어권 학습자의 B 이동 마디 단계 분포이다.

<표 47> 중국어권/베트남어권 학습자의 B 이동 마디 분포

| | 이동 마디와 단계 | 중국어권 | 베트남어권 |
|------|-------------------|------------------|-----------------|
| B | 수치/결과 나열 | | |
| B1 | 구체적인 정보 추출 | | |
| B1-1 | 연대기적 나열 | 10(16.9%) | 5(10%) |
| B1-2 | 시각적 데이터 순서대로 나열 | 44(74.5%) | 36(72%) |
| B2 | 의도적 그룹화 | | |
| B2-1 | 수치가 높은/낮은 순서대로 나열 | 36(61%) | 27(54%) |
| B2-2 | 수치적 비교 | 16(27.1%) | 41(82%) |
| B2-3 | 특성에 따른 그룹화 | 53(89.8%) | 44(88%) |
| B2-4 | 제일 큰/작은 수치에 주목 | 26(44%) | 16(32%) |
| B2-5 | 특이성/중요 지점 지적 | 10(17%) | 17(34%) |
| B3 | 시각적 데이터 언급 표지 사용 | 40(67.7%) | 47(94%) |
| B4 | 수치의 의미 제시 | 52(88.1%) | 50(100%) |
| | 전체 데이터 설명문 편수 | 59 | 50 |

중국어권 학습자와 베트남어권 학습자는 모두 'B1-2. 시각적 데이터 순서대로 나열', 'B2-3. 특성에 따른 그룹화', 'B3. 시각적 데이터 언급 표지 사용', 'B4. 수치의 의미 제시'의 네 단계를 60% 이상 사용함으로써 두 언어권 집단이 고정적으로 사용하는 단계를 보여주었다. 다만 중국어권 학습자는 'B2-1. 수치가 높은/낮은 순서대로 나열'을 추가적으로 더 고정적으로 사용하며, 베트남어권 학습자 집단은 'B2-2. 수치적 비교'에서 추가로 82%의 분포를 보여주고 있다. 이렇듯 두 집단의 단계 선호와 고정성은 B2-1과 B2-2의 차이를 제외하고는 모두 유사한 것으로 나타났다. 따라서 이들 B1-2, B2-3, B3, B4 사용은 언어권과 관계없이 데이터 설명문에 공통적으로 사용하는 이동 마디 단계라 할 수 있다.

B2-1은 B1-2와 유사하나 순서를 한 번 더 생각해야 하기 때문에 B2로 배정되었으며, 주어진 시각적 데이터를 보면서 수치가 높은/낮은 항목을 찾아 배열하기 때문에 B1처럼 낮은 인지적 노력을 요한다. 따라서 B2-1을 선호하는 중국어권 학습자들은 한국어 사용에 조금 더 인지적 노력을 사용해야 하는 숙달도가 낮은 학습자일 가능성이 있다. 실제로 B2-1이 나타난 중국어권 학습자 데이터 설명문 36편 중 22편은 중급 학습자가 작성하였다.

B2-2는 주어진 시각적 데이터 내의 수치의 관계를 파악하고 이를 서로 비교하여 간단한 산술을 통해 몇 배인지 등을 제시해야 하기 때문에 작업 기억에 상대적 여유를 확보할 수 있는 한국인 대학생들이 주로 사용한다고 분석한 바 있다. 베트남어권 대학생들이 중국어권 대학생들보다 작업 기억을 더 사용하기 위해서는 한국어 숙달도가 높거나, 고학년이어서 산술적 비교 및 계산과 같은 데이터 처리 방식에 익숙해 제2언어 사용이든, 인지적 사용이든 어느 측면으로든 자동화되어 이 이동 마디를 선호한다고 볼 수 있을 것이다. 실제로 베트남어권 학습자 데이터 설명문에서 B2-2 이동 마디를 사용한 학습자 41명 중 37명은 3~4학년으로 고학년군에 속한다. 숙달도도 이에 영향을 미칠지 다음에서 다시 확인할 것이다.

그렇다면 이들 집단의 이동 마디 단계 선택 빈도 차이는 분포와는 다른

지를 알아보도록 하겠다. 학습자 언어권별 빈도수 평균 차이는 다음 <표 48>과 같다.

<표 48> 중국어권/베트남어권 학습자의 B 이동 마디 선택 빈도 평균 t-test 결과

| | 빈도 | | 평균 | | 표준편차 | | t | p |
|------|-----|-----|-------------|-------------|------|------|--------|------|
| | 중국 | 베트남 | 중국 | 베트남 | 중국 | 베트남 | | |
| B1_1 | 24 | 8 | 0.41 | 0.16 | 1.08 | 0.50 | 1.48 | 0.14 |
| B1_2 | 174 | 83 | 2.95 | 1.66 | 2.61 | 1.49 | 3.20* | 0.00 |
| B2_1 | 115 | 59 | 1.95 | 1.18 | 2.66 | 1.45 | 1.89 | 0.06 |
| B2_2 | 34 | 111 | 0.58 | 2.22 | 1.11 | 1.49 | -6.38* | 0.00 |
| B2_3 | 142 | 107 | 2.41 | 2.14 | 1.61 | 1.39 | 0.91 | 0.36 |
| B2_4 | 51 | 22 | 0.86 | 0.44 | 1.14 | 0.73 | 2.25* | 0.03 |
| B2_5 | 14 | 20 | 0.24 | 0.40 | 0.59 | 0.60 | -0.61 | 0.54 |
| B3 | 106 | 154 | 1.80 | 3.08 | 1.65 | 1.76 | -3.88* | 0.00 |
| B4 | 142 | 177 | 2.41 | 3.54 | 1.57 | 1.66 | -3.61* | 0.00 |

*p<0.05

B1-1은 전체 학습자군에서 모두 선택 빈도가 낮는데, 연도별이나 시간별로 제시하는 것보다는 그룹으로 묶어 추이를 확인하는 것을 더 선호하기 때문인 것으로 보인다. 두 학습자군에서 모두 다수 사용돼 빈도 차이가 유의하지 않은 B2-1과 B2-3의 경우는 수치의 배열을 정비하나 시각적 데이터에서 이들을 확인하며 그중 특수한 것들을 골라내 배열한다는 점에서 B1 이동 마디의 단계들보다는 조금 더 인지적 노력을 통해 저자의 의도를 반영할 수 있도록 하지만, 그 노력이 시각적 데이터를 보면서 확인할 수 있다는 데서 조금 더 연산이 필요한 ‘B2-2. 수치적 비교’ 처럼 많이 들지는 않는다. 또한 ‘B2-5. 특이성/중요 지점 지적’ 역시 두 학습자군 모두에서 선택 빈도수가 낮다. B2-5의 선택 빈도가 낮은 것은 이 단계가 갖는 인지적 어려움 때문이라 본다. 이 단계를 구성하기 위해서는 데이터의 특이한 지점을 지적하고 그에 따라 논의를 전개하거나 다른 현상을 연결하는 복잡한 인지적 노력이 필요하다. 외국인 학습자군에서는 이를 한국인 대학생에 비해 유의하게 다수 사용했으나 언어적으로

는 선택 빈도 차이는 없어 다음 숙달도별 분석에서 다시 확인해 보도록 하겠다.

한편 'B1-2. 시각적 데이터 순서대로 나열'과 'B2-4. 제일 큰/작은 수치에 주목'의 사용 빈도는 두 단계 모두 중국어권 학습자가 유의하게 높았다. B1-2는 인지적 노력이 적게 들어 주로 중급인 중국어권 학습자가 사용하는 것으로 보인다. 또한 저학년군(1~2학년)의 사용이 174개 중 134개로 학술적 데이터 설명문에 익숙하지 않거나 제한된 학술성에만 노출되어 그러할 수도 있다는 추론을 하게 한다. 이는 뒤에서 중급 학습자들의 사용 경향을 확인하여 검토해 보도록 한다. B2-4는 외국인 대학생 전체 빈도 평균은 유의하게 높지 않았는데, 중국어권 학습자에게서 그 빈도가 높게 나타난다. 이는 중국어권 학습자의 특성으로 인지적 노력은 B2-1과 유사하게 주어진 시각적 데이터에서 제일 큰 또는 작은 수치를 발견해서 지적하는 것이라 큰 어려움은 없을 듯하다. 다만 그 뒤에 이어지는 내용이 데이터 설명문이 지향하는 주제 또는 목표를 뒷받침하는 것으로 추가될 수 있어야 인지적 노력이 더 수반된 것이라 할 수 있다.

베트남어권 학습자는 'B2-2. 수치적 비교', 'B3. 시각적 데이터 언급 표지 사용', 'B4. 수치의 의미 제시'를 유의하게 더 많이 사용함을 확인할 수 있다. B2-2는 인지적 노력이 더 필요하다는 측면에서 이 단계의 빈번한 사용이 숙달도 및 학년과 같은 다른 인지적 요인과 관계가 있는 것이 아닌지 앞에서 확인하기로 하였다. 학년 요인은 빈도에서도 확인할 수 있는데, 베트남어권 학습자가 산출한 B2-2 이동 마디 111개 중 104개가 3~4학년의 고학년군이 작성한 것임을 알 수 있다. B3은 데이터 설명문의 텍스트적 메타 기능을 수행하는 부분으로 데이터 설명문의 구조의 응집성을 담당하며 독자에게 저자가 원하는 곳을 볼 수 있도록 인도한다는 측면에서 대인적 메타 기능도 지닌다. 이러한 텍스트의 대인적 메타 기능을 이해하고 사용하는 것은 TOPIK과는 다른 학술적 데이터 설명문 구조에 대한 노출이 있었으며, 그에 대한 지식도 있음을 뜻한다. 즉, 고학년의 사용이 높음을 확인할 수 있다(3~4학년 작문에서 177개 중 123개 B3 사용). 숙달도에 대한 것은 역시 다음 절에서 알아보도록 한

다. B4는 한국인 대학생들 역시도 유의하게 높은 빈도로 사용하는 단계이기 때문에 ‘대학생군’으로 묶일 수 있는 요인이 있을 것이라 본다. 따라서 베트남어권 학습자들의 해당 단계 사용에는 인지적 능력에 영향을 끼치는 다른 변인이 개입했다고 볼 수 있다. 예를 들어, 전공 수업이 한국어의 학술적 글쓰기 노출에 영향을 주는 학년 변수라거나 데이터 설명문을 구성할 때 가용할 수 있는 어휘문법적 레퍼토리에 영향을 미치는 제2언어 숙달도 변수가 이에 속한다. 학년 변수 개입은 B2-2, B3에 이어 여기에서도 확인이 되는데, 베트남어권 학습자가 산출한 B4 이동 마디 177개 중 143개가 3~4학년 학습자가 생성한 것이기 때문이다.

3) 학습자 숙달도별 이동 마디 및 단계 분포 및 빈도

이어 학습자 숙달도별로 B 이동 마디 단계 사용 빈도와 그의 차이가 있는지 알아보도록 하겠다. 중급 학습자와 고급 학습자는 동일하게 B1-2, B2-3, B3, B4 단계를 사용한 데이터 설명문 편수가 60%가 넘어감으로써 이들 단계를 고정적으로 사용하는 양상을 드러내었다. 따라서 한국어 숙달도별로 이동 마디 단계 분포에 차이가 나타나지는 않는 것을 알 수 있다. 이들 공통 단계들은 숙달도에 관계없이 모든 외국인 학습자들이 공통적으로 사용하는 고정 단계라 할 수 있다. 특히 B1-2은 이를 사용한 후 다음에 다른 단계를 다시 사용해서 동일한 데이터를 여러 방면으로 확인한다는 측면에서 인지적 노력이 낮지만 여러 숙달도의 학습자들이 사용한다고 할 수 있다. 다만 B2-1은 중급 학습자에게 주로 분포함으로써 제2언어적 구성에 더 많이 작업 기억을 사용해야 하기 때문에 앞에서 언급했던 바와 같이 단계 구성에 필요한 인지적 노력을 아끼려는 시도를 확인할 수 있다.

<표 49> 중급/고급 학습자의 B 이동 마디 분포

| | 이동 마디와 단계 | 중급 | 고급 |
|---|-----------|----|----|
| B | 수치/결과 나열 | | |

| | | | | |
|---------------|-------------------|------------------|------------------|--|
| B1 | 구체적인 정보 추출 | | | |
| B1-1 | 연대기적 나열 | 13(17.1%) | 3(5.4%) | |
| B1-2 | 시각적 데이터 순서대로 나열 | 51(72.8%) | 38(%) | |
| B2 | 의도적 그룹화 | | | |
| B2-1 | 수치가 높은/낮은 순서대로 나열 | 42(60%) | 30(47.2%) | |
| B2-2 | 수치적 비교 | 30(42.8%) | 30(47.2%) | |
| B2-3 | 특성에 따른 그룹화 | 59(84.2%) | 50(90.9%) | |
| B2-4 | 제일 큰/작은 수치에 주목 | 36(51.4%) | 14(25%) | |
| B2-5 | 특이성/중요 지점 지적 | 16(22.8%) | 14(25%) | |
| B3 | 시각적 데이터 언급 표지 사용 | 51(72.8%) | 47(85.4%) | |
| B4 | 수치의 의미 제시 | 65(92.8%) | 52(94.5%) | |
| 전체 데이터 설명문 편수 | | 70 | 55 | |

다음으로는 중급과 고급 학습자의 이동 마디 선택 빈도 차이를 알아보도록 하겠다. 우선 빈도 차이가 유의하게 나타난 것은 B1-1, B3 두 개이다. B1-1은 중급 학습자, B3는 고급 학습자의 빈도 평균이 유의하게 높았다.

중급 학습자는 시각적 데이터의 단순 언어적 변환인 B1-1의 빈도수가 컸다. 이는 앞에서 언급한 바와 같이 인지적 노력이 크게 필요하지 않은 B1 이동 마디의 단계들이 한국어 사용으로 작업 기억을 소모하는 중급 학습자들에게 선호되는 것일 수 있다고 본다. 또한 숙달도가 낮은 학습자일수록 데이터 설명문에 노출되어도 그 서술 방식의 다양성을 언어적으로 이해하지 못할 가능성이 있기 때문에 데이터 설명문 경험이 TOPIK에 국한되어 그가 요구하는 B1 이동 마디의 단계들을 주로 사용하는 것이 아닌가 생각된다.

<표 50> 중급/고급 학습자의 B 이동 마디 선택 빈도 평균 t-test 결과

| | 빈도 | | 평균 | | 표준편차 | | t | p |
|------|----|----|-------------|------|------|------|-------|------|
| | 중급 | 고급 | 중급 | 고급 | 중급 | 고급 | | |
| B1_1 | 30 | 5 | 0.43 | 0.09 | 1.06 | 0.39 | 2.22* | 0.03 |

| | | | | | | | | |
|------|-----|-----|------|-------------|------|------|--------|------|
| B1_2 | 164 | 113 | 2.34 | 2.05 | 2.44 | 1.82 | 0.75 | 0.46 |
| B2_1 | 123 | 82 | 1.76 | 1.49 | 2.44 | 1.83 | 0.69 | 0.49 |
| B2_2 | 67 | 82 | 0.96 | 1.49 | 1.35 | 1.61 | -2.00 | 0.05 |
| B2_3 | 150 | 129 | 2.14 | 2.35 | 1.62 | 1.32 | -0.74 | 0.46 |
| B2_4 | 64 | 19 | 0.91 | 0.35 | 1.08 | 0.67 | -0.76 | 0.45 |
| B2_5 | 18 | 19 | 0.26 | 0.35 | 0.50 | 0.67 | -0.81 | 0.42 |
| B3 | 137 | 154 | 1.96 | 2.80 | 1.69 | 1.86 | -2.62* | 0.01 |
| B4 | 183 | 174 | 2.61 | 3.16 | 1.60 | 1.73 | -1.82 | 0.07 |

*p<0.05

또한 과제의 분석과 제2언어의 사용이 동일한 작업 기억 용량 내에서 동시에 수행되어야 한다면 둘 중 하나의 인지적 노력이 덜 드는 방식을 채택하는 것이 자연스러울 것이다. 한국어 표현에 인지적 노력이 많이 드는 중급 학습자가 과제의 분석 시 인지적 노력이 덜 드는 방식을 택하는 것은 당연한 일일 것이다. 따라서 중급 학습자들은 데이터 설명문 작성 시 주로 단순 시각적 데이터 제시 및 높고 낮은 수치에 집중하는 경향을 확인할 수 있다. 이는 다음과 같은 중급 학습자 작문에서 확인할 수 있다.

<표 51> 중급 학습자 B2-계열 사용 양상

(A2) 요즘 가사 노동 시간 대한 초사를 한다. 설문 응답자가 1000명 있는데 그중에 남성과 여성 각 500명 있다. (B2-4)대졸 학력이 사람은 620명으로 제일 많다. 중졸 이하가 사람은 10명만 있다. 하루 평균 가사 노동 시간은 1시간~2시간이 402명으로 제일 많아 사람이 있지만 4시간 이상이 17명만 이다. 혼인 상태는 기혼자가 높은 것 있다.

(B3) 맞벌이 상대별 하루 가사 노동 시간은 조사 결과를 보면 (B2-3) 맞벌이 종에 여성은 가사 노동 시간이 남성보다 참가한 사람을 많다. 외벌이 종에는 남성보다 여성 높은 것 편이다.

(B3) 남성과 여성의 하루 평균 가사 시간 비교는 조사 결과를 보다. (B1-1) 2019년보다 2021년까지 여성 가사 노동 시간 제일 많아 편이다. (B2-3) 그런데 남성에 하루 평균 가사 노동 시간 잠가주가 천천히 오르다

[11-중급-중국어-경제1-여]

고급 학습자는 데이터 설명문 작성에서 'B3. 시각적 데이터 언급 표지' 단계를 중급 학습자에 비해 더 많이 사용한다. B3은 데이터 설명문의 구성 요소인 '위치'에 해당하는 필수적인 요소이기도 하고, 독자에게 시각적 데이터 내 데이터에 집중할 수 있도록 시선을 견인하는 대인적 역할을 맡고 있어 데이터 설명문의 특성을 드러내는 중요한 역할을 맡고 있다. B3은 외국인 대학생과 한국인 대학生の 비교에서는 한국인 대학生の 사용 빈도가 더 높았고 언어권별 비교에서는 베트남어권 학습자들의 수치가 유의미하게 높았다. 그리고 중급과 고급의 비교에서는 고급의 사용 빈도가 유의적으로 더 높은 것으로 나타났다. 이를 볼 때 B3의 사용은 한국어 숙달도가 높을수록 비교적 그 사용 빈도가 높다고 볼 수 있다. 이는 독자를 견인하는 B3과 같은 담화 표지의 사용은 텍스트적으로 응집성을 나타내고, 독자와의 대인적 거리를 어느 정도 파악해야만 사용할 수 있는 관습적인 것이기 때문에 학술적 텍스트 구성 경험이 전제되어야 한다고 할 수 있다. 따라서 B3 빈도가 높은 것은 베트남어권의 특성이라기보다는 숙달도의 측면이 크다고 볼 수 있다.

다만 앞에서 언급한 베트남어권 학습자의 B2-2, B4 단계 선호가 언어 숙달도와 관계가 있을 것이라는 추정은 중급과 고급 차이가 유의하지 않은 것으로 보아 옳지 않음이 밝혀졌다. 따라서 B2-2, B4 단계 선호에 대해 인지적 기준으로 남은 설명은 학년뿐이라 볼 수 있다. 또한 B2-5의 외국인 대학생 빈도가 높은 점이나, 중국어권 학습자의 B1-2, B2-4 빈도가 높은 점도 숙달도로는 설명이 되지 않는다. 따라서 B2-4, B2-5는 앞서 추정된 대로 단순히 눈에 띄는 점만을 지적한 뒤 후행하는 해석을 덧붙이지 않은 저학년군의 장르성 부재로 판단할 수 있으며, B1-2 역시도 저학년군의 인지적 노력 절감을 위한 단순한 수치 나열이라 볼 수 있다. 이러한 모든 경향을 B1, B2, B3, B4로 정리하면 다음과 같다.

[B1 하위 이동 마디 정리]

첫째, 외국인과 한국인 대학생 B1-2 빈도의 평균 차이는 외국인 대학생이 유의하게 더 높다. 또한 B1-2가 외국인 대학生の 데이터 설명문에

분포하는 비율도 71%로 고정성이 높은 반면 한국인 대학생의 분포는 34%로 고정성이 높지 않다.

둘째, 중국어권 학습자는 베트남어권 학습자에 비해 B1-2의 빈도 평균이 유의하게 높았다. 다만 중국어권 학습자와 베트남어권 학습자 모두 B1-2의 분포가 70%를 넘어 두 집단 모두 고정성을 갖고 있다고 할 수 있다.

셋째, 중급 학습자는 고급 학습자에 비해 B1-1 단계의 빈도 평균에서 유의한 차이를 보였다. B1-1은 중급과 고급 모두에서 20% 미만의 낮은 빈도를 보여 고정성은 낮은 것으로 밝혀졌다. 따라서 중급 학습자 중 일부가 B1-1을 다수 사용한 것으로 보인다.

[B2 하위 이동 마디 정리]

첫째, 'B2-1. 수치가 높은/낮은 순서대로 나열'과 'B2-5. 특이성/중요 지점 지적'은 외국인 대학생의 빈도 평균 차이가 한국인 대학생에 비해 유의하게 높다. 반면 한국인 대학생은 B2-2에서 외국인 대학생에 비해 유의하게 높은 차이를 보였다. 그러나 B2-1, B2-5는 두 집단 모두에서 고정적으로 사용되지 못하고, B2-2만 한국인 대학생군에서 66%로 고정성을 나타낸다.

둘째, 'B2-2. 수치적 비교'는 중국어권 학습자에 비해 베트남어권 학습자의 빈도가 유의하게 높았고, 'B2-4. 제일 큰/작은 수치에 주목'은 반대로 베트남어권 학습자에 비해 중국어권 학습자의 빈도 평균이 유의하게 높았다. B2-2의 분포는 베트남어권에서는 82%를 차지하였지만, B2-4의 분포는 두 집단 모두에서 50% 미만이라 이들은 고정적인 단계가 아니고 일부 학습자들이 선호하는 것을 알 수 있다.

셋째, 중급과 고급 학습자는 B2 하위 단계에서는 유의한 차이를 보이는 것이 없었다. 즉, 두 그룹이 모두 유사하게 단계들을 사용한 경향을 보인다. 중급 학습자군에서 'B2-1. 수치가 높은/낮은 순서대로 나열'의 분포가 고정적이었던 것, 'B2-3. 특성에 따른 그룹화'는 중급과 고급 모두에서 분포가 고정적이라는 것은 알 수 있다.

[B3 이동 마디 정리]

첫째, 한국인 대학생과 외국인 대학생의 B3 빈도 평균은 유의한 차이가 없었다. 다른 단계들에 비해 두 집단 모두 B3을 다수 사용하였다. 분포는 두 집단 모두 70%가 넘어 고정성이 있다고 할 수 있다.

둘째, 중국어권 학습자들에 비해 베트남어권 학습자의 B3 평균이 유의하게 높았다. 또한 그 분포는 두 집단 모두에서 고정성이 있는 것으로 확인되었지만, 중국어권 학습자는 67.7%, 베트남어권 학습자는 94%로 지배적인 정도는 베트남어권 학습자가 더 컸다.

셋째, 고급 학습자의 B3 사용은 중급 학습자에 비해 유의하게 높았다. 중급과 고급 학습자 모두에서 그 분포도 70% 이상이어서 고정성도 드러난다.

[B4 이동 마디 정리]

첫째, B4 사용에 있어 한국인 대학생이 외국인 대학생에 비해 그 빈도 평균이 유의하게 높았다. 또한 두 집단 모두에서 90%가 넘는 사용 분포를 보여 고정성이 매우 큼을 알 수 있다.

둘째, 중국어권 학습자들에 비해 베트남어권 학습자의 B4 평균이 유의하게 높았다. 또한 두 집단 모두에서 B4의 분포가 80% 이상 높게 나타났다. 특히 베트남어권 학습자는 모든 작문에서 B4가 나타나는 것을 볼 수 있다.

셋째, 중급과 고급 학습자의 B4 사용에서 유의한 차이는 없었다. 다만 두 집단 모두에서 분포는 90% 이상으로 대단히 높은 고정성을 보였다. 이들의 사용 빈도도 다른 단계들에 비해 높은 것을 볼 수 있다.

(3) 이동 마디 C: 해석 및 결론

Swales & Feak(1994)에서는 학문 목적 학습자라면 데이터 설명문에서 그에 대한 논평이 반드시 나타나야 함을 지적한 바 있다. 이동 마디에서 C 유형이 자주 생략되는 것은 학습자들이 학술적 데이터 설명문에 필요한 이와 같은 지식을 명시적으로 학습하지 못하였거나, 주어진 데이터를 주어진 것 이상으로 해석하여 텍스트 전체의 의미를 이끌어낼 수 있는 능력을 갖추지 못함을 뜻한다. 본고에서 살펴본 C 이동 마디 사용 양상은 아쉽게도 Swales & Feak(1994)가 지적한 제2언어 학습자의 한계를 그대로 나타내고 있다. 이를 한국인 대학생과 비교한 후, 외국인 학습자 집단에 대해 언어권 및 숙달도별로 차이를 알아보도록 하겠다.

1) 외국인/한국인 대학생의 이동 마디 단계 분포 및 빈도

외국인/한국인 대학생의 이동 마디 단계 분포에서 외국인 대학생 집단은 고정적으로 나타나는 이동 마디 단계가 없는 것을 볼 수 있다. 반면 한국인 대학생 집단에서는 C1과 C3가 60%가 넘어 이들 단계를 고정적으로 사용하는 것을 볼 수 있다.

<표 52> 외국인/한국인 대학생의 C 이동 마디 분포

| | 이동 마디와 단계 | 외국인 | 한국인 |
|----|------------------------------|-----------|----------------|
| C | 해석 및 결론 | | |
| C1 | 주어진 데이터/상식적 지식에 근거한 해석 | 69(55.2%) | 48(99%) |
| C2 | 데이터의 비판점과 개선 방향 | 2(1.6%) | 13(26%) |
| C3 | 문제 해결 방안/ 후속 연구/ 미래 방향 제시 | 51(40.8%) | 34(68%) |
| | 전체 데이터 설명문 편수 | 125 | 50 |

특히 C1의 경우 한국인 대학생 데이터 설명문에서는 두 편을 제외하고 모두 사용한 것을 알 수 있다. 즉, 한국인 대학생은 데이터 설명문에 거

의 대부분 해석을 덧붙이는 경향이 있음을 알 수 있으며, 더 나아가 미래의 방향이나 문제 해결 방안 등을 제시하기도 한다. 이러한 경향이 이동 마디 단계 선택 빈도에서도 동일하게 나타나는지 아래 <표 53>에서 확인해 보았다.

<표 53> 한국인/외국인 대학생의 C 이동 마디 선택 빈도 평균 t-test 결과

| | 빈도 | | 평균 | | 표준편차 | | t | p |
|----|-----|-----|------|-------------|------|------|-------|------|
| | 외국인 | 한국인 | 외국인 | 한국인 | 외국인 | 한국인 | | |
| C1 | 197 | 139 | 1.58 | 2.78 | 2.19 | 2.71 | 2.77* | 0.01 |
| C2 | 2 | 20 | 0.02 | 0.40 | 0.13 | 0.85 | 4.90* | 0.00 |
| C3 | 69 | 63 | 0.56 | 1.29 | 0.77 | 1.43 | 4.19* | 0.00 |

*p<0.05

C 이동 마디의 단계 선택 빈도는 모두 한국인 대학생이 외국인 대학생에 대해 모두 95% 수준에서 유의하게 높은 빈도를 보였다. 이는 C 이동 마디 전체를 한국인 대학생이 외국인 대학생보다 더 많이 사용하는 데서 기인한 것이라 본다. 따라서 이 이동 마디는 한국인 대학생의 데이터 설명문과 외국인 대학생의 데이터 설명문을 가르는 기준이 될 수 있다.

일반적으로 C1과 C3은 빈도도 높고 차이도 유의하나 C2의 경우에는 전반적으로 모든 대학생의 선택 빈도가 낮고, 그중에서도 외국인 대학생의 선택 빈도가 현저히 적은 것으로 보인다. 한국인 대학생의 경우에는 다음의 (ㄱ)처럼 수집된 데이터의 한계나 분석에 필요한 다른 데이터를 더 추가해 분석을 개선할 것을 촉구하기도 한다.

(ㄱ) 반면 3가지 자료만으로는 성별 비혼자의 하루 평균 가사노동 시간 차이, 학력별 하루 평균 가사노동 시간 차이, 자녀유무에 따른 기혼자의 성별 하루 평균 가사노동 시간 차이 등에 대해 알아낼 수 없다는 아쉬움이 든다. 그리고 표에는 나오지 않았지만 지역별-소득수준별 가사노동시간 차이에 대해서 다루었다면 제법 흥미로운 결과가 나왔을 것이라는 생각이 든다. 그리고 비슷한 사회경제적 배경을 가진 타 국가의 하루 평균 가사노동시간에 대해 다루었다면 우리나라의 하루 평균 가사노동시간과 비교분석하여 더 나

은 대안을 모색할 수 있을 것이다.

[2-한-경영4-여]

이러한 비판과 분석은 주제로 제시된 성별 가사 노동 지표의 분석을 위해 저자가 상당 수준의 지식을 갖추었음을 알려준다. 따라서 단순히 구조 교육만으로는 채울 수 없는 전공 지식 소양이 뒷받침되어야 함을 알 수 있다. 이는 전공 지식이라는 학년 변수가 생각보다 데이터 설명문의 구조에 영향을 많이 미친다는 것을 확인할 수 있다. 또한 Swales & Feak(1994)에서 강조했던 비판적인 저자의 태도 역시도 이 단계에서 나타남을 알 수 있다.

하지만, 외국인 대학생의 경우에는 이러한 서술이 이상치에 가까울 정도로 드문 것을 볼 수 있다. 본고에서는 이 역시도 외국인 대학생이 학술적 장에서 경험한 데이터 설명문에 C2와 같은 단계를 수록했던 것이 없어서 산출을 하기 어려웠으며, 전공 지식을 구조와 관련해 발전시키는 훈련도 받지 못해 그러한 것으로 본다.

2) 학습자 언어권별 이동 마디 및 단계 분포 및 빈도

다음으로는 중국어권 학습자와 베트남어권 학습자의 C 이동 마디 단계 분포를 확인해 보도록 하겠다. 중국어권 학습자는 모든 단계에서 60%를 넘지 않아 C 이동 마디에서는 고정적으로 나타나는 단계가 없다. 반면 베트남어권 학습자에게는 C1이 80%로 나타나 이 단계가 베트남어권 학습자 집단에서 고정적으로 나타나는 것임을 알 수 있다.

<표 54> 중국어권/베트남어권 학습자의 C 이동 마디 분포

| | 이동 마디와 단계 | 중국어권 | 베트남어권 |
|----|------------------------|-----------|----------------|
| C | 해석 및 결론 | | |
| C1 | 주어진 데이터/상식적 지식에 근거한 해석 | 24(40.6%) | 40(80%) |
| C2 | 데이터의 비판점과 개선 방향 | 0 | 2(4%) |

| | | | |
|----|------------------------------|-----------|---------|
| C3 | 문제 해결 방안/ 후속 연구/ 미래 방향 제시 | 20(33.8%) | 26(52%) |
| | 전체 데이터 설명문 편수 | 59 | 50 |

한편 중국어권 학습자와 베트남어권 학습자의 선택 빈도 차이를 확인해 보면 다음 <표 55>와 같다.

<표 55> 중국어권/베트남어권 학습자의 C 이동 마디 선택 빈도 평균 t-test 결과

| | 빈도 | | 평균 | | 표준편차 | | t | p |
|----|----|-----|------|-------------|------|------|--------|------|
| | 중국 | 베트남 | 중국 | 베트남 | 중국 | 베트남 | | |
| C1 | 71 | 117 | 1.20 | 2.34 | 2.20 | 2.23 | -2.64* | 0.01 |
| C2 | 0 | 2 | 0.00 | 0.04 | 0.00 | 0.20 | -1.43 | 0.16 |
| C3 | 28 | 34 | 0.47 | 0.68 | 0.79 | 0.76 | -1.36 | 0.18 |

*p<0.05

중국어권과 베트남어권 학습자의 C 이동 마디의 단계 빈도수 차이는 ‘C1. 주어진 데이터/상식적 지식에 근거한 해석’ 단계 이외에서는 유의하지 않았다. 중국어권과 베트남어권 학습자들은 주로 데이터 설명문을 앞 이동 마디에서 분석한 데이터에 기반하여 마무리 짓는 경향이 있는 듯하다. 이러한 경향은 또한 베트남어권 학습자에게서 더욱 크게 나타난다. 중국어권 학습자의 이동 마디 단계 빈도수가 베트남어권 학습자에 비해 적은 이유는 C 이동 마디 자체를 사용하지 않는 중국어권 학습자가 많기 때문인데, 이 이동 마디를 사용하지 않는 원인에 대해서는 TOPIK의 영향과 숙달도 등으로 앞에서 추론한 바 있다.

‘C2. 데이터의 비판점과 개선 방향’은 두 언어권 학습자 모두 사용 빈도가 매우 낮아서 이들 단계에 대한 선호도 자체가 매우 떨어짐을 알 수 있다. 그리고 위에서 보았듯, 한국인 대학생에게서도 그 빈도가 높지는 않다.

3) 학습자 숙달도별 이동 마디 및 단계 분포 및 빈도

중급/고급 학습자의 C 이동 마디 단계 분포를 보면 중급 학습자군에서는 고정적으로 나타나는 단계가 하나도 없는 반면, 고급 학습자군에서는 C1 단계가 60%를 넘어 고정적으로 나타난다. 즉, 고급 학습자들은 C1을 고정적인 단계로 인식하고 있다는 것이다.

<표 56> 중급/고급 학습자의 C 이동 마디 분포

| | 이동 마디와 단계 | 중급 | 고급 |
|----|------------------------------|-----------|----------------|
| C | 해석 및 결론 | | |
| C1 | 주어진 데이터/상식적 지식에 근거한 해석 | 31(44.2%) | 38(69%) |
| C2 | 데이터의 비판점과 개선 방향 | 0 | 2(3.6%) |
| C3 | 문제 해결 방안/ 후속 연구/ 미래 방향 제시 | 22(31.4%) | 29(52.7%) |
| | 전체 데이터 설명문 편수 | 70 | 55 |

그러나 이들의 단계 선택 빈도는 분포와는 조금 다르게 나타났다. 이는 다음 <표 57>에서 확인할 수 있다.

<표 57> 중국어권/베트남어권 학습자의 C 이동 마디 선택 빈도 평균 t-test 결과

| | 빈도 | | 평균 | | 표준편차 | | t | p |
|----|----|-----|------|-------------|------|------|--------|------|
| | 중급 | 고급 | 중급 | 고급 | 중급 | 고급 | | |
| C1 | 76 | 121 | 1.09 | 2.20 | 1.09 | 2.46 | -2.79* | 0.01 |
| C2 | 0 | 2 | 0.00 | 0.04 | 0.00 | 0.19 | -1.43 | 0.16 |
| C3 | 30 | 39 | 0.43 | 0.71 | 0.71 | 0.82 | -2.03* | 0.04 |

‘C. 해석 및 결론’ 이동 마디의 단계 선택 빈도에서는 고급 학습자가 중급 학습자에 비해 C1 사용 빈도가 유의하게 높음이 검증되었다. ‘C1. 주어진 데이터/상식적 지식에 근거한 해석’ 단계는 시각적 데이터와 A,B

이동 마디에서 다른 남녀의 가사 노동 시간 차이를 극복해야 함을 주장하는 것이 대부분이다. 이 이동 마디와 단계까지 나아간 학습자군 자체가 중급에는 얼마 없기 때문에 단계의 빈도는 고급 학습자에게서 더 높게 나타나는 것으로 생각된다.

‘C3. 문제 해결 방안/ 후속 연구/ 미래 방향 제시’ 단계의 빈도 분석 결과는 중급 학습자와 고급 학습자의 사용 빈도 차이가 유의하지 않음을 알려준다. 이를 통해 다른 단계에 비해 고급 학습자도 여기까지 나아가 데이터 설명문을 마무리하려는 시도는 적다는 것을 알 수 있다.

나머지 ‘C2. 데이터의 비판점과 개선 방향’은 빈도수도 각각 0회, 2회로 대단히 적고 그 차이도 유의하지 않은 것으로 나타났다. 이는 학습자들이 데이터 자체의 비판점과 개선 방향에 대해서 서술하려는 의도는 많지 않거나, 이러한 접근을 할 줄 모르는 것으로 파악된다.

특히 C 이동 마디의 단계 유형은 그 관념적 메타 기능이나, 데이터 분석과 관련된 다양한 지식을 알고 있어야 비판이나 제안이 가능하다는 점에서 언어권보다는 숙달도의 영향이 우세한 것으로 생각된다. 위에서 중국어권 학습자들이 술직 시험을 통해 모어로 글을 구성하는 방식에 대해 이미 알고 있음을 지적한 바 있다. 따라서 중국어권 학습자들이 언어의 영향으로 C 이동 마디 유형을 누락하지는 않았을 것이라 생각한다. 다만 중급 학습자들이 데이터 설명문을 아예 C이동 마디의 단계가 없는 것으로 생각해 이를 기술하려는 생각조차 할 수 없는 것은 데이터 설명문 텍스트 구조를 조직해 본 경험이 적기 때문이거나 한국어 데이터 설명문 작성 교육을 받았을 때 해석을 하는 이 부분이 누락되었기 때문이 아닌가 사료된다. 특히 글쓰기 교육에서의 불분명한 위상 때문에 데이터 설명문 작성 교육 내용은 그다지 많이 개발되지 않았다. 따라서 학습자들이 주로 데이터 설명문을 학습하는 것은 주로 TOPIKⅡ 53번 쓰기 문제의 유형별 구조를 통해서이다. 그러나 이 문제 유형은 분량의 제한으로 저자의 해석을 쓸 공간이 없어 C 이동 마디 전부를 삭제하는 방향으로 텍스트 구조가 조직된다. 그러나 학술적 데이터 설명문 작성에서 C 이동 마디와 그 하위 단계들은 반드시 필요한 부분이며, 학습자들이 TOPIK

시험에서 더 나아가 학술적으로 구조를 파악하고 교정할 수 있도록 교육 방향이 제시되어야 할 것이다. 이들 분석을 정리하면 다음과 같다.

첫째, 한국인과 외국인 대학생의 C 이동 마디들은 세 개 모두 한국인이 유의하게 많이 사용한 것으로 나타났다. 이는 C 이동 마디 자체가 한국인 대학생에게 더 많이 생산되었기 때문이다. 이에 C1과 C3의 분포도 한국인에게서는 고정성이 나타났고, 특히 C1은 99%의 분포도를 보였지만, 외국인 대학생의 경우에는 세 단계 모두에서 분포가 60% 미만으로 고정성을 보이지 않았다.

둘째, 베트남어권 학습자들은 중국어권 학습자에 비해 C1 사용 빈도 평균이 유의하게 높았다. 베트남어권 학습자의 C1 분포가 80%가 넘는 것으로 보아 대부분의 베트남어권 학습자는 C1을 고정적으로 사용하고 있다.

셋째, 고급 학습자는 중급 학습자에 비해 C1과 C3에서 그 빈도 평균이 유의하게 높았다. 또한 고급 학습자의 C1 분포도 69%로 다수의 고급 학습자가 C1을 골고루 사용한다는 것을 알 수 있다.

2. 학습자 학술적 데이터 설명문의 어휘문법 사용

2.1. 주제 및 방향 요약(A)

(1) (소)제목을 활용한 주제 제시(A1)

2절에서는 앞에서 살펴본 이동 마디 단계 내에서 사용되는 어휘문법에 대한 분석을 진행할 것이다. 각 이동 마디 단계별 어휘문법의 사용이 한국인 대학생과 외국인 대학생, 외국인 대학생 집단 내에서도 언어권과 숙달도에 따라 어떻게 달라지는지 이를 세부적으로 파악해 제시할 것이다. 이때 어휘문법 목록은 언어권별 분류를 우선으로 하고, 이후 숙달도 별로 제시한다.

‘(소)제목을 활용한 주제 제시(A1)’ 이동 마디는 제목과 소제목으로 이루어졌기 때문에 말뭉치 양이 매우 적어 n-gram 분석을 하여도 특정한 어휘문법이 추출되지는 않았다. 특히 제목과 소제목은 실험에서 제시한 작문 과제의 제목을 그대로 딴 ‘코로나 시기 성별 가사 노동 시간 연구’류가 가장 많았고, 그다음으로는 과제에서 제시한 표와 그래프의 제목인 ‘설문 응답자 기본 정보’, ‘맞벌이 상태별 하루 가사 노동 시간’, ‘남성과 여성의 하루 평균 가사 노동 시간 비교’를 사용한 경우가 대다수였기 때문에 어휘나 구성의 특이성이 발견되지 않았다. 따라서 연구 방법에서 잠시 언급했듯 명사군 구성으로 되어 있는 경우가 대다수이고 그 수가 많지 않기 때문에 (소)제목들을 그대로 가져와 분류해 보았다. 또한 연구자가 과제로 제시한 표와 그래프 제목을 그대로 사용한 경우가 많아 중복되는 것이 많을 때에는 하나만 표기하였다.

<표 58> A1 언어권별 어휘문법 목록

| 이동 마디와 단계 | | 언어권 | | |
|-----------|------------------|---|--|---|
| | | 중국어권 | 베트남어권 | 한국인 |
| A | 주제 및 방향 요약 | | | |
| A1 | (소)제목을 활용한 주제 제시 | <ul style="list-style-type: none"> • 코로나 시기 성별 가사 노동 시간 연구 • 코로나 시기 성별 가사 노동 시간 연구 • 코로나 시기 성별 가사 노동 시간 연구 보고서 • 언제나 여자 가사 노동량 최다 | <ul style="list-style-type: none"> • 코로나 시기 성별 가사 노동 시간 연구 • 코로나 시기 성별 가사 노동 시간 연구 보고 • 코로나 시기 성별 가사 노동 시간 연구 보고서 • 설문 응답자 기본 정보 • 맞벌이 상태별 하루 가사 노동 시간 분석 • 남성과 여성의 하루 평균 가사 노동 시간 비교에 따른 분석 • 결론 | <ul style="list-style-type: none"> • 서론 • 설문 대상 선정 • 설문 응답자의 특성(에 따른 분석) • 행태 분석 • 본론 • 원인 추정 • 기혼 가정의 맞벌이 상태별 하루 가사 노동 시간의 일반적인 경향 • 맞벌이 상태별 하루 가사 노동 시간 비교 • 결론 • 남성과 여성의 하루 평균 가사 노동 시간 비교에 따른 분석 • 남성과 여성의 하루 평균 가사 노동 시간의 일반적인 경향 |

- 코로나 전후 남성과 여성의 하루 가사 노동 시간 변화 추이
- 코로나 이후 남성과 여성의 하루 가사 노동시간의 변화
- 코로나 전후 하루 평균 가사 노동 시간의 차이
- 코로나 시기 성별 가사 노동 시간 연구
- 코로나 시기 성별 하루 평균 가사 노동 시간 분석
- 코로나 시기 성별에 따른 가사 노동의 변화
- 부부 가사 노동 부담 여성에게 치우치는 경향이 강하다!
- 여성은 가사 돌봄의 주체가 아니다라고 하는 인식 확장되었지만 가사 시간 2.5배는 그대로!
- 연구 결과 보고
- 남성과 여성의 평균 가사 노동 시간의 이해는 과연 이루어질 수 있을 것인가?

<표 59> A1 숙달도별 어휘문법 목록

| 이동 마디와 단계 | | 숙달도 | |
|-----------|------------------|--|--|
| | | 중급 | 고급 |
| A | 주제 및 방향 요약 | | |
| A1 | (소)제목을 활용한 주제 제시 | <ul style="list-style-type: none"> • 코로나 시기 성별 가사 노동 시간 연구 • 코로나 시기 성별 가사 노동 시간 연구 • 코로나 시기 성별 가사 노동에 따른 연구 • 코로나 시기 성별 가사 노동에 따른 연구 결과를 보고 • 코로나 시기 성별 가사 노동에 따른 연구 결과 보고 • 코로나 시기 성별 가사 노동 시간 연구 보고서 • 설문 응답자 기본 정보 • 남성과 여성의 하루 평균 가사 노동 시간의 교에 따른 분석 • 결론 • 맞벌이 상태별 하루 가사 노동 시간 분석 | <ul style="list-style-type: none"> • 코로나 시기 성별 가사 노동 시간 연구 • 코로나 시기 성별 가사 노동 시간 연구 결과 보고서 • 남성과 여성의 하루 평균 가사 노동 시간 비교에 따른 분석 • 설문 응답자 기본 정보 • 맞벌이 상태별 하루 가사 노동 시간 분석 • 남성과 여성의 하루 평균 가사 노동 시간 비교 • 언제나 여자 가사 노동량 최다 • 결론 |

한국인 대학생의 말뭉치에서 발견된 A1 이동 마디 단계에서의 소제목 개수는 총 38개이며, 중복된 것을 제외하면 21개의 제목이 추출되었다. 이 목록에서는 ‘코로나 시기 성별 가사 노동 시간 연구’가 빈도수 5회로 가장 많았다. 중국어권 학습자의 경우 총 9건의 제목이 수집되었으나 ‘코로나 시기 성별 가사 노동 시간 연구’가 6건 중복되어 4건의 제목만을 수록하였다. 베트남어권 학습자는 총 42개 제목에서 중복되는 것들을 처리하고 7개의 제목만을 사용한 것을 확인할 수 있다. 최대 중복도를 보인 제목은 ‘코로나 시기 성별 가사 노동 시간 연구’로 14개가 출현했다.

고급 학습자 말뭉치에서는 총 28개 제목이 출현하였으나 중복된 것을 빼면 8개에 불과하였다. 여기에서는 ‘코로나 시기 성별 가사 노동 시간 연구’가 11개로 최대 중복도를 보였다. 중급 학습자는 24개 중 중복되지 않는 것만을 모았을 때 10개가 남았으며, ‘코로나 시기 성별 가사 노동 시간 연구’가 10개로 최대 중복되는 것을 볼 수 있다. 따라서 외국인 대학생들은 데이터 설명문을 자료에 종속되어 있는 것으로 보는 경향이 강하며, 이에 데이터 설명문의 제목과 자료의 제목을 동일시한다. 한국인 대학생들은 이에 비해 데이터 설명문의 제목은 시각적 데이터 제목과는 달리 붙이는 것을 발견할 수 있다. 이는 한국인 대학생들이 시각적 데이터와 데이터 설명문을 별개로 인식하고 있으며, 데이터 설명문을 구성할 때는 새로운 목적과 의도를 지닌 담화 구조를 만들어냄을 뜻한다.

목록에서 확인한 (소)제목의 어휘문법적 특성으로는 거의 모든 학생들의 제목에서는 명사 나열 구성이 대부분이었다. 이윤진(2016:248)에서는 학술적 텍스트의 제목이 다른 장르의 텍스트 제목과는 달리 학문 공동체의 관습 및 규정에 적합한 형식과 내용을 담보해야 하는 차별성이 있다고 하였다. 이에서는 객관성이 필요한 한국어 학술적 텍스트는 대부분의 제목이 명사로 끝맺어짐을 지적하고 있다. 데이터 설명문이 학술적으로 사용되는 상위 장르를 생각해 보았을 때 실험 보고서나 학술지 논문, 학위 논문 등으로 객관성을 띠는 경우가 다수이기에 이들 장르에 나타나는 경우 명사형으로 끝맺는 제목 사용이 적절하다 할 것이다. 이 때문에 학술적 데이터 설명문의 제목은 (ㄱ), (ㄴ)과 같이 명사로 끝나는 것이 일

반직임을 확인할 수 있었다. 이는 과제 제목이나 그래프 제목이 명사형 구성인 것에도 영향을 받았다 볼 수 있다.

(ㄱ) 설문 응답자 기본 정보

[6-한국어-경제4-여]

(ㄴ) 맞벌이 상태별 하루 가사 노동 시간 분석

[79-고급-베트남어-경영4-여]

반면 아래 (ㄷ)은 형용사로 문장을 종결하며, (ㄹ)은 ‘언제나’와 같이 부사로 수식하고, 객관적 수치가 아닌 평가적 어조의 ‘최다’라는 용어를 사용하기 때문에 학술성에 맞지 않다고 볼 수 있다. 이러한 구성은 학술적 텍스트의 제목이기보다는 신문 기사의 그것과 더 어울린다.

(ㄷ) 부부 가사노동 분담, 여성에게 치우치는 경향이 강하다

[15-한국어-경영1-남]

(ㄹ) 언제나 여자 가사 노동량 최다

[99-고급-중국어-경제2-여]

전반적으로 외국인 대학생은 제목 사용률이 적고, 제목으로 사용한 것 역시도 과제로 제시한 표와 그래프의 제목을 그대로 쓴 경우가 대다수였다. 따라서 중복된 제목이 많아 그 수가 적어 보인다. 중급 학습자의 경우에는 거의 대부분 표와 그래프 제목을 차용해 제목을 썼고, 그마저도 많지 않았지만, 고급 학습자의 경우에는 표와 그래프 제목에 ‘보고서, 분석’과 같은 어휘를 덧붙여 제목을 달리 구성하려는 시도가 보인다. 이는 한국인 대학생의 제목 사용에서도 나타나는 특징이다. 또한 외국인 대학생의 제목 사용에서는 아래의 예에서 볼 수 있듯, 오타가 생성된 경우도 보인다.

(ㄹ) 코로나 시기 성별 가사 노동 시간 연구[30-중급-중국어-경영3-여]

한국인 대학생의 경우 표와 그래프에서 제시한 제목을 조금씩 바꾸어 사용하는 경향이 있어 제시된 제목의 종류가 매우 다양하다. 또한 외국인 대학생들보다 소제목에 숫자를 사용해 구분하는 경향이 있다. 그럼에도 이들이 항상 외국인 대학생들보다 더 학술적인 제목을 붙이는 것은 아니다. 다음과 같은 비학술적 제목은 한국인 대학생들에게 더 많이 보인다.

(ㄴ) 여성은 가사 돌봄의 주체가 아니다라고 하는 인식 확장되었지만 가사 시간 2.5배는 그대로!

[16-한-경영1-여]

(ㄷ) 남성과 여성의 평균 가사 노동 시간의 이해는 과연 이루어질 수 있을 것인가?

[22-한-경영2-여]

위에서 언급한 (ㄷ)과 여기에서 다루는 (ㄴ)의 경우는 다른 장르의 표절 때문에 학술성에 문제가 생긴 것으로 볼 수 있다. 즉, 표절 대상 텍스트가 학술적 텍스트가 아니기 때문에 학술성에서 문제가 생긴 것이다. 이는 비단 장르성의 문제뿐만 아니라 글쓰기 윤리성에서도 문제가 있다고 보인다.

<표 60> 표절 의심 작문과 그 대상

| 표절 의심 작문 | 표절 대상 |
|----------------|---|
| [15-한국어-경영1-남] | 불평등 연구회, “부부 가사노동 분담, 무엇이 문제인가?” |
| [16-한국어-경영1-여] | 한겨레 신문, “‘여성≠가사·돌봄’ 인식 확장됐지만... 가사 시간 2.5배는 그대로”, 박고은 기자, 2022.4.19. 기사 |

다만 (ㄷ)과 같은 의문형 구성의 제목은 논의의 여지가 있다. 이슬비 (2016:129)에서는 학술적 텍스트에서 독자 유도를 위한 상호작용적 태도

로서 의문문이나 청유문을 사용한다고 하였다. 또한 이윤진(2016:259)에서도 의문형 제목이 독자에게 질문을 던져 저자의 문제 제기를 독자와 공유하고자 하는 전략을 띠고 있다고 하였다. 그러나 의문형 제목이 학술적 텍스트에서 사용되었다 하더라도 이것이 효과적인 전달 방식인가 또는 장르 내포적으로 존재하는 데이터 설명문에서 출현하는 것이 허용되는가는 또 다른 문제이다. 단정적으로 의문문이 대학생의 데이터 설명문 과제에서 학술성 여부를 드러낸다고 판단하기는 어려우나 학습자 작문 전반을 살펴보았을 때 하나밖에 없는 매우 드문 경우라고는 할 수 있다.

(2) 연구 주제와 내용 및 방법 설명(A2)

‘연구의 주제와 내용 및 방법 설명(A2)’은 주로 데이터 설명문을 시작하는 가장 첫머리에 등장한다. 이 데이터 설명문에서 알아보하고자 하는 것이 무엇인지, 어떠한 연구를 어떻게 수행하였는지, 데이터 수집 기관이나 주체가 누구인지 등을 한두 문장 정도로 간략하게 제시하는 경우가 많다. 전체 데이터 설명문의 도입 문장이 대부분 이 A2에 속하며, 위에서 확인한 바대로 대부분 집단이 A2 이동 마디 단계 분포도 60%를 넘는다. 또한 A2 단계는 학습자 집단 간 사용 빈도 차이가 크지 않은 단계이기도 하다. 즉, 거의 모든 대학생이 데이터 설명문을 구성하기 위해 유사한 빈도로 사용한다.

A2는 설문조사의 주관 기관을 인용하기 위한 ‘따르다’와 공기하는 어휘 문법적 표현을 모든 학습자군에서 사용하는 것을 알 수 있다. 그다음으로는 연구 대상을 언급하며 ‘-를 대상으로’를 사용하며, 연구 주제 언급에 ‘-에 대하여’를 사용한다. 마지막으로 연구 수행의 방법을 ‘실시하다, 통하다, 조사하다, 분석하다’ 등의 어휘 공기 표현을 통해 제시하고 있다. 위에 제시한 어휘문법들로 구성해 보자면, ‘이 연구는 (수치)명을 대상으로 가사 노동 시간에 대한 설문조사를 실시하였다.’ 와 같이 사용됨을 알 수 있다.

<표 61> A2 언어권별 어휘문법 목록

| 이동 마디와 단계 | | 언어권 | | |
|-----------|-----------------------|---|---|---|
| | | 중국어권 | 베트남어권 | 한국인 |
| A | 주제 및 방향 요약 | | | |
| A2 | 연구의 주제와 내용 및 방법 설명 | <ul style="list-style-type: none"> • (수치)년에 • (수치)명을 • (수치)-이/-이면 • 각 (수치)명 • 결과이다 • -에 대해- 조사/연구 • 조사/연구 -를 실시하-/-를 하였-/-이다 • 명을 대상으로 • 하여 조사- • 하였다 | <ul style="list-style-type: none"> • (수치)년 -부터/-까지/-에 • -에 대해-/따르- • 정보-에 관하-/-에 대해-/-와 관련 • -는 결과를 • -를 보고하/분석하/실시하-/통하-/활용하-/포함하-/알아- • 분석/비교/설명-(하고자) 한다 • -기 위하- • -고 이것이 • 이 연구는/보고서는 • 설문조사를 • 시간을 알아 보- • 조사하- • 데이터를 활용하-/-분석하고자 하- • -를 대상으로 N • 중이었 • 진행된 | <ul style="list-style-type: none"> • -고 있- • -고자 하- • -는/을 것- • -와 같- • -에 대해-/관하-/따르- • -를 분석하-/-위하-/비교하-/조사하-/실시하/대상으로- • 되었다 • 자료를 V 하였- • 해 보- |

- 해 보-
- 하였다

<표 62> A2 숙달도별 어휘문법 목록

| 이동 마디와 단계 | 숙달도 | |
|-----------|--|---|
| | 중급 | 고급 |
| A | 주제 및 방향 요약 | |
| A2 | <p>연구의 주제와 내용 및 방법 설명</p> <ul style="list-style-type: none"> • (수치)년-에/통계청 • (수치)명-을/-이 • 각 (수치)명 • -고자 한다 • -에 대해-/따르- • 관련된 • -를 대상으로 설문/조사 • -어 조사를 • -를 실시하-/대상으로/분석하-/하였- • 시간에 대해- | <ul style="list-style-type: none"> • -하여 조사를 실시 • 하였다 • -를 대상으로/알아 보- • 이 연구-는/-를/-에 대- • 결과를 분석/실시/참고 • -고자 하- • 관련되- • -는 설문 응답자/조사 • 데이터-를/-부터 분석하-/활용하-/종합하- • 따르- • -를 실시하-/위하-/통하-/비교하-/대상으 로 하였- • 있었다 |

한국인 대학생 및 외국인 대학생이 공통적으로 사용하는 어휘문법은 ‘-에 대한/관한’을 들 수 있으나, 외국인 대학생의 작문에서 ‘-에 관한’을 찾아보기는 어려웠다. ‘한국어 교육 어휘 내용 개발 4단계(2015)’에서 제시한 등급별 어휘 목록에서 ‘관하다’가 고급에 속하는 어휘이기 때문에 숙달도가 높지 않은 학습자들은 접근하기 어려울 것이라 생각할 수 있다. 또한 ‘조사’라는 단어가 모든 집단에서 출현하는데, 이는 시각적 데이터에 제시된 ‘설문 응답자 기본 정보’ 표에서 보여지는 자료들이 ‘설문 조사’로 획득된 데이터라는 사실을 파악하여 사용하는 것으로 생각된다.

또한 외국인 학습자들의 어휘문법에서는 여러 구성의 결합형이 다수 출현한다. 3-gram에서는 각기 따로 등장했으나, 다른 n-gram 구성 확인을 통해 이들이 특정 구를 형성하는 요소들이 있음을 확인하였다. 예를 들어, 중국어권 학습자들은 ‘-에 대해- 조사/연구, 조사/연구 -를 실시하-/-를 하였-/-이다’의 어휘문법이 ‘-에 대한 조사를 하였다/-에 대하여 연구를 실시하였다/-에 대한 연구이다’ 등으로 다양한 구성을 보였다. 베트남어권 학습자들은 ‘분석, 보고, 설명, 실시’와 같은 어휘 사용이 눈에 띈다. 한국인 대학생에 비해서도 이러한 용어 사용은 두드러진 것임을 알 수 있다. 숙달도에 따라서 외국인 대학생들의 어휘문법 사용 양상은 특히 분석적 어휘를 사용하는 데서 차이가 나타난다. 중급 학습자에게서는 확인할 수 없는 ‘비교하-, 활용하-, 종합하-’와 같은 어휘 사용이 나타나는 점에서 어휘 레퍼토리가 중급 학습자보다 넓음을 확인할 수 있다. 그러나 주제와 내용을 언급할 때는 중급과 고급 학습자 모두 ‘대하다’, ‘대상’, ‘조사’와 같은 공통적 어휘를 사용한다.

이들 외국인 대학생들의 A2 말뭉치에서 공통적으로 발견된 어휘문법적 표현은 전형적으로 TOPIK 문제지 등에서 제시하는 TOPIK 53번 문제의 예시 답안에서 자주 사용하는 것이다. 이태환(2019:133)에서는 순위 그래프를 쓰는 문단 자체를 다음과 같이 고정하여 제시하고 있다.

(조사 기관)에서 (대상)을 대상으로 (주제)에 대하여 조사를 하였다. 그 결과 (비교 대상 1)은 (무엇)이 (몇)%로 가장 많았다. 그 다음으로 (무엇)과 (무엇)이 뒤를 이었다. 반면 (비교 대상 2)는 (무엇)이 (몇)%로 가장 높게 나

타났다. 그 다음으로 (무엇), (무엇) 순이었다.

위 문제집에서의 예상 답안에서는 ‘대상’, ‘-에 대하여’, ‘조사’와 같은 단어가 첫 문장에 들어가 있는 것을 볼 수 있다. 한국인 대학생에 비해 이러한 단어 사용이 많은 것은 외국인 대학생들이 데이터 설명문을 작성할 때 자신이 학습했던 TOPIKⅡ 53번 글쓰기의 구조와 어휘를 이용함을 의미한다.

한국에서 대학을 다니는 외국인 유학생들은 졸업하기 전까지 TOPIKⅡ 점수가 있는 것이 보통이며, 교육과정상에는 그래프 설명문이 제시되지 않지만, TOPIK 시험을 준비하는 과정에서 이와 같이 고정된 형식으로 데이터 설명문을 작성하는 교육을 받았기 때문에 공유된 사용 양식을 재생산하는 것으로 보인다. 이는 보통 TOPIK 시험을 준비하지 않는 한국인 대학생들의 이동 마디 단계의 어휘문법과 비교해 보았을 때 차이가 두드러지는 부분이다.

개별 어휘 사용을 보았을 때 베트남어권 학습자군, 고급 학습자군, 한국인 대학생군 모두 유사한 사용을 보였고, 중국어권 학습자군 및 중급 학습자군의 어휘문법적 사용과 비교했을 때는 그 구성이 좀 더 다양함을 알 수 있다.

(3) 데이터 현상에 대한 배경 설명(A3)

‘데이터 현상에 대한 배경 설명(A3)’은 주제와 관련된 배경에 대해 학습자들이 서술한 단계를 포착한 것이다. A3은 주어진 표와 그래프에서 다루는 현상이 발생한 사회적, 맥락적, 환경적 요인을 제시한다. 예를 들어, 코로나 바이러스 감염자가 늘어나고 있다는 것, 이 때문에 사람들이 집에서 머무는 시간이 길어졌다는 것, 그 시간 동안 가사 노동의 절대량이 늘어났다는 것, 전통적으로 가사 노동은 여성의 부담이었다는 점 등을 이 단계에서는 다루고 있다.

앞에서 다룬 바 있듯, 이 단계는 한국인 대학생에 비해 외국인 대학생이 거의 작성하지 않는다. 이에 대한 원인을 한국인 대학생들과 외국인

대학생들이 ‘학술적’ 상황에서 가장 근접하게 경험한 데이터 설명문의 목적과 서술 방향이 달랐기 때문이라 진단한 바 있다. 즉, 한국인 대학생, 특히 전공 수업을 많이 듣지 않은 저학년 학생들은 대입 논술과 같이 데이터의 수치를 실제 사회와 관련지어 맥락적으로 해석하는 데이터 설명문을 가장 최근에 겪어보았고, 외국인 대학생들은 시각적 데이터의 내용을 정확히 이해하고 한국어로 전달하는 TOPIKⅡ 53번 문제가 가장 최근 겪은 데이터 설명문이다. 이 때문에 글쓰기 구조와 양식이 명확히 제시되지 않은 상황에서는 자신이 경험했던 것 중 가장 근접해 보이는 작문의 양식을 따르려는 경향이 있음을 알 수 있다. 따라서 A2의 경우에는 TOPIK 답안 가이드를 따르는 어휘문법이 빈번하게 눈에 띄을 위에서 본 바 있다.

그러나 A3은 TOPIK에서는 요구되지 않는 것이라 이를 선택한 외국인 학습자의 최근 데이터 설명문이 TOPIK 외의 다른 학술적 데이터 설명문임을 추론할 수 있다. 다만 전체적인 단계의 선택 빈도수가 많지 않아 언어권별, 숙달도별 어휘문법 빈도가 높지 않다.

A3 단계에서 한국인 대학생은 다른 학습자군에 비해 다양한 어휘와 구성을 사용하고 있다. 즉, 코로나 상황이 성별 가사 노동에 어떠한 영향을 행사했는지 논하기 위해 그 배경에 대한 설명을 비교적 다양하고 하고 있는 것이다. 특히 가정 체류 시간이 증가해 가사 노동의 부담이 늘었다는 내용을 주로 제시해서 그러한지 ‘증가하-, -가 늘어나-, -는 시간이’와 같은 어휘문법이 눈에 띈다. 이는 (ㄱ)과 같은 문장 구성에서 추출된 것이다.

(ㄱ) 최근 들어 코로나 바이러스로 인해 자가 격리, 재택 근무 등 집에 머무는 시간이 늘어나면서 성별에 따른 가사 노동 시간에 변동, 차이가 있는 것으로 보인다.

[47-한-경영2-여]

<표 63> A3 언어권별 어휘문법 목록

| 이동 마디와 단계 | | 언어권 | | | |
|-----------|---------------------|---|--|---|---|
| | | 중국어권 | 베트남어권 | 한국인 | |
| A | 주제 및 방향 요약 | | | | |
| A3 | 데이터 현상에 대한 배경 설명 | <ul style="list-style-type: none"> • -고 있다 • 많아지- • 사람들- • -에 있는 • -를 더 • 하였- | <ul style="list-style-type: none"> • 우리-는/의 • 모든 N -의/-이 | <ul style="list-style-type: none"> • 것이다 • -고 있- • -는 시간이 • -에 따르- • 되었다 • -을 것-/정도로 많 • - | <ul style="list-style-type: none"> • 비하여 • -가 늘어나- • 있는 시간 • 증가하- • -지 않- • 진행되- • 하여- |

<표 64> A3 숙달도별 어휘문법 목록

| 이동 마디와 단계 | | 숙달도 | |
|-----------|---------------------|--|---|
| | | 중급 | 고급 |
| A | 주제 및 방향 요약 | | |
| A3 | 데이터 현상에 대한 배경 설명 | <ul style="list-style-type: none"> • -게 되- • 사람들- • 있는 시간이- | <ul style="list-style-type: none"> • -고 있- • 더 많이 V |

시간이 흐름에 따라 여성의 사회 진출이 늘어나고 전통적 성역할에서의 탈피 움직임이 있다는 시대의 변화를 언급하며 ‘-에 따르-, -고 있-’과 같은 어휘문법을 사용한다.

(자) 이에 따라 가정 내에서 남성의 가사 노동 시간이 증가하게 되는 실질적인 변화와 함께 가정의 구성원으로서 모두가 분담해야 하는 일이라는 인식적 변화가 나타나게 되었다.

[11-한 경제4-여]

그에 비해 외국인 대학생들의 경우에는 사용한 어휘문법적 항목의 개수가 상대적으로 적고, 사람들이 가정에 머물러 있게 된 상황을 주로 언급하는 것으로 보인다.

(ㄷ) 사람들은 바이러스를 두려워하고, 건강을 위해 외출하는 것을 줄이고, 집에 있는 시간이 길어지면 자연스럽게 가사가 많아진다.

[41-중급-중국어-경제2-여]

특징적인 것은 베트남어권 학습자의 ‘우리’라는 표현이다. 다음 예문 (ㄷ)과 (ㄹ)을 보면 베트남어권 학습자인 저자는 문장의 주어인 자신을 1인칭 복수인 ‘우리’로 표현하고 있다.

(ㄹ) 우리는 왜 그런 변화가 일어날까 매우 궁금하다.

[76-고급-베트남어-경영3-남]

(ㄹ) 여기서 우리는 남성과 여성의 가사 노동 시간의 변화를 들 수 있다.

[103-고급-베트남어-경영3-여]

학술적 텍스트를 구성할 때 필자가 자기 자신을 표현하는 이른바 ‘필자 지칭 표현’은 주로 3인칭을 사용하여 텍스트와 필자를 분리한다. 이는 텍스트 내용에 객관성을 부여하려는 목적이다. 그러나 1인칭 복수인 ‘우리’

의 사용은 ‘이 글’과 같이 대상화해서 표현할 수 있다면 예외적으로 사용된다(이윤진, 2014). 또한 연구 내용의 일반화를 위해 저자와 독자를 공동체화하는 ‘우리’라면 이 또한 허용 가능하다고 하였다(김한샘 외, 2017:111).

따라서 ‘우리’가 허용적인 상위 장르라면 이러한 사용이 문제될 것은 없어 보인다. 그러나 (ㄴ)과 (ㄹ)에서 사용된 ‘우리’는 이윤진(2014)와 김한샘 외(2017)에서 말하는 ‘우리’와는 조금 다른 것으로 생각된다. 저자의 복수화를 나타내는 것이기 때문이다. 같은 형태를 가진 ‘우리’이지만 이들이 허용되는 특정 조건이 있기 때문에 고급 학습자이고 3학년인 학생들도 이러한 모호한 부분을 구분하지 못하는 경우가 있음을 알 수 있다. PHAN, HOANG MY THUONG(2020:82)에서는 베트남 학술적 담화 공동체에서는 개인보다 집단을 중요시하기 때문에 ‘I’보다 ‘We’에 해당하는 1인칭 대명사를 선호한다고 하였다. 이러한 측면을 보았을 때 ‘우리’의 사용은 베트남어권 학습자의 모어 전이에 의한 것일 수 있다.

반면 한국인 대학생의 A3 단계에서의 어휘문법 사용에서는 ‘우리 사회’, ‘우리 주변’과 같은 ‘우리-’ 결합 어휘는 종종 등장해도 필자 지칭 표현으로서 ‘우리’를 사용하는 경우는 1회에 머물렀다. 다음 (ㄹ) 예문을 보면, 1학년 한국인 대학생이 학술적 문체에 익숙하지 못해 필자 지칭 표현을 ‘우리’로, 종결어미를 구어로 사용하는 것을 볼 수 있다.

(ㄹ) 코로나19 팬데믹이라는 유례없는 급변화된 상황 속에서 자녀의 돌봄 공백을 채우기 위한 여성들의 고민이 늘어났고, 이러한 상황에서 우리는 여성들의 가사노동에 대한 부담을 인지하고, 이를 어떻게 해소해야하는지 진지한 고찰이 필요하게 되었습니다.

[23-한-경제1-남]

저자와 독자를 공동체화하는 ‘우리’의 사용은 상위 장르에 따라 허용 가능할 것이다. 그러나 이 ‘우리’의 사용은 23번 대학생에게는 저자와 독자의 공동체화라는 기능의 실현보다는 문체적 미숙성에 의한 것임을 알 수 있다. 또한 PHAN, HOANG MY THUONG(2020:61)에서는 베트남어권

학습자가 학문 목적 글쓰기에서 구어체를 사용하는 이유를 베트남어의 종결어미가 구어/문어를 구분하는 역할을 하지 못해 학습자들이 이를 파악하지 못하여 그렇다고 하였다. 그러나 구어체 사용은 언어권의 영향도 있겠지만, 한국인 대학생들과 베트남 외 기타 언어권 학습자도 자주 나타내는 양상이기 때문에 이는 언어권을 불문하고 학술적 텍스트에 요구되는 문어 사용을 잘 파악하지 못해 발생한 것이라 볼 수 있다.

2.2. 수치/결과 나열(B)

본격적으로 데이터 설명문의 특성을 드러내는 이동 마디인 ‘수치/결과 나열(B)’에서는 ‘구체적인 정보 추출(B1)’과 ‘의도적 그룹화(B2)’의 두 하위 이동 마디에 따라 하위 단계의 특성을 확인해 보도록 한다. 이들 이동 마디의 특징은 데이터를 직접 언급하여 수치와 단위 명사를 다수 사용한다는 것이다. 대학생들의 자세한 어휘문법 사용은 다음을 보며 확인 하겠다.

(1) 구체적인 정보 추출(B1)

우선 B1-1은 연대기적 나열로 모든 집단에서 적게 사용하였다. 따라서 이 단계의 말뭉치가 매우 작기 때문에 2회 이상 등장하는 어휘문법도 매우 적었다. 데이터를 시간 순서대로 나열하는 B1-1은 시각적 데이터를 보이는 순서대로 나열하는 B1-2와 더불어 작문에 소요되는 인지적, 언어적 노력이 크지 않은 경우가 대부분이다. 시계열적으로 제시된 시각적 데이터라면 그 시간적 순서대로 왼쪽에서 오른쪽으로 서술하면 되기 때문이다.

실제로 실험에 제시한 <그림 2>의 데이터를 해석할 때 시각적 데이터를 눈에 보이는 순서대로 여성의 가사 노동 시간 먼저, 남성의 가사 노동 시간을 나중에 제시하는 학습자와는 달리, 시계열적 정보에 주목하여 2019년, 2020년, 2021년에 남성과 여성의 가사 노동 시간 차이를 순서대

로 제시하는 학습자들은 추세를 설명하는 어휘를 사용하지 않아도 되고 단순히 '(수치)년, (변수 1)의 수치, (변수 2)의 수치'를 반복하면 된다.

B1-2 유형의 사용은 내용의 중요도나 작문 주제에 따라 데이터를 배열하는 것이 아닌 시각적 데이터를 단순히 '보고 쓰는' 유형에 속하므로 순서나 배열에 대한 인지적 부담에서 자유롭다. 그래서 제2언어 사용에 대한 부담이 더 큰 중급 학습자가 인지적 노력을 줄이기 위해서인지 자주 사용하는 경향이 있다. 시각적 데이터의 데이터에 대해 먼저 생각하고 조직과 배열을 고민하기보다는 자신의 언어 가용 정도에 맞추어 빠르게 기술해 버리는 것이다.

이러한 인지적인 장점 이외에도 논의의 전형적인 장을 형성하려는 관념적 의미로 B1-1이나 B1-2를 사용하는 경우도 있다(이지연, 2023:165). 특히 한국인 대학생들은 이들 이동 마디 단계를 통해 자세한 정보를 제시하고 그 뒤에 중요한 정보를 다시 골라 분석적으로 제시한다. 이는 이지연(2023)에서 발견한 경제학 및 경영학 학술지 전문 필자들이 채택하는 방식이기도 하다. 따라서 이 두 단계는 인지적으로든 언어적으로든 낮은 난도를 지녀 낮은 숙달도의 학습자가 쉽게 접근할 수도 있고, 높은 숙달도의 학습자도 더 자세한 데이터 해석의 배경으로서 흔히 사용한다고 할 수 있다.

<표 65> B1 언어권별 어휘문법 목록

| 이동 마디와 단계 | | 언어권 | | |
|-----------|-----------------|---|--|---|
| | | 중국어권 | 베트남어권 | 한국인 |
| B1 | 구체적인 정보 추출 | | | |
| B1-1 | 연대기적 나열 | <ul style="list-style-type: none"> (수치)년 (수치)분 | <ul style="list-style-type: none"> (수치)분-으로/-에서 (수치)년-에(는)/-까지/-부터 증가하였- | <ul style="list-style-type: none"> (수치)시간-으로/정도/-은/-을 /-이 (수치)분-이- /늘었- /-으로/-을 V/-만 (수치)년 -에는/-은/-이/때/기준/-보다 증가하- 투자하- 하였- |
| B1-2 | 시각적 데이터 순서대로 나열 | <ul style="list-style-type: none"> (수치)명 중/-이-/-이-/-으로/-을/-을 있/-으로 가장/-으로 나타나/-까지/노동/미만/부터/사이 /-에/-에서/-은/이내/이상 (수치)분 -이-/-으로/-을/-으로 나타나/-으로 드러나/-을 차지 -/-이-/-정도 (수치)년 -에/-에는 | <ul style="list-style-type: none"> (수치)명-으로/-을 차지 /-이-/-은/-이- (수치)분-이-/-을/-으로/나 오-/-이- (수치)시간-이 (수치)년-에는 -것으로 나타나-/보이- 결과에 따르- | <ul style="list-style-type: none"> (수치)시간 -으로/미만/-은/-을/-이/이상 (수치)분-으로/-이-/-의/-이/정도/-을 (수치)명-으로/-이 -ㄴ 것-/반면/차이를 -에 사용하였-/투자하였-/소요하였- |

| | | | |
|--|---|--|---|
| | <ul style="list-style-type: none"> • 가정-에서(는)/-의 경우 • 같은 경우-는/N • -ㄹ/을 경우 남성/여성 • -ㄴ 것으로 • -으로/로 나타났- • -ㄴ/은 남녀가/남성이/사람은/사람이 • 났다 • 씁니다 • -었지만 N • -에/은 (수치)명/분 • 하는 가정/경우/시간/외별이 • -에 대하여/불과하- • 상태-가/-는/별/-에 대 • 하였다 • -를 차지하-/살펴보면/졸업하- /있다/하- • 드러났다 • -이었다 • 이 조사 결과 | <ul style="list-style-type: none"> • 경우- • 구성되어- • 상태(에서)는 • 설문(에) 응답(자) • 응답하- • 시간-은/을/이 • 쓴다고 응답 • -에 따르-/소비하-/대하- • -으로 조사되- • -을 차지하- • -이었- • 하-고 있-/였-/ㄴ 하루 • 대상으로 • 되었다 • 하였- • 조사에 참여 | <ul style="list-style-type: none"> • -의/-일/-인 경우- • 하였- |
|--|---|--|---|

• 많-

<표 66> B1 숙달도별 어휘문법 목록

| 이동 마디와 단계 | 숙달도 | |
|-----------|---|---|
| | 중급 | 고급 |
| B1 | 구체적인 정보 추출 | |
| B1-1 | 연대기적 나열 <ul style="list-style-type: none"> • (수치)분 -으로/-에(서)/증가/-이- • -에(는) (수치)분/증가하- • (수치)년-에(는)/-까지/-은 • -으로 나타났- • 감소하였- • 하였- • -이 더 길- | <ul style="list-style-type: none"> • (수치)분-∅ • (수치)년 -까지/-부터/-에(는) |
| B1-2 | 시각적 데이터 순서대로 나열 <ul style="list-style-type: none"> • (수치)시간-부터/-은/이상/-에/-까지/미만 • (수치)분-을/-으로/-이/-이-/동안/-에서/만 • (수치)명-에/-으로/-은/-을/-이/-이/중 • (수치)년-에(는) • -ㄴ 것으로 나타났/드러났-나왔- • -를 차지하-/따르-/살펴보-/하- • 그 다음에, 그 뒤에, 그 중에 • -은/는/ㄴ 경우 | <ul style="list-style-type: none"> • (수치)년-에(는) • (수치)시간-이/-으로/-을/이상/-에/-은 • (수치)분-씩/-이-/정도/-에/-으로/-을 • (수치)명-을/-의/-이/-으로/-은/-만-/-에(서)/-이- • -으로 가장 높-/구성되-/나타났-/보이- • -를 차지하- • 하였- |

| | | |
|--|--|--|
| | <ul style="list-style-type: none"> • -에 대하여/따르-/불과하-/비하-/도달하- • 살펴보면 N • 하였-/되었- • 높았다 • 있-다/-습니다 • 정보-이다/-에 따르-/에 관하- • 조사- 결과를/-에 참여- • 졸업하-ㄴ/는 사람 | <ul style="list-style-type: none"> • -에 불과하- • 일하는 • 경우-(에)는/N • -에 대해/불과하- • -ㄴ다고 응답하- • -고 있- • (더)많- • 되어/-이 있- • 응답한 • 조사되었 |
|--|--|--|

다만 학습자군별로 각 단계에서의 어휘문법적 사용에는 차이가 있었다. 중국어권 학습자의 경우에는 연대기적 나열을 사용하는 학습자가 다수였음에도 불구하고 공통적으로 등장하는 어휘문법적 표현은 ‘(수치)년(수치)분’ 하나밖에 등장하지 않아 보편적으로 사용하는 어휘문법적 표현이 아직 수렴되지 않은 것으로 보인다. 베트남어권 학습자와 한국인 대학생의 연대기적 이동 마디 코드에서는 중국어권 학습자와는 달리 ‘증가하-’와 같은 추세 관련 어휘를 사용하는 점이 눈에 띈다. 즉, 이 이동 마디를 단순 나열만으로 채울 수 있으나, 언어권별로 추세에 대한 인식이 함께 등장한다고 볼 수 있다. 다만 숙달도별 언어 사용에서는 또 이와는 반대로 중급 학습자에게서 추세 관련 어휘가 등장하고, 고급에서는 연도와 사용 시간(분)에 대한 어휘문법적 표현만이 등장한다. 다만 이는 앞에서 보았듯, 고급 학습자가 B1-1 단계를 선호하지 않아 선택한 빈도가 현저히 적어 그러할 수 있다.

‘시각적 데이터 순서대로 나열(B1-2)’ 단계는 이에 비해 많은 참여자들이 선택하였기 때문에 그 특성을 더 자세히 확인할 수 있다. 모든 학습자군에서 ‘(수치)+단위 명사’의 구분이 확인되었으며, 그에 후행하는 조사와 어휘문법이 다양함을 볼 수 있다. 그러나 이들 조사 사용이 모두 문법적으로 옳은 것은 아니다.

(ㄱ) 1시간~2시간은 402명을 있다.

[48-중급-중국어-경영1-여]

(ㄴ) 그리고 혼인 상태는 비혼자가 280명과 기혼자가 720명을 나뉘었다.

[고급-82-베트남어-경영2-여]

(ㄱ), (ㄴ)과 같이 서술어와 공기하는 조사를 잘못 사용하는 경우가 있어 ‘(수치)+단위 명사’ 다음에 등장하는 다양한 조사 사용에 대해서는 긍정적으로만은 볼 수 없을 것으로 생각된다.

또한 전반적으로 모든 학습자군에서 특정 수치가 밝혀진 것을 표시하는 ‘드러나-, 나타나-’가 빈번히 사용됨을 알 수 있다. 이는 특히 ‘-ㄴ/은/는

것으로 나타났다/드러났다'의 구성으로 사용되어 수치를 제시하는 방식에서 학습자군이 공유하는 형식이 있음을 나타낸다. 이에 더해 '경우'는 각 사건들을 구분 짓는 역할을 하는데 주요하게 사용됨을 알 수 있다.

하나 특수하게 발견되는 것은 중국어권 학습자의 어휘문법적 사용에서 '습니다'가 나타난 것이다. 학술적 텍스트 전반에서는 격식적 문어로서 '-(ㄴ)다' 종결어미를 사용한다(이슬비, 2016:66). 따라서 전공 과제로서 데이터 설명문을 작성할 때도 상위 장르의 영향을 받아 종결어미를 '-(ㄴ)다'로 구성하는 것이 일반적이다. 그러나 학술적 텍스트 작성에 대한 훈련이 되어 있지 않은 참여자들에게서는 격식적 구어인 '-습니다'체 종결어미가 다수 사용되는 것을 관찰할 수 있다. 이는 한국인 대학생에게서도 나타난다.

연구 참여자들은 전공 과제의 형식을 다른 학술적 텍스트보다는 더 자유롭게 생각하고 있을 가능성이 높다. 또는 실험 요건에서 작성하는 텍스트를 '전공 과제'라 명시하였음에도 불구하고 참여자들은 단일 독자인 연구자에게 말을 거는 형식으로 구어체를 사용하였을 가능성도 있다. 지정된 독자를 대상으로 할 때는 거리의 간격을 좁히기 위함인지 문어를 사용해야 하는 부분에서도 구어를 사용하는 경향이 있다. 그러나 웅웁티 후웅센(2015:108-109)에서는 문어 사용 상황에서의 구어 사용은 사회언어학적 기능을 충족하지 못한 한국어능력시험 하위 성적 집단의 특성이라 한 바 있다. 따라서 학술적 데이터 설명문 작성에서 구어 사용이 상위 장르적 특성에 비추어 볼 때 받아들여지기 어렵다는 점을 학습자에게 교수해야 한다.

(2) 의도적 그룹화(B2)

B2 단계에는 '수치가 높은/낮은 순서대로 나열(B2-1)', '수치적 비교(B2-2)', '특성에 따른 그룹화(B2-3)', '제일 큰/작은 수치에 주목(B2-4)', '특이성/중요 지점 지적(B2-5)'의 다섯 개 하위 단계가 있다. 이들 B2 단계의 특징은 시각적 데이터를 단순히 언어적 기호로 변환하는 것이 아닌 이들의 선택과 표현에 필자의 의도를 개입한다는 특징이 있다. 이러한

필자 의도적 특성은 대인적 메타 기능에 해당한다. 즉, 이 이동 마디에서는 논의 전개를 위해 시각적 데이터로 제시된 데이터 중 무엇을 선택할 것인가, 어떤 순서로 배열할 것인가, 어떤 방식으로 기술할 것인가를 모두 포함한다. 따라서 B1단계보다는 인지적 노력이 더 필요한 부분이다.

이들 B2 이동 마디의 단계는 보통 논의의 전형적인 장이 마련된 후 진행되는 경향이 있다. 이러한 논의의 장은 앞서 제시되었던 B1의 단계가 선행하는 경우도 있지만, 시각적 데이터를 게시하고 아래에 주석과 같은 형식으로 내용을 제시하는 방식도 있다. 즉, 이 단계는 앞에 데이터 자체에 집중할 수 있게 하는 장이 미리 제시되어야 한다는 특징이 있다.

이러한 측면에서 B2-1은 B1과 B2의 특성 사이를 오가는 단계라 할 수 있다. 주어진 수치의 순서를 파악해야 한다는 점에서 인지적 노력이 더 필요하기는 하나, 시각적 데이터를 확인하면서 뽑아 쓸 수 있기 때문에 다른 하위 단계들보다는 데이터에 쏟는 노력이 덜하다. 데이터를 전반적으로 제시하는 B1 이동 마디에서도 그 순서를 연대기적 순서나 시각적 데이터 순서가 아닌 더 큰 수치나 작은 수치 순서로 전체적인 내용을 배열하는 것이 특이하거나 이상하지는 않기 때문이다. B2-1과 같은 배열은 어찌되었든 해당 시각적 데이터의 전체 데이터를 모두 언급하게 되어 있기 때문에 더욱 그러하다. 특정 시각적 데이터의 전체 데이터를 그 안에서의 수치 크기 순서를 통해 모두 배열해야 한다는 특성은 이 이동 마디에서의 어휘문법적 사용에서도 드러난다. 한국인 대학생은 B2-1을 비교적 덜 선택하기 때문인지 외국인 학습자에 비해 어휘문법 목록이 짧다. 학습자 언어권과 숙달도 모두에서는 ‘순서’와 관련된 어휘문법이 공통적으로 도출된다. ‘가장/제일 높-/낮-/적-’이나 ‘차지하다, 순으로, 다음으로, 잇다, 뒤’와 같은 것들이 그러하다. 이에 더해 B1에서 확인한 바와 같은 ‘(수치)+단위 명사’ 구성과 ‘나타나-’류의 모든 데이터 서술 관련 표현이 등장하는 것도 이 둘이 유사한 특성을 갖고 있음을 말해 준다.

(ㄱ) 하루 평균 가사 노동 시간은 1시간~2시간이 제일 많았으며 그다음으로는 1시간 미만, 2시간~3시간 순으로 많았다.

[103-고급-베트남어-경영3-여]

<표 67> B2 언어권별 어휘문법 목록

| 이동 마디와 단계 | | 언어권 | | |
|-----------|----------------------|--|---|--|
| | | 중국어권 | 베트남어권 | 한국인 |
| B2 | 의도적 그룹화 | | | |
| B2-1 | 수치가 높은/낮은 순서대로 나열 | <ul style="list-style-type: none"> • (수치)명 -으로 (가장/나타나)/-이/있- • (수치)분 -으로/-이다 • -가 가장/제일 많았-/높-/많-/적- • 그 다음은/그 다음으로/그 뒤를 잇- • 나타났- • -는 사람-들/-이/-은 • -라고 대답하-/응답하- • -를 차지하-/살펴보- • 불과하- • (수치)시간 미만/까지/부터/사이/에서/은/을/이/이상 • -에 따르면 • -에서는 | <ul style="list-style-type: none"> • (수치)시간 순으로/-이/-은/이상 • (수치)명의로 가장/제일 높-/많-/이 • 것으로 나타나- • 그 다음으로(는) • 그 뒤를 • -를 이었-/차지하- • 많았- • 응답한 • 순으로 많-/잇- | <ul style="list-style-type: none"> • (수치)시간 미만/이상/사이/-을/-이/-에서/-은/-의 • (수치)명-으로/-이/-은/순으로 • -것으로 나타나- • -는 경우- • 뒤를 잇- • 다음으로는 |

| | | | | |
|------|--------|---|---|--|
| | | <ul style="list-style-type: none"> • -으로 가장 많-/나타났- • 이었다 • -한/하는 결과/사람 • 하였다 • 이었다(원형-있다)/-이었다 • -이 있- | | |
| B2-2 | 수치적 비교 | <ul style="list-style-type: none"> • (수치)분-이/-으로 • 가장 A • 것-으로 보-/-이다 • 경우에는 • (수치)배-가/-이-/정도 • 시간이 • -ㄹ 하는- • -ㄹ 수 있- • 이 돈을 • 높게 나타났- | <ul style="list-style-type: none"> • (수치)분 이상/-으로/-이/증가/-에/줄어들-/-에 불과/ • (수치)년-부터/-까지/-에/-에는 • (수치)명-이 • (수치)배-이/이상/이하/-에/-를/-로 줄어들/에 가깝- • 가장 적- • -것으로 나타나-/조사 • 경우 그 차이/-에는/남성(남편)의 • -는 것/ 경우/시간이 • (더) 많- • 되었- • -ㄹ 수 있-/정도로 | <ul style="list-style-type: none"> • (수치)시간-이/-을/-과/늘- • (수치)분 내외/이내/-으로/-이/순/-과/정도/줄어들-/증가하-/-이-/-을/더/-에서/-의/중 • (수치)배-가량/이상/차이/-에(서)/-의/많-/정도/높-/더/-로/-를/-까지/수준 • (수치)년-에(는,도,서)/-까지/간/-으로 • -는 것으로 나타났-/드러났-/파악-/확인- • -ㄴ/은/는/-의 경우-에/-에는/-에도/중/-와/-라/N • -고 있- • 그 차이/비율-은/는/이/가 |

| | | | | |
|------|---------------|---|--|--|
| | | | <ul style="list-style-type: none"> • -를 보이고/하는 • -ㅁ을 알- • 보다 (약) (수치)배 • 보이고 있- • 비하여 • 시간(의) 차이/-이/-을 • -에 달하/비하- • 증가하였- • 하는 경우 • 할 때 • 하였다 • 많이 하는 • -어야 하- | <ul style="list-style-type: none"> • 알 수 있- • -에 비하여 |
| B2-3 | 특성에 따른 그룹화 | <ul style="list-style-type: none"> • (수치)분-으로/-에서/-이 • (수치)년-부터/-의/-에/-까지 /동안 • (수치)명-이/-으로 • -가 일하-/가사(를)하는 • -의/-는 경우- • 감소하-, 늘어나- • ㄴ-(것)으로 나타나-/(꾸준 | <ul style="list-style-type: none"> • (수치)분-으로/-에서/-을 • (수치)년-동안-부터/-에/-에 는/-까지/(수치)분/(수치)분 • (수치)명-씩/-이/-으로/-의 /-이- • -까지 지속적 • 가장/더 많- • 감소하였- | <ul style="list-style-type: none"> • (수치)년-부터/-에(서)/-은/- 을/-까지/-의/-보다 • (수치)시간-이/-이-/이하/이상 /-의 • (수치)분-이/-이-/에(서)/-을 /-까지/-가량/이내/줄어들-/동 안/ • (수치)명-을/-이/-으로/-의/- |

| | | | | |
|------|---------|---|---|---|
| | | <ul style="list-style-type: none"> 히) 증가하-/상승하- • 까지 남성의/여성의/증가하- • -은 하루에 • 늘- • 으로 소폭 감소/늘어나 • 제일 높- • -하-였-/-는-/-ㄴ 것 • 훨씬 더 많- | <ul style="list-style-type: none"> • 는 것으로 나타나-/보- • 경우-도/에는/-는 • -고 있- • 꾸준히 증가하- • 나타났다 • 는 것으로/경우에서/경향을 • 늘어나- • 반면(에) • -ㄹ 때에 • -를 하- • 버는 가족 • 보였다 • 시간-은/-이 • -으로/-은 증가하-/감소하- • -을 분석하-/하는 • -의 경우 • 이었다 • 지속적으로 • -하-/-는 것/반면 • 하였다 | <ul style="list-style-type: none"> • -ㄴ 경우-에(는,도)/-이 -/N/adv • 꾸준-증가- • -ㄴ 것/-음을 알수 있-/-을 확인할 수 있-/볼 수 있- • -에 비하여 • 하였- |
| B2-4 | 제일 큰/작은 | <ul style="list-style-type: none"> • -으로 가장 많-/적-/높- | <ul style="list-style-type: none"> • (가장) 많- | <ul style="list-style-type: none"> • (수치)시간-에서/이상/미만/- |

| | | | | |
|------|--------------|---|--|---|
| | 수치에 주목 | <ul style="list-style-type: none"> • -ㄴ 사람은/-이 • 높게 나타나- • 대부분이- • 많- • (수치)명-으로 가장/제일 • 시간에 대- • 알 수- | <ul style="list-style-type: none"> • -게 나타났- • -ㄹ/을 알 수 있- • -으로/이 가장 높- • -ㄹ/을 수 있- • 경우- 이/-에서 가장 | <p>으로/정도/-을/-의</p> <ul style="list-style-type: none"> • 가장 많-/높 • -를 보이-/차지하- |
| B2-5 | 특이성/중요 지점 지적 | <ul style="list-style-type: none"> • -ㄴ/는 시간이 • (수치)년-에 | <ul style="list-style-type: none"> • (수치)년-에는 • 것을 보- • -ㄹ/을 수 있- | <ul style="list-style-type: none"> • -하는 경우 |

<표 68> B2 숙달도별 어휘문법 목록

| 이동 마디와 단계 | | 숙달도 | |
|-----------|----------------------------|--|---|
| | | 중급 | 고급 |
| B2 | 의도적 그룹화 | | |
| B2-1 | 수치가 높은/낮은 순서대로 나열 | <ul style="list-style-type: none"> • (수치)시간 미만/이상/-에서/-부터/-이/-까지/-과 • (수치)명 중/-이/-에/정도/-만/-이/-으로/-의 • 가장/제일 높-/낮-/많-/적- • -으로/-이/-게 나타났- | <ul style="list-style-type: none"> • (수치)시간 미만/정도/이상/이내/-에서/-부터/순/-이/-까지/사이/-씩/-의/-은/-을 • (수치)명-으로/-이/-순/-이/-과/-을 • -으로 가장/제일/더 높-/많-/적- |

| | | | |
|------|-----------|---|--|
| | | <ul style="list-style-type: none"> • 그 뒤를 잇-/N • 그 다음-으로/-은 • -를 차지하- • 하-였/-는 • 대답하- | <ul style="list-style-type: none"> • 순으로 많- • -에 따르- • -를 잇-/차지하- • 그 뒤를 • -(하)는 사람이 |
| B2-2 | 수치적 비교 | <ul style="list-style-type: none"> • (수치)분-에/-이/-이-/높- • (수치)배-로/-이/-를/-으로/더/정도/-가 • (수치)년 동안 • 마지막은 • 은 비율을 • 많았다 • 보내- • -에 달하- | <ul style="list-style-type: none"> • (수치)배 더/-이-/이상/이하/정도/-로/가까이/-나 /-에 • (수치)년-에(는)/-의/-부터/-까지/ • (수치)분-으로/-이-/증가하-/-에/-의/줄어들-/-에 다다르- • (수치)명-이 • 그 차이-가/-는 • -ㄴ/ㄹ 경우 N/-에는 • -ㄹ 정도로 감소되-/격차가 줄-/좁혀지-/줄어들- • 가장 적었- • 것으로 보-/나타나- • -음/것을 알- • -는 것이-/반면 • (더) 많- • -ㄹ 수 있- • -를 하는/알- |

| | | | |
|------|------------------|--|--|
| B2-3 | 특성에 따른 그룹화 | <ul style="list-style-type: none"> • (수치)명-이-/중/-에서/-은/-을 • (수치)년 사이/-부터/-에/-의/이후/-까지/-은/동안/-만에 • (수치)시간-은/-에/-을/-이/이상/-씩/-도 • (수치)분-으로/이상/-부터/-에서/-이-/-까지/-을 • -으로 (꾸준히/계속) 증가하-/상승하-/감소하- • -ㄴ 것으로 나타나- • 반면(에) • 하였- • 그 중/중(-오타)에 • -ㄹ 수 있- • 살펴보면 • -에 따르면 • 하는 경우 | <ul style="list-style-type: none"> • 비하여 • (수치)년-에(는, 서)/-부터/-까지/-은/-의/동안/이후 • (수치)시간-에서/-으로/-은/-의/-이 • (수치)분-부터/-에서/-으로/-을/-이/-이-/-씩 • (수치)명-이/중/-으로/-과/-의/-이-/-을/-에(서) • (수치)배 -나/-에 • 가장 높-/낮-/많-/적- • (-까지/꾸준히/-으로) 감소하였-/증가하였-/지속적 • -의 경우 N/-는/-도/-에(는, 서)/중 • -는 것으로 • 더 많이 하- • 나타났- • -ㄹ 수 있- • ㄹ 때-에/N • 반면 -에/N • -보다 N/adv 높-/더- • 추세-를/-이- • 하였- • -고 있- |
|------|------------------|--|--|

| | | | |
|------|-----------------------------------|---|--|
| | | | <ul style="list-style-type: none"> • 불과하였- • 비하여 • -에 따르- • 있었으며 • 하는 것/N |
| B2-4 | <p>제일 큰/작은 수치에 주목</p> | <ul style="list-style-type: none"> • (수치)명-이/-이-/-만/중/-으로/-있- • (수치)시간-에서/-이/미만/-까지/이내/이상 • -으로 가장/제일 많-/적-/높- • -것으로/-게 나타났- • -에 따르면 | <ul style="list-style-type: none"> • (수치)시간-은/-에서/-이 • (수치)명-밖에/-으로/-과 • 가장 많-/작-/적-/높- • -게/으로 나타났- • 이상-은/을/이 • -는 것으로 |
| B2-5 | <p>특이성/중 요 지점 지적</p> | <ul style="list-style-type: none"> • 것을 발견- • 하였다 | <ul style="list-style-type: none"> • (수치)년-부터/까지 • (수치)분-으로 • 줄어들었다 • 시간 차이가 |

B2-2에서는 수치의 비교가 나타난다. 수치의 비교는 기준 수치가 있고 그에 따라 여러 수치들을 비교하는 경우가 대부분이다. 따라서 기준으로 삼고자 하는 데이터에 대한 전반적 내용이 앞에 제시된다. 다만 이 단계의 선택 빈도가 낮은 것은 수치 계산이라는 인지적 노력이 필요하기 때문에 제2언어 사용에 인지적 용량을 주로 사용하는 학습자들에게는 매력적인 선택이 되지 못한다. 또한 이동 마디 단계 빈도에서 볼 수 있듯, 대부분의 학습자들이 시각적 데이터 순서에 따라 제시하거나 시계열로 제시, 또는 수치가 높은 순서대로 전반적인 데이터를 제시하고 난 후에는 비교나 특수성에 대한 논의가 필요 없다고 보는 듯하다. 그러나 이 단계까지 진입한 대학생들은 논하고자 하는 것의 수치를 비교하여 표현하는 ‘(수치)배’, ‘-보다’, ‘비하여’, ‘더’와 같은 어휘문법을 자주 사용한다. 중국어권 학습자군을 제외한 나머지 학습자군에서는 이를 더욱 강조하기 위해 ‘차이, 폭, 비율, 격차’와 같은 어휘를 더 선택해 사용하는 경향이 있다. 그리고 중국어권 학습자와 중급 학습자는 이 이동 마디 선택 빈도가 낮아 말뭉치가 많이 구축되지 않아서 그러하겠지만, 다른 말뭉치에서 수치의 변화를 나타내는 어휘가 ‘가깝다, 증가하다, 줄어들다, 높다, 많다, 좁혀지다’로 비교적 풍부한데 비해 각각 ‘높다’와 ‘많다’ 하나밖에 없다. 이와 같은 더 다양한 어휘 선택의 차이는 학습자들의 어휘적 숙달도와 관련이 있음을 알 수 있다.

또한 B2-2부터의 이동 마디의 단계들은 대인적 메타 기능을 수반한다. 즉, 이 단계에서는 독자에게 전달하고자 하는 저자의 가치가 나타난다. 데이터 배열을 객관적으로 제시해 논의의 장을 만드는 B1 이동 마디 및 그와 성격이 유사한 B2-1과는 달리 B2-2에서부터는 ‘-ㄴ/을 정도로, 불과하다, 수준, 달하다’와 같은 가치 개입적 어휘문법 요소가 눈에 띈다. 이 단계에서는 저자가 특정한 기준을 정해 놓고 그를 지향점으로 다른 데이터를 평가한다. ‘-ㄴ/을 정도로’는 기준에 비해 과도한 또는 과소한 취급을 받았을 때 그의 부당함을 나타내기 위한 의도로 주로 사용된다. 또한 ‘불과하다’, ‘수준’, ‘달하다’ 역시 데이터의 수치가 특정 수준에 도달하지 못해 아쉬운 필자의 태도를 드러내는 단어라 볼 수 있다.

(ㄴ) 이는 특히 남편이 외벌이를 하는 경우 남성의 하루 가사 노동 시간은 채 50여분이 되지 않는 반면, 여성의 가사노동 시간은 250분에 다다를 정도로 약 8배에 가까운 가사 노동시간의 차이를 보이고 있는 것이다.

[6-한-경제4-여]

(ㄷ) 반면 남성의 가사 노동 시간은 하루에 거의 50 분에 불과하여 여성의 약 5배정도 낮다.

[79-고급-베트남어-경영4-여]

(ㄹ) 2019년 여성의 가사 노동 시간은 남성의 4배 수준으로 나타났다.

[102-고급-베트남어-경영3-여]

(ㄹ) 그런데 기혼 응답자 수량이 비혼 응답자 수량의 2.5배에 달하는 것으로 나타난다.

[33-중급-중국어-경영3-여]

이러한 비교는 궁극적으로 판단하고자 바가 있다. 이 판단하고자 하는 바를 밝혀냈음을 제시하는 어휘문법으로는 ‘-(ㄹ/것)을 알 수 있다’를 들 수 있다. 아래의 (ㄷ)과 같은 경우에도 단순히 시계열적으로 가사 노동 시간 증가를 나타내는 것이 아니라 여성의 가사 노동 시간과의 비교를 통해 남성의 가사 노동 시간이 높아진 것을 일단 긍정하는 가치적 지향을 드러내고 있다.

(ㅁ) 2019년에 비교하여 남성은 2021년에 평균 30분의 가사 노동 시간이 높아진 것을 알 수 있고, 여성은 6분 높아진 것을 볼 수 있다.

[62-중급-러시아-경영1-여]

‘특성에 따른 그룹화(B2-3)’ 역시도 데이터가 드러내는 현상에 대한 저자의 가치적 개입을 나타낸다. 구체적으로는 저자가 비교하고자 하는 그룹의 특성에 초점을 맞춘다는 것에서 저자의 관심이 어느 그룹을 중심으로 하는지를 표시한다. 다만 이 단계는 그룹화에 해당하는 집단들을 저자가 애써 만들 필요 없이 시각적 데이터에 제시된 대로 선택할 수 있다

는 점에서 B2-1과 유사하게 인지적 부담이 비교적 적다. 따라서 이 단계는 모든 집단의 학습자가 선호하고 사용하는 어휘문법도 큰 차이가 없다. 이 단계는 그룹 내의 수치 추이를 좇는다는 점에서 개별 변인을 비교하는 B2-2와 약간은 다른 어휘문법적 표현이 다수 드러난다.

모든 학습자 집단에서 나타나는 ‘꾸준히/계속/지속적’과 같은 어휘는 한 그룹의 수치 변화를 좇을 때 주로 사용되고, ‘더, -보다, 비하여, 반면’이 그룹 간 비교에 사용된다.

(사) 반면 남녀 모두 맞벌이를 하는 경우 여성의 가사 노동 시간이 크게 줄어들고(129분) 남성의 가사노동 시간도 줄어들지만 전반적으로는 크게 줄어들지 않는다.

[66-중급-베트남어-경영3-여]

이 단계를 사용하는 학습자들의 어휘문법 차이는 집단별로 크지 않고 유사한 경향을 나타낸다. 고급 학습자의 ‘추세’와 중국어권 학습자의 ‘소폭’, 베트남어권 학습자의 ‘경향’ 정도가 다른 어휘라 할 수 있다. 다만 한국인 대학생의 사용 빈도는 외국인 대학생에 대해 상대적으로 적어서 그러한지 특수하게 나타나는 어휘문법이나 구 구성이 눈에 띄지 않았다.

그룹화라는 행위 자체가 이미 특정 가치에 따라 집단을 나누는 것이어서 어휘문법적 요소에까지는 가치적 측면이 잘 드러나지는 않는다. 하나 드러나는 가치적 어휘문법이 있다면 중국어권 학습자의 강조가 덧붙는 ‘훨씬 더 많-’ 정도가 있다.

(오) 여성이 일하는 것과 상관 없이 모두 가사 노동 시간이 남성보다 훨씬 더 많다는 것을 알 수 있다.[1-고급-중국어-경제4-여]

남성과 여성, 외벌이 가정과 맞벌이 가정이라는 그룹은 이미 시각적 데이터에 제시되어 있기도 하여 학습자들이 쉽사리 그룹화를 할 수 있다. 그래서 그러한지 학습자들은 곱하고 나누어야 하는 수치적 비교 방식을 통해 가치적 어휘문법을 사용하기보다는 우선 그룹을 나누고 그 그룹에

서 나타나는 데이터의 추이에 초점을 맞춘다. 따라서 그룹을 특징하는 ‘경우’가 모든 집단의 어휘문법에 나타난다.

(자) 그다음은 외벌이고 아내가 일하는 경우가 106명이다.

[21-중국-경영2-여]

(카) 구체적으로 살펴보면 맞벌이의 경우 남성의 하루 가사 노동 시간이 27분에 불과하였으나 여성의 하루 가사 노동 시간이 129분이었다.

[3-고급-베트남-경제4-여]

(크) 코로나가 발생한 2019년에는 남성과 여성의 하루 평균 가사노동시간이 각각 48분, 여성의 경우 174분을 차지하였으나...

[12-한-경제4-여]

B2-4와 B2-5는 데이터의 특수한 부분만을 지적하여 그에 대해 여러 방향으로 해석과 추론을 한 데이터 설명문에서 나타났다. 즉, 주어진 시각적 데이터를 설명할 때 데이터 중 저자가 데이터에 숨겨진 중요한 의미를 찾아내어 제시하는 부분이다. 미숙한 저자의 경우에는 지적해야 할 데이터의 특수한 부분을 알아내는 것 자체가 매우 어려울 수 있다. 그리고 알아냈다 하더라도 인지적 부담이 대단히 크게 작용하는 이러한 글쓰기를 자의로 하는 경우도 거의 없다. 따라서 이 단계를 선택한 대학생의 빈도는 매우 적다. 즉, 이들에 속하는 코드 자체가 많지 않아 단계의 어휘문법적 특성을 알아내기 어려운 측면이 있다.

다만 B2-4는 ‘제일 큰/작은 수치에 주목’하는 것이기 때문에 ‘가장/제일 많-/적-/높-/작-’과 같은 표현이 모든 학습자군에서 두드러지게 나타났고, 그러한 수치를 통해 어떠한 의미가 보여진다는 점에서 ‘알 수 있-, 나타나다, 보이다, 차지하다’ 등의 어휘를 사용한다. 이러한 어휘문법 사용은 뒤에서 다룰 B4 의미의 해석과도 유사한 부분이 있다. B4와의 차이점이라면 하나의 가장 작고 가장 큰 수치에 대해서만 해석을 진행한다는 점이다.

(ㄷ) 이런 결과에 중예 대졸인 응답자들이 620명으로 1위를 차지했으며 하루 평균 가사 노동 시간이 1시간부터 2시간까지로 가장 높게 나타났다.
[107-중급-베트남-경영4-여]

B2-5에서 드러나는 공통적 어휘문법은 그 자체로 수가 매우 적어 위에 서 볼 수 있듯, 전체를 모두 나열하여도 공통적인 것이 시각적 데이터의 수치 추이를 나타내는 것 정도이다. 그리고 그마저도 ‘(수치)년-에’, ‘(수치)년-에는’, ‘(수치)년-부터/까지’, ‘(수치)분-으로’의 네 개로 모두 달라 모든 학습자 집단에서 잘 사용하지 않고 공통점도 발견하기 어려운 단계라 볼 수 있다. 이는 데이터의 특이한 점에서부터 시각적 데이터가 함의하는 의미를 해석해 나가야 한다는 부담이 있어 기술을 하지 않은 것일 수도 있고, 특이한 점을 굳이 찾아야 할 필요성을 느끼지 못해 그러할 수도 있다.

(3) 시각적 데이터 언급 표지 사용(B3)

‘시각적 데이터 언급 표지 사용(B3)’은 위의 단계들이 하위 단계를 포함하는 것과 달리 단일 단계에 해당한다. 여기에서는 시각적 데이터를 언급하여 독자의 시선을 시각적 데이터로 이동하게 한다. 이는 시각적 데이터와 언어적 기호가 유기적인 하나의 텍스트를 만들어 내는 텍스트적 메타 기능을 표상하고, 필자의 의도에 따라 표의 특정 부분을 인지하도록 독자에게 명령하는 대인적 메타 기능을 아울러 가지고 있다.

Crismore et al.(1993:47)에서는 Hyland(2005)에서 대인적 측면에만 집중했던 메타 담화(metadiscourse) 개념을 텍스트적 메타 담화로 확장하여 그 하위 분류를 텍스트적 표지와 해석적 표지로 두었다. 텍스트적 표지로는 논리적 접속어(logical connectives), 순서어(sequencers), 안내어(reminders), 주제 부각어(topicalizers)가 있다(이슬비, 2016:75, 번역 재인용). 정혜승(2012:470)에서는 이 중 안내어(해당 논문에서는 이정표)에 대해 “텍스트의 다른 부분을 지시”하며 그 예로 “앞에서 본 바와 같이”와 같은 표지를 들 수 있다고 하였다. 그러나 본고에서는 데이터 설명문에

서 텍스트의 다른 부분을 지시하는 역할을 하는 안내 또는 이정표의 표지로 시각적 데이터 언급 표지도 들어갈 수 있다고 본다. 이들은 구체적으로 텍스트의 특정 부분으로 시선을 인도하고, 그가 지적하는 내용으로 독자를 끌어들이기 때문이다.

텍스트적 메타 기능은 텍스트의 응집성과 일관성에 기여하는 어휘문법적 요소로 표상된다. 시각적 데이터와 언어적 기호를 오가며 형성하는 텍스트의 순서와 진행 방향은 다음 작문처럼 ‘표/그림’의 배열 및 연번 순서대로 데이터를 분석하며 텍스트를 진행하는 것이 있다.

(ㄱ) 다음 <표1>와 같이 확인할 수 있는 듯이 조사 대상은 1000명이며 그 중에 여자와 남자 비율이 동등하게 각각 500명으로 되어 있으며 기혼 상태인 응답자가 720명이고 비혼 상태인 응답자가 280명이다. …(중략)…

<표1> (삽입함)

2022년 통계청의 조사에 따르면 맞벌이 상태별 하루 가사 노동 시간은 다음 <그림 1>에서 확인할 수 있다.

<그림 1>(삽입함)

위의 <그림1>에서 볼 수 있듯이 맞벌이 가정과 외벌이 가정(남편 혹은 아내)은 모두 여자가 남자보다 하루 가사 노동 시간을 더 많이 투입하고 있는 것으로 나타났다. …(후략)

[2-고급-베트남어-경제4-여]

또한 텍스트적 메타 기능은 ‘위/다음 표/그림’과 같이 시각적 데이터의 위치를 언급하거나 ‘표/그림 -에 따르면/-과 같이/-을 (살펴)보면/-에서’ 처럼 참고 자료로서의 자세한 지점을 부각해 언급하는 데에서 확인할 수 있다. 독자에게 특정 지점을 보도록 견인하는 것은 필자가 자신의 주장을 관철하기 위해 독자에게 명령하는 행위이다(이지연, 2023:166). 학습자들은 모든 집단에서 이러한 역할을 잘 이해하고 사용하는 것으로 생각된다.

(ㄴ) 표1을 보면 성별, 학력 하루 평균 가사 노동 시간하고 혼인 상태에 관하여 인원수를 통계했다.

[31-중급-중국어-경영3-여]

<표 69> B3 언어권별 어휘문법 목록

| 이동 마디와 단계 | | 언어권 | | |
|-----------|---------------------|---|--|--|
| | | 중국어권 | 베트남어권 | 한국인 |
| B3 | 시각적 데이터 언급 표지 사용 | <ul style="list-style-type: none"> • -에 대해- • (연번)과 같- • (연번)-은 (연도 수치) • 결과-를 보 • -에 따르-/의하- • -은 다음-과/그림/표 • 대하- • -을 (살펴)보면/비교하-/만들었- • -(으)면 위의 • (수치)명을 대상 • 살펴보- • 위의 그림/표 (연번) • 하였다 • -ㄹ 수 있- | <ul style="list-style-type: none"> • (표/그림 연번)-에 따르-/을 통하-/=-은/-을/-에/-에서/과/을 살펴 보- • -와/과 같-다/-은 • -에 따르면 • -고 있다 • -고자 한다 • -를 나타내고 • 다음으로 • 대한 조사- • (보/알) -ㄹ 수 있- • -에서 볼 • -는 위의 (그림/도표/통계표) • 이는 다음 • 하였다 • 일하는 | <ul style="list-style-type: none"> • (표/그림 연번)-을 (살펴)보 • -/-에서(는)/-의/-는 • -고 있- • 다음/(표/그림 연번)/(수치)-과 같- • -ㄴ/은 결과- • 다음-으로/-과/N • -ㄹ 수 있- • -를 따르-/통하-/비교하- • -에 대해-/따르- • 위(의) 표/그림(연번)-에(서)/=-은/-을/-이 |

<표 70> B3 숙달도별 어휘문법 목록

| 이동 마디와 단계 | | 숙달도 | |
|-----------|------------------|---|---|
| | | 중급 | 고급 |
| B3 | 시각적 데이터 언급 표지 사용 | <ul style="list-style-type: none"> • (표/그림 연번)과 같-/-에/-은/-을/-과 • (표/그림 연번)/결과-에 따르-/-에 의하-/-에 대해-/-를 보면 • (수치)년에 • -를 (살펴)보-/비교하- • -를 통하- • 다음 그림(연번)/-과 같- • 위의 그림/표 • -고 있- • ㄹ 수 있- • 시간 비교에/-에 대-/-을 보내- • 하는/한 결과/설문 | <ul style="list-style-type: none"> • -(표/그림 연번)과 같-/-를/-으로/-에(서)/-은 /-의/중에 • -을 (살펴)보면/통하-/나타내-/위하-/비교하- • -에 대해-/-따르-/-의하- • 위의/다음(으로) 표/그림 • 본 연구 • -ㄴ 보고에 • -는 다음 표 • 비교한 • 살펴보- • -ㄹ 수 있- • 알-/하-ㄹ 수 • -의 데이터를 • 이는 다음 |

B3은 모든 집단에서 선택 빈도가 높은 단계이기는 하나, 이 단계에 해당되는 문장 또는 구가 단순하고 짧아 어휘문법의 목록이 길게 생성되지는 않았다. 또한 집단별로도 ‘-과 같-’, ‘-에/를 따르-/통하-’와 같은 표나 그래프가 선행하는 어휘문법을 공유하며, 고급과 베트남어권 학습자군을 제외하고는 모두 ‘결과’ 어휘를 사용한다. 그러나 고급에서만 사용된 ‘중에’는 일반적인 시각 데이터 언급 표지 사용례인 (ㄷ)과 비교해 보았을 때 달리 특이한 용법을 보이고 있다.

(ㄷ) 그림 2중에 남성과 여성의 하루 평균 가사 노동시간 비교 그림을 보며, 여성이 남성보다 가사 노동 시간 훨씬 더 많지만, 2020년까지는 계속 증가하는 경향이 있고 2021년에 살짝 떨어지는 반면에 남성이 2019년부터 2021년까지는 늘 증가하는 추세다.

[4-고급-중국어-경제2-여]

(ㄹ) 위의 <그림 2>를 통해 2019년부터 2021년까지 여성이 남성보다 하루 평균 가사 노동 시간이 약 2배 이상 많은 것을 알 수 있다.

[25-한-경제1-여]

따라서 B3은 대부분의 집단에서 공유되는 짧고 지시적인 어휘문법으로 이루어져 있지만, 때로는 어색한 사용이 눈에 띈다. 이는 고급 학습자 중 중국어권 학습자 두 명에게만 나타나는 것이고, 중급 중국어권 학습자에게서는 나타나지 않아 모어의 영향인지, 단순한 실수인지, 다른 요인이 있는지 더 탐구가 필요하다.

또한 ‘-ㄷ 수 있-’의 사용이 모든 학습자군에서 나타나는 점도 주목할 만하다. 이 어휘문법은 주로 ‘표/그림에서 확인하/알/보-ㄷ 수 있-’의 구조로 구성되며, (ㄱ)과 같이 문장 중간에서 시각적 데이터를 인용하는 데 사용되거나, (ㄴ) 및 (ㄷ)과 같이 문장의 종결부에서 특정 사실을 자신의 견해가 아닌 객관적 증거가 있는 것임을 강조할 때 사용된다.

(㉑) 다음 <표1>와 같이 확인할 수 있는 듯이

[2-고급-베트남어-경제4-여]

(㉒) <그림 2>를 통해 남자의 가사 노동시간이 천천히 증가하고 있는 거
알 수 있다.

[125-고급-중국어-경영1-남]

(㉓) 특정 성별이 더 많은 가사노동을 부담하고 있는 사실은 <그림 2>에서
한번 더 확인할 수 있었다.

[19-한-경제2-남]

이는 학술적 텍스트가 갖는 대인적 속성과도 관계가 있다. 학술적 텍스트의 주장은 전문 독자의 설득을 전제하기 때문에, 저자의 견해보다는 객관성에 근거하는 경우가 다수이다. 이를 위해 객관성을 드러내는 여러 표지를 사용하는데, 시각적 데이터를 인용하여 주장의 근거로 삼는 점도 이러한 표지 중 하나라 볼 수 있다.

(4) 수치의 의미 제시(B4)

‘수치의 의미 제시(B4)’ 단계는 앞서 다루었던 수치적 내용을 수치에서 의미로 변환하는 단계이다. 이 단계에서는 특히 앞서 다루었던 시각적 데이터의 수치 결과들을 정리하여 제시한다. 이동 마디 C와 유사한 듯하지만, C의 경우는 앞에서 다룬 모든 수치들의 의미를 실제 세계와 연관 지어 해석한다는 점에서 B4와 차이가 있다.

<표 71> B4 언어권별 어휘문법 목록

| 이동 마디와 단계 | | 언어권 | | | |
|-----------|-----------|--|--|---|---|
| | | 중국어권 | 베트남어권 | 한국인 | |
| B4 | 수치의 의미 제시 | <ul style="list-style-type: none"> • (수치)년-부터/-까지/-에 • 사실을/-는 것을 알- • -를 알 수 있- • 결과를 보-/통하- • 계속/꾸준히 증가하- • -고 있- • -나다는 것 -이/갈- • -는 시간이 • 늘어나- • 돈을 벌- • -을 보면/통하여(서) • -(으)면 코로나 시기/맞벌이 부부 • 비하여 • -보다 훨씬 (더) 많-/가 | <ul style="list-style-type: none"> • (수치)년 -부터 /-에/-까지/-에는 • 것으로 나타나/보-/보이- • 것을 보여주-/알- /-볼- • 것이다 • -게 나타나- • 결론을 내리- • -고 있다- • 그 차이-가/-는 • -기는 하- • 꾸준히 증가하- • 나타-났다 / 내 고 있- • -는 것으로/ 경우 | <ul style="list-style-type: none"> • 보내고 있 • 보인다 • 불구하고 • 비하여 • 사이의 폭 • 시간은 늘어나/ 줄-/증가 • 시간을 보내 • 시간이 증가 • 시간은 여전히 • 시간을 더/보내- • 알 -게 되-/수 • 수 있다 • -어지고 • -에 따라/비하/ 있는 | <ul style="list-style-type: none"> • 것으로 드러나-/나타나-/ 해석-/확인-/보- • -고 있-(다)는/-으며/-으 나/-음을 • -게 나타나- • 경우에는 • -ㄹ/을 수 있- • 많- • -에 비하-/투입하-/있- • -를 확인하-/알- |

- 사 노동/여성이
- 상승하-
- 상관없이
- -에 따르면
- -로 보면
- -ㄴ/은 시간을/증가하-
- -이/가 더 길-/많-
- -에 더 많-
- -으로 보-/나타났-
- -은 것으로
- 많-다/-아지/-은 것/-이
- -지 않-
- 집안일을 하-
- 차지하-
- 하고 있-
- 하는 것/사람/시간/외별
- 이
- -습니다/-습니다
- 훨씬 (더) 많-

- /시간이
- 경우-에/-에서/-
- 임
- 늘어나-
- -ㄹ 수 있다-/것
- 이
- 위의
- 이를, 이에, 이러
- 한
- -다-고 있는/-는
- 것을
- 보다(도)/더/이상
- 많-
- 도출한-
- -들 사이의
- 시간에 따르-
- 때문이다
- 많은 것/노동
- 볼 수
- 인해 집에서

- -에도 불구하고
- -와/과 관계
- 위의 데이터
- -로 나타났-/나
- 왔-/보인다
- -로/을 예상하-
- -ㄴ 것으로/것을/
- 증가하-
- -을 내릴-/더 많
- 이/보내고/통하여/
- 보이-/알/예상하
- /하-
- -(음)을 알
- 일하-
- -은 (점점/점차/조
- 금씩) 줄어들-/증
- 가하-
- -지 않-
- 크게 줄-
- 통하여

| | | | |
|--|--|--|---|
| | | <ul style="list-style-type: none"> • 있-었다/-음을 | <ul style="list-style-type: none"> • 하-고 있-/-는 것 /-ㄴ다는 결론/ 는 시간-/-쓰지만 /-였- |
|--|--|--|---|

<표 72> B4 숙달도별 어휘문법 목록

| 이동 마디와 단계 | | 숙달도 | |
|-----------|-----------|---|---|
| | | 중급 | 고급 |
| B4 | 수치의 의미 제시 | <ul style="list-style-type: none"> • (수치)년-부터/-까지/-에는 • -것 -을 알-/-이다/-으로/-은 • -고 있- • 결과-를 보-/-를 통하- • -ㄴ/는 것으로/보-/나타나- • -를 알-/보- ㄹ 수 있-/하- • -에 따르-/비하-/있- • 더 많-/길 | <ul style="list-style-type: none"> • -가 지속되- • -는 것으로 나오-/나타나-/보이- • -고 있- • -보다(도) 더 많- • 보-/알-/예상하- ㄹ 수 있- • -를 보-/통하- • -에 따르- • 비하- |

- 일하는
- -보다 더 오래/하루 평균
- 은 증가하-
- -지 않-/-고 있-
- -기 때문-
- -ㄴ 차이가
- -는 경우/시간이
- 늘어나-
- -다는 것/결론-을
- 돈을 벌-
- 되었다
- 보인다
- 보내고 있-
- 상관없이
- 시간이 더
- 차지하-

- -와 상관없-/관계없-/관련없-/달리
- -게/점차 줄어들-
- (꾸준히) 증가하-
- 감소하는
- ㄴ-추세/결과-/것/것을 (알-)/것이-/것으로
- 훨씬 (더) 많-/적-
- 결론을 내리-
- 결과를 통하-
- -기 때문-
- -다는 것을
- (보다 더) 많-
- -로 도출하-
- -를 나타나-/나타내고
- -를 보내고
- 있-었다/-다는 것/-음을
- 조사하여-
- 큰 변화
- 투입하-
- 하-는 것/ 경우/시간/추세/-었다/-지 않-

‘수치의 의미 제시(B4)’는 저자의 주장에서 데이터의 의미에 대한 평가 부분이다. 데이터가 어떠한 의미인지 정리하여 설명하는 부분으로 역시 단독 단계이다. 그러나 이 단계에 대인적 메타 기능만이 할당된 까닭은 그만큼 독자를 고려해야 하는 정도가 크기 때문이다. 학술적 텍스트의 특성상 데이터의 가치를 주장하거나 단언하는 것은 독자에게 주관적이라는 인상을 심어줄 수 있다. 따라서 필자가 데이터에 근거해 그들의 가치를 확인하고자 해도 이는 학술성에 의해 제한된다. 이 과정에서 데이터를 객관적으로 표현하는 것처럼 보이면서도 그에 대한 필자의 의도를 덧씌울 때는 데이터를 인용하거나, 데이터 그 자체로 존재한다는 것을 강조해 보여준다. 이러한 기능을 하는 메타 담화를 완화 표지(hedge)라고 하는데, 신영주(2011:64)에서는 특히 학술 논문 서론에 나타난 완화 표지에 대해 Hyland(1998)의 정의를 바탕으로 “필자가 단언하지 않고 가능한 반박에 대비하기 위해 자신의 의견을 완곡하게 표현하는 것”으로 정의하고 있다. 이런 완화 표지들은 수치가 확실한 데이터 설명문에서도 저자가 확신을 부드럽게 전달하기 위해 의도적으로 사용하는 장치이다(이준호, 2012:273). 이러한 완화 기능은 데이터 설명문에서는 ‘-는 것을 보여준다/알려준다/볼 수 있다/알 수 있다’ 또는 ‘-에 따라/통해/보면’과 같이 이미 제시된 데이터를 인용해 객관성을 유지하고 독자에게 단언을 피하려는 양상으로 나타난다.

(ㄱ) 여전히 여성의 반도 안되는 수치였지만 남성이 가정에서 일하려고 노력하고 있는 것을 알 수 있다. [1-고급-중국어-경제4-여]

‘-ㄴ/는 것으로 나타나다/드러나다’와 같이 데이터가 다른 요인, 특히 필자의 의도에 영향을 받지 않고 주도적으로 존재함을 드러내는 예는 다음과 같은 것들이 있다.

(ㄴ) 이번 조사를 통해 코로나바이러스 기간 동안 남성의 가사 노동 시간은 증가했지만 여성의 가사 노동 시간은 여전히 더 많은 것으로 나타났다. [112-중급-베트남어-경영3-여]

이러한 어휘문법의 사용은 필자가 생각하는 주제와 가치에 대한 근거로서 데이터를 객관적으로 활용하고자 하는 학술성에 기인한 것으로 보인다.

B4는 모든 집단이 다수 사용하는 단계이고, 그 빈도 차이도 크게 유의하지는 않다. 그러나 한국인 대학생의 어휘문법 수가 다른 집단에 비해 현저히 적은 것을 볼 수 있다. 이는 한국인 대학생들의 어휘문법 사용이 개인별로 다양해서 중복되는 것이 몇 개 없기 때문인 것으로 생각된다. 외국인 대학생들은 (ㄱ), (ㄴ)과 같이 남성의 코로나 시기 가사 노동 시간이 증가하기는 했지만 여전히 여성이 가사 노동 부담이 크다는 수치적 정리를 간단한 문장으로 제시하기 때문에 중복되어 나타나는 어휘문법이 다수 추출되었다. 그러나 한국인 대학생들의 경우에는 같은 현상을 나타낼 때 더 다양한 어휘를 사용함으로써 중복되는 어휘문법이 적은 것으로 보인다.

(ㄷ) 간극이 점점 더 좁혀지고 있는것으로 보이나 그 수치는 아주 미미하다.

[7-한-경영4-여]

(ㄹ) 이는 곧 집 안에서 머무르는 시간이 압도적으로 길어짐에 따라 야기된 결과이다.

[35-한-경영4-여]

2.3. 해석 및 결론(C)

(1) 주어진 데이터/상식적 지식에 근거한 해석(C1)

해석 및 결론(C) 이동 마디는 ‘주어진 데이터/상식적 지식에 근거한 해석(C1), 데이터의 비판점과 개선 방향(C3)’ 세 하위 단계로 나뉜다. C 이동 마디에 속하는 말뭉치 수는 다른 이동 마디들에 비해 많지 않다. 텍

스트 구성에 미숙한 필자들이 결론을 누락하는 경향이 있기도 하고, 외국인 대학생들에게 익숙한 데이터 설명문 구조가 그래프의 내용만 정확히 기술해야 하는 TOPIKⅡ의 53번 문제이기 때문에 데이터 설명문이라면 이들을 누락해야 한다는 생각도 있어 그러하다. 누락을 하지 않은 경우에도 익숙하지 않은 이동 마디여서 그러한지 논리적 비약이 일어나거나 텍스트적으로 연결이 되지 않는 경우도 다수 있다. 그럼에도 불구하고 이 이동 마디의 C1 단계를 선택해 작성한 학습자들은 앞서의 B4 이동 마디와 유사하게 완화 표지를 사용하여 저자의 의도가 포함된 데이터를 객관적인 것처럼 제시한다. 이 과정에서 ‘-ㄴ/는 것을 알-게 되-/르 수 있-’이나 ‘나타나다, 보이다, 분석되다, 보여주다’와 같은 어휘문법적 요소가 사용된다. C1에서는 B4와는 달리 저자의 판단과 가치가 더욱 적극적으로 반영되나, 오히려 그래서 더 공격받기 쉽다고 생각하는지 B4와 같이 완화 표현을 다수 사용한다.

특히 이 단계에서는 피동 표현과 어휘 사용을 발견할 수 있다. 한국인 대학생과 중국어권 학습자 고급 학습자 집단에서는 ‘생각되다/판단되다/해석되다/사료되다/분석되다’와 같은 ‘명사+되다’의 피동 구성이 나타나 대학생들의 완화 표지 사용을 확인할 수 있다. 그러나 베트남어권 학습자군과 중급 학습자군에서는 이러한 사용이 빈번하게 드러나지 않는다. 따라서 객관성을 유지하는 ‘-되다’ 구성 역시도 지속적으로 학습이 필요하다고 본다. C1이 저자의 적극적인 해석이 개입하는 단계이기는 하지만 완화 표지는 그래서 더 필요하기 때문이다. 이러한 인식이 없는 중급 학습자들은 (ㄱ), (ㄴ)처럼 조금 더 직설적으로 자신의 의견을 개진하는 경우도 있다.

(ㄱ) 그러나 예전의 보수 사상은 완전히 없애기에는 너무 심오하다.

[127-중급-베트남어-경제4-여]

(ㄴ) 이런 현상은 정확하지 않다.이런 상황은 남녀 불평등의 현상을 반영한다.

[35-중급-중국어-경영3-여]

<표 73> C1 언어권별 어휘문법 목록

| 이동 마디와 단계 | | 언어권 | | | |
|-----------|------------------------|---|---|---|---|
| | | 중국어권 | 베트남어권 | 한국인 | |
| C1 | 주어진 데이터/상식적 지식에 근거한 해석 | <ul style="list-style-type: none"> • -ㄴ/는 것 같-/과 관련-/으로/-을/-이다 • -고 있- • -기 때문-이-/에 • -ㄹ/을 수 있- • 많- • 분석되- • 사람들이 • 생각한다 • -아/어지- • 안하- • 않- • -하여야 한다- • -에 따르- • -에 있- • 영향-을/으로 • -(으)로 보인다 | <ul style="list-style-type: none"> • -것으로 보이- • -게 되 • -고 있- • 긍정적이- • -기 때문- • -ㄴ/는 영향을/경향-/시간이 • -ㄴ/는/ㄹ 것 • -ㄴ/는 점점/점차 줄- • 늘어나- • 때문에 • -ㄹ/을 수 있- • 많아지- • 많은 시간을 • 보내는 시간 • 불구하고 | <ul style="list-style-type: none"> • -아/어지- • -어야 하- • -에 더 많-/비하 • -ㄴ/은 시간을 • -를 보내- • -를 알-/하는/할애하- • -이/가 많아 • 이었다 • 있었다 • -적이- • 줄어들- • 증가하- • 졌- • 집에 있- • 참여하- | <ul style="list-style-type: none"> • -ㄴ 것으로 보이-/생각되-/판단되-/해석되- • -것을 알게 되- • -기 때문- • -게 나타나-/되- • -고 있- • -ㄴ 경우/경향/결론 • -ㄹ/을 수 있- • -다 -(라)고/는 생각/인식 • 더 많- • -에 따르-/대하-/비하- • -의 경우 • -이/가 증가하- • -(으)를 보여주-/드러내-/나타내-/시사- • 인해 • 해석하- |

| | | | | |
|--|--|--|--|--|
| | <ul style="list-style-type: none"> • -를 하-ㄴ/여야 • -이/가 증가하- • -하는 시간 | <ul style="list-style-type: none"> • 비하여 • 시간도 증가(증가-오타) • 시간이 많- • 생각하- • 시간 차이가 | <ul style="list-style-type: none"> • 하고 있- • 하는 것 • 할 수 • 하였다 • 활동(이) 제한 • 에서 보내는 • 을 미치- | <ul style="list-style-type: none"> • 훨씬 A |
|--|--|--|--|--|

<표 74> C1 숙달도별 어휘문법 목록

| 이동 마디와 단계 | | 숙달도 | | |
|-----------|------------------------|--|--|--|
| | | 중급 | | 고급 |
| C1 | 주어진 데이터/상식적 지식에 근거한 해석 | <ul style="list-style-type: none"> • -것으로 보이- • -ㄹ 수 있- • -때문- • 많- • -고/-에 있- • -ㄴ다고 생각해- • 영향을 미치-/받- | <ul style="list-style-type: none"> • -ㄴ 것으로 보이-/분석되-/나타- • -고/-에 있- • -게 되- • -기 때문- • 때문이다 • -ㄴ/은/는 것-으로/-이 | <ul style="list-style-type: none"> • -이라고 생각한다 • 못하- • -에 비하여 • -어야 하- • -어지- • -에도 불구하고 • 있- |

- 은 집안일-
- 집에 있-
- 해야 하-

- -ㄹ/을 수 있-
- 더 많은
- -이 많아/제한되-
- 보여주-
- -를 볼-
- 많아지-
- -에 따라
- 은 줄어들지 않/증(증-오타)
가하-

- 가족을 돌보-
- ㄴ다는 (고정)관념
- 시간이 늘어나/더/증가하-/
지남
- 시대에는
- 않은 것
- -을 알-/하는
- 차이가 크-
- 하였다

또한 ‘사료, 판단’과 같은 어휘는 주로 한국인 대학생에게서 중복적으로 나타나고, ‘할애’와 같은 단어도 베트남어권 학습자군에서 확인된다. 고급에서는 ‘고정관념’이라는 어휘가 나타나 학술적 텍스트에 어울리는 고급 어휘들은 중국어권과 중급에서는 나타나지 않는 것을 알 수 있다.

(2) 데이터의 비판점과 개선 방향(C2)

<표 75> C2 언어권별 어휘문법 목록

| 이동 마디와 단계 | 언어권 | | | |
|-----------|-----------------------|-------|-----|---------------------------------|
| | 중국어권 | 베트남어권 | 한국인 | |
| C2 | 데이터의 비판점과 개선 방향 | - | - | -에 대해/따르- -리 수- 없- 어렵- |

<표 76> C2 숙달도별 어휘문법 목록

| 이동 마디와 단계 | 숙달도 | | |
|-----------|-----------------------|----|---|
| | 중급 | 고급 | |
| C2 | 데이터의 비판점과 개선 방향 | - | - |

‘데이터의 비판점과 개선 방향(C2)’은 Swales & Feak(1994)가 학술적 데이터 설명문에 반드시 들어가야 한다고 평한 ‘해설’에 속한다. 그러나 이들이 잘 드러나지 않기 때문에 역설적으로 교수의 필요성이 강조된 것이라 볼 수 있을 만큼 외국인 대학생의 작문에서는 발견하기 어렵다. 본

고의 실험에 참여한 학습자들의 작문에서도 이는 마찬가지다. 한국인 대학생의 작문에서는 그래도 확인 가능한 정도의 말뭉치가 구축되었지만, 고급 학습자에게서는 두 텍스트에서 각 한 개의 코드만 발견되었고, 중급 학습자의 텍스트에는 이 이동 마디에 해당하는 것이 존재하지 않는다.

따라서 이 단계에서는 특징적인 어휘문법도 잘 나타나지 않았다. 다만 한국인 대학생의 어휘문법을 통해 주어진 데이터의 한계를 나타내는 표현인 ‘-르 수 없다’, ‘어렵다’ 등이 쓰임을 알 수 있다. 이는 아래 (ㄱ)과 (ㄴ)에서도 확인할 수 있다.

(ㄱ) 그러나 해당 변화가 혼인 상태별로 구체적으로 집계되지 않았기 때문에 부부 이상의 단위에서 남성의 가사 노동 상대적 비중이 유의미하게 증가하고 있다고 단정지을 수는 없다.

[11-한-경제4-여]

(ㄴ) 표본은 미혼 28%와 기혼 72%로 이루어져 있는데, 같은 가사일이라면 가사일을 분담할 수 있는 기혼이 노동량이 적어야 하지만, 대개 가족 구성원이 늘어날수록 가사일도 늘어나고 집의 크기도 커진다는 것을 고려해보았을 때 혼인 상태가 1인당 가사 노동시간에 어떠한 영향을 끼치는지 이 자료만으로 정확히 알긴 어려웠다.

[10-한-경영4-여]

위의 예문에서 볼 수 있다시피 이러한 해설은 시각적 데이터들을 통합적으로 이해하는 능력이 선행되어야 한다. 또한 더 정확한 분석을 위해 어떠한 정보가 필요한지 선택할 수 있는 지식도 필요하다. 따라서 한국인 대학생의 경우에도 이러한 이동 마디를 구사하는 것은 전공 지식의 도야를 통해 데이터 이해와 분석 능력을 함양한 고학년이 되어야 가능하다. C2가 나타난 12개 데이터 설명문 중 3개를 제외하고는 모두 3,4학년이 작성한 것이 이를 뒷받침한다.

(3) 문제 해결 방안/후속 연구/미래 방향 제시(C3)

‘문제 해결 방안/후속 연구/미래 방향 제시(C3)’에는 데이터에 등장하는 현상에 대해 문제의식을 갖고 그에 대해 해결 방안을 제시하거나 그를 통해 향후 어떻게 변할 것이라는 미래적 의지가 담겨 있다. 그러나 C2와 더불어 C3도 선택한 학습자가 그다지 많지 않아 어휘문법의 목록 수가 짧다.

이 이동 마디에 속하는 텍스트를 작성한 학습자들은 전체적으로 앞에서 논한 시각적 데이터의 분석과 관련된 개선 방향이나 옳은 방향에 대한 내용을 생각해 낸다는 점에서 관념적이고 이에 인지적인 노력이 더 소모된다. 이 단계에서는 자신이 생각한 옳은 방향을 제시하기 위해 ‘-ㄴ 필요 있다, -어야 하다’ 등으로 선언을 한다. 이는 문제 해결 방안을 촉구하고 제시한다는 점에서 이러한 어조를 채택하는 것으로 볼 수 있다.

(ㄱ) 남성도 가사노동에 더 많은 시간을 할애해 여성의 가사노동 스트레스를 줄여야 한다고 생각합니다.

[32-중급-중국어-경영3-여]

(ㄴ) 이것은 분명히 여성보다 낮은 것으로 간주될 수 있으므로 가사 노동 시간을 하는 남성의 비율을 늘릴 필요가 있다.

[108-고급-베트남어-경영4-여]

(ㄷ) 이에 따라 기존에 사회에 팽배한 이러한 성 고정관념을 없애고, 남성 여성이 아니라 가족 구성원 모두가 함께 가사 노동을 분담하여 해야 한다는 인식을 기를 필요가 있다고 생각한다.

[27-한-경영3-남]

<표 77> C3 언어권별 어휘문법 목록

| 이동 마디와 단계 | 언어권 | | |
|---------------------------------------|--|---|---|
| | 중국어권 | 베트남어권 | 한국인 |
| C3 문제 해결 방안/ 후속 연구/ 미래 방향 제시 | <ul style="list-style-type: none"> • 더 많은 • 바꾸- • 발전- • 생각하- • -어야 한다- • 것으로 N • -기 때문에 | <ul style="list-style-type: none"> • 위하여(서는) • -는 것 • 대한 • -ㄹ/는 것- • -을 더 많이 • -(라)고 생각하- | <ul style="list-style-type: none"> • -기 위하- • -ㄹ 필요가 있-/수 있- • 생각하- • -어야 하- • 이러한 • -ㄴ다는 N • -ㄹ/는 것- • 위하- |

<표 78> C3 숙달도별 어휘문법 목록

| 이동 마디와 단계 | 숙달도 | |
|------------------------------------|---|---|
| | 중급 | 고급 |
| C3 문제 해결 방안/ 후속 연구/ 미래 방향 제시 | <ul style="list-style-type: none"> • 것이-다/라(고) • -기 위하여 • 기대된다 • -ㄴ 것-을/은 • ㄹ 것-이/으로 | <ul style="list-style-type: none"> • -ㄴ 다고 생각하- • -ㄹ 수 있 • -어야 하- • -ㄴ/ㄹ 것-으로/이 • 증가-시켜/해야 • 보일 것 |

마지막 이동 마디 단계에서 자신의 견해를 적극적으로 밝히기로 한 학습자들은 위 (ㄱ), (ㄴ)에서도 확인할 수 있듯, ‘생각하다’와 같은 자신을 주어로 한 서술어도 다수 사용한다. 중급과 고급, 중국어권과 베트남어권 학습자 및 한국인 대학생에게서 모두 나타난다. 또한 이 ‘생각하다’는 ‘-라고/다고 생각하다’의 형태로 사용되어 자신의 주장에 대한 완화적 표현을 제시하는 것으로 보인다. 이는 C3 단계가 독자와의 거리와 관계를 고려하는 대인적 메타 기능을 나타내고 있기 때문이기도 하다.

(ㄴ) 경제가 비약적으로 발전하는 오늘에 우리의 생각도 그에 따라 발전해야 한다고 생각한다.

[17-고급-중국어-경영4-여]

어휘의 사용이 적절하다고 느끼지 못하는 경우도 있는데, 앞서의 두 경우는 필자를 주체적으로 드러내는 능동적 구조를 사용하여 이제까지 고수해 왔던 객관적 태도와 사이에서 간극이 있는 정도만을 보여 주었다. 그러나 데이터 설명문 해석 이동 마디에 사용되는 ‘기대되다’와 같은 어휘는 상위 장르가 보도문에서는 적절할지 몰라도 학술적 데이터 설명문의 결론으로는 적절하지 못한 것처럼 보인다. 보통 아래 (ㄷ), (ㄹ)과 같은 경우에는 기대되다 대신 ‘생각되다’나 ‘예측하다’와 같은 객관적 추론의 어조를 부여한다. 그러나 ‘기대’는 이들 어휘에서 나타나는 추론보다는 저자의 미래 희망을 독자와 관계없이 제시하고 있다. 이러한 부적절한 어휘 사용은 중급 학습자들에게 주로 나타나기에 이 또한 한국어 텍스트의 학술성 판단이 부족해 발생하는 것으로 생각된다. ‘기대되다’ 사용은 중급 학습자, 중국어권 학습자 및 베트남어권 학습자에게서 모두 나타나며, 고급 학습자와 한국인 대학생에게서는 찾아볼 수 없다.

(ㄷ) 결혼함으로써 여성이 본인아니라 가족을 위해 돈을 벌어야 한다고 생각한다. 자녀 양육비, 생계비, 적립금 등 많은 부담에 처할 것이다. 그래서 가계 부담을 줄여주기 위해 남성의 가사 노동 참여에 대한 인식을 변화시킬 것으로 기대된다.

[105-중급-베트남어-경영4-여]

(ㄹ) 앞으로 남성 가정보동시간이 여성의 가사노동시간보다 더 클 것으로 기대된다.

[101-중급-중국어-경제1-여]

다만 ‘바람’을 나타내는 어조는 한국인 대학생에게서도 찾아볼 수 있는데, 이때는 ‘기대’라는 단어는 사용하지 않았다. 그럼에도 ‘-으면 좋겠다’라는 구를 통해 이를 표현하고 있는데, 이를 보았을 때 이 단계에서는 다양한 방식과 다양한 학습자가 ‘기대’의 의미를 전달함을 알 수 있다. 이 역시도 그러나 감정의 투영이기 때문에 학술적 데이터 설명문에서는 맞지 않아 보인다. (ㄷ)을 작성한 대학생은 1학년으로 학술적 텍스트의 마지막에 이러한 감정적 투사가 적절하지 않음을 알아채지 못했을 가능성이 있다.

(ㄷ) 가정 내 남편과 아내의 지속적인 대화와 공동의 노력으로 더욱 나은 가정의 변화를 끌어냈으면 좋겠습니다.

[23-한-경제1-남]

제효봉(2015:216-217)에서는 중국어권 학습자의 모어 텍스트 구조 전이를 연구하며 중국어권 학습자의 전형적 마무리 방식 중 ‘희망적 표현’이 있다고 하였다. 이를 사용하는 중국어권 학습자는 ‘-기 바라고 있다, -으면 좋겠다’와 같은 표현을 통해 텍스트를 마무리 짓는다고 하였다. 물론 제효봉(2015)에서 다룬 학습자 텍스트는 백일장에서 수집한 비학술적 장르이기 때문에 이러한 구조가 전이되었다고 본다. 그러나 다양한 언어권 학습자가 위에서 본 바처럼 동일한 구조로 끝을 맺는바, 한국어 학술적 데이터 설명문에서는 이러한 기대와 바람의 표현이 한 언어권의 텍스트 구조에서 전이되었다기보다는 학술적 장과 사용역에서의 미숙성의 문제로 생성된 것이라 보는 것이 타당할 것이다.

IV. 학습자 변인별 학술적 데이터 설명문 양상에 대한 논의와 교육적 제언

이 장에서는 III장에서 살펴보았던 대학생들의 학술적 데이터 설명문 양상을 학습자 변인별로 정리하여 학술적 데이터 설명문을 구성했을 때 나타난 학습자군별 특성을 해석할 것이다. 이러한 해석을 바탕으로 한 교육적 시사점 역시 도출할 것이다.

한국어 학습자의 학술적 데이터 설명문의 분석은 ‘이동 마디’와 관련하여 제시할 수 있는데, 본고에서는 이동 마디 순서와 이동 마디 단계 빈도, 이동 마디 단계별 어휘문법 사용의 세 기준으로 학습자의 학술적 데이터 설명문을 분석한 바 있다. 이들은 각각 학습자가 어떠한 순서로 이동 마디를 나열해 데이터 설명문을 구성하는가, 그 과정에서 이동 마디 단계를 얼마나 사용하는가, 각 이동 마디 단계에서 사용되는 어휘문법은 무엇인가를 나타낸다. 즉, 앞의 둘은 학습자 데이터 설명문의 담화 구성을, 나머지는 언어 사용을 나타내는 것이다. 따라서 아래 1절과 2절에서 다루는 내용은 각각 연구 문제 1. ‘한국어 학습자의 학술적 데이터 설명문은 학습자 변인에 따라 어떠한 구조 및 언어적 특징을 나타내는가?’의 답이 될 것이다. 이를 토대로 3절에서는 연구 문제 2. ‘한국어 학습자의 학술적 데이터 설명문 분석 결과를 토대로 도출할 수 있는 학습자 변인별 교육적 시사점은 무엇인가?’에 대해 다루도록 한다.

1. 학술적 데이터 설명문 담화 구성에서의 학습자 변인별 특성

1.1. 학습자 변인별 이동 마디 순서 유형의 특성

앞 장에서의 분석이 이동 마디와 어휘문법 사용에 대한 결과를 보여준다면, 여기에서는 분류한 학습자군별로 앞에서 다루었던 특성을 해석해 볼 것이다. 즉, 앞 장에서는 분절적으로 제시하였던 중급 학습자:고급 학

습자 및 중국어권 학습자:베트남어권 학습자의 특성을 ‘중급 중국어권: 고급 중국어권: 중급 베트남어권: 고급 베트남어권’으로 세분화하여 제시 하도록 한다.

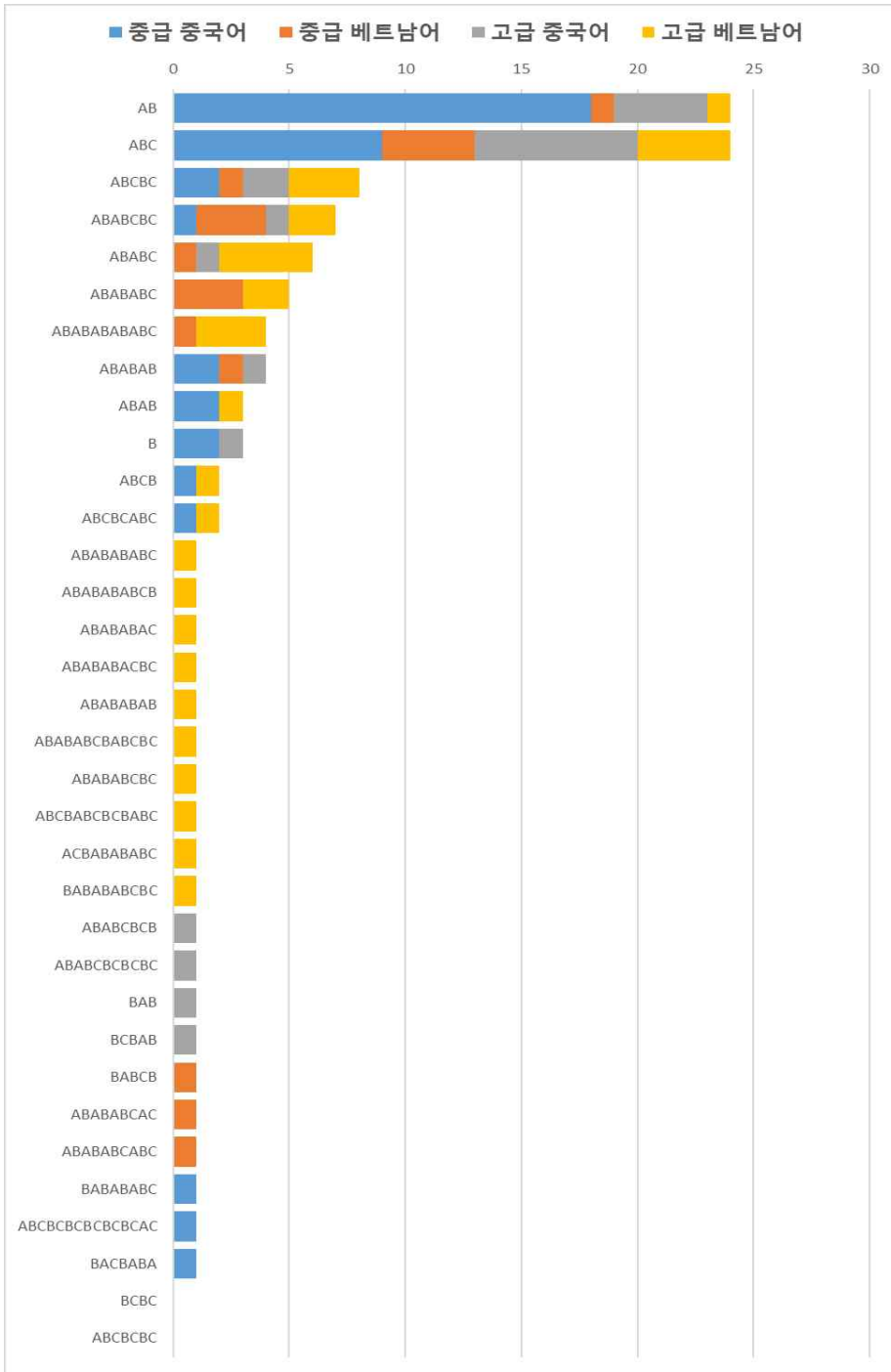
우선 학습자의 학술적 데이터 설명문이 어떠한 순서로 구성되었는지 상위 이동 마디만으로 구성된 ‘이동 마디 순서 유형’부터 보도록 하겠다. 이동 마디 순서 유형은 전체 텍스트의 구조라 할 수 있다. 데이터 설명문을 포함해 익숙하지 않은, 또는 관습적으로 정해지지 않은 텍스트 구조에 영향을 미치는 것은 모어의 텍스트 구조, 유사한 유형의 텍스트 구조 모방, 낮은 숙달도에 따른 한국어 작문 구조 인식 부족 정도 등을 들 수 있다. 이런 영향 요인들 중 어떠한 것이 수집된 학습자 데이터 설명문의 경향성을 설명하는지 자세히 파악해 보도록 한다.

연구 참여자가 생성한 이동 마디 순서 유형은 총 52개로 외국인과 한국인 집단에서 이동 마디 순서 유형을 각각 상위 5개씩 추출한 결과 총 9개⁴¹⁾ 데이터 설명문 구성 중 8개가 A-()-C 유형이었고, 이들이 각 집단에서 차지하는 비율도 외국인 대학생의 데이터 설명문에서 59.2%, 한국인 대학생 데이터 설명문 편수에서도 44%를 차지하였다. 그러나 이 9개 유형에서 하나만 C가 누락된 AB형인데, 이는 외국인 대학생 그룹에서 2위(25편, 20%)를 차지한 유형이다.

외국인과 한국인 대학생의 데이터 설명문 공통점은 A-()-C 유형이 압도적이며, 특히 ABC의 수치가 가장 높았다는 점을 들 수 있다. III장의 데이터 설명문 이동 마디 단계 분석에서도 볼 수 있듯, A-()-C 유형은 ABC 유형의 변형 중 하나로, 시각적 데이터를 각각 드러내어 분절적으로 다루는가 그렇지 않는가, 만약 분절한다면 어떠한 기준으로 나누는가(예: 표/그림1+그림2, 또는 표+그림1/그림 2)에 따라 달라지는 것을 볼 수 있다. 또한 이들 A-()-C 유형은 한국인 대학생 집단에서 상위 4개 모두에 해당되었고, 베트남어권에서도 마찬가지였다. 특히 베트남어권 학습자에게 주로 나타나는 A-()-C 유형의 복잡한 괄호 안 구성은 보고서 형식으로 연번을 나누거나 첫째, 둘째 등의 내용 순서를 구분하

41) 한국인 대학생의 경우 5위가 동수인 것이 4개 유형이어서 4위까지만 집계하였다.

는 방식 때문인 것으로 나타났다. 이러한 방식은 한국어의 보고서 쓰기와 유사한 점이 있는데, 이러한 점은 베트남어권 학습자의 한국어 숙달도가 높아 다양한 상위 장르에서의 데이터 설명문 구성에 대한 이해와 인식 정도가 높아져 그러하다고 생각된다.



<그림 6> 세부 학습자군별 이동 마디 순서 빈도

위 <그림 6> 세부 학습자군별 이동 마디 순서 빈도에서 보면, 세 번째 부터 일곱 번째까지의 A-()-C 유형은 중국어권 학습자보다는 베트남어권 학습자가 주를 차지하며, ABC와 ABABABC를 제외하고는 고급 베트남어권 학습자의 비율이 높다. 고급 중국어권도 A-()-C 유형을 자주 사용하는 경향이 있어 이들의 숙달도에 따라 데이터 설명문 구성 인식이 다를 수 있다.

다만 중국어권과 중급, 고급 학습자 집단에서는 C 누락형이 상위권에 나타났다. 특히 중국어권 학습자의 데이터 설명문에서는 C 누락형이 상위 5개 중 3개에 해당하였다. 그렇다면 이 C 누락형이 외국인 대학생 하위 그룹의 어떠한 특성과 관련이 있는지 언어권과 숙달도별로 나누어 다시 확인해 보았다. 그 결과 중국어권에서는 C 누락형인 AB 유형이 1위로 19편, B 유형이 5위로 3편 등장하였으나 베트남어권에서는 C 누락형이 순위권에 등장하지 않았다. 따라서 중국어권의 텍스트 구조 전이를 의심해 볼 수 있겠으나, 세 가지 이유로 언어 전이를 확신하기에는 어려웠다. 첫째, 중국어권에서 AB를 선택한 학습자들 19명 중 18명이 중급 학습자여서 숙달도 요인을 의심할 수 있다. 둘째, 데이터 설명문은 내포 장르로서 상위 장르의 영향을 많이 받아 전범이 되는 데이터 설명문 양식이라는 것을 설정하기 어렵다. 셋째, 데이터 설명문이 아닌 중국어의 일반적인 내러티브 글쓰기 구조에도 C에 해당하는 결론 단계는 존재한다(제효봉, 2015; 유민애, 2017).

따라서 본고의 분석 기준인 숙달도와 관련하여 AB 유형 사용을 다시 한 번 알아보면, 중급 학습자의 상위 5개 이동 마디 순서 유형 중 1위(19편, 27.1%), 고급 학습자의 상위 5개 이동 마디 순서 유형 중 2위(6편, 10.9%)를 차지하고 있다. 더불어 두 집단 모두 상위 5개 이동 마디 순서 유형 중 C 누락형은 AB 유형밖에 없었다. 따라서 AB 유형은 고급보다는 중급 텍스트에서 더 많은 비율을 차지하고 있음을 확인할 수 있다. 그러나 이 결과로도 AB 선택자 25명 중 22명이 중국어권이어서 결국 AB 유형이 숙달도가 낮아서 생기는 것인지, 중국어권이어서 생기는 것인지는 수치상으로는 알기 어렵다. 중국어권의 텍스트적 구조 전이에

대해서는 Ⅲ장에서 중국어권의 보편 텍스트인 ‘술직’의 마무리 구조라든지 내러티브의 기대와 바람 구조와 같은 C 유형을 확인한 바 있어 단순히 중국어권의 영향이라 보기는 어렵다. 이외에도 데이터 설명문의 구조에 영향을 미치는 중국의 텍스트 구조가 있다면, 중등학교급에서 훈련한 구조가 될 것이어서 중국어권 학습자들의 중등교과과정, 특히 고등학교 교과과정에서 데이터 설명문과 관련된 작문 내용이 있는지 확인해 보았다. 중화인민공화국 교육부(2017)의 2017 개정 보통고등학교 교육과정에서는 보통고등학교 학습자의 중국어 교육과 관련해 필수 과목 7개 및 선택 필수 6개 과목을 제시하고 있으며, 선택 과목도 6개를 제공하고 있다.⁴²⁾ 이들 중 대부분이 읽기와 문학에 관련된 것이고(예: 독서와 토론, 사변적 읽기와 표현, 문학 읽기 및 쓰기) 데이터 설명문과 관련이 있어 보이는 과목은 선택 과목인 ‘학술 논저 심포지엄’ 정도이다. 특히 학술 논저 심포지엄에서는 ‘학술적 소논문 쓰기’ 시간을 배정함으로써 연구 성과를 표현하고, 이론의 도출과 데이터 인용, 참고 문헌 서술 등의 학술성을 명시적으로 제시하고 있다(중화인민공화국 교육부, 2017:32). 그러나 여기에서 그래프나 표 등을 어떻게 다룰지는 나와 있지 않으며, 필수가 아닌 선택 과목이기에 중국어권 학습자들이 이를 통해 보편적으로 데이터 설명문을 학습한다고 보기도 어렵다.

따라서 학습자의 한국어 데이터 설명문 구조에 영향을 미칠 법한 학습자 모어 텍스트 구조의 영향은 확신할 수 없으며, 나머지 요인인 유사한 텍스트 구조 영향과, 낮은 숙달도 및 학년에 따른 한국어 작문 구조 인식 부족이 주원인으로 보인다. 특히 고급 중국어권 학습자들은 중급 중국어권 학습자들과는 달리 ABC 유형을 제일 많이 선택한 것으로 보아 언어권의 영향은 미미할 것이라는 추론이 타당함을 알려준다. 본고에서

42) 본고의 실험 참여자들은 20대 대학생들이 대부분이므로 가장 최근의 교육과정 개정이 적용되지 않은 세대라 보아 최근 교육과정 바로 이전의 교육과정을 사용하였다. 최근 한국의 2022 국어과 개정 교육과정(국가교육과정 정보센터, 2022.6.20. 검색) 및 중국의 2022 어언과목 개정 교육과정에서는 데이터에 대한 교수-학습 인식이 높아져 데이터 설명문 관련 내용이 반영되고 있다(중화인민공화국 교육부, 2022). 그러나 본고의 실험 참여자들은 해당 교육과정 적용 이전 세대들로, 의무 교육에서 관련 내용을 학습한 바 없으므로 이들은 제외하였다.

제시한 ‘수업 과제’와 같은 언어 사용에 관습적인 엄격함이 존재하지 않는 학술적 글쓰기 유형에서는 ‘대학과 관련된 쓰기 경험’이 영향을 가장 많이 끼칠 수밖에 없다고 본다. 외국인 학습자들에게 학술적인 한국어 데이터 설명문의 첫 경험은 TOPIK 평가 문제이다. 그러나 한국인 대학생들에게는 특히 저학년이라면 대학 입학 논술 시험이 학술적 데이터 설명문의 첫 경험이었을 가능성이 높다.

아래 <그림 7>에서 볼 수 있듯, 대학 입학에 위한 논술 시험은 주로 어떠한 관점을 논하기 위해 시각적 데이터를 제시하므로 시각적 데이터의 언어적 이해와 표현 능력을 알아보려는 TOPIK과는 그 목적과 요구 조건이 다르다. 논술 답안에서는 시각적 데이터를 통해 확인할 수 있는 자료의 수치 나열보다는 그를 어떤 관점으로 해석하여 제시할 것인지가 중요하기 때문에 시각적 데이터를 풍부하게 해석하기 위한 방식으로 자료의 배경을 사용하는 것으로 보인다. 이러한 시각적 데이터 서술 방식은 대학 수준에서의 학술적 쓰기에 아직 익숙하지 않은 한국인 대학생들에게 전범으로서 활용될 수도 있을 것이다.

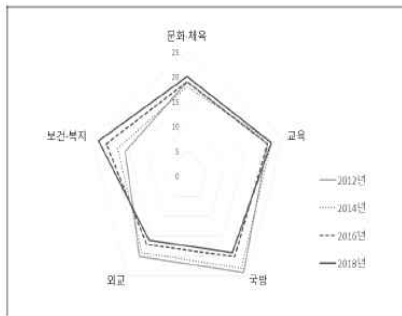
논술시험 (인문 1)

[문제2] <사례1>과 <사례2>에 각각 제시된 두 국가 A, B의 자료를 해석하고, 각 사례가 [문제1]의 어느 입장을 지지하는지 설명하십시오. (40점)

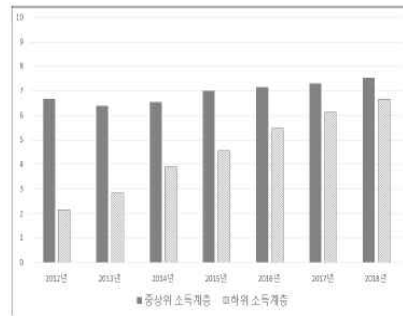
<사례1>

아래의 <그림1-1>과 <그림1-2>는 각각 A국의 예산지출의 항목별 비중과 소득계층별 행복지수를 보여준다.

<그림1-1> A국의 예산지출의 항목별 비중(%)



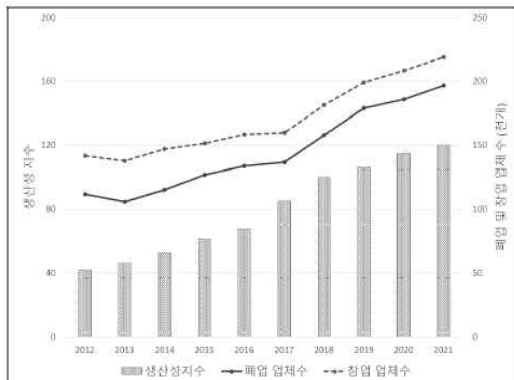
<그림1-2> A국의 소득계층별 행복지수



- 주1) 모든 연도에 A국의 예산지출 총액은 동일하고, 보건·복지, 문화·체육, 교육, 국방, 외교 이외의 예산지출 항목은 없다.
- 주2) 모든 연도에 다른 모든 조건은 동일하다.

<사례2>

아래 자료는 B국 소비자 산업의 생산성 지수와 사업자 폐업 및 창업 현황을 보여준다. B국은 2018년부터 소비자 산업 내의 신규 진입을 계약하는 인허가 및 등록 등에 관한 모든 규제를 폐지하였다.



- 주1) 생산성 지수는 노동 및 자본 투입 대비 부가가치의 비율로 정의하며 2018년을 기준연도(2018=100.0)로 하여 측정한다.
- 주2) 모든 연도에 다른 모든 조건은 동일하다.

<그림 7> 2022학년도 성균관대학교 수시모집 논술우수전형 논술 시험 기출 문제(성균관대학교 입학처, 2023.6.20. 검색)

본고에서는 이와 같은 학술적 데이터 설명문에 대한 이전의 학습자 경험에 한국인 저학년 대학생들에게는 대입 논술 문제의 영향으로 나타나고, 외국인 대학생들에게는 TOPIKⅡ 53번 평가 문제로 나타난다고 본바 있다. 따라서 이동 마디 순서 유형에 선택에 대한 세부 학습자군의 변인별 특성은 다음과 같이 정리할 수 있다.

1) 중급 중국어권: AB 유형 구성을 선호한다. 보편적인 중국어권 내러티브 텍스트, 또는 술직 텍스트에 C 유형이 반드시 존재하는 것으로 보이는 중국어 텍스트 구조의 영향은 아닌 것으로 판단된다. 주로 저학년 학습자들이 AB 유형을 사용해 숙달도가 낮고 학술적 텍스트 노출도가 낮은 학습자들이 TOPIKⅡ 53번 유형을 따라 쓰고 있는 것으로 볼 여지가 많다.

2) 고급 중국어권: A-()-C 유형 구성을 선호한다. 다만 AB를 포함한 C 누락 유형도 일부 사용하나, 이는 저학년 학습자에게서 주로 나타난다. 한국어로 제출하는 ‘과제’에 C 이동 마디가 존재함을 알고 사용하려는 것으로 보인다.

3) 중급 베트남어권: C 누락형은 3건 정도 나타나나 일반적으로 A-()-C 유형 구성을 선호한다. 이러한 양상은 베트남어권 학습자들이 주로 고학년이어서 한국어 과제에 대한 노출도가 큰 데서 기인한 것으로 보인다. 다만 숙달도의 한계 때문에 다양한 텍스트 구성 유형을 시도하지는 않는다.

4) 고급 베트남어권: A-()-C 유형 구성을 선호하고 그 안에서 다양한 유형을 시도한다. 이 학습자군은 이동 마디 순서 유형의 분포가 다양해 개인차가 있다. 숙달도가 높고 학년도 높은 학생들이 많기 때문에 상위 장르의 엄격함이 존재하지 않는다면 새로운 유형의 학술적 텍스트 구조를 시도해 보며 이에 대한 탐색이 이루어지는 것으로 보인다.

따라서 학습자의 언어권 영향 요인에 대한 처치보다는 한국어 숙달도가 낮고, 학술적 텍스트 구조에 대한 경험이 부족한 저학년 학습자들에게 다양한 상위 장르에 내포된 데이터 설명문 구조를 명시적으로 노출할 필요가 있다. 특히 대부분의 외국인 학부생들은 한국 대학 입학 요건 때문에 TOPIK 53번 문제에 대한 지식을 이미 갖고 있어 이들 ‘시각적 데이터 순서대로 수치 나열’이 다른 장르에서는 어떠한 의미를 갖고 어떻게 변용되는지를 맥락 내에서 명시적으로 지적해야 한다.

또한 숙달도가 높은 학습자에 대해서는 ‘A-()-C’ 유형에서 괄호에 들어갈 수 있는 구성이 상위 장르에서 허용 가능한 것인지 확인할 수 있도록 교수-학습할 필요가 있다.

1.2. 학습자 변인별 이동 마디 사용 빈도의 특성

(1) 주제 및 방향 요약에서 언어권별 차이의 해석

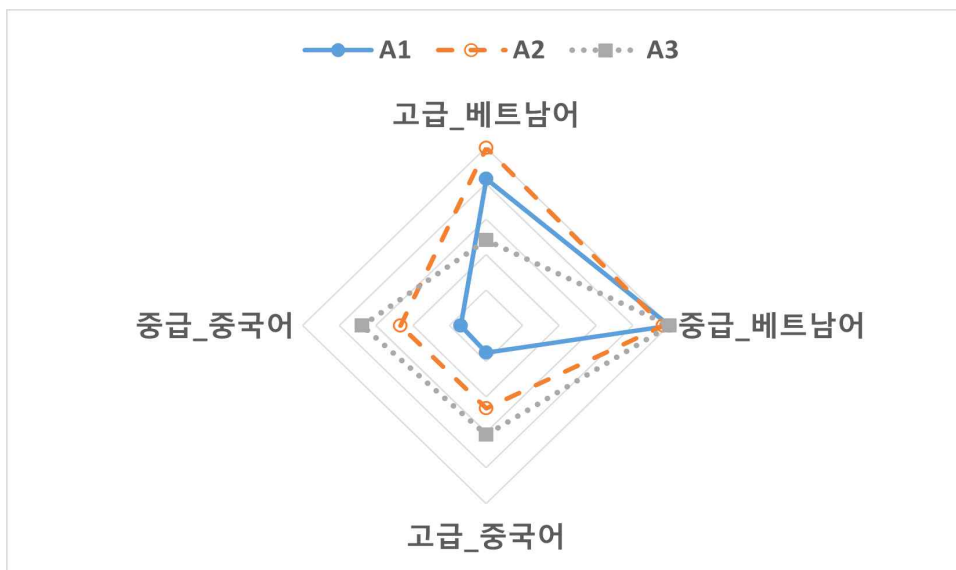
학습자 변인별 집단에서 이동 마디 단계 사용 빈도에 유의한 차이가 있는 것들을 앞서 III장에서 t-test를 통해 확인한 바 있다. 그 결과를 유의한 것들만을 뽑아 중급 중국어권, 고급 중국어권, 중급 베트남어권, 고급 베트남어권 학습자군으로 나누어 사용 빈도의 평균을 정규화하였다. 즉, 가장 높은 빈도를 보인 값을 1로 두고 나머지를 그 최빈값으로 나누어 정규화하였다. 이는 모수와 빈도가 서로 다른 두 집단을 같은 기준으로 비교하기 위한 것이다. 즉, 그래프의 끝에 수치가 가 있다 하더라도 이는 절대적인 값이 높다가보다는 다른 집단과의 비교를 했을 때 가장 높은 값을 가진 것으로 생각하면 된다. 그 결과는 아래 <그림 8>과 같이 분석할 수 있다.

1) 중국어권 학습자는 중급, 고급 모두에서 거의 A1을 사용하지 않는다. 반면 베트남어권 학습자는 A1의 선호가 극단적으로 두드러진다. 다만 앞 장에서 살펴본 바와 같이 A1 분포가 60%를 넘지 않아 고정성이 떨

어지는 것으로 보아 베트남어권 학습자 중 일부가 A1을 다수 사용함을 알 수 있다.

2) 이러한 언어권별 특성은 A2에서도 유사하게 나타난다. 베트남어권 학습자가 중급과 고급을 막론하고 A2를 모두 비교적 높은 빈도로 사용하고 있으며, 중국어권 학습자는 중급, 고급 모두 A2를 저조하게 사용하고 있음을 알 수 있다.

3) A3은 중급 베트남어권 학습자만 그 수치가 높고 나머지 세부 학습자군에서는 비교적 유사한 크기를 보이고 있다. 다만 언어권별, 숙달도별 차이에 대한 통계적 유의성이 있지는 않았기 때문에 단순히 이 네 세부 집단 중 중급 베트남어권 학습자의 A3 빈도가 제일 높았던 것이라 할 수 있다.



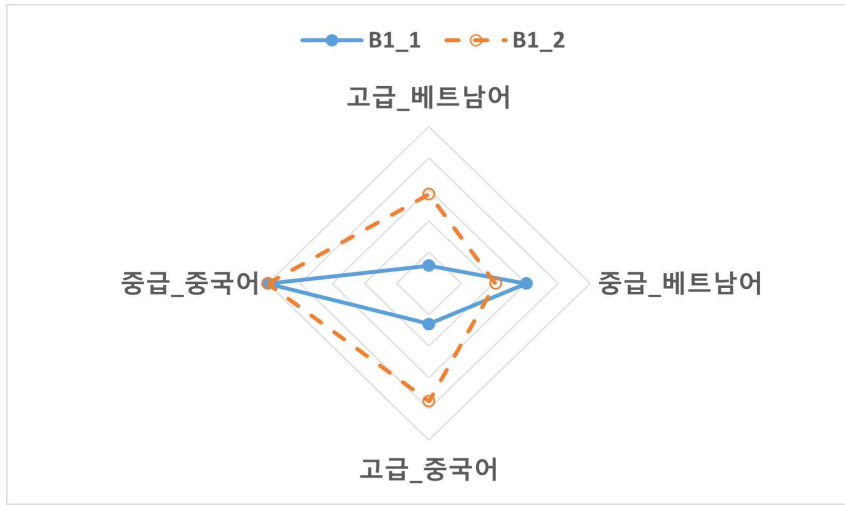
<그림 8> 세부 학습자군별 A 이동 마디 빈도

A1은 중급과 고급 베트남어권 학습자 모두에게서 다수 사용된다. 그러나 Ⅲ장의 분석에서 볼 수 있듯이 이러한 경향은 학습자 모두에게 나

타나는 것이라기보다는 일부 학습자들에게서 나타난다고 볼 수 있다. 이는 한국인 대학생들에게서도 나타나는 경향이다. 이러한 원인을 추정해 보자면 베트남어권 대학생들의 학년이 높은 것을 들 수 있다. 확인 결과, 중급 베트남어권 학습자 18명 중 17명이 3~4학년이고, 고급 베트남어권 학습자 32명 중 22명이 3~4학년이다. 전체 베트남어권 학습자 50명 중 1~2학년은 11명에 불과하다. 고학년 학생들은 저학년 학생들보다는 한국 대학에서 상대적으로 많은 학술적 텍스트를 접했을 것이라 생각된다. 이때 과제로 제출하는 텍스트에는 제목을 붙여야 한다거나, 보고서의 형식으로 소제목은 여러 개 붙여 제출해 본 경험이 있을 것으로 생각된다. 하지만 특정 언어의 영향이라면 한국어의 텍스트적 특징 역시도 제목을 붙이는 것이라고도 해석해야 할 것이다. 따라서 학년을 변수로 두고 분석하지는 않았지만, 그가 해석의 한 요인이 될 수 있을 것이다. 이러한 학년의 영향은 베트남어권 학습자의 A2의 다수 사용에서도 확인할 수 있다.

(2) 수치/결과 나열에서의 인지적 차이 확인

B 이동 마디 단계는 그 수가 매우 많아서 B1, B2, B3과 B4 세 부류로 묶어 다루겠다. 우선 B1의 하위 이동 마디인 'B1-1. 연대기적 나열'과 'B1-2. 시각적 데이터 순서대로 나열'은 인지적 노력이 적게 들어가는 단계이기도 하고, 특이한 점을 뽑아내기 위해 데이터의 전체적인 추세를 한 눈에 보여주는 데 사용하기도 한다. 따라서 고급과 중급 학습자들이 모두 선호하는 이동 마디이기도 하다. B1 단계의 학습자군별 하위 단계 선택 빈도의 차이는 다음 <그림 9>와 같이 정리할 수 있다.



<그림 9> 세부 학습자군별 B1 이동 마디 빈도

[B1의 하위 이동 마디]

1) B1-1은 고급 학습자들에게는 선호되지 않는다. 중급 중국어권 학습자의 사용 빈도가 가장 높고 그다음을 중급 베트남어권 학습자가 차지한다. 따라서 중급 학습자가 고급 학습자에 비해 B1-1의 사용에서 유의한 결과를 내는 것은 중급 중국어권 학습자의 경향성 때문이라 할 수 있다. 따라서 B1-1은 숙달도가 낮고 학년이 낮은 학습자들이 주로 사용하는 경향이 있는 것으로 보인다.

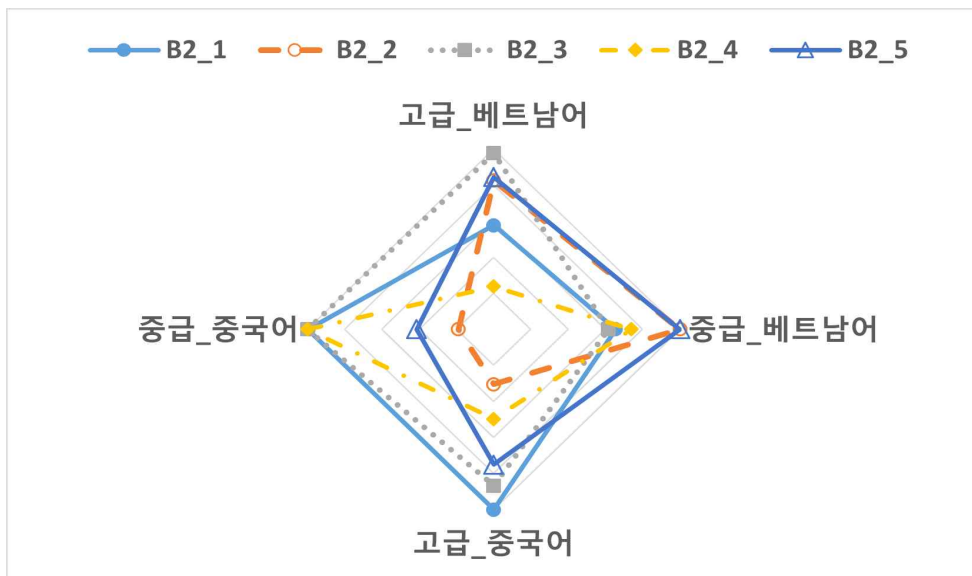
2) B1-2는 베트남어권 학습자들이 적게 사용하였고, 대체적으로 중국어권 학습자들의 사용 빈도가 비교적 높았다. 특히 중급 중국어권 학습자가 가장 높은 빈도를 보임을 알 수 있다. B1-2는 B1-1보다는 고급 학습자들의 선호가 높음을 알 수 있다.

이렇듯, 한국어 숙달도가 낮은, 그리고 학년이 낮은 학습자들이 B1 단계를 선호하는 까닭은 이들 서술 방식이 인지적으로 제일 난도가 낮기 때문에 서술 방식보다는 한국어 사용에 인지적 노력을 더 사용하기 위함

이라고 분석한 바 있다. 그러나 이들 B1 이동 마디는 또한 전체 수치를 나열함으로써 뒤에 나올 수치적 계산이나 특이점을 논하게 만드는 전형적인 장을 만들어 주는 역할을 하기 때문에, B2의 하위 단계를 사용하는 학습자들도 선제적으로 이를 사용하기도 한다. 즉, B1과 B2를 동시에 사용하는 경우도 있다.

[B2의 하위 이동 마디]

- 1) B2-1은 중국어권 학습자 전체가 선호한다.
- 2) B2-3은 중급 베트남어권을 제외하고는 모든 집단에서 다수 사용한다.
- 3) B2-4는 언어권 관계없이 중급 학습자의 선호가 두드러지나, 고급 학습자 사이에서는 베트남어보다는 중국어 화자가 더 선호하는 것으로 나타났다.
- 4) B2-2와 B2-5는 고급 학습자와 베트남어권 학습자들이 선호함을 알 수 있다.



<그림 10> 세부 학습자군별 B2 이동 마디 빈도

B1의 하위 단계들은 인지적 난도가 낮아 한국어 숙달도가 낮은 학습자들이 작업 기억의 용량을 보존하기 위한 수단으로 쓰기도 하고, 데이터 수치 나열의 다른 방법과 그 의미를 알지 못해 사용하지 못하는 경우도 있다. B2의 하위 단계들은 데이터를 한 번 더 가공해 표현한다는 점에서 B1보다는 조금은 인지적 난도가 높은 편이다. 다만 그 중에서도 ‘B2-1. 수치가 높은/낮은 순서대로 나열’, ‘B2-3. 특성에 따른 그룹화’는 시각적 데이터를 직접 보면서 수치의 높고 낮음이나 범례를 그룹명으로 쓸 수 있다는 점에서 인지적 노력이 다른 것들보다는 적은 편이고, ‘B2-4. 제일 큰/작은 수치에 주목’, ‘B2-5. 특이성/중요 지점 지적’은 뒤에 이 지적에 대한 해석이나 다른 수치 자료 연결이 있는 경우에는 인지적으로 노력을 요하나 단순 지적은 B2-1, B2-3과 같이 보고 쓸 수 있다는 점에서 비교적 난도가 낮다. B2-2는 제시된 수치들의 계산을 통해 해당 수치의 위치를 결정한다는 점에서 인지적 난도가 가장 높다고 할 수 있다.

따라서 B2-1과 B2-3은 전반적으로 모든 세부 학습자군에서 선호하고 자주 사용하나, 학년이 낮은 중국어권 학습자들이 주로 사용하는 경향이 있다. B2-2와 B2-5는 인지적 노력이 다른 하위 단계에 비해 다수 필요하다는 점에서 고학년이 많은 베트남어권 학습자와 숙달도가 높은 고급 학습자들이 주로 선택하는 것으로 보인다. B2-4는 숙달도와 학년이 낮을수록 주로 선택하는 것으로 보인다. 제일 큰/작은 수치에 주목한 뒤 이것이 어떠한 의미가 있는지, 다음 내용과 어떻게 이어지는지에 대한 서술이 있다면 B2-5와 같이 고급, 고학년 학습자가 다수 사용했겠으나, 단순히 시각적 데이터에서 두드러지는 수치를 언어로 변환만 한 것이어서 숙달도가 낮고 저학년인 학습자들이 주로 사용한 것으로 보인다.

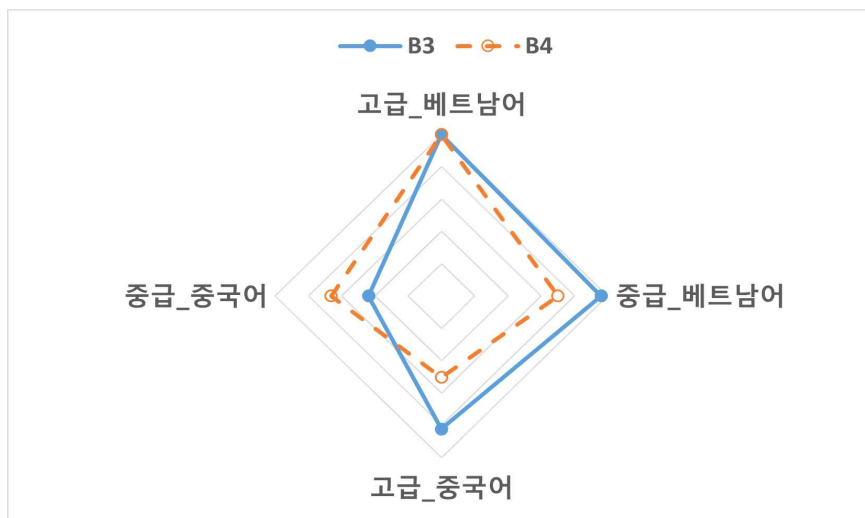
[B3과 B4]

1) B3. 시각적 데이터 언급 표지 사용

B3 사용 빈도는 베트남어권 학습자와 고급 학습자에게 높게 나타난다. 이는 고학년, 고숙달도 학습자들이 B3 사용을 자주 한다는 것을 알 수 있다. 이들은 시각적 데이터를 언급하여 텍스트적 응집성을 만들고, 독자의 시선을 움직여 가며 수치를 설명할 수 있다는 메타 기능을 인지하고 있는 것이라 할 수 있다.

2) B4. 수치의 의미 제시

한편 B4는 모든 그룹에서 비슷하게 사용되지만, 고급 베트남어권 학습자에게서 대단히 우세하게 나타난다. 숙달도와 학년이 높아지며 앞에 나열한 수치의 의미를 비수치적으로 변환하는 것을 인지하고 있다고 볼 수 있다.



<그림 11> 세부 학습자군별 B3, B4 이동 마디 빈도

(3) 해석 및 결론에서의 숙달도 요인 부각

1) C1은 중국어권 학습자들보다 베트남어권 학습자들의 사용이 더 많음을 알 수 있고, 고급 중국어권 학습자의 빈도는 중급 중국어권 학습자의 빈도보다 조금 더 높은 것을 확인할 수 있다. 이는 C1 사용의 필요성을 숙달도가 높거나 학년이 높은 학습자들은 인지하고 있다는 뜻이 된다. 그러나 숙달도도 낮고 학년도 낮은 학습자들은 이러한 구조에 대한 노출이 낮거나 노출이 되어도 모르는 경우가 있어 TOPIK에서 습득했던 C 누락 구조를 그대로 사용하는 것으로 보인다.

2) C2는 고급 베트남어 학습자에게서만 그 사용을 확인할 수 있다. 외국인 대학생의 C2 사용은 2개에 불과하여 외국인 대학생들은 데이터에 비판적 분석과 해석을 덧붙여야 한다는 것 자체를 인식하고 있지 못하는 것으로 나타났다. 한국인 대학생도 이러한 비판적 인식까지 나아가는 학습자는 얼마 되지 않았다. 상위 장르에 따라 C2가 나타나지 않아야 하는 경우도 있는데, 주로 특정 주장을 위해 자신의 데이터 또는 타인의 데이터를 선택적으로 사용한 경우가 그러하다. 따라서 이제까지 학습자들이 접했던 학술적 데이터 설명문에 C2에 해당하는 비판적 해석이 숨겨져서 이러한 구조를 인식하지 못했을 가능성이 있다.

3) C3은 고급 학습자와 중급 베트남어 학습자가 주로 사용하며, C1과 같은 모양이지만, 조금 더 범위가 넓은 것으로 보아 그 수치가 더 큼을 알 수 있다. 이 역시도 C1과 같이 해당 구조에 대한 노출이 저학년 및 저숙달 학습자에게 이루어지지 않아 선택이 저조함을 추정할 수 있다.



<그림 12> 세부 학습자군별 C 이동 마디 빈도

2. 학술적 데이터 설명문 언어 사용에서의 학습자 변인별 특성

2.1. 상위 장르의 어휘문법 레퍼토리

이 장에서는 III장에서 분석한 이동 마디 단계별 어휘문법 사용 분석을 몇 가지 부류로 묶어 정리해 보았다. 우선 상위 장르의 어휘문법 레퍼토리 부족은 데이터 설명문이 속해 있는 상위 장르의 어휘문법 사용을 파악하지 못한 데서 발생한 어휘문법적 사용들이 속한다. 본고에서는 ‘수업 과제’라는 상당히 모호한 양식의 상위 장르를 제시했기 때문이기도 하겠지만 학습자들이 인지하는 학술성의 범위는 상당히 넓었다.

A1의 경우 학술적 텍스트의 ‘제목’이기 때문에 명사형으로 끝나는 것이 바람직함을 알 수 있다(이윤진, 2016:258). 외국인 학습자들은 주로 주어

진 과제, 표, 그래프의 제목을 그대로 사용하는 경향이 높아 유사한 명사형 제목을 사용하나, 한국인 학습자의 경우에는 제목을 주어진 것과 달리 붙이려 하는 것이 눈에 띈다. 특히 소제목이 있는 경우에는 더 그러하다. 다만 명사형 제목이 아니거나 다른 장르에 적합한 제목을 붙이는 것은 외국인과 한국인을 가리지 않고 저학년에서 볼 수 있다. 이는 주어진 제목보다는 데이터를 이용해서 새로운 글을 구성하고자 하는 인식이 있으나, 학술적 제목들의 특성을 어휘문법적으로 파악하지 못해 선택의 문제가 있었던 것으로 생각된다. 일례로, 한 중국어권 학습자의 제목은 명사로 제목을 끝낸 형태이지만 앞에 부사를 함께 사용해 결국에는 문장을 명사화한 것으로 파악되었다. 이러한 구성은 단 한 번 출현했기에 중국어권 학습자의 전반적 경향이라고는 할 수 없으나, 다른 언어권 학습자에게서는 이를 확인할 수 없었다. 이러한 제목 구성을 사용한 학습자는 숙달도로서는 고급이나 학년이 2학년인 것을 고려했을 때 학술적 텍스트의 제목을 구성하는 데에는 익숙하지 못하거나, 좀 더 눈에 띄는 제목을 제시하고자 하는 의도로 다른 문체를 시도했을 가능성이 있다.

또한 ‘-ㅂ니다’의 사용은 B1 단계의 중국어권 학습자군에서 발견되었지만 모든 데이터 설명문 전반에서 발견된 흔한 종결어미 사용이다. TOPIKⅡ 글쓰기 시험을 본 학습자들은 문어체의 종결어미가 ‘-(ㄴ)다’라는 것을 알고 있지만, 시험의 상황과 과제 제출의 상황을 매우 다르게 인식했을 수 있다. 학습자들은 TOPIK 시험지를 불특정 평가자에게 제출하기 때문에 필자와 독자의 거리가 매우 사무적이고 멀다고 생각해 문어체 종결어미인 ‘-(ㄴ)다’를 사용하는 것을 수용했을 수 있다. 그러나 수업 과제물은 한 학기 동안 같은 강의실에서 자신을 가르치는 교수자에게 제출하기 때문에 과제물 작성을 교수자와 학습자의 가까운 의사소통으로 보는 학습자들은 수업 상황인 것처럼 ‘-ㅂ니다’체를 사용하는 것으로 보인다. 즉, 격식적 구어체를 사용하는 학습자들은 텍스트 구성 시 고려하는 대인적 거리를 TOPIK 시험과 수업 과제에서 달리 두고 있는 것이다. 독자가 누구인지를 고민하고 그에 맞추어 종결어미를 사용하려는 것은 학습자의 사회언어학적인 인식이 반영된 긍정적인 현상이라 보이지

만, 한편으로는 문어 학술적 텍스트 작성에서는 ‘-(ㄴ)다’ 종결어미 외에는 선택권이 없음을 숙지하지 못한 탓이기도 하다(장경현, 2006:47-54; 이슬비, 2016:66).

따라서 일반적인 학술적 문어 텍스트의 종결어미는 기본적으로 ‘-(ㄴ)다’임을 주지하게 할 필요가 있으며, 독자의 수에 따른 독자와 필자의 거리, 상위 장르에 따라 문어체와 구어체가 학술적 텍스트에서 나타나는 양상이 다양함을 학습 필자에게 제시할 필요가 있다. 이와 더불어, 과제를 제시할 때 작성해야 하는 학술적 텍스트의 종결 형태를 명시적으로 지정하는 것이 효과적일 수 있다. 다만 전체 텍스트를 구어체나 문어체로만 구성한 학습자들에게는 이와 같은 명시성이 효과적일 수 있으나, 문어체와 구어체를 한 텍스트 내에서 혼용하는 것은 단순한 필자의 부주의와 미숙성으로 여겨진다. 이러한 부주의를 방지하도록 훈련하는 것 또한 필요하다고 본다. 이를 위해서는 생성하려는 실제 상위 장르에서 사용되는 제목의 형태를 수집하거나 상위 장르 전체의 종결 어미를 확인하게 하며 그 문체의 경향성을 확인하게 할 필요가 있다.

2.2. 기학습된 어휘문법의 전이와 강화

A3 이동 마디의 어휘문법 분석에서는 베트남어권 학습자가 본인을 표현하는 ‘우리’의 사용이 확인되었다. 이는 PHAN, HOANG MY THUONG(2020)에서 언급한 것처럼 베트남의 학술적 저자 표현 방식이 저자와 독자를 동일시하는 ‘we’에 해당하기 때문일 것으로 보고 있다. 또한 중국어권의 학술적 텍스트도 1인칭 복수형을 학술 논문에서 주로 사용하고 있는데, 탄나나 외(2022)에서는 객관성을 드러내기 위한 장치로 중국어권의 언어학 학술 논문에서는 저자를 복수화하는 ‘우리’를 주로 사용한다고 하였다. 또한 III장에서 확인한 바와 같이 ‘우리’의 사용은 저자와 독자를 담화 공동체로 동일시하는 의미로 사용할 때 한국어 학술적 텍스트에서 예외적으로 허용되며, 그마저도 상위 장르에 따라 불가능한

경우도 있다.⁴³⁾ 즉, 이 ‘우리’는 그 객관화 정도와 포함되는 사람의 종류 등이 다르기는 하지만, 동일한 형태가 중국과 베트남, 한국의 학술적 주어를 나타낼 때 사용되는 것을 알 수 있다. 따라서 학습자의 첫 ‘우리’ 사용은 모어의 전이었다고 해도, 교과서와 논문 등의 학술적이라 여겨지는 자료에서 한국어로 ‘우리’ 표현이 사용되었기 때문에 그를 사용이 가능한 것으로 판단하고 형태를 강화했을 수 있다.

그러나 이윤진(2014:233)에서는 학술적 텍스트에서 학습 필자의 미숙성을 드러내는 필자 지칭 표현을 ‘1인칭 대명사 또는 그에 상응하는 표현을 과다 사용(우리, 나)’으로 제시하고 있다. 따라서 대부분의 학술적 텍스트에서는 이들 1인칭 대명사는 주관성을 나타내기 때문에 선호하지 않는다. 복수의 저자를 ‘우리’로 표현하여 결국 저자를 드러내는 것과 독자를 같은 주체로 끌어들이는 ‘우리’는 형태가 같은 까닭에 ‘우리’ 포괄 대상을 한정하여 허용되는 장르와 사용역을 인식하기 어렵다. 한국어 학술적 텍스트라고 해서 객관성을 드러내는 장치가 모두 동일한 것도 아니고, 저자와 독자의 담화 공동체로서의 거리가 같은 것도 아니기 때문이다.

이윤진(2014:231)은 필자 지칭 표현의 선택은 학술적 글쓰기 능력을 보여주는 하나의 중요한 표지이기 때문에, 학문 사회의 진입 단계에 있는 학습 필자에게는 반드시 훈련이 필요한 항목이라 하였다. 학술적 텍스트는 논리성이 필요하며, 그를 충족하기 위해 필자가 자신과 논문을 분리하여 객관성을 유지하는 데 쓰는 것이 필자 지칭 표현이기 때문이다(박나리, 2008:172-173). 따라서 이는 학술적 글쓰기에 속하는 학술적 데이터 설명문에서도 동일하게 적용된다.

따라서 이런 형태적 강화를 막기 위해 특정한 상위 장르와 그 내포 장르에서 사용되는 전이 표현들을 맥락적으로 제시하여 학습자들이 해당장에서 독자와의 거리를 판단해 필자 지칭 표현을 사용할 수 있게 해야 한다. 그렇지 않으면 다음 30번 학습자와 같이 전공 과제로서의 데이터

43) 박나리(2008:181)에서는 자연과학계열 논문에서는 독자와의 유대감을 증폭하기 위한 시도를 하지 않기 때문에 ‘우리’와 같은 대명사를 사용하지 않는 경향을 보인다고 하였다.

설명문을 작성하는 데에도 필자 지칭 표현의 선택이 혼란스러울 것이다.

(ㄱ) 우리가 코로나 시기 가사 노동 시간에 대해 연구를 하였다.

[30-중급-중국어-경영3-여]

(ㄴ) 다음에 이 보고서에서 설명하고 분석할 것이다.

[30-중급-중국어-경영3-여]

C3에서는 텍스트 마지막에 사용되는 ‘기대되다’ 어휘의 전이를 살펴 보았다. 저자의 희망으로 텍스트를 마무리 짓는 것은 중국어의 일반적인 내러티브 텍스트에서 주로 사용되는 희망적 마무리 표현이 전이된 것으로 볼 수도 있고(제효봉, 2015), 한국인 대학생과 베트남어권 대학생에게서도 나타나는 것을 보아 학습자들에게 학습된 보편적인 글 마무리 방식일 수도 있을 것이다. 글의 마지막을 저자의 희망으로 제시하는 이러한 보편성은 모어의 구조에서 처음 전이되었을 수도 있다. 문제는 이러한 비학술적 텍스트의 마무리 경험이 학술적 텍스트에서도 반복되는 것이다. 이는 저학년인 한국인 대학생에게서도 나타난다. 이들의 문제는 고정적 양식이 희미한 본고의 과제와 같은 학술적 텍스트 작성 시 비학술적 마무리 방식이 적합한지 여부를 명확히 구분하지 못한다는 것이다. 따라서 이를 지적하고 교정할 수 있도록 개입할 근거가 없어 전이가 강화된다고 볼 수 있을 것이다.

저자 기대 표현이 텍스트 말미에 사용될 수 있지만, 학술적 텍스트에서는 이것이 앞서 내용과 저자의 거리두기로 구축한 객관성을 해친다는 데 있다. 저자의 기대를 사용하고 싶다면 충분한 완화 표현을 사용해야 한다. 그리고 이러한 방식은 상위 장르와 그에 내포된 장르들의 영향을 받으므로 이들 텍스트에서 어떻게 나타나는지를 확인하여 전이를 수정할 수 있도록 해야 한다.

2.3. 어휘 선택에서의 다양성과 차이

학술적 데이터 설명문이 학술 논문과 같이 고정적 양식을 지닌 상위 장르였다면 특정 이동 마디 단계에서 나타나는 어휘문법은 공통적인 것들이 다수 출현할 것이다. 그러나 상위 장르가 고정되지 않은 상태에서 작성하는 데이터 설명문은 그보다는 더 자유로운 구성과 어휘문법을 보여준다. 이러한 어휘문법의 사용은 오류, 공통 어휘문법 수렴 정도의 차이, 레퍼토리의 차이 등 다양한 양상으로 나타난다.

앞서 지적한 대로 학습자들에게 실험 과제로 제시한 학술적 데이터 설명문 쓰기는 ‘전공 수업 과제’라는 다소 양식이 모호한 상위 장르였다. 따라서 문체, 어휘에서 어떠한 선택을 해야 하는지 학습자들이 자의적으로 판단하는 경우가 많았다. 이에 따라 다양한 어휘문법이 나타났는데, B1-2와 같이 다수 학습자군에서 선택한 어휘문법 표현 목록에서는 말뭉치가 큰 만큼 학습자의 오류 양상도 함께 담고 있다. B1-2에서는 중국어권 학습자이든 베트남어권 학습자이든 수치를 나열할 때 사용하는 조사와 서술어 공기에 오류가 있는 것을 다음과 같이 볼 수 있다.

(ㄱ) 1시간~2시간은 402명을 있다.

[48-중급-중국어-경영1-여]

(ㄴ) 그리고 혼인 상태는 비혼자가 280명과 기혼자가 720명을 나뉘었다.

[고급-82-베트남어-경영2-여]

따라서 어휘문법 목록 비교 시 공통적으로 수렴된 ‘(수치)+조사’의 결합에만 집중할 것이 아니라 그 조사와 공기하는 서술어에 대한 올바른 접근도 필요하다. 이를 위해서는 목록에 나타난 어휘문법 앞뒤로 문장과 문맥을 확인할 수 있는 장치가 필요하다.

또한 공통 어휘문법 수렴 정도의 차이도 어휘문법 목록에서 나타나는 양상 중 하나이다. 한국인 대학생은 B4 단계를 외국인 대학생 못지 않게 많이 사용한 것으로 나타났지만, 그 어휘문법 목록은 외국인 대학생에 비해 짧은 것을 알 수 있다. 이는 두 가지를 추론해 볼 수 있는데, 학습

자들이 거의 동일한 어휘문법을 사용하여 B4단계를 구성하였거나, 학습자가 선호하거나 선택하는 어휘문법의 개인차에 의한 것으로 볼 수 있다. 학습자들이 해당 단계에서 얼마 이상 반복하여 사용하는 어휘문법이 아니라면 말뭉치에서 빈도를 만족하지 못하기 때문에 목록에 들어갈 수 없다. 이에 따라 실제 말뭉치 분석을 해 보니, 한국인 대학생들이 사용하는 어휘문법의 개인차가 큰 것으로 나타났다.

(ㄷ) 이 결과를 참고하건대 여성의 전통적인 역할로 인하여 여성이 가사노동을 하는 시간이 더 큰 추세를 보인다는 것을 알 수 있었다.

[12-한-경제4-여]

(ㄹ) 하지만 코로나 시기를 거치며 가사 분담 비율이 보다 균등하게 조정되고 있음을 데이터를 통해 확인 가능하다. [1-한-경영4-여]

B2-2와 C1에서는 중국어권 학습자 및 중급 학습자에게서 수치 비교 시 사용하는 어휘의 난도가 비교적 낮게 나타나는 것을 확인하였다. 즉, 숙달도가 높은 학습자이거나 베트남어권 학습자일 때 나타나는 ‘차이, 폭, 비율, 격차 및 ’할애, 고정관념’과 같은 어휘가 보이지 않는 것이다. 이는 언어 숙달도와 전공에 대한 숙련도가 높아져감에 따라 발생한 것으로 생각된다. 따라서 어휘의 측면에서 숙달도 및 전공 숙련도가 낮은 학습자들을 위해 해당 상위 장르에서 사용되는 어휘를 모아 놓은 보조 도구를 제공하면 고급 어휘에 대한 접근성이 높아질 것이라 생각된다. 따라서 이동 마디 단계의 의미와 서술 방식, 그에 요구되는 어휘문법적 요소가 상위 장르 맥락적으로 제시될 필요가 있다.

3. 학습자의 학술적 데이터 설명문 양상에 따른 교수-학습 제언

3.1. 학습자의 학술적 데이터 설명문 양상에 따른 쓰기 교수-학습

의 필요성

이 절에서는 앞에서의 학습자 데이터 설명문 분석을 토대로 상위 장르가 달라졌을 때 학술적 데이터 설명문을 어떻게 구성할 것인지를 살펴볼 것이다. 우선, 작성하고자 하는 장르의 텍스트 경험이 필요하다. 본고의 연구 결과 중급 중국어권 학습자의 담화 구조 특징으로 AB 유형이 있음을 알 수 있다. C 이동 마디의 해석이 제거된 이러한 형태의 데이터 설명문은 TOPIKⅡ 53번 문제 유형의 구성과 유사하다. 이는 중급 중국어권 학습자의 숙달도 및 영역 지식의 수준이 낮음으로 인해 자신이 작성해야 하는 과제와 가장 근접해 보이는 학술적 텍스트를 모방했기 때문인 것으로 추론된다. 다른 학습자의 담화 구조도 마찬가지이다. 베트남어권 학습자들은 고학년 학습자들이 많아 학술적 환경에서 접해 본 학술적 텍스트의 구조가 다수임을 추정할 수 있다. 또한 한국의 대학에서 과제를 제출한 경험도 있을 것이기에 일반적인 한국어의 학술적 텍스트가 지향하는 A-()-C형 구조를 택했을 가능성이 있다. 이렇듯 학습자들은 새로운 장르를 맞닥뜨렸을 때 기존에 자신이 경험했던 텍스트 구조 중 새롭게 작성해야 하는 텍스트와 가장 유사한 것의 구조를 모방하는 경향이 있다. 따라서 학습자들이 새로운 상위 장르에서 데이터 설명문을 구성하게 하기 위해서는 그들에게 새로운 장르의 데이터 설명문 구조를 경험할 수 있도록 해 주어야 한다. 또한 이러한 경험을 학습으로 구체화하여 ‘상위 장르가 다른 데이터 설명문 담화 구조’에 대해 인식하고 자신의 담화 구조와 어떠한 차이가 있는지 이해하게 해야 한다.

둘째, 상위 장르에 따라 담화 구조가 변화하는 것을 받아들일 수 있어야 한다. 일반적인 한 편의 학술적 텍스트를 구성할 때는 보통 서론-본론-결론 형태인 ABC 순서로 진행된다는 인식이 있는 것은 좋지만, 데이터 설명문이 다른 상위 장르에 내포될 때 이동 마디 구성과 체계, 순서가 매우 많이 변화될 수 있다는 유연한 태도를 가져야 한다. 이 과정에서는 필자의 마음대로 여러 이동 마디 구성을 사용할 수 있는 것이 아닌 상위 장르의 규칙을 면밀히 살펴 그것을 따라야 할 필요성을 인지시켜야

한다. 예를 들어, 학술지 논문의 데이터 설명문은 주로 결과와 논의 부분에 집중되기 때문에 서론에 해당하는 A3와 같은 데이터 관련 배경 현상을 언급하는 경우가 거의 없다. 또한 C2와 같이 자신이 제시하는 데이터에 대한 비판적 언급을 피하는 경우가 많으며, 만약 해야 한다고 해도 결론과 같이 데이터 설명문에 해당하지 않는 별도의 장에서 제시한다. 이렇듯 상위 장르에 따라 데이터 설명문의 위치와 구조가 변화할 수도 있음을 학습자에게 먼저 제시해야 한다.

셋째, 학습자의 작업 기억에의 부하를 고려해야 한다. 중급 학습자와 중국어권 학습자들이 선택하는 B1-2의 경우에는 최소한의 인지적 노력을 들여 시각적 데이터 순서대로 나열만 하는 것이라 학습자들이 작업 기억 용량을 자신이 부족하다고 여기는 담화 구조 구성이나 어휘문법적 선택 등에 우선적으로 할당할 수 있다. 그러나 상위 장르와 그래프, 표 등의 특성에 따라서는 B1-2의 사용이 바람직하지 않을 때도 있다. 따라서 이때에는 학습자들의 이동 마디 단계 선택을 바꾸어 주어야 한다. 다만 이를 위해 단기간에 한국어 숙달도를 향상시킬 수 없기 때문에 그래프와 수치를 놓고 인지적 부하가 큰 이동 마디 단계 구성을 별도의 연습을 통해 자동화하는 방안을 생각해 볼 수 있다. 또한 유사한 상위 장르의 데이터 설명문 이동 마디 분석을 통해 각 이동 마디별 어휘문법 말뭉치를 콘코던스 목록 형식으로 제시한다. 즉, 학습자가 참조할 수 있는 어휘문법을 제공함으로써 한국어 사용에 대한 인지적 노력을 줄여줄 수 있다. 다만 이러한 목록은 실제적 어휘문법이 들어가 있기 때문에 학습자에게 제공하기 이전 교수자의 설명과 편집이 반드시 필요하다. 주제와 관련된 논문의 말뭉치일 때는 전문 용어와 주제어도 함께 다루고, 그에 대해 학습자가 미리 알아오게 해 발표하게 하는 방법도 있다. 또는 주제와 관련되지 않은 말뭉치를 사용할 때에는 전문 용어와 어휘를 모두 빼고 문법적 구성에만 초점을 맞추어 수업을 구상할 수도 있다.

넷째, 상위 장르가 전제하는 독자와의 거리를 인식해야 한다. 상위 장르의 특성상 독자와의 관계가 변화되어 어휘문법을 달리 선택해야 하거나 내포 텍스트가 데이터 특정 지점에 대한 다른 해석 기술을 포함해야 한

다면 해당 맥락 내에서 적합한 어휘문법이 필요해진다. 이 역시도 위에서 언급한 콘코던스의 선택적 사용과 교수-학습을 통해 이루어질 수 있다. 이렇게 구축된 새로운 상위 장르의 데이터 설명문 콘코던스를 통해 필자 지칭 표현의 사용, 문어/구어체의 선택, 데이터의 기술 방법 등을 구체적으로 제시할 필요가 있다.

3.2. 학술적 데이터 설명문 교수-학습을 위한 원리와 그 적용

(1) 데이터 설명문의 장르 내포적 성격에 따른 교육 원리

앞에서 살펴본 바와 같이 학습자들은 학술적 데이터 설명문의 작성에 학술적 장에서, 또는 학술적 장과 인접한 곳에서 경험해 본 가장 유사한 텍스트의 구조를 빌리는 경향이 있다. 고정되지 않은 장르 내포적 텍스트의 구성, 시각적 기호와 언어적 기호의 동시적 사용으로 갖게 되는 의미를 학술성에 맞게 정확하게 표현하는 것 등은 명시적으로 교수-학습하기 쉽지 않아 장르에 적응하는 동안 학습자들이 혼란에 휩싸인다.

이에 더해, 학술성이 갖는 분과 학문적 전문성 역시도 학술적 데이터 설명문 작성에 영향을 미치는 요인이다. 이러한 요인은 언어 교육에서 통제하기 어렵기 때문에 관련 학과목과 어떻게 연계하여 학술적 데이터 설명문 교육을 할 것인지가 고민이 되는 부분이다.

사실 어떠한 장에서 장르성과 사용역에 적응하는 것은 모어 화자이든 비모어 화자이든 혼란이 있을 수밖에 없는 일이다. Knapp & Watkins(2005:120)에서는 이러한 혼란을 거쳐 “학습자의 언어 능력이 점차 신장되면 창조적으로 텍스트를 조정할 수 있게 된다.”고 하였다. 그러나 한국어 비모어 필자가 목표어로 장르적 쓰기를 할 경우 텍스트를 조정할 수 있는 언어적 자원이 매우 제한되므로 학습자의 언어 능력이 저절로 신장되기를 바라는 이러한 접근은 학습의 방치를 불러올 수 있다. 또한 어떤 장과 사용역에서 어떠한 언어 능력을 신장하게 할 것인가 역시도 신중하게 고려되어야 한다.

이에 학문 목적 제2언어/외국어 쓰기 교육에서는 목표어 텍스트가 갖는 장르성을 통제하여 그 안에서 보편적으로 사용되는 구조와 어휘문법적 사용을 교수하는 장르 중심 교수법에 집중하기 시작하였다. 특수 목적 영어 교육에서 장르 중심 교수법은 장르에 대한 여러 학문 분과의 논의에도 불구하고 고정적 형식이 있음을 상정한다. 이러한 고정적 형식을 알아내기 위한 방법 중 하나가 실제로 사용된 언어를 모아 놓은 말뭉치를 사용하는 것이다. Polio(2012:320)는 언어가 학습자들이 만드는 어휘와 구조의 관계를 통한 용법으로 학습된다고 보았다. 이 접근에서는 현저성이나 빈도처럼 어휘와 구조의 관계에 영향을 미치는 것을 알아내고자 하였다. 즉, 말뭉치를 특정 언어학적 근거를 대는 말뭉치 근거(corpus-based) 방식이 아닌, 실제 용법이 어떻게 활용되는지에 초점을 맞춘 말뭉치 기반(corpus-driven) 방식을 통해 자연스럽게 맥락적인 언어 사용의 특성을 알아보고자 하는 것이다(남길임, 2014)⁴⁴.

다만 이러한 학습의 대상 장르는 본고에서 언급한 ‘상위 장르’에 국한되는 경우가 많다. 학위 논문의 서론, 실험 보고서, 학술지 논문의 초록 등 분량과 형식이 한 편으로 분절되어 제시될 수 있다. ‘데이터 설명문’이라는 텍스트의 특성과 상위 장르의 성질이 결합된 경우는 고정된 양식을 확인할 수 없어 장르 중심의 자료 기반 학습에서는 잘 다루지 못하는 실정이다. 물론 상위 장르별 데이터 설명문의 특성을 제시하는 방법도 있다. 예를 들어, ‘경영학 학위 논문 실험 분석 장에서의 데이터 설명문 교육’이나 ‘화학 실험 보고서 결과 해석에서의 데이터 설명문 교육’과 같은 것이 그러하다. 그러나 이러한 세부적인 장르 교육은 대상자가 매우 적고, 연속성이 있는 교육과정으로 구성되는 것이 아니라 교육 내용과 연구가 일회성이 되어 버리는 것은 아닌지 우려된다. 따라서 본고에서는 학습자를 대상으로 하는 구체적인 세부 장르에서의 내포 장르 교육 내용과 과정 개발보다는 교수자가 다양한 상위 장르에서의 내포 장르 교육

44) 남길임(2014)에서는 coropus-based를 ‘말뭉치 기반’, corpus-driven을 ‘말뭉치 주도’로 번역해 사용하였으나 제2언어/외국어교육에서는 data-driven learning을 언어 자료 기반 학습(이동주 외, 2017; 주성재, 2019; 이인숙 외, 2023), 데이터 기반 학습(신범숙, 2020b)으로 번역해 사용하는 경우가 많아 based를 ‘근거’, driven을 ‘기반’으로 사용한다.

수행할 수 있도록 그 원리와 방안을 제시하는 데 초점을 두도록 한다.

이에 본고에서는 본 장의 제목을 구체적인 장르 교수-학습의 목적을 갖는 ‘교육 내용 연구’가 아닌 학습자의 특성에 따라 다양한 상위 장르에서 어떻게 데이터 설명문을 교육할지에 대한 ‘제언’으로 정하였다. 또한 본고에서는 각 세부 내포 장르에 대한 교수자들의 교수 방법을 범용적으로 제시하기 위해 장르 내포적 데이터 설명문 교육에 자료 기반 학습(data-driven learning, DDL)을 사용할 것을 제안하고자 한다.

(2) 자료 기반 학습의 방법과 데이터 설명문의 적용

1980년대부터 시작된 자료 기반 학습(data-driven learning)⁴⁵⁾은 말뭉치 기반 견해를 실제적으로 뒷받침할 수 있는 기술적 영역의 교수-학습 방법이다. 초기의 말뭉치 관련 언어 학습은 말뭉치에서 예문을 뽑아내거나 교재를 만드는 말뭉치 근거 활동에 가까운 방식으로 수행되었다. 그러나 Johns(1991:2)에서는 학습자가 직접 말뭉치를 활용하여 자신의 학습에 사용하거나, 말뭉치를 검색할 수 있는 소프트웨어를 활용하여 언어 학습에 도움을 받는 것을 자료 기반 학습으로 명명하였다. 즉, “학습자가 데이터를 연역적 또는 귀납적으로 분석하는 과정을 통하여 언어를 학습하는 방법(김동영, 2021:37)”을 자료 기반 학습이라 부를 수 있다. 이러한 정의를 따른 자료 기반 학습에서는 주로 온라인 말뭉치나, 쓰기 교정을 위한 소프트웨어, 콘코던스(concordance)⁴⁶⁾ 등을 사용한다(Fuentes, 2018:159; Heift et al., 2017:30).

이와 같이 영어 교육에서는 맥락적 언어 자원을 활용한 장르 기반 쓰기의 관점에서 자료 기반 학습을 사용하는 경향이 있다. 자료 기반 학습 관련 연구 중 Flowerdew의 논의는 거개가 학문 목적 논문 쓰기 장르와 관련이 되어 있다. 예를 들어, Flowerdew(2015)는 이공계열 논문의 ‘논

45) 여기서의 ‘데이터’는 앞에서 언급한 데이터 설명문이 지시하는 수치적 자료의 의미가 아닌 맥락적 언어 자료를 모아 놓은 ‘말뭉치(corpus)’와 동의어이다.

46) 사용자의 편의에 따라 말뭉치를 검색할 수 있는 소프트웨어에서 특정 어휘를 검색했을 때 나타나는 어휘의 앞 뒤 문맥을 뜻한다. 보통 다음의 [그림 1]과 같은 형태로 제시된다.

의(discussion)’ 장르를 교수하기 위해 콘코던스를 이용하는 방안을 제시하고 있다. 이에서는 두 번의 장르 관련 수업 후 저널에 게재된 소논문 말뭉치를 이용해 자신의 논문 초고를 수정하는 과정을 담고 있다. 이 과정에서 학문 목적 쓰기 프로그램에 적합한 말뭉치 범위와 도구를 특정한다.

영어 교육의 장르를 기반으로 한 자료 기반 학습 연구에 힘입어 한국어 교육에서도 점차 자료 기반 학습을 도입하려는 연구가 보인다. 신범숙(2020b)에서는 말뭉치를 활용해 한국어 학습자의 학문 목적 쓰기를 주장하였다. 이 논문에서는 학문 목적 한국어 학습자를 대상으로 하향식 텍스트 분석, 학습자 전공 학술 논문으로 구성된 말뭉치 분석과 정리, 이를 통해 확인한 구조 및 표현의 효과와 의미 논의, 쓰기 활동, 피드백, 수정으로 이루어지는 학습 절차를 제시하였다.

교수 방법에 대한 논의 이후 그 효과에 대해서도 측정된 연구도 존재한다. Cotos et al.(2017)에서는 장르 중심 쓰기를 위해 영어 말뭉치와 웹 기반 자동 쓰기 평가 엔진을 이용해 학습자가 직접적으로 말뭉치를 사용하는 것이 장르 학습과 쓰기 능력 향상에 영향을 주는지 알아보았다. 그 결과 장르 지식을 잘 모르는 초급 모어/비모어 학습자가 장르 관습을 탐구하고 적용할 수 있게 되었으며, 장르 지식의 수사적, 형식적, 과정적 발달을 강화할 수 있게 되었음을 확인하였다. Fuentes(2018)에서는 학문 목적 영어 학습자에게 자료 기반 학습을 통한 수업을 제시하고 그 전후에 테스트를 시행한 결과 통계적으로 어휘 인식과 습득이 일어났음을 확인하였다. 이에 더해 면담, 교실 토론, 주관식 설문 등을 이용해 자료 기

| 앞 문맥 | 중심어 | 뒤 문맥 |
|----------------------|------|--------------------|
| ... 있다"고 규정하였으나 OECD | 데이터에 | 따르면 2014년을 기준으로 한국 |
| 그러나 이 기술은 로봇이 | 데이터가 | 있는 소설 같은 경우에만 문장을 |
| 또한 로봇이 | 데이터가 | 의존하기 때문에 소설 같은 독창 |

<그림 13> 국립국어원 한국어 학습자 말뭉치에서 나타난 ‘데이터’ 어휘 콘코던스

반 학습이 실제 학술적 쓰기에 도움이 되었는지 정성적으로도 파악해 보았다. 실험에 참여자 대부분은 자료 기반 학습이 전문 영역 쓰기에 발전적인 영향을 준다고 답변하였다.

그러나 학습자가 직접 말뭉치 소프트웨어를 사용하는 방식으로 수행되는 자료 기반 학습만이 효과적인 것만은 아니다. Heift & Vyatkina(2017:38)에서는 자료 기반 학습을 직접적(hand-on)인 것과 간접적(hand-off)인 것으로 나눈다. 직접적 자료 기반 학습은 더 자율적이고, 직선적이며 컴퓨터를 기반으로 한 것이고, 간접적 자료 기반 학습은 덜 자율적, 우회적, 종이 기반이라는 점에서 차이가 있다. 이 두 방식의 효과는 다양한 선행 연구들에서 제시하고 있는데, 이를 정리한 Heift & Vyatkina(2017:39)에서는 일반적으로 특정 언어 단위를 바로 확인할 수 있는 직접적 컴퓨터 기반 말뭉치 검색은 고급 학습자들에게만 효과적이거나, 풍부한 목표어 입력을 확인할 수 있는 간접적 종이 말뭉치 목록은 숙달도에 관계없이 모든 학습자들에게 효과적임을 밝히고 있다. 이는 학습자들의 목표어 숙달도에 따라 말뭉치 검색에 사용할 수 있는 목표어 자원이 다르고, 말뭉치의 효용이 다수의 언어 자원을 맥락적으로 접하는데 있기 때문인 것으로 분석된다. 관련 실험 연구들에서도 자료 기반 학습의 효과에 대해서는 긍정하나, 숙달도가 고급이 아닌 학습자들에게 사용하기에는 어려움이 있다는 결론을 내놓고 있다(Yoon & Hivela, 2004; Boulton, 2010; Xue-hua & Ming-ying, 2013; Fuentes, 2018).

따라서 본고에서도 학술적 데이터 설명문 작성에 말뭉치 자료가 상위 장르의 특성과 그 안에서의 데이터 설명문 양식 파악에 큰 도움이 된다고 보고 이를 사용하는 방법을 제안한다. 그러나 말뭉치를 학습자가 스스로 이용하는 데에는 무리가 있다고 판단하였다. 모범 텍스트를 단순히 학습자에게 제공하는 것만으로는 적절한 말뭉치 검색(corpus consultation)이 이루어질 수도 없고, 학습자가 무엇을 어떻게 검색해야 할지도 알 수 없기 때문이다. 즉, 모르는 단어를 검색해서 그 맥락을 알아보아야 하는데, 그 모르는 단어를 어떻게 검색할 수 있을지에 대한 대안이 없다. 이에 교수자가 개입해 말뭉치 자료를 가공, 편집해 학습자에

게 제공하는 말뭉치의 간접적(hand-off) 활용 방식이 모든 숙달도의 학습자에게 더 효과적이라 본다. 즉, 교수자가 데이터 설명문의 경계와 특징에 대해 인식하고 그를 학습자들에게 노출하며 적절한 설명과 연습을 병행해야 학습이 이루어질 것이라 판단하였다.

그러나 데이터 설명문은 단독으로 장르를 형성하는 텍스트 유형이라기 보다는 다른 장르에 내포되어 나타나는 특성이 있기 때문에 텍스트 생산자와 사용자가 안정적으로 인식할 만한 거시 구조, 정규적 순서, 또는 수사적 변화는 발견하기 어렵다(Guinda, 2011:116). 따라서 보편적으로 적용될 수 있는 데이터 설명문, 또는 데이터 설명문의 이동 마디에 천착하기보다는 이미 상위 장르에 내포된 상태를 통해 그 구조를 밝혀내어야 한다.

따라서 본고에서는 앞에서 살펴본 바와 같이 대학생인 외국인 학습자의 데이터 설명문 작성 양상을 토대로 이들이 이미 갖고 있는 텍스트 구성 지식을 새로운 상위 장르에서 변용하는 방법을 자료 기반 학습을 통해 제시해야 한다고 본다. 학습자가 이미 가지고 있는 지식을 바탕으로 새로운 영역에서 지식을 습득할 때는 이에 조력할 수 있는 교수-학습 방법이 비계로서 역할을 할 수 있기 때문이다. 본고에서는 학습하고자 하는 상위 장르의 데이터 설명문 말뭉치를 비계로 활용하는 교수-학습 방법이 학습자가 갖고 있는 내포 장르류 텍스트에 대한 경험에서 얻은 텍스트 구성 방식을 상위 장르에 맞게 적용하게 함으로써 쓰기 발달에 도움을 줄 수 있다고 본다. 즉, 학습자가 이미 갖고 있는 데이터 설명문 구성 지식을 통해 상위 장르의 관습적 속성에 접근할 수 있도록 돕는 것이다.

구체적으로, 목표로 하는 새로운 상위 장르의 데이터 설명문 구성 단계별 의미와 그의 표현 빈도를 언어 레퍼토리로서 제시하는 방법이 있다. 목표 상위 장르에서 추출한 데이터 설명문을 이동 마디별로 구분하여 말뭉치로 만든 후, 그 이동 마디의 의미와 각 이동 마디 내에서 자주 사용되는 어휘문법을 제시한다. 학습자들은 이를 자신의 데이터 설명문과 비교하여 담화 구조 및 어휘문법의 차이점을 인식하고 더 관습적인 표현에

주의를 기울이거나, 자신의 데이터 설명문을 스스로 첨삭할 수 있다. 특히 학습자가 목표로 하는 상위 장르에 익숙하지 않을 때에는 말뭉치를 전범으로 상위 장르 내에 어떠한 형식으로 데이터 설명문이 내포되는지를 파악하여 스스로 장르성을 인식할 수 있다. 이러한 교육을 위해서는 콘코던스 소프트웨어와 그를 사용할 수 있는 매체가 갖추어진 교실, 교수가 통제해 제시하는 말뭉치가 필요하다.

V. 결론

본 연구는 이제까지 한국어 교육에서 연구되지 않았던 데이터 설명문이라는 텍스트 유형을 조망하였다는 데 그 의의가 있다. 데이터 설명문은 상위 장르의 영향을 받은 장르 내포(genre embedding)의 성격을 가지고 있어 텍스트 순서를 나타내는 이동 마디(move) 단독으로는 텍스트의 특성을 밝혀내기 어렵다. 따라서 본고에서는 이동 마디에 더해 체계 기능적 다중 양식 담화 분석(Systemic functional-multimodal discourse analysis, SF-MDA) 이론과 데이터 설명문 평가에서 다루는 인지적 요인을 적용하여 구체적인 단계(step)를 파악하려 하였다. 분석의 대상인 데이터 설명문은 경제학 및 경영학 전공 학부 대학생에게 제공하는 수업 과제로 조건을 통제해 수집하였다. 이렇게 수집한 학습자의 데이터 설명문 특성을 파악하기 위해 각 학습자의 데이터 설명문에 위에서 개발한 이동 마디와 하위 단계를 부착하여 코딩한 후 반복적 내용 분석을 통해 이동 마디와 단계를 수정하였다. 그 결과 학습자들의 데이터 설명문에서도 출된 이동 마디는 'A. 주제 및 방향 요약, B. 수치/결과 나열, C. 해석 및 결론'이다.

이들 이동 마디의 순서 유형 및 이동 마디 하위 단계의 선택 빈도를 학습자 언어권 및 숙달도 변인에 따라 파악한 결과 중급 중국어권 학습자에게서 AB 순서 유형이 가장 높게 나타났고, 나머지 고급 중국어권, 중급 베트남어권, 고급 베트남어권 학습자군에서는 A-()-C 유형이 주로 나타났다. 일반적인 중국어 텍스트의 결론이 주로 C3에 해당하는 것을 볼 때 이러한 C 누락 구성에서 언어권의 영향을 추론하기는 어렵다. 따라서 C 누락 구성은 한국어의 학술적 텍스트 구조에 익숙하지 않은, 한국어 숙달도가 낮고 전문성이 낮은 학습자들이 TOPIK II 53번 유형의 글쓰기 구조를 차용하기 때문에 나타나는 것으로 생각된다. 이 추론의 근거는 고급 중국어권 학습자들은 더 이상 AB 유형을 사용하지 않는다는 데서 확인할 수 있다.

이동 마디 단계 빈도는 각 이동 마디 단계에 해당하는 구와 문장을 학

습자들이 얼마나 많이 사용했는지 알아보는 것이다. A 이동 마디와 하위 단계 분석에서는 베트남어권 학습자가 중국어권 학습자에 비해 ‘(소) 제목을 활용한 주제 제시(A1)’ 및 ‘연구의 주제와 내용 및 방법 설명(A2)’을 유의하게 많이 사용한 것으로 나타났다. 다만 ‘데이터 현상에 대한 배경 설명(A3)’은 집단 간 통계적 차이가 유의하지 않았다.

시각적 데이터를 수치적으로 제시하는 B 이동 마디의 하위 단계인 ‘연대기적 나열(B1-1)’은 중급 학습자가 유의하게 많이 사용하였고, ‘시각적 데이터 순서대로 나열(B1-2)’은 중국어권 학습자의 선택 빈도가 유의하게 높았다. 이들 B1의 하위 이동 마디는 한국어 숙달도가 낮은 중급 학습자들이 선호하는 것으로 보인다. 또한 대부분이 저학년인 중국어권 학습자들이 인지적 노력을 덜 소모하는 데이터 나열 방법을 사용하기 위해 B1을 선택했다고도 할 수 있다. 이에 더해 중급과 저학년 학습자들은 학술적 데이터에 노출된 정도가 고급 및 고학년 학습자들에 비해 적어 다른 데이터 기술 방식을 알지 못했을 가능성이 있다.

‘수치가 높은/낮은 순서대로 나열(B2-1)’은 순서를 배열해야 한다는 점에서 B1보다 인지적 노력이 조금 더 필요하지만, 시각적 데이터를 직접 보며 배열이 가능하다는 점에서 학습자들에게 인지적 부담이 크지 않은 듯하다. ‘특성에 따른 그룹화(B2-3)’ 역시도 B2-1과 마찬가지로 시각적 데이터에 나타난 범례명을 그룹으로 간주하여 그룹별 비교를 한다는 점에서 인지적 노력이 들어가지만 그 정도는 그렇게 크지 않아 보인다. 이 두 이동 마디는 데이터를 제시할 필요가 있을 때 숙달도를 막론하고 모든 집단에서 다수 사용하였고, 그 평균 차이도 유의하지 않았다. ‘제일 큰/작은 수치에 주목(B2-4)’ 단계는 중국어권 학습자의 선호가 두드러짐을 알 수 있다. 단순히 제일 눈에 띄는 수치를 제시하는 것은 역시 인지적 난도가 크지 않다. 그러나 B2-4나 ‘특이성/중요 지점 지적(B2-5)’의 경우는 이러한 두드러진 점을 지적하고 그를 해석하는 기술이 뒤따라야 하는데, 외국인 학습자의 작문에서 이러한 해석이 후행하는 경우는 거의 없었다. 따라서 인지적 노력이 필요한 부분을 누락하고 시각적으로 눈에 띄는 점만을 지적한 것을 알 수 있다. 이는 언어 숙달도가 낮아 이에 노

력을 쏟느라 인지적으로 부담이 되는 과제 수행을 누락했거나 이러한 구성에 대해 인지하지 못하여서 그러했음을 알 수 있다. ‘수치적 비교(B2-2)’는 인지적 노력이 앞의 단계들보다 더 필요하다는 점에서 고학년이 다수인 베트남어권 학습자들과 언어 숙달도가 높은 고급 학습자들이 주로 사용하였다. 이들은 한국어나 구조 사용에 있어 인지적 노력이 덜 소모되므로 데이터의 가공과 배열에 소모할 작업 기억 용량이 확보된 것으로 생각된다.

‘시각적 데이터 언급 표지 사용(B3)’은 베트남어권 학습자와 고급 학습자에게 유의하게 높게 나타난다. 이들 학습자는 시각적 데이터를 언급하여 텍스트의 응집성을 만들고, 독자의 시선을 움직여 가며 수치를 설명할 수 있다는 메타 기능을 인지하고 있기 때문에 해당 이동 마디를 적극 사용한다고 볼 수 있다. 한편 ‘수치의 의미 제시(B4)’는 모든 그룹에서 비슷하게 사용되지만, 베트남어권 학습자에게서 대단히 우세하게 나타난다. 단순 데이터 나열이 아닌 그에 대한 의미까지 언급할 수 있다는 데서 언어적 숙달도와 고학년의 영향을 확인할 수 있다.

해석 및 결론을 나타내는 C 이동 마디는 우선 중국어권 학습자의 C 누락 형태 선호로 베트남어권 학습자의 선택 빈도가 높았다. 특히 ‘주어진 데이터/상식적 지식에 근거한 해석(C1)’에서는 베트남어권 학습자의 선택 빈도가 유의하게 높았다. 또한 고급 학습자군의 C1 및 C3의 빈도 평균 역시도 중급에 대해 유의하게 높았다. 다만 ‘데이터의 비판점과 개선 방향(C2)’은 네 집단 모두 선택이 저조했고, 선택 빈도 차이 또한 유의하지 않았다.

데이터 설명문의 언어 사용은 이들 이동 마디와 하위 단계별로 말뭉치를 구축해 형태소 기준 3-gram을 추출하여 알아내었다. 각 이동 마디 단계에서 공통적으로 사용하는 어휘문법은 각 단계의 의미와 기능을 나타낸다. 이를 각 이동 마디 단계별로 나열하면 다음과 같다.

A1은 제목이어서 3-gram 추출에서 예외가 되는 항목이다. 따라서 전부를 수록하였는데, 중국어권 학습자는 주어진 표와 그래프, 과제 제목을 반복하는 경향이 크고, 그것이 반복되어 4개의 목록이 도출되었다. 베트남

남어권 학습자들은 소제목을 ‘1. 설문 응답자 기본 정보’와 같이 연번을 달아 사용하는 경향이 있으며, 표와 그래프의 제목을 그대로 차용하는 경우가 많았다. 중급 학습자와 고급 학습자도 표와 그래프 제목을 동일하게 사용한 경우가 많았는데, 중급에서는 맞춤법 오류가 일어나거나 연번을 사용한 경우가 등장했다. 이 단계의 어휘문법은 전반적으로 명사형으로 종결되는 형태를 보인다.

A2에서는 ‘(수치)년-’ ‘조사/실시/분석(하-)’와 같은 구성이 공통으로 나타났다. 중국어권, 중급, 고급 학습자에게서는 ‘-에 대해-’, ‘-를 대상으로’와 같은 어휘문법이 공통적으로 나타났다. 그러나 한국인 대학생들에게서는 이와 같은 어휘문법이 나타나지 않아 이들 목록은 외국인 대학생들이 연구의 주제와 내용을 설명할 때 자주 사용하는 것임을 알 수 있다. 또한 외국인 대학생들은 실제 수행된 설문조사나 실험이 적시하는 연도와 수치를 언급하며 조사나 분석을 행하였다. 특히 베트남어권 학습자들의 A2 이동 마디에서는 다른 학습자군보다 더 많은 공통 어휘문법이 나타난다. 따라서 이 이동 마디에서는 베트남어권 학습자들이 공유하는 어휘문법이 많고, 이를 공통적으로 다량 서술한 것을 알 수 있다.

A3 이동 마디에서는 중국어권, 고급, 한국인 대학생의 어휘문법 목록에 ‘-고 있-’이 공통으로 등장하였다. 이는 그래프에서 나타내는 현상이 현재 상황(코로나)에 대한 기술을 다수 포함하고 있기 때문이라 본다.

B1-1, B1-2는 주로 ‘(수치)년 (수치)시간/분’을 사용하여 어느 해에 몇 시간/분 가사 노동을 했는지 다루는 문장이 많다. B1-1보다 B1-2에서 나타나는 공통 어휘문법이 더 많으며, 특히 B1-2에서 중국어권의 목록이 다른 학습자군보다 긴 것을 알 수 있다.

B2 이동 마디는 대부분 ‘(수치)+단위 명사’, ‘나타나-’가 공통적으로 사용되었다. B2-1은 수치가 높거나 낮은 순서대로 나열하기 때문에 ‘그 다음, 그 뒤, -순으로’와 같이 다음에 오는 내용을 언급하는 어휘문법 및 ‘(수치)명-으로’와 같이 구체적 수치 제시를 위한 구, ‘차지하다, 있다, 나타내다’와 같은 수치 인용과 확인의 어휘가 공통적으로 사용되었다. B2-2의 경우는 수치적 비교가 들어가는 부분이어서 ‘비하다, -보다, -배’

와 같은 비교의 어휘문법이 눈에 띈다. B2-3은 특정 그룹의 수치적 특성을 지적하기 때문에 ‘(수치)+단위 명사’ 어휘문법 뒤에 수반되는 명사 및 조사가 다양하다는 특성이 있다. 또한 각 그룹의 수치적 특성에는 수치 변화도 포함되기 때문에 ‘감소하다, 늘어나다, 줄어든다, 증가하다, 상승하다’와 같은 변화를 나타내는 어휘 사용도 잦다. 또한 ‘-는 것 -을 알 수 있다/-으로 나타나다’와 같이 그래프에 나타난 정확한 수치를 인용해서 보고하는 객관적 태도를 나타내는 어휘문법을 구사한다. 외국인 대학생과 한국인 대학생의 차이는 이 객관적 태도를 나타내는 어휘문법에서 나타나는데, ‘-을 수 있-’에 선행하는 전성어미 활용과 ‘확인하다, 보다, 알다’ 등의 어휘 사용 다양성에서 차이가 있다. B2-4는 제일 큰/작은 수치에 주목하는 단계여서 ‘제일, 가장, 더 많-’과 같은 어휘문법을 공통적으로 사용하였다. B2-5는 모든 학습자군이 저조하게 선택해 ‘(수치)년’ 계열을 제외하고는 공통적으로 드러나는 어휘문법이 잘 보이지 않았다.

B3은 시각적 데이터를 인용하거나 언급하는 단계이다. 따라서 ‘-를 (살펴)보-, 나타내-, (자료 연번)과 같다’와 같은 어휘문법이 학습자군 모두에서 자주 사용되었다.

B4는 학습자 언어권별로 공통적 어휘문법이 다수 나타났다. 중국어권과 베트남어권 모두 매우 긴 목록을 갖고 있으나, 중급과 고급, 한국인 대학생은 그다지 목록이 길지 않았다. 이 단계에서는 수치의 의미를 제시하기 때문에 ‘-것을 알 수 있-, -는 것으로 보-’와 같은 수치에 대한 해석을 확인할 수 있다.

C1은 이제까지 나온 수치들과 상식적 배경 등을 종합해 필자의 해석을 제시하는 부분이다. 따라서 ‘-것으로 보이-, ‘나타내-/분석되-’와 같은 어휘문법이 공통적으로 사용되었다. C2는 말뭉치 자체가 매우 작아 공통적인 목록이 형성되지 않았다. C3는 ‘-어야 하-’를 주로 사용하는 경향이 보이나 중급 학습자는 ‘기대된다’와 같이 데이터 설명과 분석에 어울리지 않는 어휘도 사용함을 볼 수 있다.

정리하자면, 저학년이 대다수인 중국어권 학습자와 한국어 숙달도가 낮은 중급 학습자는 담화 구조에 있어 두 가지 특성을 보인다. 첫째, 학술

적 텍스트 구성에 대한 노출이 많지 않아 기노출된 학술적 데이터 설명문 유형인 TOPIKⅡ 53번의 구조를 주로 선택한다. 둘째, 인지적으로 노력을 요하지 않는 방식의 데이터 서술인 B1 및 B2-1, B2-3을 자주 사용하는 것으로 보인다. 이들의 어휘문법적 사용에 있어서는 해당 이동 마디 단계에 공통적으로 공유되는 어휘문법을 사용하나 그 목록이 상대적으로 다른 집단에 비해 길지 않은 것으로 보아 알고 있는 어휘문법의 수는 적은 것으로 생각된다. 이는 특히 중급의 경우에 고급 어휘 사용량이 부족한 것으로 확인할 수 있다.

반면 고학년이 많은 베트남어권 학습자와 한국어 숙달도가 고급인 학습자들은 상대적으로 학술적 텍스트 노출도가 높아서 그러한지 일반적인 텍스트 구조 방식인 서론-본론-결론 구조인 A-()-C 순서 유형을 주로 사용한다. 또한 베트남어권 학습자의 대부분이 고학년이어서 데이터 관련 내용의 이해나 분석이 빠른 것도 있고, 고급 학습자들의 한국어 숙달도가 높아 표현에 있어 상대적으로 수월하기 때문에 인지적 노력이 필요한 데이터 기술 방법을 사용하는 것으로 생각된다. 예를 들어, 이들 집단에서는 B2-2와 B2-5, B4와 같이 인지적 노력을 요하는 이동 마디 단계들이 유의하게 많이 사용된다. 베트남어권/고급 학습자들에게서는 또한 세부적으로 상위 장르에서 용인되지 않는 어휘문법을 사용하는 경향이 있다. 이는 절대적인 어휘문법의 부족이라기보다는 저자-독자 간 거리, 학술적 텍스트의 객관성 등을 나타내는 상위 장르의 맥락적 어휘문법 자원이 부족한 것 때문으로 생각된다.

위와 같은 학습자의 데이터 설명문 구성 양상을 볼 때, 학습자들이 학술적 데이터 설명문을 작성할 때는 조금 더 세밀한 장르 내포적 특성을 파악해 서술에 반영할 필요가 있음을 알 수 있다. 이를 위해 특정 상위 장르에 내포된 데이터 설명문의 특성을 파악하고, 상위 장르의 양식과 규범에 맞게 데이터 설명문을 구성할 수 있는 교수-학습 방법을 구안해야 한다. 이에 본고에서는 상위 장르의 맥락적 확인을 통해 데이터 설명문을 작성할 수 있는 자료 기반 학습(data-driven learning, DDL)을 제안하였다.

이상과 같이 본 연구는 학술적 텍스트의 내포 장르로서 데이터 설명문에 대한 분석을 처음 시도하였다는 의의를 지닌다. 그러나 장르 내포 자체에 대한 여러 현상 및 다양한 상위 장르 내에서의 데이터 설명문의 변용, 구체적인 교육 방안에 대해 다루지 못했다는 한계가 있다. 따라서 다음과 같은 후속 연구가 진행된다면 상위 장르에 따라 양식이 변화하는 내포 장르 그 자체에 대한 논의와 내포 장르 중 하나로서 데이터 설명문에 대한 논의가 더욱 발전할 수 있을 것이라 본다.

첫째, 상위 장르 변화에 따른 데이터 설명문의 변화를 확인해 보아야 한다. 즉, 다른 학문 분과 및 상위 장르에서도 본고에서 제시한 학술적 데이터 설명문의 이동 마디와 어휘문법의 특성이 공유되는지, 다르다면 어떤 부분이 다른지도 확인해 보아야 한다. 이는 학술적 텍스트 내에서도 여러 분과의 데이터 설명문을 분석해 보는 방법으로 추후 연구가 필요하다. 특히 영어권의 데이터 설명문 연구에서 주 분석 대상으로 삼았던 중학문(hard science) 계열 학술적 텍스트를 분석함으로써 본 연구에서와의 차이를 확인해 볼 필요가 있다. 한국어 교육에서 중학문 계열 연구의 필요성이 떨어진다는 판단을 장르 비교와 같은 다른 맥락에서 다시 생각해 볼 필요가 있는 것이다.

둘째, 데이터 설명문뿐만 아니라 장르 내포적 성격을 갖는 다른 텍스트 유형에 대해서도 고찰이 필요하다. 어떠한 요소가 학술적 텍스트에서, 또는 다른 장르의 텍스트에서 장르 내포적 성격을 갖는지 본고의 논의와 같이 상위 장르가 변화하였을 때 기존의 텍스트 지식을 기반으로 비계를 통해 다른 장르의 관습성을 습득할 수 있는지 여부를 확인해 보아야 한다. 이 과정에서 문자로만 구성되지 않은 다른 텍스트들을 어떻게 받아들일지에 대한 것도 후속 연구가 필요하다.

셋째, 학문 목적 학습자들의 전공 교육과 한국어 교육이 어떻게 융합되어야 할지를 더 고민해 내포 장르 텍스트의 연속적인 교육 과정을 구성해야 한다. 본고의 분석에서도 논의했다시피 대학생들의 작문 차이를 야기하는 영향력 있는 변인 중에는 학년이 포함되어 있다. 전공 지식의 심화와 사용역에 대한 노출도를 생각했을 때 이는 당연한 결과이겠으나,

대학생들의 전공 학년의 변화에 따른 한국어 교육 방법은 아직 논의된 바 없다. 연속적인 내포 장르 교육을 위해서는 다양한 전공과 학년의 외국인 학습자들에게 어떠한 한국어 글쓰기가 필요한지에 대한 연구도 마련되어야 한다. 연구의 마지막에 다루었던 자료 기반 학습에 대한 구체적인 적용 방안도 이러한 논의를 통해 상세화할 수 있다.

참 고 문 헌

- 교육부. (2019). 외국인 유학생 및 어학연수생 표준업무처리요령(2019 개정).
- 교육안전정보국 교육통계과. (2021). 2021년 교육기본통계 주요 내용.
- 구본관. (2014). 경제학 학술 논문 텍스트의 특성과 창의성. 텍스트언어학, 36, 41-77.
- 권혁승·정채관·김재훈. (2018). 코퍼스 언어학 기초. 한국문화사.
- 김금숙. (2019). 한국어능력시험 (TOPIK) 쓰기 완성형 문항 분석 및 개선 방안. 외국어로서의 한국어 교육, 55, 57-86.
- 김동영. (2021). 2021 KATE 한국영어교육학회 SIG Proceedings, 31-40.
- 김민혜. (2014). 학위 논문 <선행 연구>의 텍스트 구조와 서술 표현 사용 양상 연구 -한국어 모어 화자와 한국어 학습자의 비교를 중심으로-. 연세대학교 대학원 석사학위논문.
- 김진웅. (2019). 한국어 교육 분야 국문초록의 이동마디 구성 분석. 리터러시 연구, 10(2), 377-400.
- 김호정·강남옥·신필여·박민신. (2021). 한국어 교육에서의 ‘매개(mediation)’ 영역 설정의 필요성과 방향. 언어와 문화. 17(3), 51-87.
- 김현민. (2020). 한국어 학습자의 학술 보고서 인용 표현 사용 양상 연구. 서울대학교 대학원 석사학위논문.
- 남길임. (2013). 한국어 정형화된 표현의 분석 단위에 대한 연구. 담화와 인지, 4, -136
- _____. (2014). 이론으로서의 말뭉치언어학에 대한 연구 현황과 쟁점. 한국어 의미학, 46, 163-187.
- 남길임·최준. (2014). 학술적 텍스트에 나타난 핵심 구문의 추출-국어국문학 학술논문 말뭉치를 중심으로. 어문론총, 60, 65-92.
- 문은식·박선환·정미경. (2007). 교육 심리학. 공동체.
- 박나리. (2008). 학술논문의 텍스트성 분석-의도성(intentionality)을 중심으로. 텍스트언어학, 25, 163-192.
- _____. (2013). 학문목적 한국어 쓰기 학습자를 위한 학술논문 각주 분석 -텍스트 기능과 기능별 언어표현을 중심으로-. 작문연구, 17(0), 227-269.
- 박수연. (2015). “선행연구 검토(Literature Review)” 장르의 내용 구조에 관한 연

- 구 -한국어 교육 박사학위 논문을 중심으로-. 외국어로서의 한국어 교육, 43(0), 59-95.
- 박은선. (2006). 한국어 학위논문 서론의 장르분석적 연구 -한국어 모어화자와 한국어 학습자를 대상으로. 한국어 교육, 17(1), 191-210.
- 박지희. (2016). 학문목적 한국어 쓰기 교육을 위한 경영학 학위논문의 장르분석: 결론 부분을 중심으로, 이화여자대학교 대학원 석사학위논문.
- 소지영. (2020). 학문 문식성 신장을 위한 기능적 메타언어 관점의 문법 교육 연구. 서울대학교 대학원 박사학위논문.
- 송은정. (2018). 외국인 대학생의 한국어 담화 통합 쓰기 과제 표상의 양상 연구. 외국어교육연구, 32(2), 87-112.
- 송유경. (2021). 데이터 리터러시 향상을 위한 데이터 기반 토론 수업 모형 개발. 서울대학교 대학원 석사학위논문.
- 신범숙. (2020a). 한국어 (KSL) 학습자의 학업 문식성 향상을 위한 교과 언어 교육 연구-역사 교과를 중심으로. 서울대학교 대학원 박사학위논문.
- _____. (2020b). 데이터기반학습을 통한 학문 목적 쓰기 교육 연구. 새국어교육, 125, 135-161.
- 신영주. (2011). 한국어 화자와 중국인 한국어 학습자의 학위 논문 서론의 완화 표지 사용 양상 비교. 담화와 인지, 18(1), 63-77.
- 우서양. (2016). 중국인 학습자를 위한 한국어 설명문 쓰기 교육 방안 연구: 학문 목적 한국어 학습자를 중심으로. 중앙대학교 석사학위논문.
- 유기웅·정중원·김영석·김한별 (2018). 질적 연구방법의 이해. 박영스토리.
- 유민애. (2017). 맥락 중심의 한국어 담화문법 교육 연구-중국인 학습자의 내러티브 분석을 중심으로-. 서울대학교 대학원 박사학위논문.
- 윤세롬. (2019). 한국어 교육학 학위논문 서론의 이동마디별 다단어단위 분석 연구. 언어과학연구, 91, 353-373.
- 이관규·김서경·노하늘·성수진·신희성·유상미·이현주·정려란·정지현·정혜현. (2021). 체계 기능 언어학 개관. 사회평론아카데미.
- 이동주·김미숙·정보람. (2017). 언어자료기반학습(Data-Driven Learning)을 적용한 중학생 어휘 학습 효과 연구. 학습자중심교과교육연구, 17(18), 559-588.
- 이명화. (2017). 한국어 학습자의 토픽(TOPIK) 쓰기 오류 양상 - 53번 설명문 쓰기 중심으로 -. 한국언어문화학. 14(2), 231-254.

- 이선영. (2017). 외국인 학부생 대상 담화통합 과제 중심 쓰기 교재 개발 방안. 이
중언어학, 66, 49-77.
- 이선진·계은진·백준오. (2023). 과제 복잡성과 학습자의 인지적 요인 연구. 새국어교
육, 134, 133-169.
- 이슬비. (2016). 한국어 학술적 텍스트의 필자 태도 표현 교육 연구. 서울대학교 대
학원 박사학위논문.
- 이윤진. (2014). 학술논문에 나타난 “필자 지칭 표현”의 사용 양상. 작문연구. 제20
집, 231-266.
- _____. (2016). 외국인 유학생의 학술적 텍스트 제목 작성 양상에 관한 연구. 언어
와 정보 사회, 28, 247-274.
- 이인숙·이동주. (2023). 언어자료기반 학습법을 적용한 초등학교 6학년 영어 문장 쓰
기 학습 효과 연구. 외국어교육, 30, 33-55.
- 이준호. (2012). 학술적 텍스트에 나타난 한국어 헤지 표현 선정 연구. 이중언어학,
49(0), 269-297.
- 이지연. (2014). 이유 표현 난이도에 따른 한국어 문법 교육 내용 연구. 서울대학교
석사학위논문.
- _____. (2022). 한국어 데이터 설명문 교육을 위한 이론적 접근. 한국언어문화교
육학회 제34차 춘계학술대회 발표자료집.
- _____. (2023). 체계 기능적 다중 양식 담화 분석을 통한 경제학 및 경영학 학술
지의 데이터 설명문 담화 모형 연구. 사회언어학, 31(1), 149-172.
- 이효인. (2019). 텍스트구조 표지를 활용한 한국어 쓰기 교육 연구: 한국어능력시험
(TOPIK) 쓰기 영역을 중심으로. 학습자중심교과교육연구, 19(8), 665-684.
- 이태환. (2019). TOPIK II 합격레시피. 한글파크.
- 장경현. (2006). 국어 문장 종결부의 문체 특성 연구. 서울대학교 박사학위논문.
- 정혜승. (2012). 대화적 문식성 교육을 위한 상위담화(metadiscourse) 범주 구성
연구. 국어교육학연구, 43(0), 459-485.
- 제효봉. (2015). 중국어권 한국어 학습자의 쓰기 텍스트에 나타난 모국어 영향 연구.
서울대학교 박사학위논문.
- 조성해. (2022). 한국어 인문 사회 계열 석사 학위 논문 결론 담화의 구조와 응집성
분석 연구. 한국어 교육, 33(2), 227-257.
- 주성재. (2019). 초등학교 영어교육에서 자료기반학습의 적용 효과. 서울교육대학교

박사학위논문.

- 중화인민공화국 교육부. (2017/2022). 보통고등학교 어언과목표준.
- 최준·송현주·남길임. (2010). 한국어의 정형화된 표현 연구. *담화와 인지*, 17(2), 163-190.
- 최지희. (2022). 한국어 학술 정형표현 연구. 연세대학교 대학원 박사학위논문.
- 최혜민. (2017). 문항 특징을 반영한 쓰기 평가 기준의 설정 방향 - 한국어능력시험Ⅱ 쓰기영역 53번 문항의 `내용 및 과제 수행`을 중심으로. *한성어문학*, 36, 135-160.
- 탄나나·현영희. (2022). 한·중 학술 논문에 나타난 ‘필자지칭표현’의 사용 양상 연구 -언어학 분야의 학술 논문을 대상으로-. *외국어로서의 한국어 교육*, 67, 157-183.
- 한강우. (2022). 한국어능력시험(TOPIKⅡ) 작문형 글쓰기 문항 답안의 개선 방안에 대한 탐색. *문화와 융합*, 44(4), 23-54.
- 허희정. (2016). 4개 분야 한국어 학술텍스트에 나타난 정형표현 비교. *언어사실과 관점*, 39(0), 305-334.
- 홍혜란·박지순(2020). 한국어 교육을 위한 텍스트 장르별 메타 담화 표지 연구. *국제어문*, 84(0), 79-114.
- PHAN, HOANG MY THUONG. (2020). 베트남인 한국어 학습자를 위한 학문 목적 쓰기 교육 내용 연구 -인문사회계열 보고서 분석을 중심으로. 이화여자대학교 박사학위논문.
- Bartolic, S. (1981). Interpretation of information transfer from a diagram. In L. Selinker, E. Tarone, & V. Hanzeli (Eds.), *Proceedings of the Second Midwest Conference on Linguistics and Language Learning* pp. 187-201. Newbury House.
- Bawarshi, A., & Reiff, M. J. (2010). *Genre: An introduction to history, theory research and pedagogy*. Parlor Press LLC./정희모·김성숙·김미란 외 8인 역. (2015). 장르: 역사 이론 연구 교육. 경진출판.
- Bhatia, V. (1997). Genre-Mixing in Academic Introductions, *English for specific purposes*. 16 (3), 81-195.
- Biber, D., Conrad, S., & Cortes, V. (2004). If You Look at: Lexical Bundles in University Teaching and Textbooks. *Applied linguistics*, 25(3), 371-

405.

- Biber, D. (2006). Stance in Spoken and Written University Registers. *Journal of English for Academic Purposes*, 5(2), 97-116.
- Biber, D., Connor, U., & Upton, T. (2007). *Discourse on the move : Using corpus analysis to describe discourse structure*. Amsterdam ; Philadelphia : John Benjamins.
- Boulton, A. (2010). Data-driven Learning: Taking the Computer out of the Equation. *Language learning*, 60(3), 534-572.
- Bunton, D. (2002). Generic moves in PhD thesis introductions. In J. Flowerdew. (Ed.). *Academic discourse*. 57-75. Routledge.
- Byrd, P., & Coxhead, A. (2010). On the other Hand: Lexical Bundles in Academic Writing and in the Teaching of EAP. *University of Sydney Papers in TESOL*, 5(5), 31-64.
- Cotos, E., Link, S., & Huffman, S. (2017). Effects of Technology on Genre Learning. *Language Learning & Technology*, 21(3), 104-130.
- Cortes, V. (2015). Situating lexical bundles in formulaic language spectrum. in D. Biber, E. Csomay, & V. Cortes. (2015). *Corpus-based research in applied linguistics studies in honor of Doug Biber*. John Benjamins.
- Council of Europe. (2020). *Common European framework of reference for languages: Learning, teaching, assessment- Companion Volume*. Strasbourg: Council of Europe Publishing. available at www.coe.int/lang-cefr.
- Crismore, A., Markkanen, R., & Steffensen, M. (1993). Metadiscourse in Persuasive Writing: A Study of Texts Written by American and Finnish University Students, *Written Communication*, 10, 39-71.
- DeKeyser, R. M. (2001), Automaticity and automaticization, In P. Robinson (Ed.). *Cognition and second language acquisition*, Cambridge; New York, N.Y.:Cambridge University Press.
- DePoy, E., & Gitlin, L. (2016). *Introduction to Research 5th ed*. Elsevier.
- Eggins, S. (2004). Introduction to systemic functional linguistics./ 김서형·유혜원·이동혁·이유진·정연주 역. (2021). 체계 기능 언어학의 이해. 역락.

- Fei, V. L. (2004). Developing an integrative multi-semiotic model. In K. O'Halloran. (2004). *Multimodal discourse analysis : Systemic functional perspectives*. 220-246. Continuum.
- Flowerdew, L. (2015). Using Corpus-based Research and Online Academic Corpora to Inform Writing of the Discussion Section of a Thesis. *Journal of English for Academic Purposes*, 20, 58-68.
- Fuentes, A. C. (2018). DDL Integration in Different EAP Scenarios of Academic Writing. *Revista Encuentro*, 27, 157-172.
- Grice, P. (1989). *Studies in the way of words*. Harvard University Press.
- Guinda, C. S. (2011). Integrating approaches to visual data commentary: An exploratory case study. In V. Bhatia, P. Sánchez Hernández, P. Pérez-Paredes (Eds.). *Researching specialized languages*. 115-138. John Benjamins.
- Guo, L. (2004). Multimodality in a biology textbook. In K. O'Halloran (Ed.), *Multimodal discourse analysis: Systemic functional perspectives*, pp. 196-219. Continuum.
- Heift, T., & Vyatkina, N. (2017). Technologies for teaching and learning L2 grammar. *The handbook of technology and second language teaching and learning*, 26-44.
- Hyland, K. (1998). *Hedging in scientific research articles*. Amsterdam/Philadelphia: John Benjamins.
- _____. (2004). *Genre and second language writing*. /이수미·이소연 역.
- _____. (2019). 장르와 제2언어 글쓰기. 파주: 하우.
- _____. (2005). *Metadiscourse: Exploring interaction in writing*, Bloomsbury Publishing.
- _____. (2018). Genre and second language writing. *The TESOL encyclopedia of English language teaching*, 1-6.
- _____. (2021). *Teaching and researching writing*. Routledge.
- Johns, A. M., & Davies, F. (1983). Text as a Vehicle for Information: The Classroom Use of Written Texts in Teaching Reading in a Foreign Language. *Reading in a Foreign Language*. 1(1), 1-20.

- Johns, T. (1991). Should you be persuaded- Two examples of data-driven learning. T. F. Johns, & P. King (Eds.). *Classroom concordancing*. Birmingham: ELR.
- Kanoksilapatham. (2005). Rhetorical Structure of Biochemistry Research Articles. *English for specific purposes*, 24, 269-292.
- Knapp, P., & Watkins, M. (2005). Genre, text, grammar: Technologies for teaching, and assessing writing./주세형·김은성·남가영 역.(2019). 장르, 텍스트, 문법: 작문 교육을 위한 테크놀로지로서의 문법. 사회평론아카데미.
- Koester, A. (2010). Building small specialised corpora. In A. O'Keeffe, & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics*. 94-107. Routledge.
- Kress G., & van Leeuwen, T. (2006). *Reading Images: The Grammar of Visual Design*. Routledge.
- Lemke, J. (1998). Multiplying meaning: Visual and verbal semiotics in scientific text. In J. R. Martin, & R. Veel (Eds.), *Reading science: Critical and functional perspectives on discourses of science*. 87-113. Routledge.
- Miller, T. (1998). Visual Persuasion: A Comparison of Visuals in Academic Texts and the Popular Press. *English for Specific purposes*, 17(1), 29-46.
- Nelson, M. (2010). Building a written corpus: what are the basics? In A. O'Keeffe, & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics*. 81-93. Routledge.
- Nordrum, L., & Eriksson, A. (2015). Data commentary in science writing. In S. Götz, M. Callies (Eds.). *Learner corpora in language testing and assessment*, pp. 59-83. John Benjamins.
- O'Halloran, K. L. (1996). *The discourses of secondary school mathematics*. Doctoral thesis. Murdoch University.
- _____. (2008a). Systemic Functional-Multimodal Discourse Analysis (SF-MDA): Constructing Ideational Meaning Using Language and Visual Imagery. *Visual Communication*. 7(4), 443-475.

- _____. (2008b). *Mathematical discourse : Language, symbolism and visual images*. Continuum.
- Paltridge, B. (2013). 18. Genre and English for specific purposes. *The handbook of English for specific purposes*, 347.
- Polio, C. (2012). The Relevance of Second Language Acquisition Theory to the Written Error Correction Debate. *Journal of second language writing*, 21(4), 375-389.
- Roth, W. M. (2013). Data Generation in the Discovery Sciences – Learning from Practices in an Advanced Research Laboratory. *Research in Science Education*, 43(4): 1617-1644.
- Robinson, P. (2001), *Cognition and second language instruction*, New York.: Cambridge University Press.
- Skehan, P. (1998), *A cognitive approach to language learning*, New York: Oxford University Press.
- Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.
- Swales, J. M., & Feak, C. B. (1994). *Academic writing for graduate students*. University of Michigan Press.
- _____. (1995). From information transfer to data commentary. In A. M. Miller (Ed.), *Genre and second language writing*, pp. 64-76. Addison Wesley Longman.
- Valentina, G. (2016). Graph Description as an Issue in L2 Academic English Writing. *Journal of Language and Education*, 2(4), 46-54.
- Widdowson, H. G. (1973). *An applied linguistic approach to discourse analysis*. Doctoral dissertation. University of Edinburgh.
- _____. (1979). *Explorations in applied linguistics*. Oxford University Press.
- Xue-hua, A., & Ming-ying, X. (2013). An Empirical Research on DDL in L2 Writing. *US-China Education Review*, 693-701.
- Yang, H. C. (2012). Modeling the Relationships Between Test-Taking Strategies and Test Performance on a Graph-Writing Task-

- Implications for EAP. *English for Specific Purposes*, 31(3), 174-187.
- _____. (2016). Describing and Interpreting Graphs: The Relationships Between Undergraduate Writer Characteristics and Academic Graph Writing Performance. *Assessing Writing*, 28, 28-42.
- Yoon, H., & Hivela, A. (2004). ESL Student Attitudes Toward Corpus Use in L2 Writing. *Journal of second language writing*, 13(4), 257-283.

[소프트웨어]

- Fuoli, M. (2019). UAM Corpus Tool Workshop. Retrieved from osf.io/5hq8g.

[웹사이트]

- 국가교육과정 정보센터. <https://ncic.re.kr/mobile.dwn.ogf.inventoryList.do#> (2022.6.20. 검색)
- 교육부 웹사이트 <https://www.moe.go.kr/boardCnts/viewRenew.do?boardID=350&lev=0&statusYN=W&s=moe&m=0309&opType=N&boardSeq=77521> (2023.6.18. 검색)
- 국립국어원 한국어 학습자 말뭉치 <https://kcorpus.korean.go.kr> (2023.6.15. 검색)
- 꼬꼬마 품사 태그 kkma.snu.ac.kr (2023.4.20. 검색)
- 브리태니커 영어 사전 <https://www.britannica.com/dictionary/commentary> (2023.6.7. 검색)
- 성균관대학교 입학처 홈페이지. <https://admission.skku.edu/admission/html/rolling/questionView.html> (2023.6.20. 검색)
- 스케치엔진 sketchengine.eu (2023.6.14. 검색)
- 아이엘츠 공식 웹사이트 <https://ielts.idp.com/results/scores/writing> (2023.4.19. 검색)
- 엘스비어 웹사이트 <https://www.elsevier.com/authors/tools-and-resources/graphical-abstract> (2023.7.13. 검색)
- 중화인민공화국 교육부 웹사이트 <https://www.ncct.edu.cn/Uploads/File/2022/04/28/%E6%99%AE%E9%80%9A%E9%AB%98%E4%B8%AD%E8%AF%AD%E6%96%87%E8%AF%BE%E7%A8%8B%E6%A0%87%E5%87%86%EF%B>

C%882017%E5%B9%B4%E7%89%88%EF%BC%89.20220428174646.pdf

(2023.6.27. 검색)

한국어능력시험 공식 웹사이트 <https://www.topik.go.kr/HMENU0/HMENU00017.d>

o (2023.4.29. 검색)

[부록] 학습자 대상 데이터 설명문 과제

1. 지시문

당신은 수업 과제로 ‘**코로나 시기 성별 가사 노동 시간 연구**’ 보고서를 써야 합니다. 보고서를 쓰기 위해 위의 데이터를 이용할 것입니다. 이 데이터를 필요한 부분에 배치하여 설명, 분석하는 보고서를 쓰세요.

<조건>

1. 글자수 기준 **600자 이상** 쓰세요.
2. <표 1>, <그림 1>, <그림 2> 데이터를 **모두** 이용하세요.
3. **데이터를 설명 및 분석하여 주제에 맞는** 보고서를 쓰세요.
4. 실제 자신의 보고서를 쓰는 것처럼 **데이터 분석에 따른 자신의 견해나 결론**을 써 주세요.
5. <표 1>, <그림 1>, <그림 2>를 **실제 필요한 부분에 ‘<데이터>(삽입함)’이라고 넣어** 글을 구성해 주세요.

[예]

이는 다음과 같다.

<표 1>(삽입함)

위의 <표 1>은……

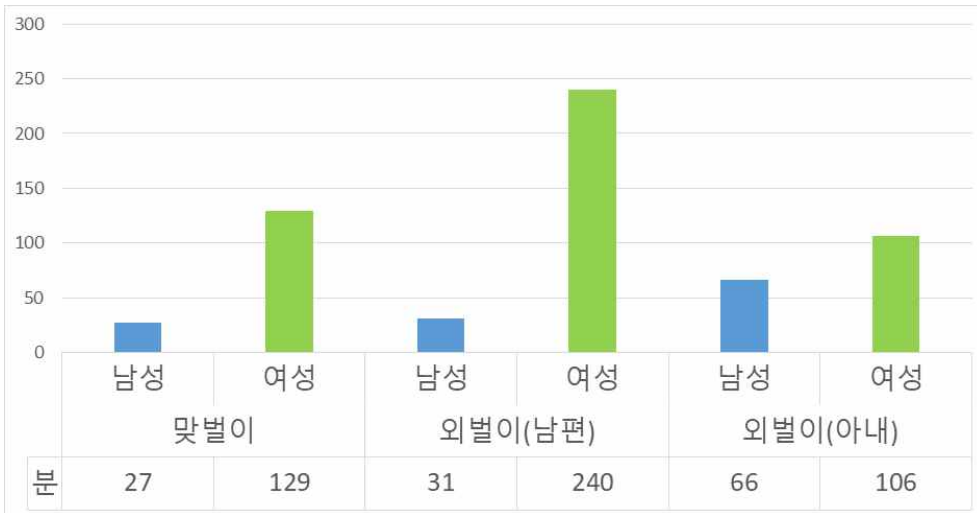
2. 시각적 데이터

<표 1> 설문 응답자 기본 정보(통계청, 2022)

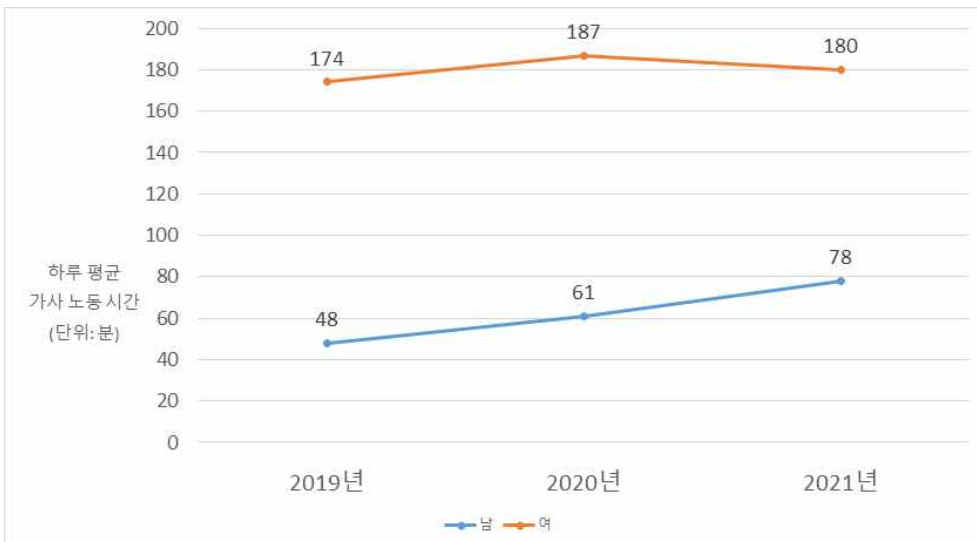
(단위: 명)

| | | |
|----------------|------------------|-----|
| 성별 | 남 | 500 |
| | 여 | 500 |
| 학력 | 중졸 이하 | 10 |
| | 고졸 | 260 |
| | 대졸 | 620 |
| | 대학원졸 | 110 |
| 하루 평균 가사 노동 시간 | 1시간 미만 | 280 |
| | 1시간~2시간 | 402 |
| | 2시간~3시간 | 227 |
| | 3시간~4시간 | 74 |
| | 4시간 이상 | 17 |
| 혼인 상태 | 비혼(미혼, 사별, 이혼 등) | 280 |
| | 기혼 | 720 |

<그림 1> 맞벌이 상태별 하루 가사 노동 시간(통계청, 2022)



<그림 2> 남성과 여성의 하루 평균 가사 노동 시간 비교(통계청, 2022)



Abstract

A Study on Academic Data Commentary Writing by Korean Language Learners

Lee, Jeeyeoun

Department of Korean Language Education

College of Education

The Graduate School of Seoul National University

Data commentary is indispensable when representing objectivity in academic texts. Visual materials, which contain numerical values and connect ideas, combine linguistic symbols that play a central role in supporting the writer's contention. Therefore, it is necessary for university students who are starting their academic writing in Korean as a second language.

However, data commentary has received little attention in the literature on academic writing for two reasons. First, genre-based instruction, used for academic text analysis and instruction, often excludes mini-genres such as visual materials, footnotes, captions,

and acknowledgements. This exclusion hinders the study of data commentary, which includes visual materials and linguistic symbols.

The second reason is the unjustifiable status of data commentary when categorizing academic texts. Genre-based instruction regards texts created by expert writers as model ones that provide fine constructions and linguistic rules for the genre, contrasting with data commentary.

Data commentary is embedded within upper genres such as theses and experiment reports. For example, data commentary is commonly found in the results and discussion section of journal articles or in result section of experiment reports. These characteristics make data commentary a variable and challenging text type to study due to its lack of fixed composition. Few studies have revealed the specific construction and linguistic features of data commentary, especially when visual materials and linguistic symbols are involved.

Authentic materials, however, encompass these variable aspects. The purpose of academic writing is to practice expert writing skills for the future. Therefore, it is more important to analyze students' data commentary to understand their schema, traits, and features than to provide clear and fixed forms for writing instruction.

This study collected data commentaries from university students, including 125 non-native speakers of Korean (70 intermediate level, 55 advanced level) and 50 native speakers of Korean online. The data commentary task consisted of one table and two graphs: one bar graph and two line graphs. Students were asked to write more than 600 characters in Korean.

The collected data commentaries were analyzed in two parts: construction and linguistic features. Move analysis was conducted based on Systemic Functional-Multimodal Discourse Analysis(SF-MDA) and

data mediation in the Common European Framework of Reference(CEFR). Linguistic features were identified using n-grams based on the moves. The moves were analyzed in three parts: A) summarizing the topic and methodology, B) representing figures and numbers, and C) analyzing and commenting on results.

The data commentaries of the students were analyzed with respect to moves. The order of the moves was compared and the frequency of each move and/or step were tested using the chi-square test based on the students' L1 and Korean proficiency level. Linguistic features of the data commentaries were analyzed for each move and/or step using 3-grams.

As a result, discourse and language use of users were identified as below: First, students with low proficiency tend to follow the similar data commentary like construction what they have experienced before. Thus, most of them use A-B type constructions for their data commentary. On the other hand, students with higher proficiency and grades tend to choose A-B-C and variation of that for their data commentary writing. Second, each moves and steps were used statically meaningful depending on the groups. Especially on move B, it was inferred that lower level of proficiency and grade make students to concentrate more on second language than constructions that require cognitive effort. Lastly, genre specific language use is challenging for both lower an higher proficiency and grade learners. Context based linguistic repertoires may help students to use discourse markers properly in consideration of the distance between writer and reader, and objectivity of academic text. Based on these results, the study suggests that Data-Driven Learning(DDL) can be a valuable learning and teaching method for data commentary in various upper genres within the academic field. Data-driven learning

uses corpus which contains abundant genre specific markers in the text that they can help students to write academically proper data commentary.

The contribution of this research lies in its investigation of data commentary, which has been challenging to study. Furthermore, the traits of data commentary were identified through the analysis of students' writing.

Keywords : data commentary, mini-genre, systemic functional multimodal discourse analysis(SF-MDA), move, data-driven learning(DDL)

Student Number : 2016-30523