creative commons

COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

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

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 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

dCollection

교육학박사학위논문

중국 배경 중도입국 청소년의 한국어 발달 연구

- 문어와 구어의 복잡성과 정확성을 중심으로 -

2023년 8월

서울대학교 대학원

국어교육과 한국어교육전공

조청화(ZHAO QINGHUA)

중국 배경 중도입국 청소년의 한국어 발달 연구

- 문어와 구어의 복잡성과 정확성을 중심으로 -

지도교수 김 호 정

이 논문을 교육학 박사학위 논문으로 제출함.
2023년 4월

서울대학교 대학원
국어교육과 한국어교육전공
조 청 화(ZHAO QINGHUA)

조청화의 교육학 박사학위 논문을 인준함.
2023년 7월

위 원 장 구 본 관 (인)

부위원장 배 은 영 (인)

위 원 강 남 욱 (인)

위 원 박 민 신 (인)

위 원 김 호 정 (인)

국문 초록

본 연구는 중국 배경 중도입국 청소년의 문어와 구어에 나타난 복잡성과 정확성 발달을 고찰하는 것을 목적으로 한다. 중도입국 청소년은 부모의 이주를 따라 청소년기에 한국에 입국하여 학업을 수행하고 한국 사회에 정착하게 되는데 그 과정에서 가장 문제가 되는 것 중의 하나가 한국어 능력이다. 중도입국 청소년의 한국어 교육을 지원하기 위해 다양한 대책들이 마련되고 있지만 이들의 언어 발달을 직접 고찰한 연구는 아직 미흡하다. 이에 본 연구에서는 중국 배경 중도입국 청소년 학습자의 문어와 구어 발달 과정에 나타나는 특징들을 다각적 측면에서 밝히고자 하였다.

이를 위해 중국 배경 중도입국 청소년 학습자 3명을 대상으로 1년 6개월 간 문어와 구어 자료를 수집하여 분석하였다. 본 연구에서는 복잡성을 통사 복잡성과 어휘 복잡성으로 구분하고 통사 복잡성은 단위 길이, 단위당 절 수, 단위 확장의 유형을 통해 측정하고 어휘 복잡성은 어휘 다양도, 어휘 밀도, 어휘 정교도를 통해 측정하였으며 정확성은 무오류절의 비율과 단위당 무오류절 수, 단위당 오류 수를 통해 측정하였다.

이러한 측정 지표들을 통해 학습자들의 문어와 구어 발달을 살펴본 결과, 학습자들의 통사 복잡성은 시간이 지남에 따라 증가하고 있었으나 어휘 복잡성과 정확성의 발달은 뚜렷하게 나타나지 않았다. 구체적으로 살펴보면 문어와 구어 모두 단위 길이와 단위당 절 수는 시간의 흐름에 따라 꾸준히 증가하고 있었으며 특히 부사절의 증가가 가장 뚜렷하였다. 어휘 복잡성 중에서 어휘 다양도는 문어와 구어 모두 1~6개월과 7~12개월 사이에는 뚜렷하게 증가하다가 13~18개월에 이르러 정체를 보였는데 이는 입국 이후 3년 정도 되는 시점이었다. 어휘 밀도는 문어와 구어 모두 기간별로 유의한 변화를 보이지 않았고 어휘 정교도는 문어에서는 기간별로 유의한 차이가 관찰되지 않은 반면 구어에서는 13~18개월에 들어서서 큰 폭으로 발달하여 문어와 비슷한 수준에 도달하였다. 정확성에 있어서는 문어와 구어 모두 시간이 지날수록 무오류절 수와 오류 수가 동반 상승하며 정확성이 향상되지 못하는 양상을 보였다.

산출 양식에 따른 복잡성과 정확성을 비교해 본 결과, 단위 길이, 단위당 절 수, 어휘 다양도는 모두 문어에서 더 높은 것으로 나타났고 어휘 밀도와 어휘 정교도는 문어와 구어 사이에 유의한 차이가 없는 것으로 나타났다. 그리고 정확성은 무오류절의 비율, 단위당 무오류절 수, 단위당 오류 수 모두 문어와 구어 간의 차이가 유의하지 않았으나 문어에서는 누락과 오형태 오류가 더 빈번하고 구어에서는 대치와 첨가 오류가 더 빈번하여 산출 양식에 따른 오류 현상의 차이를 확인할 수 있었다.

복잡성과 정확성의 관계는 측정 지표에 따라 서로 상반되는 결과를 보이고 있었다. 문어와 구어 모두 통사 복잡성과 무오류절 수, 통사 복잡성과 오류 수가 뚜렷한 정적 관계를 보이고 있어 지지 혹은 경쟁의 고정적인 관계를 확인할 수 없었으나 상관계수의 크기를 비교해 보면 문어에서는 통사 복잡성이 무오류 수와 더 밀접한 관계를 보이고 구어에서는 통사 복잡성과 오류 수와 더 밀접한 관계를 보이고 있었다.

이러한 분석 결과를 통해 도출한 시사점을 정리하면 다음과 같다. 첫째, 중국 배경 중도입국 청소년 학습자들은 별도의 한국어 교육이 없는 자연스러운 공교육 환경에서도 시간이 지나면 문장의 구조적 복잡성은 지속적으로 발달한다. 그러나 문법 항목의 다양성과 정확도 측면에서는 한계를 보이며 한정적인 항목을 반복 사용하는 양상이 나타났다. 이를 개선하기 위하여 부사절과 명사절, 인용절을 중심으로 한 다양한 어미의 의미 및 형태 교육이 좀 더 명시적으로 이루어질 필요가 있을 것이다. 둘째, 학습자들은 자연적인 입력만으로는 어휘 다양도나 어휘 정교도 발달에 한계를 나타냈다. 따라서 학습자들의 어휘 복잡성 발달을 위해서는 기존에 습득한 어휘를 바탕으로 합성이나 파생과 같은 어휘 형성의 원리 교육을 통해 어휘의 지속적인 확장을 촉진할 필요가 있을 것이다. 셋째, 자연적인 환경에서 학습자들의 정확성은 시간이 지난다고 해서 무조건적으로 향상되는 것은 아니었다. 이러한 견지에서 조사, 선어말 어미, 연결어미의 정확한 사용을 지속적으로 교육하고 어휘력 확장을 통해서 어휘 오류를 개선할 필요가 있다. 넷째, 학습자들은 산출 양식에 따라 통사와 어휘 복잡성에서 차이를 보였으며 오류 영역과 오류 현상에서도 차이를 나타냈다. 따라서 문어와 구어의 균형잡힌 발달을 위해서는 산출 양식을 고려한 맞춤형 교육적 처치가 이루어져야 할 것이다. 다섯째, 학습자들은 문어에서 구어체를 사용하는 현상이 빈번한바, 문어와 구어의 사용역에 따른 언어적 특징에 대한 명시적인 교육이 필요한 것으로 보인다.

본 연구는 종적 자료를 수집하여 중국 배경 중도입국 청소년 학습자의 문어·구어 발달 과정을 고찰하였다. 본고에서는 자료에 나타난 복잡성과 정확성을 분석하였을 뿐만 아니라 이 두 구인 사이의 관계를 탐구함으로써 시간의 흐름에 따른 언어 발달을 다각적으로 고찰하였다는 데 의의가 있다. 또한 문어와 구어를 비교함으로써 산출 양식에 따른 복잡성과 정확성의 차이를 구체적으로 밝혔다는 점에서도 의의를 찾을 수 있다. 이러한 결과는 향후 중도입국 청소년의 한국어 발달을 연구하고 교육 방안을 설계하는 데에도 중요한 기초 자료가 될 수 있을 것이다.

주요어 : 중도입국 청소년, 문어, 구어, 통사 복잡성, 어휘 복잡성, 정확성
학　번 : 2018-31140

목 차

I. 서론 1

1. 연구의 목적 및 필요성 1
2. 선행연구 검토 3
   2.1. 중도입국 청소년 대상 한국어 교육 연구 3
   2.2. 복잡성 관련 연구 7
   2.3. 정확성 관련 연구 12
   2.4. 복잡성과 정확성 관계 연구 14
   2.5. 문어와 구어의 복잡성과 정확성 관계 연구 15
3. 연구 범위 및 절차 19

Ⅱ. 이론적 고찰 20

1. 제2언어 습득에서의 복잡성과 정확성 20
2. 복잡성의 개념 및 측정 22
   2.1. 통사 복잡성의 개념 및 측정 26
   2.2. 어휘 복잡성의 개념 및 측정 32
3. 정확성의 개념 및 측정 39
   3.1. 정확성의 개념 39
   3.2. 정확성의 측정 43
4. 복잡성과 정확성의 관계 45
   4.1. 경쟁 관계: 제한된 주의력 모델 46
   4.2. 지지 관계: 다자원 주의력 모델 48
   4.3. 역동적 관계: 역동 체계 이론 50
5. 문어와 구어의 관련성 53

III. 연구 방법 60

1. 연구 참여자 60
2. 연구 자료 65
3. 자료 분석 기준 71
   3.1. 통사 복잡성 측정 지표 및 분석 기준 71
   3.2. 어휘 복잡성 측정 지표 및 분석 기준 87
   3.3. 정확성 측정 지표 및 분석 기준 96
4. 자료 분석 절차 및 방법 105

4.1. 자료 수집 및 입력 105
4.2. 자료 분석 및 통계 108

IV. 연구 결과 111

1. 통사 복잡성의 발달 111
1.1. 문어의 통사 복잡성 발달 111
1.2. 구어의 통사 복잡성 발달 130
1.3. 산출 양식에 따른 통사 복잡성 비교 150
2. 어휘 복잡성의 발달 162
2.1. 문어의 어휘 복잡성 발달 162
2.2. 구어의 어휘 복잡성 발달 171
2.3. 산출 양식에 따른 어휘 복잡성 비교 180
3. 정확성의 발달 186
3.1. 문어의 정확성 발달 187
3.2. 구어의 정확성 발달 199
3.3. 산출 양식에 따른 정확성 비교 212
4. 복잡성과 정확성의 관계 221
4.1. 문어에 나타난 복잡성과 정확성의 관계 222
4.2. 구어에 나타난 복잡성과 정확성의 관계 225

V. 논의 및 제언 230

1. 복잡성과 정확성의 통시적 발달 230
1.1. 통사 복잡성의 발달 230
1.2. 어휘 복잡성의 발달 240
1.3. 정확성의 발달 243
2. 산출 양식에 따른 복잡성과 정확성의 관계 245
2.1. 산출 과정 측면에서 나타난 차이 246
2.2. 발달 경로 측면에서 나타난 차이 249
3. 산출 양식과 통시적 흐름에 따른 복잡성과 정확성의 발달 관계 251
3.1. 전체 기간에 나타난 복잡성과 정확성의 관계 251
3.2. 시간의 흐름에 따른 복잡성과 정확성의 관계 변화 254
4. 복잡성과 정확성의 발달과 학습자 변인 255
4.1. 학습자 간 발달 경로와 발달 수준 차이 255
4.2. 학습자별 문어와 구어의 발달 차이 258
5. 교육적 제언 260

VI. 결론 265
참고문헌 270
Abstract 287

표 목차

<표 1> 연구 참여자 정보 61
<표 2> 연구 참여자 진단 및 성취도 검사 결과 62
<표 3> <한국어(KSL) 교육과정>의 3단계와 4단계 말하기·쓰기 성취기준 62
<표 4> 연구 참여자의 입국 전과 입국 후 언어 사용 비율 64
<표 5> 문어와 구어 산출 과제 예시 70
<표 6> 문어와 구어 산출 주제 70
<표 7> 수집된 언어 자료의 크기 71
<표 8> 의미 기능에 따른 보조 용언 목록(고영근·구본관, 2018) 85
<표 9> 복문의 분류 체계 86
<표 10> 한국어의 품사 분류(구본관 외, 2015:162) 89
<표 11> 본 연구의 오류 분류 체계 101
<표 12> 복잡성과 정확성 측정 지표 104
<표 13> 문어 자료 전사 예시 106
<표 14> 구어 자료 전사 예시 107
<표 15> 기간별 문장 길이 113
<표 16> 학습자별 문장 길이 113
<표 17> 기간별 문장당 절 수 114
<표 18> 학습자별 문장당 절 수 115
<표 19> 6개월 단위 문장당 대등절 수와 내포절 수 116
<표 20> 기간별 문장당 대등절과 내포절 수 117
<표 21> 학습자별 문장당 대등절 수와 내포절 수 117
<표 22> 유형별 내포절의 기간별 사용 양상 118
<표 23> 유형별 내포절의 학습자별 사용 양상 119

<표 24> 대등절 관련 문법 항목의 기간별 사용 양상 120
<표 25> 명사절 관련 문법 항목의 기간별 사용 양상 121
<표 26> 관형사절 관련 문법 항목의 기간별 사용 양상 123
<표 27> 부사절 관련 문법 항목의 기간별 사용 양상 124
<표 28> 인용절 관련 문법 항목의 기간별 사용 양상 129
<표 29> 기간별 AS-UNIT 길이 132
<표 30> 학습자별 AS-UNIT 길이 132
<표 31> 기간별 AS-UNIT당 절 수 134
<표 32> 학습자별 AS-UNIT당 절 수 134
<표 33> 6개월 단위 AS-UNIT당 대등절 수와 내포절 수 135
<표 34> 기간별 AS-UNIT당 대등절 수와 내포절 수 136
<표 35> 학습자별 AS-UNIT당 대등절 수와 내포절 수 136
<표 36> 유형별 내포절의 기간별 사용 양상 138
<표 37> 유형별 내포절의 학습자별 사용 양상 139
<표 38> 대등절 관련 문법 항목의 기간별 사용 양상 140
<표 39> 명사절 관련 문법 항목의 기간별 사용 양상 141
<표 40> 관형사절 관련 문법 항목의 기간별 사용 양상 142
<표 41> 부사절 관련 문법 항목의 기간별 사용 양상 143
<표 42> 인용절 관련 문법 항목의 기간별 사용 양상 149
<표 43> 기간별 문어와 구어의 단위 길이 151
<표 44> 학습자별 문어와 구어의 단위 길이 152
<표 45> 기간별 문어와 구어의 단위당 절 수 152
<표 46> 학습자별 문어와 구어의 단위당 절 수 153
<표 47> 기간별 문어와 구어의 단위당 대등절 수와 내포절 수 154
<표 48> 학습자별 문어와 구어의 단위당 대등절과 내포절 수 155
<표 49> 기간별 문어와 구어의 유형별 내포절 수 157
<표 50> 산출 양식에 따른 대등절 관련 문법 항목 사용 양상 158
<표 51> 산출 양식에 따른 명사절 관련 문법 항목 사용 양상 159
<표 52> 산출 양식에 따른 관형사절 관련 문법 항목 사용 양상 159

<표 53> 산출 양식에 따른 부사절 관련 문법 항목 사용 양상 160
<표 54> 산출 양식에 따른 인용절 관련 문법 항목 사용 양상 161
<표 55> 기간별 문어 어휘 다양도 163
<표 56> 기간별 문어 어휘 유형 수와 어휘 구현 수 164
<표 57> 체언과 용언의 기간별 어휘 다양도 165
<표 58> 학습자별 문어 어휘 다양도 166
<표 59> 기간별 문어 어휘 밀도 167
<표 60> 학습자별 문어 어휘 밀도 168
<표 61> 기간별 문어 어휘 정교도 170
<표 62> 학습자별 문어 어휘 정교도 171
<표 63> 기간별 구어 어휘 다양도 172
<표 64> 기간별 구어 어휘 유형 수와 어휘 구현 수 173
<표 65> 체언과 용언의 기간별 어휘 다양도 174
<표 66> 학습자별 구어 어휘 다양도 174
<표 67> 기간별 구어 어휘 밀도 176
<표 68> 학습자별 구어 어휘 밀도 177
<표 69> 기간별 구어 어휘 정교도 178
<표 70> 학습자별 구어 어휘 정교도 179
<표 71> 기간별 문어와 구어의 어휘 다양도 181
<표 72> 학습자별 문어와 구어의 어휘 다양도 182
<표 73> 기간별 문어와 구어의 어휘 밀도 183
<표 74> 학습자별 문어와 구어의 어휘 밀도 183
<표 75> 기간별 문어와 구어의 어휘 정교도 185
<표 76> 학습자별 문어와 구어의 어휘 정교도 185
<표 77> 기간별 문어 무오류절 비율 188
<표 78> 학습자별 문어 무오류절 비율 188
<표 79> 기간별 문장당 무오류절 수 189
<표 80> 학습자별 문장당 무오류절 수 190
<표 81> 기간별 문장당 오류 수 191

<표 82> 학습자별 문장당 오류 수 191
<표 83> 문어 자료에 나타난 조사 오류의 기간별 빈도 193
<표 84> 문어 자료에 나타난 어미 오류의 기간별 빈도 194
<표 85> 문어 자료에 나타난 어휘 오류의 기간별 빈도 197
<표 86> 문어 자료에 나타난 표현 및 기타 오류의 기간별 빈도 198
<표 87> 기간별 구어 무오류절 비율 200
<표 88> 학습자별 구어 무오류절 비율 201
<표 89> 기간별 AS-UNIT당 무오류절 수 202
<표 90> 학습자별 AS-UNIT당 무오류절 수 202
<표 91> 기간별 AS-UNIT당 오류 수 203
<표 92> 학습자별 AS-UNIT당 오류 수 204
<표 93> 구어 자료에 나타난 조사 오류의 기간별 빈도 206
<표 94> 구어 자료에 나타난 어미 오류의 기간별 빈도 207
<표 95> 구어 자료에 나타난 어휘 오류의 기간별 빈도 210
<표 96> 구어 자료에 나타난 표현 및 기타 오류의 기간별 빈도 211
<표 97> 기간별 문어와 구어의 무오류절 비율 213
<표 98> 학습자별 문어와 구어의 무오류절 비율 213
<표 99> 기간별 문어와 구어의 단위당 무오류절 수 214
<표 100> 학습자별 문어와 구어의 단위당 무오류절 수 214
<표 101> 기간별 문어와 구어의 단위당 오류 수 215
<표 102> 기간별 문어와 구어의 단위당 오류 수 216
<표 103> 산출 양식에 따른 조사 오류 빈도 217
<표 104> 산출 양식에 따른 어미 오류 빈도 218
<표 105> 산출 양식에 따른 어휘 오류 빈도 220
<표 106> 문어 복잡성과 정확성의 관계 222
<표 107> 기간별 복잡성과 정확성(문장당 오류 수)의 관계 223
<표 108> 기간별 복잡성과 정확성(문장당 무오류절 수)의 관계 224
<표 109> 학습자별 문어 통사 복잡성과 정확성의 관계 224
<표 110> 구어 복잡성과 정확성의 관계 225

<표 111> 기간별 복잡성과 정확성(AS-UNIT당 오류 수)의 관계 226
<표 112> 기간별 복잡성과 정확성(AS-UNIT당 무오류절 수)의 관계 227
<표 113> 학습자별 구어 통사 복잡성과 정확성의 관계 227
<표 114> 선행연구의 통사 복잡성 측정 값 239
<표 115> 학습자별 복잡성과 정확성의 변동폭(1~6개월과 13~18개월 비교) 255
<표 116> 학습자 간 문어와 구어의 복잡성과 정확성 비교 256
<표 117> 학습자별 문어와 구어의 복잡성과 정확성 비교 259

그림 목차

<그림 1> 복잡성 구인의 분류 (BULTÉ & HOUSEN, 2012:22-25) 24
<그림 2> 문법 복잡성의 단계(BULTÉ & HOUSEN, 2012:27) 27
<그림 3> 어휘 복잡성의 단계(BULTÉ & HOUSEN, 2012:28) 34
<그림 4> KELLOG의 쓰기 절차 모델(KELLOGG, 1996:59) 55
<그림 5> 학습자 언어 자료 전사 화면 107
<그림 6> 대등절과 내포절 코딩 화면 108
<그림 7> 시점별 문장 길이와 6개월 단위 발달 추이 112
<그림 8> 시점별 문장당 절 수와 6개월 단위 발달 추이 114
<그림 9> 유형별 내포절 수의 시점별 빈도와 6개월 단위 발달 추이 118
<그림 10> 시점별 AS-UNIT 길이와 6개월 단위 발달 추이 131
<그림 11> 시점별 AS-UNIT당 절 수와 6개월 단위 발달 추이 133
<그림 12> 시점별 유형별 내포절 수와 6개월 단위 발달 추이 137
<그림 13> 문어와 구어의 단위당 유형별 내포절 수 156
<그림 14> 시점별 문어 어휘 다양도와 6개월 단위 발달 추이 163
<그림 15> 문어 자료에 나타난 체언과 용언의 6개월 단위 발달 추이 164
<그림 16> 시점별 문어 어휘 밀도와 6개월 단위 발달 추이 167
<그림 17> 시점별 문어 어휘 정교도와 6개월 단위 발달 추이 169
<그림 18> 기간별 사용 어휘의 등급 분포 170

<그림 19> 시점별 구어 어휘 다양도와 6개월 단위 발달 추이 172
<그림 20> 구어 자료에 나타난 체언과 용언의 6개월 단위 발달 추이 173
<그림 21> 시점별 구어 어휘 밀도와 6개월 단위 발달 추이 175
<그림 22> 시점별 구어 어휘 정교도와 6개월 단위 발달 추이 178
<그림 23> 기간별 사용 어휘의 등급 분포 179
<그림 24> 산출 양식에 따른 체언과 용언의 어휘 다양도 181
<그림 25> 산출 양식에 따른 유형별 내용어 수 182
<그림 26> 산출 양식에 따른 등급별 어휘 분포 186
<그림 27> 시점별 문어 무오류절 비율과 6개월 단위 발달 추이 187
<그림 28> 문장당 무오류절 수의 시점별 발달 추이와 6개월 단위 발달 추이 189
<그림 29> 시점별 문장당 오류 수와 6개월 단위 발달 추이 190
<그림 30> 문어 자료에 나타난 오류 영역 및 오류 현상의 기간별 빈도 192
<그림 31> 시점별 구어 무오류절 비율과 6개월 단위 발달 추이 200
<그림 32> 시점별 AS-UNIT당 무오류절 수와 6개월 단위 발달 추이 201
<그림 33> 시점별 AS-UNIT당 오류 수와 6개월 단위 발달 추이 203
<그림 34> 구어 자료에 나타난 오류 영역 및 오류 현상의 기간별 빈도 205
<그림 35> 산출 양식에 따른 영역별 및 현상별 오류 빈도 217
<그림 36> 산출 양식에 따른 표현 및 기타 오류 빈도 220

I. 서론

1. 연구의 목적 및 필요성

본 연구의 목적은 중국 배경 중도입국 청소년의 한국어 발달 과정을 고찰하여 교육적 함의를 도출해 내는 데 있다. 이를 위해 중국 배경 중도입국 청소년의 문어와 구어 산출 자료를 수집하고 자료에 나타난 복잡성과 정확성을 분석하여 이들의 한국어 능력 수준 및 발달 양상을 살펴보고자 한다.

중도입국 청소년은 내국인과 재혼한 이주여성이 본국에서 출산했거나 양육하던 자녀, 외국인 부모와 동반 입국한 청소년, 학업이나 근로를 목적으로 청소년기에 입국한 외국 출신의 무연고 청소년, 그리고 국제결혼가정의 자녀 중 외국인 부모의 본국에서 성장하다가 청소년기에 입국한 청소년, 북한이탈주민과 외국인 사이에서 출생하여 제3국에서 성장한 청소년 등을 포함한다(임선일 외, 2021:1). 중도입국 청소년의 정의에서도 나타나듯이 중도입국 청소년은 그 유형과 범주가 매우 다양하다. 그러나 이들은 공통적인 특징을 지니고 있는데 바로 '한국이 아닌 외국의 언어와 문화 속에서 성장하고 교육 받아 오다가 중도에 한국에 입국하여 새롭게 한국어와 한국의 문화에 적응해야 하는 학습자(원진숙, 2015)'라는 점이다.[1]

중도입국 청소년들은 대부분 출신국에서 언어, 생활, 학습, 문화를 익힌 뒤 입국하기 때문에[2] 국내출생 다문화가정 자녀와 달리 한국어를 제2언어로 학습하게 되고 따라서 한국어 수준이 높지 않은 것은 물론 문화적 정체성도 상이하다. 또한 이들은 청소년들이고 한국 공교육에 진입하여 학업을 수행해야 한다는 점에서 인지적 수준, 동기, 학습 환경, 학습 목적이 일반 성인 학습자와 상이하다. 그동안 중도입국 청소년의 한국어 학습을 지원하기 위한 목적으로 교육과정이나 교재에 대한 연구들이 꾸준히 이루어져 왔으나 이들의 한국어 발달을 직접 고찰한 연구는 매우 미흡한 실정이다.

언어 발달 연구는 구체적인 교육 내용이나 교수 방안을 제시하지는 않으나 '그간의 교수 학습을 평가하고 향후의 발전적인 언어 교수 학습을 계획하기 위한 기초자료(김

---

[1] 한국 사회에서는 다문화학생 혹은 이주배경 청소년 등의 용어도 비슷한 의미로 사용되고 있는데 다문화학생은 '국제결혼 가정 자녀(국내출생 자녀+ 중도입국 자녀)+ 외국인 가정 자녀'로 정의되고(교육부 기본통계) 이주배경 청소년은 '부모 또는 본인이 국경을 넘는 이주의 경험을 지닌 만 24세 이하의 연령에 속하는 자'로 정의되는(양계민, 2021:13) 것으로 중도입국 청소년은 다문화학생, 다문화가정 자녀, 이주배경 청소년의 하위 유형이다.

[2] 현재 국내에 입국하고 있는 중도입국 청소년 중 다수는 7세 이상의 후기 청소년으로, 유소년기를 출신국에서 보낸 이들이 많다(여성가족부, 2016).

호정, 2007:89)'라는 점에서 언어교육 분야에서 큰 중요성을 가진다. 이미 교육과정과 교재가 개발되어 사용되고 있는 현 시점에서 이들의 한국어 발달을 살피는 것은 그동안 시행된 한국어 교육의 효과를 검증하고 앞으로의 교육에 시사점을 제공하여 더 발전적인 교수를 계획하는 데 기여할 수 있을 것이다.

학습자의 언어 발달을 살펴보는 가장 직접적이고 효과적인 방법은 이들의 언어를 분석하는 것이다. 그동안 한국어교육 분야의 언어 발달 연구들은 일반적으로 문어 혹은 구어 중 어느 한 쪽에만 초점을 맞춰 왔으며 같은 학습자를 대상으로 문어와 구어의 발달을 다룬 연구는 많지 않았다. 그러나 모어화자가 일반적으로 구어를 먼저 습득하고 문어의 발달이 구어에 의존하는 단계를 거쳐(Kroll, 1981) 결과적으로 문어와 구어의 균형적인 발달을 이루는 것과 달리 제2언어 학습자는 학습 환경이나 학습 목적에 따라 구어보다 문어를 먼저 습득하거나 구어와 문어를 동시에 습득하기도 하며 문어와 구어의 발달이 균형적이지 않은 경우가 많다. 또한 문어와 구어는 같은 지식 자원을 바탕으로 하지만 산출 과정에서 겪는 인지적 처리 과정의 차이로 인해 산출 양식에 따라 다른 성취를 보이기도 한다. 중도입국 청소년들은 한국에서 생활을 하면서 학업도 수행해야 한다는 점에서 쓰기와 말하기 능력에 대한 요구가 일반 학습자에 비해 더 높으며 단순한 의미 전달이 아닌 더 정교하고 정확한 언어를 산출할 필요가 있다. 따라서 본고에서는 학습자들의 한국어 발달을 문어와 구어로 나누어 살펴보고자 한다.

언어를 분석하기 위해서는 언어 발달의 수준을 가늠할 수 있는 객관적인 측정 지표가 필요한데 그동안 제2언어 학습자의 중간언어 발달 연구에서 가장 많이 사용되어 온 구인은 복잡성, 정확성, 유창성(complexity, accuracy, fluency, CAF)으로 여러 선행연구들(Skehan, 1998; Ellis, 2003; Housen & Kuiken, 2009)에서 세 구인이 학습자의 제2언어 숙달도를 구분할 수 있음을 언급하였다. 이 중에서 복잡성은 '과제 수행 시 산출된 언어의 정교하고 다양한 정도(Ellis, 2003:340)'이고 정확성은 '목표어의 규칙 체계에 맞는 언어를 생성하는 능력(Skehan, 1996b:23)'을 의미한다. 그리고 유창성은 시간이나 실시간성이 중요한 변수로 작용하는 구인으로 '원어민과 같은 신속성(Lennon, 1990:390)'을 나타내는 구인이다. 머뭇거림 없는 유창한 언어를 산출하는 것 역시 학습자 언어 능력의 중요한 지표가 되겠지만, 본고에서는 신속성보다는 문장 구성의 복잡성과 정확성에 초점을 두고 복잡성과 정확성만 다루고자 한다.

학습자의 언어 발달은 횡적으로 살펴볼 수도 있고 종적으로 살펴볼 수도 있다. 횡적으로 비슷한 조건에 있는 학습자의 자료를 다량으로 수집하여 분석하는 경우 보편적인 언어 발달 단계를 밝힐 수 있다는 장점이 있다. 그러나 제2언어 발달은 고정된, 선형의

발달 단계가 아닌 끊임없는 연속과 불연속, 진보와 후퇴를 반복하며 학습자의 내적, 외적 요인에 따라 다양한 변이성(variability)을 갖는다는 것이 여러 연구를 통해 밝혀지고 있다. 따라서 횡적인 자료로는 학습자 언어의 가변적인 특성이나 시간의 흐름에 따른 발달을 파악하는 데 한계가 있다. 이에 본 연구는 세 명의 중국 배경 중도입국 청소년 학습자를 대상으로 1년 6개월 간의 문어와 구어 산출 자료를 구축하여 다음과 같은 연구문제를 살펴보고자 한다.

1. 중국 배경 중도입국 청소년의 문어·구어의 복잡성과 정확성은 시간의 흐름에 따라 어떤 발달 양상을 보이는가?
2. 중국 배경 중도입국 청소년의 복잡성과 정확성은 산출 양식에 따라 어떤 차이가 있는가?
3. 중국 배경 중도입국 청소년의 문어·구어의 복잡성과 정확성은 어떤 관련성을 보이는가?
4. 중국 배경 중도입국 청소년의 문어·구어의 복잡성과 정확성은 학습자별로 어떤 차이가 있는가?

2. 선행연구 검토

본격적인 연구를 진행하기에 앞서 그동안 중도입국 학습자를 대상으로 진행된 한국어 교육 분야의 연구, 그리고 복잡성과 정확성 측정을 통해 제2언어 학습자의 언어 발달을 연구한 선행 연구들을 살펴보고자 한다. 중도입국 청소년의 한국어 발달 연구가 많지 않은 관계로 학습자 범위를 확장하여 살펴보되 본고의 연구 내용과 직접적인 관련이 있는 중도입국 자녀의 교육과정, 교육 내용, 교재 개발 연구나 문어와 구어의 복잡성과 정확성을 다룬 연구들을 보다 상세히 고찰한다.

2.1. 중도입국 청소년 대상 한국어 교육 연구

2000년대부터 시작된 다문화 가정 자녀에 대한 연구는 2006년 '다문화가정 자녀 교육지원 대책' 발표와 함께 본격적으로 확대되기 시작하였다. 그동안 한국어 교육 분야에서 진행된 다문화가정 자녀에 대한 연구는 학습자의 음운 인식 능력, 언어 환경, 이중언어 능력, 학업 성취 향상, 교재 개발, 교육 방안, 교육과정 등 다양한 분야에서 진행되어 왔다. 그러나 대부분의 연구가 결혼이주민 가정의 자녀인 국내 출생 다문화 가정 자녀에 주목하면서 유아기의 언어 발달이나 언어 발달의 지체 현상(배소영 외, 2009; 최진숙·고은, 2017), 다문화 아동과 부모의 언어교육 방법(오소정, 2015; 박영옥·이진

경, 2018), 일반 가정의 아동과 다문화 가정 아동의 언어 발달 비교 연구(김의수·김혜림, 2012)들이 주를 이루고 있었으며 중도입국 자녀에 중점을 둔 연구는 찾아보기 어려웠다.

중도입국 자녀는 초기에 '다문화가정 자녀'라는 큰 범주에 포함되어 크게 주목을 받지 못하다가 그 수가 점점 증가하면서 문화적으로 이질적 환경에서 성장해 온 중도입국 자녀와 국내 출생 다문화 자녀를 구분해야 한다는 주장(김명정, 2011)들과 함께 본격적으로 연구되기 시작하였다(윤새은임·유진이, 2017:63). 한국어 교육 분야에서 중도입국 자녀에 대한 논의는 주로 교육과정과 교재, 교수법이나 교육자료를 다루고 있는데, 그중에서도 교육과정, 교재 관련 논의가 가장 많았다.

먼저 한국어 교육과정과 관련된 논의는 김영란(2014), 손희연·조태린(2015), 이효인(2015), 홍종명(2016) 등이 있었는데 김영란(2014)에서는 '다문화'라는 용어의 무분별한 사용과 모호한 범위를 지적하면서 다문화 가정 자녀를 대상으로 한 한국어 교육과정 대상에 한국어를 제1언어로 학습하는 학생과 제2언어로 학습하는 학생이 혼재하는 점을 지적하였다. 손희연·조태린(2015)에서는 한국어 교육과정의 교육 대상인 '다문화 배경을 가진 학생'을 '한국어 학습' 학생으로 재설정하고 '생활 한국어'와 '학습 한국어'로 이원화된 내용 체계를 '학교 의사소통'을 위한 과제 중심 체계로 구성할 것을 제안하였다. 이효인(2015)에서는 전남지역에서의 중도입국 자녀의 한국어 교육 현황과 문제점을 살펴보아 중도입국 자녀 대상 한국어 교육의 과제를 제시하였고 홍종명(2016)에서는 한국어 교육 대상자로서 중도입국 청소년의 특징과 한국어 교육과정 운영상의 문제점을 지적하였다. 그리고 강수정(2018)에서는 여전히 중도입국 학습자를 위한 독립적인 교육과정 개발에 대한 상세한 논의가 부족함을 언급하면서 요구 분석을 통해 중도입국 학습자를 대상으로 한 통합적 교육과정 구축 방안을 제안하였다.

교재 관련 연구는 주로 이미 개발된 한국어 교재에 대한 분석을 통한 문제점 도출과 개선방안 제시가 많았는데 이민아(2013), 이경숙(2015)에서는 『표준 한국어』의 구성 내용을 분석한 뒤 문제점 지적과 함께 개선방안을 제시하였고 이루리(2018)에서는 실제 『표준 한국어』를 사용하여 수업을 진행하고 있는 다문화 예비학교의 교사와 학습자를 대상으로 설문조사와 면담 분석 결과를 바탕으로 교재에 대한 개선방안을 도출하기도 하였다. 그리고 이현경(2020)에서는 『중학생을 위한 표준 한국어 학습 도구』와 개정된 <한국어(KSL) 교육과정(2017)>을 비교 분석하여 교육과정의 핵심 논점과 수정사항이 새로 개발된 교재에서 구현된 정도를 파악하고 실제 현장에서 교재를 사용하는 학습자와 교사의 의견을 분석하여 교재의 활용도와 문제점 및 개선점을 제시하였

다.

학습 자료와 관련해서는 중도입국 학습자가 어려움을 느끼는 교과목과 한국어 교육을 연계한 학습 한국어 자료 개발이나 교수법에 대한 연구들이 진행되고 있었다. 신영권(2014)에서는 중고등학생 207명을 대상으로 한 설문조사 결과를 바탕으로 중도입국 자녀가 실제 학교생활을 하는데 필요한 문화 적응 관련 언어 재료에 필요한 내용을 도출하였고 오영훈 외(2012), 노정은·장미정(2019)에서는 중도입국 자녀가 한국어 능력이 매우 부족한 상태에서 한국 사회에 진입하기 때문에 교과 수업을 따라가기에 많은 어려움이 있음을 언급하면서 한국인 학생들과의 학습 격차를 줄이기 위한 읽기 자료 개발 방안을 제시하기도 하였다. 또한 김연희(2018)에서는 중도입국 자녀들이 한국 사회나 한국 문화에 대한 이해 부족으로 국어와 사회 역사 과목에서 더 어려움을 느낀다(원진숙, 2008:278)는 점을 언급하면서 초등학교 국어 교과서 내의 언어 문화 항목[3]을 추출하고 중도입국 자녀에게 필요한 언어문화 교육 내용을 선별하여 디딤돌 학습[4]으로 활용할 수 있는 교육 내용을 선정하기도 하였다. 그리고 이소현(2016)은 중도입국 자녀가 수학 과목에 어려움을 느낀다는 점에 주목하여 초등학교4-6학년과 중학교 1학년 수학 교과서의 텍스트를 바탕으로 수학 학습에 필요한 학습 어휘를 선별하였고 진송수(2020)에서는 중도 입국 학습자의 학교 적응을 지원하기 위하여 내용 중심 교수법을 기반으로 주제 중심 어휘 교수를 설계하고 직접 실행하여 그 효과를 검증하기도 하였다. 이 외에도 한국어 능력과 문화 이해를 위한 교육 항목과 교육 방안을 제안한 연구(이찬우, 2015), 학업 문식성 신장을 위한 교재 개발(이소라, 2012), 국어교과 학습을 위한 한국어 교재 개발(심나빈, 2013) 등 구체적인 기능이나 과목의 학업 수행 능력을 높이기 위한 교재 관련 연구도 많이 진행되었다.

교육과정이나 교재 외에 중도입국 학습자의 개인적 변인과 한국어 능력 간의 관계에 주목한 연구들도 있었는데 주로 학습자의 동기와 한국어 성취도를 연결시킨 연구가 많았다. 중도입국 학습자의 동기 관련 연구는 김영주 외(2012), 홍종명(2014), 권미경·이소연(2017), 김미강·김영주(2021) 등이 있는데 김영주 외(2012)에서는 국내 북한이탈청년과 중도입국 청년을 대상으로 문법 및 화용 능력을 측정하여 한국어 습득에 영향을 주는 학습자 변인을 살펴보았고 홍종명(2014)에서는 중도입국 학습자의 학습동기에 주목하여 자기결정성 동기 유형의 양상과 성별, 연령, 한국 거주기간 등 개인 변인

[3] 추출한 언어문화 항목은 존댓말과 반말, 준언어·비언어, 관용표현, 속담, 명언 등이다.

[4] 심혜령(2015:331)에서 처음으로 '디딤돌 단계'를 언급하였다. KSL 학습자를 위해 필요한 학습 한국어의 개념으로서 KSL 학습자의 교과교육을 위해서는 이 디딤돌 단계에서 다루어지는 언어들이 재료가 되어야 한다는 것이다.

과의 상관관계를 조사하였다. 그 결과, 중도입국 청소년은 낯선 환경에서 오는 정서적 어려움 등으로 인해 학습 동기 형성이 활발하지 않을 것이라는 통념과 달리 한국어 학습에 대한 동기 수준이 높다는 결과를 얻었다. 권미경·이소연(2017)에서는 질적연구 방법으로 중도입국 청소년의 한국어 학습 동기(통합적 동기, 도구적 동기)가 학업성취도에 미치는 영향을 살펴보았고 김미강·김영주(2021)에서는 중도입국 청소년의 정의적 요인(자아존중감, 자기효능감, 불안)과 한국어 숙달도와의 상관관계를 분석하였다. 이러한 연구들은 모두 중도입국 학습자들의 개인적, 환경적 변인과 학업을 결부시켜 큰 범위에서의 교육적 틀을 제공하고 있지만 실제적으로 학습자가 산출한 언어 자체에 주목한 연구는 거의 없다는 것을 알 수 있다.

중도입국 학습자의 언어 발달에 주목한 연구에는 장병순(2019), 한하림(2019a), 한하림(2019b), 조청화(2020), 조청화(2022)가 있는데 장병순(2019)에서는 중도 입국 중학생 20명의 쓰기 자료에 나타난 오류를 분석하여 유형화 시킨 뒤 중도 입국 자녀 학습자의 쓰기 교수-학습 지도안을 제안하였는데 학습자들에게서 과거 시제와 현재 시제의 대치 오류, 연결어미의 의미나 문법적 제약에 대한 이해 부족으로 인한 오류, 조사의 대치 오류가 빈번하게 나타남을 확인하였다. 한하림(2019a)에서는 17개월 동안 수집한 3명의 중도입국 아동의 일기 자료에 나타난 복문 산출 양상을 산출 길이, 문장 확장 유형, 복문의 빈도를 통해 고찰하였다. 그 결과 중도입국 아동의 복문 습득 과정이 한국어 모어화자 아동이나 여성결혼이민자 자녀와 다르며 비슷한 환경일지라도 개인에 따라 학습자 간의 통사 발달 편차가 크다는 것을 밝혔다. 한하림(2019b)에서는 한하림(2019a)과 같은 자료를 바탕으로 KSL아동의 연결어미 발달 과정을 질적으로 고찰하였다. 연구에서는 자료에 나타난 연결어미의 의미, 형태, 통사, 화용의 네 가지 측면을 분석함으로써 연결어미의 출현 시점 및 의미 범주별 발달 양상을 살펴보았는데 그 결과 제2언어 학습자의 중간언어는 역동적이며 비선형적인 발달 과정을 보이며 학습자 간 발달 양상에도 차이가 있지만 거시적 관점에서 볼 때, 한 번도 산출하지 않은 연결어미의 유형이 일치하거나 침묵기를 거친 다음 새로운 연결어미의 산출을 시도하는 등의 공통점이 있다고 하였다. 조청화(2020)에서는 중국 배경 중도입국 청소년의 쓰기 자료를 3개월 간 수집하여 중도입국 청소년의 연결어미 사용과 연결어미에 나타난 오류들을 분석하였다. 연구에서는 동영상 작문과 일기라는 두 가지 과제로 나누어 학습자 간에 나타난 차이와 과제 간에 나타난 차이를 세부적으로 비교하였다는 데 의의가 있으나 연구 참여자의 수가 적고 관찰 기간이 짧아 시간의 흐름에 따른 변화 발달 과정을 살피는 것에는 한계가 있었다. 조청화(2022)에서는 공교육에 진입한 3명의 중도입

국 청소년의 구어 자료를 7개월 간 수집하여 자료에 나타난 복잡성과 정확성을 종적으로 고찰하였다. 그 결과 학습자들은 시간이 지날수록 통사 복잡성과 어휘 복잡성을 동시에 발달시키지 못하며 명시적 교육이 없이 자연스러운 환경에서는 정확성이 향상되기 힘들다는 것을 발견하였다. 이 외에 중도입국 자녀의 언어 자료를 수집하여 분석한 연구로는 강란숙 외(2017), 김해영(2013)이 있으나, 언어 발달 자체에 주목하기보다 교수-학습 방법의 효용성 검증을 위한 과정이었다는 점에 본 연구와는 차이가 있다.

종합해 보면, 중도입국 청소년 학습자의 학업이나 한국어 습득에 관심을 둔 연구는 증가하고 있는 추세이지만 실제 중도입국 학습자의 언어 산출물을 바탕으로 한국어 수행을 분석한 연구는 아직 부족한 실정이며 대부분의 선행연구에서는 쓰기 자료만 수집하여 분석하고 있음을 확인할 수 있었다. 서론 부분에서도 언급하였듯이 학습자의 맞춤형 언어 교수를 위해서는 학습자 언어에 대한 이해가 선행되어야 한다. 그러므로 지금까지 진행된 교육과정이나 교재, 교수법의 효과를 검증하고 또 앞으로 더 개선된 교수 학습 방법을 개발하기 위해서는 실제 학습자의 언어 발달 과정을 살펴보는 것이 매우 중요하며, 지속적인 연구를 통해 그 결과가 축적될 필요가 있다.

2.2. 복잡성 관련 연구

한국어 중간언어 발달 연구에서 복잡성(complexity)에 대한 측정을 통해 학습자의 언어 발달을 기술하고자 하는 시도는 이영근(2005)에서 시작되었다. 이영근(2005)에서는 한국어 중간언어 발화의 형태 통사적 복잡성을 측정하기 위한 분석 도구를 소개하고 한국어 교육에서 관련 연구의 필요성을 주창하였다. 이후 박지순·서세정(2009)을 비롯한 여러 연구들에서 통사 복잡성을 중심으로 한국어 학습자의 언어 발달 양상을 밝히고자 하였으며 그중에서 가장 활발히 진행된 연구는 숙달도별 통사 복잡성의 발달에 대한 연구이다.

먼저 박지순·서세정(2009)에서는 영어권에서 통사적 숙달도 측정 도구로 사용되어 왔던 T단위 분석(T-unit analysis)을 한국어의 첨가어적인 특성에 맞게 T-unit 대신 문장을 단위로 하고 문장 구조 관련 지표를 추가하여 한국어 언어 발달 통사적 분석 지표 18가지를 제안하고 측정 지표들의 숙달도 변별 가능 여부를 검증하였다. 그리고 본격적으로 서세정(2009)에서는 이전의 한국어 발달 연구에서 모어 전이(L1 transfer)에만 초점을 두고 학습자가 생산한 오류를 중심으로 평가되어 왔던 것을 지적하면서 학습자의 중간언어를 통사적 구조의 관점에서 분석하였다. 연구에서는 초, 중, 고급 학습자의

작문 114편을 수집한 뒤 이를 한국인 대학생의 작문 38편과 비교 분석하였다. 그 결과 학습자의 숙달도에 따라 문장당 어절 수, 문장당 절 수, 문장당 형태소 수가 학습자의 숙달도에 따라 유의미하게 증가하였으나 숙달도가 높아짐에 따라 통사적 오류의 비율 또한 높아짐을 밝혔다. 이 연구는 한국어 교육 분야에서 최초로 학습자의 산출물을 통사적 하위 체계로 분류하여 분석하였다는 데 의의가 있다.

이후 복잡성 연구들이 점차 활발하게 나타나기 시작하였는데 연구자들은 주로 숙달도에 따른 복잡성을 분석함으로써 학습자들의 언어 발달을 가늠하거나 어떤 측정 지표가 숙달도 변별에 유용한지를 검증하기도 하였다. 숙달도와 복잡성을 연결시켜 분석한 연구들을 살펴보면 김목아(2010), 서신웨(2021), 남주연·김영주(2015), 남주연 외(2016), 최예슬(2018) 등이 있었다. 김목아(2010)에서는 초급과 중급의 중국인 학습자의 작문자료 60부와 한국인 대학생의 작문자료 15부에 나타난 '평균 발화 길이(mean length of utterance: MLU)[5]', 어휘 사용 양상, 복문 사용 양상을 분석하여 중국인 학습자 중간언어의 상대적 발달 정도를 측정하였다. 결과에 따르면 작문당 평균 낱말 길이, 평균 형태소 길이, 어휘 다양도, 복문 사용 정도 등 여러 측정지표가 초급, 중급, 한국인 대학생 순으로 증가하였으며 숙달도가 올라갈수록 다양한 연결어미를 사용한다는 것을 확인하였다. 서신웨(2021)에서는 미얀마 대학교 2학년~4학년에 재학 중인 한국어 전공 학습자 74명을 대상으로 통사적 복잡성 발달 양상을 분석하고 측정지표들이 학년별로 유의미한 차이를 보이는지 고찰하였다. 연구 결과 문장 길이와, 문장 확장 측정 지표의 평균값은 학년이 올라감에 따라 증가하는 양상을 보였지만 통계적 검증 결과 3학년과 4학년 집단 간의 차이는 유의미하지 않은 것으로 나타나 통사적 복잡성의 발달 정도가 미약해진다고 보았다.

그리고 김현아(2016)에서는 재미동포 청소년 학습자의 작문 자료를 통해 학습자들의 숙달도[6] 및 연령에 따른 통사 복잡성의 발달 양상과 관련 문법 항목의 사용 양상을 분석하였다. 그 결과 문장 길이, 문장 확장의 양, 문장 확장의 유형과 같은 통사 복잡성 측정 지표들은 숙달도 변인에서만 유의미한 차이를 보였고 연령 변인에서는 유의미하지 않다고 하였다. 최수완(2019)에서는 호주 교포 학습자의 작문 132부와 한국어 모어 화자 작문 117부를 분석하여 호주 교포 청소년 학습자의 통사적 복잡성 양상을 분석하

[5] 평균 발화 길이를 뜻하는 MLU는 일반적으로 구어 자료에서 분석척도로 사용하지만 김목아(2010)에서는 발화 단위를 문장 단위로 바꾸어 사용하였다.

[6] 김현아(2016)에서는 만 5세 이전에 이주하였거나 미국에서 출생한 재미동포 학습자를 대상으로 하였으며, 숙달도는 한글학교의 급을 TOPIK급수로 치환하여 3급~6급으로 수준을 나누었다.

였는데 전체적으로 학년 군에 따라 통사적 복잡성이 증가하는 양상을 보이지만 문장당 부사절 수와 문장당 인용절 수가 학년군에 따라 뚜렷한 차이를 나타내지 않았으며 모어화자가 학년별로 단계적으로 발달이 이루어지는 것에 비해 호주 교포 학습자는 그렇지 않음을 보여주었다.

통사 복잡성으로 숙달도나 발달을 측정하려는 연구들이 축적되면서 여러 측정 지표 가운데 숙달도를 잘 예측할 수 있는 지표를 검증하고자 하는 연구들이 진행되었는데 대표적으로 남주연·김영주(2015), 남주연 외(2016)과 최예슬(2018)이 있다. 남주연·김영주(2015)에서는 구어 자료를 바탕으로 숙달도(mini-TOPIK)와 영역별 점수에 따른 통사 복잡성의 차이와 상관관계를 검증하였는데 숙달도가 높을수록 발화 단위의 길이와 발화 단위당 절 수가 더 높았으며 숙달도 총점과 높은 상관성을 보이는 지표는 AS-unit당 절 수, AS-unit당 내포절 수, 대등지수임을 밝혔다. 또한 이 연구에서는 대등접속이 숙달도를 설명할 수 없음도 밝히고 있다. 남주연 외(2016)에서는 문어 자료를 바탕으로 숙달도(Mini-TOPIK) 등급에 따른 통사 복잡성 측정 값의 차이, 숙달도 총점 및 영역별 점수와 통사 복잡성의 상관관계를 분석하여 문어에서 숙달도를 예측할 수 있는 측정 지표를 찾고자 하였다. 그 결과 T-unit은 모든 측면에서 문장에 비해 나은 결과를 내지 못함을 발견하였고 문장당 내포절 수, 절 당 내포절 수, 문장 당 어절 수, 문장당 단어 수가 숙달도 변별에 유의한 지표임을 밝혔다. 그리고 내포절 중에서는 부사절만이 통사적 복잡성을 구하는 지표로 적절하며 특히 내포절 관련 지표는 어휘·문법, 쓰기, 듣기, 읽기 모든 영역과 상당한 상관성을 보인다는 결론을 제시하였다. 한편, 최예슬(2018)에서는 남주연 외(2016)의 연구 대상자가 대부분 중급에 속해있다는 점과 사후 검증 결과를 해석한 방식에 의문을 제기하며 그동안 선행연구에서 사용되어 온 측정 지표들을 정리하고 종합하여 23개의 분석 지표를 설정한 뒤 주제, 과제, 수행 환경과 숙달도가 통제된 학습자 100명의 자료를 바탕으로 어떤 지표가 숙달도에 대한 변별성이 있는지 확인하였다. 그 결과, 그동안 사용해 온 많은 통사적 복잡성 측정 지표 가운데 작문당 절 수와 관형사절 수의 값만이 숙달도 수준을 가늠할 수 있다는 결과를 제시하기도 하였다. 원미진 외(2017)에서는 어휘 다양도, 어휘 밀도, 어휘 세련도가 쓰기 능력을 측정하는 지표로 사용될 수 있는지를 검증하고자 한 연구이다. 연구에서는 중국인 중·고급 학습자의 작문 자료 40편을 수집하여 분석한 결과, 어휘 밀도와 어휘 정교도는 등급 간 유의미한 차이를 보였으나 어휘 다양도는 쓰기 점수와 관련이 없음을 밝혔다.

어휘 복잡성에 대한 연구들은 주로 어휘 다양도를 중심으로 이루어지고 있었는데 숙

달도에 따른 어휘 복잡성을 비교하거나 어휘 복잡성 지표와 쓰기 점수와의 관계를 밝힘으로써 어휘 복잡성과 쓰기 능력을 어느 정도 설명하는지 밝히고자 하는 연구들이 있었다. 대표적으로 안경화(2003), 진대연(2006), 배도용(2012), 이경(2015), 원미진 외(2017), 강정희(2020) 등이 해당된다. 안경화(2003)에서는 어휘 다양도가 어휘 구사력을 추측할 수 있는 지표임을 주장하며 숙달도 단계별로 어휘 다양도가 증가하고 있음을 밝히며 어휘 다양도가 어휘 구사력을 추측할 수 있는 지표이지만 텍스트 크기의 영향을 많이 받는다는 점에서 적절한 크기의 텍스트나 같은 크기의 텍스트를 대상으로 사용할 것을 제언하고 있다. 진대연(2006)에서는 일본어와 중국어권 학습자를 대상으로 쓰기 점수와 어휘량, 어휘 다양도[7], 어휘 수준, 어휘 빈도 간의 관계를 살펴보았다. 그 결과 일반적으로 모어화자의 경우 어휘 다양도가 높은 텍스트가 높은 평가를 받는 것과 달리 한국어 학습자의 글은 어휘 다양도가 낮은 글이 더 높은 평가를 받는 경향이 있음을 밝혔다. 배도용(2012)에서는 언어권별 학습자의 어휘 다양도와 어휘 밀도를 비교하였는데 중국어권 학습자의 어휘 다양도가 가장 높고 베트남어권 학습자의 어휘 다양도가 가장 낮아 언어권별로 어휘 다양도가 다르게 나타나며 중국어권 학습자는 내용어에서, 일본어권 학습자는 기능어에서 TTR이 더 높은 경향이 있음을 보고하였다. 그러나 동시에 분석 자료가 문어 텍스트임에도 불구하고 학습자들의 어휘 밀도가 40%를 넘지 못하였음을 지적하고 있다. 토픽(TOPIK)시험의 작문 문항 답안지에 나타난 어휘 복잡성을 분석한 이경(2015)에서는 어휘 다양도와 어휘 정교도는 학습자의 숙달도와 유의한 상관관계가 있지만 어휘 밀도는 숙달도와 큰 상관성을 보이지 않는다는 결과를 보고하였다.

통사 복잡성 관련 연구가 축적되면서 숙달도와 복잡성 외에도 학습자 요인과 통사·어휘 복잡성의 관계를 다룬 연구들도 나타나기 시작하였다. 강정희(2020)에서는 학습자의 어휘 능력에 영향을 주는 요인을 탐색하고자 학습자의 숙달도, 언어권, 성취도, 장르에 따른 어휘 복잡성을 분석하였다. 그 결과, 어휘 능력은 숙달도에 따라 증가하며 언어권, 언어 능력, 장르에 따라 차이를 보인다는 것을 발견하였는데 어휘 다양도와 어휘 밀도는 중국어권 학습자가 더 높았으나 어휘 정확도는 일본어권 학습자가 더 높다는 것, 어휘 다양도와 어휘 세련도는 언어 능력과 정적 상관을 보이지만 어휘 밀도는 언어

[7] 진대연(2006)에서는 '어휘 다양도'가 아닌 '어휘 밀도'라는 어휘를 사용하였지만 해당 연구에서 어휘 밀도는 어휘 유형 수(type)와 구현 수(token)의 비율로 측정하고 있다. 그러나 어휘 유형 수와 구현 수의 비율은 대부분의 연구에서 '어휘 다양도'로 표시하는 개념이고 본고에서도 '어휘 다양도'로 칭하고 있기 때문에 이해의 편의를 위해 '어휘 다양도'로 수정하여 제시하도록 한다.

능력(듣기, 읽기)과 부적 상관을 보인다는 것, 설명문에서는 어휘 다양도가 더 높고 주장문에서는 어휘 밀도와 세련도가 더 높다는 것을 밝혔다. 김영주 외(2017)에서는 구어 통사 복잡성과 학습자의 심리적 요인, 사회적 요인 및 숙달도 간의 관계에 대해 다루었는데 한국어 듣기 활동 시간이 명사절, 관형절, 인용절의 사용을 촉진시킨다는 점과 함께 구어 기반의 사회 활동보다 복합도가 높은 언어적 입력이 복잡성의 발달에 더 도움이 될 수 있다고 제언하였다. 그리고 통사 복잡성과 어휘 복잡성을 함께 다루며 그 사이의 관계를 밝히려고 한 연구도 있었는데 남주연(2015)에서는 한국어 학습자 55명과 모어화자 10명의 구어 자료를 분석하여 통사 복잡성, 어휘 복잡성과 숙달도의 관계, 통사 복잡성과 어휘 복잡성 간의 관계를 분석하였다. 분석 결과, 숙달도와 상관성이 높은 복잡성 지표는 발화 단위(AS-unit)당 단어 수, 발화 단위 당 내포절 수, 어휘 다양도과 어휘 정교도인 것으로 나타났다.

이 외에 유사 종단적 연구가 개개인 학습자의 발달 특성을 살펴보기 어렵다는 한계를 지적하며 적은 수의 학습자를 대상으로 학습자의 개별적 특징에 초점을 맞춘 종단적 연구들도 있었다. 먼저 음승민(2012)은 한국어 교육기관의 중국인 학습자 7명이 1년간 산출한 작문 자료를 분석하여 학습 단계에 따른 구문의 변화 과정을 양적으로 고찰하였다. 연구 결과에 따르면 학습 단계가 향상됨에 따라 문장의 길이가 길어지고 복문의 비율, 복문당 절 수 역시 증가하지만 단계별로 문장 길이에 영향을 미치는 문법 단위가 달라지고 개별 학습자마다 복문의 구조가 전 단계에 비해 크게 복잡해지는 단계가 다르거나 문장 길이의 증가 량이 다르게 나타나는 등 학습자 간의 차이가 있음을 밝혔다. 최유진 외(2022)에서는 구어 통사 복잡성 발달을 관찰한 연구로 이 연구에서는 초급 한국어 학습자 8명을 대상으로 5개월 간 7회의 자유 대화 말뭉치를 구축하고 이를 분석하여 통사 복잡성의 발달 양상을 살펴보았다. 이 연구는 학습자의 통사 복잡성 발달에 나타난 언어적 특성을 각 학습자의 개인적 변인(학습 목적 및 동기, 태도, 한국어 모어화자 접촉 빈도 및 한국어 사용 빈도, 오류에 대한 민감성, 한국어 수준에 대한 자가진단, 한국 문화 및 대중문화 콘텐츠 접촉 빈도 등)과 연결시켜 질적으로 분석하였다. 해당 연구는 초급 학습자가 단문 발화에서 처음 통사 복잡성이 발달하는 시점을 관찰하였다는 점, 통사 복잡성의 발달을 학습자의 개인적 변인과 연결시켜 분석하였다는 점에서 그 의의가 크다. 다만 질적 분석의 특성상 학습자의 수준을 가늠할 만한 수치에 대한 제시가 부족하여 개별 학습자의 발달을 직관적으로 알기에는 어려움이 있었다.

2.3. 정확성 관련 연구

정확성은 오류와 직결되는 개념이다. 따라서 정확성 관련 연구들은 주로 오류 분석의 차원에서 이루어지거나 복잡성과 연결하여 복잡성과 정확성의 관련성을 살펴보는 연구가 대부분이다. 이 절에서는 먼저 오류 연구에 대해 살펴보고 복잡성과 정확성을 함께 살펴본 연구는 다음 절에서 자세히 살펴보도록 한다. 한국어 교육 분야의 오류 분석 연구들은 초반에는 언어권별 오류의 현상이나, 오류 유형, 오류의 원인을 분석하는 연구들이 주를 이루었으나 최근에는 말뭉치를 기반으로 복잡성과 함께 발달의 관점에서 무오류절이나 오류의 비율을 측정한 연구들이 늘어나고 있는 추세이다. 우선 한국어 학습자의 오류의 범주나 현상, 오류의 판정에 대한 세부적인 기준을 제시하고 유형별 오류를 구분하고 분석한 연구에는 이정희(2002), 조철현(2002), 김유미(2002), 석주연·안경화(2003), 이정연(2017), 국립국어원(2018) 등이 있다.

이정희(2002)에서는 K대학교 언어교육원 학습자들의 쓰기와 말하기에 나타난 오류를 분석하여 국적별, 학습 단계별 차이를 살펴보아 학습자의 오류 유형은 대치, 누락, 첨가 순으로 나타나며 격 조사 오류, 연결어미 오류, 전성어미의 오류 모두 고급으로 갈수록 오류율이 낮아짐을 밝혔다. 특히 이 연구는 이전의 한국어 오류 연구가 미리 예상되는 오류 항목을 설정하고 예상 결과에 맞추어 연구 자료를 한정한 것을 지적하면서 한국어 문법에 알맞는 오류의 범주(발음, 어휘, 문법), 오류 현상(대치, 첨가, 누락), 오류 정도(전체, 부분)에 따른 분류 방법을 제시하였다는 점에 의의가 있다. 조철현(2002)에서도 각 대학교의 언어교육원 학습자들의 쓰기 자료를 바탕으로 자료에 나타난 오류를 어휘, 문법, 혼합, 철자의 언어학적 범주로 구분하고 다시 생략, 대치, 첨가, 환언, 형태 오류로 오류 현상을 세분화하여 각 오류 유형을 귀납적으로 제시하였다. 김유미(2002)에서는 대규모 말뭉치를 바탕으로 오류의 빈도와 유형, 오류의 원인을 세분화하여 분석한 연구인데 이 연구에서는 오류의 유형은 누락, 첨가, 대치로 나누고 오류의 원인은 모어의 간섭, 목표어 자체의 복잡성에 의한 과잉일반화, 심리적 원인으로 인한 회피로 범주화하기도 하였다.

석주연·안경화(2003)에서는 오류문에 대해 그 오류 정도와 학습자의 숙달도에 따라 교정 여부를 다르게 하며 최소한도의 오류를 기술할 것을 건의하였고 이정연(2017)에서는 오류에 대해 거시적 오류(구성, 문단, 주체, 문체)와 미시적 오류(어휘, 문법, 철자, 문장 부호, 대치, 누락, 첨가 등)로 나누어 살펴보며 더 넓은 범위에서의 오류 분류 기준을 제시하기도 하였다. 마지막으로 국립국어원(2018)에서는 학습자 말뭉치에 나타난 오류를 분석하였는데 오류 주석 체계를 오류 위치(명사, 형용사, 조사, 어미 등)와 오류

층위(발음, 형태, 통사, 담화), 그리고 오류 양상(누락, 첨가, 대치, 오형태)으로 나누고 이 외에 분석 불가능 범주를 설정하였다.

다음으로, 중간언어적 관점에서 오류를 분석한 연구를 살펴보면 차숙정·송향근(2006), 이정희(2008), 김호정·강남욱(2010a), 김호정·강남욱(2010b), 서세정·어지혜(2011), 강현화(2017) 등이 있다. 차숙정·송향근(2006)에서는 중급 수준의 중국어 학습자의 쓰기에 나타난 어휘 오류를 분석하였는데 형태적 오류, 어미 오류, 통사적 오류, 학습 전략 오류 중 형태적 오류가 가장 높게 나타났으며 한국어 어휘와의 의미적 차이로 인한 유의어 오류와 모국어의 간섭으로 인한 모국어 직역 오류의 빈도가 매우 높다는 것을 밝혔다. 이정희(2008)에서는 중국인 초·중·고급 학습자의 작문 자료 500편을 수집하여 어휘 오류에 나타난 오류의 원인을 분석하였다. 그 결과 모국어 전이에 의한 오류의 비중은 숙달도가 올라감에 따라 증가하며 발달 과정에 의한 오류는 초급에서, 전략 사용으로 인한 오류는 중급에서 가장 높게 나타난다고 하였다. 서세정·어지혜(2011)에서는 중국어, 일본어, 러시아어권 학습자의 작문 자료에 나타난 연결어미의 정확성을 분석하였다. 분석 결과, 모든 숙달도 단계에서 의미적 오류가 가장 높게 나타났으며 숙달도가 높아질수록 통사적 오류의 비율이 높아진다고 보고하였다. 강현화(2017)에서는 국립국어원의 1급~6급의 대규모 문어와 구어 말뭉치 자료를 사용하여 중국어권 학습자의 문어와 구어에 나타난 중간언어 양상을 분석하였다. 이 연구는 중국어권 학습자의 오류를 전체 학습자의 오류와 비교하여 중국인 학습자의 기능 어휘의 오류율이 높다는 점과 다른 언어권 학습자에 비해 오형태 오류가 지속적으로 높으며 초급에서는 누락 오류가 높고 중급과 고급에서는 대치 오류가 높아진다는 것을 확인하였다. 그리고 종적으로 오류의 변화 과정을 관찰한 연구도 있었는데 김호정·강남욱(2010a, 2010b)에서는 2년 동안 수집한 학습자 문어 자료를 바탕으로 연결어미 '-고'와 '-어서', 조사 '이/가', '은/는'에 나타난 오류의 시간의 흐름에 따른 변화 발달 과정을 추적하기도 하였다.

이 외에도 세부 문법 항목의 오류 양상을 살핀 연구들이나 대규모 말뭉치를 바탕으로 한 오류 분석 연구들이 꾸준히 진행되며 학습자의 오류 양상을 다양한 관점에서 밝히고 있었다. 그러나 대부분 연구가 문어 자료에 집중되어 있고 문어와 구어 자료의 비율이 동일하지 않거나 언어교육기관 학습자들의 자료를 바탕으로 한 연구가 많았다. 이에 원미진(2019)에서는 언어 교육 기관 밖의 학습자들을 대상으로 한 연구가 많이 이루어지지 않았음을 지적하기도 하였다. 본고에서는 앞에서 살펴본 오류 연구들을 바탕으로 본고의 오류 분석 체계를 수립하고 또 선행 연구의 오류 분석 결과와 본고의 분

석 결과를 비교하여 중국 배경 중도입국 청소년 학습자의 오류에 나타나는 특징들을 밝히고자 한다.

2.4. 복잡성과 정확성 관계 연구

앞에서 살펴본 것과 같이 제2언어로서의 한국어 습득 분야에서 복잡성과 정확성에 대한 연구가 축적되면서 최근에는 복잡성이나 정확성을 독립적인 지표로 인식하기보다 복잡성과 정확성의 관련성에 주목하여 언어 발달의 다각적인 측면을 고찰하려는 목적으로 복잡성과 정확성을 함께 살피는 연구들이 늘어나고 있다. 이러한 연구들에는 복잡성과 정확성의 발달을 종단적으로 살핀 김아름(2017), 주월랑(2018)이나 유사종단적 방법으로 장르에 따른 복잡성과 정확성을 살핀 주우동(2019), 서신웨(2022) 등이 있다.

김아름(2017)에서는 중국인 학습자 20명의 구어 발달을 10개월 간 4회에 걸쳐 종적으로 고찰하였다. 분석 결과, 학습자들의 정확성은 꾸준한 발달을 보였지만 복잡성과 유창성은 6개월이 지난 시점부터 퇴행과 지연이 나타났으며 기간별로 발달 속도에 차이가 있음을 발견하였다. 연구에서는 학습자들의 언어 능력이 개별 화행에 따라서 다르게 나타남도 밝히고 있는데 요청 화행은 10개월 간 꾸준히 발달하였으나 거절 화행은 6개월이 지난 시점에서 퇴행하였다고 보고하였다. 이 연구는 관찰 기간이 비교적 짧지만 과제에 따라 수행이 다르게 나타날 수 있다는 점을 보여주었다는 점에서 의의가 있다. 주월랑(2018)에서는 7명의 결혼이민자를 대상으로 8개월 동안 수집한 쓰기 자료를 분석하여 거주 기간과 복잡성 및 정확성 간의 상관관계를 살펴보았다. 그 결과, 여성결혼이민자들의 통사 복잡성은 시간의 흐름에 따라 양적으로 증가하는 추세를 보이지만 어휘 다양도와 어휘 밀도는 감소하였으며 오류율이 지속적으로 증가하며 정확성 역시 낮아지는 추세를 보였다. 특히 시간이 지날수록 오류가 어휘보다 문법에 집중되어 나타났는데 이러한 결과에 대해 자연습득 환경에서 주로 한국어를 습득하는 여성결혼이민자는 교실 내 학습자와 달리 통사와 어휘가 동시에 발달하기 어렵다는 것을 의미한다고 하였다.

주우동(2019)에서는 중국인 학습자의 문어 산출물 338편을 분석하여 숙달도와 장르에 따른 복잡성과 정확성의 차이를 살펴보았다. 연구 결과, 학습자들은 숙달도가 높아짐에 따라 통사 복잡성과 어휘 정교도는 지속적으로 높아지지만 어휘 다양도와 밀도는 숙달도 간에 거의 차이가 없으며 정확도 역시 숙달도에 따라 높아지는 것으로 나타

났다. 또한 설명문에서 문장당 연결어미 수가 더 높고 논술문에서 어휘 복잡성(다양도, 밀도, 정교도)이 더 높게 나타나 학습자의 언어 수행이 장르나 주제에 따라서 다르게 나타날 수 있음을 밝혔다. 그러나 이 연구에서는 복잡성과 정확성을 따로 제시하고 있어 두 요인 간의 관련성을 직관적으로 확인하는 것에 한계가 있었다.

우성숙(2021)은 중국인 성인 학습자의 구어 발달을 고찰하였는데 5개월 간 3명의 전화 한국어 수업의 음성 자료에 나타난 복잡성과 정확성, 유창성을 분석하여 학습자들의 단위 길이는 지속적으로 상승과 하락을 반복하였고 오류율은 시간의 흐름에 따라 감소하는 추세를 보이지만 학습자마다 세 구인의 발달이 동시에 나타나지 않으며 학습자 간 차이가 있다는 결과를 보고하였다.

마지막으로 서신웨(2022)에서는 토픽 3급~6급 수준의 학습자 109명의 문어 자료를 수집하여 숙달도와 장르에 따른 통사 복잡성, 어휘 복잡성과 정확성을 분석하고 복잡성과 정확성의 관계를 고찰하였다. 연구 결과, 어휘 밀도를 제외한 복잡성과 정확성의 모든 지표는 숙달도가 높아짐에 따라 향상되는 경향을 보이지만 유의미하게 증가하는 시기가 다른 것으로 나타났다. 그리고 장르에 따라 통사 복잡성과 어휘 복잡성의 지표들은 장르 간에 유의한 차이를 보인다는 것을 통해 장르 변인 복잡성에 영향을 미치는 중요한 변인임을 주장하였다. 또한 이 연구에서는 숙달도와 장르 변인에 따른 복잡성과 정확성의 관계를 밝히면서 과제와 숙달도 변인에 따른 한국어 학습자의 문어 산출을 다각적으로 고찰하였다.

지금까지 살펴본 복잡성과 정확성 연구들을 종합해 보면, 수많은 연구에서 다양한 국적의 학습자를 대상으로 숙달도나 과제의 차이에 따른 언어 수행을 비교하거나 시간의 흐름에 따른 언어 발달을 살펴보아 한국어 학습자들의 언어 발달에 나타난 복잡성과 정확성을 고찰하고 있음을 확인할 수 있다. 그러나 대부분의 연구에서 문어 자료에 초점을 두고 있어 구어 발달에 대한 연구는 드물다는 것을 알 수 있다. 다음 장에서는 본고와 비슷하게 산출 양식에 따른 복잡성과 정확성을 분석한 연구들을 살펴보도록 한다.

2.5. 문어와 구어의 복잡성과 정확성 관계 연구

한국어교육 분야에서 학습자의 문어와 구어를 동시에 살핀 연구는 많지 않으며 모두 성인 학습자를 대상으로 이루어져 왔다. 연구가 진행된 시간 순서에 따라 각 연구의 실험 방법과 결과에 대해 정리해 보면 다음과 같다.

김영주 외(2013)에서는 130명의 학습자를 대상으로 mini-TOPIK을 기반으로 숙달도를 구분한 뒤 숙달도별 문어와 구어의 정확성, 복잡성, 유창성을 살펴보았다. 그 결과 문어와 구어의 언어 산출은 모두 숙달도와 유의한 상관성을 보이며 발달하지만 문어의 정확성이 유창성이나 복잡성에 비해 굴곡이 더 많다는 것을 발견하였다. 문어와 구어의 지표를 비교해볼 때, 정확성은 문어에서 더 높게 나타나고 복잡성은 구어에서 더 높게 나타나지만 문어와 구어의 발달이 긴밀한 상관관계에 있으며 연쇄적으로 발달한다는 점을 밝혔다.

김정운 외(2017a)에서는 자연 학습 환경에 있는 한국어 학습자의 체류 기간에 따른 문어와 구어의 통사 발달을 학습자의 인지적, 심리적, 사회적 변인과 연결시켜 고찰하였다. 연구 결과, 모어화자와의 능동적인 상호 교류는 문어와 구어의 복잡성 발달에 기여하지만 정확성 발달에는 기여가 없음을 발견하였으며, 자기 주도적인 태도를 가질 경우 공식적 교육 없이도 문어와 구어의 복잡성과 유창성 발달이 가능하지만 정확성 발달은 어려움이 있음을 밝히면서 교육적 개입이 필요함을 주장하였다. 그리고 문어와 구어의 언어 산출을 비교하여 보면 통사 복잡성은 문어에서 더 높게 나타나고 정확성은 구어에서 더 높게 나타난다는 점을 밝히고 있다. 김정운 외(2017b)에서는 김정운 외(2017a)에서와 동일하게 자연 학습 환경의 학습자를 대상으로 같은 측정 방법을 사용하였으나 15개월 간 종적인 방법으로 학습자 6명의 언어 발달을 추적하였다. 연구에 따르면 체류 기간이 증가함에 따라 문어와 구어의 정확성이 향상되고, 문장 확장에 따른 복잡성의 발달 역시 뚜렷하게 나타났으나 그 발달 경로가 선형적이지는 않았음을 제시하고 있다. 학습자의 변인 중에서는 언어 적성이 구어의 통사 복잡성을 가장 잘 예측하는 변인이며, 사회적 변인 중 쓰기 활동량은 문어의 정확성, 말하기 활동량은 문어의 정확성을 예측하는 변인으로 나타나 자기 주도적인 언어 접촉을 통한 활동은 문어와 구어의 경계가 없이 언어 발달에 도움을 줄 수 있음을 발견하였다. 이 연구의 결과는 같은 측정도구를 사용한 김정운 외(2017a)에서 자연 학습 환경의 학습자의 경우 정확성의 발달이 유의하지 않다고 한 점과는 반대되는 결과이다.

남주연·김영주(2015)와 남주연 외(2016)는 같은 학습자들을 대상으로 한 등급별 구어 산출과 문어 산출에 대한 연구이다. 두 연구에서는 중급 수준의 한국어 학습자 55명의 언어 산출물을 분석하고 있는데 두 연구의 결과를 비교해 보면 내포절의 사용량은 문어와 구어가 비슷하지만 대등절의 사용량은 구어에서 더 높았으며 절 길이 또한 구어에서 문어에 비해 길다는 것을 확인할 수 있다. 또한 두 연구 모두 내포절과 단위당 절 수가 숙달도에 따른 차이를 보여준다는 점을 제시하고 있어 단위당 내포절 수와 단

위당 절 수가 산출 양식에 관계 없이 숙달도를 구분할 수 있는 유효한 지표임을 알 수 있었다.

이복자(2015)에서는 유사 종단 연구가 개별 학습자의 발달 과정을 일반화하여 학습자 간의 발달에 나타나는 차이를 관찰하는데 한계가 있음을 지적하며 1년 2개월 간 3명의 학습자의 쓰기와 말하기 자료를 바탕으로 개별 학습자의 시간의 흐름에 따른 복잡성, 정확성, 유창성의 발달 경로와 요인 간의 변화 발달 과정을 분석하였다. 이 연구에서는 같은 학습 과정을 거친 학습자일지라도 그 발달 양상은 각자 다르게 나타난다는 점과 복잡성, 정확성, 유창성 간에 경쟁-지지 관계를 반복하는 복합적인 관계가 나타난다는 점을 밝히면서 학습자의 언어 발달 경로는 고정된 선형의 패턴을 보이지 않는다고 설명하고 있다. 이 연구는 개별적인 학습자의 언어 발달 과정을 면밀히 보여주고 있으나 학습자 간의 차이점과 공통점, 문어와 구어의 발달 경로의 차이점이나 공통점에 대한 비교를 보여주지는 않고 있었다.

안의정(2017)과 남주연 외(2017)는 문어와 구어의 어휘 복잡성에 초점을 두고 연구를 진행하였다. 안의정(2017)에서는 다양한 장르로 구성된 한국어 말뭉치를 분석하여 문어와 구어의 어휘 풍요도를 비교하여 각 장르의 어휘 다양도와 어휘 밀도 모두 구어보다 문어에서 더 높은 것으로 나타난다는 점과 영어와 달리 대부분의 구어 텍스트의 어휘 밀도도 문어와 비슷하게 40%이상을 유지한다는 점을 제시하고 있었다. 그리고 한국어 학습자의 어휘 복잡성을 분석한 남주연 외(2017)에서도 비슷한 결과를 제시하고 있는데 한국어 학습자의 문어와 구어에 나타난 어휘 다양도를 숙달도별로 비교 분석한 결과, 숙달도가 높아짐에 따라 문어와 구어의 어휘 다양도는 모두 증가하지만 문어의 어휘 다양도가 더 높게 나타났다고 밝히고 있다.

황현경(2018)에서 역시 문어와 구어의 복잡성과 정확성을 숙달도에 따라 분석하였는데 통사 복잡성은 고급으로 갈수록 증가하였으나 정확성은 문어와 구어에서 다르게 나타났는데 구어에서는 숙달도가 증가하면서 발음 오류와 문법 오류가 줄어들지만 어휘 오류가 고급 단계에서 증가하는 반면 문어에서는 문법 오류와 어휘 오류가 모두 숙달도가 증가함에 따라 줄어들고 있다는 점을 제시하면서 말하기와 쓰기는 같은 표현 영역이지만 정확성과 복잡성이 동일하게 발달하지 않으며 비선형적인 양상을 나타낸다고 하였다.

지금까지 한국어 교육 분야에서 이루어져 온 복잡성과 정확성에 대한 연구들을 살펴본 결과, 다음과 같은 연구 성과들을 확인할 수 있었다. 첫째, 측정 지표의 유의미성을 검증하려는 연구들을 통해 학습자들의 언어 능력을 예측할 수 있는 유효한 측정 지표

를 찾을 수 있었다. 둘째, 복잡성과 정확성은 학습자의 개인적 변인, 학습 단계, 숙달도에 따라 다르게 나타나며 산출 과제나 산출 양식에 의해서도 영향을 받을 수 있음을 알 수 있었다. 셋째, 학습자의 문어와 구어, 복잡성과 정확성은 독립적으로 발달하는 것이 아니며 서로 연결되어 영향을 주면서 상호작용 속에서 발달한다.

그러나 동시에 선행 연구들의 제한점을 도출할 수 있었다. 첫째, 대부분의 연구가 한국어교육센터에서 공부하는 성인 학습자들을 대상으로 하고 있어 연구 대상에 다양성이 부족하였다. 자연 학습 환경이나 결혼이주여성, 중도입국 청소년과 같은 특수 목적 한국어 학습자의 경우 일반 성인 학습자와는 다른 환경에 처해있다는 점에서 이들의 한국어 발달 과정을 밝히는 연구들이 필요하다. 둘째, 여전히 대부분의 연구가 유사 종단 연구의 방식으로 학습자의 숙달도와 언어 산출 간의 관계를 살피고 있어 동일한 학습자의 언어 발달 과정을 지속적으로 관찰한 연구가 드물다. 여러 통시적 연구에서도 밝히고 있듯이 언어의 각 측면은 선형적으로 꾸준히 발달하지 않으며 서로 다른 발달이나 퇴행을 보인다. 그러므로 제2언어 학습자의 중간언어 발달 과정을 다각도로 관찰하기 위해서는 동일 집단의 언어 발달을 지속적으로 관찰하는 것이 필요하다(김정운 외, 2017b). 셋째, 학습자 언어 발달의 복잡성, 정확성, 유창성 각각의 요인에 대한 연구는 많지만 대부분 연구가 숙달도별 복잡성 양상이나 오류 양상을 기술하는 데 그치고 있으며 복잡성과 정확성의 관계를 살핀 연구가 부족하다. 언어의 각 측면은 서로 연결되어 있기 때문에 각 요인의 발달뿐만 아니라 요인들 사이의 상관관계를 밝히는 연구가 더 이루어질 필요가 있다. 넷째, 대부분의 연구가 문어의 통사 복잡성과 정확성 발달에 집중되어 있고 문어와 구어를 함께 다루거나 발달 과정을 비교한 연구가 미진하다. 앞에서도 살펴보았듯이 같은 학습자라고 할지라도 문어와 구어는 서로 다른 산출을 보인다. 따라서 문어와 구어의 발달을 비교한 연구들이 더 이루어진다면 기존의 연구 성과들을 더 보완하는 정보를 제공해줄 수 있을 것으로 기대한다.

이에 본 연구에서는 중도입국 청소년 학습자의 한국어 발달 과정을 다각도로 고찰하기 위해 종적 연구의 방식으로 문어와 구어의 발달 과정을 모두 살펴봄과 동시에 복잡성과 정확성의 상관관계를 분석하여 중도입국 청소년의 한국어 발달 과정에 나타나는 특징을 분석함으로써 한국어 교육을 위한 시사점을 도출하고자 한다.

3. 연구 범위 및 절차

본 연구는 중국 배경 중도입국 청소년의 문어·구어에 나타난 복잡성과 정확성을 분석함으로써 시간의 흐름에 따른 한국어 발달 과정을 고찰하고 산출 양식에 따른 복잡성과 정확성의 차이, 복잡성과 정확성의 상호관계를 탐색하는 연구이다. 본 연구의 연구 절차는 다음과 같다.

1장 서론에서는 연구의 목적 및 필요성을 밝히고 연구문제를 제시하였다. 그리고 그동안 중도입국 학습자의 한국어교육에 주목한 연구들을 살펴보아 본 연구의 필요성을 확인하고 한국어 학습자의 복잡성과 정확성의 발달을 다룬 연구, 복잡성과 정확성의 관계를 고찰한 연구, 문어와 구어의 복잡성과 정확성을 동시에 다룬 연구들을 검토하여 선행연구들의 제한점을 도출하고 본 연구의 결과가 기여할 수 있는 부분을 밝혔다.

2장에서는 기존의 선행연구들에서 정의되어 온 복잡성과 정확성의 개념 및 측정 방법들을 살펴보고 본고에서 정의하는 복잡성과 정확성의 범위를 규정한다. 복잡성과 정확성의 범위는 연구의 목적에 따라 그 범위를 다양하게 나눌 수 있는데 본 연구에서는 복잡성은 통사적 복잡성과 어휘 복잡성으로, 정확성은 무오류와 오류의 비율로 살피고자 한다. 또한 문어와 구어의 언어 산출 모델과 복잡성과 정확성의 상관관계에 대한 이론들을 살펴보아 본 연구의 연구 가설을 세운다.

3장에서는 연구를 수행하기 위한 구체적인 연구 방법을 제시한다. 먼저 학습자의 문어와 구어 산출물 수집을 위한 과제를 선정하기 위해 과제의 복잡성 구성 요소를 살펴보아 학습자들의 수준에 적합한 과제를 선정한다. 다음 자료 분석에 사용될 통사 복잡성, 어휘 복잡성, 정확성의 측정 지표를 선정하고 구체적인 분석 기준을 제시한다. 마지막으로 본고의 자료 분석의 절차와 통계 분석 방법을 제시할 것이다.

4장에서는 3장에서 설정한 측정 지표 및 분석 기준을 바탕으로 수집된 자료들을 분석하여 그 결과를 제시한다. 이 과정에서 시간의 흐름에 따른 복잡성과 정확성의 발달 양상을 살펴보고 더 나아가 문어와 구어의 복잡성과 정확성의 비교를 통해 산출 양식에 따른 복잡성과 정확성의 공통점과 차이점을 고찰하며 마지막으로 복잡성과 정확성의 상호관계를 살펴볼 것이다.

5장에서는 4장의 연구 결과를 바탕으로 중국 배경 중도입국 청소년의 복잡성과 정확성의 발달 및 산출 양식에 따른 공통점과 차이점, 구인 사이의 상호 관련성에 대해 기존의 선행연구들과 비교하여 논의를 진행하고 중국 배경 중도입국 청소년 학습자를 위한 교육적 함의를 도출하는 것으로 연구를 마무리하도록 한다.

Ⅱ. 이론적 고찰

1. 제2언어 습득에서의 복잡성과 정확성

모국어 습득이나 제2언어 습득 분야에서 지속적으로 논의가 되고 있는 것은 '학습자가 무엇을 알고 있는지'를 측정하는 것이다. 학습자가 알고 있는 것, 즉 학습자의 수준에 대한 정확한 평가는 단순히 숙달도를 구분 짓는 것을 넘어서 학습자 변인별 습득 양상을 살피거나 교육적 처치의 효과를 검증하고 발달을 추적함에 있어 가장 기본적인 요소이다. 이러한 필요성에 의해 제2언어 습득의 발달 지수(developmental index)를 찾기 위한 학자들의 노력은 오래전부터 이루어져 왔다.

초기의 제2언어 습득 연구에서는 아동의 모국어 발달을 측정하는 데 사용되는 MLU를 사용하여 제2언어 학습자의 언어 발달 수준을 가늠하기도 하고 필수 경우 분석(obligatory occasion analysis: Brown, 2013)이나 빈도 분석(frequency analysis: Cancino et al., 1978)을 적용하여 특정 언어 형태의 습득을 살펴보거나 습득의 순서를 밝히기도 하였다. 그러나 아동의 모국어 발달 측정 지표를 인지적으로 성숙한 성인 외국어 학습자에게 적용하는 것은 적합하지 않다는 지적(Larsen-Freeman & Strom, 1977)에 따라 '객관적인' 측정을 통해 제2언어 발달 수준을 결정할 수 있는 발달 지수 개발 필요성이 제기되었다. 이러한 발달 지수 개발의 중요한 목적 중 하나는 연구에서 학습자의 발달 수준을 보다 정확하게 설명하는 것인데 이는 연구 간의 비교 가능성을 높임과 동시에 언어 능력 시험에 접근할 수 없는 제2언어 교사들도 사용할 수 있다는 장점이 있다(Wolfe-Quintero 외, 1998:3).

CAF 중 가장 먼저 출현한 것은 정확성이었는데 이는 형태소 습득 연구에서 형태소 사용의 정확성을 점수화하여 습득 순서를 결정하던 것(Dulay & Burt, 1973)과 관련이 있다. 즉 학습자가 정확하게 사용한 것이 곧 습득된 것이라고 보았기 때문에 학습자의 형태적 정확성으로 숙달도를 측정하고자 한 것이다. 이에 Larsen-Freeman & Strom(1977)에서는 Hunt(1965)에서 모국어 발달 측정에 사용했던 T-unit을 제2언어 발달 연구에 적용하여 무오류 T-unit의 평균 길이를 측정하는 것으로 제2언어 학습자가 산출한 언어의 길이와 정확성을 측정하여 학습자의 숙달도를 가늠하고자 하였다. 그러나 언어 학습에서 형태 중심이 아닌 의미 중심의 의사소통 교수법이 대두됨에 따라 학습자의 언어 능력에 대한 초점이 '목표어의 표준적인 언어 기준'을 지키는 형태적 정확성에서 '멈추거나 머뭇거림이 없이 모어화자와 같은 속도로 언어를 산출하는

(Lennon, 1990)' 유창성으로 옮겨갔다. 그러나 또 다시 유창성만으로는 정확한 의사소통이 어렵다고 느끼면서 1980년대부터는 '유창성'과 '정확성'의 개념을 구분하여 제2언어 교수학습 활동을 계획하기도 하고 학습자의 언어 수행을 평가하기 시작하였다. 대표적으로 Brumfit(1984:52-57)에서는 교실 활동을 '유창성 지향 활동(fluency-oriented activities)'과 '정확성 지향 활동(accuracy-oriented activities)'로 나누어 제시함으로써 정확성과 유창성의 개념을 '의사소통의 유창성'과 '문법과 언어 형태의 정확성'으로 구분하였다. 이는 처음으로 유창성과 정확성의 개념을 양분하여 제시한 연구이다(Housen & Kuiken, 2009:1).

이어서 마지막으로 출현한 복잡성은 과제 중심 교수법의 등장과 관련이 있다. 과제 중심 교수법이 등장하면서 학자들은 '더 나은 언어 산출을 이끌어 내는 과제를 찾는 것'과 '과제의 난이도를 구분'하려는 노력(Long, 1989)을 하게 되었고 이 과정에서 자연스럽게 난이도, 즉 '복잡성'의 개념이 등장하게 되었다. 복잡성 중 언어적 측면에서의 복잡성은 과제 변인에 따른 학습자의 언어 수행을 측정하면서 시작되었는데 정확성과 유창성만으로는 과제 수행에 나타난 학습자의 언어를 제대로 측정하기에는 어려움이 있었다. 따라서 무오류 T-unit당 단어 수(words per error-free T-unit)나 발화 당 담화표지(discourse markers per utterance)와 같은 정확성과 유창성 지표 외에도 어휘 유형 수와 구현 수의 비율(type token ratio, TTR), T-unit당 내포절 수(subordinate clause), 발화당 내포절 수(Crookes, 1989) 등의 지표를 사용하여 내용에 대한 조직성이나 구문의 다양성과 정교함을 측정하기 시작하였다. 이후 Skehan(1996, 1998)에서 이러한 측정 지표들을 복잡성, 정확성, 유창성의 세 차원으로 묶어 숙달도 모델(proficiency model)을 도입하면서 세 구인의 구분이 시작되었다.

Foster & Skehan(1996:303)에서는 복잡성과 정확성은 모두 형태에 집중하지만 주안점에는 상당한 차이가 있다고 지적하였다. 특히, 복잡성은 산출한 내용의 조직성과 언어의 정교함과 다양함에 주목하는 것으로 이는 학습자가 더 복잡한 하위 체계를 발전시키는 과정에서 나타나는 재구조화와 위험 감수 의지와 관련되는 것이고 정확성은 복잡성과 달리 상대적으로 간단하더라도 잘 통제된 언어 형식을 사용하여 목표어와 가까운 언어 사용을 하는 것과 관련된다고 하였다. 즉 복잡성은 중간언어 체계의 변화, 발전과 성장 기회와 관련되는 것인 반면 정확성은 특정 중간언어 체계에서의 통제 능력과 관련이 있다(Foster & Skehan, 1996:304).

언어적 측면에서 복잡성의 등장은 정확성과 유창성만으로는 설명할 수 없었던 학습자 언어의 원숙함, 깊이, 다양성, 정교함과 세련됨(Housen 외, 2012)을 설명할 수 있게

하였다. 1990년대부터 복잡성, 정확성, 유창성은 제2언어 학습자의 언어 수행을 측정하는 독립적인 위치를 차지하는 3개의 언어적 차원(dimension)으로 이론화 되었다(Larsen-Freeman, 2006:582). 이후 복잡성, 정확성, 유창성의 개념은 제2언어 습득과 사용에 관한 수많은 연구에 사용되어 왔다. 비록 CAF는 그 자체만으로 이론을 구성하거나 연구 항목이 되지는 않지만 CAF 3차원 구인의 확립은 학습자의 언어를 '언어 체계, 언어 형태, 언어 속도'라는 세 가지 서로 다른 측면에서 측정함으로써 경험적이고 모호하던 평가를 수치로 명확하게 나타낼 수 있게 하였다. 연구자들은 CAF 세 구인을 학습자 수준이나 과제 특성과 같은 독립변수의 변동을 평가하기 위한 종속변수로 사용하기도 하고(Pallotti, 2009:590) 사전 사후 실험 결과를 측정하여 발달의 변화를 검토하거나 종적인 관점에서 학습자의 발달 경로를 추정하기도 하며(Michel, 2017:2) 연구에 다양하게 활용하고 있다. 이같이 다양한 연구들이 실행되어 오면서 이제 CAF 세 구인은 학습자의 언어 발달 수준을 정확히 설명할 수 있는 가장 기본적이면서 확실한 척도로 자리매김을 하였다.

언어 발달 연구에서는 이 세가지 구인을 함께 사용하기도 하고 연구자의 목적에 따라 따로 사용하기도 하는데 연구자는 수집된 자료나 연구의 목적에 따라 CAF 구인을 선택하여 적용하는 것이 중요하다. CAF 세 구인 중 유창성은 시간 및 실시간성이 중요한 변수로 작용하는 구인이기 때문에 일반적으로 문어 연구보다는 구어 연구에 많이 사용된다. 본고에서는 문어와 구어 자료를 모두 다루고 있다는 점에서 복잡성, 정확성과 더불어 유창성을 다룰 수도 있으나, 학습자들의 문장 구성 능력에 초점을 두고 있으므로 복잡성과 정확성만 다루도록 한다.

한편, 정확성과 복잡성은 많은 연구에서 논의가 된 만큼 그 정의와 측정 범위, 측정 방법이 매우 다양하다. 이 중 정확성은 상대적으로 명확히 정의할 수 있는 개념이지만 복잡성은 그 정의가 비교적 모호하고 범위가 넓다. 따라서 다음 절에서는 복잡성과 정확성의 개념과 측정 방법들을 살펴보고 본고에서 정의하는 복잡성과 정확성의 범위와 개념을 밝히도록 한다.

2. 복잡성의 개념 및 측정

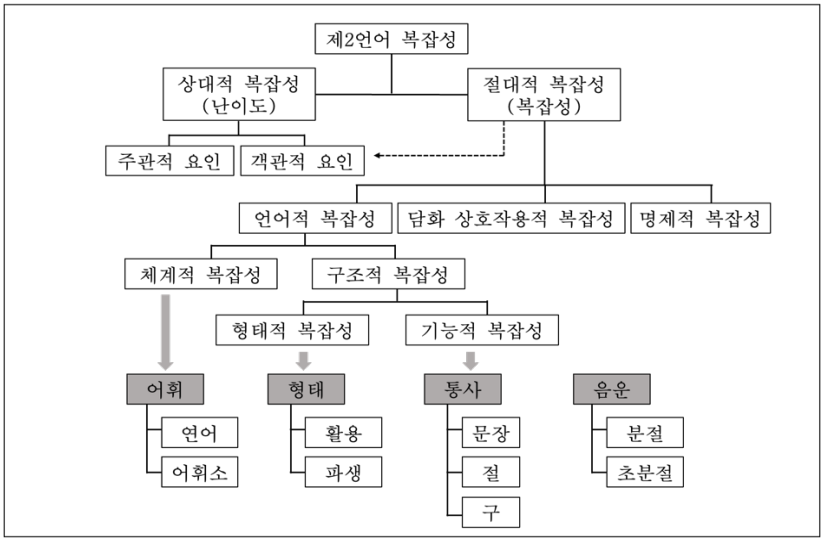
복잡성은 복잡성, 정확성, 유창성 중 가장 늦게 나타났지만 가장 다양한 관점에서 많이 논의되어 온 구인이기도 하고 개념을 정의하기가 가장 어려운 구인이기도 하다. 앞에서 살펴보았듯이 '복잡성'이라는 용어는 다차원적인 것으로 언어 산출 과제 자체의

복잡성을 나타내는 과제 복잡성과 학습자의 언어 수준을 가늠하는 지표로서의 언어 복잡성으로 나눌 수 있다. 본고에서는 학습자의 언어 발달을 고찰하기 때문에 언어적 복잡성에 초점을 두어 그 개념 정의와 측정 방법을 살펴보도록 한다.

복잡성에 대한 정의는 여러 연구들에서 찾아볼 수 있는데 초기에는 주로 학습자 언어의 재구조화 과정에서 나타나는 복잡하고 다양한 정도라는 측면에서 접근하여 포괄적으로 정의하였다. Skehan(1998)에서는 복잡성에 대해 '과제 수행에 있어서 학습자가 생산하는 언어의 정교함과 다양함'으로 정의하였고 Skehan & Foster(1999:96-97)에서는 '자신의 언어를 효과적으로 조절하지 못하더라도 더 발전된 언어를 사용하기 위하여 모험 시도를 하는 능력으로 이는 위험을 감수하려는 의지, 그리고 재구조화(restructuring)와 관련된다'고 하였으며 Wolfe-Quintero 외(1998:4)에서는 '제2 언어 지식을 확장하고 재구조화하는 범위를 드러내는 것'으로 정의하였다. 즉 복잡성은 학습자 중간언어 체계의 변화, 발달 중에 나타나는 재구조화와 관련되는 것으로 재구조화 과정에서 나타나는 언어의 정교함과 다양함을 가리킨다고 볼 수 있다. 그러나 언어는 다양한 하위 체계들을 내포하고 있으므로 언어의 정교함과 다양함을 어떤 측면에서 정의할 것인지를 결정해야 한다. 이는 복잡성에 대한 정의가 실증 연구에서의 조작적 정의, 즉 측정 지표에 영향을 주기 때문이다.

제2언어 복잡성의 각 측면과 구성 요소에 대해 비교적 체계적이고 포괄적으로 정리한 연구는 Bulté & Housen(2012)을 들 수 있다. 이 연구는 제2언어 복잡성을 구성하는 요소들을 세분화하여 체계적으로 정리하고 1995년부터 2008년 사이에 진행된 40편의 실증적 과제 중심 연구들을 분석하여 언어 복잡성 연구의 유형과 측정 방법을 조사하였다. Bulté & Housen(2012)에서 정리한 복잡성의 구인 및 구인 사이의 관계는 다음의 <그림 1>과 같다.

<그림 1> 복잡성 구인의 분류 (Bulté & Housen, 2012:22-25)



Bulté & Housen(2012)는 제2언어 복잡성에 대해 언어 학습자가 느끼는 상대적인 어려움인 인지적 복잡성[8]과 언어 자체의 난이도 즉 언어적 요소에 대한 복잡성으로 구분하여 이를 상대적 복잡성(relative complexity)과 절대적 복잡성(absolute complexity)으로 명명하였다.

상대적 복잡성은 언어 사용자의 관점에서 정의되는 것으로 특정 언어 요소를 처리하고 내재화하기 위해 언어 사용자가 쏟아야 하는 정신적 노력, 즉 난이도(difficulty)와 관련된다. 그리고 상대적 복잡성은 다시 태도, 동기, 언어 적성, 기억 용량 등 개인적 차이를 발생시키는 주관적 요인(subjective determinants)과 언어 입력 중 제2언어 특징에 대한 지각 현저성(perceptual saliency)과 출현 빈도, 의사소통적 부담 등을 나타내는 객관적 요인(objective determinants)으로 나뉘는데 이러한 객관적 요인은 또 언어 자체의 복잡성인 절대적 복잡성도 포함한다. 절대적 복잡성은 제2언어 자체의 언어적 요소가 가지는 고유한 복잡성을 의미하는데 언어 체계를 구성하고 있는 구성 요소의 수량이나 구성 요소 간의 연결 수량과 같은 양적이고 객관적인 복잡성을 의미한다. 그렇기 때문에 절대적 복잡성은 계량화하여 객관적으로 측정하는 것이 가능하며 L2 학습자의 언어 수행을 측정하기 위해 쓰이는 '복잡성'과 동일한 개념이다(주월랑, 2018:12).

[8] 인지적 복잡성은 상대적이고 주관적인 개념으로 제2언어 수행 및 제2언어 학습 중 언어 요소가 처리되는 상대적인 어려움을 의미하며 이는 학습자의 개별적 배경인 적성, 동기, 제2언어 발달 단계, 모국어 등 요인에 의해 부분적으로 결정된다(Housen 외, 2012:4).

절대적 복잡성은 다시 언어적(linguistic) 복잡성, 담화-상호작용적(discourse-interactional) 복잡성과 명제적(propositional) 복잡성으로 구분된다.

이 중에서 언어적 복잡성은 절대적 복잡성의 첫 번째 구성 요소로 학습자의 언어 체계에서 확장된 정교함과 풍부함의 정도를 의미한다. 그리고 언어적 복잡성은 다시 학습자 L2 체계의 전체적(global) 혹은 체계적(system) 복잡성과 학습자의 L2 체계를 구성하는 개별 언어 항목, 구조 또는 규칙의 보다 안정적인 속성을 가리키는 국소적(local) 또는 구조적(structure) 복잡성으로 나눌 수 있다.

체계적 복잡성은 학습자 언어의 폭(breadth), 즉 크기와 관련되는 것으로 얼마나 폭넓고 풍부한 언어 체계를 가지고 얼마나 넓은 범위에서 다양한 구조나 언어 항목을 사용할 수 있는지를 의미한다. 따라서 체계적 복잡성은 학습자가 음운 체계나 어휘, 문법 등에 얼마나 숙달되었으며 그것을 얼마나 잘 조절할 수 있는지를 보여 준다. 반면, 구조적 복잡성은 폭(breadth)이나 범위(range)보다는 깊이(depth)와 관련된 것으로 다시 형태적 복잡성(formal complexity)과 기능적 복잡성(functional complexity)으로 세분된다. 기능적 복잡성은 언어 구조의 의미나 기능의 수량, 언어의 형태와 의미/기능 간 대응(mapping)의 명료성이나 다양성 정도를 가리키고 형태적 복잡성은 목표어의 형태를 구성하는 데 필요한 개별 요소의 수량, 즉 기본형에 결합되는 언어 요소들의 개수를 가리킨다.[9] 그리고 이와 같은 언어적 복잡성의 각각의 하위 체계는 또다시 어휘, 음운, 형태, 문법의 네 영역으로 나뉘고 각 영역 안에서 활용이나 파생의 복잡성, 구, 절, 문장의 복잡성으로 나뉠 수 있다.

이 도표를 바탕으로 Bulté & Housen(2012:25-26)에서는 제2언어 복잡성을 살필 때, 학습자 L2 체계의 음운 체계, 어휘 체계, 형태 체계, 통사 체계 등 측면에서 전체적 정교도 혹은 체계 복잡성을 다룰 수도 있고 제2언어 학습자의 언어 체계를 구성하는 제2언어 체계의 각각의 통사, 형태, 음운, 어휘의 국부적 기능이나 형태 복잡성을 측정할 수도 있다고 하였다. 즉, 음운적, 형태적, 어휘적, 통사적 측면에서 보는 복잡성 또는 특정 언어 항목, 문법 규칙, 언어 구조의 측면에서 보는 복잡성은 모두 제2언어 복잡성에 해당한다(남주연, 2015:35). 그러나 본고에서는 학습자의 제2언어 체계 발달을 전체적인 관점에서 살피고자 하므로 특정 언어 항목이 아닌 학습자 언어 체계의 폭, 즉 체계적 복잡성을 다루고자 한다.

지금까지의 선행연구들을 살펴보면 언어적 복잡성 중에서 어휘 복잡성과, 통사 복잡

[9] 예를 들면, 현재 완료형이나 과거형을 표시하는 형태적 요소의 수나 능동형에서 피동형으로 파생될 때 필요한 형태적 요소의 수 등이 형태적 복잡성에 해당된다(Bulté & Housen, 2012:25).

성을 중심으로 연구들이 이루어져 왔다. 음운과 형태 역시 언어에서 간과할 수 없는 중요한 요소이지만 언어에서 어휘는 의미 전달의 가장 기본적인 요소이고 통사는 이러한 어휘들을 연결하고 배치하여 문장을 구성하는 능력이라는 점에서 이 두 가지 능력으로 학습자의 언어 산출을 측정하기에 적합하다고 볼 수 있다. 이에 따라, 본고에서는 어휘와 통사 복잡성으로 연구 범위를 한정하고자 한다. 아래 통사 복잡성과 어휘 복잡성의 개념에 적합한 측정 방법을 살펴보도록 한다.

2.1. 통사 복잡성의 개념 및 측정

2.1.1. 통사 복잡성의 개념

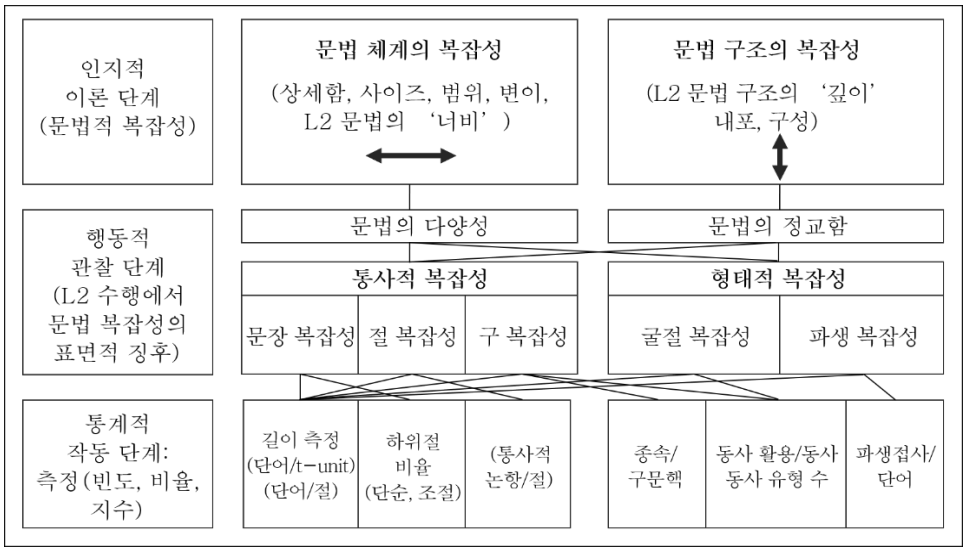
언어 복잡성 중에서 통사적 복잡성은 '언어 발화에서 나타나는 다양한 형태의 범위와 그 형태들의 정교한(sophistication) 정도(Ortega, 2003:492)'로 통사적 숙달도(syntactic proficiency) 혹은 통사적 성숙도(syntactic maturity)[10]라고 불리기도 한다. 또한 통사적 복잡성은 주로 문법 구조의 정교함이나 다양성 정도를 통해 측정되므로 문법적 복잡성(grammatical complexity)과 비슷한 개념으로 볼 수도 있다.[11]

Bulté & Housen(2012)에서는 통사 복잡성을 문법 복잡성의 하위 분류로 정의하고 문법 복잡성에 대해 이론 단계(theoretical level), 관찰 단계(observational level), 통계적 측정 단계(operational level)로 구분하여 문법 복잡성의 분류부터 측정 방법까지의 단계들을 <그림 2>와 같이 도식화하여 제시하고 있다.

[10] Gaies(1980:53)에서는 통사적 성숙도에 대해 '많은 양의 의미 혹은 정보를 적은 단어로 압축적으로 표현할 수 있는 능력'으로 정의하였다.

[11] Wolfe-Quintero 외(1998)에서는 학습자들이 다양하고 정교한 구나 문법 구조를 얼마나 생산하는지를 통해 복잡성을 측정하였고 Lu(2011:36)에서는 제2언어 쓰기에서 나타난 문법의 다양성과 정교함, 혹은 학습자가 생성한 통사 구조의 범위와 그러한 구조의 정교함의 정도를 통사 복잡성이라고 하였다.

<그림 2> 문법 복잡성의 단계(Bulté & Housen, 2012:27)



먼저, 이론 단계를 살펴보면 앞 절의 언어적 복잡성의 구분에 따라 문법 복잡성을 체계적 복잡성과 구조적 복잡성으로 구분한 것을 확인할 수 있다. 앞서 살펴보았듯이 체계적 복잡성은 '너비', 즉 다양성과 관련이 있는 것으로 학습자가 사용하는 문법의 범위, 크기, 변이 등을 가리키고 구조적 복잡성은 '깊이', 즉 정교함과 관련이 있는 것으로 문법 구조가 내포하고 있는 형태나 구성을 가리킨다.

이어서 Bulté & Housen(2012)에서는 문장, 절, 구에 나타나는 문법의 다양성, 즉 T-unit당 단어 수, 절당 단어 수 등과 같은 길이로 측정되거나 하위절의 비율로 측정될 수 있는 것들을 통사적 복잡성으로, 굴절이나 파생과 같이 구문핵당 종속 요인, 동사당 동사 굴절 형태, 단어당 파생 접사 등으로 측정되는 것들을 형태적 복잡성으로 구분하면서 통사적 복잡성을 '단어나 구, 절과 같은 언어 구성 성분을 다양하고 복잡한 문법으로 연결하여 자신이 표현하고자 하는 의미를 나타낼 수 있는 능력(Bulté & Housen, 2012:35)'으로 정의하였다. 그러므로 학습자 쓰기와 말하기 자료에서의 통사 복잡성은 텍스트에 나타난 문법 복잡성 중 문장이나, 절, 구 단위에서 언어 구성 요소들 간의 접속 관계, 상관관계를 나타내는 문법 요소들의 다양한 정도와 범위라고 정의할 수 있다.

통사 복잡성에 대한 이 같은 정의는 기존의 한국어 교육 분야에서도 동일하게 사용되어 왔는데 한국어교육 분야의 여러 연구들을 살펴보면 '문장 구성 요소들 사이의 관계의 복잡 정도(음승민, 2012)', '통사적으로 정교하고 다양한 구조(남주연, 2015)', '통사적으로 정교하고 다양한 문법 형태(주우동, 2019)', '구, 절 문장 단위를 다양하게 사

용하여 단위의 범위나 크기를 확장할 수 있는 능력(서신웨, 2022)' 등으로 정의하고 있음을 확인할 수 있다. 이러한 논의들을 참고하여 본고에서는 통사 복잡성에 대해 '다양한 문법 요소들을 사용하여 길고 복잡한 문장을 산출할 수 있는 능력'으로 정의하도록 한다.

2.1.2. 통사 복잡성의 측정

통사적 복잡성을 측정하는 방법은 연구의 목적에 따라 다양하지만 일반적으로 통사적 구조의 범위, 단위의 길이, 특정 통사적 구조의 구조적 복잡성 정도나 정교도(sophistication), 대등(coordination)이나 내포(subordination)[12] 관계의 양과 유형을 측정하는 방법을 사용한다.[13] 이러한 방법들은 크게 특정(specific) 측정 방법과 포괄적(general) 측정 방법으로 구분할 수 있는데(Skehan, 2003; Norris & Ortega, 2009) 특정 측정은 수동형이나 조사, 관계절, 시제, 지시어, 접속부사, 단수와 복수, 형용사 사용 등과 같은 특정 언어 구성 요소나 문법 항목에 초점을 두고 이에 대한 사용 정도를 측정하는 방법이고(Kuiken & Vedder, 2012:145) 포괄적 측정 방법은 분석 단위를 설정한 뒤, 단위의 길이를 측정하거나 단위당 대등절이나 내포절의 양, 즉 접속관계의 비율이나 빈도를 측정하는 방법이다.

일반적으로 통사 복잡성 연구에서는 특정 측정 방법보다는 일반 측정 방법을 많이 사용하는데 특정 측정 방법은 특정 항목이나 구조를 자세히 살펴볼 수 있다는 장점이 있지만 연구자가 확인하고자 하는 문법 구조나 항목을 학습자가 사용하게끔 유도해야 한다는 점에서 자연스러운 언어 사용을 방해할 수 있고, 자유로운 언어 산출에서는 특정 구조나 항목이 나타나지 않을 수도 있으며 전체적인 언어 발달을 살피기 어렵다는 단점이 있다. 반면 일반 측정 방법은 학습자의 언어 산출물의 통사 복잡성 수준을 수치화하여 보여줌으로써 학습자의 언어 발달을 전체적으로 조망할 수 있다는 장점이 있다. 따라서 학습자의 숙달도를 구분하거나 특정 변수의 차이에 따른 언어 산출물의 변화를 살피거나 언어 발달을 관찰하는 연구에서는 일반 측정 방법을 사용하는 것이 대부분이며, 특정 측정 방법은 일반 측정 방법의 보완적인 수단으로 사용되는 경우가 많다

---

[12] 'subordination, subordinate clause'는 '종속, 종속절'로 번역하는 경우도 있지만 본고에서는 포유문에서 안긴 절을 모두 '내포절'로 칭하고 있으므로 '내포, 내포절'로 번역하도록 한다.

[13] 통사적 복잡성 측정은 일반적으로 통사적 구조의 범위, 단위의 길이, 특정 통사적 구조의 구조적 복잡성 정도나 정교도(sophistication), 대등(coordination), 내포(subordination)의 양과 유형 중 하나 이상을 수치화하는 것을 목표로 한다(Bulté & Housen, 2012:35).

(Vercellotti, 2012:15).

앞에서 언급하였듯이 일반 측정 방법은 크게 단위 길이를 측정하는 방법과 단위당 접속 관계의 비율이나 빈도, 즉 단위 확장 정도를 측정하는 방법으로 나뉘며 한국어교육 분야의 통사 복잡성 연구들(서세정, 2009; 음승민, 2012; 남주연, 2015; 이복자, 2015; 서신웨, 2022)에서도 이러한 구분을 따라 왔다. 이에 본고 역시 학습자들의 언어 자료에 나타난 길이와 접속 관계를 통해 학습자들의 통사 복잡성을 측정하고자 한다.

1) 길이에 따른 통사 복잡성 측정

산출 길이에 따른 복잡성은 MLU, 즉 평균 발화 길이라고도 불리는데 이는 '각 발화에서 쓰인 어휘들의 평균 비율을 형태소, 단어, 어절, 절 등의 단위에서 살펴보는 것'이다(김명광, 2016:454). 길이에 기반한 통사 복잡성 측정은 제1언어 발달 연구에서 먼저 사용되었으나 현재는 제1언어 발달뿐만 아니라 제2언어 습득 분야에서도 학습자의 언어 발달을 측정하는 대표적인 방법이다. 길이에 따른 측정 방법은 통상적으로 숙달도가 높은 학습자가 숙달도가 낮은 학습자에 비해 더 긴 문장을 산출할 것이라는 통념에 기반한 것으로 통사 복잡성을 측정함에 있어 가장 기본적이고 가장 많이 사용되는 방법이다.

단위 길이를 통한 복잡성 측정에서는 먼저 기본 분석 단위를 설정한 뒤 해당 단위에 포함된 형태소, 단어, 어절 수를 측정한다. 이때 분석 단위는 분석 대상 자료가 문어 자료인지 구어 자료인지에 달라지게 되는데 일반적으로 문어 자료 분석에서는 T-unit이나 문장을 단위로 사용하고 구어 자료 분석에서는 AS-unit이나 C-unit을 사용한다.

전통적으로 문어 복잡성 연구에서 가장 많이 그리고 오래 사용되어 온 단위는 Hunt(1965)에서 제안한 T-unit이다. T-unit은 최소 종결 가능 단위(minimal terminable unit)의 약자로[14] 하나의 주절(독립절)과 그 주절에 붙어 있거나 안겨 있는 내포절(subordinate clause) 혹은 구(phrase)를 포함하는 단위로 정의된다(Hunt, 1970:4). T-unit은 그동안 영어 모어화자나 ESL, EFL 쓰기 텍스트의 통사적 발달을 분석하는 기본 단위로 사용되어 왔다. 그러나 한국어 교육에서 통사적 복잡성 분석을 제

[14] '최소 종결 가능 단위(minimal terminable unit)'라는 명칭에 대해 Hunt(1970:4)에서는 일반적으로 한 문장에 반드시 하나의 주절이 있어야 하지만 주절에는 하나 이상의 내포절이나 다양한 종류의 구가 첨부되거나 포함될 수 있기 때문에 문장을 T-unit으로 자르는 것은 문법적으로 문장 부호를 붙일 수 있는 가장 짧은 단위로 자르는 것이므로 종결 가능한 최소 단위라고 설명하고 있다.

안한 박지순·서세정(2009)에서는 영어는 문장이 여러 개의 주술 관계를 갖춘 최소의 문장 단위로 나누어질 수 있기 때문에 문장보다는 작고 절보다는 큰 T-unit과 같은 분석 단위가 필요하지만 한국어는 첨가어적인 특성이 있어 T-unit이 문장과 일치하는 경우가 많다고 지적하면서 실용적인 관점에서 문장 단위를 사용하는 것이 바람직하다고 제안하였다.

이 뒤로도 여러 연구(음승민, 2012; 이복자, 2015; 남주연 외, 2016; 주월랑, 2018; 주우동, 2019; 최수완, 2019)들에서도 문장을 기본 단위로 사용하며 그 타당성을 입증하였다. 남주연 외(2016)에서는 문장당 단어 수와 문장당 어절 수에 의한 분석은 총점뿐만 아니라 어휘·문법, 쓰기, 듣기, 읽기 모든 영역과 상관성을 보인 반면 T-unit당 단어, T-unit당 어절은 유의미하지 않게 나타났다는 점과 T-unit당 내포절 수보다 문장당 내포절 수가 더 유의미하게 등급 간 차이를 보여 준다는 점을 통해 T-unit보다 문장 단위로 분석하는 것이 더 타당함을 확인하였다. 그리고 김영주 외(2013:216)에서 역시 한국어는 T-unit 분석이 쉽지 않고 낮은 숙달도 학습자의 작문 분석에서 어려움이 있으므로 구두점을 분석 단위로 삼는 것이 용이하다고 언급하였다. 이러한 점들을 근거로 본고 역시 문장을 문어 자료의 기본 분석 단위로 삼고자 한다.

한편, 구어는 문어와 달리 실시간으로 언어 산출이 이루어져 반복, 자기 수정, 휴지가 빈번히 나타나고 완결된 문장보다는 구 단위의 발화가 많이 나타나 문어와 같은 분석 단위를 사용하기가 어렵다. 이에 구어 자료 분석에서는 C-unit과 AS-unit이 자주 사용된다. C-unit은 의사소통 단위를 가리키는 것으로 화용적 의미를 가진 단어나 구, 문장, 문법적·비문법적 발화 단위를 포함한다(Pica 외, 1989). C-unit은 주어나 서술어의 생략, 혹은 기타 성분이 누락된 불완전한 문장이나 하나의 어휘로 표현되는 감탄사까지도 분석 대상에 포함시킨다는 점에서 T-unit이나 문장에 비해 구어 자료의 특성을 반영할 수 있다는 장점이 있다. 그러나 C-unit은 의사소통 단위라는 개념이 모호한 문제, 또 여전히 구어의 특징인 억양이나 휴지 등을 충분히 반영하지 못한다는 문제점들이 지적되었다. 이에 최근에 자주 사용되는 구어 분석 단위로는 AS-unit이 있다. AS-unit은 Foster 외(2000)에서 제시한 것으로 기본적으로 통사 단위이지만 억양과 휴지를 포함하는 구어 분석 단위이다.

AS-unit은 '내포절(subordinate clause)과 함께 나타나는 독립절 또는 절 이하의 단위(sub-clausal unit)로 구성된 단일 화자의 발화(Foster 외, 2000:365)'로 정의되며 T-unit과 달리 그 자체로 완전한 문장을 이루지 못하더라도 의미상으로 절의 복원이 가능하다면 구 단위의 대답도 하나의 단위로 분석한다. 또한 같은 주어로 두 개의 절이 연

결되어 있더라도 선행절과 후행절 사이에 상승이나 하강의 억양이 발생하여 절이 단절되거나 혹은 후행절과의 사이에 기준 이상의 휴지가 발생한다면 새로운 독립된 절의 시작으로 처리하는 등 구어의 특징을 충분히 고려하여 발화 단위를 구분할 수 있다는 장점이 있다.[15] 이에 그동안 AS-unit은 이복자(2015), 남주연(2015), 남주연·김영주(2015), 김정운 외(2017a, 2017b), 주월랑(2020)등 한국어교육 분야의 선행연구에서 자주 사용되며 그 타당성을 입증해 왔으며 남주연(2015:209)에서는 AS-unit당 내포절 수가 C-unit당 내포절 수보다 숙달도(총점)와 더 높은 상관성을 보인다는 점을 확인하였다.[16] 이러한 점을 고려하여 본고 역시 구어 자료 분석의 단위로 AS-unit을 사용하고자 한다. 그러므로 본고에서는 문어에서는 문장의 길이를 측정하고 구어에서는 AS-unit의 길이를 측정하는 것으로 길이에 따른 통사 복잡성의 정도를 측정하도록 한다.

2) 단위 확장에 따른 통사 복잡성 측정

단위 길이에 따른 통사 복잡성 측정은 그 유효성에 대해 이견이 거의 없지만 단순히 단어를 나열하는 것만으로도 단위 의 길이는 길어질 수 있기 때문에 단위 길이 외에 구성 요소 간의 접속 관계를 통해 문장의 구조적 복잡성을 살필 필요가 제기되었다. 접속 관계에 의한 측정에서는 더 복잡하고 많은 접속 관계를 통해 문장을 확장할수록 통사 복잡성이 높다고 판단하게 된다. 그러나 단위 길이의 측정에 비해 단위 확장의 측정에서는 측정 지표가 다양하고 그 유효성에 대해 합의가 이루어지지 않고 있다. 단위 확장의 측정에서 보편적으로 사용되는 방법은 대등(coordination)과 내포(subordination) 관계의 양을 측정하는 방법이다. 이는 기본이 되는 분석 단위를 설정한 뒤 단위 안에서 나타나는 절의 수를 측정하거나 절과 절 사이의 접속 관계를 유형화하여 빈도로 나타내는 방법이다. 자주 사용되는 방법으로는 복문의 비율, 단위당 절 수, 단위당 대등절이나 내포절 수, 혹은 대등접속지수(coordination index)등이 있다.

한국어교육 분야의 선행연구들을 살펴보면 작문을 단위로 작문당 절 수, 복문 수를 측정하는 경우도 있고 문장을 단위로 문장당 절 수, 대등절 수, 내포절 수를 측정하기도 하며 절을 단위로 절당 대등절 수나 내포절 수를 측정하기도 한다. 이 중 문장을 기본 단위로 하여 문장당 절 수와 문장당 대등절 수, 내포절 수나 연결어미 수를 지표로

[15] AS-unit은 독립절을 기본으로 하는 T-unit 분석에 의미 단위의 측정인 C-unit을 적용하고 거기에 억양과 휴지라는 구어의 특성을 추가로 반영한 것이다(남주연, 2015:47).

[16] 남주연(2015)에서는 AS-unit을 단위로 분석한 자신의 연구 결과를 C-unit을 단위로 분석한 김영주 외(2013)의 결과와 비교하여 이와 같은 결론을 도출하였다.

사용한 연구가 가장 많았는데 그중에서도 문장당 절 수가 숙달도별 차이를 잘 보여 준다는 연구들이 가장 많았다. 서세정(2009:44)에서는 문장당 절 수가 초, 중, 고급 및 대학생 간의 차이를 더 잘 보여 준다는 것을 통하여 문장당 절 수가 한국어 통사적 숙달도의 구조적 복합도 진단 도구로서 그 효용성이 크다고 하였고 김현아(2016)에서는 문장당 절 수가 3~5급 사이의 숙달도별 차이를 유의미하게 나타낸다고 하였으며 서신웨(2022)에서 역시 문장당 절 수가 중급과 고급 사이의 차이를 유의미하게 구분한다고 하다고 보고하였다. 또한 학습자의 구어 자료를 분석한 남주연(2015)에서도 AS-unit 당 절 수가 전체 숙달도 사이에서 유의미한 증가를 보인다는 점에서 학습자의 숙달도를 예측할 수 있는 효과적인 통사 측정 방법임을 증명하였다. 이러한 결과들을 참고하여 본고에서는 문어와 구어 모두 단위당 절 수로 단위 확장의 양을 측정하도록 한다.

3) 단위 확장의 유형에 따른 통사 복잡성 측정

단위 길이와 단위 확장의 양을 측정하는 것으로도 학습자 언어 산출물의 통사적 복잡성 정도를 수치화하여 측정할 수 있지만 이러한 문장 확장이 어떤 문법 항목에 의해 이루어졌는지를 추가적으로 살피는 것 또한 학습자 언어 발달의 중요한 정보를 제공해 준다. 즉 학습자가 어떤 종류의 문법 항목을 얼마나 사용했는지를 측정하여 학습자가 사용한 문법의 다양성 정도나 난이도를 살펴봄으로써 학습자의 통사 복잡성을 좀 더 면밀히 관찰하는 것이다.

한국어의 문장은 단문과 복문으로 나뉘고 복문은 다시 접속문과 포유문으로 나뉜다. 이때 접속문은 대등절에 의해 구성되며 포유문은 명사절, 관형사절, 부사절, 인용절과 같은 다양한 내포절에 의해 구성된다. 그리고 각각의 절들은 다양한 연결어미와 전성어미를 통해 실현된다. 따라서 단위 확장의 유형은 학습자들이 구체적으로 어떤 어미들을 사용하여 문장을 확장하는지 살피는 것이다. 따라서 본고에서는 학습자의 언어 자료에 나타난 대등절, 명사절, 관형사절, 부사절과 같은 세부적 항목의 빈도를 측정함으로써 학습자의 통사 복잡성을 더 면밀히 고찰하고자 한다. 통사 복잡성의 각 측정 지표에 대한 구체적인 분석 기준은 3장에서 자세히 기술하도록 한다.

2.2. 어휘 복잡성의 개념 및 측정

2.2.1. 어휘 복잡성의 개념

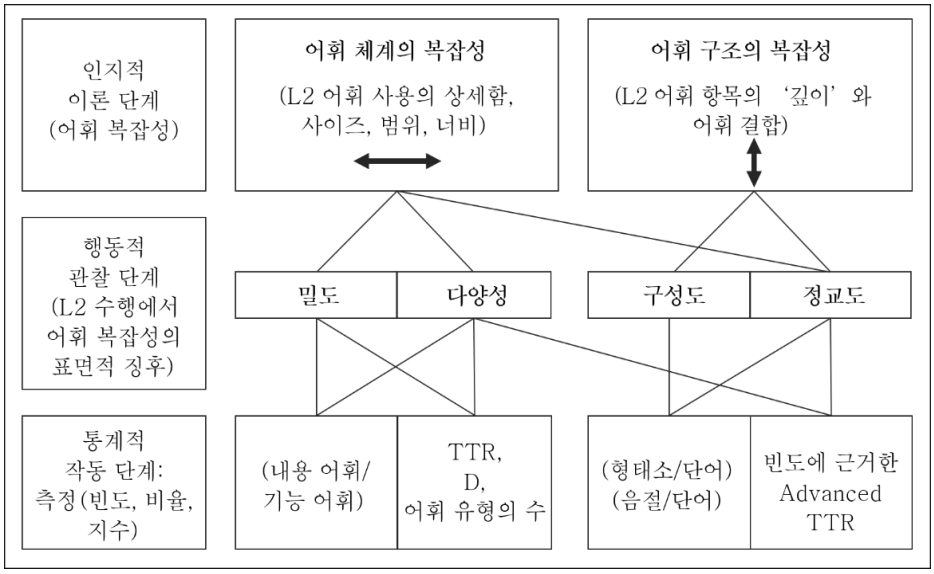
어휘는 개념 및 사고 표현의 도구로서 의사소통에서 가장 중요한 요소이며(임지룡, 1998:3) 문법과 함께 제2언어 학습자의 언어 능력을 나타내는 매우 중요한 요소로 간주되어 왔다. Laufer(1994:21)에서는 어휘의 질과 쓰기의 질은 서로 연결되어 있기 때문에 어휘의 발달을 통해 쓰기의 발달을 측정할 수 있다고 하였고Laufer & Nation(1995)에서는 모어화자와 제2언어 학습자의 명백한 차이 중 하나는 쓰기에 사용된 어휘에 있다고 하기도 하였다.

제2언어 습득에서 어휘 연구는 문법에 비해 상대적으로 늦게 시작되었는데 본격적인 어휘 연구는 1980년대 중반에 학습자 말뭉치를 구축하고 분석하면서 시작되었다. 연구자들은 학습자의 어휘 능력을 측정하기 위한 여러 척도들을 연구하기 시작하였고 이 과정에 '어휘 복잡성', '어휘 풍요도'와 같은 용어를 사용하여 학습자의 어휘 사용 능력을 정의하기 시작하였다. 어휘 복잡성은 학습자가 얼마나 다양한 어휘를 사용하는지, 그리고 얼마나 정교하게 내용을 전달하는지로 설명된다(Bulté & Housen, 2012). 즉 통사 복잡성이 문법 사용의 폭과 깊이에 관련된 것이라면 어휘 복잡성은 어휘 사용의 폭과 깊이에 관련된 척도로 학습자가 어휘를 얼마나 풍부하고 정교하게 사용하는지를 평가하기 위한 것이다. 그러므로 어휘 복잡성 측정에서는 학습자가 사용한 어휘에 대한 양적인 측정뿐만 아니라 질적인 측정 또한 간과할 수 없다.

그동안 선행연구들을 살펴보면 주로 어휘 다양도(diversity), 어휘 밀도(density), 어휘 정교도(sophistication) 3가지 지표를 활용하여 학습자의 어휘 사용 능력, 즉 어휘 복잡성을 평가해 왔다. 이러한 측정 방법의 근거는 Laufer(1991), Read(2000)와 Bulté & Housen(2012) 등에서 찾아볼 수 있는데 초반의 연구에서는 어휘의 풍요도(lexical richness)라는 개념으로 학습자 어휘 사용의 양과 질을 나타낸 것을 확인할 수 있다. 먼저 Laufer(1991)에서는 어휘 사용의 풍요도에 대해 어휘 독창성(lexical originality), 어휘의 밀도(lexical density), 어휘도(lexical sophistication), 어휘의 변이(lexical variation)로 정의하였다. 반면, Read(2000)에서는 제2언어 학습자의 좋은 글에 나타나는 어휘적 특징들을 어휘 풍요도라고 정의하면서 제2언어 학습자의 글쓰기에는 다양한 종류의 형태, 의미 오류가 발생하기 때문에 어휘 오류 수(numbers of errors) 역시 좋은 글을 가늠하는 중요한 척도이며 이를 어휘 풍요도의 측정 지표에 포함할 것을 주장하였다. 그리하여 Read(2000:200-201)에서는 어휘 풍요도를 어휘 다양도(Lexical diversity), 어휘 밀도(Lexical density), 어휘 정교도(Lexical sophistication), 오류 수(number of errors)로 정의하였다. 그리고 Bulté & Housen(2012)에서는 어휘의 복잡성을 어휘 체계의 복잡성과 어휘 구조의 복잡성으로 나누고 그 하위 요소로 밀도

(density), 다양도(diversity), 구성도(compositionality)와 정교도(sophistication)를 제시하였다. 여기에서 어휘 구성도는 어휘의 구조적 측면과 관련된 것으로 형식과 의미 구성 요소, 즉 단어당 형태소, 단어당 음절수로 측정되는 것이다. 다음의 <그림 3>은 Bulté & Housen(2012:28)에서 제시한 어휘 복잡성의 단계이다.

<그림 3> 어휘 복잡성의 단계(Bulté & Housen, 2012:28)



정리하면 어휘 풍요도나 어휘 복잡성은 학습자의 어휘 사용 능력을 나타내는 개념이고 어휘 다양도, 어휘 밀도, 어휘 정교도 어휘 독창성은 이 능력을 측정하는 구체적 요소인 것이다. 다음 장에서는 어휘 복잡성 나타내는 여러 측정 요소들을 살펴보고 본 연구에 적합한 측정 요소를 선정하고자 한다.

2.2.2. 어휘 복잡성의 측정

1) 어휘 다양도(lexical diversity)

텍스트나 말뭉치(corpora)에서 단어(유형)가 의미 전달을 위해 사용되는 것을 살펴보면 일부 단어(특히 문법적 단어)는 종종 반복되며, 다른 단어들은 단지 몇 번만 사용되는 것을 알 수 있다. 전체적으로 텍스트 또는 말뭉치가 광범위한 어휘를 사용하는지 또는 재활용되는 제한된 범위의 어휘 항목만 사용하는지 측정하기 위해 어휘 다양도

통계를 계산할 수 있다(Brezina, 2018:57).

어휘 다양도는 학습자 어휘 사용에 나타난 다양성 정도를 측정하는 척도로 어휘 다양도가 높다는 것은 한편의 글이나 말에서 나타난 어휘의 종류는 다양하지만 같은 단어의 중복은 적음을 의미한다. 그러므로 높은 어휘 다양도는 어휘의 크기와 그것을 효과적으로 사용할 수 있는 능력의 조합을 나타낸다(Malvern 외, 2004:6). 어휘 다양도는 처음에 아동의 제1언어 발달을 측정하기 위해 고안된 지표였으나 이후 제1언어뿐만 아니라 제2언어 습득 분야에서도 문어와 구어 발달의 범용적 척도로 사용되어 오며 그 유효성을 입증하고 있다.

Jarvis(2002)에서는 핀란드와 스웨덴 청소년 영어 학습자 210명과 모어화자 청소년 66명의 쓰기 텍스트를 분석한 결과 어휘 다양도는 교육의 양과 긍정적인 상관관계가 있음을 발견하였고 Bulté 외(2008:279)에서는 어휘 다양도는 제2언어 학습자의 어휘 지식의 크기, 학습자의 어휘 숙달도를 보여 준다고 하였다. Yu(2010)에서는 어휘 다양도가 쓰기와 말하기 점수를 예측하는 중요한 변수임을 증명하였다. 한국어 교육 영역에서도 진대연(2006)과 배도용(2012)에서는 언어권별 어휘 다양도의 특징을 구분하였고 안경화(2003:177)에서는 같은 분량의 작문을 평가할 때 어휘면의 평가 항목에서는 적절하고 다양한 어휘를 쓴 글이 더 잘 쓴 글이라고 할 수 있다고 하면서 다양한 어휘를 사용하는 능력은 숙달도와 어느 정도 상관관계가 있을 것이라고 하였다. 어휘 다양도를 쓰기의 한 평가 항목으로 사용한 박정은·김영주(2014:5)에서는 다양한 어휘를 사용하는 것은 좋은 글의 충분한 조건은 될 수 없지만 좋은 점수를 받기 위한 필요 조건임을 주장하고 있다. 이처럼 다양한 어휘를 사용하는 능력은 학습자의 어휘 능력뿐만 아니라 언어 숙달도를 나타내는 중요한 척도가 된다. 또한 어휘 다양도는 문어와 구어를 구분해주기도 하는데 안의정(2017)에서는 한국어 문어와 구어 텍스트들을 분석한 결과 문어 텍스트의 어휘 다양도가 구어 텍스트에 비해 높음을 확인하였다.

한편, 어휘 다양도는 학습자가 산출한 전체 어휘 중에서 중복되지 않은 새로운 어휘가 얼마나 많은 지로 측정되며 어휘 유형(type)와 어휘 구현(token)이라는 개념을 기본으로 한다. 이와 같이 학습자 어휘 유형(type)의 빈도 분포에 기반한 어휘 다양도 측정 방법은 20세기부터 여러 학자들에 의해 제기되었는데 대표적으로 Thomson & Thompson(1915)에서 하나의 샘플 텍스트를 통해 이 학습자의 어휘량을 예측할 수 있는지를 연구하면서 시작되었다(Malvern 외, 2004:31). 이후 서로 다른 단어의 수를 세는 방법인 NDW(number of different words), 고정된 길이의 언어 샘플에서 발생하는 총 단어의 수를 측정하는 TNW(total number of words), 어휘의 유형과 구현의 비율을

측정하는 TTR(type-token ratio), 어휘 구현의 수를 표준화하여 표준화된 어휘 구현의 수에서의 어휘 유형을 측정하는 MSTTR(mean segmental type-token ratio), 이동창을 설정하여 각 구간의 평균 TTR을 구하는 MATTR(moving average type-token ratio), G값 지표(guiraud index), 그리고 D값 등이 제시되었다.

정리하면, 어휘 다양도가 학습자의 언어 능력을 반영할 수 있음에는 모두가 동의하지만 효과적인 어휘 다양도 지표가 무엇인지에 대해서는 의견이 분분하다. 그러므로 어휘 다양도 측정에서 가장 중요한 것은 길이의 영향을 덜 받는 측정지표를 선정하는 일이 될 것이다. 본고에서는 여러 어휘 다양도 측정 지표 중에서 비교적 영향력이 있는 지표로 인정받고 있으며 한국어교육의 여러 선행연구에서 자주 사용되어 온 CTTR을 사용하고자 한다. 이 외에도 어휘 다양도 측정에서 고려해야 할 문제는 단어의 정의에 대한 것인데 같은 단어족에서 파생된 파생어, 혹은 같은 단어의 다른 활용형을 같은 어휘로 간주해야 하는지 다른 어휘로 간주해야 하는지와 같은 것이다. 이와 관련된 구체적인 데이터 가공 방법은 3장에서 제시하도록 한다.

어휘 다양도가 학습자 어휘 능력과 어휘 숙달도를 나타내는 척도임은 분명하다. 그러나 어휘 다양도는 학습자가 어떤 유형의 단어를 사용했는지와 같은 단어의 질 보다는 자신이 알고 있는 어휘를 얼마나 잘 표현할 수 있는지를 나타내는 지표(Laufer & Nation, 1995:310)이기 때문에 다양성 정도 외에 실제로 사용한 어휘의 수준을 확인할 수 없다는 단점이 있다. 만약 두 학습자가 같은 어휘 다양도 값을 가진다고 할 때 사용한 어휘의 수준이 어떤 수준에 있는지 까지는 확인할 수 없기 때문이다. 따라서 어휘의 질을 측정하기 위한 추가적인 지표가 필요하게 되는데 아래에서 살펴볼 어휘 밀도와 어휘 정교도가 대표적으로 사용되는 어휘 질의 측정 지표이다.

2) 어휘 밀도(lexical density)

어휘 밀도는 텍스트에 나타난 내용어의 비율을 가리키는 것으로 Ure(1971)에서 처음 제기하고 측정하였다. 대부분의 언어에서 어휘는 ‘문법어(기능어)’와 ‘내용어’로 구분하는 것이 가능한데(Nunan, 1999:101-102) 내용어는 명사, 동사, 형용사, 부사 등 실질적 의미를 전달하는 어휘를 가리키고, 기능어는 실질적 의미보다는 문법적인 기능을 가지고 있는 어휘를 가리킨다. 그러므로 텍스트에서 어휘 밀도가 높다는 것은 내용어의 비율이 높다는 것을 의미하며 내용어의 비율이 높다는 것은 더 많은 양의 정보를 전달한다는 것을 의미한다(Johansson, 2008; Ryoo, 2018:64). 따라서 어휘 밀도는 가독성과도 관련되며, 어휘 밀도가 높은 텍스트일수록 읽거나 쓰기가 어려우며 더 높은

문식성을 가지고 있다고 본다(Read, 2000:196). 어휘 밀도는 내용어 구현 수와 총 어휘 구현수의 비율, 혹은 내용어 구현 수와 기능어 구현 수의 비율로 측정되는데 그 공식은 다음과 같다(Bulté & Housen, 2012:31).

$$LD= \frac{\textit{내용어 구현 수 X 100}}{\textit{총 어휘 구현 수}}$$

또한 어휘 밀도는 문어와 구어를 구별하는 한 가지 특징이기도 하다. 이는 문어에서 구어보다 정보나 생각 더 집중된 방식으로 제시되기 때문이다(Read, 2000:200). Ure(1971)에서는 문자 텍스트에서는 내용어의 비율이 40%이상인 반면 구어 텍스트에서는 40% 미만임을 발견하였고 이 외에도 행위 동반 언어[17]와 보고 텍스트, 독백과 대화 사이에도 어휘 밀도에 차이가 있음을 발견하였다. 한국어에서도 비슷한 연구 결과를 찾아볼 수 있는데 여러 사용역의 한국어 텍스트를 분석한 안의정(2017)에 따르면 강연을 제외한 모든 구어 텍스트의 밀도가 문어 텍스트보다 낮게 나타났음을 보고하였다. 그러나 Ure(1971)의 연구 결과와 다른 점은 구매 대화와 수업 대화를 제외하면, 구어 텍스트의 밀도도 40% 이상을 유지한다는 점이었다.

어휘 밀도가 숙달도와 관련이 있다고 보는 연구들도 있다. Waller(1993)에서는 원어민과 제2언어 학습자의 어휘 복잡성 비교할 때 어휘 밀도가 50% 이상인 텍스트는 실제로 원어민이 작성했거나 원어민 독자가 그렇게 인식한다는 결과를 얻기도 하였다(Read, 2000:201 재인용). 또 남주연(2015)에 따르면 총 어휘당 내용어 유형 수는 전체 숙달도 안에서 유의미한 차이가 있었고 총 어휘당 내용어 구현 수는 가장 높은 등급과 모어화자 간에서 유의미한 차이를 보여줄 수 있음을 확인하였다. 이상의 연구 결과들을 통해, 어휘 밀도는 문어와 구어, 모어화자와 제2언어 학습자, 제2언어 숙달도를 구별할 수 있는 척도임을 알 수 있다.

어휘 밀도를 측정하기 위해서는 먼저 내용어와 기능어에 대한 분류가 선행되어야 한다. 한국어에서는 명사, 동사, 형용사, 부사 등이 내용어가 되겠고 조사나 어미가 기능어가 될 것이다. 따라서 어휘 밀도는 총 어휘 중에서 내용어가 차지하는 비율로 측정하도록 한다.

[17] 행위 동반 언어는 요리 대화와 같이 행동에 수반되는 언어(language in action)를 가리킨다.

3) 어휘 정교도 (lexical sophistication)

어휘 다양도와 어휘 밀도도 학습자의 어휘 수준을 가늠하는 유효한 척도이지만 두 척도 모두 실제로 학습자가 어떤 어휘를 사용하는지를 알 수 없다는 단점이 있다. 어휘 정교도는 양보다는 텍스트에 나타난 어휘의 질을 가늠하는 척도이다. 어휘 정교도는 어휘 세련도라고 부르기도 하는데 학습자가 흔하지 않고 고급 수준에 해당되는 어휘를 적절한 문맥에서 사용하는 능력을 나타내는 척도이다. 어휘 정교도에 대한 학자들의 정의를 살펴보면 Laufer & Nation(1995:309)에서는 어휘 정교도에 대해 '텍스트에 나타난 '고급(advanced) 어휘의 비율'로 정의하고 있고 Read(2000:203)에서는 이를 '학습자의 언어 텍스트에서 나타난 특이하거나 고급스러운 단어의 비율을 나타내는 척도'라고 정의하면서 '희소성(rareness)'이라고 부르기도 하였다. 여기에서 고급 어휘는 일반적인 어휘가 아닌 글의 주제와 스타일에 적합한 저빈도 어휘를 가리키며 전문 용어 혹은 학습자가 의미를 정교하게 표현할 수 있도록 하는 비일반적인 단어가 포함된다(Read, 2000:200). 그리고 구체적인 어휘 목록은 연구자가 직접 정의할 수 있다(Laufer, 1994:22; Laufer & Nation, 1995:309).

정리하면 어휘 정교도는 학습자 텍스트에 나타난 어휘 중 일상적으로 사용 빈도가 상대적으로 낮으며 학습자가 표현하고자 하는 의미를 정확하고 정교하게 표현할 수 있는 고급 어휘의 비율이라고 정의할 수 있다. 그러므로 어휘 정교도가 높다는 것은 학습자가 저빈도 어휘, 즉 일상생활에서 많이 사용되지 않는 어려운 어휘를 많이 알고 있으며 이를 주제와 상황에 맞게 정교하게 사용할 수 있음을 의미한다. 이는 학습자의 언어 숙달도가 높아질수록 더 전문적이고 추상적인 단어를 사용하여 좀 더 복잡한 언어로 자신의 생각을 표현할 수 있을 것이라는 통념에서 비롯된 것이다. 실제로 어휘 정교도는 학습자의 언어 숙달도를 구분할 수 있는데 Arnaud(1984)와 Linnarud (1986)평균적으로 제2외국어 학습자가 모국어 학습자보다 덜 다양하고 정교한 어휘를 사용한다는 것을 발견하였고 Laufer(1991, 1994)[18]에서는 어휘 발달에서 가장 유의미하고 명확한 발달을 보인 수치가 어휘 정교도, 즉 저빈도 어휘의 발달이라고 하였다. 어휘 정교도는 고급 어휘 구현 수와 총 어휘 구현수의 비율로 측정되며 다음과 같은 공식으로 나타낸다.

---

[18] Laufer(1991, 1994)에서는 이스라엘 영어 학습자들을 대상으로 어휘 발달을 측정하였으며 저빈도 어휘 목록으로 사용한 것은 '대학 어휘 목록'이었다. 이는 연구 대상자들이 고등학교를 졸업하고 대학교 1학년 학생들이기 때문이다.

$$LS = \frac{\text{고급 어휘 구현 수} \times 100}{\text{총 어휘 구현 수}}$$

앞에서 언급하였듯이 '고급 어휘'의 목록은 학습자의 특성, 교육과정에 따라 연구자가 정의하게 된다. 일반적으로 학생에게 고급 어휘는 마지막 두 학년의 어휘 요강(lexical syllabus)에 있는 어휘를 가리키지만 대학생에게는 꼭 그러하지 않으며 다른 교육 목표를 가진 다른 나라일 경우 같지 않을 수 있다(Laufer & Nation, 1995:310). 어휘 정교도는 어휘의 등급 목록에 따라 달라지기 때문에 어휘 정교도 측정에 있어 가장 핵심적인 부분은 고급 어휘의 목록이 될 것이다. 따라서 학습자의 교육 과정과 수준에 맞는 '고급 어휘'를 선정하는 것이 가장 중요하다. 본 연구의 학습자는 중도입국 청소년으로 모두 <한국어(KSL) 교육과정>을 바탕으로 제작된 '의사소통 한국어' 교재를 사용하여 한국어를 학습하였다. 그러므로 '고급' 어휘 선정에 있어 <한국어(KSL) 교육과정>을 참고할 필요가 있다. 따라서 본고에서는 <한국어(KSL) 교육과정>의 초, 중급 어휘 목록을 바탕으로 어휘 등급을 나누어 어휘 정교도를 측정하도록 한다.

정리하면, 본고에서는 자신이 아는 어휘를 잘 사용하는 능력인 어휘 다양도, 고급스러운 어휘를 사용할 수 있는 능력인 어휘 정교도, 풍부한 내용어를 사용할 수 있는 능력인 어휘 밀도의 세 측면으로 나누어서 어휘 복잡성을 측정하도록 한다.

3. 정확성의 개념 및 측정

3.1. 정확성의 개념

정확성은 복잡성, 정확성, 유창성 구인 중 가장 오래되고 내적으로 일관된 구인으로(Housen & Kuiken, 2009:463) 오래 전부터 제2언어 학습자의 언어 발달 수준을 기술하기 위한 주요 구인으로 사용되어 왔다. 연구자들은 학습자가 정확하게 사용한 것을 곧 습득한 것으로 보고 정확성을 통해 숙달도를 측정하기도 하고 언어 습득의 순서를 연구하기도 하였다.

정확성에 대한 정의를 살펴보면, Skehan & Foster(1999:96)에서는 '학습자가 언어를 산출할 때 오류(error)를 피하는 능력으로 이는 언어에 대한 더 높은 수준의 통제 혹은 오류를 유발할 수 있는 도전적인 구조를 회피하는 보수적인 지향성'이라고 하였고 Pallotti(2009:592)에서는 '특정 규범에 대한 접합성의 정도'라고 설명하였다. 그리고

Housen 외(2012:4)에서는 정확성을 '일반적으로 제2언어 수행이 규범으로부터 벗어나는 정도'라고 하였다. 즉 정확성은 '규범에 맞는 언어를 산출할 수 있는 능력'이며 규범에서 벗어난 부분은 '오류'가 된다. 그러므로 학습자 언어에 나타난 정확성을 측정하기 위해서는 '규범'과 '오류'에 대한 정의가 필요하다.

1) 규범의 정의

제2언어 학습에서 '규범'이 어떻게 정의되어야 하는지 논쟁이 있는데 규범은 연구목적에 따라 해당 목표어의 원어민 수준의 언어가 되거나 다른 중간언어 학습자의 언어, 혹은 학습자 자신의 이전, 이후 단계의 중간언어가 될 수 있다(Towell, 2012:52-53). 그러나 일반적으로 제2언어 학습은 '모어화자와 같은 언어를 구사하는 것'을 목표로 하기 때문에 제2언어습득 분야에서 '규범'은 일반적으로 목표어의 규칙 체계, 즉 모어화자의 언어를 가리키는 경우가 많다. 이는 오류에 대한 정의에서도 나타나는데 Lennon(1991:182)에서는 오류를 '같은 맥락과 유사한 상황에서 모어화자에게서는 생산되지 않는 언어적 형태나 결합된 형태'로 정의하였다. 한국어교육 분야에서도 마찬가지인데 배두본(2002:348)에서는 정확성은 교실에서 언어의 사용과 관련되는 개념이 아니라 언어의 쓰임과 관련된 언어 형태의 정확함을 지칭하는 것으로 그 표준은 교양이 있는(well-educated) 원어민 화자가 표현하는 말이라고 하였다. 이러한 정의를 참고하여 본고 역시 '규범'을 모어화자, 즉 한국어 원어민 화자의 언어로 정의하도록 한다.

2) 오류의 정의

오류의 1차적인 정의는 '규범에서 벗어난 부분'이지만 실제 연구에서 사용되는 오류의 분석 기준이나 오류에 대한 수용 정도는 연구의 목적에 따라 달리 설정되기도 한다. 선행연구를 살펴보면 오류에 대한 정의는 매우 다양하다. 예를 들면, Odlin(1989:27)에서는 오류에 대해 '목표 언어와 이미 습득된 언어 사이에서의 유사성과 상이성에서 나오는 결론들의 영향'으로 정의하였고 Brown(2000:277)에서는 '원어민의 성인 문법을 기준으로 보았을 때 무엇인가 잘못된 것, 이탈 형태(deviation)로서 학습자의 언어 능력을 반영하는 것'이라고 정의하였다.

그러나 오류를 무엇으로 볼 것인지에 정하기에 앞서 필요한 것은 바로 언어의 본질, 즉 학습자의 언어 능력에 대한 정의이다. 외국어 학습자의 언어 능력에 대한 정의는 황

적륜(1975:31)에서 찾아볼 수 있는데 이 연구에서는 그동안 변형생성문법론에서 외국어 학습자의 언어 능력을 규정함에 있어 '문법성(grammaticality)'에만 초점을 두었던 것을 비판하며 실제의 언어 사용은 문법 규칙에 의해서만 규제되는 것이 아니라 여러 가지 사회적 요건에 의해 이루어지는 사회언어학적 규칙에 의해 규제된다'고 하며 언어 운용 능력 즉 사회적 '용인 가능성(acceptability)'을 포함해야 함을 주장하였다. 여기에서 '용인 가능성'은 '사회적으로 타당한 표현'을 가리킨다. 이와 비슷하게 Corder(1981:72-76)에서는 적형성(well-formedness)과 적절성(appropriateness)을 오류 분석의 기준으로 삼으면서 여기에서 적형성은 문법성과 같은 개념으로 기호체계의 범주, 즉 형태적으로 정확한 것을 가리키고 적절성은 기호체계의 적절한 사용 즉 문맥에서의 적절한 언어 사용을 가리킨다고 하였다. 그리고 적절성을 다시 지시적 적절성(referential appropriateness)과 사회적 적절성(social appropriateness)으로 나누었는데 지시적 적절성은 발화의 구체적인 진위가[19]에 의해 판단되는 것이고 사회적 적절성은 사회적 상황에 맞는 언어의 사용 영역과 관련이 있다고 하였다.

한편, 이정희(2002)에서는 황적륜(1975)의 논의를 참고하여 오류 판정에 대한 기준을 '문법성'과 '용인가능성(acceptability)'으로 규정하고 있으나 이 연구에서의 용인가능성은 Lyons(1977)의 교정가능성(corrigibility)에 가깝다고 밝히고 있다. 교정가능성에 대한 Lyons(1977)의 정의를 확인해 보면 '비문법적인 발화를 모국어 화자가 용인 불가능하다고 인식은 하지만 교정은 가능한 것'으로 정의하고 있다. 문법성과 용인성에 대한 정의는 Chomsky(1965)와 Hughes & Lascaratou(1982)에서도 찾아볼 수 있는데 Chomsky(1965:10-11)에서는 용인성을 '언어 수행과 관련된 개념으로 자연스럽고 이해 가능함'을 나타낸다고 하였고 Hughes & Lascaratou(1982:175)에서는 규칙의 기본성(basicness of the rules)과 해독가능성(intelligibility)을 오류 판정의 기준으로 들면서 외국어 교육의 목표가 의사소통이라면 그 평가의 기준은 해독가능성이어야 한다고 주장하였다.

위의 논의들을 정리해 보면 오류를 판정함에 있어 '적형성'이나 '문법성', 즉 형태적 정확성에 대한 기준은 동일하며 이견이 없음을 알 수 있다. 그렇다면 남은 것은 '용인성(acceptability)'의 문제인데 황적륜(1975)에서 주장하는 '사회적 용인가능성'과 Corder(1981)에서 주장하는 '적절성'은 사회언어학적 관점에서 상황과 맥락에 적절한

[19] Corder(1981)에서의 진위가(眞僞價)는 논리적이고 사실적인가에 대한 판단을 가리키는데 예를 들면 학습자가 'I have a hat on my head'라고 말했지만 실제적으로 학습자가 캡(cap)을 쓰고 있거나 대머리인 경우 이 발화가 허용은 가능하지만 부적절하다고 판단하는 것이다.

언어인지를 판단하는 기준이다.[20] 즉 학생이 교사에게 '밥 먹었어?'라고 발화하는 경우 문법적으로는 문제가 없지만 사회적으로 용인되지 않는 발화임에는 확실하다. 그러나 Chomsky(1965), Hughes & Lascaratou(1982), 이정희(2002)에서의 '용인성'은 '이해가능성' 즉 '화자의 발화 의도에 대한 이해가능여부'를 가리킨다고 볼 수 있다. 그러므로 '해독가능성' 관점에서의 '용인성'은 비문법적일지라도 용인가능성이 있거나 용인가능성이 높다고 볼 수 있다.

오류에 대한 정의가 이처럼 다양한 이유는 오류의 정의가 연구의 목적에 따라서 달라지기 때문이다. 학습자의 오류를 연구하는 목적은 이론적인 것과 실용적인 것으로 나눌 수 있는데 이론적인 관점에서 학습자의 오류를 분석하는 것은 언어 학습의 과정을 관찰하기 위한 것으로 특정 시기의 목표어에 대한 학습자의 지식을 기술하기 위한 것이다. 반면, 실용적인 관점에서 학습자의 오류를 분석하는 것은 오류를 교정하기 위한 방법을 찾는 것이다(Corder, 1981:83). 따라서 오류를 교정하기 위한 목적이라면 문법성에 약간 어긋나더라도 용인가능성, 이해가능성이 높다면 오류로 보지 않고 교정 처방을 내리지 않을 수도 있다. 그러나 학습자의 오류를 기술하기 위한 목적의 연구에서는 사소한 오류라고 할지라도 모두 기록하고 분석할 필요가 있다. 또한 해석 가능성이나 용인 가능성은 판단하는 입장에 따라서도 크게 차이가 날 수 있다는 단점이 있다. 본 연구에서 오류를 바라보는 관점은 한국어에 대한 학습자의 지식을 기술하기 위한 것으로 '용인가능성'과 '해석가능성'은 오류의 판정 기준으로 사용하기에 적합하지 않다.

정리하면, 본고에서 오류문 판정에 필요한 기준은 형태적 정확성인 적형성(문법성)과 맥락적 정확성인 적절성이다. 그리고 적형성, 적절성 판단의 참조 기준은 '성인 모어화자의 언어'가 된다. 따라서 본고에서 정의하는 정확성은 '성인 한국어 모어화자가 같은 의미를 표현하기 위해 사용하는 언어를 기준으로 적형성과 적절성에 위배되지 않는 문장'이다. 그리고 이러한 적형성과 적절성에 위배되는 문장 중에서 화자의 의도를 파악하기 힘들거나 여러 오류가 혼합되어 있어 '교정가능성(해석가능성, 재구성가능성)이 없다고 판단되는 문장들은 구체적인 오류 분류가 불가능한 문장으로 '총체적 오류'로 구분하도록 한다. 오류의 판단은 문장 단위를 기본으로 하며 지시어나 접속어의 오류 판단이 필요할 경우 선행문과 후행문의 연결을 고려하여 오류를 판단한다.

---

[20] Corder(1981: 39-40)에서는 정형성(well-formedness)을 나타내는 적합성(acceptability)과 맥락과의 관계에서 결정되는 적절성(appropriateness)이라는 개념으로 오류를 규정하였다. 즉 적합하지 않거나 적절하지 않은 문장을 모두 오류로 보았다.(석주연·안경화, 2003:192)

3) 오류와 실수

오류를 판정함에 있어 오류의 기준을 정하는 것만큼이나 중요한 것은 오류(errors)와 실수(mistake)를 구분하는 것이다(Corder, 1967; Ellis, 1997; Brown, 2000). Corder(1967:166-167)에서는 오류에 대해 단순히 불완전한 학습의 산물이 아닌 학습자만의 언어 체계의 규칙에 의한 것으로 보았다. 따라서 기저 언어 지식을 드러내는 체계적인 오류(systematic error)는 학습자의 특정 언어 체계를 보여주는 증거이지만 발화 실수와 같은 우연적이고 비체계적인 오류(nonsystematic error)는 별로 중요하지 않다고 보며 언어 수행의 오류(errors of performance)는 실수(mistake)로, 언어 능력의 오류(errors of competence)는 오류(error)로 구분하기도 하였다.

Brown(2000:278)에서 역시 오류와 실수는 구별되어야 한다고 주장하며 실수는 학습자가 언어 체계를 알고 있으나 사용하는 데 실패한 경우, 오류는 학습자의 언어 체계가 정확하지 않은 경우로 구분하였다. 그러나 동시에 오류와 실수를 구별하는 것의 어려움을 언급하며 학습자의 자기수정 과정을 관찰하는 것이 좋지만 학습자의 자기수정 과정을 관찰 불가능하다면 이탈 형태의 빈도를 확인할 수 있다고 하였다. 그러나 실수와 오류는 구분하는 데에는 연구자의 주관적인 판단이 개입될 수 있는 부분이다. 이에 대해 이정희(2002:43)에서는 학습자 자료의 오류와 실수를 제대로 구분하기 위해서는 연구자가 학습자 개개인의 언어 숙달도 및 학습자들의 개인적인 언어 습득 과정에 대해 잘 이해하고 있어야 한다고 하였다. 이 점에서 본고는 종적 연구로 같은 학습자의 자료를 지속적으로 수집하여 학습자의 언어 능력에 대해 비교적 잘 파악하고 있으므로 자료에서 나타나는 오류와 실수를 잘 판단할 수 있다는 장점이 있다.

실수는 특히 맞춤법이나 종결 어미에서 많이 나타나는데 철자 오류의 경우, 한 편의 작문에서 여러 번 쓰였으나 한 차례의 오기가 나타난 경우, 혹은 이전에는 항상 정확하게 써 왔으나 이후 일차적인 오기가 발생한 경우는 실수로 판단하였다. 실수는 학습자 본인이 오류임을 인식하고 정확히 수정할 수 있다는 점에서 언어 능력의 결함은 아니다. 따라서 문장의 정확성 여부를 판단함에 있어 실수는 오류로 보지 않는다. 또 구어 발화 중에 학습자가 자기 수정을 하여 최종적으로 정확한 형태로 발화하였다면 학습자 스스로 오류임을 인지하고 수정을 한 것이므로 실수로 판단하도록 한다.

3.2. 정확성의 측정

‘규범’과 ‘오류’에 대해 정의가 확정되면 정확성을 측정할 수 있다. 정확성을 측정하

는 방법 역시 통사 복잡성 측정과 마찬가지로 전체적 측정 지표와 구체적 측정 지표로 나눌 수 있다. 전자는 무오류절수(error-free clause), 100단어당 무오류절 수, 100단어당 오류 수와 같은 방법들로 다양한 언어 과제에 사용할 수 있다는 장점이 있으나 특정 항목의 정확도를 측정하는 데 제한이 있고 후자는 특정 언어 형식에 대한 정확도를 측정하기 위한 방법으로 학습자가 특정 언어 항목이나 구조를 생산하도록 설계된 유도 과제, 구조집중형과제(focused task)에서 널리 사용된다.

이 중 특정적 측정은 특정 언어 항목의 정확성을 살핌으로써 해당 언어 항목의 오류 양상이나 발달 양상을 관찰할 수는 있지만 이를 학습자 언어 발달의 지표로 삼기에는 무리가 있다. 또한 특정 측정은 측정하고자 하는 언어 항목을 선정한 뒤 학습자가 해당 항목을 정확하게 사용할 수 있다면 다른 것들도 사용할 수 있을 것이라는 가정을 바탕으로 한다. 그러나 학습자는 모든 문법을 동시에 습득할 수 없고 일부 기능은 일찍 습득되고 일부 기능은 늦게 습득되며, 또 학습자의 모국어가 특정 기능의 학습을 쉽게 만들 수 있기 때문에 특정 문법 항목을 통해 정확성을 측정하는 것은 위험한 일이다(Ellis & Barkhuizen, 2005:151).

반면, 일반적 측정은 특정 언어 형태를 유도할 필요가 없이 자연스러운 맥락에서 학습자의 언어 산출을 살필 수 있다는 장점이 있다. 또한 특정 측정에 비해 실험 조건 간의 차이를 감지하는 데 더 민감하며(Skehan & Foster, 1999:106) 더 폭넓은 정확성 양상을 제공할 수 있어 집단 간, 학습자 간의 차이를 비교할 수 있고 비구조화된 과제에 적합하다(이복자, 2015:39). 그러므로 본고와 같은 언어 발달 연구에서는 특정 측정보다는 일반적 측정이 더 적합하다고 볼 수 있다. 일반 측정 방법은 100단어당 오류 수나 100단어당 무오류절 수와 같은 빈도 측정, 그리고 무오류절의 비율과 같은 비율 측정으로 구분할 수 있는데 Wolfe-Quintero 외(1998:38)에 따르면 비율 측정이 빈도 측정에 비해 더 높은 신뢰도를 확보할 수 있음을 확인하였다.

한국어교육 분야의 선행연구들을 살펴보면 정확성은 일반적으로 무오류절의 비율(박지순·서세정, 2009; 김영주 외, 2013; 서신웨, 2022), 단위당 무오류절 수(서세정, 2009; 김영주 외, 2013; 서신웨, 2022)나 단위당 오류 수(이복자, 2015; 주월랑, 2018; 주우동, 2019a)로 측정되어 왔다. 또한 연구에 따라 차이가 있지만 세 가지 지표 모두 숙달도별로 유의미한 차이를 나타내고 있음을 확인할 수 있다. 기존의 연구에서는 무오류를 통한 측정이나 오류를 통한 측정 중 한 가지를 선택하여 측정하기도 하였으나 세 가지 지표는 단위를 통한 측정과 비율을 통한 측정, 무오류를 통한 측정과 오류를 통한 측정이라는 점에서 정확성의 서로 다른 측면을 나타내는지표이기도 하다. 본고에서

는 학습자의 정확성을 다각적으로 살펴보기 위한 목적으로 세 가지 측정 지표를 모두 사용하도록 한다.

그러나 일반적 측정만을 사용할 경우, 구체적으로 어떤 오류가 발생하는지에 대한 정보를 얻을 수 없게 된다는 단점이 존재한다. 그러므로 가장 이상적인 것은 무오류절의 비율과 같은 일반적 측정과 특정 측정을 함께 사용하는 것이다(Ellis & Barkhuizen, 2005:151). 이러한 혼합 측정의 경우 일반적인 측정에 특수 항목을 보충하여 보다 세부적인 정보를 제공함으로써 언어 발달의 깊이 있는 고찰을 가능하게 한다는 장점이 있다(Norris & Ortega, 2009). 그러나 본 연구는 학습자의 언어 발달을 전체적인 관점에서 살펴보는 연구로 모든 언어 항목의 오류를 살피기에는 한계가 있다. 따라서 특정 측정 대신 오류를 유형별로 나누어 유형별 오류를 측정하는 것으로 특정 측정을 대신하여 좀 더 세부적인 정보를 얻고자 한다.

4. 복잡성과 정확성의 관계

복잡성, 정확성, 유창성이라는 세 가지 측정 구인이 확인되면서 제2언어 습득 영역의 연구는 새로운 단계로 진입하였다. 학습자 언어에 대해 더 이상 모호하고 경험적인 평가가 아니라 구체적인 측정 지표를 활용한 실증적 분석의 단계로 진입한 것이다. 학습자의 언어 수행을 수치화할 수 있게 되면서 연구자들은 측정 지표를 다차원적으로 세분화하여 학습자의 언어 지식이나 과제의 특징, 인지 능력과 같은 요인이 언어 수행에 미치는 영향들을 탐색하기 시작하였고 그 과정에서 세 구인 사이의 관계에도 관심을 가지게 되었다.

제2언어습득에서 학습자 언어의 복잡성, 정확성, 유창성 세 구인 간의 관계에 주목한 연구들은 유창성의 증가가 정확성과 복잡성의 발달에 영향을 줄 것이라고 추측하거나(Ellis, 1994:107) 복잡성-정확성-유창성의 순서로 언어가 발달한다고(Housen 외, 2012) 하기도 하였다. 복잡성과 정확성은 언어 발달의 각각의 다른 측면을 나타내지만 복잡성은 '제2언어 지식에 대한 확장과 재구조화의 정도'이고 정확성은 '제2언어 지식과 목표어의 일치 정도'이다(Wolfe-Quintero 외, 1998:4). 따라서 이들은 같은 언어 수행과 숙달도의 다른 측면이며 모두 학습자의 인지처리 과정에 따른 결과라는 점에서 서로 연관되고 서로 영향을 주게 된다. 또한 복잡성은 학습자가 명시적으로 습득된 중간언어의 규칙을 점차 내면에서 절차화하려는 시도에서 발생하고 정확성은 학습자의 형태에 대한 의식적인 집중과 관련이 있는 것으로 학습자의 선언적(declarative) 또는

명시적(implicit)인 중간언어 지식에 의해 부분적으로 결정된다. 따라서 정확성과 복잡성은 명시적인 선언적 지식과 암시적인 절차적 지식과도 부분적으로 연결되어 있다(Housen 외, 2012:5-6). 즉 언어 처리 기제 안에서 복잡성과 정확성은 모두 언어 지식으로서 상호 연결되어 있다. 그러므로 학습자 언어 수행 과정에서 복잡성과 정확성이 서로 관련성이 있으며 서로 영향을 줄 것이라는 점에서는 대부분의 연구자들이 같은 입장을 보였다.

그러나 구체적으로 복잡성, 정확성, 유창성 간에 어떤 관계가 존재하는지에 대해서는 접근하는 관점에 따라 다양한 주장들이 존재하며 서로 상반되는 입장들도 있다. 먼저 인지적 관점에서 인간의 주의력의 한계로 인해 구인들 사이에 경쟁 관계가 존재하며 한 구인의 발달이 다른 구인의 희생을 동반한다고 주장하는 관점과 이와 반대로 과제 조건에 따라 세 구인이 동반 증가 혹은 동반 감소할 수 있다고 주장하는 관점이 존재한다. 그리고 사회적 관점에서는 CAF 각각의 구인을 큰 체계 속에서 서로 연결된 하위 체계로 보고 구인들 사이의 관계는 고정적인 것이 아니라 부단히 변화하면서 발전한다고 보기도 한다.

이와 같이 언어 발달의 각 측면에 대해 분리하여 보지 않고 언어 발달을 학습자의 인지적 능력과 처리 제약의 복합적인 상호작용의 결과로 보는 대표적 이론에는 Skehan(1996a)의 '제한된 주의력 모델의 맞교환 가설(trade-off hypothesis)'과 Robinson(2001a, 2005)의 '다자원 주의력 모델의 인지 가설(cognitive hypothesis)', 그리고 언어의 각 측면 사이의 영향 관계가 고정적이지 않고 서로 영향을 주면서 역동적으로 발달한다고 보는 역동체계이론(dynamic system theory)이 있다.

이 중 맞교환 가설과 인지 가설은 인지 처리 기제(mechanisms)와 CAF사이의 관계에 대해 서로 상반되는 입장을 보이지만 모두 과제 기반 연구에서 제기된 것으로, 학습자의 복잡성과 정확성, 유창성 사이의 관계를 설명하는데 주의력, 작업 기억, 자동화, 추론 및 기타 인지 처리 기제의 역할에 초점을 맞추고 있다(Housen 외, 2012:6). 그리고 역동 체계 이론은 어느 하나의 이론을 지지하는 것이 아닌 전체적인 관점에서 언어 발달을 바라보며 환경과 환경 속에서 변화 발달하는 과정에 주목한다. 아래, 각 이론에서 주장하는 복잡성과 정확성의 관계에 대해 살펴보도록 한다.

4.1. 경쟁 관계: 제한된 주의력 모델

Skehan(1996, 1998)의 제한된 주의력 모델(limited attentional capacity model)은

정보처리 이론에 기반하여 인간의 인지적 처리 능력에는 제한이 있으며 이러한 제한된 작업 기억(working memory)으로 인해 학습자는 입력과 산출 과정에서 형태와 의미에 균형적으로 집중할 수 없다고 보았다. Skehan(1998, 2001)은 과업 수행 시 학 습자들의 중점적 주의 집중이 의미와 언어 형태에 동시에 투여될 수 없다고 지적하고, 이를 상쇄효과(trade-off effect)로 설명한다. 즉 한 영역에 대한 주의력 자원 집중은 다른 영역의 주의력 자원에 대한 손실을 일으키게 되어 결국 언어 수행 영역 간에 상호 맞교환이 이루어진다고 본다.

그러므로 인지 처리 부담이 높은 복잡한 과제를 수행하는 경우, 학습자들의 주의력 자원은 과제의 의미 처리 과정에 할당되어 언어 형태에 대해 주의를 기울이지 못하게 한다. 즉 복잡한 과제에서 학습자의 주의력은 과제의 복잡한 내용과 그에 상응하는 복잡한 언어 형태에 할당되기 때문에 언어의 복잡성은 높아지는 반면, 언어의 정확성은 낮아지게 된다고 보는 것이다. 반대로, 단순한 과제에서는 과제를 수행하는데 있어 복잡한 언어 구조를 필요로 하지 않으므로 언어의 복잡성은 낮아지지만 정확성은 높아진다는 것이다. 여기서의 언어 복잡성은 복잡한 시제, 하위절, 내포절 등과 같이 과업 수행 시 필요한 언어적 요소를 의미한다.

이처럼 제한된 주의력 모델의 관점에서 볼 때 복잡성과 정확성은 한정된 주의력 자원에 대해 서로 경쟁하는 관계로, 한 가지 구인이 향상되면 다른 한 구인은 하락하는 경쟁 관계에 놓여 있고 학습자 언어 수행의 결과는 학습자가 자신의 주의를 어디에 두는지와 과제의 유형, 과제 수행 조건과 관련된다(Skehan & Foster, 1999:96-97). Skehan(1996a:49-50)에서는 정확성과 복잡성의 관계에 대해 다음과 같이 설명하고 있다.

> 정확성에 초점을 맞추면 중간언어에서 변화가 일어날 가능성이 낮아지고, 말이 느려질 가능성이 높아지며, 주의력 자원의 상당 부분을 소비하게 된다. 복잡성과 재구조화 과정에 초점을 맞추면 새로운 형태가 중간언어 체계에 통합될 가능성이 높아지고, 위험 감수를 촉진하며, 새로운 언어 형태의 구성에 주의를 기울이게 된다(Foster & Skehan, 1994). 마지막으로, 유창성에 초점을 맞추면 언어가 더 빨리 산출되고 접근성에 중점을 두게 되지만(Bygate, 1988) 언어의 정확성이나 새로운 형식에 대한 중시 정도가 낮아지게 된다(Skehan, 1996a:49-50).

요약하자면, 제한된 주의력 모델의 주요 주장은 인간의 주의력은 한정적이며 학습자들은 과제 수행에서 어느 한 부분에 주의를 집중하면 그 주변 부분에 대한 주의 집중이

감소하게 되며 이러한 현상이 정확성과 복잡성이 동시에 나타나지 못하도록 한다고 본다. 다음 장에서 제시할 인지 가설을 포함한 여러 실증적 연구들을 통해 이 주장이 항상 정확한 것이 아님이 증명되었지만 제한된 주의력 모델은 복잡성, 정확성과 유창성의 세 구인을 연결시켜 바라본 첫 연구로 이후의 CAF 관련 실증적 연구나 이론적 연구의 계기가 되었다.

4.2. 지지 관계: 다자원 주의력 모델

Robinson(2001b, 2005)이 제안한 다자원 주의력 모델(multiple resources attentional model)의 인지 가설(cognition hypothesis) 역시 과제 복잡성을 기반으로 학습자의 언어 수행과 언어 발달을 설명하는 이론이다. 그러나 인지 가설은 맞교환 가설과 반대로 주의력 자원이 제한적이지 않다고 주장한다.

Robinson의 인지 가설은 Wickens(1980)의 다중 자원 모형(multiple resource model)에 기초한 것이다. 다중 자원 모형에 따르면, 만약 추가적으로 요구되는 주의집중이 동일한 차원(dimension)에 있지 않고 시각적 자원과 청각적 자원과 같이 다른 차원에 있다면 인간의 주의적 자원(attentional resources)은 한정적이지 않다. 이에 인지 가설은 인간의 주의력 자원은 제한적이지 않으며 언어 학습의 과정에서 발생하는 주의력의 배분 역시 서로 경쟁하지 않는다고 설명한다.

인지 가설에서는 학습자는 근본적으로 다차원적이고 비경쟁적인 주의력 자원에 동시에 접근할 수 있으며, 결과적으로 L2의 복잡성과 정확성은 함께 발달할 수 있다고 본다. 그리고 정확성과 복잡성은 서로 연계되어 있기 때문에 복잡한 과제일수록, 학습자가 더 복잡한 생각을 표현해야 할수록 과제 수행에 더 많은 주의력을 기울여야 하기 때문에 두 구인이 함께 발달한다고 주장한다(Robinson, 2003:47). 특히 다양한 인지적 요소를 포함한 더 복잡한 과제가 학습자의 높은 언어적 출력을 이끌어 최종적으로 학습자로 하여금 구조적 복잡성과 정확성을 성취하게 된다는 입장이다. 그러나 동시에 Robinson은 복잡한 과제에서 학습자의 언어 산출은 더 복잡하고 정확해지지만, 같은 과제의 더 간단한 버전에 비해 유창성은 저하된다고 주장하기도 하였다(Housen 외, 2012:148).

정리하면, Robinson의 인지 가설은 복잡성과 유창성, 정확성과 유창성 사이의 경쟁 관계에 대한 입장은 동일하지만 복잡성과 정확성의 관계에 대한 입장은 상반됨을 알 수 있다. 즉 Robinson의 인지 가설은 Skehan의 제한된 주의력 모델과 반대로 언어 수

행에서 복잡성과 정확이 동시에 높게 나타날 수 있다고 주장하는 것이다. 그러나 비록 맞교환 가설과 인지 가설은 복잡성, 정확성, 유창성의 상호 관계에 대한 입장은 서로 반대되지만 모두 인지심리학적 관점에서 언어 발달의 각 측면을 연결된 것으로 바라보고 구인 사이의 상호 관계를 밝히려고 했다는 점에서는 동일하며 제2언어습득 연구에서 복잡성, 정확성, 유창성에 대해 구인 각각에 주목하였던 것으로부터 구인 사이의 관계에 주목하게 하였다는 점에 큰 의의가 있다. 그러나 두 이론 모두 복잡성과 정확성 사이의 고정된, 보편적인 관계를 밝히려고 했다는 점에서 언어의 역동적인 측면을 고려하지 않았다.

Skehan과 Robinson의 주장이 제기된 이후 여러 연구들에서는 이 두 가설을 검증하려는 시도들을 하였다. 그러나 종속변수(즉 복잡성, 정확성, 유창성)의 개념적(conceptual) 조작적(operational) 명확성의 부족함으로 인해 입증되지 못하고 있으며 지금까지 이 두 모델을 명확히 지지하는 경험적 증거들은 충분하지 않다(Housen 외, 2012:6). Ahmadian & Tavakoli(2011)의 연구에서는 복잡성과 정확성은 함께 발달할 수 있지만 복잡성과 유창성, 정확성과 유창성은 경쟁 관계에 있는 것으로 나타났고 Vercellotti(2012)의 종적 연구에서는 복잡성, 정확성과 유창성 사이에 고정된 경쟁이나 지지 관계가 나타나지 않음을 확인하였다. 또한 여러 축적된 연구들은 제2언어습득에서 복잡성, 정확성, 유창성의 구인은 동시에 발달하지도 선형적인 발달 과정을 보이지도 않으며 상호 연계되어 다변인적이고 역동적인 양상을 나타낸다는 것을 발견하였다(Larsen-Freeman, 2006; Spoelman & Verspoor, 2010). 즉 이 세 구인은 동일하게 발달하는 것이 아니라 때로는 상호 지지적이고 때로는 서로 경쟁적인 관계를 나타내는 것으로 나타났다(Housen 외, 2012:7).

특히 이러한 주장들은 2010년을 기점으로 많이 제기되었는데 이는 제2언어습득에 대한 연구가 심화되면서 연구자들이 보편적인 규칙으로는 설명하기 힘든 여러 가지의 현상들을 발견하였고 이를 학습자 개개인의 언어 발달 과정을 추적하여 밝히고자 한 종적 연구들이 늘어난 것과 관련이 있다. 이러한 종적 연구들에서는 언어 발달의 많은 측면은 비선형적(non-linear)이고 CAF 세 구인은 다변량적(multivariate)이고 역동적(dynamic)이며(Spoelman & Verspoor, 2010:547) 각각의 하위 구성 요인은 모두 기타 구성 요인과 상호작용을 하면서 자신만의 발달 경로를 나타낸다는 것(Larsen-Freeman, 2006; Norris & Ortega, 2009)을 발견하였다.

이처럼 종적 연구를 바탕으로 학습자 개개인의 발달 과정과 발달 과정에서 나타난 구인 간의 상호 관계, 상호 관계의 변화에 주목한 연구들은 대부분 다음에서 제시할 역

동 체계 이론을 바탕으로 한 연구들이다. 역동 체계 이론은 앞의 두 이론과 마찬가지로 CAF 구인의 연결성과 상호작용에 주목하지만 이와 더불어 이들 사이의 상호작용 방식은 시간에 따라 변한다는 사실에 주목한다.

4.3. 역동적 관계: 역동 체계 이론

역동 체계 이론은 복잡 체계를 바탕으로 하는 이론이다. 여기서 복잡 체계는 단순 체계(simple systems)와 대응되는 개념으로 서로 독립적이고 제한적인 요소들로 구성되고 조직화된 규칙이 있어 단순하게 예측이 가능한 체계(예: 교통 신호 체계)와 달리 다양한 성질과 다양한 층위의 수많은 요소들이 복잡하게 연결되고 서로 끊임없이 상호작용을 하여 예측이 불가능한 복잡한 체계(예: 하나의 큰 숲 체계)를 가리킨다.

복잡 체계에 대한 관심은 1940년대 생물학자 Conrad Waddington의 연구로 거슬러 올라갈 수 있는데 그는 당시의 '유전자가 유기체의 형태에 대한 완전한 설명을 전달한다'라는 주장을 반대하며 '유전자는 배아 형성의 시작점일 뿐이며 배아가 형성되기 시작하면 발달 과정의 각 단계가 다음 발달 단계의 조건이 됨'을 제시하였다. 그 뒤로 생물학자 Von Bertalanffy(1950)에서는 '일반 체계 이론(general systems theory)'을 제기하여 복잡한 질서(complex order)가 어떻게 생성되는지를 해석하면서 하나의 독립체(entity)를 각 부분의 합으로 보는 환원주의를 반대하고 각 부분과 전체의 관계를 이해하는 체계적 접근법(systems approach)을 주장하였다(Larsen-Freeman & Cameron, 2008:3). 이 뒤로도 화학, 수학, 기상학 등 분야에서 복잡 체계에 대해 연구가 계속되었고 복잡 체계에 대한 복잡한, 역동적인, 비선형적인, 자기조직적인, 개방적인, 창발적인, 가끔은 혼돈적이며 자기적응적인 특징에 대해 다루었다(Larsen-Freeman, 1997).

복잡 체계는 현재 물리학, 생물학, 수학뿐만 아니라 사회과학, 관리학, 경제학, 의학, 교육, 문학 등 다양한 분야에서 적용되고 있는 개념으로 다양한 층위(level)의 사실에도 적용할 수 있다(Larsen-Freeman & Cameron, 2008:1). 복잡 체계 이론을 제2언어 습득에 적용하면 언어의 발달을 하나의 복잡 체계로 보고 그 하위 체계를 학습자 요인, 언어 요인, 환경 요인으로, 언어 요인 체계의 하위 요인으로 음운, 어휘, 문법, 담화로, 더 나아가서 문법 체계를 더 작은 하위 요인으로 나눌 수 있는데 상위 체계를 이루는 하나의 요인은 동시에 이를 이루는 하위 요인들로 구성된 하나의 체계가 되는 것이다. 복

잡 체계에서는 이를 프랙탈(fractal)[21]이라고 부른다.

복잡 체계는 '복잡 체계(complexity systems)', '역동 체계(dynamical systems)', '복잡 적용 체계(complex adaptive systems)' 등 여러 용어로 불리는데 각각의 용어는 강조하는 속성이 조금씩 차이가 있다. '복잡 체계'는 어휘의 의미 그대로 체계의 복잡성을 강조하는 용어로 한 체계의 내부에는 다양한 층위의 복잡한 하위 체계가 존재하며 체계의 변화는 그 체계 내의 복잡하고 다양한 구성 요소의 상호작용으로 인한 결과라는 것이다. '역동 체계'는 이러한 복잡한 체계의 구성 요소와 구성 요소 간의 상호작용이 시간의 흐름에 따라 끊임없이 변화한다는 '역동성'을 강조하는 용어이며 '복잡 적응 체계'는 이러한 복잡한 체계 내에서 적응과 학습이 이루어진다는 측면을 강조한 용어이다. 즉 역동 체계는 복잡 체계를 기반으로 하면서 체계의 역동적인 속성을 강조한 이론이다.

역동 체계 이론은 복잡하고 역동적인 체계의 변화 과정, 즉 복잡한 체계 중에서 상호작용하는 각 부분은 체계에 어떤 집합 행동(collective behavior)을 일으키며 체계는 어떻게 이를 둘러싼 환경과 상호작용을 하며 역동적으로 변화하는지를 해석하고 묘사하기 위한 이론이다. 응용언어학의 영역에서는 Larsen-Freeman(1997)에서 복잡 체계 이론을 제2언어 습득에 처음 적용하였다. Larsen-Freeman(1997:151)에서는 제2언어 습득의 복잡하고 역동적이며 비선형적이고 개방적인 특징들을 설명하면서 그동안 이원론적으로 보아왔던 제2언어 습득의 과정을 연구하는 새로운 관점을 제기하였다.

역동 체계 이론의 관점에서 볼 때 인지 자원은 제한적이다. 그러나 또 서로 연결되어 있으며 서로 보충해 주기도 한다(De Bot, 2008:170). 이러한 체계 간의 상호 연결은 하나의 체계의 변화가 다른 체계 내부의 기타 하위 체계에 영향을 주게 된다. 즉 언어를 하나의 체계로 볼 때 음운, 어휘, 문법과 같은 하위 체계들이 서로 연결되어 있으며 한 체계의 변화가 다른 체계의 변화에 영향을 미치는 것이다. 이러한 상호 관계는 자원이나 시간의 흐름에 따라 변화하며 이는 결국 학습자 언어발달이 단순한 누적이나 합으로 이루어진 선형적인 발달 과정이 아닌 비선형적이고 역동적인 발달 과정으로 나타내게 한다. 이는 전통적인 관점에 기반한 대부분의 언어 습득 연구에서는 제2언어 발달 과정은 '입력'과 '산출'이라는 인지적 처리 과정을 거쳐 '목표어'를 향해 발전하는 선형적인 과정이며 대부분의 학습자가 비슷한 발달 과정을 보인다는 가정으로 접근하여 언

[21] 프랙탈(fractal)이란 전체를 부분으로 쪼갰을 때 부분 안에 전체의 모습이 담겨 있는 기하학적 도형을 의미하는 말로 프랙탈 도형은 '자기유사성(self-similarity)'과 '순환성(recursiveness)'이라는 속성을 가지고 있다(이강원·손호웅, 2016).

어 발달에 나타나는 '일반적'이고 '대표적'인 규칙을 찾는 것에 집중한 것과는 다른 접근법이다.

그러므로 역동 체계 이론의 관점에서 볼 때 복잡성, 정확성은 한정된 자원 내에서 서로 연결되어 시간의 흐름 속에서 변화하고 발달하는 체계가 된다. 그러나 체계들은 항상 제한된 자원을 경쟁적으로 사용하는 것이 아니라 일부 연결된 성장자(connected growers)들은 서로 지지적 관계에 있으며, 이 경우 두 개의 성장자보다 더 적은 자원을 필요로 하기도 한다(Spoelman & Verspoor, 2010:535). 이때 두 체계가 서로 지지 관계에 놓여 있으면 지지적 성장자(cooperative growers)라고 부르고 두 체계가 서로 경쟁 관계에 있어 동시에 발달하지 못할 때 경쟁적 성장자(competitive growers)라고 부른다. 예를 들어 복잡성과 정확성, 명사구와 문법 간의 경쟁적 관계(Larsen-Freeman, 2006), 어휘 다양도과 문장 길이의 경쟁관계(Verspoor 외, 2008), 어휘 복잡성과 명사구, 어휘 복잡성과 통사 복잡성 간의 지지적 관계(Spoelman & Verspoor, 2010)와 같은 것들이다. 그러나 이러한 경쟁 관계, 지지 관계 역시 고정적인 것이 아니라 시간의 흐름에 따라 변화한다는 것이 역동 체계 이론의 중심 주장이다.

제2언어습득에 역동 체계 이론을 적용한 Larsen-Freeman은 복잡성, 정확성, 유창성 간의 단일한 고찰로는 세 구인 사이의 상호작용 및 상호 관계의 시간의 흐름에 따른 변화를 알 수 없다(Larsen-Freeman, 2009:582)고 하며 제2언어 학습자의 언어 발달을 나타내는 세 구인인 복잡성, 정확성, 유창성을 관찰할 때 이들을 하나의 복잡한 역동 체계로 보아 각 하위 체계의 발달과 하위 체계 간의 관계에 대한 더 깊은 탐구가 필요함을 강조하였다. 또한 Spoelman & Verspoor(2010:551)에서 역시 정확도와 몇 가지 복잡성의 발달 패턴이 피크와 회귀, 진행과 퇴행, 변수 간의 복잡한 상호작용으로 특징지어진다고 하며 제2언어 발달은 분명히 비선형적이라고 하며 개별 학습자의 변이성은 역동 체계의 원칙에 부합하며 정확성과 복잡성의 하위 체계의 서로 다른 수준의 개인 간 변이성을 분석하는 것이 제2언어 발달 과정에 대한 중요한 통찰력을 제공할 수 있다고 주장하였다. 즉 역동 체계 이론의 관점에서 복잡성과 정확성의 발달을 살피는 것은 두 구인 사이의 상호 관계에 대한 더 깊은 고찰을 가능하게 할 수 있다는 것이다.

위에서 살펴본 세 가지 이론은 복잡성과 정확성의 상호 관계에 대한 관점은 서로 다르지만 구인 간에 서로 연결되어 있음을 인식하고 그 관계를 해석하려고 했다는 점에서는 공통점을 가진다(이복자, 2015:47). 이를 통해, 두 요인은 개별적으로 발달하는 것이 아니라 서로 연결되어 상호작용을 한다는 점을 확인할 수 있었다. 따라서 본 연구에서도 복잡성과 정확성 간의 관계가 어떻게 나타나는지, 문어와 구어에서 어떻게 다르

게 나타나는지, 이러한 관계가 고정적인 것인지, 변화한다면 시간의 흐름에 따라 어떻게 변화하는지를 분석해 보고자 한다.

5. 문어와 구어의 관련성

제2언어 발달 분야에서 문어와 구어의 어휘와 통사 복잡성에 대한 연구(Polio, 1997; Wolfe-Quintero 외, 1998; Ortega, 2003; Bulté & Housen, 2012)들은 매우 많았지만 문어와 구어의 발달을 직접 비교한 연구는 거의 없었다. 이는 문어와 구어가 비록 동일한 자원에 의존하지만(Levelt, 1993) 실제로 생성 및 인지 방식(예를 들면, 소리와 문자의 배열)이 매우 다르기 때문이다(Chan 외, 2015:302).

제2언어습득 분야뿐만 아니라 제1언어습득 분야에서도 문어와 구어의 연결을 직접 탐구한 연구는 많지 않다. 제1언어습득 분야에서 연구자들은 기본적으로 학습자들이 문어보다 구어를 훨씬 더 일찍 습득할 것이라고 가정하거나(Kang, 2020) 쓰기를 단순히 말하기의 결과로 보고 문어의 발달은 구어의 발달에 크게 의존할 수 있다고 가정하기도 하였으며(Kroll & Vann, 1981) 심지어 문어를 구어의 기록(speech-written-down)으로 간주하기도 하였다(Kroll, 1981). 그러나 보통 구어가 먼저 발달하고 나서 이후 문자 교육을 받기 시작하는 모어화자와 달리, 제2언어 학습자는 일반적으로 문어와 구어를 동시에 습득하게 되며 환경적 요인에 따라 문어만 사용하는 경우도 있고 구어만 사용하는 경우도 있다. 즉 제2언어 학습자는 반드시 쓰기와 말하기 능력을 균형적으로 습득하지는 않는다. 따라서 말하기에 능숙한 L2 학습자라고 할지라도 반드시 글을 잘 쓰거나 읽을 수 있는 것이 아니며(Milton & Hopkins, 2006) 말하기에 능숙하지 않으나 이미 읽고 쓸 수 있는 사람도 있다(Hudelson, 1984). 그러므로 문어와 구어 사이의 선형 관계는 특히 L2 학습자에게 명확하지 않을 수 있다(Kang, 2020).

Halliday(1987)에서는 문어와 구어의 관계를 가장 자연스럽고(natural), 자동적(spontaneous)이며, 의식의 통제를 받지 않는(un-self-conscious) 언어와 가장 자기 점검적(self-monitoring)이고, 통제적(controlled)이며 자기 의식적(self-conscious)인 언어를 양 극단으로 하는 연속선 위에 존재하는 서로 다른 두 양식으로 규정하였다. 즉, 쓰기와 말하기는 같은 지식 자원을 바탕으로 이해가 아닌 산출을 한다는 점에서는 동일하며 긴밀히 연결된 것으로 볼 수 있으나 산출 양식이 다르기 때문에 같은 지식 자원을 바탕으로 할지라도 인지적 처리 과정에 차이가 있는 것이다. 따라서 문어와 구어의 관계를 알기 위해서는 쓰기와 말하기의 산출 과정에 나타나는 인지적 처리 과정을 살

펴볼 필요가 있다. 쓰기와 말하기 과정에 대한 모델은 Flower & Hayes(1981), Kellogg(1996), Levelt(1993)를 참고할 수 있다. 두 모델 모두 L1을 바탕으로 개발된 것이지만 L2의 문어와 구어 산출을 설명하는 데에도 자주 사용된다.

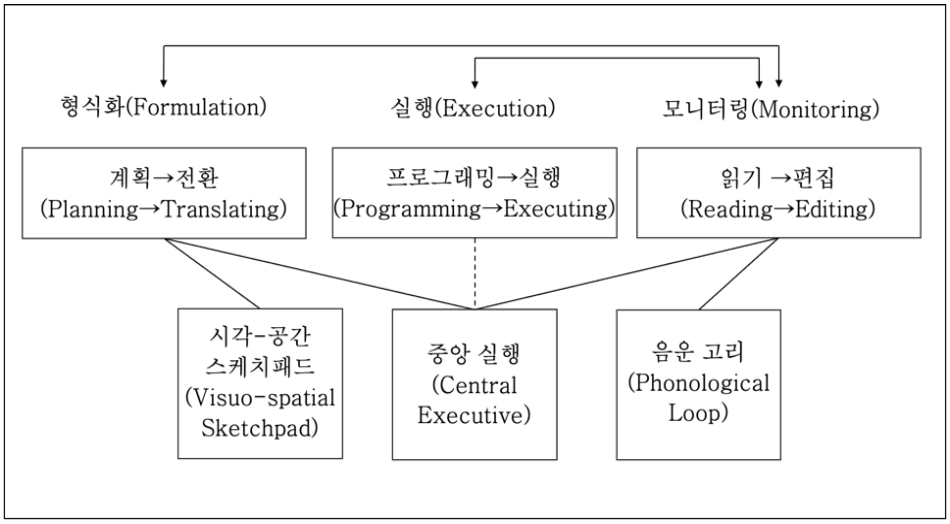
Levelt(1993:23-28)의 말하기 모델을 살펴보면 3단계의 연속된 절차로 제시하고 있는데 각각 개념화(conceptualizing), 형식화(formulating), 조음(articulating) 단계이다. 그리고 모니터링(monitoring) 절차가 있다. 각 단계를 자세히 살펴보면 개념화 단계는 메세지의 생성과 관련된 것으로는 화자의 의도나 목적에 맞는 정보나 메세지를 생각하는 과정, 즉 말하고자 하는 내용을 계획하는 단계이다. 이 과정에서 생성된 정보는 '전언어적 메세지(preverbal message)라고 한다. 전언어적 메세지는 언어보다 명제적 성격이 강한 것으로 개념화의 단계에서는 목표에 초점을 맞추는 거시적 계획과 목표를 실현하는데 필요한 내용을 선택하고 구성하는 미시적 계획도 포함된다. 다음, 형식화(formulator) 단계는 개념적 구조를 언어 구조로 전환하는 단계로 전언어적 메세지를 발화 계획으로 전환한다. 이 전환 과정은 문법 인코딩과 음운 인코딩의 두 가지로 나눌 수 있는데 먼저 화자의 심성 어휘집에서 발화 전 메세지와 일치하는 어휘소(lemas)를 식별하고 화자가 의미에 맞는 어휘소를 선택하면 어휘 형태가 활성화된다. 어휘소(lemas)의 선택은 또 학습자에게 메세지의 표면 구조를 구성하는 데 필요한 구문 정보를 제공한다. 또한 이러한 과정이 진행되는 동안 어휘소(lema)와 관련된 형태-음운론적 정보도 활성화되어 표면 구조에 통합된다. 이 과정의 결과는 음성 계획(phonetic plan)으로 출력된다. 그리고 출력된 음성 계획은 조음 단계를 거쳐 실제 발화로 변환되어 산출된다고 설명한다. 또한 세 단계의 산출 과정 외에 모니터링 과정이 있다. 모니터링은 화자가 자신의 음성 계획과 발화를 듣고 문제점을 발견하고 문제가 되는 요소들을 스스로 수정하는 과정으로 산출 과정과 상호작용하면서 발생할 수 있는 모든 실수를 식별한다.

정리하면 말하기 과정은 전언어적 메세지(preverbal message)를 생성하는 개념화 단계, 문법 및 음운을 인코딩하는 형식화 단계, 음성이 산출되는 조음 단계를 거치게 되며 이 과정이 진행되는 동안 모니터링이 작동하는 것이다. 다음 쓰기 절차를 살펴보도록 한다.

쓰기 절차 모델은 Kellogg(1996)와 Flower & Hayes(1981)에서 제시하고 있다. Kellogg(1996)의 쓰기 절차 모델은 문어 텍스트 생산을 형식화(formulation), 실행(execution), 모니터링(monitoring)의 과정으로 나누어 제시하고 있다. 아래의 <그림 4>에서 확인할 수 있듯이 각 단계는 두 가지 구성 요소 또는 절차를 포함한다. 이 세 단

계의 절차는 구분하는 방식에 있어서는 Levelt(1993)의 말하기 모델과 약간 차이가 있으나 산출 과정에 필요한 절차들은 거의 동일함을 확인할 수 있다. 형식화(formulation) 단계는 계획(planning)과 작성(translating)을 포함한다는 점에서 Levelt(1993)의 개념화(conceptualizing) 단계와 형식화(formulating) 단계를 통합하는 개념임을 알 수 있다. 그리고 실행(execution) 단계는 프로그래밍(programming)과 수행(executing)을 포함하는 것으로 Levelt(1993)의 문법 및 음운 코딩과 조음 단계에 해당하는 부분으로 볼 수 있다.

<그림 4> Kellog의 쓰기 절차 모델(Kellogg, 1996:59)



이러한 쓰기 절차는 Flower & Hayes(1981)에서도 비슷하게 제시되고 있다. Flower & Hayes(1981)는 쓰기 과정을 계획하기(planning), 전환하기·작성하기(translating), 검토하기(reviewing)로 이루어진다고 보았다. 이 중 계획 단계에서 필자는 아이디어를 생성하고 정리하며 글쓰기 목표를 설정하는데 이 세 부분은 선형적으로 이루어지는 과정이 아니라 서로 상호작용하는 과정이라고 설명하고 있다(Flower & Hayes, 1981:386-387).

위에서 살펴본 문어와 구어의 산출 과정을 비교해 보면 계획(개념화)과 전환(형식화), 산출(실행) 그리고 모니터링(자기 수정)이라는 매우 비슷한 단계를 거치고 있음을 알 수 있다. 또한 문어와 구어는 모두 산출의 화제(topic)뿐만 아니라 청자/독자(receiver)도 고려해야 하는 의사소통 기능을 가지고 있다는 점, 모두 산출 목표에 따라 정보를 조

직하고 언어를 계획하여 산출 방식을 결정한다는 점에서도 공통점이 존재한다(Chan 외, 2014:56).

그러나 두 과정은 여러 가지 차이점도 존재하는데 그중 가장 중요한 차이점은 '실시간성'에서 나타난다. 먼저, 말하기는 실시간으로 산출이 이루어지고 일반적으로 크게 멈추기가 힘든 반면 쓰기는 일반적으로 시간의 제약을 받지 않는다는 점에서 필자가 계획을 세우고 언어로 인코딩한 뒤 모니터링하는 과정이 재귀적으로 이루어질 수 있다(Kormos & Trebits, 2012:446). 이 때문에 쓰기 과정에서는 말하기에 비해 문법이나 어휘 인코딩에 대한 부담이 덜할 것으로 예상할 수 있다. 둘째는 쓰기는 시각적인 기록을 남긴다는 점에서 실체가 보이지 않는 청각적인 기록과는 달리 점검이 용이하다는 장점이 있다. 이는 실수가 눈에 더 잘 띄게 함으로써 모니터링 절차가 더 용이하게 이루어질 수 있다. 따라서 문어에서 구어보다 자기 교정이 더 쉽게 이루어진다고 볼 수 있다.

말하기 쓰기와 관련된 단계와 말하기, 쓰기의 공통점과 차이점은 L2에서도 동일한 과정으로 작동될 것이다. 그러나 L2로 말하기와 쓰기를 수행하려면 개념화 단계에서 계획된 메세지를 형식화 단계에서 L2언어로 인코딩해야 하며 이를 L2언어로 산출해야 한다. 특히 개념화 단계에서 거시적 계획은 언어에 국한되지 않지만 미시적 계획은 언어에 국한되며(De Bot, 2020) 모어화자는 형식화 단계와 조음 단계를 특별한 주의 없이 자동적으로 수행하지만 L2학습자는 통제된 처리를 통해 장기 기억에 존재하는 언어 지식을 활성화하고 실행해야 하기 때문에 작업 기억 용량이 심각하게 제한되어 많은 인지적 요구가 가해지게 되는데 이는 학습자로 하여금 모니터링의 여유가 없어지게 한다. 또한 쓰기에 비해 말하기는 실시간으로 산출되어야 한다는 점에서 숙달도가 높지 않고 어휘와 문법에 대한 절차적 지식(procedural knowledge)이 부족한 제2언어 학습자의 경우 말하기에서 어려움을 겪을 가능성이 더 높다. 따라서 통제된 과정을 통해 언어 자원에 접근할 수 있는 시간이 더 많은 쓰기에서 더 나은 성과를 거둘 가능성이 높다(Ellis & Yuan, 2005). 이러한 논의를 통해 초보적으로 제2언어 학습자는 문어에서 더 나은 수행, 즉 더 복잡한 언어를 산출하고 구어에서 덜 복잡하고 오류가 더 많은 언어를 산출할 것으로 예상해볼 수 있다.

그러나 다른 한편으로 쓰기는 문자의 형태를 산출해야 한다는 점에서 문자에 대한 자동화가 부족한 학습자의 경우 쓰기 수행 자체가 인지적 자원을 많이 소모할 가능성이 크다는 연구들도 존재한다. 이는 Bourdin & Fayol(1994), Grabowski(2004)의 연구 결과를 참고할 수 있는데 두 연구에서는 단어 목록 연속 회상 과제를 통해 성인과 아동

학습자의 쓰기와 말하기 수행을 비교하였다. 그 결과 쓰기에 익숙한 성인 학습자의 경우 문어와 구어의 수행에서 차이를 보이지 않은 반면 쓰기에 익숙하지 않은 초등학생 학습자의 경우 2학년과 4학년 집단 모두 구어에서 뚜렷하게 나은 수행을 보인다는 것을 발견하였다. 이에 Grabowski(2004:3)에서는 쓰기가 자동화되지 않은 학습자의 경우 쓰기 수행 과정이 인지적으로 큰 부하를 차지하기 때문에 문어 생산이 오히려 구어 생산보다 더 어려우며 그 어려움은 철자법(orthographic)이나 쓰기 행동(grapho-motoric)과 같은 낮은 수준의 프로세스에서 나타난다고 보았다. 이에 단순 어휘 목록 연속 회상 과제에서도 이러한 차이가 나타난다는 것은 장기 기억과 관련되는 더 복잡한 쓰기와 말하기에서는 산출 양식에 따른 수행 결과가 훨씬 다를 수 있음을 언급하였다. 이러한 결과는 앞서 살펴본 것과 반대로, 제2언어 학습자는 쓰기 행위에 대한 부담으로 인해 오히려 구어에서 더 나은 수행을 보일 가능성도 있음을 의미한다. 다음으로, 실제 문어와 구어 발달을 비교한 연구들을 살펴보도록 한다.

Yu(2010)에서는 D값을 활용하여 25명의 고급 영어 학습자의 문어와 구어의 어휘 다양도를 비교하였다. 그 결과, 산출물에 대해 계획하고 구성할 기회가 더 많은 문어에서 어휘 다양도가 더 높을 것이라는 예상과 달리 문어와 구어의 어휘 다양도는 차이가 크지 않아 거의 동일한 수준으로 났으며 서로 유의미한 정적 상관관계를 나타내고 있었다. 이에 Yu(2010)에서는 시간적 압박은 어휘 다양도에 영향을 주지 않는다는 결론과 함께 구어 발화에서 학습자들이 뒤섞인 발화('scrambled' utterances)가 어휘 다양도를 높이는 요인이었을 수도 있음을 언급하였다. 그러나 대학생 영어 전공 학습자를 대상으로 준비 시간 제공 여부와 관련하여 문어와 구어의 복잡성과 정확성을 비교한 Ellis & Yuan(2005)에서는 어휘 복잡성은 준비 시간의 여부와 관계 없이 문어의 어휘 다양도가 훨씬 높다는 결론을 제시하기도 하였다.

Ellis & Yuan(2005)에서는 통사 복잡성과 정확성에 대해서도 비교를 하였는데 시간적 압박이 있을 경우 문어에서 더 복잡한 언어를 산출하지만 계획 시간이 주어질 경우 산출 양식에 따른 복잡성의 차이가 유의하지 않다는 점을 제시하면서 통사 복잡성은 허용되는 추가 시간의 영향을 받는다고 하였다. 또한 정확성 역시 준비 시간과 관계 없이 쓰기의 정확성이 더 높은 것으로 나타났다. 이에 대해 Ellis & Yuan(2005)에서는 숙달도가 낮은 학습자일수록 자동화가 어렵기 때문에 시간적 부담이 없을 때 통제된 처리를 통해 언어적 자원에 더 잘 접근하기 때문에 더 나은 수행을 보인다고 설명하였다. 과제 복잡성에 따른 중급 이탈리아어 학습자의 문어와 구어 복잡성 정확성을 비교한 Kuiken & Vedder(2011)에서는 두 가지 과제 모두 통사 복잡성과 어휘 다양도는 문어

에서 더 높지만 정확성은 구어에서 더 높다는 결과를 제시하며 실시간 산출이 이루어지는 구어에서는 '압박감'이 더 크며 이러한 압박감은 단순화된 언어로 이어진다고 설명하고 있다. 500시간 정도의 학습 경험이 있는 프랑스어 학습자의 문어와 구어 산출을 비교한 Granfeldt(2008)에서는 Kuiken & Vedder(2011)과 유사한 결과를 제시하고 있는데 이 연구에 따르면 정확성은 구어가 더 높으나 어휘 다양도는 문어에서 더 높았으며 통사 복잡성은 문어와 구어의 차이가 뚜렷하지 않았다고 하고 있다.

이 외에 문어와 구어의 발달을 종적으로 비교한 연구들을 살펴보면 또 다른 결과들을 찾아볼 수 있는데 소수의 학습자의 자료를 종적으로 살펴본 연구에서는 문어와 구어의 차이는 고정적인 것이 아니라 시기별, 학습자별로 차이가 있음을 밝히고 있다. 대표적으로 초급 청소년 영어 학습자(Common European Framework of Reference level A2~B1)인 일란성 쌍둥이의 8개월 간의 문어와 구어 발달을 비교한 Chan 외(2014)과 Chan 외(2015)에서는 학습자의 발달 단계가 다른 시점에서 나타나며 단계에 따라 문어와 구어의 복잡성이 서로 전환되기도 한다는 것을 발견하였다. 어휘 다양도(D값)와 어휘 크기(vocabulary size)[22]를 측정한 Chan et al.(2014)에서는 두 학습자 모두 구어에서 어휘 다양도는 더 낮지만 어휘 난이도는 낮지 않다는 결과를 도출하였고 구문 발달을 비교한 Chan 외(2015)에서는 한 명은 T-unit길이, T-unit당 내포절(subordinate clause), T-unit당 대등절 수 모두 구어에서 더 높게 나타난 반면, 다른 학습자는 초반에는 구어 복잡성이 더 높던 것에서 문어 복잡성이 더 높은 것으로 전환되며 두 학습자가 서로 다른 발달 패턴을 보이고 있다는 결과를 도출하였다. 이에 대해 Chan 외(2015)에서는 일란성 쌍둥이임에도 불구하고 대조적인 발달을 보여주는 것은 학습자들의 입력이 동일하더라도 각 학습자는 자신만의 발달 경로를 보인다는 결론을 내리기도 하였다.

연구마다 학습자의 숙달도나 과제, 측정 지표가 다르기 때문에 단순 비교를 하는 것에는 무리가 있지만 앞선 연구들을 종합해 보면 전반적으로 통사 복잡성과 어휘 다양도는 문어에서 더 높게 나타난다는 공통적인 결론을 보이며 정확성에 있어서는 서로 반대되는 결과들이 존재함을 알 수 있다. 이러한 결과들을 앞서 살펴본 문어와 구어의 산출 절차와 연결시켜볼 때, 숙달도가 높지 않은 본고의 학습자의 경우 복잡성의 측면에서는 언어 자원에 접근할 시간이 더 많은 문어에서 더 나은 수행을 보일 것이지만 정

[22] V-size(vocabulary size)는 다양한 빈도 대역(frequency bands)에 속하는 단어 수를 기반으로 학습자의 어휘 크기를 추정하기 위해 개발되었다. V-size의 기본 가정은 학습자의 숙달도가 높을수록 글에 빈도가 낮은 단어가 더 많다는 것으로 학습자의 예상 어휘 크기 또는 글에 나타난 어휘 난이도의 수준을 가늠할 수 있는 점수를 생성한다(Chan 외, 2014:57).

확성의 측면에서는 언어 산출의 실시간성에 대한 부담과 문자 생산의 부담의 크기에 따라 다른 결과를 보일 것으로 예상해볼 수 있다. 또한 시간의 흐름에 따라 문어와 구어의 발달을 비교할 때 이러한 차이가 고정적인 것이 아니며 학습자별, 시기별로 다르게 나타날 수도 있음을 예측해 볼 수 있다.

Ⅲ. 연구 방법

1. 연구 참여자

본 연구의 목적은 비슷한 배경을 가진 중국 배경 중도입국 청소년들이 공교육에 적응하는 과정에서 나타나는 한국어 발달 과정을 연구하는 것이다. 중국 배경 중도입국 청소년 중에는 이미 본국의 조선족 학교에서 조선어로 공부했던 학생들도 있는데 이러한 경우 한국어로 의사소통을 하거나 학업을 수행하는 데 있어 크게 어려움이 없다. 따라서 본고에서는 본국에 있을 때 공식적으로 조선어[23]를 배운 적이 없으며 입국한 뒤에 한국어를 제2언어로 학습하기 시작한 중도입국 청소년을 연구 대상으로 선정하였다. 일반적으로 중도입국 청소년은 한국에 입국한 이후 다문화 예비학교를 통해 기본적인 한국어와 한국 문화를 학습한 후 거주지 인근 학교에 서류를 제출하여 공교육으로 편입학을 진행한다. 본고에서는 공교육 입학 이후에 나타나는 한국어 발달을 살펴보고자 하였으므로 다음과 같은 연구 참여자 선정 기준을 설정하였다.

① 본국에서 공식적으로 조선어/한국어를 배운 적이 없는 학생
② 한국 입국 이후 예비학교 등 한국어 교육기관을 거쳐 중학교에 입학한 학생
③ 한국에서 공교육에 입학한 지 1년 미만인 학생
④ 연구 기간 동안 본국에 귀국할 계획이 없는 학생

연구자는 위의 연구 대상자 선정 기준을 통해 연구 참여자를 모집하였다. 연구가 시작된 시점인 2021년 10월에는 9명의 학습자가 연구에 참여하였으나 중도에 다른 학교로 전학을 가거나 연구 참여를 포기한 학습자, 연구에 불성실하게 참여한 학습자를 제외한 뒤 최종 3명의 자료를 분석 대상으로 선정하였다. 본 연구에서는 이들을 각각 A, B, C학습자로 명명하고자 한다. 연구 참여자의 기본 정보는 다음의 <표 1>과 같다.

[23] 본고에서 사용하는 '조선어(朝鲜语)'라는 명칭은 중국에 거주하고 있는 재중동포들이 사용하고 있는 민족어로서의 한국어를 가리킨다.

<표 1> 연구 참여자 정보

학습자	성별	연령[24]	부모 국적	입국시기	지역/민족	가정 내 사용언어
A	여	만(14)세 2007년 생	부: 중국 모: 중국	2019.12	흑룡강성/조선족	중국어/조선어
B	남	만(15)세 2006년 생	부: 중국 모: 중국	2019.9	길림성/조선족	중국어/조선어
C	남	만(16)세 2005년 생	부: 중국 모: 중국	2020.1	길림성/조선족	중국어/조선어

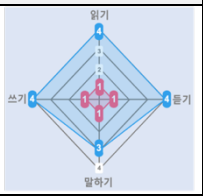
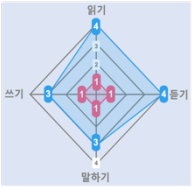
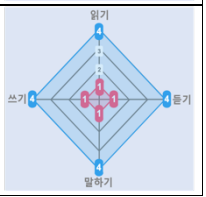
세 학습자는 한국에 입국한 시간에는 약간의 시간 차이가 있었지만[25] 모두 본국에서 공식적인 조선어 교육을 받은 적이 없는 학습자들로 입국 이후 2020년 5월부터 2021년 1월까지 경기도 소재의 다문화예비학교에서 한국어를 학습하였다. 이들의 한국어 수업에 사용한 교재는『중고등학생을 위한 표준 한국어 의사소통』이었으며 세 학습자 모두 1권부터 4권까지 학습을 완료하였다. 학습자들의 한국어 수준은 '한국어능력 진단-보정 시스템'[26]을 통해 측정되었는데 세 학습자가 예비학교에 입학할 때와 수료할 때 진행한 한국어능력 평가 결과는 다음의 <표 2>와 같다.

[24] 연구 참여자의 연령은 연구를 시작한 시점을 기준으로 계산된 것이다.

[25] 중도입국 청소년의 경우 일반 목적이나 학문 목적 한국어 학습자와 달리 입국 이후 상당 기간 동안 제대로 된 교육을 받지 못한 채 가정 내에 방치되는 경우가 많다. 이는 부모의 한국어 능력 부족이나 정보 부족으로 인해 적절한 교육 기관을 찾지 못하는 경우가 많고 또 학력 심의 서류 미비로 인해 입학 시간이 지연되는 경우 등 여러 원인이 존재한다. 그러므로 이들의 한국어 능력을 판단함에 있어 입국 시간보다는 한국어 교육을 받기 시작한 시점이 더 중요하게 작용할 것이다. 경기도가족여성연구원(2012)의 조사에 따르면 중도입국청소년이 한국에 와서 처음 3개월 간 한 일은 '집에 있었다(55.2%)', '학교에 다녔다(29.1%)', '지원센터(외국인센터/종교기관)에 다녔다.(9.0%)' 순으로 나타났으며 양계민·조혜영(2011)의 중도입국 청소년 실태조사에 따르면 중도입국 청소년은 한국 입국 이후 평균 5.53개월의 공백기가 생긴다고 나타났다.

[26] 세 학습자가 한국어를 학습한 다문화예비학교에서는 입학과 수료 시에 '한국어능력 진단-보정 시스템'을 이용하여 학습자들의 한국어 숙달도를 진단하고 성취도를 평가하였다. '한국어능력 진단-보정 시스템(https://ko-ls.or.kr)'은 한국어에 익숙하지 못한 초·중·고등학생들을 대상으로 한국어능력을 진단하고 보정학습 자료를 제공하는 시스템이다. '한국어능력 진단-보정 시스템'에서 판정하는 등급은 <한국어(KSL) 교육과정(2017)>의 단계별 성취 기준을 바탕으로 한다.

<표 2> 연구 참여자 진단 및 성취도 검사 결과

학습자	영역	성취 수준			
		회차별 성취 단계			전후 비교[27]
		1회차 (2020.5.20)	2회차 (2020.8.20)	3회차 (2020.12.22)	
A	읽기	1	3	4	
	듣기	1	3	4	
	말하기	1	3	3	
	쓰기	1	3	4	
B	읽기	1	3	4	
	듣기	1	3	4	
	말하기	1	3	3	
	쓰기	1	3	3	
C	읽기	1	3	4	
	듣기	1	3	4	
	말하기	1	1	4	
	쓰기	1	3	4	

위 표에서 확인할 수 있듯이 세 학습자는 2020년 5월 20일 입학 시에 듣기, 말하기, 읽기, 쓰기 네 영역에서 모두 가장 낮은 1단계에서 출발하여 2020년 12월 22일에는 네 영역 모두 3~4단계에 도달하여 비슷한 수준의 한국어 숙달도를 보였다. <한국어(KSL) 교육과정(2017)>에 따르면 학습자가 『중고등학생을 위한 표준 한국어 의사소통1~4』을 학습하고 도달할 수 있는 수준은 중급(3~4단계)으로 세 학습자는 다문화예비학교 수료 시 모두 교육과정에서 요구하는 성취기준에 도달하였으며 3단계와 4단계의 말하기와 쓰기의 성취기준은 다음의 <표 3>과 같다.[28]

<표 3> <한국어(KSL) 교육과정>의 3단계와 4단계 말하기·쓰기 성취기준

말하기	3단계	① 하루 일과를 시간적 순서에 따라 말할 수 있다. ② 도움을 요청하거나 제안하는 말을 할 수 있다. ③ 간단한 전화 대화를 할 수 있다. ④ 교실 표현 및 기본적인 학습 도구 표현을 말할 수 있다.

[27] 분홍색은 입학 시 진단 수준을 의미하고 파란색은 수료 시의 진단 수준을 의미한다.
[28] 기타 영역이나 기타 등급의 성취기준은 <한국어(KSL) 교육과정(2017)>에서 확인할 수 있다.

	4단계	① 상황에 맞게 적절한 대화를 할 수 있다. ② 사실에 관해 이유를 묻고 답할 수 있다. ③ 지나간 일과 앞으로 할 일에 대해 묻고 답할 수 있다. ④ 사고 도구 어휘와 범용 지식 어휘 표현을 사용하여 간단히 말로 설명할 수 있다.
쓰기	3단계	① 문장을 어려움 없이 받아쓸 수 있다. ② 일상생활에서 필요한 내용을 간단하게 기록할 수 있다. ③ 하루의 일과를 간단한 여러 개의 문장으로 쓸 수 있다. ④ 교실 표현 및 기본적인 학습 도구 표현을 사용하여 짧은 글을 쓸 수 있다.
	4단계	① 겪은 일이 잘 드러나게 짧은 글을 쓸 수 있다. ② 친숙한 주제에 대한 짧은 글을 읽고 요약하여 쓸 수 있다. ③ 자신과 가족을 소개하는 글을 쓸 수 있다. ④ 사고 도구 어휘와 범용 지식 어휘를 사용하여 한두 문단 수준의 글을 쓸 수 있다.

세 학습자는 다문화예비학교 수료 이후 2021년 3월 중학교에 입학하였다. 그러나 연구 참여자들이 입학한 중학교는 교내에 예비학급을 운영하여 입학한 학생의 한국어 능력이 원적학급 수업에 참여하기에 부족하다고 판단될 경우 매주 4일 간은 오전의 4교시 모두 한국어 수업을 듣고 오후에는 원적학급으로 돌아가며, 매주 1일만 정상적으로 원적학급의 수업에 참여한다. 그리고 학기마다 한국어 테스트를 진행하여 일정 수준에 도달하면 더 이상 예비학급의 한국어 교육을 받지 않고 완전히 원적학급 수업에만 참여한다. 세 학습자는 모두 입학 첫 학기인 2021년 3월~2021년 6월까지 예비학급에서 추가로 한국어를 학습하였다. 그러나 연구가 시작된 시점인 2021년 10월에는 세 학습자 모두 원적학급의 수업에만 참여하고 있었으며 학교생활 외에 별도의 한국어 교습을 받지 않고 있었다. 각 학습자의 언어 사용 상황을 살펴보면 다음과 같다.

A학습자의 부모님은 아버지는 조선족이고 어머니는 한족이다. 아버지는 조선어에 익숙하지만 학습자가 3살 때 한국에 온 관계로 학습자는 조선어를 배우지 못하였다. 학습자는 한국 입국 전까지 어머니와 중국에서 거주하다가 2019년에 어머니와 함께 한국에 입국하였다. 그러나 아버지는 타 지역에서 근무를 하기 때문에 월 1~2회 정도만 집에 돌아오며 가정 내에서 조선어를 거의 사용하지 않는다고 하였다. A학습자는 입학 이후 첫 학기에는 예비학급에서 대부분의 시간을 보내며 한국인 친구를 사귀지 못하고 같은 중국 배경 중도입국 청소년들과 관계를 맺고 있었으나 입학 1년 차가 지난 시점에서는 한국인 친구들을 사귀고 가깝게 지내기 시작하였다. A학습자는 전반적으로 원만한 교우 관계와 성실한 학습 태도로 학교 생활에 잘 적응하고 있었다.

B학습자는 부모님이 모두 조선족으로 조선어에 능숙하지만 학습자가 3살 때 한국에 오게 되면서 학습자는 2019년 한국 입국 전까지 부모님과 떨어져서 보호자와 함께 생활하였다. 보호자는 중국의 한족 부부로 조선어를 몰랐으며 학습자 역시 한족 학교에

서 공부하였다. 평소 가정 내 언어 사용을 살펴보면, B학습자는 아버지와 중국어로 대화를 하고 어머니와 조선어로 대화를 하는데 어머니와 하루 평균 5분 정도의 대화를 한다고 하였다. B학습자는 중학교 입학 초기부터 여러 나라의 친구들을 사귀면서 원만한 학교 생활을 하였으며 한국 친구 3~4명과 가까이 지냈다. 그러나 시간이 지나면서 개인적인 이유로 한국인 친구들과 사이가 멀어지게 되었다고 한다. B학습자는 입학 초기에는 학교생활에 적극적이었으며 교사와의 관계도 좋은 편이었으나 시간이 지남에 따라 학교 결석을 빈도가 높아지며 학교 밖의 친구들과 어울리는 시간이 많았다.

C학습자 역시 부모님 모두 조선족이며 조선어에 능숙한 편이다. C학습자는 출생 이후 6살까지 조부모님과 생활하다가 초등학교 생활 6년 간 타 지역에 있는 부모님과 함께 생활하였다. 그리고 학습자가 중학교 1학년에 입학할 때 부모님이 한국에 입국하면서 학습자는 다시 고향에 돌아가 조부모님과 생활하게 되었다. C학습자는 입국 이후에도 부모님이 타 지역에서 근무하는 관계로 공교육에 입학한 지 1년 정도 되는 시점까지 거의 혼자 지내다가 2022년 3월부터 외할머니와 함께 거주하기 시작하였다. C학습자의 언어 사용 상황을 살펴보면, 학습자는 부모님과 대화할 때에는 서로 중국어를 사용하고 외할머니와 대화할 때에는 외할머니가 조선어로 질문하고 학습자 본인은 중국어로 대답을 한다고 하였다. C학습자는 학교의 규칙을 어기거나 결석을 하는 등의 문제가 없이 원만한 학교생활을 이어 나가고 있었으나 관찰이 끝나는 시점까지 한국인 친구들과 관계를 맺지 않았고 입국 초기에 사귄 중국인 친구들과 관계를 이어오고 있었다. 학습자 본인과 부모님들을 통해 파악한 세 학습자의 입국 전과 입국 후 언어 사용 비율을 표로 정리하면 다음과 같다.

<표 4> 연구 참여자의 입국 전과 입국 후 언어 사용 비율

학습자	입국 전			입국 후		
	가정	학교	친구	가정	학교	친구
A	중국어	중국어	중국어	중국어(95%) 조선어(5%)	한국어	중국어(80%) 한국어(20%)
B	중국어	중국어	중국어	중국어(95%) 조선어(5%)	한국어	중국어
C	1~6살: 중국어(50%) 조선어(50%) 7~12살: 중국어(90%) 조선어(10%) 13~15살: 중국어(80%) 조선어(20%)	중국어	중국어	중국어(95%) 조선어(5%)	한국어	중국어

중국에 거주하고 있는 조선족은 초기에 이주하였을 때에는 동북 지역에 집거하면서 민족어 교육을 중시하고 잘 유지해 왔으나 1992년 한중수교 이후 대량의 조선족이 중국의 산동반도와 같은 연해도시나 베이징 상하이와 같은 대도시로 이동하면서 그 자녀들은 중국어를 제1언어로 학습하고 사용하며 한족학교에 입학하는 경우가 많아지게 되었다. 또한 이 외에도 조선족과 한족의 통혼으로 인해 조선어를 학습하지 않는 경우도 점점 늘어나고 있다(최미화, 2017; 方香玉·甘善權, 2022). 본고에서 연구 대상으로 삼은 세 명의 학습자 역시 호적 상으로는 '조선족'으로 표기되어 있지만 여러 가지 이유로 인해 조선어를 제대로 학습한 적이 없으며 조선족 학교에 다닌 적이 없는 학생들로 한국 입국 이후에 한국어를 학습하기 시작하였다. 앞에서 언급하였듯이 본국에 있을 때 조선족 학교에서 교육을 받았던 학생들은 한국어를 사용함에 있어 크게 어려움이 없다. 따라서 세 학습자는 한국어 사용에 어려움을 겪는 통상적인 중국 배경의 중도입국 청소년 학습자 유형에 해당한다고 볼 수 있다.

2. 연구 자료

언어 연구를 위한 자료 수집에서 과제가 매우 중요한 변인으로 작용한다는 점은 여러 연구들을 통해 밝혀진 바 있다. 그러므로 학습자의 언어 발달을 제대로 살펴보기 위해서는 학습자의 언어 능력을 충분히 드러낼 수 있는 과제를 부여하는 것이 매우 중요하다. Skehan(1998)에서는 인간의 인지적 자원이 한정적이라는 입장에서 과제 수행 자체에 많은 인지 자원이 투입되면 목표 언어 항목에 대한 주의가 낮아진다고 보았다. 즉 과제 수행 과정에 주의력, 기억력 등의 인지 능력이 많이 필요하게 되면 L2 발화의 유창성, 정확성, 복잡성에 영향을 준다는 것이다. 또한 Vanpatten(1990:288)에서는 인간의 인지 작용의 특성상 언어의 형태보다는 내용에 먼저 주의를 기울이게 되기 때문에, 내용이 복잡해지면 형태적으로는 더 단순한 언어를 사용하게 된다고 하였다. 즉, 학습자의 수준에 맞지 않고 지나치게 어려운 과제를 설정하는 것은 학습자의 언어 산출에 방해가 될 수 있다. 이 때문에 언어 자료의 수집에 있어 과제의 복잡성을 적절히 통제하는 것은 양질의 언어 산출을 이끌어 내는 데 중요한 요소로 작용하게 된다. 그렇다면 과제의 복잡성은 어떤 요인들에 의해 결정이 되는 것인지 알아볼 필요가 있다.

Skehan & Foster(2001:195)에서는 학습자의 수행에 영향을 미치는 과제의 요인들을 기호적 복잡성(code complexity), 인지적 복잡성(cognitive complexity), 의사소통 부담(communicative stress)으로 범주화하여 제시하였다. 여기에서 기호적 복잡성은

언어적 복잡성과 다양성, 어휘량과 어휘 다양성을 포함하고 인지적 복잡성은 인지적 친숙도와 인지 처리를 포함하는데 인지적 친숙도는 화제에 대한 친숙도, 담화 장르에 대한 친숙도, 과제에 대한 친숙도를 가리키고 인지 처리는 정보 조직, 인지 처리, '계산(computation)'의 양, 제공된 정보의 명확성과 충분성 등을 가리킨다. 의사소통 부담은 시간 압박(time pressure), 참여자 수나 텍스트 길이와 같은 규모(scale), 산출 양식(modality), 과제수행 부담(stakes), 과제 진행에 대한 조절 기회(opportunity for control) 등을 포함한다. 아래 요인별로 자세히 살펴보도록 한다.

먼저, 기호적 복잡성은 과제에서 필요로 하는 구문이나 어휘의 수준으로 학습자의 언어 수준과 관련한 복잡성이다. 그러므로 과제에서 필요로 하는 문법 구조나 시제, 어휘의 난이도와 학습자의 실제 언어 수준을 고려할 필요가 있다. <표 2>에서 확인하였듯이 연구에 참여한 학습자들은 공교육 진입 전 『중고등학생을 위한 표준 한국어 의사소통1~4』까지의 학습을 모두 마치고 진단평가에서 3~4등급의 수준에 도달하였다. 따라서 하루 일과나 자신이 겪은 일, 자신이나 가족을 소개하는 글을 쓰는 정도의 과제를 수행할 수 있는 수준이다. 그러므로 쓰기나 말하기의 내용은 학습자 자신이나 주변 인물, 일상 생활이나 본인이 경험한 일을 중심으로 제시하는 것이 적절할 것이다.

다음, 인지적 복잡성은 주제, 담화 장르, 과제 수행 방법에 대한 친숙도와 관련된 것이다. 즉 내용적으로, 형태적으로 학습자에게 친숙한 과제이거나 과제 수행 과정에서 많은 계산(computation)을 필요로 하지 않는 과제일수록 인지적 부담을 적게 느낀다는 것이다. 예를 들면 수수께끼를 알아맞히는 것보다 간단한 이야기를 서술하는 것이 더 적은 주의력을 필요로 하며 여행에 대한 설명의 경우 시작, 중간, 끝이 분명하고 참고할 수 있는 서사적 단서가 있기 때문에 인지적 복잡성이 상대적으로 낮다. 또 학습자가 과제 완성에 필요한 지식을 이미 보유하고 있는지 여부 역시 인지적 복잡성에 영향을 미치는데 과제가 학습자 자신이나 자신의 삶에 대해 이야기하도록 요구하는 것은 상상된 사람이나 상황에 대한 견해를 표현하는 것보다 난이도가 낮다. 그리고 과제를 수행하기 위해 제공되는 정보 역시 명확하고 충분한 것이 과제의 난이도를 낮춰줄 수 있다(Skehan & Foster, 2001:196). 이는 쓰기와 말하기 과제에서는 주제와 장르, 그리고 지시문으로 귀결될 수 있다.

쓰기나 말하기에서 주제는 언어 산출에 가장 큰 영향을 미치는 요인으로 특정한 배경지식을 필요로 하는 주제의 경우 개인의 언어 능력과 무관하게 산출의 양과 수준에 영향을 줄 수 있다. Cooper(1984)에서는 쓰기 주제가 다르면 쓰기 과제의 난이도도 다르기 때문에 학생들의 쓰기 점수는 쓰기 주제에 따라 상당한 차이를 보인다고 하였으

며 서수현(2008)에서는 쓰기 과제에 적합한 좋은 주제는 학생의 경험 내에서 의미 있고 뚜렷한 목적을 수반한 것이어야 하며 따라서 필자의 지식 차원에 대한 공통적인 특질을 고려한, 추상적이고 이론적인 것보다 구체적인 것이 더 바람직하다고 주장하였다. 이는 말하기에서도 마찬가지인데 노아실(2007)에서는 주제에 대한 관심도가 높을수록 휴지나 망설임, 자기 수정의 저해 요인은 낮아지는 반면, 발화의 양, 담화표지, 의사소통 전략 등은 높아진다고 보고하였고 진대연(2006)에서는 쓰기의 주제에 대한 친밀도가 어휘 다양도에 직접적인 영향을 미친다는 점을 주장하며 이는 학습자가 자신의 평소 경험이나 생각을 표현할 경우에 글을 쓰는데 부담을 덜 느끼기 때문인 것으로 해석하고 있다.[29] 그러므로 쓰기와 말하기의 주제는 학습자들의 공통적인 지식 차원과 경험을 고려하여 친숙한 주제로 선정할 필요가 있다. 본고의 학습자들은 모두 중학생이며 중국에서 생활한 적이 있고 현재 한국에서 중학교를 다니고 있는다는 공통점이 있다. 그러므로 쓰기와 말하기의 주제는 이들의 관심사나 경험을 고려한 학교, 친구, 부모님, 게임, 선물 등과 내용을 선정하였다.

장르 역시 언어 산출에 영향을 주는 주된 요인으로, 그동안 학문적 맥락과 비학문적 맥락의 장르 연구에서 널리 사용되어 온 장르는 서사(narrative)와 설득(expository) 장르이다(Ortega, 2015:88). 장르와 언어 산출에 대한 연구들을 살펴보면 학습자들이 기술적(descriptive) 글쓰기에서 더 긴 절을 생성하고 설득적(persuasive) 글쓰기에서 T-unit당 더 많은 절을 생성할 수 있다(Beers & Nagy, 2011)거나 서사적 (narrative)글쓰기보다 논증적(argumentative) 글쓰기에서 문장당 절 수가 더 높게 나타난다(Lu, 2011)는 결과들을 찾아볼 수 있다. 따라서 학습자들의 통사 복잡성을 이끌어내기 위해서는 설득적, 논증적 장르가 더 적합한 것으로 볼 수 있다. 그러나 논증적 글쓰기와 서사적 글쓰기는 쓰기의 목적이나 담화 기능이 다르며 필요한 배경지식이나 수사학적 구조 역시 차이가 있다. 이로 인해 두 장르는 글쓰기의 난이도에 있어서도 차이가 있는데, 서사적 글쓰기는 일반적으로 화자의 경험을 배경지식으로 하며 상대적으로 간단한 문장 구조를 사용하여 이야기의 배경이나 등장인물, 줄거리를 시간의 흐름에 따라 서술하는 글(김수미, 2011:31)로 친숙도가 높고 비교적 쉽게 쓸 수 있는 반면, 설득적·논증적 장르의 글은 어떤 사실이나 상황에 대한 이슈, 그 이슈에 대한 배경지식, 필자의 주장, 그 주장을 뒷받침할 관련 근거 및 주장과 반박을 포함하는 글(Oshima & Hogue,

[29] 진대연(2006)에서는 일본어권과 중국어권 학습자들을 대상으로 쓰기에 나타난 어휘 능력을 살폈다. 그 결과 한국의 경제와 관련된 글보다 한일 관계와 관련될 글의 어휘 다양도가 더 높게 나타났으며 이는 통계적으로 유의미한 차이임을 확인하였다.

2009:142-146)로 학습자들이 충분한 논거를 가지고 논리적으로 글을 전개해야 한다. 따라서 설득적·논증적 장르의 경우 학습자 개인의 배경지식과 논리성에 대한 요구가 비교적 높고 더 복잡한 수사학적 구조를 필요로 하게 되므로 글쓰기의 난이도가 높아 고급 단계의 학습자에게 많이 사용된다. 그러므로 본 연구의 학습자의 한국어 수준과 연령을 고려하였을 때 설득적·논증적 장르의 글쓰기에는 어려움이 있다고 판단하여 상대적으로 친숙한 기술적·서사적 장르를 선정하였다.[30]

지시문 역시 쓰기와 말하기의 산출에 영향을 미치는 요인이다. Powers & Fowles(1998:3)에서는 쓰기 지시문을 쓰기 과제를 구성하는 하나의 요소로 보았다. 이 연구에는 과제(task)와 지시문(prompt), 화제(topic)를 구분하고, 쓰기 지시문과 화제를 과제에 포함시키고 있다. 주제가 '무엇'에 대한 것이라면 지시는 '어떻게'에 대한 것으로 과제의 범위를 제한하거나 과제 수행을 유발하고 과제 수행을 안내하는 역할을 한다(김효은·김가연, 2018). 과제에서 지시문은 과제의 방향과 범위를 제공해준다는 점에서 인지 처리 과정에서의 명확하고 충분한 정보를 제공함으로써 쓰기와 말하기의 수행에 도움을 줄 수 있다. 지시문은 단순 지시문(bare prompt), 규정 지시문(framed prompt), 자료 기반 지시문(text-based or reading based prompt)으로 나눌 수 있는데(Kroll & Reid, 1994:233-234) 단순 지시문은 수행할 과제를 단순하고 명료하게 한 두 문장으로 지시하는 것이고, 규정 지시문은 특정 상황 맥락을 제시하면서 그 배경에서의 과제 수행을 좀 더 구체적으로 지시하는 것이며, 자료 기반 지시문은 한 단락 이상의 긴 읽기 자료를 제시한 뒤 자료의 논지를 파악하고 의견을 제시하도록 하는 것이다.[31] 이 중에서 규정 지시문은 일반적으로 특정한 상황과 맥락을 제시하는 글과 함께 학습자의 입장이나 의견을 요구하는 질문을 포함하는데 논증적 장르에서 많이 제공된다. 또한 긴 읽기 자료를 읽고 논지를 파악하는 것 역시 서사적 글쓰기에 필요하지 않으며 한국어 수준이 높지 않은 학습자에게 읽기 부담을 줄 수 있다. 특히 본고에서 사용한 주제는 학습자들의 개인적인 경험과 일상적인 내용을 다루고 있으므로 상황과 맥락에 대한 자료를 제시하기가 어렵다. 그러나 너무 단순한 지시문은 학습자의 말하기나 쓰기 수행을 유발하기 힘들며 쓰기와 말하기의 방향이나 범위를 제공하기 힘들다는 단점이 있다. 이에 본고에서 상황 맥락이나 읽기 자료가 없는 단순 지시문을 사용하되 주제와 관련하여 다룰 수 있는 내용을 질문의 형식으로 제시함으로써 학습자들이 내용

[30] 실제로 중도입국 중학생 6명을 대상으로 '학생은 화장을 해야 하는 것인가'와 'TV의 좋은 점과 나쁜 점'을 주제로 글쓰기와 말하기 실험을 진행하였는데 학습자들 30-50자 정도의 간단한 대답 정도만 써서 제출하여 복잡성이나 정확성을 분석하기 어려운 수준이었다.

[31] 세 가지 지시문의 구체적인 설명과 예시문은 Kroll & Reid(1994:233-234)에서 찾을 수 있다.

을 구상하는 데 방향을 제시하고자 하였다.

마지막으로, 의사소통 부담은 과제 수행 시간의 제한이나 상호작용의 실시간성, 과제 수행의 부담과 관련이 있는 것으로 주어진 시간 안에 과제를 완성하도록 하거나 과제 수행 시간을 매우 짧게 제한할 경우 인지적 부담에 높아지게 되는데 특히 구어 과제에서는 상호작용 과제인지 독백과제인지에 따라 통사 복잡성이 달리 나타나기도 한다. 상호작용 과제의 경우 좀 더 자연스러운 발화를 수집할 수 있다는 장점이 있지만 실시간 소통이 이루어져야 한다는 점에서 학습자의 의사소통 부담이 크게 된다. 반면 독백과제는 실시간 상호작용의 부담이 발생하지 않는다는 점에서 한국어 수준이 높지 않은 학습자에게 적절한 과제로 볼 수 있다. 또한 독백과제(monologic tasks)는 대화식 수행(dialogic performances)보다 유창성이 덜 나타나지만 통사적으로 더 길고 복잡한 발화가 나타난다(Ferrari, 2012)는 점과 독백 과제에서 산출된 발화가 대화식 수행에 비해 숙달도와 더 뚜렷한 상관관계를 보인다는 점(Gilabert 외, 2011)들을 고려하여 본고에서는 구어 과제로 독백 과제를 사용하였다. 과제 수행 시간 역시 언어 산출에 영향을 주는 요인인데 관련 선행연구(Crookes, 1989; Foster & Skehan, 1996; Ortega, 1999)들을 살펴보면 사전 계획 시간을 부여할 때 언어 복잡성이 증가하는 것으로 나타나며 특히 Skehan & Foster(1997)는 개인적 정보 교환(personal information exchange) 과제와 서사적(narrative) 과제에서 사전 계획 시간의 부여는 정확성을 향상시킬 수 있다는 결과를 보여주고 있다. 따라서 본 연구에서는 연구 참여자 개인이 준비 시간과 산출 시간을 통제할 수 있도록 과제 수행에 시간 제한을 두지 않는다. 이상의 논의들을 참고하여 본고에서는 문어와 구어 산출 과제 선정에 다음과 같은 기준을 선정하였다.

① 학습자의 언어 수준을 고려하여 일상이나 개인적 경험을 서술하는 과제로 선정한다.
② 학습자들의 공통적인 지식 차원과 경험을 고려한 친숙한 주제를 선정한다.
③ 과제 수행에 방향 제시의 역할을 하는 질문 형식의 지시문을 제시한다.
④ 문어 산출 과제는 작문 쓰기, 구어 산출 과제는 주제에 대한 독백을 실시한다.
⑤ 문어와 구어 과제 모두 과제 수행의 준비/산출 시간에 제한을 두지 않는다.

본고에서는 문어와 구어 산출을 비교하기 위해, 문어와 구어 산출의 전체 주제는 동일하게 하되 주제의 순서를 조정하여 같은 주제 사이에 최소 2주에서 최대 6주 간의 간격을 둘 수 있도록 하였다. 언어 산출 과제에 사용된 주제들은 『중고등 학생을 위한 표준 한국어 의사소통』 1권~4권의 단원 주제와 국립국어원의 중도입국 청소년용 말뭉치 수집에 사용된 주제를 참고하여 선정하였다. 또한 세 학습자에게 동일한 시점에 동

일한 주제를 부여하여 학습자 간에 과제 차이가 없도록 하였다. 다음의 <표 5>는 문어와 구어 산출에 사용된 과제의 예시이고 <표 6>은 언어 산출 과제의 전체 주제 목록이다.

<표 5> 문어와 구어 산출 과제 예시

받고 싶은 선물 어떤 선물을 받고 싶어요? 왜 그 선물을 받고 싶어요? 지금까지 받은 선물 중에 가장 좋은 선물이 뭐예요? 선물에 대해서 글을 쓰세요/말해 보세요.

<표 6> 문어와 구어 산출 주제

회차	문어	구어	회차	문어	구어
1	우리 가족	나의 취미	19	중국학교	거짓말
2	내가 좋아하는 계절	나의 친구	20	나의 고민	스트레스
3	최근에 있었던 일	주말이야기	21	거짓말	행복한 일과 슬픈 일
4	나의 취미	우리 가족	22	학교에서 있었던 일	나의 고민
5	나의 친구	내가 좋아하는 계절	23	행복한 일과 슬픈 일	후회하는 일
6	주말이야기	최근에 있었던 일	24	잊지 못할 추억	학교에서 있었던 일
7	받고 싶은 선물	나의 식생활	25	후회하는 일	칭찬
8	내가 가장 아끼는 물건	받고 싶은 선물	26	한국의 학교 생활	잊지 못할 추억
9	나의 식생활	나의 공부 방법	27	칭찬	부모님께 혼난 일
10	바꾸고 싶은 생활습관	내가 가장 아끼는 물건	28	기억에 남는 사람	한국의 학교 생활
11	나의 공부 방법	나의/친구의 생일	29	부모님께 혼난 일	다치거나 아팠던 일
12	여행	바꾸고 싶은 생활습관	30	할머니/할아버지	기억에 남는 사람
13	나의/친구의 생일	나의 한국 생활	31	다치거나 아팠던 일	친구와 싸운 일
14	명절이야기	여행	32	좋아/싫어하는 선생님	할머니/할아버지
15	나의 한국 생활	나의 미래 모습	33	친구와 싸운 일	닮고 싶은 사람
16	보고싶은 사람	명절이야기	34	게임	좋아/싫어하는 선생님
17	나의 미래 모습	중국학교	35	닮고 싶은 사람	영화/책
18	스트레스	보고싶은 사람	36	영화/책	게임

위에서 설정한 과제를 통해 학습자들은 1년 6개월 간 약 2주 간격으로 36차례의 쓰기와 말하기 과제를 수행하였으며 연구자는 학습자별로 72건(작문 자료 36편, 음성 자료 36건), 총 216건의 언어 자료를 수집하였다. 본 연구는 동일한 학습자를 대상으로 반복적으로 자료를 수집해야 하는 연구로 학교생활 외의 추가적인 한국어 교육은 자연스러운 언어 발달에 영향을 줄 수 있다. 학습자들은 자료 수집 기간 동안 기본적인 학교 생활 이외에 한국어를 포함한 어떤 추가적인 교육도 받지 않았으며 연구자 역시 자료 수집 이후 학습자가 산출한 자료에 대해 피드백을 제공하지 않았다. 또한 수집한 자료 외에 학습자들에게 기타 과제를 부여하거나 교육을 진행하지 않았으므로 사전 학습에 의해 특정 주제의 언어 산출이 영향을 받았을 가능성은 배제할 수 있을 것이다. 또한 연구자는 개강이나 방학과 같은 학교 일정에 관계 없이 일정한 간격으로 자료를 수집하였다. 본고에서 수집한 자료의 구체적인 정보는 다음의 <표 7>과 같다.

<표 7> 수집된 언어 자료의 크기

학습자	문어			구어		
	텍스트	어절 수	문장 수	텍스트	어절 수	AS-unit 수
A	36	6,000	900	36	4,833	701
B	36	5,721	765	36	4,999	718
C	36	4,201	681	36	5,589	908
합계	108	15,922	2,346	108	15,421	2,327

3. 자료 분석 기준

2장에서 살펴보았듯이 복잡성과 정확성 측정에 사용되는 측정 지표는 매우 다양하며 연구마다 사용하는 지표가 상이하다. 그러므로 정확한 분석을 위해, 본 연구에 알맞는 분석 지표와 분석 기준을 설정할 필요가 있다. 아래에서는 자료 분석을 위한 분석의 단위를 설정하고 복잡성과 정확성 측정을 위한 분석 지표를 선정한 뒤 각 분석 지표에 대한 세부적인 분석 기준을 제시하도록 한다.[32]

3.1. 통사 복잡성 측정 지표 및 분석 기준

[32] Ellis & Barkhuizen(2005)에서는 복잡성, 정확성, 유창성의 분석 수행 과정을 1) 분석 단위 설정, 2) 분석 지표 설정, 3) 분석 지표를 바탕으로 측정, 4) 복잡성, 정확성, 유창성 계산 5) 분석 결과 해석의 5단계로 나누어서 제시하였다.

통사 복잡성 측정을 위해서는 우선적으로 측정 지표를 선정해야 한다. 선행연구에서 살펴보았듯이 그동안 한국어 통사 복잡성 관련 연구는 각 학습 단계의 통사적 발달 특징을 찾기 위한 연구와 함께 한국어의 특징에 알맞는 복잡성 측정 지표를 찾기 위한 연구들도 지속적으로 있어 왔다. 이러한 연구들은 어떤 지표가 한국어의 숙달도 차이를 유의미하게 구분해 주는지 또 어떤 지표가 숙달도의 차이를 구분하지 못하는지를 보여주었다. 그러나 통사 복잡성 측정 지표는 아직 명확히 합의된 바가 없으며 연구마다 각각 다른 지표를 사용하거나 같은 지표일지라도 그 기준이 상이한 경우도 많고 또 지나치게 많은 수의 지표를 중복 측정하는 경우들도 존재한다.

기존의 연구들에 사용되어 온 통사 복잡성 측정 지표를 정리하면 크게 단위 길이, 단위 확장의 양, 그리고 단위 확장의 유형의 3가지로 범주화 할 수 있다. 이는 모두 분석 단위를 설정하고 단위의 길이와 단위 안에서 나타난 접속 관계의 양, 그리고 접속 관계의 유형을 살펴보는 것이다. 본 연구에서는 선행연구의 결과들을 참고하여 각 측면에서 숙달도 구분에 유용하다는 결과를 얻은 1~2개의 지표를 선정하여 통사 복잡성의 발달을 측정하고자 한다.

3.1.1. 기본 분석 단위

복잡성이나 정확성 분석을 위해서는 기본적인 측정 단위를 설정해야 한다. 2장에서 제시하였듯이 본고에서는 문어 자료에서는 문장을 기본 분석 단위로 삼고 구어 자료에서는 AS-unit을 기본 분석 단위로 삼고자 한다.

문장을 기본 분석 단위로 설정하기 위해서는 문장에 대한 정의가 선행되어야 한다. 문장에 대한 학자들의 정의를 살펴보면 구본관 외(2015:209)에서는 '생각이나 감정을 말과 글로 표현할 때, 완결된 내용을 나타내는 최소의 독립적 형식 단위'로 정의하고 있고 임지룡 외(2005:237)에서도 이와 비슷하게 '생각이나 감정을 완결된 내용으로 표현하는 최소의 언어 형식'으로 정의하고 있다. 또한 문장을 구분하는 표지로는 주어와 서술어라는 기본 구성 요소와 문장이 끝났음을 나타내는 문장 부호가 있으며 한국어의 특성 상 종결 어미가 문장 종결의 표지가 되기도 한다. 그러므로 한국어에서 문장은 '주어와 서술어를 갖추고 있으며 문장 부호 혹은 종결 어미로 종결된, 생각이나 감정을 표현하는 독립적이고 완결된 의미 단위'로 정리할 수 있다. 이는 비교적 모호한 정의일 수 있으나 다양한 형태를 띠고 있는 학습자들의 언어 산출물을 분석하기 위해 이러한 정의를 사용하도록 한다. 실제로 수집한 학습자 자료를 살펴보면, 의미적으로는 문장이

맞으나 학습자의 의도에 의해 서술어가 생략된 언어 단위들을 찾아볼 수 있다. 이와 같이 언어 능력의 부족이 아닌 학습자의 의도에 의해 완결되지 않은 언어 단위에 대해서는 하나의 문장으로 인정하여 분석하도록 한다.[33] 예를 들면 다음과 같은 문장이다.

(1) 가. 예: 라면, 만두, 烙饼(√전병) 등. (AW9)
나. 사과에서 연필로 바꿔(√바뀌고), 연필은 가방으로 바꿔(√바뀌고), 가방은 돈으로 바꿔(√바뀌고)... (AW27)
다. 그리고 너를 봐라. 성적이도(√성적도) 안 좋고 매일 계속 핸드폰을 놀고. (AW29)

이어서 구어의 기본 분석 단위인 AS-unit의 판단 기준을 살펴보도록 한다. Foster 외(2000:365)에서는 AS-unit을 '하나의 독립절(independent clause) 혹은 하위절(sub clausal) 단위, 그리고 이와 결합된 내포절(subordinate clause)로 구성된 구성된 단일 화자의 발화'로 정의하고 있다.[34] 그러나 T-unit과 다른 점은 독립 하위절(independent sub-clausal) 단위를 허용한다는 것인데 독립 하위절은 '문장이 아닌(nonsentences)' 작은 발화까지 포함된다. 따라서 의사소통 과정 중에 일어나는 짧은 구나 대답 역시 하나의 AS-unit으로 분석된다. 따라서 다음과 같은 문장도 하나의 AS-unit으로 분석할 수 있다.

(2) |이거예요|이거 이거|(2AS-unit) (CS8)

다음, 대등절과 내포절에 대한 것이다. 대등절과 내포절에 대한 처리 방법은 T-unit과 동일한데 내포절은 모절과 같은 단위에 포함되어 하나의 AS-unit으로 분석되는 반면 대등절(coordinate clause)은 선행절과 후행절의 주어가 같은 경우와 다른 경우로 나뉜다. 이때 Foster 외(2000:365)에서는 후행절에서 선행절과 같은 주어가 생략되고 동작이 대등하게 연결된 절은 하나의 AS-unit으로 구분하지만 선행절과 후행절의 주어가 다를 경우 대등하게 연결되었을지라도 서로 다른 AS-unit으로 분석한다고 설명하고 있다.[35] 그러나 영어의 대등절은 'and'와 같은 등위접속사로 연결되는 두 문장이

[33] 표준국어대사전에서는 문장은 생각이나 감정을 말과 글로 표현할 때 완결된 내용을 나타내는 최소의 단위로 주어와 서술어를 갖추고 있는 것이 원칙이나 때로 이런 것이 생략될 수도 있음을 언급하며 글의 경우, '철수는 몇 살이니?', '세 살.', '정말?' 등을 예로 들고 있다.
[34] 분석 자료에서 AS-unit은 '| |'로 표시하고 절 사이의 경계는 '::'로 표시한다.
[35] T-unit에서도 대등하게 연결된 주절은 다른 T-unit으로 분석하기도 하지만 주어가 같고 후행절의 주어가 생략되어 대등하게 연결된 동사구는 하나의 T-unit으로 보기도 한다(Lennon,

지만 첨가어인 한국어의 경우 등위접속사의 역할을 하는 문법 요소는 '-고, -거나, -지만'과 같은 연결어미로, 두 개의 절을 연결해 준다는 점은 영어와 동일하지만 문장의 종결을 나타내는 종결 어미나 시제를 나타내는 선어말 어미가 후행절의 서술어에 한 번만 나타나는 등 선행하는 절과 후행하는 절이 통사적으로 관련을 맺고 있으므로 서로 다른 단위로 나누어질 수 없다(박지순·서세정, 2009:5). 또한 본고에서는 문어와 구어의 길이를 서로 비교한다는 점에서 문어에서 선행절과 후행절을 한 단위로 보고 구어에서는 선행절과 후행절을 분리하여 계산한다는 것은 적절하지 않다. 따라서 본고에서는 선행절과 후행절의 주어가 다른 대등절이라고 할지라도 하나의 AS-unit으로 분석하도록 한다. 그러므로 다음의 예문(3)과 예문(4)는 모두 하나의 AS-unit으로 분석된다.

(3) 가. | 집에 와서 :: 시간이 있으면 :: 요리를 해요.| (1AS-unit) (AS4)
나. | 나는 코로나 끝나구 :: 중국에 가서 :: 할아버지하고 할머니 볼 거예요.| (1AS-unit) (CS18)

(4) 가. | 저는 스트레스를 받을 때 가야금을 연습하거나 :: 책을 읽어요.| (1AS-unit) (AS20)
나. | 그 계란의 흰색 부분은 액체이고 :: | 노란색 부분은 고체예요.| (1AS-unit) (AS14)

그러나 AS-unit은 억양과 휴지를 고려하는데 선행절과 후행절의 주어와 관계 없이 절(동사구)의 시작이 하강 혹은 상승의 억양으로 나타나고 절의 끝에 0.5초 이상의 휴지가 발생하는 경우 각각 다른 AS-unit으로 구분한다. 이때 0.5초 이상의 휴지 길이는 특별한 기구의 측정 없이도 듣고 분명히 인지할 수 있는 길이이다. 따라서 다음의 예문(5)는 3개의 AS-unit으로 분석된다.

(5) | 영어 좋아하는 이유는 영어 재미도 있구(0.6) ::| (어~ 어 또 너무 쉽지 쉽지 않구) 너무 쉽지 않구(1.0) :: | 단어 외우는거 :: 좋아하기 때문에 영어를 좋아해요.| (3AS-unit) (CS9)

이 외에도 AS-unit은 의미 없는 음절(반복, 간투사, 재구성 이전의 발화)을 모두 제외한 후 의미 단위로 분석한다. 구어, 특히 제2언어로서의 구어 자료에서는 잘못된 시작이나 중복, 수정이 자주 나타나는데 이때, 잘못된 시작은 발화가 시작되었으나 완전히 버려지거나 혹은 다른 방식으로 다시 표현하는 것을 가리킨다. 만약 하나의 AS-unit이 메시지가 버려지기 이전에 산출된 것이라면 AS-unit의 표준에 부합되는 발화 부분

1990).

은 AS-unit으로 간주하고 다른 부분은 잘못된 시작으로 기록한다. 자기 수정이 나타날 경우, 마지막 버전을 AS-unit으로 간주하며 그 이전의 버전은 제외한다.

(6) 가. |(고민이 있으면...) 예를 들면 내가 돈이 없으면 고민 나와요.|(1AS-unit) (BS22)
나. |점심은 (그...) 간식을 (많이 많이) 많이 먹어서 :: (점심을 음) 점심을 안 먹어요.|(1AS-unit) (CS12)

3.1.2. 단위 길이

길이에 따른 복잡성의 측정은 T-unit, AS-unit, 문장과 같은 단위의 평균 길이를 측정하는 방법인데 이는 통사적으로 숙달된 학습자는 숙달도가 낮은 학습자에 비해 길이가 비교적 긴 문장을 산출할 것이라는 일반적인 판단에 근거한다(서세정, 2009:39). 산출 길이에 따른 복잡성은 MLU, 평균 발화 길이라고 불리기도 하는데 이는 각 발화에서 쓰인 어휘들의 평균 비율을 형태소, 단어, 어절, 절 등의 단위에서 살펴보는 방법이다(김명광, 2016:454). 한국어 교육의 연구에서는 일반적으로 문장당 단어 수, 형태소 수로 문장의 길이를 측정하거나 절당 단어 수, 절당 형태소 수로 절의 길이를 측정하는 방법을 사용하고 있다. 기존의 선행연구들을 살펴보면 단위당 단어 수는 박지순·서세정(2009), 남주연 외(2016), 주우동(2019)에서 숙달도의 차이를 구분할 수 있음으로 밝혀졌고 단위당 어절 수는 서세정(2009), 남주연 외(2016)에서 숙달도에 따라 증가한다고 하였으며 단위당 형태소 수는 서세정(2009), 음승민(2012), 서신웨(2022) 등 연구에서 숙달도에 따라 증가함을 확인하였다. 즉 여러 연구마다 다른 지표를 사용하여 길이를 측정하고 있으나 대부분 비슷한 결과를 보이고 있었다. 이에 주우동(2019:67)에서는 한국어의 특성 상 문장당 어절 수, 어휘 수, 형태소 수는 대동소이하고 문장당 어절 수가 많으면 어휘 수, 형태소 수가 많기 때문에 유사 지표를 불필요하게 많이 사용할 필요가 없다고 하기도 하였다. 그러므로 여러 가지 지표로 중복 측정을 하는 것보다 연구의 특성에 알맞는 지표를 적용하는 것이 중요하다.

주요하게 사용되는 세 가지 지표인 단어 수, 어절 수, 형태소 수를 살펴보면 '단어'는 그 정의와 범위에 대해 논쟁의 여지가 있으며 연구자마다 측정 방법이 상이하여 그 비교가 어렵다는 단점이 있고[36] 단위당 어절 수는 서세정(2009), 남주연 외(2016)를 포함

[36] 음승민(2012), 남주연 외(2016), 최예슬(2018)에서는 조사를 단어에 포함한 반면 이복자(2015)에서는 조사를 포함하지 않고 있다.

한 여러 연구에서 숙달도 사이에서 유의미한 차이를 보여 준다고 하였으나 어절을 기준으로 할 경우 띄어쓰기가 기준이 된다는 점에서 학습자의 조사 사용 여부나 다양한 어미들의 복합적인 사용을 섬세하게 측정하지 못하게 된다. 특히 본고의 연구 대상자는 초·중급 수준의 학습자라는 점에서 학습자들의 조사 사용이 익숙하지 않은 경우가 많다. 따라서 학습자들의 문장 길이를 더 정교하게 측정하기 위해 단위당 형태소 수를 통해 길이를 측정하도록 한다.

3.1.3. 단위 확장의 양

문장의 길이도 통사 복잡성을 대표하는 하나의 중요한 척도이지만 그 문장이 얼마나 복잡한 구조로 이루어졌는지를 살펴보는 것도 통사 복잡성에 대한 중요한 정보를 제공해 준다. 문장은 기본적으로 주어와 서술어로 이루어진다. 이때, 주어와 서술어의 관계가 몇 번 나타나는지에 따라 단문과 복문으로 나눌 수 있는데 주어와 서술어의 관계가 한 번만 나타나는 단문(=홑문장, simple sentence)에 비해 주어와 서술어의 관계가 두 번 이상 나타나는 복문(=겹문장, complex sentence)이 구조적으로 더 복잡하다고 보는 것이다. 같은 맥락으로 한 문장 안에서 주어와 서술어의 관계를 여러 번 사용하여 구조적으로 더 복잡한 문장을 산출할 수록 통사 복잡성이 더 높다고 보는 것이다. 따라서 단위 확장의 양은 하나의 문장에 나타난 주어와 서술어의 구조를 측정하게 된다. 앞서 2장에서 제시한 바와 같이 본고에서는 단위당 절 수로 단위 확장의 양을 측정하도록 한다.

그러나 단위당 절 수를 측정함에 있어서는 절에 대한 정의가 연구마다 차이가 있는데 절을 문장보다 작은 단위로 보는 입장과 문장이 절보다 크거나 같다고 보는 입장으로 나뉜다. 먼저 절을 문장보다 작은 단위로 보는 입장은 박지순·서세정(2009), 음승민(2012), 주월랑(2018) 등이 있다. 박지순·서세정(2009), 서세정(2009)에서는 절은 '주어와 술어를 갖추었으나 독립하지 않고 다른 문장의 한 성분으로 쓰이는 단위'로 정의하고 명사절, 관형사절, 부사절, 인용절, 서술절 및 대등절의 선행절만 측정하였다. 그리고 음승민(2012)에서는 선행절과 후행절, 내포절과 모절과 같이 모든 절을 세는 방법을 사용하였지만 복문에서만 절을 측정하여 단문은 절로 계산하지 않고 있었으며 주월랑(2018)에서도 음승민(2012)에서와 마찬가지로 '절은 주어와 술어의 관계를 가지되 문장보다 작은 단위'로 정의하고 복문만을 대상으로 선행절과 후행절, 내포절과 모절의 수를 측정하였다. 반면, 남주연(2015)과 이복자(2015), 남주연 외(2016)에서는 문장이 절보다 크거나 같다고 보는 입장을 취하여 내포절과 모절, 선행절과 후행절뿐만

아니라 독립된 단문 역시 하나의 절로 처리하였다.

우선, 박지순·서세정(2009), 서세정(2009)에서와 같이 내포절과 선행절만을 측정할 경우 문장당 절 수는 결국 문장당 대등절·내포절 수와 같게 된다. 본고에서는 문장 확장의 유형에서 대등절 수와 내포절 수로 따로 살펴볼 것이기 때문에 이렇게 측정할 경우 중복 측정을 하게 된다. 또한 절을 '주어와 서술어로 이루어진 구성'이라고 할 때 단문의 주어와 서술어 조합은 절로 보지 않고 복문에서 선행절, 후행절, 내포절, 모절 모두를 절로 보는 것은 일관성이 없다. 따라서 본고에서는 남주연(2015), 이복자(2015), 남주연 외(2016)에서와 같이 문장이 절보다 크거나 같다는 입장에 따라 독립된 단문 역시 절로 처리하도록 한다. 아래의 예문(7가)와 같이 주어와 서술어의 관계가 한 번만 나타나는 경우, 단문일지라도 하나의 절로 보고 (7나)와 같이 대등절로 구성된 문장은 선행절과 후행절을 모두 절로 보고 2개의 절로 계산한다. 그리고 (7다)와 같이 둘 이상의 절로 구성된 문장의 경우 내포절과 모절 모두 절로 보고 3개의 절로 계산한다.

(7) 가. 하지만 소녀는 이 영화의 마지막 말처럼 자신의 행동을 후회하지 않아요. (AW26)
나. 매일 피곤하고 :: 힘들어요. (BW25)
다. 얘가 말하면 :: 주변에 있는 :: 친구 다 같이 웃어 줘요. (CW28)

그러므로 본고에서 정의하는 단위당 절 수는 하나의 텍스트에 나타난 총 절의 개수를 해당 텍스트에 나타난 총 단위 수로 나눈 것으로 문어 자료에서는 문장당 절의 개수가 될 것이고 구어 자료에서는 AS-unit당 절의 개수가 될 것이다.

3.1.4. 단위 확장의 유형

학습자의 통사 복잡성 발달은 얼마나 많은 단위 확장을 이루어 내는지도 매우 중요한 지표가 되겠지만 구체적으로 어떤 문법 항목을 사용하여 구조적으로 복잡한 언어를 산출하는지에 대한 검토도 함께 이루어질 필요가 있다.

한국어의 문장은 절로 구성되는데 하나의 절로 이루어진 문장은 단문(=홑문장, simple sentence)이라고 하고 둘 이상의 절로 이루어진 문장을 복문(=겹문장, complex sentence)이라고 한다(구본관 외, 2015:254). 그리고 복문은 다시 접속문(이어진 문장)과 포유문(안은 문장)으로 나눌 수 있는데 접속문은 선행절(접속절)과 후행절이 의미상 각각 독립적이며 문법적으로도 서로 연관되지 않는 채 별개의 절로 대등하게 독립해 있는 것을 가리키고 포유문은 한 절이 다른 절을 문장 성분의 일부로서 안고 있는 문장

을 가리킨다(구본관 외, 2015). 아래 본 연구에서의 접속문과 포유문에 대한 구체적인 분석 기준을 제시하도록 한다.

(1) 접속문

접속문(이어진 문장)은 대등적으로 이어진 문장도 있고 종속적으로 이어진 문장도 있다. 이에 연구자에 따라 대등접속과 종속접속으로 나누는 경우도 있다. 그러나 종속적으로 이어진 문장은 부사절을 안은 문장은 그 차이가 뚜렷하지 않다는 점에서 외국어로서의 한국어 문법을 다루고 있는 국립국어원(2005)의 『외국인을 위한 한국어 문법』에서는 접속문에서 대등접속과 종속접속(부사절)을 구분하지 않고 모두 연결어미로 칭하고 있다.

이에 대해 구본관 외(2015)에서 역시 같은 입장을 확인할 수 있는데 구본관 외(2015:259)에서는 대등적으로 이어진 문장과 종속적으로 이어진 문장을 구별할 통사적 현상은 상당히 많지만 종속적으로 이어진 문장과 부사절을 구별할 통사적 현상은 그리 많지 않고 단편적이라고 하면서 종속적 연결어미로 다루어졌던 어미들을 모두 부사형 어미로 다룰 것을 제안하였다. 이러한 점들을 참고하여 본고 역시 종속적 연결어미를 따로 구분하지 않고 모두 부사형 어미에 포함시켜 내포절로 분류하고자 한다.

대등접속을 이루는 연결어미에는 '나열·순접'의 '-고, -요, -(으)며/-(으)면서, -자', '대조·역접'의 '-(으)나, -지만, -ㄴ데', '선택'의 '거나' 등이 있고 종속접속을 이루는 연결어미로 '계기적 사건'의 '-고, -자', '동시적 사건'의 '-(으)며/-(으)면서', '배경 상황 제시'의 '-는데' 등이 있다. 그러나 본고에서는 종속접속을 부사절에 포함시켜 분석하므로 종속접속의 연결어미는 부사형 어미에서 다루도록 한다.

(2) 포유문

포유문은 한 절이 다른 한 절을 안고 있는 문장을 가리키는데 안고 있는 절을 모절(matrix clause) 혹은 안은절이라고 하고 안겨 있는 절을 내포절(embedded clause) 혹은 안긴절이라고 부른다. 내포절은 다양한 방식으로 안기는데 안긴 절의 성격에 따라 명사절, 관형사절, 부사절, 인용절, 서술절로 나눌 수 있다. 그러나 어떤 문장을 포유문으로 볼 것인지에 대해서 역시 학자마다 견해 차이가 있으며 연구 목적에 따라 그 구분이 달라지기도 한다. 아래 본고에서 내포절을 판단하는 기준을 제시하도록 한다.

ㄱ. 명사절

명사절은 절이 명사화되어 문장 성분의 일부로 쓰이는 경우로 절에 명사형 전성어미 '-(으)ㅁ, -기'가 결합되어 이루어지는 것이 가장 보편적이지만 '-느냐, -(으)냐, -는가/-(으)ㄴ가, -는지/-(으)ㄴ지, -(으)ㄹ지'와 같은 의문형 종결 어미가 결합되어 이루어지기도 하고 이 외에도 문법적으로 명사절이 되기 어려운 특정 형식의 구성이 명사(구) 혹은 명사절로 쓰이기도 한다(구본관 외, 2015:260-261).[37] 이 중 의문형 종결 어미를 포함한 간접 의문절을 명사절로 볼 것인가에 대해서는 여러 의견들이 있지만 구본관 외(2015:261)에서는 다음의 예문(8)을 예로 들며 이러한 간접 의문절에 목적격 조사 '을/를'이 결합된다는 점에서 선행절이 명사절의 자격임을 알 수 있다고 하였다. 이는 남기심·고영근(2011:86)에서도 같은 입장을 보인다. 따라서 본고 역시 간접 의문절을 명사절로 인정하도록 한다.

(8) 가. 누가 고양이 목에 방울을 다느냐가 문제이다.
나. 우리는 그 제안을 받아들일지를 아직 결정하지 못했다.

위에 제시한 명사절 외에도 관형사형 전성어미와 의존명사 '것'이 결합한 형태인 '-(으)ㄴ/는 것, -ㄹ 것, -던 것'이 이끄는 절을 명사절로 보는 연구자들도 있다. 그러나 이 역시 학자마다 의견이 다른데 고영근·구본관(2018:330)에서는 '-(으)ㄴ 것', '-는 것'에 포함된 '-(으)ㄴ', '-는'은 관형사형 어미로 명사절을 만드는 명사형 어미와 같은 층위이기 때문에 '-(으)ㄴ 것', '-는 것'의 구성은 '관형사절+ 것' 으로 보아야 한다고 주장하지만 남기심·고영근(2011:386)에서는 '관형사형+ 것'을 명사절로 인정하고 있다. 한편, '-는 것, -(으)ㄴ 것, -(으)ㄹ 것'은 '것'이 꾸며주는 대상에 따라 또 다른데 구본관 외(2015:264)에서 제시한 아래의 예문(9가)와 같이 '것'이 꾸며주는 대상이 명확히 존재하는 경우가 있고 (9나)와 같이 밑줄 부분인 '관형사절+ 것' 전체가 명사절처럼 기능하는 경우가 있다.

(9) 가. 저기 하늘에 뜬 것이 애드벌룬이다.
나. 병이 나으려면 <u>이 약을 먹는 것</u>이 좋다.

[37] 구본관 외(2015:261)에서는 '그는 열정에 사로잡히면 일이 어렵고 쉽고를 가리지 않는다'의 밑줄 친 부분과 같이 특별한 장치가 없더라도 명사절이 되는 경우를 예시로 들고 있다. 그러나 본고에서 수집한 학습자 자료에서는 이와 같은 명사절이 나타나지 않으므로 자세한 설명을 생략한다.

그러나 고영근·구본관(2018:330), 구본관 외(2015:264), 임지룡 외(2005:253)에서도 보여주듯이 '-(으)ㅁ, -기'는 '-는 것, -ㄴ 것'으로 대치할 수 있다는 점에서는 명사절 기능을 하는 것을 부정할 수 없다. 또한 본 연구에서는 구어 자료 역시 분석하기 때문에 있고 특히 구어체인 경우 '-(으)ㅁ'보다는 '것'을 사용하는 것이 더 자연스럽다는 점(고영근·구본관, 2018:330)에서 '관형사형+ 것'에서 '것'이 꾸며주는 대상이 구체적인 경우에는 관형사절로 판단하지만 그 대상이 추상적인 경우 명사절로 판단하기로 한다. 구어체에서의 명사 '것'과 주격 조사 '이'가 축약된 '게', 의존명사 '것'과 목적격 조사 '을'이 축약된 '걸' 등도 의존명사 '것'에 실질적인 의미가 없는 경우 명사절로 인정하도록 한다.

ㄴ. 관형사절

관형사절은 절이 관형사화되어 관형어로 쓰이는 것으로 서술어가 관형사형 어미 '-(으)ㄴ', '-는', '-(으)ㄹ', '-던'을 취함으로써 이루어진다. 관형사절은 동격 관형사절(동격절)과 관계 관형사절(관계절)로 분류된다. 동격 관형사절이란 한 문장의 모든 필수 성분을 완전하게 갖추고 있는 관형사절이고 관계 관형사절이란 수식하는 명사와 동일한 성분이 빠져 있는 관형사절을 가리킨다(구본관 외, 2015:265). 아래의 예문(10)은 각각 주어, 목적어, 부사어가 생략된 관계절이다.

(10) 가. [e 밥을 먹은] 동생이 낮잠을 잔다.
나. [내가 e 본] 드라마는 한국 드라마이다.
다. [우리가 어제 e 간] 맛집은 이 지역의 명물이다.
라. [e 넓은] 초원이 펼쳐져 있다.

(구본관 외, 2015:267)

이때 (10가~다)의 밑줄 부분이 관형사절임은 이견이 없지만 (10라)의 밑줄 부분을 관형사절로 볼 수 있을지에 대해서는 연구마다 다른 입장을 보인다. 구본관 외(2015:267)에서는 '넓은'은 표면적으로 절의 형식을 갖추고 있지 않으나 주어 대상인 '초원'이 꾸밈을 받는 체언인 '초원'과 동일하기 때문에 생략된 것이며 '초원이 넓다'로 재구성할 수 있기 때문에 이를 관형사절로 보아야 한다고 하고 있으며 외국인을 위한 한국어 문법을 다룬 국립국어원(2005)에서 역시 (10라)와 같은 문장을 관형사절을 안은 문장에 포함시켜 제시하고 있다. 비록 동사나 형용사의 서술어에 관형사형 어미가 결합되는 것은 형태적으로나 기능적으로 관형사절을 형성하고 있음은 부정할 수 없으

나 문장 확장을 판단함에 있어 하나의 단어만으로 체언을 꾸며 주는 것이 통사적으로 복잡한 구조를 형성하였다는 보기는 어려운 것은 사실이다(남주연, 2015:99; 김현아, 2016:59; 서신웨, 2022:140). 또한 '절은 주어와 서술어를 갖춘 언어 형식'이라는 점에서도 이와 같은 형태는 절로 보기 힘들다.[38] 본 연구 역시 학습자의 문장 확장 능력을 측정하기 위한 목적으로 절을 분석하기 때문에 주어가 생략되고 관형사형 어미가 결합된 하나의 요소만으로 형성된 관형사절은 문장 확장으로 보지 않고 동사나 형용사의 어간에 관형사형 어미가 결합되어 있는 구조 앞에 논항이나 추가적인 수식이 있는 경우에만 관형사절로 인정하도록 한다.

동격절은 형태적으로는 필수 성분이 모두 포함되어 절의 형태를 띠고 특정 명사를 수식하고 있지만 피수식어와 관형절의 동격 관계가 성립되지 않는 경우가 있다. (11가)와 같이 동격절이 자립 명사를 수식하는 경우에는 뒤의 명사를 수식하고 한정하는 기능을 하지만 (11나)와 (11다)와 같은 밑줄을 친 부분의 동격절이 '것'이나 '줄'을 수식한다고 보기에는 힘들다는 것이다.

(11) 가. <u>우리가 이겼다는</u> 소문을 들었다.
나. <u>우리가 이긴</u> 것이 믿기지 않는다.
다. <u>우리가 이길</u> 줄 알았다.
라. <u>상대가 한 눈을 판</u> 사이에 우리가 이겼다.

(11다), (11라)와 비슷하게 '관형사형 전성어미+ 명사/의존명사+ 조사'의 구성으로 선행절과 후행절을 연결시켜 주는 역할을 하는 문법 항목에는 '-(으)ㄴ/는 가운데(에), -(으)ㄴ 결과(로), -(으)ㄴ/는/(으)ㄹ 경우(에)'와 같은 형태나 '-을 것이다, -ㄴ/는 것 같다, -ㄹ 수 있다'와 같이 양태적 기능을 하는 덩어리 표현들도 있다. 이와 같은 덩어리 표현들은 한국어교육에서는 문형이나 문법 표현으로 지칭하고 이들의 문법화한 문법적 기능에 주목하여 의미 기능을 교수해 왔다(한송화, 2015a:215). 한송화(2015:216)에서는 전형적인 관형절 외에 관형사형 어미가 의존명사와 결합하는 경우에 대해 아래와 같은 세 가지 경우로 나누어서 구분하고 있다.

(12) 가. 준명사형 관형절 예: <u>축구 선수가 되는 것</u>은 어렵다
나. 문장연결형 관형절 예: <u>도착하는</u> 대로 전화 주세요.

[38] 역대 학교 문법에서는 주어가 생략된 관형사절의 예로 (10가)와 같은 문장만 다루어 왔고 현행 학교 문법에서도 그러하다(구본관 외, 2015:267).

다. 양태형 관형절　　예: 난 테니스를 칠 수 있다.

위의 세 경우에서 (12가)의 경우 본고에서는 명사절로 분석하고 (12나) 역시 두 절을 연결해주는 역할을 한다는 점에서 관형절로 분석한다. 그러나 (12다)와 같은 문장은 관형사절을 생략하여도 주요 의미에는 변화가 없으며 양태적 의미만을 첨가하는 역할을 한다. 이러한 구조는 문장 확장으로 보기 어렵다고 판단하여 문장 확장으로 판단하지 않도록 한다.

ㄷ. 부사절

부사절은 절에 부사형 어미가 결합하여 부사화되어 부사어로 쓰이는 것을 가리킨다. 부사절은 기본적으로'-이, -게, -듯(이), -(으)ㄹ수록'에 의해 만들어 지지만 앞에서 언급하였듯이 본고에서는 종속절과 부사절을 구분하지 않고 종속적 연결어미가 이끄는 절을 모두 부사절에 포함시켜 분석하도록 한다. 이에 따라 종속적 연결어미로 분류되던 어미 모두 부사형 어미로 간주할 수 있게 된다.

구본관 외(2015:271)에서는 이러한 어미를 의미 범주에 따라 '동시적 사건'의 '-(으)면서/-(으)며', '계기적 사건'의 '-자, -자마자, -고(서), -아(서)/어(서)', '사건 전환'의 '-다가', '이유·원인'의 '-아(서)/어(서), -(으)니까, -느라(고), -다가, -기에, -길래', '조건·가정'의 '-(으)면, -거든, -아야/어야', '인정·양보'의 '-아도/어도, -(으)ㄹ지라도, -더라도, -ㄴ데, -아야/어야, -(으)ㄴ들', '목적·의도의' '-(으)러, -(으)려고, -고자', '배경·상황 설명'의 '-는데, -(으)ㄴ데, -(으)니'로 구분하여 제시하였다. 그리고 전통적으로 부사절로 분류되었던 어미에 대해서는 '결과·도달'의 '-게', '점차 심해짐'의 '-(으)ㄹ수록', '비유'의 '-듯(이)', 그리고 '달리, 같이, 없이' 세 형태로 한정되어 나타나는 '-이'로 구분하고 있다.

부사절 역시 기본적인 의미 범주 구분 외에도 부사절의 범위를 어디까지 인정할 것인지에 대한 논의가 필요하다. 먼저, 관형사절과 마찬가지로 부사형 어미가 결합하였지만 한 어절로 이루어진 경우를 부사절로 볼 것인가 하는 문제이다. 특히 부사형 어미 '-게'는 다른 부사형 어미와 동일하게 주어와 다른 성분을 갖추어 부사절을 이루기도 하지만 주어가 생략되어 보통의 부사어처럼 쓰이기도 한다. 구본관 외(2015:272)에서는 다음과 같은 (13가)와 같은 예를 들면서 관형사절을 안은 문장과 마찬가지로 '아름답게'의 주어인 '꽃이'가 생략된 것으로 볼 수 있다고 하며 부사절로 인정하였다. 그러나 앞서 관형사절에서도 설명하였듯이 부가적인 요소 없이 하나의 어절로 이루어진 형태는 절보다는 부사어의 성격이 강하고 문장 확장으로 보기 힘들다는 점에서 한 어절

로 이루어진 부사절은 문장 확장으로 분석하지 않고 (13나)와 같이 주어와 다른 성분을 갖춘 경우에만 부사절로 분석하도록 한다.[39]

(13) 가. 꽃$_i$이 [e_i 아름답게] 피었다.
나. 엄마는 아이에게 혀가 닳게 타일렀다.

ㄹ. 인용절

인용절이란 어떤 문장이 남의 말이나 글, 생각 따위를 따오듯이 표현하는 방식으로 상위문에 안긴 절을 가리키는데(구본관 외, 2015:274) 주어진 문장에 인용의 부사격 조사 '라고, 하고'나 부사형 어미 '-고'가 붙어서 이루어진다. 인용절에 대해서는 명사절로 보는 견해도 있고 부사절로 보는 견해도 있다. 그러나 인용절은 '화자 자신이나 다른 사람의 말이나 생각을 옮기기'위해 사용된다는 독특한 기능을 가지고 있으므로 따로 다루는 것이 유용하다(고영근·구본관, 2018:338)는 점에서 본고 역시 인용절을 부사절이나 명사절과 구분하여 다루도록 한다.

인용은 직접 인용과 간접 인용으로 구분할 수 있는데 직접 인용의 경우에는 '(이)라고, 하고'가 붙고[40] 그 이외에는 '-고'가 붙는다. 간접 인용의 경우에는 인용되는 문장의 유형에 따라 인용되는 부분의 종결 어미가 달라지는데 평서형에서는 '-다, -라'가 붙고 의문형에는 '-느냐, -(으)냐', 명령형에는 '-라', 청유형에는 '-자'가 붙는다(남기심·고영근, 2011:395). 따라서 본고에서는 직접 인용절에서는 부사격 조사 '(이)라고'와 '-하고'를 표지로 삼고 간접 인용절은 부사형 어미[41] '-다고, -(느)냐고, -(으)라고, -자고'를 표지로 삼아 판단하도록 한다.

ㅁ. 서술절

[39] 한 어절로 이루어진 관형사절과 부사절을 절로 인정할 것인지에 대해서는 연구마다 다른 입장을 취하고 있다. 예를 들면 남주연(2015), 김현아(2016)에서는 '예쁜 꽃'와 '빨간 꽃', '아름답게 피었다'와 같이 한 어절로 이루어진 경우 절로 인정하지 않고 있지만 최예슬(2017), 최수완(2019)에서는 이를 부사절로 인정하지 않을 근거 또한 부족하다고 하여 절로 인정하고 있다.

[40] 구본관 외(2015:275)에서는 일반적으로 '하고'는 아직 문법화가 완전히 이루어지지 않아 인용 동사 '하다'의 활용형으로 본다고 하였지만 인용절을 만든다는 점에서 편의에 의해 직접 인용의 조사로 간주하도록 한다.

[41] '(이)라고'와 '-고'에 대해서는 학자마다 지칭하는 방법이 다른데 본고에서는 구본관 외(2015)의 견해에 따라 직접 인용의 '(이)라고'는 부사격 조사로, 간접 인용의 '-다고, -냐고, -(으)라고, -자고'를 부사형 어미로 지칭하고자 한다.

한국어 문장 중에는 주어가 두 번 실현되는 것처럼 보이는 문장들이 있다. 예를 들면 '토끼가 앞발이 짧다'와 같은 문장들이다. 그러나 이러한 문장에 대해서는 '이/가'를 가진 명사구를 대주어와 소주어로 보는 입장, 주제-설명으로 보는 입장, 심층에서의 다른 격 조사가 표면에서 '이/가'를 취한 것으로 보는 입장, 서술절을 가진 안은 문장으로 보는 입장 등 매우 다양한 입장들이 존재하며 아직 의견의 일치를 보지 못하고 있다(고영근·구본관, 2018:341).

학교 문법에서는 이러한 문장을 서술절을 가진 안은 문장으로 보고 있지만 명사절이나 관형사절 등과 달리 어미나 조사와 같은 뚜렷한 표지가 없다는 점이나 서술절의 서술어가 오히려 상위절의 주어와 일치하는 경우[42] 등을 해석하기 어렵다는 비판이 여전히 존재한다(고영근·구본관, 2018:342). 그러므로 본 연구에서는 그 판단 근거의 모호성으로 인하여 서술절은 인정하지 않는다.

ㅂ. '본용언+ 보조 용언' 구성

절을 판별하는 데 있어 '본용언+ 보조 용언' 역시 고려 대상이 된다. 한국어에는 본동사가 '-아, -게, -지, -고'나 관형사형 전성어미를 통해 보조 용언과 결합하여 서술어의 역할을 하는 경우가 있는데 이때 보조 용언은 본용언이 갖는 기본 의미에 부정, 진행, 완료, 가능 등의 의미를 더하는 부가적인 기능을 한다.

보조 용언 구문에 대해서는 학계에서 분명한 합의가 이루어지지 않아 학문 문법에서는 보조 용언 구문에 대해 필수 부사절을 내포절로 요구하는 문장으로 보는 견해도 있지만 고영근·구본관(2018)에서는 보조 용언은 자립성이 결여되어 있거나 아주 희박하며 본용언에 문법적인 기능을 더해준다는 점에서 영어의 조동사와 비슷한 기능을 가진다고 하였다. 구본관 외(2015)에서 역시 보조 용언 구문에 대해 '내포절과 모절이 주어가 반드시 같아야만 하는 문장만'으로 정의하고 '-어/아, -게, -지, -고, -고야'를 보조적 연결어미로 구분하여 이를 본용언과 보조 용언을 매개하여 하나의 서술어로 기능하게 하는 역할을 한다고 보아 '본용언+ 보조적 연결어미+ 보조 용언'구성을 하나의 서술어로 보고 있다. 또한 한국어 교육에서는 이러한 '본용언+ 보조 용언' 구성을 고정된 표현으로 교육하는 경우가 많기 때문에 문장 확장으로 보기 어렵다. 따라서 본고에서 역시 '본용언+ 보조 용언'에 대해서 하나의 서술어로 보고 복문으로 분석하지 않도록 한다. '본용언+ 보조 용언' 구성의 판단은 고영근·구본관(2018:103-106)에서 제시한 표

[42] 고영근·구본관(2018:342)에서는 '할아버지가 돈이 많이 있으시다'를 예로 들면서 '-시-'가 서술절의 주어가 아닌 상위절의 주어와 일치하는 현상을 해석하기 어렵다고 하였다.

를 참고하되 표에 제시되어 있지 않지만 '본용언+ 보조 용언'의 구성을 보이는 문장에 대해서는 추가적으로 국립국어원(2005a:292)[43], 김영태(1997:25-26)를 참고하여 판단하도록 한다. 고영근·구본관(2018)의 보조 용언 목록은 다음의 <표 8>과 같다.

<표 8> 의미 기능에 따른 보조 용언 목록(고영근·구본관, 2018)

구분	의미 기능	목록
보조 동사	완료	-(고) 나다, -(어/아) 내다, -(어/아) 버리다
	진행	-(어/아) 가다, -(어/아) 오다, -(어/아 있다/계시다
	보유	-(어/아) 두다, -(어/아) 놓다, -(어/아) 가지다
	시행	-(어/아) 보다
	사동	-(게) 하다, -(게) 만들다
	피동	-(어/아) 지다, -(게) 되다
	부정	-(지) 아니하다(않다), -(지) 못하다, -(지) 말다
	당위	-(어야/아야) 하다, -(어/아) 보이다
	시인	-(기는) 하다
	강세	-(어/아) 대다
	봉사	-(어/아) 주다/드리다
보조 형용사	상태 지속	-(어/아) 있다
	부정	-(지) 아니하다(않다), -(지) 못하다
	희망	-(고) 싶다
	추측	-(ㄴ)가/-는가/-나 보다, -는가/-나/-(으)ㄹ까 싶다
	시인	-(기는) 하다

그러나 위의 표에서 '-게 하다'와 '-게 만들다' 사동 표현은 '내포절과 모절이 주어가 반드시 같아야만 하는 문장'이라는 보조 용언 구문의 정의에 부합하지 않으며 또 본용언만으로는 문장이 성립되지 않는다(고영근·구본관, 2018)는 문제점이 있다. 그러므로 본고에서는 '-게 하다'와 '-게 만들다' 사동문에 대해서는 부사형 어미 '-게'를 포함한 부사절로 분석하도록 한다.

ㅅ. 오류문에 대한 분석:

통사 복잡성을 단위당 어절 수나 절 수, 또는 단위 확장의 표지로 측정할 때 오류문에 대해 어떻게 판단할 지에 대한 고려가 따른다. 정확하지 않지만 긴 문장을 산출하거나 정확하지 않은 연결어미나 전성어미를 사용하여 문장 확장을 시도한 것에 대해 통사적으로 복잡하다고 판단해야 할지에 대한 문제이다. 이 문제에 대해 김현아(2016)에

[43] 국립국어원(2005a)에서는 이와 같은 구성을 양태 표현에 포함시켜 제시하고 있다.

서는 Skehan & Foster(1999)의 복잡성에 대한 정의[44]를 언급하면서 형태 혹은 의미적 오류가 있는 문법 항목을 사용하였을지라도 학습자가 통사적으로 문장을 확장하려는 의지를 가지고 확장을 시도하였다면 그렇지 않은 학습자 혹은 전략적으로 회피한 학습자에 비해 복잡성 측면에서는 더 높은 발달을 이루어 낸 것으로 보아야 한다고 하였다. 본고 역시 이와 같은 입장을 지지하는 바이다. 또한 본고에서는 정확성 역시 함께 살펴보기 때문에 오류에 대해서는 정확성 분석 단계에서 다시 살필 수 있다. 이에 본고에서는 문장 확장을 기본적으로 연결어미나 전성어미, 조사로 판단을 하며 (14나)의 밑줄 부분과 같이 의미나 형태적으로 오류가 있지만 형태적으로 연결어미를 사용했다고 판단되는 경우에는 문장 확장으로 판단하고 (14가)와 같이 문장을 인용하였으나 인용 조사를 사용하지 않은 등 확장의 표지가 없는 경우 문장 확장으로 판단하지 않는다.

(14) 가. 그리고 일어나고 내가 '오늘 꼭 일직 잘거야!'(√라고) 생각해요.
나. 나도 뭐 먹고 싶어는대(√싶었지만) 그냥 안 먹었어요.

이상의 접속문과 포유문에 대한 논의에 근거하여 본고에서는 다음과 같은 접속문과 포유문의 분류 체계[45]를 설정하였다.

<표 9> 복문의 분류 체계

복문	분류	세부 내용
대등절	나열·순접	-고, -며, -면서, -자
	대조·역접	-ㄴ데, -나, -지만, -아도/어도, -면서, -고
	선택	-거나, -든지/든가[46], -든

[44] Skehan & Foster(1999:96-97)에서는 복잡성에 대해 '자신의 효과적으로 통제하지 못할지라도 더 발전된(advanced) 언어를 사용할 수 있는 능력'으로 정의하면서 이는 위험을 감수하면서도 통제되지 않은 언어체계를 사용하려는 의지와 중간언어 체계를 변화 발전시키는 재구조화(restructuring)와 관련된다고 하였다.

[45] 각 절을 이끄는 문법 항목은 기본적으로 구본관 외(2015)를 참고하여 분류 및 나열하였으며 구본관 외(2015)에서 제시되지 않았지만 학습자의 자료에서 나타난 어미들에 대해서는 국립국어원(2005a)을 참고하여 각 범주에 추가하였다. 본고에서 제시한 어미나 조사 외에도 각 절을 이끄는 어미나 조사가 더 있지만 자주 사용하지 않거나 본 연구의 학습자 자료에서 나타나지 않은 어미나 조사는 표에 제시하지 않았다.

[46] 구본관 외(2015)에서는 '-든지/든가'를 연결어미 보다는 내포 어미로 분류하는 것이 더 타당하며 표현문형 즉 덩어리 표현(chunk expression)으로 기술하는 것이 유용하다고 하였지만 '-든지/든가'는 앞의 것이나 뒤의 것 중에서 하나를 선택함을 나타내며 두 절을 연결해준다는 점에서 본고에서는 국립국어원(2005a)의 분류에 따라 선택의 대등절에 포함시키도록 한다.

내포절	명사절	명사형 전성어미	-ㅁ, -기
		관형사형 전성어미+ 것[47]	-ㄴ 것, -는 것, -ㄹ 것, -던 것
		의문형 종결 어미	-ㄴ지/는지, -ㄹ지, -ㄹ까, -냐/느냐, -ㄴ가/는가, -던가
	관형사절	관형사형 전성어미	-ㄴ, -는, -ㄹ, -던
	부사절	동시적 사건	-며, -면서
		계기적 사건	-자, -자마자, -고(서), -아(서)/어(서)
		사건의 전환	-다(가)
		이유·원인	-아(서)/어(서), -니/니까, -느라(고), -다가, -므로, -라(서), -기에/길래
		조건·가정	-면, -거든, -아야(지)/어야(지), -려면, -라면, -다면, -자면, -냐면
		인정·양보	-아도/어도, -ㄹ 지라도, -더라도, -라도, -ㄴ데, -아야/어야, -ㄴ들
		목적·의도	-러, -려고, -고자, -도록, -게
		배경·상황 설명	-ㄴ데/는데, -니, -더니
		수단·방법	-아(서)/어(서), -고
		방식·정도	-게, -도록, -이
		점진적	-ㄹ 수록
		비유	-듯(이)[48], -다시피
		사역	-게
	인용절	직접 인용	-(이)라고, 하고
		간접 인용	-다고/(이)라고, -냐고, -(으)라고, -자고

3.2. 어휘 복잡성 측정 지표 및 분석 기준

어휘 복잡성을 측정하기 위해서는 '어휘'에 대한 개념 정의가 선행되어야 한다. 어휘는 어휘, 단어, 낱말 등 여러 용어들이 구분되지 않고 사용되기도 하는데 구본관 외(2015:104)에서는 단어는 '규칙에 의해 발화 과정에서 새롭게 만들어지는 문장과 달리 머릿속 어휘부(mental lexicon)에 저장되어 있는 단위'로 정의하고 '어휘'는 개별 단어가 아닌 단어 집합의 의미로 많이 사용된다(구본관 외, 2015:107)고 하였다. 그리고 김

[47] '것'이 가리키는 대상이 추상적인 경우

[48] '-이'는 '달리, 같이, 없이' 세 가지로 한정되어 있는 부사형 전성어미를 가리킨다. 구본관 외(2015:272)에서는 '-이'는 결합 제약이 심해 접미사로 볼 수도 있으나 절에 결합한다는 점을 더 중시하면 부사 파생 접미사 보다는 부사형 어미로 보는 편이 더 합리적이라고 하였다.

광해(1993:39)에서도 어휘는 '일정한 범위 안에서 사용되는 단어의 집합'이고 어휘를 이루는 하나하나의 단어들은 어휘소라고 하였으며 임지룡(1998:3)에서도 단어는 개별적인 단위인 반면, 어휘는 단어의 총체로서 집합적인 성격을 띠고 있다고 하였다. 즉 단어는 문장을 이루는 하나 하나의 단위이고 어휘는 단어의 집합으로 정리할 수 있다.

그러므로 '어휘'에 대해 정의하기 위해서는 '단어'에 대한 정의가 선행되어야 한다. 한국어에서 단어의 정의는 조사와 어미를 단어에 포함시킬 것인가에 따라 그 범위가 달라진다. 조사와 어미를 단어에 포함시킬지는 학자마다 의견이 다른데 주시경을 비롯한 초기의 문법가들은 단어를 최대한으로 분석하려는 관점에서 조사와 어미의 일부까지 단어로 인정한 반면 정렬모나 장하일, 이숭녕 등 역사 문법가들은 조사와 어미 가 자립성과 어휘적 의미를 갖지 않는다는 점에서 조사와 어미 모두를 단어로 보지 않기도 한다(고영근·구본관, 2018:34). 한국어의 조사나 어미는 모두 선행하는 어휘 요소에 결합하여 문법적인 기능을 실현하며 라틴어와 같은 굴절어에서 굴절 요소가 단어에 결합하여 굴절 요소를 포함한 전체가 단어로 기능하는 것과 달리 구에 결합한다는 점에서 조사나 어미가 결합한 단위를 단어로 보기 곤란하다는 특징이 있다. 한편, 조사와 어미는 자립성 측면에서 차이가 있는데 어미와 조사는 모두 선행하는 어휘 요소에 결합하여 문법적인 기능을 실현하지만 자립성의 기준에서 보면 어미 앞에 오는 용언 어간은 자립성이 없지만 조사 앞에 오는 체언은 자립성이 있다는 점이 다르다(고영근·구본관, 2018:106).[49] 이에 최현배 등 학자들은 조사는 단어로 인정하고 어미는 단어로 인정하지 않는 절충적 방안을 주장하였는데 현재는 학교 문법에서도 이를 받아들여 어미는 조사에 포함하지 않지만 조사는 단어에 포함시켜 9품사 체계를 적용하고 있다. 이러한 논의를 참고하여 본 연구에서도 어미는 단어에 포함하지 않지만 조사는 단어에 포함하여 분석하도록 한다.

단어는 또 문법적 성질의 공통성에 따라 분류하는데 이것을 품사라고 부른다(남기심·고영근, 1985). 품사를 분류할 때에는 그 분류 기준을 대체로 형태, 기능, 의미의 순서로 적용하는데 먼저 형태의 기준으로는 활용을 하는 단어와 그렇지 않은 단어로 나누

[49] 고영근·구본관(2018:32)에서는 단어에 대해 가장 널리 알려진 정의가 '최소자립형식(a minimum free form)임을 언급하면서 이러한 자립성의 기준을 엄격하게 적용하면 조사는 단어의 자격을 가지기 어렵고 체언에 조사가 결합한 것과 어간에 어미가 결합한 것은 동일한 형식으로 보아 모두 그 전체를 단어로 인정해야 하지만 조사에 선행하는 체언의 자립성이 어미에 선행하는 용언 어간의 자립성보다 훨씬 높다는 것이 어미와 달리 조사를 단어로 취급하는 중요한 이유라고 하였다.

어 각각 가변어와 불변어로 지칭한다. 다음 기능에 따라 불변어는 체언, 수식언, 독립언, 관계언으로, 가변어는 용언으로 나누어 진다. 그리고 마지막으로 의미에 따라 체언은 '명사, 대명사, 수사'로 수식언은 '관형사, 부사'로, 용언은 '동사, 형용사'로 나눌 수 있으며 독립언에는 감탄사, 관계언에는 조사가 포함된다. 이와 같이 한국어의 품사는 흔히 9품사로 나누어지는 것이 일반적이지만 이 외에도 적게는 5품사에서 많게는 13품사까지 나누기도 한다(구본관 외, 2015:162). 본고에서는 학교 문법의 분류에 따라 9품사 체계를 따르도록 한다.

<표 10> 한국어의 품사 분류(구본관 외, 2015:162)

분류		세부 내용
불변어	체언	명사
		대명사
		수사
	수식언	관형사
		부사
	독립언	감탄사
	관계언	조사
가변어	용언	동사
		형용사

어휘를 분석함에 있어 데이터를 가공하는 방법 역시 분석 결과에 큰 영향을 미친다. 본고에서는 어휘 분석에 있어 다음과 같은 기준을 적용하였다.

① 기본적인 어휘 구분은 <표준국어대사전>에 등록되어 있는 표제어를 기준으로 하여 합성어의 경우 <표준국어대사전>에서 하나의 단어로 제시되어 있으면 하나의 단어로 간주한다.
② 접두사와 접미사는 통합하여 어근과 함께 하나의 어휘로 취급하여 측정한다. 예를 들면 '무의식적'과 같은 명사의 경우 '무의식적' 전체를 하나의 단어로 처리한다. 단 '들, 끼리, 간, 네, 님, 씩, …'과 같은 단어 형성에 관여하지 않는 접미사는 따로 구분한다.
③ 보조 용언의 경우 띄어 쓴 것을 원칙으로 하여 각각의 어휘로 인정한다.
④ 어휘 다양도 분석을 위해 '은/는, 이/가, 을/를'과 같이 자음 뒤 이형태와 모음 뒤 이형태가 교체하는 경우 자음 뒤 이형태를 대표형으로 처리하여 같은 어휘 유형으로 분석한다.
⑤ 동형어 구분을 위해 같은 형태를 가지고 있지만 의미가 다른 단어들을 다르게 태깅하

되 그 의미는 국립국어원을 따른다.

⑥ 학교 이름이나 지역 이름은 고유 명사로 인정하지만 사람의 이름은 어휘 분석에 포함하지 않는다.

⑦ 어휘에 나타난 철자 오류, 띄어쓰기, 활용형이나 이형태의 오류는 바른 형태로 수정하여 분석한다.[50]

⑧ 한국어에서 외래어로 사용되지 않는 외국어에 대해서는 모두 외국어(SL)로 분석하며 분석이 불가능한 어휘는 'NA'로 구분한다.

⑨ '근데' 혹은 '널', '넌'과 같은 준말 형태는 본딧말로 복원하여 '그런데', '너+를', '너+는'과 같이 분석하며 구어에서 자주 나타나는 '그리고-그리구', '-려고, -려구'와 같은 자모음의 변이와 같은 이형태는 모두 표준어로 복원하여 계산한다.

⑩ 구어에서 학습자들이 머뭇거리는 과정에서 발생하는 '어, 음, 이, 그, 저 …'와 같은 주저어는 분석에서 제외하며 학습자가 의도적으로 발화한 경우만 계산에 포함한다.

3.2.1. 어휘 다양도

어휘 다양도는 일반적으로 텍스트에 나타난 어휘 유형 수와 어휘 구현 수의 비율로 측정한다. 어휘 유형 수와 어휘 구현 수의 비율이 높다는 것은 반복된 어휘를 적게 사용하고 새로운 어휘를 많이 사용한다는 것을 의미한다. 이는 다양한 어휘를 많이 사용하는 것이 한정된 어휘를 반복적으로 사용하는 것보다 어휘 지식이 더 발달된 것으로 볼 수 있기 때문이다.

어휘 다양도를 측정하는 가장 오래 된 지표는 TTR(type-token ratio)이다(Johnson, 1944). 여기서 어휘 유형 수는 텍스트에 나타난 서로 다른 형태나 의미를 가진 단어의 유형 수를 가리키고 어휘 구현 수는 텍스트에 나타난 총 단어의 수를 가리킨다. 따라서 TTR값이 높을수록 더 다양한 어휘를 사용하였다고 볼 수 있다. 그러나 TTR은 길이가 서로 다른 텍스트의 어휘 다양도를 측정함에 있어서는 제한점이 존재하였다. TTR로 어휘 다양도를 측정할 때 전체 텍스트의 길이가 길어질수록 어휘 구현 수는 높아지는 반면 새로운 어휘가 나타날 기회는 낮아지므로 어휘 유형 수는 상대적으로 덜 높아지게 되어 어휘 구현 수가 많은 학습자가 어휘 구현 수가 적은 학습자에 비해 다양도가 낮게 나타난다는 것이다(Daller 외, 2003:199). 이는 말을 하거나 글을 쓸 때 같은 단어들을 반복해서 사용하기 때문에 어휘 유형 수의 분포는 텍스트의 길이와 선형적인 관계

[50] 어휘 복잡성 분석은 정확성 분석과는 별개의 측정 지표이기 때문에 정확한 분석을 위하여 올바른 형태로 수정하여 분석하고 어휘의 오류는 정확성 분석 부분에서 다시 살펴보도록 한다.

를 나타내지 않기 때문이다(Miller & Biber, 2015:36). 그러므로 단순 TTR은 동일한 길이의 텍스트를 비교하는 데에만 사용할 수 있다는 단점이 있다(Brezina, 2018:58). 이에 텍스트 길이의 영향을 줄이고자 위해 기존의 TTR방식에서 좀 더 변형된 MSTTR(mean segmental TTR), Guiraud(1960)에서 제안한 RTTR(root TTR)[51], CTTR(corrected TTR), MATTR, D값 등이 개발되었다.

MSTTR은 STTR(standardized type token ratio)에 기반한 것으로 Johnson(1944)에서 제시되었다(Covington & McFall, 2010:96). MSTTR은 텍스트를 표준 크기(예: 500단어, 1000단어)로 분할한 뒤 각 부분의 TTR를 계산한 다음 평균 값을 취하는 방법이다. 그러나 모든 텍스트가 같은 크기로 분할이 될 수 없기 때문에 표준 크기보다 짧은 마지막 부분은 계산에서 제외되어야 한다. 또 짧은 텍스트(100개 미만 단어)의 경우 더 작은 표준 크기(예: 30개 단어로 분할)를 선택해야 하지만 너무 짧은 크기로 나눌 경우 결과를 왜곡하는 것으로 보고되었다(Malvern & Richards, 2002:88).

MATTR은 Covington & McFall(2010)에서 제시한 것으로 이동 평균 TTR을 의미한다. MATTR 역시 STTR에 기초한다는 점에서 MSTTR과 유사하지만 MATTR은 텍스트를 정확히 나눈 뒤 평균을 구하는 대신 범위가 서로 겹치면서 이동하는 이동창을 사용한다. MATTR은 표준 크기를 설정한 뒤 이 창이 이동하는 위치에 따라 해당 위치에서의 TTR을 계산한 뒤 평균 값을 구하는 방식이다. MATTR은 텍스트의 모든 부분을 고려한다는 점에서 MSTTR에 비해 더 정확한 지표라고 할 수 있다(Brezina, 2018:58).

다음으로, 수학적 공식을 통해 표본 크기에서 나타나는 차이를 보상하기 위한 시도들이 있었으며 대표적으로 Guiraud(1960)가 소개한 RTTR(root TTR)과 Carroll(1964)이 소개한 CTTR(corrected TTR)이 있다. RTTR과 CTTR은 모두 TTR이 어휘 구현 수의 증가에 따라 어떻게 하락하는지를 수학적으로 표현하는 비교적 간단한 방법을 찾고 이를 승수(multiplier) 또는 다른 수식(modifier)으로 사용하여 전체 샘플에서 TTR이 상수로 작용하게끔 수학적으로 변형하는 방법을 시도하였다. 그 결과 Guiraud(1960)와 Carroll(1964)은 모두 TTR의 하강과 어휘 구현수의 루트가 정비례한다고 가정하여 어휘 구현 수를 조절한 공식을 다음과 같이 제시하였다(Malvern 외, 2004:26).

$$\text{RTTR} = \frac{\text{어휘 유형 수}}{\sqrt{\text{어휘구현 수}}} \qquad \text{CTTR} = \frac{\text{어휘 유형 수}}{\sqrt{2} \times \text{어휘 구현 수}}$$

[51] RTTR(root TTR)은 Guiraud(1954, 1960)에서 제안한 것으로 G값 지표(guiraud index)라고 불리기도 한다.

RTTR과 CTTR은 여러 변형된 TTR 중에서 가장 영향력이 있는 지표로 여러 연구(Broeder et al., 1993; Vermeer, 2000; Daller et al., 2003)들을 통해 그 타당성이 증명되기도 하였다. 그리고 어휘 다양도 측정 지표 중에는 가장 최근에 개발된 D값(D-value)이 있다. D값은 Richards & Malvern(2000)에서 제시한 것으로 텍스트 길이가 길어질 때 새로운 어휘가 나타날 가능성을 수학적 그래프로 모형화 하여 어휘 구현수의 증가에 따라 감소하는 TTR의 매개변수 값 D를 어휘 다양도로 규정하는 방법이다. D값 역시 정확한 어휘 다양도 측정 지표로 평가 받고 있으나 100-2,000 어휘의 범위를 벗어나는 샘플에서는 텍스트 길이의 영향을 벗어나기가 어렵다(McCarthy & Jarvis, 2007)는 단점이 있다. 본고의 분석 자료들 중에는 어휘 수가 100개 이하인 샘플들이 존재하는 관계로 정확성을 위해 D값보다는 CTTR을 사용하도록 한다.

3.2.2. 어휘 밀도

어휘는 실질적으로 의미를 전달하는 어휘인 내용어(content words)와 문법적인 기능을 하는 기능어(fuction word)로 나눌 수 있다. 임지룡(1998:15)에서는 어휘 사전에서는 조사, 어미와 같은 기능어는 기능어는 소수이며 언어 습득의 초기에 완성되어 늘어날 가능성이 거의 없다는 점에서 폐쇄적인 반면 내용어는 다수이며 그 수치나 의미가 가변적이라는 점에서 개방적이기 때문에 어휘량을 늘린다거나 어휘력을 평가한다고 할 때는 내용어가 중심이 되어야 한다고 하였다. 특히 텍스트에서 내용어의 비율은 생각이나 정보 전달의 집중도를 나타내기도 한다.[52] 이처럼 어휘 밀도는 텍스트에서 내용어가 차지하는 비율로 측정이 되는데 기능어와 내용어의 비율로 측정하기도 하고 전체 어휘와 내용어의 비율로 측정하기도 한다.

그러므로 어휘 밀도를 측정하기 위해서는 내용어와 기능어의 범주를 구분하는 것이 중요하다. 먼저 영어권의 연구를 살펴보면 명사, 동사, 형용사, 부사(Laufer & Nation, 1995; Read, 2000)를 내용어로 보고 있는 것을 확인할 수 있다. 반면 한국어교육 연구에서는 내용어와 기능어의 범주 설정이 연구자마다 약간 상이한데 배도용(2012), 강정희(2020), 김민수(2022)에서는 일반 명사, 동사, 형용사, 일반 부사 4가지를 내용어로 선정하고 있는 반면 강범모·김흥규(2009), 남주연(2015), 류한레이(2016) 등 연구에서는 관계언인 조사(격 조사, 보조사, 접속 조사)를 제외한 일반명사, 의존명사, 고유 명

[52] Read(2000)에서는 Ure(1971)의 연구에서 쓰기 자료의 내용어 비율은 40%이상인 반면 구어 자료의 내용어 비율이 40%미만인 것을 언급하며 이는 서면 언어에서 생각이나 정보가 더 집중적으로 제시됨을 반영한다고 하였다.

사, 대명사, 수사, 동사, 형용사, 관형사, 일반 부사, 접속 부사, 감탄사, 지정사, 보조 용언까지 내용어에 포함시키기도 하였다. 그리고 주우동(2019)이나 주월랑(2018)에서는 좀 더 세부적으로 구분하여 조사 외에도 문법적 성격이 강하다고 판단되는 의존 명사, 지시 관형사, 접속 부사 등을 기능어로 분류하기도 하였다.

그러나 이런 다양한 분류 방법과 시도에도 불구하고 어휘 밀도가 숙달도에 따라 높아진다는 이론적 주장과 다르게 한국어에서 숙달도별 어휘 밀도를 측정한 여러 연구들에서는 어휘 밀도가 숙달도 구분에 큰 의미가 없다는 결과들이 대다수이다. 특히 어휘 밀도를 내용어당 전체 어휘수로 측정한 배도용(2014), 남주연(2015), 주월랑(2018), 주우동(2019) 등 연구들은 어휘 밀도와 숙달도와의 관계가 유의미하지 않음을 밝히고 있다. 주월랑(2018)에서는 실질 어휘 유형 및 구현 수는 시간의 축에 따라 양적으로 증가하였으나 총 어휘 대비 실질 어휘 유형 수는 감소하였다고 보고하였고 주우동(2019)는 어휘 밀도가 숙달도와 유의미한 차이를 보여주지 못한다고 하였으며 강정희(2020), 서신웨(2022)에서는 성취도가 낮거나 숙달도가 낮은 학습자의 어휘 밀도가 반대로 더 높게 나타났다고 보고 하였다.

이러한 현상에 대해 Hwang(2012)과 배도용(2014)은 짧은 텍스트를 산출하는 학습자가 긴 텍스트를 산출하는 학습자보다 어휘 밀도가 더 높게 나타날 수 있다는 것을 언급하며 어휘 밀도가 텍스트 길이의 영향을 받을 가능성을 주장하기도 하였다. 그러나 아직 어휘 밀도와 텍스트 크기의 상관성이 명확히 밝혀진 바는 없다(안의정, 2017:354). 본고에서는 '생각이나 정보 전달'이라는 내용어의 정의에 초점을 두고 앞에서 살펴본 연구에서 공통적으로 내용어로 간주하여 분석한 일반 명사, 동사, 형용사, 일반 부사만 내용어로 보고 어휘 밀도를 분석하도록 한다.

3.2.3. 어휘 정교도

어휘 정교도는 어휘 다양도나 밀도와 달리 학습자가 사용한 어휘의 수준을 측정하는 지표로 텍스트에 나타난 어휘 중에서 비일상적이고 전문적인 어휘의 비율로 측정된다. 비일상적이고 전문적인 어휘는 저빈도 어휘라고 부르기도 하는데 저빈도 어휘는 일상적으로 사용하는 고빈도 어휘를 제외한 어휘를 가리키는 것으로 Kuiken & Vedder(2012:146)에서는 일상적으로 사용하는 고빈도 어휘 2,000개를 제외한 어휘를 저빈도 어휘로 규정하기도 하였다. 이는 한국어에서도 비슷한데 김광해(1993)에서는 일상생활에 있어 필수적인 단어로 1,000~2,000개를 기초 어휘로 규정하였고 국립국제교육원의 한국어능력시험(TOPIK) 홈페이지에서 공개하고 있는 초급 어휘 역시

1,847개로 2,000개 정도임을 확인할 수 있다.

어휘 정교도는 학습자가 산출한 텍스트 자체만 분석하는 내부적 측정이 아닌 외부적 어휘 목록에 기댄 외부 측정이므로[53] 기준이 되는 '어휘 목록'이 중요한 변수로 작용한다. 특히 제2언어 학습자의 어휘 정교도 측정은 연구 대상자인 학습자가 처한 숙달도 수준보다 높은 단계의 어휘를 사용할 수 있는 능력을 측정한다는 점에서 연구자는 연구 대상자의 언어 수준과 교육과정을 고려하여 학습자에게 적절한 '저빈도 어휘'를 선정하는 것이 가장 중요하다.

기존의 연구들을 살펴보면 한국어 교육 연구에서 자주 사용되는 어휘 목록은 조남호(2003), 국제 통용 한국어 교육 표준 모형, 국립국어원(2012, 2013, 2014), 국립국어원(2015)[54]이 있는데 그중 가장 많이 사용된 어휘 목록은 국립국어원에서 개발한 '한국어 교육 어휘 목록'이다. '한국어 교육 어휘 목록'은 한국어능력시험(TOPIK) 자료와 한국어 교재 말뭉치, 국제 통용 한국어 교육 표준 모형의 어휘 목록 등을 참고하여 개발된 어휘 목록으로 초급 어휘[55] 1,835개, 중급 어휘 3,855개, 고급 어휘 4945개 단어를 포함하고 있다. 이에 성인 한국어 학습자를 연구한 이복자(2015), 류한레이(2016), 원미진 외(2017), 주우동(2019a), 서신웨(2022) 등 연구에서 이 목록을 사용하여 학습자의 어휘 정교도를 측정하였다.

그러나 중도입국 청소년의 한국어 교육은 교육 목표나 교육 내용, 학습 환경에 있어 일반 한국어 학습자와 큰 차이가 있다. 2장에서 제시하였듯이 본고에서 연구 대상으로 삼고 있는 학습자는 모두 <한국어(KSL) 교육과정(2017)>을 바탕으로 개발된 '의사소통 한국' 교재로 한국어를 학습하였다. <한국어(KSL) 교육과정(2017)>은 중도입국 학생이나 외국인가정 자녀 등과 같이 한국에서 태어나지 않았거나 한국어가 아닌 다른 언어를 제1언어로 하여 한국어 의사소통 능력이 없는 학령기 학생을 대상으로 개발된 한국어 교육과정으로(교육부, 2017:1) 초등학교, 중학교, 고등학교 단계로 나누어 단계별로 학습이 필요한 주제, 의사소통 기능, 텍스트 유형, 문법, 어휘 목록을 제시하고 있다. 이 중 어휘 목록은 '생활 한국어 어휘'와 '학습 한국어 어휘'로 나누어서 제시하고

[53] Skehan(2009a:514)에서는 어휘 측정을 텍스트 자체 내에서 계산된 텍스트 내부 측정과 텍스트 외부 측정으로 나누고 있는데 이에 따르면 어휘 다양도와 어휘 밀도는 내부 측정에 해당되고 어휘 정교도는 외부 측정에 해당된다.

[54] 국립국어원(2012, 2013, 2014)과 국립국어원(2015)은 '한국어 교육 어휘 내용 개발'의 1~4단계 보고서로 같은 어휘 목록이다.

[55] 여기에서 '초급 단계'는 한국어능력시험(TOPIK)의 초급 1급과 2급, 즉, 한국어 학습 400시간에 해당하는 수준을 의미한다.

있는데 ‘생활 한국어 어휘’는 다시 ‘일상생활 어휘’와 ‘학교생활 어휘’로 나누고 ‘학습 한국어 어휘’는 ‘학습 도구 한국어 어휘’와 ‘교과 적응 한국어 어휘’로 나뉜다. 본고에서는 연구 대상이 중학생이고 또 언어 산출에 사용된 쓰기와 말하기 과제가 모두 일상생활을 주제로 하고 있으므로 어휘 정교도 측정에 중학교 학습자료 중 ‘일상생활 어휘’목록을 사용하도록 한다.

한편, 어휘 정교도 측정은 같은 어휘 목록을 사용하더라도 그 측정 방법에서 차이가 있는데 첫째는 초급 어휘, 즉 고빈도 어휘 목록을 정한 뒤 고빈도 어휘를 제외한 나머지 어휘들을 저빈도 어휘로 보고 그 비율을 측정하는 방법이고 다른 한 가지 방법은 중·고급 어휘 목록을 선정한 뒤 학습자의 텍스트에서 나타난 중·고급 어휘의 비율을 측정하는 방식이다. 학습자의 텍스트에 나타난 중·고급 어휘 목록에 해당되는 어휘의 비율을 측정하는 방법은 비교적 직관적으로 중급이나 고급 어휘의 비율을 측정할 수 있다는 장점이 존재하지만 어휘 목록에 포함되지 않은 어휘에 대해서는 측정이 불가하다는 단점이 존재한다. 이는 선행연구에서도 확인할 수 있는데 주월랑(2018)에서는 국립국어원(2015)의 초, 중, 고급 어휘 목록을 사용하고 모든 목록에 나타나지 않는 어휘를 X등급으로 설정하였고 주우동(2019a)역시 같은 어휘 목록을 사용하였으나 어휘 목록에 나타나지 않은 어휘에 대해서는 최고급 어휘로 따로 분류하기도 하였다. 특히 중도입국 청소년 학습자는 한국의 일반 중학교의 수업에 참여하고 있다는 점에서 이들의 어휘 습득이 교육과정의 순서를 그대로 따르지 않을 것은 예상 가능한 부분이다. 그러므로 중급이나 고급과 같이 이미 한정된 목록에 해당되는 어휘의 비율을 측정하는 방법보다는 고빈도 어휘를 제외하는 방법이 좀 더 적합할 것으로 보인다. 또한 현재 <한국어(KSL) 교육과정(2017)>의 일상생활 어휘 목록은 초급과 중급 어휘만 제시되어 있으므로 초급 어휘를 제외한 나머지 어휘를 저빈도 어휘로 보는 것이 실제적인 방법일 것이다.

<한국어(KSL) 교육과정(2017)>에서 제시한 <중학교 언어 재료>의 ‘일상생활 어휘 목록’에서는 초급 어휘 1,795개와 중급 어휘 3,717개를 제시하고 있다. 따라서 본고에서는 학습자의 텍스트에 나타난 어휘 중 ‘중학교 일상생활 초급 어휘’에 해당되지 않는 어휘의 비율을 측정하는 것으로 학습자의 어휘 정교도를 살펴보고자 하며 목록에 포함되지 않는 어휘는 ‘X등급’으로 분류하도록 한다.

3.3. 정확성 측정 지표 및 분석 기준

3.3.1. 오류 판정 기준

2장에서 제시하였듯이 본고에서는 정확성을 일반적 측정인 무오류절의 비율, 단위당 무오류절 수, 단위당 오류 수로 측정하고자 한다. 이때, 무오류절의 비율은 오류가 없는 절의 수를 전체 절 수로 나눈 것이고 단위당 무오류절 수는 오류가 없는 절의 수를 전체 단위 수(문장 수/AS-unit)로 나누어서 측정한다. 마지막으로 단위당 오류 수는 오류의 유형과 관계없이 언어 자료에 나타난 오류의 수를 전체 단위 수로 나누어 측정하도록 한다.

학습자가 산출한 문장에 대해 정확성을 판정하기 위해서는 오류에 대한 정확한 판단이 선행되어야 한다. 즉 오류 혹은 무오류를 판정하기 위해서는 기준이 되는 정확한 문장이 있어야 한다. 본고에서는 앞서 2장에서 제시한 바와 같이 '성인 모어화자의 언어'를 기준으로 적형성(문법성)과 적절성의 위배 여부에 따라 정확성을 판단하도록 한다. 그러나 오류를 정의하고 분석하는 것은 분석자의 관점에 따라 차이가 있으며 같은 한국어 모어화자라고 할지라도 오류에 대한 판단 기준은 매우 주관적이며 상이하게 나타난다.[56] 그러므로 정확성 분석에서 연구자의 개인적 주관에 의해 오류 판단이 달라지는 것을 방지하기 위해서는 오류 판정의 기준을 명시적으로 제시할 필요가 있다. 본고의 오류문 교정의 원칙은 다음과 같다.

① 오류 교정은 학습자가 산출한 형태를 가능한 한 훼손하지 않는 전제에서 학습자의 표현 의도를 고려하여[57] 교정하는데 이 때 최소교정을 원칙으로 하며 교정한 문장은 '한국어 성인 모어화자의 언어'에 따른다.

② 학습자의 의도를 파악할 수 없거나 문장의 전면 교정이 필요한 문장은 교정을 하면 연구자의 해석이 지나치게 반영될 수 있으므로 분석 불가능 범주에 포함시킨다.

---

[56] 석주연·안경화(2003)에서는 한국어 모어화자 20명을 대상으로 오류의 정도가 각기 다른 한국어 학습자의 오류문을 보여준 뒤, 연구참여자의 직관에 따라 각 예문의 오류 여부를 판단하게끔 하였다. 그 결과 같은 예문이라도 오류와 비오류의 판단이 서로 엇갈림을 확인하였다.

[57] 연구자는 1:1로 자료 수집을 진행하였기에 학습자가 작문 쓰기를 완료한 직후 바로 내용을 확인하여 학습자가 표현하고자 하는 바가 명확하지 않은 부분에 대해서 학습자의 모국어로 질문을 하고 학습자가 표현하고자 하는 의도를 따로 메모해 두었다. 구어 역시 녹음이 끝난 직후, 잘 들리지 않은 내용이나 연구자가 이해하기 힘든 문장에 대해 학습자의 표현 의도를 확인하여 메모해 두었다.

③ 맥락에 따라 교정하되 내용어보다 기능어 교정을 우선으로 한다.
④ 높임법은 동일 문장 내 높임 표현의 호응에 따라 필수적인 경우에만 오류로 삼는다.[58]
⑤ 문장 부호는 오류에 포함하지 않는다.
⑥ 오류의 교정은 기본적으로 문장 단위에서 이루어진다.[59] 그러나 담화 오류에 해당하는 지시어나 접속어의 경우 선행문과 후행문의 의미적 연결을 고려해야 오류 교정이 가능하므로 앞, 뒤 문장을 고려하여 오류를 판정하고 교정한다.

3.3.2. 오류 분류 기준

모국어 화자이든 제2언어 학습자이든 언어를 학습하는 과정에서 오류는 피할 수 없는 것이다. 특히 제2언어 학습자들은 학습 단계별, 언어권별, 그리고 연령이나 성별에 따라서 오류의 양상이 다르게 나타남은 그동안 여러 연구들을 통해 잘 알려진 사실이다. 따라서 학습자가 얼마나 많은 오류를 생산하는지 보다 더 중요한 것은 학습자가 어떤 오류를 생산하는지를 관찰하고 설명하는 것이다. 오류에 대해 설명을 하기 위해서는 오류에 대한 타당한 기술에서 출발해야 한다. 오류의 기술을 위한 첫 단계는 오류의 종류를 범주화하는 것이다. 그동안의 선행연구를 살펴보면 오류에 대한 분류 기준은 오류를 바라보는 관점에 따라 달라 여러 가지로 나뉘는데 크게 오류의 원인에 따른 분류와 결과의 판정에 따른 분류로 구분할 수 있다.

오류의 원인에 따른 분류는 Richards(1971), Selinker(1972)와 Brown(2000) 등이 있다. Richards(1971:205)에서는 오류를 모국어 간섭 오류(interference error), 언어 내 오류(intralingual error), 발달 오류(developmental error) 세 가지로 구분하고 언어 내 오류와 발달 오류는 특히 특정 단계에서의 학습자의 능력과 언어 습득의 일반적인 특성 중 일부를 설명한다고 하였다. Selinker(1972:216-217)에서는 제2언어 습득 과정에서 학습자가 생산하는 오류를 포함한 학습자 언어를 중간언어(interlanguage)로

[58] 한국어의 높임법은 주체 높임, 상대 높임, 어휘에 의한 객체 높임 등 담화 층위에서 화자와 청자의 관계에 따라 실현되는 것으로 일상생활에서 모어화자들도 지키지 않는 경우가 많아 높임 표현에 대한 적절성과 용인성의 기준을 일관적으로 규정하기가 어렵다(국립국어원, 2018:198). 본고에서는 높임법의 오류 판정은 문장 내의 호응을 고려하여 판정하도록 한다. 예를 들면 주격 조사 '께서'를 사용하였지만 서술어에서 선어말 어미 '-으시-'를 생략한 경우처럼, 문장 내에서 한 부분이라도 높임법을 사용한 경우에는 그 문장을 높임법을 적용한 문장으로 교정하여 오류를 판단한다.

[59] 담화 층위는 문장 단위를 넘어서 발생하는 오류를 다루는 것으로 그 범위가 넓고 체계화와 판정 기준이 쉽지 않다. 따라서 본 연구에서는 담화표지, 지시, 접속과 같이 표지가 비교적 분명하고 판정 기준이 명확한 항목만을 교정의 대상으로 삼는다.

보고 이러한 오류를 일으키는 원인으로 첫째, 모국어 전이로 인한 오류. 둘째, 교육 과정(training procedure)에서 발생하는 교육 전이(transfer-of-training)로 인한 오류. 셋째, 교육자료에 대한 접근에서 발생하는 제2언어 학습 전략으로 인한 오류. 넷째, 목표어 모국어 화자와의 의사소통 과정에서 이루어지는 제2언어 의사소통 전략으로 인한 오류. 다섯째, 목표언어의 규칙과 의미에 대한 과잉일반화(overgeneralization)로 인한 오류로 구분하였다. 마지막으로 Brown(2000:284-287)에서는 오류의 원인을 크게 언어 간 전이(interlingual transfer), 언어 내 전이(interlingual transfer), 학습의 맥락(context of learning), 의사소통 전략(production strategy)의 네 가지로 나누고 언어 간 전이를 긍정적 전이와 부정적 전이로 나누고 오류는 부정적 전이로 인한 것으로 보았다.

결과의 판정에 따른 오류 분류에는 Burt(1975), Dulay 외(1982)이 있는데 Burt(1975:56-57)에서는 오류가 영향을 주는 범위에 따라 전체적 오류(global error)와 부분적 오류(local error)로 나누었다. 여기에서 전체적 오류는 전체 문장 조직에 영향을 주는 오류를 가리키고 부분적 오류는 의사소통에 크게 영향을 주지 않는 단일 요소인 명사, 동사, 굴절관사, 조동사, 수량사 등 오류를 가리킨다. 그리고 Dulay 외(1982:Ch7)에서는 학습자의 오류를 크게 언어학적 범주 유형(linguistic category classification)과 표면 전략적 유형(surface strategy taxonomy)으로 분류하였는데 언어학적 범주는 언어적 층위(level)인 음운, 형태소, 문법 차원 및 그 하위의 언어 항목(class)인 동사, 명사, 전치사, 관사 등으로 나누고 표면 전략 유형은 누락(omission), 첨가(addition), 잘못된 변형(misformation), 잘못된 어순(misordering) 등 오류의 표면 양상으로 나누었다. Carl James에서는 이를 수용하고 변형하여 생략(omission), 과대포함(overinclusion), 잘못 선택(misselection), 잘못된 어순(misordering)로 구분하기도 하였다.

오류를 원인에 따라 분류하는 것과 결과에 따라 분류하는 것은 모두 각각의 이점이 있지만 이는 연구 목적에 따라 다를 수 있다. 이론적인 목적을 위한 연구는 학습자의 제2언어 습득 과정을 밝히기 위한 것으로 학습자의 중간언어(interlanguage) 전반에 대해 총체적이면서도 기술적인 방법으로 이루어진다. 이와 달리 교육적인 목적을 위한 연구는 교육 현장에서 오류분석의 결과를 활용하기 위한 것으로 오류의 객관적 분석보다는 학습자의 오류를 확인하고 유형을 기술하며 원인을 추적하고 평가하여 효율적으로 오류를 수정하는 것에 초점을 맞춘 실용적인 분석이다(석주연·안경화, 2003:189-190). 언어 교육 현장에서 오류가 발생하는 원인을 분석하고 수정하는 것은 매우 중요

한 일이지만 그보다 선행되어야 할 것은 학습자의 제2언어습득 과정을 이해하고 오류를 예방하는 것이다. 물론 본 연구 역시 오류의 유형을 기술하고 교육적 함의를 찾는다는 점에서 교육적인 목적을 완전히 배제할 수는 없으나, 연구의 목적이 중도입국 청소년의 한국어 발달 과정을 고찰하기 위한 것이므로 이론적인 목적을 위한 연구에 가깝다고 볼 수 있겠다. 따라서 오류의 원인보다는 오류의 양상을 기술하는 것에 초점을 두고자 한다.

또한 오류의 원인에 따른 분류는 그 분류 기준을 명시적으로 제시하기 어려우며 연구자의 주관적인 판단이 크게 개입될 수 있다는 단점이 있다. 반면 결과에 따른 분류는 한국어의 언어 체계를 기본적인 틀로 그 위치나 층위에 나타난 오류 현상을 유형화한다는 점에서 비교적 명시적인 기준을 제시할 수 있다는 장점이 있다. 그러므로 본 연구에서는 오류의 결과를 기준으로 학습자의 오류를 유형화하여 살펴보고자 한다.

한국어교육 분야에서 학습자의 오류 전반에 대해 관찰하고 분류한 연구는 대표적으로 이정희(2002), 고석주 외(2004) 조철현(2002), 국립국어원(2018)이 있는데 이정희(2002:49)에서는 오류의 유형을 내용적인 측면과 형식적인 측면 두 가지로 나누고 내용적인 측면은 오류의 원인에 따른 유형 분류, 형식적인 측면은 오류의 판정에 따른 분류로 구분하였다. 그리고 결과의 판정에 따른 분류를 다시 범주별(발음, 어휘, 문법 등) 오류, 현상(대치, 누락, 첨가)에 따른 오류, 정도(전체, 부분)에 따른 오류로 구분하였다. 그리고 조철현(2002)과 고석주 외(2004)에서는 오류의 유형을 크게 어휘 오류, 문법적 오류, 오류의 혼합, 철자 오류로 나누고 문법적 오류의 하위 분류로 조사 오류, 어미 오류, 어순 오류, 표현 오류, 문장 전체 오류로 구분하였다. 또한 오류 현상으로는 생략, 잘못 씀, 첨가, 대치, 환언, 형태 오류로 나누었다. 마지막으로 국립국어원(2018)에서는 오류 주석 체계로 오류 위치(명사, 형용사, 조사, 어미 등)와 오류 층위(발음, 형태, 통사, 담화), 오류 양상(누락, 첨가, 대치, 오형태)로 나누고 이 외에 분석 불가능 범주를 설정하고 있다.

네 연구를 비교해 보면 오류의 문법 영역별 분류에 있어서는 크게 차이가 없이 어휘, 문법, 표현, 맞춤법, 그리고 오류의 혼합이나 분석 불가 범주로 나누고 있음을 알 수 있다. 그러나 이정희(2002)에서는 표현 오류를 따로 구분하지 않고 모두 어휘 오류에 포함시키고 있다는 점에서 차이가 있었다.[60] 오류의 현상의 분류에 있어서는 연구마다 차이가 있는데 이정희(2002)는 오류 현상을 대치, 누락, 첨가로 구분하였고 국립국어원

[60] 이정희(2002:83)에서는 어휘 오류의 범주 한정에 대한 어려움을 언급하면서 '형태통어적 구성으로 볼 수 있는 '-기 때문에', '-(으)ㄹ 수 있다' 등은 모두 어휘 오류로 처리한다'고 하였다.

(2018)에서는 대치, 누락, 첨가 외에 오형태를 추가하였다. 그리고 조철현(2002)과 고석주 외(2004)에서는 동일 범주 대치, 다른 범주 대치, 다른 것으로 환언, 생략, 형태 오류, 첨가로 나누고 있다. 조철현(2002), 고석주 외(2004)의 분류를 이정희(2002)와 비교하면 오류 현상에 있어 조철현(2002)에서 '형태 오류'와 '환언 오류' 범주가 더 많음을 알 수 있다. 조철현(2002)에서 정의하는 환언 오류는 어휘, 조사, 어미 등에서 적절한 한국어 표현이 있음에도 불구하고 다른 적절하지 않은 어휘나 조사, 어미로 환언하여 쓰는 경우를 가리키는데 그 예로 다음과 같은 문장을 제시하고 있다.

(15) 가. 알지 않는(√모르는) 사람에게 인사를 했어요.
나. 단 특급이나 급행은 탈 수 없다. 그 표를 써서(√표로) 자주 여행을 했었지.

조철현(2002: 51, 56)

그러나 이를 환언 오류로 볼 경우 수정할 수 있는 방안이 여러 가지로 나뉠 수 있으며 이 과정에서 학습자의 의도를 곡해하고 불필요한 교정을 할 수 있다. 위의 예문에서 '알지 않는'은 '알지 못하는'으로 수정할 수도 있고 '그 표를 써서'는 '그 표를 사용해서'로 수정할 수도 있으나 오류로 보지 않을 수도 있는 부분이다. 따라서 조철현(2002)에서 정의하는 환언 오류는 문법성보다는 적절성의 문제에 더 가까운 것으로 보인다. 또한 환언 오류는 결국 대치에 의한 오류이므로 본고에서는 '환언 오류' 범주는 따로 설정하지 않기로 한다.

'형태 오류' 범주는 이정희(2002)에서는 설정하지 않았고 조철현(2002), 고석주 외(2004)와 국립국어원(2018)에서는 설정한 오류 현상인데 같은 어휘를 사용하고 있지만 포함하고 있는 오류 유형이 다름을 확인할 수 있었다. 먼저 조철현(2002)에서는 형태 오류에 대해 주로 음운적 환경에 따라 이형태가 있는 경우 적절한 형태를 선택하지 못한 경우를 가리키고 맞춤법 오류는 형태 오류에 포함하지 않은 반면 국립국어원(2018:177)에서는 형태 오류에 대해 '오철자 오류, 용언 활용, 조사 이형태 곡용, 어미 활용 등 형태를 잘못 활용한 경우를 포함한다'고 규정하며 맞춤법 오류도 오형태에 포함시키고 있다. 이와는 또 다르게 이정희(2002)에서는 활용이나 곡용에 나타난 이형태 오류는 대치 오류에 포함시키고 맞춤법 오류 범주를 따로 구분하고 있다. 그러나 같은 문법의 이형태 선택에 오류가 생기는 것과 완전히 다른 형태로 대치를 하는 것은 분명히 다른 오류이다. 예를 들면 조사 오류의 경우 '나는 밥를 먹었어요'와 '나는 밥이 먹었어요'를 비교하면 학습자는 전자는 목적격 조사 선택에는 문제가 없으나 앞의 자모음

에 따른 형태 변환에만 오류를 보인다고 할 수 있으나 '밥이 먹었어요'로 산출할 경우 주격 조사와 목적격 조사를 헷갈리는 것이기 때문에 오류의 성격이 다르다고 볼 수 있다. 따라서 본고에서는 범주 간 대치 오류와 형태 오류를 서로 구분하고자 한다. 물론 어미나 조사에 나타난 이형태 오류 역시 일종의 대치 오류로 볼 수도 있고 또 문법의 의미 기능에 대한 이해와 사용에 문제가 없다는 점에서 대치나 누락, 첨가와 동등한 층위에 배치하기 어렵다는 문제가 존재할 수도 있지만 형태 오류는 단순 철자 오류와는 큰 차이가 있다는 점에서 대치, 누락, 첨가와 함께 하나의 오류 현상으로 구분하여 학습자의 오류를 좀 더 면밀히 분석하고자 한다.

이 외에 본고에서는 발음 오류와 철자 오류는 다루지 않는다. 발음과 철자 오류를 분석하지 않는 첫 번째 이유는 본 연구는 학습자의 문장 구성 능력의 발달을 중심으로 보는 것으로 발음과 철자 오류를 포함할 경우 학습자 개개인의 발음 습관이나 특성에 따라 각 오류의 비율과 오류율에 영향을 주게 되기 때문이다. 또한 본 연구가 문어와 구어의 오류율을 비교한다는 점에서 철자와 발음 오류를 포함할 경우 발음 오류와 철자 오류를 같은 층위에 놓게 된다는 문제가 존재한다. 이러한 이유로 인해 본고에서는 철자와 발음 오류를 제외한 문법과 어휘 오류만 다루도록 한다. 위의 선행연구들을 살펴본 결과를 바탕으로 본고에서는 오류 영역별 분류로 어휘 오류, 문법 오류, 맞춤법 오류로 나누고 오류 현상별로 대치, 누락, 첨가, 오형태로 나누고자 한다. 구체적인 분류는 다음의 <표 11>과 같다.

<표 11> 본 연구의 오류 분류 체계

<table>
<tr><th colspan="3">오류 영역</th><th>오류 구분</th><th>오류 양상</th></tr>
<tr><td rowspan="10">문법 오류</td><td colspan="2" rowspan="3">조사오류</td><td>격 조사</td><td rowspan="9">대치
누락
첨가</td></tr>
<tr><td>보조사</td></tr>
<tr><td>접속 조사</td></tr>
<tr><td rowspan="5">어미 오류</td><td rowspan="3">어말 어미</td><td>종결 어미</td></tr>
<tr><td>연결어미</td></tr>
<tr><td>전성어미</td></tr>
<tr><td rowspan="2">선어말 어미</td><td>시제 선어말 어미</td></tr>
<tr><td>높임 선어말 어미</td><td>오형태</td></tr>
<tr><td colspan="2">표현 오류</td><td colspan="2"></td></tr>
<tr><td colspan="2">어순 오류</td><td colspan="2"></td></tr>
<tr><td colspan="3">어휘 오류</td><td colspan="2"></td></tr>
<tr><td colspan="3">기타 오류</td><td colspan="2"></td></tr>
</table>

본고에서 정의하는 오류 현상별 정의와 이에 포함되는 오류는 다음과 같다.

1) **대치 오류:** 어휘나 문법의 의미, 용법을 제대로 알지 못하여 의미적으로 서로 다른 어휘나 문법을 바꾸어서 쓴 오류를 가리킨다.
2) **누락 오류:** 완전한 문장 또는 발화에서 필수적으로 있어야 할 형태가 빠져 있는 오류를 가리킨다.
3) **첨가 오류:** 완전한 문장이나 발화에서 불필요한 요소가 추가적으로 나타나거나 같은 요소를 중복적으로 사용한 오류를 가리킨다.
4) **오형태 오류:** 어휘나 문법의 조합 양상이나 활용 과정에 나타난 형태상의 오류로 조사와 어미의 이형태 오류, 어휘의 파생이나 합성 오류 등을 가리킨다.

아래, 각 영역에 나타나는 오류의 대표적인 유형을 제시하여 오류 분류의 일관성을 높이고자 한다.

1) 조사 오류

한국어의 조사는 주로 체언에 첨가되어 격을 표시해 주는 기능을 한다. 그러나 문법적인 기능 외에 다양한 의미를 나타내는 조사도 있다. 본고에서는 조사에 대해 격 조사, 보조사, 접속 조사로 나누어 오류를 살펴보도록 한다. 조사의 오류는 그 판정 기준이 비교적 분명한 편이다. 조사의 대치 오류는 다른 조사로 바꾸어서 쓴 경우로, 누락 오류는 필요한 조사를 생략한 경우로, 첨가 오류는 불필요한 조사를 첨가하거나 두 개 이상의 조사를 연속적으로 사용한 경우로, 그리고 오형태 오류는 선행하는 체언의 자모음에 따른 이형태 선택에서 오류가 발생한 경우로 나누었다.

2) 어미 오류

한국어의 어미는 어말 어미와 선어말 어미로 구분할 수 있고 어말 어미는 다시 종결어미, 연결어미, 전성어미로 구분된다. 본고 역시 이 기준에 따라 어미 오류를 분류하였다. 선어말 어미는 높임을 나타내는 선어말 어미 '-시-'와 시제를 나타내는 '-는-, -었-, -겠-'으로 구분하였다. 어미의 대치 오류는 다른 어미로 바꾸어서 쓴 경우이고 누락 오류는 필요한 위치에서 어미가 생략된 경우, 어미 첨가 오류는 불필요한 어미를 첨가하거나 두 개 이상의 어미를 반복적으로 사용한 경우이다. 어미의 오형태는 용언의 받침 유무에 따라 어미의 이형태 선택이 달라지는 경우로 이형태가 있는 어미에서 형태 선택이 정확하지 않은 경우이다. 즉 '힘들습니다(√힘듭니다), 가아요(√가요)'와 같이 필

수적으로 탈락을 시켜야 하는 경우임에도 탈락시키지 않은 경우와 '쉬요(√쉬어요), 아름답니다(√아름답습니다)'와 같이 과도한 탈락으로 인한 축약 모두 오형태로 처리한다. 마찬가지로 종결 어미 '아요/어요'에서 '요'만 쓴 경우, 'ㅂ니다/습니다'에서 '니다' 또는 '입니다'를 잘못 사용한 경우도 종결 어미 오형태 오류로 처리한다.

3) 표현 오류

표현 오류는 경우에 따라 어휘 오류로 분류될 수도 있고 특정 문법의 오류로 분류될 수도 있다. 그러나 본고의 학습자의 경우 '-ㄹ 수 있다, -ㄹ 것 같다, -ㄴ/는 것 같다'와 같은 양태적 표현의 전체적 누락이나 '왜냐하면 …기 때문이다'와 같은 호응 표현 중 '-기 때문이다'의 누락, '본용언 + 보조 용언'의 구성에서 보조적 연결어미나 보조 용언을 누락하거나 잘못 사용하는 경우가 빈번히 발생하고 있었다. 따라서 이와 같이 둘 이상의 요소의 결합 구성으로 이루어진 문형에서 나타난 오류는 표현 오류로 구분하도록 한다.

4) 어순 오류

한국어는 어순이 자유로운 편이나 조사나 부사 어순 혹은 수량 명사구의 어순은 고정된 것이기 때문에 오류가 발생하기도 한다. 즉 어순 오류는 한국어의 통사 구조에 맞지 않는 방식으로 문장이 배열된 경우를 가리킨다.

5) 어휘 오류

어휘 오류[61]는 명사, 동사, 형용사, 부사와 같은 실질 어휘에 나타난 오류를 가리킨다. 어휘의 오류에는 다른 어휘로의 대치, 불필요한 어휘를 첨가하는 첨가 오류, 필요한 어휘를 누락한 누락 오류, 용언의 활용과 체언의 파생·합성 과정에 나타난 오형태 오류가 있다. 사동과 피동은 경우에 따라 다른 범주에서 다루어지게 되는데 본고에서는 '-이/히/리/기/우/구/추-'에 의한 사동사와 '-이/히/리/기-' 에 의한 피동사는 어휘 범주에서 다루고, '-아/어지다'와 '-게 되다'에 의한 피동은 표현의 범주에서 다루며 '-게 하다'는 어미의 범주에서 다루도록 한다.[62]

---

[61] 본고에서 지칭하는 어휘는 실질 어휘를 가리킨다.

[62] '-게 되다'와 '-게 하다'의 오류 분류가 다른 것은 앞서 통사 복잡성 분석 기준에서 제시한 바와 같이 본 연구에서 '-게 되다'는 보조 용언의 범주에서 다루지만 '-게 하다'는 부사형 어미 '-게'를 포함한 부사절로 분석하기 때문이다.

6) 기타 오류

위의 오류 분류에 포함되지 않는 오류는 기타 오류로 분류하였다. 기타 오류에는 완결되지 않은 문장이나 일부 절이 중국어로 작성된 문장, 의미 해석이 어렵거나 여러 오류가 혼재되어 있어 분석이 불가능한 오류, 문장 성분의 누락이나 첨가로 인한 오류 등을 포함한다.

7) 구어 오류 판정을 위한 추가적 기준

구어의 오류 판정은 문어와 마찬가지로 적형성과 적절성의 기준을 적용한다. 그러나 구어는 문어에 비해 조사 생략이 자유롭고 구어 현실 발음이 존재한다는 차이점이 있다. 따라서 구어에서 자연스러운 조사 생략은 오류로 판정하지 않는다. 그러나 자연스러운 조사 생략에 대한 기준은 사람마다 상이하게 나타난다. 따라서 구어의 조사 오류는 연구자와 연구자 이외의 한국어 모어화자 두 명이 함께 판단하였으며[63] 세 명의 합의가 도출되지 않은 일부 문장에 대해서는 제3의 모어화자를 통해 판단하였다. 또한 생략이 가능한 위치라고 하더라도 해당 위치에 학습자가 틀린 조사를 사용하였다면 대치 오류를 우선적으로 적용하였다. 또 구어에는 구어의 특성이나 표현 의도에 따른 특징적인 발음이나 모어화자들에게서도 자주 나타나는 현실 발음[64]이 있는데 이러한 발음들은 오류로 처리하지 않도록 한다. 이 외에 구어에서는 준말 역시 자주 사용되는데 예를 들면 '맘(√마음)'과 같은 말들이다. 준말에 대해서는 <표준국어대사전>을 바탕으로 준말로 등재되어 있는 것은 오류로 보지 않고 등재되지 않은 형태는 오류로 처리한다. 지금까지 살펴본 복잡성과 정확성의 측정 지표를 정리하면 다음의 <표 12>와 같다.

<표 12> 복잡성과 정확성 측정 지표

구분	구인	측정 지표	계산 방법
문어	통사 복잡성	문장 길이	총 형태소 수/총 문장 수
		문장당 절 수	총 절 수/총 문장 수
	어휘 복잡성	어휘 다양도	어휘 유형 수/어휘 구현 수×2의 제곱근
		어휘 밀도	내용어 수/총 어휘 수
		어휘 정교도	'중학교 일상생활 초급 어휘 목록'에 해당되지 않는 어휘 수/총 어휘 수
	정확성	무오류절 비율	총 무오류절 수/총 절 수

[63] 두 명의 한국어 모어화자는 연구자와 같은 학교에 재학 중인 대학원생으로 모두 중도입국 청소년을 대상으로 2년 정도의 한국어 교육 경험을 가지고 있다.

[64] 예를 들면 '할려고(하려고), 그리구(그리고), 좀(쫌), 쪼금(조금)'과 같은 발음들이다.

		문장당 무오류절 수	총 무오류절 수/ 총 문장 수
		문장당 오류 수	총 오류 수/총 문장 수
구어	통사 복잡성	AS-unit 길이	총 형태소 수/총 AS-unit 수
		AS-unit당 절 수	총 절 수/총 AS-unit 수
	어휘 복잡성	어휘 다양도	어휘 유형 수/어휘 구현 수×2의 제곱근
		어휘 밀도	내용어 수/총 어휘 수
		어휘 정교도	‘중학교 일상생활 초급 어휘 목록’에 해당되지 않는 어휘 수/총 어휘 수
	정확성	무오류절 비율	총 무오류절 수/총 절 수
		AS-unit당 무오류절 수	총 무오류절 수/ 총 AS-unit수
		AS-unit당 오류 수	총 오류 수/총 AS-unit 수

4. 자료 분석 절차 및 방법

연구를 위해 3장의 1절과 2절에 서술한 기준으로 연구 대상을 모집하고 언어 자료를 수집한 뒤, 3절에서 제시한 측정 지표와 분석 기준에 따라 분석을 진행하였다. 자료의 분석은 수집된 자료의 입력, 자료의 분석 및 코딩, 그리고 통계 분석의 순서로 진행되었으며 각 단계별 세부 내용은 다음과 같다.

4.1. 자료 수집 및 입력

1) 자료 수집

본 연구의 자료 수집은 2021년 10월~2023년 3월까지 진행되었다. 연구 참여자와 언어 산출 과제에 대해서는 3장의 1,2절에서 자세히 서술하였다. 본 연구의 자료 수집은 1:1로 진행되었으며 연구자가 학습자의 학교를 방문하여 빈 교실에서 진행하거나 비대면 화상회의(zoom)를 통해 진행되었다. 앞에 제시하였듯이 과제 수행 준비나 산출 시간에는 제한을 두지 않았으나 학습자의 자연스러운 언어를 수집하기 위해 사전이나 인터넷을 사용한 검색은 하지 못하도록 하였다. 쓰기와 말하기는 주제가 서로 다르기 때문에 정해진 순서 없이 학습자가 원하는 순서로 진행하였으나 2~3차례를 제외하고 학습자들은 쓰기를 먼저 수행하는 것을 선택하였다. 대면으로 자료 수집을 진행할 경우, 연구자는 학습자가 쓴 작문을 사진을 찍어서 보관하고 말하기는 녹음을 진행하였다. 화상회의를 통해 자료 수집을 진행할 경우, 학습자 본인이 자신의 작문을 사진으로 연구자에게 위챗(Wechat)을 통해 전송하였으며 말하기는 zoom의 녹화 기능을 사용하

여 기록되었다. 연구 참여자의 개인정보보호를 위해 녹화 이후 영상 자료는 즉시 삭제하였으며 음성 자료는 연구자의 개인 노트북에 보관하였다.

2) 자료 입력

연구자는 학습자의 사진 자료와 음성 자료를 바탕으로 모든 내용을 전사하는 과정을 거쳤다. 전사는 기본적으로 학습자가 산출한 내용을 그대로 옮겨 적는 것을 원칙으로 하며 학습자의 오류 또한 그대로 반영한다. 문어와 구어의 전사는 워드 2019(Microsoft Word 2019) 프로그램을 통해 이루어졌으며 문어 자료의 전사는 다음과 같은 원칙을 바탕으로 이루어졌다.

① 전사 과정에서는 학습자가 산출한 원문을 그대로 반영한다.
② 문어 자료에서 띄어쓰기는 원문에서 오류가 있더라도 연구자가 띄어쓰기 원칙에 맞게 수정하여 입력한다.
③ 문어 자료의 한글 이외의 한자, 영어, 문장 부호 등은 원문 그대로 입력한다.
④ 키보드에 없는 문자나 식별이 불가한 문자의 경우 거꾸로 된 물음표(¿)를 사용하여 입력한다.
⑤ '예쁘ㅏ요'와 같이 한글에 없는 문자 역시 거꾸로 된 물음표(¿)를 사용하여 입력한다.
⑥ 오류가 있는 문장에 대해서는 오류를 그대로 적고 그 뒤 괄호 안에 체크(√) 표시와 함께 맞는 표현을 기재한다.

<표 13> 문어 자료 전사 예시

내가 좋아하는 과목은 수학이에요. 왜냐하면 수학가(√수학은) 너무 재미있어요(√재미있기 때문이에요). 그리고 挑战性를(√도전성이) 있어요. 모르는 문제를 풀으면(√풀면) 기분가(√기분이) 너무 좋아요.(AW11[65])

다음은 구어 자료에 대한 전사이다. 구어 자료에서 학습자의 발화를 그대로 담아 내기 위해서는 학습자가 산출한 모든 언어를 전사해야 한다. 전사는 담화표지나 주저어, 간투사를 포함하여 모든 내용을 전사하도록 하며 전사 원칙은 다음과 같다.

① 전사 과정에서는 학습자가 산출한 원문을 그대로 반영한다.

[65] 첫 알파벳 A, B, C는 학습자를 가리키고 두 번째 알파벳은 산출 양식을 가리키는 'W'는 문어 자료, 'S'는 구어 자료를 가리킨다. 그리고 알파벳 뒤의 숫자는 자료 수집의 회차를 의미한다.

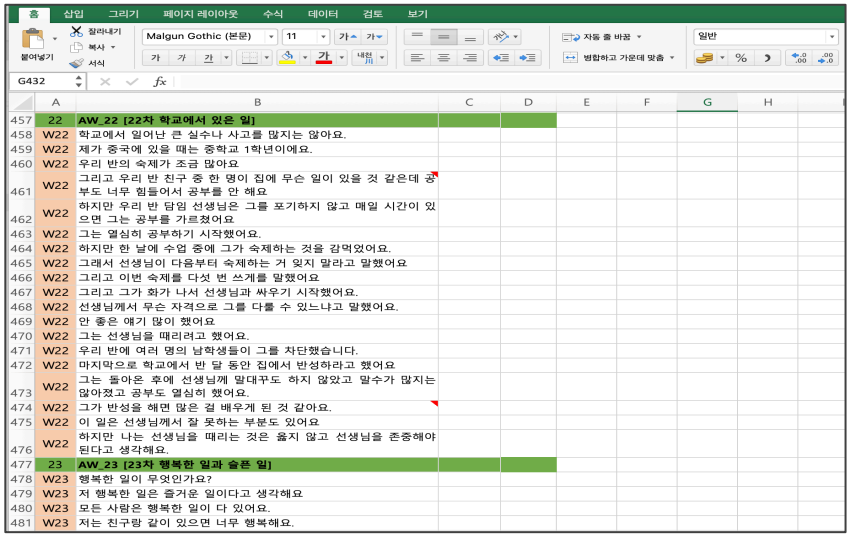
② 외래어는 발음에 따라 한글로 표기하고 외래어가 아닌 외국어에 대해서는 그대로 전사한 뒤 괄호 안에 한국어 발음을 표시한다.
③ 제대로 들리지 않은 음절은 음절 수만큼 X로 표시한다.
④ 화자가 길게 발음한 부분은 '~'로 표기한다.
⑤ 0.5초 이상의 휴지는 ( )안에 표시하며 0.1초 단위까지 표시한다.
⑥ 오류가 있는 문장에 대해서는 오류를 그대로 적고 그 뒤 괄호 안에 체크(√) 표시와 함께 맞는 표현을 기재한다.

<표 14> 구어 자료 전사 예시

그런데 어 엄 다른 국가 사람(√사람들이) 많이 있는데 种族(√종족) 차이 별로 없어요. 우리 학교가(√학교는)~ 공부에 스트레스 별로 없어요. 어 배우는 거 다(0.6) 쪼금 쉬워요. 그런데 보통 보통 한국 사람(√사람은) 어 쉬우는데(√쉬운데) 제가 어 외국인이니까 언어 문제 조금 있어요.(BS28)

전사된 파일은 분석을 위해 엑셀 2019(Microsoft Excel 2019) 프로그램으로 옮겨졌다. 엑셀 프로그램으로 옮기는 과정에서는 한 셀당 하나의 문장이나 하나의 AS-unit을 입력하여 분석에 용이하도록 하였다. 이 때, 구어 자료는 휴지를 참고하여 기본 분석 단위를 구분하여 옮기며 분석에 필요하지 않은 간투사나 문장 부호 등은 제외시킨다. 이러한 과정을 통해 구축된 학습자 언어 자료는 다음의 <그림 5>와 같다.

<그림 5> 학습자 언어 자료 전사 화면



	A	B
457	22	AW_22 [22차 학교에서 있은 일]
458	W22	학교에서 일어난 큰 실수나 사고를 많지는 않아요.
459	W22	제가 중국에 있을 때는 중학교 1학년이에요.
460	W22	우리 반의 숙제가 조금 많아요
461	W22	그리고 우리 반 친구 중 한 명이 집에 무슨 일이 있을 것 같은데 공부도 너무 힘들어서 공부를 안 해요
462	W22	하지만 우리 반 담임 선생님은 그를 포기하지 않고 매일 시간이 있으면 그는 공부를 가르쳤어요
463	W22	그는 열심히 공부하기 시작했어요.
464	W22	하지만 한 날에 수업 중에 그가 숙제하는 것을 감먹었어요.
465	W22	그래서 선생님이 다음부터 숙제하는 거 잊지 말라고 말했어요
466	W22	그리고 이번 숙제를 다섯 번 쓰게를 말했어요
467	W22	그리고 그가 화가 나서 선생님과 싸우기 시작했어요.
468	W22	선생님께서 무슨 자격으로 그를 다룰 수 있느냐고 말했어요.
469	W22	안 좋은 얘기 많이 했어요
470	W22	그는 선생님을 때리려고 했어요.
471	W22	우리 반에 여러 명의 남학생들이 그를 차단했습니다.
472	W22	마지막으로 학교에서 반 달 동안 집에서 반성하라고 했어요
473	W22	그는 돌아온 후에 선생님께 말대꾸도 하지 않았고 말수가 많지는 않아졌고 공부도 열심히 했어요.
474	W22	그가 반성을 해면 많은 걸 배우게 된 것 같아요.
475	W22	이 일은 선생님께서 잘 못하는 부분도 있어요
476	W22	하지만 나는 선생님을 때리는 것은 옳지 않고 선생님을 존중해야 된다고 생각해요.
477	23	AW_23 [23차 행복한 일과 슬픈 일]
478	W23	행복한 일이 무엇인가요?
479	W23	저 행복한 일은 즐거운 일이다고 생각해요
480	W23	모든 사람은 행복한 일이 다 있어요.
481	W23	저는 친구랑 같이 있으면 너무 행복해요.

4.2. 자료 분석 및 통계

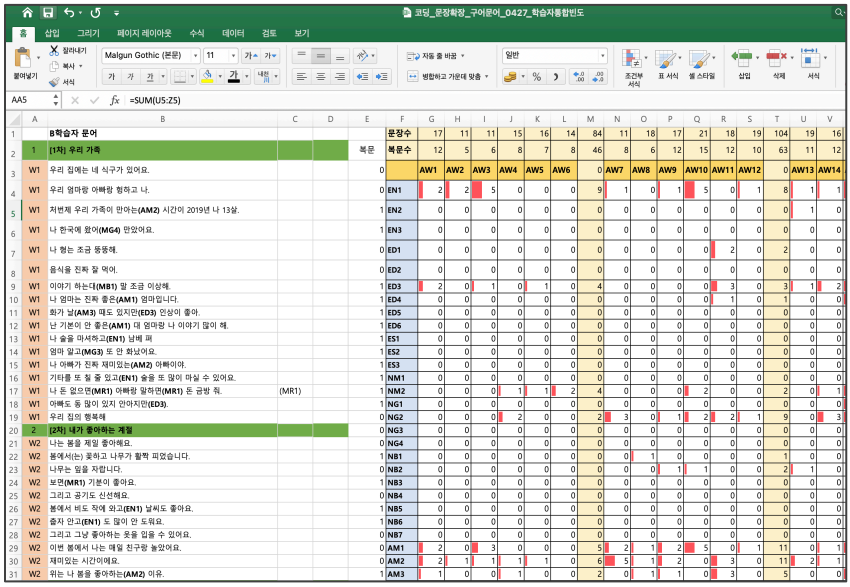
1) 자료 분석 및 코딩

(1) 통사 복잡성

통사 복잡성은 단위 길이와 단위당 절 수, 단위 확장의 유형으로 나누어 측정한다. 단위 길이는 하나의 분석 단위에 나타난 형태소 수로 측정하기 때문에 문어 자료에서는 문장당 형태소 수, 구어에서는 AS-unit당 형태소 수로 측정한다. 형태소 분석은 ‘키위 형태소 분석기v0.11.0’을 사용하였으며 하나의 텍스트에 나타난 총 형태소의 수를 문장/AS-unit의 수로 나누어 단위당 형태소 수를 계산하였다.

단위 길이 외에 단위당 절 수와 유형별 내포절을 측정하기 위해서는 각 텍스트에 나타난 유형별 절을 계산해야 한다. 본고에서는 각 대등절과 내포절의 문법 표지마다 코딩 기호를 부여하여 연구자가 직접 미리 구축된 엑셀 파일에 입력한 뒤, 각 코딩 기호의 빈도를 세는 방법으로 문장 확장의 양과 유형을 분석하였다. 다음의 <그림 6>은 유형별 대등절과 내포절의 코딩 화면이다.

<그림 6> 대등절과 내포절 코딩 화면



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
1		B학습자 문어				문장수	17	11	11	15	16	14	84	11	18	17	21	18	19	104	19	16
2	1	[1차] 우리 가족			복문	복문수	12	5	6	8	7	8	46	8	6	12	15	12	10	63	11	12
3	W1	우리 집에는 네 식구가 있어요.			0		AW1	AW2	AW3	AW4	AW5	AW6	0	AW7	AW8	AW9	AW10	AW11	AW12	0	AW13	AW14
4	W1	우리 엄마랑 아빠랑 형하고 나.			0	EN1	2	2	5	0	0	0	9	1	0	1	5	0	1	8	1	1
5	W1	저번제 우리 가족이 만나는(AM2) 시간이 2019년 나 13살.			1	EN2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
6	W1	나 한국에 왔어(MG4) 만았어요.			1	EN3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	W1	나 형는 조금 뚱뚱해.			0	ED1	0	0	0	0	0	0	0	0	0	0	0	2	0	2	0	0
8	W1	음식을 진짜 잘 먹어.			0	ED2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	W1	이야기 하는다(MB1) 말 조금 이상해.			1	ED3	2	0	1	0	1	0	4	0	0	0	0	3	0	3	1	2
10	W1	나 엄마는 진짜 좋은(AM1) 엄마입니다.			1	ED4	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0
11	W1	화가 날(AM3) 때도 있지만(ED3) 인상이 좋아.			1	ED5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	W1	난 기분이 안 좋은(AM1) 대 엄마랑 나 이야기 많이 해.			1	ED6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	W1	나 술을 마셔보고(EN1) 남배 피			1	ES1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	W1	엄마 알고(MG3) 또 안 화났어요.			1	ES2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	W1	나 아빠가 진짜 재미있는(AM2) 아빠이야.			1	ES3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	W1	기타를 또 칠 줄 있고(EN1) 술을 또 많이 마실 수 있어요.			1	NM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	W1	나 돈 없으면(MR1) 아빠랑 말하면(MR1) 돈 금방 줘.	(MR1)		1	NM2	0	0	0	1	1	2	4	0	0	0	2	0	0	2	0	1
18	W1	아빠도 동 많이 있지 안아지만(ED3).			1	NG1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	W1	우리 집의 행복해			0	NG2	0	0	0	2	0	0	2	3	0	1	2	2	1	9	0	3
20	2	[2차] 내가 좋아하는 계절			0	NG3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	W2	나는 봄을 제일 좋아해요.			0	NG4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	W2	봄에서(는) 꽃하고 나무가 활짝 피었습니다.			1	NB1	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0
23	W2	나무는 잎을 자랍니다.			0	NB2	0	0	0	0	0	0	0	0	0	1	1	0	0	2	1	0
24	W2	보면(MR1) 기분이 좋아요.			1	NB3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	W2	그리고 공기도 신선해요.			0	NB4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	W2	봄에서 비도 작에 와고(EN1) 날씨도 좋아요.			1	NB5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	W2	춥자 안고(EN1) 도 많이 안 도워요.			1	NB6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	W2	그리고 그날 좋아하는 옷을 입을 수 있어요.			0	NB7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	W2	이번 봄에서 나는 매일 친구랑 놀았어요.			0	AM1	2	0	3	0	0	0	5	2	1	2	5	0	1	11	0	1
30	W2	재미있는 시간이에요.			0	AM2	2	1	1	1	1	0	6	5	1	2	0	3	0	11	2	1
31	W2	위는 나 봄을 좋아하는(AM2) 이유.			1	AM3	1	0	0	1	0	0	2	0	1	1	0	3	0	5	0	0

(2) 어휘 복잡성

어휘 복잡성은 어휘 다양도, 어휘 밀도, 어휘 정교도를 통해 측정하였다. 어휘 다양도는 CTTR인 '어휘 유형 수/어휘 구현 수*2의 제곱근'으로 측정되었다. 어휘 유형 수를 구하기 위해 먼저 형태소 분석기의 분석 결과에서 어휘에 해당되지 않는 부분을 제거한 뒤, 남은 부분에 대해 어휘의 유형 수와 구현 수를 측정하였다.

어휘 밀도는 어휘 다양도와 마찬가지로 형태소 분석기의 결과를 바탕으로 분석한다. 형태소 분석 결과에서 어휘에 해당되지 않는 부분을 제거한 후, 남은 부분에서 내용어에 해당하는 어휘의 빈도를 총 어휘 수로 나눈 뒤 100을 곱하여 계산하였다.

마지막으로 어휘 정교도는 AntWord Profiler프로그램을 통해 분석하였다. 이 프로그램은 2개 이상의 어휘 목록 간의 일치도를 확인할 수 있는 프로그램으로, 기준 어휘 목록과 분석 대상 어휘 목록을 입력하면 기준 어휘의 목록과 일치하는 어휘들을 보여준다. 본고에서는 기준 어휘 목록으로 <한국어(KSL) 교육과정(2017)>의 <중학교 언어 재료>의 '일상생활 어휘 목록'을 입력하였으며 초급 어휘 목록과 일치하지 않는 어휘의 비율에 100을 곱하여 어휘 정교도를 계산하였다.

(3) 정확성

정확성은 무오류절의 비율, 단위당 무오류절 수, 단위당 오류 수로 측정되며 추가적으로 오류의 유형을 살펴보았다. 무오류절의 비율과 무오류절 수는 미리 구축된 엑셀 파일에 오류가 없는 절을 직접 코딩한 뒤 무오류절 수와 총 절 수의 비율, 단위당 무오류절 수를 계산하였다. 오류 수와 오류 유형은 유형별 내포절을 측정하는 것과 같은 방법을 사용하였다. 먼저 각 오류 유형별로 코딩 기호를 부여한 뒤, 연구자가 엑셀 파일에서 오류가 발생한 위치에 해당 코딩 기호를 입력한다. 그리고 각 코딩 기호의 빈도를 계산하는 방법으로 측정하였다.

2) 통계 분석

본고에서는 추세선과 그래프를 통해 각 요인의 변화 발달의 전체적인 추이를 살펴보고 6개월을 단위로 묶어 기간별 발달을 비교하였다. 그러나 단순한 비교로는 그 차이가 유의한 것인지 확인할 수 없기 때문에 통계 분석의 방법을 사용하여 학습자 간 차이나 기간별 차이, 문어와 구어의 차이를 비교 분석하였다.

통계 분석은 IBM SPSS Statistics 24프로그램을 사용하였다. 두 집단 간의 차이, 즉

문어와 구어의 차이를 비교하는 경우 대응표본 t검정을 실시하였고 학습자별, 기간별로 세 집단의 차이를 검증하는 경우 일원배치 분산분석(One-Way ANOVA)과 쉐페(Sheffe) 사후분석을 실시하였으며, 복잡성과 정확성의 관계를 살펴보기 위해 상관관계 분석(Correlation analysis)을 실시하였다. 일원배치 분산분석에서 등분산 가정을 만족하기 않을 경우, 이분산 분석인 웰치 테스트(Welch-test)를 실시하고 사후분석은 게임스 하웰(Games-Howell) 사후분석을 실시하였다.

Ⅳ. 연구 결과

본 연구에서는 수집한 18개월의 자료를 6개월을 단위로 1~6개월, 7~12개월, 13~18개월의 세 구간으로 나누어 세 구간의 평균을 비교하는 방식으로 발달을 측정하였다. 일반적으로 종적으로 언어 발달을 고찰하는 연구에서는 3개월이나 6개월을 간격으로 살펴보는 경우가 많다. 본고의 분석 결과를 3개월을 간격으로 1~3개월, 4~6개월, 7~9개월, 10~12개월, 13~15개월, 16~18개월의 6개 구간으로 나누었을 때, 측정 지표에 따라 약간의 차이가 있지만 대부분의 지표에서 1~3개월과 7~9개월, 4~6개월과 10~12개월 혹은 1~3개월과 10~12개월, 7~9개월과 16~18개월과 같이 인접한 두 구간 사이에는 유의미한 차이를 보이지 않고 최소 6개월 정도의 간격을 두고 통계적으로 유의미한 차이를 보이는 것을 확인하였다.[66] 따라서 본고에서는 학습자들이 최소 6개월 정도 시간이 지나야 유의미한 발달을 보인다는 점을 확인하고 6개월을 간격으로 구간을 나누었다.

1. 통사 복잡성의 발달

본 절에서는 학습자가 산출한 문어와 구어 자료를 바탕으로 학습자들의 통사 복잡성의 발달 양상을 살펴보고자 한다. 우선 단위 길이와 단위당 절 수를 통해 양적인 측면의 통사 복잡성을 살펴본 뒤 단위 확장의 유형 및 단위 확장에 관여하는 문법 항목들의 빈도를 측정함으로써 학습자들이 구체적으로 어떻게 문장을 확장하고 있는지 살펴보도록 한다.

1.1. 문어의 통사 복잡성 발달

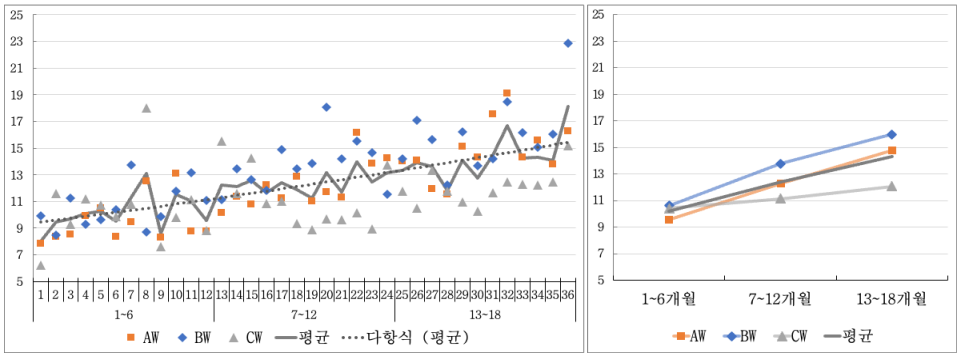
1.1.1. 문장 길이

문장의 길이는 학습자의 통사 복잡성을 가늠하는 가장 기본적인 측정 지표이다. 이는 학습자의 언어 숙달도가 높아질수록 긴 문장을 산출할 것이라는 가정에서 출발하였

[66] 문장당 형태소 수의 길이를 예로 들면, 1~3개월과 7~9개월, 10~12개월, 13~15개월, 16~18개월(1~3개월<7~9개월, 10~12개월, 13~15개월, 16~18개월) 사이, 4~6개월과 13~15개월 사이(4~6개월<13~15개월), 7~9개월과 16~18개월(7~9개월<16~18개월)사이에만 유의미한 차이를 보이며 최소 6개월, 최대 9개월을 간격으로 통계적으로 유의미한 차이를 보였다.

으며 여러 선행연구들을 통해 그 유효성이 검증되었다. 앞장에서 설명한 것과 같이 본고에서는 문장당 형태소 수로 문장 길이를 측정하였다. 세 명의 학습자는 1년 6개월 간 36회에 걸쳐 108개의 작문 자료에서 총 2,346개의 문장과 29,393개의 형태소를 산출하였으며 문장당 형태소 수의 평균은 12.53개였다. 각 시점별 학습자의 문장 길이와 6개월을 단위로 살펴본 문장 길이의 발달은 다음의 <그림 7>과 같다.

<그림 7> 시점별 문장 길이와 6개월 단위 발달 추이



세 학습자의 평균을 통해 학습자들의 문장 길이 발달 추이를 살펴보면 큰 폭과 작은 폭의 등락을 반복하지만 전반적으로 직선에 가까운 안정적인 상승 추세를 보이며 문장 길이가 점점 길어지지는 양상을 보이고 있었다. 이는 6개월 단위 발달 추이에서도 잘 나타나는데 학습자들의 문장 길이는 일정하고 안정적인 폭으로 증가하고 있음을 확인할 수 있다. 문장 길이의 기간별 증가폭을 확인해 보면 7~12개월에는 1~6개월보다 2.19개, 13~18개월에는 7~12개월보다 1.90개 증가하고 있었다.

기간별 문장 길이의 증가가 통계적으로 유의한지를 검증하기 위해 일원배치 분산분석(One-Way ANOVA)을 실시하였다. 그 결과 <표 15>와 같이 F=27.471, p=.000으로 나타나 기간별 문장 길이의 차이가 통계적으로 유의하다는 것을 확인할 수 있었다. 사후검정 결과, 세 구간 사이에 모두 유의한 차이(1~6개월<7~12개월<13~18개월)가 발견되었다. 즉 학습자들의 문장 길이는 시간이 흐름에 따라 안정적이고 유의미하게 증가하고 있었다.

<표 15> 기간별 문장 길이

측정 지표	기술통계량				F/p	사후검정
	기간(개월)	N	Mean	SD		
문장 길이	1~6	36(a)	10.19	2.13	27.471/.000***	a<b<c
	7~12	36(b)	12.38	2.20		
	13~18	36(c)	14.28	2.66		

*p<.05 **p<.01 ***p<.001

학습자별로 발달을 살펴보면 세 학습자 모두 평균과 마찬가지로 안정적인 속도로 문장 길이를 늘려가고 있었지만 A, B학습자가 유사한 발달 속도를 보이는 것과 달리 C학습자는 상대적으로 느린 속도로 증가하여 시간이 흐를수록 세 학습자의 문장 길이의 차이가 커지고 있었다. 세 학습자의 문장 길이의 평균은 B>A>C순으로 나타났는데 학습자 간 문장 길이의 차이가 통계적으로 유의한지 확인하기 위해 일원배치 분산분석을 실시한 결과 <표 16>과 같이 F=6.055, p=.003으로 학습자 간에 유의한 차이가 있는 것으로 나타났다. 그리고 사후검정 결과 B학습자의 문장 길이가 C학습자에 비해 유의하게 긴 것으로 나타났다.

<표 16> 학습자별 문장 길이

측정 지표	기술통계량				F/p	사후검정
	학습자	N	Mean	SD		
문장 길이	AW	36	12.20	2.84	6.055/.003**	BW>CW
	BW	36	13.45	3.06		
	CW	36	11.20	2.27		

*p<.05 **p<.01 ***p<.001

정리하면, 문장 길이의 발달은 시간의 흐름에 따라 증가하는 추세를 보이며 6개월을 간격으로 유의미한 증가를 보였다. 그러나 학습자별로 문장 길이의 증가 속도는 다르게 나타났으며 시간이 흐를수록 학습자 간의 차이가 커지고 있었다.

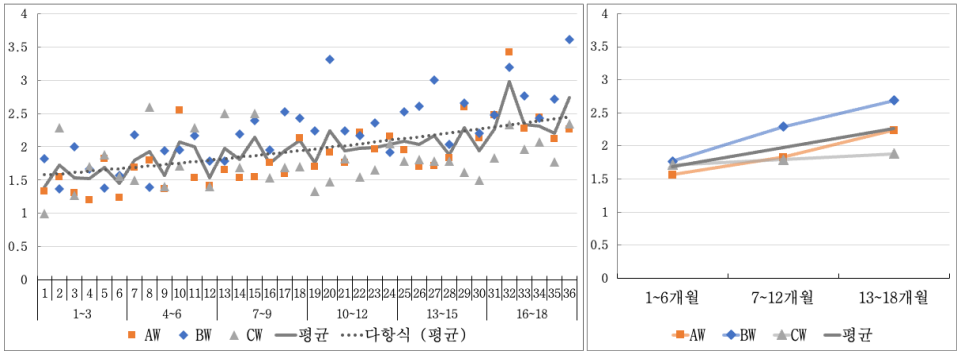
1.1.2. 문장 확장

1.1.2.1. 문장당 절 수

시간의 흐름에 따른 문장당 절 수의 변화 발달을 살펴보면 <그림 8>과 같다. 학습자

들의 문장당 절 수는 끊임없는 등락을 반복하지만 등락의 폭이 비교적 일정하며 시간의 흐름에 따라 약간 증가하는 추세를 보이고 있었다. 2차 다항식 추세선은 거의 직선에 가까운 모양을 띠고 있는데 이는 문장당 절 수의 증가 속도가 안정적임을 의미한다.

<그림 8> 시점별 문장당 절 수와 6개월 단위 발달 추이



6개월을 단위로 문장당 절 수의 평균을 살펴보면, 7~12개월에는 1~6개월보다 0.29개 증가하고 13~18개월에는 7~12개월보다 0.30개 증가한 것으로 나타났다. 이는 문장 길이의 발달과 같은 양상으로 학습자들이 안정적인 속도로 문장의 구조를 차츰 확장해 나가고 있음을 의미한다.

기간별 문장당 절 수의 증가가 통계적으로 유의한지를 검증하기 위해 일원배치 분산분석(One-Way ANOVA)을 실시하였다. 그 결과 <표 17>과 같이 F=16.170, p=.000으로 기간별 문장당 절 수의 차이가 통계적으로 유의하다는 것을 확인할 수 있었다. 사후검정 결과, 세 기간 사이에 모두 유의미한 차이(1~6개월<7~12개월<13~18개월)가 있었다. 즉 학습자들은 시간이 지날수록 한 문장에서 더 많은 절을 사용하게 되며 이러한 변화는 유의미한 것임을 알 수 있다.

<표 17> 기간별 문장당 절 수

측정 지표	기술통계량				F/p	사후검정
	기간(개월)	N	Mean	SD		
문장당 절 수	1~6	36(a)	1.68	0.39	16.170/.000***	a<b<c
	7~12	36(b)	1.97	0.41		
	13~18	36(c)	2.27	0.51		

*p<.05 **p<.01 ***p<.001

학습자별로 문장당 절 수의 변화 추이를 살펴보면 세 학습자 모두 평균과 마찬가지로 일정한 속도로 문장당 절 수를 확장하는 추세를 보이지만 학습자별로 증가 속도에는 차이가 있었다. 세 학습자 중 B학습자가 가장 빠른 속도로 문장당 절 수를 확장하고 있었고 C학습자는 A학습자에 비해 높은 출발을 보였으나 증가 속도가 완만하여 13~18개월에는 세 학습자 중 가장 낮은 절 수를 보이고 있었다. 세 학습자의 문장당 절 수의 평균은 B>A>C순으로 나타났는데 학습자별로 문장당 절 수의 차이가 통계적으로 유의한지 확인하기 위해 일원배치 분산분석을 실시한 결과 <표 18>과 같이 F=9.952, p=.000으로 학습자 간에 유의한 차이가 있는 것으로 나타났다. 사후검정 결과 B학습자의 문장당 절 수가 A, C학습자에 비해 유의하게 높은 것으로 나타났다.

<표 18> 학습자별 문장당 절 수

측정 지표	기술통계량				F/p	사후검정
	학습자	N	Mean	SD		
문장당 절 수	AW	36	1.88	0.46	9.952/.000***	BW>AW,CW
	BW	36	2.25	0.53		
	CW	36	1.80	0.37		

*p<.05 **p<.01 ***p<.001

학습자들의 문장당 절 수와 문장 길이는 모두 기간별로 안정적이고 유의미한 증가 추세를 보이고 있었다. 이는 학습자들이 시간이 흐를수록 한 문장 안에 더 많은 절을 사용하여 더 긴 문장을 산출하며 통사 복잡성을 발달시키고 있음을 의미한다. 그러나 학습자별로 발달 속도나 정도에는 차이가 있었는데 문장 길이와 문장당 절 수를 종합해 볼 때, B학습자의 통사 복잡성 발달이 가장 빠르고 C학습자의 통사 복잡성 발달이 가장 느린 것으로 나타났다.

1.1.3. 대등절·내포절의 유형

1.1.3.1. 문장당 대등절 수와 내포절 수

대등절과 내포절은 문장을 확장하는 방식에 있어 차이가 있다. 문장당 대등절 수와 문장당 내포절 수의 발달을 나누어 살펴보면 다음의 <표 19>와 같다.

<표 19> 6개월 단위 문장당 대등절 수와 내포절 수

학습자	대등/내포	1~6개월	증가폭	7~12개월	증가폭	13~18개월	평균
A	대등	0.07	+0.08	0.16	+0.07	0.23	0.15
	내포	0.49	+0.17	0.66	+0.34	1.00	0.72
B	대등	0.15	+0.16	0.31	+0.16	0.48	0.31
	내포	0.60	+0.35	0.95	+0.22	1.18	0.91
C	대등	0.14	+0.05	0.20	+0.04	0.23	0.19
	내포	0.57	+0.02	0.59	+0.05	0.64	0.60
평균	대등	0.12	+0.10	0.22	+0.09	0.31	0.22
	내포	0.55	+0.18	0.74	+0.20	0.94	0.74

먼저, 학습자들의 평균을 살펴보면 기간별 문장당 대등절 수의 평균은 1~6개월에 0.12개, 7~12개월에 0.22개, 13~18개월에 0.31개이고 기간별 문장당 내포절 수의 평균은 각각 1~6개월에 0.55개, 7~12개월에 0.74개, 13~18개월에 0.94개로 학습자들은 꾸준히 대등절보다 내포절을 더 많이 사용하고 있었으며 증가폭 역시 내포절이 더 큰 것을 확인할 수 있다. 대등절과 내포절의 증가 속도를 비교하면 대등절은 7~12개월에 1~6개월보다 0.10개, 13~18개월에 7~12개월보다 0.09개 증가하여 증가 속도가 약간 느려지는 것에 반해 내포절은 7~12개월에 1~6개월보다 0.18개, 13~18개월에 7~12개월보다 0.20개 증가하여 증가 속도가 더 빨라지는 것으로 나타났다. 학습자별로 살펴보면 B학습자의 내포절을 제외하고 모두 평균과 같은 추세를 나타고 있었다.

학습자들의 문장당 대등절과 내포절의 평균 차이가 통계적으로 유의한지 알아보기 위해 일원배치 분산분석을 실시한 결과 <표 20>과 같이 대등절은 F=11.784, p=.000, 내포절은 F=12.090, p=.000으로 대등절과 내포절 모두 기간별로 유의한 차이가 있는 것으로 나타났다. 사후 검정 결과 대등절은 1~6개월과 7~12개월, 16~18개월 사이에 차이가 있었고 내포절은 1~6개월, 7~12개월과 16~18개월 사이에 유의한 차이가 있는 것으로 나타났다. 이를 통해 학습자들의 대등절은 1~6개월과 7~12개월 사이에 유의미하게 증가하지만 이후 뚜렷하게 증가하지 않고 내포절은 7~12개월과 13~18개월 사이에 유의미하게 발달함을 알 수 있다. 이는 학습자들이 초반에는 대등절이 더 빨리 발달하지만 12개월이 지난 시점부터는 대등절보다 내포절의 발달 속도가 빠름을 의미한다.

<표 20> 기간별 문장당 대등절과 내포절 수

측정 지표	기술통계량				F/p	사후검정
	기간(개월)	N	Mean	SD		
문장당 대등절 수	1~6	36(a)	0.12	0.14	11.784/.000***	a<b,c
	7~12	36(b)	0.22	0.16		
	13~18	36(c)	0.31	0.19		
문장당 내포절 수	1~6	36(a)	0.55	0.32	12.090/.000***	a,b<c
	7~12	36(b)	0.74	0.29		
	13~18	36(c)	0.94	0.38		

*p<.05 **p<.01 ***p<.001

학습자 간 문장당 대등절 수와 내포절 수의 차이가 통계적으로 유의한지 확인하기 위해 일원배치 분산분석을 실시한 결과 <표 21>과 같이 대등절은 F=8.809, p=.000, 내포절은 F=7.463, p=.001로 대등절과 내포절 모두 학습자 간에 유의한 차이가 있는 것으로 나타났다. 사후검정 결과 대등절은 B학습자와 A, C학습자(B>A,C) 사이에 차이가 있었고 내포절은 B와 C학습자(B>C) 사이에만 유의한 차이가 있는 것으로 나타났다.

<표 21> 학습자별 문장당 대등절 수와 내포절 수

측정 지표	기술통계량				F/p	사후검정
	학습자	N	Mean	SD		
문장당 대등절 수	AW	36	0.15	0.15	8.809/.000***	BW>AW,CW
	BW	36	0.31	0.19		
	CW	36	0.19	0.16		
문장당 내포절 수	AW	36	0.72	0.37	7.463/.001***	BW>CW
	BW	36	0.91	0.38		
	CW	36	0.60	0.27		

*p<.05 **p<.01 ***p<.001

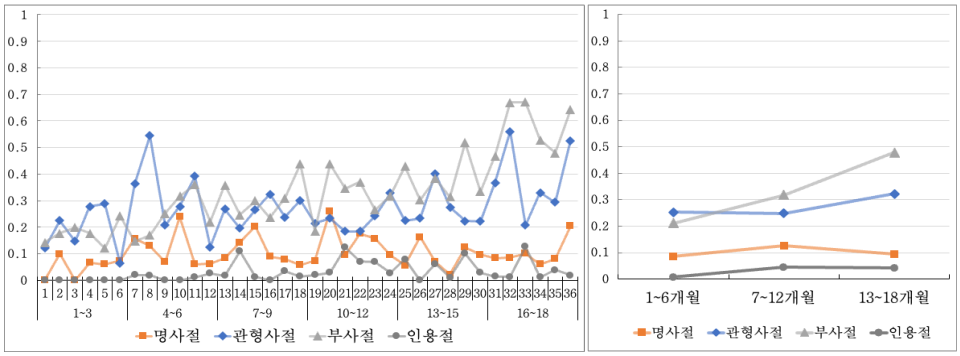
정리하면, 대등절과 내포절은 모두 시간이 지남에 따라 증가하지만 대등절은 시간이 지날수록 증가 속도가 완만해지고 내포절은 시간이 지날수록 증가 속도가 빨라지는데 대등절은 7~12개월, 내포절은 13~18개월에 들어 유의미한 증가가 나타난다.

1.1.3.2. 내포절의 유형

내포절은 크게 명사절, 관형사절, 부사절, 인용절로 나눌 수 있다. 아래 <그림 9>는 세 학습자들의 평균으로 살펴본 유형별 내포절의 변화 그래프이다. 그래프에서 확인할

수 있듯이 높은 빈도와 함께 지속적이고 뚜렷한 발달을 보이는 것은 부사절이었다. 6개월을 단위로 묶어 살펴보면 부사절만이 꾸준히 증가하고 있고 관형사절은 7~12개월에 퇴행 이후 13~18개월에 다시 발달, 명사절과 인용절은 7~12개월에 발달 이후 13~18개월에 다시 퇴행이 나타나고 있었다.

<그림 9> 유형별 내포절 수의 시점별 빈도와 6개월 단위 발달 추이



이와 같은 각 유형별 내포절의 발달이나 퇴행이 기간별로 유의미한 차이가 있는지 확인하기 위해 일원배치 분산분석을 실시한 결과 <표 22>와 같이 명사절과 관형사절은 기간별로 유의미한 차이가 발견되지 않았고 부사절은 F=20.369, p=.000, 인용절은 F=9.551, p=.000으로 기간별로 유의미한 차이가 있는 것으로 나타났다. 사후검정 결과 부사절은 세 기간 사이의 차이(1~6개월<7~12개월<13~18개월)가 모두 유의하였고 인용절은 1~6개월과 나머지 두 기간 사이(1~6개월<7~12개월, 13~18개월)에만 유의한 차이를 보였다. 이러한 결과는 학습자들이 부사절은 꾸준히 발달시키고 있으나 명사절과 관형사절의 발달은 뚜렷하지 않으며 인용절은 7~12개월까지 발달을 하지만 이후에 지속적인 발달이 이루어지지 못하고 있는 것으로 해석할 수 있다.

<표 22> 유형별 내포절의 기간별 사용 양상

측정 지표	기술통계량				F/p	사후검정
	기간(개월)	N	Mean	SD		
명사절	1~6	36(a)	0.08	0.10	1.653/.197	-
	7~12	36(b)	0.13	0.12		
	13~18	36(c)	0.10	0.08		
관형절	1~6	36(a)	0.25	0.22	2.190/.120	-
	7~12	36(b)	0.25	0.12		
	13~18	36(c)	0.32	0.18		

부사절	1~6	36(a)	0.21	0.14	20.369/.000***	a<b<c
	7~12	36(b)	0.32	0.14		
	13~18	36(c)	0.48	0.21		
인용절	1~6	36(a)	0.01	0.02	9.551/.000***	a<b,c
	7~12	36(b)	0.04	0.07		
	13~18	36(c)	0.04	0.07		

*p<.05 **p<.01 ***p<.001

각 유형별 내포절 수의 학습자 간 차이가 통계적으로 유의한지 확인하기 위해 일원배치 분산분석을 실시한 결과 <표 23>과 같이 명사절, 부사절, 인용절의 사용에서만 유의미한 차이가 있는 것으로 나타났다. 사후검정 결과 명사절은 B학습자와 C학습자(B>C) 간에 차이가 있었고 부사절은 B학습자와 A, C학습자(B>A, C) 간에 차이가 있었으며 인용절은 A, B학습자와 C학습자(A, B>C) 간에 유의한 차이가 있는 것으로 나타났다.

<표 23> 유형별 내포절의 학습자별 사용 양상

측정 지표	기술통계량				F/p	사후검정
	학습자	N	Mean	SD		
명사절	AW	36	0.09	0.09	3.823/.027*	BW>CW
	BW	36	0.14	0.13		
	CW	36	0.08	0.06		
관형절	AW	36	0.28	0.15	1.077/.344	-
	BW	36	0.30	0.18		
	CW	36	0.24	0.21		
부사절	AW	36	0.30	0.20	6.666/.002**	BW>AW,CW
	BW	36	0.43	0.21		
	CW	36	0.28	0.14		
인용절	AW	36	0.05	0.07	13.001/.000***	AW,BW>CW
	BW	36	0.04	0.06		
	CW	36	0.003	0.01		

*p<.05 **p<.01 ***p<.001

정리하면, 학습자들은 주로 관형사절과 부사절을 통해 문장을 확장하고 있으며 명사절과 인용절의 사용은 상대적으로 저조하였다. 또한 부사절은 시간이 지남에 따라 사용 빈도 측면에서 꾸준한 발달을 보이지만 관형사절은 7~12개월, 명사절과 인용절은 13~18개월에 퇴행이 나타나기도 하였다. 인용절은 모어화자의 언어에서도 자주 사용되지 않는 문법 항목이기 때문에 학습자의 언어 습관이나 주제의 영향을 많이 받을 수

있지만 명사절이 퇴행이 나타나는 점은 특이하다고 할 수 있다.

1.1.4. 대등절과 내포절 관련 문법 항목

대등절과 내포절은 사용 빈도도 중요하지만 어떤 문법 항목을 얼마나 사용하여 문장을 확장하는지 살펴보는 것 또한 발달에 대한 중요한 정보를 제공한다. 이 부분에서는 대등절과 내포절을 생성하는 각 문법 항목의 빈도를 통해 학습자들의 문장 구조 확장을 구체적으로 살펴보고자 한다. 이를 위해 먼저 학습자들이 산출한 각 문법 항목의 표준화 빈도[67]를 산출하였다. 그리고 산출된 표준화 빈도를 다시 6개월을 단위로 묶어 기간별 변화 발달 양상을 살펴보았다.

먼저 대등절을 살펴보도록 한다. 학습자들의 문어 자료에 나타난 대등절 관련 문법 항목의 총 빈도와 전체 절에서 차지하는 비율을 살펴보면 1~6개월에는 61회(17.0%), 7~12개월에는 158회(21.5%), 13~18개월에는 285회(23.5%)로 총 사용 빈도뿐만 아니라 비율 또한 증가하고 있었다. 각 의미 범주별 문법 항목의 기간별 사용 양상은 다음의 <표 24>와 같다.

<표 24> 대등절 관련 문법 항목의 기간별 사용 양상

구분		문법 항목	1~6개월	7~12개월	13~18개월
대등절	나열 · 순접	-고	8.46	14.99	19.44
		-며/-면서	-	0.13	-
	대조 · 역접	-ㄴ데	0.55	1.13	2.78
		-지만	2.21	3.53	5.46
	선택	-거나	-	0.25	0.60

학습자들은 관찰 기간 동안 총 5가지의 문법 항목을 사용하여 대등절을 구성하고 있었다. 그중에서 학습자들이 가장 많이 사용한 것은 '나열·순접'의 '-고', '대조·역접'의 '-지만'이었으며 이는 세 기간 모두 동일하게 나타났다. 특히 '나열·순접'의 '-고' 사용의 빈도가 다른 항목에 비해 압도적으로 높게 나타나 학습자들이 주로 '-고'를 사용하여 대등절을 구성하고 있음을 알 수 있었다.

기간별 발달을 살펴보면 모든 문법 항목의 사용 빈도가 시간이 흐름에 따라 증가하

[67] 표준화 빈도는 각 문법 항목의 사용 빈도를 문장 수로 나눈 뒤 100을 곱한 수치로 문장 수가 서로 다른 자료의 빈도를 산출하는 데 적합하다. 선행연구에서는 김현아(2016), 최수완(2019), 서신웨(2022)에서 사용되었다.

며 발달하고 있었다. 그러나 각 문법 항목들의 출현 시기에는 차이가 있었는데 6가지의 대등절 중 '나열·순접'의 '-고', '대조·역접'의 '-ㄴ데', '-지만'은 관찰 시작 시점인 1~6개월부터 꾸준히 출현하고 있었으나 '나열·순접'의 '-며/-면서'는 7~12개월에만 출현하였고 '선택'의 '-거나'는 7~12개월에 들어서서 출현하기 시작하여 비교적 늦게 습득됨을 확인할 수 있었다. 이 외에 한 번도 사용되지 않은 문법 항목에는 '대조·역접'의 '-나', '-라', '선택'의 '-든지/든가' 등이 있었다. 대등절을 형성하는 문법 항목의 사용 양상을 살펴보면 예문(16)과 같다.

(16) 가. 시간이 있으면 우리는 같이 여행하고 공부하고 自助餐(√뷔페)을 먹었어요.(AW5)
나. 행복하면서 마음이 힘들게 10일 살았어요.(BW20)
다. 학생회는 재미있지만 일이 너무 많아요.(CW25)
라. 부모님이 나를 욕하거나 때리면 나도 위험한 행동 할 수 있는데...(BW29)

다음 명사절을 살펴보도록 한다. 학습자들의 문어 자료에 나타난 명사절 관련 문법 항목의 총 빈도와 전체 절에서 차지하는 비율을 살펴보면 1~6개월에는 46회(12.8%), 7~12개월에는 95회(12.9%), 13~18개월에는 90회(7.4%)로 7~12개월에는 1~6개월보다 사용 빈도와 비율이 증가하였으나 13~18개월에는 빈도와 비율 모두 하락하며 지속적인 증가를 보이지 않았다. 명사절 관련 문법 항목의 기간별 사용 양상을 세부적으로 살펴보면 다음의 <표 25>와 같다.

<표 25> 명사절 관련 문법 항목의 기간별 사용 양상

구분		문법 항목	1~6개월	7~12개월	13~18개월
명사절	명사형 전성어미	-ㅁ	-	0.13	-
		-기	3.13	3.27	1.98
	관형사형 전성어미+것	-ㄴ/는 것	4.04	7.68	4.66
		-ㄹ 것	0.18	-	0.20
		-던 것	-	-	0.10
	의문형 종결어미	-ㄴ지/는지	0.37	0.50	1.69
		-ㄹ지	0.55	0.25	0.20
		-ㄹ까	0.18	0.13	0.10

표에서 확인할 수 있듯이 학습자들은 관찰 기간 동안 총 8가지의 문법 항목을 사용하여 명사절을 산출하고 있었다. 그중에서 가장 많이 사용된 항목은 '관형사형 전성어미+ 것'인 '-ㄴ/는 것'이고 두 번째로 많이 사용된 항목은 '명사형 전성어미'인 '-기'로

세 기간 모두 이 두 가지의 문법 항목이 압도적으로 높은 빈도로 나타나 학습자들이 주로 '-ㄴ/는 것'과 '-기'를 사용하여 명사절을 구성하고 있음을 알 수 있었다.

기간별 발달을 살펴보면, 각 문법 항목별로 서로 다른 발달을 보이고 있었는데 먼저 모든 기간에서 빈도가 꾸준히 증가하고 있는 항목은 '의문형 종결 어미'인 '-ㄴ지/는지' 한 항목이었다. 이 외에 '명사형 전성어미'인 '-기', 관형사형 전성어미+ 것'인 '-ㄴ/는 것'은 7~12개월에는 1~6개월보다 빈도가 증가하였으나 13~18개월에 다시 감소하는 퇴행을 보였고 '의문형 종결 어미'인 '-ㄹ지', '-ㄹ까'는 시간이 지남에 따라 그 빈도가 꾸준히 감소하고 있었다. 또한 '명사형 전성어미'인 '-ㅁ'은 7~12개월에만 낮은 빈도로 출현하며 거의 사용되지 않고 있었으며 '관형사형 전성어미+ 것'인 '-던 것'은 13~18개월에 처음 출현하였으나 올바르게 사용하지 못하고 있었다. 그리고 관찰 기간 동안 한 차례도 출현하지 않은 문법 항목에는 '의문형 종결 어미'인 '-냐/느냐', '-ㄴ가/는가', '-던가' 등이 있었다.

즉 문어에서 명사절의 발달은 뚜렷하지 않았으며 각 문법 항목별로 습득의 시기나 발달, 퇴행이 다르게 나타났다. 또한 학습자들은 시간이 지나면서 새로운 문법 항목을 습득하기도 하지만 사용 빈도 면에서는 7~12개월에 증가 이후 13~18개월에 다시 감소를 보이며 퇴행하는 양상이 나타나 꾸준한 발달을 이루지 못하고 있었는데 가장 눈에 띄는 것은 사용 빈도가 높은 '-ㄴ/는 것'과 '-기'의 감소였다. 명사절을 형성하는 문법 항목의 사용 양상을 살펴보면 예문(17)과 같다.

(17) 가. 나는 피아노 치<u>는 것</u>* 좋아해요.(BW4)
나. 나의 고모는 상하이에 있<u>기</u> 때문에 상하이에 갔어요.(CW12)
다. 우리* 재료를 사고 어떻게 해야 하<u>는지</u> 함께 공부했어요.(AW23)
라. 그래서 다치거나 아팠<u>던 것</u> 다 보통이에요.(BW30)

계속하여 문어 자료에 나타난 관형사절의 발달을 살펴보도록 한다. 자료에 나타난 관형사절 관련 문법 항목의 총 빈도와 전체 절에서 차지하는 비율을 살펴보면 1~6개월에는 127회(35.4%), 7~12개월에는 196회(26.6%), 13~18개월에는 311회(25.6%)로 사용 빈도는 꾸준히 증가하고 있으나 비율은 지속적으로 감소하고 있었다. 이는 문장 수의 증가로 인한 표준화 빈도의 감소로 인한 것일 수도 있고 기타 절의 증가로 인한 상대적인 비율 감소일 수도 있으나 앞서 내포절의 유형에서 확인하였듯이 시간이 지날수록 부사절의 빈도가 급격히 증가한다는 점을 고려할 때, 이는 관형사절의 사용에 의존하던 문장 확장이 부사절이나 다른 내포절 유형으로 분산되면서 나타나는 현상

일 수도 있을 것이다. 관형사절 관련 문법 항목의 기간별 사용 양상을 세부적으로 살펴보면 <표 26>과 같다.

<표 26> 관형사절 관련 문법 항목의 기간별 사용 양상

구분		문법 항목	1~6개월	7~12개월	13~18개월
관형절	관형사형 전성어미	-ㄴ/는	20.59	17.00	19.05
		-ㄹ	2.76	6.93	11.41
		-던	-	0.76	0.40

관형사절은 관형사형 전성어미를 통해 구성된다. 표에서 확인할 수 있듯이 학습자들이 관형사절을 구성하는 데 가장 많이 사용한 문법 항목은 '-ㄴ/는'과 '-ㄹ'이며 그중 '-ㄴ/는'의 사용 빈도가 압도적으로 높게 나타났다. 기간별 발달을 살펴보면, '-ㄴ/는'의 빈도는 1~6개월에 비해 7~12개월에 약간 감소하고 13~18개월에 다시 증가하며 전체적으로 큰 변화를 보이지 않는 반면 '-ㄹ'의 빈도는 1~6개월에 비해 7~12개월에 증가하고 7~12개월에 비해 13~18개월에 더 높은 빈도를 보이며 꾸준히 사용 빈도가 높아지고 있음을 확인할 수 있었다. 이는 학습자들이 관형사절을 구성함에 있어 초반에는 '-ㄴ/는'만 집중적으로 사용하던 것에서 '-ㄹ'을 병행하여 사용하게 되면서 나타난 현상으로 해석할 수 있다. 반면 '-던'은 1~6개월에는 한 차례도 출현하지 않았다가 7~12개월에 처음 출현하였으나 13~18개월에 다시 빈도가 낮아지며 거의 사용하지 못하는 것을 확인할 수 있었다. 이러한 현상은 명사절 관련 문법 항목의 발달에서도 비슷하게 나타났던 것으로 학습자들이 과거형 관형사형 전성어미 '-던'의 사용을 비교적 어려워하며 늦게 습득하는 것으로 해석할 수 있다. 관형사절을 형성하는 문법 항목의 사용 양상을 살펴보면 예문(18)과 같다.

(18) 가. 나는 야채를 먹지 않고 간식을 많이 먹<u>는</u> 습관을 고치고 싶어요.(CW10)
나. 그런데 나도 나쁜 거짓말 <u>한</u> 적이 있어요.(BW21)
다. 10급 결과가 나왔<u>을</u> 때 너무 즐거웠어요.(AW24)
라. 우리가 중국에 같이 놀<u>던</u> 일 아직도 생각나요.(BW28)

계속하여 부사절을 살펴보도록 한다. 학습자들의 문어 자료에 나타난 부사절 관련 문법 항목의 총 빈도와 전체 절에서 차지하는 비율을 살펴보면 1~6개월에는 120회(33.4%), 7~12개월에는 251회(34.1%), 13~18개월에는 478회(39.4%)로 시간이 흐름에 따라 사용 빈도와 비율 모두 지속적으로 상승하며 꾸준한 발달을 나타내고 있었다.

부사절 관련 문법 항목의 기간별 사용 양상을 세부적으로 살펴보면 다음의 <표 27>과 같다.

<표 27> 부사절 관련 문법 항목의 기간별 사용 양상

구분		문법 항목	1~6개월	7~12개월	13~18개월
부사절	동시적 사건	-며/-면서	0.74	1.39	0.89
	계기적 사건	-고(서)	4.60	7.05	11.41
		-아(서)/어(서)	2.02	3.15	5.95
	사건의 전환	-다(가)	-	0.38	0.10
	이유 · 원인	-아(서)/어(서)	3.68	3.65	5.46
		-니/니까	-	0.63	5.26
		-다가	0.18	-	-
		-라(서)	-	0.13	-
	조건 · 가정	-면	8.46	10.08	10.32
		-아야/어야(지)	-	0.13	0.20
		-려면	-	0.13	0.10
		-라면	-	-	-
		-다면	-	-	0.10
	인정 · 양보	-아도/어도	0.55	1.26	1.39
		-ㄴ데	-	-	-
	목적 · 의도	-러	0.37	0.88	1.19
		-려고	-	0.25	0.30
	배경 · 상황 설명	-ㄴ데	0.37	0.50	1.98
		-니	-	-	-
		-더니	-	0.25	-
	수단 · 방법	-아(서)/-어(서)	-	-	0.20
		-고	0.74	0.38	0.50
	방식 · 정도	-게	0.37	0.88	0.79
		-이	-	0.13	0.20
	사역	-게	-	0.38	1.09

표에서 확인할 수 있듯이 학습자들은 관찰 기간 동안 총 22가지의 문법 항목을 사용하여 부사절을 산출하고 있었다. 단일 문법 항목으로 볼 때, 다른 문법 항목에 비해 뚜렷하게 많이 사용된 항목은 '조건·가정'의 '-면', '계기적 사건'의 '-고(서)', '이유·원인'의 '-아(서)/어(서)', 그리고 '계기적 사건'의 '-아(서)/어(서)'였다. 그러나 구체적인 빈도 순위는 기간별로 다르게 나타났는데 1~6개월과 7~12개월에는 모두 '-면'>'-고(서)'>'이유'의 '-아(서)/어(서)'>'계기'의 '-아(서)/어(서)'의 순으로 나타난 반면 13~18개월에는 '-고(서)'>'-면'>'계기'의 '-아(서)/어(서)'>'이유'의 '-아(서)/어(서)'의 순으로 나타나 13~18개월에 '계기적 사건'의 '-고(서)'와 '-아(서)/어(서)'의 사용 빈도가 급

격히 증가하는 것을 확인할 수 있었다.

의미 범주별로 사용 빈도 순위를 살펴보면 '계기적 사건'>'조건·가정'>'이유·원인'>'인정·양보'>'목적·의도'>'배경·상황 설명'>'동시적 사건'>'방식·정도' >'수단·방법'>'사역'>'사건의 전환'의 순으로 나타났다. 아래 의미 범주를 단위로 각 문법 항목의 사용 양상을 세부적으로 살펴보도록 한다.

시간과 관련된 의미 범주인 '동시적 사건', '계기적 사건', '사건의 전환'을 나타내는 문법 항목을 살펴보면, '동시적 사건'을 나타내는 문법 항목은 '-며/-면서'로 1~6개월보다 7~12개월에 빈도가 증가하지만 13~18개월에 다시 감소하였고 '계기적 사건'을 나타내는 문법 항목은 '-고(서)'와 '-아(서)/어(서)' 두 가지만 사용되었는데 두 항목 모두 시간이 지날수록 빈도가 빠르게 증가하며 발달을 보였다. 그러나 '-자', '-자마자' 등의 항목은 한 차례도 나타나지 않았다. '사건의 전환'을 나타내는 문법 항목은 '-다(가)' 한 항목으로 1~6개월에는 산출이 없다가 7~12개월에 처음으로 사용하기 시작하였으나 13~18개월에 다시 빈도가 감소하는 것을 확인할 수 있었다. 시간 관련 선후 순서를 나타내는 문법 항목의 사용 양상을 살펴보면 다음의 예문(19)와 같다.

(19) 가. 집에서 애니메이션 보고 딸기 먹으면서 하루 보냈어요.(BW14)
나. 우리 할아버지하고 할머니 지금도 중국에 있어요. 그래서 나는 이번 여름 방학 때 중국에 가서 볼 거에요*.(CW16)
다. 그리고 나는 담임선생님* 말하고 병원에 갔어요.(CW22)
라. 같이 놀다가 잠깐 후에 얘가 안 하고 싶어다고* 말했어요.(AW25)

'이유·원인'의 의미를 표현하는 문법 항목으로는 '-아(서)/어(서)', '-니/니까', '-다가', '라(서)'가 사용되었는데 그중 '아(서)/어(서)'의 사용 빈도가 압도적으로 높았고 '-니/니까'는 7~12개월부터 나타나기 시작하였으나 그 빈도가 빠르게 증가하여 13~18개월에는 '-아(서)/어(서)'와 비슷한 빈도로 나타나기도 하였다. 이 외에 '-다가'와 '-라(서)'는 그 빈도가 매우 낮고 꾸준한 산출을 보이지는 않았는데 '-다가'는 1~6개월에만 나타났고 '-라(서)'는 7~12개월에만 나타났다. '이유·원인'의 의미를 표현하는 문법 항목 중 한 차례도 출현하지 않은 것은 '-느라(고)', '-므로', '-기에/길래' 등이 있었다. 학습자가 산출한 실제 문장을 살펴보면 예문(20)과 같다.

(20) 가. 하지만 SQI큐브는 조금 비싸서 나는 안 샀어요.(AW8)
나. 우리 학교에 좋은 선생님 많이 있으니까 당연히 잘 지내지.(BW24)

다. 나의* 11살 때 농구 하다가 다리랑 손을 다쳤어요.(CW4)
라. 한 명씩 들어가서 보는 시험이라 너무 긴장해요*.(AW24)

'조건·가정'의 의미를 나타내는 문법 항목으로는 '-면', '아야/어야(지)', '-려면', '-다면'이 사용되었다. 그중에서 '-면'의 사용 빈도가 두드러지게 높았으며 시간의 흐름에 따라 꾸준히 빈도가 증가하고 있는 것을 확인할 수 있었다. '-면'을 제외한 기타 문법 항목은 출현 시기가 비교적 늦었으며 사용 빈도 또한 매우 낮았는데 '-아야/어야(지)'와 '-려면'은 7~12개월에 처음 출현하였고 '-다면'은 13~18개월에 처음 출현하였다. 그리고 '-라면', '-자면', '-냐면'과 같은 문법은 한 차례도 나타나지 않았다. 이는 학습자들이 문장을 확장함에 있어 '조건·가정'의 부사절을 매우 빈번하게 사용하지만 주로 '-면'에 기대어 의미를 표현하고 있으며 다양한 문법 항목의 습득과 사용이 이루어지지 못하고 있는 것으로 해석할 수 있다. '조건·가정'의 문법 항목과 관련하여 학습자가 실제 산출한 문장은 다음의 예문(21)과 같다.

(21) 가. 밖에서 놀면 기분이 좋아요.(CW2)
나. 시험이 끝난 후 선생님은 시간이 지나야 결과가 나온다고 말했어요.(AW24)
다. 고민을 해결하려면 한국어 공부를 열심히 하고 한국 친구를 많이 사귀고...(AW20)
라. 얘가 게임을 같이 하자라고 말하지 않았다면 이렇게 되지 않았을 거예요.(AW25)

'인정·양보'의 의미를 나타내는 문법 항목은 '-아도/어도'와 '-ㄴ데' 두 가지가 사용되고 있었는데 그중 '-아도/어도'가 주로 사용되고 있었으며 시간이 지날수록 그 빈도가 증가하고 있었다. '-ㄴ데'는 13~18개월에 처음 사용되며 '-아도/어도'에 비해 늦게 나타나고 있었다. '인정·양보'의 의미를 표현할 수 있는 문법 항목 중 한 차례도 나타나지 않은 항목으로는 '-ㄹ지라도', '-더라도', '-라도', '-아야/어야' 등이 있었다. 그러나 모어화자의 자료[68]에서도 '인정·양보'의 의미는 거의 '-아도/어도' 위주로 출현하고 있다는 점을 참고할 때 이는 자연스러운 현상이기도 하다. '인정·양보'를 나타내는 문법 항목의 사용 양상을 살펴보면 다음의 예문(22)와 같다.

(22) 가. 어떻게 생각해도 내가 잘못했어요.(BW21)
나. 그리고 저는 게임을 해도 돈을 벌 수 있어요.(AW29)

[68] 본고에서 비교한 모어화자의 자료는 김현아(2016)와 최수완(2019)이며 자세한 비교는 5장에서 다시 다루도록 한다.

다. 아무 말도 안 하는데 마음이 조금 따뜻해 졌어요.(BW30)

'목적·의도'를 나타내는 문법 항목으로는 '-러'와 '-려고'가 사용되고 있었는데 '-러'의 경우 기간별로 빈도가 꾸준히 증가하며 발달을 보이고 있었고 '-려고'의 경우 7~12개월에 처음 출현하며 '-러'에 비해 늦게 사용되기 시작하였으나 13~18개월에 산출 빈도가 증가하며 발달을 보이고 있었다. '목적·의도'를 표현하는 문법 항목 중 한 차례도 나타나지 않은 항목에는 '-고자', '-도록', '-게' 등이 있었다. 그러나 '목적·의도'를 표현하는 문법 항목 역시 모어화자의 자료에서도 '-러'와 '-려고'에 의존하고 있다는 점에서 본고의 학습자와 동일하였다. 학습자가 산출한 문장을 살펴보면 예문(23)과 같다.

(23) 가. 한국에 생활을 하러 왔어요.(BW15)
나. 우리는 롯데월드에* 다 놀고 밥 먹으러 갔어요.(CW23)
다. 할머니가 저와 사이가 좋아지려고 매일 열심히 중국말을 배웠어요.(AW30)

'배경·상황 설명'의 의미를 표현하는 데 사용된 문법 항목으로는 '-ㄴ데'와 '-더니' 두 가지가 있었다. 그중 '-더니'는 7~12개월에만 나타났고 그 이전과 이후에는 사용되지 않고 있었으며 '배경·상황 설명'을 나타내는 '-(으)니'는 한 차례도 나타나지 않았다. 학습자가 산출한 문장을 살펴보면 예문(24)와 같다.

(24) 가. 방에 들어갔더니 중간에 중국 가야금은* 한 대 놓여 있었어요.(AW24)
나. 수행평가* 쓰고 있는데 선생님이 오고* 'OO아, 잘 하고 있어'라고 했어.(BW27)

'수단·방법'을 의미하는 문법 항목으로는 '-고'와 '-아(서)/어(서)'를 들 수 있다. 학습자들은 '-고'와 '-아(서)/어(서)'를 모두 사용하고 있었으나 '-고'는 세 기간 모두에서 산출된 반면 '-아(서)/어(서)'는 13~18개월에 이르러서 처음 사용되었다. 이러한 현상은 본고와 비슷한 주제로 작성된 모어화자 청소년의 작문을 분석한 김현아(2016)나 최수완(2019)에서 모두 '수단·방법'의 '-고'와 '-아(서)/어(서)' 중 '-아(서)/어(서)'가 더 높은 비율로 나타나고 있다는 점과 대조된다. 수단·방법'과 관련한 문법 항목의 사용 양상은 다음의 예문(25)와 같다.

(25) 가. 그리고 우리는 지하철* 타고 집으로 돌아갔어요*.(CW7)

나. 모두 끝나고 우리가 장화를 신고 장갑을 끼고 조개를 잡으러...(BW12)
다. 의자에 앉아서 오늘 배운 내용을 정리하고 오이 한 개 가지고 아침 밥으로 먹었어요.(AW32)

'방식·정도'를 표현하는 문법 항목으로는 '-게'와 '-이' 두 가지가 사용되고 있었는데 '-게'는 1~6개월부터 관찰 기간 동안 꾸준히 출현한 반면 '-이'는 7~12개월에 처음 출현하여 '-게'보다 늦게 나타났으며 '-게'에 비해 사용 빈도 또한 낮았다. 그리고 '방식·정도'를 의미하는 문법 항목 중 출현하지 않은 항목으로는 '-도록'이 있었다. '방식·정도'와 관련한 문법 항목의 사용 양상은 다음의 예문(26)과 같다.

(26) 가. 우리는 1년 넘게 못 만났어요.(AW16)
나. 그래서 행복하면서 마음이 힘들게 10일 살았어요.(CW13)
다. 내가 부모님 옆에 있을 때 스트레스 없이 살 수 있어요.(BW29)

'사역'의 의미를 나타내는 '-게'는 1~6개월에는 사용되지 않다가 7~12개월에 처음 출현하였고 이후 13~18개월에 그 사용 빈도가 급격하게 증가하면서 빈번하게 사용되고 있었다. '사역' 관련 문법 항목의 사용 양상을 살펴보면 예문(27)과 같다.

(27) 가. 좋은 거짓말은 사람을 기쁘게 하고 행복하게 해요.(AW21)
나. 거짓말을 잘못 사용하면 사람을 기분 나쁘게 할 수 있어요.(CW21)
다. 너무 짜증 나* 때 하면 기분을 좋게 해줄 수 있어요.(CW34)

학습자들의 부사절 관련 문법 항목의 산출 양상을 정리하면 다음과 같다. 먼저 부사절 관련 문법 항목의 발달과 관련해서 사용 빈도가 가장 높은 항목은 '조건·가정'의 '-면', '계기적 사건'의 '-고(서)', '이유·원인'의 '-아(서)/어(서)', '계기적 사건'의 '-아(서)/어(서)'였고 세 기간 동안 꾸준히 빈도가 상승하는 항목은 '계기적 사건'의 '-고(서)'와 '-아(서)/어(서)', '조건·가정'의 '-면', '인정·양보'의 '-아도/어도', '목적·의도'의 '-러', '배경·상황 설명'의 '-ㄴ데'였다. 반면 사용 빈도가 감소하는 문법들도 있었는데 7~12개월에 빈도가 감소하는 문법 항목은 '이유·원인'의 '-아(서)/어(서)'와 '-다가', '수단·방법'의 '-고'이고 13~18개월에 빈도가 감소하는 문법 항목은 '동시적 사건'의 '-며/-면서', '사건의 전환'의 '-다(가)', '이유·원인'의 '-라(서)', '조건·가정'의 '-려면', '배경·상황 설명'의 '-더니', '방식·정도'의 '-게'가 있었다.

다음으로, 발달 시점과 관련해서 살펴보면 전체 부사절 관련 문법 항목 중 1~6개월에 사용되지 않다가 7~12개월에 새로 나타나기 시작한 문법 항목으로는 '사건의 전환'의 '-다(가)', '이유·원인'의 '-니/니까'와 '-라(서)', '조건·가정'의 '-아야/어야(지)'와 '-려면', '목적·의도'의 '-려고', '배경·상황 설명'의 '-더니', '방식·정도'의 '-이', '사역'의 '-게'가 있었고 1~12개월까지 사용되지 않다가 13~18개월에 들어서서 새로 나타난 문법 항목에는 '조건·가정'의 '-다면', '인정·양보'의 '-ㄴ데', '수단·방법'의 '-아(서)/어(서)'가 있었다.

마지막으로 인용절 관련 문법 항목의 사용 양상을 살펴보도록 한다. 문어 자료에 나타난 인용절 관련 문법 항목의 총 빈도와 전체 절에서 차지하는 비율을 살펴보면 1~6개월에는 5회(1.4%), 7~12개월에는 36회(4.9%), 13~18개월에는 49회(4.0%)로 7~12개월에는 1~6개월보다 사용 빈도와 비율이 큰 폭으로 증가하였고 13~18개월에는 사용 빈도 측면에서는 증가하였으나 비율은 약간 하락하는 양상을 보이고 있었다. 이는 문장 수의 증가에 따른 표준화 빈도의 퇴행으로도 볼 수 있고 다른 절의 비율 증가에 따른 상대적 퇴행일 수도 있다. 인용절 관련 문법 항목의 기간별 사용 양상을 세부적으로 살펴보면 다음의 <표 28>과 같다.

<표 28> 인용절 관련 문법 항목의 기간별 사용 양상

구분		문법 항목	1~6개월	7~12개월	13~18개월
인용절	직접 인용	-(이)라고	0.18	1.89	1.39
		-하고	-	-	0.20
	간접 인용	-다고/-(이)라고	0.74	2.39	3.17
		-냐고	-	0.25	-
		-(으)라고	-	-	0.10

인용절은 '직접 인용절'과 '간접 인용절'로 나눌 수 있는데 표에서 살펴보면 세 기간 모두 '간접 인용'의 문법 항목을 더 많이 사용하고 있음을 알 수 있다. 문법 항목별로 살펴보면 학습자들은 관찰 기간 동안 총 5가지의 문법 항목을 사용하여 인용절을 산출하고 있었는데 '직접 인용'에서 많이 사용된 문법 항목은 '-(이)라고'이고 '간접 인용'에서 가장 많이 사용된 문법 항목은 '-다고/-(이)라고'였다. 그중에서 '-다고/-(이)라고'가 더 많이 사용되었으며 1~6개월보다 7~12개월에 더 많이 사용하고 7~12개월보다 13~18개월에 더 많이 사용하며 그 빈도가 점점 증가하고 있었다. 반면 '직접 인용'의 '-(이)라고'는 7~12개월에는 1~6개월보다 대폭 증가하였으나 13~18개월에 다시 감소하는 양상을 보이고 있었다. 이 외에 '-하고'와 '-냐고', '-(으)라고'는 앞의 두 항목

에 비해 늦게 출현한 데다가 사용 빈도가 매우 저조하여 인용절은 주로 '-(이)라고'와 '-다고/-(이)라고'에 의해 구성되고 있음을 알 수 있었다. 학습자들이 산출한 인용절을 살펴보면 예문(28)과 같다.

(28) 가. OO이 갑자기 '내일 어떻게 보내고 싶어?'라고 저한테 물어봤어요. (BW14)
나. B가 A한테 '사람이 이렇게 많을* 것을 봐라...'라고 말했어요.(AW33)
다. 그냥 나는 운수 좋은 사람이라고 생각하면 스트레스 조금 없어져요. (BW35)
라. B는 A가 거짓말을 하고 있다고 생각해요. (AW33)
마. 엄마는 나한테 춥냐고 물어봤어요. (AW21)

문어 자료에 나타난 대등절과 내포절의 유형을 정리해 보면, 학습자들은 대등절에서 5가지, 명사절에서 8가지, 관형절에서 3가지, 부사절에서 23가지, 인용절에서 5가지의 문법을 사용하여 다양한 절들을 구성하며 문장을 확장해 나가고 있었다. 기간별로 살펴보면 가장 안정적으로 발달하는 것은 대등절 관련 문법 항목이고 지속적인 발달에 어려움을 겪는 항목은 명사절 관련 문법 항목이었다. 특히 명사절은 7~12개월에는 빠르게 발달하지만 이후 13~18개월에는 사용 빈도 측면에서 뚜렷한 퇴행이 나타났으며 특히 사용 비중이 가장 높은 '-기'와 '-ㄴ/는 것'의 사용 빈도가 오히려 낮아지고 있었다. 관형사형 전성어미에서는 과거형 관형사형 전성어미 '-던'의 사용이 비교적 늦게 나타났으며 사용 빈도 또한 매우 저조하여 학습자들이 '-던'의 습득을 어려워하는 것으로 나타났다. 그러나 전반적으로 학습자들은 시간이 지남에 따라 더 다양한 문법 항목을 사용하여 문장을 확장하려는 시도를 보이고 있었으며 대부분의 문법 항목은 사용 빈도가 증가하면서 학습자들이 점점 더 복잡하고 정교한 문장을 구성해가고 있음을 확인할 수 있었다.

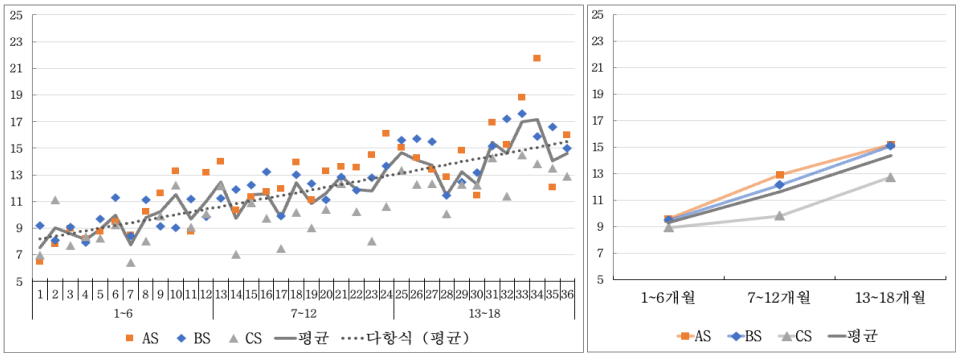
1.2. 구어의 통사 복잡성 발달

구어 자료는 AS-unit을 단위로 분석하였기 때문에 AS-unit을 기본 분석 단위로 AS-unit당 형태소 수, AS-unit당 절 수로 양적인 측면의 통사 복잡성 확장을 측정하였다. 그리고 문어 자료의 분석과 마찬가지로 대등절 수와 내포절 수, 유형별 내포절 및 단위 확장에 관여하는 각 문법 항목들의 사용 빈도를 살펴보도록 한다.

1.2.1. AS-unit 길이

구어의 단위 길이 역시 문어와 마찬가지로 형태소 수로 측정하였다. 그러나 구어의 기본 분석 단위는 AS-unit을 사용하였기 때문에 단위 길이는 AS-unit당 형태소 수로 표시하도록 한다. 세 명의 학습자는 1년 6개월 간 36회에 걸쳐 108개의 음성 자료에서 총 2,327개의 AS-unit과 27,510개의 형태소를 산출하였으며 단위당 형태소 수의 평균은 11.82개였다. 각 시점별 AS-unit 길이와 6개월을 단위로 살펴본 AS-unit 길이의 발달은 다음의 <그림 10>과 같다.

<그림 10> 시점별 AS-unit 길이와 6개월 단위 발달 추이



세 학습자의 평균을 통해 학습자들의 AS-unit 길이 발달을 살펴보면 눈에 띄는 퇴행 구간이 나타나기도 하지만 전반적으로 안정적이고 직선에 가까운 상승 추세를 보이며 시간의 흐름에 따라 길이가 점점 길어지는 것을 확인할 수 있었다. 6개월을 단위로 발달을 살펴보면 7~12개월에는 1~6개월보다 2.31개, 13~18개월에는 7~12개월보다 2.71개 증가하여 시간이 지날수록 증가 속도가 미세하게 더 빨라지는 양상을 보였다.

기간별 AS-unit 길이의 증가가 통계적으로 유의한지를 검증하기 위해 일원배치 분산분석(One-Way ANOVA)을 실시하였다. 그 결과 <표 29>와 같이 F=55.376, p=.000으로 기간별 단위 길이의 차이가 통계적으로 유의한 것으로 나타났다. 사후검정 결과, 세 구간 사이에 모두 유의한 차이(1~6개월<7~12개월<13~18개월)가 발견되었다. 즉 학습자들의 구어 단위 길이는 시간이 흐름에 따라 꾸준하고 뚜렷하게 증가하고 있었다.

<표 29> 기간별 AS-unit 길이

측정 지표	기술통계량				F/p	사후검정
	기간(개월)	N	Mean	SD		
AS-unit 길이	1~6	36(a)	9.35	1.68	55.376/.000***	a<b<c
	7~12	36(b)	11.66	1.97		
	13~18	36(c)	14.36	2.36		

*p<.05 **p<.01 ***p<.001

학습자별로 발달을 살펴보면 세 학습자는 거의 같은 수준에서 출발하여 모두 평균과 마찬가지로 전체적으로 단위 길이가 길어지는 추세를 보이지만 A, B학습자가 유사한 속도로 길이를 늘려가는 것과 달리 C학습자는 7~12개월까지 느린 속도로 발달하다가 13~18개월에 들어선 뒤 A, B학습자 비슷한 증가 속도를 보이고 있었다. 세 학습자의 AS-unit 길이의 평균은 A>B>C순으로 나타났는데 학습자별 AS-unit 길이의 차이가 통계적으로 유의한지 확인하기 위해 일원배치 분산분석을 실시한 결과 <표 30>에서와 같이 F=5.923, p=.004로 학습자 간에 유의한 차이가 있는 것으로 나타났다. 사후검정 결과 A, B학습자의 AS-unit 길이가 C학습자에 비해 유의하게 긴 것으로 나타났다.

<표 30> 학습자별 AS-unit 길이

측정 지표	기술통계량				F/p	사후검정
	학습자	N	Mean	SD		
AS-unit 길이	AS	36	12.59	3.23	5.923/.004**	AS,BS>CS
	BS	36	12.26	2.70		
	CS	36	10.51	2.23		

*p<.05 **p<.01 ***p<.001

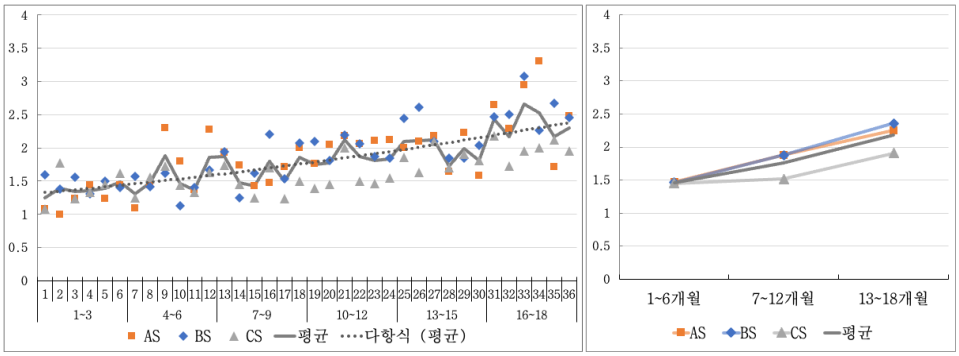
정리하면, 학습자들의 AS-unit 길이는 문장 길이와 마찬가지로 시간의 흐름에 따라 안정적인 증가 추세를 보이며 6개월을 단위로 단계적으로 유의미하게 증가하고 있었다. 학습자 간 증가 속도는 서로 다르게 나타났지만 이러한 차이는 후반으로 가면서 좁혀지는 추세를 보이고 있었다.

1.2.2. AS-unit 확장

1.2.2.1. AS-unit당 절 수

시간의 흐름에 따른 AS-unit당 절 수의 변화 발달을 살펴보면 <그림 11>과 같다. 학습자들의 AS-unit당 절 수는 작은 폭의 등락을 반복하지만 시간의 흐름에 따라 약간 증가하는 추세를 보이고 있었다. 2차 다항식 추세선은 후반으로 갈수록 미세하게 가팔라지는 모양을 띠고 있는데 이는 시간이 지날수록 증가 속도가 빨라짐을 의미한다.

<그림 11> 시점별 AS-unit당 절 수와 6개월 단위 발달 추이



AS-unit당 절 수의 기간별 증가폭을 확인해 보면 7~12개월에는 1~6개월보다 단위당 0.29개의 절을 더 산출하였고 13~18개월에는 7~12개월보다 단위당 0.42개의 절을 더 산출하며 시간이 흐를수록 증가폭이 커지는 것을 확인할 수 있다. 이는 AS-unit 길이의 발달과 같은 현상으로 절 수가 늘어날 수록 길이가 길어지는 한국어의 보편적인 현상이기도 하다.

기간별 AS-unit당 절 수의 증가가 통계적으로 유의한지를 검증하기 위해 일원배치 분산분석(One-Way ANOVA)을 실시하였다. 그 결과 <표 31>과 같이 F=35.869, p=.000으로 기간별 AS-unit당 절 수의 차이가 통계적으로 유의함을 확인할 수 있었다. 사후검정 결과, 세 기간 사이에 모두 유의미한 차이(1~6개월<7~12개월<13~18개월)가 있는 것으로 확인되었다. 즉 학습자들은 시간이 지날수록 하나의 AS-unit에 더 많은 절을 사용하며 문장을 확장시켜 나가고 있었다.

<표 31> 기간별 AS-unit당 절 수

측정 지표	기술통계량				F/p	사후검정
	기간(개월)	N	Mean	SD		
AS-unit당 절 수	1~6	36(a)	1.46	0.29	35.869/.000***	a<b<c
	7~12	36(b)	1.76	0.30		
	13~18	36(c)	2.18	0.42		

*p<.05 **p<.01 ***p<.001

학습자별로 나누어 살펴보면 세 학습자 모두 평균과 마찬가지로 전체적으로 AS-unit당 절 수가 증가하는 추세를 보이지만 A와 B학습자가 거의 일치된 발달 추이를 보이고 있는 것과 달리 C학습자는 7~12개월까지 거의 정체되어 있다가 13~18개월에 들어서 A, B학습자와 비슷한 속도로 발달하고 있었다. 세 학습자의 AS-unit당 절 수의 평균은 B>A>C순으로 나타났는데 그 차이가 통계적으로 유의한지 확인하기 위해 일원배치 분산분석을 실시한 결과 <표 32>와 같이 F=6.015, p=.004로 학습자 간에 유의한 차이가 있는 것으로 나타났다. 그리고 사후검정 결과 A, B학습자의 AS-unit당 절 수가 C학습자에 비해 유의하게 높은 것으로 나타났다. 앞에서 살펴본 AS-unit 길이와 비교해볼 때, C학습자의 절 수가 가장 낮은 것은 동일하지만 AS-unit 길이는 A학습자가 B학습자에 비해 긴 것과 달리 AS-unit당 절 수는 B학습자가 더 높다는 것은 A학습자가 B학습자에 비해 더 긴 절을 산출하고 있음을 알 수 있는 부분이다.

<표 32> 학습자별 AS-unit당 절 수

측정 지표	기술통계량				F/p	사후검정
	학습자	N	Mean	SD		
AS-unit당 절 수	AS	36	1.87	0.52	6.015/.004**	BS,AS>CS
	BS	36	1.90	0.46		
	CS	36	1.63	0.28		

*p<.05 **p<.01 ***p<.001

정리하면, AS-unit당 절 수의 발달은 AS-unit 길이와 마찬가지로 시간이 흐름에 따라 점차 증가하고 있었다. 이는 학습자들이 시간이 흐를수록 하나의 AS-unit 안에 더 많은 절을 사용하여 긴 문장을 산출하며 통사 복잡성을 발달시키고 있음을 의미한다. 그러나 문장 길이와 문장당 절 수의 발달 속도가 시간이 지나도 일정함을 유지하는 것과 달리 AS-unit 길이와 AS-unit당 절 수는 시간이 지날수록 증가폭이 커지는 현상을 발견할 수 있었다. 이는 한국 거주 기간이 늘어날수록 구어의 통사 복잡성 발달 속도가

점점 더 빨라진다는 것을 의미한다.

1.2.3. 대등절·내포절의 유형

1.2.3.1. AS-unit당 대등절 수와 내포절 수

앞서 문장당 대등절 수와 내포절 수에서도 확인하였듯이 대등절과 내포절의 발달은 서로 다르게 나타난다. 구어 자료의 AS-unit당 대등절 수와 내포절 수의 발달을 나눠서 살펴보면 다음의 <표 33>과 같다.

<표 33> 6개월 단위 AS-unit당 대등절 수와 내포절 수

학습자	대등/내포	1~6개월	증가폭	7~12개월	증가폭	13~18개월	평균
A	대등	0.07	+0.12	0.19	+0.07	0.26	0.07
	내포	0.38	+0.30	0.68	+0.30	0.98	0.38
B	대등	0.14	+0.06	0.21	+0.14	0.35	0.14
	내포	0.30	+0.35	0.66	+0.31	0.97	0.30
C	대등	0.05	+0.04	0.09	+0.14	0.23	0.05
	내포	0.40	+0.02	0.42	+0.24	0.66	0.40
평균	대등	0.09	+0.08	0.16	+0.12	0.28	0.09
	내포	0.36	+0.23	0.58	+0.28	0.87	0.36

학습자들의 평균을 통해 AS-unit당 대등절 수와 내포절 수의 발달을 살펴보면 AS-unit당 대등절 수의 평균은 1~6개월에 0.09개, 7~12개월에 0.16개, 13~18개월에 0.28개였고 기간별 AS-unit당 내포절 수의 평균은 각각 1~6개월에 0.36개, 7~12개월에 0.58개, 13~18개월에 0.87개로 학습자들은 관찰 기간 내내 대등절보다 내포절을 더 많이 사용하고 있었으며 증가폭 역시 내포절이 더 높게 나타났다. 이는 문어와 일치하는 결과이며 한국어의 대등절 관련 문법 항목에 비해 내포절 관련 문법 항목이 압도적으로 많은 것에 기인하는 것으로 학습자들은 시간이 지날수록 자연스럽게 내포절을 더 많이 학습해 나가는 것으로 해석할 수 있다.

대등절과 내포절의 기간별 증가폭을 살펴보면 대등절은 7~12개월에 1~6개월보다 0.08개, 13~18개월에 7~12개월보다 0.12개 증가하고 내포절은 7~12개월에 1~6개월보다 0.23개, 13~18개월에 7~12개월보다 0.28개 증가하여 대등절과 내포절 모두 시간이 흐를수록 증가 속도가 빨라지는 것을 확인할 수 있었다.

단위당 대등절과 내포절의 평균 차이가 통계적으로 유의한지 알아보기 위해 일원배치 분산분석을 실시한 결과 <표 34>와 같이 대등절은 F=23.576, p=.000, 내포절은 F=28.931, p=.000으로 대등절과 내포절 모두 기간별로 유의한 차이가 있는 것으로 나타났다. 사후 검정 결과 대등절과 내포절 모두 6개월을 단위로 유의한 차이(1~6개월<7~12개월<13~18개월)를 보였다. 이는 구어에서 대등절과 내포절 수가 지속적이고 유의미하게 증가하고 있음을 의미한다.

<표 34> 기간별 AS-unit당 대등절 수와 내포절 수

측정 지표	기술통계량				F/p	사후검정
	기간(개월)	N	Mean	SD		
AS-unit당 대등절 수	1~6	36(a)	0.09	0.09	23.576/.000***	a<b<c
	7~12	36(b)	0.16	0.12		
	13~18	36(c)	0.28	0.14		
AS-unit당 내포절 수	1~6	36(a)	0.36	0.25	28.931/.000***	a<b<c
	7~12	36(b)	0.58	0.26		
	13~18	36(c)	0.87	0.33		

*p<.05 **p<.01 ***p<.001

학습자별로 AS-unit당 대등절 수와 내포절 수의 차이가 통계적으로 유의한지 확인하기 위해 일원배치 분산분석을 실시한 결과 <표 35>에서와 같이 대등절은 F=5.308, p=.006, 내포절은 F=3.910, p=.025로 대등절과 내포절 모두 학습자 간 유의한 차이가 있는 것으로 나타났다. 그러나 사후검정 결과 대등절은 B학습자와 C학습자(B>C) 사이에 차이가 있었으나 내포절은 학습자 간 차이를 통계적으로 변별하지 못하였다. 이는 구어의 내포절은 학습자들 간에 뚜렷한 차이가 없이 비슷한 수준으로 발달하고 있음을 보여주는 결과이다.

<표 35> 학습자별 AS-unit당 대등절 수와 내포절 수

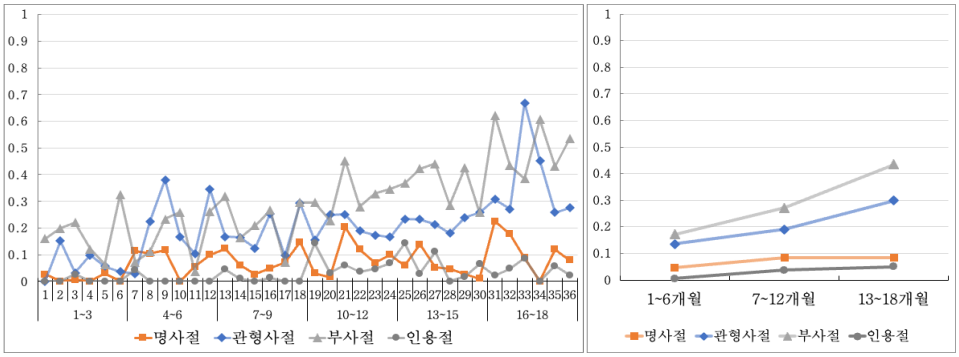
측정 지표	기술통계량				F/p	사후검정
	학습자	N	Mean	SD		
AS-unit당 대등절 수	AS	36	0.17	0.16	5.308/.006**	BS>CS
	BS	36	0.23	0.14		
	CS	36	0.13	0.11		
AS-unit당 내포절 수	AS	36	0.68	0.42	3.910/.025*	-
	BS	36	0.64	0.36		
	CS	36	0.49	0.23		

*p<.05 **p<.01 ***p<.001

1.2.3.2. 내포절의 유형

이어서 각 유형별 내포절의 발달을 살펴보도록 한다. 아래 <그림 12>는 세 학습자들의 평균으로 표시한 유형별 내포절의 시점별 그래프와 6개월 단위 기간별 발달 추이이다. 그래프에서 확인할 수 있듯이 부사절과 관형사절은 시간의 흐름에 따라 뚜렷한 발달을 보였고 명사절은 7~12개월에는 약간 증가하였지만 13~18개월에는 정체가 나타났으며 인용절은 작은 폭이지만 발달을 이어 나가고 있었다.

<그림 12> 시점별 유형별 내포절 수와 6개월 단위 발달 추이



각 유형별 내포절의 발달이 기간별로 유의미한 차이가 있는지 확인하기 위해 일원배치 분산분석을 실시한 결과 <표 36>과 같이 명사절을 제외한 기타 세 가지 절은 기간별로 유의한 차이가 있는 것으로 나타났다. 사후검정 결과 부사절은 세 기간의 차이(1~6개월<7~12개월<13~18개월)가 모두 유의하였고 관형사절은 1~6개월, 7~12개월과 13~18개월(1~6개월, 7~12개월<13~18개월) 간의 차이가 유의하였으며 인용절은 1~6개월과 13~18개월(1~6개월<13~18개월) 간에 유의한 차이를 보였다.

이러한 결과는 여러 내포절의 발달 시기가 서로 다르다는 것을 보여주는 결과로 부사절의 사용은 초반부터 안정적이고 빠르게 증가하고 있는 반면 관형사절은 13~18개월에 이르러서야 뚜렷한 발달이 나타나고 인용절은 증가 속도가 매우 완만하여 1년 정도 간격을 두고 유의미한 발달이 일어나며 명사절은 관찰 기간 내내 뚜렷한 발달을 보이지 않는 것으로 해석할 수 있다. 그러나 6개월을 단위로 살펴본 구어의 각 유형별 내포절은 퇴행이 나타나는 구간 없이 지속적으로 증가하고 있다는 점에서 문어의 발달과 차이가 있었다. 문어와 구어의 각 유형별 내포절의 기간별 산출 빈도를 살펴보면 세 구간 모두에서 구어의 산출 빈도가 더 낮은 것을 확인할 수 있는데 이는 구어에서는 실시

간으로 언어를 산출하면서 새로운 시도보다는 문어에서 익숙해진 항목들을 위주로 사용하게 되고 이러한 현상이 더 안정적인 증가로 나타났을 수도 있을 것이다.

<표 36> 유형별 내포절의 기간별 사용 양상

측정 지표	기술통계량				F/p	사후검정
	기간(개월)	N	Mean	SD		
명사절	1~6	36(a)	0.05	0.08	2.297/.106	-
	7~12	36(b)	0.08	0.10		
	13~18	36(c)	0.09	0.09		
관형절	1~6	36(a)	0.14	0.17	8.385/.000***	a,b<c
	7~12	36(b)	0.19	0.13		
	13~18	36(c)	0.30	0.21		
부사절	1~6	36(a)	0.17	0.16	26.675/.000***	a<b<c
	7~12	36(b)	0.27	0.14		
	13~18	36(c)	0.43	0.16		
인용절	1~6	36(a)	0.01	0.02	13.775/.001***	a<c
	7~12	36(b)	0.04	0.07		
	13~18	36(c)	0.05	0.07		

*p<.05 **p<.01 ***p<.001

유형별 내포절 수의 사용의 학습자 간 차이가 통계적으로 유의한지 확인하기 위해 일원배치 분산분석을 실시한 결과 <표 37>과 같이 관형사절과 인용절의 사용에서만 학습자 간에 유의한 차이가 있는 것으로 나타났다. 사후검정 결과 관형사절과 인용절 모두 A학습자와 C학습자(A>C) 간에만 유의한 차이가 있었다. 이를 통해 구어에서 명사절과 부사절은 학습자 간 차이가 없이 비슷한 수준으로 발달하고 있음을 알 수 있다.

이러한 현상이 나타난 원인을 찾기 위해 문어 자료에 나타난 학습자별 각 유형별 내포절의 빈도와 구어 자료에 나타난 내포절의 빈도를 비교해 보았다. 그 결과, 명사절과 부사절은 모두 문어에서 B학습자가 다른 학습자에 비해 유의하게 높게 나타났던 절이지만 구어에서는 다른 두 학습자는 문어와 비슷한 수준의 명사절과 부사절을 산출하고 있는 반면 B학습자는 문어에 비해 구어에서 현저히 적은 빈도로 산출하게 되며 학습자 간의 차이가 줄어들게 된 것이 가장 직접적인 원인이었다. 관형사절 역시 문어에서는 세 학습자의 관형사절 산출이 비슷하였으나 구어에서 B학습자와 C학습자의 관형사절 산출이 문어에 비해 현저히 낮아지면서 A학습자의 관형사절이 유의하게 높은 것으로 나타났다. 이는 B와 C학습자가 A학습자에 비해 문어와 구어의 불균형이 더 강함을 의미하며 학습자마다 산출 양식이나 내포절의 유형에 따라 발달 수준이 다르게 나타날

수 있음을 알 수 있다.

<표 37> 유형별 내포절의 학습자별 사용 양상

측정 지표	기술통계량				F/p	사후검정
	학습자	N	Mean	SD		
명사절	AS	36	0.06	0.08	.579/.562	-
	BS	36	0.09	0.11		
	CS	36	0.07	0.08		
관형사절	AS	36	0.29	0.24	6.531/.003**	AS>CS
	BS	36	0.19	0.15		
	CS	36	0.14	0.11		
부사절	AS	36	0.27	0.20	1.243/.293	-
	BS	36	0.33	0.19		
	CS	36	0.28	0.17		
인용절	AS	36	0.06	0.08	14.309/.001***	AS>CS
	BS	36	0.03	0.06		
	CS	36	0.01	0.02		

*p<.05 **p<.01 ***p<.001

1.2.4. 대등절과 내포절 관련 문법 항목

이 절에서는 구어 자료에 나타난 대등절과 내포절 관련 문법 항목 및 사용 빈도를 통해 학습자들의 구어 통사 복잡성의 발달에 관여하는 문법 항목들의 발달 양상을 자세히 살펴보도록 한다. 문어 자료와 마찬가지로 각 문법 항목의 표준화 빈도를 산출한 뒤 6개월을 단위로 평균 표준화 빈도를 비교해 보았다.

먼저 대등절을 살펴보도록 한다. 학습자들의 구어 자료에 나타난 대등절 관련 문법 항목의 총 빈도와 전체 절에서 차지하는 비율을 기간별로 살펴보면 1~6개월에는 48회(20.3%), 7~12개월에는 135회(22.4%), 13~18개월에는 232회(23.2%)로 총 사용 빈도뿐만 아니라 비율 또한 증가하고 있음을 확인할 수 있었다. 구체적인 문법 항목의 기간별 사용 양상은 다음의 <표 38>과 같다.

<표 38> 대등절 관련 문법 항목의 기간별 사용 양상

구분		문법 항목	1~6개월	7~12개월	13~18개월
대등절	나열 · 순접	-고	7.02	13.04	19.34
		-며/-면서	-	0.12	-
	대조 · 역접	-ㄴ데	0.51	0.97	1.86
		-지만	0.68	2.05	3.50
	선택	-거나	-	0.12	0.55

표에서 확인할 수 있듯이 학습자들은 관찰 기간 동안 총 5가지의 문법 항목을 사용하여 대등절을 구성하고 있었다. 그중에서 사용 빈도가 가장 높은 것은 '나열·순접'의 '-고'와 '대조·역접'의 '-지만'이었으며 이는 세 기간 모두 동일하게 나타났다. 그러나 '나열·순접'의 '-고' 사용의 빈도가 '-지만'이나 다른 항목에 비해 압도적으로 높게 나타나 학습자들의 대등절 형성은 주로 '-고'에 의존하고 있음을 알 수 있었다.

기간별 발달을 살펴보면 모든 문법 항목의 사용 빈도가 시간이 흐름에 따라 증가하며 발달을 나타내고 있었으나 각 문법 항목들의 출현 시기에는 차이가 있었다. 7가지의 대등절 중 '나열·순접'의 '-고', '대조·역접'의 '-ㄴ데', '-지만'은 관찰 시작 시점인 1~6개월부터 꾸준히 출현하고 있었으나 '나열·순접'의 '-며/-면서'는 7~12개월에만 출현하였고 '선택'의 '-거나'는 7~12개월에 들어서서 출현하기 시작하였다. 이 외에 한 번도 사용되지 않은 문법 항목에는 '대조·역접'의 '-나', '-라', '선택'의 '-든지/든가' 등이 있었다. 대등절을 형성하는 문법 항목의 사용 양상을 살펴보면 예문(29)와 같다.

(29) 가. 산기슭에는 꽃도 많이 있고 나무도 많이 있어요.(AS14)
나. 뚱뚱하지만 잘 생겼어요.(BS4)
다. 나* 하지 말라고 했는데 왜 또 했어.(CS24)
라. 얘네 부모님가* 나의 친구한테* 때리거나 화를 내요.(BS27)

다음 명사절을 살펴보도록 한다. 학습자들의 문어 자료에 나타난 명사절 관련 문법 항목의 총 빈도와 전체 절에서 차지하는 비율을 살펴보면 1~6개월에는 22회(9.3%), 7~12개월에는 69회(11.5%), 13~18개월에는 70회(7.0%)로 7~12개월에는 1~6개월보다 사용 빈도와 비율이 비교적 큰 폭으로 증가하였으나 13~18개월에는 빈도는 정체되고 비율이 하락하며 약간의 퇴행을 나타내고 있었다. 이와 같은 13~18개월의 퇴행은 문어와 구어에서 공통적으로 나타나는 현상으로 앞서 유형별 내포절의 발달에서도 확인한 바 있다. 명사절 관련 문법 항목의 기간별 사용 양상을 세부적으로 살펴보면 다

음의 <표 39>와 같다.

<표 39> 명사절 관련 문법 항목의 기간별 사용 양상

구분		문법 항목	1~6개월	7~12개월	13~18개월
명사절	명사형 전성어미	-ㅁ	-	-	-
		-기	1.37	1.93	1.53
	관형사형 전성어미+것	-ㄴ/는 것	2.40	6.28	5.25
		-ㄹ 것	-	-	0.11
		-던 것	-	-	-
	의문형 종결어미	-ㄴ지/는지	-	-	0.66
		-ㄹ지	-	0.12	0.11
		-ㄹ까	-	-	-

표에서 확인할 수 있듯이 학습자들은 관찰 기간 동안 총 5가지의 문법 항목을 사용하여 명사절을 산출하고 있었다. 그중에서 가장 많이 사용된 항목은 '관형사형 전성어미+ 것'인 '-ㄴ/는 것'이고 두 번째로 많이 사용된 항목은 '명사형 전성어미'인 '-기'로 세 기간 모두 이 두 가지의 문법만 주로 나타나고 있어 학습자들은 구어에서도 문어와 마찬가지로 주로 '-ㄴ/는 것'과 '-기'를 사용하여 명사절을 구성하고 있음을 알 수 있었다.

기간별 발달을 살펴보면, 가장 높은 빈도로 사용되는 '명사형 전성어미'인 '-기'와 관형사형 전성어미+ 것'인 '-ㄴ/는 것'은 7~12개월에는 1~6개월보다 빈도가 증가하였으나 13~18개월에 다시 감소하는 퇴행을 보였다. 그리고 7~12개월에 '의문형 종결 어미'인 '-ㄹ지'가 새롭게 출현하고 13~18개월에는 '관형사형 전성어미+ 것'인 '-ㄹ 것'과 '의문형 종결 어미'인 '-ㄴ지/는지'가 새롭게 출현하며 문법 항목의 종류를 늘려가고 있었다. 그러나 학습자들이 문어 자료에서는 총 8가지의 문법 항목을 사용하였다는 점과 문어 자료에서는 '-ㄹ 것'과 '-ㄴ지/는지', '-ㄹ지', '-ㄹ까' 모두 1~6개월부터 꾸준히 나타났고 '-던 것' 역시 13~18개월에 출현하였던 것과 비교해 보면, 구어에서 명사절 관련 문법 항목의 발달이 문어에 비해 늦게 나타난다고 볼 수 있다. 구어 자료에서 한 차례도 출현하지 않은 문법 항목은 '명사형 전성어미' '-ㅁ'과 '관형사형 전성어미+ 것'인 '-던 것', '의문형 종결 어미'인 '-ㄹ까', '-냐/느냐', '-ㄴ가/는가', '-던가' 등이 있었다. 이 중에서 '-ㅁ'은 구어에서 잘 사용되지 않는 문법이라는 점을 감안하더라도 구어에서 더 적은 종류의 명사절을 사용하고 있음을 알 수 있는 부분이다.

정리하면 학습자들은 문어와 마찬가지로 구어에서도 명사절의 발달이 뚜렷하게 나타나지 않았다. 학습자들은 시간이 지나면서 새로운 문법 항목을 습득하기도 하지만

문어와 마찬가지로 7~12개월에 증가 이후 13~18개월에 다시 감소하는 양상을 보이며 꾸준한 발달을 이어 나가지 못하는 것으로 확인되었다. 명사절 관련 문법 항목의 사용 양상을 살펴보면 예문(30)과 같다.

(30) 가. 그리고 선생님과 싸우기 시작했어요.(AS24)
나. 나는 주변 친구랑 많이 안 싸웠기 때문에 나쁜 친구를 많이 안 사귀었어요.(CS31)
다. 내가* 얘를* 왜 화났는지 잘 모르겠어요.(BS31)
라. 그래서 우리* 한* 일을 할 때 어떻게 해야 할지 신중하게 생각해야 해요.(AS23)

계속하여 구어 자료에 나타난 관형사절 관련 문법 항목을 살펴보도록 한다. 자료에 나타난 관형사절 관련 문법 항목의 총 빈도와 전체 절에서 차지하는 비율을 살펴보면 1~6개월에는 65회(27.5%), 7~12개월에는 147회(24.4%), 13~18개월에는 249회(24.9%)로 사용 빈도는 꾸준히 증가하고 있으나 비율은 1~6개월 이후 약간 감소하였음을 확인할 수 있다. 이는 앞서 문어에서와 마찬가지로 부사절과 같은 다른 절의 빈도가 높아지면서 발생하는 현상으로 해석할 수 있다. 관형사절 관련 문법 항목의 기간별 사용 양상을 세부적으로 살펴보면 다음의 <표 40>과 같다.

<표 40> 관형사절 관련 문법 항목의 기간별 사용 양상

구분		문법 항목	1~6개월	7~12개월	13~18개월
관형절	관형사형 전성어미	-ㄴ/는	10.10	12.56	18.14
		-ㄹ	1.03	4.71	9.07
		-던	-	0.48	-

표에서 확인할 수 있듯이 학습자들이 관형사절을 구성하는 데 가장 많이 사용한 문법 항목은 '-ㄴ/는'과 '-ㄹ'이며 그중 '-ㄴ/는'의 사용 빈도가 압도적으로 높게 나타났다. 기간별 발달을 살펴보면, '-는/는'과 '-ㄹ'의 빈도는 1~6개월, 7~12개월, 13~18개월 세 기간 거쳐 꾸준히 증가하면서 발달을 보이고 있었다. 이는 문어에서 '-ㄴ/는'이 1~6개월에 매우 높은 빈도로 나타나다가 7~12개월과 13~18개월에 감소하는 것과는 다른 양상이다. 한편, '-던'은 1~6개월에는 한 차례도 출현하지 않았다가 7~12개월에 처음 출현하였으나 13~18개월에는 다시 한 차례도 사용되지 않으며 문법의 사용이 지속적으로 유지되지 않았다. 이러한 현상은 문어에서 13~18개월에 들어선 뒤 '-던'의 비율이 약간 감소하는 것과 비슷한 것으로 학습자들은 과거형 관형사형 전성어미 '-던'을 7~12개월 사이에 습득하지만 지속적인 사용은 어려워함을 의미한다. '-던'이나 '-

던 것'의 습득의 어려움은 문어에서도 동일하게 나타난 현상으로 교육 현장에서 더 중점적으로 다루어 질 필요가 있음을 시사한다. 구어 자료에 나타난 관형사절의 사용 양상을 살펴보면 예문(31)과 같다.

(31) 가. 나* 제일 보고 싶은 사람은 할아버지하고 할머니이에요*.(CS12)
나. 내가* 고민* 별로 없는 사람이에요.(BS22)
다. 지금 저는 몸이 아플 때 우리 외할아버지를* 보고 싶어요.(AS30)
라. 중국에서 다녔던 학교는 초등학교 중학교 고등학교* 같이 있는 학교이에요*. (AS23)

계속하여 부사절을 살펴보도록 한다. 학습자들의 문어 자료에 나타난 부사절 관련 문법 항목의 총 빈도와 전체 절에서 차지하는 비율을 살펴보면 1~6개월에는 96회(40.7%), 7~12개월에는 221회(36.7%), 13~18개월에는 404회(40.4%)로 시간이 흐름에 따라 사용 빈도는 증가하지만 비율은 약간 감소하는 것을 확인할 수 있다. 부사절 관련 문법 항목의 기간별 사용 양상을 세부적으로 살펴보면 다음의 <표 41>과 같다.

<표 41> 부사절 관련 문법 항목의 기간별 사용 양상

구분		문법 항목	1~6개월	7~12개월	13~18개월
부사절	동시적 사건	-며/-면서	0.17	0.24	0.98
	계기적 사건	-고(서)	3.60	4.95	9.95
		-아(서)/어(서)	2.40	3.86	5.57
	사건의 전환	-다(가)	-	-	-
	이유 · 원인	-아(서)/어(서)	2.74	4.11	6.45
		-니/니까	-	0.48	2.84
		-다가	-	-	-
		-라(서)	-	-	-
	조건 · 가정	-면	5.82	9.54	13.01
		-아야/어야(지)	-	-	0.11
		-려면	-	0.24	-
		-라면	-	0.12	-
		-다면	-	-	-
	인정 · 양보	-아도/어도	0.17	1.21	0.98
		-ㄴ데	0.17	-	-
	목적 · 의도	-러	-	0.36	0.55
		-려고	-	0.12	-
	배경 · 상황 설명	-ㄴ데	1.20	0.72	1.09
		-니	-	-	0.11
		-더니	-	-	-

	수단 · 방법	-아(서)/-어(서)	0.17	-	0.66
		-고	-	0.36	0.87
	방식 · 정도	-게	-	0.24	0.44
		-이	-	-	0.33
	사역	-게	-	0.12	0.22

표에서 확인할 수 있듯이 학습자들은 관찰 기간 동안 총 20가지의 문법 항목을 사용하여 부사절을 산출하고 있었다. 단일 문법 항목으로 볼 때, 다른 문법 항목에 비해 뚜렷하게 많이 사용된 항목은 '조건·가정'의 '-면', '계기적 사건'의 '-고(서)', '이유·원인'의 '-아(서)/어(서)', 그리고 '계기적 사건'의 '-아(서)/어(서)'였다. 그리고 이러한 순서는 세 기간 모두 동일하게 나타났는데 '-면'>'-고(서)'>이유의 '-아(서)/어(서)'>계기의 '-아(서)/어(서)'의 순으로 나타났다. 그리고 의미 범주별로 빈도 순위를 살펴보면 '계기적 사건'>'조건·가정'>'이유·원인'>'배경·상황 설명'>'인정·양보'>'수단·방법'>'동시적 사건'>'목적·의도'>'방식·정도'>'사역'>'사건의 전환'의 순으로 나타났다. 아래에서는 의미 범주를 단위로 각 문법 항목의 사용 양상을 세부적으로 살펴보도록 한다.

시간과 관련된 의미 범주부터 살펴보면, '동시적 사건'을 나타내는 문법 항목인 '-며/-면서'와 '계기적 사건'을 나타내는 문법 항목은 '-고(서)'와 '-아(서)/어(서)' 3 가지가 사용되었는데 세 항목 모두 시간이 지날수록 빈도가 점점 증가하며 발달하고 있었다. 그러나 문어와 마찬가지로 '-자', '-자마자' 등의 항목은 한 차례도 나타나지 않았다. '사건의 전환'을 나타내는 문법 항목은 구어에서 한 차례도 사용되지 않았는데 이는 문어에서 '-다(가)'가 7~12개월에 처음 출현하여 13~18개월에도 꾸준히 사용되었던 것과 대조되는 부분이다. 그러나 문어에서도 사용 빈도가 높지 않고 13~18개월에 빈도가 다시 감소하였다는 점을 고려할 때 학습자들이 '-다(가)'의 사용에 익숙하지 않음을 알 수 있다. '-다가'는 일상에서 비교적 자주 사용되는 문법 항목으로 모어화자의 언어에서는 비교적 높은 빈도로 나타나는 문법 항목이다. 그러나 모어화자의 구어와 문어에 나타난 연결어미 빈도를 비교한 이민경(2016)에 따르면 '-다가'는 구어보다는 문어에서 더 높은 빈도로 나타나는 항목이기도 하다. 따라서 본고에서는 문어에서도 낮은 빈도로 출현했던 '-다가'가 구어에서는 사용되지 않은 것으로 볼 수도 있을 것이다. 학습자가 산출한 시간 관련 부사절의 실제 문장들을 살펴보면 예문(32)와 같다.

(32) 가. 나는 이 앨범 들으면서 책을 많이 읽었어요.(CS35)
나. 저는 이 말을 듣고 너무 감동했어요.(AS32)
다. 나랑 나* 친구는 바다에 가서 수영하고 물놀이도 했어요.(CS6)

라. 어느 날에 우리 친구 몇 명 한강에 가서 놀았어요.(BS31)

'이유·원인'의 의미를 표현하는 문법 항목으로는 '-아(서)/어(서)', '-니/니까'가 사용되었는데 그중 '아(서)/어(서)'의 사용 빈도가 압도적으로 높았고 '-니/니까'는 7~12개월부터 나타나기 시작하였다. '이유·원인'의 의미를 표현하는 문법 항목 중 한 차례도 출현하지 않은 것은'-다가', '-느라(고)', '-므로', '-기에/길래', '-라(서)' 등이 있었다. 이 중에서 '-므로'나 '-기에'의 경우 구어보다는 문어에서 주로 사용되는 문법이라는 점에서 구어 자료에 나타나지 않은 점은 자연스럽지만 '-라(서)'의 경우 구어에서 사용되는 문법 항목임에도 불구하고 문어 자료에는 나타났던 반면 구어 자료에 나타나지 않았다는 점은 학습자들이 문어체와 구어체의 구분에 대한 인식이 부족하다는 점을 엿볼 수 있는 부분이기도 하다. '이유·원인' 관련 문법 항목을 사용한 실제 문장들을 살펴보면 예문(33)과 같다.

(33) 가. 야채는 맛이 없어서 야채를 안 먹어요.(CS12)
나. 근데 많이 먹어서 비록 맛 없지만 친구랑 이야기하면서 줄도 안 서도 되고 먹는 거가 자유하고* 마음이 더 편해요.(BS26)
다. 그 때 나* 옆에 부모님 없으니까 치료를 잘 안 받았어.(BS29)

'조건·가정'의 의미를 나타내는 문법 항목으로는 '-면', '아야/어야(지)', '-려면'이 사용되었다. 그중에서 '-면'의 사용 빈도가 두드러지게 높았으며 시간의 흐름에 따라 꾸준히 빈도가 증가하고 있는 것을 확인할 수 있었다. '-면'을 제외한 기타 문법 항목은 출현 시기가 비교적 늦었으며 사용 빈도 또한 매우 낮아 거의 사용되지 않고 있었는데 '-아야/어야(지)'는 13~18개월에 처음 출현하였고 '-려면'과 '-라면'은 7~12개월에 출현하였으나 꾸준히 사용되지 못하고 13~18개월에는 나타나지 않았다. 그리고 '-거든', '-자면', '-냐면', '-다면'과 같은 문법들은 한 차례도 나타나지 않았다. 이는 학습자들이 문장을 확장함에 있어 '조건·가정'의 부사절을 매우 빈번하게 사용하지만 주로 '-면'에 기대어 의미를 표현하고 있으며 다양한 문법 항목의 습득과 사용이 이루어지지 못하고 있음으로 해석할 수 있는 것으로 문어와 구어에서 동일하게 나타나는 현상이었다. 이 외에 문어와 구어에서 출현 시기에 차이가 나는 문법 항목은 '-아야/어야(지)'와 '-다면'이었는데 문어에서 '-아야/어야(지)'가 7~12개월에 출현하였고 '-다면' 역시 13~18개월에 사용하기 시작하였다는 점을 고려할 때, 구어에서 문어보다 새로운 문법 항목의 사용이 좀 더 늦게 나타나는 것으로 볼 수 있다. '조건·가정'을 나타내는 부

사절을 안은 문장들을 살펴보면 예문(34)와 같다.

(34) 가. 집에 와서 시간이 있으면 요리도 해요.(AS4)
나. 사람 엄청 많아니까* 줄 엄청 길게 서야 밥을 먹을 수 있어요.(BS26)
다. 이 고민을 해결하려면 한국어를* 공부를 열심히 하고 ...(AS22)
라. 그 둘이는* 알지 못했다라면* 이렇게* 일은 없을 것 같아요*.(AS23)

'인정·양보'의 의미를 나타내는 문법 항목은 '-아도/어도'와 '-ㄴ데' 두 가지가 사용되고 있었는데 그중 '-아도/어도'는 시간이 지날수록 그 빈도가 증가하고 있었지만 '-ㄴ데'는 '-아도/어도'와 마찬가지로 1~6개월부터 사용되었으나 이후의 7~12개월과 13~18개월에는 출현하지 않았다. '인정·양보'의 의미를 표현할 수 있는 문법 항목 중 한 번도 나타나지 않은 항목으로는 '-더라도', '-라도', '-아야/어야' 등이 있었다. 그러나 앞에서도 언급하였듯이 '인정·양보'의 의미는 모어화자 역시 거의 '-아도/어도' 위주로 사용하고 있다는 점에서 자연스러운 현상으로 볼 수 있다. 학습자들의 사용 양상을 살펴보면 예문(35)와 같다.

(35) 가. 여름에 입은(√입는) 옷을 입어도 너무 더워요.(AS14)
나. 나는 친구 없어도 혼자 잘 살 수 있어요.(CS22)
다. 먹는데 많이 안 먹어요.(BS7)

'목적·의도'를 나타내는 문법 항목으로는 '-러'와 '-려고'가 사용되고 있었는데 두 항목 모두 사용 빈도가 높지 않았으며 1~6개월에는 사용되지 않다가 7~12개월에 처음 출현하였다. 출현 이후 '-러'는 그 빈도를 늘려가고 있는 반면 '-려고'는 7~12개월에 사용된 이후 13~18개월에는 나타나지 않았다. 문어와 비교해볼 때 '-러'의 출현 시기가 문어에 비해 늦었는데 문어에서는 1~6개월부터 꾸준히 사용되며 빈도를 늘려가는 것과 달리 구어에서는 7~12개월에 처음 출현하였다. '목적·의도'의 의미를 나타내는 문법 항목의 사용 양상을 살펴보면 예문(36)과 같다.

(36) 가. 그 여자는 그 친구 만나러 농구장에 왔어요.(CS26)
나. 그래서 우리가 선생님이* 찾으러 갔었어요.(BS24)
다. 선생님은 밥을 일찍 먹으려고 5분 전에 시험을 끝났어요*.(AS21)

‘배경·상황 설명’의 의미를 표현하는 데 사용된 문법 항목으로는 ‘-ㄴ데’와 ‘-니’ 두 가지가 있었다. 그중 ‘-니’는 13~18개월에 처음 사용되었고 ‘-ㄴ데’는 비교적 높은 빈도로 꾸준히 사용되고 있었으나 1~6개월에 가장 높은 빈도를 보인 뒤 7~12개월에 감소, 13~18개월에 다시 증가하였으나 여전히 1~6개월보다 낮은 빈도로 나타나고 있었다. 그리고 문어와 비교할 때 ‘-더니’는 한 차례도 나타나지 않았다. 그러나 ‘-더니’는 모어화자의 자료에서도 나타나지 않은 항목으로 주제의 영향으로 볼 수도 있을 것이다. ‘배경·상황 설명’과 관련된 문법 항목의 사용 양상은 예문(51)과 같다.

(37) 가. 맨날 잠을 자<u>는데</u> 나* 옆에서 자요.(BS4)
나. 우리 외할아버지는 40년 동안 의사 생활을 했<u>는데</u> 마지막에 병이* 때문에 돌아가셨어요.(AS30)
다. 싸우고 나<u>니까</u> 성적이 올라갔지만 저는 싸우는 방법이 좀 안 좋다고 생각했어요*.(AS27)

‘수단·방법’을 의미하는 문법 항목으로는 ‘-고’와 ‘-아(서)/어(서)’가 있다. 학습자들은 ‘-고’와 ‘-아(서)/어(서)’를 모두 사용하고 있었으나 ‘-고’는 7~12개월과 13~18개월에만 나타났고 ‘-아(서)/어(서)’는 1~6개월과 13~18개월에서만 사용되고 있었다. 그러나 두 항목 모두 13~18개월에는 비교적 높은 빈도를 보이며 학습자들이 시간이 지날수록 두 항목을 비교적 익숙하게 사용하고 있음을 보여주었다. ‘수단·방법’을 나타내는 문법 항목의 사용 양상을 살펴보면 다음의 예문(38)과 같다.

(38) 가. 이 귤을 다 먹은 씨를 심으면 내년에 집에서 자기의 돈을 쓰지 않<u>고</u>...(AS32)
나. 유명한 흑인 피아니스트가 백인 운전사를* 차 타<u>고</u> 피아노 공연하러 가요.(BS35)
다. 부모님* 이 일을 들어서* 빨리 뛰<u>어서</u> 농구장에 왔어요.(CS27)
라. 그러니까 게임는* 시간이* 맞<u>춰서</u> 해야 돼요.(BS36)

‘방식·정도’를 표현하는 문법 항목으로는 ‘-게’와 ‘-이’ 두 가지가 사용되고 있었는데 ‘-게’는 7~12개월에 처음 나타난 뒤 13~18개월에 그 빈도가 증가하고 있었고 ‘-이’는 13~18개월에 들어서서 처음 사용되었다. 두 항목의 첫 출현 시기를 문어와 비교해볼 때 ‘-게’와 ‘-이’ 모두 구어에서 6개월 더 늦게 출현하였다는 것을 알 수 있었다. 이 외에 ‘방식·정도’를 의미하는 문법 항목 중 출현하지 않은 항목으로는 ‘-도록’이 있었는데 이는 문어와 구어 모두 동일한 부분이다. ‘방식·정도’와 관련된 문법 항목의 사용 양상

을 살펴보면 예문(39)와 같다.

(39) 가. 우리는 2년 넘게 못 만났어요.(AS18)
나. 우리 반에 다 모여서 노래도 부르고 소리 크게 이야기도 해요.(CS28)
다. 그냥 이유 없이 욕했어요.(BS30)

'사역'의 의미를 나타내는 '-게'는 문어와 동일하게 1~6개월에는 사용되지 않다가 7~12개월에 처음 출현하였고 이후 13~18개월에 그 사용 빈도가 약간 증가하는 양상을 보이고 있었다. '-게'의 사용 양상을 살펴보면 예문(40)과 같다.

(40) 가. 제가 친구 A를 좀 귀찮게 해요.(AS23)
나. 그들은 말이* 나쁘게 하고 나* 주변 환경도 나쁘게 해* 것 같아서 너무 싫어해요*.(CS31)

구어 자료에 나타난 부사절 관련 문법 항목의 사용 양상을 정리하면 다음과 같다. 먼저 부사절 관련 문법 항목의 발달과 관련해서 사용 빈도가 가장 높은 항목은 '조건·가정'의 '-면', '계기적 사건'의 '-고(서)', '이유·원인'의 '-아(서)/어(서)', 계기적 사건의 '-아(서)/어(서)'로 문어와 동일하였다. 그리고 세 기간 동안 꾸준히 빈도가 상승하는 항목은 '계기적 사건'의 '-고(서)'와 '-아(서)/어(서)', '이유·원인'의 '-아(서)/어(서)', '조건·가정'의 '-면'이었다. 반면 사용 빈도가 감소하는 문법들도 있었는데 7~12개월에 빈도가 감소하는 문법 항목은 '인정· 양보'의 '-ㄴ데', '배경·상황 설명'의 '-ㄴ데', '수단·방법'의 '-아(서)/-어(서)'이고 13~18개월에 빈도가 감소하는 문법 항목은 '조건·가정'의 '-려면', '-라면', '인정·양보'의 '-아도/어도', '목적·의도'의 '-려고' 였다.

다음으로, 발달 시점과 관련해서 살펴보면 전체 부사절 관련 문법 항목 중 1~6개월에 사용되지 않다가 7~12개월에 새로 나타나기 시작한 문법 항목으로는 '동시적 사건'의 '-며/-면서', '이유·원인'의 '-니/니까', '조건·가정'의 '-려면'과 '-라면', '목적·의도'의 '-러', '-려고', '수단·방법'의 '-고', '방식·정도'의 '-게', '사역'의 '-게'가 있었고 1~12개월까지 사용되지 않다가 13~18개월에 들어서서 새로 나타난 문법 항목에는 '조건·가정'의 '-아야/어야(지)', '배경·상황 설명'의 '-니', '방식·정도'의 '-이'가 있었다.

각 문법 항목의 문어와 구어의 첫 출현 시기를 비교해볼 때 구어에서 먼저 출현한 문법은 '인정·양보'의 '-ㄴ데'와 '수단·방법'의 '-아(서)/-어(서)'였고 문어에서 먼저 출현한 문법은 '동시적 사건'의 '-며/-면서', '조건·가정'의 '아야/어야(지)', '목적·의도'의 '-러', '수단·방법'의 '-고', '방식·정도'의 '-게'와 '-이'였다. 이 외에 구어에서 사용되었으

나 문어에서 사용되지 않은 문법 항목은 '조건·가정'의 '-라면', '배경·상황 설명'의 '-니'였고 문어에서 사용되었으나 구어에서 사용되지 않은 문법 항목은 '사건의 전환'의 '-다(가)', '이유·원인'의 '-라(서)', '-다가', '조건·가정'의 '-다면', '배경·상황 설명'의 '-더니'였다.

마지막으로 인용절 관련 문법 항목의 사용 양상을 살펴보도록 한다. 문어 자료에 나타난 인용절 관련 문법 항목의 총 빈도와 전체 절에서 차지하는 비율을 살펴보면 1~6개월에는 5회(2.1%), 7~12개월에는 30회(5.0%), 13~18개월에는 44회(4.4%)로 7~12개월에는 1~6개월보다 사용 빈도와 비율이 큰 폭으로 증가하였으나 13~18개월에는 사용 빈도 측면에서는 증가하였으나 비율은 약간 하락하는 양상을 보이고 있었다. 인용절 관련 문법 항목의 기간별 사용 양상을 세부적으로 살펴보면 다음의 <표 42>와 같다.

<표 42> 인용절 관련 문법 항목의 기간별 사용 양상

구분		문법 항목	1~6개월	7~12개월	13~18개월
인용절	직접 인용	-(이)라고	0.34	0.48	1.20
		-하고	0.17	-	0.11
	간접 인용	-다고/-(이)라고	0.34	2.90	3.39
		-냐고	-	0.12	0.11
		-(으)라고	-	0.12	-

인용절은 '직접 인용절'과 '간접 인용절'로 나눌 수 있다. 표에서 살펴보면 1~6개월에는 '직접 인용'을 더 많이 사용하였으나 7~12개월과 13~18개월에는 '간접 인용'을 훨씬 더 많이 사용함을 알 수 있다. 문법 항목별로 살펴보면 학습자들은 총 5가지의 문법 항목을 사용하여 인용절을 산출하고 있었는데 '직접 인용'에서 많이 사용된 문법 항목은 '-(이)라고'이고 '간접 인용'에서 가장 많이 사용된 문법 항목은 '-다고/-(이)라고'였다. 두 항목의 빈도를 비교해 보면 1~6개월에는 같은 빈도로 사용되었으나 이후 '-다고/-(이)라고'의 빈도가 빠르게 증가하며 '직접 인용'의 '-(이)라고'에 비해 훨씬 더 높은 빈도로 사용되고 있음을 확인할 수 있었다. 그리고 '직접 인용'의 '-(이)라고' 역시 시간이 흐름에 따라 꾸준히 빈도를 높여 가고 있었다. 이 외에 '-하고'와 '-냐고', '-(으)라고'는 사용 빈도가 매우 저조하여 학습자들이 주로 '-(이)라고'와 '-다고/-(이)라고'를 사용하여 인용절을 형성하고 있음을 알 수 있었다. 학습자들이 구어에서 산출한 인용절을 살펴보면 예문(41)과 같다.

(41) 가. 선생님가* '다 먹어야 되요*'라고 했어요.(BS7)
나. 나는 '알겠습니다'하고 그냥 교실에 갔어요.(CS3)
다. 선생님께서 무슨 작겨를* 있으냐고* 물어봤어요.(AS24).
라. 이 사건이 끝난 후에 학교에서* 반 달 동안 집에서 반성하라고 말했어요.(AS24)

구어 자료에 나타난 대등절과 내포절의 유형을 정리해 보면 학습자들은 대등절에서 5가지, 명사절에서 5가지, 관형절에서 3가지, 부사절에서 20가지, 인용절에서 5가지의 문법을 사용하여 다양한 절들을 구성하며 문장을 확장해 나가고 있었다. 기간별로 살펴보면 가장 안정적으로 발달하는 것은 부사절 관련 문법 항목이고 지속적인 발달에 어려움을 겪는 항목은 명사절 관련 문법 항목이었다. 구어 자료에서 명사절은 문법 항목의 종류 측면에서는 약간의 발달을 보였으나 빈도 측면에서는 13~18개월에 퇴행이 나타나고 있었는데 문어와 마찬가지로 사용 빈도가 가장 높은 '-기'와 '-ㄴ/는 것'의 사용이 감소하고 있었다. 관형사형 전성어미에서는 문어와 마찬가지로 '-ㄴ/는'을 이용한 관형사절 구성이 집중적으로 나타났으며 과거형 관형사형 전성어미 '-던'은 비교적 늦게 나타나며 사용 빈도 또한 매우 저조하였다. 학습자들은 구어에서도 전반적으로 시간이 지남에 따라 대부분 문법의 빈도가 증가하고 더 다양한 문법 항목을 사용하여 문장을 확장하려는 시도를 보이고 있었다. 그러나 동일한 문법 항목이라고 할지라도 문어보다 구어에서의 출현 빈도가 더 낮은 경우가 많았으며 문어에 비해 구어에서 6개월 정도 더 늦게 출현하는 문법들이 여럿 관찰되기도 하였다. 이러한 현상은 문어에 비해 구어에서 문법 사용이 약간 지체되는 것으로 해석할 수 있으며 문어와 구어에서 새로운 문법이 나타나는 순서를 비교한 뒤 구어에서 새로운 문법의 사용이 더 늦게 나타난다고 주장하는 Weissberg(2000)의 연구 결과와 일치한다. 그러나 본고의 문어와 구어 산출의 주제가 평균 4주 정도 간격을 두고 선후 순서가 바뀐다는 점에서 주제의 영향을 받았을 가능성을 배제할 수는 없을 것이다.

1.3. 산출 양식에 따른 통사 복잡성 비교

앞에서 문어와 구어의 통사 복잡성 발달을 각각 단위 길이, 단위 확장의 양과 유형으로 나누어 살펴보았다. 본 절에서는 학습자들의 통사 복잡성이 산출 양식에 따라 어떠한 차이가 있는지 알아보고자 한다. 앞절에서 확인하였듯이 문어와 구어는 기간별로 발달 수준이나 속도에 차이가 있다. 따라서 전체 기간에 나타난 문어와 구어의 차이 외에도 6개월을 간격으로 각 기간별 문어와 구어의 산출을 비교해 보도록 한다.

1.3.1. 단위 길이의 비교

산출 양식에 따른 단위 길이의 차이를 알아보기 위해 학습자들의 문장당 형태소 수와 AS-unit당 형태소 수를 비교하였다. 전체 기간에 나타난 문어와 구어의 단위 길이 평균을 비교해 보면 문장당 형태소 수의 평균은 12.28이고 A-unit당 형태소 수의 평균은 11.79로 학습자들은 문어에서 더 긴 문장을 산출하고 있었다. 그러나 기간별로 나누어 살펴보면 평균과 다른 양상을 보이는데 1~6개월과 7~12개월에는 문장 길이가 더 길지만 13~18개월에는 AS-unit 길이가 미세하게 더 길어지는 것을 확인할 수 있다. 즉 구어의 단위 길이는 문어에 비해 낮은 시작을 보이지만 더 빠른 속도로 증가하여 13~18개월에 들어서는 구어의 단위 길이가 더 길어지는 것이다.

문어와 구어의 단위 길이에 나타난 차이가 통계적으로 유의한지 확인하기 위해 t검정을 실시한 결과 <표 43>에서와 같이 전체 기간(t=2.079, p=.040)에서만 유의한 차이가 있는 것으로 나타났다. 이는 기간별 문어와 구어의 단위 길이 차이는 약간의 변화를 보이지만 그 차이가 아직은 통계적으로 유의미할 정도가 아님을 의미한다. 그러나 구어의 단위 길이 증가 속도를 고려해 볼 때 관찰 기간이 더 길어진다면 AS-unit길이가 문장 길이보다 유의미하게 길어질 수도 있음을 유추할 수 있다.

<표 43> 기간별 문어와 구어의 단위 길이

측정 지표	기간	N	문어		구어		t/p
			Mean	SD	Mean	SD	
단위당 형태소 수	1~6개월	36	10.19	2.13	9.35	1.68	1.967/.057
	7~12개월	36	12.38	2.20	11.66	1.97	1.838/.075
	13~18개월	36	14.28	2.66	14.36	2.36	-.194/.847
	전체 기간	108	12.28	2.87	11.79	2.87	2.079/.040*

*p<.05 **p<.01 ***p<.001

문어와 구어의 단위 길이의 차이를 학습자별로 나누어 살펴보면 다음의 <표 44>와 같다. 표에서 확인할 수 있듯이 세 학습자 중 B학습자(t=3.024, p=.005)만이 산출 양식에 따른 단위 길이의 차이가 유의한 것으로 나타났다. 이는 A학습자와 C학습자가 문어와 구어에서 비슷한 길이의 문장을 산출한다는 것을 의미한다. 또한 A학습자는 AS-unit 길이(M=12.59)가 문장 길이(M=12.20)에 비해 높게 나타나기도 하였다.

<표 44> 학습자별 문어와 구어의 단위 길이

측정 지표	기술통계량				t/p	r/p
	학습자	N	Mean	SD		
단위 길이	AW	36	12.20	2.84	-1.077/.289	.750/.000***
	AS	36	12.59	3.23		
	BW	36	13.45	3.06	3.024/.005**	.673/.000***
	BS	36	12.26	2.70		
	CW	36	11.20	2.27	1.553/.130	.296/.080
	CS	36	10.51	2.23		

*p<.05 **p<.01 ***p<.001

1.3.2. 단위 확장의 비교

1.3.2.1. 단위당 절 수

산출 양식에 따른 문장 확장의 정도를 확인하기 위해 학습자들의 문장당 절 수와 AS-unit당 절 수를 비교하였다. 전체 기간에 나타난 문어와 구어의 단위당 절 수의 평균을 비교해 보면 문장당 절 수의 평균은 1.97이고 AS-unit당 절 수의 평균은 1.80으로 학습자들은 문어에서 단위당 더 많은 절을 사용하여 문장을 확장하고 있었다. 6개월을 단위로 나누어 살펴보아도 동일한 결과를 확인할 수 있는데 학습자들은 세 기간 모두 문어에서 더 많은 절을 산출하고 있었는데 이는 학습자들이 구어보다 문어에서 구조적으로 더 복잡한 문장을 산출할 수 있음을 의미한다.

문어와 구어의 단위당 절 수에 나타난 차이가 통계적으로 유의한지 확인하기 위해 t 검정을 실시한 결과 <표 45>에서와 같이 1~6개월(t=2.994, p=.005)과 7~12개월(t=2.859, p=.007), 그리고 전체 기간(t=4.023, p=.000) 내에서 유의한 차이가 있는 것으로 나타났다. 한편, 13~18개월(t=1.191, p=.242)에는 유의한 차이를 보이지 않았는데 이는 시간이 지날수록 문어와 구어의 차이가 좁혀짐을 의미하기도 한다.

<표 45> 기간별 문어와 구어의 단위당 절 수

측정 지표	기간	N	문어		구어		t/p
			Mean	SD	Mean	SD	
단위당 절 수	1~6개월	36	1.68	0.39	1.46	0.29	2.994/.005**
	7~12개월	36	1.97	0.41	1.76	0.30	2.859/.007**
	13~18개월	36	2.27	0.51	2.18	0.42	1.191/.242
	전체 기간	108	1.97	0.50	1.80	0.45	4.023/.000***

*p<.05 **p<.01 ***p<.001

이러한 현상에 대해 Kuiken & Vedder(2012)에서는 구어의 실시간 산출이라는 특징으로 인해 '압박감'이 더 크기 때문에 구어에서 더 단순한 구조를 산출하게 되며 이러한 결과가 구어의 통사 복잡성이 감소하는 결과로 이어진다고 설명하고 있는데 전체 기간을 바탕으로 하는 문어와 구어의 통사 복잡성의 비교에 대한 본고의 결과는 이러한 주장을 뒷받침한다고 볼 수 있다. 그러나 단위 길이에서는 13~18개월에 들어선 뒤 구어에서 더 높아지는 현상이 나타났고 단위당 절 수는 13~18개월에 들어선 뒤 문어와 구어의 차이가 통계적으로 유의미하지 않을 정도로 좁혀졌다는 점은 구어의 통사 복잡성이 항상 더 낮지는 않으며 이는 학습자의 숙달도나 환경적 특징에 따라서 변화할 수 있음을 드러낸다.

산출 양식에 따른 단위당 절 수의 차이를 학습자별로 나누어 살펴보면 다음의 <표 46>과 같다. 표에서 확인할 수 있듯이 산출 양식에 따라 통계적으로 유의한 차이를 보이는 학습자는 B학습자(t=4.883, p=.000)와 C학습자(t=2.494, p=.018)이며 두 학습자 모두 문장당 절 수가 AS-unit당 절 수보다 높아 문어에서 단위당 더 많은 절을 사용하고 있는 것으로 나타났다. 반면 A학습자는 문장당 절 수와 AS-unit당 절 수가 같은 것으로 나타나 문어와 구어의 절 사용에 유의한 차이가 없음을 알 수 있었다. 앞서 살펴본 단위 길이의 차이와 연결시켜 볼 때 세 학습자 중 A학습자는 문어와 구어를 비교적 균형적으로 발달시키고 있고 B학습자는 문어와 구어의 통사 복잡성의 간극이 매우 큰 편이며 C학습자는 길이에 있어서는 문어와 구어의 차이가 뚜렷하지 않지만 구조적 복잡성은 구어에서 뚜렷하게 낮다는 것을 알 수 있었다.

<표 46> 학습자별 문어와 구어의 단위당 절 수

측정 지표	기술통계량				t/p	r/p
	학습자	N	Mean	SD		
단위당 절 수	AW	36	1.88	0.46	.086/.932	.558/.000***
	AS	36	1.88	0.52		
	BW	36	2.25	0.53	4.883/.000***	.628/.000***
	BS	36	1.90	0.46		
	CW	36	1.80	0.37	2.494/.018*	.261/.124
	CS	36	1.63	0.28		

*p<.05 **p<.01 ***p<.001

1.3.3. 대등절과 내포절의 유형 비교

1.3.3.1. 단위당 대등절과 내포절 수에 나타난 차이

문어와 구어에 나타난 단위당 절의 사용을 대등절과 내포절로 나누어 비교해 보았다. 우선 전체 기간에 나타난 문어와 구어의 단위당 대등절 수와 내포절 수를 비교해 보면 문장당 대등절 수의 평균은 0.22개, AS-unit당 대등절 수의 평균은 0.18개이고 문장당 내포절 수의 평균은 0.74개, AS-unit당 내포절 수의 평균은 0.60개로 대등절과 내포절 모두 문어에서 더 높은 빈도로 사용되고 있는 것으로 나타났다. 이를 다시 6개월을 단위로 묶어서 비교해 보면 전체 기간과 마찬가지로 세 기간 모두 문어에서 더 많은 대등절과 내포절을 사용하고 있는 것으로 나타났다.

문어와 구어의 단위당 대등절 수와 내포절 수에 나타난 차이가 통계적으로 유의한지 확인하기 위해 t검정을 실시한 결과 <표 47>에서와 같이 대등절은 전체 기간(t=2.435, p=.017)에서만 유의한 차이를 보였고, 내포절은 단위당 절 수와 마찬가지로 1~6개월(t=3.136, p=.003), 7~12개월(t=2.871, p=.007)과 전체 기간(t=4.029, p=.000)에서만 차이를 보여 시간이 지날수록 문어와 구어의 내포절의 차이가 줄어들고 있음을 알 수 있었다. 단위 길이와 단위당 절 수, 단위당 내포절 수의 이러한 결과는 학습자들이 초반에는 문어에서 더 높은 통사 복잡성을 보이지만 시간이 지날수록 구어의 복잡성이 더 빨리 발달하면서 문어와 구어의 통사 복잡성의 차이가 좁혀지고 있음을 알 수 있는 부분이다.

<표 47> 기간별 문어와 구어의 단위당 대등절 수와 내포절 수

측정 지표	기간	N	문어		구어		t/p
			Mean	SD	Mean	SD	
단위당 대등절 수	1~6개월	36	0.12	0.14	0.09	0.09	1.506/.141
	7~12개월	36	0.22	0.16	0.16	0.12	1.780/.084
	13~18개월	36	0.31	0.19	0.28	0.14	.975/.336
	전체 기간	108	0.22	0.18	0.18	0.14	2.435/.017*
단위당 내포절 수	1~6개월	36	0.55	0.32	0.36	0.25	3.136/.003**
	7~12개월	36	0.74	0.29	0.58	0.26	2.871/.007**
	13~18개월	36	0.94	0.38	0.87	0.33	1.095/.281
	전체 기간	108	0.74	0.36	0.60	0.35	4.029/.000***

*p<.05 **p<.01 ***p<.001

산출 양식에 따른 단위당 대등절 수와 내포절 수의 차이를 학습자별로 나누어 살펴보면 다음의 <표 48>과 같다. 대등절의 경우 A학습자(t=-.652, p=.519)는 문어보다 구어에서 더 많은 대등절을 사용하고 있었으나 그 차이가 통계적으로 유의하지 않았고 B

학습자(t=2.622, p=.013)와 C학습자(t=2.265, p=.030)는 구어보다 문어에서 더 많은 대등절을 사용하였으며 그 차이가 유의한 것으로 나타났다. 그리고 내포절의 경우, 세 학습자 모두 문어에서 구어보다 더 많은 내포절을 사용하였으나 A학습자(t=.575, p=.000)와 B학습자(t=.531, p=.001)는 그 차이가 유의하였고 C학습자(t=.229, p=.180는 문어와 구어의 내포절의 사용의 차이가 유의하지 않은 것으로 나타났다. 즉 B, C학습자는 문어에서 더 많은 대등절을 사용하고 있었고 A, B학습자는 문어에서 더 많은 내포절을 사용하고 있었으며 A학습자는 대등절에서, C학습자는 내포절에서 산출 양식 간의 차이가 유의하지 않았다.

<표 48> 학습자별 문어와 구어의 단위당 대등절과 내포절 수

측정 지표	기술통계량				t/p	r/p
	학습자	N	Mean	SD		
단위당 대등절 수	AW	36	0.15	0.15	-.652/.519	.358/.032*
	AS	36	0.17	0.16		
	BW	36	0.31	0.19	2.622/.013*	.426/.010**
	BS	36	0.23	0.14		
	CW	36	0.19	0.16	2.265/.030*	.229/.180
	CS	36	0.13	0.11		
단위당 내포절 수	AW	36	0.72	0.37	.575/.000***	.614/.543
	AS	36	0.68	0.42		
	BW	36	0.91	0.38	.531/.001***	4.501/.000***
	BS	36	0.64	0.36		
	CW	36	0.60	0.27	.229/.180	2.083/.045*
	CS	36	0.49	0.23		

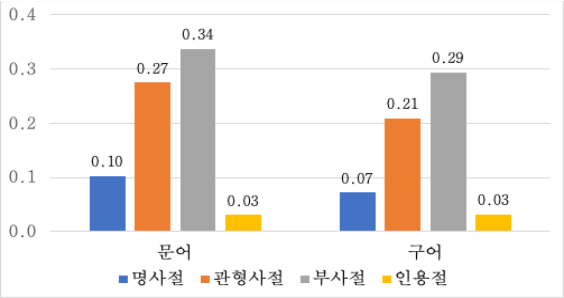
*p<.05 **p<.01 ***p<.001

1.3.3.2. 유형별 내포절 수에 나타난 차이

한국어의 내포절은 여러 유형이 있으며 각 내포절은 문장에서 다양한 기능을 한다. 문어와 구어에서 각 유형별 내포절의 사용에 차이가 있는지 확인하기 위해 산출 양식에 따른 각 유형별 내포절 수의 단위당 산출 빈도를 비교해 본 결과 <그림 13>과 같았다. 각 유형별 내포절의 사용 빈도 순서를 살펴보면 문어와 구어에서 동일하게 모두 '부사절>관형사절>명사절>인용절'의 순으로 나타나 학습자들이 문어와 구어에서 모두 부사절을 가장 많이 사용하고 인용절을 가장 적게 사용한다는 것을 알 수 있었다. 이는 학습자들은 산출 양식과 관계 없이 주로 부사절과 관형사절을 사용하여 문장을 확장하고

있음을 의미한다.

<그림 13> 문어와 구어의 단위당 유형별 내포절 수



전체 기간에 나타난 각 유형별 내포절의 평균을 비교해 보면 부사절, 관형사절, 명사절은 모두 문어에서 더 높게 나타나고 있었고 인용절은 문어와 구어에서 같은 빈도로 나타나 산출 양식에 따른 차이를 보이지 않았다. 6개월을 단위로 묶어 기간별로 살펴보면 13~18개월의 인용절을 제외하고 네 가지 내포절은 모든 기간에서 구어보다 문어에서 더 높은 빈도로 나타나고 있었다. 즉 학습자들은 문어에서 더 많은 명사절, 관형사절, 부사절을 사용하고 있는 것이다.

위에서 살펴본 문어와 구어에서의 유형별 내포절 수의 사용에 나타난 차이가 통계적으로 유의한지 알아보기 위해 t-검정을 실시한 결과 다음의 <표 49>와 같이 전체 기간에서는 명사절(t=2.412, p=.018), 관형사절(t=2.917, p=.004), 부사절(t=2.271, p=.025)은 산출 양식에 따라 유의한 차이가 있었으나 인용절(t=-.097, p=.923)은 문어와 구어 간에 유의한 차이를 보이지 않았고 기간별로 나누어 살펴보면 관형사절만이 1~6개월(t=2.678, p=.011)과, 7~12개월(t=2.156, p=.038)에 유의한 차이가 있었다. 그러나 관형사절 역시 13~18개월에는 문어와 구어 간의 차이가 좁혀지면서 통계적으로 유의하지 않은 상태로 전환되는 것을 확인할 수 있다. 즉 전체 기간 내에서는 인용절을 제외한 각 유형별 내포절이 문어에서 더 높은 빈도로 사용되지만 각 기간에서는 그 차이가 유의하지 않으며 초반에 유의한 차이가 있다고 하더라도 시간이 지나면서 그 차이가 좁혀진다는 것을 알 수 있다.

<표 49> 기간별 문어와 구어의 유형별 내포절 수

측정 지표	기간	N	문어		구어		t/p
			Mean	SD	Mean	SD	
단위당 명사절 수	1~6개월	36	0.08	0.10	0.05	0.08	1.791/.082
	7~12개월	36	0.13	0.12	0.08	0.10	1.678/.102
	13~18개월	36	0.10	0.08	0.09	0.09	.548/.587
	전체 기간	108	0.10	0.10	0.07	0.09	2.412/.018*
단위당 관형절 수	1~6개월	36	0.25	0.22	0.14	0.17	2.678/.011*
	7~12개월	36	0.25	0.12	0.19	0.13	2.156/.038*
	13~18개월	36	0.32	0.18	0.30	0.21	.522/.605
	전체 기간	108	0.27	0.18	0.21	0.18	2.917/.004**
단위당 부사절 수	1~6개월	36	0.21	0.14	0.17	0.16	1.155/.256
	7~12개월	36	0.32	0.14	0.27	0.14	1.574/.124
	13~18개월	36	0.48	0.21	0.43	0.16	1.222/.230
	전체 기간	108	0.34	0.20	0.30	0.19	2.271/.025*
단위당 인용절 수	1~6개월	36	0.01	0.02	0.01	0.02	.103/.918
	7~12개월	36	0.04	0.07	0.04	0.07	.439/.663
	13~18개월	36	0.04	0.07	0.05	0.07	-.703/.487
	전체 기간	108	0.03	0.06	0.03	0.06	-.097/.923

*p<.05 **p<.01 ***p<.001

이러한 각 유형별 내포절의 문어와 구어의 차이를 학습자별로 살펴보면 단위당 명사절은 B학습자(t=2.082, p=.045), 단위당 관형사절은 B학습자(t=3.163, p=.003)와 C학습자(t=2.507, p=.017), 단위당 부사절은 B학습자(t=3.084, p=.004)에서만 통계적으로 유의미한 차이를 보였고 단위당 인용절은 세 학습자 모두 문어와 구어 간에 차이가 없는 것으로 나타났다. 그리고 A학습자는 모든 유형의 내포절에 있어 산출 양식에 따른 유의미한 차이를 보이지 않았다.

정리하면, 전체 기간을 바탕으로 살펴본 학습자들의 단위 길이, 단위당 절 수, 단위당 대등절 수, 단위당 내포절 수, 그리고 인용절을 제외한 각 유형의 내포절은 모두 문어에서 더 높게 나타나며 이러한 차이는 통계적으로 유의미하였다. 그리고 이와 더불어 인용절의 사용은 산출 양식의 영향을 받지 않는다는 점을 확인할 수 있었다. 그러나 학습자 간에는 약간의 차이가 있었는데 B학습자는 인용절을 제외한 복잡성의 모든 측정 지표에서 문어가 유의미하게 높은 것으로 나타났고 C학습자 역시 모든 측정 지표에서 문어가 구어에 비해 높았으나 통계적으로 유의미하게 높은 것은 문장당 절 수 와 문장당 대등절 수, 문장당 관형사절 수였다. 반면 A학습자는 단위 길이는 구어에서 유의미하게 더 높았고 단위당 내포절 수는 문어에서 유의미하게 더 높았으나 각 유형별 내포

절은 문어와 구어 간의 차이가 유의하지 않았다. 또한 단위당 대등절 수와 관형사절 수, 인용절 수는 구어에서 더 높게 나타나기도 하였다. 이러한 결과는 B와 C학습자는 구어보다 문어에서 더 길고 복잡한 언어를 산출하는 반면, A학습자는 전체적으로 구어에서 좀 더 긴 문장을 산출하고 문어에서는 구조적으로 더 복잡한 문장을 산출하지만 이러한 복잡도가 절의 유형에 따라 다르게 나타난다고 해석할 수 있다.

1.3.4. 대등절과 내포절 관련 문법 항목 비교

이 장에서는 대등절과 내포절을 구성하는 다양한 문법 항목들의 사용 빈도와 비율을 세부적으로 살펴보아 문어와 구어의 문장 확장이 구체적으로 어떤 차이가 있는지 살펴보도록 한다. 비교를 위해 학습자들의 문어 자료와 구어 자료에 나타난 각 문법 항목의 표준화 빈도와 비율을 산출하였다. 먼저 문어와 구어 자료에 나타난 대등절 관련 문법 항목의 사용 양상을 비교해 보면 다음의 <표 50>과 같다.

<표 50> 산출 양식에 따른 대등절 관련 문법 항목 사용 양상

구분		문법 항목	문어 표준화 빈도		비율	구어 표준화 빈도		비율
대등절	나열 · 순접	-고	15.39	15.43	21.9%	14.01	14.05	22.6%
		-며/-면서	0.04			0.04		
	대조 · 역접	-ㄴ데	1.71	5.75		1.20	3.48	
		-지만	4.05			2.28		
	선택	-거나	0.34	0.34		0.26	0.26	

전체 절 중에서 대등절이 차지하는 비율을 비교해 보면 문에서는 21.9%, 구어에서는 22.6%를 차지하여 구어에서의 대등절 비율이 문어에 비해 높은 것을 알 수 있었다. 그러나 의미 범주별 단위당 빈도를 살펴보면 문어의 빈도가 구어에 비해 높은 것을 알 수 있다. 이는 문어의 단위당 절 수가 더 높은 것에 기인한 것으로 대등절의 사용 빈도는 문어에서 더 높지만 비율은 구어에서 더 높게 나타난 것이다. 구체적인 문법 항목을 비교해 보면, 거의 모든 문법 항목이 문어에서 더 빈번하게 사용되고 있었다. 그러나 앞서 시간의 흐름에 따른 대등절과 내포절의 발달에서도 확인하였듯이 시간이 지날수록 대등절보다는 내포절이 더 빠르게 증가한다는 점을 고려할 때, 구어의 대등절의 비율이 높은 것은 오히려 구어의 통사 복잡성이 낮음을 다시 한 번 확인할 수 있는 부분이기도 하다. 다음, 명사절 관련 문법 항목의 사용 양상을 비교해 보면 <표 51>과 같다.

<표 51> 산출 양식에 따른 명사절 관련 문법 항목 사용 양상

구분		문법 항목	문어 표준화 빈도		비율	구어 표준화 빈도		비율
명사절	명사형 전성어미	-ㅁ	0.04	2.73	10.0%	-	1.63	8.8%
		-기	2.69			1.63		
	관형사형 전성어미+것	-ㄴ/는 것	5.54	5.71		4.90	4.94	
		-ㄹ 것	0.13			0.04		
		-던 것	0.04			-		
	의문형 종결어미	-ㄴ지/는지	0.98	1.41		0.26	0.34	
		-ㄹ지	0.30			0.09		
		-ㄹ까	0.13			-		

전체 절 중 명사절의 비율을 살펴보면 문어에서 10.0%, 구어에서 8.8%로 문어에서 더 높은 비율의 명사절을 사용하는 것으로 나타났다. 의미 범주별 빈도를 살펴보면, 명사절 역시 모든 의미 범주에서 문어의 빈도가 구어보다 높게 나타나 문어에서 단위당 더 많은 명사절을 사용하고 있음을 알 수 있었다. 이는 각 문법 항목별 빈도에서도 동일하게 나타났는데 명사절 관련 문법 항목은 모두 구어보다 문어에서 더 높은 빈도로 사용되고 있었다. 뿐만 아니라 구어에서는 '-ㅁ', '-던 것', '-ㄹ까'가 한 차례도 사용되지 않아 문어에서 더 다양한 명사절을 더 높은 빈도로 사용하고 있음을 알 수 있었다. '-ㅁ'은 앞에서 언급하였듯이 구어에서 잘 사용되지 않는 문법임을 감안하더라도 '-던 것'과 같이 학습자들이 어려움을 겪는 문법이 문어에서는 사용되었으나 동일한 주제를 산출한 구어에서는 한 차례도 나타나지 않았다는 점을 고려해 보면 같은 문법 항목이라도 문어에서 어느 정도 숙달된 이후 구어에서 출현하게 되는 것으로 추측해 볼 수 있다. 계속하여 관형사절 관련 문법 항목의 사용 양상을 비교해 보도록 한다. 비교 결과는 <표 52>와 같다.

<표 52> 산출 양식에 따른 관형사절 관련 문법 항목 사용 양상

구분		문법 항목	문어 표준화 빈도		비율	구어 표준화 빈도		비율
관형절	관형사형 전성어미	-ㄴ/는	18.71	27.02	27.5%	14.14	19.81	25.0%
		-ㄹ	7.89			5.50		
		-던	0.43			0.17		

문어와 구어 나타난 관형사절의 비율을 살펴보면 문어에서는 27.5%, 구어에서 25.0%로 문어에서의 비율이 구어에 비해 높은 것을 확인할 수 있다. 또한 빈도 측면에서도 모든 문법 항목의 출현 빈도가 구어 자료보다 문어 자료에서 더 높게 나타나 학습자들은 구어보다 문어에서 더 빈번하고 더 높은 비율로 관형사절을 사용하고 있음을 알 수 있었다. 부사절 관련 문법 항목의 사용 양상을 비교해 보면 다음의 <표 53>과 같다.

<표 53> 산출 양식에 따른 부사절 관련 문법 항목 사용 양상

구분		문법 항목	문어 표준화 빈도		비율	구어 표준화 빈도		비율
부사절	동시적 사건	-며/-면서	1.02	1.02	36.8%	0.52	0.52	39.2%
	계기적 사건	-고(서)	8.35	12.45		6.57	10.74	
		-아(서)/어(서)	4.09			4.17		
	사건의 전환	-다(가)	0.17	0.17		-	0.00	
	이유 · 원인	-아(서)/어(서)	4.43	6.99		4.68	5.97	
		-니/니까	2.47			1.29		
		-다가	0.04			-		
		-라(서)	0.04			-		
	조건 · 가정	-면	9.80	10.06		9.97	10.14	
		-아야/어야(지)	0.13			0.04		
		-려면	0.09			0.09		
		-라면	-			0.04		
		-다면	0.04			-		
	인정 · 양보	-아도/어도	1.15	1.15		0.86	0.90	
		-ㄴ데	-			0.04		
	목적 · 의도	-러	0.90	1.11		0.34	0.39	
		-려고	0.21			0.04		
	배경 · 상황 설명	-ㄴ데	1.11	1.19		0.99	1.03	
		-니	-			0.04		
		-더니	0.09			-		
	수단 · 방법	-아(서)/-어(서)	0.09	0.60		0.30	0.77	
		-고	0.51			0.47		
	방식 · 정도	-게	0.72	0.85		0.26	0.39	
		-이	0.13			0.13		
	사역	-게	0.60	0.60		0.13	0.13	

문어 자료와 구어 자료의 전체 절 중에서 부사절이 차지하는 비율을 살펴보면 문어에서는 36.8%, 구어에서는 39.2%로 구어에서 문어보다 더 높은 비율로 부사절을 사용하고 있음을 알 수 있었다. 앞에서 살펴본 명사절과 관형사절의 사용 비율이 모두 구어에서 더 낮았던 점과 연결해 보면 학습자들이 구어에서 부사절은 다른 절에 비해 비교적 잘 사용할 수 있다는 것을 의미하기도 하지만 동시에 구어에서 부사절을 통한 문장 확장에 대한 의존도가 높음을 의미하기도 한다.

이어서 의미 범주별로 세분화하여 비교해 보면 다른 절과 마찬가지로 대부분의 의미 범주에서 문어의 사용 빈도가 더 높게 나타났는데 구어에서 더 높은 빈도로 나타난 의미 범주는 '조건·가정'과 '수단·방법'의 두 가지였다. 구체적인 문법 항목의 사용 빈도를 비교해 보면 구어에서 더 높은 빈도로 나타난 문법 항목으로는 '계기적 사건'의 '-아(서)/어(서)', '이유·원인'의 '-아(서)/어(서)', '조건·가정'의 '-면', '수단·방법'의 '-아(서)/-어(서)'였다. 즉 학습자들은 구어에서 '-아(서)/어(서)'의 형태를 더 빈번하게 사

용하고 있는 것이다. 이 외에 문어와 구어에서 사용되는 문법 항목에서도 차이가 있었는데 구어에만 나타난 문법 항목으로는 '조건·가정'의 '-라면', '인정· 양보'의 '-ㄴ데', '배경·상황 설명'의 '-니'가 있었고 문어에만 나타난 문법 항목으로는 '사건의 전환'의 '-다(가)', '이유·원인'의 '-다가', '-라(서)', '조건·가정'의 '-다면', '배경·상황 설명'의 '-더니'가 있었다. 마지막으로 인용절 관련 문법 항목의 문어와 구어 산출 양상을 비교해 보면 다음의 <표 54>와 같다.

<표 54> 산출 양식에 따른 인용절 관련 문법 항목 사용 양상

구분		문법 항목	문어 표준화 빈도		비율	구어 표준화 빈도		비율
인용절	직접 인용	-(이)라고	1.28	1.36	3.9%	0.73	0.82	4.3%
		-하고	0.09			0.09		
	간접 인용	-다고/-(이)라고	2.34	2.47		2.45	2.58	
		-냐고	0.09			0.09		
		-(으)라고	0.04			0.04		

인용절은 문어에서 3.9%, 구어에서 4.3%를 차지하며 구어에서 더 높은 비율로 나타나고 있었다. 또한 문어와 구어에서 모두 '직접 인용'보다 '간접 인용'의 문법 항목을 더 많이 사용하고 있었다. 그러나 인용절은 인용의 방식에 따라 문어와 구어의 빈도에 차이가 있었는데 직접 인용은 문어 자료에서 더 높은 빈도로 나타난 반면 간접 인용은 구어 자료에서 더 높은 빈도를 보이고 있었다. 즉 문어에서는 구어보다 더 많은 직접 인용을 사용하고 구어에서 문어보다 더 많은 간접 인용을 사용하고 있는 것이다.

정리하면, 전체적인 비율로 볼 때 대등절과 부사절, 인용절은 구어에서 더 높은 비율로 나타나고 명사절과 관형사절은 문어에서 더 높은 비율로 나타나고 있었다. 이는 문어에서 상대적으로 더 다양한 절을 사용하는 것과 달리 구어에서는 대등절과 부사절을 집중적으로 사용하고 있음을 의미한다. 여러 문법 항목 중에서[69] 구어에서 더 높은 빈도로 나타난 항목은 '조건·가정'의 '-면'과 '계기적 사건'의 '-아(서)/어(서)', '이유 원인'의 '-아(서)/어(서)', '수단 방법'의 '-아(서)/-어(서) 4가지였다.

한국어의 연결어미 중에는 구어에서 자주 사용되는 연결어미도 있고 문어에서 더 자주 사용되는 연결어미가 있으며 문어에서만 사용되는 연결어미도 있다. 대부분의 연결어미는 구어체와 문어체의 명확한 구분이 없는 경우도 있지만 모어화자의 언어 자료를 분석하여 비교한 연구들을 살펴보면 일부 구어성이 강한 연결어미와 문어성이 강한 연결어미들을 구분할 수 있다. 한국어 모어화자와 중국어권 학습자의 구어와 문어 산출

[69] 문어와 구어에서 모두 사용된 문법 항목들을 기준으로 비교한 결과이다.

자료를 비교한 이민경(2016)의 연구 결과에 따르면 '-아(서)/어(서)'는 모어화자의 언어 중에서도 구어에서 더 높은 비율로 사용되고 있다. 따라서 본고의 학습자들이 구어에서 '-아(서)/어(서)'를 더 높은 빈도로 사용하는 것은 한국어의 자연스러운 특징으로 볼 수 있다. 그러나 본고의 학습자의 자료에서 구어에서 더 높은 빈도로 나타난 '-면'은 이민경(2016)에서 문어에서 더 빈번하게 사용되는 항목인 것으로 나타났다. 이 외에도 구어성이 매우 강한 '-니까'의 경우 본고에서는 문어에서 구어의 2배 가까이의 빈도로 나타나기도 하는 것을 볼 때, 학습자들이 아직 사용역에 따라 연결어미의 쓰임이 다르다는 점에 대한 인식이 부족함을 알 수 있었다. 이 외에 이민경(2016)에서 구어에서 빈도가 더 높은 것으로 나타난 문법 항목에는 '-며/면서, -아야/어야(지), -려면, -려고' 등이 있다. 그러나 이러한 항목들은 본고에서 모두 구어에서 더 낮은 빈도로 사용되고 있는데 이러한 현상이 나타난 원인에는 전반적으로 학습자들이 구어에서 더 적은 내포절을 사용한다는 점과 문어에 비해 구어에서 다양한 연결어미를 사용하는 것에 익숙하지 않은 탓도 있었을 것이다.

2. 어휘 복잡성의 발달

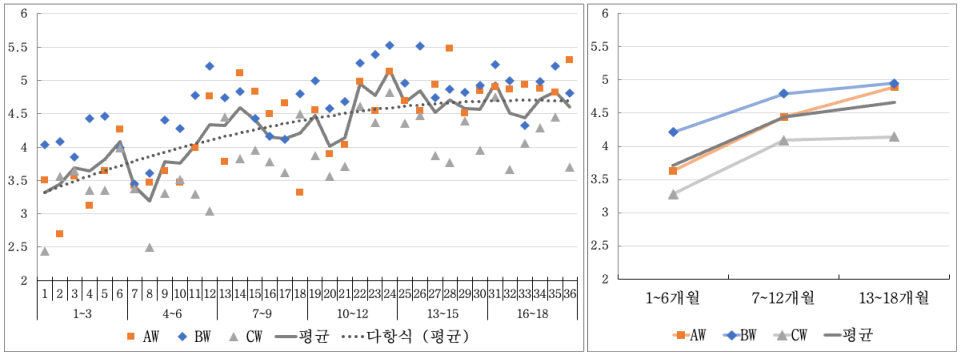
2.1. 문어의 어휘 복잡성 발달

2장에서 제시한 바와 같이 어휘 복잡성의 발달은 어휘 다양도, 어휘 밀도, 어휘 정교도 세 가지 지표를 통해 살펴보도록 한다. 세 명의 학습자는 1년 6개월 간 36회에 걸쳐 108개의 작문 자료에서 총 21,349개의 어휘를 산출하였으며 문장당 평균 어휘 수는 9.1개였다.

2.1.1. 어휘 다양도

어휘 다양도는 CTTR로 측정하였다. 먼저 각 시점별 학습자의 어휘 다양도와 6개월을 단위로 묶어 살펴본 어휘 다양도의 변화 추이는 다음의 <그림 14>와 같다. 세 학습자의 평균을 통해 학습자들의 어휘 다양도 변화 추이를 살펴보면 그래프는 전체적으로 상승의 추세를 띠고 있지만 이러한 상승은 시간이 흐를수록 완만해지면서 후반에는 점점 수평에 가까워지고 있었다. 6개월을 단위로 증가폭을 확인해 보면 7~12개월에는 1~6개월보다 0.74 증가한 것에 반해 13~18개월에는 7~12개월보다 0.22 증가하여 증가폭이 확연히 감소한 것을 확인할 수 있었다.

<그림 14> 시점별 문어 어휘 다양도와 6개월 단위 발달 추이



기간별 어휘 다양도의 발달이 통계적으로 유의한지 검증하기 위해 일원배치 분산분석(One-Way ANOVA)을 실시한 결과 <표 55>와 같이 F=29.521, p=.000으로 통계적으로 유의미한 차이가 있었다. 그러나 사후검정 결과, 1~6개월과 7~12개월, 13~18개월(1~6개월<7~12개월, 13~18개월) 간에만 유의한 차이를 보이고 있었는데 이는 어휘 다양도가 12개월이 지난 이후부터 뚜렷한 발달을 보이지 못하며 약간의 정체가 나타나고 있음을 보여준다.

<표 55> 기간별 문어 어휘 다양도

측정 지표	기술통계량				F/p	사후검정
	기간(개월)	N	Mean	SD		
CTTR	1~6	36(a)	3.71	0.61	29.521/.000***	a<b, c
	7~12	36(b)	4.44	0.56		
	13~18	36(c)	4.66	0.48		

*p<.05 **p<.01 ***p<.001

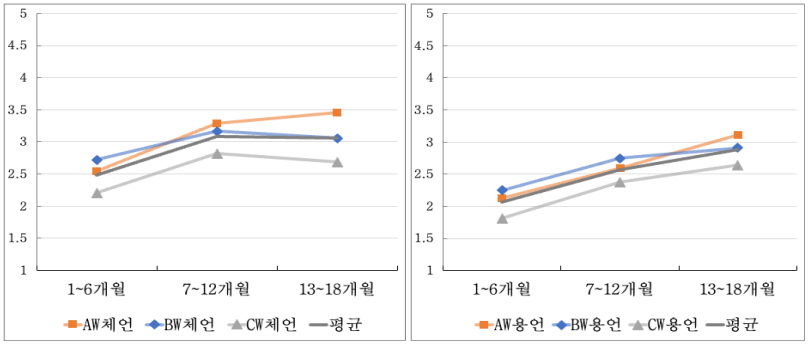
어휘 다양도는 어휘 유형 수(type)와 어휘 구현 수(token)로 산출된다. 따라서 어휘 유형 수와 어휘 구현 수의 기간별 변화를 확인해 보면 다음의 <표 56>과 같다. 표에서 확인할 수 있듯이 어휘 구현 수와 유형 수는 모두 시간이 지남에 따라 증가하지만 어휘 구현 수는 시간이 지날수록 증가폭이 커지는 것과 달리 어휘 유형 수는 구현 수와 반대로 시간이 지날수록 증가폭이 감소하고 있었다. 이는 학습자들이 시간이 지날수록 더 많은 어휘를 사용하지만 그에 상응하는 어휘 유형 수의 확장이 이루어지지 못하고 있으며 이러한 현상이 결국 어휘 다양도의 정체로 이어졌음을 알 수 있는 부분이다.

<표 56> 기간별 문어 어휘 유형 수와 어휘 구현 수

구분	1~6개월	증가폭	7~12개월	증가폭	13~18개월
type	55.1	+32.8	87.9	+23.4	111.3
token	111.3	+84.5	195.8	+90.3	286.0

어휘는 품사로 분류할 수 있으며 각 품사는 문장에서 서로 다른 기능을 한다. 한국어의 어휘는 기능에 따라 크게 체언, 용언, 수식언, 독립언, 관계언으로 나눌 수 있는데 체언(명사, 대명사, 수사)은 문장에서 주어, 목적어, 보어 등으로 쓰이고 용언(동사, 형용사)은 서술어로 쓰여 문장을 구성하는 데 중요한 역할을 하며 어휘의 상당 부분을 차지하고 있다.[70] 이에 본고에서는 어휘 중에서 체언과 용언을 추출하여 각각의 어휘 다양도의 변화를 살펴보았다. 체언과 용언의 어휘 다양도의 변화를 6개월을 단위로 살펴보면 다음의 <그림 15>와 같다. 그림에서 확인할 수 있듯이 용언의 어휘 다양도는 체언의 어휘 다양도에 비해 전반적으로 낮지만 꾸준히 증가하고 있는 반면 체언의 어휘 다양도는 7~12개월까지 증가한 뒤 13~18개월에는 정체 및 하락하는 양상이 나타나고 있었다. 6개월을 단위로 증가폭을 확인해 보면 용언은 7~12개월에는 1~6개월보다 0.51 증가하고 13~18개월에는 7~12개월보다 0.31증가한 것과 달리 체언은 7~12개월에는 1~6개월보다 0.60증가하였으나 13~18개월에는 7~12개월에 비해 오히려 0.02 하락하는 퇴행을 보이고 있었다.

<그림 15> 문어 자료에 나타난 체언과 용언의 6개월 단위 발달 추이



체언과 용언의 어휘 다양도의 기간별 변화가 통계적으로 유의한지를 검증하기 위해

[70] 어휘 구현 수를 바탕으로 계산할 때, 체언과 용언이 전체 어휘에서 차지하는 비율은 문어에서 61.2%, 구어에서 60.0%였다.

일원배치 분산분석을 실시한 결과 <표 57>과 같이 체언의 어휘 다양도(F=15.600, p=.000)와 용언의 어휘 다양도(F=43.038, p=.000)의 기간별 차이가 통계적으로 유의미하다는 것을 확인할 수 있었다. 그러나 사후검정 결과, 용언의 어휘 다양도는 세 기간 사이에 모두 유의미한 차이(1~6개월<7~12개월<13~18개월)를 보이며 발달하고 있었으나 체언의 어휘 다양도는 1~16개월과 13~18개월, 7~12개월(1~6개월<13~18개월, 7~12개월) 사이에만 유의한 차이를 보이고 있었다. 통계적 결과로도 확인할 수 있듯이 12개월이 지난 시점부터 어휘 다양도가 정체되는 현상은 용언의 증가세가 약간 둔화된 영향도 있지만 주로 체언 확장의 정체 및 퇴행이 큰 원인으로 작용하였음을 알 수 있다.

<표 57> 체언과 용언의 기간별 어휘 다양도

측정 지표	기술통계량				F/p	사후검정
	기간(개월)	N	Mean	SD		
체언 다양도	1~6	36(a)	2.49	0.52	15.600/.000***	a<c,b
	7~12	36(b)	3.09	0.55		
	13~18	36(c)	3.06	0.48		
용언 다양도	1~6	36(a)	2.06	0.40	43.038/.000***	a<b<c
	7~12	36(b)	2.57	0.38		
	13~18	36(c)	2.88	0.37		

*p<.05 **p<.01 ***p<.001

학습자별 어휘 다양도의 발달을 살펴보면 각 학습자의 어휘 다양도 수준이나 발달 경로는 약간의 차이가 있지만 세 학습자 모두 평균과 마찬가지로 13~18개월에 어휘 다양도의 증가가 둔화되고 있었다. 학습자 간 어휘 산출에 차이가 있는지 확인하기 위해 일원배치 분산분석을 실시한 결과 <표 58>과 같이 전체 어휘 다양도(F=20.044, p=.000), 체언의 어휘 다양도(F=9.420, p=.000), 용언의 어휘 다양도(F=6.158, p=.003) 모두 학습자 간에 유의한 차이가 있는 것으로 확인되었다. 사후검정 결과 세 지표 모두 A, B학습자의 어휘 다양도가 C학습자에 비해 유의하게 높고 A학습자와 B학습자 간의 차이는 유의하지 않은 것으로 나타났다.

<표 58> 학습자별 문어 어휘 다양도

측정 지표	기술통계량				F/p	사후검정
	학습자	N	Mean	SD		
CTTR	AW	36	4.32	0.70	20.044/.000***	BW,AW>CW
	BW	36	4.65	0.52		
	CW	36	3.84	0.57		
체언 CTTR	AW	36	3.09	0.65	9.420/.000***	AW,BW>CW
	BW	36	2.98	0.45		
	CW	36	2.57	0.50		
용언 CTTR	AW	36	2.61	0.53	6.158/.003**	BW,AW>CW
	BW	36	2.63	0.46		
	CW	36	2.27	0.46		

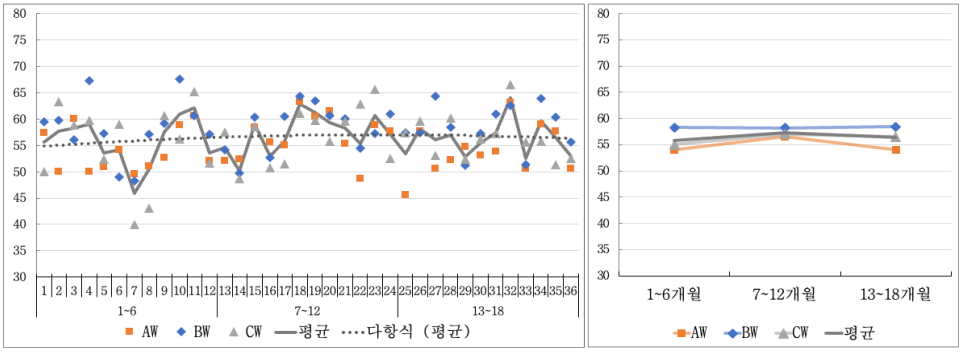
*p<.05 **p<.01 ***p<.001

정리하면, 학습자들의 문어 어휘 다양도는 시간의 흐름에 따라 지속적으로 증가하지만 증가 속도가 점점 느려 지면서 관찰 시작 12개월이 지난 시점부터 정체에 가까운 상태를 보이고 있었다. 이는 어휘 구현 수의 증가폭 확대와 어휘 유형 수의 증가폭 감소로 인한 것으로 확인되었으며 더 자세히 살펴보면 체언의 어휘 다양도의 퇴행이 매우 큰 원인으로 작용하는 것을 알 수 있었다. 또한 이러한 현상은 세 학습자 모두 비슷하게 나타나고 있었다.

2.1.2. 어휘 밀도

어휘 밀도는 텍스트에 나타난 전체 어휘 중 실질적 의미를 전달하는 내용어의 비율을 가리키는 것으로 학습자들이 얼마나 밀도 있게 내용을 전달하는지를 보여주는 척도이다. 본 연구에서는 일반명사, 동사, 형용사, 일반부사를 내용어로 보고 분석하였다. 먼저 각 시점별 학습자의 어휘 밀도와 6개월을 단위로 살펴본 어휘 밀도의 변화 추이는 다음의 <그림 16>과 같다. 세 학습자의 평균을 통해 어휘 밀도의 변화 추이를 살펴보면 그래프는 전체적으로 수평에 가까우며 미세하게 양쪽이 낮은 형태를 띠고 있다. 이는 어휘 밀도가 관찰 기간 내내 큰 발달이나 퇴행이 일어나지 나타나지 않았음을 의미한다. 6개월을 단위로 기간별 변화 추이를 살펴보면 세 학습자의 평균은 7~12개월이 가장 높고 1~6개월과 13~18개월이 낮은 것으로 나타났다. 기간별 어휘 밀도의 변화폭을 확인해 보면 7~12개월에는 1~6개월보다 1.55 증가하였으나 13~18개월에는 7~12개월보다 0.95 감소한 것으로 나타났다.

<그림 16> 시점별 문어 어휘 밀도와 6개월 단위 발달 추이



어휘 밀도는 내용어 구현 수와 총 어휘 구현 수로 계산된다. 총 어휘 구현 수와 내용어 구현 수의 기간별 차이를 살펴보면 총 어휘 구현 수는 7~12개월과 13~18개월에 각각 84.5개와 90.3개 증가하였고 내용어 구현 수는 7~12개월과 13~18개월에 각각 49.94개와 47.69개 증가하였다. 이러한 기간별 어휘 구현 수와 내용어 수, 어휘 밀도의 변화가 통계적으로 유의한지를 검증하기 위해 일원배치 분산분석(One-Way ANOVA)을 실시하였다. 그 결과 <표 59>와 같이 총 어휘 구현 수(F=73.794, p=.000)와 내용어 수(F=72.057, p=.000)는 모두 기간별로 유의한 차이를 보였으나 어휘 밀도는 F=.802, p=.451로 기간별로 유의한 차이가 없는 것으로 나타났다. 이를 통해 학습자들의 총 어휘 구현 수와 내용어 구현 수 모두 시간이 지남에 따라 유의미하게 증가하지만 그 증가 속도가 비슷하여 내용어의 비율이 높아지지는 않고 있음을 알 수 있었다.

<표 59> 기간별 문어 어휘 밀도

측정 지표	기술통계량				F/p	사후검정
	기간(개월)	N	Mean	SD		
총 어휘 구현 수	1~6	36(a)	111.3	49.0	73.794/.000***	a<b<c
	7~12	36(b)	195.8	71.5		
	13~18	36(c)	286.0	60.5		
내용어 구현 수	1~6	36(a)	62.44	29.3	72.057/.000***	a<b<c
	7~12	36(b)	112.4	42.1		
	13~18	36(c)	160.1	30.7		
어휘 밀도	1~6	36(a)	55.74	6.23	.802/.451	-
	7~12	36(b)	57.29	4.62		
	13~18	36(c)	56.34	4.67		

*p<.05 **p<.01 ***p<.001

학습자별로 어휘 밀도의 발달을 살펴보면 세 학습자의 어휘 밀도는 그 정도는 다르지만 모두 수평에 가깝거나 13~18개월에 감소하는 양상을 보이며 뚜렷한 변화를 보이지 않고 있었다. 세 학습자의 어휘 밀도에 차이가 있는지 확인하기 위해 일원배치 분산분석을 실시한 결과 <표 60>과 같이 세 학습자의 총 어휘 구현 수(F=6.455, p=.002), 내용어 수(F=6.081, p=.003)와 어휘 밀도(F=4.143, p=.019)는 모두 학습자 간에 유의한 차이가 있는 것으로 나타났다. 사후검정 결과 총 어휘 구현 수와 내용어 구현 수는 A, B학습자와 C학습자(A, B>C) 간에 유의한 차이가 있었고 어휘 밀도는 B학습자와 A학습자(B>A) 간에 유의한 차이가 있었다. 즉 A학습자는 내용어 수와 총 어휘 수 모두 세 학습자 중 가장 높지만 어휘 밀도는 세 학습자 중 가장 낮은 것으로 나타났다. 또한 문장 길이나 문장당 절 수, 어휘 다양도와 같은 복잡성 지표가 가장 낮았던 C학습자의 어휘 밀도가 오히려 A학습자에 비해 높게 나타난 점을 고려하면 한국어에서 어휘 밀도는 기타 복잡성 측정 지표와 관련성이 크지 않음을 유추할 수 있다.

<표 60> 학습자별 문어 어휘 밀도

측정 지표	기술통계량				F/p	사후검정
	학습자	N	Mean	SD		
총 어휘 구현 수	AW	36	228.9	103.3	6.455/.002**	AW, BW>CW
	BW	36	208.5	88.1		
	CW	36	155.6	74.4		
내용어 구현 수	AW	36	125.4	56.12	6.081/.003**	AW, BW>CW
	BW	36	121.6	51.09		
	CW	36	87.86	42.87		
어휘 밀도	AW	36	54.88	4.45	4.143/.019*	BW>AW
	BW	36	58.29	4.80		
	CW	36	56.19	5.84		

*p<.05 **p<.01 ***p<.001

숙달도별 어휘 밀도를 측정한 기존의 연구들 역시 본 연구와 비슷한 결과를 도출하였는데 Hwang(2012), 강정희(2020), 서신웨(2022)에서는 텍스트의 길이가 짧은 학습자나, 숙달도가 낮은 학습자의 어휘 밀도가 오히려 높게 나타났다고 보고하고 있다. 이러한 결과는 한국어에서는 어휘 밀도가 높은 것이 꼭 복잡성과 직결되지 않는 것일 수도 있음을 시사한다.

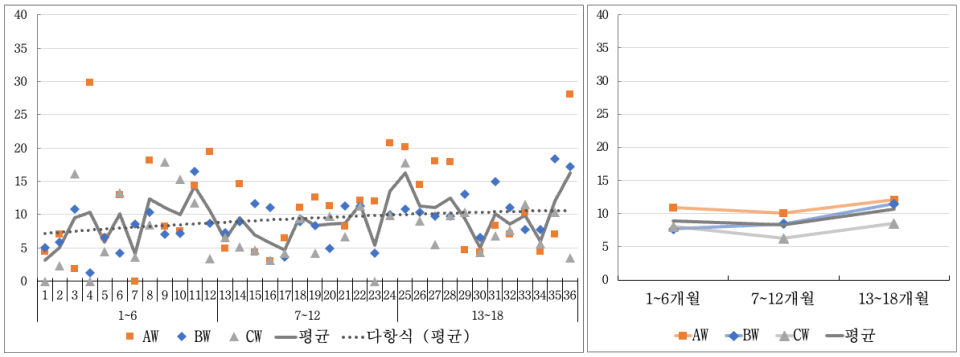
정리하면, 학습자들의 문어 어휘 밀도는 시간이 흐름에 따라 미세하게 증가하고 있지만 그 증가 속도가 매우 완만하며 어휘 밀도는 다른 복잡성 지표와 같은 결과를 보이

지 않았다. 그러나 이러한 결과는 텍스트 길이나 숙달도 외에도 학습자들의 조사 생략과도 관련이 있는 것으로 추측해볼 수 있는데 본고의 연구 대상 학습자들을 살펴볼 때, 세 학습자 중 A학습자는 조사를 거의 생략하지 않는 반면 B와 C학습자는 조사를 생략하는 오류가 매우 빈번하게 발생하고 있었다. 조사가 기능어에 포함된다는 점을 고려하면 조사 생략이 어휘 밀도에 영향을 줄 가능성도 생각해볼 수 있는 부분이다.

2.1.3. 어휘 정교도

어휘 정교도는 학습자가 사용한 어휘의 수준을 측정하는 지표로 텍스트에 나타난 어휘 중에서 비일상적이고 전문적인 어휘의 비율로 측정된다. 본고에서는 <한국어(KSL) 교육과정(2017)>의 <중학교 언어 재료>의 '일상생활 어휘 목록'을 바탕으로 초급 어휘에 포함되지 않는 어휘의 비율로 어휘 정교도를 측정하였다. 각 시점별 학습자의 어휘 정교도와 6개월을 단위로 살펴본 어휘 정교도의 변화 추이는 다음의 <그림 17>과 같다. 학습자들의 평균을 통해 어휘 정교도의 변화 추이를 살펴보면 어휘 정교도는 각 시점별로 지속적인 등락을 반복하지만 뚜렷한 상승 추세를 보이지는 않고 있었다. 6개월을 단위로 변화 추이를 확인하면 7~12개월에 오히려 정교도가 하락하는 것을 확인할 수 있는데 기간별 변화폭을 확인해 보면 7~12개월에는 1~6개월보다 0.59 감소하고 13~18개월에는 7~12개월보다 2.42 증가하여 하락 후 다시 증가하는 양상을 나타내고 있었다.

<그림 17> 시점별 문어 어휘 정교도와 6개월 단위 발달 추이



어휘 정교도의 기간별 변화가 통계적으로 유의한지 검증하기 위해 일원배치 분산분석(One-Way ANOVA)을 실시하였다. 그 결과 <표 61>과 같이 F=1.922, p=.151로 기

간별 차이가 유의하지 않은 것으로 나타났다. 이는 학습자들의 어휘 정교도가 관찰 기간(18개월) 동안 뚜렷한 변화가 일어나지 않았음을 의미한다.

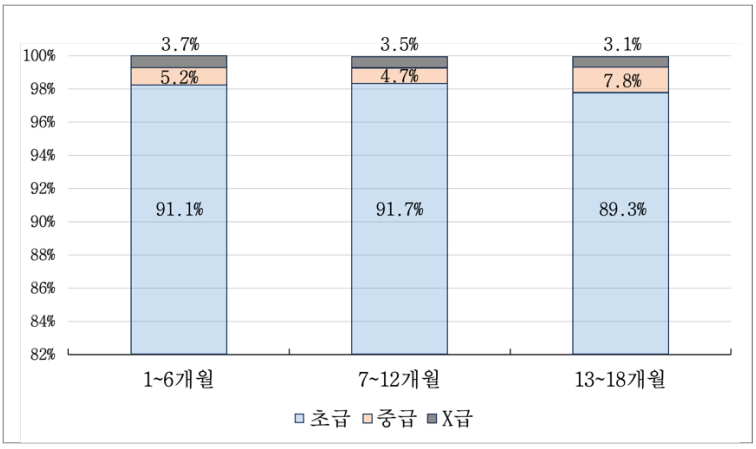
<표 61> 기간별 문어 어휘 정교도

측정 지표	기술통계량				F/p	사후검정
	기간(개월)	N	Mean	SD		
어휘 정교도	1~6	36(a)	8.86	6.56	1.922/.151	-
	7~12	36(b)	8.27	4.07		
	13~18	36(c)	10.69	5.46		

*p<.05 **p<.01 ***p<.001

6개월을 단위로 각 등급별 어휘의 사용 비율을 살펴보면 학습자들은 7~12개월에는 1~6개월보다 초급 어휘의 비율이 늘어나면서 퇴행을 보이다가 13~18개월에 들어서서 초급 어휘의 비율이 감소하고 중급 어휘의 비율이 높아지면서 약간의 발달을 보이고 있었다. 그러나 16~18개월에도 여전히 초급 어휘가 90%에 가까운 비율로 나타나고 있어 학습자들이 중·고급 어휘를 확장시키지 못하고 있음을 확인할 수 있었다.

<그림 18> 기간별 사용 어휘의 등급 분포



학습자별 어휘 정교도의 발달 추세를 살펴보면 B학습자는 미세하지만 꾸준한 증가를 보이는 반면 A, C학습자는 평균과 같은 추이를 나타내고 있었다. 학습자 간 어휘 정교도에 차이가 있는지 확인하기 위해 일원배치 분산분석을 실시한 결과 <표 62>와 같이 세 학습자의 어휘 정교도는 F=3.020, p=.056으로 학습자 간에 유의한 차이를 보이

지 않았다.

<표 62> 학습자별 문어 어휘 정교도

측정 지표	기술통계량				F/p	사후검정
	학습자	N	Mean	SD		
어휘 정교도	AW	36	11.02	7.03	3.020/.056	-
	BW	36	9.19	3.80		
	CW	36	7.61	4.78		

*p<.05 **p<.01 ***p<.001

어휘 다양도, 어휘 밀도, 어휘 정교도의 분석 결과들을 종합해 보면, 학습자들의 어휘 다양도는 1~16개월과 7~12개월 사이에는 뚜렷하게 발달하지만 12개월 이후에 발달이 정체되는 현상이 나타나고 어휘 밀도와 어휘 정교도는 유의미한 증가를 보이지 않았는데 이는 학습자들의 통사 복잡성의 각 측정 지표가 시간의 흐름에 따라 증가하는 것과 대조되는 결과이기도 하다.

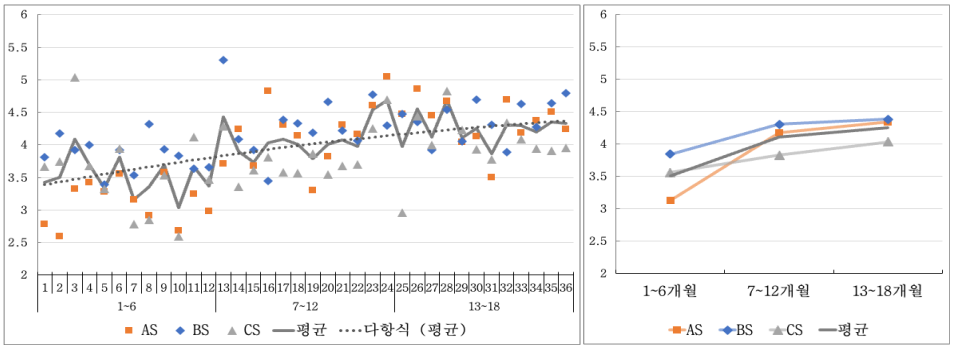
2.2. 구어의 어휘 복잡성 발달

구어 어휘 복잡성 역시 문어와 마찬가지로 어휘 다양도, 어휘 밀도, 어휘 정교도 세 가지 지표를 통해 살펴보았다. 세 명의 학습자는 1년 6개월 간 36회에 걸쳐 108건의 음성 자료에서 총 20,408개의 어휘를 산출하였으며 하나의 AS-unit당 평균 8.8개의 어휘를 산출하였다.

2.2.1. 어휘 다양도

각 시점별 어휘 다양도와 6개월 단위 평균 어휘 다양도의 변화 추이를 살펴보면 다음의 <그림 19>와 같다. 그림에서 확인할 수 있듯이 어휘 다양도의 평균 추세선은 후반으로 갈수록 상승세가 약간 둔화되는 곡선을 나타내고 있었다. 시점별 어휘 다양도 수치를 살펴보면 12개월(24차)까지는 큰 폭의 등락을 반복하지만 전체적으로 상승하는 추세인 반면, 12개월이 지난 시점부터 등락은 반복되지만 상승의 변화 추세를 보이지 않고 있었다. 이는 6개월 단위의 증가폭을 통해서도 확인이 가능한데 학습자들의 어휘 다양도는 7~12개월에는 1~6개월보다 0.59 증가한 것에 반해 13~18개월에는 7~12개월보다 0.15 증가하여 증가폭이 확연히 감소하고 있었다.

<그림 19> 시점별 구어 어휘 다양도와 6개월 단위 발달 추이



기간별 어휘 다양도의 증가가 통계적으로 유의한지를 검증하기 위해 일원배치 분산분석(One-Way ANOVA)을 실시한 결과 <표 63>에서와 같이 F=24.641, p=.000으로 기간별 차이가 통계적으로 유의미하였다. 그러나 사후검정 결과, 1~6개월과 7~12개월, 13~18개월(1~6개월<7~12개월, 13~18개월) 간에만 유의한 차이를 보여 위의 어휘 다양도 발달 그래프에서 보여지는 바와 같이 학습자들의 어휘 다양도가 12개월이 지난 시점부터 뚜렷한 발달이 이루어지지 못하며 약간의 정체가 나타나고 있음을 확인할 수 있었다.

<표 63> 기간별 구어 어휘 다양도

측정 지표	기술통계량				F/p	사후검정
	기간(개월)	N	Mean	SD		
CTTR	1~6	36(a)	3.51	0.53	24.641/.000***	a<b,c
	7~12	36(b)	4.10	0.49		
	13~18	36(c)	4.25	0.40		

*p<.05 **p<.01 ***p<.001

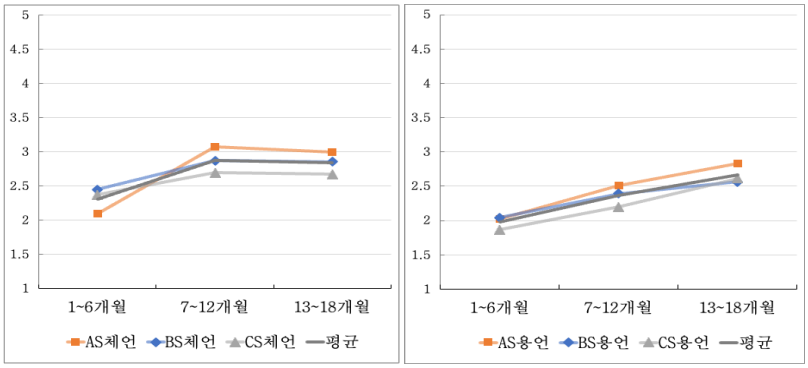
어휘 유형 수와 어휘 구현 수의 증가를 살펴보면 다음의 <표 64>와 같다. 표에서 확인할 수 있듯이 어휘 유형 수와 어휘 구현 수 모두 기간별로 증가하고 있으나 증가폭은 모두 7~12개월에 비해 13~18개월에 감소하고 있었다. 또한 어휘 구현 수의 증가폭에 비해 어휘 유형 수의 증가폭이 매우 적은 것을 알 수 있는데 이는 학습자들이 시간이 지날수록 더 많은 어휘를 사용하지만 그에 상응하는 정도의 어휘 유형을 늘려가지 못하고 있음을 의미한다.

<표 64> 기간별 구어 어휘 유형 수와 어휘 구현 수

구분	1~6개월	증가폭	7~12개월	증가폭	13~18개월
type	51.9	+28.2	80.1	+16.8	96.9
token	109.9	+82.5	192.4	+72.2	264.6

계속하여 체언과 용언의 어휘 다양도의 변화를 살펴보면 다음의 <그림 20>과 같다. 그림에서 확인할 수 있듯이 용언의 어휘 다양도는 꾸준히 증가하고 있는 반면 체언의 어휘 다양도는 7~12개월까지 증가한 뒤 13~18개월에는 미세하게 하락하는 퇴행을 보이고 있었다. 6개월을 단위로 증가폭을 확인해 보면 용언은 7~12개월에는 1~6개월보다 0.39, 13~18개월에는 7~12개월보다 0.30 증가한 것과 달리 체언은 7~12개월에는 1~6개월보다 0.57 증가하였으나 13~18개월에는 7~12개월에 비해 오히려 0.04 하락하는 퇴행을 보이고 있었다.

<그림 20> 구어 자료에 나타난 체언과 용언의 6개월 단위 발달 추이



이와 같은 체언과 용언의 어휘 다양도의 기간별 변화가 통계적으로 유의한지를 검증하기 위해 일원배치 분산분석을 실시하였다. 그 결과 <표 65>와 같이 체언의 어휘 다양도(F=16.791, p=.000)와 체언의 어휘 다양도(F=37.187, p=.000)의 기간별 차이가 통계적으로 유의미하다는 것을 확인할 수 있었다. 그러나 사후검정 결과, 용언의 어휘 다양도는 세 기간 사이에 모두 유의미한 차이(1~6개월<7~12개월<13~18개월)를 보이며 발달하고 있는 것과 달리 체언의 어휘 다양도는 1~16개월과 13~18개월, 7~12개월(1~6개월<13~18개월, 7~12개월) 사이에만 유의한 차이를 보이고 있었다. 이러한 결과는 문어와 동일한 것으로 12개월이 지난 시점부터 어휘 다양도가 정체되는 현상은 용언보다는 체언 다양도 확장의 정체 및 퇴행의 원인이 크다고 해석할 수 있다.

<표 65> 체언과 용언의 기간별 어휘 다양도

측정 지표	기술통계량				F/p	사후검정
	기간(개월)	N	Mean	SD		
체언 다양도	1~6	36(a)	2.31	0.51	16.791/.000***	a<c,b
	7~12	36(b)	2.88	0.48		
	13~18	36(c)	2.84	0.41		
용언 다양도	1~6	36(a)	1.97	0.30	37.187/.000***	a<b<c
	7~12	36(b)	2.37	0.42		
	13~18	36(c)	2.67	0.30		

*p<.05 **p<.01 ***p<.001

어휘 다양도의 발달 경로는 학습자 간에 큰 차이가 없이 모두 평균과 비슷하게 12개월까지 증가하다가 13~18개월에 전체 어휘 다양도의 증가 속도는 감소하고 체언의 다양도는 정체 혹은 퇴행, 용언의 다양도는 지속적으로 증가하고 있었다. 학습자 간 어휘 산출에 차이가 있는지 확인하기 위해 일원배치 분산분석을 실시한 결과 <표 66>과 같이 전체 어휘 다양도는 F=6.016, p=.004로 학습자 간 유의한 차이가 있는 것으로 나타났으나 체언과 용언의 어휘 다양도는 학습자 간의 차이가 유의하지 않은 것으로 확인되었다. 또한 전체 어휘 다양도 역시 B학습자와 C학습자 간에만 유의한 차이가 있었다. 이는 문어에서 전체 어휘 및 체언과 용언의 어휘 다양도 모두 학습자 간 차이가 유의하였던 것과 대조되는 결과로 문어보다 구어에서는 학습자 간의 어휘 다양도에 차이가 크지 않음을 의미한다.

<표 66> 학습자별 구어 어휘 다양도

측정 지표	기술통계량				F/p	사후검정
	학습자	N	Mean	SD		
CTTR	AS	36	3.88	0.67	6.016/.004**	BW>CW
	BS	36	4.18	0.42		
	CS	36	3.81	0.53		
체언 CTTR	AS	36	2.72	0.62	.881/.417	-
	BS	36	2.73	0.46		
	CS	36	2.58	0.51		
용언 CTTR	AS	36	2.45	0.46	2.484/.088	-
	BS	36	2.33	0.33		
	CS	36	2.22	0.50		

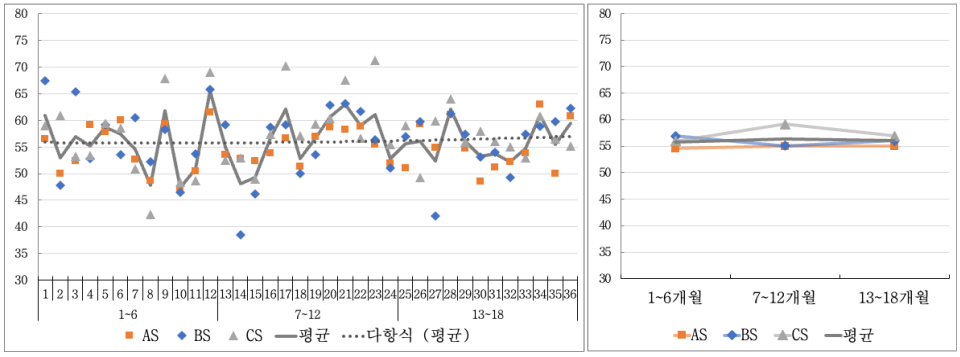
*p<.05 **p<.01 ***p<.001

정리하면, 구어의 어휘 다양도 역시 시간의 흐름에 따라 지속적으로 증가하지만 증가폭이 점점 감소하는 양상을 나타내고 있었으며 이는 어휘 유형 수의 증가폭 감소로 인한 것이었다. 그리고 체언과 용언의 어휘 다양도를 세부적으로 살펴본 결과 이는 체언의 어휘 다양도의 퇴행과 깊은 관련이 있을 것으로 예상할 수 있었다. 이러한 결과는 문어 자료 분석 결과와 동일한 것으로 학습자들의 어휘 다양도는 문어와 구어에서 모두 12개월이 지난 시점부터 발달이 정체되는 것으로 해석할 수 있다. 즉 중도입국 청소년들의 어휘 다양도는 기본적인 학교 생활만으로 발달할 수 있으나 이러한 발달에는 한계가 있음을 알 수 있다.

2.2.2. 어휘 밀도

구어 역시 문어와 마찬가지로 일반명사, 동사, 형용사, 일반부사를 내용어로 어휘 밀도를 분석하였다. 각 시점별 학습자의 어휘 밀도와 6개월을 단위로 살펴본 어휘 밀도의 변화 추이는 다음의 <그림 21>과 같다. 세 학습자의 평균을 통해 학습자들의 어휘 밀도 변화 추이를 살펴보면 거의 수평에 가까운 것을 확인할 수 있다. 6개월을 단위로 기간별 변화폭을 확인해 보면 7~12개월에 1~6개월보다 0.59증가하였다가 13~18개월에는 다시 0.43 감소하여 문어와 마찬가지로 미세한 상승 이후 다시 하락하며 거의 변화를 보이지 않고 있었다.

<그림 21> 시점별 구어 어휘 밀도와 6개월 단위 발달 추이



총 어휘 구현 수와 내용어 구현 수의 6개월 단위 변화 추이를 살펴보면 총 어휘 구현 수는 7~12개월과 13~18개월에 각각 82.5개와 72.2개 증가하였고 내용어 구현 수는 7~12개월과 13~18개월에 각각 48.31개와 38.47개 증가하였다. 이러한 기간별 어휘

구현 수와 내용어 수, 어휘 밀도의 변화가 통계적으로 유의한지를 검증하기 위해 일원배치 분산분석(One-Way ANOVA)을 실시하였다. 그 결과 <표 67>과 같이 총 어휘 구현 수(F=50.156, p=.000)와 내용어 구현 수(F=43.938, p=.000)는 기간별로 유의한 차이가 있었으나 어휘 밀도(F=.093, p=.911)는 기간별로 유의한 차이가 없는 것으로 나타났다. 즉 구어 역시 문어와 마찬가지로 총 어휘 구현 수와 내용어 구현 수는 모두 기간별로 유의하게 증가하지만 두 항목의 증가 속도가 비슷하여 내용어의 비율이 뚜렷하게 증가하는 현상은 나타나지 않고 있었다.

<표 67> 기간별 구어 어휘 밀도

측정 지표	기술통계량				F/p	사후검정
	기간(개월)	N	Mean	SD		
총 어휘 구현 수	1~6	36	109.9	53.1	50.156/.000***	a<b<c
	7~12	36	192.4	71.1		
	13~18	36	264.6	70.9		
내용어 구현 수	1~6	36	61.44	29.70	43.938/.000***	a<b<c
	7~12	36	109.75	44.30		
	13~18	36	148.22	42.46		
어휘 밀도	1~6	36	55.83	6.63	.093/.911	-
	7~12	36	56.42	6.31		
	13~18	36	55.99	4.80		

*p<.05 **p<.01 ***p<.001

학습자별로 어휘 밀도의 발달을 살펴보면 A, C학습자는 평균과 마찬가지로 7~12개월에 증가 이후 13~18개월에 감소하는 양상을 보이는 반면 C학습자는 기타 두 학습자와 반대로 7~12개월에 약간 감소하였다가 13~18개월에 다시 증가하였다. 그러나 세 학습자 모두 변화의 폭이 매우 작아 거의 변화를 보이지 않았다. 세 학습자의 어휘 밀도에 차이가 있는지 확인하기 위해 일원배치 분산분석을 실시한 결과 <표 68>과 같이 세 학습자의 총 어휘 구현 수(F=.395, p=.675), 내용어 구현 수(F=67.938, p=.518)와 어휘 밀도(F=1.537, p=.220)는 모두 학습자 간에 유의한 차이가 없는 것으로 나타났다. 이러한 결과는 문어에서 총 어휘 구현 수, 내용어 구현 수 및 어휘 밀도가 학습자 간에 유의한 차이를 보였던 것과 대조되는 것으로 문어에 비해 구어에서 학습자들 간의 어휘 밀도 수준이 더 비슷하다는 것을 의미한다.

<표 68> 학습자별 구어 어휘 밀도

측정 지표	기술통계량				F/p	사후검정
	학습자	N	Mean	SD		
총 어휘 구현 수	AS	36	187.39	92.21	.395/.675	-
	BS	36	180.28	73.08		
	CS	36	199.22	105.85		
내용어 구현 수	AS	36	103.22	52.21	67.938/.518	-
	BS	36	100.92	41.23		
	CS	36	115.28	63.16		
어휘 밀도	AS	36	54.91	4.21	1.537/.220	-
	BS	36	55.99	6.62		
	CS	36	57.34	6.51		

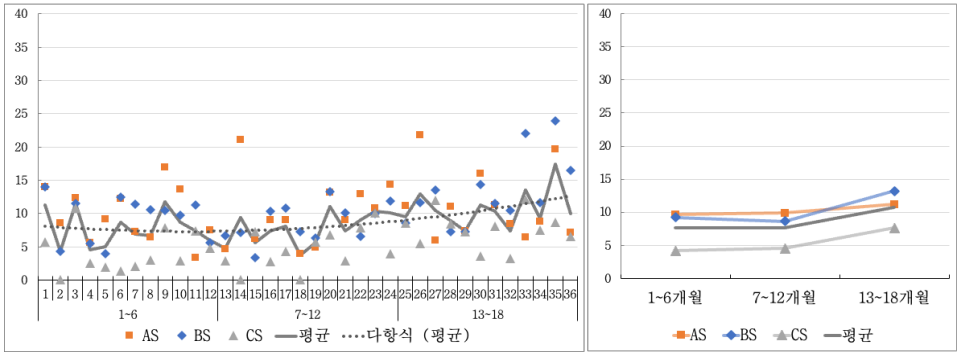
$^{*}p<.05$ $^{**}p<.01$ $^{***}p<.001$

정리하면, 학습자들의 문어와 구어의 어휘 밀도는 관찰 기간(18개월) 동안 뚜렷한 발달이 나타나지 않았다. 또한 구어의 어휘 밀도 측정 결과 역시 문어와 마찬가지로 다른 복잡성 측정 지표들과 다른 결과를 보였다. 어휘 밀도에 대한 기존의 연구들을 살펴보면 Wolfe-Quintero 외(1998)에서는 숙달도가 낮은 학습자들은 내용어보다 기능어를 적게 사용하기 때문에 어휘 밀도가 높게 나타날 수 있다고 하였고 Hwang(2012)에서는 긴 텍스트를 산출하는 학습자의 어휘 밀도가 짧은 텍스트를 산출하는 학습자에 비해 낮다고 하고 있다. 즉 숙달도가 높아지면서 더 많은 텍스트를 산출할수록 어휘 밀도가 낮아질 수 있다는 것이다. 동시에 이는 한국어에서 어휘 밀도는 복잡성의 측정에 적합하지 않음(서신웨, 2022:210)을 지지하는 결과이기도 하다.

2.2.3. 어휘 정교도

구어의 어휘 정교도는 문어와 마찬가지로 <한국어(KSL) 교육과정(2017)>의 <중학교 언어 재료>의 '일상생활 어휘 목록' 중 초급 어휘에 포함되지 않는 어휘의 비율로 측정하였다. 우선 각 시점별 학습자의 어휘 정교도와 6개월을 단위로 어휘 정교도의 변화 추이를 그래프로 표시해 보면 다음의 <그림 22>와 같다. 학습자들의 어휘 정교도의 추세선을 살펴보면 가운데가 약간 낮고 양쪽이 높은 모양으로 어휘 정교도가 감소하였다가 다시 증가하는 것을 확인할 수 있다. 이러한 과정은 6개월 단위 변화 추이에서도 나타나는데 기간별 변화폭을 살펴보면 1~6개월보다 7~12개월에 0.02 감소하였다가 13~18개월에 다시 3.00 증가하는 것으로 나타났다.

<그림 22> 시점별 구어 어휘 정교도와 6개월 단위 발달 추이



어휘 정교도의 기간별 변화가 통계적으로 유의한지 검증하기 위해 일원배치 분산분석(One-Way ANOVA)을 실시하였다. 그 결과 <표 69>와 같이 F=5.166, p=.007로 기간별로 유의한 차이가 있는 것으로 나타났다. 사후검정 결과 1~6개월, 7~12개월과 13~18개월(1~6개월, 7~12개월<13~18개월) 사이에 유의한 차이가 발견되었다. 이는 구어의 어휘 정교도가 13~18개월에 이르러 뚜렷한 발달이 일어남을 의미한다.

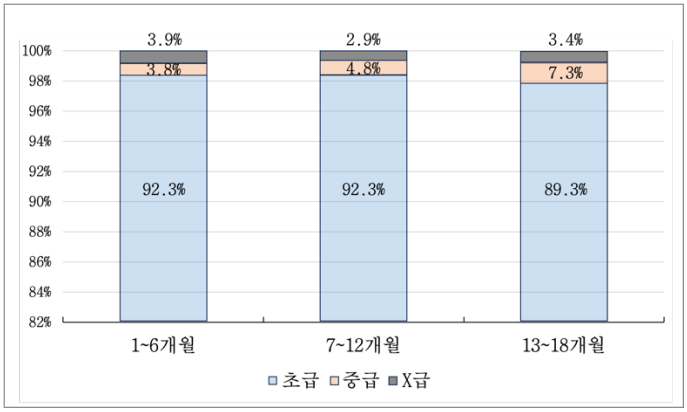
<표 69> 기간별 구어 어휘 정교도

측정 지표	기술통계량				F/p	사후검정
	기간(개월)	N	Mean	SD		
어휘 정교도	1~6	36(a)	7.72	4.28	5.166/.007**	a,b<c
	7~12	36(b)	7.71	4.33		
	13~18	36(c)	10.71	5.04		

*p<.05 **p<.01 ***p<.001

6개월을 단위로 각 등급별 어휘의 사용 비율을 살펴보면 학습자들은 7~12개월에는 1~6개월보다 중급 어휘의 비율이 늘어나지만 초급 어휘의 비율에는 변화가 없었다. 이후 13~18개월에는 초급 어휘의 비율이 감소하고 중급 어휘의 비율이 높아지면서 어휘 정교도가 증가하고 있으나 문어와 마찬가지로 16~18개월에도 여전히 초급 어휘가 90%에 가까운 비율을 차지하면서 중·고급 어휘를 뚜렷하게 확장시키지 못하고 있었다.

<그림 23> 기간별 사용 어휘의 등급 분포



학습자별 어휘 정교도의 발달 추세를 살펴보면 B, C학습자는 평균과 마찬가지로 13~18개월에 뚜렷한 증가를 보이고 A학습자는 세 기간 사이에 큰 변화를 보이지 않고 있었다. 학습자 간 어휘 정교도에 차이가 있는지 확인하기 위해 일원배치 분산분석을 실시한 결과 <표 70>과 같이 세 학습자의 어휘 정교도는 F=16.398, p=.000으로 학습자 간에 유의한 차이가 있는 것으로 나타났다. 사후 검정 결과 A, B학습자와 C학습자(B, A>C) 사이에 차이를 보이고 있었다.

<표 70> 학습자별 구어 어휘 정교도

측정 지표	기술통계량				F/p	사후검정
	학습자	N	Mean	SD		
어휘 정교도	AS	36	10.31	4.68	16.398/.000***	BS,AS>CS
	BS	36	10.37	4.40		
	CS	36	5.46	3.32		

*p<.05 **p<.01 ***p<.001

문어와 구어의 어휘 복잡성 발달을 종합해 보면, 어휘 다양도는 문어와 구어에서 비슷한 경로로 발달하며 관찰 시작 12개월이 지난 시점부터 발달이 정체되는데 특히 체언의 정체 및 퇴행이 뚜렷하게 나타나 어휘 다양도의 지속적인 증가에 영향을 주고 있었다. 그리고 어휘 밀도는 관찰 기간 내내 뚜렷한 발달이나 퇴행을 보이지 않고 있었다. 마지막으로 어휘 정교도는 문어와 구어 모두 퇴행과 발달을 반복하지만 문어는 1~6개월부터 8% 정도의 정교도에서 출발하여 완만하게 증가하는 것과 달리 구어에서는 1~12개월까지 7% 정도에 머무르다가 13~18개월에 문어와 비슷한 수준으로 뒤늦게

발달이 일어나고 있었다. 그러나 문어와 구어 모두 관찰이 끝나는 시점까지 90%에 가까운 초급 어휘의 비율을 유지하며 중·고급 어휘의 확장이 일어나지 못하고 있었다.

한편으로 어휘의 산출은 주제와 밀접한 연관이 있다는 점에서 학습자 어휘 다양도가 정체되는 현상이나 초급 어휘를 집중적으로 사용하는 현상이 일상생활과 관련된 친숙한 주제들을 중심으로 하는 과제의 영향을 받았을 가능성을 배제할 수는 없을 것이다. 그러나 통사 복잡성의 지표들은 지속적으로 증가하였다는 점이나 어휘 다양도에서 용언의 다양도는 지속적으로 증가하는 반면 체언의 다양도는 오히려 퇴행을 보였다는 점, 고급 어휘가 아닌 중급 어휘의 비율도 매우 낮다는 점들을 고려해 볼 때 과제의 영향보다는 학습자의 어휘 능력의 한계도 큰 원인으로 작용하는 것으로 보인다. 이러한 결과를 바탕으로, 중도입국 청소년은 별도의 한국어 교육이 없는 공교육 환경에서 통사 복잡성은 꾸준히 발달할 수 있지만 어휘 복잡성의 지속적인 발달에는 어려움이 있다고 추측해볼 수 있다.

2.3. 산출 양식에 따른 어휘 복잡성 비교

2.3.1. 어휘 다양도의 비교

전체 기간을 단위로 문어와 구어의 어휘 다양도의 평균을 비교해 보며 문어의 어휘 다양도는 4.27, 구어의 어휘 다양도는 3.96으로 문어에서 더 높은 어휘 다양도를 보이고 있었다. 이를 다시 6개월 단위로 나누어 비교해 보면 세 기간 모두 문어의 어휘 다양도가 구어 어휘 다양도에 비해 높으며 그 차이가 점점 커지는 것을 확인할 수 있다. 이는 학습자들은 시간이 지날수록 구어보다 문어에서 더 다양한 어휘를 사용할 수 있음을 의미한다.

문어와 구어의 어휘 다양도에 나타난 차이가 통계적으로 유의한지 확인하기 위해 t 검정을 실시한 결과 <표 71>에서와 같이 1~6개월(t=1.717, p=.095)에는 유의한 차이가 없으나 7~12개월(t=3.906, p=.000), 13~18개월(t=4.427, p=.000)과 전체 기간(t=5.520, p=.000)에서 통계적으로 유의미한 차이가 있는 것으로 나타났다. 즉 구어의 어휘 다양도는 문어에 비해 늦게 발달하며 증가 속도 또한 문어에 비해 느리다는 것을 알 수 있다. 이는 통사 복잡성의 발달에서 구어의 통사 복잡성이 문어에 비해 늦게 나타났으나 문어에 비해 더 빠른 속도로 발달하여 13~18개월에는 산출 양식에 따른 통사 복잡성의 차이가 유의하지 않았던 것과 대조되는 현상으로 통사 복잡성과 어휘 다양도의 발달 경로가 동일하지 않음을 의미하기도 한다.

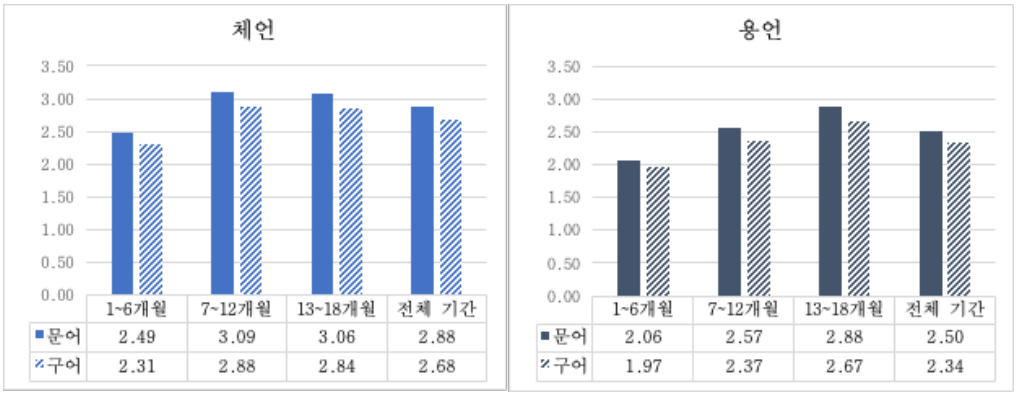
<표 71> 기간별 문어와 구어의 어휘 다양도

측정 지표	기간	N	문어		구어		t/p
			Mean	SD	Mean	SD	
CTTR	1~6개월	72	3.71	0.61	3.51	0.53	1.717/.095
	7~12개월	72	4.44	0.56	4.10	0.49	3.906/.000***
	13~18개월	72	4.66	0.48	4.25	0.40	4.427/.000***
	전체 기간	216	4.27	0.68	3.96	0.57	5.520/.000***

*p<.05 **p<.01 ***p<.001

전체 어휘 다양도 외에 체언과 용언의 어휘 다양도를 산출 양식에 따라 비교해 보면 다음의 <그림 24>와 같다. 그림에서 확인할 수 있듯이 체언과 용언의 어휘 다양도 역시 전체 어휘의 어휘 다양도와 마찬가지로 6개월 단위의 각 기간과 전체 기간 모두 문어의 다양도가 더 높을 뿐만 아니라 시간이 지날수록 문어와 구어의 차이가 점점 커지고 있었다. 한국어 학습자와 모어화자의 구어 산출물에 나타난 어휘 다양도를 비교한 남주연·김영주(2014)에서는 모어화자와 제2언어 학습자의 전체 어휘 다양도가 같음에도 체언과 용언의 어휘 다양도를 따로 분석한 결과 두 집단 간에 뚜렷한 차이가 있음을 발견하고 체언과 용언은 학습자의 숙달도가 상당한 수준에 이르더라도 모어화자와는 큰 차이가 있음을 밝히고 있다. 동시에 이는 체언과 용언이 어휘 다양도 발달에서 갖는 중요성을 엿볼 수 있는 연구 결과로 문어와 구어의 균형적인 어휘 발달을 위해서는 구어에서 체언과 용언의 산출을 촉진할 필요가 있음을 알 수 있다.

<그림 24> 산출 양식에 따른 체언과 용언의 어휘 다양도



문어와 구어의 어휘 다양도의 차이를 학습자별로 나누어 살펴보면 다음의 <표 72>와 같다. 표에서 확인할 수 있듯이 A, B학습자는 문어의 어휘 다양도가 유의하게 더 높

지만 C학습자는 문어와 구어의 어휘 다양도에 차이가 없는 것으로 나타났다.

<표 72> 학습자별 문어와 구어의 어휘 다양도

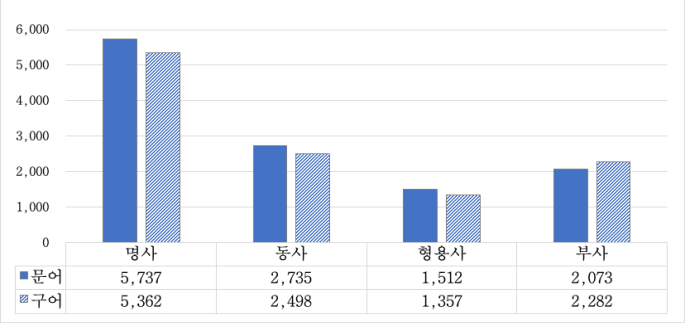
측정 지표	기술통계량				t/p	r/p
	학습자	N	Mean	SD		
CTTR	AW	36	4.32	0.70	4.872/.000***	.693/.000***
	AS	36	3.88	0.67		
	BW	36	4.65	0.52	5.309/.000***	.361/.030*
	BS	36	4.18	0.42		
	CW	36	3.84	0.57	.305/.763	.395/.017*
	CS	36	3.81	0.53		

*p<.05 **p<.01 ***p<.001

2.3.2. 어휘 밀도의 비교

2장에서 살펴보았듯이 어휘 밀도는 문어와 구어를 구별하는 하나의 특징으로 간주되어 왔으며 문어 텍스트의 어휘 밀도가 구어보다 높다는 여러 연구들이 있었다. 산출 양식에 따른 어휘 밀도를 비교하기 이전에 학습자들이 산출한 문어와 구어의 내용어 구현 수를 비교해 보면 학습자들은 문어에서 12,057개, 구어에서 11,499개의 내용어를 산출하고 있었다. 이를 품사별로 나누어 구체적으로 살펴보면 다음의 <그림 25>와 같다. 그림에서 확인할 수 있듯이 문어와 구어에서 많이 사용된 내용어의 유형은 모두 '명사>동사>부사>형용사'순으로 나타났다. 산출 양식에 따른 유형별 내용어 빈도를 비교해 보면 문어에서는 명사, 동사, 형용사를 더 많이 사용하고 구어에서는 문어보다 부사를 더 많이 사용하는 것으로 나타났다.

<그림 25> 산출 양식에 따른 유형별 내용어 수



문어와 구어의 어휘 밀도를 비교해 보면, 먼저 전체 기간을 단위로 볼 때 문어의 어휘 밀도는 56.45, 구어의 어휘 밀도는 56.08로 문어의 어휘 밀도가 약간 더 높았고 기간별로 나누어 살펴보면 1~6개월에는 구어의 어휘 밀도가 더 높았으나 7~12개월과 13~18개월에는 문어의 어휘 밀도가 더 높아지는 것을 확인할 수 있었다. 문어와 구어의 어휘 밀도에 나타난 차이가 통계적으로 유의한지 확인하기 위해 t검정을 실시한 결과 <표 73>에서와 전체 기간과 6개월 단위의 세 기간 및 전체 기간의 평균 모두에서 통계적으로 유의미한 차이를 보이지 않았다. 이는 학습자들의 어휘 밀도는 산출 양식에 따른 차이가 크지 않음을 의미한다.

<표 73> 기간별 문어와 구어의 어휘 밀도

측정 지표	기간	N	문어		구어		t/p
			Mean	SD	Mean	SD	
어휘 밀도	1~6개월	72	55.74	6.23	55.83	6.63	.060/.952
	7~12개월	72	57.29	4.62	56.42	6.31	.707/.484
	13~18개월	72	56.34	4.67	55.99	4.80	.296/.769
	전체 기간	216	56.45	4.78	56.08	5.06	.498/.619

*p<.05 **p<.01 ***p<.001

문어와 구어의 어휘 밀도의 차이를 학습자별로 나누어 살펴보면 다음의 <표 74>와 같다. 표에서 확인할 수 있듯이 A와 C학습자는 구어의 어휘 밀도가 더 높고 B학습자는 문어의 어휘 밀도가 더 높았다. 그러나 세 학습자 모두 문어와 구어의 어휘 밀도의 차이가 유의하지 않은 것으로 나타났다.

<표 74> 학습자별 문어와 구어의 어휘 밀도

측정 지표	기술통계량				t/p	r/p
	학습자	N	Mean	SD		
어휘 밀도	AW	36	54.88	4.45	-.025/.980	-.106/.540
	AS	36	54.91	4.21		
	BW	36	58.29	4.80	1.532/.135	-.223/.191
	BS	36	55.99	6.62		
	CW	36	56.19	5.84	-.895/.377	.234/.170
	CS	36	57.34	6.51		

*p<.05 **p<.01 ***p<.001

산출 양식에 따른 어휘 밀도의 비교 결과는 문어의 어휘 밀도가 구어에 비해 더 높다

는 기존의 연구 결과들과 차이가 있었다. 영어의 경우 어휘 밀도는 구어 텍스트에서 평균 40% 미만, 문어 텍스트에서 평균 40% 이상으로 보고되고 있고(Ure, 1971; McCarthy, 1990) 한국어에서 사용역에 따른 어휘 밀도를 비교한 안의정(2017)의 연구[71]에서는 문어 텍스트의 평균 어휘 밀도는 52.1%, 구어 텍스트의 평균 어휘 밀도는 47.2%(독백: 48.1%, 대화: 46.5%)임을 밝히고 있다. 그러나 본고의 학습자의 평균 어휘 밀도는 문어에서 56.45, 구어에서 56.08로 안의정(2017)에서 제시한 문어 텍스트 중 어휘 밀도가 가장 높은 교과서 텍스트(55.0)와 구어 텍스트 중 어휘 밀도가 가장 높은 강연(50.5)보다도 높은 것으로 나타났다. 이는 어휘 밀도가 숙달도를 가늠하는 척도로 사용되기가 힘들다는 점을 방증하는 것으로 볼 수 있을 것이다.

2.3.3. 어휘 정교도의 비교

학습자들의 어휘 정교도가 산출 양식에 따라 어떤 차이가 있는지 알아보기 위해 문어와 구어의 어휘 정교도를 비교하였다. 전체 기간에 나타난 문어와 구어의 어휘 정교도를 비교해 보면 문어의 평균 어휘 정교도는 9.27, 구어의 평균 어휘 정교도는 8.71로 문어의 평균 정교도가 더 높은 것으로 나타났다. 이를 기간별로 살펴보면 1~6개월과 7~12개월에는 문어의 어휘 정교도가 더 높지만 13~18개월에 이르러서는 구어의 어휘 정교도가 미세하게 더 높아지는 것을 확인할 수 있다. 이는 앞에서 살펴본 기간별 발달과도 관련이 있는데 문어의 어휘 정교도는 기간별로 유의한 증가를 보이지 않는 것과 달리 구어의 어휘 정교도는 13~18개월에 유의미한 증가를 보였다. 즉 문어의 어휘 정교도는 초반에는 구어에 비해 높은 수준으로 출발하였으나 시간이 지나도 뚜렷한 증가를 보이지 않은 반면 구어의 어휘 정교도는 1~6개월과 7~12개월에는 문어에 비해 낮은 수준에 머물러 있다가 13~18개월에 빠르게 증가하여 문어와 비슷한 수준에 도달하는 것이다. 이를 통해 구어의 어휘 정교도는 문어에 비해 늦게 나타나지만 발달 속도가 더 빠르다고 정리할 수 있다.

문어와 구어의 어휘 정교도 차이가 통계적으로 유의미한 지 확인하기 위해 t검정을 실시한 결과 <표 75>와 같이 전체 기간 및 6개월 단위 세 기간 모두 통계적으로 유의한 차이를 보이지는 않았다. 이는 학습자들의 문어와 구어의 정교도는 출발 수준과 발

[71] 안의정(2017)에서는 한국어의 여러 장르의 문어와 구어 텍스트의 어휘 밀도를 분석하여 비교하였는데 본고와 마찬가지로 '일반명사, 동사, 형용사, 일반부사'를 내용어로 어휘 밀도를 분석하였다는 점에서 비교가 가능하다.

달 속도에 차이가 있지만 통계적으로 유의한 수준의 뚜렷한 차이가 아님을 의미한다.

<표 75> 기간별 문어와 구어의 어휘 정교도

측정 지표	기간	N	문어		구어		t/p
			Mean	SD	Mean	SD	
어휘 정교도	1~6개월	72	8.86	6.56	7.72	4.28	.885/.382
	7~12개월	72	8.27	4.07	7.71	4.33	.705/.486
	13~18개월	72	10.69	5.46	10.71	5.04	.017/.987
	전체 기간	216	9.27	5.51	8.71	4.74	.889/.376

$^{*}p<.05$ $^{**}p<.01$ $^{***}p<.001$

학습자별로 산출 양식에 따른 어휘 정교도를 비교해 보면 A, C학습자는 문어의 정교도가 더 높은 반면 B학습자는 구어의 어휘 정교도가 더 높게 나타나 일관된 차이를 보이지는 않았다. 이러한 차이가 통계적으로 유의미한지 t-검정을 실시해 보면 다음의 <표 76>과 같다. 표에서 확인할 수 있듯이 C학습자(t=-2.601, p=.014)만이 문어와 구어의 어휘 정교도에 유의한 차이가 있는 것으로 나타났다.

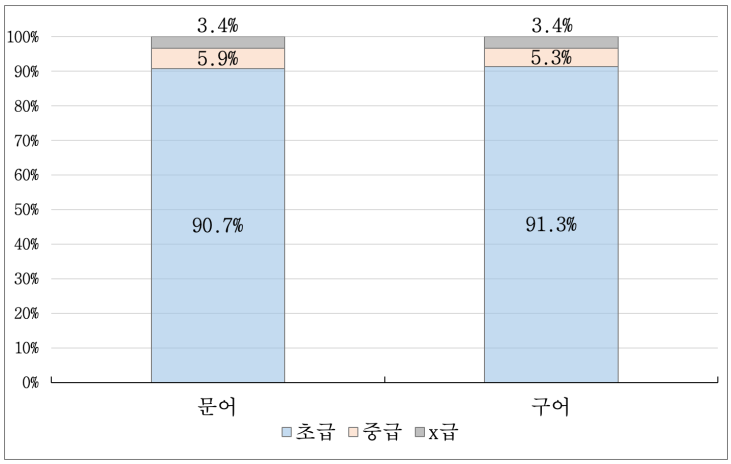
<표 76> 학습자별 문어와 구어의 어휘 정교도

측정 지표	기술통계량				t/p	r/p
	학습자	N	Mean	SD		
어휘 정교도	AW	36	11.02	7.03	.482/.633	-.089/.605
	AS	36	10.31	4.68		
	BW	36	9.19	3.80	-1.439/.159	.297/.079
	BS	36	10.37	4.40		
	CW	36	7.61	4.78	2.601/.014*	.291/.085
	CS	36	5.46	3.32		

$^{*}p<.05$ $^{**}p<.01$ $^{***}p<.001$

등급별 어휘 사용 비율을 살펴보면 문어와 구어에서 어휘 목록에 포함되지 않는 어휘의 사용 비율은 동일하게 나타났다. 그러나 초급과 중급 어휘의 비율에는 차이가 있었는데 문어에서는 90.7%가 초급에 해당하는 어휘인 것으로 나타났고 구어에서는 91.3%가 초급 어휘에 해당하는 것으로 나타났다. 즉 구어에서 문어보다 초급에 해당하는 어휘를 더 많이 사용하고 중급에 해당하는 어휘를 덜 사용하는 것이다. 그러나 문어와 구어 모두 초급 어휘의 비율이 90% 이상을 차지하여 큰 차이를 보이지는 않았다.

<그림 26> 산출 양식에 따른 등급별 어휘 분포



앞의 결과들을 통해 중국 배경 중도입국 청소년들은 전체적으로 문어에서 더 다양한 어휘를 사용하고 시간이 지날수록 문어의 어휘 다양도가 구어에 비해 더 빠르게 발달하지만 어휘의 밀도나 어휘의 정교도에 있어서는 문어와 구어에 큰 차이가 없음을 확인할 수 있었다. 이는 일란성 쌍둥이 학습자의 어휘 발달을 종적으로 연구하여 구어에서 어휘 다양도는 더 낮지만 어휘 난이도(정교도)는 낮지 않다는 것을 발견한 Chan 외(2014)와 같은 결과로 어휘 다양도는 산출 양식의 영향을 받지만 어휘 정교도와 밀도는 산출 양식에 크게 영향을 받지 않는다는 결론을 얻을 수 있다. 이는 또 같은 과제를 사용하여 문어와 구어의 어휘 다양도를 비교한 뒤 문어의 어휘 다양도가 더 높다는 결론을 얻은 Kormos & Trebits(2012)와도 같은 결과이다. 이러한 결과는 문어 산출에서 시간적 압박이 덜하고 구어 산출이 더 자동적이기 때문에 쓰기의 인지적 부하가 더 적고, 이는 쓰기에서 특정 의미를 표현하는데 더 적절한 단어를 찾을 수 있다는 Halliday(1989)의 주장을 뒷받침한다고 볼 수 있다.

3. 정확성의 발달

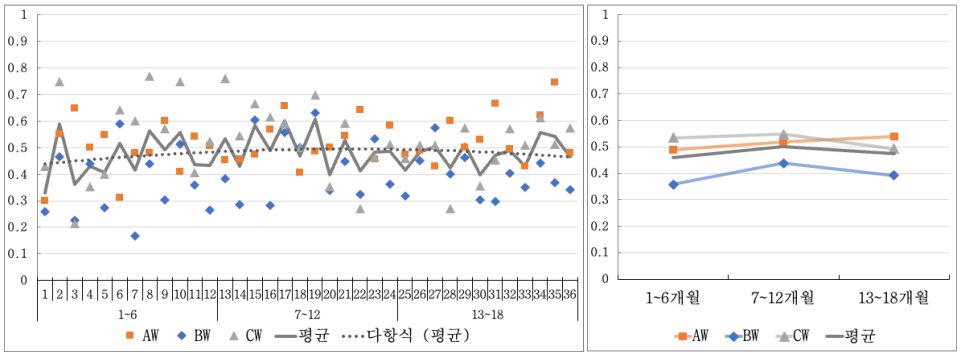
본 절에서는 학습자들의 문어와 구어에 나타난 정확성 발달을 살펴본다. 정확성은 3장에서 제시한 바와 같이 무오류절의 비율, 단위당 무오류절 수, 단위당 오류 수와 각 유형별 오류의 빈도를 통해 측정하고, 각 측정 지표의 시간의 흐름에 따른 발달을 살펴본 뒤 산출 양식에 따른 차이를 비교하도록 한다.

3.1. 문어의 정확성 발달

3.1.1. 무오류절의 비율

학습자들의 무오류절 비율의 시점별 변화 추이와 6개월 단위 평균을 살펴보면 다음의 <그림 27>과 같다. 세 학습자의 평균 추세선을 통해 시간의 흐름에 따른 무오류절의 변화를 살펴보면 미세하게 상승하였다가 다시 하락하는 모습을 보인다. 특히 정확성의 하락은 10개월(20차)을 분기점으로 선명한 차이를 보이는데 이는 무오류절의 비율이 관찰 시작 10개월을 기점으로 하락하고 있음을 의미한다. 6개월 단위의 변화폭을 살펴보면 7~12개월에는 1~6개월보다 0.04 증가한 것에 반해 13~18개월에는 7~12개월보다 0.03 감소하여 정확성이 약간 증가하였다가 다시 감소하는 것을 확인할 수 있다.

<그림 27> 시점별 문어 무오류절 비율과 6개월 단위 발달 추이



무오류절의 기간별 변화가 통계적으로 유의한지 검증하기 위해 일원배치 분산분석(One-Way ANOVA)을 실시하였다. 그 결과 <표 77>과 같이 학습자들의 기간별 무오류절의 변화는 F=.891, p=.414로 통계적으로 유의하지 않은 것으로 나타났다. 이는 학습자들의 무오류절 비율은 관찰 기간 동안 약간의 증가와 감소는 있었지만 통계적으로 유의미할 정도의 변화는 발생하지 않았음을 의미한다.

<표 77> 기간별 문어 무오류절 비율

측정 지표	기술통계량				F/p	사후검정
	기간(개월)	N	Mean	SD		
무오류절 비율	1~6	36(a)	0.46	0.16	.891/.414	-
	7~12	36(b)	0.50	0.12		
	13~18	36(c)	0.47	0.11		

*p<.05 **p<.01 ***p<.001

학습자별로 살펴보면 세 학습자 중 B, C학습자는 평균과 같은 변화 추이를 보이고 A 학습자만이 기간별로 미미한 증가를 보이고 있었다. 세 학습자의 무오류절의 비율 차이가 통계적으로 유의한지 확인하기 위해 일원배치 분산분석을 실시한 결과 <표 78>과 같이 F=13.341, p=.000으로 학습자 간에 유의한 차이가 있는 것으로 나타났다. 사후검정 결과 A, C학습자의 무오류절 비율이 B학습자에 비해 유의하게 높은 것으로 나타났다.

<표 78> 학습자별 문어 무오류절 비율

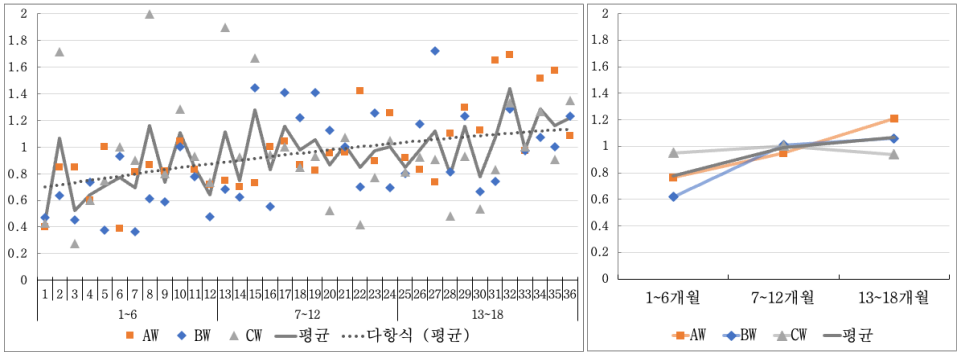
측정 지표	기술통계량				F/p	사후검정
	학습자	N	Mean	SD		
무오류절 비율	AW	36	0.52	0.09	13.341/.000***	AW,CW>BW
	BW	36	0.40	0.12		
	CW	36	0.52	0.14		

*p<.05 **p<.01 ***p<.001

3.1.2. 문장당 무오류절 수

학습자들의 문장당 무오류절 수의 시점별 변화 추이와 6개월 단위 평균을 살펴보면 다음의 <그림 28>과 같다. 2차 다항식 추세선을 통해 살펴보면 세 학습자의 문장당 무오류절 수는 전체적으로 시간의 흐름에 따라 완만하게 상승하는 추세를 보이지만 무오류절의 비율과 비슷하게 10개월을 기점으로 선명한 감소를 보이다가 16개월(32차)에 다시 증가하고 있었다. 6개월을 단위로 문장당 무오류절 수의 변화폭을 살펴보면 7~12개월에는 1~6개월보다 0.21개 증가하였고 13~18개월에는 7~12개월보다 0.08개 증가하여 지속적인 증가를 보이지만 13~18개월의 증가는 매우 미미하여 정확성의 발달이 거의 정체되어 있음을 알 수 있었다.

<그림 28> 문장당 무오류절 수의 시점별 발달 추이와 6개월 단위 발달 추이



문장당 무오류절 수의 기간별 변화가 통계적으로 유의한지 검증하기 위해 일원배치 분산분석(One-Way ANOVA)을 실시하였다. 분석 결과 <표 79>와 같이 학습자들의 기간별 무오류절의 변화는 F=7.405, p=.001로 통계적으로 유의한 것으로 나타났으며 사후검정 결과 1~6개월과 7~12개월, 13~18개월(1~6개월<7~12개월, 13~18개월) 사이에 유의한 차이를 보였다. 이는 그래프에서 확인한 것과 마찬가지로 학습자들의 정확성이 12개월까지는 발달하지만 이후의 13~18개월 구간에서는 정체됨을 의미한다.

<표 79> 기간별 문장당 무오류절 수

측정 지표	기술통계량				F/p	사후검정
	기간(개월)	N	Mean	SD		
문장당 무오류절 수	1~6개월	36(a)	0.78	0.35	7.405/.001**	a<b,c
	7~12개월	36(b)	0.99	0.33		
	13~18개월	36(c)	1.07	0.31		

*p<.05 **p<.01 ***p<.001

학습자별로 살펴보면 세 학습자는 7~12개월까지는 그 정도에 차이가 있지만 모두 문장당 무오류절 수가 증가하는 것과 달리 13~18개월에는 서로 다른 발달 경로를 보이고 있었는데 B학습자는 거의 정체된 양상을 보이고 C학습자는 하락하고 있었으며 A학습자만이 증가를 보이고 있었다. 세 학습자의 문장당 무오류절 수의 평균은 A>C>B 순으로 나타났는데 학습자별 무오류절의 비율 차이가 통계적으로 유의한지 확인하기 위해 일원배치 분산분석을 실시한 결과 <표 80>과 같이 F=.544, p=.582로 학습자 간 문장당 무오류절 수의 차이는 유의하지 않은 것으로 나타났다.

<표 80> 학습자별 문장당 무오류절 수

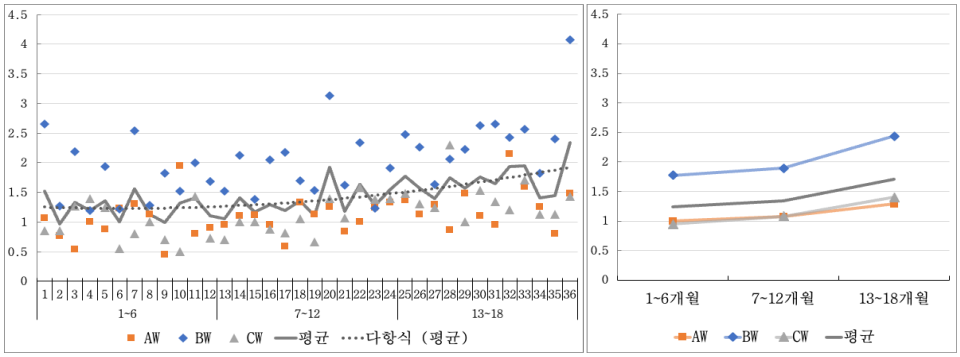
측정 지표	기술통계량				F/p	사후검정
	학습자	N	Mean	SD		
문장당 무오류절 수	AW	36	0.97	0.31	.544/.582	-
	BW	36	0.90	0.34		
	CW	36	0.96	0.40		

*p<.05 **p<.01 ***p<.001

3.1.3. 문장당 오류 수

다음으로, 문장당 오류 수 측정을 통해 학습자들의 정확성을 살펴보도록 한다. 학습자들의 문장당 오류 수의 시점별 변화 추이와 6개월 단위 평균을 살펴보면 다음 <그림 29>와 같다. 우선 세 학습자의 평균 2차다항식 추세선을 통해 시간의 흐름에 따른 문장당 오류 수의 변화를 살펴보면 처음에는 완만하지만 갈수록 가팔라지는 곡선 모양을 띠고 있는 것을 확인할 수 있다. 이는 학습자들의 문장당 오류 수가 시간이 지날수록 빠른 속도로 증가하고 있음을 의미한다. 6개월을 단위 변화폭을 확인해 보면 이러한 현상은 더 뚜렷하게 나타나는데 문장당 오류 수는 7~12개월에는 1~6개월보다 0.11 증가하고 13~18개월에는 7~12개월보다 0.36 증가하여 후반으로 갈수록 증가 속도가 빨라지고 있었다.

<그림 29> 시점별 문장당 오류 수와 6개월 단위 발달 추이



문장당 오류 수의 기간별 변화가 통계적으로 유의한지 검증하기 위해 일원배치 분산분석(One-Way ANOVA)을 실시한 결과 <표 81>과 같이 학습자들의 기간별 무오류절의 변화는 F=6.173, p=.003으로 통계적으로 유의한 것으로 나타났다. 사후분석 결과

1~6개월, 7~12개월과 13~18개월(1~6개월, 7~12개월<13~18개월) 사이에 유의한 차이를 보여 학습자들의 오류 수가 13~18개월에 유의미하게 증가하는 것으로 나타났는데 이는 정확성이 13~18개월에 유의미하게 감소하였음을 의미하기도 한다.

<표 81> 기간별 문장당 오류 수

측정 지표	기술통계량				F/p	사후검정
	기간(개월)	N	Mean	SD		
문장당 오류 수	1~6	36(a)	1.24	0.56	6.173/.003**	a,b<c
	7~12	36(b)	1.35	0.53		
	13~18	36(c)	1.71	0.68		

*p<.05 **p<.01 ***p<.001

학습자별로 살펴보면 세 학습자는 그 정도의 차이가 있으나 모두 평균과 같은 변화 추세를 보이고 있었다. 세 학습자의 문장당 오류 수의 평균은 B>A>C순으로 나타났는데 학습자별 문장당 오류 수의 차이가 통계적으로 유의한지 확인하기 위해 일원배치 분산분석을 실시한 결과 <표 82>와 같이 F=33.926, p=.000으로 학습자 간 유의한 차이가 있는 것으로 나타났다. 사후검정 결과 B학습자의 문장당 오류 수가 A, C학습자에 비해 유의하게 높은 것으로 나타났다.

<표 82> 학습자별 문장당 오류 수

측정 지표	기술통계량				F/p	사후검정
	학습자	N	Mean	SD		
문장당 오류 수	AW	36	1.32	0.45	33.926/.000***	BW>AW,CW
	BW	36	1.80	0.56		
	CW	36	1.06	0.32		

*p<.05 **p<.01 ***p<.001

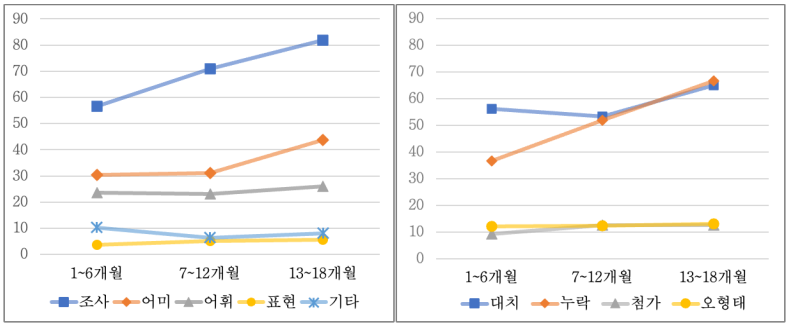
세 가지 지표를 통해 살펴본 정확성의 발달을 정리하면, 무오류절의 비율을 통해 살펴본 학습자들의 정확성 발달은 기간별로 유의한 차이가 없어 관찰 기간(18개월) 동안 정확성에 뚜렷한 변화가 없는 것으로 나타났고 문장당 무오류절 수로 살펴본 학습자들의 정확성은 1~6개월에서 7~12개월 사이에는 유의하게 발달하지만 이후 정확성 발달이 정체되는 것으로 나타났다. 그리고 문장당 오류 수를 통해 살펴본 정확성은 1~6개월과 7~12개월 사이에는 차이가 없으나 13~18개월에는 뚜렷하게 하락하는 것으로 나타났다. 즉 전반적으로 학습자들의 문어 정확성은 측정 지표에 따라 다르게 나타나는

데 7~12개월에는 먼저 무오류절 수가 빠르게 증가하지만 이후 13~18개월에는 오류 수가 빠르게 증가하며 시간이 지날수록 정확성이 하락하는 것으로 정리할 수 있다. 학습자별로 정리하면, A학습자는 무오류를 통한 측정에서 세 학습자 중 가장 높은 측정값을 보이는 반면 B학습자는 오류를 통한 측정에서 가장 높은 측정값을 보이고 있었으며 C학습자는 두 학습자의 중간 정도 수준을 보이고 있었다.

3.1.4. 오류 유형

정확성은 오류와 직결된다. 따라서 본 절에서는 학습자들의 정확성 발달에 영향을 주는 오류들을 유형별로 살펴보고자 한다. 먼저 학습자들의 언어 자료에 나타난 오류들을 오류 유형과 오류 현상에 따라 나누고 6개월을 단위로 묶어 시간의 흐름에 따른 오류의 변화를 분석하도록 한다. 비교의 편의를 위해 오류 빈도 역시 표준화 빈도를 사용하였다. 유형별 오류를 자세히 살펴보기 이전 학습자들의 문어 자료에 나타난 오류들을 오류 영역과 오류 현상에 따라 6개월 단위로 묶어 전체적인 변화 양상을 표시하면 다음의 <그림 30>과 같다.

<그림 30> 문어 자료에 나타난 오류 영역 및 오류 현상의 기간별 빈도



우선 오류 영역별로 살펴보면 '조사 오류>어미 오류>어휘 오류>기타 오류>표현 오류' 순으로 나타나 학습자들이 조사에서 가장 많은 오류를 범하고 있음을 확인할 수 있었다. 또한 각 영역별 오류 중에서 기타 오류를 제외한 다른 오류들은 모두 시간이 흐름에 따라 점점 증가하고 있었는데 그중에서 조사 오류의 증가가 가장 가파르게 나타났고 어미 오류와 어휘 오류는 7~12개월까지 큰 변화가 없다가 13~18개월에 급격히 상승하였다. 오류 현상별로 살펴보면 '대치 오류>누락 오류>오형태 오류>첨가 오류' 순으로 나타나고 있었으나 각각의 오류 현상은 시간의 흐름에 따라 서로 다른 변화를

보이고 있었다. 대치 오류는 7~12개월에 감소하였다가 13~18개월에 다시 증가하였고 누락 오류는 뚜렷하고 가파르게 상승하고 있었다. 그리고 첨가 오류와 오형태 오류는 미세하게 상승했다가 다시 감소하였으나 일정한 수준을 유지하며 큰 변화를 보이지 않았다. 이어서 6개월을 단위로 각 영역별 오류의 변화 양상을 살펴보도록 한다. 먼저 조사 오류의 기간별 변화 양상을 살펴보면 다음의 <표 83>과 같다.

<표 83> 문어 자료에 나타난 조사 오류의 기간별 빈도

구분		1~6개월	7~12개월	13~18개월	합계
조사	대치	25.00	28.46	31.94	85.41
	누락	24.26	34.89	41.87	101.02
	첨가	3.68	5.54	4.96	14.18
	오형태	3.86	2.14	2.98	8.98

앞에서 살펴보았듯이 조사 오류는 오류 영역 중 가장 높은 빈도로 나타났다. 조사 오류 중 가장 빈번하게 나타난 것은 누락 오류와 대치 오류였는데 누락과 대치 오류 모두 시간이 지날수록 점점 증가하고 있었으며 그중에서도 누락 오류의 빈도가 급격히 증가하고 있었다. 첨가와 오형태 오류는 비교적 낮게 나타났는데 첨가 오류는 7~12개월까지 증가하였으나 13~18개월에 약간 감소하였고 오형태 오류는 7~12개월에 감소하였으나 13~18개월에 다시 증가하는 모습을 보이고 있었다. 조사 오류 중 가장 많이 나타난 오류는 주격 조사 '이/가'와 목적격 조사 '을/를', 보조사 '은/는'의 오류였다. 주격 조사는 대치 오류가 가장 많았고 목적격 조사와 보조사는 누락 오류가 가장 많이 발생하였다. 그리고 부사격 조사는 주로 처소의 부사격 조사의 누락과 대치 오류가 빈번하게 발생하고 있었다. 아래 예문(42가)와 (42나)는 보조사와 목적격 조사를 주격 조사로 대치한 오류이고 예문(42다)는 목적격 조사 누락 오류, 예문(42라)는 보조사 누락 오류의 예시이다.

(42) 가. 제가(√저는) 후회하는 일 별로 없는 사람지만* 진짜 후회하는 일 하나 있어요.(BW28)
나. OO는 성격이 내성적이에요. OO는 게임가(√게임을) 너무 잘해요.(AW5)
다. 학생을 가르치는 방법이* 많아요. 그래서 학생을 때리는 것 필요 없어요. 학생들(√학생들을) 자꾸 때리면 나쁠게 해요*.(CW32)
라. 한국 생활(√생활은) 좋지만 재미없는 것도 많아요.(CW13)

한편, 중국인 한국어 학습자의 글쓰기 자료를 분석한 김선정·김성수(2009)에서는 본

고와 다르게 누락 오류보다 대치 오류가 더 많은 것으로 보고하고 있는데 이 연구에서는 한국어 교육기관에 재학 중인 학습자를 대상으로 했다는 점에서 학습자들이 구어와 문어의 특성에 대해 인지하고 문어에서 조사를 누락하지 않았을 것으로 분석하였다. 이와 달리 본고의 학습자는 초반에 잘 사용하던 조사도 후반에 가서는 누락하는 현상이 자주 나타나고 있었다. 따라서 학습자들의 이러한 현상은 관찰 시작 시점에서는 한국어 교실에서 학습한 대로 조사를 사용하려고 노력하지만 구어 중심의 환경에 지속적으로 노출됨과 동시에 사용역에 따른 언어 사용에 대한 교육이 동반되지 못하면서 자연스럽게 문어에서도 조사를 생략하게 되는 것으로 해석할 수 있다.

다음은 어미에서 나타난 오류를 살펴보도록 한다. 한국어에서 어미는 문장의 서법과 경어법, 시제, 문장의 내포와 접속 등 많은 기능을 담당하고 있다. 또한 그 종류와 의미 기능이 다양하여 어미는 한국어 학습자들이 가장 어려워하는 문법 중의 하나이다(김선정·김성수, 2009:151). 본고에서 어미 오류는 어말 어미, 선어말 어미, 연결어미와 전성어미로 나누어 살펴보았다. 각 영역별 오류 현상과 빈도를 정리하면 다음의 <표 84>와 같다.

<표 84> 문어 자료에 나타난 어미 오류의 기간별 빈도

구분			1~6개월	7~12개월	13~18개월	합계
어미	어말어미	대치	6.07	3.65	3.08	12.79
		누락	0.37	-	0.20	0.57
		첨가	-	-	-	-
		오형태	2.02	1.39	1.29	4.70
	선어말어미	대치	-	-	0.10	0.10
		누락	7.35	11.84	17.96	37.15
		첨가	3.86	3.27	3.77	10.90
		오형태	-	-	-	-
	연결어미	대치	3.31	4.03	6.94	14.28
		누락	0.18	0.13	0.60	0.91
		첨가	-	0.38	0.40	0.77
		오형태	1.10	1.51	2.18	4.80
	전성어미	대치	4.41	4.28	5.95	14.65
		누락	1.29	0.38	0.99	2.66
		첨가	0.37	-	-	0.37
		오형태	0.92	0.76	0.69	2.37

어말 어미는 오류가 비교적 적게 나타난 영역이었는데 그중 가장 높게 나타난 오류

는 어말 어미의 대치 오류였다. 어말 어미의 오류는 주로 같은 회차의 자료에서 '-ㅂ니다'와 '-아/어요', '-아/어'를 섞어서 사용하면서 발생한 대치 오류가 많았다. 그러나 어말 어미는 대치, 누락, 첨가, 오형태 오류 모두 시간이 흐름에 따라 감소하는 양상을 보이고 있었다. 이는 어말 어미의 오류는 자연스럽게 교정될 수 있는 부분임을 의미한다. 아래의 예문(43)은 어말 어미의 대치 오류의 예시이다.[72]

(43) 가. 왜냐하면 수학 너무 어려워요. 진짜 잘할 수 없어요. 그런데 너무 싫어하는 거 아니에요. 잘하면 좋아합니다(√좋아해요). 못하면 싫어요.(BW9)

선어말 어미는 어미 오류 중 가장 높은 오류 빈도를 보이는 항목이었는데 가장 빈번하게 발생하는 것은 선어말 어미의 누락과 첨가 오류였으며 주로 오류가 발생하는 항목은 과거 시제 선어말 어미 '-았/었'의 누락과 첨가로 인한 것이었다. 특히 선어말 어미의 첨가 오류는 1~6개월보다 7~12개월과13~18개월에 약간 감소하는 반면 선어말 어미의 누락 오류는 시간이 흐름에 따라 오류 빈도가 급격히 증가하고 있었는데 첨가에 비해 누락 오류가 3배 이상 많게 나타나며 자연스럽게 개선되지 못하고 있었다. 아래의 예문(44가)와 (44나)는 시제 선어말 어미의 첨가 오류이고 (44다)와 (44라)는 누락 오류의 예시이다.

(44) 가. 중국 가야금을 배운 지 8년이 되었었어요(√되었어요).(AW4)
나. 지금 코로나 때문에 할머니를 못* 보러 갔어요(√가요).(AW30)
다. 나는 좋은 사람을 많이 만나지만(√만났지만) 나쁜 사람도 많이 만났어요.(CW28)
라. 할아버지랑 할머니* 돌아가시지만(√돌아가셨지만) 할아버지 할머니랑 함께 있던 추억이 계속 나요*.(BW30)

연결어미와 전성어미 오류는 비슷한 빈도로 나타났는데 연결어미의 오류가 전성어미에 비해 조금 더 높은 빈도로 발생하고 있었다. 연결어미의 오류 중 가장 빈번하게 발생하는 것은 대치 오류로, 연결어미의 대치 오류는 13~18개월에 급격히 증가하고 있었다. 또한 대치 오류 외에 누락, 첨가, 오형태 오류 모두 시간이 흐를수록 점점 증가하

[72] '-ㅂ니다'나 '-아/어요', '-아/어' 사이의 대치 오류는 높임의 오류로도 볼 수 있다. 그러나 학습자들은 주로 '해요체'를 사용하면서도 동일한 자료에서 여러 종결어미를 섞어서 사용하는 경우들이 발생하고 있어 같은 텍스트 내에서 종결어미를 통일시켜야 한다는 것을 인지하지 못한 것으로 판단하여 종결어미의 대치 오류로 분류하였다.

는 추세를 보이고 있었다. 이는 학습자들의 문장당 절 수는 증가하지만 그에 따른 다양한 연결어미를 사용하지 못하거나 연결어미의 의미나 형태에 대한 이해 부족으로 발생한 오류들이다. 특히 빈번하게 발생한 오류는 '계기적 사건'의 '-고(서)'와 '-아(서)/어(서)'의 대치 오류였고 다른 연결어미가 필요한 부분에 '-고(서)'로 대치한 오류도 빈번하게 나타나고 있었다. 예문(45가~다)는 연결어미에 나타난 대치 오류이고 예문(45라)는 선행하는 자모음에 따른 연결어미의 이형태 선택에 오류가 나타난 경우이다.

(45) 가. 나도 갑자기 화나고(√화나서) 그 학생* 때렸어요.(BW29)
나. 검사 결과* 나와서(√나왔는데) 나는 코로나가* 걸렸어요.(CW31)
다. 그렇게 한 개월* 게임 안 하고(√하다가) 또 하고 싶었어요*.(BW34)
라. 모르는 문제를 풀으면(√풀면) 기분가* 너무 좋아요.(AW11)

전성어미 역시 대치 오류가 가장 빈번하게 발생하고 있었는데 누락이나 첨가, 오형태 오류는 시간이 흐름에 따라 약간 감소하는 반면 대치 오류는 13~18개월에 급격히 증가하고 있었다. 연결어미와 전성어미의 이러한 오류 양상은 학습자들이 다양한 어미를 활용하여 점점 더 긴 문장을 산출하는 과정에서 의미나 형태에 대한 교육이 동반되지 못하여 생긴 결과로 해석할 수 있다. 전성어미 오류 중 가장 높은 빈도로 나타난 오류는 '관형사형 전성어미'인 '-ㄴ/는'의 대치 오류였다. 아래 예문(46가)는 형용사에 사용하는 '-ㄴ'과 동사에 사용하는 '-는'을 대치하여 나타난 오류이고 (46나)는 현재형 시제를 표시하는 '는'과 과거형 시제를 표시하는 '-ㄴ'을 대치하여 사용한 오류이며 (46다) 역시 시제 혼동으로 인해 발생한 대치 오류이다.

(46) 가. 저는 A와 B하고* 잘못하는(√잘못한) 점이 다 있다고 생각해요.(AW33)
나. 저는 이 말을 듣는(√들은) 후에 감동했어요.(AW30)
다. 무슨 일이나 다 말할 수 있고 거짓말도 한(√할) 필요 없어요.(BW26)

계속하여 어휘 사용에 나타난 오류를 살펴보도록 한다. 어휘 오류의 분석 결과는 다음의 <표 85>와 같다. 어휘 오류 중 가장 높은 빈도로 나타난 것은 어휘의 대치 오류와 오형태 오류였는데 그중 대치 오류의 빈도가 압도적으로 높게 나타나고 있었다.

<표 85> 문어 자료에 나타난 어휘 오류의 기간별 빈도

구분		1~6개월	7~12개월	13~18개월	합계
어휘	대치	18.01	12.22	16.27	46.50
	누락	0.37	1.26	1.39	3.02
	첨가	1.65	3.27	2.68	7.61
	오형태	3.68	6.05	5.16	14.88

기간별로 살펴보면 대치 오류는 1~6개월에 가장 높게 나타나고 7~12개월에 감소하였다가 13~18개월에 다시 증가하는 양상을 보이고 첨가와 오형태 오류는 7~12개월에 늘어났다가 13~18개월에 다시 감소하고 있었으며 누락 오류는 빈도가 높지는 않으나 꾸준히 증가하고 있었다. 어휘의 대치 오류는 의미 대치, 외국어 대치, 품사 대치, 피사동 대치로 나누었는데 가장 높은 빈도로 나타난 것은 어휘의 의미 대치 오류였다.

의미 대치 오류 중 눈에 띄는 것은 '나'와 '우리'의 대치로 '나 엄마, 나 아빠, 나 집'과 같이 한국어의 '우리'가 아닌 중국어의 습관에 따라 '나'라는 표현을 사용하는 오류였다. 이 외에 '덥다-따뜻하다, 지내다-사귀다, 기억하다-기억나다, 특별하다-독특하다, 좋다-좋아하다, 돌아가다-되돌리다, 화나다-화내다'와 같이 형태나 의미가 유사한 어휘의 의미 혼동으로 인한 오류나 '시험을 받다(√시험을 치다), 같이 있다(√사귀다), 병이 있다(√병에 걸리다), 골절이 나다(√골절하다)'와 같은 중국어식 표현으로 인한 오류들도 빈번하게 발생하고 있었다. 그리고 이 외에 '건강(√건강하다) 변화(√변화하다), 불쌍(√불쌍하다)와 같이 '어근+ 하다'의 형태에서 '하다'를 누락하는 오류들도 비교적 자주 발생하고 있었다. 이는 중국어에 동일한 형태의 단어가 동사와 명사로 동시에 쓰이는 동명사 어휘들의 영향을 받은 것으로 명사가 동사의 기능을 한다고 인식하여 발생하는 오류들이다.

(47) 가. 지금이* 그때보다 한국어* 잘하고 선생님하고 학생도 친구처럼 사겼어요(√지내요). (BW24)
나. 바람이 조금 불지만* 나의 마음이 너무 더웠어요(√따뜻했어요).(BW24)
다. 건강지만(√건강하지만) 맛이 없어요.(CW10)

다음 어휘 오류 중 자주 발생하는 오류로는 용언의 활용 과정에 단어의 기본형을 인지하지 못한 채 어미를 결합하는 오류와 파생이나 합성의 규칙을 잘못 사용하여 한국어에 존재하지 않는 단어를 만들어 내거나 중국어의 한자를 그대로 번역하여 사용하면서 나타난 오류들이 자주 관찰되었다. 아래 (48가)와 (48나)는 '아파요'나 '와요'와 같

은 활용형을 자주 접하면서 연결어미와 결합 시 어간을 제대로 인지하지 못하여 생긴 오류이다. (48다)는 중국어의 '一天'을 그대로 번역하여 사용한 오류이며 (48라)는 '꾸미다'에 '하다'를 결합하여 한국어에 없는 단어를 만들어 낸 오류이다.

(48) 가. 봄에는 비도 작게* 와고(√오고) 날씨도 좋아요.(BW2)
나. 많이 하면 손을도 아팔(√아플) 수 있어요.(BW6)
다. 한 날(√어느 날)에 할머니가 귤을 사 주셨어요.(AW30)
라. 우리는 먼저 GYM3을 꾸며했어요(√꾸몄어요).(CW3)

마지막으로 표현 오류와 기타 오류를 살펴보도록 한다. 기간에 따른 표현 오류와 기타 오류의 표준화 빈도는 다음의 <표 86>과 같다.

<표 86> 문어 자료에 나타난 표현 및 기타 오류의 기간별 빈도

구분		1~6개월	7~12개월	13~18개월	합계
표현	대치	-	0.25	1.09	1.34
	누락	2.76	3.90	3.47	10.13
	첨가	-	0.25	0.60	0.85
	오형태	0.37	0.63	0.79	1.79
기타(어순·총체적)		9.74	6.30	8.13	24.17

표현 오류는 다른 오류에 비해 비교적 낮은 빈도로 나타나고 있었으나 시간이 지남에 따라 전반적으로 증가하고 있었다. 표현 오류 중 가장 많이 발생한 것은 누락 오류였으나 대치, 첨가, 오형태 오류도 꾸준히 상승하고 있었다. 이는 학습자들이 통사 복잡성을 구축하는 과정에서 더 복잡한 의미를 표현하려는 시도를 하지만 다양한 표현을 습득하지 못하면서 나타난 결과로 해석할 수 있다. 표현의 누락 오류 중 가장 빈번하게 나타난 오류는 이유 표현에서 '-기 때문이다'의 누락이었다. '-기 때문이다'의 누락은 모어의 영향이 비교적 큰 것으로 중국어에서 사용되는 이유 표현은 한국어의 '왜냐하면'에 해당하는 부분만 있고 '-기 때문이다'에 해당하는 부분이 없다. 따라서 학습자들은 중국어의 습관에 따라 '-기 때문이다'를 자주 누락하는 것으로 보인다. 이 외에 눈에 띄는 것은 보조 용언의 누락이나 대치 오류였다. 아래의 예문(49가)와 (49나)는 '-기 때문이다'를 누락하여 생긴 오류이고 (49다)와 (49라)는 '-고 싶다'와 '-지 못하다'를 누락하여 발생한 오류이다.

(49) 가. 햄버거랑 마라탕을 좋아하는 이유는 맛있어요(√맛있기 때문이에요).(CW9)
나. 왜냐하면 우리 중국인도 기억하고 도와줬어요(√도와줬기 때문이에요).(BW35)
다. 찾은(√찾고 싶은) 문제가* 가끔 못 찾아요.(AW11)
라. 중국 가야금 10급 시험 전날 나는 밤에 잠을 못했습니다*. 저는 내일* 10급 시험을 통과할까 봐(√통과하지 못할까 봐) 너무 걱정이에요*.(AW24)

기타 오류에는 총체적 오류와 어순 오류 및 문장 성분의 첨가나 누락 오류가 포함되어 있다. 총체적 오류는 1~6개월에 매우 높게 나타났으나 7~12개월에 감소하였고 이후 13~18개월에 다시 증가하였지만 1~6개월에 비해 낮은 빈도를 보이며 약간 개선되는 것을 확인할 수 있었다.

학습자들의 문어 자료에 나타난 오류를 영역별로 정리해 보면, 조사에서는 누락 오류와 대치 오류가 빈번하게 발생하고 선어말 어미는 누락 오류, 연결어미와 전성어미는 대치 오류, 그리고 어휘 역시 대치 오류가 가장 많이 발생하고 있었다. 그리고 이러한 오류들 중 어휘의 대치 오류를 제외하고 모든 오류가 13~18개월에 가장 높은 빈도로 나타나고 있어 학습자들의 정확성 구축에 영향을 주고 있었다. 특히 조사의 누락 오류는 후반으로 갈수록 빠른 속도로 증가하고 있었으며 학습자들은 초반에 잘 사용하던 조사도 후반으로 갈수록 누락하는 경우가 점점 많아지고 있었다. 이는 학습자들이 구어 환경에 노출되는 시간이 길어지면서 자연스럽게 문어에서도 구어의 습관을 그대로 적용하여 나타난 현상으로 추측해볼 수 있다. 이러한 현상은 앞서 연결어미의 사용에서도 언급한 바와 같이 학습자들이 문어와 구어의 사용역을 구분하지 못하여 생긴 결과로 볼 수 있으며, 이들을 위한 한국어 교육에서 사용역에 따른 언어적 특징을 명시적으로 교육할 필요가 있음을 보여주는 부분이다.

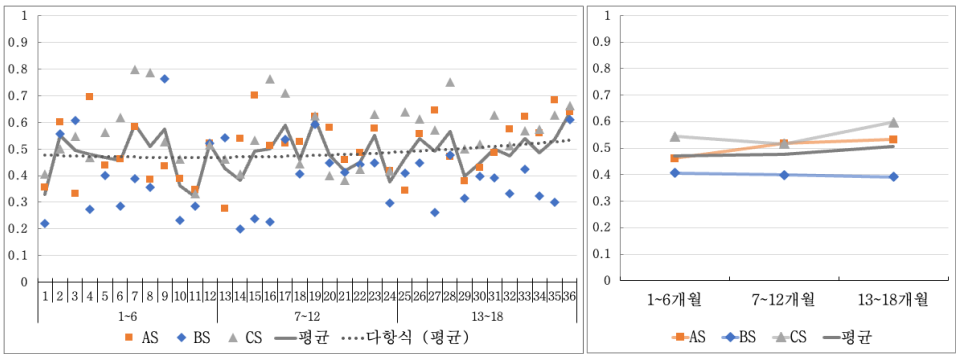
3.2. 구어의 정확성 발달

3.2.1. 무오류절의 비율

구어 자료에 나타난 무오류절 비율의 시점별 변화 추이와 6개월 단위 평균을 살펴보면 <그림 31>과 같다. 세 학습자의 평균 추세선을 통해 시간의 흐름에 따른 무오류절의 변화를 살펴보면 수평에 가까운 곡선이지만 후반으로 갈수록 미세하게 상승하는 추세를 띠고 있었다. 이는 문어 무오류절의 비율이 13~18개월에 하락하는 것과는 대조되는 양상이다. 구어 무오류절의 비율을 6개월 단위로 나누어 살펴보면 7~12개월에는

1~6개월보다 0.01, 13~18개월에는 7~12개월보다 0.03 증가하면서 미미하지만 정확성의 발달이 가속되는 것으로 나타났다.

<그림 31> 시점별 구어 무오류절 비율과 6개월 단위 발달 추이



이러한 무오류절의 기간별 증가가 통계적으로 유의한지 검증하기 위해 일원배치 분산분석(One-Way ANOVA)을 실시하였다. 분석 결과 <표 87>과 같이 학습자들의 기간별 무오류절의 변화는 F=.698, p=.500으로 통계적으로 유의한 차이를 보이지 않았다. 이는 학습자들의 구어 무오류절의 비율은 문어에서와 달리 미세하게 증가하며 약간의 발달을 보이지만 이 역시 통계적으로 유의미할 정도의 변화가 아니라는 것을 의미한다.

<표 87> 기간별 구어 무오류절 비율

측정 지표	기술통계량				F/p	사후검정
	기간(개월)	N	Mean	SD		
무오류절 비율	1~6개월	36(a)	0.47	0.15	.698/.500	-
	7~12개월	36(b)	0.48	0.13		
	13~18개월	36(c)	0.51	0.13		

*p<.05 **p<.01 ***p<.001

학습자별로 살펴보면 세 학습자의 무오류절 비율은 서로 다른 변화 추이를 보이고 있는데 B학습자의 무오류절 비율은 하락하고 있었고 C학습자의 무오류절 비율은 7~12개월에 하락 이후 13~18개월에 다시 증가하였으며 A학습자만이 작은 폭이지만 지속적으로 무오류절의 비율을 높여가고 있었다. 세 학습자의 무오류절 비율의 평균은 C>A>B순으로 나타났는데 이러한 차이가 통계적으로 유의한지 확인하기 위해 일원배

치 분산분석을 실시한 결과 <표 88>과 같이 F=15.066, p=.000으로 학습자 간 유의한 차이가 있는 것으로 나타났다. 사후검정 결과 A, C학습자의 무오류절 비율이 B학습자에 비해 유의하게 높은 것으로 나타났다.

<표 88> 학습자별 구어 무오류절 비율

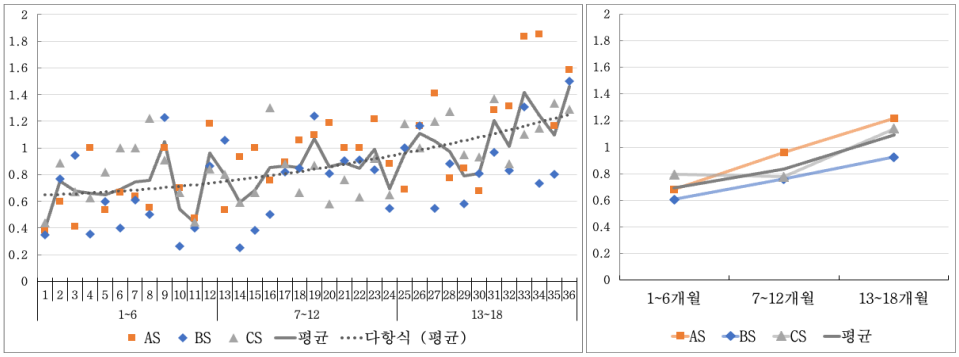
측정 지표	기술통계량				F/p	사후검정
	학습자	N	Mean	SD		
무오류절 비율	AS	36	0.50	0.11	15.066/.000***	CS,AS>BS
	BS	36	0.40	0.13		
	CS	36	0.55	0.12		

*p<.05 **p<.01 ***p<.001

3.2.2. AS-unit당 무오류절 수

학습자들의 구어 AS-unit당 무오류절 수의 시점별 변화 추이와 6개월 단위 평균을 살펴보면 <그림 32>와 같다. 세 학습자의 문장당 무오류절 수는 무오류절 비율과 유사하게 전체적으로 시간의 흐름에 따라 상승세가 가속되는 양상을 보였으며 상승의 폭이 무오류절의 비율에 비해 뚜렷하였다. 6개월을 단위로 AS-unit당 무오류절 수의 변화 폭을 살펴보면 7~12개월에는 1~6개월보다 0.14 증가하였고 13~18개월에는 7~12개월보다 0.26 증가한 것을 확인할 수 있었다.

<그림 32> 시점별 AS-unit당 무오류절 수와 6개월 단위 발달 추이



AS-unit당 무오류절 수의 기간별 변화가 통계적으로 유의한지 검증하기 위해 일원배치 분산분석(One-Way ANOVA)을 실시한 결과 <표 89>와 같이 학습자들의 기간별

무오류절의 변화는 F=19.320, p=.000으로 통계적으로 유의한 것으로 나타났다. 사후검정 결과 1~6개월, 7~12개월과 13~18개월(1~6개월, 7~12개월<13~18개월) 사이에 차이를 보였다. 이는 그래프에서 확인한 것과 마찬가지로 학습자들의 정확성이 13~18개월에 이르러 뚜렷하게 발달함을 의미하는데 문어의 문장당 무오류절 수가 12개월 이후 13~18개월 구간에서 발달이 정체되는 것과 상반되는 현상이다.

<표 89> 기간별 AS-unit당 무오류절 수

측정 지표	기술통계량				F/p	사후검정
	기간(개월)	N	Mean	SD		
AS-unit당 무오류절 수	1~6	36(a)	0.69	0.27	19.320/.000***	a,b<c
	7~12	36(b)	0.83	0.24		
	13~18	36(c)	1.09	0.32		

*p<.05 **p<.01 ***p<.001

학습자별로 살펴보면 A, B학습자는 정확성이 꾸준히 발달하고 C학습자는 7~12개월에 약간의 퇴행을 보이지만 전반적으로 세 학습자 모두 1~6개월에 비해 13~18개월에 무오류절 수가 뚜렷하게 증가하는 양상을 보이고 있었다. 세 학습자의 AS-unit당 무오류절 수의 평균은 A>C>B순으로 나타났는데 학습자별 무오류절의 비율 차이가 통계적으로 유의한지 확인하기 위해 일원배치 분산분석을 실시한 결과 <표 90>에서와 같이 F=3.481, p=.034로 학습자 간에 유의한 차이를 보였다. 사후검정 결과 A학습자와 B학습자(A>B) 간에만 유의한 차이가 있는 것으로 나타났다.

<표 90> 학습자별 AS-unit당 무오류절 수

측정 지표	기술통계량				F/p	사후검정
	학습자	N	Mean	SD		
AS-unit당 무오류절 수	AS	36	0.95	0.37	3.481/.034*	AS>BS
	BS	36	0.76	0.31		
	CS	36	0.90	0.26		

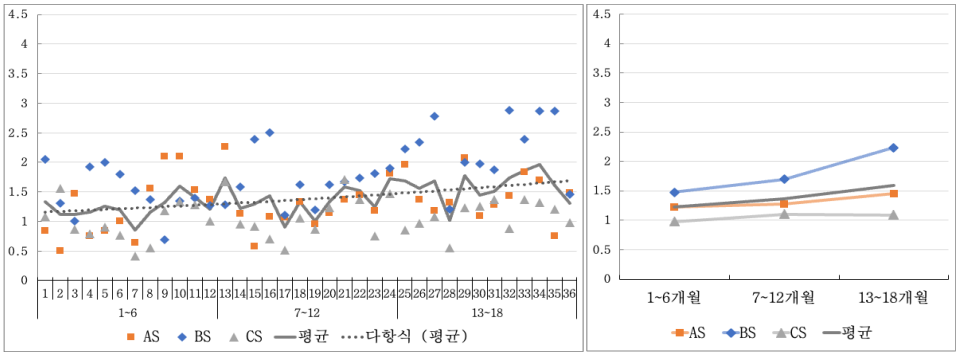
*p<.05 **p<.01 ***p<.001

3.2.3. AS-unit당 오류 수

AS-unit당 오류 수의 시점별 변화 추이와 6개월 단위 평균을 살펴보면 <그림 33>과 같다. 우선 세 학습자의 AS-unit당 오류 수의 평균 2차 다항식 추세선을 살펴보면 직선

에 가까운 상승세를 보이고 있다. 이는 학습자들의 AS-unit당 오류 수가 시간이 지날수록 점점 증가하고 있음을 의미한다. 6개월을 단위로 증가폭을 확인해 보면 7~12개월에는 1~6개월보다 0.13개 증가하고 13~18개월에는 7~12개월보다 0.23개 증가하여 후반으로 갈수록 증가 속도가 더 빨라지는 것을 확인할 수 있다.

<그림 33> 시점별 AS-unit당 오류 수와 6개월 단위 발달 추이



문장당 오류 수의 기간별 변화가 통계적으로 유의한지 검증하기 위해 일원배치 분산분석(One-Way ANOVA)을 실시한 결과 <표 91>과 같이 학습자들의 기간별 무오류절의 변화는 F=4.409, p=.014로 통계적으로 유의한 것으로 나타났다. 사후분석 결과 1~6개월과 13~18개월(1~6개월<13~18개월) 간에만 유의한 차이를 보였다. 이는 학습자들의 오류 수의 증가가 비교적 완만하지만 유의미하게 증가하고 있음을 의미하며 동시에 이는 학습자들의 정확성이 하락함을 의미하기도 한다.

<표 91> 기간별 AS-unit당 오류 수

측정 지표	기술통계량				F/p	사후검정
	기간(개월)	N	Mean	SD		
AS-unit당 오류 수	1~6	36(a)	1.23	0.47	4.409/.014*	a<c
	7~12	36(b)	1.36	0.48		
	13~18	36(c)	1.59	0.63		

*p<.05 **p<.01 ***p<.001

학습자별로 살펴보면 C학습자는 7~12개월에 오류 수가 증가한 뒤 13~18개월에 더 이상 증가하지 않고 있었으나 나머지 두 학습자는 7~12개월과 13~18개월에 모두 오류 수가 증가하며 평균과 같은 추세를 나타내고 있었다. 세 학습자의 AS-unit당 오류

수의 평균은 문어와 마찬가지로 B>A>C순으로 나타났는데 학습자별 오류 수의 차이가 통계적으로 유의한지 일원배치 분산분석을 통해 확인한 결과 <표 92>에서와 같이 F=21.149, p=.000으로 학습자 간에 유의한 차이가 있는 것으로 나타났다. 사후검정 결과 세 학습자 사이에는 모두 유의한 차이(B>A>C)가 있는 것으로 나타났다.

<표 92> 학습자별 AS-unit당 오류 수

측정 지표	기술통계량				F/p	사후검정
	학습자	N	Mean	SD		
AS-unit당 오류 수	AS	36	1.32	0.45	24.149/.000***	BS>AS>CS
	BS	36	1.80	0.56		
	CS	36	1.06	0.32		

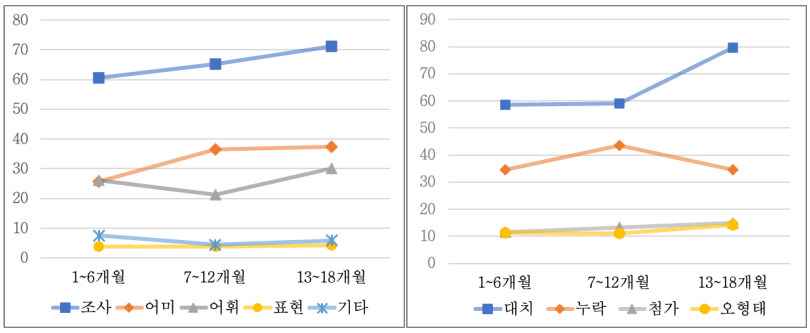
*p<.05 **p<.01 ***p<.001

세 가지 측정 지표를 통해 살펴본 정확성의 발달을 정리하면, 구어 역시 무오류절의 비율을 통해 살펴본 학습자들의 정확성 발달은 문어와 마찬가지로 기간별로 유의한 차이를 보이지 않아 관찰 기간(18개월) 동안 정확성에 변화가 없는 것으로 나타났고 단위당 무오류절 수로 살펴본 학습자들의 정확성은 1~7개월에서 7~12개월 사이에는 뚜렷한 발달이 없으나 7~12개월과 13~18개월 간에 뚜렷한 차이를 보이며 정확성이 발달하는 것으로 나타났다. 마지막으로 AS-unit당 오류 수를 통해 살펴본 정확성의 발달은 인접한 두 기간 사이에는 차이가 없었으나 1~6개월보다 13~18개월의 오류 수가 유의미하게 높은 것으로 확인되며 정확성이 점점 하락하는 것으로 나타났다. 이와 같이 구어의 정확성은 측정 지표에 따라 다른 결과를 보여주고 있었다. 따라서 학습자들의 구어는 시간이 지날수록 무오류절 수도 증가하지만 동시에 오류 수도 지속적으로 증가하고 있다고 정리할 수 있다.

3.2.4. 오류 유형

이 절에서는 학습자들의 구어 자료에 나타난 오류들을 세부 유형으로 나누어 학습자들이 구어에서 어떤 오류들을 범하고 있는지를 살펴보도록 한다. 구어 자료의 오류 유형 역시 문어와 마찬가지로 표준화 빈도를 사용하였으며 6개월을 단위로 묶어 시간의 흐름에 따른 오류의 변화를 분석하도록 한다. 유형별 오류를 자세히 살펴보기 이전에 학습자들의 구어 자료에 나타난 오류들을 오류 영역과 오류 현상에 따라 6개월 단위로 묶어 전체적인 변화 양상을 살펴보면 <그림 34>와 같다.

<그림 34> 구어 자료에 나타난 오류 영역 및 오류 현상의 기간별 빈도



오류 영역별 빈도를 살펴보면 '조사 오류>어미 오류>어휘 오류>기타 오류>표현 오류' 순으로 나타나 학습자들은 구어에서도 조사 사용에서 가장 많은 오류를 범하고 있음을 확인할 수 있었다. 각 영역의 오류 빈도의 변화 추이를 살펴보면 조사 오류는 꾸준히 가파르게 증가하고 있었고 어미 오류는 7~12개월에 급격히 증가한 뒤 13~18개월에는 7~12개월과 비슷한 수준을 유지하고 있었으며 어휘 오류는 어미 오류와 다르게 7~12개월에 하락한 뒤 13~18개월에 다시 증가하고 있었다. 그리고 기타 오류와 표현 오류는 미세하게 하락 이후 다시 증가하고 있었으나 변화폭이 크지 않았다. 특히 구어에서는 어미와 어휘 오류의 증감 추이가 상반되게 나타나고 있는데 이는 문어 자료에서 어미와 어휘 오류의 증감 추이가 비슷하게 나타났던 것과 대조되는 현상이다.

오류 현상별 빈도를 살펴보면 '대치 오류>누락 오류>첨가 오류>오형태 오류' 순으로 나타나고 있었는데 각각의 오류 현상은 시간의 흐름에 따라 서로 다른 변화를 보이고 있었다. 대치 오류는 7~12개월에 미세하게 감소하였다가 13~18개월에 대폭 증가하는 양상을 보였고 누락 오류는 7~12개월에 증가하였다가 13~18개월에 대폭 하락하는 추세를 보이고 있었다. 그리고 첨가 오류와 오형태 오류는 큰 변화를 보이지 않았는데 자세히 살펴보면 첨가 오류는 지속적으로 상승하고 있고 오형태 오류는 7~12개월에 미세하게 하락한 뒤 13~18개월에 다시 1~6개월보다 높은 수준으로 상승하였다. 오류 현상 중에서 문어와 가장 큰 차이를 보인 것은 누락 오류였다. 문어에서는 13~18개월에 누락 오류가 큰 폭으로 증가한 것과 달리 구어 자료에서는 큰 폭으로 하락하며 누락 오류가 감소하고 있었다. 아래 각 영역별로 기간에 따른 오류의 변화 양상을 살펴보도록 한다. 먼저 조사에 나타난 오류 현상별 빈도를 살펴보면 다음의 <표 93>과 같다.

<표 93> 구어 자료에 나타난 조사 오류의 기간별 빈도

구분		1~6개월	7~12개월	13~18개월	합계
조사	대치	27.91	30.43	42.08	100.42
	누락	20.72	25.72	18.36	64.80
	첨가	6.34	4.35	4.92	15.60
	오형태	5.65	4.71	5.90	16.26

앞에서 살펴보았듯이 조사 오류는 오류 영역 중 가장 높은 빈도를 보이는 오류이다. 조사 오류 중 가장 빈번하게 나타난 것은 누락 오류와 대치 오류였는데 누락 오류는 1~6개월보다 7~12개월에 증가하지만 이후 13~18개월에 대폭 감소하는 반면 대치 오류는 그 빈도가 지속적으로 높아지고 있었다. 이는 문어에서 조사 누락 오류가 점점 더 증가하는 것과 대조되는 현상이다. 첨가와 오형태 오류는 대치와 누락 오류에 비해 낮게 나타났는데 두 오류 모두 7~12개월에 감소하였으나 13~18개월에 다시 증가하고 있었다. 조사 오류 중 가장 많이 나타난 오류는 주격 조사와 보조사, 목적격 조사의 오류였다. 오류 현상을 살펴보면 주격 조사와 목적격 조사는 대치 오류가 가장 높았고 보조사는 누락 오류와 대치 오류가 거의 비슷한 빈도로 나타나고 있었다. 그중 주격 조사의 대치 오류는 보조사를 주격 조사로 대치한 오류가 압도적으로 많이 나타나고 있었는데 '내가'와 '나는'의 대치 오류가 빈번하게 관찰되었다. 아래 예문(50가)는 보조사를 주격 조사로 대치한 오류이고 (50나)는 목적격 조사 대치 오류, (50다)는 보조사 누락 오류의 예시이다.

(50) 가. 한달 후에 시험 점수를* 나왔어요. 제가(√저는) 시험 점수를 보고 기분이 너무 좋아요*. (AS26)
나. 그래서 그 친구는 고민을(√고민이) 많이 있어요.(CS22)
다. 이 세상에 후회하는 일은 너무 많아요. 后悔药도 없어요. 그래서 우리(√우리는) 한* 일을 할 때는 어떻게 해야 할지 신중하게 생각해야 해요.(AS23)

'이/가', '을/를', '은/는'의 오류는 조사 오류 중에서도 매우 빈번하게 발생하는 오류로(고석주 외, 2004; 김정숙 외, 2002) 중국어권 고급 학습자의 조사 사용을 고찰한 양희연(2010:15-34)에서는 '이/가'를 '은/는'이나 '을/를'로 대치하는 오류가 가장 높은 것으로 조사되었다. 그러나 선행연구를 살펴보면 이는 중국어권 학습자나 초급 학습자만의 오류는 아님을 알 수 있는데 대학원 학문목적 중·고급 학습자의 쓰기를 분석한 김호정·강남욱(2010b)에서도 비교적 쉬울 것으로 예상되는 '이/가'의 오류가 '은/는'의 오

류에 비해 압도적으로 높게 나타나고 있음을 밝히고 있다. 즉 조사의 대치 오류는 중·고급 단계까지도 지속적으로 나타나는 오류로 지속적인 지도가 필요한 부분임을 알 수 있다. 다음으로, 어미에서 나타난 오류를 살펴보도록 한다. 어미 오류는 어말 어미, 선어말 어미, 연결어미와 전성어미로 나누어 살펴보았으며 그 결과는 아래 <표 94>와 같다.

<표 94> 구어 자료에 나타난 어미 오류의 기간별 빈도

구분			1~6개월	7~12개월	13~18개월	합계
어미	어말어미	대치	5.31	3.99	1.53	10.82
		누락	-	0.24	-	0.24
		첨가	-	-	-	-
		오형태	1.54	1.33	0.22	3.09
	선어말어미	대치	-	0.12	0.44	0.56
		누락	8.90	12.92	11.80	33.63
		첨가	2.23	6.04	5.14	13.40
		오형태	-	-	-	-
	연결어미	대치	3.42	4.59	8.63	16.65
		누락	0.17	0.24	0.22	0.63
		첨가	-	0.24	0.98	1.23
		오형태	0.68	0.24	0.44	1.36
	전성어미	대치	3.08	5.19	6.99	15.27
		누락	0.17	0.85	0.33	1.34
		첨가	-	0.12	0.11	0.23
		오형태	0.17	0.36	0.66	1.19

문어에서와 마찬가지로 구어에서도 어말 어미는 오류 빈도가 비교적 낮은 영역이었으며, 가장 빈번하게 발생한 오류는 같은 회차의 자료에서 '-ㅂ니다'와 '-아/어요', '-아/어'를 섞어서 사용하면서 발생한 대치 오류였다. 그러나 어말 어미는 대치, 누락, 오형태 오류 모두 시간이 흐름에 따라 빠르게 감소하며 자연스럽게 오류가 교정되는 것을 확인할 수 있었다.

선어말 어미 오류는 어미 오류 중에서 가장 높은 빈도를 보이는 영역으로 누락과 첨가 오류가 빈번하게 발생하고 있었다. 선어말 어미 중 학습자들이 주로 범하는 오류는 과거 시제 선어말 어미 '-았/었'의 누락과 첨가에 의한 것이었다. 구어의 선어말 어미 누락과 첨가 오류는 1~6개월보다 7~12개월에 큰 폭으로 증가하였으나 13~18개월에 다시 약간 감소하는 양상을 보였는데 이러한 변화 추이는 문어 자료에서 선어말 어미

누락 오류는 지속적으로 증가하고 첨가 오류는 7~12개월에 감소 이후 13~18개월에 다시 증가하는 것과는 대조되는 현상이다. 다음의 예문(51)은 선어말 어미 '-았/었'의 누락과 첨가 오류의 예시이다.

(51) 가. 고양이가 너무 많았는데 그 고양이가 제일 귀여워(√귀여웠어요).(BS6)
나. 나는 한국에 처음에 왔을 때 한국어 거의 못해요(√못했어요)(CS25)
다. 우리 학교 한 친구는 체육 수업을 했을(√할) 때 목 너무 말라서 물 마시고 싶어요*.(CS24)
라. 그리고 지금은 스트레스를 받았을 때 저는 노래를 듣고 친구하고 맛있는 음식을 먹고 게임을 하고 쇼핑도 했어요(√해요).(AS20)

연결어미와 전성어미 오류는 비슷한 빈도로 발생하고 있었는데 연결어미의 오류가 전성어미에 비해 약간 더 높은 빈도로 나타나고 있었다. 연결어미의 오류 중 가장 빈번하게 발생하는 것은 대치 오류로 시간이 흐를수록 증가하고 있었으며 특히 13~18개월에 빠르게 증가하는 것으로 나타났다. 이는 13~18개월에 빠르게 증가하는 내포절의 사용과 관련이 있을 것으로 보인다. 이 외에 누락 오류는 7~12개월에 증가하였다가 다시 감소하고 첨가 오류는 7~12개월에 처음 나타난 뒤 13~18개월에 빠르게 증가하고 있었다. 그러나 구어에서는 문어와 달리 오형태 오류가 거의 나타나지 않고 거의 대부분의 오류가 대치 오류에 집중되어 있었는데 문법 항목으로 살펴보면 주로 '계기적 사건'의 '-고(서)', '-아(서)/어(서)', '조건·가정'의 '-면'의 대치 오류가 가장 빈번하였다. '-고(서)'와 '-아(서)/어(서)'는 두 항목 간의 대치 오류도 빈번하였으나 다른 문법 항목을 사용해야 할 위치에 '-고(서)'나 '-아(서)/어(서)'를 사용하는 경우가 자주 관찰되었다. 아래 예문(52)는 연결어미의 대치 오류의 예시이다.

(52) 가. 집 가고(√가서) 고양이 만나면 엄청 행복해요.(BS21)
나. 근데 내가 노력하고(√노력하면) 얘보다 잘할 수 있어요.(BS21)
다. 하지만 나는 친구랑 싸워기* 싫어서 대답 안 해서(√하고) 그냥 나 혼자 하고 싶은 거 했어요.(CS36)
라. 나는 우리 집 더 좋아서(√좋다고) 생각했어요.(CS27)

전성어미 역시 대치 오류가 가장 빈번하게 발생하고 있었는데 누락과 첨가 오류는 13~18개월에 가장 낮은 빈도를 보이며 감소하는 추세를 보이는 반면 대치 오류와 오

형태 오류는 점점 증가하는 추세를 나타내고 있었다. 구어의 전성어미 오류 중 가장 높은 빈도로 나타난 오류는 문어와 마찬가지로 '관형사형 전성어미'인 '-ㄴ/는'의 대치 오류였다. 아래의 예문(53)은 '-ㄴ/는'의 대치 오류의 예시이다. 예문에서 확인할 수 있듯이 학습자들은 18개월 차에 이르러서도 '-ㄴ/는'의 동사와 형용사의 현재형을 혼동하거나 현재, 과거, 미래형 관형사형 전성어미를 서로 대치하여 사용하는 오류가 지속적으로 나타나고 있었다.

(53) 가. 저의 공연복은 중국에* 옛날에 입은(√입던) 옷이에요.(AS24)
나. 선생님은 안 오는(√온) 학생들의 의자에 앉아서 오늘* 배운(√배울) 내용을 정리하고 아침밥을 먹어요.(AS34)
다. 제가 좋아한(√좋아하는) 게임은 롤이에요.(BS36)

관형사형 전성어미의 대치 오류는 문어와 구어에서 모두 빈번하게 나타나는 오류로 이는 학습자들의 모어인 중국어에서 그 원인을 찾을 수 있다. 한국어의 관형사형 어미가 '-은(과거)', '-을(미래)', '-은/는'(현재), '-던(회상)으로 관형사형 어미 자체가 시제를 의미하거나 동사나 형용사에 따라 '-ㄴ'과 '-는'이 구분되는 것과 달리 중국어에서는 일반적으로 관형사절과 피수식어인 체언 사이가 시제와 관계 없이 모두 '的'로 연결되며 '的'는 문장에 '내일, 지금, 다음, 얼마 전'과 같은 시간을 표현하는 어휘나 미래 시제를 나타내는 '要', 완료를 나타내는 '了'를 사용함으로써 시제를 표현한다. 따라서 중국인 학습자는 시제나 품사에 따라 다른 형태를 사용하는 것은 익숙하지 쉽지 않다.

연결어미와 전성어미의 오류 양상을 통해 구어에서도 문어와 마찬가지로 문장의 확장과 관련된 문법 항목의 종류나 빈도 측면에서는 발달이 일어나지만 정확한 의미나 형태에 대한 지식 부족으로 난이도가 높은 연결어미나 전성어미를 사용하는 과정에서 오류가 지속적으로 발생하고 있음을 알 수 있었다. 주월랑(2018:86)에서는 자연적 습득 환경에 장기간 거주한 여성결혼이민자의 경우 모어화자와의 접촉으로 인해 문법 사용의 기능적인 측면은 어느 정도 습득할 수 있으나 문법 사용의 통사 및 의미 제약 습득에 한계가 존재한다고 지적하였다. 이와 마찬가지로 중도입국 청소년 학습자들 역시 자연스러운 학교 생활을 통해 다양한 연결어미나 전성어미를 습득할 수는 있으나 추가적인 한국어 교육이 없이는 그 형태나 의미를 바르게 습득하고 사용하는 것에는 한계가 있음을 알 수 있다. 계속하여 어휘 사용에 나타난 오류를 살펴보도록 한다. 어휘 오류의 현상별 빈도 분석 결과는 다음의 <표 95>와 같다.

<표 95> 구어 자료에 나타난 어휘 오류의 기간별 빈도

구분		1~6개월	7~12개월	13~18개월	합계
어휘	대치	18.66	13.77	19.34	51.78
	누락	1.37	1.33	0.87	3.57
	첨가	2.91	1.93	3.17	8.01
	오형태	3.08	4.23	6.67	13.98

어휘 오류 중 가장 높은 빈도로 나타난 것은 어휘의 대치 오류와 오형태 오류였는데 그중 대치 오류의 빈도가 다른 오류에 비해 압도적으로 높았다. 각 현상별로 살펴보면 어휘의 대치와 첨가 오류는 7~12개월에 감소하였다가 13~18개월에 다시 대폭 상승하는 양상을 보이고 오형태 오류는 지속적으로 증가하고 있었으며 누락 오류는 지속적으로 감소하고 있었다.

구어 역시 어휘의 대치 오류(의미 대치, 외국어 대치, 품사 대치, 피사동 대치) 중 가장 높은 빈도로 나타난 것은 어휘의 의미 대치 오류였다. 구어에서도 문어와 마찬가지로 '나'와 '우리'의 대치 오류가 많았지만 추가적으로 교사나 부모님을 '얘' 혹은 '얘네'로 지칭하는 오류가 다수 발생하고 있었다. 이는 학습자들이 주변 친구들의 언어를 통해 어휘를 습득하지만 정확한 의미 기능에 대한 교육이 동반되지 못하였음을 보여준다. 또한 이 외에도 '물건-건물, 아무-아무것, 내일-다음 날, 이제-지금, 작다-적다, 지나다-지내다'와 같이 의미나 형태적으로 유사한 어휘의 대치, '감동-감동적, 화나다-화내다, 좋다-좋아하다, 싫다-싫어하다'와 같이 같은 어근에서 파생된 어휘의 혼동이 빈번하였다. 또한 문어에서는 중국어식 표현이 많았던 반면 구어에서는 '챙겨서 가다(데리고 가다), 과일을 사다(가져오다), 문제가 나다(발생하다), 생각이 있다(들다), 허리가 크다(굵다), 도움이 있다(되다), 시험을 하다(치다), 설날을 하다(보내다)'처럼 구체적인 어휘를 사용해야 하는 곳에 '있다, 하다, 사다'와 같은 일반적인 어휘로 대치하는 오류도 빈번하게 나타나고 있었다. 다음의 예문(54)는 어휘 대치 오류의 예시이다.

(54) 가. 나는 신발하고 옷 샀어요. 그 친구는 돈 없어서 아무도(√아무것도) 안 샀어요(CS16)
나. 하지만 나는 이(√이런) 고민을* 없어요.(CS22)
다. 근데 요즘에 새로운 스트레스* 나왔어요(√생겼어요).(BS20)
라. 그래서 많은 사람들이 고민을 너무 싫어요(√싫어해요).(AS22)

대치 오류 다음으로 빈번하게 발생한 오류는 문어와 마찬가지로 용언이 어미와 결합

하는 과정에서 발생한 활용 오류나 한자어를 어근으로 하는 어휘의 파생이나 합성 규칙의 잘못된 사용으로 인해 한국어에 없는 어휘를 만들어내는 오류였다. 예문(55가)와 (55나)는 '못해요', '귀여워요'와 같은 활용형을 자주 접하면서 어간에 대한 인지가 부족하여 발생한 오류이고 예문(55다)와 (55라)는 '후회(后悔), 자유(自由)'와 같이 어근이 중국어에서 하나의 단어로 사용되는 어휘들에서 나타난 파생 오류이다.

(55) 가. 한국말을 못해면(√못하면) 나쁜 점이 너무 많아요.(AS22)
나. 그 여자는 너무 귀여워고(√귀엽고) 성격도 좋아요.(CS26)
다. 그래서 저는 너무 후회예요(√후회해요).(AS27)
라. 줄도 안 서도 되고 먹는 거가* 자유하고(√자유롭고) 마음이 더 편해요.(BS26)

마지막으로 표현 오류와 기타 오류를 살펴보도록 한다. 표현 오류와 기타 오류의 기간에 따른 표준화 빈도는 다음의 <표 96>과 같다.

<표 96> 구어 자료에 나타난 표현 및 기타 오류의 기간별 빈도

구분		1~6개월	7~12개월	13~18개월	합계
표현	대치	0.17	0.85	0.55	1.56
	누락	3.25	2.17	2.95	8.38
	첨가	0.17	0.60	0.55	1.32
	오형태	0.17	0.12	0.22	0.51
기타(어순·총체적)		7.53	4.47	5.90	17.90

표현 오류는 다른 오류에 비해 비교적 낮은 빈도로 나타나고 있었는데 표현 오류 중 가장 많이 발생한 것은 누락 오류였다. 누락 오류는 1~6개월에 가장 높게 나타나고 7~12개월에 감소하였다가 13~18개월에 다시 약간 증가하고 있었다. 이 외에 대치 오류나 첨가, 오형태 오류도 증가와 감소를 반복하며 전체적으로는 증가하는 양상을 보이고 있었다. 구어 역시 '-기 때문이다'의 누락과 보조 용언의 결합 과정에서 나타난 오류가 가장 빈번하였다. 표현의 누락 오류의 예시는 다음의 예문(56)과 같다.

(56) 가. 왜냐면 내가* 좋은 학생입니다(√학생이기 때문이에요).(BS17)
나. 엄마는 들으면 피곤해요(√피곤해 해요).(AS4)
다. 나* 엄마 나* 샴페인 두 개 샀어요(√사 줬어요).(CS11)
라. 햄버거랑 치킨 간식* 같은 음식을 많이 먹으면 너무 뚱뚱해요(√뚱뚱해져요).(CS12)

학습자들의 구어 자료에 나타난 오류를 영역별로 정리해 보면, 조사에서는 대치 오류와 누락 오류가 빈번하게 발생하고 선어말 어미는 누락 오류, 연결어미와 전성어미는 대치 오류, 그리고 어휘 역시 대치 오류가 가장 많이 발생하고 있었다. 그리고 이러한 오류들 중에서 어휘 대치 오류는 7~12개월에 감소하였다가 13~18개월에 다시 증가하고 있었고 조사의 누락 오류와 선어말어미의 누락 오류는 7~12개월에는 증가하였으나 13~18개월에 다시 감소하였다. 그리고 세 기간에 거쳐 지속적으로 증가하는 오류는 조사 대치 오류와 연결어미, 전성어미의 대치 오류였다. 즉 구어에서 학습자들의 누락 오류는 증가하였다가 다시 감소하기도 하지만 대치 오류는 꾸준히 증가하고 있었다. 이러한 결과를 앞에서 살펴본 문어와 비교해볼 때, 문어와 구어에서 모두 증가하는 오류는 조사 대치, 연결어미 전성어미 대치 오류이고 구어에서 자연스럽게 감소하는 오류는 조사의 누락과 선어말 어미의 누락 오류이며 문어에서 자연스럽게 감소하는 오류는 어휘 대치 오류라고 정리할 수 있다.

3.3. 산출 양식에 따른 정확성 비교

2장에서 살펴보았듯이 쓰기와 말하기의 큰 차이점 중 하나는 '실시간성'에서 나타난다. 그리고 이러한 실시간성은 모니터링과 교정 과정에 직접적인 영향을 미치게 되어 정확성을 떨어뜨린다고 보았다. 따라서 본 절에서는 문어와 구어에 나타난 정확성을 비교하여 앞에서 살펴본 이론들을 검토하고 실제로 학습자들에게서 나타나는 오류의 유형이 문어와 구어에서 어떻게 다르게 나타나는지를 살펴보도록 한다.

3.3.1. 무오류절의 비율 비교

전체 기간에 나타난 무오류절의 비율을 통해 문어와 구어의 정확성을 비교해 보면 문어의 무오류절 비율은 0.48, 구어의 무오류절 비율은 0.49로 구어의 무오류절 비율이 미세하게 더 높은 것을 확인할 수 있었다. 이를 다시 기간별로 나누어 살펴보면 1~6개월에는 구어의 무오류절 비율이 더 높았다가 7~12개월에는 문어의 무오류절 비율이 더 높아지고 13~18개월에는 다시 구어의 무오류절이 더 높아지며 일관된 차이를 보이지 않고 있었다. 그러나 그 차이가 매우 미미하여 문어와 구어가 거의 동일한 수준에 머물러 있었다. 이는 통계적 검증에서도 드러나는데 문어와 구어의 무오류절 비율에 나타난 차이는 <표 97>에서와 같이 전체 기간과 6개월 단위의 세 기간 모두 통계적으로 유의하지 않았다. 즉 무오류절의 비율을 통한 정확성에서는 문어와 구어의 정확

성에 큰 차이가 없는 것을 알 수 있었다.

<표 97> 기간별 문어와 구어의 무오류절 비율

측정 지표	기간	N	문어		구어		t/p
			Mean	SD	Mean	SD	
무오류절 비율	1~6개월	36	0.46	0.16	0.47	0.15	-.309/.759
	7~12개월	36	0.50	0.12	0.48	0.13	.998/.325
	13~18개월	36	0.47	0.11	0.51	0.13	-1.279/.209
	전체 기간	108	0.48	0.13	0.49	0.14	-.395/.694

*p<.05 **p<.01 ***p<.001

계속하여 각 학습자의 문어와 구어의 무오류절 비율을 비교해 보면 다음의 <표 98>과 같다. 표에서 확인할 수 있듯이 A학습자는 문어 무오류절의 비율이 더 높고 B학습자는 무어와 구어의 무오류절 비율이 동일하며 C학습자는 구어의 무오류절이 더 높게 나타났으나 세 학습자 모두 차이가 크지 않았으며 통계적으로 유의미한 차이를 보이지 않았다. 따라서, 무오류절의 비율은 산출 양식에 따른 차이와 학습자 간의 차이 모두 뚜렷하지 않다는 결과를 확인할 수 있었다.

<표 98> 학습자별 문어와 구어의 무오류절 비율

측정 지표	기술통계량				t/p	r/p
	학습자	N	Mean	SD		
무오류절 비율	AW	36	0.52	0.09	.497/.622	.067/.700
	AS	36	0.50	0.11		
	BW	36	0.40	0.12	-.094/.925	-.101/.558
	BS	36	0.40	0.13		
	CW	36	0.52	0.14	-.995/.326	.192/.262
	CS	36	0.55	0.12		

*p<.05 **p<.01 ***p<.001

3.3.2. 단위당 무오류절 수 비교

전체 기간에 나타난 문어와 구어의 단위당 무오류절 수를 비교해 보면 문장당 무오류절 수의 평균은 0.95, AS-unit당 무오류절 수의 평균은 0.87로 문어의 정확성이 약간 더 높은 것으로 나타났다. 기간별로 나누어 살펴보면 문어와 구어의 단위당 무오류절 수는 모두 시간이 지남에 따라 증가하고 있지만 1~6개월과 7~12개월에는 문장당 무

오류절 수가 더 높았다가 13~18개월에는 AS-unit당 무오류절 수가 더 높은 것을 확인할 수 있다. 이는 앞에서 살펴본 단위 길이와 같은 현상으로 구어의 정확성이 더 늦게 나타나지만 더 빨리 발달한다고 해석할 수 있다.

문어와 구어의 단위당 무오류절 수의 차이가 통계적으로 유의한지 확인하기 위해 t 검정을 실시한 결과 <표 99>에서와 같이 문어와 구어의 단위당 무오류절 수는 7~12개월에서만 t=2.447, p=.020으로 통계적으로 유의하며 기타 기간이나 전체 기간에서는 유의하지 않은 것으로 나타났다. 이는 단위당 무오류절 수를 통한 정확성 측정에서는 문어와 구어의 정확성에 큰 차이가 없음을 의미한다.

<표 99> 기간별 문어와 구어의 단위당 무오류절 수

측정 지표	기간	N	문어		구어		t/p
			Mean	SD	Mean	SD	
단위당 무오류절 수	1~6개월	36	0.78	0.35	0.69	0.27	1.306/.200
	7~12개월	36	0.99	0.33	0.83	0.24	2.447/.020*
	13~18개월	36	1.07	0.31	1.09	0.32	-.338/.738
	전체 기간	108	0.95	0.35	0.87	0.32	1.826/.071

*p<.05 **p<.01 ***p<.001

각 학습자의 문어와 구어의 단위당 무오류절 수의 차이를 비교해 보면 다음의 <표 100>과 같다. 표에서 확인할 수 있듯이 세 학습자 모두 문장당 무오류절 수가 더 높은 것으로 확인되었지만 그 차이가 통계적으로 유의하지는 않았다.

<표 100> 학습자별 문어와 구어의 단위당 무오류절 수

측정 지표	기술통계량				t/p	r/p
	학습자	N	Mean	SD		
단위당 무오류절 수	AW	36	0.97	0.31	.357/.724	.410/.013*
	AS	36	0.95	0.37		
	BW	36	0.90	0.34	1.870/.070	.171/.320
	BS	36	0.76	0.31		
	CW	36	0.96	0.40	.856/.398	.191/.264
	CS	36	0.90	0.26		

*p<.05 **p<.01 ***p<.001

3.3.3. 단위당 오류 수 비교

전체 기간에 나타난 문어와 구어의 단위당 오류 수 비교해 보면 문장당 오류 수의 평균은 1.43개, AS-unit당 오류 수의 평균은 1.39개로 문장당 오류 수가 약간 더 높은 것으로 나타났다. 이는 문어의 정확성이 구어에 비해 더 낮음을 의미하는 것으로 앞에서 살펴본 단위당 무오류절 수를 통한 정확성 측정과 상반되는 결과이다. 기간별로 나누어 살펴보면 1~6개월에는 구어의 오류 수가 더 낮고 7~12개월에는 문어의 오류 수가 더 낮다가 13~18개월에는 다시 구어의 오류 수가 더 낮게 나타나며 복잡한 관계를 보이고 있었다. 이러한 차이가 통계적으로 유의한지 확인하기 위해 t검정을 실시한 결과 <표 101>에서와 같이 전체 기간 및 6개월 단위의 기간 모두에서 문어와 구어의 단위당 오류 수의 차이가 통계적으로 유의하지 않은 것으로 나타났다. 이는 무오류절의 비율을 통한 측정, 단위당 무오류절 수를 통한 측정과 같은 결과로 세 가지 지표를 통한 측정 모두에서 문어와 구어의 정확성에 큰 차이가 없다는 결론을 내릴 수 있다.

<표 101> 기간별 문어와 구어의 단위당 오류 수

측정 지표	기간	N	문어		구어		t/p
			Mean	SD	Mean	SD	
단위당 오류 수	1~6개월	36	1.24	0.56	1.23	0.47	.163/.872
	7~12개월	36	1.35	0.53	1.36	0.48	-.131/.897
	13~18개월	36	1.71	0.68	1.59	0.63	1.006/.321
	전체 기간	108	1.43	0.62	1.39	0.55	.679/.499

$^{*}p<.05$ $^{**}p<.01$ $^{***}p<.001$

각 학습자의 문어와 구어의 단위당 오류 수의 차이를 비교해 보면 다음의 <표 102>와 같다. 표에서 확인할 수 있듯이 A학습자는 구어의 오류 수가 더 높고 B학습자와 C학습자는 문어의 오류 수가 더 높았다. 이러한 차이가 통계적으로 유의미한 지 검증하기 위해 t검정을 실시한 결과 A학습자(t=-2.383, p=.023)의 오류 수만이 문어와 구어 사이에 유의한 차이가 있는 것으로 나타났다.

<표 102> 기간별 문어와 구어의 단위당 오류 수

측정 지표	기술통계량				t/p	r/p
	학습자	N	Mean	SD		
단위당 오류 수	AW	36	1.12	.35	-2.383/.023*	.255/.133
	AS	36	1.32	.45		
	BW	36	2.04	.60	1.807/.079	.117/.497
	BS	36	1.80	.56		
	CW	36	1.14	.37	1.057/.298	-.018/.918
	CS	36	1.06	.32		

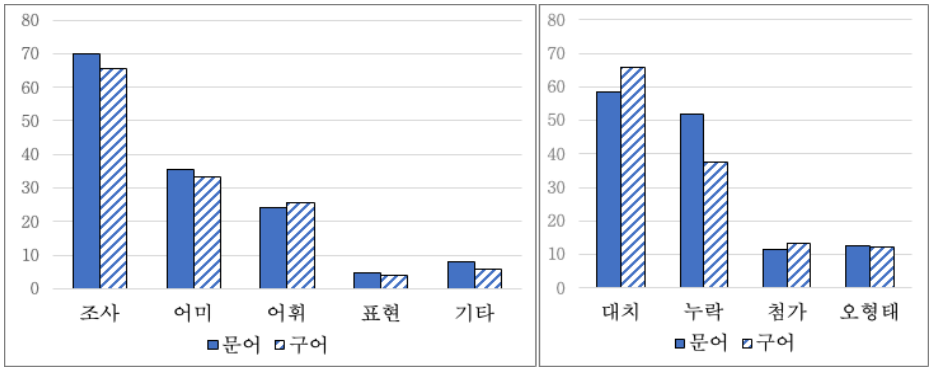
*p<.05 **p<.01 ***p<.001

정리하면, 세 가지 측정 지표를 통한 문어와 구어의 정확성을 비교한 결과, A학습자의 단위당 오류 수를 제외한 모든 측정에서 문어와 구어의 정확성은 모두 유의한 차이를 보이지 않았다. 이는 문어의 산출 과정에서 모니터링과 교정이 더 잘 이루어진다는 점에서 문어의 정확성이 더 높을 것이라는 예상과는 다른 결론으로 문어의 정확성이 더 높다고 한 Ellis & Yuan(2005)이나 구어의 정확성이 더 높다는 결과를 얻은 Kuiken & Vedder(2011), Granfeldt(2008)와도 다른 결과이다. 이러한 결과는 본고에서 문어와 구어 모두 충분한 준비 시간을 주었다는 점에서 나타난 결과일 수도 있고 또한 학습자들이 평소에 쓰기보다 말하기를 더 많이 사용하는 것과 관련이 있을 수도 있다. 그러나 앞서 문어와 구어의 정확성 발달에서도 나타났듯이 문어와 구어의 오류는 구체적인 오류 영역이나 오류 현상에 따라 차이가 존재한다. 아래에서는, 각 오류 영역과 오류 현상별 비교를 통하여 문어와 구어의 오류에 어떤 공통점과 차이점이 있는지 더 자세히 살펴보도록 한다.

3.3.4. 오류의 유형 비교

이 부분에서는 학습자들의 문어와 구어에 나타난 오류들을 영역별로 나누어 빈도를 비교하고 각 영역에서 오류가 빈번하게 나타나는 항목들을 비교해 봄으로써 문어와 구어에서 나타난 오류가 어떤 공통점과 차이점이 있는지 살펴보도록 한다. 비교의 편의를 위해 오류 항목의 빈도는 모두 표준화 빈도를 사용하였다. 우선 문어와 구어의 영역에 따른 오류 빈도와 오류 현상에 따른 빈도를 비교해 보면 다음의 <그림 35>와 같다.

<그림 35> 산출 양식에 따른 영역별 및 현상별 오류 빈도



문어와 구어의 영역별 오류를 비교해 보면, 어휘 오류를 제외한 모든 오류가 문어에서 더 빈번하게 나타나고 있었다. 이는 학습자들이 구어에서 문법적으로 더 정확한 언어를 사용하고 문어에서 더 정확한 어휘를 사용한다는 것을 의미한다. 오류 현상을 비교해 보면 대치 오류와 첨가 오류는 구어에서 더 높게 나타나고 누락과 오형태 오류는 문어에서 더 빈번하게 발생하고 있었다. 이는 산출 양식이 오류의 영역과 오류 현상에 영향을 줄 수 있음을 의미한다. 이어서 각 영역별로 나누어 살펴보도록 한다. 먼저 문어와 구어 자료에 나타난 조사 오류를 비교해 보면 다음의 <표 103>과 같다.

<표 103> 산출 양식에 따른 조사 오류 빈도

구분		문어 표준화 빈도	총 빈도(비율)	구어 표준화 빈도	총 빈도(비율)
조사	대치	29.16	72.34 (49.4%)	34.38	66.44 (48.8%)
	누락	35.42		21.57	
	첨가	4.86		5.07	
	오형태	2.90		5.41	

조사 오류는 문어와 구어에서 모두 가장 빈번하게 나타난 오류로 문어에서는 49.4%, 구어에서는 48.7%의 비율을 차지하며 문어와 구어에서 모두 전체 오류의 50%에 가까운 비율을 차지하고 있었다. 또한 조사 오류는 산출 양식에 따라 오류의 현상에 차이가 있었는데 문어에서는 누락 오류가 대치 오류에 비해 높은 것에 반해 구어에서는대치 오류가 누락 오류보다 높게 나타났고 문어에서는 첨가 오류가 오형태 오류보다 높지만 구어에서는 반대로 나타나고 있었다. 산출 양식에 따른 오류 현상의 빈도를 비교해 보면 조사의 누락 오류를 제외한 모든 오류가 구어에서 더 높은 빈도로 나타나고 있었다. 이는 구어는 실시간 언어 산출이 진행된다는 점에서 조사의 의미나 형태에 주의를 기

울이지 못한 결과로 해석할 수 있다. 이 외에 누락 오류가 구어에서 더 낮게 나타난 것은 한국어는 구어에서 조사 생략이 비교적 자유롭다는 점에 기인한 것으로 볼 수 있다.

문어와 구어에 높은 빈도로 나타난 조사 오류를 살펴보면 문어에서는 '주격 조사 대치>목적격 조사 누락>주격 조사 누락'의 순으로 나타났고 구어에서는 '주격 조사 대치>목적격 조사 대치>보조사 누락'의 순으로 나타났다. 즉 문어와 구어 모두 주격 조사의 대치 오류가 가장 많이 발생한다. 그리고 문어에서는 목적격 조사의 누락 오류가 높은 반면 구어에서는 목적격 조사의 대치 오류가 빈번히 발생하고 문어는 보조사 누락 오류보다 주격 조사 누락 오류가 높지만 구어에서는 보조사 누락 오류가 더 빈번히 나타났다. 특히 조사 누락 오류 중에서 문어에 비해 구어의 목적격 조사 누락이 높지 않은 것을 확인할 수 있는데 이는 목적격 조사의 경우 동사가 목적어를 취할 수 있는지 없는지의 여부만 판단할 수 있으면 쓰는 데 어려움이 없으며 생략한다고 해도 원어민들은 대체로 거슬릴 정도의 오류문으로 인지하지 않기 때문이다. 다음, 문어와 구어 자료에 나타난 어미 오류의 빈도와 비율을 비교해 보았다. 그 결과는 <표 104>와 같다.

<표 104> 산출 양식에 따른 어미 오류 빈도

구분			문어 표준화 빈도	총 빈도(비율)	구어 표준화 빈도	총 빈도(비율)
어미	어말 어미	대치	3.96	5.63 (3.8%)	3.35	4.38 (3.2%)
		누락	0.17		0.09	
		첨가	-		-	
		오형태	1.49		0.95	
	선어말 어미	대치	0.04	17.09 (11.7%)	0.21	16.42 (12.0%)
		누락	13.43		11.47	
		첨가	3.62		4.73	
		오형태	-		-	
	연결 어미	대치	5.12	7.46 (5.1%)	5.89	7.00 (5.1%)
		누락	0.34		0.21	
		첨가	0.30		0.47	
		오형태	1.71		0.43	
	전성 어미	대치	5.03	6.73 (4.6%)	5.37	6.36 (4.7%)
		누락	0.85		0.47	
		첨가	0.09		0.09	
		오형태	0.77		0.43	

어미별 오류 현상을 살펴보기 이전에 영역별 오류 비율을 살펴보면 문어와 구어에서 모두 '선어말 어미 오류>연결어미 오류>전성어미 오류>어말 어미 오류' 순으로 나타나 학습자들이 선어말 어미의 사용에서 가장 많은 오류를 범하고 어말 어미를 가장 정

확하게 사용할 수 있음을 알 수 있었다. 산출 양식에 따른 빈도를 비교해 보면 모든 어미의 오류 빈도가 문어에서 더 높은 것을 확인할 수 있는데 이는 조사의 사용에서 누락 오류를 제외하고 문어의 오류가 더 낮았던 것과 대조되는 현상이다. 어미는 문장 확장과 밀접하게 연결되는 문법 요소라는 점과 문어의 단위당 절 수가 더 높았던 점을 고려해볼 때 이러한 결과가 나타난 원인은 문어에서 더 많은 문장 확장이 이루어진 것과 관련이 있을 것이다.

어미별로 나타난 오류를 살펴보면 어말 어미는 문어와 구어 모두 대치 오류와 오형태 오류가 주로 발생하고 있었으며 모든 오류 현상에 있어 문어의 오류 빈도가 구어에 비해 높게 나타나고 있었다. 이는 학습자들이 구어에서 더 정확한 어말 어미를 사용하고 있음을 의미한다. 선어말 어미에서는 문어와 구어 모두 누락 오류와 첨가 오류가 주로 나타나고 있었으며 대치 오류는 거의 발생하지 않고 있었다. 산출 양식에 따라 비교해 보면 선어말 어미의 대치 오류, 첨가 오류는 구어에서 더 높게 나타나고 누락 오류는 문어에서 더 높게 나타나고 있었는데 선어말 어미의 대치 오류는 모두 선어말 어미 '-았/었-'을 '-겠-'으로 대치하여 사용한 오류였고 첨가와 누락 오류는 거의 대부분이 '-았/었-'으로 인한 오류였다.

산출 양식에 따른 연결어미의 오류를 비교해 보면 문어와 구어 모두 대치 오류가 가장 높게 나타나고 있었다. 그러나 문어에서는 비교적 높은 빈도로 발생하고 있는 오형태 오류가 구어에서는 매우 낮은 빈도를 보이고 있었는데 이는 문어에서 형태적으로 부정확한 연결어미를 더 많이 산출하고 있음을 보여준다. 그러나 이는 문어에서 더 많고 다양한 어미를 사용하는 것과도 관련이 있을 것이다. 문어와 구어를 비교해 보면 누락 오류와 오형태 오류는 문어에서 더 높게 나타나고 대치 오류와 첨가 오류는 구어에서 더 빈번하게 발생하고 있었다. 구체적인 문법 항목을 비교해 보면 문어에서 가장 많이 발생한 연결어미 오류는 '계기적 사건'의 '-고(서)' 대치 오류, '조건·가정'의 '-면'의 오형태 오류, '계기적 사건'의 '-아(서)/어(서)'의 대치 오류였고 구어에서 가장 많이 발생한 연결어미 오류는 '계기적 사건'의 '-고(서)' 대치 오류, '계기적 사건'의 '-아(서)/어(서)'의 대치 오류, '조건·가정'의 '-면'의 대치 오류였다.

전성어미 역시 연결어미와 마찬가지로 문어와 구어 모두 주로 대치 오류에 집중되어 있었는데 누락 오류와 오형태 오류는 문어에서 더 높았고 대치 오류는 구어에서 더 높게 나타났다. 즉 연결어미와 전성어미 모두 구어에서 대치 오류가 더 많이 발생하고 문어에서 누락과 오형태 오류가 더 많이 나타나는 것이다. 전성어미 중 문어 자료에서 가장 많이 나타난 오류는 '관형사형 전성어미'인 '-ㄴ/는'의 대치 오류, '명사형 전성어미'

인 '-기'의 누락 오류, '관형사형 전성어미+ 것'인 '-는 것'의 대치 오류였고 구어 자료에서 가장 많이 나타난 오류는 '관형사형 전성어미'인 '-ㄴ/는'의 대치 오류, '관형사형 전성어미+ 것'인 '-는 것'의 대치 오류, '관형사형 전성어미'인 '-ㄹ'의 대치 오류였다. 계속하여 문어와 구어 자료에 나타난 어휘 오류의 빈도와 비율을 비교해 보면 다음의 <표 105>와 같다.

<표 105> 산출 양식에 따른 어휘 오류 빈도

구분		문어 표준화 빈도	총 빈도(비율)	구어 표준화 빈도	총 빈도(비율)
어휘	대치	15.30	24.17 (16.5%)	17.19	25.91 (19.0%)
	누락	1.11		1.16	
	첨가	2.64		2.66	
	오형태	5.12		4.90	

어휘 오류의 비율을 살펴보면 문어에서 16.8%, 구어에서 19.0%로 구어에서 더 많은 어휘 오류를 범하고 있음을 알 수 있었다. 어휘 오류는 문어와 구어 모두 대치 오류가 가장 높았고 오형태 오류가 그 뒤를 잇고 있었다. 산출 양식에 따른 오류 현상의 빈도를 비교해 보면 오형태 오류를 제외하고 모든 오류 현상이 구어에서 더 높은 빈도로 나타나고 있었다. 어휘의 대치 오류를 의미 대치, 외국어 대치, 품사 대치, 피사동 대치로 나누고 어휘의 오형태 오류를 활용 오류와 파생합성 오류로 나누어 가장 많이 발생한 어휘 오류를 살펴보면 문어와 구어 모두 '의미 대치 오류>용언 활용 오형태 오류>어휘 첨가 오류' 순으로 나타나고 있었다.

<그림 36> 산출 양식에 따른 표현 및 기타 오류 빈도

구분		문어 표준화 빈도	총 빈도(비율)	구어 표준화 빈도	총 빈도(비율)
표현	대치	0.47	4.99 (3.4%)	0.81	4.43 (3.3%)
	누락	3.50		2.90	
	첨가	0.43		0.38	
	오형태	0.60		0.34	
기타(어순·총체적)		8.06	8.06(5.5%)	5.75	5.75(4.3%)

표현 오류의 비율을 비교해 보면 문어에서 3.4%, 구어에서 3.3%로 문어와 구어에서 비슷한 수준으로 발생하고 있었다. 또한 문어와 구어에서 모두 누락 오류가 가장 높은 빈도로 발생하고 있었는데 이는 앞서 문어와 구어의 정확성 발달 양상에서 살펴본 바와 같다. 오류 현상을 비교해 보면 누락과 오형태 오류는 문어에서 더 높은 빈도로 나

타나고 첨가 오류는 구어에서 더 높은 빈도로 나타나며 대치 오류는 문어와 구어에서 비슷한 수준으로 발생하고 있었는데 이는 연결어미나 전성어미와 같은 현상으로 문어에서 더 복잡한 문장을 산출하려는 시도를 더 많이 하는 것에서 비롯된 것으로 볼 수 있다.

문어와 구어의 오류 유형별 비교를 정리해 보면, 비율적으로는 문어에서 조사 오류, 어말 어미 오류, 선어말 어미의 오류가 더 많이 발생하고 구어에서는 어휘 오류가 더 많이 발생하며 연결어미, 전성어미, 표현 오류는 문어와 구어에서 비슷한 비율로 발생하고 있었다. 그리고 빈도를 비교해 볼 때, 어휘 오류를 제외한 모든 오류가 문어에서 더 높게 발생하고 있어 문법적 정확성은 구어에서 더 높지만 어휘적 정확성은 문어에서 더 높다는 결론을 내릴 수 있었다. 특히 구어의 어휘 오류 중 눈에 띄게 높은 것은 어휘 대치 오류였는데 여러 대치 오류 중에서도 의미 대치만이 문어에 비해 높은 것으로 나타났다.[73] 이러한 결과가 나타난 원인은 앞서 오류 유형에서 살펴본 바와 같이 구어에서는 '있다, 하다, 사다'와 같이 단순한 어휘로 어려운 어휘를 대치하는 오류가 문어에 비해 더 빈번하게 발생하는 것과 관련이 있는 것으로 추측할 수 있다.

4. 복잡성과 정확성의 관계

제2언어습득 분야에서 복잡성과 정확성이 서로 연결되어 영향을 주며 발달한다는 점은 널리 받아들여지고 있으나 구체적으로 복잡성과 정확성이 서로 지지적인 관계에 있는지 경쟁적인 관계에 있는지에 대해서는 아직 합의가 이루어지지 않았으며 여러 다양한 주장들이 존재한다. 이에 본 연구에서도 복잡성과 정확성의 상호 관련성을 살펴본 뒤 선행 연구들의 결과와 비교하여 중국 배경 중도입국 청소년의 한국어 학습에서 복잡성과 정확성은 어떤 관계로 발달하는지를 살펴보고자 한다.

복잡성과 정확성의 관련성을 탐구함에 있어서, 전체 자료를 바탕으로 복잡성과 정확성의 관련성을 살펴보는 것도 중요하지만 학습자의 언어는 어떤 상태에 머물러 있는 것이 아니라 항상 역동적으로 변화하는 것이며(황종배, 2004:44) 본 연구에서 수집한 자료 역시 종적 자료라는 점에서 시간의 흐름에 따라 변화하는 복잡성과 정확성의 관

[73] 문어와 구어의 종류별 어휘 대치 오류의 표준화 빈도를 살펴보면 다음과 같다.

구분	의미 대치	외국어 대치	품사 대치	피사동 대치
문어	12.11	2.09	0.90	0.21
구어	15.47	1.12	0.60	0.00

계를 살펴보는 것 또한 학습자들의 발달을 살피는 데 도움이 될 것이다. 따라서 이 부분에서는 먼저 전체 자료를 대상으로 복잡성과 정확성의 관계를 살핀 이후 6개월을 단위로 기간별 복잡성과 정확성을 살펴봄으로써 시간의 흐름에 따라 변화하는 복잡성과 정확성의 관계를 살펴보도록 한다.

4.1. 문어에 나타난 복잡성과 정확성의 관계

문어 자료에 나타난 복잡성과 정확성의 각 측정지표에 대해 상관분석(Pearson correlation coefficient)을 실시한 결과 다음의 <표 106>과 같다. 문장 길이와 문장당 절 수는 높은 상관관계를 가지며 거의 비슷하게 발달하기 때문에 통사 복잡성은 문장 길이만 비교하도록 한다. 그리고 문어의 무오류절 비율은 어휘 밀도와의 약한 부적 상관관계를 제외하고 나머지 지표와 유의미한 상관관계를 보이지 않았기 때문에 표로 제시하지 않았다.

<표 106> 문어 복잡성과 정확성의 관계

구분		N	문장당 오류 수(r/p)	문장당 무오류절 수(r/p)
통사 복잡성	문장 길이	108	.566/.000***	.617/.000***
어휘 복잡성	어휘 다양도	108	.380/.000***	.164/.090
	어휘 밀도	108	.188/.051	-.083/.393
	어휘 정교도	108	.121/.212	-.073/.453

*p<.05 **p<.01 ***p<.001

통사 복잡성과 정확성을 살펴보면 문장당 오류 수는 문장 길이와 r=.566, p=.000으로 뚜렷한 정적 상관관계를 나타내고 있었다. 이는 통사 복잡성이 증가하면 문장당 오류 수가 증가함을 의미한다. 한편, 문장당 무오류절 수 역시 문장 길이와 r=.617, p=.000으로 뚜렷한 정적 상관을 나타내고 있었는데 이는 통사 복잡성이 증가할 수록 문장당 무오류절 수 역시 증가함을 의미한다.

즉 통사 복잡성과 정확성의 관계는 측정 지표에 따라 다른 관계를 나타내고 있었는데 통사 복잡성이 증가하면 오류 수도 증가하지만 무오류절 수 역시 동시에 증가하는 현상을 발견할 수 있었다. 어휘 복잡성과 정확성의 관계를 살펴보면, 문장당 오류 수는 어휘 다양도와 r=.380, p=.000으로 유의한 정적 상관관계를 나타내고 있지만 어휘 밀도, 어휘 정교도와는 유의한 관계를 나타내지 않았다. 이는 어휘 다양도가 증가할 수록 오류 수 역시 증가함을 의미하는 것으로 문어에서 어휘 다양도와 정확성은 서로 경쟁

관계에 있으며 동시에 발달하기 어려움을 의미한다.

즉 학습자들은 쓰기에서 통사 복잡성이 증가하면 문장당 오류 수가 증가하지만 동시에 문장당 무오류절 수 또한 증가하는 현상이 나타나 통사 복잡성과 정확성의 관계가 측정 지표에 따라 다르게 나타나고 있어 일관된 경쟁 혹은 지지 관계를 확인하기 어려웠다. 그러나 어휘 복잡성은 문장당 무오류절 수와는 유의미한 관계를 보이지 않았고 어휘 다양도만이 문장당 오류 수와 정적 상관관계를 보이고 있었다. 이어서 앞에서 살펴본 복잡성과 정확성을 다시 기간별로 나누어 살펴보아 시간의 흐름에 따라 복잡성과 정확성의 관계에 변화가 있는지 살펴보도록 한다. 문장당 오류 수와 복잡성 간의 기간에 따른 상관관계의 분석 결과는 다음의 <표 107>과 같다.

<표 107> 기간별 복잡성과 정확성(문장당 오류 수)의 관계

구분		N	문장당 오류 수(r/p)		
			1~6개월	7~12개월	13~18개월
통사 복잡성	문장 길이	36	.390/.019*	.492/.002**	.589/.000***
어휘 복잡성	어휘 다양도	36	.363/.030*	.321/.056	.208/.224
	어휘 밀도	36	.176/.304	.201/.240	.228/180
	어휘 정교도	36	-.070/.683	.140/.416	.173/.313

*p<.05 **p<.01 ***p<.001

표에서 확인할 수 있듯이 문장당 오류 수와 문장 길이는 세 기간 모두 유의한 정적 상관관계를 나타내고 있었다. 이는 문장 길이가 증가할수록 오류가 증가함을 의미하는 것으로 학습자들의 통사 복잡성과 정확성이 동시에 발달하지 못하고 있음을 보여준다. 또한 이러한 상관관계는 시간이 흐를수록 상관계수가 높아지고 있었는데 이는 후반으로 갈수록 통사 복잡성과 정확성의 경쟁 관계가 심해진다고 해석할 수 있다. 반면 어휘 복잡성과 문장당 오류 수는 뚜렷한 상관관계를 보이지 않았는데 1~6개월에만 어휘 다양도와 문장당 오류 수가 r=.363, p=.030으로 약한 정적 관계를 띠고 있었다. 다음은 무오류절 수를 통한 복잡성과 정확성의 기간별 상관관계를 살펴보도록 한다. 분석 결과는 <표 108>과 같다.

<표 108> 기간별 복잡성과 정확성(문장당 무오류절 수)의 관계

구분		N	문장당 무오류절 수(r/p)		
			1~6개월	7~12개월	13~18개월
통사 복잡성	문장 길이	36	.528/.001***	.568/.000***	.564/.000***
어휘 복잡성	어휘 다양도	36	-.300/.075	.043/.805	.156/.363
	어휘 밀도	36	-.091/.597	.132/.442	.220/.198
	어휘 정교도	36	-.056/.748	-.109/.526	-.213/.212

*p<.05 **p<.01 ***p<.001

표에서 확인할 수 있듯이 문장당 무오류절 수는 세 기간 모두 문장 길이와 유의한 정적 상관관계를 유지하고 있었다. 이는 문장 길이가 길어지면 무오류절의 수도 증가함을 의미하는 것으로 무오류절을 통한 정확성과 복잡성은 함께 발달하는 지지적 관계에 있음을 보여준다. 문장당 무오류절 수와 통사 복잡성의 두 지표 간의 상관계수는 꾸준히 0.5 이상의 뚜렷한 상관관계를 유지하고 있지만 7~12개월에 상관계수가 증가한 뒤 13~18개월에는 약간 하락하는 것을 확인할 수 있다. 한편, 문장당 무오류절 수는 세 기간 모두 어휘 다양도, 어휘 밀도, 어휘 정교도와 유의미한 관계를 보이지 않고 있었다. 앞에서 어휘 다양도가 문장당 오류 수와 유의미한 관계가 있었음을 고려해볼 때 문어에서 어휘 다양도는 무오류보다는 오류와 더 밀접한 관계가 있음으로 해석할 수 있다. 계속하여 상관관계가 비교적 뚜렷하게 나타난 통사 복잡성과 정확성의 관계가 학습자별로 어떻게 나타나는지 살펴보면 다음의 <표 109>와 같다.

<표 109> 학습자별 문어 통사 복잡성과 정확성의 관계

구분		N	문장당 오류 수(r/p)	문장당 무오류절 수(r/p)
AW	문장 길이	36	.584/.000***	.810/.000***
BW	문장 길이	36	.693/.000***	.598/.000***
CW	문장 길이	36	.184/.284	.710/.000***

*p<.05 **p<.01 ***p<.001

세 학습자의 문어 통사 복잡성과 정확성의 관계는 서로 다른 양상을 보이고 있는데 A학습자는 문장 길이와 오류 수, 무오류절 수가 모두 정적 관계를 보이지만 무오류절 수의 상관관계가 더 높았고 B학습자 역시 문장 길이와 오류 수, 무오류절 수 모두 정적 관계를 보이지만 오류 수와 관계가 더 밀접한 것으로 나타났다. 그리고 C학습자는 문장 길이와 문장당 무오류절 수만이 높은 정적 관계를 보였다. 즉 A학습자는 통사 복잡성과 정확성이 경쟁·지지적 관계에 있으나 지지적 관계가 더 강하게 나타났고 B학습자

역시 경쟁·지지적 관계에 있지만 경쟁 관계가 더 뚜렷하였으며 C학습자는 복잡성과 정확성이 지지적으로 발달하고 있었다. 이는 앞서 복잡성과 정확성의 발달에서 확인한 바와 같이 B학습자는 가장 복잡하지만 오류가 가장 많은 문장을 산출하고 A학습자는 복잡성이 비교적 높고 정확성 역시 비교적 높은 문장을 산출하며 C학습자는 복잡성이 낮지만 정확성이 높은 문장을 산출한다는 점을 뒷받침하는 결과이며 학습자마다 집중적으로 발달시키는 구인이 다르다는 것을 보여주는 결과이기도 하다.

4.2. 구어에 나타난 복잡성과 정확성의 관계

문어와 마찬가지로 구어 자료에 나타난 복잡성과 정확성의 각 측정지표에 대해 상관분석(Pearson correlation coefficient)을 실시한 결과는 <표 110>과 같다. 구어 자료 역시 무오류절의 비율은 통사 복잡성과 어휘 복잡성의 모든 지표와 유의한 상관관계를 보이지 않아 표로 제시하지 않았다.

<표 110> 구어 복잡성과 정확성의 관계

구분		N	AS-unit당 오류 수 (r/p)	AS-unit당 무오류절 수 (r/p)
통사 복잡성	AS-unit 길이	108	.603/.000***	.519/.000***
어휘 복잡성	어휘 다양도	108	.233/.015*	.270/.005**
	어휘 밀도	108	-.109/.261	.265/.006**
	어휘 정교도	108	.326/.001***	.073/.451

*p<.05 **p<.01 ***p<.001

통사 복잡성과 정확성을 살펴보면, AS-unit당 오류 수는 AS-unit 길이와 r=.603, p=.000으로 통사 복잡성과 뚜렷한 정적 상관관계를 보이고 있었다. 이는 통사 복잡성이 증가하면 AS-unit당 오류 수가 증가함을 의미한다. 한편, AS-unit당 무오류절 수 역시 AS-unit 길이와 r=.519, p=.000으로 뚜렷한 정적 상관관계를 보였는데 이는 통사 복잡성이 증가할 수록 AS-unit당 무오류절 수 역시 증가함을 의미한다. 즉 구어의 통사 복잡성과 정확성의 관계는 문어와 마찬가지로 측정 지표에 따라 다른 관계를 나타내고 있었으며 통사 복잡성이 증가하면 오류 수도 증가하지만 무오류절 수 역시 동시에 증가하는 현상이 나타나고 있었다.

어휘 복잡성과 정확성의 관계를 살펴보면, AS-unit당 오류 수는 어휘 다양도와 r=.233, p=.015, 어휘 정교도와 r=.336, p=.001로 약한 정적 관계를 보이고 있었으며 어휘 밀도와는 유의한 상관관계를 보이지 않았다. 이는 어휘 다양도와 정교도가 증가

할 수록 문장당 오류 수 역시 증가한다는 것을 의미하며 어휘 복잡성 중의 어휘 다양도와 정교도는 정확성과 서로 경쟁 관계에 있으며 동시에 발달할 수 없음을 의미한다. 그리고 AS-unit당 무오류절 수는 어휘 다양도와 r=.270, p=.005, 어휘 밀도와 r=.265, p=.006로 약한 정적 관계를 보이고 있었다. 이는 어휘 다양도나 어휘 밀도가 증가하면 AS-unit당 무오류절 수가 증가한다고 해석할 수 있다.

즉 학습자들은 구어에서도 문어와 마찬가지로 통사 복잡성이 증가하면 오류 수가 증가하지만 동시에 무오류절 수 또한 증가하는 현상이 나타나 통사 복잡성과 정확성의 관계가 측정 지표에 따라 다르게 나타나고 있어 복잡성과 정확성에 대해 일관된 결론을 도출하기가 어려웠다. 또한 문어와 달리 구어에서는 어휘 복잡성 역시 정확성과 일관된 관계를 보이지 않고 있었는데 어휘 다양도는 오류 수 및 무오류절 수와 약한 정적 관계를 나타내고 있었고 어휘 밀도는 무오류절 수와 유의한 정적 관계를 보였으며 어휘 정교도는 오류 수와 정적 관계를 보였다. 정리하면, 어휘 다양도는 정확성과 지지적·경쟁적 관계, 어휘 밀도는 정확성과 지지 관계, 어휘 정교도는 정확성과 경쟁 관계에 있는 것으로 나타났다.

이어서 앞에서 살펴본 복잡성과 정확성을 다시 기간별로 나누어 살펴보아 시간의 흐름에 따라 복잡성과 정확성의 관계에 변화가 있는지 살펴보도록 한다. AS-unit당 오류 수와 복잡성 간의 기간에 따른 상관관계의 분석 결과는 다음의 <표 111>과 같다.

<표 111> 기간별 복잡성과 정확성(AS-unit당 오류 수)의 관계

구분		N	AS-unit당 오류 수(r/p)		
			1~6개월	7~12개월	13~18개월
통사 복잡성	AS-unit 길이	36	.570/.000***	.619/.000***	.615/.000***
어휘 복잡성	어휘 다양도	36	.130/.450	.162/.345	.078/.651
	어휘 밀도	36	.067/.698	-.260/.125	-.172/.317
	어휘 정교도	36	.337/.045*	.115/.503	.332/.048*

*p<.05 **p<.01 ***p<.001

표에서 확인할 수 있듯이 AS-unit당 오류 수와 AS-unit 길이는 세 기간 모두 유의한 정적 상관관계를 나타내고 있었다. 이는 AS-unit 길이가 길어질수록 오류가 증가함을 의미하는 것으로 학습자들의 통사 복잡성과 정확성이 동시에 발달하지 못한다는 것을 보여주는 결과이다. 또한 이러한 상관관계는 시간이 흐를수록 상관계수가 높아지며 후반으로 갈수록 통사 복잡성과 정확성의 경쟁 관계가 점차 심해지는 것을 확인할 수 있었다. 어휘 복잡성과 AS-unit당 오류 수는 규칙적인 관계를 보이지 않았는데 AS-unit

당 오류 수와 어휘 정교도만이 1~6개월에 r=.337, p=.045, 13~18개월에 r=.332, p=.048로 약한 정적 상관관계를 보이고 있었다. 다음으로, 무오류절 수를 통한 복잡성과 정확성의 기간별 상관관계를 살펴보도록 한다. 분석 결과는 <표 112>와 같다.

<표 112> 기간별 복잡성과 정확성(AS-unit당 무오류절 수)의 관계

구분		N	AS-unit당 무오류절 수(r/p)		
			1~6개월	7~12개월	13~18개월
통사 복잡성	AS-unit 길이	36	.134/.435	.094/.585	.422/.010**
어휘 복잡성	어휘 다양도	36	-.131/.446	.215/.209	.008/.965
	어휘 밀도	36	.242/.155	.355/.034*	.369/.027*
	어휘 정교도	36	-.293/.083	.181/.292	-.119/.488

*p<.05 **p<.01 ***p<.001

AS-unit당 무오류절 수와 통사 복잡성의 상관관계는 기간별로 다르게 나타나고 있었는데 세 기간 중 13~18개월(r=.422, p=.010)에만 유의미한 정적 관계를 보이고 있었다. 이는 1~12개월에는 통사 복잡성과 무오류절 수 사이에 뚜렷한 관계가 없다가 13~18개월에 복잡성과 정확성 사이에 지지적 관계가 나타난다고 해석할 수 있다. 한편, AS-unit당 무오류절 수는 세 기간 모두 어휘 다양도, 어휘 밀도, 어휘 정교도와 유의미한 관계를 보이지 않으며, 어휘 다양도와 정확성 사이에 뚜렷한 상관관계를 보이지 않았다. 학습자별로 복잡성과 정확성의 관계를 살펴보면 다음의 <표 113>과 같다.

<표 113> 학습자별 구어 통사 복잡성과 정확성의 관계

구분		N	문장당 오류 수(r/p)	문장당 무오류절 수(r/p)
AS	AS-unit 길이	36	.567/.000***	.765/.000***
BS	AS-unit 길이	36	.704/.000***	.378/.023*
CS	AS-unit 길이	36	.594/.000***	.471/.004**

*p<.05 **p<.01 ***p<.001

표에서 확인할 수 있듯이 세 학습자의 구어 역시 통사 복잡성과 정확성의 관계가 서로 다른 양상을 보이고 있는데, A학습자는 경쟁·지지적 관계에서 지지적 관계가 더 강하고 B학습자는 경쟁·지지적 관계에서 경쟁 관계가 더 뚜렷하게 나타나며 문어와 같은 양상을 보였으나 C학습자는 문어와는 달리 통사 복잡성과 오류 수의 관계가 더 밀접한 것으로 나타났다. 이는 A와 B학습자의 복잡성과 정확성의 관계는 산출 양식의 영향을 크게 받지 않지만 세 학습자 중 숙달도가 가장 낮은 C학습자의 경우 문어에서는 복잡

성과 정확성이 지지적 관계에 있지만 구어에서는 경쟁적 관계가 더 강하게 나타남을 알 수 있었다. 위에서 살펴본 문어와 구어에 나타난 복잡성과 정확성의 관계를 정리하면 다음과 같다.

먼저 전체 기간으로 볼 때, 문어와 구어의 통사 복잡성 측정 지표와 정확성 측정 지표는 모두 상관계수 0.5 이상의 뚜렷하고 유의미한 정적 관계를 보였다. 그러나 이는 측정 지표에 따라 반대되는 결론으로 이어지게 되는데 문어와 구어 모두 통사 복잡성이 증가할수록 단위당 오류 수와 단위당 무오류절 수가 동시에 증가하는 것으로 나타나 통사 복잡성과 정확성의 일관된 관계를 도출하기가 어려웠다. 그러나 상관계수를 비교해 보면 문어에서는 무오류 수와의 상관계수가 오류 수와의 상관계수보다 높은 반면 구어에서는 문어와 상반되는 경향을 보이고 있음을 확인할 수 있다. 이를 통해 문어와 구어의 통사 복잡성과 정확성은 경쟁과 지지 관계를 동시에 보이고 있지만 문어에서는 지지적 관계가 좀 더 강한 반면 구어에서는 경쟁 관계가 좀 더 강한 것으로 이해할 수 있다.

어휘 복잡성은 문어와 구어에서 지표에 따라 다르게 나타났다. 문어에서는 어휘 다양도와 문장당 오류 수만 정적 관계를 보여 어휘 다양도와 정확성이 동시에 발달하지 못하고 경쟁하는 모습을 보인 반면, 구어에서는 어휘 다양도가 오류 수와 무오류절 수 모두와 약한 정적 관계를 나타내고 있었고 어휘 밀도는 무오류절 수와 정적 관계, 어휘 정교도는 오류 수와 정적 관계를 나타내고 있었다. 따라서 문어에서는 어휘 다양도와 정확성은 경쟁 관계에 있고는 구어에서는 어휘 다양도는 정확성과 경쟁·지지 관계, 어휘 밀도는 정확성과 지지 관계, 어휘 정교도는 정확성과 경쟁 관계에 있는 것으로 정리할 수 있다.

기간별로 나누어 살펴본 결과를 정리하면, 문어 자료에서는 세 기간 모두 통사 복잡성과 정확성의 지표가 유의하고 뚜렷한 정적 관계를 보이며 시간이 흐를수록 상관계수가 증가하며 밀접한 관계를 보였고, 구어 자료에서는 단위당 오류 수는 세 기간 모두 통사 복잡성과 정적 관계를 보였지만 단위당 무오류절 수는 13~18개월에만 유의한 상관관계를 보였다. 즉, 기간별로 살펴본 문어와 구어의 통사 복잡성과 정확성의 관계는 전체 기간과 마찬가지로 측정 지표에 따라 반대되는 관계를 보이고 있었다.

그러나 통사 복잡성과 정확성의 두 지표와의 상관계수를 세부적으로 비교해 보면 우선 문어에서는 1~6개월, 7~12개월에는 무오류절 수와의 상관계수가 오류 수와의 상관계수보다 비해 높았으나 13~18개월에는 오류 수와의 상관계수가 더 높아지는 것을 확인할 수 있고 구어에서는 세 기간 모두 무오류절 수와의 상관계수에 비해 오류 수의

상관계수가 더 높은 것을 확인할 수 있다. 즉 구어에서는 전체 관찰 기간 동안 통사 복잡성과 정확성의 경쟁 관계가 지지 관계에 비해 더 뚜렷하며 문어 역시 관찰 시작 12개월이 지난 시점인 13~18개월에 들어서면서 지지 관계보다 경쟁관계가 더 뚜렷해짐을 알 수 있다.

이를 통해 문어와 구어 모두 복잡성과 정확성의 관계는 일정하지 않고 시간의 흐름에 따라 지속적으로 변화하고 있지만 두 요인이 지지적으로 발달한다고 보기는 어렵다는 결론을 내릴 수 있다. 기간별 어휘 복잡성과 정확성은 문어에서는1~6개월에 어휘 다양도와 오류 수만이 유의한 정적 관계를 보였고 기타 기간에는 어휘 복잡성과 정확성의 유의한 관계가 발견되지 않았다. 구어에서는 문어보다 더 복잡한 관계가 나타났는데 1~6개월과 13~18개월에는 어휘 정교도와 오류 수가 정적 관계를 나타냈고 7~12개월, 13~18개월에는 어휘 밀도와 무오류절 수가 정적 관계를 보이고 있었다.

종합해 보면, 학습자들의 문어와 구어는 모두 통사 복잡성과 정확성 간에 뚜렷한 관계를 보이지만 측정 지표에 따라 경쟁-지지의 상반되는 관계를 나타내고, 어휘 복잡성은 지표나 기간에 따라 다르게 나타나는데 전반적으로 문어에서는 어휘 다양도와 정확성이 관계를 보이고 구어에서는 어휘 밀도와 정교도가 정확성과 유의미한 관계를 보인다.

V. 논의 및 제언

본 연구는 중국 배경 중도입국 청소년의 한국어 문어와 구어의 복잡성과 정확성 발달 양상과 문어와 구어의 차이에 더불어 복잡성과 정확성의 관계를 살펴보는 것에 목적을 두었다. 분석 결과, 학습자들의 통사 복잡성은 시간의 흐름에 따라 꾸준히 증가하지만 어휘 복잡성은 12개월이 지난 시점에서 정체가 발생하였고 정확성은 전체 관찰 기간 동안 뚜렷한 발달을 보이지 않고 있었다. 또한 문어와 구어는 비슷한 경로로 발달하였지만 문장의 확장 유형이나 문법 항목에 따라 차이를 보이기도 하였고 복잡성과 정확성은 경쟁과 지지 관계를 반복하며 고정적인 관계를 보이지 않았다. 아래 4장의 분석 결과를 바탕으로 서론에서 상정한 네 가지 연구 문제에 대해 논의를 해보도록 한다.

1. 복잡성과 정확성의 통시적 발달

본고의 첫 번째 연구 문제는 공교육에 진입한 중국 배경 중도입국 청소년의 한국어 문어와 구어의 복잡성과 정확성이 어떤 발달 양상을 보이는지 살펴보는 것이다. 이를 위해 본 연구에서는 먼저 세 명의 학습자들이 산출한 문어와 구어 자료를 각각 6개월 단위로 묶은 뒤 평균을 산출하여 6개월 간격으로 복잡성과 정확성의 각 측정 지표의 변화 추이를 살펴보았다. 그리고 기간별로 일원배치 분산분석(One-Way ANOVA)을 실시하여 기간별 차이의 유의성과 각 지표가 유의한 발달이 나타나는 시점을 살펴보았다. 아래에서는 본고에서 분석한 결과를 바탕으로 다른 연구의 결과들과 비교하여 본고의 학습자의 언어 발달 경로, 발달 시점, 발달 수준에 대해 논의하도록 한다.

1.1. 통사 복잡성의 발달

1.1.1. 통사 복잡성의 발달 경로

본고에서는 통사 복잡성을 단위 길이, 단위당 절 수, 단위 확장의 유형을 통해 측정하였다. 그 결과, 학습자들의 문어 자료에 나타난 문장 길이는 시간의 흐름에 따라 일정한 폭으로 증가하며 안정적인 발달을 보이고 있었으며 이는 구어 자료 역시 동일하였다. 학습자들은 시간이 지날수록 하나의 AS-unit에 더 많은 형태소를 사용하며 단위의 길이를 점점 늘려가고 있었다. 그리고 단위당 절 수 역시 단위 길이와 마찬가지로 문

어 자료와 구어 자료에서 모두 시간의 흐름에 따라 일정한 폭으로 증가하면서 안정적인 발달을 보이고 있었다. 또한 학습자들의 문어와 구어의 단위 길이와 단위당 절 수는 1~6개월, 7~12개월, 13~18개월 사이의 차이가 모두 통계적으로 유의미한 것으로 나타나 학습자들은 산출 양식에 관계 없이 시간이 지날수록 점점 더 길고 구조적으로 복잡한 언어를 산출한다는 것을 확인할 수 있었다.

이러한 결과는 대학기관 학습자의 한국어 통사 복잡성 발달을 종적으로 살핀 음승민(2012)이나 이복자(2015), 여성결혼이민자의 쓰기 발달을 살핀 주월랑(2018), 자연 학습 환경에 있는 학습자의 거주 기간에 따른 문어와 구어의 통사 복잡성을 분석한 김정운 외(2017a), 같은 자연 학습 환경의 학습자의 한국어 발달을 종적으로 연구한 김정운 외(2017b) 등의 연구들과 일치하는 결과로 중국 배경 중도입국 청소년 학습자들 역시 한국 거주 기간이 길어짐에 따라 문어와 구어 모두 길이가 길어지고 통사적으로 복잡해진다는 결론을 내릴 수 있다.

통사 복잡성이 학습자의 숙달도를 보여주는 매우 중요한 요인임은 여러 선행연구들에서 이미 밝혀진 바 있다. 박지순·서세정(2009), 서세정(2009), 남주연(2015), 남주연·김영주(2015), 남주연 외(2016), 주우동(2019), 서신웨(2022) 등 연구에서는 숙달도에 따른 한국어 학습자들의 언어 자료를 분석하여 숙달도가 높아질수록 단위 길이와 단위당 절 수가 유의미하게 증가한다는 것을 밝혔으며, 남주연(2015)에서는 AS-unit당 절 수가 학습자의 숙달도 총점뿐만 아니라 어휘·문법, 쓰기, 듣기, 읽기 영역과 높은 정적 상관성을 보인다는 것을 발견하고 한 번의 발화에서 더 많은 절을 사용하여 통사적 복잡성을 높일 수 있는 능력은 한국어의 이해 영역뿐만 아니라 쓰기와 어휘·문법 영역까지도 영향을 미치는 것으로 해석하였다. 이러한 결과들과 연결 지어 볼 때, 본고의 연구 대상인 중국 배경 중도입국 청소년은 시간이 지날수록 문어와 구어에서 모두 더 길고 복잡한 언어를 산출하며 언어 숙달도를 높여 가고 있다는 결론을 내릴 수 있다.

이어서, 본고에서는 절의 확장을 대등절과 내포절 및 유형별 내포절로 나누어 살펴보았다. 우선 대등절과 내포절의 기간별 증가폭을 확인해 보면 문어와 구어 모두 내포절 수가 대등절 수에 비해 높으며 증가폭 역시 내포절이 더 큰 것을 확인할 수 있었다. 이는 여성결혼이민자를 대상으로 한 주월랑(2018)이나 대학 기관 학습자를 대상으로 한 주우동(2019), 남주연·김영주(2015), 혹은 자연 학습 환경의 학습자의 통사 발달을 고찰한 김정운 외(2017a, 2017b)와도 일치하는 결과이다. 내포절을 형성하는 문법 항목이 대등절을 형성하는 문법 항목에 비해 종류가 압도적으로 높아 내포절을 활용하여 더 정교하고 다양한 의미를 표현할 수 있는 한국어의 특징을 고려할 때, 이는 한국어 발

달의 보편적인 현상으로 추측할 수 있다. 또한 이러한 결과를 대등절의 사용은 숙달도와 관련성이 없으나 내포절은 숙달도에 따라 증가한다고 밝힌 남주연(2015), 남주연 외(2016), 서신웨(2022) 등의 결과와 연결해 볼 때, 학습자들이 유의미하게 통사 복잡성을 높여가고 있음을 보여주는 긍정적인 현상으로 해석할 수 있다.

증가폭 외에도 대등절과 내포절은 유의미한 증가가 나타나는 구간에 차이가 있었다. 문어에서 대등절은 1~6개월과 7~12개월 사이에만 유의한 증가가 나타나고 내포절은 7~12개월과 13~18개월 사이에만 유의한 증가를 보이고 있었다. 그러나 구어에서는 대등절과 내포절 모두 1~6개월, 7~12개월, 13~18개월 사이에 유의한 차이를 보이며 문어에 비해 더 안정적인 폭으로 꾸준히 증가하는 것으로 나타났다. 이는 학습자들이 문어에서는 대등절의 발달이 먼저 일어나고 이후 내포절의 발달이 일어나며 기간별로 발달 속도에 차이가 있는 반면 구어에서는 대등절과 내포절이 일정한 속도로 발달한다는 것을 의미한다.

내포절을 다시 세분화하여 유형별 내포절의 발달을 살펴보면, 부사절은 문어와 구어에서 모두 꾸준히 발달하고 명사절은 문어와 구어 모두에서 발달이 일어나지 않으며 관형사절, 인용절은 문어와 구어에서 서로 다른 발달을 보이고 있었다. 명사절과 관형사절의 발달에 대해 논의해 보면, 구어에서 명사절은 13~18개월에 이르러 정체가 나타나고 문어에서는 13~18개월에 빈도가 감소하며 퇴행을 보이고 있었다. 그러나 숙달도별 내포절의 발달을 살펴본 남주연 외(2016)나 서신웨(2022)에서는 숙달도가 올라감에 따라 명사절이 꾸준히 증가한다고 밝히고 있고 재외동포 청소년 학습자의 한국어 발달을 고찰한 김현아(2016)에서는 학습자의 숙달도나 연령에 따라 꾸준히 증가하는 절은 명사절과 부사절이라고 밝히고 있다. 또한 최수완(2019)에서도 계승어 학습자와 모어화자 모두 학년이 올라감에 따라 문장당 명사절, 관형사절, 부사절이 꾸준히 증가한다고 한 것과 달리 본고의 학습자들은 문어와 구어에서 명사절의 정체, 퇴행 현상을 보인다는 점에서 그 의미를 숙고할 필요가 있다.

숙달도에 따른 명사절의 사용을 분석한 한송화(2015b)에 따르면 '-ㅁ, -기'에 의한 순명사절과 '-ㄴ/는 것'에 의한 준명사절, '-ㄴ/는 지, -ㄹ까, ㄴ/는 가' 등에 의한 의사명사절 중에서 초, 중, 고급 간에 모두 유의한 차이를 보이는 것은 준명사절이라고 밝히고 있는데 본고의 학습자들은 문어와 구어 모두 '-기'와 '-ㄴ/는 것'의 사용이 눈에 띄게 줄어들고 있었다. 명사절은 중국어에서는 문법 표지 자체가 없거나 '的, 之, 所'와 같은 조사, 대명사 '其', 그리고 일부 단순 명사나 굳어진 표현에 의해서 이루어지기도 하는 등 한국어의 명사절과 달리 뚜렷한 형태적 특징을 보이지 않는다. 따라서 학습자

들은 한국어에서 관용 표현으로 학습한 '-ㄴ/는지 알다/모르다'나 '-기 어렵다/쉽다'와 같은 한정된 명사절에 의존하는 경향이 강하며 이로 인해 더 다양한 명사절의 확장을 이루어내지 못하고 있는 것으로 해석할 수 있다. 한송화(2015b)에서 역시 중국어권 학습자의 명사절 사용률이 비교적 낮은 데다가 과잉사용 오류 역시 다른 언어권에 비해 높은 것으로 보고하고 있다는 점을 볼 때 명사절 습득의 어려움은 중국어권 학습자의 특징으로 해석할 여지가 있다. 이 외에도 본고에서는 한송화(2015)와 달리 '-ㄴ/는 것이다'와 같은 양태적 기능을 하는 '-것'의 형태는 명사절에 포함시키지 않고 있었는데 이러한 차이가 '-ㄴ/는 것'의 출현 빈도를 감소시키는 원인이 되었을 수도 있을 것이다.

다음, 관형사절은 문어에서는 기간별로 유의미한 발달을 보이지 않은 것과 달리 구어에서는 7~12개월과 13~18개월 사이에 발달을 보였다. 이러한 결과는 구어 자료를 바탕으로 분석한 남주연(2015)과 문어 자료를 분석한 남주연 외(2016)의 결과를 참고해볼 수 있는데 남주연(2015)에서는 부사절과 관형사절이 숙달도에 따라 유의미하게 증가했다고 밝힌 반면 남주연 외(2016)에서는 부사절만 숙달도에 따른 발달 양상이 나타났다는 결과를 보고하였다. 본고에서 역시 문어에서는 기간별 유의한 차이를 보이지 않던 관형절이 구어에서는 7~12개월과 13~18개월에 유의한 차이가 나타났다는 점에서 문어보다는 구어에서 관형사절의 발달이 더 뚜렷하게 나타난다고 할 수 있다. 그러나 이는 발달의 유의성에 대한 것으로 구어의 관형사절 사용 빈도가 문어에 비해 높음을 의미하지는 않는다. 문어와 구어의 기간별 관형사절 수를 비교해 보면 문어는 구어에 비해 높은 출발을 보이고 있으며 1~6개월과 7~12개월에는 구어의 단위당 관형사절 수가 문어에 비해 유의하게 낮게 나타나다가 13~18개월에 들어 차이가 유의하지 않게 된다. 이는 명사절이나 부사절, 인용절에서는 나타나지 않는 현상으로 관형사절만이 구어에서 초반의 지체 현상이 나타나고 있음을 뜻한다.

1.1.2. 문법 항목의 사용 비율

단위 확장에 관여하는 각각의 문법 항목은 모어화자의 언어에서도 높은 빈도로 사용되는 문법 항목이 있는 반면 거의 사용되지 않는 항목도 있다. 또한 이러한 문법 항목은 장르나 주제, 문어나 구어와 같은 사용역에 따라 달라지기도 한다. 그러므로 학습자가 사용한 각 문법 항목의 비율을 모어화자의 언어와 비교해 본다면 학습자의 언어 발달의 정도를 가늠해 볼 수 있을 것이다. 그러나 중도입국 청소년 학습자는 청소년이라는 점에서 인지 발달이 성인 학습자와 차이가 있다. 또한 서론에서 언급하였듯이 중도입국 청소년 학습자들은 공교육에 진입하여 한국 학생들과 함께 학업을 수행해야 한다

는 점에서 이들의 언어 발달은 비슷한 연령대의 모어화자와 비교되어야 할 것이다.[74] 따라서 본고에서는 본고의 학습자와 비슷한 연령대의 모어화자의 언어 자료를 분석한 연구와 비교해 보고자 한다. 단위 길이와 단위당 절 수에서 확인하였듯이 학습자들은 시간이 지날수록 통사 복잡성이 증가하여 13~18개월에 가장 다양한 문법 항목을 높은 빈도로 산출하고 있었다. 그러므로 학습자들이 13~18개월에 산출한 문법 항목을 기준으로 모어화자 청소년들과 비교하도록 한다.

선행연구를 살펴보면 한국인 청소년 모어화자의 통사 복잡성을 분석한 연구들은 대부분 재외동포 아동의 언어 발달 수준을 모어화자 청소년의 언어와 비교하는 과정에서 이루어지고 있었다. 대표적으로 김현아(2016)와 최수완(2019)의 연구가 있는데 김현아(2016)에서는 한국인 중·고등학생을 묶어 모어화자의 통사 복잡성을 분석하였고 최수완(2019)에서는 한국인 청소년을 4~5학년, 6~7학년, 8~9학년으로 나누어 분석하였다. 두 연구 모두 본 연구와 마찬가지로 서사적 쓰기[75]를 수집하였고 대등절과 내포절 관련 문법 항목의 사용 빈도를 명시하고 있어 비교가 가능하였다.[76] 본고의 연구 대상자 중 A학습자는 공교육 편입 시 1학년으로 입학하였고 B, C학습자는 2학년으로 편입하였다. 따라서 연구 시작 13~18개월 차에 세 학습자는 중학교 2학년과 3학년에 재학하고 있었다. 이에 본고에서는 김현아(2016)에서 제시한 중·고등학생의 자료, 최수완(2019)에서 제시한 8~9학년의 자료를 본 연구의 분석 결과와 비교하여 각 문법 항목의 사용 비율이 같은 연령대의 모어화자의 언어와 비교하여 적절한지, 특정 문법 항목에 과도하게 치중되어 있지 않은지 살펴보았다. 비교는 단위 확장의 유형과 동일하게 대등절과 내포절(명사절, 관형사절, 부사절, 인용절)로 나누어 살펴보았다.

먼저 대등절 관련 문법 항목을 비교해 보면, 가장 눈에 띄는 것은 '대조·역접'의 '-ㄴ데'의 사용이었다. '-ㄴ데'는 모어화자의 쓰기 자료에서는 2% 정도로 나타나는 것에 비

---

[74] 중도입국 청소년 학습자의 자료를 같은 연령대의 모어화자와 비교하는 이유는 각 문법 항목의 사용 비중이 일반 모어화자와 비교하여 적절한 지, 특정 문법에 과도하게 치중되어 있지 않은 지 살펴보기 위함이다.

[75] 김현아(2016)에서 수집한 쓰기 자료의 주제는 나의 소개, 겨울방학, 크리스마스, 우리 가족, 나의 꿈, 내가 좋아하는 운동 등 9가지 서사적 주제 중 자유롭게 선택하여 쓴 작문이었고 최수완(2019)에서 수집한 쓰기 자료의 주제는 '방학 중 가장 기억에 남는 일'이었으며 두 연구 모두 집단별로 40편의 자료를 분석하였다.

[76] 대등절과 내포절을 구성하는 문법 항목의 분류는 연구마다 약간 상이하다. 이 경우 본고와 다른 분류들은 본고의 기준에 맞게 수정하여 수치를 비교하였다. 예를 들면 '배경'의 '-아(서)/어(서)'는 본고에서 '계기적 사건'에 포함시키고 있으므로 최수완(2019)에서 제시한 '선후'의 '-아(서)/어(서)'와 '배경'의 '-아(서)/어(서)'를 합친 수치를 비교하였다.

해 본고의 학습자는 문어에서 9.8%, 구어에서 7.4%[77]로 나타나 학습자들이 '-ㄴ데'를 과도하게 사용하고 있음을 알 수 있다. 또한 '-ㄴ데'는 문어보다는 구어에서 많이 사용되는 연결어미로 연세대학교 언어정보연구원의 구어와 문어 말뭉치를 바탕으로 분석한 전영옥·남길임(2005)의 결과를 확인해 보면 '-ㄴ데'는 구어에서는 4번째로 많이 사용되는 반면 문어에서는 14번 째로 많이 사용되는 항목으로 나타났고 한국어 모어화자의 연결어미를 분석한 이민경(2016)에서도 문어에서 6%, 구어에서 11.5%를 차지하는 것을 확인할 수 있다.[78] 이와 달리 본고의 학습자는 오히려 문어에서 '-ㄴ데'를 더 높은 빈도로 사용하고 있었다는 점은 학습자들이 문어체와 구어체의 표현을 아직 구분하지 않고 있음을 알 수 있다. 이 외에도 김현아(2016)에서는 '-고'와 '-지만'의 빈도만을 제시하여 알 수 없었으나 최수완(2019)과 비교할 때, '-나, -라, -든지/든가'와 같은 문법 항목은 본고의 학습에게서는 나타나지 않은 반면, 한국어 모어화자의 작문에서는 낮은 빈도이지만 다양하게 출현하고 있었다. 이는 다양한 대등절 관련 문법 항목의 사용을 가르쳐 특정 항목에 대한 의존도를 줄일 필요가 있음을 시사한다.

다음, 명사절 관련 문법 항목을 살펴보면 가장 눈에 띄는 것은 '-ㄴ/는 것'과 '-ㄴ지/는지'의 비율이었다. '-ㄴ/는 것'은 모어화자 청소년들의 자료에서는 50% 미만을 차지하는 것에 비해 본고의 학습자들은 문어에서 52.2%,구어에서 68.6%의 비율로 사용하고 있었고 '-ㄴ지/는지'는 모어화자 청소년들의 언어에서는 5%미만으로 나타나는 것과 달리 본고의 학습자는 문어에서 18.9%, 구어에서 8.6%의 비율로 사용하고 있었다. 그리고 이와 달리 '-ㄹ 것'과 '-던 것'은 모어화자 청소년에 비해 현저히 낮은 비율로 나타나고 있었다. 명사형 전성어미 '-ㅁ'과 '-기'를 살펴보면 '-기'의 비율에서는 본고의 학습자와 모어화자 사이에 큰 차이가 없었으나 '-ㅁ'은 모어화자 청소년에게서는 2~4% 정도의 낮은 비율이지만 두 연구에서 모두 관찰된 것에 반해 본 연구에서는 문어 자료에서도 한 차례도 나타나지 않았다. 이러한 결과를 통해 학습자들의 명사절 사용에서 관찰된 '-ㄴ/는 것'과 '-ㄴ지/는지' 두 항목에 대한 과도한 의존도는 결국 맥락에 따라 적절한 명사절을 사용하지 못하는 것에서 기인하는 것으로 볼 수 있다. 그러므로 '-ㄹ 것, -던 것, -ㅁ'과 같이 모어화자에 비해 사용 비율이 현저히 낮거나 사용하지 못하는 문법 항목들에 주목하여 교육을 진행할 필요가 있을 것이다.

---

[77] 각 문법 항목별 사용 비율은 앞서 4장에서 제시한 기간별 문법 항목의 사용 빈도 중 13~18개월의 빈도를 바탕으로 산출된 백분율이다.

[78] 전영옥·남길임(2005)과 이민경(2016)에서는 의미 기능에 따라 구분하지 않고 '-ㄴ데' 형태의 빈도를 산출하였다는 점에서 본고와 차이가 있지만 '-ㄴ데' 형태가 문어보다 구어에서 더 많이 사용됨을 알 수 있다.

이어서 관형사절 관련 문법 항목의 사용 양상을 비교해 보면 본고의 학습자와 모어화자의 가장 큰 차이점은 관형사형 어미 '-던'의 사용에서 나타났다. 최수완(2019)의 쓰기 주제가 '방학에 있었던 일'이라는 점에서 학습자들이 자연스럽게 '-던'을 더 많이 사용하였을 것이라는 점을 감안하더라도 모어화자 청소년의 경우 '-던'에 의한 관형사절이 전체 관형사절의 9~15% 정도 차지하는 것에 반해 본고의 학습자는 문어에서는 1.3%의 비율로 나타나고 구어에서는 나타나지 않고 있었다. 이러한 현상은 김현아(2016)와 최수완(2019)뿐만 아니라 여성결혼이민자의 한국어 발달을 연구한 주월랑(2018)에서도 동일하게 제기하고 있는 문제로 중도입국 청소년 학습자만의 문제이기보다는 한국어 관형사형 어미 '-던' 자체의 어려움에서 오는 문제인 것으로 보인다. 관형사형 어미 습득의 어려움은 이진경(2006), 한송화(2015)등 연구에서도 동일하게 제시된 것으로 김수정·박덕유(2017)에서는 다른 관형사형 어미에 비해 '-던'이 낮은 사용률과 높은 오류율을 보이는 것은 의미 기능에 대한 명확한 인지 부족으로 사용을 회피하거나 유사 어미와의 혼동으로 인한 것으로 해석하고 있다.

이와 같이 과거 관형사형 어미는 언어권이나 학습 연령을 불문하고 습득의 어려움을 제기하는 항목이기는 하지만 4장에서도 언급하였듯이 시제나 선행하는 품사에 관계없이 일반적으로 '的'로 통일되는 중국어가 모어인 학습자가 구분하여 사용하기는 더 쉽지 않을 것이다. 따라서 각 관형사형 전성어미가 가진 의미 기능의 유사성이나 차이점을 명시하여 명확히 인지할 수 있도록 교육을 할 필요가 있으며 특히 '-던, -았던'과 같은 관형사형 어미는 자주 결합하는 빈도 부사나 시간 부사어와 연결하여 지도하는 등의 방법을 고민해볼 수 있을 것이다.

부사절 관련 문법 항목의 사용을 모어화자 청소년과 비교해볼 때 눈에 띄는 차이점은 '계기적 사건'의 '-고(서)'와 '이유·원인'의 '-니/니까', '조건·가정'의 '-면'에 대한 집중도였다. 특히 '계기적 사건'의 '-고(서)'는 모어화자의 언어에서 5~10% 정도의 비율로 나타나는 것에 반해 본고의 학습자들은 문어에서 24.1%, 구어에서 22.5%로 2배 이상의 비율로 나타나고 있어 학습자들이 해당 문법 항목에 대한 의존도를 가늠할 수 있었다. '이유·원인' 관련 문법 항목에서는 '-니/니까'는 문어에서 11.1%, 구어에서 6.4%의 비율로 나타나며 모어화자의 언어에서 3%미만으로 나타나는 것과 대조적이었다. 흥미로운 점은 '-니/니까'와 달리 같은 '이유·원인'을 표현하는 '-아(서)/어(서)'의 비율은 모어화자에 비해 낮게 나타났는데 특히 '-아(서)/어(서)'의 비율이 모어화자에 비해 낮은 만큼 '-니/니까'의 비율이 더 높은 것을 확인할 수 있다. 동시에 '-니까'는 '-아(서)/어(서)'와의 대치 오류가 빈번히 관찰되는 항목으로, 이는 학습자들이 이유·원인을

표현할 때 '-니/니까'나 '-아(서)/어(서)'를 사용한다는 점은 인지하고 있지만 두 항목의 의미 차이를 제대로 구분하지 못하고 있음을 의미한다. 이정희(2002)에서도 '이유·원인'의 의미를 나타내는 연결어미 중 가장 많은 오류를 보이는 항목이 '-니까'와 '-아(서)/어(서)'임을 밝히고 있는데 이는 문법 항목의 교육에 있어 형태나 의미 기능 외에도 유사 연결어미의 차이점을 중요하게 다루어야 할 필요성이 있음을 보여준다. 이 외에 '조건·가정'의 '-면' 역시 모어화자 청소년의 2배 정도 비율로 나타나고 있어 학습자들이 많이 의존하고 있는 문법 항목임을 알 수 있었다. 그러나 '-면' 역시 집중도가 높은 타 문법 항목과 마찬가지로 대치 오류와 이형태 오류가 매우 빈번하였다.

반대로, 모어화자 청소년에 비해 뚜렷하게 낮은 비율로 사용되고 있는 어미로는 '동시적 사건'의 '-며/-면서'나 '사건의 전환'의 '-다(가)', '수단·방법'의 '-아(서)/-어(서)', '방식·정도'의 '-게'와 '-이'가 있었다. 전체적인 부사절 관련 문법 항목의 종류를 비교해 보면 13~18개월 기준 문어에서 19가지, 구어에서 16가지 문법 항목을 사용하고 있는 반면 모어화자 청소년은 최대 28가지의 문법 항목을 사용하고 있었다.

정리하면, 비슷한 연령대의 모어화자와 비교할 때 본고의 학습자들은 부사절을 사용하는 과정에서 특정 2~3가지 문법 항목에 대한 의존도가 특별히 높은 것으로 나타났으며 이러한 현상은 구어에서 더 강하게 나타나고 있었다. 또한 학습자들의 의존도가 높은 항목들은 대치 오류나 오형태 오류를 범하는 일이 빈번한 항목이기도 하였는데 이는 학습자들이 다양한 연결어미를 습득하지 못하였거나 유사 연결어미 간의 차이점을 인지하지 못한 상태에서 문장을 확장하면서 발생한 현상이라고 할 수 있다. 이러한 결과들은 중도입국 청소년 학습자들의 구조적 복잡성은 양적인 측면뿐만 아니라 질적인 측면에서 제고되어야 할 필요성이 있음을 시사한다.

마지막으로 인용절 관련 문법 항목의 사용에서 눈에 띄는 점은 직접 인용 '-(이)라고'의 사용 비율이었다. 모어화자 청소년은 낮게는 75% 이상, 높게는 90% 가까이 간접 인용을 사용하고 있는 반면 본고의 학습자는 간접 인용절의 사용이 문어에서 67.3%, 구어에서 70.5%로 모어화자 청소년에 비해 낮고 그 대신 직접 인용절을 모어화자에 비해 높은 비율로 사용하고 있었다. 이는 호주 교포 학습자에게서도 비슷하게 나타나는 현상으로, 최수완(2019)에서는 8~9학년의 교포 학습자의 직접 인용 '-(이)라고'의 사용 비율이 14.28%이라고 보고하고 있다.

한국어 문법에서 직접 인용은 큰 따옴표 안에 다른 사람의 말이나 글을 그대로 따온 뒤 '-(이)라고' 혹은 '-하고'와 같은 조사를 붙여 이루어지지만 간접 인용은 원래의 문장 형식이 그대로 옮겨지지 않고 화자의 현재 관점에서 기술된다는 점에서 직접 인용

과 달리 인칭 대명사나 시간 표현이 달라져야 한다는 차이가 있다. 따라서 학습자들의 직접 인용 사용 비율이 모어화자에 비해 높은 것은 문법 사용의 난이도에 기인하였을 것으로 추측해볼 수 있다. 그러나 인용 표현의 경우, 맥락에 따라 직접 인용이 자연스러운 경우도 있고 간접 인용이 자연스러운 경우도 있기 때문에 학습자가 편의에 의해 특정 항목을 사용하지 않도록 명시적인 교육이 필요할 것으로 보인다.

1.1.3. 통사 복잡성의 발달 수준

앞에서 문어와 구어의 통사 복잡성이 어떤 경로로 발달하는지, 어떤 문법 항목들을 통해 발달하는지 살펴보았다. 그러나 단순히 발달 경로나 문법 항목의 사용 비율로는 학습자들의 복잡성이 어느 수준에 도달하였는지 확인하기가 어렵다. 이에 본고에서는 같은 측정 지표를 사용하여 숙달도별, 연령별 통사 복잡성을 측정한 기타 연구들의 결과와 본고의 연구 결과를 비교하여 본 연구에서 대상으로 한 중국 배경 중도입국 청소년 학습자들의 통사 복잡성 발달 수준을 가늠해 보고자 한다.

선행연구들을 살펴보면 숙달도와 통사 복잡성의 관계를 살핀 연구들은 여럿 있으나 구체적인 측정 지표가 본 연구와 다르거나 숙달도 구분을 명시하지 않은 경우도 있고 대학 기관의 등급, 혹은 자체 제작 문항으로 숙달도를 구분한 연구들이 대부분이었다. 이에 본고에서는 여러 선행연구들 중 김현아(2016), 최수완(2019), 서신웨(2022)의 연구 결과와 비교해 보고자 한다. 김현아(2016)와 최수완(2019)의 연구는 앞에서 설명한 바와 같이 청소년 모어화자 및 재외동포 학습자의 언어를 분석하였다는 점에서 본고의 학습자와 비슷한 연령대의 모어화자의 언어 수준을 알 수 있고 서신웨(2022)는 대학기관 학습자의 쓰기 자료를 수집하여 분석한 연구로 본고와 1:1비교가 어렵지만 토픽 시험 점수를 기준으로 숙달도를 구분하였다는 점에서 학습자들의 숙달도 수준을 가늠하는 참고할 수 있을 것으로 판단하였다. 또한 세 연구 모두 본고와 같은 측정 지표를 사용하여 비교가 가능하였다. 다음의 <표 114>는 김현아(2016), 최수완(2019), 서신웨(2022)에서 제시한 문장 길이와 문장당 절 수의 측정 값을 정리한 표이다.

<표 114> 선행연구의 통사 복잡성 측정 값

숙달도(TOPIK)		3급	4급	5급	6급	-	-	-	-
서신웨 (2022)	문장당 형태소 수	18.3	22.3	25.0	26.8	-	-	-	-
	문장당 절 수	2.74	3.29	3.94	4.00	-	-	-	-
숙달도(한글학교)		3급	4급	5급	6급	초	중	고	NS[79]
김현아 (2016)	문장당 형태소 수	10.7	12.5	14.2	15.1	11.1	12.8	14.1	25.9
	문장당 절 수	0.47	0.78	1.17	1.26	0.56	0.85	1.10	2.84
숙달도(학년)		HS4~5	HS6~7	HS8~9	NS4~5	NS6~7	NS8~9	-	-
최수완 (2019)	문장당 형태소 수	12.67	13.31	18.69	17.58	21.27	26.35	-	-
	문장당 절 수	0.93	1.02	1.93	1.86	2.46	3.34	-	-

비교에 앞서 본 연구의 기간별 문장당 형태소 수를 살펴보면 1~6개월에 10.19개, 7~12개월에 12.38개, 13~18개월에 14.28개였다. 이를 서신웨(2022)의 결과와 비교해 보면, 본고의 연구 참여자의 13~18개월의 문장 길이는 서신웨(2022)에서 제시한 3급 학습자의 평균인 18.3개보다 낮은 수준임을 알 수 있다. 그리고 김현아(2016)의 재미동포 청소년 학습자들과 비교해 보면, 1~6개월에 숙달도 3급, 7~12개월에 숙달도 4급, 13~18개월에 숙달도 5급과 비슷한 수준이며, 연령대별로 비교해 보면 7~12개월에는 재미동포 중학생과 비슷한 수준, 13~18개월에는 재미동포 고등학생의 평균 보다 약간 높은 수준을 보이고 있었다. 그러나 이는 김현아(2016)에서 제시한 청소년 모어화자의 평균인 25.9개와는 큰 격차가 있었으며 최수완(2019)에서 제시한 초등학교 4~5학년의 문장당 형태소 수인 17.58개보다 낮은 수치였다.

계속하여 문장당 절 수를 비교해 보도록 한다. 본고의 학습자의 문장당 절 수는 1~6개월에 1.68개, 7~12개월에 1.97개, 13~18개월에 2.27개였다. 이를 서신웨(2022)에서 제시한 결과와 비교해 보면 13~18개월의 문장당 절 수가 토픽 3급 학습자보다 낮은 것을 알 수 있다. 그리고 김현아(2016)에서는 문장당 절 수를 모절을 제외한 문장당 대등절·내포절 수로 측정하고 있는데 이는 본고에서 모절을 포함하여 문장당 절 수를 측정한 것과 차이가 있다. 따라서 본고 역시 문장당 대등절·내포절 수를 산출하면 1~6개월에 0.68개, 7~12개월에 0.96개, 13~18개월에 1.25개였다. 이를 김현아(2016)에서 제시한 결과와 비교해 보면 1~6개월에는 재외동포 학습자의 숙달도 3~4급 수준, 7~12개월에는 4~5급 수준, 13~18개월에는 6급 수준에 도달하고 나이를 기준으로 하면 7~12개월에 재외동포 학습자의 중학교 수준보다 높고 13~18개월에는 고등학교 수준보다 높은 것으로 나타났다. 그러나 여전히 모어화자와는 2배 이상의 차이가 있으며

[79] 재외동포 청소년과 비슷한 연령대의 중·고등학교 모어화자.

모어화자 4~5학년 청소년보다 낮은 수준임을 추측할 수 있었다.

본고의 학습자들이 2019년 9월~2020년 1월 사이에 입국하였다는 점을 고려할 때, 관찰 시작 13~18개월 차는 학습자들이 입국한 지 3년 정도 되는 시점이고 공교육에 입학한 지 2년 정도 되는 시점이다. 따라서 본고의 연구 참여자인 중국 중도입국 청소년 학습자는 입국 3년, 공교육 입학 2년 정도 되는 시점에 토픽 3급 정도 수준의 통사 복잡성을 보이며 모어화자 청소년에 비해서 매우 부족한 수준의 한국어를 구사하고 있다고 볼 수 있다[80]. 오히려 학습자들의 통사 복잡성은 같은 연령대의 재외동포 청소년 학습자와 비슷한 수준을 보이고 있었는데 이는 학습자들이 모어화자와 동일한 환경에서 교육을 받기에는 어려움이 클 것임을 예상할 수 있는 부분이다.

1.2. 어휘 복잡성의 발달

어휘 복잡성은 어휘 다양도, 어휘 밀도, 어휘 정교도로 나누어 측정하였다. 문어 자료의 어휘 다양도를 살펴보면 1~6개월과 7~12개월 사이에는 뚜렷하게 증가하지만 7~12개월과 13~18개월 사이에는 그 증가폭이 크게 감소하며 발달이 정체되는 현상이 나타나고 있었다. 그리고 이러한 현상은 구어 자료에서도 동일하게 나타나고 있었다. 일원배치 분산분석 결과 문어와 구어의 어휘 다양도는 1~6개월과 7~12개월 사이에는 통계적으로 유의한 차이가 있었으나 7~12개월과 13~18개월 사이에는 유의한 차이가 나타나지 않았다. 즉 학습자들은 관찰 시작 12개월이 지난 시점부터 관찰이 끝나는 시점까지 문어와 구어 모두 어휘 다양도 발달에 정체가 나타나고 있었다.

이는 김선정(2012), 남주연(2015), Polat & Kim(2014)에서 어휘 습득이 문법보다 용이하다거나 입력이 제공되면 어휘 다양도가 향상된다고 하는 결과와 차이점을 보이며 결혼이민자의 한국어 발달 연구에서 거주 기간이 5년 이상 된 이민자는 어휘 발달이 정체된다고 한 김경령(2013)이나 여성결혼이민자의 경우 초급에서 중급으로 갈 때에는 새로운 어휘가 습득되면서 실질 어휘의 사용 비율이 높아지지만 중급에서 고급으로 가는 과정에서는 새로운 어휘의 습득이 감소한다고 보고한 허은혜(2010)와 비슷한 결과이다. 또한 한국 거주 기간이 4년 이상인 여성결혼이민자의 어휘 발달을 6개월간 살펴본 주월랑(2018) 역시 5명의 연구 참여자 중 1명을 제외하고 모두 어휘 다양도 값이 하

[80] 앞에서 설명하였듯이 세 연구의 결과와 비교를 진행한 것은 본고의 학습자들의 대략적인 한국어 수준을 가늠해보기 위한 것으로 각 연구마다 환경 변인, 과제 변인, 연령 변인에 차이가 있기 때문에 해당 결과가 학습자들의 실제 통사 복잡성 수준을 대표하는 것이 아님을 밝힌다.

락한다는 비슷한 결과를 보고하며 이주 초기의 학습자들은 사용 가능한 언어 수단 중 접근이 가장 쉬운 어휘를 집중적으로 사용하면서 언어능력을 향상시킬 것이므로 어휘가 문법보다 빠르게 습득될 것이지만 거주 기간이 어느 정도 지난 이민자는 새로운 어휘를 습득하기보다, 반복적인 실질 어휘만을 사용하면서 어휘 습득이 정체된다고 해석하고 있다. 그러나 이러한 해석은 미국에 이주한 지 2년 6개월이 된 이민자의 언어 발달을 관찰하고 통사 복잡성은 눈에 띄는 발달이 없지만 어휘 다양도는 큰 폭으로 발달하였다는 결과를 보고한 Polat & Kim(2014)의 결론에는 적용되지 않는다.

반면 대학기관 학습자의 어휘 발달을 분석한 연구들에서는 이와 다른 결과를 보고하고 있는데, 대학기관 학습자의 숙달도별 어휘 다양도를 측정한 남주연(2015)이나 서신웨(2022)에서는 중급과 고급 사이에서 유의미한 발달을 보인다고 하였고 안경화(2003), 남주연 외(2017) 역시 숙달도 단계별로 어휘 다양도가 지속적으로 증가함을 밝히고 있으며 TOPIK의 답안지를 분석한 이경(2015)에서는 어휘 다양도가 전반적인 숙달도 점수와 강한 양의 상관관계를 보인다고 하였다. 즉 대학기관 학습자들은 숙달도에 따라 어휘 다양도가 꾸준히 증가하는 것이다.

이러한 결과가 나타난 원인에 대해서는 Laufer(1991), Polat & Kim(2014)의 연구를 참고해 볼 수 있다. Laufer(1991)에서는 영어영문학과에 입학하여 원어민 교사의 수업에 참여한 학습자의 어휘 복잡성 발달을 3학기에 걸쳐 관찰한 뒤 어휘 다양도가 큰 폭으로 발달한 학습자는 입학 시험에서 학급 평균보다 낮은 수준을 보인 학습자들이었으며 처음에 높은 다양도를 보였던 학습자들은 거의 발달을 보이지 않거나 수준이 하락하기도 하였다는 점을 발견하고 평균보다 낮은 학습자들은 학과의 요구 사항에 맞춰서 평균 수준에 도달하기 위해 어휘를 발달시켰으나 높은 어휘 수준을 보였던 학습자들은 '안전한 플레이(playing it safe)'라는 전략을 채택하면서 자신에게 가장 익숙한 단어를 제한적으로 사용하였을 가능성을 제기하였다. 그리고 이를 통해 어휘 발달은 'i+1'이라는 이해가능한 입력에 의해서보다는 학습자의 필요에 의해 진보한다는 주장을 제기하였다. 이런 관점에서 볼 때, Polat & Kim(2014)의 연구 참여자가 이민국에서의 거주 기간이 꽤나 지났음에도 불구하고 어휘 다양도가 지속적으로 발달한 것은 연구 참여자가 직장인이었다는 점에서 원어민과의 필수적 의사소통의 필요성에 의한 것으로 해석할 수 있다.

이러한 결과들을 연결시켜 볼 때, 대학기관 학습자의 경우 지속적인 입력과 더불어 단어 시험, 말하기나 쓰기 활동의 필요에 의해 어휘 다양도가 지속적으로 발달하지만 여성결혼이민자나 중도입국 청소년의 경우, 일상생활에 필요한 기본적인 어휘를 익힌

이후에는 자연스러운 입력만으로 어휘 복잡성의 지속적인 발달에 한계가 나타날 수 있으며 이는 주로 규칙적인 입력의 부재와 수준 높은 언어 산출에 대한 필요성이 높지 않기 때문인 것으로 해석할 수 있다. 따라서 어휘 다양도를 지속적으로 증가시키기 위해서는 자연스러운 입력 외에 지속적이고 의식적인 어휘 교육이 이루어져야 할 것이다.

다음은 어휘 밀도의 발달에 대해 논의하도록 한다. 어휘 밀도는 실질적 의미를 나타내는 내용어의 비율을 가리키는 것으로 어휘 밀도가 높은 텍스트일수록 더 높은 문식성을 가지고 있다(Read, 2000:196)고 보는 연구도 있지만 이와 반대로 어휘 밀도가 숙달도가 다른 집단 사이나 제2언어 쓰기 능력을 변별하지 못한다고 주장하는 연구들(Wolfe-Quintero 외, 1998; Lu, 2011)도 존재한다. 본 연구에서 측정한 학습자들의 어휘 밀도는 문어와 구어에서 모두 7~12개월에 1~6개월보다 미세하게 상승 이후 13~18개월에 다시 하락하는 양상을 변동을 보였으나 그 차이가 미미하여 통계적으로 유의미할 정도의 변화가 아니었다. 즉 학습자들의 어휘 밀도는 시간의 흐름에 따라 유의미한 변동을 보이지 않고 있었다. 이는 기타 대부분의 복잡성 측정 지표가 6개월을 단위로 유의미한 변화를 보였던 점과는 다른 양상으로 중도입국 청소년의 문어와 구어의 어휘 밀도는 통사적 복잡성이나 기타 어휘 복잡성과 다른 경로로 발달한다고 해석할 수 있을 것이다.

한편, 한국어교육 분야에서 어휘 밀도를 측정한 연구들을 살펴보면 대부분 어휘 밀도가 숙달도를 구분하는 데 유용하지 못하다는 결론들을 다수 찾아볼 수 있는데 숙달도에 따른 어휘 밀도를 측정한 이경(2015), 주우동(2019), 강정희(2020), 서신웨(2022)에서는 숙달도가 올라감에 따라 어휘 밀도가 오히려 감소한다고 하였으며 남주연(2015)에서는 내용어/총 어휘 수는 등급에 따른 유의한 차이를 설명하지 못한다고 하였다. 그러나 본고를 포함한 어휘 밀도 관련 여러 연구들은 내용어와 기능어, 어휘의 범주 설정에 각각 차이가 있고 또 앞에서 살펴본 선행연구들은 본고의 학습자와 연령이나 학습 목적이 모두 다른 학습자들이기 때문에 어휘 밀도와 한국어 숙달도 간의 관계를 판단하기 위해서는 동일한 학습자를 대상으로 동일한 지표를 사용한 더 많은 연구가 이루어질 필요가 있을 것이다.

마지막으로 어휘 정교도의 발달에 대해 논의하도록 한다. 어휘 정교도는 학습자가 사용한 어휘의 난이도, 즉 어휘 수준을 보여주는 지표이다. 문어와 구어 자료에 나타난 어휘 정교도는 모두 7~12개월에 약간 하락 이후 13~18개월에 증가하는 양상을 나타내고 있었다. 또한 구어의 어휘 정교도는 7~12개월까지 문어에 비해 낮다가 13~18개월에 들어 문어보다 약간 높아지며 비슷한 수준에 도달하였다. 이는 어휘 다양도가

13~18개월에 정체되는 것과 대조되는 현상으로 학습자들은 13~18개월에 어휘 다양도가 정체되는 대신 어휘의 정교도에 더 집중하고 있는 것으로 해석할 수 있다. 그러나 문어와 구어 모두 13~18개월에 이르러서도 전체 어휘의 90% 가까이가 초급 어휘로 구성되어 있어 학습자들의 어휘 정교도가 낮은 수준임을 알 수 있었다. 원미진 외(2017)에서는 어휘 정교도가 토픽 3급 학습자와 6급 학습자의 차이를 유의미하게 구분해주는 것으로 보고하였고 강정희(2020)에서는 어휘 정교도가 중급 내 및 중급과 고급 숙달도 간에 유의한 차이가 있음을 밝히고 있다. 그리고 주월랑(2018)에서는 한국 거주 기간이 길어짐에 따라 학습자들의 저빈도 어휘가 증가하지만 습득에 소요되는 시간이 길어 발달이 느리다는 점을 지적하고 있다. 본고의 학습자 역시 어휘 정교도가 점점 증가하지만 입국 3년 정도 되는 시점까지 여전히 초급 어휘 위주로 산출하고 있음을 볼 때, 현재 수준보다 난이도가 높은 어휘의 입력과 산출을 강화하여 어휘 정교도의 발달을 촉진할 필요가 있음을 시사한다.

물론 4장에서 언급하였듯이 본고의 학습자가 지속적으로 초급 수준의 어휘를 주로 사용하는 것은 본고에서 사용된 쓰기와 말하기의 주제가 영향을 미쳤을 가능성을 배제할 수 없다. 그러나 앞에서 살펴본 통사 복잡성의 발달과 비교해 볼 때 어휘 복잡성의 발달이 전반적으로 느리거나 정체된다는 점을 확인할 수 있으며 이는 대학기관 학습자들의 통사 복잡성과 어휘 복잡성이 동반 발달하는 것과는 대조되는 양상이다.[81] 따라서 통사와 어휘의 균형적인 발달을 위해서는 Laufer(1991)에서 제기한 임계값(threshold) 이상의 어휘를 이끌어 내는 것이 중요하며 이를 위해서는 수준 높고 다양한 어휘를 활성화 할 수 있는 일관되고 체계적인 입력과 산출 활동이 이루어져야 할 것이다.

1.3. 정확성의 발달

정확성의 발달은 무오류절의 비율과 단위당 무오류절 수, 단위당 오류 수 3가지 측정 지표를 사용하여 측정하였으나 문어와 구어 모두 정확성이 증가한다고 보기는 어려웠다. 문어에서는 1~6개월과 7~12개월 사이에는 무오류절 수가 유의하게 증가하고 7~12개월과 13~18개월 사이에는 오류 수가 유의하게 증가하여 후반으로 갈수록 오류 수가 우세하였고 구어에서는 1~6개월과 7~12개월 사이에는 모든 지표에서 뚜렷한

[81] 남주연(2015)에서는 대학기관 학습자들의 통사 복잡성과 어휘 복잡성이 밀접한 상관관계를 보이며 서로 상승 작용을 한다는 것을 밝혔고 이복자(2015)에서는 통사 복잡성과 어휘 다양도가 서로 경쟁과 지지관계를 반복하지만 대체적으로 함께 증가한다는 것을 밝혔다.

발달이 없다가 13~18개월에 들어 무오류절 수와 오류 수가 동시에 증가하며 측정 지표에 따라 서로 다른 결과를 보이고 있었다. 이는 통사 복잡성의 발달이 문어와 구어에서 모두 시간의 흐름에 따라 안정적인 증가를 보였던 것과 다른 양상이다.

숙달도나 시간의 흐름에 따른 정확성 발달을 살펴본 기존의 연구 결과들 역시 통사 복잡성과 달리 일치되는 결론을 보이지 않는다. 숙달도별 정확성을 분석한 박지순·서세정(2009), 서세정(2009), 주우동(2019), 서신웨(2022) 등 연구에서는 숙달도가 높아질수록 학습자들의 정확성이 증가한다고 밝히고 있는 반면, 종적으로 단위당 오류 수를 살펴본 이복자(2015)의 경우 쓰기와 말하기에서 모두 정확성이 발달하지 않는다고 하였고, 쓰기의 문장당 오류 수를 측정한 주월랑(2018)에서 역시 문장이 길어지거나 확장될 때 정확성은 낮아진다고 보고하고 있다. 또 김정운 외(2017b)에서는 거주 기간이 늘어남에 따라 쓰기의 무오류절 비율은 약간 증가하는 추세이지만 말하기 무오류절 비율은 뚜렷한 증가를 보이지 않고 상승과 하락을 반복한다고 하였다. 그리고 Ortega(1999:137)에서 역시 제2언어 발달의 모든 요소가 동등하게 발달하지 않는다는 점을 강조하면서 중간언어 발달에서 정확성에 비해 복잡성과 유창성이 선형적으로 발달한다는 점을 언급하고 있다.

이러한 결과들을 종합해볼 때, 정확성은 숙달도를 바탕으로 구분할 때에는 어느 정도 차이를 보일 수 있으나 이어지는 시간 속에서 종단적으로 살필 경우 등락이 매우 빈번하여 오류의 감소가 명확히 드러나기 힘든 것으로 추측해볼 수 있다. 특히 이러한 추측은 자체적으로 개발한 문항을 바탕으로 2등급~9등급의 숙달도 단계를 구분하고 문어와 구어 모두 복잡성은 거의 선형적으로 증가하지만 정확성은 일시적으로 주춤하거나 퇴행하는 단계가 빈번히 발생한다는 점을 발견한 김영주 외(2013)의 연구 결과나 6개월 간 4차례의 종적 발달을 살핀 뒤 복잡성, 유창성과 달리 정확성에 더 큰 굴곡이 있음을 발견한 Larsen-Freeman(2006)의 연구, 6개월에 걸쳐 학습자의 쓰기 과제를 관찰한 뒤 학습 기간이 지속됨에 따라 정확성은 뚜렷하게 향상되지 않았지만 담화 수준 및 언어적 복잡성은 향상되었다고 보고한 Benevento & Storch(2011), 그리고 일본어 초급 학습자 7명을 대상으로 2주 간격으로 총 4회 구어를 분석한 뒤 유창성과 통사 복잡성은 시간이 지남에 따라 향상되었지만 정확성과 어휘 복잡성은 개별 학습자마다 상승과 하락의 곡선이 달랐다고 보고한 Fukuda(2014) 등의 연구 결과들을 참고해볼 때, 복잡성에 비해 정확성의 발달이 더 어려우며 단기간에 뚜렷한 차이를 보이기 힘든 것으로 해석할 수 있다.

그러나 같은 종단연구임에도 불구하고 대학 기관의 학문목적 학습자의 말하기 능력

을 10개월 간 살핀 김아름(2017)에서는 오류 수가 점점 감소한다는 결과를 보고하고 있으며 앞에서 살펴본 종단연구들은 이복자(2015)를 제외하고 모두 대학기관 학습자들을 대상으로 한 것이 아님을 발견할 수 있는데, 자연 학습 환경의 학습자를 대상으로 거주 기간에 따른 통사 발달을 살펴본 김정운 외(2017a)에서는 문어와 구어의 복잡성은 발달했으나 정확성의 발달은 유의하지 않았다는 점을 지적하면서 자기 주도적인 학습은 정확성의 발달을 가져오는 데 역부족이라고 지적하기도 하였다. 따라서, 본고의 정확성 분석 결과는 명시적인 입력과 피드백을 받는 대학기관의 학습자들의 경우 정확성이 빠르게 발달할 수 있으나 결혼이민자나 자연 학습 환경에 있는 학습자들은 정확성의 발달이 이루어지기 어렵다고 해석할 수도 있을 것이다.

정리하면, 학습자들의 정확성과 복잡성은 항상 동반적으로 발달하지는 않으며 별도의 한국어 교육 없이 공교육 환경에 있는 중국 중도입국 청소년 학습자들의 경우 통사 복잡성은 자연스럽게 발달할 수 있지만 정확성은 그렇지 않다. 그리고 이러한 정확성 발달의 정체 현상은 주로 명시적 교육 및 교정의 부족과 의미 전달 위주의 의사소통에서 그 원인을 찾을 수 있을 것이다.

그러나 한편으로 정확성의 발달은 처음에는 오류가 없다가 오류가 대폭적으로 증가하는 과정을 거쳐 다시 목표어에 점점 근접하는 정확한 언어를 생산하는 U자형 발달 패턴을 보인다(Lightbown, 1983; Kellerman, 1985)는 점과 복잡성의 발달은 정확성과 반대인 Ω형 발달을 보인다(Wolfe-Quintero 외, 1998)는 일반적인 언어 발달 패턴을 고려할 때, 관찰 기간 동안 본고의 학습자들은 정확성보다 복잡성의 발달에 집중하는 단계로 복잡성이 증가하고 정확성이 하락하는 단계에 처해있으며 시간이 더 흐른다면 다른 정확성이 자연스럽게 향상될 수 있는 가능성을 배제할 수는 없다. 그럼에도 불구하고 중도입국 청소년의 경우 빠르게 한국어 수준을 높여 학업에 적응을 해야 한다는 점과 언어 발달에서 교정이 제대로 이루어지지 않으면 화석화(fossilazaiton)가 발생하게 된다는 점들을 고려할 때, 학습자들의 정확성을 촉진하기 위한 명시적인 교육이 이루어질 필요가 있을 것이다.

2. 산출 양식에 따른 복잡성과 정확성의 관계

본고의 두 번째 연구 문제는 학습자들의 복잡성과 정확성이 산출 양식에 따라 차이를 보이는지, 차이를 보인다면 어떤 차이를 보이는지를 살펴보는 것이다. 이를 위해 본 연구에서는 전체 기간(18개월) 및 6개월 단위의 각 구간별로 산출 양식에 따른 복잡성

과 정확성의 차이를 비교하고 대응표본 t검정(t-test)을 실시하여 그 차이가 통계적으로 유의한지 검증하였다. 아래에서는 본고에서 분석한 결과를 바탕으로 산출 양식에 따른 언어 수행의 차이에 대해 논의하도록 한다.

2.1. 산출 과정 측면에서 나타난 차이

과제 복잡성이 언어 산출에 영향을 미친다는 것은 널리 알려진 사실이다. 과제 복잡성은 인지적 복잡성, 과제 수행 조건(참여자, 개방형/폐쇄형 과제)이나 학습자 요인(태도, 동기, 불안) 등 다양한 요인의 영향을 받는다. 그러나 과제 복잡성은 과제를 수행하는 방식에 따라서도 달라질 수 있다(Kuiken & Vedder, 2011). 2장의 이론적 배경에서 서술한 바와 같이 문어와 구어는 긴밀히 연결되어 있어 서로 영향을 주는 언어 능력으로 볼 수도 있으나 구어는 자동적(spontaneous)이며, 의식의 통제를 받지 않는(un-self-conscious) 자연스러운(natural) 언어이고 문어는 자기 점검적(self-monitoring)이고, 통제적(controlled)이며 자기 의식적(self-conscious)인 언어라는 점에서는 양 극단의 양식이기도 하다(Halliday, 1987). 또한 문어와 구어의 산출 절차는 비슷한 단계로 구성되어 있지만 동시에 차이가 존재하며 가장 큰 차이는 산출의 실시간성과 산출 형태에서 나타난다. 그리고 산출의 실시간성은 언어 자원에 접근할 시간과 모니터링의 기회에 영향을 미치며 산출 형태는 '쓰기 행동'에 대한 인지적 부담으로 나타나게 된다. 따라서 본고에서는 기본적으로 구어보다 문어에서 더 나은 수행을 보일 것이지만 실시간성과 언어 산출에 대한 부담의 크기에 따라 다른 결과가 나타날 수 있을 것으로 예상하였다.

이러한 가정을 확인하기 위해 전체 기간을 단위로 문어와 구어의 통사 복잡성과 어휘 복잡성, 정확성을 비교해본 결과 통사 복잡성(단위 길이, 단위당 절 수)과 어휘 복잡성(어휘 다양도)은 문어에서 더 높다는 결과를 얻었고 정확성은 무오류절의 비율, 단위당 무오류절 수, 단위당 오류 수 세 가지 측정 지표 모두 문어와 구어 사이에 유의한 차이가 없는 것으로 나타났다.

선행연구를 살펴보면 어휘 다양도 측면에서는 대부분의 연구에서 일관된 결론을 찾아볼 수 있는데 Ellis & Yuan(2005), Granfeldt(2008), Kuiken & Vedder(2011), Kormos & Trebits(2012), Chan 외(2014) 등 여러 연구에서 모두 문어에서의 어휘 다양도가 구어보다 높음을 밝히고 있다. 이에 초급 영어 학습자 일란성 쌍둥이의 어휘 발달을 비교한 Chan 외(2014)에서는 문어와 구어의 어휘 난이도(정교도)는 차이가 없었

으나 문어의 어휘 다양도가 구어보다 높다는 결과에 대해, 초급 학습자일수록 수행 시간이 더 긴 쓰기에서 다양도가 더 높게 나타나지만 쓰기에 더 많은 시간을 할애한다고 해서 어휘 사용의 난이도가 높아지지는 않는다고 해석하기도 하였다. 이러한 결과는 본 연구의 결과와도 일치하는 것으로 본 연구의 학습자 역시 아직 초급 단계라는 점에서 쓰기에 주어지는 추가적인 수행 시간이 더 다양한 어휘를 산출하는 데 도움이 되었지만 이러한 추가적인 수행 시간이 언어의 정교도에는 크게 영향을 주지 않은 것으로 예상해볼 수 있다. 이 외에도 한국어 균형 말뭉치를 사용하여 구어와 문어의 어휘 다양도를 비교한 안의정(2017)이나 중급 혹은 중급 이상 수준의 학습자를 대상으로 연구한 Granfeldt(2008) 등 기타 연구들에서도 문어의 어휘 다양도가 더 높다는 결과들을 참고해 볼 때, 문어의 어휘 다양도가 구어에 비해 더 높게 나타나는 현상은 숙달도나 언어권에 상관없이 제2언어 학습자에게 보편적으로 나타나는 현상으로 볼 수 있으며 동시에 어휘 다양도는 계획 시간의 영향을 많이 받는 요인임을 의미하기도 한다.

산출 양식에 따른 어휘 복잡성에 대한 연구들 대부분이 학습자들이 문어에서 더 나은 수행을 보인다는 일관된 결과를 보이는 것과 달리 산출 양식에 따른 통사 복잡성이나 정확성을 비교한 연구들은 연구마다 서로 다른 결과를 보이고 있었다. 먼저 Granfeldt(2008)에서는 어휘 다양도는 구어보다 문어에서 유의미하게 높았지만 문법 복잡성은 그렇지 않았으며 정확성은 오히려 구어에서 더 높게 나타났다는 결과를 보고하면서 말하기에 비해 쓰기 활동이 통제, 계획, 모니터링 가능성이 더 높기 때문에 문어에서 더 높은 복잡성과 정확성을 보일 것이라는 예상과 다르다는 점을 보고하였고 Sarah Yazdani(2018)에서는 초·중·고급 모두 문어의 통사 복잡성이 높다는 결과를 보이기도 하였다.

한편, 문어와 구어의 복잡성과 정확성을 비교한 연구들 중에는 과제의 특성이 쓰기와 말하기의 산출 능력에 영향을 끼칠 것이라고 보는 연구들도 있었다. Kuiken & Vedder(2011)와 Kormos & Trebits(2012)는 복잡한 과제와 단순한 과제로 나누어 과제를 제시하고 산출 양식에 따른 복잡성과 정확성을 비교하였다. 그 결과, Kuiken & Vedder(2011)에서는 단순한 과제와 복잡한 과제 모두 통사 복잡성과 어휘 다양도는 문어에서 더 높지만 정확성은 구어에서 더 높다는 결과를 도출하였다. 동시에 이 연구에서는 과제 복잡성이 문어의 통사 복잡성에는 큰 영향을 끼치지 않지만 구어에서는 복잡한 과제에서 덜 복잡한 문장을 생성한다는 점을 발견하고 이러한 이유로 L2 학습자는 말을 할 때 단순화하게 된다고 보았다. 반면 Kormos & Trebits(2012)에서는 단순한 과제와 복잡한 과제에서 모두 쓰기의 어휘 다양도와 정확성이 더 높지만 통사 복

잡성에는 큰 차이가 없으며 정확성은 단순한 과제에서만 차이를 보인다(문어>구어)는 점을 밝히며 Kuiken & Vedder(2011)과 반대되는 결과를 도출하기도 하였다. 이러한 결과에 대해 Kormos & Trebits(2012)에서는 단순한 과제에서 쓰기의 정확성이 말하기에 비해 유의미하게 높은 것은 연속된 만화를 보고 내용을 서술하는 단순한 과제보다 순서가 정해져 있지 않은 그림들을 보면서 자유롭게 이야기를 만들어내는 개념화 요구가 더 높은 과제에서는 학습자들이 L2 형태를 인코딩하고 모니터링하는 것보다 내용 계획에 주의 자원을 할애하기 때문인 것으로 보았다. 두 연구는 차이점도 존재하지만 과제가 복잡해질 때 말하기보다 쓰기에서 내용에 더 주의를 기울이게 되고 이것이 통사 복잡성을 증가시킴과 동시에 정확성에 대한 주의를 감소시킨다고 보는 입장은 동일하다. 즉 이론적으로는 실시간 산출인 말하기보다 쓰기가 정확성에 대해 모니터링할 수 있는 특성이 높지만, 이런 쓰기의 순환적 특성은 개념화의 요구가 없고 특정 콘텐츠의 언어적 부호화가 필요한 만화 설명 과제와 같은 과제에서만 정확도를 높일 수 있다는 것이다. 이는 그림을 보고 자유롭게 이야기를 만들어내는 개방형 과제에서 순서가 정해져 있는 폐쇄형 과제에 비해 복잡성은 더 높지만 정확성이 더 낮은 언어를 산출한다고 주장한 Foster & Skehan(1996)과 Skehan & Foster(1997)의 연구 결과와 상통한다.

이런 관점을 바탕으로 본고의 연구 결과를 논의해 볼 때, 본고에서 제시한 과제는 주제를 제시하고 주제에 따라 자유롭게 내용을 구상해야 하는 과제로 개념화에 대한 요구가 비교적 높은 과제로 볼 수 있다. 따라서 이러한 본고의 과제 특성이 학습자들의 쓰기 통사 복잡성과 어휘 복잡성을 높이는 것에는 긍정적인 역할을 하였을 수는 있지만 구어의 복잡성을 낮추는 결과를 초래했을 수 있다.[82] 그러나 낮아진 구어의 복잡성이 정확성으로 이어지지는 않았다. 이는 앞에서 살펴본 것과 같이 말하기의 실시간성으로 인한 모니터링 부족에서 기인한 것으로 볼 수 있는데 특히 구어에서 대치 오류가 문어에 비해 높게 나타난 점은 모니터링 기회의 부족으로 인한 오류의 증가를 어느 정도 설명할 수 있을 것이다. 정리하면, 본고의 학습자는 복잡성 측면에서는 문어에서 더 나은 수행을 보이면서 제2언어 학습자들은 언어 자원에 접근할 수 있는 시간이 더 많은 쓰기에서 더 나은 성과를 거둘 가능성이 높다고 한 Ellis & Yuan(2005)의 주장을 지지하는 결과를 보여주었으나 정확성 측면에서는 기존의 연구와 일치하는 결론을 보이지 않

[82] Kuiken & Vedder(2011)에서는 연구 결과에 대해 온라인 과제(구어)에서는 오프라인 과제(문어)에 비해 '압박감(pressure)'이 더 크다고 가정하는 것이 합리적이라는 점을 밝히며 이러한 '압박감' 때문에 L2 학습자는 말을 할 때 단순화하도록 강요받으며, 이는 구어에서 문법적 복잡성이 감소하는 결과로 이어진다고 주장한다.

았다.

그러나 본고에서 언급한 요인 이외에도 문어와 구어의 수행을 비교한 대한 연구들은 학습자가 형태론적 구성을 처리할 때 '양식 선호도(modality preferences)'가 있어 구어나 문어 중 한 가지를 더 선호할 수 있다는 주장(Weissberg, 2000)이나 과제에 따라 언어 인코딩 부하가 더 높거나 개념화 부하가 더 높기 때문에 각각의 인지적 부담 형태가 산출 양식에 미치는 영향을 고려해야 한다는 주장(Kormos & Trebits, 2012), 쓰기 행위 자체가 차지하는 인지적 부담 때문에 문어의 수행이 더 낮을 수 있다는 주장(Bourdin & Fayol, 1994; Grabowski, 2004) 등 다양한 견해들을 찾아볼 수 있다. 따라서 수행해야 할 과제의 특성이나 학습자의 수준, 학습자의 특성을 고려하지 않은 채 문어와 구어의 수행에 대한 일반적인 특성을 도출하기에는 어려움이 있을 것이다.

2.2. 발달 경로 측면에서 나타난 차이

산출 양식에 따른 언어 수행의 차이는 전체 기간을 단위로 비교하는 것도 의미가 있지만 문어와 구어의 발달 경로에 차이가 있다는 점에서 6개월을 단위로 묶어 각 기간에 나타난 문어와 구어의 차이를 비교해 본 결과 다음과 같은 결과를 확인할 수 있었다.

먼저, 통사 복잡성에서 단위 길이의 경우, 1~6개월에는 구어의 단위 길이가 문어에 비해 낮은 수준이었으나 시간이 지날수록 구어 단위 길이의 증가 속도가 문어에 비해 빨라지면서 13~18개월에 이르러서는 구어의 단위 길이가 문어에 비해 길어지고 있었다. 단위당 절 수 역시 1~6개월에는 구어에서 더 낮은 수준을 보였으나 더 빠른 속도로 증가하여 시간이 지날수록 문어와 구어의 차이가 점점 좁혀지고 있었다. 즉 통사 복잡성은 전체 기간의 평균에서는 문어에서 더 높게 나타났으나 구어에서 발달 속도가 더 빠른 것으로 나타났다. 다음, 어휘 복잡성은 통사 복잡성과 다른 경향을 보였는데 어휘 다양도는 1~6개월에 문어와 구어 간의 차이가 유의하지 않던 데로부터 7~12개월과 13~18개월에는 문어가 유의하게 앞서고 있었으며 후반으로 갈수록 그 차이가 커지고 있었다. 그리고 어휘 밀도와 어휘 정교도는 세 기간 모두 유의한 차이를 보이지 않았다. 마지막으로 정확성의 경우, 무오류절의 비율에 있어서는 세 기간 사이에 뚜렷한 증가나 감소 없이 동일한 수준을 유지하고 있었으며 문어와 구어 간에도 차이를 보이지 않았다. 그리고 단위당 무오류절 수와 단위당 오류 수는 모두 시간이 흐름에 따라 증가하였지만 7~12개월에 단위당 무오류절 수가 문어에서 구어에 비해 유의하게 높은 것을 제외하고 6개월을 단위로 한 각 기간과 전체 기간 모두 문어와 구어의 차이가 유의하

지 않은 것으로 나타나 전체적으로 정확성에 있어서는 문어와 구어에 차이가 없는 것으로 나타났다. 즉, 학습자들은 초기에는 구어보다 문어의 통사 복잡성이 더 높지만 한국 거주 기간이 늘어날수록 구어 통사 복잡성이 더 빠르게 발달하여 문어와 비슷하거나 더 높은 수준에 도달한다. 그러나 어휘 복잡성은 통사 복잡성과 달리 한국 거주 기간이 길어질수록 문어와 구어 사이의 간극이 점점 더 커지고 있었다.

숙달도나 거주 기간에 따른 문어와 구어의 산출을 살핀 연구들을 살펴보면, 숙달도를 바탕으로 문어와 구어의 산출을 비교한 김영주 외(2013)에서는 정확성은 모든 숙달도 수준에서 문어의 정확성(무오류절 비율)이 더 높았으며 복잡성은 5급을 제외한 모든 숙달도에서 구어의 복잡성(종속절 비율[83])이 더 높다고 보고하였다. 이와 다르게 자연 학습 환경의 학습자의 거주 기간에 따른 통사 발달을 관찰한 김정운 외(2017a)에서는 모든 집단에서 문어의 통사 복잡성이 구어에 비해 높게 나타났고 구어의 정확성이 문어에 비해 높다는 결론을 보여 김영주 외(2013)와 완전히 상반되는 결론을 제시하고 있다. 반면 동일하게 자연 학습 환경의 학습자의 거주 기간에 따른 통사 발달을 종적으로 고찰한 김정운 외(2017b)에서는 문어와 구어의 복잡성(단위당 절 수) 차이는 서로 등락을 반복하며 일정한 차이를 보이지 않지만 구어의 정확성(단위당 무오류절 수)이 더 높다는 결론을 제시하여 김정운 외(2017a)와 비슷한 결과를 제시하였다.

이 연구들은 모두 문어와 구어의 산출을 함께 다루고 있다는 점에서는 본고와 동일하지만 학습 언어, 숙달도 수준, 학습 환경이나 실험 과제 등 면에서 모두 차이가 있기 때문에 단순 비교를 하는 것에는 무리가 있다. 그럼에도 불구하고 공통적인 경향성을 찾아보자면 김영주 외(2013)를 제외하고 대부분의 연구에서 통사 복잡성은 문어에서 더 높게 나타나고 정확성은 구어에서 더 높게 나타나고 있다는 점이다. 앞에서 제시하였듯이 본 연구의 학습자는 문어의 복잡성이 더 높게 나타났고 문어와 구어의 정확성의 차이가 유의하지 않았다는 점에서 복잡성에 있어서는 김영주 외(2013)와 반대되지만 김정운 외(2017a)와 같은 결론을 보이고 정확성에 있어서는 세 연구 모두와 다른 결론을 보였다.

위의 결과들을 종합해 볼 때, 중국 배경 중도입국 청소년들은 언어 지식에 더 충분히 접근할 시간이 주어지는 문어에서 더 복잡한 언어를 산출하지만 문어와 구어에서 비슷한 수준의 정확성을 보여주고 있었다. 그러나 4장에서 확인하였듯이 문어와 구어의 정확성을 저해하는 요인은 서로 다르다. 먼저, 문어에서는 문법 오류가 더 높았으나 구어

[83] 김영주 외(2013)에서는 본고에서 내포절로 본 명사절, 관형사절, 부사절, 인용절을 모두 종속절로 구분하고 있으며 서술절도 종속절로 포함시켜 분석하였다.

에서 어휘 오류가 더 높았고, 어휘 영역을 제외한 모든 누락 오류, 모든 영역의 오형태 오류가 문어에서 더 높았으며 어말 어미를 제외한 모든 영역의 대치 오류가 구어에서 더 높았다는 점을 고려할 때, 산출 양식은 단순히 정확도가 아닌 오류의 영역 및 현상과 구체적으로 연관됨을 예상할 수 있다. 동시에 6개월을 간격으로 살펴본 문어와 구어의 차이는 학습자들의 문어와 구어는 동일한 속도로 발달하는 것이 아니며 각 요인의 발달 속도도 다르다는 점을 보여준다.

3. 산출 양식과 통시적 흐름에 따른 복잡성과 정확성의 발달 관계

본고의 세 번째 연구 문제는 중국 배경 중도입국 청소년의 한국어 발달에 나타난 복잡성과 정확성의 관계를 고찰하는 것이다. 이를 위해 본고에서는 상관분석(Pearson correlation coefficient)을 통해 전체 기간에 나타난 복잡성과 정확성의 관계와 6개월 단위의 각 구간에 나타난 복잡성과 정확성의 관계를 분석하여 복잡성과 정확성의 관계가 어떻게 변화·발달하는지, 이러한 관계가 문어와 구어에서 어떻게 다르게 나타나는지 살펴보았다.

3.1. 전체 기간에 나타난 복잡성과 정확성의 관계

제2장에서 살펴본 것과 같이 복잡성과 정확성의 관계에 대한 고찰은 매우 오래 전부터 이루어져 왔으며 학자들은 CAF구인이 서로 연결되어 있음에 주목하여 세 구인이 어떤 관계로 발달하는지를 탐구하기 위해 많은 연구들을 진행해 왔다. 그러나 복잡성과 정확성의 관계에 대한 대표적인 두 이론은 학습자가 과제를 수행하는 과정에서 주의력 자원(attentional resource)을 분배하는 방식에 대해 상반되는 견해를 보이는데 ‘제한된 주의력 모델(Skehan, 1996a, 1998)’에서는 인간의 인지적 자원은 제한적이며 복잡성이 증가하면 정확성이 하락하는 상쇄 효과가 나타난다고 주장하는 반면 ‘다자원 주의력 모델(Robinson, 2001a, 2005)’에서는 학습자가 복잡성과 정확성에 동시에 집중할 수 있다고 주장한다. 두 모델은 모두 복잡한 과제, 즉 인지적 부하가 높은 과제는 더 복잡한 언어를 이끌어낼 수 있다는 점에 동의하지만 복잡한 내용에 대한 인지적 자원의 분배가 정확성에 어떤 영향을 미치는지에 대한 입장은 서로 상반된다. 즉 ‘제한된 주의력 모델’에서는 두 요인이 서로 경쟁 관계에 있다고 보고 ‘다자원 주의력 모델’에서는 두 요인이 서로 협력하는 관계에 있다고 보는 것이다.

이러한 주장과 관련된 본고의 연구 결과를 확인해 보면, 전체 기간을 단위로 볼 때 학습자들의 통사 복잡성과 정확성은 문어와 구어에서 모두 복잡성이 증가하면 무오류절 수와 오류 수가 동시에 증가하면서 지지적 관계와 경쟁적 관계가 동시에 나타나고 있었다. 그리고 어휘 복잡성은 문어에서는 어휘 다양도와 문장당 오류 수만이 정적 상관관계를 나타내고 있었고 구어에서는 어휘 다양도와 오류 수 및 무오류절 수가 약한 정적 관계를 보여 통사 복잡성과 마찬가지로 정확성과 경쟁적·지지적 관계를 보였으며 어휘 밀도는 무오류절 수와 약한 정적 관계, 어휘 정교도는 오류 수와 약한 정적 관계를 보였다. 따라서 전반적으로 학습자들은 복잡성과 정확성이 동반 발달한다고 보기는 어려웠다.

과제 복잡성 측면에서 복잡성과 정확성의 관계를 밝히고자 한 연구들을 살펴보면 복잡한 과제가 더 복잡한 언어를 이끌어 낸다는 결과(Song, 2005; Oh & Lee, 2012)는 대부분 동일하지만 복잡성과 정확성의 관계에 대해서는 다양한 결과들이 존재함을 확인할 수 있다. Cho(2018)와 Afshar & Tofighi(2021)에서는 Skehan(1998)의 결과와 동일하게 복잡성과 정확성이 상호 경쟁 관계를 보인다고 하였고 Awwad & Tavakoli(2022)와 Zalbidea(2017)에서는 Robinson(2005)과 마찬가지로 지지 관계를 보인다고 하였으며 Kormos & Trebits(2011)나 Hyun & Lee(2018)에서는 과제 복잡성에 따라 언어 산출에 차이를 보이지 않는다고 하였다. 그리고 Oh & Lee(2012)에서는 단방향(독백) 복잡과제는 복잡성을 증가시키지만 정확도를 증가시키지는 않으며 학습자 언어의 정확성은 복잡성과 지지나 경쟁 관계를 보이지 않는다고 해석하며 언어의 정확성은 과제 내용이나 복잡성보다는 양방향(대화) 과제에서 상호작용에 의해 높아지며 과제의 형식에서만 차이를 보인다고 하였다. 이처럼 복잡성과 정확성의 관계에 대해서는 다양한 연구들이 존재하지만 각 연구들에서 설정한 과제 복잡성이나 과제 조건들이 상이하여 일관된 견해를 찾아보기는 어려웠으며 과제 복잡성과 복잡성 및 정확성의 관계는 다양한 요인의 영향을 받는다는 점을 알 수 있었다.

Skehan(2009a)에서는 과제 복잡성 증가에 따른 복잡성과 정확성의 관계에 대한 연구들이 서로 다른 결과를 보인다는 점과 일부 연구에서 주장하는 복잡성과 정확성의 관계가 단순히 과제의 영향보다는 준비 시간과 같은 기타 요인의 영향일 수 있는 가능성, 그리고 인지가설에서 주장하는 과제의 복잡성을 구분하는 방법에 대한 정의에 의문을 제기하며 Levelt(1993, 1999)의 음성 생산 모델에 기대여 '제한된 주의력 모델' 가설을 강화하기도 하였다. 이 연구에서는 제1언어 말하기 모델인 Levelt(1993, 1999)와의 연계를 통해 과제의 복잡성을 언어 산출 과정과 연계하여 다시 정리하여 인지적

복잡성이나 작업 기억에 대한 요구는 '개념화' 단계에 대한 부담을 높이고 어려운 어휘에 대한 요구나 과제의 협상 불가성(non-negotiability)은 어휘 인코딩에 대한 부담을 높이며 시간적 압박이나 방대한 입력은 문법 인코딩에 대한 부담을 높인다고 보았다. 즉 과제의 복잡성을 개념적 영역과 언어적 영역으로 구분하고 개념적 영역이 높을 경우 언어적 영역에 대한 주의가 감소한다고 보는 것이다.

이러한 관점을 본 연구의 결과와도 일맥상통하는 부분이 있다. 본고의 쓰기와 말하기 과제는 추론이나 논쟁을 하는 과제가 아니라는 점에서 논증적 쓰기나 주장하기와 같은 과제에 비해서는 개념화에 대한 요구가 낮은 것으로 볼 수 있으나 일반적으로 과제 복잡성의 영향을 측정하는 과제에서 자주 사용되는 '그림을 보고 서술하기'와 같은 과제에 비해서는 개념화 단계에 대한 요구가 높다. 그리고 이러한 인지적 부담은 학습자들의 언어의 복잡성은 높였을 수 있지만 어휘나 문법에 대한 자동화가 되어 있지 않은 제2언어 학습자의 경우 '개념화'에 대한 인지적 부담으로 인해 '형식화'가 중단되어 병렬처리가 불가능해지고(Kormos, 2006; Skehan, 2009b) 결과적으로 복잡성과 정확성이 동시에 증가하지 못하는 결과를 낳은 것으로 해석할 수 있다.

또한 이와 같이 복잡성과 정확성의 관계를 언어 산출 과정과 연결시키는 주장은 산출 양식에 따른 복잡성과 정확성의 관계 차이에 대해서도 해석이 가능하다. 본고의 연구 결과, 문어의 통사 복잡성과 무오류절 수의 상관계수(.617)는 오류 수와의 상관계수(.566)보다 높은 것에 반해 구어의 통사 복잡성과 오류 수의 상관계수(.603)는 무오류절 수와의 상관계수(.519)보다 더 높게 나타나 문어에서는 지지적 관계가 좀 더 강한 반면 구어에서는 경쟁 관계가 좀 더 강한 것으로 나타났다. 이는 문어보다 구어에서 복잡성과 정확성의 경쟁 관계가 더 뚜렷할 것이라는 예상과 일치하는 결론으로 문어와 구어는 개념화와 형식화, 실행의 동일한 단계를 거치지만 문어의 실시간성으로 인한 '모니터링' 단계에 대한 주의 부족이 구어의 복잡성과 정확성의 경쟁 관계를 더 뚜렷하게 했을 것으로 해석할 수 있다.

그러나 한편으로 주의력 자원, 기억, 처리 능력과 같은 학습자의 인지적 요인도 학습자가 주의력 자원을 총괄하고 분배하는 방법에 영향을 미칠 수 있다고 보는 연구(Robinson & Gilabert, 2007:168)나 학습자마다 형태의 다른 측면에 집중할 수 있다고 보는 주장(Skehan, 2009a)들도 다양하게 존재하기 때문에 복잡성과 정확성의 관계는 단순히 일반화하기는 어려우며 과제의 복잡성이나 산출 양식, 학습자의 인지적 요인과 같은 다양한 변수를 통제함으로써 더 자세히 확인해야 할 것이다.

3.2. 시간의 흐름에 따른 복잡성과 정확성의 관계 변화

이어서 학습자들의 복잡성과 정확성의 관계의 시간의 흐름에 따른 변화에 대해 논의해 보도록 한다. 6개월을 단위로 살펴본 복잡성과 정확성의 관계는 전체 기간과 마찬가지로 통사 복잡성과 무오류절 수, 오류 수가 모두 정적 상관을 보이고 있었다. 그러나 각 기간별로 상관계수를 기간별로 상관계수를 살펴보면 흥미로운 점을 발견할 수 있는데 문어에서는 1~6개월, 7~12개월에는 통사 복잡성과 무오류절 수의 상관계수가 오류 수와의 상관계수에 비해 높지만 13~18개월에는 오류 수와의 상관계수가 더 높아지고 구어에서는 세 기간 모두 통사 복잡성과 오류 수의 상관계수가 무오류절 수와의 상관계수에 비해 더 높으며 상관계수가 점점 높아지는 것을 확인할 수 있다. 이를 통해 통사 복잡성과 정확성은 문어에 비해 구어에서 경쟁 관계가 더 강하게 나타나며 구어에서는 6개월을 단위로 한 세 기간 모두에서 경쟁 관계가 더 뚜렷하고 문어에서 역시 시간이 지날수록 경쟁 관계가 더 뚜렷해진다는 것을 알 수 있었다.

기간별 어휘 다양도와 정확성은 문어에서만 1~6개월에 정확성과 경쟁한 뒤 정확성과 유의미한 관계를 나타내지 않았으며 어휘 정교도는 구어에서만 1~6개월과 13~18개월에 정확도와 경쟁 관계, 어휘 밀도는 구에서만 7~12개월과 13~18개월에 정확성과 지지 관계로 발달하고 있었다. 이는 다양한 어휘의 사용은 문어와 구어에서 모두 정확성에 크게 영향을 주지 않지만 구어에서 어려운 어휘를 사용하려는 시도는 정확성에 영향을 주게 된다고 해석할 수 있다. 그리고 이를 다시 쓰기와 말하기의 산출 절차와 연결시켜 볼 때, 문어보다는 실시간으로 언어 자원에 접근해야 하는 구어에서 어려운 어휘의 복잡성에 대한 시도는 정확한 의미나 형태에 대한 주의 집중을 방해하게 되는 것으로 해석할 수 있다.

그러나 학습자들이 복잡성과 정확성을 동시에 발달시키지 못하는 이유는 문어와 구어에서 차이가 있었는데 정확성의 발달을 가장 크게 저해하는 요인은 문어와 구어 모두 조사 오류였지만 조사 오류를 제외하면 문어에서는 어미 오류가 13~18개월에 크게 증가하였고 구어에서는 어휘 오류가 13~18개월에 크게 증가하며 정확성의 구축에 영향을 주고 있었다. 또한 크게 보았을 때 문어에서는 문법 오류가 더 높게 나타나고 구어에서는 어휘 오류가 더 높게 나타났다는 점에서 산출 양식에 따라 정확성 교육이 다르게 이루어질 필요가 있음을 시사하기도 한다.

기간별로 살펴본 복잡성과 정확성의 관계와 관계의 변화는 언어의 발달이 고정적이거나 선형적인 것이 아니며 이 과정에서 언어의 각 차원은 서로 지지와 경쟁 관계를 반

복하며 끊임없이 변한다고(Bulté & Housen, 2017) 주장하는 역동체계이론의 관점을 지지하는 것으로 볼 수 있다. 또한 이러한 현상은 앞서 언급한 복잡성과 정확성의 발달 경로와도 관련이 있을 수 있는데 복잡성은 Ω형 발달을 보이고 정확성은 U자형 발달 패턴을 보인다는 점과 복잡성이 먼저 출현한 뒤에 정확성이 증가한다는 Housen 외(2012:7-8)의 주장들을 연결시켜 볼 때, 아직 본고의 학습자는 통사 복잡성이 발달하는 단계에 있으며 이러한 단계적 특성으로 인해 복잡성과 정확성의 경쟁 관계가 점점 심해지는 것으로 해석할 수 있다.

4. 복잡성과 정확성의 발달과 학습자 변인

본고의 네 번째 연구문제는 학습자별 복잡성과 정확성의 발달에 어떤 공통점과 차이점이 있는지를 비교하는 것이다. 이를 위해 일원배치 분산분석(One-Way ANOVA)을 실시하여 학습자 간의 복잡성과 정확성의 차이를 확인하였다. 이후 학습자별로 문어와 구어의 발달이 어떻게 다르게 나타나는지 비교하기 위해 대응표본 t-검정(t-test)을 실시하여 각 학습자의 문어와 구어에 나타난 복잡성과 정확성을 비교하였다. 아래 4장에서 분석한 결과를 종합하여 제시하고 분석 결과에 대해 논의를 진행하도록 한다.

4.1. 학습자 간 발달 경로와 발달 수준 차이

학습자 간 복잡성과 정확성의 차이를 확인하기 전에 세 학습자의 복잡성과 정확성의 발달 속도를 비교하고자 1~6개월과 13~18개월 사이의 변동폭을 아래 <표 115>로 정리하였다.

<표 115> 학습자별 복잡성과 정확성의 변동폭(1~6개월과 13~18개월 비교)

구분	측정 지표	문어			구어		
		A학습자	B학습자	C학습자	A학습자	B학습자	C학습자
통사 복잡성	단위 길이	5.26	5.38	1.65	5.63	5.62	3.80
	단위당 절 수	0.68	0.92	0.17	0.78	0.90	0.46
	단위당 내포절 수	0.51	0.57	0.07	0.60	0.67	0.26
어휘 복잡성	다양도	1.27	0.73	0.86	1.21	0.54	0.74
	밀도	0.13	0.17	1.49	0.46	-0.91	0.91
	정교도	1.22	3.79	0.48	1.53	4.00	3.43

정확성	단위당 무오류절 수	0.45	0.44	-0.01	0.54	0.32	0.35
	단위당 오류 수	0.29	0.66	0.46	0.23	0.76	0.11

먼저 문어의 발달을 살펴보면, 세 학습자 중 통사 복잡성이 가장 빠른 속도로 발달하는 학습자는 B학습자이고 어휘 다양도가 가장 빨리 발달하는 학습자는 A학습자였으며 어휘 정교도가 가장 빠른 속도로 발달한 학습자는 B학습자였다. 그리고 정확성의 구축에 있어서는 A학습자가 가장 나은 수행을 보이고 있었는데 A학습자는 무오류절 수의 증가폭이 가장 크고 오류 수의 증가폭이 가장 낮은 학습자였다. 한편, C학습자는 어휘 밀도를 제외하고 모든 측정 지표에서 다른 학습자들에 비해 작은 폭의 증가만 보이고 있었는데 특히 무오류절 수는 거의 증가하지 않은 반면 오류 수는 A학습자보다 더 큰 폭으로 증가하여 복잡성뿐만 아니라 정확성의 구축에서도 다른 학습자에 비해 낮은 성취를 보였다. 구어의 발달을 살펴보면, A학습자의 단위 길이를 제외하고 문어와 큰 차이가 없었다. B학습자는 구어 역시 다른 학습자들에 비해 통사 복잡성과 어휘 정교도의 발달이 빨랐으며 A학습자는 어휘 다양도의 발달이 가장 빨랐다. 그러나 B학습자는 통사 복잡성과 어휘 정교도에서 다른 학습자들에 비해 빠른 발달을 보이는 것과 동시에 단위당 오류 수가 눈에 띄게 빠른 속도로 증가하고 있었으며 어휘 다양도는 문어와 구어 모두 세 학습자 중 가장 낮은 발달을 보이고 있었다. 이러한 학습자별 구인들의 발달 속도는 학습자들 간의 수준 차이로 이어지게 되는데 학습자 간의 복잡성과 정확성의 차이를 비교해 보면 다음의 <표 116>과 같다.

<표 116> 학습자 간 문어와 구어의 복잡성과 정확성 비교

구분	측정 지표	문어	구어
통사 복잡성	단위 길이	BW>CW	AS,BS>CS
	단위당 절 수	BW>AW,CW	BS>AS,CS
	단위당 내포절 수	BW>CW	-
어휘 복잡성	다양도	BW,AW>CW	BW>CW
	밀도	BW>AW	-
	정교도	-	BS,AS>CS
정확성	무오류절 비율	AW,CW>BW	CS,AS>BS
	단위당 무오류절 수	-	AS>BS
	단위당 오류 수	BW>AW,CW	BS>AS>CS

표에서 확인할 수 있듯이 문어와 구어에서 모두 B학습자의 통사 복잡성이 가장 높고 C학습자의 통사 복잡성이 가장 낮았다. 구어의 단위 길이에서 A학습자의 단위 길이가 B학습자에 비해 약간 더 높기는 하였지만 통계적으로 유의미한 수준은 아니었다. 눈에 띄는 점은 구어에서 단위당 내포절 수가 학습자 간에 차이를 보이지 않는다는 점인데 이는 복잡성이 가장 높은 B학습자가 구어에서 문어에 비해 적은 내포절을 산출한 것과 달리 C학습자는 산출 양식에 따른 내포절 수의 차이가 뚜렷하지 않은 것에 기인한 것이었다. 어휘 복잡성 역시 학습자 간 차이가 없는 항목을 제외하고 대부분 B학습자의 복잡성이 가장 높은 것을 확인할 수 있다. 특히 B학습자는 어휘 다양도의 발달이 속도가 세 학습자 중 가장 낮음에도 불구하고 가장 높은 다양도를 보이고 있었는데 이는 B학습자가 초반부터 높은 어휘 다양도에서 출발한 것과 관련이 있다. 즉 크게 발달을 하지는 않지만 가장 높은 어휘 다양도를 유지하고 있는 것이다. 어휘 정교도는 구어에서만 학습자 간의 차이가 유의하였는데 이는 다른 학습자에 비해 C학습자가 구어에서 낮은 정교도를 보이면서 나타난 현상이었다. 마지막으로 정확성을 살펴보면, 무오류절을 통한 측정에서는 B학습자가 가장 낮은 수준을 보였고 오류를 통한 측정에서는 B학습자가 가장 높은 점수를 보였다. 이는 앞에서 살펴본 것과 같이 B학습자의 오류 수가 통사 복잡성과 함께 빠른 속도로 증가한 것과 관련이 있다.

정리하면, 세 학습자는 관찰 시작 시점에서 통사 복잡성은 거의 비슷한 수준이었으나 어휘 복잡성에는 차이가 있었다. 그러나 시간이 지나면서 B학습자는 통사 복잡성을 빠르게 발달시키는 대신 어휘 다양도와 정확성의 발달을 이루지 못하고 있었고 A학습자는 통사 복잡성과 어휘 다양도, 정확성을 비교적 안정적으로 고루 발달시켜 나가고 있었다. 그리고 C학습자는 통사 복잡성과 어휘 복잡성 모두 다른 학습자에 비해 느리게 발달하여 다른 학습자들과 차이가 벌어지고 있었으나 그 대신 오류가 적은 언어를 산출하는데 집중하고 있었다.

본고에서 살펴본 세 명의 중도입국 중국 학습자는 비슷한 시기에 한국에 입국하여 같은 한국어 교육을 마친 뒤 같은 시점에 중학교에 입학한 학습자였다. 그러나 학습자들의 복잡성과 정확성의 발달 속도는 서로 상이하였으며 각각 통사 복잡성에 더 집중하기도 하고 짧지만 더 정확한 언어를 산출하는 것에 집중하기도 하며 분명한 차이점도 드러내고 있었다.

이러한 결과는 대학기관 학습자의 한국어 발달을 종적으로 살피고 학습자 개개인의 발달 경로는 서로 다르다고 주장한 이복자(2015), 결혼이민자의 언어 발달을 살핀 주월랑(2018)과 동일한 결과로 결국 각 학습자의 언어 발달 경로나 발달 속도는 동일하

지 않음을 의미한다. 이러한 현상에 대해 Larsen-Freeman(2006)에서는 학습자 간의 차이는 외부의 사회적 요인에 기인하는 측면도 있지만, 언어 학습 체계 내에서 발생하는 내부적 재구조화 과정에서도 발생할 수 있음을 주장한다. 대표적으로 같은 언어 환경에 처해있는 일란성 쌍둥이의 어휘와 통사 발달을 살핀 Chan 외(2014); (2015)에서는 쌍둥이 중 한 명은 어휘 발달에 더 집중한 반면 다른 한 명은 통사 발달에 더 집중하는 등의 차이가 존재함을 밝히면서 동일한 환경 내에서 유사한 입력에 노출된 비슷한 성격과 관심사를 가진 학습자들이라 할지라도 서로 다른 발달 경로를 보일 수 있음(Chan 외 2015:337)을 주장하였다. 본고의 이러한 결론은 학습자들의 언어 발달은 공통적인 경향을 보이기도 하지만 개개인의 언어의 발달은 역동적이며 언어의 여러 하위 체계는 서로 다른 속도로 발달한다고 하는 역동 체계 이론의 관점을 지지하는 결과라고 볼 수 있다.

그러나 이러한 차이에도 불구하고 학습자들은 공통적으로 시간의 흐름에 따라 통사 복잡성은 꾸준히 발달하지만 어휘 복잡성은 초반에 발달하다가 같은 시점에서 정체가 발생하고 정확성은 뚜렷한 발달을 보이지 않는 등의 비슷한 발달 경로를 보이고 있었으며 특히 A와 B학습자는 복잡성과 정확성의 수준에는 차이가 있었지만 매우 유사한 발달 경로를 보이고 있었다. 이는 여러 종단 연구에서 학습자들은 모두 시간의 흐름에 따라 언어가 점점 복잡해진다는 결과나 Caspi(2010)에서 네 명의 학습자 중 세 명의 학습자가 어휘와 통사 발달에서 비슷한 순서를 보인다고 한 것, Van Dijk 외(2011)에서 서로 다른 스페인어 학습자들이 영어 부정 동사 구문을 습득할 때 비슷한 순서를 밟는다는 것을 발견한 것과 같은 맥락으로 학습 환경이나 학습 연령이 비슷한 학습자들 사이에는 유사한 발달 경로가 존재한다는 것을 의미하는 결과이기도 하다.

4.2. 학습자별 문어와 구어의 발달 차이

마지막으로 시간의 흐름에 따른 발달 외에도 산출 양식에 따른 학습자별 문어와 구어의 수준을 비교해 보기 위해 전체 기간을 바탕으로 문어와 구어 간에 유의미한 차이를 보이는 지표들을 정리해 보았다.

<표 117> 학습자별 문어와 구어의 복잡성과 정확성 비교

구분	측정 지표	A학습자	B학습자	C학습자
통사 복잡성	단위 길이	AS>AW	BW>BS	-
	단위당 절 수	-	BW>BS	CW>CS
	단위당 내포절 수	AW>AS	BW>BS	-
어휘 복잡성	다양도	AW>AS	BW>BS	-
	밀도	-	-	-
	정교도	-	-	CW>CS
정확성	무오류절 비율	-	-	-
	단위당 무오류절 수	-	-	-
	단위당 오류 수	AS>AW	-	-

세 학습자의 통사 복잡성을 비교해 보면 B학습자는 통사 복잡성의 모든 지표에서 문어가 우세하고 있었고 A학습자는 단위 길이는 구어에서 더 높은 반면 단위당 내포절 수는 문어에서 더 높은 것으로 나타나 구어에서 더 긴 문장을 산출하지만 문어에서 더 복잡한 언어를 산출하고 있었다. 그리고 C학습자는 단위당 절 수를 제외하고 통사 복잡성에서 산출 양식에 따른 차이를 크게 보이지 않고 있었다. 어휘 복잡성에 있어서는 A와 B학습자는 어휘 다양도에서만 차이를 보였고 C학습자는 어휘 정교도에서만 차이를 보였다. 그리고 정확성에서는 A학습자의 단위당 오류 수를 제외하고 차이를 보이지 않았다. 정리하면, 학습자별로 차이가 존재하지만 A학습자의 단위 길이를 제외하고 학습자들은 문어에서 더 나은 발달을 보였으며 이러한 산출 양식 간의 차이는 주로 통사적 측면에서 나타나고 있었다.

이 결과를 위의 <표 115>에서 나타난 학습자별 문어와 구어의 각 지표의 발달 속도(증가폭)와 연결시켜 보면 흥미로운 점을 발견할 수 있는데 <표 115>에 따르면 세 학습자 모두 통사 복잡성의 발달 속도는 구어에서 더 빠르고 어휘 다양도의 발달은 문어에서, 어휘 정교도의 발달은 구어에서 더 빠르다는 것을 알 수 있다. 다시 말해, 학습자들은 관찰 기간 동안 문어에 비해 구어의 통사 복잡성이 더 빠르게 발달했음에도 불구하고 여전히 문어의 통사 복잡성이 더 높다는 것이다.

이는 본고의 학습자가 아직 청소년이고 또 목표어 환경에 노출되어 있음에도 불구하고 Kroll & Vann(1981)에서 주장하는 '글쓰기가 본질적으로 문어의 발달이 구어의 발달에 크게 의존할 수 있다거나 글쓰기가 '말을 받아 적는' 단계부터 시작한다고 보는 모

어화자 아동의 문어와 구어 발달 모형(Kroll & Vann, 1981)과는 큰 차이가 있음을 의미한다. 따라서 본고의 결과는 이미 모어의 문해력이 발달한 L2 학습자의 경우 문어의 발달은(특히 초기 단계에서) 구어 발달에 기초하는 것이 아니라 별도의 숙달도로 발달한다고 보는 Weissberg(2000)의 주장을 지지한다. 그러므로 중도입국 청소년 학습자의 문어와 구어 발달은 모어화자 아동의 언어 습득과는 차이가 있을 것이며 습득 연령이나 환경 변인을 고려할 때 성인 학습자와도 분리되어야 하므로 이들의 언어 발달 경로에 대한 탐구가 더 많이 이루어져야 할 것이다.

5. 교육적 제언

이 장에서는 지금까지의 논의를 바탕으로 중국 배경 중도입국 청소년의 한국어 문어와 구어 산출에 나타난 복잡성과 정확성의 발달이 한국어 교육에 시사하는 점이 무엇인지를 논의해보도록 한다.

복잡성과 정확성의 발달을 분석한 결과, 중국 배경 중도입국 청소년 학습자들은 한국 거주 기간이 늘어남에 따라 문어와 구어의 통사 복잡성은 꾸준히 증가하였지만 어휘 복잡성은 그렇지 못하였다. 어휘 복잡성 중 어휘 다양도는 문어와 구어 모두 관찰 시작 13~18개월에 정체가 발생하였고 어휘 정교도 역시 뚜렷한 발달을 보이지 못하고 있었다. 또한 통사 복잡성이 높아질수록 무오류절 수도 증가하지만 오류 수 역시 지속적으로 증가하면서 정확성의 발달을 이루어내지 못하고 있었다. 여러 선행연구에서도 나타나듯이 언어의 각 측면이 항상 동시에 발달할 수는 없지만 복잡성과 정확성을 저해하는 요인들을 이해한다면 학습에 어느 정도 도움을 줄 수 있을 것이다. 아래에서는 통사 복잡성, 어휘 복잡성, 정확성의 각 측면에서 학습자들의 발달을 저해하는 요인들에 대해 논의해 보도록 한다.

통사 복잡성은 문장을 길고 복잡하게 발화할 수 있는 능력을 가리키는 것이다. 학습자들은 한국 거주 기간이 늘어날수록 문어와 구어 모두 점점 복잡한 발화를 산출하고 있었으나 세부적으로 살펴보면 발달의 정체나 퇴행을 보이는 통사적 구조들이 존재하였다. 그중 사용 빈도가 증가하지 못하는 항목은 명사절이었으며 부사절이나 관형사절은 빈도 측면에서는 꾸준한 발달을 보이지만 특정 항목에 대한 의존도가 비교적 높은 것을 확인할 수 있었다. 이는 다양한 의미 관계를 표현하는 어미를 습득하지 못한 채 한정적인 어미에 기대어 구조를 확장하면서 나타난 현상으로, 결국 오류로 이어지는 경우가 빈번하였다. 한국어의 연결어미와 전성어미는 그 수가 많아 학습자가 자연스럽게

다양한 어미들을 습득하고 미세한 의미 차이를 구분하기가 쉽지 않다. 따라서 통사 복잡성의 양적인 발달뿐만 아니라 질적인 발달을 도모하기 위해서는 연결어미와 전성어미의 의미와 형태에 대한 교육이 지속적으로 이루어져야 하며 이해를 바탕으로 한 산출의 연습이 동반되어야 할 것이다.

어휘 복잡성은 어휘의 다양하고 정교한 사용 능력과 관련이 있다. 특히 어휘 다양도는 학습자의 어휘력과 직결되는 것으로 어휘의 다양도 발달이 정체된다는 것은 어휘의 습득이 정체되고 있음을 의미하기도 한다. 학습자들은 관찰 시작 12개월이 지난 시점부터 어휘 다양도의 발달이 정체되기 시작하였으며 특히 체언의 퇴행이 뚜렷하게 나타났다. 이는 학습자들이 공교육에 진입한 지 1년 6개월이 지난 시점이며 한국에 입국한 지 3년 전후의 시점이었다. 이러한 어휘 다양도의 정체는 결국 어휘 정교도의 발달과 이어져 학습자들은 연구가 끝나는 시점까지 전체 어휘의 90% 정도가 초급 수준에 머물러 있었다. 조현용(2000)에서는 어휘 저장소인 어휘부(lexicon)는 어휘 사전, 어휘 형성부, 어휘 해석 부로 나눌 수 있으며 각 부분을 효과적으로 활용하여 학습 단계에 따른 어휘 교육을 진행할 필요가 있다고 주장한다. 이에 따르면 초급에서는 어휘 사전을 적극 활용할 수 있는 모국어와의 일대일 대응을 통한 기계적인 암기가 효과적이고 중급에는 어휘 파생이나 합성과 관련된 어휘 형성부의 비중을 높여 어휘를 생성하는 원리를 가르쳐야 하며 고급 단계에서는 이미 구축된 어휘를 이용하여 더 진전된 어휘 형성 원리를 교육할 필요가 있다는 것이다. 중도입국 청소년의 경우, 일상생활에서 기초적인 어휘는 습득이 비교적 용이하다는 장점이 있다. 따라서 사용 빈도가 높은 접두사나 접미사를 활용한 다양한 파생어나 합성어 교육, 기존에 학습한 어휘의 유의어나 동음이의어를 통한 어휘 확장, 상황이나 문맥을 통한 의미 파악과 같은 활동들을 통해 학습자들의 어휘 다양도와 정교도를 지속적으로 발달시킬 수 있는 방안을 강구할 필요가 있다. 이 외에도 어휘 복잡성의 발달은 입력보다는 '필요성'에 따라 발달한다는 Laufer(1991)의 주장을 참고할 때 학습자들에게 규칙적인 입력뿐만 아니라 의사소통 과제를 부여하거나 필수적으로 의사소통을 필요로 하는 환경을 제공해 주어 언어의 산출의 필요성을 늘리는 것도 하나의 방법이 될 것이다.

다음은 정확성의 발달에 대해 논의하도록 한다. 정확성은 오류 없이 정확한 언어를 사용하는 능력이다. 그러나 학습자들의 정확성은 복잡성과 동반 상승·하락을 하면서 줄곧 뚜렷한 발달을 보이지 않고 있었는데 정확성의 발달을 저해하는 가장 큰 요인은 조사와 어휘 오류, 그리고 어미 오류 중의 선어말 어미의 오류인 것으로 밝혀졌다. 특히 조사 오류 중 문어에서는 주격 조사, 목적격 조사, 보조사의 오류가 가장 높았으며 주

격 조사는 대치 오류, 목적격 조사와 보조사는 누락 오류가 가장 많았다. 그리고 구어에서는 주격 조사, 보조사, 목적격 조사의 오류가 가장 많았는데 주격 조사와 목적격 조사는 대치 오류가 가장 높았고 보조사는 누락과 대치 오류가 비슷한 빈도로 발생하고 있었다.

문어와 구어에서 공통적으로 높게 나타나는 오류는 주격 조사의 대치와 보조사의 누락/대치였고 특히 학습자들은 보조사의 사용에 익숙하지 않아 주격 조사나 목적격 조사를 사용하는 경우가 매우 빈번하였다. 조사는 언어권이나 학습 목적, 숙달도를 불문하고 한국어 학습자에게 가장 빈번하게 오류가 발생하는 항목이라는 점은 여러 연구에서 증명되었다. 그러나 모국어에 대응하는 조사 체계가 없는 중국어권 학습자들은 조사의 습득에 더 큰 어려움을 겪을 수 밖에 없다. 따라서 중국 배경 중도입국 청소년을 대상으로 한 한국어 교육 역시 전반적으로 조사의 오류를 예방하고 교정하기 위한 노력이 필요하며 특히 주격 조사와 보조사의 교육이 절실히 필요함을 알 수 있었다.

보조사는 기본적으로 화제, 대조, 강조의 의미로 분리하여 기술하고 교육하고 있지만 이는 학습자들이 '은/는'의 통합적 의미를 파악하는 것과 언어 사용에서의 화자의 함축적 의미나 태도, 담화에서의 기능을 연계하여 파악하는 것을 어렵게 하며 다양한 실생활에서 활용할 수 있는 응용 능력을 기르는 것에도 도움이 되지 않는다(한송화, 2017:108). 특히 보조사 '은/는'과 주격 조사 '이/가'의 의미 기능 차이는 매우 미묘하며 '주격', '주어', '주제' 개념 등과 관련해서 여러 가지 설명이 시도되고 있지만 여전히 다양한 설명 방식과 관점이 혼재되어 학습자들이 혼동과 어려움을 겪고 있는 부분으로(김호정·강남욱, 2010b) 인지적 능력이 아직 발달 단계에 처해있는 청소년 학습자의 경우 성인 학습자에 비해 문법에 대한 분석적 접근이 더 어렵기 때문에 사전적인 의미 제시만으로는 추상적인 의미가 담긴 보조사의 기능을 이해하기가 더욱 어려울 것이다. 따라서 여러 가지 의미 기능을 가지고 있는 보조사의 경우 다양한 담화 상황과 함께 담화 맥락에서의 사용을 통해 교육을 하는 것이 더 필요하다고 할 수 있겠다.

어휘 오류 중에서 가장 높게 나타난 것은 의미 대치 오류와 오형태 오류였다. 이 중에서 의미 대치는 대부분 어휘력 부족으로 인한 것으로 정교한 의미를 표현해야 하는 것을 빈도가 높고 일상적인 상위어로 대치하는 오류나 유사 의미를 가진 어휘로 대치하는 오류가 빈번하게 발생하고 있었다. 그리고 오형태 오류 중 가장 눈에 띄는 것은 어휘의 기본형을 인지하지 못한 채 어미와 결합하면서 발생하는 활용 오류와 파생 합성을 잘못 사용하여 한국어에 없는 어휘를 만들어 사용하는 오류였다. 또한 어휘 오류 중에는 '명사+ 하다' 형태에서 접미사 '하다'를 생략하는 오류가 빈번하게 나타났는데 이

러한 오류는 주로 한자어에서 많이 발견되었다. 이 같은 오류는 중국어에서 동사와 명사로 동시에 쓰이는 2음절 '동명사' 어휘를 그대로 직역함으로써 중국어에 없는 접미사 '하다'가 누락된 형태를 그대로 사용하면서 나타난 것이었다. 한국어는 한자어가 많은 언어라는 점에서 중국인 학습자에게 있어 어휘 습득이 용이하다는 이점이 분명히 존재한다. 그러나 두 언어 간에는 형태와 의미가 유사한 어휘들도 많지만 형태는 유사하지만 의미가 다른 경우도 있고 의미는 유사하지만 형태가 다른 경우도 있으며 형태와 의미가 유사하지만 사용 상의 제약이 있는 어휘들도 존재하는 만큼 어휘의 의미와 형태 정보뿐만 아니라 구체적인 문맥을 통한 통사와 화용 교육이 필요할 것이다.

어미 오류는 문어와 구어 모두 시제 선어말 어미의 '-었-'누락 오류가 가장 높게 나타났는데 특히 문어에서는 시간이 지남에 따라 오류가 점점 증가하고 있었다. '-었-'은 초급 단계에 학습되지만 중국어권 학습자뿐만 아니라 비중국어권 학습자들에게서도 공통적으로 빈번하게 나타나는 오류 현상이다. 이숙(2013)에서는 '-었-'의 누락 오류가 빈번한 이유로 한국어에서 현재 시제 형태인 '-는-'이나 미래 시제 형태인 '-겠-'과 같이 시제 형태가 나타나지 않고도 시제 의미가 실현되는 것과 비슷하게 과거 시제도 '-었-'의 형태가 나타나지 않아도 과거의 의미가 실현될 수 있을 것이라는 인식이 이러한 오류를 발생시키는 것으로 보았다. 이 외에도 '-었'은 중국어의 '了'와 비슷하게 인식되지만 한국어의 '-었-'은 형용사와 결합이 가능한 반면 중국어에는 '-었-'에 해당하는 과거 시제를 나타내는 형태가 없고 한국어에는 '-었었-'의 형태도 가능하다는 점에서 여러 차이가 존재한다. 따라서 과거 시제 선어말 어미를 가르침에 있어 미래나 현재 시제와 달리 필수적으로 형태를 요구한다는 사실을 이해할 수 있게 하는 명시적인 교육과 함께 다양한 상황 맥락에서의 결합 규칙과 제약에 대한 교육을 강화할 필요가 있을 것이다.

산출 양식에 따른 언어 사용 측면을 보면, 통사 복잡성에서는 인용절을 제외한 모든 지표가 문어에서 더 높게 나타났고 어휘 복잡성에서는 어휘 밀도와 어휘 정교도는 산출 양식에 따른 차이가 나타나지 않았지만 어휘 다양도는 문어에서 뚜렷하게 더 높은 것으로 나타났다. 그리고 통사 복잡성의 발달에서 문어와 구어의 차이가 점점 좁혀지는 것과 달리 어휘 다양도의 발달은 문어와 구어 사이의 차이가 점점 커지고 있었다. 정확성은 전체적인 오류율에서는 산출 양식에 따른 차이가 나타나지 않았으나 문어에서는 문법 오류가 더 빈번한 것과 달리 구어에서는 어휘 오류가 더 빈번하였으며 누락과 오형태 오류는 문어에서 더 높고 대치와 첨가 오류는 구어에서 더 높게 나타났다. 이는 1차적으로 문어와 구어에서 교육적 처치가 다르게 이루어져야 함을 시사하지만 이러

한 차이는 언어 산출 모델의 관점에서는 제2언어 학습자의 어휘나 문법에 대한 '자동화'의 부족으로 인한 것으로 해석할 수 있다. 따라서 문어와 구어의 동반 발달을 위해서는 어휘와 문법에 대한 다양한 입력과 반복적인 연습을 통해 언어 산출 과정의 '형식화(formulating)' 단계의 자동성을 높일 필요가 있을 것이다.

마지막으로, 복잡성과 정확성의 관계를 살펴본 결과, 학습자들은 문어와 구어에서 모두 복잡성과 정확성에 일관된 관계를 보이지 않고 지표에 따라 다르게 나타나는데 통사 복잡성이 증가하면 무오류절 수와 오류 수가 동시에 증가하고 있었다. 그러나 문어에서는 복잡성과 무오류절 수의 상관계수가 더 높고 구어에서는 복잡성과 오류 수의 상관계수가 더 높게 나타나 구어에서 복잡성과 정확성이 동시에 발달하는 것이 더 어려움을 확인할 수 있었다. 또한 기간별로 나누어 살펴보면 문어의 복잡성 역시 초반에는 무오류절 수와의 상관계수가 더 높지만 후반에 들어서는 오류 수와의 상관계수가 더 높아지면서 결국 문어와 구어 모두 복잡성과 정확성의 경쟁 관계가 지지 관계보다 더 뚜렷하게 나타나고 있었다. 이러한 결과는 공교육에 진입한 중도입국 청소년 학습자들의 효과적인 한국어 습득을 위해서는 학습자들의 통사 복잡성 발달 과정에서 정확성에 대한 교육이 집중적으로 이루어져야 하며 특히 구어에서의 정확성에 더 초점을 두어야 함을 시사한다.

연구를 통해 일정 수준의 한국어 교육을 마치고 이미 공교육 내에 진입한 중도입국 청소년이라고 할지라도 별도의 한국어 교육 없이 자연스러운 입력만으로는 복잡성과 정확성의 발달에 한계가 있음을 알 수 있었다. 따라서 중도입국 청소년의 지속적인 한국어 발달을 꾀하기 위해서는 명시적이고 지속적인 한국어 교육 지원이 이루어질 필요가 있다.

VI. 결론

본 연구는 한국에 입국하여 예비학교에서 한국어 학습을 마친 뒤 공교육에 진입한 중국 배경 중도입국 청소년의 한국어 문어와 구어에 나타난 복잡성과 정확성의 발달을 시간의 흐름에 따라 종적으로 고찰하였다.

중도입국 청소년의 수가 늘어나면서 이들의 한국어교육과 관련된 연구들이 꾸준히 이루어지고 있으나 대부분의 연구가 교재나 교육과정만을 다루고 있으며 학습자의 실제 언어 산출물을 분석한 연구가 매우 부족하다. 그러나 교육과정이나 교재를 논하기 위해서는 학습자들의 언어에 대한 이해가 선행되어야 한다. 이에 본고에서는 중도입국 청소년 학습자의 언어 산출물을 분석하여 이들의 한국어 발달 과정을 고찰하고자 하였다. 본고에서는 문어와 구어가 비록 같은 지식 자원을 바탕으로 하지만 인지적 처리 과정에 차이가 있으며 항상 균형적으로 습득되지는 않는다는 점과 학습자들이 한국에서 학업을 수행해야 한다는 점에 주목하여 문어와 구어의 발달을 모두 살펴보고 더 나아가 산출 양식에 따른 언어 수행과 언어 발달 경로의 차이를 알아보고자 하였다. 이를 위해 학습자들의 문어와 구어에 나타난 통사 복잡성, 어휘 복잡성, 정확성을 측정하였다. 본고에서는 통사 복잡성의 하위 구인으로 단위 길이, 단위당 절 수, 단위 확장 유형을 살펴보았고 어휘 복잡성의 하위 구인으로 어휘 다양도, 어휘 밀도, 어휘 정교도를 살펴보았다. 그리고 정확성의 하위 구인으로 무오류절의 비율, 단위당 무오류절 수, 단위당 오류 수, 오류의 유형을 살펴보았다. 이에 따른 분석 결과를 정리하면 다음과 같다.

시간의 흐름에 따른 복잡성과 정확성의 발달을 살펴보면, 학습자들의 통사 복잡성은 시간이 지남에 따라 꾸준히 발달하고 있었으나 어휘 복잡성은 측정 지표에 따라 다른 발달을 보였으며 정확성은 무오류와 오류가 모두 증가하여 뚜렷한 발달이나 퇴행으로 정의하기 어려웠다 . 어휘 복잡성과 정확성의 발달을 지표별로 나누어 살펴보면 어휘 다양도는 문어와 구어 모두 7~12개월까지 증가하다가 1년이 지난 시점인 13~18개월에 정체가 발생하였고 어휘 밀도는 문어와 구어 모두 관찰 기간 동안 뚜렷한 발달이나 퇴행을 보이지 않았으며 어휘 정교도는 문어와 구어 모두 7~12개월까지 발달을 보이지 않다가 13~18개월에 약간 향상되고 있었다. 정확성의 발달은 측정 지표에 따라 상반되는 결과를 보였는데 무오류절의 비율은 문어에서는 7~12개월에 상승 이후 13~18개월에 다시 하락하였고 구어에서는 7~12개월까지 정체되어 있다가 13~18개월에 미세하게 증가하였다. 그리고 단위당 무오류절 수는 문어에서 7~12개월까지 증가 이후 13~18개월에 정체를 보이는 반면 구어에서는 시간의 흐름에 따라 완만하게 증가하는

양상을 보이고 있었다. 단위당 오류 수는 그 정도에 차이가 있으나 문어와 구어에서 모두 시간이 흐름에 따라 증가하고 있었으며 후반으로 갈수록 증가 속도가 가속되는 양상을 보이고 있었다.

측정 지표별로 유의한 증가가 나타난 시점을 살펴보면, 단위 길이와 단위 확장은 문어와 구어 모두 6개월을 간격으로 지속적인 증가를 보이고 있었고 어휘 다양도는 문어와 구어에서 모두 1~6개월과 7~12개월 사이에만 뚜렷한 증가를 보였으며 어휘 정교도는 구어에서만 7~12개월과 13~18개월 사이에 유의한 증가를 보였다. 정확성 측정 지표에서 무오류절 수는 문어에서는 1~6개월과 7~12개월 사이에 뚜렷한 발달이 있었고 구어에서는 7~12개월과 13~18개월 간에 차이를 보였다. 그리고 정확성 측정 지표 중 오류 수는 문어에서는 7~12개월과 13~18개월 사이에 유의한 증가가 있었고 구어에서는 1~6개월과 13~18개월 사이에 유의하게 증가하는 것으로 나타났다.

산출 양식에 따른 복잡성과 정확성의 차이를 분석한 결과, 학습자들은 복잡성 측면에서 구어보다 문어에서 더 나은 수행을 보이고 있었다. 구어보다 문어에서 유의미하게 높은 산출을 보인 지표는 단위 길이, 단위당 절 수, 단위당 대등절 수, 단위당 내포절 수(명사절, 관형사절, 부사절), 어휘 다양도였으며 정확성은 문어와 구어 사이에 유의미한 차이를 보이지 않았다. 이를 다시 기간별로 나누어 문어와 구어의 차이를 비교해본 결과, 학습자들은 초반에 문어에서 통사적으로 훨씬 더 복잡한 언어를 사용였으나 시간이 지날수록 구어의 통사 복잡성 발달이 더 빨라져 문어와 구어의 차이가 점점 좁혀지고 있었으며 단위 길이는 13~18개월에 들어 구어의 AS-unit 길이가 문장 길이보다 길어지기도 하였다. 반면, 어휘 다양도는 전반적으로 문어에서 더 높게 나타났으며 초반에는 뚜렷하지 않았던 문어와 구어의 차이가 점점 커지고 있었다. 즉 초기에는 학습자들이 문어에서 더 높은 복잡성을 보이지만 시간이 지날수록 통사 복잡성은 구어에서 더 빨리 발달하고 어휘 다양도는 문어에서 더 빠르게 발달하고 있었다. 한편, 정확성은 기간별로 나누어 보아도 유의한 차이가 나타나는 기간이 없어 문어와 구어의 정확성에 차이가 없음을 다시 한번 확인할 수 있었다.

복잡성과 정확성의 관련성을 분석한 결과, 문어와 구어 모두 정확성 측정 지표에 따라 복잡성과 정확성 간에 서로 상반되는 관계가 나타나고 있었는데 통사 복잡성이 증가하면 무오류절 수와 오류 수가 동시에 증가하고 있었다. 그러나 상관계수의 크기를 비교해 보면 문어에서는 통사 복잡성과 무오류절 수의 상관계수가 오류 수와의 상관계수보다 높고 구어에서는 통사 복잡성과 오류 수의 상관계수가 무오류 수와의 상관계수보다 더 높아 문어는 무오류절 수와 더 밀접한 관계가 있고 구어는 오류 수와 더 밀접

한 관계를 보인다는 점을 확인할 수 있었다. 기간별로 비교해 보면, 문어에서는 1~6개월과 7~12개월에는 무오류절 수와의 상관계수가 더 높던 것에서 13~18개월에는 오류 수와의 상관계수가 더 높아졌으며 구어에서는 세 기간 모두 오류 수와의 상관계수가 더 높아졌다. 이를 통해, 문어에서는 정확성과 지지적 관계가 더 강하던 것에서 시간이 지날수록 경쟁관계가 더 강해지게 된 반면, 구어에서는 지속적으로 정확성과 경쟁 관계가 더 강하게 나타나고 있음을 확인하였다. 그리고 어휘 복잡성은 지표별로 정확성과 다른 관계를 맺고 있었는데 어휘 밀도는 7~12개월과 13~18개월에 정확성과 지지적 관계, 어휘 정교도는 1~6개월과 13~18개월에 정확성과 경쟁적 관계에 있는 것으로 나타났다. 이러한 결과를 통해 복잡성과 정확성의 관련성은 지표에 따라 다르게 나타나며 시간의 흐름에 따라 변화하기도 한다는 것을 확인하였다.

마지막으로 세 학습자의 언어 발달을 비교해본 결과, 학습자마다 집중적으로 발달시키는 구인이 달랐으며 복잡성이나 정확성을 발달시킴에 있어서도 발달이 집중적으로 나타나는 지표가 서로 다르다는 것을 확인할 수 있었다. 그리고 이러한 학습자 간의 차이는 외부의 사회적 요인 외에도 학습자 체계 내에서 발생하는 재구조화 과정에서도 발생할 수 있으며 이러한 차이가 학습자마다 다른 언어 발달 경로를 만들어 낸다고 보는 '역동 체계 이론'을 일부 지지하는 것으로 볼 수 있었다. 그러나 본고의 연구에 참여한 3명의 학습자들은 통사 복잡성이나 어휘 복잡성, 정확성 모두 큰 틀에서 비슷한 순서로 발달을 하고 비슷한 지점에서 발달이나 정체를 보인다는 점에서 제2언어 학습자들 사이에 유사한 발달 경로가 존재한다는 점을 부인할 수 없었다.

이러한 분석 결과를 바탕으로 본고에서는 다음과 같은 시사점을 제시하였다. 중국 배경 중도입국 청소년은 한국에 거주하면서 교육을 받는 시간이 길어지면서 문어와 구어의 통사 복잡성이 뚜렷하게 발달한다. 그러나 어휘 복잡성은 공교육 입학 1년 6개월에서 2년 사이, 즉 한국 입국 3년 정도 되는 시점에서 정체가 나타나기 시작하며 정확성은 발달을 보이지 않는다. 따라서 기초적인 한국어 교육을 받고 공교육에 이미 진입한 중도입국 청소년이라고 할지라도 지속적인 교육적 지원이 이루어질 필요가 있다. 구체적으로, 통사 복잡성의 측면에서는 다양한 부사절 관련 어미의 의미 변별과 형태에 대한 교육이 필요하며 일상생활에서 자주 접하기 어려운 명사절과 인용절의 학습과 사용을 촉진시킬 필요가 있다. 어휘 복잡성 측면에서는 다양한 입력 및 출력 강화를 통한 어휘의 다양도와 정교도 교육이 이루어져야 한다. 또한 복잡성과 달리 학습자들의 정확성이 지속적으로 개선되지 않는다는 점을 고려할 때 정확성 교육에 더 주목할 필요가 있으며 특히 정확성의 구축에 가장 큰 영향을 주는 조사 오류와 선어말 어미 오류

를 줄이기 위해서는 해당 문법에 대한 교육이 중요하게 다루어져야 한다. 또한 어휘 다양도와 정교도를 높임으로써 높은 빈도로 나타나는 어휘 오류 역시 개선될 필요가 있다. 이 외에도 본 연구에서는 학습자들의 언어가 문어와 구어라는 두 가지 산출 양식에 따라 서로 다르게 발달하는 것을 확인하였다. 따라서 문어와 구어의 고른 발달을 위해서는 구어의 어휘 복잡성과 정확성 측면에 더 주목할 필요가 있으며 최종적으로 문어와 구어 모두 정확성을 바탕으로 한 지속적인 한국어 교육이 이루어질 필요가 있다.

본 연구의 의의는 크게 세 가지로 정리할 수 있다. 우선, 본 연구는 그동안 자세히 다루어진 적이 없는 공교육에 진입한 중국 배경 중도입국 청소년의 한국어 발달 과정을 종적으로 살펴보았다. 이어서, 본 연구는 이들의 한국어 수행을 문어와 구어로 나누어 산출 양식에 따른 발달을 비교하였다. 한국어 교육 분야에서 문어와 구어의 발달을 함께 다룬 연구들은 있었으나 직접적으로 그 차이에 대해 비교와 논의를 진행한 연구가 드물다는 것을 고려할 때, 본 연구는 산출 양식에 따른 언어 발달을 비교하여 문어와 구어의 통사 복잡성과 어휘 복잡성, 오류의 영역 및 오류 현상에 차이가 있다는 점, 문어와 구어의 발달 경로에 차이가 있음을 밝혔다는 점에서도 의의가 있을 것이다. 마지막으로 본 연구에서는 문어와 구어의 비교뿐만 아니라 복잡성과 정확성에 나타난 관계를 살펴봄으로써 중도입국 청소년의 한국어 발달 과정과 그 과정에 나타난 문제점들을 면밀히 고찰하였다. 본 연구의 결과는 중국 배경 중도입국 청소년의 한국어 발달 과정을 이해하고 향후 중도입국 청소년을 위한 한국어 교육 지원을 하는 데 있어 유용한 기초 자료가 될 수 있을 것을 기대한다.

그러나 위와 같은 의의에도 불구하고 본 연구는 몇 가지 한계점을 지니고 있다.

첫째, 다수의 종적 연구가 지닌 한계점으로서 연구 참여자 수가 제한적이라는 것이다. 본 연구에서는 비슷한 시기에 한국에 입국하여 같은 시기에 중학교에 입학한 같은 언어권의 학습자를 대상으로 연구를 진행하여 학습자 간의 변인 차이를 최대한 통제하였으나 참여자 수가 제한적이라는 점에서 연구 결과를 일반화하는 데에 한계가 있다. 이에 후속 연구에서는 더 많은 종적, 횡적 데이터를 수집하여 중도입국 청소년 학습자의 한국어 발달을 더 명확히 파악할 필요가 있을 것이다.

둘째, 학습자들 간의 언어 발달을 비교함에 있어 개인적 특성을 반영하지 못했다는 점이다. 학습자들의 언어 발달은 학습 환경이나 교수법과 같은 입력이 중요한 역할을 하지만 학습 동기나 언어 적성, 지능과 같은 개인적인 변인 역시 언어 학습 결과에 영향을 미치는 중요한 요인이다. 따라서 후속 연구에서는 학습자들의 개인적 특성과 연결시킨 언어 발달 연구들이 활발히 이루어지기를 기대해 본다.

셋째, 본 연구에서는 중국 출신의 중도입국 청소년 학습자만을 대상으로 연구를 진행하였다는 점에서 언어권에 한계가 있었다. 현재 한국에는 점점 더 다양한 언어권의 중도입국 청소년들이 유입되고 있는 만큼 중국어권을 제외한 다른 언어권 중도입국 학습자의 한국어 발달에 대한 연구도 활발히 이루어져 이들을 위한 교육 설계에 기반이 되는 자료가 더 많이 축적될 수 있기를 기대해 본다.

참 고 문 헌

강란숙·김수영·박지은·이동만·허용(2017), 중도입국 청소년 대상 한국어 쓰기 수업에서 시각자료 활용기반 유도 작문 교수-학습 효과에 관한 시론적 연구, 다문화와 평화 11(1), 성결대학교 다문화평화연구소, 162-193.

강범모·김흥규(2009), 한국어 사용 빈도: 1500만 어절 세종형태의미분석말뭉치 기반, 서울: 한국문화사.

강수정(2018), 중도입국 한국어 학습자 관련 연구에 대한 비판적 고찰과 향후 과제, 한글 79(4), 한글학회, 895-937.

강정희(2020), 한국어 학습자의 어휘 능력 연구: 쓰기 텍스트의 어휘 풍요도를 중심으로, 동국대학교 대학원 박사학위논문.

강현화(2017), 한국어 학습자 말뭉치 연구 및 구축 사업, 국립국어원 보고서.

경기도가족여성연구원(2012), 다문화가정 중도입국 자녀 초기적응 지원방안 연구, 수원: 경기도교육청

고석주·김미옥·김제열·서상규·정희정·한송화(2004), 한국어 학습자 말뭉치와 오류 분석, 서울: 한국문화사.

고영근·구본관(2018), <개정판>우리말 문법론, 서울: 집문당.

구본관·박재연·이선웅·이진호·황선엽(2015), 한국어 문법 총론 I, 서울: 집문당.

국립국어원(2005), 외국인을 위한 한국어 문법 1, 서울: 커뮤니케이션북스.

국립국어원(2012), 한국어교육 어휘 내용 개발(1단계) 보고서.

국립국어원(2013), 한국어교육 어휘 내용 개발(2단계) 보고서.

국립국어원(2014), 한국어교육 어휘 내용 개발(3단계) 보고서.

국립국어원(2015), 한국어교육 어휘 내용 개발(4단계) 보고서.

국립국어원(2018), 2018년 한국어 학습자 말뭉치 연구 및 구축 사업 보고서.

국립국어원(2019), 중고등학생을 위한 표준 한국어 의사소통 1, 서울: 마리북스.

국립국어원(2019), 중고등학생을 위한 표준 한국어 의사소통 2, 서울: 마리북스.

국립국어원(2019), 중고등학생을 위한 표준 한국어 의사소통 3, 서울: 마리북스.

국립국어원(2019), 중고등학생을 위한 표준 한국어 의사소통 4, 서울: 마리북스.

권미경·이소연(2017), 중도입국청소년의 한국어학습 동기와 학업성취에 관한 질적 연구, 새국어교육(110), 한국국어교육학회, 147-181.

김경령(2013), 결혼이민자의 거주 기간별, 언어권별에 따른 한국어 능력 -10개국 출신의 전국 거주자를 대상으로, 외국어로서의 한국어교육 38, 연세대학교 언어연구교육원 한국어학당, 1-27.

김광해(1993), 국어 어휘론 개설, 서울: 집문당.

김명광(2016), 언어 숙달도 평가에서 MLU 측정법의 해석 타당성에 대한 일고, 새국어교육(109), 한국국어교육학회, 453-484.

김명정(2011), 특집논문: 동반, 중도입국 자녀들을 위한 다문화교육, 교육문화연구 17(2),

인하대학교 교육연구소, 55-76.
김목아(2010), 중국인 한국어 학습자의 작문에 나타난 중간언어 양상 연구, 계명대학교 대학원 석사학위논문.
김미강·김영주(2021), 중도입국 청소년의 정의적 요인이 한국어 숙달도에 미치는 영향, 國語敎育學硏究 56(1), 국어교육학회, 5-41.
김민수(2022), 일본인 한국어 학습자의 작문에 나타나는 어휘 풍요도 연구, 학습자중심교과교육연구 22(2), 학습자중심교과교육학회, 263-284.
김선정·김성수(2009), 중국인 한국어 학습자의 글쓰기에 나타난 오류 연구, 언어와 문화 5(3), 한국언어문화교육학회, 139-160.
김수미(2011), Twin texts의 읽기 순서가 학생들의 독해력에 미치는 영향-서사적 텍스트와 설명적 텍스트를 중심으로, 고려대학교 대학원 석사학위논문.
김아름(2017), 한국어 학습자의 말하기 능력 발달 연구: 국내 중국인 학습자를 대상으로, 경희대학교 대학원 박사학위논문.
김연희(2018), 중도입국학생을 위한 효율적인 한국어 교육방안-언어문화 교육과정 중심으로, 한말연구(48), 한말연구학회, 31-56.
김영란(2014), KSL "한국어 교육과정"에 대한 비판적 고찰, 國語敎育學硏究 49(3), 국어교육학회, 60-92.
김영주·고윤석·김아름·김소현(2012), 국내 이주배경 청년의 한국어 습득 예측 변인 연구: 북한이탈 청년과 중도입국 청년을 중심으로, 한국어 교육 23(4), 국제한국어교육학회, 31-64.
김영주·남주연·이선진·이선희(2013), 한국어 숙달도와 산출 능력 간의 상관관계 연구, 국어국문학(164), 국어국문학회, 209-244.
김영주·백준오·이선진·오진희·정혜원(2017), 한국어 학습자의 심리적, 사회적 요인과 구어 통사 복합도 간의 상관관계, 한국어 교육 28(2), 국제한국어교육학회, 1-36.
김영태(1997), 현대국어 보조용언 연구, 대구: 문창사.
김유미(2002), 학습자 말뭉치를 이용한 한국어 학습자 오류 분석 연구, 외국어로서의 한국어교육 27, 연세대학교 한국어학당, 141-168.
김의수·김혜림(2012), 다문화가정 아동과 일반가정 아동의 구어 문장 비교 연구, 우리어문연구 42, 우리어문학회, 31-62.
김정운·김영주·이선진(2017a), 자연 학습 환경에서의 L2 한국어 통사 발달과 학습자 변인에 대한 횡적 고찰, 이중언어학 67, 이중언어학회, 63-101.
김정운·김영주·이선진(2017b), L2 한국어 통사 발달과 학습자 변인에 대한 종적 고찰: 자연 학습 환경의 예, 한국어 교육 28(4), 국제한국어교육학회, 1-38.
김해영(2013), 다문화 가정 학생을 위한 주제중심 초급 한국어 표현교육 지도방안, 광주여자대학교 대학원 석사학위논문.
김현아(2016), 재미동포 청소년 학습자의 한국어 통사 복잡성 발달 연구, 서울대학교 대학원 석사학위논문.
김호정(2007), 재외동포의 한국어 사용 실태 연구 방향에 대한 일고찰 -한국어 교육 정책

수립을 위한 기초 연구 관점에서-, 한국어 교육 18(1), 국제한국어교육학회, 73-100.
김호정·강남욱(2010a), 한국어 학습자의 문법 습득 양상 연구(1)- "-고"와 "-어서"를 중심으로, 국어교육(132), 한국어교육학회, 295-323.
김호정·강남욱(2010b), 한국어 학습자의 문법 습득 양상 연구 (Ⅱ) - 조사 {이/가}와 {은/는}의 사용 양상을 중심으로, 국어국문학(156), 국어국문학회, 5-41.
김효은·김가연(2018), 지시문 유형에 따른 쓰기 양상 연구, 학습자중심교과교육연구 18(7), 학습자중심교과교육학회, 237-257.
남기심·고영근(1985), 표준국어문법론, 서울: 塔出版社.
남기심·고영근(2011), 표준국어문법론, 서울: 탑.
남주연(2015), 한국어 학습자의 구어 복잡성 연구: 통사 및 어휘 복잡성을 중심으로, 경희대학교 대학원 박사학위논문.
남주연·김양희·김영주(2016), L2 한국어 문어 산출에서의 통사 복잡성 측정, 한국어 의미학 51, 한국어의미학회, 21-56.
남주연·김영주·김양희(2017), 한국어 학습자의 문어 어휘 다양성 연구-구어 어휘 다양성과의 비교를 중심으로, 한국어 의미학 55, 한국어의미학회, 1-28.
남주연·김영주(2014), 한국어 학습자의 구어에 나타난 어휘 다양성 측정-D값을 중심으로, 한국어 의미학 45, 한국어의미학회, 69-97.
남주연·김영주(2015), 한국어 학습자의 구어 산출에 나타난 통사 복잡성 발달, 한국어 교육 26(1), 국제한국어교육학회, 59-87.
노아실(2007), 주제에 대한 관심과 숙달도가 한국어 말하기 능력에 미치는 영향, 이화여자대학교 대학원 석사학위논문.
노정은·장미정(2019), 중도 입국 자녀 대상의 한국어 읽기 자료 개발 방안 연구: 교과서 기반 초등학교 고학년 학습 자료를 중심으로, 리터러시 연구 10(3), 한국 리터러시학회, 621-662.
류한레이(2016), 중국인 한국어 학습자의 중간 언어 연구: 어휘를 중심으로, 연세대학교 대학원 박사학위논문.
박영옥·이진경(2018), 다문화 배경 유아를 위한 언어교육 방법 모색: “언어자각(Language Awareness)” 프로그램을 중심으로, 이중언어학 73, 이중언어학회, 153-175.
박정은·김영주(2014), 한국어 고급 학습자의 작문에 나타난 어휘의 다양성, 한국어 교육 25(2), 국제한국어교육학회, 1-32.
박지순·서세정(2009), 쓰기 텍스트 분석을 통한 한국어 학습자의 통사적 숙달도 측정 연구, 언어와 문화 5(2), 한국언어문화교육학회, 151-173.
方香玉·甘善權(2022), 중국 조선족 학교 학생들의 한국어 교육 실태 조사 연구, 한국어교육연구(17), 한국어교육연구학회, 61-78.
배도용(2012), 한국어 학습자의 쓰기에 나타난 어휘 다양도 및 어휘 밀도 연구, 언어과학 19(1), 한국언어과학회, 99-117.
배도용(2014), 한국어 학습자 문어 산출물에 나타난 어휘 풍요도 측정을 위해 해결해야 할

몇 가지 문제, 우리말연구 36, 우리말학회, 133-156.
배두본(2002), 영어교육학 총론, 서울: 한국문화사.
배소영·곽금주·김근영·정경희·김효정(2009), 다문화가정 어머니와 발달지원자 설문을 통해 본 아동의 언어환경 및 언어발달 실태, 言語治療硏究 18(4), 한국언어치료학회, 165-184.
서세정(2009), 통사적 숙달도 진단을 통한 한국어 학습자의 중간언어 발달 연구: 학습자 작문 텍스트의 분석을 바탕으로, 연세대학교 대학원 석사학위논문.
서세정·어지혜(2011), 한국어 학습자의 숙달도별 연결어미 정확도 변이 양상 연구, 한국어 교육 22(1),123-143.
서수현(2008), 쓰기 과제 구성 요소의 설정에 대한 연구, 國語敎育學硏究 33, 국어교육학회, 449-472.
서신웨(2021), 미얀마인 한국어 학습자의 학년에 따른 통사 복잡성 연구, 외국어교육연구 35(1), 한국외국어대학교 외국어교육연구소, 105-129.
서신웨(2022), 한국어 중·고급 학습자의 문어 산출에 나타난 복잡성과 정확성 연구: 서사와 논증 장르를 중심으로, 서울대학교 대학원 박사학위논문.
석주연·안경화(2003), 한국어 학습자 표현 오류분석의 몇 가지 문제, 한국어 교육 14(3), 국제한국어교육학회, 189-215.
손희연·조태린(2015), 한국어교육학 : <한국어 교육과정>의 개선 방향에 대한 소고 -교육 대상과 내용 체계를 중심으로, 새국어교육(104), 한국국어교육학회, 327-361.
신영권(2014), 중도입국학생의 학교생활 적응을 위한 언어재료 연구, 한국어 교육 25(3), 국제한국어교육학회, 185-205.
심나빈(2013), 중도입국 청소년을 위한 국어 교과학습용 한국어 교재 개발, 이화여자대학교 대학원 석사학위논문.
심혜령(2015), 한국어교육학 : KSL 학습자를 위한 내용중심한국어교육 기반 교과교육 방안 연구 -과학교과를 중심으로, 새국어교육(105), 한국국어교육학회, 325-347.
안경화(2003), 중간언어 어휘론 연구의 관제와 전망, 이중언어학 23, 이중언어학회, 167-186.
안의정(2017), 한국어 텍스트의 어휘 다양도와 어휘 밀도 분석, 언어사실과 관점 41, 연세대학교 언어정보연구원(구 연세대학교 언어정보개발원), 349-365.
양계민(2021), 이주배경 아동·청소년 행정자료 연계 및 법제화방안 연구, 한국청소년정책연구원 연구보고서, 한국청소년정책연구원, 1-122.
양계민·조혜영(2011), 중도입국 청소년 실태조사, 한국청소년정책연구원.
여성가족부(2016), 2015년 전국다문화가족실태조사 분석, 한국여성정책연구원.
오소정(2015), 다문화가정 아동 언어평가 및 언어치료교육의 의미와 실제: 언어치료 전공 전문가 심층면담 연구, 예술인문사회융합멀티미디어논문지 5(4), 사단법인 인문사회과학기술융합학회, 583-600.
오영훈·박봉수·허숙·팜티휀짱(2012), 중도입국청소년을 위한 초급 한국어 교재 개발 방안

연구, 다문화교육 3(2), 한국다문화교육연구학회, 51-73.
우성숙(2021), 전화 한국어 수업에서 나타난 말하기 능력의 발달 양상에 대한 연구: 복잡성·정확성·유창성을 중심으로, 경인교육대학교 대학원 석사학위논문.
원미진(2019), 한국어 학습자 언어 자료 분석의 방법과 실제, 서울: 한국문화사.
원미진·왕억문·주우동·왕호(2017), 한국어 학습자의 쓰기에 나타난 어휘 풍요도 연구-숙달도 측정 도구로써 어휘 풍요도 측정 가능성을 중심으로, 어문론총 71, 한국문학언어학회(구 경북어문학회), 33-55.
원진숙(2008), 다문화 시대의 초등학교 국어과 교육, 國語敎育學硏究 32, 국어교육학회, 269-303.
원진숙(2015), 2012 한국어 교육과정의 성격과 KSL 프로그램 운영 방안, 국어교육(149), 한국어교육학회(구 한국국어교육연구학회), 207-240.
윤새은임·유진이(2017), 중도입국청소년에 관한 연구동향 분석, 다문화아동청소년연구 2(1), 다문화아동청소년연구원, 57-93.
음승민(2012), 중국인 한국어 학습자의 구문 발달에 대한 양적 연구-학습 단계 향상에 따른 문장의 구조 및 길이 변화를 중심으로, 고려대학교 대학원 석사학위논문.
이강원·손호웅(2016), 지형 공간정보체계 용어사전, 서울: 구미서관.
이경(2015), 담화 차원에서의 한국어 어휘 유창성 신장 방안 연구, 고려대학교 대학원 박사학위논문.
이경숙(2015), 중학교 KSL 학습자의 학습 한국어 능력 향상을 위한 『표준 한국어』 교재 개선 방안 연구, 인하대학교 대학원 석사학위논문.
이루리(2018), 다문화 예비학교의 '중학생을 위한 표준 한국어' 교재 활용 실태 및 교재 개선 방안 연구, 고려대학교 대학원 석사학위논문.
이민경(2016), 구어와 문어 텍스트에 나타난 중국어권 한국어 학습자의 응결장치 연구, 이화여자대학교 대학원 박사학위논문.
이민아(2013), 중등학교 다문화가정 학생을 위한 표준 한국어 교재 분석, 연세대학교 대학원 석사학위논문.
이복자(2015), 한국어 학습자의 쓰기와 말하기에 나타난 복잡성, 정확성, 유창성의 역동적 발달 연구, 연세대학교 대학원 박사학위논문.
이선진·계은진·백준오(2023), 과제 복잡성과 학습자의 인지적 요인 연구, 새국어교육,133-169.
이소라(2012), 중도입국 청소년의 학업 문식성 신장을 위한 한국어 교재 개발 연구, 이화여자대학교 대학원 석사학위논문.
이소현(2016), 중도입국청소년들의 수학 학습과 연계한 한국어 학습 어휘 선정 연구 -초등학교 4-6학년 수학 교과서 분석을 바탕으로, 언어와 언어학(70), 한국외국어대학교 언어연구소, 459-496.
이숙(2013), 중국어권 한국어 학습자의 "-었-" 사용 오류의 특성, 외국어로서의 한국어교육 39, 연세대학교 언어연구교육원 한국어학당, 137-163.
이영근(2005), 한국어 학습자 중간언어 발화의 정확성과 복잡성을 측정하기 위한 분석도

구에 관한 소고, 국제한국어교육학회 학술대회논문집, 국제한국어교육학회, 171-175.
이정연(2017), CIS 지역 재외동포 학습자들의 한국어 쓰기 특성 연구 -요구 조사와 오류 분석을 중심으로, 언어와 언어학 76, 한국외국어대학교 언어연구소, 99.
이정희(2002), 한국어 학습자의 표현 오류 연구, 경희대학교 대학원 박사학위논문.
이정희(2008), 중국어권 한국어 학습자의 어휘 오류 연구 －원인 분석을 중심으로－, 한국어 교육 19(3), 국제한국어교육학회, 403-425.
이진경(2006), 한국어 학습자의 관형사형 어미 사용 연구, 연세대학교 대학원 석사학위논문.
이찬우(2015), 중도입국자녀를 위한 한국어 교재의 문화 교육 연구-학령기 중도입국자녀를 중심으로, 영남대학교 대학원 석사학위논문.
이현경(2020), '한국어 교육과정(2017)'과 『중학생을 위한 표준 한국어-학습 도구』 내용 분석, 연세대학교 대학원 석사학위논문.
이효인(2015), 중도입국 자녀 대상 한국어 교육 연구 -전남지역의 현황과 과제를 중심으로, 호남문화연구(58), 전남대학교 호남학연구원, 307-331.
임선일·김기영·한경은·이용민(2021), 중도입국청소년의 학교생활, 서울: 학이시습.
임지룡(1998), 어휘력 평가의 기본 개념, 국어교육연구 30(1), 국어교육학회, 1-41.
임지룡·이은규·김종록·송창선·황미향(2005), 학교 문법과 문법 교육, 서울: 박이정.
장병순(2019), 중도 입국 자녀 중학생들의 쓰기 오류 분석 연구: 자유 작문 분석을 중심으로, 중부대학교 대학원 석사학위논문.
조남호(2003), 한국어 학습용 어휘 선정 결과 보고서, 국립국어원 보고서.
조철현(2002), 한국어 학습자의 오류 유형 조사 연구, 문화관광부.
조청화(2020), 중국 중도입국 청소년의 한국어 쓰기 통사적 복잡성 발달 양상 연구, 이중언어학 80, 이중언어학회, 271-302.
조청화(2022), 중국 이주배경 청소년의 한국어 구어 사용에 나타난 복잡성과 정확성 연구, 이중언어학 86, 이중언어학회, 259-291.
조현용(2000), 한국어 어휘교육 연구, 서울: 박이정.
주우동(2019), 중국인 한국어 학습자의 언어 발달 양상 연구-복잡성과 정확성을 중심으로, 연세대학교 대학원 박사학위논문.
주월랑(2018), 여성결혼이민자의 중간언어 발달 연구-쓰기 자료의 복잡성과 정확성을 중심으로, 연세대학교 대학원 박사학위논문.
주월랑(2020), 여성결혼이민자의 거주기간에 따른 구어 복잡성 발달 연구: 통사 복잡성을 중심으로, 문화와 융합 42(10), 한국문화융합학회, 749-772.
진대연(2006), 한국어 쓰기능력 구성요소로서의 어휘에 대한 연구, 이중언어학 30, 이중언어학회, 379-411.
진송수(2020), 중도 입국 학습자 대상 학교 적응 지원 한국어 어휘 교수 설계 연구, 이화여자대학교 대학원 박사학위논문.
차숙정·송향근(2006), 중국인 한국어 학습자의 쓰기에서 나타나는 어휘 오류, 한국어 교

육 17(1), 국제한국어교육학회, 415-436.
최미화(2017), 한인 디아스포라로서의 재중동포(조선족) 사회의 언어 유지와 발전 과정 연구, 상명대학교 대학원 박사학위논문.
최수완(2019), 한국어 계승어 학습자의 통사적 복잡성 발달 양상 연구: 호주 교포 청소년 학습자의 작문을 중심으로, 서울대학교 대학원 석사학위논문.
최예슬(2018), 한국어 학습자의 수준별 통사적 복잡성 비교 연구: 학습자 작문 자료 분석을 통하여, 연세대학교 대학원 석사학위논문.
최유진·이선영·심지연·김효은(2022), 초급 한국어 학습자의 변인과 통사 복잡성 발달에 대한 질적 연구, Journal of Korean Culture 58, 한국어문학국제학술포럼, 5-50.
최진숙·고은(2017), 다문화가정 유아의 쓰기 특성에 대한 연구, 言語治療硏究 26(3), 한국언어치료학회, 49-64.
한송화(2015a), 한국어 학습자의 관형절과 관형사형 어미 사용 양상 연구 -관계관형절과 동격관형절을 중심으로, 문법 교육 24, 한국문법교육학회, 211-243.
한송화(2015b), 한국어 학습자의 명사절 사용 양상과언어 발달 연구 -숙달도와 언어권의 차이를 중심으로, 어문론총 66, 한국문학언어학회(구 경북어문학회), 59-93.
한송화(2017), 구어 담화에서 보조사 '은/는'의 의미와 담화 기능 -한국어교육 문법을 위하여, 한국어 의미학 55, 한국어의미학회, 81-111.
한하림(2019a), 제2언어로서의 한국어 아동 학습자의 복문 산출 양상 연구, 한국어학 82, 한국어학회, 155-184.
한하림(2019b), KSL 아동의 연결어미 발달 양상 연구, 한국어 교육 30(2), 국제한국어교육학회, 355-382.
허은혜(2010), 여성 결혼 이민자들의 작문에 나타난 어휘 사용 빈도 연구: 학습교재와의 비교를 중심으로, 한국외국어대학교 대학원 석사학위논문.
홍종명(2014), 중도입국 청소년 한국어 학습 동기 연구, 語文論集 60, 중앙어문학회, 421-446.
홍종명(2016), 한국어(KSL) 교육과정 운영 체계와 중도입국 학습자의 특징, 국제한국어교육학회 학술발표논문집, 국제한국어교육학회, 72-76.
황적륜(1975), 社會言語學과 外國語敎育-Communicative Competence의 문제, 應用言語學 7(2), 서울大學校 語學硏究所, 127-142.
황종배(2004), 제2언어 습득론 개관, 서울: 경진문화사.
황현경(2018), 말하기와 쓰기에 나타난 중간언어 발달 양상 연구, 한국어와 문화 24,183-213.
Afshar, S. H., & Tofighi, S.(2021), The role of task complexity and working memory capacity in L2 accuracy and fluency. *Two Quarterly Journal of English Language Teaching and Learning University of Tabriz 13*(27), 389-416.
Ahmadian, M. J., & Tavakoli, M.(2011), The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity, and fluency in EFL learners' oral production. *Language Teaching Research 15*(1), 35-59.

Awwad, A., & Tavakoli, P.(2022), Task complexity, language proficiency and working memory: Interaction effects on second language speech performance. *International Review of Applied Linguistics in Language Teaching 60*(2), 169-196.

Beers, S. F., & Nagy, W. E.(2011), Writing development in four genres from grades three to seven: syntactic complexity and genre differentiation. *Reading and Writing 24*(2), 183-202.

Benevento, C., & Storch, N.(2011), Investigating writing development in secondary school learners of French. *Assessing Writing 16*(2), 97-110.

Bourdin, B., & Fayol, M.(1994), Is written language production more difficult than oral language production? A working memory approach. *International journal of Psychology 29*(5), 591-620.

Brezina, V.(2018), *Statistics in corpus linguistics: A practical guide*, Cambridge, MA: Cambridge University Press.

Brown, H. D.(2000), *Principles of language learning and teaching* (4th), White Plains, NY: Pearson Longman. [이흥수·박주경·이병민·이소영·최연희·차경환·이성희 공역(2015), 외국어 학습·교수의 원리, 서울: 피어슨에듀케이션코리아.]

Brown, R.(2013), *A first language: The early stages*, Cambridge, MA: Harvard University Press.

Brumfit, C.(1984), *Communicative methodology in language teaching: the roles of fluency and accuracy*, Cambridge, MA: Cambridge University Press.

Bulté, B., & Housen, A.(2012), Defining and operationalising L2 complexity. In A. Housen, F. Kuiken & I. Vedder(Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA,*(pp. 21-46). Amsterdam: John Benjamins.

Bulté, B., & Housen, A.(2017), Syntactic complexity in L2 writing: Individual pathways and emerging group trends. *International Journal of Applied Linguistics 28*(1), 147-164.

Bulté, B., Housen, A., Pierrard, M., & Van, D. S.(2008), Investigating lexical proficiency development over time—the case of Dutch-speaking learners of French in Brussels. *French Language Studies 18*(3), 277-298.

Burt, M. K.(1975), Error analysis in the adult EFL classroom. *Tesol Quarterly 9*(1), 53-63.

Bygate, M.(1988), Units of oral expression and language learning in small group interaction. *Applied linguistics 9*(1), 59-82.

Cancino, H., Rosansky, E., & Schumann, J.(1975), The acquisition of English negatives and interrogatives by native Spanish speakers. *Tesol Quarterly 9*(4), 421-430.

Carroll, J. B.(1964), *Language and thought*, Englewood Cliffs, NJ: Prentice-Hall.

Caspi, T.(2010), A dynamic perspective on second language development, University of Groningen Doctoral dissertation.

Chan, H. P., Lowie, W., & De Bot, K.(2014), *A case study of lexical development of writing and speaking in identical twins*. 16èmes Rencontres Jeunes Chercheurs (RJC 2013): Modèles et modélisation dans les sciences du langage.

Chan, H. P., Verspoor, M., & Vahtrick, L.(2015), Dynamic development in speaking versus writing in Identical twins. *Language learning 65*(2), 298-325.

Cho, M.(2018), Task complexity, modality, and working memory in L2 task performance. *System 72*, 85-98.

Chomsky, N.(1965), *Aspects of the theory of syntax*, Cambridge, MA: Cambridge University Press.

Cooper, P. L.(1984), *The assessment of writing ability: A review of research*, Princeton, NJ: Educational Testing Service.

Corder, S. P.(1967), The significance of learner's errors. *IRAL 5*(4), 161-170.

Corder, S. P.(1981), *Error analysis and interlanguage*, Oxford: Oxford University Press. [신승용·김지민·이순영·역(2011), 중간언어와 오류 분석, 서울: 박이정.]

Covington, M. A., & McFall, J. D.(2010), Cutting the Gordian knot: The moving-average type-token ratio(MATTR). *Journal of Quantitative Linguistics 17*(2), 94-100.

Crookes, G.(1989), Planning and interlanguage variation. *Studies in second language acquisition 11*(4), 367-383.

Daller, H., Van Hout, R., & Treffers-Daller, J.(2003), Lexical richness in the spontaneous speech of bilinguals. *Applied linguistics 24*(2), 197-222.

De Bot, K.(2008), Introduction: Second language development as a dynamic process. *The Modern Language Journal 92*(2), 166-178.

De Bot, K.(2020), A bilingual production model: Levelt's' speaking'model adapted. In L. Wei(Ed.), *The bilingualism Reader,*(pp. 384-404). New York, NY: Routledge.

Dulay, H. C., & Burt, M. K.(1973), Should we teach children syntax? *Language learning 23*(2), 245-258.

Dulay, H. C., Burt, M. K., & Krashen, S. D.(1982), *Language two*, Oxford: Oxford University Press.

Ellis, R.(1994), *The study of second language acquisition*, Oxford: Oxford University Press.

Ellis, R.(1997), *Second language acquisition*, Oxford: Oxford University Press. [박경자·장미경·오은진 공역(2001), 제2언어습득, 서울: 박이정.]

Ellis, R.(2003), *Task-based language learning and teaching*, Oxford: Oxford University Press.

Ellis, R., & Barkhuizen, G.(2005), *Analysing learner language*, Oxford: Oxford University Press.

Ellis, R., & Yuan, F.(2005), The effects of careful within-task planning on oral and written task performance. In R. Ellis(Ed.), *Planning and task performance in a second language,*(pp. 167-192). Amsterdam: John Benjamins.

Ferrari, S.(2012), A longitudinal study of complexity, accuracy and fluency variation in second language development. In A. Housen, F. Kuiken & V. I.(Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA,*(pp. 277-297). Amsterdam: John Benjamins.

Flower, L., & Hayes, J. R.(1981), A cognitive process theory of writing. *College composition and communication 32*(4), 365-387.

Foster, P., & Skehan, P.(1996), The influence of planning and task type on second language performance. *Studies in second language acquisition 18*(3), 299-323.

Foster, P., Tonkyn, A., & Wigglesworth, G.(2000), Measuring spoken language: a unit for all reasons. *Applied linguistics 21*(3), 354-375.

Gaies, S. J.(1980), T-unit analysis in second language research: applications, problems and limitations. *Tesol Quarterly 14*(1), 53-60.

Gilabert, R., Barón, J., & Levkina, M.(2011), Manipulating task complexity across task types and modes. In P. Robinson(Ed.), *Second language task complexity,*(pp. 105-136). Amsterdam: John Benjamins.

Grabowski, J.(2004), Speaking, writing, and memory span performance: Replicating the Bourdin and Fayol results on cognitive load in German children and adults. In L. Allal & J. Dolz(Eds.), *Proceedings writing*. Geneva: Adcom Productions.

Granfeldt, J.(2008), *Speaking and writing in L2 French: Exploring effects on fluency, accuracy and complexity*. Proceedings of the conference on complexity, accuracy and fluency in second language use, learning and teaching, Koninklijke Vlaamse Academi Brussels.

Guiraud, P.(1960), *Problèmes et méthodes de la statistique linguistique*, Paris: Presses universitaires de France.

Halliday, M. A. K.(1987), Spoken and written modes of meaning. In R. Horowitz & S. J. Samuels(Eds.), *Comprehending oral and written language*. San Diego, CA: Academic Press.

Halliday, M. A. K.(1989), *Spoken and written language*, Oxford: Oxford University Press.

Housen, A., & Kuiken, F.(2009), Complexity, accuracy, and fluency in second language acquisition. *Applied linguistics 30*(4), 461-473.

Housen, A., Kuiken, F., & Vedder, I.(2012), Complexity, accuracy and fluency: Definitions, measurement and research. In A. Housen, F. Kuiken & I. Vedder(Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA,*(pp. 1-20). Amsterdam: John Benjamins.

Hudelson, S.(1984), Kan yu ret an rayt en ingles: Children become literate in English as a second language. *Tesol Quarterly 18*(2), 221-238.

Hughes, A., & Lascaratou, C.(1982), Competing criteria for error gravity. *ELT Journal 36*(3), 175-182.

Hunt, K. W.(1965), *Grammatical structures written at three grade levels* (No.8). National Council of Teachers of English.

Hunt, K. W.(1970), Syntactic maturity in schoolchildren and adults. *Monographs of the Society for Research in Child Development 35*(1), iii-67.

Hwang, E.(2012), Korean EFL College Learners' Linguistic Features in Narrative and Argumentative Writings in terms of CAF, 숙명여자대학교 대학원 박사학위논문.

Hyun, J., & Lee, J.(2018), The effects of task complexity and working memory on Korean adult learners' English speaking performance. *English Teaching 73*(1), 115-134.

Jarvis, S.(2002), Short texts, best-fitting curves and new measures of lexical diversity. *Language Testing 19*(1), 57-84.

Johansson, V.(2008), Lexical diversity and lexical density in speech and writing: A developmental perspective. *Lund University, Department of Linguistics and Phonetics 53*, 61-79.

Johnson, W.(1944), Studies in language behavior: A program of research. *Psychological Monographs 56*(2), 1-15.

Kang, J.(2020), Speaking and writing connections in L2: The roles of multimodal teaching and learning. *어학연구 56*(2), 263-286.

Kellerman, E.(1985), If at first you do succeed. In S. Gass & C. Madden(Eds.), *Input in second language acquisition,*(pp. 345-353). Rowley, MA: Newbury House.

Kellogg, R. T.(1996), A model of working memory in writing. In C. M. Levy & S. Ransdell(Eds.), *The science of writing: theories, methods, individual dfferences and applications,*(pp. 57-72). Mahwah, NJ: Lawrence Erlbaum.

Kormos, J.(2006), *Speech production and second language acquisition*, Hillsdale, NJ: Lawrence Erlbaum.

Kormos, J., & Trebits, A.(2011), Working memory capacity and narrative task

performance. In P. Robinson(Ed.), *Second language task complexity: researching the cognition hypothesis of language learning and performance,*(pp. 128). Amsterdam: John Benjamins.

Kormos, J., & Trebits, A.(2012), The role of task complexity, modality, and aptitude in narrative task performance. *Language learning 62*(2), 439-472.

Kroll, B., & Reid, J.(1994), Guidelines for designing writing prompts: Clarifications, caveats, and cautions. *Journal of second language writing 3*(3), 231-255.

Kroll, B. M.(1981), Developmental relationships between speaking and writing. In B. M. Kroll & R. J. Vann(Eds.), *Exploring speaking-writing relationships: Connections and contrasts,*(pp. 32-54). Urbana, IL: National Council of Teachers of English.

Kroll, B. M., & Vann, R. J.(1981), *Exploring speaking-writing relationships: Connections and contrasts*, Urbana, IL: National Council of Teachers of English.

Kuiken, F., & Vedder, I.(2011), Task complexity and linguistic performance in L2 writing and speaking. In P. Robinson(Ed.), *Second language task complexity,*(pp. 91-104). Amsterdam: John Benjamins.

Kuiken, F., & Vedder, I.(2012), Syntactic complexity, lexical variation and accuracy as a function of task complexity and proficiency level in L2 writing and speaking. In A. Housen, F. Kuiken & I. Vedder(Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA,*(pp. 143-170). Amsterdam: John Benjamins.

Larsen-Freeman, D.(1997), Chaos/complexity science and second language acquisition. *Applied linguistics 18*(2), 141-165.

Larsen-Freeman, D.(2006), The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied linguistics 27*(4), 590-619.

Larsen-Freeman, D.(2009), Adjusting expectations: The study of complexity, accuracy, and fluency in second language acquisition. *Applied linguistics 30*(4), 579-589.

Larsen-Freeman, D., & Cameron, L.(2008), *Complex systems and applied linguistics*, Oxford: Oxford University Press.

Larsen-Freeman, D., & Strom, V.(1977), The construction of a second language acquisition index of development. *Language learning 27*(1), 123-134.

Laufer, B.(1991), The development of L2 lexis in the expression of the advanced learner. *The Modern Language Journal 75*(4), 440-448.

Laufer, B.(1994), The lexical profile of second language writing: Does it change over time? *RELC Journal 25*(2), 21-33.

Laufer, B., & Nation, P.(1995), Vocabulary size and use: Lexical richness in L2 written production. *Applied linguistics 16*(3), 307-322.

Lennon, P.(1990), Investigating fluency in EFL: A quantitative approach. *Language learning 40*(3), 387-417.

Lennon, P.(1991), Error: Some problems of definition, identification, and distinction. *Applied linguistics 12*(2), 180-196.

Levelt, W. J.(1993), *Speaking: From intention to articulation*, Cambridge, MA: MIT Press.

Levelt, W. J.(1999), Language production: A blueprint of the speaker. In C. Brown & P. Hagoort(Eds.), *Neurocognition of language,*(pp. 83-122). Oxford: Oxford University Press.

Lightbown, P.(1983), Exploring relationships between developmental and instructional sequences in L2 acquisition. In H. Seliger & M. H. Long(Eds.), *Classroom oriented research in second language acquisition,* 217-243. Rowley, MA: Newbury House.

Lu, X.(2011), A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *Tesol Quarterly 45*(1), 36-62.

Lyons, J.(1977), *Semantics*, Cambridge, MA: Cambridge University Press.

Malvern, D., & Richards, B.(2002), Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language Testing 19*(1), 85-104.

Malvern, D., Richards, B., Chipere, N., & Durán, P.(2004), *Lexical diversity and language development: Quantification and assessment* (1st), New York, NY: Palgrave Macmillan

McCarthy, M. J.(1990), *Vocabulary*, Oxford: Oxford University Press. [김지홍 역(2003), 어휘, 서울: 범문사.]

McCarthy, P. M., & Jarvis, S.(2007), Vocd: A theoretical and empirical evaluation. *Language Testing 24*(4), 459-488.

Michel, M.(2017), Complexity, Accuracy and Fluency(CAF). In S. Loewen & M. Sato(Eds.), *The Routledge Handbook of Instructed Second language Acquisition,*(pp. 1-27). New York, NY: Routledge.

Miller, D., & Biber, D.(2015), Evaluating reliability in quantitative vocabulary studies: The influence of corpus design and composition. *International Journal of Corpus Linguistics 20*(1), 30-53.

Milton, J., & Hopkins, N.(2006), Comparing phonological and orthographic vocabulary size: Do vocabulary tests underestimate the knowledge of some learners? *Canadian Modern Language Review 63*(1), 127-147.

Norris, J. M., & Ortega, L.(2009), Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied linguistics 30*(4), 555-578.

Nunan, D.(1999), *Second language teaching & learning*, Boston, Mass: Heinle & Heinle Publishers.

Odlin, T.(1989), *Language transfer: cross-linguistic influence in language learning*, Cambridge, MA: Cambridge University Press.

Oh, M., & Lee, H.(2012), The effects of task complexity and task condition on learner language. *응용언어학 28*(4), 39-71.

Ortega, L.(1999), Planning and focus on form in L2 oral performance. *Studies in second language acquisition 21*(1), 109-148.

Ortega, L.(2003), Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied linguistics 24*(4), 492-518.

Ortega, L.(2015), Syntactic complexity in L2 writing: Progress and expansion. *Journal of second language writing 29*, 82-94.

Oshima, A., & Hogue, A.(2009), *Writing academic english: Level 4* (4th), New York, NY: Pearson Longman.

Pallotti, G.(2009), CAF: Defining, refining and differentiating constructs. *Applied linguistics 30*(4), 590-601.

Pica, T., Holliday, L., Lewis, N., & Morgenthaler, L.(1989), Comprehensible output as an outcome of linguistic demands on the learner. *Studies in second language acquisition 11*(1), 63-90.

Polat, B., & Kim, Y.(2014), Dynamics of complexity and accuracy: A longitudinal case study of advanced untutored development. *Applied linguistics 35*(2), 184-207.

Polio, C. G.(1997), Measures of linguistic accuracy in second language writing research. *Language learning 47*(1), 101-143.

Powers, D. E., & Fowles, M. E.(1998), Test takers'judgments about GRE writing test prompts. *ETS Research Report Series*(2), i-23.

Read, J.(2000), *Assessing vocabulary*, Cambridge, MA: Cambridge University Press.

Richards, B., & Malvern, D.(2000), Accommodation in oral interviews between foreign language learners and teachers who are not native speakers. *Studia Linguistica 54*(2), 260-271.

Richards, J. C.(1971), A non-contrastive approach to error analysis. *ELT Journal 25*(3), 204-219.

Robinson, P.(2001a), Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied linguistics*

22(1), 27-57.
Robinson, P.(2001b), *Cognition and second language instruction*, Cambridge, MA: Cambridge University Press.
Robinson, P.(2003), The cognition hypothesis, task design, and adult task-based language learning. *Second language studies 21*(2), 45-105.
Robinson, P.(2005), Cognitive complexity and task sequencing: studies in a componential framework for second language task design. *IRAL - International Review of Applied Linguistics in Language Teaching 43*(1), 1-32.
Robinson, P., & Gilabert, R.(2007), Task complexity, the cognition hypothesis and second language learning and performance. *International Review of Applied Linguistics in Language Teaching 45*(3), 161-176.
Ryoo, Y.(2018), Comparing lexical diversity and lexical sophistication in Korean EFL writing: Topic and text length. *Multimedia-Assisted Language Learning 21*(3), 63-87.
Selinker, L.(1972), INTERLANGUAGE. *IRAL - International Review of Applied Linguistics in Language Teaching 10*(3), 209-231.
Skehan, P.(1996a), A framework for the Implementation of task-based Instruction. *Applied linguistics 17*(1), 38-62.
Skehan, P.(1996b), Second language acquisition research and task-based instruction. In J. Willis & D. Willis(Eds.), *The challenge and change in language teaching*. Oxford: Heinemann.
Skehan, P.(1998), *A cognitive approach to language learning*, Oxford: Oxford University Press.
Skehan, P.(2001), Tasks and language performance assessment. In M. Bygate, P. Skehan & M. Swain(Eds.), *Researching pedagogic tasks: Second language learning, teaching, and testing,*(pp. 167-185). Harlow, Essex: Pearson Education.
Skehan, P.(2003), Task-based instruction. *Language Teaching 36*(1), 1-14.
Skehan, P.(2009a), Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied linguistics 30*(4), 510-532.
Skehan, P.(2009b), Models of speaking and the assessment of second language proficiency. In A. Benati(Ed.), *Issues in second language proficiency,*(pp. 202-215). London, NY: Continuum publishing.
Skehan, P., & Foster, P.(1997), Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research 1*(3), 185-211.
Skehan, P., & Foster, P.(1999), The influence of task structure and processing

conditions on narrative retellings. *Language learning 49*(1), 93-120.

Skehan, P., & Foster, P.(2001), Cognition and tasks. In P. R. Robinson(Ed.), *Cognition and second language instruction,*(pp. 183-205). Cambridge, MA: Cambridge University Press.

Song, J.(2005), Task-processing conditions as influences on spoken language performance. *English Teaching(영어교육) 60*(3), 117-137.

Spoelman, M., & Verspoor, M.(2010), Dynamic patterns in development of accuracy and complexity: A longitudinal case study in the acquisition of Finnish. *Applied linguistics 31*(4), 532-553.

Towell, R.(2012), Complexity , accuracy and fluency from the perspective of psycholinguistic second language acquisition research. In A. Housen, F. Kuiken & I. Vedder(Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA,*(pp. 47-69). Amsterdam: John Benjamins.

Ure, J.(1971), Lexical density and register differentiation. *Applications of linguistics 23*(7), 443-452.

Van Dijk, M., Verspoor, M., & Lowie, W.(2011), Variability and DST. In M. Verspoor, K. de Bot & W. Lowie(Eds.), *A dynamic approach to second language development methods and techniques,*(pp. 55-84). Amsterdam: John Benjamins.

Vanpatten, B.(1990), Attending to form and content in the input: An experiment in consciousness. *Studies in second language acquisition 12*(1), 287-301.

Vercellotti, M. L.(2012), Complexity, accuracy, and fluency as properties of language performance: The development of the multiple subsystems over time and in relation to each other, University of Pittsburgh Doctoral dissertation.

Verspoor, M., Lowie, W., & Van Dijk, M.(2008), Variability in second language development from a dynamic systems perspective. *The Modern Language Journal 92*(2), 214-231.

Waller, T.(1993), Characteristics of near-native proficiency in writing. In H. Ringbom(Ed.), *Near-native proficiency in English* (pp. 183-293). Abo, Finland: Akademi University.

Weissberg, B.(2000), Developmental relationships in the acquisition of English syntax: Writing vs. speech. *Learning and instruction 10*(1), 37-53.

Wickens, C. D.(1980), The structure of attentional resources. *Attention and performance VIII 8*, 239-257.

Wolfe-Quintero, K., Inagaki, S., & Kim, H. Y.(1998), *Second language development in writing: Measures of fluency, accuracy & complexity*, Honolulu, Hawaii:

Second Language Teaching & Curriculum Cener, University of Hawai'i Press.

Yu, G.(2010), Lexical diversity in writing and speaking task performances. *Applied linguistics 31*(2), 236-259.

Zalbidea, J.(2017), 'One task fits all'? The roles of task complexity, modality, and working memory capacity in L2 performance. *The Modern Language Journal 101*(2), 335-352.

Abstract

A Study on the Korean Language Development of Chinese Immigrant Youth

- Focusing on Complexity and Accuracy in Written and Spoken Language

Zhao, Qinghua

Korean Language Education as a Foreign Language
Department of Korean Language Education
The Graduate School
Seoul National University

This research aims to investigate the development of complexity and accuracy in both written and spoken language among Chinese immigrant youth in Korea.

Chinese immigrant youth relocate to Korea with their parents in their adolescents to settle in Korean society and pursue education. One significant challenge they face is attaining proficiency in the Korean language. However, there is still a lack of research that directly examines the Korean language development of Chinese immigrant youths, despite various measures being taken to support their Korean language education. Therefore, the purpose of this study is to explore the various aspects of Korean language development among Chinese immigrant youths and reveal their distinctive characteristics.

For this purpose, the written and spoken data collected from three Chinese immigrant youths over a period of 1 year and 6 months were analyzed. In this study, complexity was categorized into syntactic complexity and lexical complexity. Syntactic complexity was measured by

unit length, the number of clauses per unit, and the type of unit expansion. Lexical complexity was measured by lexical diversity, lexical density, and lexical sophistication. Accuracy was measured using the ratio of error-free clauses, the number of error-free clauses per unit, and the number of errors per unit.

Based on these measurement criteria, an examination of learners' written and spoken language development revealed that their syntactic complexity increased over time, while the development of lexical complexity and accuracy did not show clear patterns. Specifically, both written and spoken language exhibited a consistent growth in unit length and the number of clauses per unit over time, with a particularly noticeable increase in adverbial clauses.

In terms of lexical complexity, lexical diversity showed a notable rise during the periods of 1-6 months and 7-12 months for both written and spoken language. However, it leveled off around 13-18 months, which aligned with approximately three years after immigration. Lexical density did not show significant changes across different time periods for both written and spoken language. The lexical sophistication did not show significant differences over time in written language, whereas in spoken language, it experienced significant growth around 13-18 months, reaching a level similar to written language. Regarding accuracy, both written and spoken language demonstrated a simultaneous increase in the number of error-free clauses and the number of errors as time progressed, showing a lack of improvement in accuracy.

Comparing complexity and accuracy based on the output formats, it was observed that unit length, the number of clauses per unit, and lexical diversity were all higher in written language compared to spoken language. However, there were no significant differences in lexical density and lexical sophistication between written and spoken language. Regarding accuracy, there were no significant differences in the ratio of error-free clauses, the number of error-free clauses per unit, and the number of errors per unit

between written and spoken language. However, it was found that omissions and misformation errors were more frequent in written language, while substitutions and additions errors were more frequent in spoken language, indicating differences in error patterns based on the output format.

The relationship between complexity and accuracy showed contrasting results depending on the measurement criteria.

Both written and spoken language exhibited distinct positive correlations between syntactic complexity and the number of error-free clauses and between syntactic complexity and the number of errors. Therefore, it was not possible to determine a fixed supportive or competitive relationship between them. However, by comparing the sizes of correlation coefficients, it was observed that in written language, syntactic complexity was more closely related to the number of error-free clauses, while in spoken language, syntactic complexity was more closely related to the number of errors.

Based on these results, the following implications have been derived. First, Chinese immigrant youth in the natural public education environment without special Korean language education continue to develop the structural complexity of sentences over time. However, they show limitations in terms of grammatical item diversity and accuracy, often repeating the use of limited items. To address this, more explicit education focused on the meanings and forms of various endings, particularly adverbial clauses, noun clauses, and quotation clauses, is necessary. Second, learners demonstrated limitations in developing lexical diversity and lexical sophistication solely through natural input. To promote continuous expansion of learners' lexical complexity, there is a need for education on vocabulary formation principles, such as compounding and derivation, based on the previously acquired vocabulary. Third, learners' accuracy did not necessarily improve over time in the natural environment. Therefore, continuous education on accurate usage of particles, prefinal endings, and connective endings, along with vocabulary enhancement, is required to

reduce lexical errors. Fourth, learners showed differences in syntactic and lexical complexity, as well as differences in error areas and error phenomena. based on the output modality. To achieve balanced development in written and spoken language, individualized pedagogical treatment considering the output formats needs to be provided. Fifth, learners frequently use colloquial language in written texts, indicating a need for explicit education on linguistic features based on the functional usage of written and spoken language.

This study collected longitudinal data to investigate the development of written and spoken language in Chinese immigrant youths. The significance of this study lies not only in analyzing the complexity and accuracy in the data but also in exploring the relationship between these two aspects to comprehensively investigate language development over time. Additionally, by comparing written and spoken language, it sheds light on the specific differences in complexity and accuracy based on the output formats. These findings can serve as a valuable foundation for future research on the Korean language development of immigrant youths and the design of educational approaches.

Keywords : immigrant youth, written language, spoken language, syntactic complexity, lexical complexity, accuracy

Student Number : 2018-31140