

수학능력시험 영어 읽기 지문의 응집성과 문항 난이도 간의 상관관계 분석

Analysis of Correlation Between Cohesion and Item Difficulty in English Reading Section of CSAT

황이수, 이제영
전주대학교 영어교육과

Leesu Hwang(hi880807@gmail.com), Je-Young Lee(jylee@jj.ac.kr)

요약

본 연구의 목적은 수학능력시험 영어 영역 문항의 정답률에 영향을 미치는 응집성 관련 텍스트 요인을 찾기 위함이다. 이를 위해, 2014~2018학년도 수능 및 수능 모의고사의 읽기 지문 299개를 수집하여 응집성 관련 지표를 코메트릭스를 통해 분석하고 이를 실제 문항 정답률과 상관분석하였다. 이를 통해 얻어진 결론은 다음과 같다. 첫째, 문장 사이의 특정 논항이 반복되어 응집성이 강해지면 학습자의 언어적 부담이 줄어들면서 문항 정답률이 상승하는 경향이 있었다. 둘째, 잠재 의미 분석의 8개 하위 영역과 문항 정답률 사이에는 통계적으로 유의한 상관관계가 발견되지 않았다. 마지막으로 본 연구의 교육적 함의와 후속연구를 위한 제언을 논의하였다.

■ 중심어 : | 대학수학능력시험 | 코메트릭스 | 지시적 응집성 | 잠재의미분석 | 문항 난이도 |

Abstract

The purpose of this study was to investigate the cohesion-related text factors that affected the correct answer rate of the English reading section in the Korean CSAT. To this end, reading passages of the 2014-2018 CSAT and CSAT mock test were collected and analyzed in terms of cohesion. And then, correlation analysis was implemented between the results regarding cohesion and item difficulty. The conclusions were as follows. First, the correct answer rate tended to increase as the linguistic burden of learners was reduced, when specific arguments between sentences were repeated. Second, there was no statistically significant correlation between the 8 sub-area of latent semantic analysis and the correct answer rate. Finally, the pedagogical implications of this study and suggestions for further study were discussed.

■ keyword : | CSAT | Coh-Metrix | Referential Cohesion | Latent Semantic Analysis | Item Analysis |

I. 서론

대학수학능력시험(이하 수능)은 학생의 대학 진학 및 진로 설정에 매우 큰 영향을 미치는 대표적인 고부담시

험(high-stakes test)이다. 따라서 수능의 난이도를 적절한 수준으로 유지하는 것은 고등학교 교육과정을 정상적으로 운영하고 수험자 및 관계자의 혼란을 최소화하기 위해 꼭 필요하다.

접수일자 : 2020년 03월 11일
수정일자 : 2020년 04월 20일

심사완료일 : 2020년 04월 23일
교신저자 : 이제영, e-mail : jylee@jj.ac.kr

특히 2018학년도 수능부터는 영어 영역이 절대평가의 형태로 시행되고 있으며, 이는 성취기준에 따른 난이도의 설정과 예측이 더욱 중요해진다고 할 수 있다. 한국교육과정평가원[1]에서는 수능 영어 영역의 절대평가의 범위는 영어 I과 영어 II의 성취기준에 근거한다고 발표하였다. 또 2015 개정 교육과정부터는 과정중심 평가와 같은 성취평가가 강조되고 있다[2].

수능 영어 영역의 난이도 조절을 위해서는 문항 난이도에 영향을 미치는 변인을 탐색할 필요가 있으며, 이러한 요인은 텍스트 내적 변인(텍스트 자체의 특성)과 외적 변인(학습자 요인 등)으로 구분할 수 있다[3]. 이중 수능 영어 읽기 지문의 텍스트 내적 변인을 탐색하는 연구는 상당수 존재한다[4-8]. 하지만 대부분의 연구들은 수능 읽기 지문의 텍스트 요인과 문항 정답률 사이의 상관관계를 직접적으로 밝히는 연구는 수능 읽기 지문의 통사적 특징과 정답률 사이의 관계를 상관분석한 황이수, 이제영의 연구[8] 외에는 찾기 힘들다.

이에 본 연구에서는 코메트릭스로 분석할 수 있는 텍스트 요인 중 응집성(cohesion)과 관련된 지시적 응집성(referential cohesion)과 잠재 의미 분석(latent semantic analysis) 지표의 분석 결과와 실제 수능 문항의 정답률 사이의 상관관계를 분석하고자 한다. 이를 위해 다음과 같은 연구 문제를 설정하였다.

첫째, 지시적 응집성의 하위 지표 중 수능 영어 읽기 문항의 정답률과 통계적으로 유의한 상관관계를 갖고 있는 것은 무엇이며, 그 크기는 어떠한가?

둘째, 잠재 의미 분석의 하위 지표 중 수능 영어 읽기 문항의 정답률과 통계적으로 유의한 상관관계를 갖고 있는 것은 무엇이며, 그 크기는 어떠한가?

II. 선행연구

1. 수능 난이도 관련 연구

수능 영어 영역 평가의 문항 난이도 관련 연구로는 다음과 같은 것들이 있다. 먼저 지슬기, 김해동[9]은 지문 친숙도와 문항 유형이 문항 난이도에 미치는 영향을 살펴보았다. 그 결과 지문 친숙도보다 문항 유형의 해

당 문항의 난이도에 더 큰 영향을 미치는 것으로 밝혀졌다. 이경숙[10]은 지문 친숙도가 높은 영어 능력을 갖춘 학생들의 듣기와 읽기 시험 수행에 영향을 준 반면, 영어 실력이 낮은 학생들의 시험 수행에는 영향을 주지 않는다고 주장하였다.

박희진 외[11]는 수능 및 모의고사의 영어 영역 지문을 수집하여 토픽모델링을 실시하였고, 이를 통해 얻어진 토픽과 문항 정답률과의 상관관계를 살펴보았다. 그 결과 사회나 가정을 주제로 한 지문이 포함된 문항에서는 정적 상관관계, 즉 학생들의 정답률이 높았으며, 전문 지식 관련 지문이 포함된 문항에서는 부적 상관관계가 발견되었다.

황이수, 이제영[8]은 2014~2018학년도의 수능 및 수능 모의고사의 영어 지문의 기초 산출치, 통사적 복잡성, 통사적 패턴 밀집성을 코메트릭스를 통해 분석한 후, 이를 해당 문항 정답률과 상관분석 하였다. 그 결과 기초산출치 부분에서는 단어 내 평균 음절 수, 단어내 평균 알파벳 수, 문장 내 평균 단어 수, 텍스트의 총 단어 수는 문항 정답률과 통계적으로 유의한 수준의 부적 상관관계를 나타내었다. 한편 통사적 복잡성과 통사적 패턴 밀집성 부분에서는 문단 내 연속하는 문장들 사이에서 단어 거리와 부정 표현 빈도 등에서 부적 상관관계가 발견되었다.

2. 코메트릭스 관련 연구

2009년 중학교 1학년 영어 교과서의 코퍼스 언어학적 특성을 분석한 전문기, 임인재[12]의 연구 이후로 코메트릭스를 분석 도구로 활용한 국내 연구가 활발히 진행되고 있다. 이러한 선행연구들을 분석 대상을 기준으로 살펴보면, 크게 영어 교과서 관련 연구[12-17], 수능과 공무원 시험 등 평가 지문 관련 연구[4-8], 영어 학습자의 작문 결과물 관련 연구[18-20] 등으로 구분할 수 있다.

위의 선행연구 중 본 연구와 같이 수능 영어 영역 문항의 읽기 지문을 코메트릭스로 분석한 연구를 중심으로 그 내용을 살펴보면 다음과 같다. 고나은, 신정아[4]는 EBS 교재와 수능의 연계 정책 실행 이후로 수능 시험 영어 지문의 난이도에 변화가 있는지 코메트릭스를 통해 비교 및 분석하였다. 그 결과 어휘적 측면, 통사적

측면, 이독성 측면에서 연계 정책 실시 이후 난이도가 증가했다. 다만 응집성 측면에서는 연계 전후 통계적으로 유의한 차이가 발견되지 않았다.

이혜진, 이재영[6]은 한·중·일의 대입 시험인 수능, 가오카오, 센터시험의 이독성을 코메트릭스를 통해 분석하여 비교하였다. FKGL(Flesch Kincaid Grade Level)을 통한 이독성 비교 결과, 한국 10.3학년, 중국 8.3학년, 일본 7.8학년의 순으로, 한국 수능 영어 지문의 난이도가 일본과 중국의 대입 시험에 비해 높으며 미국의 10학년 수준에 해당하는 수준임을 밝혔다.

최민주, 김정렬[7]은 코메트릭스를 활용해 수능 영어 영역의 문항 유형별로 응집성과 어휘정보에 차이가 있는지 비교하였다. 그 결과 응집성 측면에서는 명사중복 및 어간중복 측정치에서, 어휘정보 부분에서는 대명사 발생 정도, 습득나이, 내용어의 구체성, 심상성 등에서 문항 유형에 따라 통계적으로 유의한 수준의 차이가 있음을 보고하였다.

III. 연구 방법

1. 분석 대상

분석 대상은 크게 수능 영어 영역의 읽기 지문과 각 읽기 지문을 통해 출제된 문항의 정답률이다. 먼저 수능 읽기 지문은 2014-2017학년도 4년간 실시된 수능 시험과 2014-2018년 6월과 9월에 2회 실시된 수능 모의평가에서 수집하였다. 읽기 지문 수집 기간이 2018년 8월이었기 때문에 2018년 자료는 6월 모의고사의 지문만 포함되었다. 지문 수집 과정 중 빈칸 채우기 문항의 경우 빈칸에 해당하는 내용을 추가하였고, 완전한 문장 형태가 아닌 그래프와 광고 등이 포함된 실용문은 분석 대상에서 제외하였다. 각 회차에는 25개의 읽기 지문이 있으며, 문장 형태가 아닌 2개의 지문을 제외한 23개의 지문을 분석 대상에 포함시켰고, 이러한 과정을 통해 13회차 총 299개의 지문을 분석에 활용하였다.

수집된 읽기 지문이 포함된 해당 문항의 정답률은 EBSi 웹사이트에서 제공하는 '풀서비스'를 통해 수집하였다. 이 서비스는 수험생이 각 문항에 대해 자신이 선

택한 답을 입력하면 자신의 등급을 확인할 수 있게 해주는 서비스이다. 참고로 분석 대상에 포함된 수능 및 수능 모의고사에 대해 성적을 입력한 학생은 6,648~18,321명이다.

2. 분석 도구

본 연구에서는 코메트릭스(Coh-Metrix)를 활용해 각 지문의 지시적 응집성과 잠재 의미 분석을 실시하였다. 코메트릭스는 미국 뎀퍼스 대학의 Arthur C. Graesser와 Danielle S. McNamara가 개발한 언어 분석 시스템으로 특정 텍스트의 다양한 특징을 분석하여 수치로서 제시해주는 기능을 갖고 있는데[12], 특히 기존의 텍스트 분석 도구와 달리 텍스트의 응집성에 대한 정보를 제공하는 것이 특징이다[6]. 코메트릭스에서 제공하는 정보는 기초 산출치(Descriptive Indices), 어휘적 다양성(Lexical Diversity) 등 총 11개 항목의 106개의 하위 지표에 대한 정량적 정보를 제공한다[21]. 이 중 본 연구에서는 지시적 응집성과 잠재 의미 분석 부분에 집중하였다.

2.1 지시적 응집성

지시적 응집력은 근접한 문장과 동일지시 사이에서 내용어가 얼마나 중복되는지를 의미하는 지표이다. 응집력을 구분하는 기준은 크게 국지적 응집성(local cohesion)과 포괄적 응집성(global cohesion)으로 구분할 수 있다[22][23]. 국지적 응집성은 연결사, 인접 문장 사이의 중복어 등을 통해 측정하며, 포괄적 응집성은 단락에 포함된 모든 문장 사이의 중복어를 측정한다.

한편 중복대상의 구체적인 정도에 따라서 명사 중복, 논항 중복, 어간 중복, 내용어 중복을 구분하는 방식으로 지시적 응집성을 측정할 수 있다. 명사 중복, 논항 중복, 어간 중복, 내용어 중복을 측정하는 지표의 경우, 선행 문장에 중복어가 있는지 여부를 통해 0 또는 1로 표시하며, 내용어의 중복을 측정하는 지표들(CRFCW01, CRFCW01d, CRFCW0a, CRFCW0ad)은 중복비율로 표시된다. 특히 어휘적 공지시(lexical co-reference)는 글을 이해하는 능력과 읽는 속도에 영향을 주는 응집력을 나타내는 중요한 지표이다.

표 1. 지시적 응집력 하위 지표

범주	내용
CRFNO1	중복 명사 비율(인접 문장)
CRFAO1	대명사 중복비율(인접 문장)
CRFSO1	내용어와 중복(인접 문장)
CRFNOa	중복 명사 비율(텍스트 내)
CRFAOa	대명사 중복비율(텍스트 내)
CRFSOa	내용어와 중복(텍스트 내)
CRFCWO1	인접문장간 내용어 중복 비율
CRFCWO1d	인접문장간 내용어 중복 비율(표준편차)
CRFCWOa	인접 문장간 조응소 중복
CRFCWOad	텍스트 내 조응소 중복

2.2 잠재 의미 분석

잠재 의미 분석은 단어 및 문서 분석에 사용되는 의미와 개념 분석 모델로, 문장 간, 단락 간의 의미적 중복 정도를 분석하여 제시한다[21][24]. 코메트릭스에서는 잠재 의미 분석과 관련하여 총 8개의 지표를 제공한다. 각각의 지표는 0(낮은 응집성)부터 1(높은 응집성)로 나타나며, 한 개의 분명한 뜻을 가진 단어와 그 의미와 관련되거나 유사한 단어들 사이의 의미적 중복을 측정한다. 코메트릭스에서는 잠재 의미 분석 지표를 이용해서 특정 정보가 앞의 문맥에서 주어진 정보인지, 혹은 새로 등장하는 정보인지를 추적한다. 이때 기존에 주어진 정보가 다시 등장하는 빈도가 높을수록 독자의 인지적 부담을 덜어주며, 따라서 글을 보다 쉽게 이해할 수 있게 한다.

표 2. 잠재 의미 분석 하위 지표

범주	내용
LSASS1	인접 문장 간 의미적 응집력(평균)
LSASS1d	인접 문장 간 의미적 응집력(표준편차)
LSASSp	단락 내 모든 문장 사이의 의미적 응집력(평균)
LSASSpd	단락 내 모든 문장 사이의 의미적 응집력(표준편차)
LSAPP1	단락 별 의미적 응집력 평균
LSAPP1d	단락 별 의미적 응집력 표준편차
LSAGN	각 단락에 신정보가 이전 단락의 내용과 비교하여 얼마나 주어졌는지에 대한 측정치(평균)
LSAGNd	각 단락에 신정보가 이전 단락의 내용과 비교하여 얼마나 주어졌는지에 대한 측정치(표준편차)

3. 자료 처리 방법

본 연구의 자료 처리는 크게 2가지 구분된다. 첫째, 코메트릭스를 통해 능숙 영어 읽기 지문을 분석하여 각 지문의 지시적 응집력과 잠재 의미 분석 관련 지표의 결과를 얻고, 각각의 평균과 표준편차를 산출하였다. 둘

째, 코메트릭스를 통해 얻어진 결과와 EBSi 풀서비스에서 수집한 정답률을 SPSS 18에 문항별로 입력하여 상관분석(Pearson의 적률상관계수)을 실시하였다.

IV. 연구 결과

1. 지시적 응집성

지시적 응집성은 명사, 대명사, 내용어, 그리고 조응소의 인접해있는 문장, 또는 텍스트 내의 중복비율로 결정된다. 이는 공통의 논항을 포함하고 있는 인접 문장쌍들의 수를 전체 인접 문장 쌍들의 수로 나눈 비율로써 산출된다[12].

다음의 [표 3]은 지시적 응집력의 각 하위 지표와 정답률과의 상관계수를 나타낸 것이다.

표 3. 지시적 응집력 하위 지표별 결과 (N=299)

범주	평균	표준편차	상관계수	유의도
CRFNO1	.37	.26	.002	.967
CRFAO1	.54	.25	.162**	.005
CRFSO1	.48	.27	-.079	.175
CRFNOa	.34	.23	.015	.791
CRFAOa	.50	.22	.191**	.001
CRFSOa	.45	.24	-.046	.423
CRFCWO1	.11	.08	.231**	.000
CRFCWO1d	.10	.04	.123**	.033
CRFCWOa	.10	.07	.272**	.000
CRFCWOad	.10	.03	.182**	.002

분석 결과 지시적 응집성과 관련한 10개의 하위 지표 중 문항 정답률과 통계적으로 유의한 수준의 상관관계를 보인 항목은 총 6개였으며, 이들은 모두 정적 상관관계였다. 이들을 상관관계의 정도로 구분하면 다음과 같다.

먼저 상관계수가 0.2 이상에 해당하는 요인은 CRFCWOa(인접 문장간 조응소 중복)과 CRFCWO1(인접 문장간 내용어 중복 비율)이었다. 조응소(anaphora)란 앞에 나온 단어나 어구를 지칭하는 대명사를 의미한다[25]. 따라서 선행 문장에 포함된 단어를 대명사로 받거나 해당 내용어가 계속 반복되는 중복 비율이 높을수록 해당 지문의 난이도가 하락하며, 이는 높은 정답률과 연결되는 것으로 해석할 수 있다.

다음으로 0.1~0.2에 해당하는 상관관계를 보인 요인에는 CRFAOa(인접 문장간 조음소 중복), CRFCWOad(텍스트 내 조음소 중복), CRFAO1(인접 문장간 대명사 중복비율), CRFCWO1d(인접 문장간 내용어 중복 비율의 표준편차)순으로 나타났다. 이와 같은 결과는 지시적 응집성이 높아질수록 읽기 속도 및 이해의 정도가 증가할 뿐만 아니라[26][27], 문항 난이도에도 상당한 영향을 준다는[28][29] 기존 선행연구 결과와 일치한다.

2. 잠재 의미 분석

잠재 의미 분석은 단어, 문장, 단락과 같은 두 텍스트 요소 사이의 의미적 유사성을 산출하는 방법으로, 하나의 단어와 의미적으로 관계가 있는 다른 단어들이 얼마나 중복되어 있는가를 나타내는 지표이다[12][30]. 이경남[31]에 따르면 이 분석은 독자의 의미 표상 과정에서 잠재 변수(latent variable)가 영향을 주는데, 잠재 변수는 통계 용어로 직접 관찰할 것은 아니지만 눈에 보이지 않는 영향을 주는 변수가 있다고 가정하는 변수이다. 즉, 어떤 현상이 발생하였지만 눈에 보이지 않는 영향을 주는 의미와 관련한 변수가 있다고 가정하고 분석하는 방법이 잠재적 의미 분석이다. 예를 들면, home이라는 단어는 인접한 문장, 혹은 단락 내 다른 문장 안에 있는 house나 table과 같은 단어와 상대적으로 높은 의미적 관계에 있다고 말할 수 있다[21].

다음의 [표 4]는 잠재 의미 분석의 각 하위 지표와 정답률과의 상관계수를 나타낸 것이다.

표 4. 잠재 의미 분석 하위 지표별 결과 (N=299)

범주	평균	표준편차	상관계수	유의도
LSASS1	.23	.12	.034	.560
LSASS1d	.15	.05	-.055	.343
LSASSp	.21	.11	.029	.613
LSASSpd	.15	.04	-.014	.811
LSAPP1	.01	.05	.088	.130
LSAPP1d	.00	.01	.051	.379
LSAGN	.28	.06	.005	.925
LSAGNd	.15	.04	.056	.335

분석 결과 잠재 의미 분석의 8개의 하위 지표 중 정답률과 통계적으로 유의한 수준의 상관관계가 발견된 지표는 없었다. 즉, 수능 읽기 지문에서 의미가 서로 유

사한 단어들이 중복되어 등장하는 비율과 실제 문항 난이도와는 의미 있는 상관관계는 없다고 해석할 수 있다. 이러한 결과는 잠재 의미 분석 결과와 문항 난이도 사이에서 통계적으로 유의한 수준의 상관관계를 발견하지 못한 Aldabe와 Maritxalar[32]의 연구 결과와는 일치하는 결과이다. 다만, 문항 난이도와 직접적으로 관련된 연구는 아니지만, 단답형 문항의 채점에 잠재 의미 분석을 활용한 결과 채점자가 직접 채점한 결과와 높은 수준의 상관관계($r=.94$)를 기록한 Lavoie, Parker, Legree, Ardison, & Kilcullen[33]의 연구 결과를 볼 때 잠재 의미 분석을 평가에 활용할 수 있는 방안에 대한 연구가 보다 활발히 이어져야 할 것으로 보인다.

V. 결론

본 연구는 수능 영어 영역 문항의 정답률에 영향을 미치는 응집성 관련 텍스트 요인을 찾기 위해 2014~208학년도 수능 및 평가원 수능 모의고사의 지문을 수집하여 코메트릭스를 통해 응집성 관련 요인들을 분석하고 이를 실제 문항 정답률과 상관분석하였다. 이를 통해 얻어진 결론은 다음과 같다.

첫째, 지시적 응집성과 관련한 하위 지표 중 인접 문장 간 조음소 중복, 인접 문장 간 내용어 중복 등에서 정답률과 통계적으로 유의한 수준의 정적 상관관계가 발견되었다. 이는 문장 사이의 특정 논항이 반복되어 응집성이 강해지면 학습자의 언어적 부담이 줄어들면서 문항 정답률이 상승하는 경향이 있음을 나타낸다.

둘째, 잠재 의미 분석의 8개 하위 지표에서는 절대값 기준으로 .014에서 .088의 상관계수가 도출되었으며, 이는 모두 통계적으로 유의한 수준의 결과가 아니었다. 잠재 의미 분석이 의미적으로 서로 유사한 단어들의 중복 정도를 나타낸다는 것을 고려할 때, 수사적 조직(rhetorical organization)을 통해 발생하는 응집성과 문항의 난이도 사이에는 뚜렷한 상관관계는 없는 것으로 나타났다.

2019년 11월 교육부는 '대입제도 공정성 강화 방안'을 통해 2023학년도 대입까지 16개 주요 대학의 정식

비율, 즉 수능을 통한 학생 선발 비율을 40% 이상이 되도록 추진할 것이라고 발표했다. 이처럼 수능의 중요성이 꾸준히 유지되고 있는 현실에서 수능 평가의 적절한 난이도 설정과 일관된 난이도 유지의 중요성은 두말할 나위 없이 중요하다. 이러한 측면에서 본 연구의 결과가 수능 영어 영역 문항의 정확한 난이도 예측을 위한 기초 자료로서 활용되기를 기대한다. 또한 기존의 텍스트 요인 분석 연구에서 한발 나아가 읽기 지문의 텍스트 요인과 실제 문항 난이도 사이의 상관관계를 직접적으로 밝히는 연구가 보다 다양한 측면에서 진행될 필요가 있다.

* 본 연구는 주저자의 학위논문 일부를 재구성 및 보완한 것임

참 고 문 헌

- [1] <http://www.suneung.re.kr/sub/info.do?m=0401&s=suneung>
- [2] 조성준, “성취평가중심 영어수업 활성화를 위한 고등학교 성취평가 현황 분석 연구,” 한국콘텐츠학회논문지, 제18권, 제4호, pp.550-566, 2018.
- [3] 이종승, 김성훈, 김재철, 손형정, 박문환, 장경숙, “대학 수학능력시험 문항난이도 추정모형 개발: 언어 영역, 수리 영역, 영어 영역을 중심으로,” 교육평가연구, 제16권, 제2호, pp.1-24, 2003.
- [4] 고나은, 신정아, “Coh-Metrix를 이용한 수능 영어 읽기 영역 지문 난이도 비교: EBS-수능 연계 정책 전후,” 중등영어교육, 제10권, 제4호, pp.3-24, 2017.
- [5] 신유선, 정영경, “코메트릭스(Coh-Metrix)를 활용한 공무원 7급과 9급 영어 시험 지문 분석: 어휘 다양성, 통사적 복잡성, 응집성과 이독성 지수를 중심으로,” 영미어문학, 제129호, pp.143-161, 2018.
- [6] 이혜진, 이혜영, “한-중 일 대학 입학시험 영어 지문의 이독성 비교 분석,” 인문사회21, 제9권, 제5호, pp.441-453, 2018.
- [7] 최민주, 김정렬, “수능 영어 문항 유형간 응집력과 어휘정보 분석,” 한국콘텐츠학회논문지, 제17권, 제12호, pp.378-385, 2017.
- [8] 황이수, 이혜영, “수학능력시험 영어 지문의 텍스트 요인과 문항 난이도의 상관관계 분석: 통사적 복잡성을 중심으로,” 영어영문학연구, 제46권, 제1호, pp.265-283, 2020.
- [9] 지슬기, 김해동, “문항 유형과 지문 친숙도가 영어 읽기 평가 결과에 미치는 영향력 비교,” 외국어교육, 제21권, 제1호, pp.215-239, 2014.
- [10] 이경숙, “문제 수, 지문길이, 지문 친숙도가 영어 청해와 독해시험에 미치는 영향,” 영어교육, 제54권, 제4호, pp.327-351, 1999.
- [11] 박희진, 장경애, 이운호, 김우제, 강필성, “텍스트마이닝을 이용한 대학수학능력시험 영어영역 정답률 예측,” 대한산업공학회 춘계공동학술대회 논문집, pp.2277-2288, 2015.
- [12] 전문기, 임인재, “코메트릭스(Coh-Metrix)를 이용한 중학교 1학년 개정 영어 교과서의 코퍼스 언어학적 비교 분석,” 영어교육연구, 제21권, 제4호, pp.265-292, 2009.
- [13] 김정렬, 양지윤, “Coh-Metrix를 통한 초중등 영어 교과서 연계성 분석,” 영어교육, 제67권, 제2호, pp.319-341, 2012.
- [14] 배지영, “중학교 1-2학년 2015 개정 영어교과서 읽기 자료의 가독성 분석과 아동문학 선정 연구,” 영어교과교육, 제18권, 제2호, pp.117-141, 2019.
- [15] 송해성, “고등학교 영어 교과서 듣기 과업의 학년별 연계성 연구,” 영어어문교육, 제19권, 제4호, pp.279-305, 2013.
- [16] 이승환, “코메트릭스(Coh-Metrix)를 이용한 고등학교 영어1과 영어2 교과서 문어 텍스트의 코퍼스 언어학적 난이도 분석,” 외국어교육연구, 제27권, 제2호, pp.131-148, 2013.
- [17] 함미영, 이정원, “영어 텍스트 수정이 텍스트 난이도와 응집성에 미치는 영향,” 영어어문교육, 제19권, 제2호, pp.335-355, 2013.
- [18] 안수진, “한국인 대학생과 영어 원어민 대학생의 영어 설명문, 논설문의 코메트릭스 분석을 통한 비교,” 새한영어영문학, 제60권, 제3호, pp.177-205, 2018.
- [19] H. J. Kim and J. Lee, “Investigating linguistic differences in adult EFL learners’ writing test performance,” Journal of British and America Studies, No.35, pp.205-234, 2015.
- [20] J. Lee, “A Coh-Metrix analysis of lexical, syntactic and discourse aspects in the newspaper articles of Korean and British

university students,” Modern English Education, Vol.19, No.4, pp.17-26, 2018.

[21] D. S. McNamara, A. C. Graesser, P. M. McCarthy, and Z. Cai, Automated evaluation of text and discourse with Coh-Metrix, Cambridge University Press, 2014.

[22] 이슬기, “토픽 모델링에 따른 고등학교 논설문의 응결성과 응집성의 상관분석,” 국어교육학연구, 제52권, 제3호, pp.136-173, 2017.

[23] S. A. Crossley and D. S. McNamara, “Predicting second language writing proficiency: The role of cohesion, readability, and lexical difficulty,” Journal of Research in Reading, No.35, pp.115-135, 2012.

[24] 이동욱, 백서현, 박민지, 박진희, 정혜옥, 이지형, “LSA를 이용한 문장 상호 추천과 문장 성향 분석을 통한 문서 요약,” 한국지능시스템학회논문지, 제22권, 제5호, pp.656-662, 2012.

[25] Y. H. Choi, H. Ahn, and J. Lee, “Anaphora resolution strategies in L2 reading,” Korean Journal of English Language and Linguistics, Vol.18, No.4, pp.395-422, 2018.

[26] S. Haviland and H. Clark, “What’s new? Acquiring new information as a process in comprehension,” Journal of Verbal Learning and Verbal Behavior, Vol.13, No.5, pp.512-521, 1974.

[27] L. Manelis and F. Yekovich, “Repetitions of propositional argument in sentence,” Journal of Verbal Learning and Verbal Behavior, Vol.15, No.3, pp.301-312, 1976.

[28] G. Jones, “Lexical cohesion as a predictor of item difficulty,” Paper presented European Association for Language Testing and Assessment, University of Turku, 4-7 June, 2009.

[29] I. Kirsch, The international adult literacy survey: Understanding what was measured (RR-01-25), Educational Testing Service, 2001.

[30] T. K. Landauer, LSA as a theory of meaning. n T. Landauer, D. McNamara, S. Dennis and W. Kintsch (Eds.), Handbook of latent semantic analysis (pp.3-34), Erlbaum, 2007.

[31] 이경남, “잠재적 의미 분석(LSA)을 활용한 독해 과정에서 추론 양상 분석,” 독서연구, 제42호, pp.107-131, 2017.

[32] I. Aldabe and M. Maritxalar, “Automatic distractor generation for domain specific texts,” NLP 2010: Advances in Natural Language Processing, pp.27-38, 2010.

[33] N. LaVoie, J. Parker, P. J. Legree, S. Ardison, and R. N. Kilcullen, “Using latent semantic analysis to score short answer constructed responses: Automated scoring of the consequences test,” Educational and Psychological Measurement, Vol.80, No.2, pp.399-414, 2019.

저 자 소 개

황 이 수 (Leesu Hwang)

정회원



- 2013년 2월 : 전북대학교 영어영문학과(문학사)
- 2019년 2월 : 전주대학교 영어교육과(교육학석사)

〈관심분야〉 : 코퍼스 언어학, 영어 교수-학습 방법

이 제 영 (Je-Young Lee)

정회원



- 2013년 2월 : 한국교원대학교 영어교육학과(교육학박사)
- 2014년 3월 ~ 2017년 8월 : 세한대학교 영어교육과 교수
- 2017년 8월 ~ 현재 : 전주대학교 영어교육과 교수

〈관심분야〉 : 학문목적영어, 코퍼스언어학, 연구종합 등