

<http://dx.doi.org/10.17703/JCCT.2021.7.3.433>

JCCT 2021-8-50

TAALES 프로그램을 활용하여 한국 대학생이 작성한 에세이에 나타난 어휘의 정교화 특성 비교

The Relationship between Lexical Sophistication Features and English Proficiency for Korean College Students using TAALES Program

이영주*

Young-Ju Lee*

요약 본 연구는 자동화된 어휘 분석 프로그램인 TAALES를 활용하여 영어능력 수준과 어휘의 정교화 지표간의 관계를 살펴본다. 본 연구에서는 ICNALE 코퍼스에 포함된 한국인 대학생이 작성한 에세이 600개를 분석하였다. 본 연구에서는 영어 수준별로 어휘의 정교화 특성이 통계적으로 유의미하게 다르게 나타나는가를 살펴보기 위해 다변량 분산분석이 수행되었다. 다변량 분산분석 결과 어휘의 정교화 지표가 통계적으로 유의미한 것으로 나타났으며, 이는 상. 중. 하의 세 개의 집단에 따라 어휘의 정교화 지표의 평균백터에 유의미한 차이가 있고 상. 중. 하의 세 집단은 어휘의 정교화 지표의 사용이 유의미하게 다르다는 것을 보여준다. 세 집단 간에 가장 큰 차이를 보인 어휘의 정교화 지표는 내용어의 빈도, 내용어의 친숙도, 기능어의 어휘 결정 반응 평균 시간, 상위어 동사, 기능어의 워드 네이밍 반응시간, 내용어의 어휘 습득 연령으로 나타났다.

주요어 : 어휘의 정교화, TAALES 프로그램, 쓰기 지도

Abstract This study investigates the relationship between lexical sophistication features and English proficiency for Korean college students. Essays from the ICNALE(International Corpus Network of Asian Learners of English) corpus were analyzed, using TAALES program. In order to examine whether or not there are statistically significant differences in lexical sophistication features across three groups, MANOVA was conducted. Results showed that the lexical sophistication features were significantly affected by English proficiency level. Essays written by Korean students with different English proficiency levels can be differentiated in terms of various lexical sophistication features including content words frequency, content words familiarity, lexical decision mean reaction time function words, hypernymy verbs, word naming response time function words, age of acquisition content words.

Key words : Lexical sophistication, TAALES program, Writing pedagogy

*정회원, 한밭대학교 영어영문학과 교수 (주저자)
접수일: 2021년 7월 31일, 수정완료일: 2021년 8월 7일
게재확정일: 2021년 8월 9일

Received: July 31, 2021 / Revised: August 7, 2021

Accepted: August 9, 2021

*Corresponding Author: yjulee@hanbat.ac.kr

Dept. of English Language and Literature, Hanbat National University, Korea

I. 서 론

말하기나 쓰기에서의 어휘 오류는 총체적인 오류(global errors)에 이를 수 있고 의사소통의 흐름을 방해할 수 있기 때문에, 어휘 지식은 의사소통 능력의 중요한 요소가 된다 [1]. 예를 들면 “The grammar is the basement of every language.”라는 문장에서 basement라는 어휘의 총체적 오류 때문에 전체 문장의 의미를 파악할 수 없게 된다 [2].

어휘 지식은 텍스트 응집력이나 통사구문의 복잡성보다도 에세이 점수를 유의미하게 예측하는 변인이 되며 [3], 어휘의 다양성과 에세이 점수 간에는 긍정적인 상관관계가 있는 것으로 나타났다 [4]. 어휘 지식이 쓰기 능력의 수행에 미치는 영향은 매우 크고 높은 어휘 지식은 고급 수준의 단어, 즉 낮은 빈도로 활용되는 정교화된 단어를 사용하는 것으로 간주된다 [5].

어휘의 정교화는 전통적으로 학습자의 말하기나 에세이 샘플에 포함되어 있는 고급 어휘(advanced words)의 상대적인 비율로 정의되어 왔으며 [6], 고급 어휘는 일반적으로 빈도수가 낮은 단어로 분류되었다 [7]. 고급 어휘 즉 정교화된 어휘는 빈도수 이외에 덜 구체적이고, 덜 친숙하며 다양하지 않은 한정된 문맥에서만 등장하고 어휘 결정 과업에서 느린 반응시간을 수반하며 학문적 상황에만 등장하는 단어로 구분되었다 [5]. 또한 두 개 또는 그 이상의 단어를 여러 개의 단어로 이루어진 단위(multiword units)로 결합하는 것을 아는 것이 어휘 지식의 핵심요소라는 이론에 착안하여, 정교화된 어휘는 여러 개의 단어로 이루어진 단어를 포함하는 것으로 확대되었다 [5].

본 연구에서는 자동화된 어휘 분석 프로그램인 TAALES (the Tool for the Automatic Analysis of Lexical Sophistication)를 활용하여 한국 대학생이 작성한 에세이에 나타난 어휘의 정교화와 영어 능력간의 관계를 탐색하고자 한다. 본 연구에서 살펴보고자 하는 구체적인 연구질문은 다음과 같다. 첫째, 어휘의 정교화 특성이 한국 대학생의 영어 수준별로 통계적으로 유의미하게 다르게 나타나는가? 둘째, 상·중·하의 세 집단 간에 차이가 크게 나타난 어휘의 정교화 지표는 무엇인가?

가장 최신의 TAALES 프로그램은 TAALES 2.2 이고 현재 TAALES 2.8은 개발 중에 있다. TAALES 프로그램에서 어휘의 정교화는 크게 5개의 요소로 구성되

고 어휘의 빈도, 어휘의 범위, 한 단어(unigram), 두 단어(bigram), 세 단어(trigram) 조합의 n-gram, 학문어휘 리스트(Academic Word List), 심리언어학적 특성이 포함된다. TAALES 프로그램과 Coh-Metrix 프로그램의 가장 큰 차이점은 Coh-Metrix 프로그램에서는 어휘의 범위, n-gram, 학문어휘 리스트와 같은 지표는 분석할 수 없다는 것이다.

II. 선행연구

어휘의 정교화를 살펴본 대부분의 국내외의 선행연구 [8, 9, 10, 11, 12, 13, 14, 15]는 Coh-Metrix 프로그램을 활용하였으며 TAALES 프로그램이 개발되면서 이를 활용한 연구가 해외 [5, 16, 17]에서만 수행되었다.

Crossley와 McNamara[13]는 원어민 대학생이 작성한 에세이와 스페인어가 모국어인 비원어민 학생이 작성한 에세이에서 어휘적 특성이 어떻게 다르게 나타나는가를 Coh-Metrix 프로그램을 활용하여 살펴보았다. 이를 위해 관별분석을 수행했으며, 분석결과 원어민이 작성한 에세이와 비원어민 학생이 작성한 에세이 간에는 어휘적 지식, 어휘의 정교화, 어휘적 응집력에 있어서 차이가 있는 것으로 나타났다. 연구결과를 구체적으로 살펴보면, 원어민과 비원어민의 두 집단으로 분류하는 유의미한 어휘적 특성은 상위어(hypermym), 동작 동사 (motion verbs), 단어 빈도, 다의어(polysemy), 어휘 습득 연령(age of acquisition) 이다. 즉, 원어민은 추상적인 단어, 다의어, 빈도수가 낮은 단어, 공간개념을 수반하는 동작 동사(예를 들면, advance, boom)를 많이 사용한 것으로 나타났다.

국내에서는 Coh-Metrix 프로그램을 활용하여 교과서의 학년 간 연계성을 분석한 연구 [8, 9, 10, 11, 12]가 많이 수행되었다. Jeon[11]은 Coh-Metrix를 활용하여 중학교 영어 교과서의 학년 간 연계성을 단어 빈도, 가독성 지표, 통사적 복잡성, 어휘의 다양성의 측면에서 분석하였다. 어휘와 관련된 연구결과를 살펴보면, 중학교 1학년에서 2, 3학년으로 올라갈수록 단어 수가 유의미하게 증가했으며, 이는 연계성이 잘 성립되어 있음을 보여준다. 단어 빈도의 경우 중학교 2학년과 3학년 간에는 통계적으로 유의미한 차이가 나타났는데, 중학교 3학년 교과서에는 2학년에 비해 덜 빈번하게 나오는 어휘가 더 많이 포함되어 있는 것으로 나타났다. 어휘 다

양성의 경우 전체 학년 간에 타입 토큰 비율 값이 유의미한 차이를 보였다. 타입 토큰 비율 값은 특정 텍스트가 다양한 어휘를 포함할 가능성을 나타내는데, 예상과는 달리 중학교 1학년의 타입 토큰 비율 값이 가장 크게 나타났다. 이는 중학교 1학년 교과서의 읽기 지문이 2, 3학년에 비해 학습해야 할 많은 단어가 제시되어 있음을 시사하며 어휘의 다양성 측면에서는 학년 간 연계가 잘 이루어지지 않았다는 것을 알 수 있다. 단어의 구체성, 이미지 형성 가능성, 어휘 습득 연령의 경우 전체 학년 간에 통계적으로 유의미한 차이가 있었으며, 학년이 올라갈수록 덜 구체적이고 이미지를 형성하기 어려운 정교화된 어휘와 언어습득 연령이 높은 단어가 교과서에 포함되어 있다는 것을 보여준다.

어휘의 정교화만을 분석하는 TAALES 프로그램이 개발되면서 해외에서 이를 활용하여 분석한 연구 [5, 16, 17]가 최근에 증가하고 있다. Kim et al.[5]은 어휘의 정교화 지표를 통합하여 12개의 요소(components)로 만들고 이를 독립변인으로 활용하여 쓰기 능숙도, 어휘 능숙도, 장기적인 어휘 발달을 얼마나 잘 예측할 수 있는지를 살펴보았다. 비슷한 특성을 갖는 어휘의 정교화 지표는 주성분분석(principal component analysis)을 통해 요인으로 통합되었으며, 새롭게 생성된 요소가 얼마나 타당한가를 알아보기 위해 실제 예제 자료를 분석하였다. 즉, 상호 연관성이 높은 어휘의 정교화 지표가 같은 요소로 통합되었으며, 예를 들면 상위어(hyponymy) 지표와 빈도(frequency) 지표가 통합되어 단어의 구체성(word specificity) 요소로 범주화되었다.

통합된 요인에는 두 단어와 세 단어 조합(bigram and trigram strength of directional association), 내용어 특성(content word properties), 두 단어 조합의 상호 정보(bigram mutual information), 두 단어와 세 단어 조합의 비율(bigram and trigram proportions), 단어 구체성(word specificity)이 있다. 이러한 5개의 요소는 쓰기 능숙도의 16.1%, 어휘 능숙도 변인의 31%를 차지하는 것으로 나타났다. 즉, 강한 연관성을 갖는 두 단어와 세 단어 조합을 많이 쓰고, 덜 친숙하고 덜 빈번히 나오며 늦은 나이에 학습이 일어나는 단어와 같은 고급 내용어의 특징을 포함한 예제는 높은 쓰기 능숙도와 어휘 능숙도를 갖는 것으로 예측되었다.

또한 두 단어와 세 단어 조합의 비율(bigram and

trigram proportions), 단어 습득 특성(word acquisition properties), 내용어 빈도(content word frequency), 두 단어 조합의 빈도와 범위(bigram frequency and range), 내용어 특성(content word properties), 기능어의 빈도와 범위(function word frequency and range)는 일 년 동안의 어휘 능력의 장기적인 발전을 유의미하게 예측할 수 있는 것으로 나타났다. Kim et al.[5]의 연구결과는 어휘의 정교화는 다차원적인 성질(multidimensional nature)을 가지고 있음을 시사한다.

선행연구 결과를 요약하면, 어휘의 정교화를 살펴본 대부분의 국내외의 선행연구는 Coh-Matrix 프로그램을 활용하였으며 국내에서는 교과서의 학년별 연계성을 분석한 연구가 대다수이고 에세이를 분석한 연구는 상대적으로 많지 않다. 어휘의 정교화만을 분석하는 TAALES 프로그램이 개발되면서 이를 활용한 연구가 해외에서만 수행되었으며 국내에서는 TAALES 프로그램을 활용하여 한국 학생들이 작성한 에세이를 분석한 연구가 거의 없다. 따라서 본 연구는 자동화된 어휘 분석 프로그램인 TAALES를 활용하여 한국 대학생이 작성한 에세이를 영어 수준별로 분석하여 어휘의 정교화 특성이 영어 수준별로 다르게 나타나는가를 고찰하고자 한다.

III. 연구방법

3.1 연구대상

본 연구에서 분석한 한국인 대학생 에세이는 ICNALE (International Corpus Network of Asian Learners of English) 코퍼스에서 추출했으며, 300명의 학생이 파트 타임 직업의 중요성과 레스토랑에서의 흡연 금지에 대한 주제에 대해 한 단락의 길이로 40분 동안 작문했다. 본 연구에 사용된 에세이는 파트 타임 주제 에세이 300개, 흡연 주제 에세이 300개로 총 600개가 분석되었다.

ICNALE 코퍼스는 학생들의 영어 능숙도를 유럽공동 표준기준인 CEFR(Common European Framework of Reference)에 의해 제시한다. 300명의 학생 중에서 A2 수준에는 75명, B1 수준에는 149명, B2 수준에는 76명이 속한다. A2 수준은 토익 성적의 545 미만에 해당하며, B1 수준은 550 이상 670 미만, B2 수준은 670 이상 800 미만에 속하기 때문에, 75명은 하 수준, 149명은 중 수준, 76명은 상 수준에 속한다.

3.2 어휘의 정교화 지표 추출

TAALES 프로그램이 생성한 138개의 어휘의 정교화 지표 중에서 정규 분포를 따르지 않는 지표를 삭제하였다. 또한 다중공선성(multicollinearity)을 확인하여 어휘의 정교화 지표 간의 상관계수가 0.7 이상의 값을 갖는 지표를 삭제하였다. 본 연구에서는 71개의 어휘의 정교화 지표가 통계분석에 투입되었다.

3.3 자료분석

본 연구에서는 SPSS 26.0의 프로그램을 활용하여 통계분석이 수행되었다. 영어 수준별로 어휘의 정교화 특성이 통계적으로 유의미하게 다르게 나타나는가를 살펴보기 위해 다변량 분산분석이 수행되었다.

IV. 연구결과

CEFR의 A2, B1, B2의 영어 수준별로 어휘의 정교화의 특성이 통계적으로 유의미하게 다르게 나타나는가를 알아보기 위해 다변량 분산분석이 수행되었고, 결과는 <표 1>과 같다. 다변량 분산분석은 종속변수인 어휘의 정교화 지표의 선형조합에 의한 평균벡터가 세 개의 집단 간에 차이가 있는지를 검증한다.

표 1. 다변량 분산분석 결과
Table 1. Results of MANOVA

	람다 값	F	자유도	유의도
영어 수준 (CEFR A2, B1, B2)	0.542	2.702	140	0.000

<표 1>에 제시된 것처럼 71개의 어휘의 정교화 지표가 통계적으로 유의미한 것으로 나타났으며(Wilks' lambda=0.542, F=2.702, $p<0.000$) 이는 상. 중. 하의 세 개의 집단에 따라 어휘의 정교화 지표의 평균벡터에 유의미한 차이가 있고 각각의 집단은 어휘의 정교화 지표의 사용이 유의미하게 다르다는 것을 보여준다.

통계적으로 유의미한 어휘의 정교화 지표 중에서 효과크기가 큰 순서로 정리하면 <표 2>와 같다. <표 2>에서 통계적 검증력(observed power)의 값이 0.8 이상이면 수용가능한 검증력이 있는 것으로 해석하며 [18], 0.99 이상의 통계적 검증력을 갖는 어휘의 정교화 지표가 제시되어 있다.

표 2. 통계적 검증력이 0.99 이상인 어휘의 정교화 지표
Table 2. Lexical sophistication features with observed power values over 0.99

어휘의 정교화 지표	df	유의도	F	효과	통계적
				크기	검증력
내용어의 빈도	2	0.000	40.915	0.121	1.000
내용어의 친숙도	2	0.000	35.806	0.107	1.000
기능어의 어휘 결정 반응 평균 시간	2	0.000	18.773	0.059	1.000
상위어 동사	2	0.000	17.236	0.055	1.000
기능어의 워드 네이밍 반응 시간	2	0.000	14.725	0.047	0.999
내용어의 어휘 습득 연령	2	0.000	14.473	0.046	0.999

통계적 검증력이 0.99 이상으로 나타난 어휘의 정교화 지표에는 6개가 있으며, 내용어의 빈도(content words frequency), 내용어의 친숙도(content words familiarity), 기능어의 어휘 결정 반응 평균 시간(lexical decision mean reaction time function words), 상위어 동사(hypernymy verbs), 기능어의 워드 네이밍 반응시간(word naming response time function words), 내용어의 어휘 습득 연령(age of acquisition content words)으로 나타났다.

통계적 검증력이 0.99 이상으로 나타난 어휘의 정교화 지표 6개의 기술통계 분석 결과는 <표 3>과 같다.

표 3. 6개 어휘의 정교화 지표의 기술통계
Table 3. Descriptive statistics of six lexical sophistication features

어휘의 정교화 지표	상 수준	중 수준	하 수준
내용어의 빈도	5.510	5.564	5.610
내용어의 친숙도	609.88	611.09	612.38
기능어의 어휘 결정 반응 평균 시간	0.979	0.977	0.976
상위어 동사	1.703	1.638	1.570
기능어의 워드 네이밍 반응 시간	123.953	122.459	121.309
내용어의 어휘 습득 연령	4.484	4.429	4.388

하 수준에서 상 수준으로 갈수록 내용어의 빈도, 내용어의 친숙도의 평균은 낮아졌으며, 이는 상 수준 학생들은 에세이에서 빈도가 낮고 친숙도가 낮은 고급 어휘를 많이 사용했다는 것을 보여준다. 기능어의 어휘 결정 반응 평균 시간과 기능어의 워드 네이밍 반응시간의 경우 하 수준에서 상 수준으로 갈수록 평균이 소폭 증가했으며, 이는 상 수준 학생들은 원어민이 실제 영어 단어인지 판단하는 과업과 워드 네이밍을 하는 어휘

과업에서 긴 시간이 걸린 정교화된 기능어를 에세이에서 많이 사용했음을 알 수 있다. 또한 상위어 동사와 내용어의 어휘 습득 연령 평균도 하 수준에서 상 수준으로 갈수록 높게 나타났으며, 이는 상 수준 학생들은 상위어 동사와 비교적 높은 연령대에서 습득되는 내용어와 같은 고급 어휘가 에세이에 많이 포함되어 있다는 것을 시사한다.

V. 결 론

어휘 지식이 쓰기 능력의 수행에 미치는 영향은 매우 크며 어휘 지식이 높으면 고급 수준의 단어, 즉 정교화된 단어를 사용하게 된다. 본 연구에서는 자동화된 어휘 분석 프로그램인 TAALES를 활용하여 한국 대학생이 작성한 에세이에 나타난 어휘의 정교화와 영어 쓰기 능력간의 관계를 살펴보았다. 본 연구에서 살펴본 두 개의 연구질문은 어휘의 정교화 특성이 한국 대학생의 영어 수준별로 통계적으로 유의미하게 다르게 나타나는가와 상·중·하의 세 집단 간에 차이가 크게 나타난 어휘의 정교화 지표는 무엇인가 이다. 본 연구에서는 SPSS 26.0의 프로그램을 활용하여 다변량 분산분석이 수행되었다.

연구분석 결과를 요약하면 다음과 같다. 첫째, 다변량 분산분석 결과 어휘의 정교화 지표가 통계적으로 유의미한 것으로 나타났으며, 이는 상·중·하의 세 개의 집단에 따라 어휘의 정교화 지표의 사용이 유의미하게 다르다는 것을 보여준다. 즉, 영어 수준이 높을수록 내용어 빈도가 낮고, 내용어의 친숙도가 낮으며 어휘 결정 반응 평균 시간과 워드 네이밍 반응시간이 오래 걸리는 기능어, 상위어 동사, 어휘습득 연령이 높은 단어를 사용한다는 것을 알 수 있다.

본 연구는 영어 수준별로 어휘의 정교화의 특성이 어떻게 다르게 나타나는가를 살펴보았고, 언어습득과 영어 교육에서 결과가 활용될 수 있다. 첫째, 본 연구의 분석 결과 상 수준 학생의 경우 빈번하게 나오는 어휘는 많이 사용하지 않고, 어휘 습득 연령이 높은 정교화된 어휘를 많이 사용한 것으로 나타났는데 하 수준 학생의 경우 어휘 습득 연령이 높고 덜 빈번히 나오는 단어를 지도할 필요가 있다. 즉, 교사는 하 수준 학생들의 경우 어휘의 정교화 특성이 상 수준 학생들과 어떻게 다르게 나타나는가에 대한 정보를 통해 어휘 교육에서

활용할 수 있다.

둘째, 상 수준 학생의 경우 구체적인 동사를 많이 사용한 것으로 나타났는데, 하 수준 학생의 경우 구체적인 동사를 익히고 연습할 기회를 많이 제공해야 한다. 듣기와 읽기를 통한 단어학습과 같은 의미에 중점을 둔 언어 입력(meaning-focused input)을 통해 동사의 형태와 의미를 학습할 수 있다. 또한 말하거나 쓰기를 통한 연습기회의 제공과 같은 의미에 중점을 둔 언어출력(meaning-focused output)을 통해 어휘 유창성을 개발시킬 수 있다 [19].

References

- [1] Ellis, R. Modified oral input and the acquisition of word meanings. *Applied Linguistics*, 16, pp. 409-435. 1995. DOI : 10.1093/applin/16.4.409
- [2] Brown, D. *Teaching by Principles: An interactive approach to language pedagogy* (2nd edition). Pearson Longman. 2010.
- [3] Jung, Y., Crossley, S. A., & McNamara, D. S. Linguistic features in MELAB writing task performances. *CaMLA Working Papers*. 2013.
- [4] Engber, C. A. The relationship of lexical proficiency to the quality of ESL compositions. *Journal of Second Language Writing*, 4, pp. 139-155. 1995.
- [5] Kim, M., Crossley, S. A., & Kyle, K. Lexical sophistication as a multidimensional phenomenon: Relation to second language lexical proficiency, development, and writing quality. *The Modern Language Journal*, 102(1), pp. 120-141. 2018. DOI : 10.1111/modl.12447
- [6] Nation, P. *Learning vocabulary in another language*. Cambridge, England: Cambridge University Press. 2013.
- [7] Laufer, B., & Nation, P. Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16, pp. 307-322. 1995. DOI : 10.1093/applin/16.3.307
- [8] Kim, Jeong-ryeol. Continuity problems of elementary and secondary English education: Sequence analysis of English textbooks. *Language Research*, 50(1), pp. 161-184. 2014.
- [9] Song, Hae Sung. Exploring the connectedness of listening tasks according to grades in high school English textbooks. *English Language & Literature Teaching*, 19(4), pp. 279-305. 2013.
- [10] Jeon, Moongee & Lim, Injae. A Corpus-based analysis of middle school English 1 textbooks

- with Coh-Metrix. *English Language Teaching*, 21(4), pp. 265-292. 2009.
- [11]Jeon, Moongee. An analysis of the continuity among middle school English textbooks with an automated language analysis program. *Modern English Education*, 16(1), pp. 195-218. 2015.
- [12]Jeong, Mi-Kyung & Kim, Nam Gook. An analysis of the linguistic features on the corpus of Korean EFL learners and native English speakers with Coh-Metrix. *Studies in Linguistics*, 33, pp. 373-395. 2014.
- [13]Crossley, S. A., & McNamara, D. S. Computational assessment of lexical differences in L1 and L2 writing. *Journal of Second Language Writing*, 18, pp. 119-135. 2009. DOI : 10.1016/j.jslw.2009.02.002
- [14]Crossley, S. A., & McNamara, D. S. Predicting second language writing proficiency: The role of cohesion, readability, and lexical difficulty. *Journal of Research in Reading*, 35, pp. 115-135. 2012. DOI : 10.1111/j.1467-9817.2010.01449.x
- [15]Guo, L., Crossley, S. A., & McNamara, D. S. Predicting human judgments of essay quality in both integrated and independent second language writing samples: A comparison study. *Assessing Writing*, 18, pp. 218-238. 2013. DOI : 10.1016/j.jslw.2013.05.002
- [16]Kyle, K., & Crossley, S. Automatically assessing lexical sophistication: Indices, tools, findings, and applications. *TESOL Quarterly*, 49(4), pp. 757-786. 2015. DOI : 10.1002/tesq.194
- [17]Kyle, K., & Crossley, S. The relationship between lexical sophistication and independent and source-based writing. *Journal of Second Language Writing*, 34, pp. 12-24. 2016. DOI : 10.1016/j.jslw.2016.10.003
- [18]Yang, Byug-Hwa. *Understanding Multivariate Data Analysis*. Seoul: Communication Books. 2006.
- [19]Coxhead, A. The academic word list 10 years on: Research and teaching implications, *TESOL Quarterly*, 45(2), pp. 355-362. 2011. DOI : 10.5054/tq.2011.254528