Research Article

Generating Lexical Estimates in Academic Writing in English Using Technology: An Exploratory Study of Adult Indian Learners

Lina Mukhopadhyay[1] & Sreelekshmi S. Pillai[2]

[1]Professor and Head, Dept. of Training & Development, The English and Foreign Languages University, Hyderabad

Email- linamukhopadhyay@efluniversity.ac.in

[2]Research Scholar, A14, Akka Mahadevi Hostel for Women, The English and Foreign Languages University, Hyderabad

Abstract

In this paper, we demonstrate how teachers can use a digital tool like Lexical Tutor to generate estimates or assessments of adult learners' productive lexical knowledge in academic writing in English as a Second Language (ESL). The digital tool-generated lexical estimates are complemented by teacher feedback to illustrate the inductive method of language learning. In the first step, lexical estimates are created from a few samples of academic writing of adult ESL learners using Lexical Tutor. In the second step, it is employed to identify frequently used collocations in learners' writing samples and then to create an illustrative bank of concordances based on available corpora. This sample of concordances can later be used to give feedback on usage rules of occurrences of lexical items within a variety of syntactic contexts. The feedback, it is hoped, would improve the content and coherence in academic writing by adding lexical richness and variety to the writing. Furthermore, the use of technology would help learners notice lexical features and improve the quality of writing in a self-regulated manner. This would also serve as an example of the inductive method of language learning.

Keywords: *academic writing, lexical estimates, Lexical Tutor, data-driven learning, vocabulary size, lexical density, lexical diversity*

Technology And L2 Writing Assessment

The National Education Policy (2020) endorses the utilization and integration of technology to enhance several aspects of education like learning, assessment, planning and administration. This endorsement is preceded by developments in educational technologies that have impacted every aspect of the language classroom, ranging from teaching and assessment to feedback to promote self-directed learning across the globe. Specifically designed language-based tools have also brought about an insurgence of corpora research and Natural Language Processing (NLP), the results of which have been used to advance ESL/EFL learning in the instructional context (Tribble, 1990; McCarthy & Carter, 1995). The ontogenesis of corpora and NLP research has resulted in the development of software and web tools, which have in turn helped in creating auto-generated lexical estimates of texts uploaded to these portals. This has opened opportunities for technology-assisted language assessment. In this paper, in a novel attempt we explore how auto-generated lexical estimates from a relevant web tool called Lexical Tutor (in short Lextutor) developed

by Cobb in 2004 can be accessed by language teachers and used to give feedback and scaffold academic writing skills of adult learners.

The focus of this paper will be the application of Lextutor to give feedback to academic writing in English using lexical concordances. We shall demonstrate how teachers can give feedback and direct learners' attention to improve their vocabulary use in academic writing using corpus-based concordance illustrations selected from web sources and generated using Lextutor (Cobb & Morris, 2004).

The current insurgence of corpora research has resulted in Data-Driven Learning (DDL) as large databases of English texts provide contextual material to understand how words and phrases occur in the sentential contexts and how they are used in real-life, social, and academic communication. DDL helps in promoting an inductive method of language learning because, through a relevant corpus of concordances, learners can get illustrations of the appropriate use of lexical items and deduce relevant morpho-syntactic rules and patterns from such web-based large linguistic corpora. Thereafter, they can engage in self-correction and attend consciously to the development of their ESL proficiency (McCarthy & Carter, 1995; Todd, 2001). This method supports the inductive approach to language learning which demands from learners a conscious effort to identify patterns in language samples by paying attention and then consciously deriving rules from these patterns (Shaffer, 1989). Concordance software helps in isolating recurrent patterns extracted from authentic language samples. Concordances "make the invisible visible" (Tribble, 1990). Research has shown that a statistically significant transfer of word knowledge to academic writing tasks can be achieved using concordance programs (Kaur & Hegelheimer, 2005). They provide samples of target language items to aid in consciousness-raising (Hadley, 2002) and provide feedback for written errors (Gaskell & Cobb, 2004).

Technology-based feedback on learner language

In human evaluation, lexical estimates are perception-based. So, teachers often assess written responses based on estimates they think are good and aligned with the requirements of specific academic tasks. However, this perceptual estimate is limiting in nature, as it is not possible or practical for a teacher to provide fine-grained in-depth analysis like text length, frequency and range of word types, families and tokens. They would also need to align learners' usage to standardized proficiency levels of the Common European Framework of References (CEFR 2001, 2020), which is not an easy task without being trained on how to apply the framework in teaching and assessment of different levels of language abilities. So technological innovations like Text Inspector (Bax, 2017) or Lextutor (Cobb, 2004) can be employed to get detailed lexical estimates and align them to standardized levels like that of CEFR levels. Teachers can then provide instances of learning through examples they generate in a data-driven mode, like identifying concordances of academic words to improve lexical usage.

The use of technology such as DDL to provide feedback during formative assessments of writing has received mixed reactions from teachers and researchers (Boulton, 2017). Usually, ESL classrooms that provide feedback using technological tools use it in combination with teacher feedback. A recent case study on Online Formative Assessments and Feedback Practices of ESL Teachers in three countries, it was

found that all three participants from each country used Google Forms to provide automated feedback. This was supplemented with recorded audio and videos as feedback from the teacher participants (Mahapatra, 2021). Though recent research shows some useful techniques in generating and using automated feedback, there is still a lack of consensus on the independent use of technology to provide feedback on written documents. It is because the usefulness of technological tools, the long-term impact on learners, and the impact of feedback provided using these tools on the development of learners' writing skills are yet to be empirically verified. This paper, in a novel attempt demonstrates how technological tools can be used to generate lexical estimates and used by teachers to give meaningful feedback on learners' lexical knowledge and help them improve the quality of their writing by attending to lexical richness and lexical substitutions and alongside this minimize phrasal repetitions. However, this is an exploratory study and the results would have to be later validated by further research of larger population and the impact of such feedback on long-term academic language use.

For teacher-mediated feedback to be efficient, three conditions need to be fulfilled: one, it is important that learners too are engaged in the feedback; two, it is provided in a non-threatening environment, and three, it is time efficient. Carless and Winstone, in 2020, first proposed this kind of a tripartite feedback framework. The three dimensions of feedback in this framework are -- *the design dimension* (the teacher designs the assessment, helps learners in making evaluative judgments, and uses technology for learners' feedback uptake), *the relational dimension* (a non-threatening relationship exists between the teacher and the student), and *the pragmatic dimension* (the degree of authenticity and practicality of the feedback in terms of time and human resources). Technology can be used at each stage. As stated above, the goal of this paper is to demonstrate a systematic use of technology to create estimates of lexis used in different types of academic writing required in higher education and complement it with teacher-mediated feedback to increase lexical variety and corresponding syntactic patterns in writing.

Academic Language and Writing Assessment

What makes some articles academic and some others not so academic - is it the lexis or the content? The answer would be that both the vocabulary used and the subject matter being discussed in the articles lend an academic nature to it. The lexis used in these articles, however, plays an important role in setting the tone of the author and the scope of the paper. Writing an academic article requires the author to use Cognitive Academic Language Proficiency (CALP) (Cummins, 2017) and the appropriate usage of words from the academic word list (AWL) (Coxhead, 2000). The degree of use of CALP and frequency of use of AWL words determines a person's academic success. How would one possibly determine whether the learner has acquired sufficient CALP? A person with advanced academic language proficiency would have an extensive vocabulary, as well as use academic words. To help estimate the academic range of lexis used, Averil Coxhead (2000) created the Academic Word List (AWL) consisting of 10 sub-lists, excluding the first 2000 most common words or the General Service List word list. In this paper, we will provide estimates of learner knowledge based on the percentage of tokens of AWL words they use in their writing samples.

Research has shed light on the positive correlation between vocabulary measures and lexical proficiency of learners' writing. The correlation serves as a useful indicator of their academic writing and language development (Maamuujav, 2021). Lexical features present in academic writing thus become paramount in determining the quality of writing. The situation calls for a discussion on what constitutes vocabulary knowledge, how it impacts writing proficiency and how can vary across different task types. Let us now look at a few vocabulary measures commonly used in research that will also be used in this paper to build lexical estimates through Lextutor.

Vocabulary Measures to Build Lexical Estimates

Production-based vocabulary knowledge can be measured with respect to its size, lexical diversity, and lexical density.

1. Vocabulary size: It is defined as the number of words a person knows, either passively or actively. It is measured in relation to the frequency of standard word lists. Two major lists frequently used by researchers are AWL and the British National Corpus (BNC). The frequency of words used from such lists helps in assessing the vocabulary knowledge of ESL/EFL learners (Cobb, 2004); the presence of less frequent BNC words (K3 levels and above) indicates higher levels of lexical sophistication.

2. However, the problem with estimating vocabulary size based on production is that only 16 per cent of the receptive vocabulary is known productively at the 5,000-frequency level and 35 per cent at the 2,000 level (Laufer, 2005 as cited in Schmitt, 2014). Moreover, studies suggest that only one-half to three-quarters of receptive vocabulary is known productively (Fan, 2000). Therefore, written samples do not give a comprehensive view of the size of a participant's entire vocabulary; rather, it is only an estimation of his/her productive vocabulary. Nevertheless, estimating production vocabulary is crucial to giving feedback for the development of content and coherence in writing.

3. Lexical diversity: It is the variety of vocabulary used in a text, measured as the ratio of the total number of new words over the total number of words in a text or the Type-Token Ratio (TTR). However, this measure is vulnerable to the text length because when tokens are less in number the TTR can be high, and conversely, in longer texts the types may be repeated and TTR may be low. The TTR value can be reliable if the range of total words used is comparable across texts written by a group of learners and is of a minimum text length of 150 words and above (Cummins, 2017).

4. Lexical density: It is the ratio of the number of lexical words (content, noun, adverb, and adjective) over the total number of words (Ure, 1971). This measure indicates linguistic complexity because it shows how informationally dense a text is. It is, therefore, assumed that the quality of writing is directly proportional to its lexical density.

The present paper uses Lextutor as a digital tool to help teachers estimate lexical knowledge and design data-driven feedback on lexis use in academic writing samples of adult ESL learners based on the following two research questions:

RQ 1. How can Lextutor be used to buildl exical estimates from learners' academic writing samples across different text types?

RQ 2. How can lexical estimates be used to prepare data driven feedback on lexis use in academic writing?

The Study

Four ESL learners (female 3; male 1) within the age group of 25 to 30 years enrolled in the first year of a doctoral program in English Language Education (ELE) at a university in Hyderabad, India, participated in the study. They had 15 years of exposure to English as a medium of instruction during the time of data collection, and they participated in the study voluntarily.

The participants had completed a course in academic writing as part of their doctoral program. They were requested to submit four types of academic writing from the submissions they had made during the course:

I. a research proposal based on their doctoral study;

II. a report of a small-scale study;

III. a critical review of concepts; and

IV. an argumentative paper to establish a claim.

All the texts dealt with concepts and research claims in English Language Education (ELE). Note that each task type has a slightly different scope in terms of its communicative purpose, though all of them fall under the broader category of analytical and research writing. So, they are quite homogenous in terms of their structural presentation and include language of argumentation. Earlier research that studied the complex relationship between task types and their influence on writing performance, has shown that differences in task types impact aspects like fluency, accuracy, and complexity (Pourdhana et al., 2013). Similarly, in ESL research, where reading for summary recall is examined through different task types, the text types make different demands on working memory, attention, content knowledge, and linguistic competence (Patil, 2022). This leads us to examine whether task types influence lexical estimates in academic writing to better understand this link and use it to fine-tune feedback.

Note that for the purposes of the study, only four learners' writing samples were considered. So, in all, we considered sixteen writing samples (four from each participant) because we intended to provide in-depth analysis and feedback on lexis use across the four text types and see the extent to which this feedback made sense to the participants and whether they were willing to revise their writing samples based on the feedback they received. So, for purposes of this small-scale exploratory study, the sample selection was considered adequate and sufficient.

Method of data analysis:

*To analyze the lexical quality of the texts, Lextutor with the program Vocabulary Profiler (VP) was used (*https://www.lextutor.ca/vp/eng*/). It identified -*

- macro measures of fluency (text length), lexical diversity (TTR), and lexical density (ratio of content words over text length); and
- micro measure of vocabulary size based on percentage frequency of BNC from K1, K2 (most frequently used 2000 words) versus K3 and above levels and words from the AWL list.

This software was used because it is accessible and easy to use. Furthermore, it is a useful tool in understanding lexis use and the academic performance of ESL learners (Cobb & Morris, 2004).

Another program called VP Concordances (https://www.lextutor.ca/conc/sent/) was used to identify frequently used collocations in the learner texts. This was used by the teacher-researcher to design a feedback plan and raise awareness of the role of collocations in building content and coherence.

Findings & Discussion

In this section, we present the findings on the lexical analysis of the texts and explain the pedagogical implications of the findings.

To answer the first research question, we present the descriptives of lexical knowledge across four text types in Table 1:

Table-1: Descriptive of lexical estimates of four types of academic writing

Estimates	Research Proposal		Small Scale Study		Critical Review		Argumentative Paper	
	M	SD	M	SD	M	SD	M	SD
Text length (fluency)	1172.25	696.75	1149	446.69	1102	414.29	1478.75	621.11
TTR	0.38	0.09	0.35	0.04	0.38	0.03	0.37	0.07
Lexical Density	0.60	0.02	0.61	0.02	0.60	0.01	0.60	0.04
AWL %	9.93	6.66	10.58	1.52	13.2	1.90	12.13	2.06
BNC K1 %	69.1	2.20	68.30	4.12	68.2	3.59	68.38	5.43
BNC K2 %	12.28	0.69	15.27	2.42	11.4	2.78	13.28	2.56
BNC K3 & above %	14.23	6.42	8.63	5.73	19.63	6.41	12.78	1.36

The average range of text length from each task is within 1100 to 1500 words, indicating that the participants have been able to compose analytical texts with a fair amount of fluency. We are able to use TTR as a lexical diversity measure in a reliable manner because all the texts are comparable and have 1000-plus token lengths. It is observed that the TTR of the texts is comparable, ranging from 0.35 to 0.38. Further, we found that the average lexical density of each task type is 0.60 to 0.61, which implies that content word use is at 60 per cent of the text length across all four text types.

At a micro level of analysis of vocabulary size, we considered the use of low-frequency words (from K3 level and above) from BNC and AWL words. Though the use of K1 and K2 level words (or 2000 most frequently occurring words) comprise 82 per cent of the content, the learners have also been able to use low-frequency words from K3 and above at 16 per cent to 20 per cent. It indicates that their lexical knowledge is well spread across different word frequencies, and they can be placed at C1 level of proficiency (CEFR, 2020, p. 110). Furthermore, academic words range from 10 per cent to 13 per cent, which is a rather good performance and supports their C1 level of performance. The reason for the high use of AWL words could be because all the tasks demand that research work

is reported in a formal standardized manner, which these participants have been able to achieve.

Impact of text types on lexical knowledge of participants

Note that as the text length and TTR of all four participants do not show much variation across tasks, this indicates that task type does not seem to have any distinct impact on fluency, lexical diversity, and density in the writing knowledge of the participants. This contradicts the earlier research findings on the relationship between task types and fluency from Pourdhana et.al, 2013. A reason for this could be the high degree of relatedness of the text types used – they all belong to research writing, and analytical and argumentation skills had to be used to compose these texts. Another reason for getting stable lexical estimates in the group could be because all four participants are advanced users of the language and have quite high academic writing skills required to complete the four tasks. Furthermore, their knowledge of topics chosen to respond to each task type may have had a strong impact on the content and homogeneity in the range of lexis used. However, this might not be true if larger and more heterogeneous groups are taken into consideration with more diverse text types, as suggested by Zhang (2022).

Feedback on writing using Lextutor inputs

To answer the second research question, a feedback plan was prepared as part of the formative assessment of writing. Previous research has reported that learners believe that editing would be more efficient if teachers indicated 'where' the editing should take place (McGarrell, 2015). Therefore, preparing a feedback plan was deemed to be effective for the participants. First, an academic word list generated by using the Phrase Profiler feature in Lextutor was accessed (https://www.lextutor.ca/vp/collocs/). Then a few frequently used collocations of some of those academic words were identified from each text type of the four participants. A few representative examples are presented in Tables 2a-d to illustrate the choice of collocation on which feedback was prepared. For each lexical item (in bold), the preceding and following syntactic items are underlined to show the context of the use of the collocation or its concordance. In the column on 'comments, ' the usage of each lexical item with variations in syntactic form(s) and meaning are listed.

Table-2: Frequent collocations in Research Proposals

Frequent Collocations	Context of use	Comments
collect data **(V+ N.uncount)** **BNC levels** **collect : K1** **data: K3**	the parental questionnaire will collect data in four main areas [PrP_F_P2019_RP] The study uses information from various statistics, feedbacks and questionnaires to collect data. [ChS_F_P2019_RP]	NP, +LOC (PP) only NP

Table-3: Frequent collocations in Small-Scale Study

Frequent Collocations	Context of use	Comments
follow instructions (V+ N.pl) BNC levels follow: K1 instructions: K2	Most of them can follow instructions in at least three languages. [ChS_F_P2019_SSS] - Reading research in the present times is interactive and goal driven, for ex – when you follow the instructions to install and start an app. [DeS_F_P2019_SSS]	generic use specific use

Table-4: Frequent collocations in Critical Reviews

Frequent Collocations	Context of use	Comments
socioeconomic status **(adj+N)** **BNC levels** **socio-economic: K11** **Status: K3**	Multimodal composition recognizes that the use of multiple modes in the classroom by the learner is always influenced by the learners' culture, history, and socioeconomic status. [MNN_M_P2019_CR] In a multilingual ESL context like India, English may not necessarily be present in the home literacy environment which is related to the socioeconomic status of the family. [PrP_F_P2019_CR]	generic use specific Use
visual image **(adj+N)** **BNC Levels** **visual: K3** **image: K2** **imagery: K2**	Kress and Van Leeuwen's (2001) research indicates digital technology enables second language learners to extract meaning from visual images, audio, and video without the constraints of language. [PrP_F_P2019_CR] Studies of mental representations have included the visual imagery of mental models of representation. [DeS_F_P2019_CR]	generic use specific use, morphological change

Table-5: Frequent collocations in Argumentative Paper

Frequent Collocations	Context of use	Comments
individual difference **(adj + N)** **BNC level** **individual: K2** **difference: K1**	In a multi-lingual, multi-cultural context the diversity of our social experiences produces immense individual differences in terms of our shared experiences and exposure to different practices. [DeS_F_P2019_AP] The stories should not be too complex for the learners to work with on their own and one must consider the needs and abilities of the learners, and also the individual differences while choosing stories. [PrP_F_P2019_AP]	generic Use plural NP specific use plural NP

In Tables 3-5, we find that the learners have been able to use the corresponding lexical items in two contexts, one for *generic use* and another for *specific use* with the definite article 'the'. We concur that owing to a higher level of writing proficiency and adequate content knowledge, the participants were able to distinguish the use between the two syntactic contexts and use them appropriately in their writing. This is an example of learners' lexical depth or various uses of the same root. Conceptually lexical depth is close to lexical density as it shows how dense a text is.

For preparing a feedback plan, this selection was a necessary step as the teacher would be able to draw learner's attention to what they have been able to achieve in their writing. It would serve as a positive indicator of learner performance to show them instances of lexical depth (as seen in Tables 2b-d) and motivate them to prepare well in future. Furthermore, this kind of selection could be taught to the learners to make them notice such information independently and follow a similar style of writing in future. Thus, this is a good example of data-driven learning as it makes the learners self-reliant.

Lastly, incorrect use of lexical items with concordances was also identified. It is through illustrations that feedback can be planned as an inductive approach to language learning. In Table 6 we present an example of incorrect use of the lexical item 'previous research':

Table-6

previous research **(adj+N.uncount)**: BNC level: previous: K2 research: K2	The hypothesis underpinning this study was based on previous researches (*) about the influence of CLIL teaching technique in acquiring language and content comprehensively. [ChS_F_P2019_RP]
explanation	'research' as a noun is uncountable. 'researches' can be used only as a verb with first person singular use: He researches the source for his paper. The teacher regularly researches for the lesson. https://grammarhow.com/researches/

As a second step to provide feedback, we can try to bring learners' attention to the appropriate use of the lexical item in an inductive manner. To this effect, a concordance list was prepared from an internet-based corpus with the use of the lexical item 'research' as a noun and as an adjective, as listed below:

The current research utilised strong research methods as it was observational. (N+V)

Research shows needle exchanges work against the spread of disease. (N+V)

Research chimps are trained to open their mouths for medical exams. (adjN+v)

MBF also supports health, conservation and marine research efforts. (adjN)

A concordance list for the term 'research' as a noun was also created using Lextutor (https://www.lextutor.ca/cgi-bin/conc/write/):

Table-7: Concordance list of the term 'research'

1.	the BBC's vision of our journalism is also based on	RESEARCH	about what our audience wants.
2.	such as VHAI are working to overcome these problems through	RESEARCH	and by strengthening channels of information.
3.	the development of Namibia's own cultural heritage through	RESEARCH	and comprehensive coverage of our people's artistic....
4.	Research and development give an indication of group	RESEARCH	and development activities, if any.

Through such illustrations, we hope learner attention would be drawn to appropriate use of lexical items through use of concordance corpus-based examples. From such examples, learners can extract the rules of appropriate usage and revise their text, as well as learn to use it appropriately in the future.

However, a word of caution for teachers: sometimes the software can generate the wrong output, as was found with the concordance list of 'research' as a noun with a plural form.

Table-8: Concordance list of wrong output with the term 'researches'

1 .	to reconsider.... Will you allow me to display your pre-war	RESEARCHES	and your drawings and your photographs? I was loo
2 .	diverse nature of the IT theses being covered by students'	RESEARCHES	meant that very small numbers of students were fr
3 .	seventeenth and eighteenth centuries, drawing largely on the	RESEARCHES	of a range of Soviet scholars, but particularly t

So, the syntactic rule that can be extracted from all the examples of concordances of the lexical item 'research' listed above is as follows:

- 'Research' as an adjective and noun can only be used in its singular form.
- if it is used as a verb, then the tense morphology will allow the use of present and past tense morphemes accordingly.
- so, the use of research as a third-person singular verb usage should not be confused with plural usage of the lexical item as a noun, which is not allowed in the language as research as a concept is uncountable.

From the incorrect concordance example above, we can draw teachers' attention to the importance of their discretion and intervention in the use of technology to provide DDL because the auto-generated output may result in listing examples of wrong usage. Thus, technology-driven feedback has to be motivated and complemented by teacher feedback.

Conclusion

The lexical estimates generated by Lextutor and presented above must be correctly interpreted by the writing

teacher to understand the level of writing knowledge of her learners, specifically their production vocabulary. This would help in evaluating the learners and placing them on a particular band of performance (CEFR, 2020) and checking the amount of heterogeneity in group performance. Furthermore, by gaining access to this kind of analysis from the Lextutor, teachers can encourage learners to increase their TTR instead of merely increasing text length (Cobb, 2004). Alongside this, the teacher can advise learners to use a higher range of academic words and phrases to increase the lexical quality or diversity in their writing. Furthermore, concordance analysis and illustrations can be designed to make learners pay attention to the use of lexical items in different syntactic contexts and improve content and coherence in writing. Future research can showcase the viability of using Lexical Tutor as a feedback and assessment tool in a long-term setting. The study provides a tangible illustration of the recommendations of NEP (2020) to use technology in the realm of teaching and assessment with a specific focus on lexical development. The paper also provides a sustainable and equitable model for assessment and self-assessment, which can be taken forward by future researchers.

References

Bax, S. (2017). Text Inspector. Available: https://textinspector.com/

Boulton, A. (2017). Data-driven learning and language pedagogy. In S. Thorne & S. May (eds.), *Language, education and technology: Encyclopedia of language and education*. New York: Springer. DOI 10.1007/978-3-319-02328-1_15-1

Carless, D., & Winstone, N.E. (2020). Teacher feedback literacy and its interplay with student feedback literacy. *Teaching in Higher Education*, 1-14. DOI:10.1080/13562517.2020.1782372

Cobb, T., & Morris, L.V. (2004). Vocabulary profiles as predictors of the academic performance of Teaching English as a Second Language trainees. *System, 32,* 75-87.

Cobb, T. (2004). The compleat lexical tutor for data-driven learning on the web. [Web-based suite for programs]. Available: http://lextutor.ca/

Coxhead, A. (2000) A New Academic Word List. *TESOL Quarterly, 34*, 213-238. https://doi.org/10.2307/3587951

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment.* Cambridge, U.K: Press Syndicate of the University of Cambridge. http://www.coe.int/lang-cefr

Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume.* Cambridge, U.K: Press Syndicate of the University of Cambridge. http://www.coe.int/lang-cefr

Cummins J. (2017) BICS and CALP: Empirical and Theoretical Status of the Distinction. In: Street B., May S. (eds) Literacies and Language Education. *Encyclopedia of Language and Education* (3rd ed.). Springer, Cham. https://doi.org/10.1007/978-3-319-02252-9_6

Fan, M. (2000). How big is the gap and how to narrow it? An investigation into the active and passive vocabulary knowledge of L2 learners. *RELC Journal, 31*, 105–119. https://doi.org/10.1177/003368820003100205

Gaskell, D., & Cobb, T. (2004). Can learners use concordance feedback for writing errors. *System, 32*, 301-319. https://doi.org/10.1016/j.system.2004.04.001

Government of India. (2020). *National Educational Policy, 2020*. https://www.education.gov.in/sites/ upload_files/mhrd/files/NEP_Final_English_0.pdf

Hadley, G. (2002). An introduction to data driven learning. *RELC Journal, 33*(2) https://doi.org/10.1177/003368820203300205

Kaur, & Hegelheimer. (2005). ESL students' use of concordance in the transfer of academic word knowledge: An exploratory study. *Computer Assisted Language Learning - COMPUT ASSIST LANG LEARN, 18* (4). 10.1080/09588220500280412.

Maamuujav, U. (2021). Examining lexical features and academic vocabulary use in adolescent L2 students' text-based analytical essays. *Assessing Writing, 49* https://doi.org/10.1016/j.asw.2021.100540

Mahapatra, S.K. (2021). Online Formative Assessment and Feedback Practices of ESL Teachers in India, Bangladesh and Nepal: A Multiple Case Study. *Asia-Pacific Edu Res, 30*, 519–530 https://doi.org/10.1007/s40299-021-00603-8

McCarthy, M., Carter, R., 1995. Spoken grammar: what is it and how can we teach it? *ELT Journal, 49*, 207-218 https://doi.org/10.1093/elt/49.3.207

McGarrell, H. M. (2015). Developing autonomy through awareness of textual features in academic texts. *OLBI Working Papers, 7*, 41-52 https://doi.org/10.18192/olbiwp.v7i0.1360

Patil, V. K. (2022). Whole text reading comprehension: An application of task-based language assessment. In N. P. Sudharshana & L. Mukhopadhyay (Eds.), Task-based language teaching and assessment, *contemporary reflections from across the world* (pp. 323–350). Springer. https://doi.org/10.1007/978981-16-4226-5_16.

Pourdhana, N. & Karimi B., Sayyed M. (2013). Task types in EFL context: Accuracy, fluency, and complexity in assessing writing performance. *International Journal of Social Science and Humanity.* 80-84. https://doi: 10.7763/IJSSH.2012.V2.73

Schmitt, N. (2014). Size and Depth of Vocabulary Knowledge: What the Research Shows. *Language Learning*, *64*(4), 913–951. https://doi.org/10.1111/lang.12077

Skehan, P. (1998). *A cognitive approach to language learning.* Oxford University Press.

Todd, R. (2001). Induction from self-selected concordances and self-correction. *System 29* 91-102. 10.1016/S0346-251X(00)00047-6.

Tribble, C. (1990). Concordancing in an EAP writing programme. *CAELL Journal, 1* (2), 10-15

Ure, J. (1971). Lexical density and register differentiation. *Contemporary Educational Psychology, 5*, 96-104. doi: 10.12691/education-1-2-3.

Zhang, X. (2022). The relationship between lexical use and L2 writing quality: A case of two genres. *International Journal of Applied Linguistics https://doi.org/10.1111/ijal.12420*