Technology Integration and SkELL: A Novelty in English Foreign Language Teaching and Learning

Juliana CYFEKU
“Fan S.Noli” University

Abstract: The development of technology and its web-applicative tools in reference a variety of fields of contemporary interest, has flourished even the horizons of education. Due to the necessity of the rapid growth of Technology, its integration was considered as a solution to the covid-19 pandemic worldwide situation that caught humanity unskilled on such frequent and easy-to-follow use of ‘learning by doing’ when dealing with knowledge. On purpose, the present study introduces the implementation of SkELL (Sketch Engine for Language Learning) in English Foreign Language (EFL) learning and teaching as a rich source of free online linguistic data, provided the use of authentic texts (hereby to corpus data). All of these constituents autonomously yield both sides: learners and teachers to learner-learner interaction, learner-teacher interaction. Throughout the exploitation of its various web-applicative tools such as: Word Sketch, Thesaurus, Wordlist, Concordance, and Visualization among a range of others made possible and available EFL learning, and assisted teaching during the hard educational situation. It consequently may serve as a pre-requisite for language practitioners, curriculum designers to facilitate overall EFL teaching and learning process.

Keywords: Technology integration, SkELL, EFL, Corpus linguistics, Web-applicative tools.

Introduction

Foreign language learning and teaching has undergone significant changes as the result of the today’s demands that all kinds of social, cultural, economic, political, technological developments are evolving during the past decades. Crucially it has become a necessity in today’s educational system majoring various general, vocational learning and teaching of human endeavor. Technology integration especially in the field of education has affected overall learning and has empowered positively students’ outcomes in terms of: raising the dependent variables’ efficacy especially in foreign language learning: motivation, engagement, skills, proficiency. Technology integration clearly explained from the website Edutopia (http://www.edutopia.org/), refers to the use of technology resources such as: PCs, mobile devices like smartphones and tablets, social media platforms and networks like Goggle classroom, Moodle, Microsoft Teams, software applications to implement in language learning like Wordsmith, Compleat Lex Tutor, Academic Word List, SKELL, the internet, … etc, and all of these serving the scope of enriching our daily classroom practices and scaffolding the so called “skills of the 21st century”. In addition, Friginal (2018) embraces the view that successful technology integration depends on the cases when technology: a) has become routine and transparent b) is accessible and readily available for the task at hand, and c) is supporting curricular goals and helping teachers/learners effectively reach their goals (Friginal, 2018, p.30). Yet what he puts emphasis more, is the fact that a well-integrated use of technology equipment and resources requires well-trained teachers to adopt all of these constituents to gain as the final outcome well-prepared students. In essence this view reports to be utopic in underlying that technology integration can cot lead to unified outcomes due to the varied tools, practices, relationships, curriculum goals and constructs and in dependency to teachers’ and learners’ involvement and willingness to embrace change in this ongoing process.

As technology continues to evolve it requires continual learning and shifts in paradigm from both teachers and learners. According to the Brookings Institute (1916) research, technology integration in education especially at higher education improves academic achievement. Yet the prerequisites that ought to be followed are: the real-time use, personalized instruction, and mastery-based progression. Grounded on these determining elements to
yield to progression in the final outcomes, our study builds on the research statement that technology once incorporated with deliberate skill from the language instructors it will allow the latter to customize language learning by creating varying levels of scaffolding support from the students involved in the process. Specifically, the present study constructs on the factual incorporation of Corpus Linguistics and its wide range of relevant applications to foreign language teaching and learning. Specifically, McEnery & Hardie (2012) confine this approach in SLA which focuses upon a set of procedures, or methods, for studying language. It has the potential to reorient our entire approach to the study of language by refining and defining a range of theories of language that we may construct for a better understanding and adequate use of the latter from attested language use and findings drawn from it. Corpus linguistics can be reasonably defined as dealing with some set of machine-readable texts deemed to give possible answers to a specific set of research questions. Yet it is not a consensually agreed set of methods and procedures for the exploration of foreign language. This underpins McEnery & Hardie’s (2012) statement that corpus linguistics is a heterogeneous field in which differences prevail due to the varying approaches that may be applied to the use of corpus data. Typical of the corpus analysis are the generalizations inferred from the large scale.

Methodology

The advancement of technology has made possible its extension in use and extent via internet, regardless space and time limits. The more it advances, the greater the users will add and considerable benefits would be gained in all spheres of human activities. Moreover, its influence is notably encountered in the field of education especially during the pandemic period which found educators and students unskilled and lacking the most essential means of technology: computer, devices and internet to fill the long gap created from the physical interruption learning process. In various developed countries technology integration was viewed as a counterproductive means of transmitting knowledge, service, and information prior the pandemic, during and long after it. It has reported to be a variable that became an integral part of the learning process and as such positively impacted student learning. Del Vecchio & Loughney (2006) define E-learning throughout technological devises as beneficial to education since it applied to all types of learners due to its flexibility features i.e. saves time, informs learners regardless space, assists them share information and activities, and fulfills commitments in cases of in-class physical absences, and above all it produces measurable, reliable and prompt results for each learner. It allows both parts: learners and teachers to record, view, analyze, correct and seek for improvements in the overall learning commitments so as to meet the demands that the today’s speed of developments require.

Yet, despite the benefits we all obtain, the limitation to technology integration underlies in the fact that learners need to have computer skills starting from basic to proceed to more advanced level. The latter is acquired to deal especially with word processing programs, internet browsers, web sites that provide contemporary tools to analyse language and/or information data. They need to gradually pass in a self-paced style from basic inquiry to more conceptual, analytic inquiry vs. learning outcomes. It is the teacher’s main objective to fit the triangle: learner - technology – learning to support what Fitriah’s (2018) statement that E-learning, is a denominator in more conceptual, analytic inquiry vs. learning outcomes. It is the teacher’s main objective to fit the triangle: learner - technology – learning to support what Fitriah’s (2018) statement that E-learning, is a denominator in more conceptual, analytic inquiry vs. learning outcomes.

SkELL (Sketch Engine for Language Learning)

Grounded on Corpus Linguistics, corpora, corpus data, the present study highlights the use of free online web-applicative tools that these huge bodies of electronic texts offer for teachers, learners and researchers to conduct all kinds of linguistic analysis of the authentic corpus-based data. Provided the fact that they are free and online it is worth using new approaches especially to English Foreign Language teaching and learning. This is made possible by implementing their direct use so as to contribute to the process the benefits that Friginal (2018) cites
from Leech 1997’s study outlining the use of computers, internet, and online tools in language learning which aligned with Computer-Assisted Language Learning (CALL) principles of:

1) automatic searching, sorting and scoring;
2) promoting a learner-centered approach;
3) open-ended supply of language data; and as such enabling the learners to be tailored.

To the time of being introduced the benefits of them were not so satisfactorily in classroom use provided that computers were large, the language data they stored was moderately accounting thousands of word usages, and internet access was limited to certain areas for restricted operational use. Today the rapid development of technology and its sophisticated devices has made possible the application of corpus-based tools and materials to which virtually all kinds of learners have global access. Moreover, hardware and software have become accessible online as it is the case of examining linguistic variation in the broader exploration of lexico-syntactic characteristics of spoken and written language through SkELL (Sketch Engine for Language Learning). Being a tool to show how words are used by real speakers of the language, it supports more than 95 languages aiming to assist linguists, translators, interpreters, lexicographers, dictionary companies and more currently is used by language teachers and students as well. Users can even build their own specialized language database (text corpora) and analyze the latter by means of built-in text analysis tools such as Word Sketch, Thesaurus, Wordlist, Concordance, and Visualization.

Implementing web-applicative tools in EFL teaching and learning: Word Sketch, Thesaurus, Wordlist, Concordance, and Visualization. Sketch Engine is the ultimate tool to explore how language works. Its algorithms analyze authentic texts of billions of words as corpora lines display to identify instantly what is more typical in a language and what is rare, unusual or recently and developmentally emerging in use. Pavel Rychlík is a computer scientist and researcher in natural language processing. Since his PhD on indexing text corpora, he has turned to efficient large-scale text processing. Pavel is the main software architect of Sketch Engine and the original author of many of its components, mainly the Manatee corpus indexing system.

Michael Rundell is a linguist and lexicographer. As a dictionary editor since 1980, he has designed and managed numerous dictionary projects. He is a leader in the field of pedagogical English dictionaries. He is currently Editor-in-Chief of macmillandictionary.com, having started a dictionary development programme at Macmillan in the late 1990s. He has been involved in the Sketch Engine development since its very beginning as a consultant and nowadays his main responsibility is running the Lexicom training courses.

Sketch engine is free and easy to use and the initial steps you need to follow are: 1) register so as to have free access of the search for 30 days 2) start by (Figure 1) logging in and switch to the new interface 3) select the English language among a range of others to proceed with the language exploration. If not properly competent in using it, click on the right side of the page on the quick start tutorial (Figure 2) to observe how to start using the program.

![Figure 1. SkELL log in interface](image-url)
Once you log in, Sketch Engine interface displays in the gadget that you use for language learning. A range of innovative tools (Figure 3) appear to assist all kinds of language learners in exploring corpora, authentic texts, language behavior output from native speakers of English product of various social communicative texts across different registers. It boosts the learners’ interest in exploring the site itself and its web-applicative tools. Moreover, it fosters autonomous learning and web-based inquiry of any encountering linguistic ambiguity or language hypothesis that goes beyond the disconnected lines of concordance text.

The selected tools from the SkEILL resulted as novelty in our EFL and ELT Master students’ classes. They provided a considerable number of linguistic data that demanded the students’ attention on what linguistic analysis to initially pursue in terms of vocabulary and grammar. The following tools, figures of illustration are a clear evidence of the actual and novel web-tasks that were generated by exploring and later on implementing this corpus search website in English language classrooms. For ease of application and understanding the targeted search word ‘opinion’ was applied in the search icon box with each respective tool so to apprehend the linguistic use, significance and insights this very word can gain in reference semantics, morphology, grammar and syntax as well.

**Word Sketch**

This tool furnishes language users with a variety of linguistic information (morphological, semantic, syntactic) viewing the word’s searched variance in terms of contextual collocations, frequency use across registers, and semantic use by having a concrete display of it in a larger and authentic use in concordance lines.
The specific and useful information that we get from this tool all in one try and in one interface is considerable provided the modifying use of the given word example “opinion” and the dozens of examples it gives for a clear understanding of it attributed in different parts of speech in contextual use. For instance, as modifiers of ‘opinion’ we notice the various collocations such as:

Adjective + opinion: the majority opinion, expert opinion, humble opinion, dissenting opinion … etc providing so connotative meanings and use of how and in what communicative occurrences ‘opinion’ can be.

Opinion + Noun: opinion polls, opinion piece, opinion formers, opinion editorials, opinion column.

Verb + opinion: express/ voice/ differ / form / reflect / issue / share opinion.

The semantic meaning, we specifically get from this type of collocation is that throughout the contextual examples we get informed for the formal use of the word “opinion” and its associate verbs in this register. If we replace the informal counterpart meaning of “opinion” with the informal “thought” we notice other verbs (few in numbers) to collocate with it such as: have / think/ clarify/ collect/ give a thought.

In the third and fourth column we may also note the use of the word “opinion” having various syntactic collocations as displayed in the Figure 4. as subject and as object to modify a specific verb in such different functions. For instance: reflect/concur opinion versus the opinion polls; opinions vary ... etc

**Thesaurus**

Thesaurus in Sketch Engine is an automatically generated list of synonyms or words belonging to the same category (semantic field). The list is produced on the exact context in which the words appear in the selected corpus. Only nouns, adjectives, verbs and adverbs are supported in most corpora. Provided the following Figures 5&6 of Thesaurus interface, the information you get in an instant is of great learning assistance. It gives learners the possibility of observing all the synonyms of the searched word without carrying any kind of dictionary or thesaurus to deal with such linguistic inquiry. Moreover, it is suitable to both types of language learners: Basic or Advanced. In the given example of the searched word “opinion” we are provided by a range of synonyms such as: belief, statement, comment, thought, view, advice, decision, perspective, argument, knowledge, etc. The score given for each synonym indicates the percentage of shared collocates. For a better comprehension of a word’s contextual implication you can use the local menu at the Change View Option to see the examples of real text use, in addition to the collocations, listed number of synonyms, clusters, counts of use in the discourse of the key word with other content words. The obtained metadata is instantly elicited in terms of frequency, similarity to indicate that the scores given for each synonym indicate the percentage of shared collocates.

As viewed, thesauri are a good tool in English language learning as it assists learners in expanding, enriching and reinforcing vocabulary. For the most it is easily and freely used on the web because no manual work and no reference book are involved. The list of synonyms generated from this simple search provides a sufficient number of authentic native occurrences found in any targeted corpus.
Wordlist

The Wordlist tool is used to generate frequency lists of all kinds: lists of words, lemmas, nouns, verbs, tags, words containing or not containing certain characters etc. Word lists typically enable the searcher to examine the frequency of a word or a phrase in different contexts as an important part of its description. Various word lists that are based to some degree on word frequency in a corpus exist especially in the English language teaching (ELT) context. Word lists are a good starting point for subsequent searches of individual items at concordance level and can be useful in the comparison of different corpora. Moreover, they can be generated to account for individual items or for recurrent sequences of two or more items. Lemmatized frequency lists (Figure 7.) all group together words from the same lemma. For example, McEnery & Hardie (2012) define the meaning of lemma as a group of word forms that are related by being inflectional forms of the same base or term throughout a simple example of the English words: destroy, destroys, destroying and destroyed as part of the verb lemma ‘destroy’; each of which will automatically be counted separately in annotated texts. It should be noted that other derive words of the stem word belonging to various parts of speech such as: destruction, destructive are both considered as separate lemmas. These are related to the stem lemma ‘destroy’ by derivational rather than inflectional processes. Lemmatization can be done manually using an alphabetical frequency list, or in an automated way.

Different forms of the same lemma tend to vary significantly in terms of their overall frequency, with one particular form tending to be more frequent than others. The wordlist works on the token level. The default settings will produce a list of words because non-words are excluded automatically. The wordlist can also be limited by frequency, by setting the minimum and maximum limit.
Frequencies can be obtained for a list of concrete words. Use the: *from this list: option on the advanced tab and input the items for which frequencies should be calculated from the selected corpus. Regular expressions can be used to define complex criteria for the words that should be included in the frequency list. For a more concrete and contextual understanding, let us try to find corpus authentic uses (Figure 8) of the word “opinion” by using the Wordlist tool and selecting another corpus for linguistic analysis such as The British National Corpus (BNC).

Concordance

Concordance may be considered as the most powerful tool as it consists of a variety of search options. Throughout its search option we can find and elicit words, phrases, documents, text types or corpus structures all of this linguistic data displayed in the form of a concordance (Figure 9). As Friginal (2018) defines concordances are reference books comprised in alphabetical listings of all significant content words such as: prepositions, articles, adverbial phrases. Besides they can add to the primary list from the source text, a secondary list of words that co-occur before or after the primary word in association to the text enabling searchers to understand the contextual meanings of each word as extracted from the genuine source of its occurrence. Concordances derived from digital text files of actual language in oral and written discourse can provide comparative and quantitative linguistic data useful in characterizing the shared meanings that refer to a specific group of people involved in communicative settings. Specifically, concordances can be utilized to identify the different usages and frequency of a content word, examine word collocations, and explore key words in context indexes to define the subtle nuances and the semantic meanings intended by the speakers or writers of the texts.
We may further use view options (Figure 10) for tags and other attributes. Here we can see the frequency of the word “opinion” in different websites. If the concordance is very large you can work with smaller numbers of randomly selected lines. A random sample with the same number of lines from the same concordance will always produce exactly the same concordance lines. This corpus-based data behavior is intentional so that different users (e.g. students) do not get confused when examining such huge amount of data. It is advisable to follow the same steps and consequently arrive at the same result.

Provided the various intentional commands we give on the search icon box, we generate a different sample for each attempt with different scores in dependency to frequency which is the base measure to examine whether the targeted word is frequent or infrequent in use.

We can order the lines by the command to yield to:

- the standard KWIC view that easily enables the context/s observation to the right and left;
- a rather advanced research by eliciting linguistic corpus-data for the token that comes before KWIC which is the left context or after it which is the right text.

As viewed from the (Figure 11) the concordance can be sorted, filtered, counted and processed further to obtain the desired result. Despite being the most powerful tool, the concordance used with large corpora may find so many results that it can tedious to analyze and interpret them.

The view options allow displaying additional information such as lemmas, tags and other attributes, text types (metadata) or corpus structures. Yet such careful decisions should be made so that introduce the new concepts and tasks in the language classroom starting from basic to more advanced tasks in English language classroom.
Visualization

The SkELL allows English foreign learners to provide word distributions obtained from the Visualization icon (Figure 12) to get a better understanding of the range of synonyms or similar words that the targeted word may have and all of them are identified automatically. Notably, as displayed from (Figure 13.) a list of the synonyms of the word “opinion” is orderly obtained in a raw coming one after another in terms of proximity of semantic meaning as well as frequency use.

The adequate explanation is based on the theory of distributional semantics which states that in a nutshell, words occurring in similar contexts are also similar in meaning. The visualization display as seen from the (Figure 13.) gives the list of words in reference the targeted word ‘opinion’ in terms of synonyms and common similar words. Precisely the semantic meaning of the presented words (such as: idea, view, decision, belief, statement…etc) both in lines and through graphical image presenting them in various font sizes with the
implication of the highest semantic meaning and use of the possible synonyms of the word ‘opinion’. To continue with the five least similar words in reference, and as consequently presented at the last lines such as: issue, review, evidence, research, rule, discussion and that visualized in small font size letters when compared to the large font visualization of the strongest synonyms in use of it.

Further visualizations (Figure.14) can be obtained if interested in eliciting the uses of each of the synonyms presented throughout corpus examples, with the use of the latter as different parts of speech and behaving in contextualized examples with various functions.

**Conclusion**

The consecutive and affect-like constituents: the development of technology, technology integration, corpora, corpus data and various issue-specific web-applicative tools whose connection and interaction in the researcher’s English teaching classes was due to the covid-19 pandemic worldwide situation especially in the field of education. The main impetus of the present study is to reveal that their implication in EFL classes can overcome and facilitate teaching and learning no matter how unpredictable the educational settings or conditions may be. The main focus of the present study concerned vocabulary instruction in a modern, real and virtual learning environment for the solid reason that the knowledge of a word goes beyond the knowledge of its dictionary definition as it embraces knowing the word’s spelling, parts of speech (POS), morphology, variant meanings (Semantics), collocations, and specific uses across registers and genres (Corino & Onesti, 2019).

Indeed corpora and the implementation of SkELL (Sketch Engine for Language Learning) assisted the present study’s ELT Master students to exploit and master different aspects in EFL language learning and teaching: lexical information, patterns of textualization (viewing the targeted word in larger contexts), with the intent on noticing the genre-structuring features of the words selected for linguistic investigation. All of these relevant tasks and activities were introduced and guided under considerable information to the foreign language learner regarding the appropriateness and acceptability of particular linguistic choices to foster autonomous learner-teacher interaction and learner-learner interaction in individual genres. As Leech (1997) argues, by learning to interact with corpora, “students find themselves learning a great deal about language, and how to study language. They learn about the kind of query questions that can be usefully put forward and answered by reference to corpus data” (Leech, 1997, p. 23). Notably, some pieces of information are not to be found either in paper or in e-dictionaries (tools that almost all learners have been using and still do so far), whereas more detailed information on lexico-grammatical features such as: syntactical markedness, annotation and nuances in meaning of similar or common synonyms is made possible through the use of corpus linguistics and the plethora of corpora consisting of a range of corpus-based tools to process linguistic data for text authentication and native-like assimilation of the English language.

Among the current, efficient, available ‘ready-to-use’ corpora with a teaching and learning purpose, SkELL is certainly a good source of information, a fine center of online language instructor that promotes Data Language Learning (DDL) on the web, a term coined for the first time by Johns (1991) in the field of Second Language Acquisition. Johns describes microcomputer-based approach as an incentive to foster the learners to “discover”
the foreign language under the provision of authentic texts from the language teacher so that lead the learners towards autonomous developing strategies vs discovery strategies through which they can “learn how to learn” (Johns, 1986, 1988). This kind of search uses a special algorithm to select occurrences from a large multi-billion samples of text, providing good KWIC (Key Word in Context) examples of the word or phrase useful for language learners especially of English, German, Italian, Czech, and Russian (Johns, 1991).

Moreover, the present study puts to practicality and highlights the importance in EFL learning and teaching the statement of Thomas task-based study on Sketch Engine (2015) that DDL is based on the principle of “cutting out the middle man” with the insinuation of learning from the genuine language rather than from mediated resources such as textbooks, grammars or dictionaries. It is exactly the case as in our study that language users have just to type the word they want to investigate in any of the selected tools: Word Sketch, Thesaurus, Wordlist, Concordance and Visualization among a range of others and the software returns a set of occurrences, the word sketch with POS (Parts of Speech) relationships to other words, and the synonyms or semantically related words. The frequency of a word or phrase in different contexts is an important part of its description. Various word lists that are based on word frequency, a measure of corpus data exist especially in ELT context. They can be generated to account for individual items or for recurrent sequences of two or more items. Yet it is to be stressed that the larger the corpora the better the results obtained for linguistic analysis and in compiling in-classroom corpus-based tasks. Corpora consisting of oral and written version of the stored and displayed texts make it possible to develop and integrate the necessary 21st skills in a multi-modal DDL practice, thus answering and strongly supporting the Cobb & Boulton (2015) question when they speculate on “multi-modal or multimedia corpora can bring to the table and their impact on speaking and listening skills.

All in all Corpus Linguistics being a new discipline in SLA acquisition is already known from the current EFL students but not from all language educators. On this point, it is worth recalling what Hutter et al. (2009) label teacher education as an ‘interface of theory and practice’, suggesting training future teachers, in-service EFL teachers and practitioners to work with corpora. Continuously being under the exert of technological developments, it is the time that these language educators deal with corpus texts and tools within an applied corpus-based/corpus-driven linguistics framework so to make way for them mediate and gain new insights on the course of exploration to English language teaching and learning practice. The implication of the above-mentioned tools of Sketch Engine search software resulted to the present ELT Master students in a consciousness-raising outcome, leading them to a long-term global assessment of English. The fact of using technological tools, devices and integrating technology in English foreign language education stimulated group work and crafted the collaborative problem-solving and understanding, affected student inclusion in corpus linguistic query and increased their motivation and involvement to gain ownership of the overall English foreign language learning.

Recommendations

The implementation of SkELL in EFL classes at university level reported to be a rich resource, suiting student’s level either Basic or Advanced when exploring and learning language on the web. The present study revealed the impact it exerted on both sides of the SLA process. It was targeted to show the richness of knowledge, additional linguistic explanation they gained while thinking critically to go beyond the hidden lines of corpora texts and gain insights in vocabulary transformations and context-like occurrences that words may suffer by implementing the right tools for autonomous inquiry. The tools it presented in real course endeavors were moderate in number when compared to the range that the search itself consists of. Further implications and insights to facilitate teaching and learning and make it globally unique among English language learners is to address other tools for classroom implications. It is of great importance in dealing not only with vocabulary matters but further exploiting and disseminating grammar by using structure strategies for self-discovery of corpus-data regularities when compared to the students’ irregularities in oral/and or written expression in English. The crucial focus of future case-specific study would be to investigate upon the inductive strategies the EFL learners employ to induce patterns or rules from the authentication of English in corpora lines.

Scientific Ethics Declaration

The author declares that she is solely responsible for the scientific, ethical, and legal aspects of the paper published in EPSTEM.
Acknowledgements or Notes

* This article was presented as a virtual presentation at the International Conference on Technology (www.icontechno.net) held in Antalya/Turkey on November 16-19, 2022.

* Brookings Institute (1916). Is a non-profit public policy organization based in Washington, DC. Their mission is to conduct in-depth research that leads to new ideas for solving problems facing society at the local, national and global level. The research agenda and recommendations of Brookings’s experts are rooted in open-minded inquiry and their 300+ scholars represent diverse points of view. Research topics cover foreign policy, economics, development, governance and metropolitan policy. Retrieved from: https://www.hoover.org/research/beyond-disruption-technologys-challenge-governance

References


Author Information

Juliana Çyfeku
University of Korcë, Albania
Address: Bulevardi, Korcë,, Albania
Contact E-mail: juliana.cyfeku@gmail.com

To cite this article: