

## Full Paper

# Incorporating Digital Corpus Analysis Technology into English for Academic Purposes (EAP) Lesson Design for the Sri Lankan Tertiary Context: A Self-experimentation by a 'Novice'

Zareena D. Hussain

Department of English Language Teaching, Faculty of Arts, University of Colombo, Colombo 00300, Sri Lanka

Email Correspondence: [zareenadh@gmail.com](mailto:zareenadh@gmail.com) (Z. D. Hussain)

Received: 9 December 2021; Revised: 6 January 2022; Accepted: 9 January 2022; Published: 15 January 2022

### Abstract

This paper reports on an observational case study conducted to investigate the possibilities available to language teachers, non-expert, or novice in corpus analysis, in integrating corpus analysis technology to design language learning activities. Despite the availability of corpus analysis technology and large amounts of studies in corpus analysis of texts, studies on corpus-based language teaching and learning, development of language teaching-learning material incorporating corpus analysis technology and techniques by teachers who are non-experts in the field of corpus analysis for classroom teaching has remained an exception. This paper records the personal experience of the researcher as a language practitioner using corpus analysis technology in designing teaching learning material to achieve the objectives of a language course for undergraduates of low English proficiency (LEP) studying in an English medium instruction (EMI) context. Although this exercise posed challenges for the language teacher who is a 'non-expert' to the field, the researcher documents its positive and promising outcomes as proof to support more extensive 'non-expert' teacher-driven, student-participatory corpus based English language learning methodologies.

**Keywords:** EAP/ELT material design, Corpus analysis technology, Self- experimentation

---

### Introduction

Today, designing of English language teaching and learning activities is intrinsic to the duties of language teachers. The teacher is no longer purely a transmitter of content from textbooks to a class of students. A teacher may also be faced with the need to design content and adapt available content to suit the learning needs of the students. In such situations the internet and computer-based technologies have been of great assistance to teachers in designing activities as well as being a rich resource of ideas and content.

One of the technology-based language analysis processes that are informing the field of language teaching and material design is corpus analysis. Regarding corpus analysis Römer (2011) observes that although this field of language study has developed, its impact on second language teaching has yet to be clearly felt. This observation reveals that corpus analysis which has the potential to contribute to second language teaching is yet to do so in a significant manner. 'Few teachers and learners know about the availability of useful resources and get their hands on corpus computers or concordances themselves' [1] Thus, one of the questions that arises is what the vast and fast developing field of technology-based corpus analysis has to offer to the language teacher cum material designer who does not possess an expert understanding in the field of corpus analysis.

The use of technology has become widespread in the recent past and especially so with the global Covid-19 pandemic calling for remote or distant computer-based teaching and learning. Teachers who may never have used technology and the internet have been compelled to get familiar with various online platforms such as Zoom and Microsoft Teams as well as learn to create and deliver lesson activities via these platforms as well as utilize Learning Management Systems (LMS) to facilitate student learning. More teachers are experimenting with online learning platforms such as Kahoot, Mentimeter and Padlet to mention a few, to make their technology-based classes more interactive and productive.

With teachers adapting to this technology-based learning so quickly in the pandemic conditions, the question arises whether teachers will be able to employ another technology resource which is relatively less known or used as part of a teacher's toolbox, that is technology-based corpus analysis techniques. With the climate ripe for introducing new technologies into teaching and learning, this paper explores the potential of adding technology-based corpus analysis techniques as a teaching-learning tool. It does this through an activity designing endeavour to design classroom teaching learning material. What is noteworthy in this endeavour is that all three stakeholders involved that is the target students, the target teachers and the researcher, are novices or non-experts to the field of corpus analysis. The researcher who is also an English teaching professional introduces the dual role aspect of a 'non-expert' teacher cum material designer in this endeavour. This research focused on the experience of the researcher in the role of a language practitioner cum activity designer as well as a non-expert in corpus linguistics or analysis, engaging in an effort to design corpus analysis techniques-based activities to be used by teachers and students who are also non-experts in corpus linguistics and analysis. Another noteworthy factor is that the target students are undergraduates with low English proficiency following a degree programme delivered in the English medium. Thus, this research examined the possibilities available to exploit corpus analysis techniques for language material designing within a context which is challenging, restrained and not conducive in more ways than one.

### **Corpus Linguistics and Analysis**

A corpus is an electronically stored collection of naturally occurring instances of texts, either spoken or written (corpora: plural). Corpus analysis involves processes that examine such texts in order to isolate and identify language features such as words, word groups and language patterns for linguistic purposes. Generally, corpora can be said to be compiled, accessed and analysed by linguists.

Corpus linguistics is a study and analysis of a 'principled collection of language' [2]. Thus, the collection would be either from a particular genre or time or have another defining commonality depending on the linguist's purpose and factors being explored. With the development of computer or digital technology and the facility to store corpora in electronic format and be analysed using analysis technology, it was found that corpus linguistics lends itself well to the study of language for language teaching and learning purposes. Lexicography or the compiling of dictionaries is one such application of corpus analysis. The Collins Corpus and Bank of English (each approx. 2.5 billion words), and the Oxford English Corpus and the Cambridge English Corpus (each approx. 2 billion words in size) [2] are a few examples of corpora that were prepared for this purpose. A noteworthy fact is that the techniques and digital technologies or tools used for linguistic analysis on corpora have been identified as directly exploitable to meet the needs of language teaching and learning. This research explores the possible applications to meet the language needs of undergraduates with low English proficiency (LEP) in English medium higher education. A main reason for exploring such applications is the inherent features of digital corpus analysis technology which

show potential to aid LEP undergraduates in EMI to quickly master the immediately needed language to function well in their academic setting [3].

### **Corpus Analysis and the Language Classroom**

In the area of language teaching, as Rómer (2011) highlights, there has been a degree of application of corpora and the results of such corpus analysis have led to syllabus decisions that is what to teach and when to teach. A more direct application of corpus analysis would be in which learners and teachers in the language classroom access corpora and apply the corpus analysis processes in the classroom for language teaching and learning purposes [1]. However as already mentioned, this practice of classroom application is a rarity although EFL classes have been described as ‘one of the most applicable fields for corpus database’ [4]. This situation could change with more initiative on the part of teachers to get to know these concepts and explore the possible applications.

As ‘...corpus can present statistically proven evidence of the language actually used’ [5] it is a good platform to collect and present data of the authentic language in an analyzable form. This has obvious relevance and advantages to the language learning context of undergraduates with low English proficiency (LEP) engaging in English medium academic studies. The technology used in corpus analysis contributes to understanding ‘the behavior of various lexical and grammatical features’ present in authentic language [5]. This aspect holds great potential to influence classroom teaching and lead to productive English for Academic Purposes (EAP) teaching pedagogies which is the focus of this research.

The general need for more initiation on the part of English language teachers, students and material/activity designers into incorporating and exploiting corpus analysis is a reality in the field of English language teaching in Sri Lanka as well. A challenging ground situation that may discourage implementing this could be that mostly all stakeholders involved, that is activity designers, students and teachers are absolute novices to the field of corpus analysis. Thus, this paper attempts to look at a way to bring in corpus analysis into the language classroom breaking the main ground reality or barrier of “noviceness” on the part of the stakeholders involved.

Application of corpus analysis in English language teaching would include the use of corpus technology, that is, the actual text selection as well as the software and related techniques to be used on electronically stored corpora (selected texts) to conduct analysis. For instance, it would involve students and teachers directly engaging and actively working with electronically stored texts, corpora and analysing them with available technology such as specialised software. This would also call for material design efforts to incorporate this technology and its related processes into meaningful language learning activities [6, 7]. Corpus analysis technology seems a highly promising tool in the field of foreign language learning as it ‘allows for compiling frequency lists, particular necessary specifications of textual features, ..., grammatical patterns’ [2] This research probes into these possibilities against the backdrop of the above-described challenge of noviceness on the part of all stakeholders involved.

### **Introducing Corpus Analysis Techniques and Technology to Undergraduates Who Struggle with English Medium Instruction**

Students in English medium instruction (EMI) constantly interact with academic English texts, both spoken and written. Thus, this research intended to explore the possibilities available for activity designers to

exploit selected techniques of corpus analysis to design learning activities that isolate and teach language features present in the very texts these students interact with. Thus, this research attempts to exploit corpus analysis techniques and technology to create activities that assist LEP undergraduates to use their academic texts for language learning and gain more access to their respective academic texts through an improved comprehension of the language.

Cobb (1997b) had employed corpus analysis as a vocabulary learning tool in a context similar to that of the current research that is 'a university in a developing country (Sultan Qaboos University, in Oman) that wished to use English as the medium of instruction but whose students were seriously underprepared for such a venture.' [3] One problem the students faced was a deficient vocabulary for academic reading. Here the goal was 'to use corpus and concordance as a way for these students to meet and learn a relatively large number of words, for use in reading comprehension, in a relatively short time.' [8, p. 485]. The rationale for using a corpus-based approach was the students' low English proficiency and the limited time available to prepare students for their English medium academic work ahead. Cobb's rationale was, firstly, the students' low English proficiency made meta-language use that is using English language to describe language, an unnecessary complexity to the LEP students. Secondly, the time restraint did not allow for enough encounters with words in contexts for natural learning to occur. In this context Cobb selected to use a corpus-based approach with language content which had immediate relevance to the students EMI context. 'The corpus was a digitization of all the ESL materials that the students were using to prepare them for forthcoming English-medium study.' [3].

This research too supports the comprehensive 'digitization' or conversion into a digital format of the language content of the students' English medium academic courses in order to readily apply digital corpus analysis technology. Such a digitized corpus could serve as a corpus of language teaching to use to meet the students' immediate linguistic needs in their academic courses. Application of corpus analysis technologies on this 'principled' corpus will allow for the manipulation of this specialized corpus to identify the language features in it and exploit them for language learning. The identified features can then be meaningful incorporated into the into language teaching-learning activities at the material/course design stage. For instance, in Cobb and Boulton's research (2015) the experimental subjects were taught vocabulary using the concordances lines (Figure 1) generated by corpus analysis to infer word meanings and confirmed by dictionary use. Although the findings of Cobb and Boulton's research were not generalizable, the advantages of using concordance-based activities for vocabulary expansion were observed in this particular tertiary context.

lysts, who focused on reported earnings and price earnings ratio  
our ability to generate EBITDA (earnings before interest, taxes,  
y off loans than on accounting earnings. In the past, these mea-  
ndication of true value than do earnings per share. These anal-  
on reported earnings and price earnings ratios. However, today  
mpany's negative reported earnings, the company's cl

**Figure 1.** Concordances (or concordance lines) - a list of language samples created by the corpus analysis function called 'keyword search'. (Software used 'AntConc'; Keyword 'earnings')

### **Advantages of Using Corpus Analysis Techniques and Teaching English for Academic Purposes (EAP)**

The corpus linguistics practice of selecting a corpus (a collection of texts) and analysing it for linguistic features seems to be in line with the objectives of EAP which is focused on the content of a particular discipline and centered on language features and functions of that discipline [8].

The trademark characteristics of corpus analysis technology such as generating concordance lines 'providing abundant example sentences and presenting keywords in context' [9], used by Cobbs and others, is garnering increasing attention in the field of English for Academic Purposes (EAP) teaching. Its other characteristics such as isolating words, word patterns, collocation information within an authentic linguistic context or environment can effectively draw the learners' attention to these linguistic features. Such characteristics are ideal for the university settings where undergraduates in the English medium are expected to comprehend as well as produce subject specific texts [2]. Thus, this technology can be exploited to teach these linguistic features in context which would be relevant to LEP students in EMI receiving EAP language training; to develop the necessary linguistic competency to 'understand language in context' as well as use language to achieve an academic purpose [2].

Further, using corpus-based technology, students can be trained to discover language features (vocabulary to grammar structures) in authentic texts relevant to their academic studies. In his "data-driven approach" or "Data-Driven Learning" (DDL) [9] Johns advocates that student 'should be guided to detect the underlying rules or patterns in language use and to draw conclusions from clues in the data [9]. The corpus-based technology and its related techniques such as the concordance lines are tailor made for "Data-Driven Learning" [2]. Further, these mostly computer and web-based techniques can become a productive self-learning tool to help students analyse and comprehend language in academic contexts.

One such well known computer-based tool is AntConc. AntConc, a corpus analysis software developed by Laurence Anthony is described as 'a lightweight, simple, and easy to use corpus analysis toolkit that has been shown to be extremely effective in a classroom context' [10]. Khairas (2019) reports on an experiment with 51 students from the Engineering Department, Politeknik Negeri Jakarta which employed AntConc as a language learning tool. It reports that 'Antconc software can be an alternative method to enhance students learning of English by using some technology on its software' (p.189) In this study, participants were first trained in the use of AntConc functions and their views of its value as a language learning tool were gathered through a questionnaire [11]. A considerable number of participants appreciated the value

of AntConc to learn vocabulary and 'the capability of AntConc Software to analyse the electronically saved authentic language data gathered from engineering e-books (the corpus) and for these analyses to be used as a media for learning vocabulary' leading to gaining the vocabulary of Engineering [11]. This paper examines the feasibility of incorporating the same technology and the same techniques into classroom activities for a Sri Lankan tertiary setting.

The sources of language in EMI higher education include lectures, textbooks, handouts, assignment tasks and examination questions. These pose a huge challenge to the LEP undergraduates contributing to poor comprehension. However, corpus linguistics holds potential to help teachers teach and students learn 'how language works in specific contexts' [12] especially in the students' respective disciplines of academic study. This research attempts to test all of the above advantages in this technology.

### **Introducing Corpus Analysis Technology to EAP Teachers and Students**

It has been observed that 'the lack of application of computerized corpus technology in the language classrooms brings about the problem of how we can successfully integrate computerized applications into language teaching.' [13]. In order to change this situation as well as any reluctance from the teachers there is a need to train or inform the teachers on the value of corpus analysis techniques and technology as useful resources and to initiate the practice of working with corpus, computers and concordances.' Mukherjee observes that 'the number of those who do so is relatively few, if not at all [13]. Mukherjee further notes 'the influence of computerized corpus-linguistic research on the actual practice of language teaching is still relatively limited' [13]. A similar observation is made by Seidlhofer [13], who states that 'the awareness level of teachers and students is very little towards the influence corpus linguistics has on the preparation of the language teaching materials and reference technology'. The widespread lack of active recognition and interest in the value of corpus analysis and its related techniques towards language learning is well captured by Braun who correctly observes that "corpora are still far from being part of overall teaching practice in the language classroom" [13].

The application of corpus analysis to English language teaching, especially for academic purposes can be considered a novel idea in the Sri Lankan university context and there may exist an unfamiliarity among English language teaching staff with regard to this field adding to the reluctance to venture into the field of incorporation of corpus analysis techniques into teaching learning. In fact, even the field of EAP in local universities may yet be at a 'premature infant stage'. [14].

Taking the above observations into consideration, this research chooses to focus on the possibilities of an English language teacher, non-expert in corpus analysis engaging, educating oneself on this technology and applying this knowledge in EAP material/activities design and in turn introducing this technology and techniques to other non-expert teachers and students. This research tracks the experience of the researcher, an English teacher by profession, a novice to corpus analysis, in the researcher's attempts to self-understand corpus analysis technology and techniques and design language teaching-learning activities incorporating this technology.

Corpus-based classroom materials is still not widespread and if there is an effort to apply corpus analysis techniques to classroom language teaching it will call for teachers adjusting them suitably to for classroom use. This will of course add an extra workload on the teachers. Further, before any successful implementation of corpus techniques-based classroom materials, teachers who will be involved in

administering such content should be ‘aware of corpus-based approaches and their successful implementations in their classrooms’ [13]. Özbay and Kayaoğlu, (2015) acknowledge that the teachers’ ‘knowledge of how they can integrate corpus technology into their teaching, what corpus-based methods to use and when they should use them are directly related to their perceptions of corpus.’ (p. 87) [13].

Thus, this study which promotes the use of corpus technology-based language teaching lessons among ‘novice’ teachers and students, aims to introduce the concepts of corpus analysis and make its incorporation as user friendly and clear as possible for both teachers and students. In this way it hopes to create a positive impression of the value of corpus analysis technology for English language teaching in higher education.

Özbay and Kayaoğlu, (2015), quote Seidlhofer who observes ‘the awareness level of the language teachers towards the huge potential of corpus to describe language and to prepare the language teaching materials is a significant factor that has to be considered seriously’ (p.87). This self-experimentation research explores the material preparation aspect highlighted by Seidlhofer by taking into consideration a possibility of even teachers’, novice in corpus analysis technology to realize and utilize the huge potential of corpus analysis for language teaching, that is through self-education and self-development of corpus analysis-based teaching material.

Figure 2 shows the schedule of a program to introduce Corpus analysis and techniques and technology to students in another study carried out by Özbay (2017) [15]. This study examined the potential of corpus-based instruction to improve students’ knowledge word collocations to improve reading. EFL students in tertiary were the subjects of this research. The findings of the study showed that corpus-based instruction on collocation had a positive effect on the awareness level and reading performances of the students. In order to integrate corpus-based instruction into the lessons, the students were introduced to established corpora such as Karadeniz Technical University Corpus of Learner English (KTUCLE), British National Corpus (BNC) and Corpus of Contemporary American English (COCA) and the corpus analysis software AntConc 3.0.1.

Duration	Systematization	Procedures
Week 1	Introduction	Corpus defined, Word lists Concordances, Frequency data, BNC, COCA.
Week 2	Exploitation	BNC search words and the use of collocation function
Week 3	Exploitation	BNC- target head words and their collocates, finding collocations with all matches to the query,
Week 4	Exploitation	COCA-target head words (make, do, take )and their collocates
Week 5	Exploitation	Dialogue completion task with the most probable collocation hits from BNC and COCA
Week 6	Exploitation	Dialogue completion task with the most probable collocation hits from BNC and
Week 7	Exploitation	Matching collocations based on the hits from BNC and COCA
Week 8	Transformation	Discussion of the benefits of using corpus concordances for collocation search and how we can use these findings while reading.

**Figure 2.** The corpus-based collocation instruction scheme on a weekly basis [15]

The above study chose to provide an initiation to the students before starting to use corpus analysis techniques and technology for language learning. If teachers too are novices to these concepts, it is clear that teachers too need an introduction before being able to apply these concepts in the class. The current

research, however, examines how corpus analysis techniques and technology can be seamlessly introduced to both teachers and students through the classroom activities themselves by employing simple and clear guidance made available commonly to both teachers and students, built into the students' module. This will circumvent the need for a separate training while moving directly into the lessons which are designed to teach language content while introducing the corpus concepts to both teachers and students.

The activities designed in this research also aim to empower non-expert teachers to become future activity designers confident in researching corpus analysis techniques and incorporating them into language teaching learning material design. Further, the activities designed intend to encourage teachers to educate themselves on this relatively new field and become confident in the use of this technology in language teaching. The activities designed also encourage students, all who are non-experts in this rather technical field, to become language analysts themselves using the corpus analysis technology to analyse academic English texts. As Boulton (2016) highlights, this will enable students '...rather than relying on experts to pre-digest the language' [16] and, '... learners could chew on the language data themselves.' [16] Thus it is hoped that knowledge of these techniques and technology will help the low English proficiency (LEP) students to handle the difficult language in their academic texts and improve comprehension as well as help teachers to help their LEP students to engage meaningfully with their academic texts.

The aim of this research reverberates well with the words of Johns in Boulton (2016) regarding Data Driven Learning (DDL). Johns describes Data Driven Learning as "the attempt to cut out the middleman as far as possible and to give the learner direct access to the data, the underlying assumption being that effective language learning is a form of linguistic research" [16]. Through this self-experimentation and the related activity design research, this research intends to introduce digital corpus analysis technology as a tool for language learning, teaching and material development. Thus, as Johns mentions, it hopes to encourage and empower all three stakeholders, namely students, teachers and material developers to become researchers and learners and directly access and unleash the benefits of corpus analysis technology in order to deal with LEP in EMI and avoid the delay of purely waiting upon experts or the 'middleman' to bring home the benefits.

### **Benefits of Self-Experimentation**

The researcher-as-subject or the self-experimentation approach, is a 'first-person methodology' which allows an in-depth investigation on the researcher leading to valuable insights and the creation of systematic knowledge [17]. According to Corti, Reddy, Choi and Gillespie (2015) participating as both observer and subject within one's own social-psychological experiment provides researchers the following advantages: (1) access to the subjective experience of social phenomena; (2) improved mental models of social phenomena, (3) potentially stimulating new research questions; and (4) an enhanced ability to be reflexive about the given experiment.

This type of experiential research captures answers to the question 'What is it like to be...?' In investigating a phenomenon in this reflective and subjective manner leads to an understanding of the phenomenon rooted in both description and experience. This research in particular wanted to explore the questions 'What is it like for a teacher cum material designer, a novice to corpus analysis, to integrate corpus analysis techniques into English language teaching activities?' 'What is it like for such an individual to design such activities for low English proficiency undergraduates in EMI?' and 'What is it like for such an individual to design such activities for the use of teachers and students who are novice to corpus analysis technology



and techniques?’

In the current research, the researcher had the ability to directly experience two factors, namely, being a non-expert language practitioner engaging in an effort to explore corpus analysis technology and techniques, and secondly forming a knowledge base and incorporating it into lessons for preliminary level students. The four challenges faced by the researcher were

1. the researcher being non-expert in the field of corpus linguistics and analysis
2. the students being non-experts in the field of corpus linguistics and analysis
3. the course delivering teachers being non-experts in the field of corpus linguistics and analysis
4. making the content as user friendly as possible for both teacher and students.

This was a subjective and reflective investigation of this specific phenomenon that of being a teacher cum material designer who attempts to integrate corpus analysis techniques into subject content-based activities designed for low English proficiency undergraduates in EMI. This investigation leads to a descriptive understanding of the process and the final outcome, i.e, a record of the design experience as well as the teaching learning material created.

Another advantage foreseen in self-reflective research is that it makes way for future experimentation by others. This research is directed at English language practitioners in higher education who are non-experts in corpus linguistics to encourage them to copy this experience and thus gain and contribute more insights into this practice of bringing corpus analysis into the local English language classes for undergraduate students. Interestingly, Roberts [18] highlights that ‘Self-experimentation lasting months or years seems to be a good way to generate plausible new ideas.’ Although the current study was conducted over a limited period of time of about 30 days, its driving purpose too was to generate new and plausible ideas, presently not in common use in Sri Lanka, but which can benefit both teachers and students, relating especially to LEP students in EMI higher education. Thus, this study intends to encourage both teachers and students of EAP to venture into the relatively new field of corpus analysis technology and techniques for language teaching and learning.

Roberts (2004) also adds ‘Self-Experimentation helped produce plausible new ideas not only by generating ideas but also by testing and developing them.’ [18] Ideas that initially seemed not plausible or too challenging to be practiced or promoted among others, through self- testing and development became more feasible. On the value of self-experimentation, Lubarta and Mouchirouda (2004) comment ‘This research method is valuable in particular because the researcher’s high level of participation in the search for a solution fosters the involvement of the necessary cognitive skills and conative traits’ [19] Lubarta and Mouchirouda (2004) here highlight the impact of self-experimentation on the search for solutions driven by both the researcher’s ideas (cognitive) and the intention or effort to achieve a goal (conative). This present research too intends to introduce a solution to the currently present problem of LEP in English Medium degree programs through the incorporation of corpus analysis technology as a meaningful English language support.

On his own self-experimentation work, Roberts (2010) makes the following observation:

Why did it work so well? First, my position was unusual. I had the subject-matter knowledge of an insider, the freedom of an outsider, and the motivation of a person with the problem [20]

This is reflective of the researcher's position in the current research. The researcher, a non-expert in corpus analysis became self-taught on some of its principals, techniques and technology. However, similar to an outsider, being the material designer there was no direct requirement placed upon the researcher to teach the designed material to the target students. Yet, the researcher was aware that the material designed will be implemented on preliminary level first year undergraduates in preparation for English medium academic work and being an English language practitioner by profession at the tertiary level, the researcher was concerned with the problem of LEP students' struggles in English medium higher education. Thus, in this research there is a clear attempt to exploit the above-mentioned positive traits of self-experimentation research in order to produce a productive and feasible solution.

### **A Critical View of Self-Experimentation**

With many of its benefits, self-experimentation also has drawbacks that researchers should be wary of. Here are a few of the weaknesses highlighted by Roberts and Neuringer (1998) [21].

1. Expectations may influence results
2. Generality across subjects unclear
3. Limited subject matter
5. Interferes with daily life

As the researcher is driven by a certain belief and a desire to produce an outcome, one's views may impact on the outcome. Although the researcher allows for the critical self-voice discussed later, the researcher may be blind to certain weaknesses in the undertaken endeavour. Secondly the generalisability of the research outcomes needs to be established with further experimentation in order to establish its value. The researcher being the main and only subject of the research limits the perspectives and therefore calls for further experimentation to establish the findings of self-research. As self-experimentation can interfere with daily life and vice versa, it may be hard to maintain researcher discipline and keep to a routine in the experimentation process.

### **Application of Conversational Learning (An Experiential Learning Theory)**

The main reason for this research to draw on experiential learning was that it fit well with the teacher's experience or situation in today's ELT context being expected to create and adapt material to suit the students' immediate needs. The focus of this research was to create knowledge through the researcher's personal experience in an attempt to incorporate computer-based corpus analysis techniques into English language teaching material production. This was particularly challenging as the material was aimed at undergraduates with low proficiency in English embarking on English medium tertiary education. In order to draw insights from this effort to inform future teacher-initiated attempts to apply corpus analysis to academic English assistance or EAP material development, this research applied the following concepts from experiential learning into the research process.

This research chose to be guided by a type of experiential learning called Conversational Learning to capture the insights and outcomes of this self-experimentation endeavour. Conversational learning is 'a process whereby learners construct meaning and transform experiences into knowledge through conversations.' [22]. In this form of learning there is an attempt at creating truth or understanding and in the process creating knowledge by engaging in conversation. 'Learners move through the cycle of experiencing, reflecting, abstracting and acting.' [22] The process involving these steps is cyclic which is launched by a concrete experience. This concrete experience provides the basis for reflection and

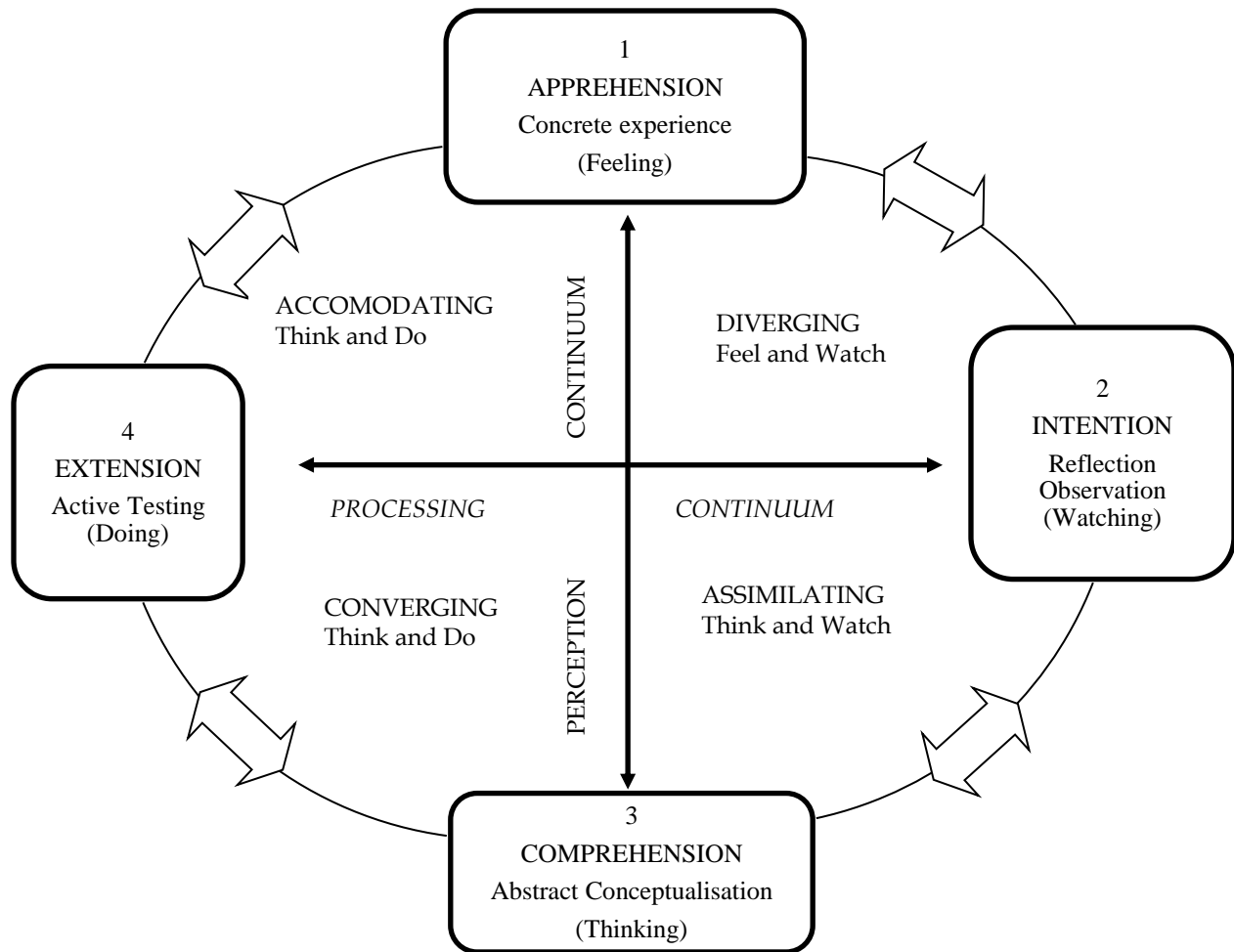
observation. The reflections become implications for action. These implications are next tested leading to the creation of more new knowledge. This cycle is named Kolb's cycle after David Kolb who analysed this model and published it in 1984.

### **Application of Kolb's Cycle**

The researcher chose to apply Kolb's cycle in order to examine if novice stakeholders, specifically the novice teacher cum activity designer, played by the researcher, can take on the task of adopting corpus analysis technology into language teaching material design in order to encourage mainly teachers working with undergraduates to take on the challenge of learning about, incorporating and introducing new concepts and technologies that they see as having potential to support their students. Corpus analysis has been mainly a domain of specialists and teachers had to wait on such specialists to guide them on the use of corpus analysis in language teaching. This creates a delay and even a complete failure in or ignoring of incorporating corpus linguistics concepts and technologies in the local language teaching scenarios. However, if the teacher is empowered to explore new concepts the chances of introducing new technologies and techniques in the local ELT fields such as in higher education will be much faster and the benefits to the students more direct. This research suggests Kolb's cycle as a process to follow to educate oneself of the above-mentioned corpus analysis concepts and generate knowledge on incorporating them meaningfully into language teaching.

The conversation described above can be with oneself and with others. This self-experimentation research selected to focus on and draw on the self-conversation aspect to explore the individual development and capacity building that can happen through internal conversation and reflections on one's own practice. Applying key concepts of Kolb's cycle (see Fig 3) to the conversation, the research follows the four stages of understanding experience; namely the stages of **apprehension** or concrete experiences and **comprehension** which is abstract conceptualization and the two stages of transforming the experiences, namely, **intension** which is reflective observation and **extension** which is active experimentation.

This research was initiated by the researcher's activity designing endeavour for a future language course for low English proficiency undergraduates studying in the English medium (concrete experience). Examining the syllabus and course objectives showed a possibility of introducing the practices of language analysis through computer-based corpus analysis into the course. It was observed that it was possible to both introduce and build these practices into the course activities through clear guided instructions (Reflections and Observations). Reflecting on the students' proficiency and the teachers and students being novice to corpus analysis called for using clear uncomplicated step by step procedures with a lot of support in order to achieve this possibility (Reflections and Observations). It was understood that simplicity and clarity had to be the key to successfully introducing this new concept as a language learning resource (Comprehension). With this in mind the researcher engaged in designing language learning activities incorporating corpus analysis technology. In this manner the researcher attempted to create knowledge through one's own personal experience, on the use of corpus analysis technology by a novice material designer for the use of novice teachers, and novice students even those of low English proficiency in order to achieve academic language teaching and learning outcomes. By engaging in self conversation, applying Kolb's cycle of experiential learning, this research attempts to capture the insights gained through this self-experimentation attempt to shed light on future directions in the application of corpus analysis technology in local higher education ELT scenarios.



**Figure 3.** The four repetitive stages in Kolb’s cycle of experiential learning development

There are many ways to make meaning of one’s personal learning journey, some of which are blogging, essay writing, journaling, storyboarding, dialoging, dancing, singing, texting etc. [23]. This research chose to capture and reflect on the researcher’s experience through dialogic journaling. Here the researcher plays both a subjective and objective role recording experiences and reflecting and questioning oneself on the observations and actions taken and thus creating new knowledge and improving on it.

As shown through the arrows in the above diagram, one may enter the learning experience process at any point and move to any one of the stages shown. In this research too it can be observed that whenever the learner/researcher engages in this research there is movement between and across the four stages shown. These movements are captured in the journal entries.

**The Reflective Internal Dialogue or Conversational-Self Approach Informing Design**

The conversational self-approach is informed by Pask’s conversation theory (1976) [24], where the inherent value of maintaining and evaluating multiple feedback loops is highlighted. Effective feedback, rather than be information about something, needs to be ‘a closed loop’ or having the capacity to affect decision making [25]. Glanville (2008) discusses the value of Pask’s conversation theory and its significance for generating

new knowledge in design research.’ Design and conversation are seen as analogous, so that design can be understood as a conversation held, generally, with the self (via paper and pencil).’ [26] This research, apart from its main purpose of testing the feasibility of the challenging task of a language practitioners’ ‘novice’ in corpus analysis, designing corpus analysis technology-based language teaching material for the consumption of a similarly ‘novice’ and language-challenged audience, was also an opportunity to explore the feasibility of employing of Pask’s conversational theory informing design to achieve this challenging task.

The Conversation theory of Pask shares characteristics with the present study. Pask’s conversational theory is described as ‘an attempt to investigate the learning of realistically complex subject matter under controlled conditions’ [27]. This characteristic is found in this study which is an attempt to understanding the relatively unfamiliar concepts and techniques in corpus analysis, identify their value for language teaching and learning in an EAP context and apply these concepts and techniques in EAP material design. This research was restricted to the self-experimentation method. Another characteristic in common with this research is the generation of new knowledge through design research. Shumack (2010) who draws on Pask’s theory refers to the structured reflection by a designer through journal writing. The researcher designer’s carefully journalled internal conversations or reflective talkback are used to create and expand perspectives. Design decisions are made and contested through self-reflective self-voice. This reflective dialogue named ‘reflective internal talkback’ [28] is recorded as journal entries. These journal entries are created as the design evolves and takes shape and these reflections play an important role in the designing process. In this research the design process was propelled forward, and the design evolved and took shape through insights of the reflective talkback.

The design process described in this paper follows the conversational self-approach, as it unfolds over time. In this conversational scheme the designer seeks to capture a variety of viewpoints about the design space being explored based on the following elements. One is, the use of alternating written entries by the researcher playing multiple roles of designer, research subject and critique. Another element is the use of conversational pairings for writing about any aspect of the design process.

### **The Journaling Process**

The basis of the self-conversation was the journaling practiced during the activity designing process which was a source of support and direction to this designing endeavour. The journal entries of the researcher record the self-reflections self-inquiries on the design process leading to progress and outcomes in the design process. The journaling also captures a variety of viewpoints about the design effort being explored allowing for the extraction of both subjective and objective views of the researcher. As the researcher engaged in a mental verbal conversation, the mental exchanges were journalled and these journal entries captured the stages characteristic of Kolb’s cycle namely apprehension, comprehension, intension and extension.

In the process of journaling, the language practitioner employed the following two questions to self-reflect:

1. Can corpus analysis-based activities be presented to non-expert teachers and students in a user friendly format?
2. How can I make this corpus analysis-based activity more user friendly to non-expert teachers and students?

As the researcher engaged in learning (corpus analysis technology and techniques) and designing (corpus analysis technology-based ESL /EAP material) the researcher carefully journalled the experience on a day-by-day basis. This created a dialogue about the design. This dialogue informed the design process and led to subsequent actions. These dialogues formed multiple feedback loops which included repeated cycles of the process of designing reflecting and adjusting the design. These dialogues drove the whole designing process forward and greatly contributed to the forming of the final design of corpus analysis techniques based English language teaching material.

In this research as the research subject was the researcher, this research did not require any clearances and permissions with regard to data collection. The researcher gives full consent to the publishing of the journal data, all the personal reflections of the researcher cum research subject.

### **Summary**

This research subscribed to related methodologies, namely (1) self as experiment subject and (2) self-reflection informing design.

This project was undertaken to explore the value of corpus analysis technology to the field of English language teaching particularly the field of English for Academic Purposes in Sri Lanka. This research explores the possibilities of harnessing corpora and corpus analysis technology and techniques to assist the target students, LEP undergraduates in EMI to develop their academic language proficiency. Employing self-experimenting and self-reflection this research attempted to understand corpus analysis technology as well as probe the experience of the researcher, novice to the field of corpus linguistics, in harnessing this new technology in developing language teaching material.

Following this research model of self-reflection informing design as well as by employing knowledge building techniques such as Kolb's cycle in partnership with Pask's conversation theory to understand new concepts and produce new designs based on the newly understood concepts this research attempts to encourage corpus analysis application among English language teachers for EAP classroom use. This research promotes the use of corpus analysis applications with LEP undergraduates to build their EAP language repertoire to meet the language demand of English medium higher education. Further, this research mainly aims to encourage self-exploration of corpus analysis among English language practitioners in Sri Lanka to go towards EAP material design incorporating corpus analysis techniques.

### **Experimental Section/Materials and Methods**

The Journaling was done during a period of 30 days, and it captured the internal dialogue within the researcher leading to the development of the design and the resulting knowledge creation. The following is the diary entry for day 1 followed by its analysis and the analysis of the day 6 diary entry. Each analysis has followed the experiential learning interpretation of the journal entries according to Kolb's cycle (see Fig.3)

#### **Day 1 (17/06/2021)**

Are there any free web-based corpus analysis technology that will be not too complicated for teachers and students to understand and use?

Google search for

- web based corpus analysis software

- web based free corpus analysis technology
- free corpus analysis software

Word clouds are easy to understand, can be used for understanding word frequencies, and contextual views of words. But word clouds don't give the frequencies of each and every word.

This online tool <https://voyant-technology.org/?corpus=63e424f91158b74a433bc7c5bb290326> was located at this webpage [http://fedora.clarin-d.uni-saarland.de/teaching/corpus\\_Linguistics/corpusTechnology.html](http://fedora.clarin-d.uni-saarland.de/teaching/corpus_Linguistics/corpusTechnology.html) which lists both downloadable and web-based corpus analysis technology:

It can be used to alert students to word frequencies. However, it didn't highlight word categories, semantic categories referred to as tagging in corpus analysis jargon.

More technology was listed on <https://allaboutcorpora.com/corpus-software-2>. One of which is AntConc. This webpage directs one to the webpage <http://www.laurenceanthony.net/software.html> on which the software is available for free download. AntConc provided the complete list of all the vocabulary used in a particular text. It required that the text be saved in .txt format as it didn't process texts saved as word documents.

The Free CLAWS web tagger <http://ucrel-api.lancaster.ac.uk/claws/free.html> listed on <https://corpus-analysis.com/>

Dissected a text into syntactic (grammar components) and tagged or provided a syntactic or grammar tag for each component.

The syntactic tags were like abbreviations. Being a novice to corpus analysis novice I couldn't understand these abbreviations. So, a google search was done for

- corpus analysis tags
- corpus analysis tags list

a pdf document available on <https://link.springer.com/content/pdf/bbm%3A978-94-015-9273-4%2F1.pdf> was located. It listed the abbreviation with their meanings.

To see if the tags were easy to follow, it was checked on the Free CLAWS web tagger. Using one of the learning outcomes for the course, that is, to identify gerunds. It was found the tag found in the tag list for gerunds in the pdf document and the one assigned by Free CLAWS web tagger matched they were the same. When google search was conducted for text analysis, the tagging used for gerund. These articles were discovered:

**Table 1-Day 1**

Entry Details	Subject Self Voice	Object Self- Voice
Apprehension/ Reflection		Are there any free web based corpus analysis technology that will be not too complicated for teachers and students to understand and use?

Action	<p>Google search for</p> <ul style="list-style-type: none"> <li>• web based corpus analysis software</li> <li>• web based free corpus analysis technology</li> <li>• free corpus analysis software</li> </ul>	
Comprehension	<p>Word clouds are easy to understand, can be used for understanding word frequencies, and contextual views of words.</p>	<p>But word clouds don't give the frequencies of each and every word</p>
Reflection	<p>This online tool <a href="https://voyant-technology.org/?corpus=63e424f91158b74a433bc7c5bb290326">https://voyant-technology.org/?corpus=63e424f91158b74a433bc7c5bb290326</a> was located at this webpage <a href="http://fedora.clarin-d.uni-saarland.de/teaching/corpus_Linguistics/corpusTechnology.html">http://fedora.clarin-d.uni-saarland.de/teaching/corpus_Linguistics/corpusTechnology.html</a> which lists both downloadable and web based corpus analysis technology: It can be used to alert students to word frequencies.</p>	<p>However it didn't highlight word categories, semantic categories referred to as tagging in corpus analysis jargon</p>
Action	<p>More technology were listed on <a href="https://allaboutcorpora.com/corpus-software-2">https://allaboutcorpora.com/corpus-software-2</a> One of which is AntConc. This webpage directs one to the webpage <a href="http://www.laurenceanthony.net/software.html">http://www.laurenceanthony.net/software.html</a> on which the software is available for free download.</p>	
Comprehension	<p>AntConc provided the complete list of all the vocabulary used in a particular text. It required that the text be saved in .txt format as it didn't process texts saved as word documents.</p>	
Action	<p>The Free CLAWS web tagger <a href="http://ucrel-api.lancaster.ac.uk/claws/free.html">http://ucrel-api.lancaster.ac.uk/claws/free.html</a> listed on <a href="https://corpus-analysis.com/">https://corpus-analysis.com/</a></p>	
Comprehension	<p>The Free CLAWS web tagger dissected a text into syntactic (grammar components) and tagged or provided a syntactic or grammar tag for each component.</p> <p>The syntactic tags were like abbreviations.</p>	<p>Being a novice to corpus analysis novice I couldn't understand these abbreviations.</p>
Action	<p>So a Google search was done for</p> <ul style="list-style-type: none"> <li>• corpus analysis tags</li> <li>• corpus analysis tags list</li> </ul> <p>a pdf document available on <a href="https://link.springer.com/content/pdf/bbm%3A978-94-015-9273-4%2F1.pdf">https://link.springer.com/content/pdf/bbm%3A978-94-015-9273-4%2F1.pdf</a> was located. It listed the abbreviation with their meanings. To see if the tags were easy to follow, it was checked on the Free CLAWS web tagger.</p>	
Comprehension	<p>using one of the learning outcomes for the course, that is , To identify gerunds. It was found the tag found in the tag list for gerunds in the pdf document and the one assigned by Free CLAWS web tagger matched. They were the same text analysis, the tagging used for gerund . These articles were discovered: (Note- not clear )</p>	

The analysis of the day 1 diary entry shows the four stages of understanding from Kolb's cycle of experiential learning at work.

On Day 1 the **Apprehension/Reflection** initiates the design process with the searching question 'Are there any free web-based corpus analysis technology that will be not too complicated for teachers and students



to understand and use?’ Being self aware of the limitations in ones knowledge in this area as well as those of the other stakeholders the researcher takes **action** to seek out answers to this reflective question. The awareness of the need for the corpus analysis method to be user friendly in order to increase its useability is reflected in the search items, in the keywords ‘free’ and ‘web based’. A free and web based tool called Wordcloud (Fig 4) was identified and found to be accessible to the envisioned stakeholders and conducive to help one understand the basic corpus analysis concept of word frequency. However on reflecting (**Reflection**) this tool was found to be uinadequate in providing sufficient linguistic or grammar information. On searchiong further the researcher came upon the software AntConc. AntConc was found to possess the earlier sought characteristics, namely, free, and providing complete word frequency data about the selected language text (**Comprehension**). But, the need for a corpus analysis software that provided grammar information remained yet unfulfilled. As the search continued (**Action**), this need was fulfilled by the next identified software ‘The Free CLAWS web tagger’ which possessed all the previously listed desired characteristics, namely, free, web-based and providing adequate grammar information for a selcted text. However, this software seemed to pose a challenge to the novice researcher, that is whether the tags or grammar labels it employed were universally used across all such tagging software. The subsequent search (**Action**) proved the tags were universal.

In the Day 1 diary entry we find the distinction between **Subject Self-Voice** and **Object Self-Voice**. The subject self-voice reports the progression in the design process. On the other hand the object self-voice takes a crtical stand on the process. However, the two voices complement each other. The object self-voice, although highlighting roadblocks to progress it initiates circumventing action and encourages progress.

**Table 2-Day 6**

<b>Entry Details</b>	<b>Subject Self-Voice</b>	<b>Object Self-Voice</b>
Action	Designing first introductory activity with Free CLAWS web tagger	The tags are quite comprehensive But how do I stop overwhelming the students The word Tag itself can be overwhelming Tagger, is a new technical word! for both teacher and students. How can I introduce it?
Reflection	New activity showing pictures of price tags added.	This made me make even the morphgological parts activity more visually appealing and clear and understandable. and used a very every day term ‘small’ word parts to describe the morphological parts (suffixes)
Comprehension		I can highlight an important factor in the chunking activity, that the verb is the essential foundation of the sentence.
Action	Designed a diagram to show how important how central VERB is to the sentence. A tree without roots collapses. Designed a sentence correction activity. The given sentences have no verb. Also a loud recitation repetition activity to drill into the students how indispensable a correct verb is to a sentence. Then to show how central it is used a picture to show that the verb is to a sentence what a foundation is to a house. Designed an activity (comparing to house building) a sentence building around verbs. Added a simple ‘Lets tag’ activity to introduce the word tag, can you think of another word for tag?	

In the Day 6 entry the two stages of **Reflection** and **Comprehension** are observed again. This day's installment to the whole design process begins with the **action** of embarking on designing a learning activity incorporating the identified software, Free Claws web Tagger. Yet, the objective self-voice promptly alerted the designer to a roadblock to future implementation of the designed activity. The researcher/designer is alerted to the need to firstly introduce the key corpus analysis function of tagging or annotating (labeling) a language text for its parts of speech (POS) which this activity intends to exploit. As a result, '**action**' was taken to design preparatory activities to introduce the term 'tag' and its function in everyday terms. **Reflecting** on this part of the design process of preparing preparatory activities helped the researcher/designer understand the value of adding visuals to the learning activities especially to those incorporating the explored technology to improve comprehension and interest in the future consumers/stakeholders (the teachers and students). At this juncture, the researcher further **comprehended** the need to highlight the verb as the central element to understanding sentence level parts of speech. This comprehension led to the designing of several activities to draw the students' attention to the verb and its role in a sentence. Having thus laid the foundation, the day's contribution ended with the designing of an activity to introduce the corpus analysis term 'tag' and its function.

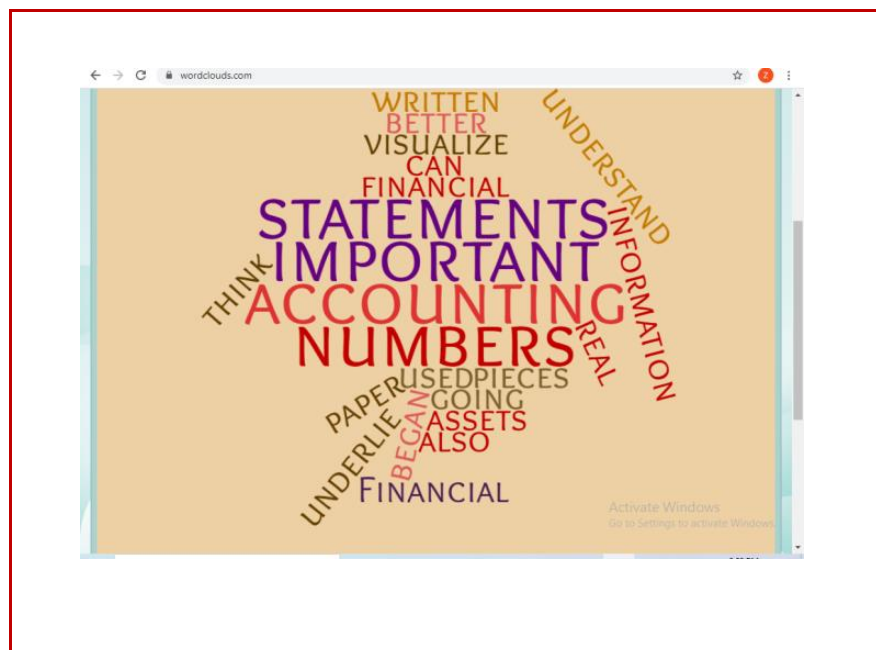
## **Results**

One reason for conducting this self-experimentation research was to empower the individual English teacher in university to self-enlighten oneself on new methodologies and to initiate their applications to their teaching settings to address the students' needs. Referring to the advantageous of self-experimentation, the researcher was able to (1) access the subjective experience of social phenomena. The researcher gained entry into the 'novice' experience of self-informing oneself and identifying applications of the corpus analysis technology and techniques for language teaching in a higher education setting. Through this, the researcher found circumventions to challenges or roadblocks that were raised in the design process. In the process the researcher gained (2) improved mental models of social phenomena. For instance, the use of Kolb's cycle implemented as self-conversation with oneself alerted the researcher to the need for adequate preparatory stages or activities leading up to the main corpus analysis technology-based activities. It was identified that such preparatory activities would be vital to help the target audience of LEP undergraduates grasp the advantageous of the corpus analysis technology more fully. Similarly, the need for visual appeal and clear instructions provided commonly to both the teachers and students were identified as assisting to keep the whole process of introducing and using these novel technology-based tools a feasible, comprehensible and a collaborative learning experience. One of the (3) potentially stimulating new research questions that naturally arose was regarding the teachers and students' feedback on future implementations. Other questions that deserve exploration relate to improving researcher's comprehension of the technology and identifying further applications of corpus analysis technology to language learning needs of LEP undergraduates. The (4) enhanced ability to be reflexive about the given experiment kept the researcher open to questions, gaps in the self-experimentation process, corpus analysis technology applications as well as the material design process. One such identified gap would be the bias and the limitations on insights contributing towards the design process.

The advantageous highlighted by Roberts (2010) were experienced first-hand. The design period benefited from 'the subject-matter knowledge of an insider, the freedom of an outsider, and the motivation of a person with the problem.' [20]. The researcher, being an English teacher to LEP undergraduates at a local

university, was familiar with the LEP undergraduates' challenges in EMI and was able to identify the benefits and relevance of the corpus analysis technology to the target audience. This knowledge motivated the researcher to continue with the project despite challenges. Further, having being in the role of a material developer, the researcher had the freedom to reach a reasonable comprehension of the technology and incorporate it into the learning material design. Further, as the researcher was working independently, the researcher had the freedom to see this effort to a conclusion undisturbed by opposing views of others who did not share the same understanding or views of the technology. However, objective self-voice in the conversational-self technique followed, provided an objective and critical view of the process without hindering the progress of the design process. However, the design process can benefit by addressing the following drawbacks present in this research such as researcher bias, non-generalisable data, limited subject matter input due to the lack of expert input on the technology and lack of multiple outsider views on the design process. Addressing these issues, it is believed, could in fact enhance the applications of corpus analysis technology and the resulting benefits to the identified audience.

The following figures are a sample of the type of activities created through this attempt to incorporate digital corpus analysis technology into material design. Figure 4 is from an activity involving generation of word clouds using a web-based software, available on [www.wordcloud.com](http://www.wordcloud.com). The words in the word cluster created appear in different sizes depending on the words importance to the text and how often it occurs in the text. This was found to be ideal to introduce the concept of frequency yet inadequate to introduce them to true frequency information as generated by a more specialised software such as AntConc, available on <https://www.laurenceanthony.net/>



**Figure 4.** Using word cloud software to introduce the concept of word frequency

Figures 5 and 6 depict the visuals used in the preparatory activities created to precede the introduction of the corpus analysis function called parts of speech tagging (POS tagging).

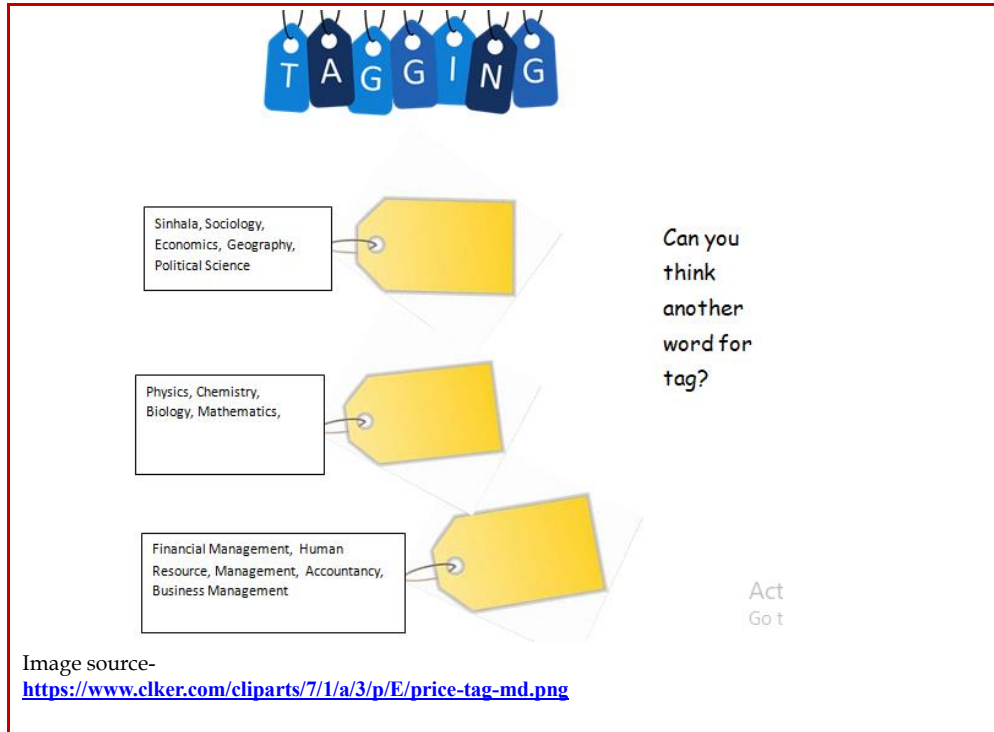


Figure 5. Introducing the corpus analysis term 'Tag'

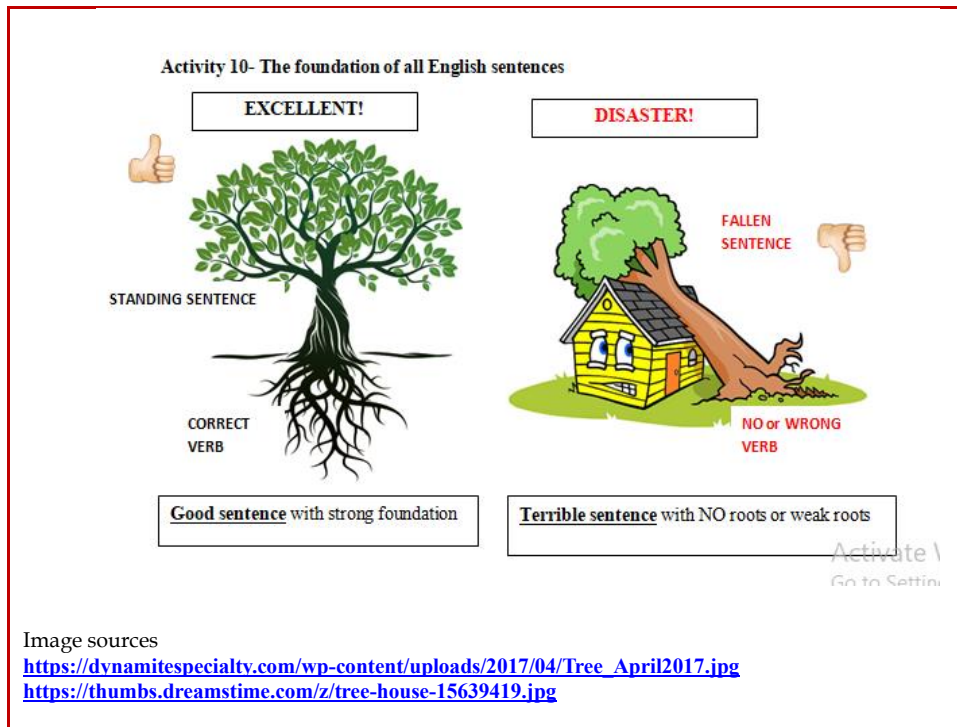


Figure 6. Preparatory visuals to introduce the corpus analysis technique of tagging to locate verbs

Figure 7 depicts the simple instruction and images provided to guide the students on how to use the web-based software called Free Claws web tagger available on <http://ucrel-api.lancaster.ac.uk/claws/free.html> to employ its POS tagging function to analyse and mark a text indicating its grammar structure. Figures 8 and 9 are examples of the POS tagging analysis generated by Free Claws web tagger in two orientations both vertical and horizontal.

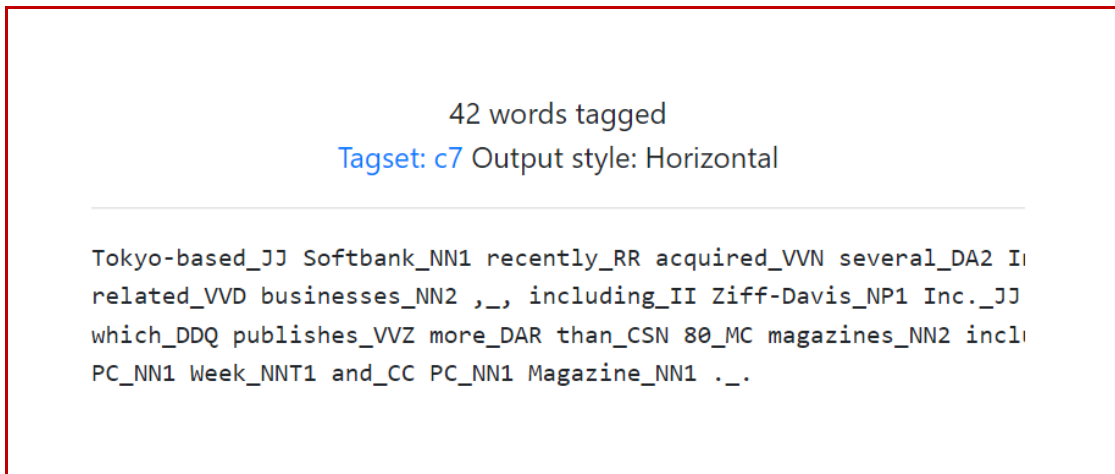
**Figure 7.** Incorporating Free CLAWS tagger to analyse texts for its grammar (POS) components

42 words tagged  
Tagset: c7 Output style: Vertical

---

0000001 002 -----	
0000005 010 Tokyo-based	33 [JJ/78]
0000005 020 Softbank	06 [NN1/79]
0000005 030 recently	03 RR
0000005 040 acquired	93 [VVN/54]
0000005 050 several	93 DA2
0000005 060 Internet	06 [NP1/56]
0000005 070 related	93 [VVD/46]
0000005 080 businesses	03 NN2

**Figure 8.** Part of Speech analysis/Vertical (Free Claws web tagger)



**Figure 9.** Part of Speech analysis/ Horizontal (Free Claws web tagger)

The following example activity is an attempt to provide instructions to guide students on how to use the AntConc software to conduct a word search for past tense verbs in a text from the student's textbook.

#### Example activity

Activity- Let's search for –ed verbs

Follow the following instructions. Try to finish as quickly as possible!

1. Copy the completed passage from the above activity (given below) and save as a .txt file
2. name this .txt file as 'Story'
3. Upload this .txt document to the AntConc software
4. Make sure your uploaded .txt documents are shown under 'corpus Files'
5. Click on the 'Concordance' tab
6. On the search bar at the bottom
  - a. type \*ed and search for words ending with –ed

#### Discussion

The self-experimentation research helped the researcher to explore the possibility of introducing digital corpus analysis techniques as tools for developing English proficiency in undergraduates with LEP in an EMI setting. This was achieved by following a process self-exploring, learning and discovery as set out in Kolb's cycle of experiential learning. This process assisted the researcher, although a non-expert in corpus analysis technologies, to gain a working knowledge or understand of the technologies, identify and incorporate them into activities designed for the above-mentioned target group. Although, the endeavour was challenging the Kolb's cycle stages such as reflection, action and comprehension, assisted to self-reflect and seek circumventions to the resulting challenges and move ahead in the design process. The final

outcome is digital corpus analysis-based language teaching activities which are ready for implementation. The effectiveness and impact of these activities with the target audience, that is, LEP undergraduates in EMI as well as English teachers in higher education who work with such students, are yet to be explored. It is hoped that the results of this self-experimentation will motivate other English practitioners at the university level in Sri Lanka who are working with undergraduates with LEP to explore this technology and other such tools to bring the benefits of such useful technologies and techniques into the classroom and to their students.

In the material designing attempt, it was found that with the researcher/language teacher's initiative it was possible to a reasonable extent to locate web based as well as downloadable corpus analysis software, master the basic techniques of corpus analysis and design activities incorporating them. In this 30-day endeavour the free software located and incorporated, were Word cloud, AntConc, and Free CLAWS web tagger. The corpus analysis techniques incorporated were word frequency, tagging and word searching based on word particles (see Example activity above).

Further it was possible to design activities that employed the students' use of corpus analysis techniques to isolate target language in order to understand their use. It was possible to bring in the aspect of 'user friendliness' by grading the activities and leading the students to the expected outcome through progressive steps. User friendliness for both the teacher and students was built in through the use of suitable visuals and instructions. The instructions were made common to both the students and teachers, as both groups, the target students and the target teachers were 'nonexperts' in corpus analysis. These activities were incorporated into the students' modules as language exercises. The classroom teachers' and students' feedback on the implementation of these activities needs to be investigated.

This self-experimentation endeavour points to the need for further investigation on the perceptions of teachers and students, on how to improve perceptions and identify and select the most suitable corpus techniques and technology that should be integrated into the actual classroom situations. In the light of the material that was developed through this self-experimentation it is clear that the researcher/English language practitioners' efforts towards the integration of corpus technology in actual classroom teaching must be promoted and supported. Such investigations on this integration of corpus as a language teaching methodology gives an opportunity to understand how corpus technology can be exploited by teachers as well as the students. Moreover, 'a systematic familiarization with the basic foundations, implications and applications of corpus linguistics can help language teachers to bridge the gap between theory and practice.' [14]

The possibility of producing new ideas to solve observed problems through self-experimentation and the relative rarity in the use of this research approach to produce EAP solutions in the local context, makes it an advisable endeavour that can be promoted among EAP/ESL teachers for their professional development and to find solutions to students' EAP/ESL needs. However, the lack of critical feedback on the design process by a third party can blind the researcher /designer to drawbacks in the design process. This was evident in this design process as well. When the designs were checked by a second teacher, it was pointed

out that the instructions for the activities and the introductions to the novel corpus analysis techniques and software were not clear enough and needed improvement. However, after the suggested improvements were done, the activities were considered more ready and suitable for classroom use.

## Conclusion

It was found that designing corpus analysis technology-based language teaching and learning activities by non-expert teachers for non-expert teachers and students is both a feasible and a much-needed endeavour in the Sri Lanka tertiary English medium context. The researcher cum non expert English language practitioner's initiative to identify available corpus analysis technology, to understand the basic uses of corpus analysis techniques and design corpus analysis based EAP activities for LEP students while meeting course targets and building in user-friendliness for both students and teachers was found to be a possible exercise. However, as highlighted in the above discussion it is understood that self-experimentation is prone to researcher bias. Therefore, if this approach to designing new technology-based learning material is combined with other approaches such as peer evaluation, it could draw more attention to the drawbacks in the design and gain more insights and input and as a result produce improved and richer outcomes for all stakeholders involved.

## Acknowledgement

I wish to thank Ms Savitri Dias, former head, Department of English Language Teaching, University of Colombo and consultant in the AHEAD project for ELT material development, University of Ruhuna Sri Lanka, for the invaluable opportunity to participate in material development for undergraduates studying in an English as Medium of Instruction context which led to this experimental experience recorded here.

## References

- [1] U. Rómer. Corpus Research Applications in Second Language Teaching, *Annual Review of Applied Linguistics*, **2011**, 31:205 – 225.
- [2] N. Stojković. Theoretical Considerations For Teaching Foreign Languages In University Education Settings in Oleškevičienė G. V., Mockienė L. and Stojković N. *Corpus Analysis for Language Studies at the University Level*, **2021**, (pp. 8-20) Cambridge Scholars Publishing, Newcastle upon Tyne, UK
- [3] T. Cobb, A. Boulton. *Classroom applications of corpus analysis*. In D. Biber & R. Reppen (eds), *Cambridge Handbook of English Corpus Linguistics*. Cambridge: Cambridge University Press, **2015**, p. 478-497.
- [4] N. A. Bobojono'g'li. Application of Corpus Analysis in Teaching Foreign Languages." *International Journal of Progressive Sciences and Technologies* **2019**, 14 (2), 351-355 [Online Web. 3 Jan. 2022]
- [5] N. D. Almutairi. The Effectiveness of Corpus- Based Approach to Language Description in Creating Corpus-Based Exercises to Teach Writing Personal Statements, *English Language Teaching*, **2016**, 9(7), 103-111.
- [6] S. Marinov. Training ESP Students in Corpus Use - Challenges Of Using Corpus-Based Exercises With Students Of Non-Philological Studies', *Teaching English with Technology*, **2013**, 13 (4), pp. 49-76.
- [7] K. Z. Li. The Use of Concordance Programs in English Lexical Teaching in High School', *Higher Education of Social Science*, **2015**, 8 (1), pp. 60-65.
- [8] T. Dudley-Evans, M. J. St John. *Developments in English for Specific Purposes: A Multi-Disciplinary Approach*. Cambridge: Cambridge University Press, **1998**
- [9] S.-L. Lai. EFL students' perceptions of corpus-technology as writing references. In F. Helm, L. Bradley, M. Guarda, & S. Thouésny (Eds), *Critical CALL – Proceedings of the 2015 EUROCALL Conference*, Padova, Italy, **2015**, 336-341.
- [10] L. Anthony. AntConc: A Learner and Classroom Friendly, Multi-Platform Corpus Analysis Toolkit. IWLeL 2004: An Interactive Workshop on Language e-Learning, **2004**, 7-13



- [11] E. E. Khairas. Using Antconc Software as English Learning Media: The Students' Perception, *Epigram*, **2019** 16 (2),189-194
- [12] M. C. Campoy-Cubillo, B. Belles-Fortuno, M. L. Gea-Valor. Introduction to Corpus Linguistics and ELT. In M. C. Campoy-Cubillo, B. Belles-Fortuno, M. L. Gea-Valor, (Eds.), *Corpus-Based Approaches to English Language Teaching*, **2010**, 3-17, London. Continuum International Publishing.
- [13] A. Ş. Özbay, M. N. Kayaoğlu. EFL Teacher's Reflections towards the Use of Computerized Corpora As a Teaching Tool in Their Classrooms Çukurova Üniversitesi Eğitim Fakültesi Dergisi, **2015**, 44(1), 85-104.
- [14] K. M. Dissanayake, R. N. S. R. Harun. Theory and Practice of EAP in the Sri Lankan Context. *Procedia - Social and Behavioral Sciences*, **2012**, 66, 106 – 116.
- [15] A.Ş. Özbay. The Potential of Corpus-Based Collocation Instruction on the Awareness Level of the Turkish EFL Students in Terms of Reading, presented at 8th International Conference on Corpus Linguistics, Malaga, **2017**, (pp. 331-340), EPiC Series in Language and Linguistics
- [16] A. Boulton. Integrating corpus technology and techniques in ESP courses, *ASp* [Online], **2016**, URL : <http://journals.openedition.org/asp/4826> ; DOI : <https://doi.org/10.4000/asp.4826>
- [17] K. Corti, G. Reddy, E. Choi. The Researcher as Experimental Subject: Using Self-Experimentation to Access Experiences, Understand Social Phenomena, and Stimulate Reflexivity. *Integrative Psychological and Behavioral Science*, **2015**, 49, 288–308.
- [18] S. Roberts. Self-experimentation as a source of new ideas: Ten examples about sleep, mood, health, and weight, *BEHAVIORAL AND BRAIN SCIENCES*, **2004**, 27, 227–288.
- [19] T. I. Lubarta, C. Mouchiroud. Why does self-experimentation lead to creative ideas? *BEHAVIORAL AND BRAIN SCIENCES*, **2004**, 27, 269–270.
- [20] S. Roberts. The unreasonable effectiveness of my self-experimentation, *Medical Hypotheses*, **2010**, 75(6), p. 482-489.
- [21] S. Roberts, A. Neuringer. Self-experimentation. In K. A. Lattal& M. Perone (Eds.), *Handbook of research methods in human operant behavior*, **1998**, pp. 619–655. Plenum Press.
- [22] A. C. Baker, P. J. Jensen, D. A. Kolb. *Conversational Learning: An Experiential Approach to Knowledge Creation*, Westport, CT, USA, Quorum Books, **2002**
- [23] B. McDonald. *Improving Teaching and Learning through Experiential Learning*, Newcastle upon Tyne, UK, [Cambridge Scholars Publishing](#), **2020**
- [24] G. Pask. *Conversation Theory Applications In Education And Epistemology*, Elsevier Scientific Publishing Company. Amsterdam, **1974**.
- [25] P. Hoverstadt. *The Fractal Organization: Creating Sustainable Organizations with the Viable System Model*, John Wiley & Sons, Germany, **2008**.
- [26] R. Glanville. Conversation and Design. In R. Luppicini (Eds.), *Handbook of Conversation Design for Instructional Applications*, **2008**, pp. 59-79.
- [27] B. Holmberg. *Theory and Practice of Distance Education*, Routledge, New York, 2005
- [28] K. Shumack. The conversational self: Structured reflection using journal writings. *Journal of Research Practice*,**2010**, 6(2), Article M17.