An L2 Teacher-fronted and CALL program compared: An Empirical Study

1. Introduction

Bidlake (2009, p. 94) argues that "with the advancement of multimedia technologies, such as interactive video and voice recognition software, self-instructed CALL is becoming a popular alternative to more traditional self-instruction materials". Multimedia materials enable language learners to practice language through a multitude of modalities, at their own pace, choosing their own route through the learning materials. Mayer (2001) and Craig, Gholson, and Driscoll (2002) stated that when learning content is presented in a range of media forms, and when the learner can exercise control over the content, learning can occur at a deeper level, mostly because learners are more actively engaged in the learning process (Evans & Gibbons, 2007). Indeed, the control feature available in multimedia gives the learner the opportunity to choose the pace and path of the learning process, the learning content. For example, a learner can use the control feature to watch a movie either with or without subtitles, pause, and rewind it. The interactive nature of multimedia provides easy access to websites that provide authentic learning materials. It is also reported that through adjustment of the "amount and difficulty of input", working at one's own pace, for instance, and through scaffolding, learners can avoid the danger of frustration they might otherwise have faced through "information overload" (Pennington, 1996, p. 9; cf. Kao & Windeatt, 2014, p. 4).

Research regarding the effectiveness of technology-based self-paced instruction reveals that "this type of instruction improves performance and that students master the learning objectives in significantly less time than students in group-paced instruction" (Dalton, Hannafin, & Hooper, 1989; Fletcher, 1996; cf. Dobrovolny, 2006, p. 55). In addition, it promotes student-centred learning in which students assume greater responsibility for their learning. In self-paced learning, the learner has the advantage of determining the learning sequence, pace of learning, and possibly the media. "For example, in a self-paced computer-based course, two students might begin the course on the same day but one may finish days ahead of the other" (Soyemi, Ogunyinka, & Soyemi, 2011, p. 704). In other words, low-ability learners can move forward at their own pace while high-ability learners can take a non-linear path and may not follow the

chronological pattern of lessons or a slower pace, thereby using the time spent in learning more efficiently (Soyemi, Ogunyinka, & Soyemi, 2011). However, despite the great advantages that CALL may have, the field has seen many applications with traditional teaching methods that focus extensively on explicit grammar instruction, even though the need for Focus on Meaning methods in CALL applications has long been recognized (e.g. see Chapelle 2001; Long & Doughty, 2003). Moreover, the research in L2 CALL instruction suffers from important limitations. Few, if any, CALL methods have been tested and compared for effectiveness on general proficiency gains (see Grgurović, Chapelle, & Shelley, 2013). Investigations are often limited to short term interventions, variables may not be sufficiently controlled for, and most importantly studies may not be ecologically valid in that they have not compared similar students in similar educational programs (Norris & Ortega, 2000).

The current study is an exception in that it compares the effectiveness of an L2 English, truly communicative CALL application to both a teacher-fronted method with the same materials and a traditional communicative language method in a semester long experiment.

2. Pedagogical approach in the CALL program

Since the introduction of communicative language teaching (CLT), the approaches to teaching English as a second/foreign language have seen a shift, which Long (1997) likened to a revolution. CLT, giving precedence to communication over form, rejects previous structural approaches to language learning and embraces the use of authentic texts, texts that are created for genuine communicative purposes rather than for use in the language classroom. The focus is on teaching of communication through language and not the teaching of language through communication (Allwright, 1979). In designing learning materials, CLT places emphasis on using authentic materials which are "rich but comprehensible input of real spoken and written language in use" (Willis, 1996, p. 11). With reference to authentic language input, Pinner states, "There are actually numerous definitions of authenticity which have arisen overtime from the research literature" (2013, p. 148). Hong (2013, p. 18), in her study on a movie method similar to ours, refers to authentic materials as "real-life language materials, not produced for pedagogic purposes (Wallace, 1992), but for real-life communication by real people (Nuttall, 2005)."

Authentic materials, in fact, are examples of oral and written language used by native speakers in daily situations for their own purpose (Roger & Medley, 1988) and not language specifically produced and designed for language teaching. Tomlison and Masuhara view authentic language as "designed not to transmit declarative knowledge about the target language but rather to provide an experience of the language in use" (2010, p. 400).

Based on these definitions, sources from which authentic materials can be obtained for second language teaching are advertisements, newspapers, magazines, pictures, symbols, radio news, TV programmes, movies, songs, literature, and the internet (Hong, 2013). Hence, authentic materials in the L2 classroom not only help expose learners to as much language as possible, but also shows them such language is used in real communicative settings in the real world.

The use of a movie in the current study is very much in keeping with this pedagogical change, which in turn is in line with a dynamic usage-based (DUB) perspective, an approach that takes a usage based perspective on language and a dynamic approach to change and development (cf. AUTHOR & X). The main difference with CLT is that especially in the early stages of language development, the DUB approach focuses more on lexis than on grammar and more on input than on output. Movies not only provide scope for input but also contain dialogue which approximates spoken language in real life interaction (Schmitt, 2010) and many opportunities for discussion of the semantic nature of the language (Hong, 2013).

Many scholars have concluded that using movie segments for language teaching can be beneficial in many ways: Pezdek, Lehrer, and Simon (1984) state that movie fragments help enhance memory and recovery of information in reading and listening. Abbs, Cook and Underwood (1980) argue that the use of movies in second language instruction is very much in line with early communicative approaches. For example, in line with the natural approach (Krashen and Terrell, 1983) movies may provide authentic input with a focus on meaning and communication and in line with Roberts (1986) the use of authentic sounding dialogues may help develop strategic competence.

AUTHOR & X also outline some other reasons for using movies in second language teaching: in a good movie, actors act as naturally as possible, coming as close as foreign language learners can get to "real life"; movies also provide examples of cultural, social, or pragmatic issues of which the teacher can take advantage of to provide scaffolding by asking leading or probing questions to extend or elaborate the knowledge the learner already possesses;

movie segments also work as a "soap opera" creating interest and curiosity in wanting to know what happens next; repeated exposure to the movie scenes gives learners the benefit of noticing linguistic features that they may miss in single viewings.

Scaffolding is generally referred to instances in which the teacher or a more knowledgeable peer helps the learner by providing comprehensible input and thereby moving the learner into the so-called zone of proximal development (Bruner, 1978). The zone of proximal development, a concept proposed by Vygotsky is […]

[…] the difference between the child's developmental level as determined by the independent problem solving and the higher level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers. (1978, p. 86)

Even though Vygotsky's notion focuses on child language acquisition, the same philosophy can be applied to second language learning in which the main issue encountered by learners, especially beginning and lower intermediate level, is their inability to use language in an increasing range of contexts. In order to develop this ability, learners need to have an understanding of the cultural and situational contexts in which they have to use language and then have the language resources to realize the meanings appropriate to those contexts. Vacca (2008) suggests that when guided, supported, and provided with the necessary attributes, students become more responsible for their learning, more motivated, and more successful. Applebee and Langer (1983) used the notion of instructional scaffolding to describe essential aspects of formal instruction. In their view, learning is a process of gradual internalization of routines and procedures available to the learner from the social and cultural context in which the learning takes place. In instructional scaffolding, the language learner is assisted in a new task by a more skilled language user, who models the language task to be used verbally and/or in writing by asking leading or probing questions to extend or elaborate the knowledge the learner already possesses. In this manner, rather than evaluating the learner's answers, the teacher is supporting, encouraging, and providing additional props. It is suggested that as the learner's competence grows, the scaffolding should be gradually reduced until the learner is able to function autonomously in that task and generalize to similar circumstances.

Dickson, Chard, and Simmons defined instructional scaffolding as "the systematic sequencing of prompted content, materials, tasks, and teacher and peer support to optimize learning" (1993, p. 12). They believed that scaffolded instruction optimizes student learning by providing a supportive environment while facilitating student independence.

According to Herrell & Jordan (20040 second language learners particularly benefit from so called visual scaffolding, which includes images and words that can be both seen and heard. The image or key words can serve to make the input considerably more comprehensible. It is also important that scaffolding be provided to only those students who need it and only when they need it. When students learn in an individual self-paced learning environment by way of a self-instructional material-centered multimedia computer programme, for instance, instructors are expected to adapt to the environment and their scaffolding needs to be adjusted to fit this new learning medium. Hence, in the current study, learning support was provided to the students through both instructional and visual scaffolding: pictures/illustrations, movie segments (both audio and video), multiple exposure to the segments, captions (onscreen text in the same language as audio), and review questions at the end of each module or scene.

The theoretical base of the current study, a DUB approach to second language teaching (AUTHOR), takes a holistic approach in presenting language constructions (also referred to as form-use-meaning mappings) with their syntactic, semantic, pragmatic, and discourse elements synchronically in meaningful and real-life contexts. The fundamental concepts of a DUB to teaching a second language are as follows:

One of the characteristics of a DUB approach to second language teaching is the focus on input. Van Patten and Benati (2010) define input specifically as the language that "learners are exposed to, that is, language couched in communicative contexts that learners either hear or read" (p. 36) and distinguish it "from language that […]the instructor might provide as models or examples of how to do something. It is distinct from language that learners process purely for its formal features. It is also distinct from output the learners produce" (pp. 94-95).

Language input is considered a necessary and vital variable in second language development (Gass, 2013; Van Patten, 2004). What these authors stipulate is that L2 learners require extensive second language exposure as language input to build their internal linguistic systems and that input needs to be easily comprehensible and message-oriented to be processed

effectively by learners. In this connection, Krashen's input hypothesis continues to assert its influence regarding the role of language input and the necessity of exposure to comprehensible language input. The input hypothesis claims that, for second language acquisition (SLA) to take place, language learners have to have exposure to comprehensible language which is authentic, interesting and/or relevant, not grammatically sequenced, and includes language structures that are beyond their current level (i+1). Krashen, (1981) identified comprehensible language input as "the only causative variable in SLA" (p. 57) and argues that for L1 or L2 acquisition to take place there should first of all be plenty of understandable exposure in a relaxed learning context, and early output and correction should be avoided. In addition, it "must be abundant enough for the learner to abstract regularities from concrete exemplars of language use" (Zyzik, 2009, p. 14). We do not support Krashen's strong claim that input is the only thing needed to acquire a second language, but we do suggest that adequate amounts of input are of utmost importance and should come before encouraging output and interaction.

Within usage-based approaches, many studies have investigated the importance of frequency for second language acquisition. Ellis states that "(…) the acquisition of language is exemplar based. It is the piecemeal learning of many thousands of constructions and the frequency-biased abstraction of regularities within them" (2002, p. 143). According to Langacker (1987), abstraction of regularities happens through the process of "entrenchment" (p. 59). Langacker refers to entrenchment as being the result of repetitions of cognitive events, that is, by "cognitive occurrences of any degree of complexity" (p. 100). Thus, the degree of entrenchment of a cognitive or linguistic unit correlates with its frequency of use. Cognitive or linguistics units are defined as symbolic units or form-meaning mappings, conventionalized in the speech community and entrenched as language knowledge in the learner's mind (Langacker, 2008a). They can be idioms (storm in a teacup, sweep under the rug), fixed phrases (in vitro, by and large, rock 'n roll), noun compounds (olive oil, laser printer), compound verbs (take a nap, bring about), etc. Langacker also elaborates that "learning" or "exposure" should "occur in meaningful context exchanges, approximating socially and culturally normal usage events" (p. 81). Thus, if second language teaching is to be effective, it should focus on multiple exposures to conventional units (also referred to as linguistic constructions, multiword expressions, and formulaic sequences) in meaningful real-life contexts (or real life-like context).

Based on the premise that multiple exposures to conventional units are needed in L2 development, the current study focuses mainly on exposing learners to authentic usage-based events (meaningful units) through a movie (divided into short segments) on the assumption that abstract constructions will emerge through the multiple exposure to such units (Larsen-Freeman, 1976; Ellis, 2002) and the process of entrenchment. Indeed, mere exposure will not lead to the acquisition of a linguistic construction.

Following the line of thinking outlined above, a semester long CALL course was developed taking a movie as content. The movie was cut up into short segments and made into a lesson with frequent repetition of the language used and with careful scaffolding for understanding. The course was truly communicative in that the only goal was for students to comprehend the story and understand what the characters said. The emphasis was primarily on meaningful input. The same content was also used in a teacher fronted class.

3. The Study

The study compared three modes of teaching: a DUB inspired CALL computer programme (DUBc), a DUB inspired teacher-fronted program in which the PowerPoint contained the same information as in the CALL program (DUBp), and a Communicative Language Teaching inspired teacher-fronted program (tCLT) that was taught traditionally at the university, in which the four skills (reading, listening, speaking, and writing) were practiced. The research question were as follows:

1. Is a CALL program based on DUB principles more effective than a teacher fronted course based on the exact same materials?
2. Is a DUB inspired program based on a movie more effective than a traditionally taught course?

3.1 Subjects/Participants

Seven classes of students, comprising in total 228 registered students of the Faculty of Management Studies and Commerce (FMSC), University of Sri Jayewardenepura participated in the current study 2, which took place in the second semester of the academic year 2011/2012. These 7 classes were assigned randomly to three conditions: one class of 55 students to DUBc, 3 classes of 100 students in total to DUBp, and 3 classes of 73 students in total to tCLT. However, at the data analysis stage, 10 participants from the DUBp condition and one each from the tCLT and DUBc conditions were excluded from the quantitative analyses of the general English test because they had either not written their names clearly on the test paper or had not taken one of the general English proficiency tests (either the pre-test or the post-test). With regard to written feedback, data of 4 of the DUBc condition and 6 respondents of the DUBp condition were eliminated from the analysis due to illegible handwriting, leaving 51 texts of the DUBc and 94 texts of the DUBp respondents for quantification.

3.2 Teacher Participants

Nine female tutors and one male instructor/computer technician affiliated to FMSC participated. The nine tutors and the instructor whose average age was thirty years (sd?) were all experienced in their respective areas of expertise, either in teaching English as a second language at tertiary level or in computer science.

Even though the researcher had originally planned to control for teachers in the sense of getting the same teacher to teach both a DUBp and a tCLT group, it proved impossible to do so due to practical reasons. Hence, six different tutors taught the three DUBp and the three tCLT groups.

Three teachers (one of whom actually taught one of the classes used in the experiment and two other volunteers), the computer technician, and the researcher were present in the computer lab to assist the DUBc participants in case they needed any technical assistance. Most students in Sri Lanka are quite knowledgeable about computers; nevertheless, the researcher sought the assistance of the teachers and the computer technician to avoid problems of a technical nature.

3.3 Instructional Materials

3.3.1 DUBc and DUBp

The DUBc and DUBp instructional approaches were very similar in that they consisted of a pedagogical sequence based on the tenets of a dynamic usage-based approach to second language teaching (AUTHOR & X). They were both multimedia instructional approaches that provided authentic input with the movie *The Pursuit of Happyness* (2006, 1:57). This movie was chosen because it dealt with business situations that would be of interest to the Management students and had an interesting and inspirational storyline.

The main difference between the two approaches was in the mode of delivery of the instructional materials. The DUBc was delivered through a self-instructional material-centred multimedia computer programme. The DUBp was delivered by a teacher, using a face-to-face delivery mode in a PowerPoint-based multimedia format. For the DUBc experiment, the PowerPoint-based multimedia instructions of the DUBp condition were transformed into 34 e-learning lessons (modules) through CourseLab developed by the Russian company WebSoft Ltd (© WebSoft Ltd., Russia), an e-learning authoring tool, and published on CD-ROMs for easy transferral to the university computers (Yi Liu, 2012). In the DUBc condition, a computer was assigned to each participant who worked at her/his own pace in an individual learning environment.

Both instructional programmes (DUBp and DUBc) consisted of an introduction (the movie trailer) and 33 scenes of the film, and the design layout was consistent for each scene. Since the object of a DUB approach to second language teaching is that learners should understand everything, for instance everything the characters say in a movie segment, including the intended meaning of the utterances, the movie segments were shown repeatedly to the students and all the utterances were explained in detail, the rationale being that upon each viewing the students could focus on different aspects. Each movie segment was taught in six steps that would help the learners understand everything:

Step 1: The learners were asked to watch the scene without sub-titles. The goal was to have students pay attention to the images and events in the scene to get the gist of what was happening. A very general question was asked such as "*Where are the son and father going?*" to activate thinking and deduction skills.

Step 2: The learners watched the same scene with subtitles. The goal was to have students pay attention to what the characters said. This step could be repeated when needed in both conditions.

Step 3: The learners saw and heard each utterance of the scene on separate slides with explanations. The goal was to help students understand each utterance through scaffolding by means of paraphrases, illustrations, or where needed an L1 translation (Sinhala and Tamil). Where appropriate, not only the literal meaning but also the intended, pragmatic meaning was explained.

Step 4: The learners saw the whole scene again with sub-titles in the DUBp condition and with the text with items that could be clicked on for explanation in the DUBc condition. The goal was to consolidate comprehension. In the DUBp condition, the students were able to ask the teacher for explanations of the utterances that the participants found difficult to understand.

Step 5: The scene was shown once more, but now without sub-titles. The goal was to expose the learners once more to the scene and give them a sense of accomplishment in understanding the scene completely.

Step 6: The students took a brief comprehension quiz consisting of True/False questions, Ordering questions (learners were asked to listen and re-order the jumbled phrases/sentences), and (3) Fill in the blank questions. These questions ranged in difficulty from simple, general questions to very specific questions and tested different linguistic skills (reading, listening and writing). The questions were displayed on a power point slide for the DUBp participants, while the DUBc attempted them on the question screen. The goal was to revisit the language utterances once more and give students a sense of accomplishment.

Although it is not possible to control for all variables in an ecologically valid environment, the DUBc and DUBp programmes were quite similar. The main difference was that students could determine their own rate of learning in the DUBc condition as they were able to repeat or skip steps where preferred. In the DUBp condition the rate was determined by the teacher.

3.3.2   tCLT

The tCLT group took the course conducted by the Faculty for the first-year, second-semester students. As such, they used the course materials that were specially designed for the undergraduates of the faculty, based broadly on the communicative language teaching approach.

The Business Communication course is an integrated skills course which attempts to develop the learners' skills in reading, writing, speaking, listening, grammar, and vocabulary. At the end of the course, the students are expected to be able to participate in business meetings, confidently and effectively; communicate effectively in a variety of situations; take notes on the basis of auditory texts; read and respond to texts for a variety of purposes; write short formal texts and business letters; and give brief presentations (Business Communication II). The course encourages students to work both individually and in groups, and the teachers are advised to use cooperative learning as an instructional strategy. Cooperative learning is the instructional use of small groups so that students work together to make the best use of their own and each other's learning (Slavin, 2011). However, class teachers often complain that it is not possible to get the students to work in groups due to space constraints in the classrooms as a result of the large number of students. Instead, the teacher usually gives instructions, explanation of vital points, and general feedback for students (it is not always possible to give individual feedback due to class sizes) in English, Sinhala, and sometimes Tamil languages.

3.4 Measures and Procedures

The intervention lasted thirteen weeks, and improvement was assessed by means of a pre-test post-test design. All met two days a week for two hours. However, the DUBc students finished their programme faster and at different rates. Some took only 10 hours to complete the course and others 20 hours.

The same test battery was administered before and after the intervention: a general objective English proficiency test (GEP, a writing component, and a questionnaire at the end of the course in the DUBc and DUBp conditions.

The GEP was the same test that Hong (2013) used for her study and was pilot tested again in the current study. It consisted 68 items on vocabulary, grammar, pronunciation, reading, cloze, dialogue matching, and listening. (See appendix for the full test). Each correct answer was given a 1 and each wrong answer was given a 0. The maximum possible score was 68.

A writing task was administered after the GEP test under examination conditions. The aim was to measure how well a learner spontaneously wrote in the target language. The topics given were *My best friend, The most unforgettable day in my life, My goals and dreams for the future, My hometown.* The participants were free to choose one of the topics. These topics were assumed to impose little or no constraints on the participants and the open and free nature of the tasks allowed students with different levels of English proficiency to attempt the task, as they did not demand use of specific grammatical or particular lexical items. Participants were encouraged to write as much as possible. No word limit was imposed as text length also gives an indication of proficiency. However, the participants had to complete the general English proficiency test and the writing task within the stipulated time, which was 70 minutes for the whole test.

The writing texts were rated and analyzed separately as follows: The 146 texts written by the 73 participants at the pre-test and the post-test, computer typed exactly as written by the participants, were rated holistically by four independent raters. As in AUTHOR, X & X, the texts were rated on a scale ranging from 0 (for papers in which no writing was attempted) to 7 (for papers considered the best among the samples). That is, the raters were requested to give an impressionistic overall score, based on the respective strengths of the written text rather than its shortcomings. The four raters rated the texts independently. The texts were arranged in chronological order, alternating pre and post writings texts. The experimental group's writing products were rated first followed by those of the control group. The raters (four females), who were unaware of the experimental set up, were experienced second language writing teachers from the Faculty of Management Studies and Commerce, University of Sri Jayewardenepura. The only information that was given to them was that the texts were produced by first-year undergraduates who were participants in the researcher's PhD study.

Participants were also asked to provide their written feedback (views) anonymously at the end of the intervention - in either Sinhala, Tamil, or English - on the strengths and weaknesses of the intervention they had received and thus to evaluate the DUBp and DUBc programmes. There was neither a time limit nor a word limit imposed. Texts written in Sinhala were translated into English.

The 145 written comments of the DUBc (51) and DUBp (94) respondents were coded (Strauss & Corbin, 1990) by two independent coders. The coders were the researcher and an experienced second language teacher, who was not involved in the research in any respect. She

was only told that the texts were written feedback pieces obtained from the participants who took part in the investigation. She was not aware of the different treatments that the students had been exposed to.

The two independent coders did a content analysis by reading the data several times. While reading, the coders marked the data by themes by colour coding them and made notes in the margin. This made it easier to identify the key themes and the sub themes. Then, a careful check was done for overlapping or similar categories and all duplications were eliminated.

4. Analysis approach

R 3.2.3 (R Core Team, 2015) and the Statistical Package for Social Science (SPSS) Version 16.0 were used for the analyses. Initially, before subjecting the data set to analyses, Cronbach's alpha reliability analyses were performed to ensure that the general English proficiency pre and post-test scores showed internal consistency. Next, a Pearson's product-moment correlation was computed to assess the relationship between the pre-test and post-test scores of the general English proficiency and writing. Then, a one-way ANOVA was computed on the pre-test scores of the three conditions to examine if there were differences between the means of DUBp, tCLT, and DUBc conditions at the outset of the study.

To assess the influence of the type of intervention type on the general English proficiency (GEP) performance, we fitted a mixed-effects logistic regression model (Baayen, 2008, Ch. 7; Baayen, Davidson, & Bates, 2008) with as dependent variable the correctness (i.e. correct, 1, or incorrect, 0) of each answer. By using mixed-effects regression, we are able to take into account the variability per participant, class and question (i.e. these are the random-effect factors). We use logistic regression, as our dependent variable is binary (1: correct; 1: incorrect). Consequently, the estimates need to be interpreted with respect to the logit (i.e. the logarithm of the odds of observing a correct vs. an incorrect answer) scale. This means that positive estimates indicate a higher probability of observing a correct answer, whereas a negative estimate indicates the opposite. The total number of observations in our dataset was equal to 29,376 (216 participants x 68 questions x 2 tests). In our analysis, we will assess the influence of the fixed-effect factors "Condition" (the type of training program: DUBc, DUBp, and tCLT) and "Test"

(either pre-test or post-test). To obtain the best model, both with respect to the optimal random-effects structure and the best fixed-effects structure, we use AIC comparisons (Akaike, 1974). A lower AIC indicates a better fitting model, while taking into account the complexity of the model. A threshold of 2 as AIC difference was used to favour a more complex model over a simpler one. Consequently, the more complex model is at least 2.7 times more likely to be an adequate model of the data than the simpler model. Fixed-effect predictors were considered to be significant in the final model summary if their two-tailed *p*-value was lower than .05.

To assess the influence on the writing performance, a similar approach was followed as above, with the only difference that instead of a mixed-effects logistic regression model, a mixed-effects linear regression model was fitted (as raters gave a numeric score between 1 and 7). Consequently, estimates directly reflected the influence on the actual score. Furthermore, as the individual scores of all four raters were included (i.e. the total number of observations was equal to 216 participants x 4 raters x 2 tests = 1728 observations), we assessed the inclusion of the random-effect factors participant, class and rater.

The written feedback data that were obtained from the DUBc and DUBp participants were first turned into a matrix, where the rows were the units of analysis (the respondents or the individual students who provided feedback). The respondents were assigned numbers 1, 2, 3….), and the columns were the variables (the themes that emerged). The cells were the values for each unit of analysis (respondents) on each variable (Bernard, 1996, p 10). The presence of a theme for each comment was coded as "1" and the absence of a theme on each comment was coded as "0". The sum total of the variables that denoted negative and positive comments (separately) were obtained and the difference between the two variables was considered the dependent variable for an independent samples *t*-test in SPSS.).

4.1 Results

GEP scores

Consistency of the GEP scores was good with Cronbach's alpha equal to 0.88 for both the pre-test and the post-test. The correlation between the pre-test scores and the post-test scores (i.e. the total score per participant on the test) was $r = .71$ ($p < .01$).

The general English proficiency pre-test scores of the participants were inspected to determine if they were at the same level of English proficiency at the beginning of the study. For this purpose a one-way ANOVA was conducted on the pre-test GEP and writing scores. Table 4.1 presents the GEP scores and writing performance (pre-test) measures separated by the type of training program.

Table 4.1: Minimum, maximum, mean, and standard deviation of GEP and writing performance (pre-test) measures by condition

Measures	N	Minimum	Maximum	Mean	Std. Deviation
Pre-test GEP DUBc	54	24	61	43.31	8.39
Pre-test GEP DUBp	90	21	60	43.92	8.82
Pre-test GEP tCLT	72	17	61	42.92	10.89
Pre-test writing DUBc	54	1.50	4.00	2.84	.63
Pre-test writing DUBp	90	1.75	6.25	3.25	.79
Pre-test writing tCLT	72	1.50	4.75	3.02	.73

The results of the one-way ANOVA indicated that there were no statistically significant differences in the GEP pre-test variable between the means of the DUBc, DUBp, and tCLT conditions, $F(2,213) = .232$, $p = .79$. However, with respect to the writing pre-test variable, there were statistically significantly differences between the three conditions, $F(2,213) = 5.5$, $p < .005$. A Bonferroni posthoc procedure revealed that the DUBc group scored significantly lower on the writing task in the pre-test than the DUBp group ($p = .004$).

To assess our research question (i.e. the influence of the type of instruction on performance, we fitted a mixed-effects logistic regression model with as dependent variable the correctness of the answer. We considered the significance of the predictors relating to the type of

test (pre and post) and the type of training (i.e. the condition). Our initial analysis (including the predictors Test and Condition and their interaction, together with random intercepts for participant, class and question) revealed that the improvement on the post-test was higher for DUBc as compared to DUBp and tCLC. DUBp and tCLC did not differ significantly. Consequently, we created a contrast distinguishing DUBc from the other two training types. Our final model, including the optimal random effect structure (which consisted of random intercepts for participant, class and question, as well as a by-participant random slope for test, and by-question random slopes for test, the DUBc vs. other training program contrast, and their interaction) showed that while all training programmes resulted in a significant improvement in the probability of answering a question correctly from the pre-test to the post-test, the improvement for the DUBc training was largest. Table 4.2 shows the coefficients of the logistic regression analysis, including their significance.

Table 4.2: Fixed-effects structure of the logisticmixed-effects regression model predicting the probability (in terms of logits) that a participant answers a question correctly in the GEP test.

Fixed effects:	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.73	0.25	2.90	< .01
Post-test vs. pre-test	0.58	0.07	8.57	< .01
Contrast between DUBc vs. DUBp and tCLT for pre-testt?	0.07	0.45	0.15	.88
Post-test vs. pre-test increase for DUBc as opposed to DUBp and tCLT	0.37	0.15	2.53	.01

Figure 4.1 below illustrates the significant improvement of the DUBc participants on the questions of the general English proficiency post-test in comparison to the DUBp and tCLT participants. Note that the logits have been converted to probabilities for this graph. It is clear

that while the increase in probability of answering a GEP question correctly for the DUBp and tCLT participants from the pre-test to the post-test is about 11%, it is 16% for the DUBc participants.
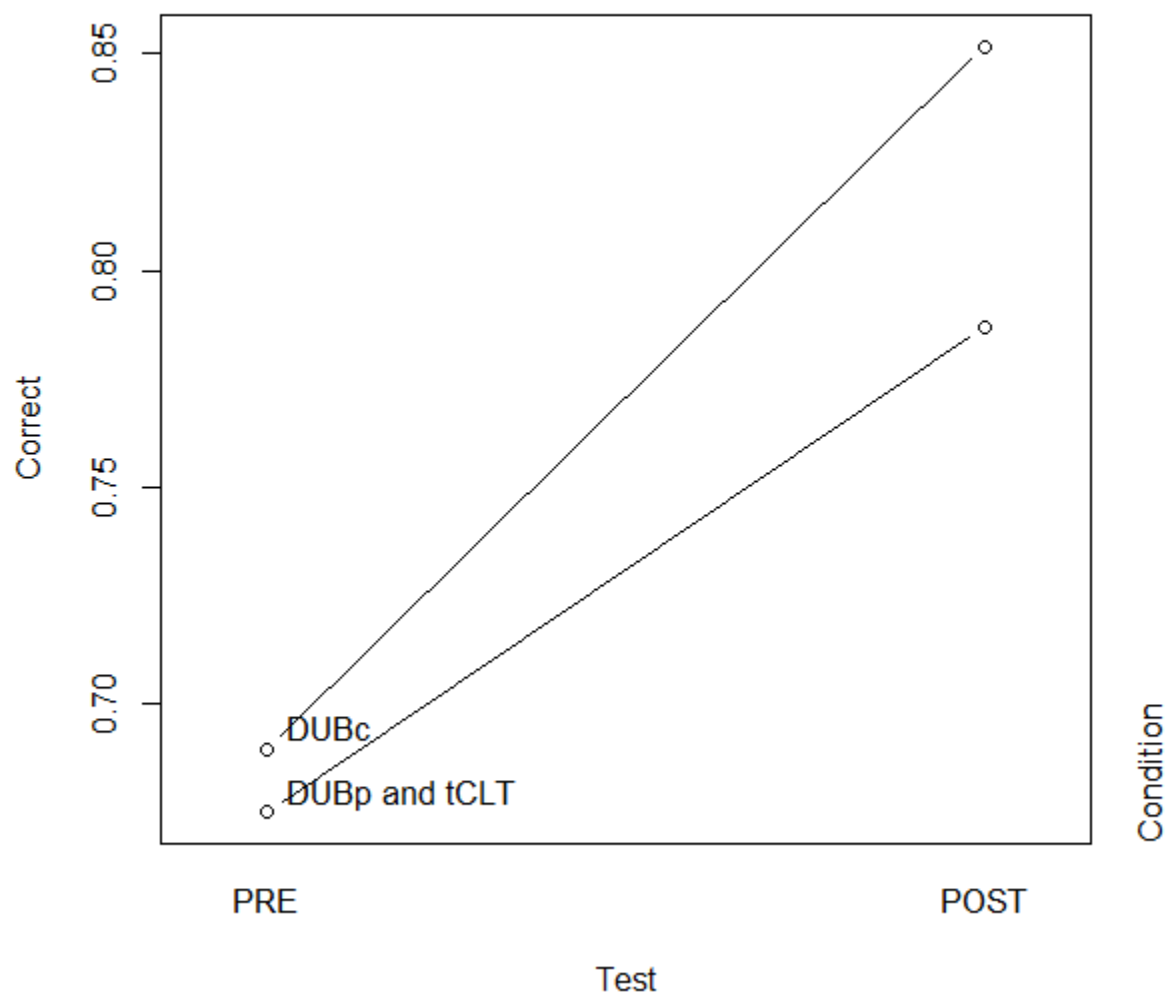

Figure 4.1: General English proficiency performance (i.e. the probability of answering a question correctly) of the pre-test and post-test for the DUBc participants vs. the DUBp and tCLT participants.

Writing scores

For the writing scores (of the four raters) consistency was acceptable (Cronbach's alpha: 0.75 for the pre-test and 0.69 for the post-test). The correlation between the (average) pre-test scores and the post-test scores (averaged was $r = .66$ ($p < .01$). With respect to the writing component of the pre-test and post-test, the optimal linear mixed-effects model indicated only a significant difference between the scores on the pre-test and post-test (with the latter being about 0.45 points higher), but no significant differences in the effect of the type of training program on this improvement. The optimal random-effects structure consisted of random intercepts for participant, class and rater, as well as a by-class random slope for test.

Table 4.3: Fixed-effects structure of the linear mixed-effects regression model predicting the writing score of a participant (for each rater)

Fixed effects:	Estimate	Std. Error	t-value	Pr(>|t|)
(Intercept)	3.04	0.18	16.55	< .01
Post-test vs. pre-test	0.45	0.13	3.47	< .01

Qualitatively, the DUBp group pointed out that grammar and writing practice was missing. The DUBc group only seemed to miss speaking practice. The DUBp group was less frequent than the DUBc group in pointing out that the programme was useful and interesting, and provided fewer remarks about the programme as a whole. The DUBc group mentioned with an overwhelming majority that the method was useful. They also pointed out more often than the DUBp group that the method was engaging and easy to learn. About 18% of the DUBc group pointed out the advantage of working at one's own pace.

The written feedback data of the DUBc and DUBp participants were also subjected to quantitative analyses. The data consisted of both positive and negative variables (comments made by the participants). A close examination of the mean scores showed that the participants, in general, were more positive than negative of the two approaches. With an overall mean score of 2.65 (positive comments), the DUBc participants were more positive than the DUBp participants whose overall mean score was 2.05.

An independent-samples t-test was conducted to examine if the difference between the two conditions was significant. The results revealed that there was a significant difference between the DUBc (M = 2.14, SD = 1.34) and DUBp (M = .73, SD = 2.38) and conditions, $t(143) = -4.5$, p < .001.

5. Discussion and conclusion

The study described in this article investigated the effect of three different interventions on the general English language proficiency of 216 undergraduates. One group received a CALL treatment in which students could work in their own pace, one group had the same materials as in the CALL programme but were taught by a teacher, and one group received a traditional communicative treatment also taught by a teacher.

The results show that the CALL programme was more effective than the teacher-fronted programmes with respect to general English proficiency. The DUBc students improved more over the course of the training program than their DUBp and tCLT peers. The DUBp and tCLT groups, whose treatment involved exposure to teacher- directed instructions based on a dynamic usage-based perspective and standard pedagogical instructions based on the current communicative language teaching approach respectively, also experienced gains in general English language proficiency from pre-intervention to post- intervention, but less substantial than the gains found in the DUBc condition.

The results of the writing tests showed that all groups performed significantly better at the post-test than at the pre-test, but there was no difference in improvement between the three groups.

Based on the written feedback the students gave, the DUBc students were significantly more positive than the DUBp students about their training programme. The self-instructional material-centred multimedia computer programme was perceived as more pleasant and interesting than the teacher-fronted alternative. Because both conditions used the exact same materials, we may conclude this was due to the mode of delivery. In the DUBc mode students could work at their own pace. In the DUBp mode, teachers have to address a heterogeneous group and some students may need less exposure than others.

As far as we know this is the first study that has compared a CALL programme with a teacher fronted programme using the same materials and DUB approach in a long-term intervention (cf. Grgurović, Chapelle, & Shelley, 2013) therefore controlling for as many variables as were possible in an ecologically valid experiment (Norris & Ortega 2000). Both were based on a movie, cut in segments, with the same steps involving repetition and scaffolding for understanding. The emphasis was on lexis and input. We believe the students improved more in the CALL programme. because they could pay attention to what they felt they needed to pay attention to and therefore work at their own pace, which was evident from the fact that some finished the programme within 10 hours whereas others finished in 20 hours, but took less time than their group-taught counterparts. This is in line with Dalton, Hannafin, & Hooper (1989), Fletcher, (1996), and Dobrovolny (2006) who claim that self-paced students may master the learning objectives faster than students in group-paced instruction. It is also in line with Soyemi, Ogunyinka, and Soyemi (2011) who argue that students will work at their own pace. The reason may be that the group=taught students may experience information overload (Pennington, 1996, p. 9; cf. Kao & Windeatt, 2014, p. 4) or get bored and distracted, which might explain the more positive responses from the DUBc students. However, we do not believe that any CALL programme would be more effective than its teacher-fronted counterpart. Despite the great advantages that CALL may have, the field has seen many applications with traditional teaching methods that focus extensively on explicit grammar instruction, even though the need for Focus on Meaning has long been advocated in CALL applications (Chapelle 2001; Long & Doughty, 2003).

This study also showed that a DUB approach with emphasis on input mainly is as effective as a traditional course inspired by CLT with explicit focus on grammar, output and interaction. In a previous study conducted at a Vietnamese university (AUTHOR & X), the students in the DUB condition performed significantly better on the same general proficiency test than their traditionally taught counterparts. There may be several reasons for differences between these groups. The Sri Lanka students scored significantly higher on the same general proficiency test than their Vietnamese counterparts and they were exposed to a different movie. Judging from the spontaneous feedback on their respective programs, the Sri Lanka DUBp students complained that they had not learned enough grammar, a remark not found among the DUBc students, so we suspect that the DUBp teachers were not wholly convinced of the method and may have

unwittingly influenced the results. Another possibility is that the tCLT treatment was quite communicative and effective in itself.

Our main result was that the CALL program was more effective in increasing students' proficiency than the two teacher-taught programs. The question is whether this would hold for all CALL programs. We do not believe so, as the current program was carefully designed on DUB principles with interesting content in the form of a movie, a great deal of exposure, repetition, and careful scaffolding using visual material where appropriate for understanding the meaning. There was no explicit focus on form or forms in the traditional sense, which usually deals with morphology only, but there was a focus on form in a non-traditional sense, in that the focus was on the whole expression, its words, phrases, and formulaic sequences. We illustrate this with an example from the program.

In one scene, the main character had just talked to his wife at the hospital and had parked his car there illegally. When he walked back to his car, he found that it had wheel clamp on, and he says "That's what happens when you're in a rush. The program shows the whole utterance in red. A part that may not be understood is also highlighted, and all that is on the screen is read out in a clearly articulated voice.
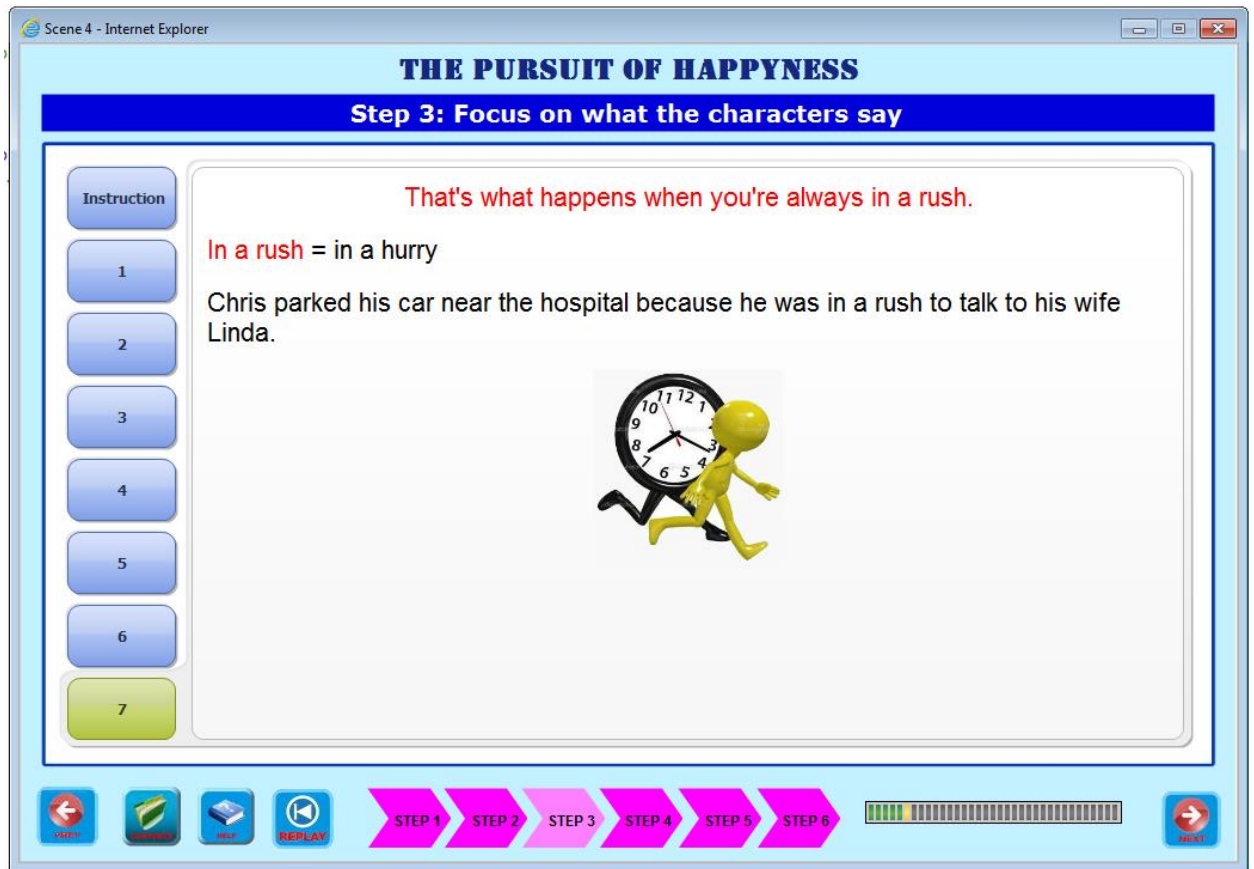
Figure …. A screenshot in Step 3.

The results of the study have practical pedagogical implications. They confirm the belief held by experts in the field of second language education that CALL programmes can be as effective or more effective than a teacher fronted program. Using a CALL program to improve general proficiency could free up teachers to spend more time on higher level skills such as academic writing.

Naturally, this study is limited in scope, interpretability, and generalizability. First, the study was a quasi-experimental one in that randomized assignment of participants to conditions was not possible, as groups of participants were chosen and assigned to different conditions from intact classrooms comprising learners with different levels of language proficiency. Hence, the researcher had to face constraints in controlling for confounding variables such as teachers, attitude and motivation of learners. Second, these students were of intermediate proficiency. From Hong (2013) we now that lower level students would benefit from this approach, but we do not know the effect on more advanced students.

We now know that the DUBc programme is more effective than the two teacher fronted approaches, but we do not know why it is more effective. Is it the control of the learning environment, the frequency of exposure, the scaffolding in understanding the expressions in a holistic manner, or merely listening enough to authentic input? These questions need further research.

References

AUTHOR & X

Continuum.

AUTHOR.

AUTHOR, X & X.

AUTHOR & X

Abbs, B., Cook, V., & Underwood, M. (1980). Authentic English for Reading1. Oxford: OUP.

Allwright, R. (1979). Abdication and responsibility in language teaching. Studies in second language Acquisition, 2, 105-121.

Applebee, A. N., & Langer, J. A. (1983). Instructional scaffolding: Reading and writing as natural language activities. Language arts, 168-175.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. Journal of memory and language, 59(4), 390-412.

Bates, D. M. (2007). Linear mixed model implementation in lme4. Manuscript, University of Wisconsin – Madison.

Bernard, H. R. (1996). Qualitative data, quantitative analysis. Field Methods, 8(1), 9-11.

Bidlake, E. (2009). Learner Experience using self-instructed CALL: Methodological and learner insights. Novitas-Royal, 3(2), 93-109.

Bruner, J. 1978. The role of dialogue in language acquisition. In A. Sinclair, R., J. Jarvelle, and W. J. M. Levelt (eds.), The Child's Concept of Language. New York: Springer-Verlag.

Chapelle, C. (2001).Computer applications in second language acquisition. Cambridge University Press.

CourseLab, http://www.courselab.com

Craig, S. D., Gholson, B., & Driscoll, D. M. (2002). Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features and redundancy. Journal of Educational Psychology, 94(2), 428-434.

Dalton, D. W., Hannafin, M. J., & Hooper, S. (1989). Effects of individual and cooperative computer-assisted instruction on student performance and attitudes. Educational Technology Research and Development, 37(2), 15–24.

Dickson, S. V., Chard, D. J., & Simmons, D. C. (1993). An integrated reading/writing curriculum: A focus on scaffolding. LD Forum, 18(4), 12-16.

Dobrovolny, J. (2006). How adults learn from self-paced, technology-based corporate training: New focus for learners, new focus for designers. Distance education, 27(2), 155-170.

Ellis, N. C. (2002). Frequency effects in language processing. Studies in second language acquisition, 24(02), 143-188.

Evans, C., & Gibbons, N. (2007). The interactivity effect in multimedia learning. Computers & Education, 49(4), 1147–1160. doi:10.1016/j.compedu.2006.01.008.

Fletcher, J. D. (1996). Does this stuff work? Some findings from applications of technology to education and training. Paper presented at the Teacher Education and the Use of Technology-Based Learning Systems Conference, Warrenton, VA.

Gass, S. M. (2013). Input interaction and the second language learner. London: Routledge.

Herrell, A., & Jordan, M. (2004). Fifty strategies for teaching English language learners. (2nd ed.). Columbus, OH: Pearson.

Grgurović, M., Chapelle, C. A., & Shelley, M. C. (2013). A meta-analysis of effectiveness studies on computer technology-supported language learning. *ReCALL*, *25*(02), 165-198.

Hong, N. T. P. (2013). A dynamic usage-based approach to second language teaching. Unpublished doctoral dissertation, Rijksuniversiteit Groningen, Groningen.

Long, M. H. (1997). Construct validity in SLA research: A response to Firth and Wagner. The Modern Language Journal, 81(3), 318-323.

Kao, P. & Windeatt, S. (2014). Low-achieving language learners in self-directed multimedia environments: Transforming understanding. In J. Son, (Ed.), Computer-assisted language learning: Learners, teachers and tools (pp.1-19). (APACALL).

Krashen, S.D. (1981). Second Language Acquisition and Second Language Learning. Oxford: Pergamon.

Krashen, S. (1982). Principles and practice in second language acquisition. Oxford: Pergamon.

Krashen, S., & Terrell, T. (1983). The natural approach: Language acquisition in the classroom. Oxford: Pergamon.

Langacker, R.W. (1997). The contextual basis of cognitive semantics. In Jan Nuyts and Eric Pederson (eds.), Language and Conceptualization (pp. 229-252). Cambridge: Cambridge University Press.

Langacker, Ronald W. 2008a. Cognitive Grammar: A Basic Introduction. Oxford: Oxford University Press.

Larsen-Freeman, D. (1976). ESL teacher speech as input to the ESL learner. Work papers in Teaching English as a Second Language, 10, 45-50.

Long, M. H., & Doughty, C. (Eds.). (2003). *The handbook of second language acquisition* (pp. 487-536). Oxford, England: Blackwell.

Mayer, R. E. (2001). Multimedia learning. Cambridge: UK: Cambridge University Press.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language learning*, *50*(3), 417-528.

Nuttall, C. (2005).Teaching reading skills in a foreign language (2nd ed.). Oxford: Heinemann.

Pennington, M. (1996). When input becomes intake: Tracing the sources of teachers' attitude change. In: D. Freeman and J. Richards (Eds), Teacher learning in Language Teaching (pp. 320-348). Cambridge: Cambridge University Press.

Pezdek, K., Lehrer, A., & Simon, S. (1984). The relationship between reading and cognitive processing of television and radio. Child Development, 2072-2082.

Pinner, R. (2013). Authenticity of Purpose: CLIL as a way to bring meaning and motivation into EFL contexts. Asian EFL Journal, 15(4).

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Roberts, J.T. (1986). The use of dialogues in teaching transactional competence in foreign languages. ELT Documents 124: The practice of communicative teaching. Oxford: The British Council/Pergamon.

Rogers, C. V., & Medley, F. W. (1988). Language with a purpose: Using authentic materials in the foreign language classroom. Foreign Language Annals, 21, 467-478.

Schmitt, N. (2010). Researching vocabulary: A vocabulary research manual. London: Palgrave Macmillan.

Slavin, R.E. (2011). Instruction based on cooperative learning. In R. Mayer (Ed.), Handbook of research on learning and instruction. London: Taylor & Francis.

Soyemi, J., Ogunyinka, O. I., & Soyemi, O. B. (2011). Integrating self-paced e-learning with conventional classroom learning in Nigeria educational system. Proceedings of the 1st International Technology, Education and Environment Conference, pp. 700-707. African Society for Scientific Research (ASSR).

Strauss, A., & Corbin, J. M. (1990). Basics of qualitative research: Grounded theory procedures and techniques. Sage Publications, Inc.

Tomlinson, B. & Masuhara, H. (Eds.). (2010). Research for materials development in language learning. London: Continuum.

Vacca, J. (2008). Using scaffolding techniques to teach a social studies lesson about Buddha to sixth graders. Journal of Adolescent and Adult Literacy, 51(8), 652–658.

Van Patten, B. (ed.). (2004). Processing Instruction: Theory, Research, and Commentary. Mahwah, NJ, Erlbaum.

Van Patten, B., & Benati, A. (2010). Key terms in second language acquisition. London:

Vygotsky, L. S. (1978). Mind in Society: The Development of Higher Psychological Processes: Harvard University Press.

Wallace, C. (1992). Reading. Oxford: Oxford University Press.

Willis, J. (1996). A framework for task-based learning. Harlow: Longman.

Yi Liu, P. (2012). CALL programme for DUB approach. Unpublished master's thesis, University of Groningen. Groningen.

Zyzik, E. (2009). The role of input revisited: Nativist versus usage-based models. L2 Journal, 1, 42–61.

Appendix

General English proficiency and writing test

PART 1

Read the sentences about going camping. Circle the most suitable underlined word for each sentence.

1. They decided / thought / felt to go camping for their holiday.

2. They wanted to stand / put / stay somewhere near the sea.

3. It had / took / got three hours to bike to the campsite.

4. They put their tent in a center / corner / back of the field.

5. They bought / chose / sent some postcards to their friends.

PART 2

Read the sentences below. Circle the best underlined word(s) for each sentence.

6. Some writers can describe things when / that / if / who they have never seen.

7. Nothing changes / was changing / has changed / changed in this town since I first visited it.

8. Hurry up! They've got only a little / much / a few / little seats left.

9. Who is going to take care up / of / after / for the children while you're away?

10. A meeting will be run / taken / held / done to discuss the matter.

PART 3

Read the description of some jobs. Write the word for each one.

The first answer has been given as an example.

Job Description	Word for the Job	Answer
Example: *I help people to learn things*.	T_ _ _ _ _ _	Teacher
11. I show customers the menu and bring them their food.	W_ _ _ _ _	
12. People come to me when they are sick.	D_ _ _ _ _	
13. I will repair your car for you.	M_ _ _ _ _	
14. If you want to change the colour of your room, I will do it for you.	P_ _ _ _ _ _	
15. I help my boss by answering the phone, making appointments, and writing letters.	S _ _ _ _ _ _	

PART 4

Circle the word with the underlined part pronounced DIFFERENTLY from that of the other words in each set.

16. A. high   B. hour   C. house   D. home

17. A. nation   B. patience   C. cancer   D. basic

18. A. lov<u>es</u>   B. practic<u>es</u>   C. chang<u>es</u>   D. watch<u>es</u>

19. A. h<u>ea</u>vy   B. <u>ea</u>sy   C. w<u>ea</u>ther   D. h<u>ea</u>d

20. A. <u>ch</u>oose   B. <u>Ch</u>ristmas   C. ar<u>ch</u>itect   D. <u>ch</u>emistry

PART 5

There is one WRONG word in each sentence. Circle the word. Then correct it.

21. Ranil is a nice guy. He always say "hi" to everyone.…………………….

22. My younger brother has worked in a bank since a long time.…………………….

23. How was your holiday? – Great! We take a lot of photos.…………………….

24. I go to watch a movie with some of my friends last weekend.…………………….

25. She's staying with her family at the summer.…………………….

PART 6

Read the passage "Trade Secrets." Circle the correct answer to Question 26.

26. What did Swarnamali say about her mother's wedding ring?

    a.   Her mother lost her own wedding ring.

    b.   As a child, she exchanged the ring for a toy.

    c.   Someone stole the ring from her mother's car.

Trade Secrets

When I was little, my friends and I always traded things. So, one day, I traded this 'treasure' I had found in my mother's car for a plastic necklace, but it turned out that the 'treasure' was my mother's wedding ring! My mother has thought for years that someone stole her ring out of her car and has no idea that it was me! I've been keeping this secret ever since, and even though I'm grown up now, I still can't tell her the truth.

Swarnamali

PART 7

Read this postcard. Write ONE word for each space. The first space has been done as an example.

Dear Ranjith,

I'm sitting on the beach at the moment. Soon, I'm 27_____ to have a swim. I arrived here three days 28_____with my family. We 29_____be on holiday together here for two weeks. It 30 _____ a beautiful place. The beach is very near 31_____hotel. The sea isn't cold and 32_____ are many interesting places to visit. Yesterday we walked 33_____ a village in the mountains. I took lots 34_____ photographs. It's 35_____ pity that you didn't come 36_____ us.

Love,

Sheela

PART 8

What does Kim say to Anita? Complete the conversation. Write the word (A, B, C, D, E, or F) in each space.

Anita: Hello, how are you doing?	A.　　Bye, see you later.
Kim: Pretty good, and you?	
	B.　　That's exactly how I used to feel.
Anitaz: I'm doing great.	
	C.　　How do you like it so far?
Kim: **37**……………..	
	D.　　You don't like it?
Anita: So, how long have you been going to this University?	
	E.　　That's great to hear.
Kim: ³⁸……………..	
	F.　　I've been going here for a couple of years now. You?
Anita: This is my first year.	
Kim: **39**……………..	
Anita: It's OK, but not great.	
Kim: **40**……………..	
Anita: I'll like it better once I finish my General Education.	
Kim: **41**……………..	

PART 9-LISTENING

Read the passage carefully for two minutes. Listen and fill in the spaces with words you hear. There is one word for each space. You will listen 3 times.

I LOVE YOU

Love makes the world [42]_____ [43]_____, not money. I agree with the centuries-old quote that says, 'Love conquers all'. It's true when you [44]_____ [45]_____ [46]_____. So much has been written about love. It must be [47]_____ [48]_____ [49]_____ most written and talked about topics ever. How many songs and poems [50]_____ [51]_____ [52]_____ love? Millions. Billions, perhaps. Love is everywhere. You can't [53]_____ [54]_____ [55]_____ day without hearing someone say 'love'. It is one [56]_____ [57]_____ [58]_____ beautiful words in any language. Your heart can melt when someone says 'I love you'. It's also very important [59]_____ [60]_____ [61]_____ you love them. You should do it every day. There are many different kinds of love and they are all important. Except perhaps when you love pizza or burgers. That's [62]_____ [63]_____.

PART 10- WRITING

Choose ONE of the following topics. Write about that topic as much as possible.

1. My best friend

2. The most unforgettable day in my life

3. My goals and dreams for the future

4. My hometown