Integrating CALL with Analytical Rubrics for Developing Speaking Skills

Abdulaziz Alfehaid (aalfehaid@iau.edu.sa)
English Language Department, Deanship of Preparatory Year & Supporting Studies,
Imam Abdulrahman bin Faisal University, Saudi Arabia

Abstract

This paper reports on the effects of integrating CALL tools with analytical rubrics for developing speaking proficiency of EFL learners in Saudi Arabia. In an 8-week intervention program in the English Language Laboratory (ELL) both meaning and form aspects of speaking were taught with the help of an analytical scale that construed speaking on 5 analytical aspects namely content, vocabulary, grammar, pronunciation and fluency. Two groups of Preparatory Year Students (PYP), Experimental Group (EG = 44) and Controlled Group (CG = 32), took part in the intervention. Speaking tests were conducted for both groups before and after the intervention. In addition, a self-satisfaction survey was also administered. The study showed a significant effect of CALL integrated intervention on the experimental group's speaking proficiency on all aspects, while it showed significant improvements for only pronunciation and grammar aspects for the controlled group. The self-satisfaction survey results also provided motivating feedback from the participants for CALL integration.

Keywords: CALL, speaking proficiency, form-focussed, meaning-focussed, analytical rubrics, group discussion

Introduction

With its interactive and integrative computer technologies, CALL has become an integral part of everyday language instruction. Consequently, class-bound teaching has gone way beyond the four walls of instruction, introducing learners to sophisticated language learning tools (Ma, 2017) that can scaffold language learning (Sun, 2017; GroB & Wolff, 2001).

This paper reports on an attempt made to study the effects of technology-based instruction on students' speaking proficiency. Tools such as listening station, text parser, pronunciation analyser for accuracy and fluency and other online pedagogical resources such as the Corpus of Contemporary American English (COCA) that provide both authentic and compressible contexts of language use were used during the intervention. Two groups of preparatory year students were taught English over a period of 8 weeks, and their speaking performances were measured on a level-specific analytical rating scale (see the appendix) that mainly described *content, pronunciation, vocabulary, grammar,* and *fluency* aspects of speaking in the rubrics. Taking into consideration the course requirements, students' entry level proficiency, Common European Framework of Reference (CEFR) scales, and the requirements of the various

academic programmes at the university the rating scales were designed.

Review of literature

Technology mediated pedagogic tools such as online corpora and concordances (Bernardini, 2004), frequency analysis tools (Nation, 2001), internet-mediated platforms (Celik, 2013), online learner dictionaries (Hanks, 2006), and other online writing tutors (Lin & Priscilla, 2014) have positively influenced classroom practices (Neri, et al., 2002). Some of the 'modern learning systems' even integrated the world outside the language classroom without tampering the reality through web-based instruction. Corpus linguistic tools such as concordances (Ballance, 2017; Johns, 1991) have recreated in fragments a range of context- specific real life instances and exposed the learners to the patterns of language use (Hoey, 2005). A number of perception studies have also studied the impact of recasts, e-portfolios and audio-blogs in language learning (Alamri & Fawzi, 2016; Azar & Molavi, 2013; Baturay & Lu, 2010; Ayres, 2002; Hsu, 2008). Ayres (2002), for example, examined the face validity CALL based spelling, writing, and grammar practice tools used by 157 non-native undergraduate speakers, and found that they were effective in promoting quality engagement.

Many research studies have also provided concurrent evidence for pedagogic technology in foreign language teaching (Buckingham & Alpaslan, 2017). Indeed, some massive online projects (Cobb, 2012; Davies, 2013) have offered a voluntary service to users by letting them access the resources free of charge. In other words, computer-assisted and online-supported language learning resources are plenty; and it is only that we have to find a way to organize our instruction. This study has attempted to integrate the freely available CALL resources in developing the speaking proficiency of learners in an EFL context.

Creating meaningful interaction conditions by effectively organizing the content around advanced computer-mediated interactive tools requires an understanding of the constructs of the skills to be improved (GroB & Wolff, 2001). Speaking is a multi-faceted skill that requires the speakers to consider a range of aspects. Therefore, studies in teaching speaking skills are varied in their focus. Some studies have focussed exclusively on production aspects such as fluency, accuracy (Tavakoli, Campbell, & McCormack, 2016; Tavakoli, 2011; de Jong & Perfetti, 2011) and pronunciation (Smotrova, 2017), while others explored the pragmatic dimensions of speech (Bygate, 1998; Taguchi, 2006), where the role-relationships between the interlocutors and their cultural backgrounds interact with each other in meaning-making. Depending on the scope and research convenience, independent research studies and larger testing agencies have defined their constructs (Luoma, 2004).

Keeping in view the course specifications, real life challenges of speech, students' level of proficiency and various constructs, level specific analytical scales for speaking were designed. Assessment rubrics and level specific descriptors for content, grammar, vocabulary, pronunciation and fluency aspects were devised. The rubrics also provided guidelines for classroom instruction and materials development while assisting the course developers in designing their learning outcomes.

One of the significant aspects of the rubrics is the inclusion of speech contexts from various socio-cultural backgrounds into the curriculum. It emphasised the need to integrate non-native models of speech and promoted diversity in terms of understanding and production. Therefore, it was necessary to consider a synchronization real-life instances with specific form-specific practices.

From its inception CALL based pronunciation practice activities have attempted to engage the users with reliable and accurate native models for practice (Pennington, 1999). Recent studies, however, have provided evidence in support of using software technology that integrated automatic speech recognition tools (ASR), voice chats, virtual discussions, and smart phone applications in the language classrooms (Golonka *et al,* 2014). These tools enabled online access to a range of non-native speaker models of language use and facilitated a comparison between various speech models, both at word and text levels. Specifically, automated speech recognition tools tendered instant feedback and encouraged learners to improve their pronunciation through pattern matching and voice recognition software (Kim, 2006; Neri, *et al.*, 2002). Some of the tools scaffolded learner practice with textual and visual models and offered constructive feedback on the quality of pronunciation they achieved (Kim, 2006).

Personalizing grammar learning has become the norm of grammar instruction. Input as output models of learning have been overthrown, and learner-centred pedagogy is prioritized. CALL has offered learners, in this regard, with a range of possibilities—from ready-made gap-filling to frozen concordances to interactive feedback tools (Bernardini, 2004; Chuo, 2007; Francis, 1995). Online corpus resources such as the Corpus of Contemporary American English (COCA) and Word and Phrase. Info (Davies, 2013) and The Compleat Lextutor (Cobb, 2012) have become the most widely used online resources. Mark Davies's freely-accessible COCA was particularly designed to assist noticing the linguistic distinctions between various genres while organizing the lexico-grammatical patterns in multiple concordances. The statistical data about the occurrence of a unit of meaning across registers such as speech, news-paper, magazines and academic texts has enabled the users to distinguish between different discourse types (Davies, 2013).

Similarly, vocabulary teaching practices have enormously benefitted from CALL. Averil Coxhead's AWL (Coxhead, 2000), the Medical Academic Word List (Wang, Liang, & Ge, 2008) and the New Academic Vocabulary List (AVL) (Gardner & Davies, 2014) promoted computer mediation in identifying the common and most frequent words of use, while the *lextutor* and COCA have directly induced CALL in vocabulary instruction. Corpus tools such as Wordsmith tools (Scott, 1999), corpus-based graded readers (Nation, 2001) and corpus-based learner dictionaries (Hanks, 2012) are a few noteworthy examples. Indeed, much of the current vocabulary teaching and testing heavily depends on research findings of CALL and corpus studies.

Speech recording tools, such as the one used in this study, have allowed the learners to organize portfolios and compare their speech samples over a period of time. With the help of speech phrasing tools, students could analyse their speech samples for content and fluency aspects.

Level-specific rubrics for teaching and testing speaking

With the advent of computer-based corpus and discourse studies (Sinclair, 1991; Hyland, 2000) speaking research such as spoken corpora (McCarthy & Carter, 2001), authenticity of spoken materials (Sinclair, 2004), and genre approaches to language teaching (Swales, 1990) the priorities for language instruction changed (Carter, 1993; Gavioli, 2005; Johns,1991). Research started emphasizing the need to analyse language in its contexts of use, and promoted authentic materials in teaching (McCarthy & Carter, 2001). Similarly, Communicative Language Teaching (CLT) and Task-based approaches (Bygate, 1998) to speaking have provided comprehensive guidelines for the improvement of speaking proficiency. The model proposed by Goh and Burns (Goh & Burns, 2012) for instance, considered knowledge of language and discourse, core speaking skills, communication and discourse strategies as the key aspects in speaking development.

The conceptualization models of speaking development encompassed narrow to broader analytical categories that holistically constituted the core sub-skills of speaking (Jong et.al., 2012). Hinkel's (2006) review, for example, identified an integrative model with 'fluency, accuracy and a sufficient lexico-grammatical repertoire constituting the core (p: 114)', while researchers such as Robinson (2001) and Yuan and Ellis (2003) proposed cognitive models of speaking development from the point of view of fluency, complexity and accuracy, and emphasized the need to 'increase the load of cognitive complexity of speaking tasks which would result in eliciting 'greater lexical variation' (McCarthy & O'Keeffe, 2004) among other aspects of production.

Language testing researchers also aimed at determining the core constituents of speaking (Goh & Burns, 2012). While some testing services aimed at broader holistic descriptive scales, some evolved analytical scales in which fluency, accuracy, and pronunciation have become key aspects (Luoma, 2004). Sawaki's (2007) Language Ability Assessment System (LAAS) speaking component, for example, consisted of 5 analytic rating scales: *pronunciation, vocabulary, cohesion, organization*, and *grammar*. These scales were beneficial 'for student placement and diagnosis', and to link assessment with instruction. Similarly, an analytic rubrics with descriptors covering range, accuracy, fluency, interaction, and coherence aspects of proficiency was used by the Council of Europe (Council of Europe, 2001). Across a range of testing contexts, test developers have preferred to use analytical scales. These analytic scales also showed us the unique variations among the different aspects of proficiency (Sawaki, 2007).

Based on the general and specific research studies cited above, the following descriptive rubrics for teaching and testing purposes was designed and standardized.

Testing Speaking

To provide comfortable and confident speaking environments, where the rater's intimidating intervention was minimal, (Fulcher, 1996) group discussion formats were used. Research in speech functional analysis on using group discussion in the assessment of conversational ability presented positive evidence (Bonk & Ockey, 2003). They referred to Glen Fulcher's (Fulcher, 1996) comparison of picture-based discussions and text-based discussions with group

discussions and suggested 'group discussion format … might be appropriate for use with learners of lower levels of proficiency' (2003: 91). Based on these observations a group discussion format was considered suitable for this study.

Students were put into groups of 5 to 7 and were invited to discuss a topic that was closely related to the themes discussed in the textbooks. They were allowed to prepare for 3 to 4 minutes and take notes using the key words. Then, each student was given 2 to 3 minutes to present their views on the topic.

Research questions

This study attempts to address the following questions:

a.   To what extent does the CALL integrated instruction have an effect on students' speaking proficiency?

b.   How did students perceive their learning in the CALL integrated language lab?

Methodology

This study used both qualitative and quantitative methods to interpret the data collected from the students. While the sample for the study was primarily convenient, to obtain student satisfaction only those who attended a minimum of 5 sessions in the language lab were considered.

Each participant's speech was carefully analysed for the five aspects mentioned in the rating scale. Specific sub-skills such as hedging, using appropriate intonation, turn-taking, using context specific vocabulary, and rate of speech were considered in assigning scores. Aspect wise sub-skills observed by the raters are shown in Table 1.

Table 1: A summary of the speaking proficiency rubrics used for the study

Content	Topic comprehension, appropriateness of content, ease of sharing content, time-bound content restructuring
Vocabulary	Range of vocabulary, use of vocabulary appropriately, use of collocations

Grammar	Choice of grammatical structure (use of be forms, complex sentences…) , grammatical accuracy, self-correction strategies
Pronunciation	Phonetic accuracy, comprehensible pronunciation, word stress appropriate intonation
Fluency	Rate of speech, communication strategies (hesitation, hedging…, turn-taking), utterance length

The scores assigned by the raters across all the groups were compared with the reliability statistics obtained by the Quality Assurance Cell of IAU. Also, the inter-rater reliability analysis scores across the university tracks were found to be significant. The following provides the details of the sample, the tools, and the research procedure.

Sample and data

To group students into proficiency levels, Imam Abdulrahman Bin Faisal University (IAU) preparatory year program relies on the standardized English Placement Test (EPT) administered by the National Center for Assessment (NCA), Saudi Arabia, and a localized speaking test designed by the Department of English. For this study, two groups seventy-six students were identified randomly from a large pool of 17 groups. All students were between the age groups of 15 and 21 and at the intermediate level.

Both groups were offered 20 hours of in-class instruction per week. While the experimental group (N=44) was offered instruction in the language laboratory in two phases—10 hours of meaning-focused group instruction and 10 hours of personalized form-focussed practice, the controlled group (N = 32) was offered regular textbook based instruction. To understand the complex nature of speaking development, an analytical scale (see the appendix) was applied. A pre-test and post-test design was used to determine performance variations between the groups.

For question 'b', 135 student responses were collected with the help of a questionnaire. The questionnaire primarily focussed on the extent of satisfaction the stakeholders felt in using the language lab.

Research procedure

Research on non-native speaker contexts has given equal importance to interaction (Bygate, 1998; Gass, 1997) (meaning-focussed language production) and language analysis (form-focused language exploration) (Nattinger & DeCarrico, 1992; Carter, 1993; Doughty, 1991). Therefore, the teaching of speaking at the ELL was both group-based and personalized. While

group-based activities encouraged learner involvement and participation in meaning-making, personalized learning targeted learner exploration of language use (Carter, 1993). Organized into three inter-connected stages, every instructional session during the intervention followed a pre-designed model of intervention based on the key ideas proposed in research.

Figure 1: Intervention model for developing speaking skills at the ELL
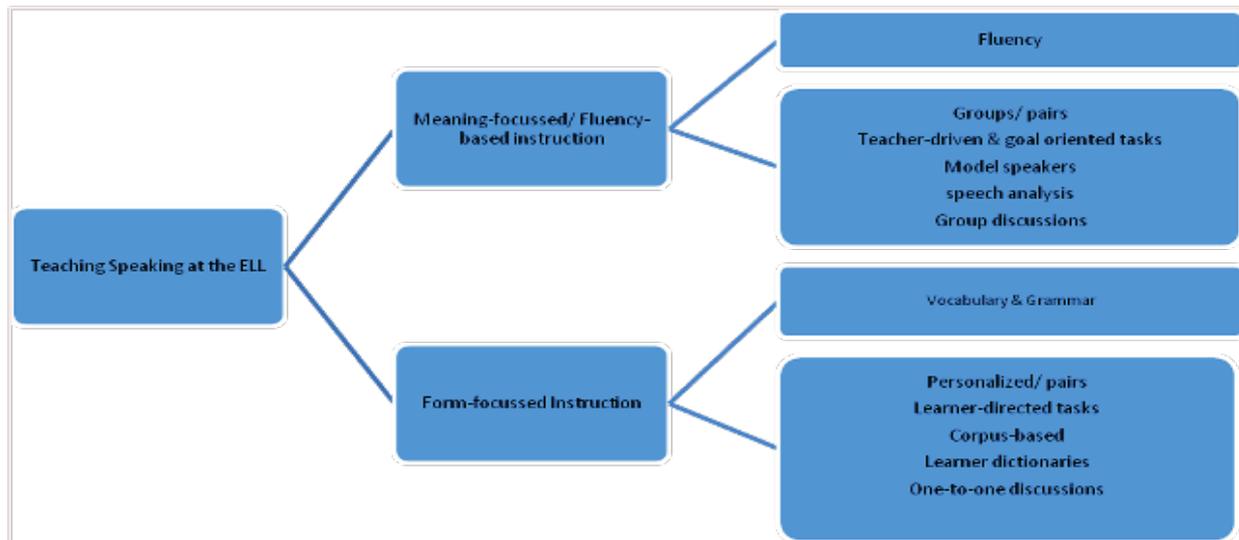
The first stage provided the participants with 'comprehensible input' through the *ampere* tools (ampere, 2011) which restructured the input to be accessed into sentences or timed sequences. Criteria such as complexity, speaker range, lexical density, contextual familiarity, discourse specificity, text length, and relevance were applied to determine the choice of the texts. Pacing of listening was determined by the individual students, and they were free to access a text more than once.

Figure 2: A meaning-focussed language learning context

The second stage was 'model learning': formed into groups of 3 to 7, the students created meaningful discourses, both as model speakers and as speakers who responded to the models. This stage was monitored by the instructor; and meaning-making was given importance. The third stage was called 'fluency practice'. During this stage, students were asked to practice extended conversations. This model of instruction aimed at providing opportunities for comprehensible input and meaningful output.

Personalized practice was as important as group-based practice. Focus of the personalized learning resource centre at the IAU was put on fostering learning through leaner exploration of language in use: learners were given access to the online resources such as COCA and BNC. Also, online vocabulary building activities and dictionaries were administered among other students to understand language use in specific contexts. Explorative practices mainly considered the lexico-grammatical patterns, collocations, use of academic vocabulary and the use of grammatical structures such as relative clauses, articles and transition words. Often their endeavours resulted in more concrete explorations that would enable them to restructure, assimilate and accommodate new patterns of language use.

Figure 3: Meaning-focussed and form-focussed intervention processes
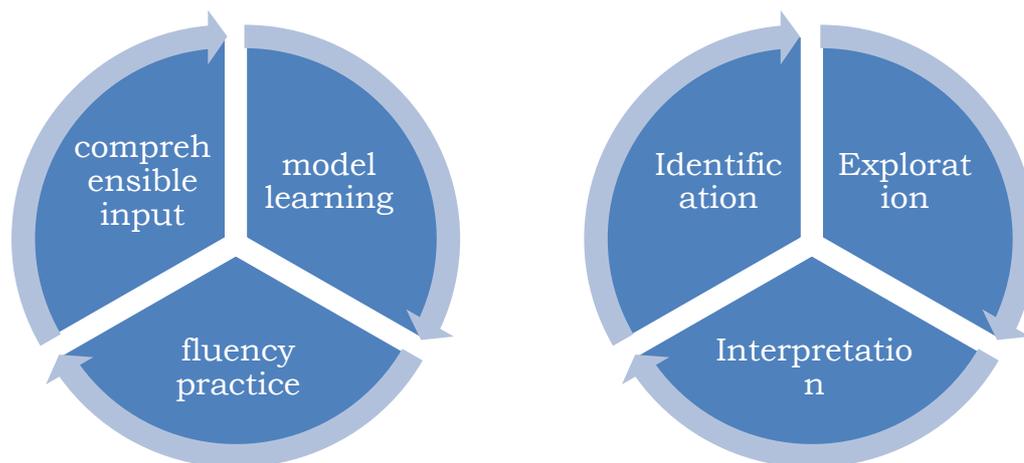
Students were guided through a set of pre-designed templates for exploring grammar and vocabulary aspects and were guided to record their generalizations and observations. The templates mainly guided students to look for the positioning, contexts, registers, and frequency of the search word/grammar category. For instance, if a student had set out a personal goal of

understanding the conditional or *if*-clauses, he was advised to look up the key word of the structure, "if", on the corpus. Similarly, for relative clauses, the use of relative pronouns, and for past perfect the use of "had" with a particle were also performed.

The model described above integrated both fluency and accuracy through meaning-focussed and form-focussed activities; and both the conscious form-focussed exploration and meaning-focussed fluency practice were given equal importance.

Findings and Discussion

The application of CALL for speaking here is methodical: both the meaning-focussed and form-focussed activities were used to help learners with their speaking skills. The focus of the activities was mainly on the five major aspects of speaking, namely content, vocabulary, grammar, pronunciation and fluency. The following outcomes discuss in detail the findings of the study.

Outcome 1

The rating scales clearly specified the grading process and guided the testers to make informed decisions. Two trained raters conducted group discussions on syllabus-based themes the students had already made themselves familiar with. In other words, the selected themes had high content validity as they were thematically analogous to the content of the course. Each student's response was objectively weighed against the rubrics, and at the end of the discussion, scores were assigned for each aspect. Table 2 and Table 3 provide the mean scores both the groups for each of the measured aspects of speaking.

Table 2: Mean scores and standard deviation of the five components of speaking proficiency of the experimental (EG)

Speaking Measures	EX- Pre	Std. Deviation	EX- Post	Std. Deviation
Content (10)	7.5568	.80128	8.0682	.94985*
Vocabulary (10)	6.9091	.67577	7.6136	.84126*
Grammar (10)	6.5909	.60302	7.2045	.85125*
Pronunciation (10)	6.6818	.60127	7.3049	.80531*
Fluency (10)	6.8523	.73595	7.5795	.81371*

Table 3: Mean and standard deviation of the five components of speaking proficiency of the controlled (CG) group

Speaking Measures	CON- Pre	Std. Deviation	CON - Post	Std. Deviation
Content (10)	7.1094	.81056	7.2500	.95038
Vocabulary(10)	6.8438	.68906	6.8594	.91787
Grammar (10)	6.1719	.65512	6.6406	.89112*
Pronunciation(10)	6.2969	.67033	6.5938	.99545*
Fluency(10)	6.5156	.71261	6.7188	.80259

Both the group's mean scores in the post test condition showed improvement in terms of securing higher grades. However, to determine whether the scores are significant, a paired sample and independent *t*-tests were administered.

The five aspects of speaking proficiency—*content*, *vocabulary*, *grammar*, *pronunciation* and *fluency*—were measured on a 10-point scale given. The scores assigned to each of the aspects along with the total score were recorded for both groups. Table 4 and table 5 below show whether the performance of the groups was significant in terms of the paired sample *t*-test.

Table 4: Experimental group's paired sample statistics of significance

Experimental Group Aspects		Mean	Std. Deviation			Sig. (2-tailed)
Pair 1	Pre_Content - Post_Content	-.51136-	1.03142	-3.289-	43	.002*
Pair 2	Pre_Vocabulary - Post_Vocabulary	-.70455-	.82348	-5.675-	43	.000*
Pair 3	Pre_Grammar - Post_Grammar	-.61364-	1.01651	-4.004-	43	.000*
Pair 4	Pre_Pronunciation - Post_Pronunciation	-.65909-	.81962	-5.334-	43	.000*
Pair 5	Pre_Fluency - Post_Fluency	-.72727-	.81021	-5.954-	43	.000*
Pair 6	Pre_Total - Post_Total	-.62045-	.48493	-8.487-	43	.000*

The paired sample statistics results showed a significant difference between the pre-test and post-test for all the five aspects. That is, *p*-value for all five aspects was recorded at 0.000 (2-taled) which is highly significant at 1% level. It indicated that the mean scores of vocabulary, grammar, pronunciation and fluency in the post-test condition were significantly higher than the pre-test.

Table 5 : Control group's paired sample statistics of significance

Controlled Group Aspects		Mean	Std. Deviation			Sig. (2-tailed)
Pair 1	Pre_content1 - Post_content1	-.14063-	.84466	-.942-	31	.354
Pair 2	Pre_vocabulary1 - Post_vocabulary1	-.01563-	.99583	-.089-	31	.930
Pair 3	Pre_Grammary1 - Post_Grammary1	-.46875-	.99139	-2.675-	31	.012*
Pair 4	Pre_Pronunciation1 - Post_Pronunciation1	-.29688-	.78143	-2.149-	31	.040*
Pair 5	Pre_Fluency1 - Post_Fluency1	-.20313-	.77104	-1.490-	31	.146
Pair 6	Pre_Total1 - Post_Total1	-.23438-	.44476	-2.981-	31	.006*

However, for the CG, the results are varied. The paired sample *t*-test of the pre- and post-test mean scores showed significant differences for grammar ($p < .012$) and pronunciation ($p < .040$) which probably seemed to influence the overall performance that showed the significance $p < .006$. That is, the *p*-values for content .354 (2-tailed), vocabulary .930, and fluency.146 were not significant.
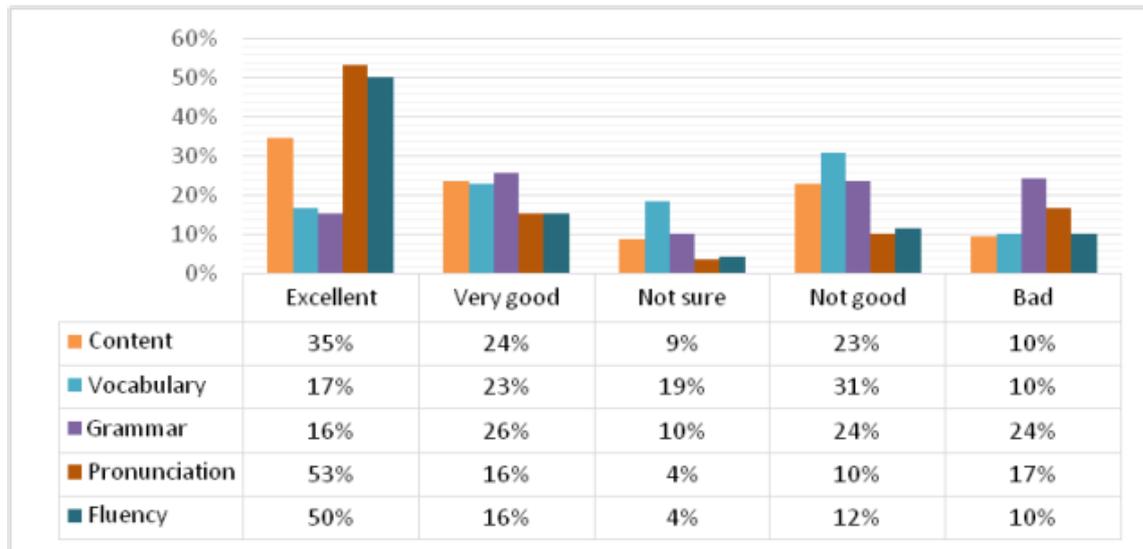
Outcome 2

Students' perception of learning

Intervention typically followed the sequence of fluency-based and meaning-focussed activities to form-focussed exploration of language use. In the first session, the participants were guided to focus on specific aspects of language use; and in the second session, they were guided to explore online pedagogic resources such as online corpora and dictionaries. During the intervention, questions pertaining to the facilities, learning activities, learner motivation, and overall satisfaction were asked to elicit open-ended responses of both the students and teachers. All student responses were elicited immediately after they had completed a session in the lab.

A cumulative analysis of student interviews revealed mixed opinions. Most participants reported that the tasks were at a higher level of complexity and required them to apply a number of skills, mainly, critical thinking skills such as analysing the responses, validating the position of the speakers, understanding the contextual factors in meaning and responding to prompts in a short time.

This study has indicated modest to highest levels of motivation from students in participating in lab activities. The figure 4 below provides information about the satisfaction levels of students for each of the five aspects of speaking.

Figure 4: Students' self-rated satisfaction for practice of the five aspects of speaking proficiency

	Excellent	Very good	Not sure	Not good	Bad
Content	35%	24%	9%	23%	10%
Vocabulary	17%	23%	19%	31%	10%
Grammar	16%	26%	10%	24%	24%
Pronunciation	53%	16%	4%	10%	17%
Fluency	50%	16%	4%	12%	10%

Respondents rated their level of satisfaction, for the five aspects of speaking, on a five point *Likert* scale where 1= Bad and 5= Excellent. Their responses, to a large extent, showed a higher level of satisfaction for pronunciation and fluency. Also, respondents pointed out that grammar and vocabulary were not addressed 'directly' in the language lab sessions. Practically, the EG participants were instructed to explore the corpus data and dictionaries based on their needs. The figure 5 plots student satisfaction of learning in the language lab for the four skills and the two aspects.

Figure 5: Trend line indicating the satisfaction of the respondents for LSRWV&G
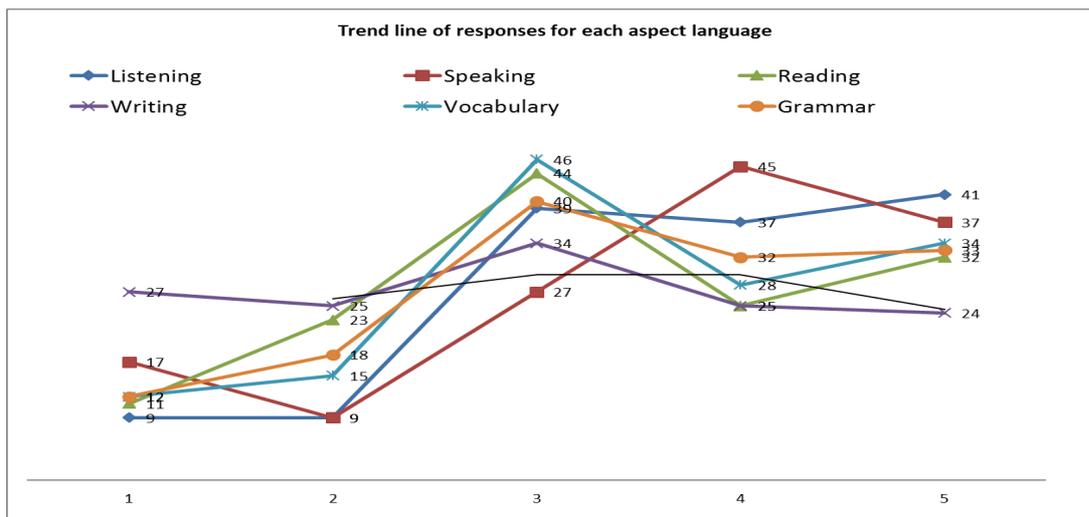
The trend-line indicated a high satisfaction rate for listening and speaking and a low satisfaction rate for writing. This could partially be attributed to the intervention methodology that

emphasised on speaking rather than writing. While the intervention integrated writing as a modality for production activities, it did not explicitly target it. Similarly, vocabulary and grammar aspects were the focus of implicit instruction and explicit practice: participants had to identify their language needs and explore the resources online on their own, based on the models for exploration and the sessions on integrating technology in language learning.

The participants shared mixed opinions about their practice in the language lab. While they expressed their satisfaction in attending the language lab sessions, they cited various reasons for their satisfaction. These reasons included noise-free environment, listening practice without any disturbance, self-paced practice, access to other multi-media resources, flexible tasks, choice of tasks, and ease of access to information among others. The following section provides a discussion of the data.

Meaning-focussed activities and their effect

The main purpose of meaning-focused intervention was to engage students in meaning-making and provide opportunities for interaction. Linguistic manifestations were mainly considered the by-products of meaning-making. Thus, higher mean scores for content and fluency could specifically be attributed to the use of meaning-focussed activities. As shown in the pictures in Figure 2, the experimental group took part in "interaction" driven tasks where the providing sound rationale/ advice/ suggestions/ reasons were prioritized over accurate production of sentences. Each participant had to adhere to the rule of "sound view" and contribute to the ongoing discussion on a topic.

With respect to the first question about the effects of the language lab-based instruction on students' speaking proficiency, the study showed significant improvement for the experimental group in all five categories. A fair amount of exposure to comprehensible input to multicultural content and the use of production/fluency activities could have caused this effect. The improvement on the five categories looks similar to the study on the effects of audio-blogs on learner's pronunciation by Hsu, Wang, & Comac, (2008).

Form-focussed activities and their effect

Form-focussed intervention was primarily individual specific. Each participant was asked to reflect on the meaning-focussed activity they took part in and identify the language resources they were required to cope with. The researcher here assisted the participants in formulating the search question to be looked up. The particpnats followed the path mentioned in the research procedure and explored the free online resources they had access to. Participants were set no upper limit to their search practices. While some preferred to look up words and their word-forms, others looked up corpus tools or dictionaries for patters and collocations. The following is an example of form-focussed exploration.

Figure 6: An excerpt of a learner exploration of online resources for form

While this study does not directly endorse the application of real life situations in understanding and acquiring languages (Beatty, 2010), it does seem to support the view that access to real life instances can influence the process of learning positively. The significance achieved by the EG for vocabulary and grammar aspects can be considered in this regard.

The act of looking up evidence in dictionaries and corpus data for specific linguistic resources that affected their performance in the fluency-based activities was a continuous process in the intervention. Depending on the performance on the meaning-focussed task, attempts were made to structure the thinking process of all students to a specific aspect of language. This seemed to reflect on their use of vocabulary and grammar in the speaking practice.

The second most important tool the students had access to, was the pronunciation practice tool where they could practice their speech both at word and sentence levels with the help of model speakers. The *ampere* tools could parse the texts at the word and sentence levels. It seemed to help students associate their speech with the models. In addition, online access to content specific videos especially TED talks, YouTube videos, and simultaneous reproduction activities of speech seemed to help students produce their speech with ease. The experimental group's rate of speech as well as pronunciation seemed to benefit from these tools and activities.

As mentioned, the access to the world outside the classroom (Chapelle, 2010) and the access to multimedia tools could have possibly brought about the change in participants' motivation.

179

While the controlled group seemed to show variation in its overall mean scores in the pre- and post-test conditions, (pre-test mean at 6.578 and post-test mean at 6.812), it is not as significant as the experimental group's test means (6.940 to 7.561).

The focus of the language lab tasks was more on problem solving and enabling students to analyse their speech. Around 69% of students expressed their extreme satisfaction about their practice of pronunciation while around 63% felt extremely satisfied about their fluency. Learners' perception CALL was found to have high face validity. Similar to the findings of Ayres (2002), this study did not provide any data which could correlate the success with specific software and tools. However, it could be considered as a limitation of the study. Longitudinal studies that capture the developments for all the five aspects mentioned over a period of time could also be helpful in designing appropriate materials for students.

Access to analytical sclaes and guided activities, freedom to pace learning, personalized learning environments, and speech analysis tools have each contributed to the holistic development of speaking. Tools such as sentence parsers, pronunciation practice, text analysis, learner dictionaries, corpus data, and voice recording, seemed to be successful in training the learners, hence, this study has provided further evidence in support of such learning environments. Most importantly, most of the tools used by the participants are available free of charge online. It is up to the language teacher to decide how the learning needs to be paced.

Conclusion

As discussed, research in speaking is diversified. Studies have probed into specific aspects such as cognitive processes (Vinthera, 2005), use of appropriate intonation, context specific grammar or vocabulary (Li, 2010), and rate of speech or fluency (Kessler, 2010). While development in one aspect can surely exert positive influence on the overall proficiency of the speaker, EFL contexts in the middle-east have not adequately studied the impact of specific innovations such as corpus and learning dictionaries on the overall improvement of speaking. The strength of this study is that it attempted to explore the impact on the holistic construct of speaking rather than any specific aspect, by using an indigenous analytical-scale of testing which we developed. Since the study categorically looked at development, it confined its exploration to finding significance rather than discussing at length or, qualitatively, the changes that occurred during the instructional period.

Research in the use of language labs and their effects on learners is limited and outdated. Moreover, much of speaking research singled out a specific aspect of speaking and studied the developmental path. In this study, an attempt has been made to study speaking as a whole with the help of an analytical scale and a standardized evaluation scheme. Studies have proved the use of scales in both teaching and assessing; and they have successfully categorized students into different proficiency levels. Here, by using standardized procedures of testing, we have looked at speaking proficiency development in a group of EFL learners. While studies that dwell on each of the analytical aspects is fruitful, studies of this type will provide the rationale for course developers to consider the potential of the resources available for instruction.

In light of the results of this study, it can be suggested that language learning programs that particularly depend on language analysis tools such as sentence parsers, speech analysis tools and pronunciation practice tools could promote better language skills among the EFL learners. EFL language programs may also benefit from the use of analytical scales that provide descriptions of the intended proficiency for the aspects of speaking. Analytical rubrics can help learners to self-evaluate their use of language while encouraging them to look for autonomous approaches to improve their skills.

References

Alamri, B., &Fawzi, H. H. (2016). Students' preferences and attitude toward oral error correction techniques at Yanbu university college, Saudi Arabia. *English Language Teaching, 9*(11), 59- 66.

ampere. (2011). *Learning spaces*: Operation Manual. ampere Inc.

Ayres, R. (2002). Learner attitudes towards the use of CALL. *Computer Assisted Language Learning*, 15(3), 241-249.

Azar, A., &Molavi, S. (2013). Iranian EFL learners' attitudes toward correction of oral errors. *The European Journal of Social and Behavioural Sciences*. doi:07/11/2016

Ballance, O. J. (2017). Pedagogical models of concordance use: correlations between concordance user preferences. *Computer Assisted Language Learning*, 30(3-4), 259-283.

Baturay, M. H., & lu, A. l. (2010). E-portfolio assessment in an online English language course. Computer Assisted Language Learning, 23(5), 413-428.

Beatty, K. (2010). *Teaching and Researching Computer-Assisted Language Learning – Second edition.* London: Longman Pearson.

Bernardini, S. (2004). Corpora in the classroom: an overview and some reflections on future developments. In J. M. Sinclair, *How to Use Corpora in Language Teaching* (pp. 15-38). Philadelphia: John Benjamins B.V.

Bonk, W. J., & Ockey, G. (2003). A many-facet rasch analysis of the second language group oral discussion task. *Language Testing, 20,* 89-110.

Buckingham, L., & Alpaslan, R. S. (2017). Promoting speaking proficiency and willingness to communicate in turkish young learners of english through asynchronous computer-mediated practice. *System, 65*, 25-   37.

Bygate, M. (1998). Theoretical perspectives on speaking. *Annual Review of Applied Linguistics, 12*, 20-42.

Carter, R. (1993). Language awareness and language teaching. In M. Hoey, *Data, Description, Discourse: Papers on the English Language in honour of John McH Sinclair* (pp. 139-150). London: Harper Collins Publishers Ltd.

Celik, S. (2013). Internet-assisted technologies for english language teaching in Turkish universities. *Computer Assisted Language Learning, 26*(5), 468 - 483.

Chapelle, C. A. (2010). The spread of computer-assisted language learning. *Language Teaching, 43*(1), 66-74.

Chuo, T.-W. I. (2007). The effects of the webquest writing instruction program on efl learners' writing performance, writing apprehension, and perception. *TESL-EJ, 11*(3).

Cobb, T. (2012). *The Compleat Lexical Tutor for Data driven Learning on the Web (v.6.2)*. Retrieved from Montreal University of Quebec: www.lextutor.ca/cgi-bin/range/texts/index.pl

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment.* Cambridge: CUP.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213-238.

Davies, M. (2013). Google scholar and coca-academic: two very different approaches to examining academic English. *Journal of English for Academic Purposes, 12*, 155 - 165.

de Jong, N., & A., P. C. (2011). Fluency training in the ESL classroom: An experimental study of fluency development and proceduralization. Language Learning, 61, 533 - 568.

Doughty, C. (1991). Second language instruction does make a difference: Evidence from an empirical study of SL relativization. *Studies in Second Language Acquisition*, 13, 431-469.

Francis, G. (1995). Corpus-driven grammar and its relevance to the learning of English in a cross-cultural situation. In A. P. (ed)., *English in Education: Multicultural perspectives.* Singapore: Unipress.

Fulcher, G. (1996). Testing tasks: issues in task design and the group oral. *Language Testing, 13*, 23-51.

Fulcher, G. (2003). *Testing Second Language Speaking.* London: Longman.

Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics, 35* (3), 305-327.

Gass, S. (1997). Input, interaction, and the second language learner. Mahwah: NJ: Lawrence Erlbaum.

Gavioli, L. (2005). *Exploring Corpora for ESP Learning.* Philadelphia: John Benjamins Publishing Co.

Goh, C. C., & Burns, A. (2012). *Teaching Speaking: A Holistic Approach.* Cambridge: CUP.

Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: a review of technology types and their effectiveness. *Computer Assisted Language Learning*, 27(1), 70-105.

GroB, A., & Wolff, D. (2001). A multimedia tool to develop learner autonomy. *Computer Assisted Language Learning, 14* (3-4), 233-249.

Hanks, P. (2012). The corpus revolution in lexicography. *International Journal of Lexicography, 25*(4), 398-436.

Hinkel, E. (2006). Current perspectives on teaching the four skills. *TESOL Quarterly, 40*(1), 109-131.

Hoey, M. (2005). Lexical Priming: A new theory of words and language. New York: Routledge.

Hui-Yin Hsu, Shiang-Kwei Wang & Linda Comac. (2008). Using audioblogs to assist English language learning: an investigationinto student perception. *Computer Assisted Language Learning, 21* (2), 181-   198.

Hyland, K. (2000). *Disciplinary Discourses.* Harlow: Pearson Education.

Johns, T. (1991). Should you be persuaded – two samples of data-driven learning materials. In T. Johns, & P. K. (Eds.), *Classroom Concordancing [ELR Journal, 4]* (pp. 1-16).

Jong, N. H., Steinel, M. P., Florijn, A. F., Schoonen, R., & Hilstijn, J. H. (2012). Facets of speaking proficiency. *Studies in Second Language Acquisition, 34*, 5-34.

Kessler, G. (2010). Fluency and anxiety in self-access speaking tasks: the influence of environment. *Computer Assisted Language Learning, 23* (4), 361-375.

Kim, I. -S. (2006). Automatic speech recognition: reliability and pedagogical implications for teaching pronunciation. *Educational Technology & Society, 9*(1), 322 - 334.

Li, J. (2010). Learning vocabulary via computer assisted scaffolding for text processing. *Computer Assisted Language Learning, 23* (3), 253-275.

Lin, S. M., & Priscilla, G. (2014). Impacts of online technology and use in second language writing: a review of the literature. *Reading Improvement*.

Luoma, S. (2004). *Assessing speaking.* Cambridge: Cambridge University Press.

Ma, Q. (2017). A multi-case study of university students' language-learning experience mediated by mobile technologies: a socio-cultural perspective. *Computer Assisted Language Learning*, 30(3-4), 183-203.

McCarthy, M., & Carter, R. (2001). Size isn't everything: spoken english, corpus, and the classroom. *TESOL Quarterly, 35*(2), 337-340.

McCarthy, M., & O'Keeffe, A. (2004). Research in the teaching of speaking. *Annual Review of Applied Linguistics, 24*, 26-43.

Nation, I. S. (2001). *Learning Vocabulary in Another Language.* Cambridge: Cambridge University Press.

Nattinger, J. & De Carrico, J. (1992). *Lexical Phrases and Language Teaching*. Oxford: OUP.

Neri, A., Cucchiarini, C., Strik, H., & Boves, L. (2002). The pedagogy-technology interface in computer assisted pronunciation training. *Computer Assisted Language Learning, 15*(5), 441-467.

Pennington, M. C. (1999). Computer-aided pronunciation pedagogy: promise, limitations, directions. *Computer Assisted Language Learning, 12*(5), 427–440.

Robinson, P. (2001). Task complexity, task difficulty, and task production: exploring interactions ina componential framework. *Applied Linguistics, 22*, 27-57.

Sawaki, Y. (2007). Construct validation of analytic rating scales in a speaking assessment:

reporting a score profile and a composite. *Language Testing, 24* (3), 355-390.

Scott, M. (1999). *Wordsmith Tools Users Help File.* Oxford: OUP.

http://lexically.net/downloads/version6/wordsmith6.pdf

Sinclair, J. (1991). *Corpu, Concordance and Collocation.* Oxford: Oxford University Press.

Sinclair, J. (2004). *Trust the Text*. London: Routledge.

Smotrova, T. (2017). Making pronunciation visible: gesture in teaching pronunciation. *TESOL Quarterly, 51*(1), 59- 89.

Sun, S. Y. (2017). Design for CALL – possible synergies between CALL and design for learning. *Computer Assisted Language Learning*, 30(6), 575-599.

Swales, J. M. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge: CUP.

Taguchi, N. (2006). Analysis of appropriateness in a speech act of request in l2 English. *Pragmatics, 16*(4), 513-33.

Tavakoli, P. (2011). Pausing patterns: differences between l2 learners and native speakers. *ELT Journal, 65*, 71-79.

Tavakoli, P., Campbell, C., & McCormack, J. (2016). Development of speech fluency over a short period of time: effects of pedagogic intervention. *TESOL Quarterly, 50*(2), 447-471.

Vinthera, J. (2005). Cognitive processes at work in CALL. *Computer Assisted Language Learning, 18*(4), 251-271.

Wang, J., Liang, S.-l., & Ge, G.-c. (2008). Establishment of a medical academic word list. *English for Specific Purposes, 27*, 442-458.

Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics, 24*, 1-27.

Appendix: Level specific rubrics designed and used in the study

	Content (10 Marks)	Vocabulary (10 Marks)	Grammar/ Accuracy (10 Marks)	Pronunciation (10 Marks)	Fluency (10 Marks)
8-10	Student response shows **mostly** comprehension of relevant lesson content. **Directly** addresses the test question. Able to communicate **easily** using relevant concepts from the lesson. Response fills the time allotted	Uses **a wide range** of vocabulary appropriate to the theme under discussion. **Nearly all** words and expressions used correctly. Uses relevant new words, expressions and collocations from textbook	**Minimal** grammatical errors. Speaker self-corrects without hesitation. Errors **never** interfere with communication	**Phonetically correct** words. Pronunciation **never** interferes with communication	Ideas are expressed with **natural pauses** and at a **natural speed**. **Minimal hesitation**. **Full** utterances. **Easily comprehensible**
6-8	Student response shows **significant** comprehension of relevant lesson content. **Mostly** addresses the test question. Shows **some limitation** in the range of ideas he/she can express on topic. Response shorter than time allotted	**Wide range** of appropriate and specific vocabulary. **Most** words and expressions are used correctly. Uses **many** new words and expressions from the textbook	**Few minor** grammatical errors that **rarely** interfere with communication. **Usually** self-corrects. **Very good** use of grammatical structures	**Phonetically correct** words. Pronounces **most but not all** words comprehensibly and appropriately	Ideas are expressed with **natural pauses** and at **a natural speed**. Speed of utterance **rarely** distracts the listener. Very little hesitation
4-6	Student response shows **some** comprehension of relevant lesson content. Addresses **some aspects of** the test question. **Main** ideas communicated are comprehensible. Response under half allotted time	**Fairly good** range of appropriate vocabulary. **Some** words and expressions are used correctly. Uses **several** new words and expressions from the textbook	**Several** grammatical errors interfere with communication. **Sometimes** self-corrects. **Fairly good** range of grammatical structures	Pronounces **many** words comprehensibly and appropriately. Pronunciation **often** interferes with communication	**Some hesitation**; searches for words. **Reasonable** speed only **sometimes** distracts the listener

2-4	Student response shows **little or no** comprehension of relevant lesson content. Shows **minimal comprehension of** the test question. **Several** ideas communicated are incomprehensible. Response under 10 sec	Uses a **basic** range of appropriate vocabulary. Uses **few** new words from relevant textbook units. Uses **many** vocabulary items incorrectly. The language produced **never** draws on lesson vocabulary	Uses a **basic range** of grammatical structures. **Very limited** self-correction. **Many** grammatical errors interfere with comprehension	Pronounces appropriately **only few** words. Pronunciation **consistently** interferes with communication	**Frequent hesitation**. **Very slow** delivery. **Incomplete utterances**
0-2	No attempt, or **incomprehensible**	No attempt, or **limited range** of vocabulary makes communication impossible. The language produced **never** draws on lesson vocabulary	No attempt, or **incomprehensible**. **Most** structures are incorrect	No attempt, or **many** pronunciation mistakes. **Many** utterances are incomprehensible	No attempt, or **constant** hesitation. **Incomprehensible**/broken delivery