International Journal of Humanities and Social Science Invention (IJHSSI)
ISSN (Online): 2319 – 7722, ISSN (Print): 2319 – 7714
www.ijhssi.org ||Volume 8 Issue 03 Ser. I || March. 2019 || PP.44-51

Developing Criteria for Evaluating CALL Software

Dr. Jaishree Umale
Associate Professor Amity School of Liberal Arts Amity University, Haryana

ABSTRACT: *CALL -- Computer Assisted Language Learning is widely used to learn a language. This has given rise to a plethora of software flooding the market. It becomes the teacher's responsibility to choose software which is able to meet the needs of the students. Checklists were used in order to judge software and it had sections like hardware information, software information, language information, summary information etc. This fixed format type of evaluation was rigid and had no scope of improvement. It was felt that theories of second language learning should be considered for evaluation; hence broadly it should focus on content, pedagogy and purpose. We had CALL software based on behaviorist theories of language, communicative approach and integrating various tasks and skills. All the approaches of evaluation viz. formative, summative, illuminative and integrative could not be found in the evaluation checklist. This paper tries to integrate all the aspects of evaluation and proposes a framework of evaluation integrating summative, formative, illuminative and integrative aspects which would be more relevant in assessing CALL software*.

Date of Submission: 20-02-2019 Date of acceptance:05-03-2019

I. INTRODUCTION

Teachers are required to evaluate their students and determine how far they have assimilated the required knowledge or skill. Educational technology has made tremendous progress and significantly transformed the teaching-learning process. With the advent of technology, teachers are also required to evaluate software student's use in the learning process. Newer software developed with excellent designs and animations do not guarantee their real worth. A thorough evaluation is necessary to determine its language learning potential. Many checklists developed for evaluation are based on methodology, technology and content but **tasks** have been ignored in the new light of SLA approach. At present, these checklists are used for holistic evaluation of the software for teaching or purchase. Research on technology, pedagogy and evaluation is revealing deeper insights that need to be integrated when developing criteria. This paper discusses and undertakes to develop criteria for evaluating software based on new researches in pedagogy and software evaluation. In section 2 literatures on CALL is reviewed, in section 3 CALL is discussed and in section 4 theory of evaluation is briefed. Based on the theories, criteria for evaluation are developed in section 5. Section 6 summarizes evaluation criteria based on different approaches. Section 7 gives the conclusion.

II. LITERATURE REVIEW

The history of CALL is very recent. It caught momentum in 1980's and lot of software emerged for educational purposes. In order to judge software, checklists were used. Strei (1983) developed six sections in his list**: title and ordering**, **target population**, **hardware information**, **software information**, **language information** and **summary information**. Each section had a series of items to tick on. Hertz (1984) also prepared yes/no type of checklist covering, general **purchase information**, **level and suitability of content**, **compatibility with text books teaching capabilities**, **hardware/ software functions** and **feedback**. Decoo (1984) criticized these fixed format questionnaires as they led to problems of **accuracy**, **compatibility** and **transferability.** The fixed format offered no scope for suggestions to improve hence, free format evaluation was recommended. Hubbard (1988) designed a framework for evaluation based on Richards and Roger's work (1986). It analyzed language teaching methods in terms of **approach, design** and **procedure**. During eighties checklists were also made by many CALL organizations like MicroSIFT, EPIE, MECC and National council of teachers of English. Heller (1991) discovered that all these checklists revolved around **technology, content** and **pedagogic** presentation**.**

In nineties, many researchers designed exercises, helping students develop language skills. Salazar (1989) suggested ideas for listening comprehension activities. Kim (1991) described the nature of effective instructional methodology and multimedia courseware design. Hutchings et al. (1992) developed evaluation of hypermedia modes of learning and instruction. Thorn (1995) proposed to include SLA to represent existing methodologies along with other areas of evaluation.

TILT project evaluation group stressed on **integrative** evaluation. Draper (1997) hinted at illuminative evaluation. Chun and Plass (1997) focused on learners integrating visual and verbal information. Chapelle (1997, 2001) discussed **judgmental** and **empirical** evaluation and prepared criteria for evaluating tasks by applying theories of SLA. Lee (2001) apart from design and methodology also discussed teacher readiness, financial concerns and after sales service.
Organizations like Technology 2002, BECTA, and ICT4LT project focused on content, pedagogy and purpose. They also included cultural insights, relevance to study, learning objectives etc.

III. COMPUTER ASSISTED LANGUAGE LEARNING (CALL)

CALL is an acronym for Computer Assisted Language Learning. Computers have been used in the past for repetitive drills and a tool to practice grammar. Now with the Internet it has endless possibilities. It has gone through the following three stages: (Warschauer and Healey, 1998)

Behaviouristic CALL,
It was based on behaviouristic theories of language learning and featured repetitive drills.

Communicative CALL
It stressed on using the language rather than focus on forms. Students were encouraged to generate original utterances. It supported cognitive theory and communicative approach.

Integrative CALL
It sought to integrate learners into authentic environments and various skills of language learning. It concentrated on task based, project-based and content based approaches.

CALL software flourished during this period as it could deliver contents for skill practice. Teachers relied on generic programmes to customize lessons and exercises for students and delivered them through the net. Other new forms like virtual reality environments, web quests, micro worlds etc. emerged stressing on collaborative learning. The potential is great and it is still evolving into newer forms.

IV. WHAT IS EVALUATION?

Evaluation is making judgments. We are always evaluating people, places etc. It is a natural part of our social life (Rea-Dickens and Germaine, 1992:4). Evaluation in education is a systematic attempt to gather information to make judgments (Lynch, 1996:2). These judgments are based on definite criteria.

4.1 Methods of Evaluation

The method of evaluation depends on what is to be evaluated. There is no direct relationship between the subject of evaluation and the method employed. (Alderson,1992: 283). **Informal evaluation** is done casually while **formal evaluation** is carried out against criteria (Heller, 1985). Analysis is done at various levels e.g. through arguments (argumentative), judgments (judgmental) and research (empirical), (Chapelle, 2001). General methods of evaluation are:

a. Quantitative evaluation
It uses standardized framework to collect data that can be presented in statistical terms.

b. Qualitative evaluation
It is used in a complex situation and includes ethnographic description or people's opinion.

c. Written methods of evaluation
It includes - tests, tick and fill evaluation forms, questionnaires etc**.**

4.2 Approaches to Evaluation

The aim of evaluation in general is to influence decision making through the provision of feedback. The approaches depend upon the aim (Draper et al. 1995). They are:

4.2.1 Formative evaluation

It is the assessment of the software with a view to modify the software in case problems are detected before its release. It entails suggestions for improvement.

4.2.2 Summative evaluation

It focuses on the product. It is the summary of the performance of the software after implementation.

4.2.3 Illuminative evaluation
It relies on observation. It discovers factors and issues that are important to the participants in that particular situation e.g. wrong answers deliberately chosen by students as it offers more involvement.

4.2.4 Integrative evaluation
It is aimed at overcoming problems encountered in CALL software by utilizing supplementary material provided by the teacher (Draper et al.1995). .

V. CRITERIA DEVELOPED FOR EVALUATION

The criteria suggested below adopt Hubbard's framework approach for its flexibility and capacity to accommodate new components. It also includes Chapelle's (2001) judgmental evaluation of the appropriateness of the tasks in SLA and new approaches of evaluation discussed by Draper et al (1995). The framework rests on the following principles adapted from Hubbard (1988) which determine form and content.

- Evaluation framework should give general information about the software along with its technical details.
- It should focus on language teaching methodology, SLA research and innovative research in evaluation.
- Evaluative framework should be flexible and accommodate newer developments in related fields.
- The framework should describe multiple dependencies for CALL components.

The major sections of the framework are given below: (See figure 1)

- General information
- Technological considerations
- Pedagogical considerations
- Evaluative considerations

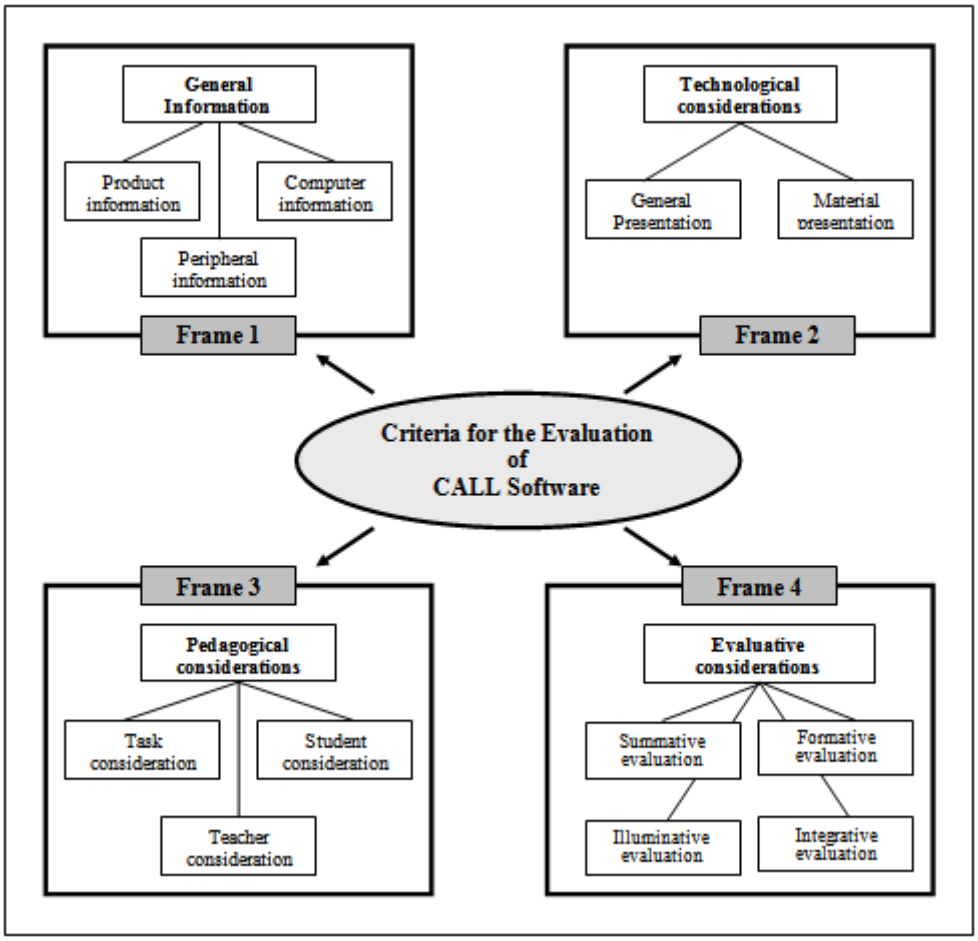

Figure1. Criteria for the development of CALL software

5.1General Information
This frame provides general information and is divided into: Product information, Computer information and Peripheral information.

5.1.1 Product information

It gives information under the following heads:

- **Product name**
- **Product type**: informs whether the software is for grammar, writing, reading etc.
- **Purpose:** gives the aim of the software
- **Media type**: gives information about the kind of software e.g. CDROM, video etc. or for a single user, lab workstation etc.
- **Level**: informs levels -- elementary, intermediate, advanced etc.
- **Activities:** states the type of activities used for language learning i.e. game, quiz, text reconstruction, text construction, problem solving and exploratory (Phillips, 1985 cited in Hubbard, 1988).

5.1.2 Computer information

This section gives information about the operating system--- Windows or Macintosh, hardware requirements, RAM, CDROM, sound and video.

5.1.3 Peripheral information

It informs about the supplementary materials like manuals, tutorials, online help, price and license.

5.2 Technological Considerations

This frame highlights the technological aspects of the software-- General presentation and Material presentation.

5.2.1 General presentation

It deals with the presentation of the material on the screen. It includes the use of colour, text and layout, use of hypertext, screen layout, use and quality of graphics, animation and video.

5.2.2 Material presentation

It deals with the presentation of the teaching material and the feedback students receive. It includes the following criteria: (Hubbard, 1988)

- **Menus and icons:** Do menus provide clear interface and various options?
- **Interaction and feedback:** Does the software reinforce correct answer with positive feedback, supply correct response, evaluate the student's answer, give explanation for incorrect answers, encourage students to obtain correct response and recommend special remediation to students. Does it give feedback on scoring for individual exercise and across exercises cumulatively (Bangert – Downs and Kozma, 1989)
- **Treatment of errors:** Does it cope with errors of format e.g. capitalization. Can the users get correct answers to move on? Is the feedback given for incorrect responses?
- **Control:** There are three potential controllers: the programme, the teacher and the learner. In a menu driven programme the student can control by choosing an activity. The teacher can control by imposing options and the programme can restrict the number of tasks to be completed (Hubbard, 1988).
- **Help options:** allow a learner to exit an exercise temporarily, view all parts of the screen and return back again.
- **Hints and clues:** in the programme help the learner complete some aspect of the activity.
- **Timings:** It is the rate at which information or stimuli appear on the screen, the amount of time the image is allowed to appear on the screen or students allowed time to respond (Hubbard, 1988).
- **Individualization:** Does the software adapt to the learner's level and ask questions according to his response? Does it allow branching or following different routes and options?
- **Record keeping:** Does it keep permanent record of the learner performance. Does it allow students to begin from they had left? Can the students work from any workstations and access records on the net?
- **Security:** Are the files secure from being tampered or accidental loss.

5.3 Pedagogical considerations

It evaluates the worth of software on the basis of: Task considerations, Teacher considerations and Student considerations.

5.3.1 Task considerations

This evaluation is based on Chapelle's (2001) judgmental evaluation of the appropriateness of the tasks. She suggests that tasks employed should result in language learning. The criteria are given below:

:

- **Language learning potential**

The task should be a language learning activity rather than an opportunity for language use. It should promote focus on form and include interactional modifications, modification of output, time pressure, modality, support etc. (Skehan 1998, cited in Chapelle 2001: 55)

- **Learner fit**

It takes into account individual differences in linguistic ability levels and non-linguistic characteristics. Learners should be given a range of target structures appropriate to their level such as willingness to communicate, age, learning style etc.

- **Meaning focus**

Learner's primary attention is directed towards meaning of the language to accomplish the task like exchanging information, making decisions etc.

- **Authenticity**

This includes the degree of correspondence between an L2 learning task and tasks that a learner is likely to encounter outside a classroom.

- **Positive impact**

The tasks should carry effects beyond language learning potential e.g. helping learners develop their Meta cognitive strategies, engage interest in target culture, help gain pragmatic abilities etc.

- **Practicality**

It refers to how easy it is for the learners and teachers to implement CALL task in the class/ language programme.

5.3.2 Teacher Considerations

It dwells on the language teaching approach and the way the computer delivers it (Hubbard, 1988)

A. Language Teaching Approach

The approach adopted is generally one out of three - **Behaviourist, Explicit/Cognitive and Acquisition/Communicative**. They have different characteristics and also differ in the treatment of errors.
These methods also discuss sequencing, role of explanations, treatment of learner differences, linguistic competence, the role of culture, motivation etc.

Language teaching approach also entails:

- **Language learning assumptions**

It is the impact of age, motivation, first language interference and other aspects of language processing. It also considers external environments like teachers, classrooms, texts etc.

- **Linguistic assumptions**

It analyzes the nature of language structure and function based on linguistic theory e.g. roles of rules, patterns, analogy in grammar and lexicon, relationship of form and function etc.

B. Computer Delivery System

Criteria for evaluation can be developed based on understanding the abilities and limitations of the computer along with the generalizations of language teaching approach e.g. criteria for explicit learning can be:

- Does the computer give meaningful rather than mechanical practice?
- Does it provide hints to lead students to correct answers?

5.3.3 Student considerations

It determines the nature of the learning environment by considering Learner variables and Syllabus (Hubbard 1988).

A. Learner Variables

They are age, native language, proficiency level, sex, learner needs, learner interests, intrinsic or extrinsic motivation etc. (Reeves, electronic article).
The following also contribute towards learner variables:

- **Learning style**

It is the perspective of learner preferences and builds on recognition, recall, comprehension, experiential learning and constructive understanding (Kemmis, 1977 in Hubbard 1988).

- **Classroom management**

It refers to the grouping of students and a teacher's involvement in controlling learners' activities. It also considers record keeping, keeping scores and time spent on specific activities.

- **Language difficulty**

Texts that are longer and more varied are difficult. Familiar and transparent texts are not difficult.

B. Syllabus

The syllabus can be structural, situational, notional-functional and content centered. It also focuses on the skills dealing with reading, writing etc. or practicing sub-skills and strategies for more effective processing, retention or production.

Other considerations are whether CALL fits into the syllabus, is it an integral part, independent part or an option?

The following also contribute towards learner variables:

- **Programme focus**

It is the linguistic objective of the activity like discourse, text, syntax, lexis, morphology, phonology etc. Programme focused at any level also involves practice at lower level as well.

- **Learner focus**

It is built around listening, speaking, reading and writing. It includes whether the software was aimed at building basic linguistic knowledge through the medium of one of these skills
(E.g. grammar, vocabulary) or practicing sub-skills.

- **Content**

The content fits the needs and interests of the learners as well as the limitations and pedagogical thrust of the syllabus e.g.

a. Is the presentation interesting and engage the attention of the students?
b. Are registers used appropriate?

VI. EVALUATION BASED ON DIFFERENT APPROACHES

Evaluation is done for different purposes but a teacher evaluating software should include illuminative as well as integrative evaluation for best results as shown below:

6.1 Formative evaluation

Concrete suggestion for the development of the software should be given.

6.2 Summative evaluation

Each criterion is given points on Likert rating scale of 1-5. **Weightage factor** of each criterion is **ascertained according to its importance in language learning**. Multiplication of rated score and weightage factor gives points scored. Maximum score for each criterion is multiplication of maximum scale rating i.e. 5 by its weightage. The total sore is taken out in percentage which gives the quantitative summative evaluation of the software (Appendix 1).

6.3 Illuminative evaluation

When software is being evaluated, the surprises encountered on close observation should be noted.

6.4 Integrative evaluation

It gives suggestion to overcome the existing deficiencies to make the best use of the software (Draper et al. 1995).

VII. CONCLUSION

It is necessary for teachers to evaluate CALL software to ascertain its language learning potential. Special attentions should be given to the task employed as it should result in language learning. While evaluating software the evaluator should offer suggestions regarding problems faced, bugs or related new

discovery. Suggestions should also be given to overcome problems to make the best use of the software. The framework type of evaluation is flexible enough to suit the requirements of a teacher. The present evaluation criteria lack **social characteristics** (Heller, 1985) such as presence or absence of competition, cooperation, humanization of the computer, moral issues and value judgments. The evaluative process is too subjective. The terms and categories are not carefully described hence; consistent evaluation of the same software cannot be obtained using the same form. Open-ended responses pose a problem as reviewers are unlikely to use the same words to describe software. Hence, making comparison is difficult. In future research should focus on various dependencies within a framework their degree of interrelatedness and the relative impact they have on language learning. New research should also focus on standardizing the criteria for CALL software.

REFERENCES

[1]. Alderson, J.C. (1992). Guidelines for the evaluation of language education. In Alderson, J.C. and Beretta, A. (Eds). Evaluating Second Language Education. Cambridge: Cambridge University Press.

[2]. Bangert – Downs, R. L. and Kozma, R. B. (1989). Assessing the design of instructional Software. Journal of Research in Computing in Education, Vol. 21(3) pp. 225-31

[3]. Chapelle, C. (1997). CALL in the year 2000: Still in search of research paradigms. Language Learning and Technology, 1/1, pp. 19-43

[4]. Chapelle, C. (2001). Computer applications in second language acquisition. UK: Cambridge University Press.

[5]. Chun, D. M. and Plass J. L. (1997). Research on text comprehension in multimedia environments. Language Learning and Technology, 1/1, pp. 60-81

[6]. Decoo, W. (1984). An application of didactic criteria to courseware evaluation. CALICO Journal, 2(2):42-46

[7]. Draper et al. (1995). Integrative evaluation: An emerging role for classroom studies of CAL. Available: http:// www.psy.gla.ac.uk/~steve/IE.html (accessed on 30th Dec 2005).

[8]. Draper, S. W. (1997). The role of open ended observation in evaluation. Available: http://www.psy.gla.ac.uk/~steve/oeo.html (accessed on 30th Dec 2005)

[9]. Heller, R. S. (1991), Evaluating software: A review of the options. Computers and Education, 17/4:285-91

[10]. Hertz, R. M. (1984). Software evaluation guide for the language arts. CALICO Journal, 1/5, pp 21-23

[11]. Hubbard, P. L. (1988). An integrated framework for CALL courseware evaluation. CALICO Journal, 6/2, pp 51-72.

[12]. Hutchings, G. Hall, W. Briggs, J. Hammond, N. Kirby, M. McKnight, C. and Riley, D. (1992). Authoring and evaluation of hypermedia for education. Computers and Education, 18/1: 171-77

[13]. Kim, W-C.M. (1991). Conceptual design of an intelligent hypermedia language learning environment. CALICO Journal, Vol. 8, No 4, pp. 25-36

[14]. Lee, K. C. (2001). Selecting and integrating CALL software programs into the EFL classrooms. Available: at http:// elc.polyu.edu.hk/conference/papers 2001/lee.htm. Accessed on 10th January 2006 last updated 21st Nov 1997

[15]. Lynch, B. (1996). Language program evaluation. Cambridge: Cambridge University Press.

[16]. Rea – Dickens, P. and Germaine, K. (1992). Evaluation. Oxford: Oxford University Press.

[17]. Reeve, T. Evaluating what really matters in computer based education. Available: http://www. Educationau.edu.au/ archives/CP/reeves.htm. Accessed on 10th January 2006 last updated 21st Nov 1997

[18]. Richards, J. and Rodgers, T. (1986). Approaches and methods in Language Teaching: A description and analysis. New York: Cambridge University Press.

[19]. Salazar, J. (1989). Interactive audio strategies for developing listening skills. CALICO Journal, 7/1:67-75.

[20]. Strei, G. (1983), Format for the evaluation of courseware used in computer – assisted language Instruction (CALL). CALICO Journal, 1/2:43-46.

[21]. Thorn, W. (1995). Points to consider when evaluating interactive multimedia. The Internet TESL Journal, 2(4) April 1995 also at http: // iteslj.org/Articles/Thorn-EvalueConsider.html

[22]. Warschauer, M. and Healey, D. (1998). Computers and language learning: An Overview. Language Teaching, 31, 57-71

Appendix 1- Evaluation Criteria

Summative Evaluation						
SN	**Criterion**	**Rating scale score (1-5)**	**Weightage**	**Points scored**	**Max. score**	**Remarks**
A.	**General Information**					
1	Product information		1		5	
2	Computer information		1		5	
3	Peripheral information		1		5	
	Total				**15**	
B1	**Technological (General)**					
1	Installation		1		5	
2	Colour		1		5	
3	Text and Layout		2		10	
4	Use of hypertext		1		5	
5	Sound		2		10	
6	Graphics		2		10	
7	Animation/video		2		10	
	Total				**55**	
B2	**Technological (Material Presentation)**					
1	Study material Presentation		3		15	
2	Menus and icons		2		10	

3	Interaction		3		15	
4	Feedback		3		15	
5	Treatment of Errors		3		15	
6	Control		3		15	
7	Help options		3		15	
8	Hints and clues		3		15	
9	Timings		1		5	
10	Individualization		3		15	
11	Record keeping		2		10	
12	Security		2		10	
	Total				**155**	
C	**Pedagogical Considerations Task Considerations**					
1	Language learning potential		4		**20**	
2	Learner fit		3		**15**	
3	Meaning focus		3		**15**	
4	Authenticity		2		**10**	
5	Positive impact		2		**10**	
6	Practicality		1		**5**	
	Teacher consideration					
7	Language teaching approach		4		20	
8	Computer Delivery system		3		15	
	Student consideration					
9	Learner variables		4		20	
10	Syllabus		4		20	
	Total				**150**	
	Grand Total				**375**	
Other Evaluation		**Remarks and Suggestions**				
Formative						
Illuminative						
Integrative						

Dr. Jaishree Umale" Developing Criteria for Evaluating CALL Software" International Journal of Humanities and Social Science Invention (IJHSSI), vol. 08, no. 3, 2019, pp. 44-51