ASSESSING LIVEMOCHA AND DUOLINGO EVALUATION TASKS: ADAPTING CHAPELLE CALL ASSESSMENT CRITERIA IN AUTONOMOUS LEANING EXPERIENCE

Irham
Universitas Islam Negeri Maulana Malik Ibrahim, Malang, Indonesia
Email: irhamaladist@gmail.com

ABSTRACT
Computer assisted language learning has been the major issue in the recent decades. Several free networking sites focusing on language learning are offered abundantly. Nonetheless, most of the extant studies seem to compare the traditional way of learning language and modern one per se. This paper, on the other hand, will focus on task evaluation of CALL system as proposed by Chapelle (2001). Six criteria of CALL assessments; language learning potential, learner fit, authenticity, meaning focus, positive impact, and practicality, will be employed to delve Livemocha and Duolingo’s language learning task on Dutch course for basic level. The researcher here plays role as the key instrument in which the data were gained through the experience of using Livemocha and Duolingo. The Dutch, as foreign language, is chosen as the subject to be studied. As the results, it is intriguing that Livemocha’ task for Dutch course is not adequately fit to beginner leaners. Meanwhile, its counterpart questionably complies authenticity point. The beginner learners may efficiently make use of computer assisted language learning tool to foster their study on foreign language. Nonetheless, they have to be aware of which tool can accelerate their purpose better.

KEY WORDS
CALL assessment, Livemocha, Duolingo, language learning

INTRODUCTION
The foreign language learning process has developed rapidly as the advance technology progresses. The extant studies, however, seem to busily compare the traditional way of learning language, that is a learning method without an assistance of technology, and the modern one that utilizes a technology in language learning (Dunkel, 1991). Some other studies, nevertheless, pursue more specifically to particular case of computer assisted language learning such as the use of Automatic Speech Recognition (ASR) and feedback giving for productive skills (Cucchiarini, Neri, Strik, 2009). Ehsani & Kyodt (1998) even argue upon the drawbacks of CALL that is still lack of theoretical framework to design and evaluate CALL itself. To that end, this paper does not intend to extend the discussion on comparative method of language learning nor to establish a rigid theoretical foundation for evaluating CALL. These topics will remain controversy and may develop simultaneously with the refinement of technology and SLA in pedagogical research.

The present study attempts to investigate two most used social networking sites in the language learning; Duolingo and Livemocha, on how the assessment or evaluation on CALL should be put into considerations. These language learning platforms have both web and app format. They offer users enormous opportunities to study foreign languages that are designed into several stages and are assigned to newbie or advanced learners. What makes them mostly chosen is the free and easy access since internet is adequately available in most places. Besides, these CALLs are also downloadable in play store so the users can make use android or tablet to enjoy studying foreign language. In addition, their prominent testing features such as automatic speech recognition (ASR) to practice or test speaking, community feedback for
writing skill, and interesting media display during the learning process are other respectable consideration that comes to the mind of the users in selecting them.

The theoretical framework on CALL assessment will be based on Chapelle’s (2001) criteria. He offers six criteria on evaluating CALL task as a system. Those are comprised of language learning potential, meaning focus, learner fit, authenticity, impact, and practicality. This framework has been applied in Hincks’ (2003) analysis on Talk to Me. The study concluded that it is less fit in practical implication due to its priced cost.

It is therefore salient to further figure out the existing free applications in regard with language learning such as Livemocha and Duolingo. Since Talk to Me is concluded to be less effective for its high cost, the latter two mentioned applications may then fit those language learners from developing countries, or those who are not capable to purchase the priced application. Besides, we can also hypothesize that these applications can be very helpful for language learners who want to understand the progress of their study. Livemocha and Duolingo in this case provide ‘evaluation’ feature that can assist them to see the progress of their study. Thus, the present study is going to specifically see the efficacy of these applications for new learners of foreign language.

**REVIEW OF RELATED LITERATURE**

**Computer Assisted Language Learning**

Before further elaborating what Livemocha and Duolingo really are, it is worth noting to present a brief definition of what is meant by CALL in several papers. Beatty (2003: 7) defines CALL as “any process in which a learner uses a computer and as a result improves his or her language”. This broad definition enables us to assume that all computers that give us avenue to study language, such as mobile phone, PC, laptop, etc., albeit not specifically designed for language learning, can be considered as CALL. No matter how the design and the purpose of the usage, insofar as it can prompt our language ability it can be regarded as CALL.

To this notion, Hubbard (2009) questions Beatty’s broad definition by elucidating how to gauge the improvement of the language learner and what the computer rudimentary stands for. For this reason, he further extends the criteria that can fit “improvement” measure. These criteria encompass “language efficiency, effectiveness, access, convenience, motivation, and institutional efficiency” (Hubbard, 2009: 2). Further study in accordance with effectiveness of CALL practice might be exhaustively exemplified by Bodnar, Cucharini, Strik, & van Hout’s (2014) paper.

Given the fact that it is hard to evaluate the improvement on the present study, this paper will not regard the development on the language learning as the core issue. The underlying reason is the limited number of time and the participants to create questionnaire for the improvement evaluation on the use of both Livemocha and Duolingo. Nonetheless, Hubbard’s (2009: 2) assertion of CALL as “any use of technology in the domain of language learning” is a worth bearing for the present paper. Thus, the present paper will stick to this notion and regard Livemocha and Duolingo as a software that are designed for language learning in a specific way. Regardless the origin of the term CALL, (see Davies, Otto, and Rüschoff, 2012 for further detail of CALL history), the latter definition from Hubbard (2009) is best fit to Livemocha and Duolingo as two tools that assist language learning via computer, in this case, website.

Further reason why Livemocha and Duolingo are chosen is due to the fact that they are relevant with Golonka, Bowles, Frank, Richardson, & Freynik’s, (2012) results who review the effectiveness of language learning technology. They recapitulate that CALL system using ASR is believed to effectively improve pronunciation. Moreover, the advent of chat in language learning can increase language development, productivity, and complexity (Golonka et al.,
Both ASR and chat are present in Livemocha and Duolingo. It eventually leads them merit to the present study.

Having known those features, efficacy, and enormous users of Livemocha and Duolingo, this paper is eager to delve how language tasks on Dutch course are given in both sites. Chapelle’s (2001) criteria on evaluating CALL will be employed to answer the problem.

**Call evaluation-Chapelle’s six criteria**

The evaluation on CALL has been paid so much attention by many scholars in the sense that CALL is dynamic not static. Hubbard (2009) even says that CALL is “both exciting and frustrating”. The rapid development of technology insists CALL to be versatile so that CALL developers and researchers have never come up with a rigid theoretical foundation in CALL research. Notwithstanding, Chapelle’s (2001) recommendation on evaluating CALL system, in regard with (second/foreign) language learning is inevitably a noticeable notion to refer. By employing these criteria, particular CALL system can be assessed. Therefore, Livemocha and Duolingo will be evaluated based on Chapelle’s (2001) criteria.

The objective is then deemed to figure out how the content of language learning is given in both apps. It is argued that the system should be kept interesting to maintain learner’s motivation in using CALL (Hubbard, 2009; Bodnar et al., 2014). Hence, feedback type of both language learning sites might be one tool to align with learners’ motivation in using Livemocha and Duolingo. To substantiate, the following table is the summary of Chapelle’s (2001) six criteria for evaluating CALL.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language learning potential</td>
<td>The degree of opportunity present beneficial focus on form</td>
</tr>
<tr>
<td>Learner fit</td>
<td>The amount of the engagement with the language under appropriate condition given learner’s characteristics</td>
</tr>
<tr>
<td>Meaning focus</td>
<td>The extent to which learner’s attention is directed toward the meaning of the language</td>
</tr>
<tr>
<td>Authenticity</td>
<td>The degree of correspondence between the learning activity and the target language activities of interest to learners out of the classroom</td>
</tr>
<tr>
<td>Positive impact</td>
<td>The positive effect of CALL activity on those who participate in it</td>
</tr>
<tr>
<td>Practicality</td>
<td>The adequacy of resources to support the use of the CALL activity</td>
</tr>
</tbody>
</table>

*Adapted from Chapelle’ (2001: 55)*

In the first criteria, Chapelle (2001) stresses much the essence of language learning activity instead of that language use. Thus, the task on a CALL system must facilitate a focus on form that can enable learner to practice language learning. Learner fit additionally refers to learner’s background knowledge that is appropriate to the current language he is studying. A good task therefore should be able to engage to learner’s possibility to work on it. Language task for beginner, for instance, should be that easy for beginner learners. Chapelle (2001) regards meaning focus as a way to direct learner’s attention to the meaning of the required task. Nevertheless, meaning focus is “not limited on oral communication task” per se (Chapelle, 2001: 55). It can cover the activity on reading or writing whenever learners intentionally project written language to construe meaning.

Authenticity refers to how most likely language learner makes use of language from the task to the world outside the classroom (outside CALL environment). It thus means that the task given in the CALL should consider the target language, where the language is going to be
practiced, and on what condition language learner may encounter it in daily life. Jamieson & Chapelle (2010) substantiate that “Language is best learned when it is used for a realistic purpose (p. 358). Additionally, what is deemed by positive impact is then defined as the effect on language learner to develop their “metacognitive strategies” after using CALL (Chapelle, 2001: 56). The task on CALL, based on this notion, should be able to foster learners’ motivation and willingness to communicate with the language they are studying. Lastly, practicality refers to “how easy it is to language learners and teachers to implement CALL task” (Chapelle, 2001: 57). The availability of the resources should be taken into account.

To recapitulate, the aforementioned criteria will be employed to assess language task in both Livemocha and Duolingo. In the following sections, the description of Livemocha and language learning, Duolingo and language learning, and the discussion of Livemocha and Duolingo’s Dutch course task for basic level are to be presented. The last two parts will deal with conclusion and suggestion for future study.

Livemocha and Language Learning

Clark & Gruba (2010) argue that Livemocha is a social networking site that “advertises itself as world’ largest community of native speakers who are eager to help and learn from each other” (p. 164). As a free social networking site, it now has over 16 million users so that Livemocha becomes one of the largest online language learning communities (Livemocha, 2015). It was firstly launched in 2007 (Livemocha, 2015) and now it offers about 35 languages to study. Livemocha incorporates an integrated language learning concept with reciprocal benefited outcome. Furthermore, it is corroborated by free interactive courses and lessons, large community of native speakers, and of course enormous foreign language learners.

The course content is comprised of several levels in which it is ranked from L1 as the basic level to L9 as the highest level. However, those levels are not necessarily the same for every languages. For instance, in Dutch course, it has only L1 level, while English, on the other hand, has been developed up to L6. Regardless of that level discrepancy, the course material for Dutch is remarkably salience in the sense that it offers some basic knowledge of daily communication.

In terms of course activity, it is structured through eight main activities in each lesson. Firstly, language learners are introduced by introduction section which enables them to overview what the materials are to study. And then the vocabularies and usage are given. To review learners’ understanding of the given vocabularies, the activity is continued by usage practice. In these three steps, learners earn points on every task they do. The points, afterwards, can be used to do other exercises such as reading and writing, reading and speaking, listening and writing, and listening and speaking (see Fig 1).

The points that learners have earned are displayed in the right corner. Another way to earn point is that by helping others exercise. As it is mentioned in its platform that Livemocha integrates free online activity in large community to help and learn from each other, user can help others in reviewing the exercise and at the same time learn from others in ‘Help others’ column. Conversely, learners’ exercise on reading, listening, speaking, and writing will be reviewed by other native communities or communities who are learning the same language.

Duolingo and Language Learning

Duolingo is an online-free language learning website which bases the study on translation method (Garcia, 2013). Language learner can learn foreign language by doing, specifically translating. Von Ahn (2013) who headed Duolingo project stated that the idea behind Duolingo is to make “people learn new language for free while simultaneously translating the web” (p. 1). This project, Duolingo is aimed at abolishing the constraint on the so called worldwide (von Ahn, 2013). Thus, von Ahn and team make this learning websites for free and it incorporates a number of languages to learn.
Two previous studies have closely investigated Duolingo in terms of its effectiveness (Vasselinov & Grego, 2012) and its contribution to education (Jašková, 2014). Both studies reveal positive attitude in a way that Duolingo gained 95% confident interval for effectiveness (Vasselinov & Grego, 2012) and perceived as potential learning assistance (Jašková, 2014). Therefore, it is worthwhile to further pick up Duolingo as subject to research on.

In Duolingo, the course on every language consists of several language skills, for instance speaking, writing, listening, etc. Furthermore, it also offers multiple choice exercises that may challenge language learner. Duolingo incorporates games within the lesson. This commitment to bring convenient and enjoyable environment of language learning can be traced since the first time we open the website (see Fig. 5).

The course is given from various levels, from basic to advanced level. In the basic level, the material is based on vocabulary mastery. The language learner will be provided with some basic vocabularies that are mostly used in daily conversation. There are two basic levels in Dutch course in which basic one encompasses five skills to pass in advance before continuing to the basic two levels. Such pattern also applies in other levels which means that language learners cannot jump to the next level unless they pass the earlier level.

METHOD

The present study is conducted within the paradigm of qualitative research which aims to explore the specific phenomena, that is the efficacy of Livemocha and Duolingo evaluation tasks. Besides, it incorporates in-depth participation and investigation in which the researcher experience directly to use both Livemocha and Duolingo to study foreign language. Employing researcher’s experience is finally fit to the characteristics of qualitative research in a way that “the entity to study is the life world as it is experienced” (Fink, 2012: 6).

The Dutch was then chosen as the sample to study for two main reasons. First, the researcher was taking his master degree in the Netherlands. Therefore, it potentially fits his necessity as the student in the Netherlands university to study Dutch. Second, the researcher’s native language is Indonesian in which both Dutch and Indonesian languages have been in contact for a long period already. It is then expected that the researcher will not encounter a huge problem.

The data were gathered from the application of Livemocha and Duolingo in which the researcher plays role as the student who studies foreign language. The researcher followed the instruction given in both apps to begin the session. All necessary details, for instance, the instructions, the structure of the course, and the evaluation model are highly recorded and noted. Afterwards, the researcher compared both Duolingo and Livemocha assessments by sticking to Chapelle’s (2011) evaluation tasks. Besides, the researcher also paid much attention on the flow of the lesson and the evaluation such as the task given in either Livemocha or Duolingo. The detailed samples of such activities are attached in the appendices.

FINDINGS AND DISCUSSION

Evaluating Livemocha

The language task in Livemocha, especially in Dutch course, is structured in the following orders; introduction, vocabulary, usage, usage practice, practice reading/writing, reading/speaking, listening/writing, and listening/speaking. The format is almost the same for all languages offered in this site. As such, learners are directed to firstly acknowledge the material and then projected to practice it. Taking this activity into account, learners are provided with activities that enable them to focus on the form and practice directly in the later step. For instance, the course on *De getallen van 1 tot 10* in L1 level is started by example of usage in the first two steps. The third step to pass is usage exercise where learners can practice the
understanding of the task. Besides, learners can also earn points in which it is useful to do the following exercises such as reading/writing or reading/speaking. Thus, language learning potential as proposed by Chapelle (2001) is accommodated in Livemocha.

In regard to the second criteria, learner fit, Dutch course in L1 offered in Livemocha does not seem to fit to beginner learners. It is not because of the task that is not relevant, but rather the language of instruction. The instruction and explanation are offered in Dutch which may constraint beginner in learning this course. Although there is a translation icon to help learners, it is not exhaustively helpful in the sense that translation appears too fast. In this case, learners are indirectly required to understand little bit about Dutch prior to their study in Livemocha. Those learners who have not been familiar with foreign language, for instance Indonesian learners, can be fervently confounded.

However, Dutch course task is arguably worth noting in terms of meaning focus. Chapelle (2010) argues that language task of particular CALL system should direct learners’ attention to the meaning of the intended task. The availability of flip card icon does help learners to eventually understand the meaning of the word. Learners are given the explanation of particular word in which they are uncertain. By so doing, Livemocha does comply meaning focus characteristic.

Additionally, authenticity deals with how the language in the task closely corresponds to target language in the real life. The Dutch course for L1 seems to fit this criterion. By providing some topics like numbers, greetings, days, etc., the materials are rudimentarily on par with beginners’ need. Furthermore, such skills and vocabularies are mainly necessitated to communicate in daily interaction. Nevertheless, it cannot be strongly argued that this course is worth in terms the authenticity for the mentioned reasons per se. Some aspects should be considered to strengthen the claim.

The fourth assessment to gauge is that how the task may impact the learners. The attractive design and structure of the task inevitably provide positive impact to language learners. Having been noted that it questionably fits beginner learners, they can build and even foster their willingness to study the language. As Chapelle (2010) mentions, the refinement of learners’ motivation may indicate that CALL task has positive impact. The only thing to worry in this case is notion of feedback giving. Since Livemocha is integrated learning, feedback is given indirectly after the exercise where learners have to wait for the response from the community or expert. This feedback giving strategy may be a drawback of Livemocha.

Concerning the practicality, this task is freely available. Learners can access this site and use it in any circumstance. Moreover, the task practice in Dutch course can be easily replicated in other fields of language learning and not limited to computer based activity. Intriguingly, language learners can implement such activity and access exercises insofar as they have internet connection.

To recapitulate, among the six criteria of CALL evaluation (Chapelle, 2001), Livemocha, in this case, the L1 course of Dutch, complies five out of six criteria. Further consideration should be put into learner fit point to meticulously evaluate whether or not the task is appropriate to beginner. Nonetheless, the rests do not necessarily mean absent from limitation. Impact and practicality can also be delved through more appropriate method and analysis to see the effectiveness of Livemocha for beginner learners of Dutch.

Evaluating Duolingo
As it is mentioned in section 4, Duolingo accounts its practices based on translation method. Thus, most of the activities require learners to listen to the instruction, choose the answer, record the Dutch pronunciation, write the English translation, and write the Dutch translation of English task. Therefore, for language learning potential, Duolingo does provide learners activities that focus on form. Learners can learn the spelling and pronunciation of Dutch. The former is exemplified by the activity on writing English phrases or sentences into.
Dutch or vice versa as in Fig 10. The latter, on the other hand, is best represented by the task of reading the Dutch sentence and recording it (see Fig 9). The site then evaluates the output of learners’ pronunciation. As such, language learners can identify their mistake in pronouncing Dutch words, phrases, or sentences. It is inevitable to say that language learning potential as defined by Chapelle (2001) is effectively employed in Duolingo by focusing on words, phrases, or sentences translation and pronunciation.

The second point is learner fit. Duolingo offers free Dutch course for beginner. The material consists of word naming which is guided by picture and voice (see Fig 6). Learners can hear the name of the picture by navigating the cursor the picture and click it. The instruction is also clear, for instance in Fig 6, learners are required to choose Dutch translation of an apple. In so doing, the software does provide three different pictures with sounds on it. If the learners know exactly what an apple look like, they will have no difficulty in choosing the right answer. However, the main point is not merely to choose the right answer, rather to give beginners learners knowledge on the instructed task, in this case the Dutch translation of an apple. To this point, this basic level task does fit to Dutch beginner leaners.

Meaning focus, as the following criterion to assess CALL, is defined as a way to direct learners’ attention to the meaning of the required task (Chapelle, 2001). Learner can adequately perform well in understanding the meaning of the word in task translation. The task is mostly segmented to word, phrase, or sentence based translation, which eventually enables learners to concern on what is being learned. Such activity projects learners’ attention to the word meaning.

In terms of authenticity, it cannot be argued that the task does completely correspondence with the target language. Nonetheless, vocabulary activity such as names of pronouns, names of fruit and milk, can be essential core tool to practice Dutch in real conversation. Pronoun, to some extent, is mostly used in speaking and interaction. Other categories such fruit, milk, and bread, are the things that attached to Dutch lifestyle. Hence, knowing those words are also pivotal.

The impact of course by no means is positive to the leaners. Having known the Dutch words and pronouns for instance, can build leaners linguistic usage. The translation task and recording foster their awareness of the language usage. Moreover, the advent of feedback on every response that learners make may be another reason that can trigger learners’ willingness to study the language in Duolingo.

The last part is dealing with practicality. It refers to “how easy it is to language learners and teachers to implement CALL task” (Chapelle, 2001: 57). In the context of language learning, it is amply practical to use Duolingo task since it is offered for free. Everybody can access insofar as leaners have access to the internet. Unlike Talk to Me in Hinck’ (2003) report that is unlikely to be practical due to the fact that Talk to Me is “reasonably priced and runs on Windows 95 or better” (p. 18), Duolingo then emanates a reasonable practicality.

In overall evaluation, Dutch course in Duolingo is designed for beginner leaners in which the task are relevant to learners with no or very limited background of Dutch. Chapelle’ (2001) criteria on evaluating CALL task has been assessed for that course. Among the six criteria, the case of authenticity seems to be lack of support in a way that the target language is not decided yet. It can be resolved and justified whenever the outcome of language use outside the task is determined in advance. Nevertheless, vocabulary translation from English to Dutch in another way around does provide worthwhile information for Dutch beginner learner.

CONCLUSION

Livemocha and Duolingo are two free online language learning sites that are noticeably worthwhile to support learners in studying foreign languages. This paper has come with the evaluation on language task of both Livemocha and Duolingo. Superficially, both sites are a wise recommendation for language learning. Nevertheless, language learner may have to
further set up his or her capability before using CALL. In fact, the notion of learner fit, for instance in learning Dutch course, Duolingo can be the much more considered than its counterparts. On the other hand, the authenticity task in Duolingo should also be highly concerned since its counterparts, Livemocha, does provide much more relevant exercise for daily practice.

Due to the fact that the present study does not incorporate questionnaire to assess the effectiveness, future study can investigate how such CALL task may be truly effective for language learners, more specifically, those who are not familiar to any foreign language. Besides, the evaluation and assessment analysis can be further modified by incorporating quantitative approach. The present paper comes up with descriptive analysis, which is to some extent, rather hard to validate. However, the meticulous description has been elaborated in this paper. Extending the number of respondents to give more holistic overview may also be considered to avoid personal bias. Therefore, the future study can be much more valuable by considering those other aspects in evaluating CALL task in, for instance, Livemocha and Duolingo.

REFERENCES


Appendix

Fig 1. Front page of Livemocha

Fig 2. Introduction section

Fig 3. Vocabulary section

Fig 4. Practice Usage

Fig 5. Main site of Duolingo

Gamification poured into every lesson.

Fig 6. Example of the exercise in basic level

Fig 7. Translation exercise

Fig 8. Feedback example on correct answer

Fig 9. Feedback example on almost correct answer
Fig 10. Feedback example on the wrong answer