



The Needs of Standardized CLIL Assessment in Indonesia

Ima Isnaini Taufiqur Rohmah^{1*}, Fitri Nurdianingsih², M. Zainudin³

^{1,2}Pendidikan Bahasa Inggris, ³Pendidikan Matematika IKIP PGRI Bojonegoro Bojonegoro

*Korespondensi Penulis. E-mail: isnainiima@ikippgribojonegoro.ac.id

Abstrak

CLIL bermanfaat dalam pendidikan terutama di pendidikan dasar, karena dengan pendekatan pengajaran dan pembelajaran fokus ganda yang digunakan di mana bahasa utama pengajaran dan pembelajaran digunakan bahasa asing untuk mempromosikan penguasaan konten dan penguasaan bahasa ke tingkat yang telah ditentukan sebelumnya. Studi ini dilakukan untuk mendalami sejauh mana praktik penilaian yang ada dapat digunakan untuk mengevaluasi penguasaan bahasa Inggris siswa. Temuan mengungkapkan bahwa penilaian yang dipraktikkan tidak memadai untuk digunakan untuk menilai kinerja bahasa Inggris siswa. Penelitian saat ini menyarankan bahwa guru harus memastikan untuk melibatkan empat keterampilan bahasa dasar dalam mengembangkan instrumen penilaian di kelas CLIL.

Kata kunci: CLIL, penilaian, standarisasi

Abstract

CLIL is a beneficial in education especially in primary education, since by its dual-focused teaching and learning approach used in which the main language of teaching and learning are used foreign language for promoting both content mastery and language acquisition to pre-defined levels. The study was aimed to investigate the extent can the assessment practiced be used to evaluate students' English performance. The findings revealed that the assessment practiced was not adequate to be used to assess students' English performance. The current research suggested that the teachers have to make sure to involves four basic language skills in developing assessment instruments

Keywords: Assessment, CLIL, Standardized

INTRODUCTION

Content and Language Integrated Learning (CLIL) is considered as an innovation in education especially in the area of language teaching and learning. It is an approach in education where dual focus of learning (content and language) is mutually beneficial for both content and language subjects. Moreover, it aims to increase learner motivation, develop learners' second or foreign language, cognitive skills, and intercultural understanding (Eurydice, 2006). Regarding its beneficial, CLIL is greatly gaining a momentum and introduced as an educational approach in the worldwide context (Rohmah, et al., 2019).

Assessment in CLIL is crucial because of its "wash-back" effect on learning; it must cover both content and language and take into consideration all aspects of CLIL communication in their specific context (Barbero, 2012). Language assessment in CLIL, as in all other education fields, must fulfil general quality criteria, two of which are essential: *validity* and *reliability* (Barbero, 2009, p. 108). Assessment must be supported by appropriate assessment tools and measuring certainly what the assessment tools intended to

assess and are in a whole, consistent with the teaching and learning goals. Moreover, assessment has to provide reliable feedback for the learner consisting of criteria, scores, and descriptors that may quantify, evaluate, and interpret the outcomes. The reliable assessment indicated the accuracy, precisely, and consistently (Hönig, 2010).

Many studies examined assessment of CLIL in European countries. For example, Wewer (2013) investigated English performance assessment in bilingual content instruction CLIL in Finland. There also some existing studies (Hönig 2010, Serragiotto 2007) as well as research in assessment concerning both language and content knowledge (Rohmah, et al., 2020; Rohmah, 2019; Lasagabaster & Sierra 2010; Sierra 2007) have mainly focused on secondary or tertiary students in the United States.

Although some schools in Asia's context also adopt CLIL in teaching and learning, the study in the CLIL assessment area was rare and only between (Fitriani, 2016) and (Rachmajanti et al., 2015). Furthermore, the needs of standardized CLIL assessment in Indonesia has not so far studied at all. Based on the background and some situation drew, my hunch is that CLIL considered as an excellent alternative way for second language acquisition, furthermore English performance assessment is a fundamental element to support the success of CLIL since we know that assessment guides learning. Students end up focusing on what they are assessed (Briggs et al., 2008).

METHOD

The current study used a case study research design to investigate the research problems. Creswell (2009) stated that a case study defined as an in-depth exploration of a bounded system. This research was carried out at 5 (three) Elementary Schools in East Java which bringing the CLIL program in teaching and learning. The subjects of the study were 10 (six) teachers consisted of (5) three subject teachers (science teachers) and 5 (three) language teachers (co-teachers). The research also involved 24 (twenty-four) students in each school. In collecting the data, this research used some to respond to the five research questions, this research applied 4 (four) instruments, i.e. observation sheet, questionnaire, interview guide, and documentation. The current study followed thematic analysis by Braun & Clarke (2006) was used to analyze qualitative data since Braun & Clarke (2006, P. 412) argued that “thematic analysis is method to identify, analyze and report themes and patterns within a data set”. In this sense thematic analysis can be also seen as a qualitative descriptive method that help the researcher in serving the core skills for conducting many other forms of qualitative analysis.

FINDINGS AND DISCUSSION

Findings

The adequacy of the written test to assess students' English performance

The adequacy of written tests to assess students' English performance was explored by observing the document of the test and also interviewing the teacher. In observing the document of the written test, the researcher used the assessment evaluation checklist for written test in CLIL and the observation assessment evaluation checklist developed by Ross (2005). The first test to be analyzed was about the adequacy of the written test used and then the adequacy of the teachers' observation to evaluate students' English performance. Here is the result of the document observation based on the assessment evaluation checklist in CLIL (*primary level*).

In analyzing the adequacy of the test used to assess students' English performance, the researcher focused on the items of the language aspects provided in the evaluation checklist. Based on the result of document observation, it can be concluded that the written

test used by the teacher was not adequate to assess students' English performance, it was based on the items appear in the document of the test that consisted of some aspects:

a. *Language structure.*

1) *The test shows awareness of grammatical of the language.*

The result of the observation checklist shows that the test does not consider the awareness of grammatical of the language. The test created by the teacher is not focused on assessing the grammatical features of the language. It can be seen from the document of the test which is identified that all the questions were used to gain the data of students' content understanding. It seems that the language was only used as a medium to serve the test. As confirmed by the teacher, the consideration of ignoring students' grammatical manner was because the function of the language (English) in CLIL was only as a medium of instruction. English is used as the language to serve the learning. It was different from the EFL assessment where the grammatical features were important to be assessed. As stated by teacher Ahmad and teacher Feni that:

Yes, it is right, in the document of the test, the element of grammatical features does not appear as the focus of assessment in the written test. It was because, in CLIL, the language was only used as a medium of instruction, not as a part of the assessment.

(Interview: Ahmad, 24/08/2021).

Yes, we are ignoring the grammatical manner of the students' language, it does not mean that grammatical of the language is not important, but we encourage the students to be brave to speak and to use English first in the schools' setting.

(Interview: Feni, 27/08/2021)

Based on the data taken from the results of document observation and also the confirmation is given by the teacher, it can be concluded that the integrated assessment by using written test practiced does not consider grammatical manner as the important component of the language to be assessed. The most common goal was assessing students' content understanding.

2) *The test identify to use appropriate language structures*

The test created by the teacher does not identify to use of appropriate language structures. There were no criteria that asked the students to use appropriate language structure to answer the questions. The test represents by CLIL teacher identified of the use test items such as *a circle (15)*, *Tick it (1)*, or less cognitively demanding completion questions, such as *write the number (4)* and *color the right (13)* where all the instruction tend to give intention that not need to use appropriate language structures to answer the question

b. *Language function.*

1) *The test involving language activities (speaking, listening, reading, writing)*

Written test created by the teacher does not involve some language activities such as speaking, listening, reading, and writing. There is no activity to assess students' language skills presented in the test. The test was not separated into speaking, reading, writing, and listening section. It consisted of a range number of questions which has a general instruction to assess students' content understanding.

The findings gained from reviewing the document of the test also confirmed by the teacher in the interview section. They confirmed that the presentation of the written test was not separated into some categories such as speaking, reading, writing, and listening. It was presented in a range number of questions with specific instruction for the students to answer the question. The common instruction used such "*give a tick, name, circle, match* and so on.

The configuration of the test was not consisted of the language skills such speaking, reading, listening or writing. It was consisted of a number of questions with various instructions to be answered by the students
(Interview: Ahmad, 27/07/2021)

No, the test does not consist of speaking, reading, listening, or writing. We think that it was difficult for the students
(Interview: Feni, 24/07/2021)

2) *CALP functions are involved in the test (CALP: Cognitive Academic Language Proficiency)*

Cognitive Academic Language Proficiency (CALP) is referred to simply as Academic proficiency. CALP was introduced into the field of language education by Cummins (1979) which consisted of two central components: knowledge of the academic language and knowledge of the specialized subject matter. Knowledge of Academic language is knowledge of the special language used (English) especially related to the ability to use the foreign language to demonstrate the content objectives. Knowledge of specialized subject matter consists of knowledge of science subject itself.

Based on the result of document observation, the test served by the teacher does not involve knowledge of the academic language. It was only covered knowledge of the specialized subject matter. The test tends to assess students' content understanding and forgetting to assess the language. The instruction presented in the test wan not drive the students to be able to use the language to demonstrate the content knowledge. It was such (to match, to give color, to It was confirmed by the teachers in the interview session dealt with the integration of CALP in the test. As stated by teacher Dani and Feni who stated that:

Yes, it is right, the focus of the test is to assess students' science knowledge, we think that it was not necessary to assess the language in the test, it will be more difficult
(Interview: Dani, 24/07/2021)

No...the main focus is not to assess academic language, we are focusing on students' content understanding towards the learning through
(Interview: Feni, 27/07/2021)

c. *Vocabulary (The test recall subject-specific vocabulary)*

In a written test created by the teacher, it was presented subject-specific vocabulary. Subject-specific vocabulary in this term is related to the specific vocabulary used in a science subject. There were numbers of vocabulary introduced and intended to be mastered by the students related to science terminology as presented in the following table 1:

Table 1. Specific science vocabulary in written test

Topic	Specific vocabulary
Dispersal	Wind dispersal
	Water dispersal
	Seed dispersal
	Animal dispersal
Photosynthesis	Glucose
	Sunlight
	Oxygen

	Carbon dioxide
	Energy
Seed	Seed coat
	Cotyledon
	Plumule
	Hypocotyl
	Embryonic root

Listed vocabulary presented in the test was a specific vocabulary related to the science subject. For example in the term of *dispersal*, there were range numbers of vocabularies related to the term, such as *water dispersal*, *wind dispersal*, *seed dispersal*, etc. In correlation to the term *seed*, it also has a range number of vocabularies that need to be mastered by the students such as *seed coat*, *cotyledon*, *plumule*, *hypocotyl*, and *embryonic root*. It can be concluded that the written test created by the teachers fulfills the need to recall a subject-specific vocabulary for the CLIL students. Though some aspects of the written test that showed features of lower validity, it cannot be claimed that they were invalid assessment instruments but it needs a serious inspection to make written test presented accommodate content and language assessment as generated by CLIL curriculum.

The adequacy of teacher's observation to assess students' English Progress

Another tool that teacher used to assess students' English performance as elaborated in the section before was teacher's observation. The focus of teacher's observation was to assess students' speaking and interaction skill, teacher observation practiced was such kind of *on-run observation*. Since it also has a vital role to assess students' English performance, the adequacy of the teacher's observation is needs to be investigated.

In investigating the adequacy of teachers' observation assessment, observing the teacher's observation assessment during the teaching and learning in the classroom practices, and also interviewing the teacher were done to gain the data of the adequacy of the assessment practiced. The observation conducted was used as the observation assessment evaluation checklist developed by Ross (2005). From the 3 (three) schools used as the research setting, the researcher only takes one school located in centre of Bojonegoro to be observed. The reason for choosing this site to be observed was because the overall assessment practiced as stated in section 4.1 was the same one and another and also the school located in Bojonegoro.

Based on the data of observation, it can be concluded that the teacher's observation assessment is not adequate to evaluate students' English performance. It can be described as follow:

First, the observation assessment conducted was unplanned before, it was incidentally conducted and flows through the teaching and learning processes. It naturally occurs during the teaching and learning situated in the classroom activities and also during the the interactions between teachers and students. It is noted that unplanned observation giving an opportunity for the teachers to naturally assess the students that emerges in the classroom activities. The finding also confirmed by teacher Ahmad, he said that:

Yes, it is right. The observation conducted was unplanned. It flows through classroom activities. I never plan it before it naturally happened.

(Interview: Ahmad, 15/08/2021)

Second, in terms of comprehensiveness, whether teacher observation assessment conducted is related to the learning outcomes. The observation assessment conducted by

the teacher does not relate to the learning outcomes. Teacher observation assessment can be used as a good and valid information sources in recording and reporting student demonstrations ability towards the learning outcomes. However, teacher observation is a source to provide valuable information towards student ability to demonstrate the learning outcomes in certain education levels. The finding was confirmed by teacher Ahmad who stated that:

The observation assessment was not comprehensively related to the learning outcomes since it was conducted incidentally. I only want to know the students' progress in the term of their speaking and interactional skill
(Interview: Ahmad, 15/08/2021)

Third, the Teacher's observation assessment should be situated within familiar learning contexts and closely related to the curriculum framework, learning experiences, and pedagogical planning. Based on the data gathered from the research setting, it can be concluded that teacher's observation assessment practiced does not always situated within familiar learning contexts and closely related to curriculum framework, learning experiences, and pedagogical planning since it incidentally happened, teacher's observation assessment was not always situated in a learning context. Sometimes it is based on the topic presented but sometimes it was outside the curriculum framework. As confirmed by teacher Ahmad in the interview session:

No, it is not, observation assessment not always conducted based on the topic presented in the curriculum and it was not always conducted during the classroom teaching and learning process
(Interview: Ahmad, 15/08/2021)

Fourth, the teacher's observation assessment conducted is deriving from a variety of situations and occasions. Based on the result of the observation checklist, the teacher's observation practiced was deriving from a variety of situations and occasions. The teacher was not only assessing students' English performance through observation in the classroom during the teaching and learning process but also outside the classroom with different topics and contexts. As confirmed by teacher Ahmad who said that:

Yes, it is right, observation assessment not always conducted inside the classroom during the learning process, but it sometimes also during students' leisure time or a class day out with various and incidental topics and context
(Interview: Ahmad, 15/08/2021)

Fifth, the teacher's observation is interesting, challenging to the students. Based on the result of the observation evaluation checklist, the observation is interesting but not challenging for students. It was interesting for the teacher since the teachers were curious to find what they have to records during the teaching, sometimes unconditioned things happened can be found as source information. It was not challenging for the students because the observation conducted was incidental and unplanned, the teacher did not create a special work or assignment that interesting and challenging. It flows naturally in the classroom setting. It also confirmed by teacher Ahmad in the interview session, he said that:

Yes, it can be said that it does not interesting and challenging enough since I am not in the way of creating a special moment or assignment for students to be observed. As I mentioned before observation is a supporting assessment and as a consideration in assessing students' English performance.
(Interview: Ahmad, 15/08/2021)

Sixth, the teacher's observation assessment conducted should emphasize relatedness and connections in learning and involving performance on complex wholes aspect of the language. Based on the result of the observation evaluation checklist, the observation conducted was emphasizing relatedness and connections in learning but it was not involving performance on complex whole aspects of the language. The aspects involved i.e speaking and vocabulary. The observation conducted were questioning and probing. In questioning, there were two aspects involved, students' speaking skills and vocabularies. The reason for not involving all aspects of language was because other aspects were assessed through a written test. As stated by teacher Ahmad:

In conducting assessments through teacher observation, we use some method to gain the data, such as questioning and probing. In questioning, there were two aspects to be paid attention are speaking and students' vocabulary, other aspects of language were assessed in written test
(Interview: Ahmad, 15/08/2021)

Seventh, the teacher's observation assessment should be recorded with the evidence. Based on the data gathered through the evaluation checklist and also interview, the observation conducted was not recorded with evidence. The teacher also did not make any note dealt with the result of observation. It was only memorized in the teacher's mind. As confirmed by teacher Ahmad:

I do not make any note or report in observing the students. I only memorizing in my mind.
(Interview: Ahmad, 15/08/2021)

Eighth, there is a standard and criteria in conducting a teacher's observation assessment. Based on the evaluation checklist, there is no standard and criteria in conducting the observation. The teacher did not compose or following a standard and criteria in conducting the observation. It was seen in the classroom situation, the observation incidentally happened. As confirmed by teacher Ahmad:

We do not have any standard or criteria in assessing the students through observation. It happened incidentally, it flows naturally.
(Interview: Ahmad, 15/08/2021)

Based on the data gathered through the evaluation checklist and interview, it can be concluded that whether unplanned teacher observation can be used as a basis for formal assessment and reporting may depend on the records that are kept. The evidence of teachers observation assessment needs to be gathered and recorded systematically.

Discussion

The finding dealt with the instrument of the written test created by the teacher does not identify to use of appropriate language structures, it also does not ask the students to use appropriate language structure to answer the questions. The finding was in contrast with the study conducted by (Pérez-Cañado, 2012) who asserted that in composing the instrument of assessment in CLIL the content and the language instruction must be balanced. Furthermore, (Coyle, Hood, & Marsh, 2010) confirmed that to get a balanced viewpoint in assessing students' content knowledge and language performance, the instruction created has to consider the use of language structure (grammatical manner).

Second, in the aspect of language function, the test created by the teacher does not involve all aspects of language activities such as speaking, listening, reading, and writing. There is no activity to assess students' language skills presented in the test. The finding was in line with Massler, Stotz, & Queisser (2014) who confirmed that the assessment instrument in CLIL has to identify the aspects of English performance such as reading, writing, speaking, listening, and vocabulary building. It also supported by the study conducted by Mehisto (2008) confirmed that in creating assessment instruments in CLIL, the curriculum developer has to be aware of the equal assessment activities provided in the test such as assessing content knowledge and language manner.

Third, one of the aspects accommodated in a written test created by the teacher is that the instrument presents subject-specific vocabulary. There were numbers of vocabulary introduced and intended to be mastered by the students related to specific terminology that is the specific vocabulary for science subject. The finding was in line with Sierra (2007) who state that a range number of specific vocabulary related to certain terminology has to be taken a consideration to be inserted in the instrument of the test in English performance assessment in CLIL.

Another finding dealt with the extent of teachers' observation conducted by the teachers was adequate to assess students' English performance especially in assessing speaking and listening skills. As Bell & Lorenzi (2004) who asserted that both learners and the teacher may be unaware of the assessment process, which is why observation ultimately serves learning facilitation and immediate decisions.

Observation assessment can be categorized and results gut instinct assessment if in the process of observation was not concerning with a specific focus and purposes. Without any specific focus, assessment through teachers' observation does not necessarily reflected into actual feedback. It is can caused the target of observation may allusive.

The findings revealed that *first*, the observation assessment conducted was unplanned and it was incidentally conducted and flows through the teaching and learning processes (on run) or incidentally. Graaff, Koopman, & Tanner (2012) confirmed that incidental observation is therefore the weakest form of teacher observation and would preferably be used only as supplementary evidence to support other forms of evidence. Relying on incidental observation alone would be unsatisfactory

The finding was in line with Aiello, Di Martino, & Di Sabato (2017) who confirmed that in observation assessment for primary students can be done accidentally, unplanned, and in any situation that appeared with systematic evidence (recorded). Lin (2015) inferred that observation is not the main assessment instrument, it was a kind of supplementary instrument to the main instrument of the assessment.

Second, the observation assessment conducted by the teacher does not relate to the learning outcomes. Since it was conducted accidentally, unplanned, and in the form of on run observation in the class. Sometimes it also does not situate within familiar learning contexts but is derived from a variety of situations and occasions. Brady, & Kennedy (2001) inferred that in teachers' observation assessment, learning outcomes that have not been demonstrated can be deliberately prompted. Assessment should be planned as well as incidental. Further Mehisto (2008) indicated that teachers have to ensure that assessment through observation is comprehensive and outcome-based.

Third, the finding whether the teacher's observation assessment is recorded with the evidence, the data showed that the observation assessment conducted was not recorded with evidence, no standard, and criteria as a guideline. The teachers did not make any note of the result of the assessment. The finding is contrary to Genesee & Upshur (2016) belief that for the implementation of teachers' observation, it has been recommended that the evidence focuses on the demonstration of learning outcomes. Evidence is documentation

that records, illustrates, or confirms student demonstrations of learning outcomes. Várkuti (2010) asserted that to get a valid assessment through observation, it must be evidence-based that recorded systematically.

CONCLUSION

The finding of the needs standardized CLIL assessment practiced can be used to evaluate students' English performance was viewed from the adequacy of the written test and teachers' observation to assess students' English performance. The finding can be concluded that written test possessed was not adequate to assess students' English performance. It was caused by some facts: a) written test developed by subject teachers tend to only assess the content understanding and forgetting less attention to assess students' English performance; b) Teachers' observation assessment practiced by language teachers to support the result of written test was done was done unplanned, no evidence-based, and not recorded. It made teachers' observation assessment practiced less of validity

REFERENCES

- Aiello, J., Di Martino, E., & Di Sabato, B. (2017). Preparing teachers in Italy for CLIL: Reflections on assessment, language performance and willingness to communicate. *International Journal of Bilingual Education and Bilingualism*, 20(1), 69–83, DOI: [10.1080/13670050.2015.104187](https://doi.org/10.1080/13670050.2015.104187).
- Barbero, T. (2012). Assessment in CLIL, in Järvinen H. (ed.), *Language in Content Instruction*, University of Turku. European Commission. <https://www.google.com/search?client=firefox-b>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>.
- Bell, N. D., & Lorenzi, D. (2004). Facilitating second language acquisition in elementary and secondary physical education classes. *Journal of Physical Education Recreation and Dance*, 75(6), 46–52. <https://doi.org/10.1080/07303084.2004.10607257>.
- Briggs, M., Woodfield A., Martin C., & Swatton P., (2008). *Assessment for Learning and Teaching*. Learning Matters, Exeter.
- Coyle, D., P. Hood & D. Marsh (2010). *CLIL. Content and Language Integrated Learning*. Cambridge: University Press.
- Creswell, J. W. (2009). *Research Design: Qualitative Quantitative and Mixed Methods Approaches*. SAGE Publication Asia-Pacific Inc. <https://doi.org/10.1017/CBO9781107415324.004>.
- Eurydice. (2006). *Content and Language Integrated Learning (CLIL) at School in Europe*. Brussels: European Commission.
- Fitriani, Ika. (2016) Grass root's Voices on the CLIL Implementation in Tertiary Education. *Dinamika Ilmu*, 16(2): 211-220.
- Graaff, R., Koopman, G.J. & Tanner, R. (2012). Integrated opportunities for subject and language learning. Implementing a rubric for cross-curricular learning activities. Quality Interfaces: Examining Evidence and Exploring Solutions in CLIL. Eds. D. Marsh and O. Meyer. Eichstaett: Eichstaett Academic Press. 154-173.
- Hönig, Ingrid (2010). *Assessment in CLIL: Theoretical and Empirical Research*. Saarbrücken: VDM Verlag.
- Lasagabaster, David & Sierra, Juan Manuel (2010). Immersion and CLIL in English: more differences than similarities. *ELT Journal*. 64 (4), 367-375. <https://doi.org/10.1093/elt/ccp082>

- Lin, A. M. Y. (2015). Conceptualising the potential role of L1 in CLIL. *Language, Culture and Curriculum*, 28(1), 74–89.
<https://doi.org/10.1080/07908318.2014.1000926>.
- Massler, U., Stotz, D., & Queisser, C. (2014). Assessment instruments for primary CLIL: The conceptualization and evaluation of test tasks. *Language Learning Journal*, 42, 137–150.
<https://doi.org/10.1080/09571736.2014.891371>.
- Mehisto, P. (2008). CLIL Counterweights: Recognizing and Decreasing Disjuncture in CLIL. *International CLIL Research Journal*, 1, 96-117. Retrieved June, 02, 2008 from <http://www.icrj.eu/11/article8.html>
- Pérez-Cañado.(2012). CLIL research in Europe: past, present, and future. *International Journal of Bilingual Education and Bilingualism*, 15(3), 315-341,
<https://doi.org/10.1080/13670050.2011.630064>
- Rachmajanti, S., Zen, E.L., Apriana A. (2015). Using English in science class as an attempt to maximizing input in second language learning. Vietnam: TESOL Conference 2015.
<http://www.vnseameo.org/TESOLConference2015/>.
- Rohmah, I.I.T., Saleh, M., Faridi, A., Fitriati, S.W. (2019) Language Assessment Pattern for Primary Education in the Content and Language Integrated Learning (CLIL) Classroom Context, *Asian EFL Journal*, 21(2.2), 101-123.
<https://www.elejournal.com/asian-efl-journal/asian-efl-journal-volume-21-issue-22-march-2019/>
- Rohmah, I.I.T. (2019). The feasibility and effectiveness of integrating content knowledge and english competences for assessing english proficiency in CLIL. *ETERNAL (English Teaching Journal)*: 10(1), 65-73.
<http://journal.upgris.ac.id/index.php/eternal/index>
- Rohmah, I. I. T., Saleh, M., Faridi, A., & Fitriati, S.W. (2020). Content and language integrated learning (CLIL):Teachers’ Point of View, Proceeding of the 2nd International Conference on Education: 594-600, <http://dx.doi.org/10.4108/eai.28-9-2019.2291094>
- Ross, S. (2005). The impact of assessment method on foreign language performance growth. *Applied Linguistics*, 26 (3), 317-342. <https://doi.org/10.1093/applin/ami011>
- Serragiotto, Graziano (2007). *Assessment and evaluation in CLIL*. In D. Marsh & D. Wolff (eds). *Diverse Contexts – Converging Goals. CLIL in Europe* (271-283). Frankfurt Main: Peter Lang.
- Sierra, Cecilia. (2007). Assessing CLIL at primary school: a longitudinal study. *International Journal of Bilingual Education and Bilingualism*, 10(5), 582-602.
<https://doi.org/10.2167/beb461.0>
- Várkuti, A. (2010). Linguistic benefits of the CLIL approach: Measuring linguistic competences. *International CLIL Research Journal*, 1(3), 67–79.
- Wewer, Taina (2013). English language assessment in bilingual CLIL instruction at primary level in Finland: Quest for updated and valid assessment methods. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 18 (2), 76-87.
<http://zif.spz.tu-darmstadt.de/jg-18-2/beitrag/Wewer.pdf>.