

Critical Analysis of Assessments to Young Language Learners in Chinese Primary School EFL Teaching Context

Jiaqi Zhao

School of Social Science & Public Policy, King's College London, London, The United Kingdom

Abstract

Most things that involve young learners are special and language assessment is not excepted. Whether we should assess EFL young language learners' academic performance and how should we do that are two prompts worth considering. This article proposes that young language learners' traits of ineffective communicating determines the necessity of assessing their study performance. It goes on to critically analyze the assessment approaches conducted in the Chinese primary school EFL teaching context and argues that the assessment tools have both merits and demerits. The merit is that formative assessment procedure adopted in primary school EFL classrooms well caters to young language learners' language acquisition stages. Demerit is that the reliability and validity of formative assessments could not be ensured because most of the formative assessments shown in the classroom are not well-designed and arranged. Furthermore, the summative assessment conducted in the teaching context does not align the teaching method shown in the classroom. At last, it concludes that teachers need to adopt the assessment tools according to their teaching method and students' study status. Only in this way, effective assessment procedures can be conducted.

Keywords

Assessment; Young Language Learners; Chinese Primary School EFL Teaching Context.

1. Introduction

Student assessment plays an important and central role in teaching and learning. It could track the progress made by students, help the teacher plan and improve their future teaching and provide information about students' academic performance to the central administration. Considering the great importance of assessment in the teaching practice, a large amount of time is spent in preparing and conducting assessments by teachers in their daily teaching. It is reported that in formal and informal teaching practice, teachers devote an average of 25% of preparation time in creating assessment instruments and observation procedures, marking, recording, and synthesizing students' assessment results [2]. To analyze assessment techniques employed by teachers in ordinary school classrooms, numerous studies are conducted [3, 4, 5]. However, less is still known about the assessment practices used by instructors of young language learners [6]. Most things that involve young learners are special and language assessment is not excepted [7]. Whether we should assess EFL young language learners' academic performance and how should we do that are two prompts worth considering. This article briefly analyzes the reasons why we need to assess young language learners (YLLs) by considering which special traits of young language learners determine the necessity of assessing young language learner's study performance. It goes on to critically analyze the approaches to assessment in my teaching context from the aspects of second language learners' acquisition process, the reliability and validity of the formative assessment, and the correspondence between the teaching approach with the assessment approach. Finally, it concludes that teachers need to adopt the assessment tools according to their teaching

method and students' study status. Only in this way, effective assessment procedures can be conducted.

2. Reason for Assessing Young Language Learners

We should assess Young Language Learners (YLLs) because assessment can serve as an effective tool for us to check students' learning process. It is reported that teachers can just get limited information about young language learners' learning status from classroom communication and observation. Young language learners are ineffective communicators. They seldom raise questions or ask for clarification even when they cannot make sense of what they were taught [8]. Children are distinctively different from adults in this regard because it takes several years for children to become equal interactant in communication and understand that each interactant has the duty of making them understood to others [9]. Even if pupils do not understand the preceding teaching content, they will continue with activities to show respect and to please their teachers. Therefore, it is not unusual to see them mouth the sentences in the textbook, pretending to complete an activity but without understanding or learning. Overall, it could be seen that teachers are not able to accurately estimate students' understanding by simply asking them whether they have understood the teaching content or by observing their classroom performance [10]. To make more accurate inference about students' study status, it is necessary for teachers to conduct various assessments on them. Formative assessments for YLLs can provide crucial information on children's mastery of the knowledge. Research has demonstrated that formative assessments conducted during the class can provide reliable information about pupil's command of English knowledge and level of proficiency [11]. It captures students' every day 'here and now' classroom performance in terms of language quality [12]. Besides, the information about students' immediate progress, a long-term description of students' performances is also needed for teachers. In that case, additional materials will be integrated into formative assessments, including self-assessment forms and profiling forms. These additional materials can exhibit both large and small milestones in each child's achievement. Therefore, it is also seen as an important mean of describing and assessing each individual's linguistic performance. However, the solely formative assessment is not sufficiently reliable. Testing from time to time should also serve as a backup to judgments about each pupil's performance. Test scores can reveal what individuals and groups are able to do, and what they need to acquire in order to progress. It can display pupils' language ability at each study stage. In a nutshell, a combination of various assessment tools could provide a well-rounded description of students' language achievement [13]. Therefore, as teachers, it is of great importance for us to implement both formative and summative assessments to YLLs. Only then, will we get more concise information about students' learning status and plan for our teaching in later stages [14]. However, although assessments are necessary, not all of the assessments conducted fit for each specific teaching context [15]. Therefore, in the following section, I will critically analyze the assessment procedures implemented in my teaching context.

3. Critical Analysis of Assessment Tools in My Teaching Context

Formative assessment procedure conducted in my teaching context well caters to young language learners' language acquisition stages. Krashen and Terrell put it that there are five stages in the second language/ foreign language acquisition process [16]. The first stage is preproduction, in this stage learners experience a silent period and their performance indicators are mostly kinesthetic in nature. The second stage is early speech. Learners in this stage can make some kinesthetic responses in accordance to the language and also some one- or two-word utterances. This stage is followed by the speech emergence, in which language learners can perform one and two-word utterances, plus phrases and simple sentences. The

fourth stage is fluency emergence. Words, phrases, and complete sentences can be uttered by language learners in this stage. The last stage is advanced fluency, in which performance indicators is the near-native level of speech. The assessment procedure that is compatible with communicative approaches to foreign language teaching mirrors the five stages of language acquisition of Krashen and Terrell [17]. In my teaching context, most of the teachers adopt the Presentation-Practice-Production teaching method, which is a weak version of the Communicative Language Teaching (CLT) approach [18]. The assessment procedure in accordance with the PPP method is the 3Rs: recognition, replication, and reorganization [19]. This assessment procedure stands in line with the five stages of second language acquisition. In the recognition part, teachers simply assess students by asking them to do some simple physical responses according to the instruction. For example, we may ask students to choose, point, mark, gesture, and act out the corresponding vocabulary to assess whether students have understood the meaning of these vocabularies. It fits the characteristic that students display in the preproduction stage. In this stage, students just have minimal comprehension. They are not able to verbalize vocabularies but they can show their understanding by drawing or pointing. The followed assessment section of replication can be divided into two parts. In the first part, teachers just require students to replicate the language by activities like "read aloud", "gap-filling" or some other simple question-answer assessments. These kinds of assessment measurements correspond to the early speech stage in the SL/FL development. In this stage, students are able to produce one- or two-word utterances and make use of some keywords and familiar phrases. After fully examine students' command of vocabularies and phrases, a teacher may ask students to make some mini-dialogue with vocabularies and sentence structures that they have already get commanded to foster their speech emergence. Although students in this stage may make some grammatical and pronunciation errors, they start to have a good understanding of the language and can produce some simple sentences. At last, students are assessed by taking part in open practice and making free conversations with others. They are required to demonstrate the ability to take given information and reorganize it into different formats. The assessment in this stage usually includes tasks that lend themselves to work in cooperate with others, such as solving a problem, analyzing and reporting the results of a survey, drawing a timeline, an outline, or a semantic map, role-play a conversation, rewriting a narrative as a dialogue and writing up the text of an oral interview. These tasks can articulate students' highly verbal responses [20]. It mirrors the last stages of FL/SL language development, which are intermediate and advanced fluency. In these stages, students gain an excellent understanding of the language and can make true communications with other competent speakers. Overall, it can be seen that the three R types of assessment we adopted in our classroom teaching roughly stand in line with students' second language acquisition stages. Recognition corresponds to the preproduction stage. Replication mirrors the early production and speech emergence stage. Reorganization is in accordance with the intermediate and advanced fluency stages. Therefore, the formative assessments we conducted in the classroom teaching fit the young language learner's language acquisition process.

The reliability and validity of formative assessments conducted in my teaching context cannot be well ensured. Although some of the formative assessment practices provide opportunities for teachers to trace students' progress in using the target language for meaningful communication, we must pay a close eye to the problems of reliability and validity of formative assessments [21]. Somehow the reliability and validity of formative assessments are "ensured" by the auditability of the procedure (leaving evidence of decision-making processes), by using multiple tasks, by training judges to use clear criteria, and by triangulating any decision making process with varied sources of data (for example, students, families, and teachers) [22]. As in all other forms of assessment, the designers and users of formative assessments must make every effort to structure the ways they design, pilot, analyze, and revise the procedures to

demonstrate and improve the reliability and validity of the formative assessment procedures. Assessments are not in and of themselves valid. Only when procedures of assessments are designed in a clear, open, honest, convincing and demonstrable way, the reliability and validity of them can be guaranteed [23]. However, in my teaching context, most of the formative assessments are not well-designed and arranged. Therefore, their reliability and validity cannot be ensured. Take the role-play performance assessment as an example, when we conduct it in our teaching practice, we always lack the effective process of making criteria, evaluating students' performance, giving feedback, and analyzing students' language achievement. In most cases, to enable learners to engage in open communication, teachers always just set few requirements for students' communication. They just ask them to pretend one student is person A, the other is person B, and then make a dialogue with each other. There is no clear criterion demonstrated about what makes an authentic dialogue and what students are expected to achieve. After students finish making their conversations, it comes to the stage of evaluation. Fluency, accuracy, and appropriacy are the three elements that contributing to an authentic conversation [24]. However, fluency and appropriacy are difficult to assess. What kind of fluency is seen as appropriate is hard to say, same as the appropriacy. As for accuracy, although it is easy for teachers to judge whether there are some grammatical errors in a conversation, the feedback is hard to give. A majority of scholars argue that teachers are not expected to correct students' mistakes in their communication because when we study our first language our parents seldom correct our pronunciation or grammar mistakes [25]. They hold the view that making mistakes is a necessary part in the learning process and a natural outcome in the development of communication skills [26, 27]. Some other scholars argue that teachers need to first judge what kind of errors students make and then decide whether we need to point them out and correct them [28]. Touchie point out that performance errors are made because of carelessness [29]. This type of error does not need teachers to correct. Learners can overcome it by themselves with little effort. In comparison, competence errors are more severe and made due to students' inadequate learning. This kind of error should be pointed out and corrected by teachers. However, the difficulty lies in teachers' hardship in finding out whether students make this mistake because of carelessness or inadequate learning. In young language learners' EFL classes, most of the students are only able to make short conversations. Consequently, most of the vocabularies in their conversations may only appear once. Under that kind of circumstance, teachers always lack the evidence to estimate whether the mistakes students make are performance errors or competence errors. Above all, it can be seen that there is a dilemma in teachers' evaluation and feedback of formative assessment in CLT classrooms. The final stage of the role play assessment is to analyze students' performance in the preceding role-play activity and make the following educational decisions based on that. The problem in this part is that in China, the class size is kind of big. As a result, it is unrealistic to assess each student's conversation. Teachers can only select some students to display their communications in front of the class. In general, students who would like to share their communications are always those who are good at English. Those students who are not confident in speaking or have some obstacles in successfully making conversations will not raise their hands and giving a performance. They just stay silent. Consequently, sometimes, advantaged students' performance may mislead the teacher to think that most of the students have already mastered the knowledge and we can go on to study other content. However, the reality might be that most students have not fully understood what the teacher has taught. Therefore, the teacher's analysis of the formative assessment may not be accurate.

The summative assessment conducted in my teaching context does not align the teaching method we adopted. The last few decades have witnessed a remarkable shift of teaching methodology in EFL teaching contexts [30]. It changes from structural teaching approaches to communicative teaching approaches. The new teaching approach recognizes that the

cultivation of students' communicative competence should be put in a central place in the English language teaching. Communicative competence is consisted of 4 parameters, which are grammatical competence, sociolinguistic competence, discourse competence, and strategic competence [31]. As a result, the assessment of students' progress and achievement in EFL learning should be designed to monitor students' communicative competence rather than solely linguistic competence. However, the summative testing in my teaching context does not experience the corresponding shift. It still focuses on testing students' mastery of linguistic accuracy and discrete language points, rather than on communicative competence. One of the pieces of evidence that can be used to prove it is that our summative language testing always lacks the evaluation of students' spoken English. There are some obstacles prevent us from conducting oral English test. First of all, in China, English teacher is responsible for too many students and we do not have enough time for testing every student's spoken English. Normally, in China, each English teacher teaches 3-5 classes. Every class is consists of 40-50 students. If we implement the spoken English test, every teacher needs to test 120-250 students. That would be very time-consuming. This problem might be tackled by using the Computer-assisted language testing [32]. And the state of computers in my school is also adequate to carry out the spoken English test. However, the problem lies in that in China we lack the domestic online oral testing system which is designed for primary school students. Some people might suggest that we may adopt the testing system developed in western countries like DIALANG test, CAL's Basic English Skills Test, and Pearson's Versant, and so on. However, these soft wares are more inclined to be not suitable to be used in the Chinese teaching context. The socioeconomic and cultural elements in the test content or procedures may pose some problems on the validity of inferences drawn about students' language competencies [33]. For example, children raised in China may be unfamiliar with the spoken topic about the Easter festival or career day that is known by most people in western countries. Another handicap in implementing oral English tests is that, in China, there is a lack of a unified criterion in assessing students' spoken ability. Without a clear set of criteria, teachers' evaluations might become very subjective. In consequence, students may not able to get effective information from the summative assessment. When they compare their performance with students who are assessed by another teacher, they may not able to get an appropriate idea about their study level among their peers. These difficulties inhibit teachers in my school from carrying out oral English tests for students, which indirectly causing the mismatch between our testing techniques and teaching approach. Apart from speaking, the listening test also does not align with the CLT approach. The communicative language teaching approach emphasizes that we learn the language for communication [34]. Therefore, when testing listening, we need to create various language environments to imitate real-life listening environments. In accordance with that, when we test the listening ability of students we may need to record the audio in the noisy background so that imitate the normal English communication environment [35]. However, in reality, when we assess young English learners' listening ability, nearly all schools only provide really clear audio without any noise. That is because they are afraid the existence of noise will affect students' score so that will do harm to the school's academic performance. Secondly, according to Louhiala & Kankaanranta, most of the English conversations are conducted between non-native speakers [36]. Therefore, as English learners, students need to be equipped with the ability to comprehend the English spoken by people who have different linguistic backgrounds [37]. As a result, in listening tests, test developers not only need to include the listening materials created by native British English speakers or American English speakers but also need to add the audios created by English speakers from other countries like Australia, India, Singapore and so on. However, in my school, most of the listening materials used in summative assessments adopt British or American English tapes, therefore lack the variety of different English accents. Further, Rost argues that teachers need to assess students' various listening

skills like the ability of constructing the main idea in a stretch of discourse, predicting subsequent parts of the discourse at conceptual levels, and inferring links between two or more propositions, etc [38]. Nevertheless, listening tests in primary schools still tend to focus on testing student's ability to decode at the phoneme, syllable, and word level, with ignorance of evaluating the top-down listening comprehension strategies [39]. Therefore, it can be concluded that in my school the speaking and listening abilities of students are not tested in the way that CLT commanded, although the teaching approach we adopt is the CLT. Assessment is closely related to instructions in all academic settings. The alignment of the teaching approach and testing procedure will bring some positive washback. However, the mismatch will lead to the negative washback effects for teaching. When students realize that they are not tested in the way they learn, they will become unwilling to cooperate in the curriculum [40]. Generally, they will ignore any teaching content that is not directly related to the test and only pay attention to the knowledge that will be tested [41]. Therefore, teachers should design the summative assessment in accordance with the teaching method.

4. Conclusion

It is necessary to test learners in all academic settings, even for young ones. Assessments of young language learners could help teachers know more about what the student has already known and what they still need to know so that make more effective instructional decisions. However, not all of the assessments conducted are suitable for evaluating young language learners. In my teaching context, our assessment tools to learners exhibit both advantages and shortcomings. The formative assessment procedure conducted in my teaching context well caters to young language learners' language acquisition stages. However, they are not well structured so that their reliability and validity cannot be well ensured. As for the summative assessment, in my teaching context, the summative assessment implemented mismatches the teaching method we adopt. Assessments in themselves are neither good nor evil. They are just simple tools. The success of assessment depends on the effective selection and proper use by teachers. Teachers must take clear eyes to look at all of these tools and carefully select the appropriate assessment techniques according to their specific teaching method and students' study status. Only in this way, the implementation of various assessments could give us the information to trace, assist and reflect upon students' study journey in their learning process.

References

- [1] A. Hasselgreen: Assessing the language of young learners. *Language testing*, Vol. 22 (2005) No.3, p. 337-354.
- [2] L. Cheng, T. Rogers and H. Hu: ESL/EFL instructors' classroom assessment practices: Purposes, methods, and procedures. *Language Testing*, Vol. 21 (2004) No.3, p.360-389.
- [3] T. Rogers: Educational assessment in Canada: evolution or extinction? *The Alberta Journal of Educational Research*, Vol. 37 (1991), No.1, p.79-92.
- [4] R. J. Wilson: Aspects of validity in large-scale programs of student assessment. *Alberta journal of educational research*, Vol. 45 (1998) No.4, p.333-343.
- [5] R. J. Wilson: A model of assessment-in-practice. In *Annual Conference of the Canadian Society for the Study of Education* (Edmonton, Alberta, May, 2000).
- [6] L. Espinosa: Assessment of Young English-Language Learners. *The encyclopedia of applied linguistics*. Vol. 24 (2012) No.5, p.113-146.
- [7] A. Hasselgreen: Assessing the language of young learners. *Language testing*, Vol. 22 (2005) No.3, p. 337-354.
- [8] L. Cameron: *Teaching languages to young learners* (Ernst Klett Sprachen, Stuttgart 2001).

- [9] R. J. Ricard: Conversational coordination: Collaboration for effective communication. *Applied Psycholinguistics*, Vol. 14 (1993) No.3, p.387-412.
- [10] E. D. Peña, A. Iglesia and C. S. Lidz: Reducing test bias through dynamic assessment of children's word learning ability. *American Journal of Speech-Language Pathology*, Vol. 10 (2001) No.1, p.38-54.
- [11] V. F. Gutiérrez-Clellen and J. Kreiter: Understanding child bilingual acquisition using parent and teacher reports. *Applied Psycholinguistics*, Vol. 24 (2003) No. 2, p.67-88.
- [12] L. Espinosa: Assessment of Young English-Language Learners. *The encyclopedia of applied linguistics*. Vol. 24 (2012) No.5, p.113-146.
- [13] J. D. Brown and T. Hudson: The alternatives in language assessment. *TESOL quarterly*, Vol. 32 (1998) No.4, p.653-675.
- [14] K. Shaaban: Assessment of young learners. *English teaching forum*, Vol. 39 (2001) No.4, p.16-23.
- [15] S. Fradd and S. Hudelson: Alternative assessment: A process that promotes collaboration and reflection. *TESOL Journal*, Vol. 5 (1995) No.1, p.5-18.
- [16] S. D. Krashen and T. Terrell: *The natural approach: Language acquisition in the classroom* (Alemany Press, Englewood Cliffs 1983).
- [17] K. Shaaban: Assessment of young learners. *English teaching forum*, Vol. 39 (2001) No.4, p.16-23.
- [18] J. Anderson: Why practice makes perfect sense: the past, present and potential future of the PPP paradigm in language teacher education. *Practice*, Vol. 19 (2016), p.14-22.
- [19] R. Olsen: Classroom questioning, classroom talk. In *Handouts given at the AUB ESL Workshop Larnaca, Cyprus*.(1996).
- [20] Olsen, R: Classroom questioning, classroom talk. (Larnaca, Cyprus, 1996).
- [21] J. D. Brown and T. Hudson: The alternatives in language assessment. *TESOL quarterly*, Vol. 32 (1998) No.4, p.653-675.
- [22] A. Huerta-Macías: Alternative assessment: Responses to commonly asked questions. *TESOL Journal*, Vol. 5 (1995) No.1, p.8-11.
- [23] J. M. Norris, J. D. Brown, T. Hudson and J. Yoshioka: *Designing second language performance assessments* (University of Hawaii Press, Honolulu 1998).
- [24] R. T. Bell: *An Introduction to Applied Linguistics: Approaches and Methods in Language Teaching* (Bastsford Academic & Educational Ltd, Bastsford 1981).
- [25] M. Nemati and M. Taghizadeh: Exploring Similarities and Differences between L1. *International Research. Journal of Applied and Basic Sciences*. Vol. 4 (2013) No.9, p.2477-2483.
- [26] J. M. Hendrickson: Error correction in foreign language teaching: Recent theory, research and practice. *Modern Language Journal*, Vol. 62 (1978) No.8, p.387-395.
- [27] D. Larsen-Freeman: *Techniques and Principles in Language Teaching*. (Oxford University Press, Oxford 1981).
- [28] R. Gefen: The analysis of pupils' errors. *English Teachers' Journal*, Vol. 22 (1979), p.16-24.
- [29] H. Y. Touchie: Second language learning errors: Their types, causes, and treatment. *JALT journal*, Vol. 8 (1986) No.1, p.75-80.
- [30] H. D. Brown: *Teaching by Principles: An Interactive Approach to Language Pedagogy* (Longman, New York 2001).
- [31] M. Canale and M. Swain: Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, Vol. 1 (1980) No.1, p.1-47.
- [32] M. M. Pathan: Computer Assisted Language Testing [CALT]: Advantages, Implications and Limitations. *Research Vistas*, Vol. 1 (2012) No.4, p.30-45.
- [33] E. D. Peña, A. Iglesia and C. S. Lidz: Reducing test bias through dynamic assessment of children's word learning ability. *American Journal of Speech-Language Pathology*, Vol. 10 (2001) No.1, p.38-54.

- [34] J. C. Richards and T. Rodgers: Approaches and methods in language teaching: A description and analysis (Cambridge University Press, Cambridge 1986).
- [35] S. Carlile: Active listening: Speech intelligibility in noisy environments. *Acoustics Australia*, Vol. 42 (2014) No.2, p.90-96.
- [36] L. Louhiala-Salminen and A. Kankaanranta: Language as an issue in international internal communication: English or local language? If English, what English?. *Public Relations Review*, Vol. 38 (2012) No.2, p.262-269.
- [37] J. Jenkins and C. Leung: From mythical 'standard'to standard reality: The need for alternatives to standardized English language tests. *Language Teaching*, Vol. 52 (2019) No.1, p.86-110.
- [38] M. Rost: *Listening in language learning* (Longman, New York 1990).
- [39] J. Field: *Listening in the Language Classroom* (Cambridge University Press, Cambridge 2008).
- [40] N. Elshawa, C. S. Heng, A. N. Abdullah and S. Rashid: Teachers' assessment literacy and washback effect of assessment. *International Journal of Applied Linguistics and English Literature*, Vol. 5 (2016) No.4, p.135-141.
- [41] R. Kiomrs, R. Abdolmehdi and N. Rashidi: On the Interaction of Test Washback and Teacher Assessment Literacy: The Case of Iranian EFL Secondary School Teachers. *English Language Teaching*, Vol. 4 (2011) No.1, p.156-161.
- [42] P. Rea-Dickins and S. Rixon: The assessment of young learners of English as a foreign language. *Encyclopedia of language and education*, Vol. 7 (1997), p.151-161.
- [43] Y. G. Butler and J. Lee: The effects of self-assessment among young learners of English. *Language Testing*, Vol. 27 (2010), p.5-31.
- [44] J. D. Brown and T. Hudson: The alternatives in language assessment. *TESOL quarterly*, Vol. 32 (1998) No.4, p.653-675.
- [45] P. Broadfoot and P. Black: Redefining assessment? The first ten years of assessment in education. *Assessment in Education*, Vol. 11 (2004), p.7-27.
- [46] A. Huerta-Macías: Alternative assessment: Responses to commonly asked questions. *TESOL Journal*, Vol. 5 (1995) No.1, p.8-11.