

## BOOK REVIEW

### Assessing Young Language Learners

Misuzu "Zuzu" Emura

#### Assessing Young Language Learners

**Author: Penny McKay**

**Series: Cambridge Language Assessment Series**

**Series Editors: J. Charles Alderson & Lyle F. Bachman**

**Cambridge University Press (2006)**

**Pp. +388 pages**

**ISBN 0-521-60123-1**

The number of children who take English lessons in their early years is increasing, especially in countries such as Japan, Korea, and Taiwan. In Japan in particular, the government has been working on the implementation of English as a mandatory subject in elementary schools, and trials have already started nationwide. With this new trend, one needs to consider important issues, particularly those related to evaluating children on their English. How should teachers evaluate children? What are some traditional and new methods? What should teachers keep in mind when evaluating children? What do teachers have to know about assessment for children? *Assessing Young Language Learners* by Penny McKay addresses these issues.

The book fills an important gap in the current literature on assessing young learners. For example, Arthur Hughes in his classic book on testing (1989) devoted only one chapter to young learner assessment. Hughes provided brief summaries of assessment approach, demand, and techniques and contains sample testing activities. Although he succeeded in providing a succinct overview of young learner assessment and useful information related to EFL settings, the chapter is rather brief. In another well known text, *Teaching English to Children* by Wendy Scott and Lisbeth Ytreberg (1990), there is only one small paragraph describing assessment. Penny McKay's new book, *Assessing Young Language Learners*, is one of the first to deal exclusively with this topic. It is a very useful book for those interested in

learning how to assess children with positive backwash. It also provides information on assessment principles and methods specific to children. It would be especially useful for those with little background knowledge about and experience in teaching children. A point that becomes clear in the book is that assessing students is not only for the students. It also provides teachers with relevant information on their teaching. Teachers can check the effectiveness of their lessons and adjust their lessons based on the information given by assessment. This book is thus about both testing and teaching. One point of caution is that information in this book is more applicable to children in ESL settings. It would be better if there are additional descriptions related to EFL young learners.

Chapter One, "A Special Case for Young Learner Language Assessment," introduces the general knowledge necessary for effective assessment of young language learners. The term "young language learners" refers to children between the ages of five and twelve who are learning a foreign or second language in language programs around the world. In terms of growth, literacy, and vulnerability, young language learners are different from adults and require special attention, and these special characteristics should be taken into consideration when assessing children. Young language learners are still "growing cognitively, socially, emotionally, and physically" (McKay 2006, p. 24) while they are learning a foreign or second language. For example, children are usually able to concentrate only 10 to 15

minutes. Children who are younger than 8 years old have not developed meta-language abilities so talking about language is not appropriate. Young language learners are developing literacy skills and understandings in their first language at the same time as they are learning a foreign or second language. Since children are particularly vulnerable, assessment experiences should “help them succeed” and “feel good about themselves” (p. 14).

The author’s explanation about children’s special characteristics reminds us of the importance of knowing about children’s physical and mental growth in teaching and testing. In my view, knowing how to nurture children is something that comes from both knowledge of nurturing and hands-on experience with children. It is experience that helps teachers deal with the assessment of young learners flexibly.

As with other educational activities with young learners, assessment should be based on “the beliefs and practices of elementary education” (p. 15). According to constructivist learning, a current theory of elementary education, “children learn through a two-way communication of ideas” (p. 16) with, for example, peers and teachers. “Learning takes place when children have the opportunity to visit and revisit the knowledge in new contexts and over time” (p. 17). The author also refers to the impact of assessment on children. It should be noticed that assessment can have a negative impact on children when, for example, it establishes power relationships between teachers and students.

Chapter Two, “Young Learners And Language Learning,” explores the characteristics of language learning by young learners. Understanding stages of language development, in addition to understanding stages in children’s development, as addressed in the previous chapter, is important for effective assessment of children’s language learning. Children usually acquire “language use ability,” or learn a foreign or second language using both sociocultural and cognitive processes. Thus, the ideal conditions for language learning are those which maximize the above two processes. “Language use

ability” is defined as “the ability to use the language for the purpose of achieving a particular goal or objective in a particular situation” (adapted from Bachman and Palmer 1996, p. 44, cited in McKay 2006, p. 27).

In the sociocultural process, children are “learning how to mean” (p. 29), or how to use the language communicatively in the new speech community, as well as learning “the cultural codes (the rules of interaction) of the language” (p. 29). This is very much akin to socialization theory (Schieffelin and Ochs, 1984, cited in Schieffelin and Ochs 2008, p. 5), which asserts that linguistic and sociocultural development processes intersect. Children acquire language in the process of learning to behave in ways that are acceptable in the society. In this manner, young learners develop new identities in their second language. Successful development also depends on the way in which their first language is treated. McKay stated in Chapter One that young language learners are still developing their identity in their first language. However, if the first language of a student is devalued in a classroom setting, his identity in his first language tends to harm the development of his second language and new identity. This demonstrates that language and identity are closely related. These factors reconfirm the need for special support for children in their language learning and assessment. The author also pointed out the significance of maintaining children’s first language skills, which teachers should encourage.

One of the cognitive processes discussed is a “dual-mode system” by Skehan (1998, cited in McKay 2006, pp. 35-38). Skehan’s “modes” are a formulaic process through which learners develop chunks of language and idioms and a rule-based process through which learners develop grammar rules and vocabulary. It is said that the optimal condition for language learning is that children use both modes in a balanced manner. In general, however, it is difficult for children in EFL settings to develop their formulaic system because they tend to be exposed to the target language only in their classrooms. Teachers in EFL settings should realize this fact and consider how to

expose the children to more English. For example, teachers may encourage parents to let their children watch quality English TV programs for children by sending home a list of recommended programs. There are programs shown worldwide such as *Sesame Street*, and in countries such as Japan, there are locally developed English TV programs for children.

Chapter Three is entitled “Research into the Assessment of Young Language Learners.” Research into young learner assessment is still a new field of study, but is more evident currently. The research is important because it identifies what is relevant, what is missing, and what needs to be improved in young learner assessment. Knowing about the research is also important for teachers and assessors. Assessment research has been hindered by “variability of programmes, lack of consensus about proficiency and variable teacher expertise” (p. 95). Both foreign and second language programs vary in, for example, the starting age when children begin to study a foreign language and the amount of knowledge about TESL that teachers have. Few language teachers for young learners are highly skilled in assessment because it has not been highly valued and opportunities for specialized training are lacking. Therefore, it is difficult to gain a common view of language proficiency.

In the remainder of the chapter, research into the assessment of young language learners and the findings are discussed under four respective research purposes: “To investigate and share information about current assessment practices, to find ways to ensure valid and fair assessment tasks and procedures, to find out more about the nature of young learner language proficiency and language growth, and to investigate and improve the impact of assessment on young language learners, their families, their teachers and their school” (p. 65). Some of the findings are very informative for classroom teachers. One of them is that the assessment task itself may limit judgment of the child’s performance due to “the characteristic of the setting, the nature of the input, the nature of the expected response, and the relationship between the input and the re-

sponse” (p. 83) in the task. For example, if a teacher selects for assessment a language task that has been used several times in her classroom, the children would have rehearsed it and so their test performance may not show their true proficiency. Using repeated tasks may be effective in teaching, but it does not work in the same way in assessment. This chapter reminds us of the importance of research and its connection to classroom practices. Researchers raise questions and provide ideas for the practitioners, and their findings can confirm, or call into question, their practices.

Chapter Four, “Assessing Language Use through Tasks,” focuses on task-based assessment which is a kind of performance assessment using tasks for assessment purposes. Performance assessment is a general term for an assessment approach which “involves either the observation of behavior in the real world or a simulation of a real-life activity” (Weigle, 2002, cited in McKay, 2006, p. 98). The goal of performance assessment is to evaluate the learner’s use of the target language. The learner’s performance tends to be observed and evaluated as a whole, including the elements of language use such as vocabulary and grammar. In other words, language use tasks, that require children to use communicative language in a particular situation, are preferred over discrete-point items. McKay does allow, however, that many language use activities including discrete-point assessment items, can be used as assessment tasks if selected and designed carefully. Her principles for the selection and use of assessment tasks procedures can be summarized as follows: “Select tasks and procedures to suit the characteristics of young learners, assess the learners’ most relevant abilities for language use, make assessment choices that ensure that assessment is valid and reliable and has a positive impact, ‘bias for best’ but maintain high expectations, engage learners intellectually, and draw from multiple sources of information” (pp. 109-111).

What should be noticed in Chapter Four is that researchers are exploring new views of validity, reliability, and impact that apply to classroom assessment. In this con-

text, validity refers to identification of children's progress and a commitment to language learning during assessment. Reliability refers to the successful collection of observation data from assessment. It is complex and difficult to reach these new, appropriate views since various factors and viewpoints are involved. I agree with the author that further research is necessary in this area.

Chapter Five, "Classroom Assessment of Language Use," explores assessment which is designed and conducted by classroom teachers and not by others who may not know the target students and their language learning abilities. Classroom assessment may be either "formative" or "summative," and the process may be either "on-the-run" or "planned." McKay defines these terms in the introduction of this chapter. Then she goes on to observe that administrative requirements, such as the use of external tests, the lack of trust in classroom assessment from parents and students, and the lack of training and experience in assessment by teachers may cause teachers to have less of a commitment to this form of assessment. I agree with the author that there is "a critical need for classroom teachers to have training in assessment as part of their pre-service education" (p. 144). Since assessment is widely embedded in teaching and learning, the presence and/or absence of the knowledge may affect the quality of the classroom lesson as well as assessment. It may be wise that teachers who do not have knowledge or experience in assessment have an opportunity to work with colleagues who do.

In general, classroom assessment consists of three phases: design, operationalization, and administration. The principles and practices of these three phases have to be modified to fit to the nature of the teacher's assessment which is "reflective, strongly driven by the teacher's beliefs about education and directed towards individual learning" (p. 174). According to Rea-Dickins (2001, cited in McKay 2006, p. 174), modification of the classroom assessment process could include some teaching practices such as "explanations and introductions, scaffolding and feedback, involvement of other teachers and

different uses of assessment information" (p. 174). This clearly illustrates that assessment for children is tied to teaching and learning.

Classroom assessment strategies include incidental observation, planned observation, conferences, portfolios, contracts of work, projects, self- and peer-assessment, and classroom tests. Teachers combine these strategies when engaging in on-the-run assessment or planned assessment. Principles for the selection of assessment tasks and procedures described in Chapter Four can be a good reference for the effective use of these strategies. A contract of work is one of the other interesting strategies mentioned in this chapter. A teacher and a student negotiate a contract, and it becomes the student's responsibility to complete the work by each due date. Both the process to complete the tasks and the outcome are assessed. This strategy nurtures students' autonomy and responsibility. The difficulty of work can be adjusted to the individual student's level since the contract is set between the teacher and individual student. This strategy helps students mature and grow as language learners, and is a good sample of how assessment can become "an essential part of teaching and learning, rather than a separate process" (p. 141). If a teacher has a large class, setting individual contracts with each student may be difficult, but it can be solved by organizing a group project and having a contract with each group.

Chapter Six, "Assessing Oral Language," is devoted to a discussion of assessment of children's oral language. It describes similarities and differences between spoken and written language. Oral language is the central component of language acquisition for young learners. They always "learn best primarily through oral language" (p. 177). Stable oral language contributes to sound literacy development of second language learners. Reflecting the demand for academic use of the target language, the trend of oral language activities is toward content-based curricula. This trend leads to the conclusion, in terms of assessment activity, that the task and its demands should be analyzed rather than the individual skills used to complete the task. The author notes

that at the age of eight or nine years old “written language becomes more supportive of oral language development” (p. 177). She also pointed out, in relation to the difficulty in oral language assessment of children, that the ability to use language for communication is gradually developed so the ability may not be fully developed when children are assessed. Such knowledge is relevant to the appropriate selection of tasks.

The scope of oral language to be assessed should be based on the curriculum. Teachers and assessors should understand the factors to be considered in assessment task selection, be aware of the quality of classroom interaction, and understand how to assess pronunciation, vocabulary, and grammar in oral language. For example, regarding the factors to be considered in the assessment task selection, it is stated that the written text should not be used in oral assessment when children’s literacy skills are not yet secured.

In the remainder of this chapter, various “types of oral language assessment tasks for young learners” (p. 198) are presented with many examples. In my view, the most interesting is portfolios, or recorded samples of students’ oral recordings. It seems effective because recordings of performance enable a detailed analysis of progress and thus thorough feedback can be given to the learners. It would be helpful to know the preferable frequency and length of recordings, but this information is not mentioned in the chapter.

Chapter Seven, “Assessing Reading and Writing,” is dedicated to the assessment of children’s reading and writing abilities. The author identifies three key elements influencing children’s literacy development: “first language literacy, cultural and background knowledge, and oral language” (p. 219). Children usually construct literacy in their second language by transferring their literacy and comprehension skills which have been already developed in their first language. Therefore, the lack of first language literacy may cause weak abilities of reading and writing in the foreign/second language. According to constructivists, when they read, children use their cultural and

background knowledge to bring meaning to the text. As explained in the previous chapter, children’s literacy skills in the target language are based on their oral language use ability which embeds their linguistic and cultural knowledge. In regard to the key elements, the author also mentions that the elements should be considered when teachers and assessors assess children’s literacy skills in the foreign/second language. Knowledge of these elements is necessary for appropriate task selection, performance judgment, and planning subsequent teaching.

As with assessment of oral language, assessment of reading and writing should be aligned with curriculum. Writing ability, since writing varies from e-mail interaction to expository essays entails different kinds of ability including social, cultural, and cognitive elements. Reading ability is characterized by three cueing systems which are used by readers: graphophonic, syntactic, and semantic cues. As reading is a receptive skill and can be assessed only indirectly through other skills, teachers and assessors should be careful not to require high levels of writing or speaking on reading assessment tasks.

Briefly referring to the assessment of reading and writing using, the author discusses its advantages and disadvantages as well as the possibility of the computer skills as a target assessment skill. This is the only section in the book in which the use of computers is mentioned. Using computers in classrooms has become common. It is probable that classroom teachers have been using computers for their language use tasks. Also, a number of computer-based language tests and activities have recently been developed and are now available for sale. More information on the use of computers in relation to young learner’s language assessment would have made this book more informative.

Chapter Eight, “Evaluating Young Learners’ Performance and Progress,” mainly discusses scoring. The author begins by briefly describing classroom formative assessment, which is one of the most common forms of assessment for young learners. Formative assessment refers to ongoing assessment in which teachers observe child-

ren's performance regularly to see how well the student is doing and offer feedback. The ultimate purpose of classroom formative assessment is to improve learning. Compared with more formal forms of assessment, the results of formative assessments tend to vary among teachers who are involved in the evaluation. One reason is that teachers are likely to rely on internalized criteria or their own constructs. Furthermore, the process of conducting formative assessment is not simple. When teachers and assessors conduct assessments, they first have to decide on the scoring method, create appropriate rubrics, and clarify the scoring procedure. Generally, the criteria should be clearly defined so that clear and objective decisions on children's performance can be made.

This process, however, cannot be fully applied to the assessment of language use because successful or unsuccessful language use depends on the interaction of different elements, not on one single element. Thus, the criteria for language use tasks should be less-specific. For example, in case of correct/incorrect criteria, "Did the child describe what she did on the class outing successfully?" (p. 268) is better than "Does the child answer 'Yes, that is a dog?'" (p. 267)

McKay provides principles and descriptions for designing quality rubrics. A wide range of factors have to be considered to make rubrics reliable. For example, anything that gives some learners an advantage over others should be excluded. Thus, a criterion expecting that children will write a page on a football game is inappropriate because boys or football fans may have an advantage. This small point may be easily overlooked even if teachers and assessors check rubrics carefully. I concur with the author in suggesting several tryouts in the process of constructing a scoring rubric. Checking by a third person may also be efficient.

Scoring rubrics can be observation checklists, criteria sheets, and holistic or analytic rating scales. The first two are commonly used for classroom assessment. Rating scales can also be used as reporting scales. Also in this chapter, the author mentioned steps to construct a scoring rubric

and procedures to maximize scoring reliability.

In Chapter Nine, "Testing Young Language Learners Through Large-Scale Tests," the author discusses large-scale standardized tests for young learners. Standardized tests are relatively easy to administer and score. Some people believe that their scores allow teachers and assessors to compare individual students as well as schools, and the results can also be a part of an official report as learning evidence. However, standardized tests often conflict with the unique characteristics of children, as well as approaches to teaching and learning, introduced in Chapter One and Two. McKay pointed out that comparing individual students has little value for children because of their "vulnerability to failure" (p. 350). Children's "lack of maturity may lead to misconceptions about the test requirements" (p. 350). Large-scale standardized tests also do not generally give children and teachers immediate feedback. Immediate feedback is important for children's further language learning and for teachers' further language teaching. In other words, unless there is immediate feedback, the teacher may not be able to adjust teaching activities to address the current language needs of the student.

Standardized tests, thus, need to be developed so that they can be "valid, fair and motivating" (p. 350) for children and so that they are consistent with the principles for the selection of assessment tasks and procedures in Chapter Four. In this chapter, the author discusses development of large-scale tests from the view point of maximizing the usefulness of the test, taking The Cambridge Young Learners English Tests as an example.

The author points out pitfalls and the impact of requiring children in ESL settings to take standardized tests designed for first language speakers. It is obvious that such tests are unfavorable to second language children since the reference points of the test are first language speakers. People involved in the education of young learners should consider and carefully judge the appropriateness of large-scale standardized tests on children. However, McKay avoids simply saying that performance-based as-

essment is appropriate and the standardized tests are inappropriate for children. Purpose and backwash of a test should always be considered also (but which classroom teachers may forget when their attention is drawn to other administrative tasks such as reporting to authorities). Furthermore, there is a need for parents to be more aware of the advantages and disadvantages of standardized tests because they tend to put trust in the results of, for example, state-wide and nation-wide standardized tests.

Chapter Ten, the last chapter, is “The Way Forward.” Young learner assessment is conceptually based on theories of general education, language education, and assessment. In its scope, young learner assessment consists of foreign and second language assessments. Therefore, what she calls the theories, frameworks, and connections need to be enhanced in order for further development of young learner assessment as a coordinated entity. The enhancement of theories, frameworks, and connections requires the reinforcement of professional understanding and research which McKay sees as the way forward for further development of the assessment field. She summarizes some of the important issues that she feels require attention some of which have described in previous chapters. One of the issues is how to connect foreign and second language assessment of children. She highlights the connection between the two in terms of research and encourages researchers working in ESL and EFL environments to find common research issues regarding general assessment of young learners. The more researchers target both foreign and second young language learners, the clearer it will be what can be shared and what

should be separated in the two fields. Identifying the shared and unshared areas between the two assessment fields is one of the keys for the further development of young language learner assessment.

To conclude, *Assessing Young Language Learners* addresses concerns in what is still a new field. Thus, consolidation of understandings and research in this area is necessary. Assessment of young language learners is tied in with children’s growth and the language teaching and learning. Therefore, teachers have to be aware of children’s characteristics and how their general growth and language learning is different from adults. Assessment should support teaching and learning, and give children positive backwash. Children should be given a chance to show what they are capable of in language use. Teachers need to collect relevant data from children’s performance in order to accurately understand children’s language use ability and adjust their teaching as needed. Classroom assessment, which can be done as a part of everyday teaching and learning, is recommended for young learners. Task-based assessment, or other kinds of performance assessments are particularly suitable for young learners.

I recommend *Assessing Young Language Learners* to any teacher who is interested in assessment of young learners. Readers will benefit from McKay’s in-depth review of the field as well as her concrete examples of rubrics, assessment tasks, and frameworks, which help readers’ understanding. Teachers would also benefit from the basic principles of assessment introduced in the book, which may serve as a useful reference for their own design and modification of classroom assessment tools.

## References

- Hughes, A. (1989). *Testing for Language Teachers*. Cambridge, UK: Cambridge University Press.
- Ochs, E & Schieffelin, B. (2008). Language socialization: An historical overview. In P. A. Duff and N. H. Hornberger (Eds.), *Encyclopedia of language and education* (2<sup>nd</sup> ed., Vol. 8, pp. 3-16). NY: Springer Science+Business Media, LLC.
- Scott, W. & Ytreberg, L. (1990). *Teaching English to Children*. NY: Longman.

[this page is intentionally blank]