

Trends, Issues, and Challenges in Assessing Young Language Learners

Marianne Nikolov

Abstract This introductory chapter aims to achieve multiple goals. The first part outlines the most important recent trends in early language learning, teaching and assessment and frames what the main issues are. The second part discusses the most frequent challenges policy makers, materials designers, test developers, researchers and teachers face. The third part introduces the chapters in the volume and explains how they are embedded in the trends. The last part suggests ideas for further research and points out some implications for educational and assessment practice.

Keywords Social dimensions • Framing assessment • Research on early language programs • Construct • Testing projects • Examinations • Uses of tests • Areas for further research

1 Introduction

The aim of this chapter is to offer insights into recent trends, emerging issues and challenges in the field of teaching and assessing young language learners and to outline which aspects the chapters in this volume highlight in various educational contexts. Recent developments are best viewed from a perspective of innovation (Davison, 2013; Davison & Leung, 2009; Kennedy, 2013). This approach to early language learning and assessment as a larger system (Markee, 2013) may allow us to understand how innovation works at various levels and how the classroom, institutional, educational, administrative, political and cultural level subsystems interact. A narrow focus on certain aspects of assessment practice is limited; innovation and change are necessary in the whole of assessment culture (Davison, 2013). The chapters in the book explore global issues and how they are embedded in local contexts. The findings may not directly translate into other situations, therefore, readers are expected to critically reflect on them and analyze how the lessons learnt can be relevant.

M. Nikolov (✉)

Institute of English Studies, University of Pécs, Pécs, Hungary

e-mail: nikolov.marianne@pte.hu

© Springer International Publishing Switzerland 2016

M. Nikolov (ed.), *Assessing Young Learners of English: Global and Local Perspectives*, Educational Linguistics 25, DOI 10.1007/978-3-319-22422-0_1

Some of the studies included in the book fall into the narrow field of language testing and share information on frameworks and the time-consuming test design and validation processes of test development. Other chapters go beyond these domains and discuss results of large-scale national studies and smaller-scale classroom projects. The common denominator in these explorations reflect stakeholders' local needs. Alternative approaches to assessment, for example, peer and self-assessment, diagnostic testing, assessment *for learning*, and ways in which young learners' individual differences interact with test results are also discussed in depth. It is hoped that a wide readership will find food for thought in the book.

Specific uses of terms are clarified in the chapters and a list of acronyms is also included at the beginning of the volume. The ages covered by the term *young learners* in the chapters range from 6 to 12 or so; children in the projects learn a foreign language in the first 6 years of their studies. The use of key terms needs clarification. In this volume we follow the widely accepted tradition of using *assessment* and *testing* interchangeably, although we are aware that *assessment* is often used "as a superordinate term covering all forms of evaluation" (Clapham, 1997, xiv). The majority of sources on young learners tends to follow this tradition and this is what authors of this volume also do.

2 Main Trends in Early Language Learning and Assessment

2.1 The Social Dimension

These days, millions of children learn a foreign language (FL), most often English (EFL), in public and private schools around the Globe. The recent dynamic increase in the number of young language learners in early language programs is embedded in larger trends. Firstly, more and more people learn English as a lingua franca, aiming to achieve useful levels of proficiency in English, the means of international communication. Today, English is increasingly perceived as a basic competence and an asset for non-native speakers of English to succeed in life. Since access to English as a commodity is often limited, early language learning has a special social dimension. Proficiency in English can empower learners and early English may offer better access to empowerment over time.

These trends have important implications for curricula, assessment and equity. On the one hand, in many countries not all children have access to equal opportunities to start learning English at a young age. It has been widely observed that parents' socio-economic status plays an important role in access to English and choices of programs. In many places around the world parents empower their children by finding earlier, more intensive and better quality programs for their offspring. For example, an article in *The Economist* (December 20th 2014, p. 83) reported that 80 % of students at international schools around the world are locals because their parents want them to study later in an English speaking country and they believe

that earlier and better quality English learning opportunities allow them to do so. “When people make money, they want their children to learn English, when they make some more money, they want them to learn in English.” As a result of high investment in children’s learning of English, highly motivated parents make sure that their children learn English in the very best programs, as is documented by the recent interest in content and language integrated learning (CLIL). This new development poses new opportunities and challenges for assessment.

Parents would like to have evidence of their children’s proficiency in English at the earliest possible stage. This need has resulted in several internationally acknowledged external proficiency examinations offering young learners opportunities to take age-appropriate exams and document their level of proficiency. How these test results are used and why may vary (see e.g., Chik & Besser, 2011). Parents who want their children to get language certificates assume that the proficiency achieved at an early stage of language learning will be automatically maintained and built on over time.

Another line of test development is documented by national and international projects implemented in more and more countries as early language learning is becoming more the norm than the exception. Certain phases and steps of the arduous process of test development are discussed in five chapters in this book. Needs vary to a large extent, as the studies indicate and the uses of test results are also very different. Some projects are initiated by policy makers in order to establish a baseline or for gatekeeping purposes, others result from more bottom up initiatives based on local needs.

2.2 An Inkblot Test or a Puzzle: ‘The Younger The Better’ vs. ‘The Slower’, or How and Why?

The boom in early language learning is due to more and more parents’ and decision makers’ belief in *‘the younger the better’* slogan; young children are expected to outsmart older starters simply by starting at a younger age. The overwhelming optimism and overconfidence characterizing early language programs is well known in research in the social sciences and behavioral economics (Kahneman, 2011). Wishful thinking is supported by evidence in favor of one’s beliefs. The approaches to interpreting data on how young learners develop and what realistic expectations are after several years of exposure to L2 can be explained by two metaphors: an inkblot test and a puzzle (Nikolov, 2013). In the first approach, interpretations are projected into what there is in the data and they are biased by emotions, expectations, beliefs, etc. In the second approach, all data contribute to a better understanding of the whole as well as the small components of the larger picture. Although the puzzle metaphor is also limited, as it supposes a single correct outcome, it represents a more objective, scientific, and interactionist approach. The chapters in this volume hopefully add meaningful pieces to the picture of early language learning.

In recent years, concerns have been voiced about early learning of a foreign language both in national and local programs, as evidence on *'the younger the slower'* has emerged (e.g., deBot, 2014; García Mayo & García Lecumberri, 2003; Muñoz, 2006; Nikolov & Mihaljević Djigunović, 2006, 2011). Many experts have emphasized that focusing on starting age as the key variable is misleading in foreign language contexts. The age factor is not the main issue. There is a lot more to success over time. The quality and quantity of early provision, teachers, programs, and continuity are more important (Nikolov, 2000; Singleton 2014). Also, it is now widely acknowledged and documented that maintaining young learners' motivation over many years is an unexpected challenge emerging in most contexts: the earlier L2 learning is introduced, the sooner typical classroom activities and topics become boring for young learners. This is one of the reasons why there is a growing interest in integrating content areas and moving towards content-based curricula, which, in turn, pose further challenges in both teaching and assessment.

More and more stakeholders realize that offering early language learning opportunities is only the starting point. Issues related to curricula, teacher education, monitoring progress and outcomes over the years, and transition across different stages of education persist and pose new challenges (e.g., Nikolov 2009a, 2009b, 2009c; Rixon, 2013). In fact, the same old challenges are reemerging in a cyclic manner, as was implicitly predicted by Johnstone (2009).

An important shift can be observed from an emphasis on the *'fun and ease'* of early language learning to standards-based measurement of the outcomes in the target language (L2; e.g., Johnstone, 2009; Rixon, 2013, 2016 in this book). The shift towards standards is not limited to foreign language programs; it is an international trend in educational assessment for accountability in public educational policies in all subjects and competences.

2.3 Research on Early Language Learning and Teaching

Test results indicating how children progress and what levels they achieve in their L2 at the end of milestones in education are often used as one of several key variables interacting in the process of early foreign language learning and teaching. In other words, it has been realized that early language learning is not at all a simpler construct than language learning of older learner. Recent research projects apply all kinds of L2 tests as one of many data collection instruments in order to answer larger research questions, as they aim to build and test models of early foreign language learning. An important area of explorations concerns how young learners' individual differences, including attitudes, motivation, aptitude, anxiety, self-perceptions, self-confidence, strategies, etc., contribute to their development in their L2 (Bacsa & Csikos, 2016; Mihaljević Djigunović, 2016; Nikolov, 2016 all in this book). Another important avenue of explorations

gaining ground looks into how learners' first (L1) and other languages interact with one another over time (e.g., Nikolov & Csapó, 2010; Wilden & Porsch, 2016 in this volume).

Yet another important line of research examines how different types of curricula contribute to early language learning. Traditional FL programs are often supplemented or substituted by early content and language integrated learning curricula (CLIL). Overall, these research studies aim to find out not only what level of proficiency children achieve in their L2, but they also want to offer explanations as to how and why. The type of curriculum has important implications for the construct as well as for the way the curriculum is implemented in the classroom. On the one hand, some recent studies focus on the relationships between contextual factors and classroom processes. Highly age-appropriate innovative approaches, including *assessment for learning* (AfL, Black & Wiliam, 1998), diagnostic (Alderson, 2005; Nikolov, 2016), peer and self-assessment are examined in ELL contexts (Butler, 2016; Hung, Samuelson & Chen, 2016 in this volume). On the other hand, some research projects aim to find out how and to what extent different curricula contribute to L2 development.

In recent years, the field of early language learning research has grown remarkably. Many new studies have been published in refereed journals. (See for example Special Issues of *English Language Teaching Journal*, 2014 (3) edited by Copland and Garton; *International Journal of Bilingualism*, 2010 (3) edited by Nikolov; and *Studies in Second Language Learning and Teaching*, 2014 (3) edited by Singleton.) A range of books and research studies are available on the early teaching and learning of modern foreign languages offering food for thought for decision makers, teachers, teacher educators and researchers. (For critical overviews see e.g., Murphy, 2014; Nikolov & Mihaljević Djigunović, 2006, 2011.) Publications on large scale surveys give insights into the big picture (e.g., Edelenbos, Johnstone, & Kubanek, 2007; Emery, 2012; Garton, Copland & Burns, 2011; Rhodes & Pufahl, 2008; Rixon, 2013, 2016 in this volume). Excellent handbooks offer classroom teachers guidance on age-appropriate methodology and assessment (e.g., Cameron, 2001; Curtain & Dahlberg, 2010; Jang, 2014; McKay, 2006; Pinter, 2006, 2011).

The growing body of empirical studies (e.g., Enever, 2011; Enever, Moon, & Raman, 2009; García Mayo & García Lecumberri, 2003; Muñoz, 2006; Nikolov 2009a, 2009b) applies some kinds of tests, as they implement quantitative or mixed research methods (Nikolov, 2009c) and analyze test results in interaction with other variables. Testing young language learners' progress over time in their classrooms and their proficiency at the end of certain periods are often the aspects of studies. Thus, the assessment of young learners has become a central issue in early language learning research and daily practice (Butler, 2009; Inbar-Lourie & Shohamy, 2009; Johnstone, 2009; McKay, 2006; Nikolov & Mihaljević Djigunović, 2011; Rixon, 2013), as chapters in the present volume indicate. As Rixon (2016) put it in the title of her chapter, these developments in assessment represent the 'Coming of Age'.

3 Challenges in Early Language Learning, Teaching, and Assessment

3.1 *The Construct and Frameworks of Assessment*

The trends outlined above have important implications for the construct. Assessment of young language learners in early learning contexts was first brought to the attention of the testing community as a bona fide domain in a special issue of *Language Testing* edited by Pauline Rea-Dickins (2000). In her editorial she emphasized an array of issues: processes and procedures teachers used in their classrooms to monitor their learners' development and their own practice, the assessment of young learners' achievement at the end of their primary education, and teachers' professional development. At that time high hopes were typical in publications on early language programs and hardly any comparative studies were available on younger and older EFL learners. However, the field was characterized by variability and diversity, as Rea-Dickins pointed out (p. 119).

Over the past 15 years, the picture has become even more complex for several reasons:

- (1) The constructs (Inbar-Lourie & Shohamy, 2009; Johnstone, 2009) cover various types of curricula;
- (2) More evidence has been found on young learners' varied achievements and on how their individual differences and contextual variables, including teacher-related ones, contribute to outcomes over time (for an overview see Nikolov & Mihaljević Djigunović, 2011).
- (3) Accountability poses a recent challenge as standards-based assessment in early language programs has been introduced in many educational contexts.

The emergence of accountability in early language learning is not an unexpected phenomenon. As Johnstone (2009, p. 33) pointed out, the third phase of early learning became a "truly global phenomenon and possibly the world's biggest policy development in education. Thus, meeting 'the conditions for generalized success' becomes an awesome challenge." The task is to establish to what extent and in what conditions early language learning can be claimed to be successful in a range of very different situations where conditions vary a lot. Stakeholders are interested in seeing results. What can young learners actually do after many years of learning their new language? An important challenge for researchers concerns what curriculum is best and what realistic age-appropriate achievement targets are included in language policy documents. Once curricula are defined, and frameworks are in place, the construct and expected outcomes have to be in line with how young learners develop and how their motivation can be maintained over years.

Although early language learning is often seen as a simple proposition (start learning early), a lot of variation characterizes models according to when programs start, how much time they allocate, what type of curriculum and method they apply, who the teachers are, and how they implement the program. In the European contexts (Edelenbos, Kubanek, & Johnstone, 2007; Johnstone, 2009), three types of curricula are popular: (1) awareness raising to languages; (2) traditional FL programs

offering one to a few classes per week, and (3) content and language integrated learning (CLIL) curricula where up to 50% of the curriculum is taught in the L2. The first type does not aim to develop proficiency in an L2; the other two usually define L2 achievement targets. CLIL programs have become popular in Europe, Asia and South America. CLIL is typically taught by non-native teachers of English, and ‘could be interpreted as a foreign language enrichment measure packaged into content teaching’ (Dalton-Puffer, 2011, p. 184). In most schools ‘CLIL students nearly always continue with their regular foreign language program alongside their CLIL content lessons’ (p. 186). What the construct is in these two programs is one of the main challenges in early language learning research. As has been indicated, the increased interest in early CLIL programs is due to growing evidence that in traditional (type 2) programs children develop at a very slow rate and many of the motivating activities lose their appeal and soon become boring. Therefore, integrating not only topics from the main curriculum (as in type 2 programs), but also teaching subjects in the target language is supposed to result in killing two problems with one stone: a focus on intrinsically motivating content also offers opportunities to acquire L2 skills in all four skills. This means that both content and language have to be assessed.

As for the construct of early language learning, Inbar-Lourie and Shohamy (2009) suggest that different types of curricula should be seen along a continuum between programs focusing on language and content. Awareness raising is at one end, FL programs somewhere in the middle, and CLIL and immersion at the other end. They propose that in early language programs language should be “a tool for gaining knowledge and meaning making and for developing cognitive processing skills” (p. 91). In this framework, L2 is closely linked to the overall curriculum and learners’ L1, and the larger view of assessment culture where assessment is a means to improve. Their proposed framework integrates widely accepted principles of age-appropriate classroom methodology as well as assessment. The challenges concern how curricula define the aims set for language and content knowledge, and cognitive and other abilities and skills.

Achievement targets in L2 tend to be modest in early language programs. Young learners are not expected to achieve native level (e.g., Curtain, 2009; Haenni Hoti, Heintzmann, & Müller, 2009; Inbar-Lourie & Shohamy, 2009). Frameworks tend to build on developmental stages in early language programs and reflect how young learners move from chunks to analyzed language use (Johnstone, 2009). Most curricula include not only L2 achievement targets, but comprise further aims. Early learning is meant to contribute to young learners’ positive attitudes towards languages, language learning, speakers of other languages, and towards learners’ own culture and identity (e.g., Prabhu, 2009). In addition to linguistic and affective aims, they often include aims related to cognition, metacognition and learning strategies. There is a controversy in the multiplicity of aims. Testing in most contexts focuses on L2 achievements and the other aims are not assessed at all or they are discussed only in a few research projects. Testing in early language learning programs is most often concerned with: (1) how learners progress in their L2 over time and (2) what levels of proficiency they achieve in some or all of the four skills

by the end of certain periods. In addition to these areas, there is a need to explore how teachers assess YLs and how classroom practices interact with children's attitudes, motivation, willingness to communicate, anxiety, self-confidence and self-perception over time.

Early language learning assessment frameworks define the main principles of teaching and assessing young learners and aim to describe and quantify what children are expected to be able to do at certain stages of their L2 development (e.g., Curtain, 2009; Jang, 2014; McKay, 2006; Nikolov, 2016 in this volume). Frameworks developed in Europe tend to use the *Common European Framework of Reference for Languages (CEFR)* (Council of Europe, 2001) as a point of departure, despite the fact that it was not designed for young learners (e.g., Hasselgren, 2005; Pižorn, 2009; Papp & Salamoura, 2009; Papp & Walczak, 2016 in this volume). In contrast, research projects on early CLIL tend to follow a different tradition unrelated to testing children or standards-based testing. They frame CLIL as an add-on to FL instruction and analyze young learners' performances along three criteria (complexity, accuracy, and fluency) used in second language acquisition research (e.g., Hausen & Kuiken, 2009). Such a framework, however, is hardly suited to document very slow development (see e.g., Bret-Blasco, 2014).

Tests for young learners have been developed for various purposes. Standards-based tests are used in national and international projects and external examinations as well as in smaller-scale research studies. The majority of national and international projects tend to apply standards aligned to levels in *CEFR*. Test construction and validation is a long and complex process. Some important work has been published on the process of developing frameworks, *can do statements*, designing and validating tests for various purposes, for example, for large-scale proficiency tests, research projects and teacher-based assessments. These areas are discussed in five chapters.

3.2 National, International and Local Testing Projects

Early language learning is compulsory in many places. In Europe, it is more the norm than the exception. National curricula typically include achievement targets and in some countries national proficiency exams are implemented annually (e.g., in Germany, Wilden & Porsch, 2016 in this volume, in Poland, Szpotowicz & Campfield, 2016 in this volume; in Slovenia, Pižorn, 2009; in Switzerland, Haenni Hoti, Heinzmann & Müller, 2009; in Hungary, Nikolov & Szabó, in press). How these tests are administered, how the test results are used and how tests impact teaching and learning raises further questions. They have to be discussed in each particular situation bearing in mind the particulars of the assessment culture.

International research projects have also been implemented to collect test data for comparative purposes and to answer questions related to the rate and level of L2 development. For example, a longitudinal study, the Early Language Learning in Europe (ELLiE) project aimed to examine what level young learners achieved in a foreign language at public schools in England, Italy, the Netherlands, Poland, Spain,

Sweden and Croatia. In addition to L2, other factors were also included to find out how they contributed to processes and outcomes in the target languages as well as in the affective domain (Enever, 2011; Mihaljević Djigunović, 2012). Researchers faced challenges similar to those in previous longitudinal studies on early language learning (Enever, 2011; García Mayo & García Lecumberri, 2003; Muñoz, 2006). The same tests were used over the years to collect valid and reliable results on participants' L2 development and a single task was used for each skill.

Assessment projects are often narrowly limited and they aim to seek answers to research questions emerging from practice. For example, how achievement tests are applied by teachers (Peng & Zheng, 2016), and how innovative assessment techniques can change classroom processes (Butler, 2016; Hung, Samuelson & Chen, 2016, both in this volume). Other projects use tests in order to build new models or to test existing ones to find out to what extent they can reflect realities in early FL classrooms (Mihaljević Djigunović, 2016; Bacsa & Csikos, 2016; see chapters in this volume).

3.3 International Language Tests for Young Language Learners

In recent years, several international examinations have been developed and made available to young language learners whose parents want them and can afford them. Three widely known exams offer certificates on children's proficiency in English: (1) Cambridge Young Learners English Tests (www.cambridgeesol.org/exams/young-learners), (2) Pearson Test of English Young Learners (www.pearsonpte.com/PTEYoungLearners); and (3) TOEFL Primary (https://www.ets.org/toefl_primary). These examinations fall somewhere in the middle of the language–content continuum with a focus on some typically taught topics young language learners can be realistically expected to know. The levels cover A1 and A2 in the *CEFR* (Council of Europe, 2001). Besides aural/oral skills literacy skills are also included. How much work is devoted to developing and validating exams is discussed in three of the chapters (Benigno & de Jong, 2016; Hsieh, 2016; Papp & Walczak, 2016). Unfortunately, hardly any studies explore how these proficiency exams impact classroom processes or how children taking them benefit from their experiences in the long run. It would also be important to know how they maintain and further develop their proficiency after taking examinations.

3.4 Assessment for Learning

Recent research on early language learning assessment has focused on how teacher-based assessment can scaffold children's development in their L2 knowledge and skills so that they can apply their learning potential (Sternberg & Grigorenko, 2002).

In this developmental framework of *assessment for learning* children should benefit from ongoing classroom testing. Teachers consider assessment as an integral part of their teaching. They build on test results to inform their teaching (Black & Wiliam, 1998; Davison & Leung, 2009; McKay, 2006). This way the teaching process can be sensitive to readiness to develop (McNamara & Roever, 2006). These are key points in teacher-based assessment: learning oriented assessment is based on these principles (Nikolov, 2011, 2016 in this volume). Very little has been published on how assessment for learning works in early foreign language contexts and how teachers apply their diagnostic competence. The “ability to interpret students’ foreign language growth, to skillfully deal with assessment material and to provide students with appropriate help in response to this diagnosis” (Edelenbos & Kubanek-German, 2004, p. 48) is definitely an area where further classroom studies are necessary.

These approaches to assessment and uses of test results definitely require teachers to reflect on their practices in a new way. The visual and written samples in Rixon’s (2016) chapter clearly document a totally different assessment culture from what one would find in classrooms where the tradition is more focused on *assessment of learning*. Three other chapters in this book discuss further aspects of learning oriented assessment. Nikolov’s (2016) account shares outcomes of a diagnostic testing project: framework, main principles, *can do statements*, topics and task types designed for young learners in the first six grades of primary school. Butler’s (2016) overview offers multiple insights into how self-assessment can be used in various domains, whereas Hung, Samuelson and Chen report on how peer-, self-, and teacher-based assessments were implemented in the EFL classroom where traditions were not in line with assessment for learning principles.

3.5 *What Tests Are Used and How*

Researching and documenting how certain tests work with young learners is time-consuming and this is an area where there is a need and a lot of room for further work. Similarly to the most brilliant age-appropriate teaching materials and tasks, the most valid and reliable tests can also be misused or abused. The chapters in this volume offer insights into some actual tests and how researchers and teachers applied them. One interesting trend needs pointing out: most of the tests discussed in the early language learning assessment literature and these chapters are similar to language tests widely used and accepted in the L2 testing literature. However, some tests and criteria for assessment are borrowed from other traditions: for example, oral production was assessed along complexity, accuracy, and fluency in Bret Blasco’s (2014) study on CLIL.

As these are key issues in assessment, a detailed and critical analysis should focus on what tests are used in assessment projects involving young learners. Often a single task is used to tap into a skill and the same test is used over the years to document development (e.g., Bret Blasco, 2014; Enever, 2011). Recently elicited

repetition has been also used to assess speaking. It is important to approach these questions from the learners' and teachers' perspectives as well and to explore how tests can be linked to offer more reliable insights into young learners' development (e.g., Nikolov & Szabó, 2012; Szpotowicz & Campfield, 2016 in this volume). There is a lot of potential in learning about the traditions in the fields of second language acquisition and language testing, and most probably both areas would benefit from a comparative analysis.

4 How This Volume Contributes to a Better Understanding of the Challenges in Young Learners' Assessment and to Advancing the Field

Assessing young learners of a FL is a complex area requiring knowledge of age-appropriate classroom methodology, including teacher- and standards-based language assessment, second language acquisition, research methodology and the actual contexts. The issues and challenges should be approached, researched and interpreted as subcomponents of innovation requiring more than change in a single aspect. The complexity of teaching and assessment results from the fact that not only the constructs vary but also because young learners' individual differences, languages, and knowledge interact with specific contextual and teacher- and parent-related variables. In what follows, let us overview what this volume comprises.

The chapters focus on various aspects of assessment in early EFL programs around the world. The first two papers draw the larger picture; Marianne Nikolov and Shelagh Rixon outline the main trends, issues and challenges and the reasons why recent international developments represent the 'coming of age'. They provide an overview on how the main points are embedded in larger trends, and discuss the construct, various frameworks for test development, international and national projects and international examinations designed to tap into children's proficiency. These two chapters offer insights also into teacher-based alternative approaches: diagnostic and self-assessment.

Chapters "The "Global Scale of English Learning Objectives for Young Learners": A CEFR-Based Inventory of Descriptors, A Framework for Young EFL Learners' Diagnostic Assessment: 'Can Do Statements' and Task Types, Examining Content Representativeness of a Young Learner Language Assessment: EFL Teachers' Perspectives, Developing and Piloting Proficiency Tests for Polish Young Learners, and The Development and Validation of a Computer-Based Test of English for Young Learners: Cambridge English Young Learners" focus on how challenges are overcome in test development. Three papers present findings on the early stages and the fourth one on how a validated paper and pencil test can go online. In chapter "The "Global Scale of English Learning Objectives for Young Learners": A CEFR-Based Inventory of Descriptors", Veronica Benigno and John de Jong give an account of how Pearson developed their first batch of CEFR-based

inventory of young learners descriptors. Chapter “[A Framework for Young EFL Learners’ Diagnostic Assessment: ‘Can Do Statements’ and Task Types](#)”, by Marianne Nikolov, discusses how a framework was developed for young EFL learners for diagnostic assessment purposes and presents *can do statements* and task types found relevant in a national project in Hungary. In chapter “[Examining Content Representativeness of a Young Learner Language Assessment: EFL Teachers’ Perspectives](#)”, Ching-Ni Hsieh offers test validity evidence for TOEFL Primary: she discusses how content representativeness was ensured at ETS by integrating teachers’ views in the process. In chapter “[Developing and Piloting Proficiency Tests for Polish Young Learners](#)”, Magdalena Szpotowicz and Dorota E. Campfield reveal how they piloted proficiency tests and used children’s feedback in a national testing project in Poland. The very first examination for young learners of English was offered by Cambridge. In chapter “[The Development and Validation of a Computer-Based Test of English for Young Learners: Cambridge English Young Learners](#)”, Szilvia Papp and Agnieszka Walczak offer insights into how a computer-based test was developed and validated to make the tests more readily available.

Chapters “[Learning EFL from Year 1 or Year 3? A Comparative Study on Children’s EFL Listening and Reading Comprehension at the End of Primary Education, A Longitudinal Study of a School’s Assessment Project in Chongqing, China, Individual Learner Differences and Young Learners’ Performance on L2 Speaking Tests](#)”, and [The Role of Individual Differences in the Development of Listening Comprehension in the Early Stages of Language Learning](#)” present five complex research projects where testing young learners’ L2 played a key part. In chapter “[Learning EFL from Year 1 or Year 3? A Comparative Study on Children’s EFL Listening and Reading Comprehension at the End of Primary Education](#)”, Eva Wilden and Raphaela Porsch intended to find out if learning EFL from the first or the third year in German primary schools was a better model by examining young learners’ EFL listening and reading comprehension at the end of their primary education. Besides the modest advantage for earlier starters, their study revealed that children’s proficiency in other languages interacted with the outcomes in important and unexpected ways. In chapter “[A Longitudinal Study of a School’s Assessment Project in Chongqing, China](#)”, Jing Peng and Shicheng Zheng compare and contrast outcomes of a longitudinal teacher-based assessment study implemented at a school in China. They discuss how children performed on two achievement tests based on two course books and triangulate their findings by interviewing teachers. In chapter “[Individual Learner Differences and Young Learners’ Performance on L2 Speaking Tests](#)”, Jelena Mihaljević Djigunović discusses the dynamic changes in the ways how young Croatian language learners’ individual differences, motivation and self-concept, contributed to their performance on EFL speaking tests over a four-year period. The aim of chapter “[The Role of Individual Differences in the Development of Listening Comprehension in the Early Stages of Language Learning](#)”, by Éva Bacsa and Csaba Csíkos, was to model how aptitude, motivation anxiety, learners’ beliefs and their parental background interacted in the development of EFL in a semester-long study involving young learners in a small town in Hungary.

The last two chapters provide insights into how peer-, self-assessment and teacher assessment interact with one another. Yuko Goto Butler, in chapter “*Self-Assessment of and for Young Learners’ Foreign Language Learning*”, offers a critical overview of research into self-assessment *of* and *for* young learners’ foreign language learning and proposes five dimension for developing further research instruments, thus linking teaching, assessment and learning. The context of the final chapter is Taiwan. Yu-ju Hung, Beth Lewis Samuelson and Shu-cheng Chen explore the relationships between peer- and self-assessment and teacher assessment of young EFL learners’ oral presentations by applying both the teacher’s and her students’ reflections for triangulation purposes.

5 Areas for Further Research and Implications for Practice

This volume outlines some of the key areas where research has been conducted. Similar inquiries would allow us to find out how results would compare in other contexts. Researchers, including classroom teachers, should consider how replication studies could offer useful information on learners’ achievements in their countries and classrooms. Data collection instruments can be of invaluable help with instructions on how to apply them. Such data repositories, for example at <http://iris-database.org/iris/app/home/index>, are available. Test development is an extremely challenging and expensive process. Questionnaires, interviews, etc. also require special expertise to develop and validate. Sharing them would allow the early language learning field to advance more rapidly.

It is also important to note which key areas are not discussed in this book in full detail or at all, and where more research is needed.

- (1) In order to answer research questions related to the larger picture on early start programs, studies should aim to find out in what domains younger learners excel over time and why this is the case. This kind of research should work towards testing models of early language learning. Studies should include proficiency tests on learners’ aural/oral and literacy skills in their L1, L2, L3. Other instruments should tap into individual differences of young learners and their teachers, and contextual variables (including characteristics of programs, materials, methods, the quality of teaching) interacting in children’s development over several years. The main benefits of an early start are most probably not in higher L2 proficiency over time and this hypothesis may have important implications for language policy, curriculum design, teacher education and classroom practice.
- (2) Hardly any studies look into the relationships between access to early foreign language learning opportunities, assessment, and equity. Do all children have equal opportunities? Research is necessary to examine how parents’ motivation, learners’ socio economic status and achievements on tests interact and how test results are used.
- (3) A recurring theme in early language teaching programs concerns transition and continuity. Studies should go beyond the early years and focus on how teachers build on what learners can do in later years and what role assessment practices play in the process. In other words, research is necessary into how children are taught and assessed, and how teachers can apply diagnostic information in their teaching.

- (4) The impact of different kinds of assessment on young language learners, their teachers, and the teaching-learning process should be explored in depth. Teachers' and learners' emic perspectives are hardly ever integrated into studies. Exploring teachers' and their learners' beliefs and lived experiences could reveal why implementing innovation often poses a major challenge. Case studies could offer insights on what it means to a child to take an external examination, what challenges learners and their teachers face due to parental pressure to produce results, and why teachers may resist change in their teaching and testing practices.
- (5) It would be essential to learn more about the ways in which achievement targets defined in curricula are assessed by teachers on a daily basis. How they balance giving children feedback on their progress in test results with maintaining their motivation and keeping their debilitating anxiety low.
- (6) Yet another avenue for classroom research for practicing teachers should explore how teachers apply traditional (assessment *of* learning) and innovative assessment techniques (assessment *for* learning, peer and self-assessment). How do they use criteria for assessing speaking and writing and keys on closed items and students' responses to open items? How do they integrate other aspects of students' behavior into their assessments, for example, their willingness to communicate, attitudes, motivation, aptitude, anxiety?
- (7) Very little is known about testing learners' knowledge and skills in CLIL programs. Exploratory classroom studies are needed to find out how teachers tease out the two domains and how they can diagnose if learners' weaknesses are in their L2 or in the subject matter.

The studies in this volume discuss various aspects of test development, outcomes of large-scale surveys, national assessment projects, and innovative smaller-scale studies. The ideas shared and the frameworks and instruments used for data collection should be of interest to both novice and experienced teachers, materials and test developers, as well as for researchers. Readers should bear in mind which of the main points are worth further explorations. It is hoped that the volume offers exciting new ideas, and result in innovation and change.

References

- Alderson, J. C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. London: Continuum.
- Bacsa, É., & Csíkos, C. (2016). The role of individual differences in the development of listening comprehension in the early stages of language learning. In M. Nikolov (Ed.), *Assessing young learners of English: Global and local perspectives*. New York: Springer.
- Benigno, V., & de Jong, J. (2016). A CEFR-based inventory of YL descriptors: Principles and challenges. In M. Nikolov (Ed.), *Assessing young learners of English: Global and local perspectives*. New York: Springer.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7–71.
- Bret Blasco, A. (2014). *L2 English learners' oral production skills in CLIL and EFL settings: A longitudinal study*. Doctoral dissertation, Universitat Autònoma de Barcelona.
- Butler, Y. G. (2009). Issues in the assessment and evaluation of English language education at the elementary school level: Implications for policies in South Korea, Taiwan, and Japan. *The Journal of Asia TEFL*, 6, 1–31.

- Butler, Y. G. (2016). Self-assessment of and for young learners' foreign language learning. In M. Nikolov (Ed.), *Assessing young learners of English: Global and local perspectives*. New York: Springer.
- Cameron, L. (2001). *Teaching languages to young learners*. Cambridge: Cambridge University Press.
- Chik, A., & Besser, S. (2011). International language test taking among young learners: A Hong Kong case study. *Language Assessment Quarterly*, 8, 73–91.
- Clapham, C. (1997). Introduction. In C. Clapham & D. Corson (Eds.), *Encyclopedia of language and education. Volume 7. Language testing and assessment* (pp. xiii–xix). Dordrecht: Kluwer Academic Publisher.
- Copland, F., & Garton, S. (Eds.). (2014). *English Language Teaching Journal. Special Issue*.
- Curtain, H. (2009). Assessment of early learning of foreign languages in the USA. In M. Nikolov (Ed.), *The age factor and early language learning* (pp. 59–82). Berlin, Germany: Mouton de Gruyter.
- Curtain, H. A., & Dahlberg, C. A. (2010). *Languages and children – Making the match: New languages for young learners* (4th ed.). Needham Heights, MA: Pearson Allyn & Bacon.
- Dalton-Puffer, C. (2011). Content and language integrated learning: From practice to principles? *Annual Review of Applied Linguistics*, 31, 182–204.
- Davison, C. (2013). Innovation in assessment: Common misconceptions. In K. Hyland & L. L. C. Wong (Eds.), *Innovation and change in English language education* (pp. 263–277). New York: Routledge.
- Davison, C., & Leung, C. (2009). Current issues in English language teacher-based assessment. *TESOL Quarterly*, 43, 393–415.
- deBot, K. (2014). The effectiveness of early foreign language learning in the Netherlands. *Studies in Second Language Learning and Teaching*. doi:10.14746/ssllt.2014.4.3.2
- Edelenbos, P., Johnstone, R., & Kubanek, A. (2007). *Languages for the children in Europe: Published research, good practice and main principles*. Retrieved from http://ec.europa.eu/education/policies/lang/doc/youngsum_en.pdf
- Edelenbos, P., & Kubanek-German, A. (2004). Teacher assessment: The concept of 'diagnostic competence'. *Language Testing*, 21, 259–283.
- Emery, H. (2012). *A global study of primary English teachers' qualifications, training and career development*. London: British Council.
- Enever, J. (Ed.). (2011). *ElliE: Early language learning in Europe*. London: British Council.
- Enever, J., Moon, J., & Raman, U. (Eds.). (2009). *Young learner English language policy and implementation: International perspectives*. Reading, UK: Garnet Education Publishing.
- García Mayo, M. P., & García Lecumberri, M. L. (Eds.). (2003). *Age and the acquisition of English as a foreign language*. Clevedon: Avon/Multilingual Matters.
- Garton, S., Copland, F., & Burns, A. (2011). *Investigating global practices in teaching English to young learners*. London: British Council.
- Haenni Hoti, A., Heinzmann, S., & Müller, M. (2009). "I can you help?" Assessing speaking skills and interaction strategies of young learners. In M. Nikolov (Ed.), *The age factor and early language learning* (pp. 119–140). Berlin, Germany: Mouton de Gruyter.
- Hasselgren, A. (2005). Assessing the language of young learners. *Language Testing*, 22, 337–354.
- Hausen, A., & Kuiken, F. (2009). Complexity, accuracy and fluency in second language acquisition. *Applied Linguistics*, 30(4), 461–473.
- Hsieh, C. (2016). Examining content representativeness of a young learner language assessment: EFL teachers' perspectives. In M. Nikolov (Ed.), *Assessing young learners of English: Global and local perspectives*. New York: Springer.
- Hung, Y.-J., Samuelson, B. L., & Chen, S.-C. (2016). The relationships between peer- and self-assessment and teacher assessment of young EFL learners' oral presentations. In M. Nikolov (Ed.), *Assessing young learners of English: Global and local perspectives*. New York: Springer.

- Inbar-Lourie, O., & Shohamy, E. (2009). Assessing young language learners: What is the construct? In M. Nikolov (Ed.), *The age factor and early language learning* (pp. 83–96). Berlin, Germany: Mouton de Gruyter.
- International schools: The new local. (2014, December 20). *The Economist*, pp. 83–84.
- Jang, E. E. (2014). *Focus on assessment*. Oxford: Oxford University Press.
- Johnstone, R. (2009). An early start: What are the key conditions for generalized success? In J. Enever, J. Moon, & U. Raman (Eds.), *Young learner English language policy and implementation: International perspectives* (pp. 31–42). Reading: Garnet Education Publishing Ltd.
- Kahneman, D. (2011). *Thinking, fast and slow*. New York: Allen Lane/Penguin Books.
- Kennedy, C. (2013). Models of change and innovation. In K. Hyland & L. L. C. Wong (Eds.), *Innovation and change in English language education* (pp. 13–27). New York: Routledge.
- Markee, N. (2013). Contexts of change. In K. Hyland & L. L. C. Wong (Eds.), *Innovation and change in English language education* (pp. 28–43). New York: Routledge.
- McKay, P. (2006). *Assessing young language learners*. Cambridge: Cambridge University Press.
- McNamara, T., & Roever, C. (2006). *Language testing: The social dimension*. Oxford: Blackwell Publishing.
- Mihaljević Djigunović, J. (2012). *Early EFL learning in context – Evidence from a country case study*. London: The British Council.
- Mihaljević Djigunović, J. (2016). Individual differences and young learners' performance on L2 speaking tests. In M. Nikolov (Ed.), *Assessing young learners of English: Global and local perspectives*. New York: Springer.
- Muñoz, C. (Ed.). (2006). *Age and the rate of foreign language learning*. Clevedon: Avon/Multilingual Matters.
- Murphy, V. A. (2014). *Second language learning in the early school years: Trends and contexts*. Oxford: Oxford University Press.
- Nikolov, M. (2000). Issues in research into early second language acquisition. In J. Moon & M. Nikolov (Eds.), *Research into teaching English to young learners: International perspectives* (pp. 21–48). Pécs: University Press Pécs.
- Nikolov, M. (Ed.). (2009a). *The age factor and early language learning*. Berlin/New York: Mouton de Gruyter.
- Nikolov, M. (Ed.). (2009b). *Early learning of modern foreign languages: Processes and outcomes*. Clevedon, UK: Multilingual Matters.
- Nikolov, M. (2009c). The age factor in context. In M. Nikolov (Ed.), *The age factor and early language learning* (pp. 1–38). Berlin, Germany/New York, NY: Mouton de Gruyter.
- Nikolov, M. (Ed.). (2010). *International Journal of Bilingualism. Special Issue*.
- Nikolov, M. (2011). Az angol nyelvtudás fejlesztésének és értékelésének keretei az általános iskola első hat évfolyamán [A framework for developing and assessing proficiency in English as a foreign language in the first six years of primary school]. *Modern Nyelvoktatás*, XVII(1), 9–31.
- Nikolov, M. (2013, August). *Early language learning: Is it child's play? Plenary talk*. EUROSLA Conference, Amsterdam. Retrieved from <http://webcolleges.uva.nl/Mediasite/Play/7883cb9b1cb34f98a21fb37534fc1ec61d>
- Nikolov, M. (2016). A framework for young EFL learners' diagnostic assessment: Can do statements and task types. In M. Nikolov (Ed.), *Assessing young learners of English: Global and local perspectives*. New York: Springer.
- Nikolov, M., & Csapó, B. (2010). The relationship between reading skills in early English as a foreign language and Hungarian as a first language. *International Journal of Bilingualism*, 14, 315–329.
- Nikolov, M., & Mihaljević Djigunović, J. (2006). Recent research on age, second language acquisition, and early foreign language learning. *Annual Review of Applied Linguistics*, 26, 234–260.
- Nikolov, M., & Mihaljević Djigunović, J. (2011). All shades of every color: An overview of early teaching and learning of foreign languages. *Annual Review of Applied Linguistics*, 31, 95–119.

- Nikolov, M., & Szabó, G. (2012). Developing diagnostic tests for young learners of EFL in grades 1 to 6. In E. D. Galaczi & C. J. Weir (Eds.), *Voices in language assessment: Exploring the impact of language frameworks on learning, teaching and assessment – Policies, procedures and challenges, Proceedings of the ALTE Krakow Conference, July 2011* (pp. 347–363). Cambridge: UCLES/Cambridge University Press.
- Nikolov, M., & Szabó, G. (in press). A study on Hungarian 6th and 8th graders' proficiency in English and German at dual-language schools. In D. Holló & K. Károly (Eds.), *Inspirations in foreign language teaching: Studies in applied linguistics, language pedagogy and language teaching in honour of Peter Medgyes*. Harlow: Pearson Education.
- Papp, S., & Salamoura, A. (2009). *An exploratory study into linking young learners' examinations to the CEFR* (Research Notes, 37, pp. 15–22). Cambridge: Cambridge ESOL.
- Papp, S., & Walczak, A. (2016). The development and validation of a computer-based test of English for young learners: Cambridge English Young Learners. In M. Nikolov (Ed.), *Assessing young learners of English: Global and local perspectives*. New York: Springer.
- Peng, J., & Zheng, S. (2016). A longitudinal study of a school's assessment project in Chongqing, China. In M. Nikolov (Ed.), *Assessing young learners of English: Global and local perspectives*. New York: Springer.
- Pinter, A. (2006). *Teaching young language learners*. Oxford: Oxford University Press.
- Pinter, A. (2011). *Children learning second languages*. Basingstoke: Palgrave Macmillan.
- Pižorn, K. (2009). Designing proficiency levels for English for primary and secondary school students and the impact of the CEFR. In N. Figueras & J. Noijons (Eds.), *Linking to the CEFR levels: Research perspectives* (pp. 87–102). Arnhem, The Netherlands: Cito/EALTA.
- Prabhu, N. S. (2009). Teaching English to young learners: The promise and the threat. In J. Enever, J. Moon, & U. Raman (Eds.), *Young learner English language policy and implementation: international perspectives* (pp. 43–44). Reading, UK: Garnet Education Publishing.
- Rea-Dickins, P. (2000). Assessment in early years language learning contexts. *Language Testing*, 17(2), 115–122.
- Rhodes, N. C., & Pufahl, I. (2008). *Foreign language teaching in U.S. Schools: Results of a national survey*. Retrieved from http://www.cal.org/projects/Exec%20Summary_111009.pdf
- Rixon, S. (2013). *British Council survey of policy and practice in primary English language teaching worldwide*. London: British Council.
- Rixon, S. (2016). Do developments in assessment represent the 'coming of age' of young learners English language teaching initiatives? The international picture. In M. Nikolov (Ed.), *Assessing young learners of English: Global and local perspectives*. New York: Springer.
- Singleton, D. (2014). Apt to change: The problematic of language awareness and language aptitude in age-related research. *Studies in Second Language Learning and Teaching*. doi:10.14746/ssl.2014.4.3.9.
- Sternberg, R. J., & Grigorenko, E. L. (2002). *Dynamic testing: The nature and measurement of learning potential*. Cambridge: Cambridge University Press.
- Szpotowicz, M., & Campfield, D. E. (2016). Developing and piloting proficiency tests for Polish young learners. In M. Nikolov (Ed.), *Assessing young learners of English: Global and local perspectives*. New York: Springer.
- Wilden, E., & Porsch, R. (2016). Learning EFL from year 1 or year 3? A comparative study on children's EFL listening and reading comprehension at the end of primary education. In M. Nikolov (Ed.), *Assessing young learners of English: Global and local perspectives*. New York: Springer.