CHAPTER 1

Assessment of Multi-tasking Wonders: Music, Olympics, and Language

Virginia C. Mueller Gathercole
ABSTRACT

These volumes address an expanding area of interest and concern in the 21st century – the assessment of bilingual speakers, both adults and children. There is a rapidly growing body of research and proposals concerning the issues surrounding the evaluation of language abilities and proficiency in multilingual speakers, and, by extension, the evaluation of any cognitive or academic abilities in such speakers. Bilingual speakers’ acquisition and knowledge of their two languages are necessarily different from acquisition and knowledge of a single language. This has ramifications for how bilingual speakers perform in a variety of tasks. If we wish to gain accurate evaluations of bilingual children’s and adults’ proficiency and abilities, we must necessarily take into account facts concerning the processes of learning, speaking, and understanding two languages. The authors in these volumes explore issues and solutions for the assessment of bilinguals. The research here comes from a variety of particular bilingual populations from around the world. The concerns expressed and the proposed solutions are relevant and applicable to bilingual populations everywhere.
You don’t get harmony when everybody sings the same note.

DOUG FLOYD (Guthrie, 2003: 41)

Imagine a world in which we saw beyond the lines that divide us, and celebrated our differences, instead of hiding from them.

WESLEY CLARK, speech, Jan. 20, 2004

If a man does not keep pace with his companions, perhaps it is because he hears a different drummer. Let him step to the music which he hears, however measured or far away.

HENRY DAVID THOREAU
(from http://shell.cas.usf.edu/~mccolm/Dquotes.html)

INTRODUCTION

Some accomplishments in life are so remarkable that we glory in them and celebrate them. For example, we are in awe of people who show excellence in more than one aspect of a given talent at the same time. Some obvious examples come from the fields of music and athletics. We have great admiration for musicians like Stevie Wonder, who not only wrote, produced, arranged, and sang “Superstition”, but also played the drum, the clavinet, and the Moog bass synthesizer for it. We applaud the accomplishments of composers like Georg Telleman, who played multiple instruments-- violín, viola da gamba, recorder, flauto traverso, oboe, shawm, sackbut and double bass—and Paul McCartney, who plays the guitar, bass guitar, piano, harmonica, recorder, banjo, mandolin, and drums.

We are equally awed by athletes that excel in not only one sport but two or three. Jackie Robinson, the first African American to play in major league baseball (for the Dodgers), had an illustrious career in baseball (e.g., winning the MVP award in 1949), but he also excelled in football, track, and basketball while he was enrolled at UCLA (see http://www.toptenz.net/top-10-multi-sport-athletes.php#ixzz1rjaZWalf). Jim Thorpe won gold medals in the pentathlon and the decathlon in the 1912 Olympics, and then he went on to play baseball for the New York Giants, the Brewers, and the White Sox. He also played
professional football and professional basketball (http://www.toptenz.net/top-10-multi-sport-athletes.php#ixzz1rjaZWaIf). We rightly applaud such multi-accomplishing individuals.

It never crosses our minds that playing the drums in addition to the guitar might detract from the musician’s accomplishments on the guitar, or that the athlete who excels in two sports might be inferior to someone who excels in only one of those sports.

We sometimes celebrate similar accomplishments in relation to language. We find it a surprise – but a delightful surprise -- that Jodie Foster speaks fluent French in addition to English (http://www.youtube.com/watch?v=c3TvLSvvKMc&feature=player_embedded); that Salma Hayek speaks Spanish, Portuguese, and Arabic; that Charlize Theron speaks Afrikaans natively (http://www.youtube.com/watch?v=2fYB9s0NyZk&feature=player_embedded); that Natalie Portman speaks Hebrew (http://www.youtube.com/watch?v=n-PDArBZrz8&feature=player_embedded); that Sandra Bullock speaks fluent German (http://www.youtube.com/watch?v=s10x38SMb-g&feature=player_embedded); and that Gwyneth Paltrow speaks Spanish fluently.

The majority of the world’s population are just like these celebrities—they speak more than one language, and they often do so fluently. But somehow the delight and awe we experience in relation to celebrities who we discover are bilingual sometimes gets diminished or turned to caution in relation to others who are bilingual, especially in relation to children who are growing up bilingually. Why might this be? It probably boils down to two things – the important role that language plays in all aspects of our lives, combined with fear associated with a lack of knowledge about how bilingual language develops.

The way in which we view the multi-tasking accomplishments and evaluate the abilities of such multi-tasking persons hinges in large part on our appreciation of the steps one takes towards those accomplishments. While we readily acknowledge those steps in the cases of musical and athletic advances, the steps in relation to language are perhaps more covert and less well understood. As we gain a fuller understanding of what it means to be a
child growing up as a bilingual or an adult who has become bilingual, our understanding of how assessments of such individuals need to take those facts into consideration is also growing.

**STEP BY STEP**

We tend to forget that people who end up being fluent bilinguals do not, of course, start out as fluent bilinguals--just as competent multi-instrumentalists or athletes who excel in multiple sports do not start life at the top of their art or their game. Everyone has to start from scratch. This means that budding musicians, athletes, and language learners all must go through multiple stages on the path towards coming to full mastery of their art. Those multiple stages involve multiple steps and periods when performance appears less than optimal. The emergent musician, athlete, and language learner alike must, as a natural course of events, pass through moments at which errors are made and during which the prognosis for the ultimate success of the endeavor may seem uncertain.

To help the budding musician, athlete, or language learner, and to gauge the level of his or her progress, we have ways of assessing whether that progress is commensurate with expectations, or whether, at points along the way, an emergent musician, athlete, or language learner may need a little extra assistance along the way. There are exams and competitions children undergo in each of these realms, exams and competitions that entail expectations at a level that is determined by our knowledge of how similar children or learners at similar stages of development have been able to perform. Knowledge of that normal level of progress at each stage is determined by years, even centuries, of experience of observing thousands upon thousands of children passing through similar stages. For musicians, norms for such expectations are used in tests for performance at local, national, or international levels; for athletes, there are meets and competitions such as the Olympics; for language learners, there are tests of language abilities, especially in relation to reading and writing (and of related
skills such as spelling), but also for oral language understanding, vocabulary knowledge, and grammatical knowledge.

Tests related to musical and athletic prowess are usually not mandatory for every child passing through a certain age group or school level. Acquisition of these skills is seen as optional, and we tend to have the attitude that excellence in them involves specialist endeavours. So an inability to perform in either realm is not usually considered detrimental to a child’s overall development, nor to have implications for any prognosis concerning their overall success in the future or as human beings.

With tests for language and language-related abilities (e.g., literacy), in contrast, the situation is quite different. All children are expected to achieve certain levels with language, and we take steps to assess their proficiency at multiple points in development. This universal testing of language occurs because of the fundamental nature of language as a key foundation on which success in a variety of areas that go well beyond language itself is built. Linguistic abilities are essential to academic success in all content areas, including not only those directly related to language (reading, writing), but also those that initially might appear to be independent of language, such as mathematical abilities. Unlike for musical talent or athletic abilities, therefore, there are high stakes associated with language abilities, as other successes appear contingent on a firm language base, which in turn is often taken as predictive of future potential in a variety of areas.

One consequence of this is that any evidence of possible difficulties with language are taken quite seriously by both parents and professionals, and any problems seem to be worthy of fairly prompt attention. For this reason, we assess children’s language abilities very early on, and we continue to do so throughout a person’s educational career. If a very young child does not appear to be talking when his or her peers are, a parent might take that child to a professional for consultation, to determine whether there are any major difficulties the child is having with language. When children enter school, they undergo tests related to reading
readiness. Then throughout school, language and reading and writing assessments are key components of the assessment of a child’s progress. Furthermore, when a child or adult begins to learn a language other than the first language, teachers administer tests to determine that person’s progress in the second language.

**EXPECTATIONS**

As noted, all of those assessments—whether for music, athletics, or language-- involve underlying assumptions or evidence on realistic expectations for how a child or student should perform – *given his or her age, level, and experience*. For example, a 4-year-old Suzuki violin player might be credited with an excellent performance for being able to play “Twinkle Twinkle” on his/her violin. No one would expect that 4-year-old to be able to play Brahms’ violin concerto, nor consider that child deficient for not being able to do so. Similarly, a 4-year-old in a tumbling class might be rewarded with gold stars for making several somersaults in a row. Again, no one would expect that child to be able to perform a back flip like Aston Merrygold of JLS can.

In the area of language, a Kindergartner who is able to read at least a few words might be lavishly praised. No one would expect that Kindergartner to be able to read a text from Shakespeare or would test him or her on the understanding of a passage from Hamlet. (At the same time, if the same assessor or teacher is presented with a high school student who cannot read more than a few words, s/he might well be concerned and would consider whether such a student was in need of extra support.)

**CHILDREN GROWING UP AS BILINGUALS**

The realistic expectations we use to evaluate performance in any of these realms come from experiences with similar children or students at similar levels along the way toward gaining mastery of the skills. These books are about what those realistic expectations might be for children growing up as bilinguals, and how we apply those realistic expectations in assessing performance. What might we expect at various stages in a bilingual child’s or
adult's progress in language? How can one tell whether a bilingual child is developing as might be expected, *given his or her age, level, and experience* with the two languages, and how can we tell whether those expectations are or are not being met?

The educational and professional communities whose job it is to assess children’s development have a good sense, and a long history of understanding, of how development occurs in monolingual children. The vast majority of the standardized tests for language have been developed with monolingual children or adults in mind (unless they are specifically designed to see if second-language learners have developed a command of their second language – e.g., the TOEFL [http://www.ets.org/toefl/]). Typical examples are receptive vocabulary tests like the Peabody Picture Vocabulary Test (Dunn & Dunn, 2007) and the British Picture Vocabulary Scales (Dunn, Dunn, Whetton, & Burley, 1997). These provide information on the normal expectations one might have for monolingual children learning English in America or in the UK. (Note, however, that the BPVS now provides bilingual norms in addition to the monolingual norms.) Scales such as the Child Development Inventories (Fenson, Marchman, Thal, Dale, Reznick, & Bates, 2007) have largely culled data from a large number of (usually monolingual) children learning the given language to allow the assessment of other children learning the same language. And standardized college placement tests such as the SAT (http://sat.collegeboard.org/home?affiliateId=nav&bannerId=h-satb) and the ACT (http://www.act.org/aap/) rely heavily on one’s knowledge of the vocabulary and grammar of the language being tested, English. E.g., the passage-based reading and sentence completion components on the SAT (see http://sat.collegeboard.org/practice/sat-practice-questions) require a knowledge of highly sophisticated vocabulary for accurate performance.

But we know that bilinguals’ language development and their knowledge of their two languages is not the same as those of monolinguals (Cummins, 1981; Grosjean, 1982). This is because, first, bilingual children are hearing two languages, often in distinct social settings
(e.g., maybe one language with grandparents, the other language at the preschool), and often with less cumulative exposure to each language than a monolingual child from either language at the given age. At the same time bilingual children are experiencing some overlap in what they are learning about each language (e.g., learning two names for the same referent—apple and manzana—or two ways of talking about an action—he ran away and salió), and even some overlap of use in the same conversations, with code-switching common among bilinguals.

The unique ways in which bilingual or multilingual children experience language affects many aspects of development and the ultimate knowledge attained. This means that the timing of development can be different from that of monolinguals, that what they know in each language can be different and complementary (e.g., they might know the word for apples in one language, but not the other), that what they attend to may be different because of the “packaging” of concepts in the two languages, and that their ultimate organization of the two languages will be different from the organization of either language in the respective monolinguals, with potential links between the two languages at multiple levels.

In addition, there can be important differences across bilingual children: Most importantly, the relative timing for when children begin each of their languages, or are exposed to each language (simultaneously from birth; beginning the second language a little later than the first, early in their preschool years; only beginning the second language on entry to school; or even beginning the second language later in life; and so on), matters for what course we might expect their language development to take. If both languages are developing simultaneously, the child’s knowledge of the two languages may be relatively “balanced”, whereas if one language is begun after the first is already somewhat established, the latter may remain “dominant” for some time, until the exposure to the second language has become more extensive.
Because language tests are designed to give some indication of whether a child is developing according to our expectations relative to a child’s or speaker’s age, level, and experience, it is not really appropriate to apply a measure that was designed for one purpose (i.e., to gauge whether realistic expectations have been met for the course and timing of development in children who have been exposed to only one language) to use for a distinct purpose (i.e., to assess language development in children who have been exposed to two or more languages). This is true whether one simply wishes to know how well the child commands this language or one wishes to judge the child’s overall linguistic abilities. In the former case, it would be un instructive, for example, to use a measure designed for someone who has had, say, 5 years’ experience learning the language with someone who has had only one or 2 years’ experience.

But worse is the latter situation. When one wishes to determine, not only how much of language X a given child knows, but, more globally, whether a child is having particular difficulties with learning language per se, the importance of the language status of a child (monolingual vs bilingual) becomes heightened. Performing poorly on a second language (or on only one of a child’s two languages) is not the same as having a systemic problem with learning language—just as the multi-instrumentalist’s abilities in music cannot be totally gauged by only observing him or her playing the drums, or the athletic abilities of a child who has been playing baseball for 4 or 5 years and has just now begun track cannot be judged by observing his/her performance in track alone. To gain a full picture of a child’s language abilities, the ideal would be to examine performance in both languages of the bilingual child to determine if there is a linguistic problem. If the child is performing up to expectations in one of the languages and not the other, that indicates that there is not a problem with learning language per se, just that the child is behind with a particular language. This is a critical distinction to make. An impairment that affects a child’s abilities to learn language should affect both (or all) languages the child is trying to speak.
GOALS OF THESE VOLUMES

The chapters in these volumes attempt to tease apart some of the multi-faceted issues related to obtaining information on or determining the best ways of assessing bilingual children’s and adults’ abilities. The questions addressed include the following:

(a) What are the normal expectations regarding patterns of development in bilingual children?
(b) What are the normal expectations regarding ultimate linguistic abilities in bilingual populations and the roles of home language experience and community language in arriving at the mature command of the language?
(c) Is it important to assess both (all) of a child’s languages?
(d) What is the best way to determine if a bilingual child has language impairment—LI or SLI?
(e) How does a professional assess language abilities in an individual if the assessor does not know the language(s) in question?
(f) How do we gauge development in a language for which no normed tests exist?
(g) Should tests be normed specifically for bilinguals?
(h) Are there potential universal means of assessing language abilities in bilinguals from distinct language populations?
(i) What effect is there on bilingual language acquisition when there is highly variable input?
(j) What sociolinguistic influences can affect bilingual language acquisition?
(k) How can we best deal with the assessment of language abilities in school children for whom the community or school language is an L2 or for children from a multitude of different language(s) spoken in the home?
(l) Are there strategies teachers can use to help improve students’ acquisition of an early or later L2?
(m) Are there strategies teachers can use to build bridges between a child’s home language and the language of instruction?

(n) What are the ultimate ramifications of educational policy on language instruction and on the use of the heritage language versus the school language?

These questions run through the chapters in the two volumes and are addressed from the perspective of the experiences of researchers and professionals working with a variety of populations around the world. This volume raises many of the ISSUES related to the assessment of bilinguals; volume 2 presents some SOLUTIONS.

**ISSUES**

Each of the chapters in this volume raises important concerns that relate to assessment and to any solutions that may be proposed for the assessment of bilinguals, addressed in volume 2. These underlying issues are at the heart of why improvements are needed in the forms and nature of assessments we apply to bilingual infants, children, and adults—whether for the identification of language impairments; for children’s progress, either in language or beyond language, in education; for L2-learners’ abilities in childhood or adulthood; for general academic performance.

The first issue to be raised, in “Why assessment needs to take exposure into account: Vocabulary and grammatical abilities in bilingual children”, by Virginia C. Mueller Gathercole, Enlli Môn Thomas, Emily J. Roberts, and Catrin Hughes, is as follows:

**Issue 1: Normed assessments of language proficiency in children are in most cases normed on monolingually developing children.** Such assessments, without norms specifically based on bilingually developing children, misrepresent the linguistic abilities of bilingual children. Assessments are needed that take into account bilingual children’s level of exposure to the language in question.
Normed vocabulary and grammar tests are commonly used to assess children’s language abilities, and, often by extension, conceptual abilities. Such assessments aim to determine whether a child is progressing in line with expectations, as determined by the performance of a large group of similar children, the norming sample. However, those samples often involve monolingual children, and the assumption that bilingual children develop in a similar fashion to monolingual children is unwarranted.

The normal course of development of vocabulary and grammar in bilingual children, and the influence of exposure are examined in relation to Welsh-English bilingual children aged 2 to 15. The chapter traces the effects of exposure on timing of development. The data support the position that initially timing of development is linked to amount of exposure, but that eventually, greater parity is achieved across groups.

Furthermore, the data are examined for evidence of crossover links, or carryover from one language to the other, especially in relation to grammatical knowledge. The authors examine children’s performance on receptive tests of 13 types of grammatical structures in their two languages, Welsh and English. On the whole, the evidence is against carryover from one language to the other, at least on a linguistic level. Any links that can be observed appear to be attributable to links on another level, perhaps cognitive, perhaps meta-cognitive, perhaps meta-linguistics.

In addition, the ramifications for such data for assessment are discussed. The model followed by the Prawf Geirfa Cymraeg (Gathercole & Thomas, 2007; Gathercole, Thomas, & Hughes, in press) is advocated. In this model, bilingual children’s performance is measured against two standards of comparison – first, relative to all children from their age range, and, second, relative to children from a similar home language background, with a similar level of exposure. Each child receives two standard scores under this model – the scores together provide a comprehensive picture of the child’s abilities in relation both to the language in question and in relation to legitimate expectations, given the child’s level of exposure to the language.
A second issue is addressed in “Assessment of language abilities in sequential bilingual children: The potential of sentence imitation tasks”, by Shula Chiat, Sharon Armon-Lotem, Theo Marinis, Kamila Poličenská, Penny Roy, and Belinda Seeff-Gabriel. The issue they raise is as follows:

**Issue 2: Given the patterns of development in bilingual children, how is it possible for speech therapists to assess language abilities and identify language impairments or language deficits in bilingual children? This is an especially intractable issue when assessments are not available in the child’s L1 (and when, often, the assessor does not speak that L1).**

Given that bilingual children’s abilities develop at a distinct pace from those of monolingual children, one cannot simply use tests designed for monolingual children to determine the language abilities (or deficits) of children who speak an L2. Low performance by the bilingual child in the L2 may reflect limited exposure to the L2, not a developmental deficit. So what is needed is some way of testing children that can reveal language *abilities* but does not require that the child have extensive exposure to the language being tested.

The ideal might be to test children in their L1. But this is not always possible: In today’s context of mass migration around the world, there are increasing numbers of children whose home languages do not match the language of their (adopted) community, and increasing numbers of L1 language backgrounds represented among those children. (A recent survey of speech and language therapists in one borough of London, for example, found that during one month of 2011, therapists treated children who spoke 67 different languages (Shah, 2011).) In this context, it is next to impossible to develop batteries of tests to cover every pair of languages that are represented among the bilingual children. So the question is whether there is a way for professionals to test such children on their L2, the
community language, and be able to distinguish the language-impaired children from those who simply have not yet had adequate exposure to the L2 to perform well.

Chiat et al. explore one alternative, the possibility of developing a language-neutral assessment measure. They examine specifically the possibility of using a sentence repetition task, in the L2, to act as such a measure. The authors point out that it has been shown that performance on non-word repetition tasks is subject to knowledge of the target language in question, but they explore whether sentence-imitation tasks might be more impervious to such effects.

The authors report on the performance of four groups of bilingual children of immigrants in three countries, speaking Russian-Hebrew or English-Hebrew in Israel, Russian-German in Germany, and Turkish-English in the UK, on sentence-imitation tasks in their developing L2 (here, Hebrew, German, and English, respectively). They compare the children’s performance with monolingual norms, and also examine the extent to which age of onset (AoO) of the L2, length of exposure to the L2 (LoE), socioeconomic (SES) level, and the structure of the test itself play roles in affecting performance.

The results show that, indeed, a majority of bilingual children performed within the monolingual ranges for the tests. At the same time, the data show that a substantial number fell below that range. AoO seems to play a role in performance: Those who began the L2 before age 2 performed particularly similar to monolinguals, but at later AoO ages, other factors seem to affect how similar the L2 children perform relative to monolinguals. SES plays a role, in that at least for some of the data, children from lower SES levels are more represented in the low-performing groups than those from higher SES levels; LoE seems perhaps less influential; and structure of the test itself is significant.

The authors propose that such sentence imitation tasks show some promise and are worthy of further research. The tasks are clearly helpful in determining children’s proficiency in the L2, but they may also prove eventually to be a valid means, at least in some settings, of
singling out children with language impairment. As these authors suggest, the effectiveness and usefulness of the procedure await verification through evidence from further testing.

In the following chapter, “Assessing Yiddish plurals in acquisition: Impacts of bilingualism”, Netta Abegov and Dorit Ravid address the challenge of the assessment of bilingual children’s performance in their L1 when that L1 is highly variable and is under continuing influence from the dominant language in the larger community. The issue they address can be formulated as follows:

**Issue 3: Does one need to take into consideration the fluidity of language patterns in a bilingual community to assess bilingual children’s acquisition of complex morpho-syntax in the language(s) they are learning?** Most assessment measures rely on there being an established norm for each given structural element being assessed, and the test examines whether a child has reached that norm. But what happens when the morpho-syntactic element being examined is in flux, undergoing changes even in the adult language? How can we assess a child’s accomplishments in regard to that element?

In this chapter, the language of interest is Hasidic Yiddish spoken in Israel, used in a context in which the language is in a fluid environment in which there is widespread contact between Yiddish and Hebrew. These researchers examine the use of plural marking, a highly complex system, on words in this dialect of Yiddish. They first elicit the forms for referents from adult speakers of the language, and they find that adults show a high level of variation in the forms they use for each word. Almost half the words take 2 or more plural forms, some words taking as many as 4 forms or more. These plurals have evolved in a sociolinguistic setting in which new Germanic-based forms, as well as Hebrew and Loshn-Koydesh forms, have evolved. These researchers then elicited 60 of these forms from children aged 3 to 17. These children’s plural forms are highly consistent with the choices in the adult language, but they also show overextensions, some revealing influence from the contact language, Hebrew.
Such data raise the question of the best way of responding to the language of bilingual children growing up in a context of highly variable linguistic input and sociolinguistically fluid environment. How should one evaluate and assess their performance? What phenomena should be considered acceptable, as following the norm? Such questions are important, since bilingual children in such a contact situation across the world are often “penalized”, or their abilities are underestimated, because the forms they use do not reflect those of the standard norm. While children’s variable choices can be seen to eventually parallel those of the adults around them, in terms of the proportions of usage of the available forms for a particular lemma, how should children’s responses at earlier ages be construed? If younger children’s performance is in line with the general choices available in the language, even if not entirely consistent with the patterns for the particular given lemma, should this perhaps be taken as positive evidence of children’s sensitivity to the range of options available in the language they are learning?

In the following chapter “Measuring grammatical knowledge and abilities in bilinguals: Implications for assessment and testing”, Rocío Pérez-Tattam, Virginia C. Mueller Gathercole, Feryal Yavas, and Hans Stadthagen-González address an issue concerning assessment of bilingual adults:

**Issue 4: Simultaneous and early bilinguals may differ in subtle ways from their monolingual counterparts on their knowledge and performance on each of their two languages. Assessments of grammatical abilities often take for granted that L2 learners may well differ from native speakers in performance, but little attention is given to whether speakers who are bilingual from birth or early in life may also differ in performance in subtle ways from their monolingual peers.**

These authors explore the use of a receptive grammar measure similar to that used by Gathrecole et al (this volume), for Welsh-English children, with Spanish-English bilingual
adults. The adults represent a population whose acquisition of their two languages can be considered to have occurred in optimal conditions -- they are highly educated and grew up in an extensively bilingual community, Miami. They either were born and grew up in Miami or immigrated to the US as children and grew up as L1 Spanish-L2 English speakers in Miami; all grew up in homes in which mostly Spanish was spoken (OSH) or both Spanish and English (ESH) were spoken. These bilinguals provide an optimal picture of the grammatical proficiency in each of the two languages of a bilingual at the end-state.

As in the study of the Welsh-English bilinguals, 13 types of structures were tested in parallel tests for Spanish and English. The tests reveal a number of results. First, all of the bilinguals, regardless of group, performed equivalently in English. That is, no matter what the initial exposure to English in the home, all groups end up as adults with parity in their understanding of the English structures tested.

With regard to Spanish, however, there are differences across the groups. First, of the bilinguals, those who grew up in homes where both English and Spanish were spoken performed below those who were L1S-L2E bilinguals. There are also some interesting differences across the groups by structure. In one case, there is evidence that the monolingual Spanish speakers pay greater attention to an important Spanish-specific morphological marker than those who come from OSH and ESH homes. The marker in question is the object marker a, used specifically in relation to animate direct objects, as in Juan la/le quería a María 'Juan loved Maria' (cf. Juan quería helado 'Juan wanted ice cream'). In other cases, there is some indication that the bilingual speakers may be carrying over a distinction in English that is less relevant to Spanish. One of these involves the comparative versus the superlative: The bilingual speakers appear to make a more definitive demarcation between these for Spanish más ADJECTIVE versus (el) más ADJECTIVE than monolingual Spanish speakers do, possibly because of the clear distinction in English between the comparative and superlative in English. These effects suggest that, while the bilinguals on the
whole perform very highly in Spanish, there are some distinctions in their processing of items that differ in subtle ways in the two languages.

In addition, even though all of the participants were themselves college students or graduates, there were differences in their backgrounds as children. When performance is examined relative to mothers’ and fathers’ education and professions, the performance on English was consistent across groups, but for Spanish, those from higher SES levels, as judged by fathers’ professions, performed better. These results are consistent with those of Stadthagen-González et al (this volume), among others. (See also Gathercole & Thomas, 2009). They differ somewhat, however, from the results of the previous chapter showing that children’s developing abilities in their two grammars are not on the whole related. This, along with the results of Stadthagen-González et al, suggest that the linking of commonalities between a bilingual’s two languages is more an emergent property of their knowledge than a developmental property.

In another chapter, “Assessment of bilinguals’ performance in lexical tasks using reaction times”, Miguel Á. Pérez, Cristina Izura, Hans Stadthagen-González, and Javier Marín raise an issue with regard to assessing the total range of knowledge in a bilingual:

**Issue 5: If assessments of bilinguals’ language abilities aim to pinpoint exactly where their performance is similar to and differs from that of monolinguals, sometimes behavioral performance on vocabulary and morpho-syntax may not be enough. Performance on such tasks may mask underlying differences in the processing of language by bilinguals and monolinguals. Most assessments of bilinguals rely exclusively on such behavioral responses, so they do not capture fully the bilinguals’ abilities.**

These authors suggest that researchers should consider employing reaction time (RT) measures, so far uncommon in standardised tests, to assess language abilities in applied settings. Reaction times constitute a well-developed assessment measure used in cognitive
and psycholinguistic research, and these could easily be adapted for use in evaluating bilingual performance. These authors use reaction time measures to assess L2 learners’ acquisition of new vocabulary items, with special attention to whether assessments reveal important information on the efficiency with which learners acquire “early” words (i.e., among the first to be learned) vs. “later” words (learned after a earlier store of L2 words have been acquired). That is, they explore whether assessments accurately reveal order of acquisition effects.

These researchers provide a useful overview of the uses to which RT measures are put in experimental psycholinguistic studies, with special reference to their use in tasks of three types—lexical decision tasks, categorization tasks (including go-no-go tasks), and naming tasks (both picture naming and word naming/reading). They propose that because of advances in technology and the availability of resources, it is now possible to consider the benefits of bringing RT measures more directly into applied assessment settings.

They demonstrate the contributions that RT measures can make with a study examining the effect of order of acquisition (OoA) on processing and acquisition of words. OoA is a measure purported to be related to the relative timing or age at which words were acquired – e.g., bread might be acquired at age 2, while earthworm might be acquired much later, maybe after age 5 or so. Theories have been developed concerning whether words learned early are acquired, processed, or remembered more easily than those learned later. In a study in which these authors trained monolingual Spanish speakers on words from an unfamiliar language (here, Welsh), they examine the accuracy and RTs of learners according to the OoA of the given words. While their results reveal no difference on accuracy for the early versus later words, RT data reveal subtle differences in performance on the early versus late words. Those subtle differences provide evidence on the ease of learning and retaining words learned early versus those learned later. These authors argue that such effects can be
exploited by those wishing to develop more sophisticated or more subtle measures of bilinguals’ and L2 learners’ abilities with language.

In the following chapter, “Assessment and instruction in multilingual classrooms”, Rebecca Burns turns to the classroom and issues related to assessment and instruction in a multilingual classroom:

**Issue 6: How can teachers assess and instruct multilingual students in their classrooms when they themselves have no knowledge of the first language(s) of their students, and when assessments usually take place in the students’ L2?** Not only does this lead to an underestimation of the child’s knowledge, but it misses a prime opportunity to use the L1 to enrich the acquisition of the L2 and bridge the cognitive gap between learning through the L1 and the L2.

Burns raises this important issue for educators. As multilingual students and students limited in proficiency in the L2 used in the classroom become more and more the norm in our ever-shrinking world, the question of how the teacher might make use of the student’s L1 for instruction and assessment becomes critical. As Burns points out, studies have shown that young learners have greater academic success when the school they attend shows an appreciation for diversity. Most importantly, teachers’ use of the L1 of the child in assessments and teaching serves to support, rather than detract from, the acquisition of the L2 as well as to provide a bridge for the transfer of cognitive knowledge in academic domains.

The author takes a very practical approach and provides some useful guides for teachers on how to both assess and support children who are L2 English learners (ELL children) in their classrooms. Her suggestions revolve around the language resources increasingly available on the internet from around the world. She demonstrates how the teacher can draw on the L1 to boost literacy skills, to bridge between the L1 and the L2, to assess those skills, to provide content instruction (including translations), and to assess content knowledge. She provides something of a “cook book” of tips for teachers, and she
provides a useful compendium of resources on the internet that can be helpful in many ways to teachers. Some of these resources can provide translations, some materials in other languages, some information on the aural pronunciation of other languages, and so forth. In short, she argues and shows that the L1 can be the teacher’s ally in the L2 child’s education and evaluations of progress.

The following chapter, by Jasone Cenoz, Eli Arozena, and Durk Gorter, “Assessing multilingual students’ writing skills in Basque, Spanish and English”, continues the investigation of assessment in education. They address the following issue:

**Issue 7: How do children fare in contexts in which bilingual education is the norm, and children are exposed to a third language, in each of their languages? Does it matter whether children are taught through the medium of the minority language of the community vs the majority language? What level of proficiency do they arrive at in each language? And what is the evidence on academic achievement under these conditions?**

These authors address these questions by examining patterns of performance by Basque-Spanish bilingual children in the Basque country. The authors provide a lucid overview of types of bilingual education programs throughout the world, and follow with a detailed description of those available in the Basque country. Three educational models are followed there, and parents have a choice in which model they wish their children to attend. In *A model programs*, Spanish is the language of instruction and Basque is studied as a second language. In *B model programs*, both Basque and Spanish are languages of instruction, each for approximately 50% of school time. Basque is also studied as a second language in this model. In *D model programs*, Basque is the language of instruction and Spanish is a school subject. In addition, although several languages are options to study as a compulsory foreign language, English is increasingly chosen.
Cenoz et al. review the evidence, first, on the academic performance of children growing up in these educational settings. All evidence shows that the children perform similar to or above their peers in other European countries on mathematics, science, and reading literacy. The authors probe further and analyze where there may be differences across the three models in children’s performance in these areas and across children with Basque versus Spanish as their first language.

In addition, they report on a study of teenagers’ compositions in their three languages, Basque, Spanish, and English. All children were enrolled in D model schools, but came from L1 Basque or L1 Spanish homes. The children’s overall achievements in Basque and Spanish were comparable. However, closer analysis of sub-skills in the writing performance showed that on the Basque compositions, while the two groups were comparable on content and organization, the L1 Basque children performed better on vocabulary, use, and mechanics. On the Spanish compositions, the L1 Basque and L1 Spanish children performed in equivalent fashion. On English, the L1 Basque speakers obtained higher scores overall than the L1 Spanish speakers, and the L1 Basque children outperformed the L1 Spanish children on vocabulary, use, and mechanics. The authors discuss their results in terms of the possible relative contributions of language status (majority/minority), relative balance of the languages in the individual bilingual, and “multilingualism” of the individual; they also comment on the relevance of these data to language assessment and argue for an approach that has a focus on “multilingualism”, involving the whole linguistic repertoire of those being assessed.

A final chapter is concerned with the assessment of bilinguals in educational settings: “Assessment of academic performance: The impact of no child left behind policies on bilingual education: A ten year retrospective”, by Stephen J. Caldas. The issue addressed here is the following:

**Issue 8: Have the treatment and stance towards bilingual children that have been implemented in relation to the No Child Left Behind policies been beneficial**
or detrimental for bilingual children’s education? What evidence is there from academic assessments on the benefits of these policies?

Caldas provides an authoritative examination of the effects of the No Child Left Behind legislation in the US on the education of bilingual children. He first provides a history of NCLB legislation, from its inception, in the Elementary and Secondary Education Act of 1965 to its emergence as the No Child Left Behind legislation in 2001. He traces the push and pull of political and philosophical preferences of the time on education practices affecting children’s education, suggesting that while the main focus has been on academic performance, there has been increasing attention with time on measurable outcomes. Caldas notes that, in principle, NCLB is not necessarily anti-bilingualism, but there is a strong push to make the attainment of English in children a priority. He notes that the NCLB has had two important effects: first, since tests of content areas are administered in English, they become tests of English proficiency (and, hence, foster educators’ emphasis on learning English), second, educational policy is being driven by testing policies.

In a revealing study, Caldas examines the performance of bilingual children (specifically ELL Hispanics) both before and after the NCLB legislation on national tests of mathematics and reading. He reports that mathematics scores increased more during the pre-NCLB period (for all groups, in fact, including non-bilinguals) than since the implementation of NCLB. The increase for Hispanic ELLs was from two and one-half to seven times greater (at grade 4 and grade 8, respectively) during the period prior to NCLB’s implementation than for the period during which NCLB was the law. For reading, scores for the ELL Hispanics showed a significant decrease in performance in the eighth-grade scores during the NCLB period. Furthermore, in fact, there was actually a statistically significant increase in the gap between non-ELL and ELL Hispanic eighth-grade reading scores between 2002 and 2009. So, gaps between Hispanic ELLs and non-ELLs actually increased significantly on the eighth-grade test.
Such results are a damning indictment of the effects of the NCLB legislation. When education draws on the resources that a child has available (e.g., through understanding in the L1), the child is more likely to succeed academically than when the child is forced to perform in an L2 for which s/he has a more tenuous grasp.

**ISSUES AND IMPORT**

These chapters together provide some important directions for movement forward towards more responsive approaches to the assessment of bilinguals. Among the principles that emerge from this work are the following:

1. If at all possible, bilingual children should be tested in both of their languages. When this is not possible, they should be tested in their dominant language.

2. When tests do not exist for a particular language, and a researcher wishes to develop a test for that language, it is important to be sensitive to (a) the structure of that language, (b) normal developmental stages children pass through in the acquisition of that language, (c) possible crossover or transfer opportunities between the bilingual’s two languages (e.g., the presence of cognates), and (d) possible ramifications for acquisitional patterns associated with growing up in a bilingual community.

3. Performance on a test for one language does not necessarily correlate with performance on any other.

4. If at all possible, tests should be normed on bilingual children from similar language backgrounds and exposure. If possible, two sets of norms may be appropriate.

5. It may be possible to discover means for testing language abilities in bilinguals by testing in the L2. Researchers here have suggested that such testing may take the form of repetition tasks, response to intervention, or reaction time measures.

6. Even fully fluent end-state bilinguals differ in subtle ways from fluent monolinguals. They may differ in the range and organization of vocabulary and in the processing of language-specific structural features, as in the case of the Spanish object marker a.
7. The child can build on knowledge of the L1 for the acquisition of L2 abilities, literacy abilities, and academic content. Teachers can make use of the L1 abilities in making bridges to the L2.

One of the resounding themes in these and other works on bilinguals is that difference does not imply deficiency. Yo-Yo Ma plays the cello differently from the way that Jacqueline du Pres did, yet both have been hailed as among the greatest instrumentalists of our time. The manner in which Fred Astaire, Gene Kelly, and Mikhail Baryshnikov danced differed dramatically, yet all excelled in their performance. Bilinguals are by definition different from monolinguals; that difference can be celebrated. The chapters in the book argue that we can only do justice to bilingual children and adults if we recognize those differences and draw on them in our treatment and assessment of their abilities.

The chapters in volume 2 attempt to incorporate many of these suggestions into new forms of assessment, to help offer solutions to the many issues raised here.
REFERENCES


Guthrie, G. (2003). *1,600 quotes & pieces of wisdom that just might help you out when you're stuck in a moment (and can't get out of it!)*. Lincoln, NE: iUniverse, Inc.