How Language Proficiency Tests Mislead Us About Ability: Implications for English Language Learner Placement in Special Education

JEFF MACSWAN
Arizona State University

KELLIE ROLSTAD
Arizona State University

*The authors argue that English language learner (ELL) language assessment policy and poor language tests partly account for ELLs' disproportionate representation in special education. Previous research indicates that many states routinely assess ELLs' first language (L1) at initial enrollment and that ELLs identified as limited in both languages have relatively high rates of identification in special education. Two common tests, the Language Assessment Scales–Oral (LAS–O) Español and the Idea Proficiency Test I–Oral (IPT) Spanish, are shown to identify 74% and 90%, respectively, Spanish-background ELLs (*N = *145) as limited L1 students, whereas a natural language measure found only 2% of participants to have unexpectedly high morphological error rates. Correlations are provided. The authors recommend changes in language testing policies and practices for ELLs.*

English language learners (ELLs) are overrepresented in special education programs, a problem that has persisted since its earliest documentation in the 1960s (Artiles & Trent, 1994; Artiles, Trent & Palmer, 2004). In a study of within-group diversity of disproportionate representation of ELL students in special education, Artiles, Rueda, Salazar, and Higareda (2005) found that ELLs identified by districts as having limited proficiency in both their native language (L1) and English (L2) showed the highest rates of identification in the special education categories investigated, were consistently overrepresented in learning disabilities and language and speech disabilities classes, and had greater chances of being placed in special

Teachers College Record Volume 108, Number 11, November 2006, pp. 2304–2328

0161-4681

education programs as compared with the other district-defined subgroups of ELLs examined in the study. The present study looks specifically at select instruments used to assess ELLs' oral L1 ability and raises questions about their validity. In this context, we discuss ELL placement in special education, suggesting that the confluence of policies and practices encouraging L1 oral language testing—used with language minority students but not with others—and poorly designed language tests disproportionately increases the chances that ELL children will be referred for special education assessment (and, ultimately, placement) because of poor performance on L1 tests.

We begin with an overview of current language testing policy for ELLs and then locate the intellectual origins of aspects of this policy in the history of deficit psychology. We then present results from a validity study of native language tests designed for ELLs in which coded speech samples of Spanish-speaking children are compared with Spanish-language test results, and we show that the language test results are dramatically misleading with respect to the actual Spanish-language ability of ELLs, as grounded in a theoretically defensible view of native language proficiency, discussed below.[1] We suggest that placement in special education is likely to be shaped by the significant limitations that we identify in these language assessment tests, and we offer recommendations for improvements.

ELLs AND LANGUAGE ASSESSMENT POLICY

An important responsibility of schools in the United States is to determine whether a child knows English sufficiently well to succeed in an all-English instructional setting. In *Lau v. Nichols* (1974), the U.S. Supreme Court interpreted Title VI of the Civil Rights Act to prohibit discrimination against language-minority children by ignoring their special language-related needs. Thus, schools must determine, for every child enrolling in school, whether the child is an English learner. For students identified as ELLs, states must offer theoretically defensible programs aimed at teaching English and provide a comprehensible school curriculum comparable with that provided for English-speaking students (Crawford, 2004).

An evaluation of children's English ability is appropriate—indeed, imperative—in light of these considerations. However, many states also require or recommend assessment of children's native language ability, with the result that numerous children are identified as "non-nons"—that is, nonspeakers of both English and their home language. In a survey of state practices and policies, Mahoney and MacSwan (2005) found that 13 states require or recommend that ELL students undergo an oral native language assessment in addition to an English assessment as part of the ELL identification process. These requirements/recommendations affect about a

Table 1. States and Other Federal Entities That Require or Recommend Native Language Assessment for English Language Learners (ELLs) as Part of the Identification Process, and the Number of ELLs in Each of These States as of 1999–2000

State	ELL Students	Total State Enrollment	Percentage of Total State Enrollment
Arizona	125,311	850,840	14.73%
Connecticut	20,190	553,993	3.64%
District of Columbia	5,177	77,194	6.71%
Hawaii	12,879	185,860	6.93%
Illinois	143,855	2,027,600	7.09%
Mississippi	1,799	500,716	0.36%
Ohio	16,841	1,836,554	0.92%
Oklahoma	28,823	627,032	4.60%
Oregon	43,845	545,033	8.04%
South Dakota	5,495	131,037	4.19%
Texas	554,949	3,991,783	13.90%
Virgin Islands	1,223	20,866	5.86%
Virginia	31,675	1,133,994	2.79%
Total	*1,043,614*	*12,482,502*	*8.36%*

quarter of the nation's 4,416,580 ELL students, who in turn make up 9.33% of the nation's total reported population of 47,356,089 students (Kindler, 2002). Relevant state-level student frequencies and percentages are shown in Table 1. Assessments of this nature may also be carried out elsewhere at the initiative of districts and schools in the absence of explicit state policy, which has been in flux in recent years under the influence of the No Child Left Behind Act (NCLB) of 2001. Table 2 presents the most commonly used tests of native language proficiency, as reported by state survey participants (Mahoney & MacSwan).

Explicit rationales for administering native language assessments to ELL students are hard to find. Although it is difficult to know the origin of the practice, the belief that minorities may have "limited language ability" and that this limited ability is related to difficulties at school has a long history in educational psychology and contemporary bilingual education theory.

Table 2. Tests Used by States for Native Language Assessment

Instrument	Number of States
Language Assessment Scale–Spanish (LAS)	11
Idea Proficiency Test–Spanish (IPT)	10
Woodcock-Munoz–Spanish	5
Bilingual Inventory of Natural Language (BINL)	1
Bilingual Syntax Measure–Spanish (BSM)	1
Peabody Picture Vocabulary Test–Spanish (PPVT)	1

Below, we explore conceptions of language ability in minority children and argue that it is conceptually indistinguishable from classical prescriptivism, the view that some language varieties are inherently inferior to others.

CONCEPTIONS OF LANGUAGE ABILITY AMONG MINORITY CHILDREN AND CLASSICAL PRESCRIPTIVISM

Dittmar (1976) traced the origin of what he termed "the Deficit Hypothesis" to Schatzmann and Strauss (1955), who had interviewed members of the lower class and middle class about their impressions and experiences after the occurrence of a disaster. They found that the former used lots of emotional language, which reputedly gave rise to what the authors called "elliptical syntax." Schatzmann and Strauss concluded that the lower classes only conveyed their meaning "implicitly," while the educated classes conveyed their meaning "explicitly."

Bereiter and colleagues (Bereiter & Engelmann, 1966; Bereiter, Engelman, Osborn & Reidford, 1966) similarly sought to explain lower educational achievement of African-American preschoolers by pointing to inherent linguistic deficiencies. These researchers reported that the four-year-olds they studied communicated by gestures, "single words," and "a series of badly connected words or phrases" (Bereiter, Engelman, Osborn & Reidford, 1966, p. 114). The authors reported that, "without exaggerating," the children in the study could "make no statements of any kind," and could not ask questions (p. 114). As one component of the assessment, children were asked to look at a picture of a squirrel in a tree and answer the question, "Where is the squirrel?" In response, children tended to answer, "In the tree," a response which Bereiter and colleagues characterized as illogical and badly formed; rather, an answer expressed as a complete sentence was required, such as "The squirrel is in the tree" (Bereiter, Engelman, Osborn & Reidford, 1966, p. 121).

As Dittmar (1976) noted, deficit theories of language ability typically define the characteristics of "better speech" in terms of those characteristics that poor people *lack*. In other words, rather than looking to a theory of language structure and acquisition to define what is linguistically well formed or developmentally appropriate, dichotomies are generated that position the language of the educated classes as the developmental goal or an improved version of the language of the unschooled.

In the context of distinguishing a *disability* from a *difference*, Artiles and Trent (1994) noted that

> the notion of disability is concerned with atypical functioning or educational performance due to biological, psychological, and/or social factors. The level of functioning for individuals with disabilities falls in

> the lower portion of the normal distribution curve. The notion of disability exists because we have established parameters to judge when a person functions anatomically, physiologically, intellectually, and/or psychosocially within the limits of what is considered typical. On the other hand, cultural diversity is not defined—at least theoretically—by a standard parameter of functioning. Although it is also concerned with the idea of difference, it is not—unlike the disability construct—inherently linked to the notion of deviance. (p. 424)

Humans learn because they are innately (biologically, anatomically) endowed with the capacity to do so. Their biological makeup interacts with a specific environment that is socially and culturally situated. With regard to language acquisition, we expect children to acquire the language of the specific speech community in which they grow up, along with whatever features of the language that might be stigmatized in the dominant social group. If a child successfully acquires the language of her speech community, we view the learner as functioning normally from a linguistic point of view. If not, there may be reason to suspect that the child has a language-related learning disability. However, whether the child's language is in any way similar to that of another speech community—for instance, the community of speakers who constitute the educated classes—is entirely irrelevant to the question of whether the child speaks her language fluently or proficiently.

Valencia (1997) defined a deficit model as a theory that posits "that the student who fails in school does so because of internal deficits or deficiencies" manifested "in limited intellectual abilities, linguistic shortcomings, lack of motivation to learn and immoral behavior" (p. 2). The transmitters of these deficits, according to Valencia, have typically been located in genetics, culture, class, and familial socialization. A linguistic deficit theory, more narrowly, attempts to legitimate the social stratification of linguistic differences by positing the existence of properties of the language system that in some way represent the socially stigmatized variety as inherently inferior to other varieties. Dichotomies such as explicit/implicit, formal/informal, and restricted/elaborated are used to label these differences, with no explicit linguistic arguments presented to justify the claim that the varieties are in some way hierarchically related. Or, as in the case of Bereiter's test, the linguistic behavior of a member of the educated classes who is knowingly performing an academic task is arbitrarily used as the standard of linguistic correctness, again offered without justification (Labov, 1970).

It is to be noted that verbal deficit theory is conceptually closely related to classical prescriptivism, the view that one or another language or variety of language has an inherently higher value than others (Crystal, 1986; Pinker, 1994; Postal, 1972). Prescriptivists have often characterized minority languages or language varieties as "inexpressive," "primitive," or lacking

complexity in comparison with their own language. The "standards" regarding English usage, which are familiar in U.S. language arts curricula and found in influential prescriptive grammars, typically draw upon Latinate analyses advanced in the late 19th and early 20th centuries and used to validate varieties of speech associated with the educated classes in England and the United States (Baugh & Cable, 1978). Linguists, by contrast, took on the project of describing a wide range of linguistic diversity using the same taxonomy for all languages; this project, spurred by leading figures of American structuralism like Leonard Bloomfield, led to the conclusion that all languages, even the so-called primitive languages, were equally complex. This research agenda seriously threatened traditional distinctions used to legitimize the concentration of privilege and social prestige in the hands of the educated classes. Newmeyer (1986) adds,

> As long as American structuralists confined their campaign to the languages of remote tribes, they did little to upset their colleagues in departments of modern and classical languages—in which almost all linguists were situated in the interwar years. But such was certainly not the case when they began crusading for the linguistic equality of *all* dialects of English and other literary languages, no matter how "substandard" they were regarded. This egalitarian view came in direct conflict with the long-seated tradition in the humanities that values a language variety in direct proportion to its literary output. (p. 42)

Verbal deficit theories are also deeply embedded in contemporary research on the education of ELLs, typified by Cummins' threshold hypothesis and well-known distinction between basic interpersonal communication skills (BICS) and cognitive-academic language proficiency (CALP). Cummins (1979) argued that "semilingualism"—later termed "limited bilingualism" (Cummins, 1981)—could be used to explain bilingual students' difficulties at school as part of the threshold hypothesis. Cummins defined semilingualism as "low level in both languages," or "less than native-like command of the vocabulary and syntactic structures" of both languages (1979, pp. 230, 238). Although the threshold hypothesis is widely accepted, no empirical evidence has been presented to support the "semilingualism" thesis embedded within it. Paulston (1983), for instance, reviewed numerous Scandinavian studies that sought linguistic evidence for the existence of semilingualism in Sweden. She concluded that "there is *no* empirical evidence to support the existence of such a language development hiatus as [semilingualism]" (p. 42). More recently, MacSwan (2000) reviewed reputed evidence from studies of language variation, linguistic structure, school achievement, and language loss, and concluded that all of it was either spurious or irrelevant to the construct.

Cummins has expressed a preference for the term *academic language* over CALP, and *interpersonal communication skills* over BICS in recent discussions of the BICS/CALP distinction, but the basic definition of the dichotomy remains unchanged (Cummins, 2000). According to Cummins (2000), "Considerably less knowledge of language itself is usually required to function appropriately in interpersonal communicative situations than is required in academic situations" (p. 35), whereas academic language generally involves "much more low frequency vocabulary, complex grammatical structures, and greater demands on memory, analysis, and other cognitive processes" (p. 36). A problem here, as with semilingualism and other varieties of the verbal deficit theory, lies in equating the language of school—and hence the language of the educated classes—with language that is inherently more complex and richer, and that places greater demands on cognitive resources. It follows from these assertions that the language of school is in some regard an improved version of the language of other contexts, a claim that is not adequately supported by the empirical and theoretical arguments presented.

Indeed, to demonstrate these claims, proponents must present some reasonable evidence that academic language actually has the properties that they attribute to it. The matter is especially compelling given the similarities with traditional prescriptivism. The attribution of special "complex grammatical structures" and greater cognitive demands to the language of the educated classes would appear to have disturbing implications for the language of the unschooled or of children of lower socioeconomic status, and hence for their cultural and linguistic identities (see MacSwan & Rolstad, 2003, for further discussion.)

Moving beyond the taxonomic analyses of the early structuralists, linguists began to study the nature of the mind/brain and its relation to acquisition. This research tradition, now well established, attributes our knowledge of language and our ability to acquire it to innately given properties of our biology, with peripheral effects of the environment visible in the form of surface-level cross-linguistic differences. As Chomsky (1965) noted,

> A consideration of the character of the grammar that is acquired, the degenerate quality and narrowly limited extent of the available data, the striking uniformity of the resulting grammars, and their independence of intelligence, motivation and emotional state, over wide ranges of variation, leave little hope that much of the structure of language can be learned by an organism initially uninformed as to its general character. (p. 58)

Chomsky's statement paints a picture of children as inwardly driven language learners who acquire their language perfectly and without instruction.

Indeed, during the most active acquisition period (ages 2–6), children learn approximately 10–12 new words a day, often on one exposure and in highly ambiguous contexts (Gleitman & Landau, 1994). Children know things about elementary aspects of sentence structure for which they have no evidence at all (Pinker, 1994), and in cases of creolization, children acquire syntactically and morphologically complex linguistic systems in accordance with principles of Universal Grammar in the presence of highly degenerate, rudimentary adult language input (pidgins) (Bickerton, 1981). In an extensive review of research on child language in the preschool years, Tager-Flusberg (1997) reported that "by the time children begin school, they have acquired most of the morphological and syntactic rules of their language" (p. 188) and possess a grammar essentially indistinguishable from adults.

Given these facts, it is surprising to find that numerous ELLs in the United States are classified as "non-nons," children with limited ability in both languages, while majority language children do not tend to be so classified, and indeed are not even tested. The practice of testing the native language ability of ELL students appears to have arisen from the expectation that such children may have inherent linguistic deficiencies, an expectation that is likely rooted in persisting deficit models in educational psychology and language minority education. If the tests used to so classify children are likewise found to be rooted in erroneous conceptions of language ability, we may conjecture that the overrepresentation of ELLs in language and speech disabilities classes observed by Artiles and colleagues (2005) is likely to be an artifact of poorly designed tests routinely administered as the result of ill-conceived language testing policy. In the next section, we explore this possibility empirically.

RESEARCH QUESTION, METHODS, AND DATA ANALYSIS

RESEARCH QUESTION

We are concerned with the following research question: Are common native language tests, used to identify many ELL children as having limited ability in the language of their own speech community, accurate measures of these children's true language abilities? As examples of common native language tests, we focus here on the two most frequently reported (Mahoney & MacSwan, 2005; see Table 2): the Language Assessment Scales–Oral-Español (LAS–O Español; De Avila & Duncan, 1994)[2] and the Idea Proficiency Test Spanish I–Oral (IPT Spanish; Williams, Ballard, Tighe, Dalton, & Amori, 1998). In referring to the "true language abilities" of these students, we intend to denote children's language abilities as understood descriptively,

taking the standard of correctness to be the language of the actual communities in which children acquired their native language rather than any other language community or context.

STUDY PARTICIPANTS AND CONTEXTS

Participants were selected on the basis of three criteria: native language background (Spanish), age (6–8), and English proficiency (nonproficient or limited proficiency, as determined by the English LAS, the test adopted by the school districts). Students included in the sample were predominantly the children of Mexican nationals, enrolled in grades 1–3, of lower socioeconomic status, and represented a distribution of both male and female students. All students meeting the selection criteria at each school were invited to participate in the study, and all who responded affirmatively to the invitation were included in the sample (estimated to be about 80% of those invited). None of the students included in the study had been identified as special education students.

Participants were situated at two urban public schools in central Arizona within separate districts with predominantly low-income and racially diverse student bodies. Six native-Spanish-speaking graduate and undergraduate research assistants (five of Mexican and one of Panamanian background) administered the language tests and interviewed the students for the purpose of obtaining the language samples. All interviews were conducted in Spanish on-site at the two schools. Research assistants were trained to administer the language assessments by a certified bilingual school psychologist who was also a graduate student researcher associated with the full range of project responsibilities.

Because of scheduling difficulties and sample attrition, we were not able to collect all data points on all study participants. For instance, although 180 students participated in some way in the study, language samples were collected on 145 of these, the LAS–O Español was administered to 161 students, and the IPT Spanish to 174. Other tests were also administered, but only the results from the two most common tests are presented here.

INSTRUMENTS

Language Assessment Scales–Oral Español (LAS–O Español)

As its name suggests, the LAS–O Español is an oral test of Spanish and is individually administered. Two forms are available, 1B (grades 1–6) and 2B (grades 7–12). The short form of Form 1B, used in the study, consists of three parts. Part 1 is intended to assess children's vocabulary and is made

up of 20 items; the examiner asks the child to identify words for places and actions illustrated on picture cards. Part 2 assesses listening; the examiner plays a tape recording illustrated with a picture and asks the child to answer 10 comprehension questions. In part 3 of the test, the child is asked to listen to a tape-recorded story, again illustrated with pictures; the child retells the story, which the examiner transcribes. Part 3 is scored holistically on the basis of a rubric moving from 0 (*no response or "I don't know"*) to 5 (*"articulate and elaborated" speech*; De Avila & Duncan, 1989, p. 17).

The test developers refer to the theory of language proficiency underlying the test construct as the probabilistic approach; the approach assumes a linear relationship between linguistic proficiency and academic achievement, and defines the passing score as the point at which children's language proficiency intersects with the 50th percentile on a test of academic achievement. The developers do not address the linguistic characteristics of the test and how the items relate to a specific theory of linguistic knowledge, except to say that the test items were selected "according to linguistic theory and prior experience [to tap] different elements of oral processing thought to be important in school" (De Avila & Duncan, 1989, p. 3). For Form 1B, the developers reported an alpha coefficient of 0.9572 for part 1 (Vocabulario), 0.886 for part 2 (Vamos a Escuchar), and interrated reliability of 0.877 and 0.837 (depending on the prompt used) for part 3 (Cuentos). The test is intended to be used for the purposes of identification, placement and reclassification of ELL students. Student scores are classified as *nonproficient Spanish speaker*, *limited Spanish speaker*, or *proficient Spanish speaker*.

The Idea Proficiency Test I—Oral, Grades K–6—Spanish, 2nd Edition (IPT Spanish)

The IPT Spanish consists of six parts, called levels A–F. Each level is designed to test a variety of skill areas and is intended to reflect an increasing level of difficulty as students progress from one level to the next. At the end of each level, a score box indicates whether a student's score within a level warrants that the student stop the test at the given level or advance to the next. In discussing theoretical considerations related to language proficiency, the test developers indicate that "theories of language acquisition and language learning have been taken into consideration in the construction of the IPT I-Oral Spanish" (Amori & Dalton, 1996, p. 3). Each test item is explicitly associated with a specific skill area (vocabulary, comprehension, syntax, and verbal expression), BICS or CALP, a placement on Bloom's Taxonomy, and a hypothesized stage of language acquisition. The test developers reported an alpha coefficient of 0.99. A table defines a student's level score as *non-Spanish speaking*, *limited Spanish speaking*, or *fluent Spanish speaking*, factoring in grade level.

In the present study, both the LAS–O Español and the IPT Spanish were scored by native-Spanish speaking research assistants, except in the case of the more subjective story retelling task of the LAS–O Español; to protect against researcher bias, this section was scored externally by a professional consulting firm recommended by the test publisher.

Natural Language Samples

Linguists concerned with the study of child language acquisition and language disabilities collect natural language samples as standard practice. A large corpus of such samples, along with tools for analysis, is maintained by Brian MacWhinney as part of the CHILDES (Child Language Data Exchange System) Project.[3] Using common methods in the study of language acquisition, the children involved in the present study were asked to interact with a native speaker of Spanish and to tell a story about a boy and a frog from a Mercer Mayer picture book with no text (Mayer, 1969). Spanish speech samples of each child telling the whole story depicted in the picture book were individually videotaped, transcribed word for word, and coded using MacWhinney's (2000) standard CHAT (Codes for the Human Analysis of Transcripts) format, as modified by Curtiss, MacSwan, Schaeffer, Kural, and Sano (2004a) and adapted to Spanish by Valadez, MacSwan, and Martínez (2002). The accuracy of transcription and coding was double-checked by a second transcriber/coder; all interviewers and coders were native speakers of Spanish. All but one transcriber was a native speaker, and the one nonnative was highly proficient and always second-checked by a native speaker. Differences of opinion regarding transcription or coding required consultation and resolution among members of the research team.

Our coding system emphasized grammatical morphology for a variety of reasons. Research on child language development, largely due to the influence of Brown (1973), has focused on the development of grammatical morphemes as an index of a child's linguistic maturity. More recently, however, linguistic theory has become increasingly focused on the role of grammatical morphemes and functional categories in syntax. In this perspective, the mapping of linguistic structure is assumed to consist of two components: (1) a lexicon, which varies across languages and, to a lesser extent, across individuals, and (2) a mental system of computational rules and principles, taken to be invariant across human languages. Differences between, say, Spanish and English, or between any two languages, relate to difference in the lexicon, mapped by the computational system into various surface forms. These linguistic differences are generally taken to be confined to the functional categories of the lexicon, which bear inflectional morphemes (sometimes abstract). This theoretical framework, adopted here, represents the

current instantiation of generative grammar as developed within contemporary linguistic theory (Chomsky, 1995, 2004).

On this perspective, knowledge of language is understood to be a purely linguistic construct, reflecting a grammatical system which consists of the rules and principles that govern syntax (word order), morphology (principles of word formation), and phonology (pronunciation), and that interface with principles of discourse, pragmatics, and semantic interpretation. Speakers and communities differ with regard to the particular form that these principles might take, resulting in the formation of distinctive varieties and conventions on language use; but each community nonetheless has a language just as rich and complex as the next (Crystal, 1986; Milroy & Milroy, 1999; Newmeyer, 1986). Because language is an inherent human ability, it becomes extremely important to distinguish it from other domains of knowledge, such as academic knowledge acquired in the specific cultural setting of school. School, like any environment, will have effects on children's language, but the specific ways in which school alters our language do not amount to qualitative differences from a linguistic point of view.

For purposes of illustration, we present the following example of a coded utterance, with translation provided in brackets:

*MAR: El niño se está durmiendo, y la rana se escapó.
[The boy is going to sleep, and the frog escaped]
%mor: DART|el D|niño REF|se IAUX|está-3Ss dormir-ido
CONJ|y DART|la D|rana REF|se IT|escapar-r3spret
%lex: N|niño N|dormir N|rana V|escapar

Errors of selection (for instance, where *la* is used when *el* is required for the morphological category DART) are prefixed with = (equal sign); errors of omission (where a category such as DART or IAUX is missing altogether) are suffixed with = 0. The coding system was developed around three functional systems known as the I-system (inflectional), the D-system (determiner) and the C-system (complementizer). Each code on the morphological tier (%mor:) is separated by a vertical line (|) from the morpheme that it classifies. Inflectional morphemes are separated from stems with a hyphen (-). The system permits the calculation of morphological error rate from the total number of functional categories and total number of errors in such categories, per transcript. A full implementation of the syntactic coding system of Curtiss and colleagues (2004b) was not needed given the specific aims of the study; instead, each utterance that evidenced an anomaly in word order was flagged as an error on the syntactic tier of the coded transcript. The syntactic error rate was calculated from the total number of utterances and total number of utterances flagged as syntactically ill formed in each transcript. A detailed description of the coding system and the

significance of other codes illustrated above may be reviewed in Curtiss and colleagues (2004a, 2004b).

It is important to note that a form was considered an error only if its presence or absence did not conform to the language of the child's speech community. For example, in an expression such as *el rana*, the article would be marked as a selection error; if missing altogether but contextually required, it would be marked as an error of omission. However, regional variation such as *pa' tras* (*para atras*, "over there") and *onde* (*donde*, "where"), among others, were not marked as errors for children whose speech communities used such forms. Determinations regarding acceptable regional variation were made in consultation with adult members of relevant speech communities and with reference to published documentation regarding linguistic variation in Spanish, principally Lipski (1994).

As described by Curtiss and colleagues (2004a), the validity of this coding system is tied to an external criterion—namely, linguistic theory—developed out of a rich history of empirical inquiry. Reliability, a necessary condition of validity (American Psychological Association and National Council on Measurement in Education, 1985), indicates the degree to which repeated coding events of the same transcript will yield similar measures. For each utterance and structure in a given transcript, coders must render a judgment regarding the grammaticality of the expression. If different coding events for the same transcript involve different grammaticality judgments on the part of coders, then scores will differ with respect to the measure of error in the respective structure or category under analysis. To guard against this threat to validity, we invoked Labov's (1975, p. 31) consensus principle ("If there is no reason to think otherwise, assume that the judgments of any native speaker are characteristic of all speakers of the language") and clear case principle ("Disputed judgments should be shown to include at least one consistent pattern in the speech community or be abandoned. If differing judgments are said to represent different dialects, enough investigation of each dialect should be carried out to show that each judgment is a clear case in that dialect").

FINDINGS

Table 3 presents frequencies of students' Spanish language proficiency levels by test. Frequencies are presented for all students in the study for whom LAS–O Español or IPT Spanish scores were available and for the subset of participants for whom we were also able to collect natural language samples. Note that the distributions are similar for both groups, with 74% ($N = 119$) of all students testing below the expected "fluent" threshold, and 73% ($N = 95$) of students with language samples available testing below the

fluent threshold on the LAS–O Español. With regard to the IPT Spanish, we see remarkably few students scoring in the fluent range, at only 10% ($N = 17$) for all students and only 9% ($N = 13$) for students for whom a natural language sample was available. In both instances, approximately 90% of students fall below the fluent threshold. These data are presented to demonstrate that large numbers of ELLs are identified as less than fluent in their native language by both tests. The subgroup of students for whom language samples were also available is presented to show that the distribution of scores for the whole group and the subgroup is similar for each test, increasing our confidence that the subgroup is representative of all the students in the study with language test scores.

Now consider Table 4, which presents frequencies of morphological and syntactic error rates for study participants by specified ranges and by test, and Table 5, which presents these frequencies for all students in the study for whom a natural language sample was available. Here one sees a dramatically different picture of the language ability of the children tested. For example, whereas the LAS–O Español identified nearly three quarters of this group as limited in their L1, the analysis of the natural language sample shows the proportion of error to be highly constrained within a very narrow range, falling at 5% or less for the vast majority of students, and at 10% or less for 97% of students for morphology, and 100% of students for syntax.

Figure 1 graphically illustrates differences in the distribution of proficiency ratings based on the LAS–O Español, the IPT Spanish, and the natural language sample. For purposes of Figure 1, we will regard children with relatively low and expected morphological error rates (below or near 10%) as proficient speakers of Spanish, and those with somewhat higher morphological error rates as "limited" on the natural language sample (we discuss the meaning of "limited" in this context below).

Our research question asks whether common tests used to identify ELLs as having limited ability in their native language measure the true language ability of these students. To address this question, we calculated correlation coefficients among the LAS–O Español, the IPT Spanish, and morphological error rate, displayed in Table 6. Error rate is inverted in the table so that higher values will indicate stronger performance, making them parallel to the proficiency levels of the LAS Español and the IPT Spanish. Although some of the correlations are statistically significant, the relationships among the measures were all found to be very weak.

DISCUSSION AND CONCLUSIONS

Let us now return to the research question we posed at the outset: Are common native language tests, used to identify many ELL children as hav-

Table 3. Frequencies of Students' Spanish Language Proficiency Levels by Language Test

Proficiency Level	*LAS–O Español*				*IPT Spanish*			
	All Students		*Subgroup with Natural Language Samples*		*All Students*		*Subgroup with Natural Language Samples*	
	Frequency	*Percent*	*Frequency*	*Percent*	*Frequency*	*Percent*	*Frequency*	*Percent*
Fluent Spanish speaker	42	26%	36	27%	17	10%	13	9%
Limited Spanish speaker	57	35%	47	36%	136	78%	111	80%
Non-Spanish speaker	62	39%	48	37%	21	12%	14	10%
Totals	161	100%	131	100%	174	100%	138	100%

Table 4. Rate of Error in Morphology and Syntax in the Spanish Natural Language Sample by Specified Range and Language Test

Percent Error	LAS–O Español				IPT Spanish			
	Morphology		Syntax		Morphology		Syntax	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
5% or less	122	93%	131	100%	129	94%	139	100%
6%–10%	5	4%	0	0%	6	4%	0	0%
12%–13%	2	2%	0	0%	2	1%	0	0%
17%	2	2%	0	0%	2	1%	0	0%
Totals	131	100%	131	100%	139	100%	139	100%

Table 5. Rate of Error in Morphology and Syntax in the Spanish Natural Language Sample for All Study Participants with Natural Language Samples

	Morphology		Syntax	
Percent Error	Frequency	Percent	Frequency	Percent
5% or less	135	93%	145	100%
6%–10%	6	4%	0	0%
12%–13%	2	1%	0	0%
17%	2	1%	0	0%
Total	145	100%	145	100%

ing limited ability in the language of their own speech community, accurate measures of these children's true language abilities?

We presented data showing that children in our sample were classified as fluent, limited, and nonspeakers of Spanish, the language of their own speech communities, on both the LAS–O Español and the IPT Spanish, two of the most widely adopted native language proficiency tests (Mahoney & MacSwan, 2005). Although the LAS–O Español classified approximately three quarters of children as less-than-fluent speakers of their L1 (that is, limited or nonspeakers), the IPT Spanish so classified approximately 90% of

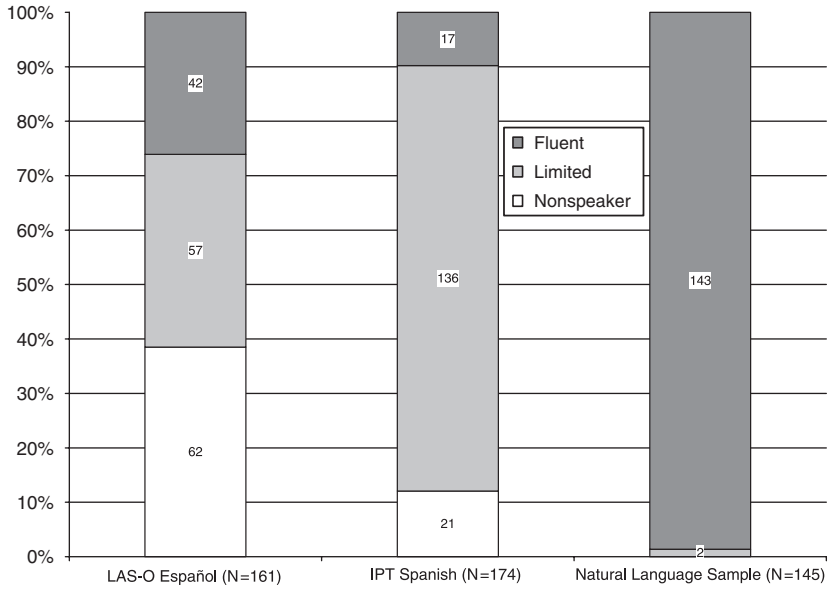

Figure 1. Proportion of Students in Each Proficiency Category for the LAS–O Español, IPT Spanish, and the Natural Language Sample

Table 6. Pearson Correlations Among the LAS–O Español, the IPT Spanish, and Morphological Error Rate (Inverted)

		Morphological Error Rate (Inverted)	LAS–O Español Proficiency Level	IPT Spanish Proficiency Level
Morphological Error Rate (Inverted)	Pearson Correlation	1.00	.279**	.235**
	Sig. (2-tailed)		.001	.005
	N	145	131	139
LAS–O Español Proficiency Level	Pearson Correlation	.279**	1.00	.092
	Sig. (2-tailed)	.001		.246
	N	131	161	161
IPT Spanish Proficiency Level	Pearson Correlation	.235**	.092	1.00
	Sig. (2-tailed)	.005	.246	
	N	139	161	174

**Correlation is significant at the 0.01 level (two-tailed).

the children we tested (see Table 3). These results are extraordinary in light of research conducted over the last half century on language acquisition, which has shown all normal children to achieve linguistically and to do so effortlessly and in the absence of instruction (Pinker, 1994). The divergence of these test results with research findings raises doubts regarding the construct validity of the test instruments, leading us to question the tests' theoretical foundations.

Indeed, neither test is constructed with an appropriate and theoretically defensible conception of language proficiency. The probabilistic approach, offered as the theory of language proficiency underlying the LAS–O Español, is introduced by the test developers with no prior status among linguists or other language researchers (De Avila & Duncan, 1989). The approach "assumes a linear relationship between linguistic proficiency and academic achievement" (p. 7). In other words, language proficiency is understood to be higher among speakers with higher academic achievement and lower among speakers with lower academic achievement, precisely the assertion of the prescriptivists who defined "better speech" in terms of those characteristics that the unschooled lacked. The prescriptivist underpinnings of the LAS–O Español are well represented in the test itself—as one sees in the scoring rubric for the story-retelling task—assigned 50% weight on the basis of its "theoretical and empirical importance" (De Avila & Duncan, 1996, p. 23). Here, a child's response is assigned a higher score because it increasingly approximates that of "a proficient speaker of *standard* [emphasis added] Spanish" (De Avila & Duncan, 1996, p. 5). The authors present no actual analysis of language proficiency as a

mental or social construct, and no attempt is made to relate the test or its underlying theory to linguistics, language acquisition, or other language-related research. (For a critical discussion of the Pre-LAS Español, a related instrument, see MacSwan, Rolstad, & Glass, 2002.)

Although the developers of the IPT Spanish make reference to published research on language acquisition, the literature is not properly represented, and theoretically divergent notions are combined in a surprising and unorthodox fashion. For instance, after describing classical "stages" of language acquisition associated with infants and toddlers, the IPT Spanish developers incorrectly indicate that these stages of acquisition continue into the elementary school years. The developers then discuss "constructs of language proficiency that exist in less visible form," such as CALP and Bloom's Taxonomy, and coin the term *higher order language skills,* which they define as "syntax, semantics, pragmatics" (Amori & Dalton, 1996, p. 3)—aspects of language that Cummins (1981) would be more likely to include in his "species minimum," or BICS. The authors do not offer an operational definition of these "less visible" forms of language proficiency; however, both CALP and Bloom's Taxonomy are usually defined to reflect language used in academic contexts so that the language of the educated classes is again privileged as developmentally superior. Indeed, just like Bereiter and colleagues, who required so-called proficient speakers to answer in complete sentences, the IPT Spanish requires children to provide answers complete with subject and predicate on several items—even in Spanish, a language that does not require overt subjects (see MacSwan, 2005, for further discussion).

Because we wished to compare the results of the LAS–O Español and IPT Spanish with a better measure of children's language ability, we additionally presented an analysis of coded language samples. The coding system attended to very narrow and specific details of children's grammatical knowledge and focused on functional categories associated with overt morphology. The focus on children's knowledge of morphology was justified in terms of current theories of the architecture of the human language faculty (Chomsky, 1995) and a long research tradition in child language acquisition and language disorders. The analysis of the natural language samples is tied to a rich theoretical and empirical tradition and involves highly detailed morpheme-by-morpheme analysis of a child's language. It involves a relatively natural use of language and respects children's home linguistic communities as the proper models of linguistic correctness.

A crucial component of our argument relies on the notion that the natural language sample is a much better indication of a child's true language ability than is either the LAS Español or the IPT Spanish. We believe that the theoretical foundation of the natural language sample and its close ties to a rich empirical tradition in research on child language acquisition and

disabilities contrast strongly with the thin and, in important respects, entirely absent theoretical foundations of the LAS Español or the IPT Spanish. Based on these considerations, we believe the natural language sample to be far superior a measure to either test.

Results of the analysis of these coded data, presented in Tables 3–5, showed about 93% of all participants to have a morphological error rate of 5% or less, and 97% of the study participants to have an error rate of 10% or less. Linguists and child language acquisition researchers assume that normal mature speakers will evidence some degree of error due to such factors as slips of the tongue and fatigue. Errors of this nature are termed "performance error," believed to result from the failure of the linguistic performance system to execute grammatical instructions due to the interference of nonlinguistic factors. These are errors of the sort each of us makes every day, errors that we often recognize ourselves as inconsistent with our knowledge of language immediately after producing them. Researchers generally estimate the range of normal error rate in typically developing mature speakers to be about 10% or less (Brown, 1973; Goodluck, 1991; Reilly, Marchman, & Bates, 1998), while the morphological error rate among language-impaired children tends to be considerably higher (Bedore & Leonard, 2005; Curtiss & Schaeffer, 1997, 2004; Leonard, 1997). Hence, we conclude that all but 4 of the 145 study participants were well within the range of fluent Spanish speaker, as determined by the natural language sample. (We will return to the 4 children with higher error rates directly.)

The proportion of children identified as fluent speakers of Spanish on the natural language measure differs considerably from what we have seen for the LAS–O Español or IPT Spanish, leading us to the conclusion that the latter instruments generate a plethora of false negatives. The correlation coefficients displayed in Table 6 evidence very weak relationships between morphological error rate and both the LAS–O Español and IPT Spanish. Furthermore, the LAS–O Español and IPT Spanish are shown to have a very low correlation, an indication that the two instruments lack concurrent validity even when compared with each other (Campbell, 1960).

The evidence considered here supports the conclusion that results of common native language tests such as the LAS–O Español and the IPT Spanish do not correctly identify the true native language abilities of ELLs. Although the natural language sample demonstrates the overwhelming majoring of children in the study to be well within the range of expected error rates, both the LAS–O Español and the IPT Spanish identified a majority of children as limited or nonspeakers of their native language.

Four of the cases in Table 5 show somewhat high morphological error rate. We suggested in our earlier discussion that normally developing linguistic minority children acquire the language of their speech community effortlessly and without instruction, just as majority language children do.

However, like majority language children, authentic cases of language-related disabilities surely exist among ELLs as well. Indeed, Leonard (1997) suggested that as many as 5% of children have a form of specific language impairment (SLI). If we suspect that these 4 children may have a language-related disability, then the proportion of children in our sample falls within the expected range of language disabilities in any population of children, majority or minority.

Indeed, the children in our sample whose error rate is somewhat high, especially the 2 children with error rates at 17%, may have such a disability. These children made frequent errors in gender agreement (*la perro,* "the dog"; *la niño,* "the child"; *el rana,* "the frog") and verb tense selection, and left out articles in obligatory contexts. If these children are language impaired, we would expect similar errors in their English. Another possibility, of course, is that the Spanish-background children with unexpectedly high error rates may be heritage speakers of Spanish who have only partially acquired Spanish but have genuine native language ability in English.

Further study would be needed to appropriately identify the specific language learning situation of the 4 children whose morphological error rates were unusually high; however, the data presented here are consistent with the view that developing children typically acquire the language of their speech community, and there is no reason to expect such children to arrive at school with limited ability in their native language any more than one would expect majority language children to do so.

We noted at the outset that Artiles and colleagues (2005) had reported that ELLs identified by districts—presumably using tests like the LAS–O Español and IPT Spanish—as having limited ability in both their L1 and L2 showed the highest rates of identification in special education categories. In light of this finding and the conclusions reached above regarding the validity of select Spanish-language tests, we offer a theory that points to institutional mechanisms to shed light on aspects of the problem of overrepresentation of ELLs, at least with respect to this specific subgroup of students. Spanish-background ELLs are more likely than any other student to undergo native language assessment because numerous states encourage districts to perform native language assessments, and tests of Spanish are readily available commercially. Because teachers might reasonably refer students identified as having low ability in both languages for special education testing, where some test or another is likely to qualify a referred student, we can expect Spanish-background ELL students to be overrepresented in special education categories. Because the Spanish-language test results are dramatically inaccurate, misidentifying a majority of ELLs as limited in their L1, these children are very likely to be inappropriately placed in special education programs in large numbers.

When the National Research Council panel convened to study the problem of overrepresentation, it asked whether overrepresentation of special education placements among ELLs and other minority students was due to "biological and social/contextual contributors to early development," whether "the school experience itself" might somehow be responsible, or perhaps a combination of both (Donovan & Cross, 2002, pp. 357–359). In the case of many ELLs, it appears that the school experience itself may be responsible, but perhaps not so much at the classroom level as at the level of state and district language testing policy. Overrepresentation appears to result in important respects from institutional factors that cannot reasonably be construed as limitations or language-related disabilities inherent in the children themselves.

We suggest here, as elsewhere (MacSwan & Rolstad, 2003; MacSwan et al., 2002), that the practice of routinely assessing the oral native language ability of minority students be abandoned. However, ELLs, like all students, should be referred for appropriate assessments when a genuine learning disability is suspected. If a language disability is suspected, the diagnosis must be carried out in an ELL's native language to avoid erroneous diagnosis of second-language speech as a disability (Paradis, 2005). Moreover, we stress an important distinction between *assessing a language* and *assessing in a language*. To determine a student's ability on nonlinguistic constructs and literacy, assessments of various kinds can and should be given in the language that the student is most comfortable using. Assessment for special education eligibility in particular should be carried out in an ELL's home language.

Although natural-language sampling of the type used in our study is a useful method for assessing language ability, the level of analysis used by Curtiss and colleagues (2005a, 2005b) and adapted here may be excessive and too time-consuming in many instances. Restrepo (1998) has stressed the usefulness of language sampling and of parent interviews as methods of appropriately and correctly identifying language impairments among bilingual children. Along with Restrepo, we urge practitioners to engage in careful analysis of actual speech samples, either immediately in an interview format or recorded for careful study, rather than relying on commercially available language tests. Such analyses must be carried out by linguistically trained and experienced practitioners with knowledge of the language under analysis.

For researchers, the challenge appears to be developing a conception of language proficiency that respects the rich resources of diverse communities. Repudiating prescriptivist dogma, which continues to thrive in the form of many language tests and theories of language proficiency, is a matter of first-order importance. Over 30 years ago, Labov (1970) issued a stark rebuke to Bereiter and his colleagues, which regrettably seems as appropriate today as it was then: "That educational psychology should be

strongly influenced by a theory so false to the facts of language is unfortunate; but that children should be the victims of this ignorance is intolerable" (p. 260). Let us hope that reforms in language-testing policies and improvements in assessment quality will be made to improve conditions for both ELLs and children with disabilities in the very near future.

The authors gratefully acknowledge research support from the Spencer Foundation. We also wish to thank the three anonymous reviewers and the editors of this special issue for providing us with careful and extensive comments and suggestions for revision. As always, we are indebted to our colleagues and students at Arizona State University for critical discussion of ideas presented here. In the end, of course, we alone are responsible for the content of the article.

Notes

1 The term *language proficiency* is used ambiguously by some authors to denote oral language ability and oral language ability plus literacy (Cummins, 1981; see also Macías, 1993); in the present discussion, we will use the term *proficiency* when summarizing the work of those who prefer this term, and we will use *language ability* to refer unambiguously to oral language ability.

2 The LAS–O Español data considered in this article are based on an earlier version of the test (De Avila & Duncan, 1994), not the most recent revision known as the LAS 2000 Español (De Avila & Duncan, 1998). Although we do not have data that speak specifically to the new version, its general similarity to the previous version of the test suggests that a study of that instrument would lead to conclusions similar to those reached here.

3 Details regarding the CHILDES Project are available online at http://childes.psy.cmu.edu/ and in MacWhinney (2000).

References

American Psychological Association, and National Council on Measurement in Education. (1985). *Standards for educational and psychological testing*. Washington, DC: Author.

Amori, B., & Dalton, E. F. (1996). *IPT I: Oral, grades K-6, Spanish, second edition. Technical manual*. Brea, CA: Ballard and Tighe.

Artiles, A. J., Rueda, R., Salazar, J. J., & Higareda, I. (2005). Within-group diversity in minority disproportionate representation: English language learners in urban school districts. *Exceptional Children*, *71*(3), 1–17.

Artiles, A. J., & Trent, S. C. (1994). Over-representation of minority students in special education. *Journal of Special Education*, *27*, 410–437.

Artiles, A. J., Trent, S. C., & Palmer, J. (2004). Culturally diverse students in special education: Legacies and prospects. In J. A. Banks & C. M. Banks (Eds.), *Handbook of research on multicultural education* (2nd ed., pp. 716–735). San Francisco: Jossey-Bass.

Baugh, A. C., & Cable, T. (1978). *A history of the English language* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Bedore, L. M., & Leonard, L. B. (2005). Verb inflections and noun phrase morphology in the spontaneous speech of Spanish-speaking children with specific language impairment. *Applied Psycholinguistics*, *26*, 195–225.

Bereiter, C., & Englemann, S. (1966). *Teaching disadvantaged children in the pre-school*. Englewood Cliffs, NJ: Prentice Hall.

Bereiter, C., Engelman, S., Osborn, J., & Reidford, P. A. (1966). An academically oriented preschool for culturally disadvantaged children. In F. M. Hechinger (Ed.), *Pre-school education today* (pp. 105–135). New York: Doubleday.

Bickerton, D. (1981). *The roots of language*. Ann Arbor, MI: Karoma Press.

Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press.

Campbell, D. T. (1960). Recommendations for APA test standards regarding construct, trait, and discriminant validity. *American Psychologist*, *15*, 546–553.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, N. (2004). Beyond explanatory adequacy. In A. Belletti (Ed.), *Structures and beyond: The cartography of syntactic structures*, (Vol. 3, pp. 104–131). Oxford, UK: Oxford University Press.

Crawford, J. (2004). *Educating English Learners: Language diversity in the classroom* (5th ed.). Los Angeles: Bilingual Educational Services.

Crystal, D. (1986). The prescriptive tradition. In D. Crystal (Ed.), *The Cambridge encyclopedia of language* (pp. 2–5). Cambridge, UK: Cambridge University Press.

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, *49*, 221–251.

Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. In *Schooling and language minority students: A theoretical framework* (1st ed.) pp. 3–49. Sacramento: Office of Bilingual Bicultural Education, California State Department of Education.

Cummins, J. (2000). *Language, power and pedagogy: Bilingual children in the crossfire*. Clevedon, UK: Multilingual Matters.

Curtiss, S., MacSwan, J., Schaeffer, J., Kural, M., & Sano, T. (2004a). GCS: A grammatical coding system for natural language data. *Behavior Research Methods, Instruments, and Computers*, *36*(3), 459–480.

Curtiss, S., MacSwan, J., Schaeffer, J., Kural, M., & Sano, T. (2004b). *GCS: Grammatical Coding System manual*. Psychonomic Society Archive of Norms, Stimuli, and Data. Retrieved January 4, 2006, from http://www.psychonomic.org/archive/

Curtiss, S., & Schaeffer, J. (1997). Syntactic development in children with hemispherectomy: The INFL-system. In E. Hughes, M. Hughes, & A. Greenhill (Eds.), *Proceedings of the 21st Annual Boston University Conference on Language Development* (pp. 103–114). Somerville, MA: Cascadilla Press.

De Avila, E. A., & Duncan, S. E. (1989). *LAS: Language Assessment Scales–Oral: Technical report, English, Forms 1C, 1D, 2C, 2D; Spanish, Forms 1B, 2B*. Monterey, CA: CTB/McGraw-Hill.

De Avila, E. A., & Duncan, S. E. (1994). *LAS: Language Assessment Scales–Oral: Administration manual, Español, Form 1B*. Monterey, CA: CTB/McGraw-Hill.

De Avila, E. A., & Duncan, S. E. (1996). *LAS: Language Assessment Scales–Oral: Scoring and interpretation manual, Spanish, Level 1B, Grades 1-6*. Monterey, CA: CTB/McGraw-Hill.

De Avila, E. A., & Duncan, S. E. (1998). *The LAS 2000 Español*. Monterey, CA: CTB/McGraw-Hill.

Dittmar, N. (1976). *A critical survey of sociolinguistics: Theory and application* (P. Sand, P. A. M. Seuren & K.Whiteley Trans.). New York: St. Martin's Press.

Donovan, S., & Cross, C. (Eds.) (2002). *Minority students in special and gifted education*. Washington, DC: National Academy Press.

Gleitman, L., & Landau, B. (1994). *The acquisition of the lexicon*. Cambridge, MA: MIT Press.

Goodluck, H. (1991). *Language acquisition: A linguistic introduction*. Cambridge, UK: Blackwell.

Kindler, A. (2002). *Summary report of the survey of the states' limited English proficient students and available educational programs and services, 2000–2001*. Washington, DC: National Clearinghouse for Bilingual Education.

Labov, W. (1970). The logic of non-standard English. In F. Williams (Ed.), *Language and poverty.* (pp. 225–261). Chicago: Rand McNally.

Labov, W. (1975). Empirical foundations of linguistic theory. In R. Austerlitz (Ed.), *The scope of American linguistics* (pp. 77–133). Lisse, Netherlands: de Ridder.

Lau v. Nichols, 414 U.S., 563, 566 (1974).

Leonard, L. B. (1997). *Children with specific language impairment*. Cambridge, MA: MIT Press.

Lipski, J. (1994). *Latin American Spanish*. New York: Longman.

Macías, R. (1993). Language and ethnic classification of language minorities: Chicano and Latino students in the 1990s. *Hispanic Journal of Behavioral Sciences*, *15*, 230–257.

MacSwan, J. (2000). The threshold hypothesis, semilingualism, and other contributions to a deficit view of linguistic minorities. *Hispanic Journal of Behavioral Sciences*, *22*, 3–45.

MacSwan, J. (2005). The "non-non" crisis and academic bias in native language assessment of linguistic minorities. In J. Cohen, K. McAlister, K. Rolstad, & J. MacSwan (Eds.), *ISB4: Proceedings of the 4th International Symposium on Bilingualism*. Somerville, MA: Cascadilla Press.

MacSwan, J., & Rolstad, K. (2003). Linguistic diversity, schooling, and social class: Rethinking our conception of language proficiency in language minority education. In C. B. Paulston & R. Tucker (Eds.), *Sociolinguistics: The essential readings* (pp. 329–340). Oxford, UK: Blackwell.

MacSwan, J., Rolstad, K., & Glass, G. V. (2002). Do some school-age children have no language? Some problems of construct validity in the Pre-LAS Español. *Bilingual Research Journal*, *26*, 213–238.

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk* (3rd ed.). Hillsdale, NJ: Erlbaum.

Mahoney, K. S., & MacSwan, J. (2005). Re-examining identification and reclassification of English Language Learners: A critical discussion of select state practices. *Bilingual Research Journal*, *29*, 31–42.

Mayer, M. (1969). *Frog, where are you?* New York: Dial Books for Young Readers.

Milroy, J., & Milroy, L. (1999). *Authority in language: Investigating Standard English*. New York: Routledge.

Newmeyer, F. J. (1986). *The politics of linguistics*. Chicago: University of Chicago Press.

Paradis, J. (2005). Grammatical morphology in children learning English as a second language: Implications of similarities with specific language impairment. *Language, Speech, and Hearing Services in Schools*, *36*, 172–187.

Paulston, C. B. (1983). *Swedish research and debate about bilingualism*. Stockholm: National Swedish Board of Education.

Pinker, S. (1994). *The language instinct: How the mind creates languages*. New York: William Morrow.

Postal, P. (1972). Language differences and prescriptivism. In R. Abrahams & R. Troike (Eds.), *Language and cultural diversity in American education* (pp. 112–116). Englewood Cliffs, NJ: Prentice Hall.

Reilly, J. S., Marchman, V., & Bates, E. A. (1998). Narratives in children with early focal brain damage. *Brain and Language*, *61*, 335–337.

Restrepo, M. A. (1998). Identifiers of predominantly Spanish-speaking children with language impairment. *Journal of Speech, Language, and Hearing Research*, *41*, 1398–1411.

Schatzmann, L., & Strauss, A. (1955). Social class and modes of communication. *American Journal of Sociology*, *60*, 329–338.

Tager-Flusberg, H. (1997). Putting words together: Morphology and syntax in the preschool years. In J. Berko-Gleason (Ed.), *The development of language* (pp. 159–209). Boston: Allyn and Bacon.

Valadez, C., MacSwan, J., & Martínez, C. (2002). Toward a new view of low achieving bilinguals: A study of linguistic competence in designated "semilinguals." *Bilingual Review*, *25*, 238–248.

Valencia, R. (Ed.). (1997). *The evolution of deficit thinking: Educational thought and practice*. London: Falmer Press.

Williams, C., Ballard, W., Tighe, P., Dalton, E., & Amori, B. (1998). *IPT I: Oral, grades K-6, Spanish* (2nd ed.). Brea, CA: Ballard and Tighe.

JEFF MACSWAN is an associate professor of language and literacy in the College of Education at Arizona State University. His research focuses on the linguistic study of bilingualism (codeswitching and language contact) and on language minority education. Recent examples of his work appear in *Bilingualism: Language and Cognition, Bilingual Research Journal, Hispanic Journal of Behavioral Sciences,* and Bhatia and Ritchie's *Handbook of Bilingualism* (Blackwell, 2004). MacSwan has served as a visiting scholar in the MIT Linguistics Department and the Center for the Study of Multilingualism at the University of Hamburg. In 2003, he was selected as a National Academy of Education/Spencer Postdoctoral Fellow.

KELLIE ROLSTAD is an associate professor of language and literacy and early childhood education at Arizona State University. Her research interests include early childhood bilingualism, language diversity, bilingual education programs and theory, and two-way bilingual immersion. Her work has appeared in the *Bilingual Research Journal, Bilingual Review, Teachers College Record*, and *Educational Policy*, among others, and she has served as a visiting scholar in the Graduate School of Education at both Harvard University and UCLA.

JEFF MACSWAN and KELLIE ROLSTAD have previously coauthored work appearing in the *Bilingual Research Journal, Hispanic Journal of Behavioral Sciences, Sozialwissenschaftliche Literatur Rundschau*, and Paulston and Tucker's *Sociolinguistics: The Essential Readings* (Blackwell, 2003). Most recently, they have coedited (with James Cohen and Kara McAlister) *ISB4*: *Proceedings of the 4th International Symposium on Bilingualism* (5 volumes; Cascadilla, 2005).