

Do Some School-Age Children Have No Language? Some Problems of Construct Validity in the Pre-LAS Español

Jeff MacSwan, Kellie Rolstad, and Gene V. Glass
Arizona State University

Abstract

This article reports the existence of a large group of students identified as “non-nons,” Spanish-background school-age children living in the United States who are reported to be non-verbal in both English and Spanish, and brings the validity of the “non-non” construct into question. In particular, the authors assess the validity of the Pre-Language Assessment Scales Español (Pre-LAS Español), an oral language assessment that purports to measure oral native language ability in children ages 4 to 6. A dataset of 38,887 students who took the Pre-LAS Español in a large urban school district in 1997 is examined, and questions are raised from internal and external evidence regarding the test’s validity. The authors conclude that there are serious concerns regarding the validity of the test, characterize the “non-non” label as an artifact of poor assessment, and recommend that districts and states reconsider current policy requiring or recommending routine oral native language assessment of language minority students.

Introduction

A common belief among teachers, policy makers, and education researchers is that some school-age children have no language. The *LA Times* reported that 6,800 children in the Los Angeles Unified School District are classified as “non-nons” and said to be “nonverbal in both English and their native language” (Pyle, 1996). Children are so classified as a result of native-language assessment instruments. One such test that is commonly used is the Language Assessment Scales-Español (LAS-Español) (DeAvila

& Duncan, 1990, 1994, 1986a, 1986b), which classifies Spanish-speaking children into the categories “fluent Spanish speaker,” “limited Spanish speaker,” and “non-Spanish speaker.”

In this article, we explore the construct validity of the Pre-LAS Español, the Spanish-language version of the LAS intended for use with children ages 4 to 6, by considering both external and internal evidence bearing on its validity.¹ Analyses reported here are based on a dataset of 38,887 subjects who took the Pre-LAS Español in a large urban school district in California in 1997. By investigating the validity of the test, we further question the legitimacy of the “non-non” label that the test applies to many Spanish-background English learners.

We begin with a discussion of some possible influences in bilingual education that may underlie current policies in many states that require or recommend assessment of oral native language ability in school-age children, and then turn to a specific analysis of the validity of one such assessment, the Pre-LAS Español. We conclude by suggesting that states should reconsider testing policies that routinely assess a language minority child’s oral native language ability.

Native Language Assessment Policy

Analysis of program alternatives for English learners has shown that academic instruction in a child’s native language is often associated with outcomes superior to those resulting from all-English instruction (Willig, 1985; Ramirez, Pasta, Yuen, Billings, & Ramey, 1991; August & Hakuta, 1998). At least two distinct explanations have been advanced to account for the advantage of native language instruction for minority students. A traditional view is that native language instruction is especially beneficial to language minority children because it makes academic content comprehensible during the years it takes them to learn English well enough to understand all-English instruction (Krashen, 1996).

Another view, which contrasts sharply with this, attributes achievement differences in language minority children to presumed ability differences in children’s native oral language. Although this view may be appealing to some, it is extremely difficult to distinguish from classical prescriptivism, in which differences between the language of the educated classes and the language of the unschooled are described in terms of levels of ability, degrees of complexity, or depth of vocabulary. We find such descriptions of children’s oral native language to be highly objectionable for empirical and theoretical reasons (for review, see Edelsky, Hudelson, Flores, Barkin, Altweger, & Jilbert, 1983; Martin-Jones & Romaine, 1986; Wiley, 1996; MacSwan, 2000; MacSwan & Rolstad, in press). More narrowly, in bilingual education, the notion of native language ability differences among minority children has been widely

used and probably owes its popularity to Cummins' (1976, 1979, 1981, 2000a, 2000b) influential Threshold Hypothesis and BICS/CALP framework. We briefly outline these ideas below with an aim to illuminate current language assessment practices in schools.

Cummins' Threshold Hypothesis originally suggested that "negative cognitive and academic effects . . . result from low levels of competence in both languages" (1979, p. 230), a view that is also known as "semilingualism" (Cummins, 1979) or "limited bilingualism" (Cummins, 1981). The original concern around which the Threshold Hypothesis was developed was a conflict in research findings on the cognitive benefits of bilingualism. Earlier studies had concluded that bilingualism adversely affects cognitive and scholastic progress, while more recent work showed "positive cognitive consequences" for bilinguals. Cummins (1976) pointed out that the studies that found a negative effect were associated with linguistic minorities, where the minority language was being replaced in some sense by the socially dominant one, while the studies that found a positive effect were associated with "additive bilingualism," a situation in which majority-language children acquire a second language. Cummins (1976) hypothesized that linguistic minorities undergo native language loss and that "the level of linguistic competence attained by a bilingual child may mediate the effects of his bilingual learning experiences on cognitive growth" (p. 24). In other words, he attempted to explain the reports of negative effects of bilingualism on "cognitive and scholastic progress" by proposing that the subject population had a low level of linguistic proficiency in its first language.² By contrast, children in the "additive" bilingual programs benefited from continued support of their first language in and out of school. As Cummins put it: "Negative cognitive and academic effects are hypothesized to result from low levels of competence in both languages or what Scandinavian researchers (e.g., Hansegård, 196[8]; Skutnabb-Kangas & Toukomaa, 1976) have termed 'semilingualism' or 'double semilingualism'" (1979, p. 230).

Despite their popularity, the Threshold Hypothesis and the embedded notion of semilingualism (or "limited bilingualism") were advanced in the absence of relevant evidence. In an extensive review of the Scandinavian literature on semilingualism, for instance, Paulston (1983) concluded that "there is no empirical evidence to support the existence of such a language development hiatus as [semilingualism]" (p. 42). MacSwan (2000) reviewed four purported sources of evidence for semilingualism—studies of language shift, school performance, linguistic structure, and language variation—and also concluded that the evidence supporting the semilingualism construct is either spurious or irrelevant to the basic claim. Conversely, considerable evidence has shown that children alleged to be semilingual or "subtractive bilinguals" actually do not differ from native speakers in terms of their linguistic competence (Cummins & Miramontes, 1989; Hakuta & D'Andrea, 1992; Valadez, MacSwan, & Martínez, 2001).

Similar criticisms have been made (Edelsky et al., 1983; Genesee, 1984; Spolsky, 1984; Troike, 1984; Martin-Jones & Romaine, 1986; Wiley, 1996; Petrovic & Olmstead, 2001; MacSwan & Rolstad, in press) regarding Cummins' distinction between basic interpersonal communication skills (BICS) and cognitive-academic language proficiency (CALP). Cummins (1980) hypothesized that, "there exists a reliable dimension of proficiency in a first language which is strongly related to cognitive skills and which can be empirically distinguished from interpersonal communication skills such as oral fluency, accent, and sociolinguistic competence" (p. 177).

Although there are clearly developmental aspects of a child's second language acquisition that persist well into the school years for many English learners, the theoretical and empirical justification for the BICS/CALP distinction is not persuasive. More specifically, in distinguishing the two kinds of language, Cummins asserts that CALP is characterized by an expanded range of vocabulary and complex grammatical structures (Cummins, 2000a, p. 63; Cummins, 2000b, pp. 35–36), an ability to make complex meanings explicit (Cummins, 2000a, p. 59), and greater demand on memory, analysis, and other cognitive processes (Cummins, 2000b, pp. 35–36). The common belief that academic language has special and enriched properties results from a long tradition of prescriptivist dogma, now propagated primarily in the academy—a tradition that has had the principal effect of justifying social inequalities in terms of "objectively assessed" deficiencies in language, culture, and behavior. However, considerable research has shown that there simply is no human language or language variety that does not have complex grammatical structures. Additionally, there is no language or language variety that does not possess the mechanisms to create new words as new situations arise or to make complex meanings explicit (Crystal, 1986; Milroy & Milroy, 1999).

Nonetheless, the practice of assessing children's oral native language proficiency has become widespread. In 1991, the Council of Chief State School Officers conducted a survey of methods used by local educational agencies to identify language minority children, and they reported that five states require, and four others recommend, that districts assess Spanish-speaking children's oral native language ability upon entry to school. Table 1 lists each state in which the education code requires or recommends native language assessment and also lists the number of children affected. In many other states, the decision to assess children's native language is made at the district level.

Table 1

States that Require or Recommend Routine Native Language Assessment for Limited English Proficient (LEP) Children

Required (as of 1991)	LEP population (as of 1997–1998)	Recommended (as of 1991)	LEP population (as of 1997–1998)
California	1,406,166	Illinois	136,186
Texas	509,282	Oklahoma	33,089
Arizona	112,522	Indiana	9,114
New Jersey	47,415	New Hampshire	1,748
Hawaii	12,869		
Total LEP children	2,088,254	Total LEP children	180,137

Note. Data reported by Council of Chief State School Officers (1991) and National Clearinghouse for Bilingual Education (2000).

It must be emphasized that the Pre-LAS Español and other native language tests used in the United States are intended to assess children's oral language ability in the native tongue, not literacy. To avoid labeling non-literates as semilingual or alingual, we should carefully distinguish language, an integral part of every person's identity, from literacy, a technological development of use to some but not all individuals and human societies (for discussion, see MacSwan, 2000; MacSwan & Rolstad, in press; Wiley, 1996).³

However, some test makers include items on their oral language instruments that assess aspects of language use that are specific to academic culture—and, in some cases, items or subparts that are not specifically related to language ability at all. Doing so in the context of oral native language assessment, and characterizing the results as an index of native language ability, enormously privileges the educated classes and recalls the classic critique of prescriptivism (Labov, 1970).

Surely there is a specific character to the way language is used in many academic contexts, and it is reasonable to suspect that the use of this academic register—or “Discourse,” as Gee (1996) prefers—will co-occur with success at school and mastery of academic content. A serious problem arises, however, when we construe this domain of language use as “more complex” or “developmentally superior” to the language children use natively at home or on the street, and consequently assess school language as indicative of “higher language proficiency.” The Pre-LAS Español, like many commercially available language assessments, purports to identify children's Spanish-speaking ability, categorizing them as “fluent Spanish speaker,” “limited Spanish speaker,” or “non-Spanish speaker.”

This way of conceptualizing language proficiency—in which academic registers are viewed as richer, more complex, and so on—not only lacks reasonable empirical and theoretical support (Martin-Jones & Romaine, 1986; Wiley, 1996; MacSwan, 2000; MacSwan & Rolstad, in press), but it also implies that entire cultures and communities that have rejected literacy and formal schooling have a “less complex” language than those in the “literate” world. The problem arises because culture-specific and language-specific characteristics are arbitrarily selected as defining characteristics of the abstract psychological construct of language ability—observed in all human communities, everywhere.

Below we consider the validity of the Pre-LAS Español in connection with these concerns. The six distinct parts of the test, which will be the focus of subsequent discussion, are described in Table 2.

Table 2

Description of the Subparts of the Pre-LAS Español
(Duncan & DeAvila, 1986a, 1986b)

Section	Name of Section	Nature of Task
Part 1	Tío Simón	Child is directed to act out commands as issued by test administrator. No cue pictures. 10 items, dichotomous scoring.
Part 2	La Casita	Child is asked to name items identified by the test administrator in a large drawing of a house. 10 items, dichotomous scoring.
Part 3	Dibujos y Frases	Child is asked to identify the picture that goes along with the phrase uttered in Spanish from a set of cues. 10 items, dichotomous scoring.
Part 4	Repetición de Frases	Child is asked to repeat sentences uttered by the test administrator. An item is wrong only if the specific underlined (targeted) portion is omitted or incorrect. 10 items, dichotomous scoring.
Part 5	Terminando Cuentitos	Child is asked to verbally finish incomplete sentences which the test administrator.
Part 6	Contando Historias	Child is told a story and asked to repeat it; the result is written down and scored as a language sample on a scale of 0-5. Two samples collected using cue cards.

Validity of the Pre-LAS Español

A psychological test such as the Pre-LAS Español is purported to measure a psychological trait—in this case, Spanish language proficiency. DeAvila and Duncan (1982) developed the LAS family of tests according to a view of language as consisting of four aspects: phonology (the sound system), the lexicon (vocabulary), syntax (principles that govern word order), and pragmatics (the use of language to satisfy particular goals) (cited in Guerrero & Del Vecchio, 1996).

Campbell (1960) pointed out that the construct validity of a test is demonstrated in part by its correlation with other known psychological variables. The test should correlate higher with variables to which it should have a theoretical relationship (convergent validity) than it does with theoretically distinct variables (divergent validity). A careful examination of the Pre-LAS Español will show that, while it may be reasonably expected that a high score on some parts reflects Spanish language proficiency, performance on other parts—which turn out to account for the greatest variance in Pre-LAS total scores—does not. These facts, we argue, raise questions about the construct validity of the Pre-LAS Español and, consequently, about the legitimacy of such designations as “semilingual” and “non-non.” As Cronbach and Meehl (1955) pointed out, the investigation of the construct validity of a test in which a theory (or “nomological network”) is used to derive a set of expectations about the test’s pattern of relationships can, at the same time, illuminate questions of the validity of both the test and the theory.

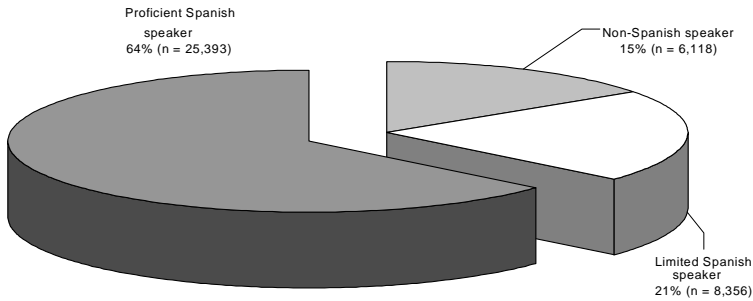
The construct validity of the test was explored in terms of both internal and external factors. Internal evidence is provided by means of intercorrelation analysis and common factor analysis, a multivariate statistical technique designed to reveal complex interrelationships among variables. External evidence is provided in an analysis of each subpart (task) on the Pre-LAS Español in terms of accepted views of language acquisition and related phenomena. We argue that the sections of the test that are overwhelmingly responsible for the “non-Spanish speaker” rating do not relate to oral language proficiency in Spanish.

The dataset for the present study consists of test scores from a group of 38,887 pre-kindergarten and kindergarten children who took the Pre-LAS Español in a large California school district in 1997.

Analysis of the Pre-LAS Español from Internal Evidence

Of the 38,887 children in our dataset, 15% ($n = 6,118$) were identified as non-Spanish speakers, 21% ($n = 8,356$) as limited Spanish speakers, and 64% ($n = 25,393$) as fluent (proficient) Spanish speakers. Thus, approximately one third of Spanish-speaking children in the dataset were identified as having limited or no ability in Spanish on the Pre-LAS Español, as illustrated in Figure 1.

Figure 1. Proportion of children identified as non-limited or fluent (proficiency) in Spanish in the Pre-LAS Español (n = 39,867).



The correlations of the Pre-LAS Español subtests with the total score on the test help explain how the total score on the test is functioning to classify speakers as proficient or not. These correlations appear in the bottom row of Table 3. For example, Part 1 correlates .282 with the Total Score, and Part 6 correlates .836 with the Total Score. These coefficients are known as “part-whole” correlations because the subtest being correlated with the total score is a constituent of that total itself; hence, the correlation is artificially inflated by the overlap (even errors of measurement would produce a positive part-whole correlation in the absence of any shared true variance whatsoever) and can give a misleading impression of the internal structure of the test battery. To correct for this overlap, correlations were recalculated by eliminating the subtest in question when correlating a subtest with the total score (e.g., Part 1 was correlated with the Total Score, minus Part 1, and so on *mutatis mutandis*). The results of this analysis appear in Table 4. (All reported correlation coefficients were significant at the 0.01 level, as would be expected with such large sample sizes.)

Table 3

Pearson Correlations of Parts 1-6 and Total Score on the Pre-LAS Español (N = 38,887)

Parts	Parts						Total score
	1	2	3	4	5	6	
1	1.000	.445	.173	.220	.201	.170	.282
2	.445	1.000	.317	.361	.327	.311	.524
3	.173	.317	1.000	.270	.253	.230	.425
4	.220	.361	.270	1.000	.375	.332	.558
5	.201	.327	.253	.375	1.000	.591	.744
6	.170	.311	.230	.323	.591	1.000	.836
Total Score	.282	.542	.425	.558	.744	.836	1.000

As expected, the removal of a Part Score from the Total Score reduces the apparent correlation between subtests and total battery score. What becomes even more apparent in Table 4 is that Parts 1–4 of the Pre-LAS Español play much less of a role in determining the Total Pre-LAS Español Score than Parts 5–6. In fact, Part 1 accounts for only about 6% of the variance in the total battery score once overlap is removed, whereas Part 6 accounts for nearly 60% of the total variance after removal of overlap. When Parts 5 and 6 are combined in the best linear combination for predicting the Total Score (with overlap), the observed multiple correlation coefficient equals .891, indicating that the Total Score on the Pre-LAS Español is almost perfectly predictable from the scores on Parts 5 and 6. (The best linear combination of Parts 1–4 accounts for slightly less than half of the Total Score variance.) Thus, Parts 5 and 6 of the Pre-LAS Español—which, we argue below, make up the most subjectively scored and theoretically ill-defined portions of the test—figure strongly in determining the Total Score, while the parts we view as somewhat better designed (Parts 1–3 and, perhaps, 4) play a much smaller role in accounting for variation in the Total Score.

Table 4

Correlations of Subtests With Total Score on the Pre-LAS Español With and Without Part-Whole Overlap

Correlation	Parts					
	1	2	3	4	5	6
Correlation With Total Score With Overlap	.282	.524	.425	.558	.744	.836
Correlation With Total Score Without Overlap	.251	.445	.350	.463	.577	.768

The six parts of the Pre-LAS Español were also subjected to a common factor analysis in an attempt to learn more about the internal structure of the battery of tests. Initial factor extraction was performed via Kaiser's Alpha factoring method, which in this instance seemed indicated by the fact that Alpha factor analysis is based on principles of psychometric internal consistency and produces results that are unaffected by test error variance (Glass, 1966). Two factors, corresponding to the latent roots 2.481 and .923 of the correlation matrix in Table 3, were extracted and subjected to oblique transformation via the direct oblimin solution. The latent roots, factor pattern and factor correlation matrix appear in Tables 5, 6, and 7.

Table 5

Initial Factor Solution for Pre-LAS Parts 1–6

Factor	Latent root	Cumulative % variance
1	2.481	41.357
2	.923	56.746
3	.860	71.082
4	.705	82.834
5	.624	93.236
6	.406	100.000

Table 6
Transformed Factor Pattern

Pre-LAS Español Subtest	Factor	
	1	2
Part 1	.309	.030
Part 2	.556	.030
Part 3	.567	-.077
Part 4	.532	.105
Part 5	.007	.812
Part 6	.023	.705

Inspection of the factor pattern reveals two quite distinct common factors among the six Pre-LAS Español subtests. The first is defined by Parts 1 through 4 and the second by Parts 5 and 6. The two factors correlate .674, as shown in Table 7. We hesitate to name these factors so as not to prejudice the arguments about validity that follow. Suffice it to say that the factor analysis results suggest that the two major portions of the Pre-LAS Español (Parts 1–4 and Parts 5–6) should be looked at quite carefully to determine why they might be functioning differentially.

Table 7
Transformed Correlation

Factor	1	2
1	1.000	.674
2	.674	1.000

Furthermore, the data show that 67.5% of “non-Spanish speakers” scored 80% or higher on Parts 1–3, and 51% of “non-Spanish speakers” scored 80% or higher on Parts 1–4. Furthermore, 20% of the “non-Spanish speakers” scoring 80% or higher on Parts 1–4 had given no response to Part 6 of the test despite their strong performance on Parts 1–4. It is, of course, impossible to do well on Parts 1–4 without knowing Spanish, so it follows that about 2/3 of all “nons” are false negatives—competent speakers of Spanish whom the test incorrectly labeled as “non-Spanish speakers.” This high proportion of false negatives raises serious questions about the construct validity of the Pre-LAS Español. We now turn to a detailed discussion of each of the subtests of the Pre-LAS Español.

Analysis of the Pre-LAS Español from External Evidence

In this section we explore the relationship between the subparts of the Pre-LAS Español and the body of empirical and theoretical work devoted to the study of language structure and acquisition. We will pay especially close attention to those sections of the test most responsible for the “non-” or “limited” category, namely, Parts 5 and 6.

As mentioned, the Pre-LAS Español (Duncan & DeAvila, 1986a, 1986b) consists of six parts. Part 1 is a receptive test of children’s ability to understand commands in Spanish; it has 10 items. In Part 2, the examiner asks the child to point to 10 specific objects inside a picture of a house. In Part 3 of the test, the child is asked to identify one of two pictures that corresponds to a spoken utterance given by the examiner (e.g., “*Hay dos animales*,” “There are two animals”). Subscales 1–3 of the Pre-LAS Español, in which each test item is scored as correct or incorrect, may reasonably be said to identify Spanish language proficiency.

In Part 4 of the test, the examiner reads 10 sentences that the child must repeat. An item is scored as incorrect if the underlined portion of the item, which is intended to target a particular grammatical construction, is omitted or “transformed.” Psycholinguists who use elicited imitation as a research tool caution that the pre-training sentences in such tasks should be of the same approximate length as the test sentences, that the test battery should begin only after the examinees show that they understand the task, that all sentences in the test battery be of nearly identical length, and that a whole test battery target a single grammatical structure while differing only in word choice (Lust, Flynn, & Foley, 1996). However, the Pre-LAS Español follows none of these cautions; thus, we do not know if the child’s incorrect responses result from failing to understand the (pragmatically anomalous) task, unexpected processing load, lexical bias, or perhaps limitations on short-term memory (Smith & van Kleeck, 1986). Hence, while it is very likely that children who do well on this part of the test are Spanish proficient, we do not know whether children who perform poorly on this section know Spanish or not. Thus, by virtue of the design of the task, we might expect that Part 4 would yield some number of false negatives.

As was shown in Table 4, Parts 5 and 6 account for an extremely large proportion of the variance of results, with Part 5 correlating .744 and Part 6 correlating .836 with the total score (.577 and .768, respectively, with the overlapping parts removed). Because of the importance of these sections in determining the overall score, very careful attention should be given to their design and underlying theoretical assumptions.

Part 5 of the test requires the child to finish five incomplete sentences. Like Part 4, this task invokes a linguistically unnatural situation that will be strange to many children. More importantly, however, the examiner has a great deal of discretion in awarding a score from zero to three for each item. Table 8 shows the scoring rubric for Part 5 of the Pre-LAS Español.

Table 8

Description of Rating Levels for Terminado Cuentitos (Part V)
(Duncan & DeAvila, 1986a, p. 3)

Level	Description
0	No response or responds in another language.
1	Clauses are awkward and/or unintelligible. Child has difficulty combining words with the facility of a proficient native speaker.
2	Response is appropriate Spanish for the age group even though it may contain "errors" in syntax commonly made by proficient speakers (e.g., children of this age are still acquiring a command of irregular verbs and may use inappropriate tense of person. Clause may be grammatically correct in and of itself, but inappropriate to the situation. There may be developmental or dialectal variations in pronunciation <i>Me voy a la escuela, me ecuelto</i> .
3	No syntactical errors; clauses, tense and person are contextually appropriate. As in a rating 2, there may be developmental or dialectal variation in pronunciation. If the response can be corrected it is not a Level 3.

The criteria for a score of zero are clear and straightforward: A child who does not respond, or who responds in another language, will receive a zero for each item of Part 5. Moreover, an appendix to the scoring manual (Duncan & DeAvila, 1986a, p. 14) lists another example of a zero for items on part 5: "*No sé*" ("I don't know").

However, consider the following test item that a child, age 4 to 6, must verbally complete:

"*Antes de vestirme* _____."

[Before getting myself dressed _____.]

A young child might easily be confused by this task, which requires significant meta-linguistic awareness, and may therefore simply fail to respond or say she doesn't know—not because she lacks knowledge of the relevant grammatical construction, but because many children in this age range may not be acquainted with sentence-completion tasks.

Moreover, because young children tend to be more sensitive to language content and its truth value than to language structure (Wexler & Culicover, 1980; Gordon, 1996), the prompt could easily be misunderstood as a request that the child speculate about what the test administrator may do before

getting dressed. Such responses would undoubtedly receive low scores. Furthermore, given the personal nature of this matter and the perceived differential of social status between the child and the adult test administrator, a child may wisely choose to say “*No sé*” (“I don’t know”) or nothing at all.

Here and elsewhere, the Pre-LAS Español directs the test administrator to render a score of zero for no response. This is an inference from the absence of evidence (the proof of a negative), a classic fallacy that one would expect to be avoided in any psychological test. Put differently, the conclusion that a child does not know Spanish, or knows it poorly, simply does not follow from the fact that no response was given to a particular test item.

According to Table 8, a score of 1 may be assigned if “clauses are awkward or unintelligible.” Unfortunately, no clear definition of “awkward” is offered, leaving the matter up to the subjective judgment of the examiner. Indeed, “awkwardness” is a literary or stylistic notion and will certainly vary considerably across communities of speakers. Thus, one sees that on this portion of the test, children may score a 1 for an item even if their response is linguistically well-formed but does not happen to appeal to the examiner on stylistic grounds.

In addition, because there is no explicit definition of “awkward” offered in the test materials, we might expect that the examiner’s response may be conditioned by dialectal differences between the examiner and the child. For instance, due to contact with English and Quechua, some varieties of Spanish have developed clitic doubling (*La vi a la mujer*, “I saw the woman”), a form also found in Greek and Romanian (Luján & Parodi, 1996).⁴ An examiner unfamiliar with this and other instances of syntactic variation might give a child a low score here due to perceived “awkwardness.”

An additional problem is that we have no idea why an utterance might be “unintelligible.” Is it inaudible? Is it so different in pronunciation from the examiner’s dialect that it cannot easily be understood? Is the child shy or nervous? No examples of “unintelligible” responses are given, nor is the term given any elaboration.

Also, note that both Level 2 and Level 3 scores express tolerance for phonological differences due to regional dialect, but it is emphasized (by italicizing “pronunciation”) that the variation must be attributable to pronunciation and not syntax. This distinction is arbitrary and unjustifiable. Consider, for instance, the use of clitic doubling in some varieties of Spanish, discussed above, or the phenomenon of *laísmo*, used in parts of Northern Spain, in which dative *le* suppletes to *la* (“*La dí el regalo*,” “I gave her the gift”). In Ecuador, many speakers of Spanish use *le* in place of *la/lo*, as in “*Yo le amo*” [I love her/him] (See Lipski, 1994 for discussion). Furthermore, Morales (1986) reports that many Caribbean varieties of Spanish make extensive use of overt subject pronouns in contexts where other Spanish dialects do not tolerate them for pragmatic reasons. However, the rubric is

written in such a way as to imply that regional variation in syntax is not tolerated at Level 3. There is of course no conceivable justification for this arbitrary provision.

Indeed, because the test does not tolerate syntactic variation due to dialectal differences, it may be said to endorse prescriptivism, the view that speakers of regional vernacular dialects have less proficiency in their language than those who speak like the educated classes. No evidence has been advanced to support this doctrine, here or elsewhere. Like elite language varieties, minority varieties are grammatically rich, systematic, and expressive (Crystal, 1986; Newmeyer, 1986; Pinker, 1994).

Another salient distinction between scores 2 and 3 in Table 8 is the “appropriateness” of the response to the situation. A Level 2 response may be “inappropriate to the situation,” but a Level 3 response yields “clauses, tense, and person” that are “contextually appropriate.” However, there is no continuity of “situation” from one sentence to the next in Part 5. Each item creates a new situation and one that is only partially spelled out. An appropriate situation is one that is consistent with the examiner’s background assumptions about a particular item. If the examinee makes the wrong guess about what the examiner is thinking, a low score may be assigned. Consider, for instance, an actual example of a set of responses from Gabriela, a 5-year-old girl who attends school in an urban California school district and speaks Spanish at home. Gabriela’s responses to Part 5 of the Pre-LAS Español are italicized below as transcribed by the examiner, with translations given in brackets:

1. “Si me levanto temprano *como*.” (3)
[If I get up early *I eat*.]
2. “Los niños tenían hambre así que *hizieron* [sic] *sopa*.” (3)
[The children were hungry so *they made soup*.]
3. “Fuimos a la fiesta y luego *compramos un pastel*.” (3)
[We went to a party and then *we bought a cake*.]
4. “Antes de vestirme *fui a una fiesta*.” (2)
[Before getting dressed *I went to a party*.]
5. “Después de jugar un rato *me siento*.” (3)
[After playing a while *I sit down*.]

(Note that *hicieron* is misspelled as “*hizieron*” in Item 2 by the examiner, not by Gabriela.) Gabriela’s total score for this part of the test was 14/15. She was marked down for her response in Item 4 presumably because one usually gets dressed up before going to a party. The same logic should lower Gabriela’s score in Item 3, since one would usually buy a cake before going to a party. Notice that the response in Item 4 is primed by the prompt in Item 3, and is a perfectly reasonable thing to say given an appropriate real-life context (for instance, you might go to a casual party before dressing up for

an evening church service). In other words, Gabriela's score suffered because she did not have the same background assumptions about the imagined situation to which her response had to be "contextually appropriate." Her final outcome on the test was, remarkably, "non-Spanish speaker," due to a score of zero on Part 6 of the test, which she obtained for offering no response to the prompt when asked to retell the story she had just heard.

Finally, as shown in Table 8, the scoring manual requires that a score of 3 be given on Part 5 of the test only if, in the judgment of the examiner, the response cannot be corrected. Of course, the untrained judgment of whether a response can be corrected may reflect a host of nonlinguistic factors, prescriptivist values and stylistic preferences among them. Indeed, virtually any sentence could be "corrected" by an overly zealous examiner.

Examinees who respond to each prompt in Part 5 might score as low as 5/15 (33%) due to the presence of regional variation or normal developmental errors in syntax, if the examiner construes syntactic differences of the sort discussed above as "awkward." We would expect children from some regions of Spain, Mexico, Ecuador, or the Caribbean to score as low as 10/15 (66%) on Part 5 of the Pre-LAS Español because the syntactic differences demonstrated in their Spanish are associated with a social stigma. In any event, because crucial terms are not defined, one would expect arbitrary scoring differences across children, according to the background assumptions of the examiners.

Finally, consider Part 6 of the test, in which the examinee's Spanish proficiency is evaluated with a language sample gathered by way of a story retelling task. Recall that Part 6 correlates .836 with the total score of the test (.768 after correcting for overlap), and is highly predictive of the total score. Because Part 6 has such a strong influence on a child's native language proficiency classification, a careful examination of the design of this section of the test is important.

Here, the examinee is told a story while looking at cue pictures, and is asked to retell the story to the examiner. Before the story begins, the examiner is directed to say, "*Ahora voy a contarte una historia con dibujos. Escucha con mucho cuidado. Después quiero que me digas lo que pasó.*" ("Now I'm going to tell you a story with pictures. Listen very carefully because afterwards I want you to tell me what happened"). The child's response is rated 0–5. This procedure is repeated with a second story.

There are three stories the examiner may choose from for this section of the test. The first story, *Pérez y Martina*, is about an ant named Martina who cooked some delicious soup in a giant pot for her ant husband, Pérez. The story relates that Pérez saw the pot of soup, climbed up a stool to peer in, and then fell into the soup. Martina called her friends to come rescue Pérez, after which all ate soup and rejoiced. The second story, *El Niño y el Lobo*, is the classic story about the boy who cried wolf. The boy cried "wolf" one day, and when the people came and saw there was no wolf, he laughed. The

next day the wolf really came; the boy shouted but nobody came, and the wolf ate all the sheep. The third story is “El Globo Amarillo,” the story of a yellow balloon that the wind carried away from its owner, a little girl. The balloon was delighted to be taking a trip, but later felt lonely and decided to return to the little girl’s house.

After hearing the story, the child is asked a series of factual recall questions in order to prompt him or her to retell the story. The story is transcribed verbatim, and scored according to the criteria indicated in Table 9.

Table 9

Description of Rating Levels for Contando Historias (Part VI)
(Duncan & DeAvila, 1986, p. 4)

Level	Description
0	No response in Spanish. Child produces no language or can only speak in a language other than Spanish.
1	The child produces only isolated words and expressions. While there are some differences across the age groups, they are very slight at this level of performance.
2	A few isolated phrases and fragments or very simple sentences are produced.
3	At Level 3, complete sentences are produced, often with systematic errors in syntax. The most salient characteristics of Level 3 are that a story line is present, however incomplete or rambling, and that the sentences, while more coherent than in Level 2, are still awkward. Thus, while the student may be able to produce sufficient vocabulary and facts necessary to retell the story, s/he has difficulty in combining the words with the same facility as that of the proficient Spanish speakers. It is also at this level that language mixing may appear.
4	At Level 4, the student responds in coherent, fluent Spanish appropriate for his/her age. While there may be errors in either syntax or vocabulary, these are errors which would not be uncommon among proficient speakers. The main difference between Level 4 and 5 is that the former is often a more limited version in terms of vocabulary and syntactical complexity. There may be developmental or dialectal variations in pronunciation.
5	At Level 5, the student produces language which is coherent and syntactical correct for his/her developmental age, and overall is an articulate, proficient Spanish speaker, although there may be some dialectal or developmental variation in pronunciation.

The scoring manual allows the child to respond off task and to tell a different story from the one provided in the prompt, which is then scored.

However, many children, unaccustomed to testing situations and the sorts of inquiries which follow, may not respond to this section of the test at all. As indicated earlier, in our sample of 38,887 examinees, 20% of children scoring 80% or higher on Parts 1–4 of the Pre-LAS Español gave no response at all to Part 6 of the test and were evaluated as “non-Spanish speakers.”

In a classic study of three communities, Heath (1983) noticed that “what” questions, like those presented by the examiner in Part 6 of the Pre-LAS Español, were typical of parent-child interactions in a mainstream community she studied, a kind of interaction also typical of elementary school culture. By contrast, parents in a working class community Heath studied did not ask questions of a child in story-telling situations. In an African-American community, children interacted with the storyline as though they were taking part in the creative process, adding sound effects and inventing new characters as events progressed.

Typically, we tell stories for the purpose of sharing experiences or relating ideas and feelings. Only children attuned to the mechanisms of evaluation, either at home or at school, will understand the purpose of basic questions that stress simple factual recall of events and characters. The structure of a narrative, as well as its social function, differs significantly from one culture to another, much as languages differ along a dialect continuum (Kaplan, 1966; Labov & Waletzky, 1967). Ability to recall storylines also appears to be related to general background knowledge. In a study conducted in Argentina, Signorini and Borzone de Manrique (1988) examined story recall in 33 third graders and found that the ability to recall any given narrative category was related to the child’s general world knowledge and previous experience. For these and other reasons, Gutierrez-Clellen and Quinn (1993) strongly caution against assessing children’s linguistic development by using narrative structure, recall, or presumed degree of “grammatical complexity” in story retelling.

Here, as in Part 5, an inference is made from the absence of evidence: Children receive a score of zero if they do not respond to the prompt. The scoring manual provides an example of a Level 1 response, offered by a 5-year-old child in response to the first story, as “*Buena sopa, no sé*” (“Good soup. I don’t know”). Although there is admittedly little to work with, there is no evidence here that the child does not speak Spanish: The construction is well formed, and a perfectly reasonable thing to say when asked a question like, “¿*Qué hizo ella un día?*” (“What did she make one day?”) (Duncan & DeAvila, 1986a, p. 15).

The definition of a Level 2 response suggests that “fragments” or incomplete sentences are indicative of a low level of Spanish language proficiency; it is contrasted with Level 3, in which “complete sentences are

produced.” This is an idea borrowed from traditional school grammar, but which has no relationship to the linguistic sciences. Not only do all normal people use fragments to interact every day, we do so in accord with specific pragmatic requirements. For instance, Grice (1975) observed that our conversational interactions are governed by a set of “conversational maxims,” one of which is the Maxim of Quantity, the provision that each participant’s contribution will be as informative as required by the situation, and no more. Ask any competent speaker of English or Spanish what Martina did for Pérez, and you’re likely to get a response such as, “Made some soup.” The response is a fragment, but a “complete sentence” such as “Martina made some soup for him” has the effect of “flouting” Grice’s maxim, and presses us to search for an interpretation beyond the compositional (literal) meaning of the response. Notice, too, that the so-called “incomplete response” requires detailed knowledge of the internal structure of the phrase out of which it is extracted. The verb “made” forms a substructure with its complement “some soup,” creating a larger constituent that excludes both “Martina” and “for him.” A truncated response such as “Martina made the soup for,” which deletes the object of the preposition, is ill-formed because the rules of deletion are sensitive to hierarchical structure in syntax (Radford, 1981).

The requirement to answer in complete sentences recalls Labov’s (1970) classic work on non-standard English. Labov reviewed work by Bereiter, Engelman, Osborn, and Reidford (1966), who developed a special preschool program to help disadvantaged African-American children. These authors conclude from a study of the language of 4-year-old children in this community that “without exaggerating . . . these 4-year-olds could make no statements of any kind” and appeared “as if [they] had no language at all.” The program Bereiter and Engelman developed presented children with questions like, “Where is the squirrel?” If the children answered “In the tree,” they would be provided with negative reinforcement of one kind or another; children who responded with the so-called “logical” response, “The squirrel is in the tree,” were to be rewarded.

Language mixing, or codeswitching, is also said to be a characteristic of Level 3, and “dialectal variation in pronunciation,” a characteristic of Level 4. However, neither language mixing nor regional variation implies linguistic incompetence. Indeed, codeswitchers have been shown to be exquisitely sensitive to extremely subtle properties of both their languages (MacSwan, 1999).

Level 5 marks a response that is overall “articulate,” even though there may be “developmental or dialectal variations in pronunciation.” Articulateness is an extremely subjective notion and, like “awkwardness,” is literary or stylistic. It is not a concept used in the linguistic sciences. Again, we see a charitable allowance for variation due to pronunciation, but what about variation due to syntax? Why should that be excluded?

Due to vagueness in the scoring manual, it is often quite difficult to apprehend differences between levels. For instance, the manual says that “the main difference between Level 4 and 5 is that the former is often a more limited version in terms of vocabulary and syntactical complexity.” Because there is no precise definition of “more limited version” or “syntactical complexity,”⁵ the distinction between 4 and 5, like the distinction between 1 and 2, is essentially arbitrary. In short, there is virtually no linguistically interesting difference among the scoring values rendered in Part 6, but we might conjecture that an examiner would give a high score to a child who was comfortable and chatty. Therefore, while a high score on Part 6 would likely indicate proficiency in Spanish, a lower score may reflect a student who is taciturn or uncomfortable with the testing situation (Levels 0–3), a tendency to codeswitch (Level 3), age-appropriate developmental patterns (Levels 4–5), regional variation in Spanish (Levels 4–5), or the ambiguity and lack of clarity in the scoring manual.

Both Tables 8 and 9 appear to have been developed by making arbitrary choices on relevant linguistic features, sometimes sanctioned by traditional prescriptivist grammar. Like criterion-referenced testing generally, both scoring rubrics are thoroughly arbitrary (see Glass, 1978, for discussion). Worse still, there is no conceivable theoretical justification for the “levels” presented, and none has any meaningful connection with linguistic and psycholinguistic research on language development.

Conclusion

Research on language acquisition tells us that all normal children acquire the language of their speech community effortlessly and flawlessly (Pinker, 1994). As Tager-Flusberg (1997, p. 188) has noted, “by the time children begin school, they have acquired most of the morphological and syntactic rules of their language,” and they possess a grammar essentially indistinguishable from adults. This view represents a consensus among researchers in child language acquisition, where the matter has been investigated for more than three decades. The results of the Pre-LAS Español provide evidence of anomalies that must be examined against the background of several decades of credible research on language acquisition. Indeed, Labov’s (1970) concluding comment in his classic work on non-standard language appears to be no less fitting today than it was more than 30 years ago: “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).

Because we know from research on child language acquisition that all normal children achieve linguistically, and because evidence for the existence of constructs like “semilingualism” has not been persuasive (Paulston, 1983;

MacSwan, 2000), we recommend that the practice of routinely testing minority language children's oral native language ability be abandoned. In the usual case, the assessment of language minority children for purposes of program placement and identification can be done with a simple home language survey, brief parent interview, and some kind of second language assessment (e.g., English, in the U.S. context).

Ability labels such as "non-non" and "semilingual," especially when they cannot be justified, stand to do considerable harm to children by subjecting them to needless "remediation" or placing them in language environments that are disadvantageous. Abandoning the practice of routinely assessing children's native language is a step toward affirming the linguistic and cultural resources of all children.

References

- August, D., & Hakuta, K. (1998). *Educating language minority students*. Washington, DC: National Academy Press.
- Bereiter, C., Engelman, S., Osborn, J., & Reidford, P. A. (1966). An academically oriented pre-school for culturally disadvantaged children. In F. M. Hechinger (Ed.), *Pre-school education today* (pp. 105–135). New York: Doubleday.
- Berko-Gleason, J. (Ed.). (1997). *The development of language*. Boston: Allyn and Bacon.
- Berwick, R. C., & Weinberg, A. (1986). *The grammatical basis of linguistic performance*. Cambridge, MA: MIT Press.
- Campbell, D. T. (1960). Recommendations for APA test standards regarding construct, trait, and discriminant validity. *American Psychologist*, 15, 546–553.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1, 1–47.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Commins, N., & Miramontes, O. (1989). Perceived and actual linguistic competence: A descriptive study of four low-achieving Hispanic bilingual students. *American Educational Research Journal*, 26 (4), 443–472.
- Council of Chief State School Officers. (1991). *Summary of state practices concerning the assessment of and the data collection about limited English proficient (LEP) students*. Washington, DC: Council of Chief State School Officers, State Education Assessment Center/Resource Center on Educational Equity.

- Cronbach, L., & Meehl, P. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52 (4), 281–302.
- 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. (1976). The influence of bilingualism on cognitive growth: A synthesis of research findings and explanatory hypotheses. *Working Papers on Bilingualism*, 9, 1–43.
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49, 221–251.
- Cummins, J. (1980). The cross-linguistic dimensions of language proficiency: Implications for bilingual education and the optimal age issue. *TESOL Quarterly*, 14 (3), 175–187.
- 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*. Sacramento: California State Department of Education.
- Cummins, J. (2000a). Putting language proficiency in its place: Responding to critiques of conversational/academic language distinction. In J. Cenoz & U. Jessner (Eds.), *English in Europe: The acquisition of a third language* (pp. 54–83). Clevedon, UK: Multilingual Matters.
- Cummins, J. (2000b). *Language, power and pedagogy: Bilingual children in the crossfire*. Clevedon, UK: Multilingual Matters.
- DeAvila, E. A., & Duncan, S. E. (1982). *A convergent approach to oral language assessment: Theoretical and technical specifications on the Language Assessment Scales (LAS), Form A*. Monterey, CA: CTB/McGraw-Hill.
- DeAvila, E. A., & Duncan, S. E. (1990). *LAS: Language assessment scales oral: Scoring and Interpretation manual, Spanish, level 2B*. Monterey, CA: CTB/McGraw-Hill.
- DeAvila, E. A., & Duncan, S. E. (1994). *LAS: Language assessment scales oral: Administration manual, Español, form 1b*. Monterey, CA: CTB/McGraw-Hill.
- DeAvila, E. A., & Duncan, S. E. (1998). *Pre-LAS 2000 Español*. Monterey, CA: CTB/McGraw-Hill.
- Duncan, S. E., & DeAvila, E. A. (1986a). *Scoring and interpretation manual for the Pre-LAS Español* (2nd ed.). Monterey, CA: CTB/McGraw-Hill.

- Duncan, S. E., & DeAvila, E. A. (1986b). *Pre-LAS Español*. Monterey, CA:CTB/McGraw-Hill.
- Edelsky, C. (1996). *With literacy and justice for all: Rethinking the social in language and education*. London: Taylor and Francis.
- Edelsky, C., Hudelson, S., Flores, B., Barkin, F., Altweger, J., & Jilbert, K. (1983). Semilingualism and language deficit. *Applied Linguistics*, 4, 1–22.
- Gee, J. P. (1996). *Social linguistics and literacies: Ideologies and discourses*. London: Taylor and Francis.
- Genesee, F. (1984). On Cummins' theoretical framework. In C. Rivera (Ed.), *Language proficiency and academic achievement* (pp. 20–27). Clevedon, UK: Multilingual Matters.
- Glass, G. V. (1966). Alpha factor analysis of infallible variables. *Psychometrika*, 31, 545–561.
- Glass, G. V. (1978). Standards and criteria. *Journal of Educational Measurement*, 15, 237–261.
- Gordon, P. (1996). The truth-value judgment task. In D. McDaniel, C. McKee, & H. S. Cairns (Eds.), *Methods for Assessing Children's Syntax* (pp. 211–232). Cambridge: MIT Press.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and Semantics 3: Speech Acts* (pp. 41–58). New York: Academic Press.
- Guerrero, M., & Del Vecchio, A. (1996). *Handbook of Spanish language proficiency tests*. Albuquerque, NM: Evaluation Assistance Center-Western Region. Retrieved March 4, 2002, from <http://www.ncbe.gwu.edu/miscpubs/eacwest/slptests.htm>
- Gutiérrez-Clellen, V. F., & Quinn, R. (1993). Assessing narratives of children from diverse cultural/linguistic groups. *Language, Speech, and Hearing Services in Schools*, 24, 2–9.
- Hakuta, K., & D' Andrea, D. (1992). Some properties of bilingual maintenance and loss in Mexican background high-school students. *Applied Linguistics*, 13, 72–99.
- Hansegård, N. E. (1968). *Tvåspråkighet eller halvspråkighet?* [Bilingualism or semilingualism?] Stockholm, Sweden: Aldus series 253.
- Heath, S. B. (1983). What no bedtime story means: Narrative skills at home and school. *Language and Society*, 11, 49–76.
- Kaplan, R. B. (1966). Cultural thought patterns in intercultural education. *Language Learning*, 1 (2), 1–20.
- Krashen, S. (1996). *Under attack: The case against bilingual education*. Culver City, CA: Language Education Associates.

- Labov, W. (1970). The logic of non-standard English. In F. Williams (Ed.), *Language and poverty*, Chicago: Rand McNally. (pp. 225–261).
- Labov, W., & Waletzky, J. (1967). Narrative analysis: Oral versions of personal experiences. In J. Helm (Ed.), *Essays on the verbal and visual arts* (pp. 12–44). Seattle: University of Washington Press.
- Lipski, J. (1994). *Latin American Spanish*. New York: Addison-Wesley.
- Lujan, M., & Parodi, C. (1996). Clitic-doubling and the acquisition of agreement in Spanish. In J. Guitierrez-Rexach and L. Silva-Villar (Eds.), *Perspectives on Spanish linguistics: Vol. 1*. Los Angeles: UCLA Program in Romance Linguistics. (pp. 119–138).
- Lust, B., Flynn, S., & Foley, C. (1996). What children know about what they say: Elicited imitation as a research method for assessing children's syntax. In D. McDaniel, C. McKee, & H. S. Cairns (Eds.), *Methods for Assessing Children's Syntax* (pp. 55–76). Cambridge, MA: MIT Press.
- MacSwan, J. (1999). *A minimalist approach to intrasentential code switching*. New York: Garland Press.
- MacSwan, J. (2000). The Threshold Hypothesis, semilingualism, and other contributions to a deficit view of linguistic minorities. *Hispanic Journal of Behavioral Sciences*, 22 (1), 3–45. (Retrieved June 6, 2002) Available at <http://www.public.asu.edu/~macswan/hjbs2000.pdf>
- MacSwan, J., & Rolstad, K. (in press). Linguistic diversity, schooling, and social class: Rethinking our conception of language proficiency in language minority education. In C. B. Paulston & R. Tucker (Eds.), *Essential Readings in Sociolinguistics*. Oxford: Blackwell.
- Martin-Jones, M., & Romaine, S. (1986). Semilingualism: A half-baked theory of communicative competence. *Applied Linguistics*, 7(1), 26–38.
- Milroy, J., & Milroy, L. (1999). *Authority in language: Investigating standard English*. New York: Routledge.
- Morales, A. (1986). *Gramáticas en contacto: Análisis sintácticos sobre el español de Puerto Rico*. Madrid: Biblioteca de Autores de Puerto Rico.
- National Clearinghouse for Bilingual Education. (2000). *Summary report of the survey of the states' limited English proficient students and available educational programs and services, 1997–1998*. Washington, DC: National Clearinghouse for Bilingual Education. Retrieved March 4, 2002, from <http://www.ncbe.gwu.edu/ncbepubs/seareports/97-98>
- Newmeyer, F. J. (1986). *The politics of linguistics*. Chicago: The University of Chicago Press.
- Paulston, C. B. (1983). *Swedish research and debate about bilingualism*. Stockholm, Sweden: National Swedish Board of Education.

- Petrovic, J. E., & Olmstead, S. (2000). Review of Jim Cummins, Language, power, and pedagogy: Bilingual children in the crossfire. *Bilingual Research Journal*, 25, 405–410.
- Pinker, S. (1994). *The language instinct: How the mind creates languages*. New York: William Morrow and Company.
- Pyle, A. (1996, June 11). Teaching the silent student. *Los Angeles Times*.
- Radford, A. (1981). *Transformational syntax*. Cambridge, UK: Cambridge University Press.
- Ramirez, D., Pasta, D., Yuen, S., Billings, D., & Ramey, D. (1991). *Final report. Longitudinal study of structured English immersion strategy, early-exit and late-exit transitional bilingual education programs for language-minority children*. (Vols. 1 & 2). San Mateo, CA: Aguirre International.
- Signorini, A., & Borzone de Manrique, A. M. (1988). Incidencia del esquema narrativo en la comprensión y el recuerdo de cuentos. *Revista Argentina de Linguística*, 4 (1–2), 91–117.
- Skutnabb-Kangas, T., & Toukomaa, P. (1976). *Teaching migrant children's mother tongue and learning the language of the host country in the context of the socio-cultural situation of the migrant family*. Helsinki, Finland: The Finnish National Commission for UNESCO.
- Smith, C., & van Kleeck, A. (1986). Linguistic complexity and performance. *Journal of Child Language*, 13, 389–408.
- Spolsky, B. (1984). A note on the dangers of terminology innovation. In C. Rivera (Ed.), *Language proficiency and academic achievement* (pp. 41–43). Clevedon, UK: Multilingual Matters.
- Tager-Flusberg, H. (1997). Putting words together: Morphology and syntax in the preschool years. In J. Berko-Gleason (Ed), *The development of language*. Boston: Allyn and Bacon.
- Troike, R. (1984). SCALP: Social and cultural aspects of language proficiency. In C. Rivera (Ed.), *Language proficiency and academic achievement* (pp. 44–54). Clevedon, UK: Multilingual Matters.
- Valadez, C. M., MacSwan, J., & Martínez, C. (2001). Toward a new view of low achieving bilinguals: A study of linguistic competence in designated “Semilinguals.” *Bilingual Review*, 25 (3), 281–291.
- Wexler, K., & Culicover, P. W. (1980). *Formal Principles of Language Acquisition*. Cambridge, MA: MIT Press.
- Wiley, T. (1996). *Literacy and language diversity in the United States*. McHenry, IL: Center for Applied Linguistics and Delta Systems.
- Willig, A. C. (1985). A meta-analysis of selected studies on the effectiveness of bilingual education. *Review of Educational Research*, 56 (3), 269–317.

Endnotes

¹ The data and content analysis considered in this article derive from the 1986 version of the Pre-LAS Español (Duncan & DeAvila, 1986b), not the recent revision of the test known as the Pre-LAS 2000 Español (DeAvila & Duncan, 1998). The new version of the test has five parts instead of six, with parts 3 and 5 of the older version eliminated, and a new part (*El Cuerpo Humano*, “The Human Body”) requires children to give names of body parts (compare Table 2). Although we do not have data that speak specifically to the new version of the Pre-LAS Español, its general similarity with the previous version of the test leads us to believe that a study of that instrument would lead us to conclusions similar to those reached here.

² A note on terminology is in order here regarding “competence” and “proficiency.” Chomsky (1965) introduced “linguistic competence” to refer to an individual’s internal knowledge of language structure and contrasted it with “linguistic performance,” one’s knowledge of language use that interacts with a host of other cognitive and external factors. In early work, Cummins (1976, 1979) used “linguistic competence” (presumably borrowed from Chomsky’s work) and “language proficiency” interchangeably, but later, following Canale and Swain (1980), defined “language proficiency” to include linguistic competence as well as aspects of language use—school literacy, in particular. Although there may be good reason to distinguish “proficiency” and “competence,” in the present context, where we focus narrowly on native oral language ability, we will use the terms interchangeably.

³ For another view of literacy in the context of language development, see Edelsky (1996, chapter 5). MacSwan and Rolstad (in press) suggest replacing the BICS/CALP distinction with SLIC, or Second Language Instructional Competence, a locally determined level of second language ability at which children are able to understand academic instruction rendered exclusively in the L2.

⁴ Clitic doubling is a syntactic phenomenon in which a pronominal clitic appears along with the noun to which it refers; thus, in the example *La vi a la mujer* (I saw the woman) the clitic *la* appears with the noun phrase *la mujer*. In more prestigious varieties of Spanish, the clitic and object noun phrase are in complementary distribution (*La vi* or *Vi a la mujer*).

⁵ There are well-defined notions of syntactic complexity available, but these notions do not appear to be related to Duncan and DeAvila’s (1986a) conception of complexity. Linguists speak of syntactic complexity in terms of parsing/production models. We might construct a theory, for instance, which associates one “cost unit” with each step in a derivation of a structure *S*, and compute the complexity of *S* by simply tallying up its cost units (known as the derivational theory of complexity). Or we might take the complexity of *S* to be a function of its “depth,” measured in terms of substructures of one type or another. There are numerous views, with distinct empirical consequences, and all make use of highly theory-internal constructs (see Berwick & Weinberg, 1986 for discussion). These ideas are part of the theory of linguistic performance, and are not generally understood as a subcomponent of linguistic competence. Just as in the case of normal language acquisition, it is doubtful that one would find any meaningful variation among normal children on theoretically defensible notions of native-language parsing/production complexity of the sort required for psychometric validation of language tests.