The Language Demands of School
Putting Academic English to the Test

Edited by ALISON L. BAILEY
University of California, Los Angeles

Introduction: Teaching and Assessing Students Learning English in School

ALISON L. BAILEY

The subtitle of this book, "Putting Academic English to the Test," is intentionally ambiguous. Each of the two possible meanings serves an important purpose. In the idiomatic use of the phrase *to put to the test*, the subtitle invites a challenge to the authors of many of the chapters: Can the research on academic language be put to service in educationally meaningful ways? What does the operationalization of the academic language concept afford us in terms of fairer and better assessment? Will documenting the hallmarks of academic language lead to greater effectiveness in curriculum and teacher professional development? Answering these questions in any great depth would take far more space than can be allotted in this one volume. Rather, the goal of this book is to illustrate what has been achieved so far and to provide frameworks that our research at the National Center for the Study of Evaluation, Standards, and Student Testing (CRESST) has produced so that others might yet yield what still needs to be achieved.

In its second meaning, *putting academic English to the test* is read literally—adding the construct "academic English language" (AEL) to assessments of English language development (ELD). This meaning captures the major focus of the book on the current state of language assessment for students learning English in schools in the United States. A review of the available ELD assessments in the spring of 2005 as I write this introduction suggests that much is in flux.

Currently, assessments of English language development aimed at the 5–18-year-old population in the United States fall into three main categories: assessments that have been available for the past 15 to 20 years that appear to tap into the language of social settings,

new assessments that have partially integrated the academic language construct into their operational definitions, and assessments under development that have embraced the notion of academic language across all four language modalities—listening, speaking, reading, and writing.

Assessment of English language learners (ELLs) in kindergarten through high school is a requirement of the No Child Left Behind Act (NCLB, 2001). This federal law mandated the yearly assessment of English language skills for ELL students beginning in the 2002–3 school year. Many states are still in the process of responding to this mandate. Many are still administering the older generation of assessments that focus on social or general uses of language rather than language aligned with the discourse of the classroom, textbooks, educational standards, and content-area assessments. Those that are in the process of developing or refining new ELD assessments are daily seeking guidance from language researchers and test developers. The chapters in this book have pulled together much of what my colleagues and I have researched and proposed to a variety of states and the assessment consortia who represent them.

In this initial chapter, I will first provide greater detail about the political and educational contexts that have led to the current activity in language assessment and in assessment and instructional practices more broadly for language minority students in the United States. I will then briefly highlight some of the new assessment development efforts across the states and the implications these efforts will have for instruction and teacher professional development. I hope that the articulation of these implications will make self-evident the need for the detailed review of what we currently think academic English language to be. This review reveals large and educationally critical gaps in our knowledge of what students need to know about language in academic settings and how we think they acquire it.

The Language Mandate

There are major research and policy problems facing the United States and other English-speaking countries with large populations of primary and secondary school students learning English in academic contexts for academic purposes. The problems can be most succinctly articulated as a lack of comprehensive information about

what language demands are placed on school-age children in general and how much English language learners can realistically be expected to learn and how quickly. While we have a large literature base about English as a second language (ESL), this research base cannot answer these questions with sufficient specificity to aid policy makers and educators faced with the creation of English language development tests and curricula.

The number of students in U.S. public schools for whom English is a second language has grown steadily during the past two decades. The latest figures show ELL student enrollment at 5.1 million (U.S. Department of Education, 2005). During this period of growth, educators have struggled to implement approaches that help to ensure both quality instruction and valid and reliable assessments of all students. Unfortunately, English language learners often enter school without the requisite English language skills to benefit from the mainstream curriculum, and thus in the past they were often excluded from accountability systems. Now, however, with the No Child Left Behind Act (2001), focus on adequate yearly progress in math, reading, and shortly (2005-6 school year) science for all students (Title I) and special emphasis on ensuring that ELLs make steady progress in acquiring English (Title III) virtually guarantees the inclusion of ELLs in state accountability systems. Under this law, ELL students must show measurable progress each year in the listening, speaking, reading, writing, and comprehension of English (many states are interpreting the fifth element as a combination of the listening comprehension and reading comprehension assessment subsections). Consequently, as already mentioned, states are urgently identifying or developing tests of English language proficiency that can help them meet this federal mandate (Olson, 2002).

Previous research in the area of rate and attainment of ELD has suggested that ELL students can take between four to eight years to achieve the English language proficiency necessary for success on academic content assessments (e.g., Collier, 1995; Cummins, 1981; Hakuta, Butler, and Witt, 2000). Students who start their U.S. schooling in upper elementary grades make more rapid progress than those who start at 12 years and older, often despite the same number of years of exposure to English and years in U.S. schooling (Collier, 1987). Differences in achievement by discipline have also been found, with quicker and greater mathematics attainment (while English language is still developing) than science, social studies (Col-

lier, 1989), or language arts attainment (Butler, Stevens, and Castellon, chapter 2; De Avila, 1997). Relatedly, research on the poorer academic achievement of students who are recent redesignated fluent English proficient (RFEP) students (compared to other students in mainstream classrooms) suggests their difficulties may be due in part to the demands of English in nonsheltered content classes and on standardized content assessments (e.g., Stack, 2002).

While this research base provides estimates of student gains and expectations for ELD performance and its impact on academic achievement, it has limitations for answering questions about the nature of AEL development. These limitations include (1) the appropriateness of existing instruments for measuring forms and uses of school language, as already mentioned above; and (2) the use of primarily cross-sectional design that does not provide longitudinal data on individual student growth and exposure to the language of school. Collectively, these pose a serious challenge to accurate and useful prediction of language and academic outcomes.

Including ELLs in accountability systems is not without challenges. For example, the language demands of content-area assessments may be so great for ELL students as to invalidate the determination of their content knowledge. Ultimately, we often have no way of knowing if the performance of ELL students primarily reflects their language abilities or their content knowledge. Thus, including ELLs in the testing process, knowing that the interpretation of test scores may be invalid, is problematic (August and Hakuta, 1997). However, to exclude ELL students is also unacceptable. If ELL students are not tested, information on their achievement is, in effect, absent from any decision making that impacts their school careers.

Educate or Accommodate?

One widely used means of achieving the goal of inclusion for all students in academic achievement assessment is the use of test accommodations with students who have not yet mastered English (e.g., Butler and Stevens, 1997; Kopriva, 2000; LaCelle-Peterson and Rivera, 1994; Rivera, Stansfield, Scialdone, and Sharkey, 2000). Test accommodations can consist of modifications made to the test itself or modifications made to the test administration procedures, including the test setting. For example, in terms of modifications to

the test, students may receive a mathematics test translated into their native language or rendered into simplified English. Accommodations made to the administration of the test include providing students with extra time to complete a test or allowing students the use of English-native language dictionaries.

Studies that have investigated the effects of language accommodations on both ELL and EO students' academic performance (e.g., Abedi, Courtney, and Leon, 2001; Abedi, Lord, Hofstetter, and Baker, 2000; Abedi, Lord, and Plummer, 1997; Castellon-Wellington, 1999; Rivera and Stansfield, 2001) leave us with mixed results as to the fairer assessment of ELL and EO students (extra time, for example, benefits all students). On the one hand, ELL student language abilities, particularly low levels of reading ability, may be a source of construct-irrelevant variance in assessments of academic achievement, such as in mathematics. That is, the mathematics learning that we hope to capture in the assessment may not be reflected; rather, we are measuring the students' inability to read the test questions. On the other hand, if, as Haladyna and Downing (2004) point out, we need to know if students can handle the complexity of language necessary to learn and convey their learning of mathematical concepts, then assessments should include the language demands typical of the instructional contexts students will encounter. Arguably, no less an assessment will allow us to know if students are ready to cope with the language of mainstream mathematics classes.

The complexities of assessing the academic achievement of ELL students are thus many. We argue that there is still much research needed to determine the effectiveness of using accommodations and even then, as others have stated before us, we need to establish clearer procedures for making accommodation decisions in an informed and systematic way (e.g., August and Hakuta, 1997; Bailey and Butler, chapter 4; Haladyna and Downing, 2004). While accommodation studies all share the same objective—to inform educational policy makers about the equitable inclusion of both ELL students and English proficient students on content-area assessments—we have suggested a change of approach that will require future inclusion efforts to begin first by ensuring equitable exposure to and learning of AEL and second by devising academic English language proficiency assessments to help gauge student readiness for mainstream instruction and content-area testing (Bailey and Butler, 2004; Bailey and Butler, chapter 4).

New Assessments of ELD

A major issue, then, at this critical policy decision-making juncture, is how to determine if the requisite English language skills for demonstrating content knowledge on assessments have been acquired by ELL students. Many existing English language proficiency tests do not assess the type of language students must understand and be able to use in the mainstream classroom and on standardized content tests. Existing language tests tend to assess social everyday language rather than the more formal academic English language of the classroom and content tests. Unfortunately for current educational testing needs, basic social language has been found to be only minimally correlated with the more demanding language of school (e.g., Collier and Thomas, 1989).

This is not to minimize the importance of social language skills for successful school and personal outcomes. Indeed, the level of linguistic sophistication necessary to navigate everyday informal situations suggests that we also need to foster student growth in the area of social language development (Bailey, in press). However, currently a student may perform well on a general language test and still not have the necessary language skills for academic tasks (see Butler et al., chapter 2 for a discussion of just such findings with eleventh-grade students). There is, then, an important assessment gap between the type of English an ELL student may know and be able to use—that tested on many current ELD tests—and the language critical to school success.

As mentioned, assessments of ELD fall into three main categories: older assessments that appear to primarily measure social language uses, new assessments that have partially integrated the academic language construct into operational definitions of ELD, and assessments still under development that have adopted the AEL construct as critical to the measurement of scholastic uses of English. For example, the Language Assessment Scales (De Avila and Duncan, 1987, 1988, 1989, 1990; Duncan and De Avila, 1990), are widely used older ELD assessments that are found to have fewer language features matching the language demands of content-area assessments (Butler et al., chapter 2; Stevens, Butler, and Castellon-Wellington, 2000) than derived versions of the LAS such as the current California English Language Development Test (CELDT, CTB/McGrawHill, 2004). The CELDT shows the influence of the California ELD stan-

dards (California State Board of Education, 1999) and includes characteristics of classroom language on the test blueprint in all four language modalities (listening, speaking, reading, and writing) (see Bailey and Butler, 2005; Sato, Lagunoff, Worth, Bailey, and Butler, 2005).

Other states are following similar initiatives and have new assessments at different stages of development incorporating varying degrees of the AEL construct.[1] For example, Wisconsin, Delaware, and Arkansas have worked together to produce the WIDA (WIDA Consortium, 2005; see also Davidson, Kim, Lee, Li, and López, chapter 6), which has developed new English language proficiency standards. These will be adopted by Teachers of English to Speakers of Other Languages (TESOL) to replace their original K-12 ELD standards (TESOL, 1997), which had served as the national ELD standards and as the foundation of many state ELD standards and test blueprints.

The English Language Development Assessment (ELDA) is the effort of member states of the Council of Chief State School Officers (CCSSO, 2005). The ELDA will include "topics" from the content areas, such as social studies and science, presumably as part of an AEL construct. Presumably the ELDA item writers will use these topics to create school-related scenarios or to provide specifications for vocabulary to be used in test items rather than to measure content-area knowledge itself.

Both the WIDA and ELDA will include all language modalities. Interestingly, one new ELD assessment responding to the NCLB Act mandate does not include all modalities within its AEL construct. The Stanford English Language Proficiency Test (Harcourt Educational Measurement, 2003), while based on the Bailey and Butler conceptual framework for AEL presented in chapter 4, has articulated speaking skills as social uses of language only and not as part of the AEL construct.

The new IPT/PEM (IDEA Proficiency Test/Pearson Educational Measurement) Title III Testing Solution (Ballard and Tighe, 2005) has also adopted the AEL construct framework as the basis for this replacement of the former, more social language-oriented IPT assessments, adding to the empirical database through analysis of videos of classroom teachers talking and various state-level ELD and content-standards documents following Bailey and Butler (2002/3; see also chapter 4) (Bailey and Luoma, 2003; Luoma, Cho, and Buck,

2004). This new suite of IPT assessments also includes K-2 reading and writing forms as well as a pre-K/K language and early literacy assessment (Bailey, Luoma, and Cho, 2005), whereas a number of the new ELD assessments being developed do not include assessment of literacy skills earlier than the second- or third-grade levels. There is some debate in the field as to what constitutes a reading and writing test with ELL students when English-only (EO) and fluent English proficient (FEP) students have also yet to learn to read and write. This, together with a concern that young children are excessively assessed, has led some states (e.g., California) to eschew the assessment of literacy skills in young ELL students (see Mayer, chapter 3 for further discussion). However, for a successful school experience, four- to five-year-old ELL students will no doubt need to participate in all aspects of their preschool or kindergarten curriculum, engaging with both teachers and peers to acquire and demonstrate knowledge of both social skills *and* nascent academic demands. Indeed, goal 1 of the National Educational Goals Panel (NEGP, n.d.) focuses on having all children start school "ready to learn." This goal specifically includes language development in its five dimensions of early development and learning.

If test developers take young children's social and cognitive developmental needs into account, then assessing both their oral and early literacy abilities can be possible. For example, using manipulatives (objects, not pictures) to engage the young ELL student during assessment allows children this age to do better on both production and comprehension tasks (e.g., Cocking and McHale, 1981), and providing verbal and gestural supports or scaffolding allows a young child to demonstrate to the assessor at what level of assistance he or she can and cannot complete a task. The latter, of course, is also information to the assessor as to what kind of instructional supports a very young ELL student may require in a preschool program.

Differences in the operationalization of language constructs found in the ELD tests outlined above may impact how students are instructed in ELD and how teachers themselves understand the construct of AEL. So what exactly is the AEL construct that these and other test developers now purport to tap with the new assessments? The next section of this chapter attempts to shed some light on this question by reviewing studies of English language acquisition and showing how these have informed our research on academic language at CRESST.

What English Language Demands Do We Place on K-12 Students?

As an entry point to understanding the AEL construct, we have found the following definition by Chamot and O'Malley (1994) to be both widely accessible and inspiring in its simplicity. Academic language is "the language that is used by teachers and students for the purpose of acquiring new knowledge and skills . . . imparting new information, describing abstract ideas, and developing students' conceptual understanding" (p. 40).

Cummins (1980, 2000) first made the crucial distinction between what is known as basic interpersonal communication skills (BICS), acquired and used in everyday interactions, and cognitive academic language proficiency (CALP), acquired and used in the academic context. BICS, or social, everyday language, most often occurs in the "here and now" of a shared conversation, allowing for use of less demanding indexical modes of communication, such as gestures and pronominalization (Schleppegrell, 2001, 2004). It is on the construct CALP that Chamot and O'Malley (1994; Chamot, 2005) based their definition of AEL and their curriculum for teaching it—the Cognitive Academic Language Learning Approach (CALLA).

As I have stressed elsewhere (Bailey, in press), we should guard against believing that there is something inherent in social language that makes it less sophisticated or less cognitively demanding than language used in an academic context. Think about the complex and highly sophisticated social applications of language that are needed to woo or deceive a loved one; and conversely, there are contexts in which AEL can be as simple as responding with a head nod to a yes or no question. Thus it is perhaps most accurate to speak of the difference between BICS and CALP as differences in the relative frequency of complex grammatical structures, specialized vocabulary, and uncommon language functions.

In some regard, it is not meaningful to conceive of *language* as either social or academic, rather it is the *situation* that is either predominantly social or academic. Even this distinction is problematic. Dewey's pedagogic creed (1897) serves as a good reminder that schools are simply one more social context or "community" and that language is always "a social instrument." Sociolinguistic studies of the language production of school-age children in both formal and informal settings (e.g., Gumperz and Cook-Gumperz, 1980; Hymes,

1974; Labov, 1972) have documented the social context of scholastic settings whereby becoming a member of a discourse community, such as that of mathematicians, is signaled by the use of certain discourse and linguistic features most commonly found in settings where "mathematics" is spoken (Pimm, 1987). Elsewhere, Johns (1997) identifies a register of English used in professional books and characterized by the specific linguistic features associated with the different academic disciplines.

While current cognitive theory has also increasingly argued that cognition and discourse are often specific to a discipline or context (e.g., Lave and Wenger, 1991), in our own work and in the literature we have found some remarkable similarities across the disciplinary discourses for teaching mathematics, science, and social studies. For example, current reform in each of these disciplines calls for students to engage in academic arguments and debates that may share the same underlying organizational structures (e.g., Bazerman, 2003). This has led us to hypothesize discipline-specific AEL as well as a common core or general AEL that cuts across disciplines (see Bailey and Butler, chapter 4 for further discussion).

We can add two additional features to Chamot and O'Malley's (1994) definition of academic language cited above. First, language as it is used in academic contexts requires students to demonstrate their knowledge by using recognizable verbal and written academic formats. For example, students must learn norms for presenting information to the teacher so that the teacher can successfully monitor their learning. The opportunity to display knowledge may also be as important as the opportunity to learn (OTL) new information, as it not only confirms acquisition for a teacher but may consolidate learning for a student (Cazden, 2001). Second, students receive fewer opportunities to negotiate meaning or to use contextual cues in the classroom setting than in many social settings. This is the decontextualized use of language that has been recognized as a key discourse-level feature of language used in school (e.g., Cazden, 2001; Menyuk, 1995; Snow, 1991).

Bailey (in press) has summarized the ability to be *academically proficient* in either a first or second language as knowing and being able to "use general and content-specific vocabulary, specialized or complex grammatical structures, and multifarious language functions and discourse structures—all for the purpose of acquiring

new knowledge and skills, interacting about a topic, or imparting information to others." Short (1994) and Lemke (1990), in particular, have documented the range of language functions found in social studies classes and science classes respectively. Most commonly language is used in service of *explanation* and *description*, and, in the case of social studies, *justification*. These findings are largely replicated in the analyses of language use in science classrooms and social studies and science textbooks reported by Bailey, Butler, Stevens, and Lord in chapter 5.

Students must be able to integrate all aspects of language that they hear and use in school and do so across all language modalities. That is, teachers often set tasks and activities that rely on students being able to process information in both the oral (speaking and listening modalities) and the print (reading and writing modalities) domains. It is the integration of the modalities themselves that Gibbons (1998) has termed *intertextual*. This use of language requires students to shift the same information back and forth across the modalities, perhaps first reading about a new science topic, then participating as both speaker and attentive audience member in a whole-group discussion, ultimately individually authoring an essay on the new information they have acquired. Oral language skills, however, are not only critical for signaling classroom participation; they also serve as a foundation upon which reading and writing skills are based, both in the initial acquisition of print skills (e.g., Allen, 1976; Dickinson, McCabe, Anastasopoulos, Peisner-Feinberg, and Poe, 2003; Stauffer, 1980) and in the ongoing process of reading and writing about new ideas and knowledge.

While there is much research on the second language acquisition process and differential outcomes for learners (see Larsen-Freeman, 2000 for review), this work has focused primarily on the development of language as a broad construct, only recently articulating the notion of English for specific purposes—for example, as the medium for acquiring discipline-specific knowledge (e.g., Douglas, 2000). However, this research has been conducted almost exclusively at the higher education level with already-educated adult ESL learners. In contrast, researchers appear to have largely ignored the K-12 arena, finding it difficult to articulate norms for even first language development in school-age children. The potential for and causes of individual variation in first and second language develop-

ment are great. Students are exposed to a wide variety of different language settings, not least by their own choice of content classes once they enter the higher grades (Nippold, 1995).

Collectively, the research cited here has documented the existence of an AEL phenomenon. However, there have been few attempts to operationalize the concept sufficiently for utilization in assessment, curriculum, or professional development. To achieve further specificity, therefore, academic language is defined in our work as language that stands in contrast to the everyday informal speech that students use outside the classroom environment. AEL can be distinguished from English in other settings on at least three key levels: the lexical, grammatical, and discourse levels, each described below.

Lexical Features of AEL

At the lexical level, following Scarcella and Zimmerman (1998), we suggest that academic vocabulary, as one component of the broader academic language construct, comprises both a general academic lexicon (e.g., containing words such as *evidence, demonstrate,* and *represent*) and specialized ones (e.g., containing content-related words such as *diameter, condense,* and *abolitionist*), each of which students must acquire in order to become fully proficient in English in the academic setting.[2] According to Nation (2001), at the tertiary education level, general academic vocabulary covers on average 8.5% of words in academic texts, and specialized or technical vocabulary covers about 5% of words in academic texts (see also Nation and Coxhead, 2001).

Beck et al. (2002), Cunningham and Moore (1993), and Stevens, Butler, and Castellon-Wellington (2000) add a third category of vocabulary to this mix—high-frequency general words used regularly in everyday contexts that contrast with the often more morphologically complex vocabulary of school, though they may have equivalent meanings in some cases (e.g., *gather* vs. *collate*) (Corson, 1997). Conversely, a single word can have multiple meanings, which helps in efficiently expanding a child's lexicon, adding one new word form but many different meanings. The English language, it is argued, is comprised of more than 60%-70% of such polysemous words (e.g., McLaughlin, August, and Snow, 2000). ELL students especially may have constrained English vocabularies in that they may know just one or two of the most frequently encountered meanings of an English word (e.g., McLaughlin et al., 2000), and the importance of vocabulary breadth and depth is further underscored by the finding that the vocabulary knowledge of fifth and sixth-grade Latino students can be the most important predictor of performance on a reading assessment (Garcia, 1991).

The primary distinction in the coding schema developed for the study of textbooks reported in Bailey et al. in chapter 5 was between academic and nonacademic usage of words. Within academic usage we further distinguished between specialized academic vocabulary and general academic vocabulary that cuts across disciplines. This distinction is important for future test development efforts that will attempt to target a broad cross-section of academic language and may therefore need to treat specialized vocabulary separately. During the course of coding we considered the sense intended in the selections and rated the form and usage of words only in the context of the given selections. Specifically, we had to be sure that the word sense intended in the textbook selections was referring to an academic concept (e.g., "Determine the centrifugal *force*" versus "Don't *force* him to do it"). This includes the specialized word sense often used in mathematics for some of the most common words in English. Prepositions, for instance, take on very precise and often unfamiliar usage in the mathematics register (e.g., Pimm, 1987, 1995). For example, the preposition *in* used in a phrase like "three in four" makes the relationship between the two numbers proportional, in this case, three-quarters or 75%. Even in the earliest grades use of a simple preposition like *by* can take on the unfamiliar meaning of *according to* in a phrase like *sort by color* (Bailey, in press).

In the analysis reported in chapter 5, if the intended meaning of a word was the same both in and out of the classroom setting, we considered the word usage to be nonacademic. Throughout this process we were careful to avoid equating unfamiliar words with academic vocabulary. While we did not classify the use of seemingly arbitrary proper names (such as *Peter* and *Anne*), most often found in mathematics word problems, as academic vocabulary, proper names related to content learning or crucial to the academic concepts of a topic were rated as specialized (e.g., *Continental Congress, Newton,* and *John Adams*). Similarly, measurement vocabulary and abbreviations for measurement and formulas were also rated as a subcategory of specialized academic vocabulary (e.g., *kilometer, km,*

and *km/hr*). Colloquialisms and idiomatic expressions (e.g., *half-joked* and *twister*) and verbatim speech formed their own separate subcategories if used to convey content information.

Grammatical Features of AEL

Turning to the grammatical or syntactic features of AEL, there are features of English grammar that are often only encountered during formal discourse or in print and for that reason are key components of the articulation of the AEL construct. In mathematics, for example, certain syntactic structures are required for conveying notions of comparative size (*X is greater/less than Y, as X as Y*) and conditional relationships (*if X, then Y*), among others (e.g., Spanos, Rhodes, Dale, and Crandall, 1988). Snow (1990) has argued that a syntactic skill that often cuts across specific content areas—writing formal definitions—requires students to combine the copula *to be*, a superordinate category, and a complement clause (e.g., *A rabbit is an animal with long ears*).

Oral and printed language often differ syntactically and place different demands on users of the language (e.g., Reppen, 2001; Schleppegrell, 2001, 2004). For example, passive voice constructions in English such as "The British were trounced by the colonials" occur less frequently in everyday conversation than in academic writing (Celce-Murcia and Larsen-Freeman, 1983), and children have been found to have more difficulty understanding passive verb forms than active verb forms (Bever, 1970; de Villiers and de Villiers, 1973; Fraser, Bellugi, and Brown, 1963). Eighth-grade students in one study, for example, were given equivalent mathematics test items with and without passive voice constructions and scored higher on the items without passive constructions (Abedi and Lord, 2001; Abedi et al., 1997).

Other syntactic features are also characteristic of written language and may increase the demands of classroom language (Halliday and Martin, 1993).[3] Relative clauses, for example, are less frequent in spoken English than in written English, so some students may have had limited exposure to them before entering school. Furthermore, relative clauses in print differ from those in spoken language (Pauley and Syder, 1983). In general, dependent clauses that embed information may make sentences more complex and difficult than coordinate clauses that do not (Botel and Granowsky,

1974; Hunt, 1965, 1977; Kemper, Jackson, Cheung, and Anagnopoulous, 1993; Lord, 2002; Wang, 1970). Conditional clauses also contribute to text difficulty, and noun phrases with several modifiers can be sources of difficulty, as Spanos et al. (1988) found in student comprehension of math test items.

Discourse Features of AEL

Turning to classroom discourse, students are required to convey the content knowledge they acquire using organizational features recognizable to teachers (Hicks, 1994). That is, academic success involves a combination of discipline-specific practices (e.g., knowing the conventions for multidigit division) and cross-cutting general academic discourse practices (e.g., knowing how to clearly present and justify an idea in any content area). This requires not only a knowledge of discourse conventions but also pragmatic knowledge that may be applied in both spoken and written contexts. The linguistic registers of the different disciplines and the uses to which language is put in content-area classrooms have also been a focus of increasing study (e.g., Bailey, Butler, LaFramenta and Ong, 2001/4; Johns, 1997; Lemke, 1990; Short, 1993). These studies show not only a wide range of functions language serves in the academic setting (e.g., *explaining, describing, comparing*) but also how the content areas are differentiated by language function (see Bailey et al., chapter 5).

Closely coupled with issues of discourse are the notions of academic language as a sociocultural and psychological phenomenon (e.g., Cazden, 2001; Corson, 1997; Erickson, 1987; Gee, 1996; Gutiérrez, 1995; Kern, 2000; Philips, 1972; Scarcella, 2003). The development of a deep conceptual understanding of discipline-specific concepts is tied to participation in the particularistic discourse practices of disciplinary communities (Greeno and Hall, 1997; Lave and Wenger, 1991; Roth, 2001) such as mathematics (Cobb, Stephan, McClain, and Gravemeijer, 2001; Lampert and Ball, 1998), science (Lehrer, Schauble, Carpenter, and Penner, 2000; Roth and Bowen, 1995), and English language arts (Cazden, 1988, 2001; Gutiérrez, 1994; Lee, 2001; Palinscar and Brown, 1984; Smagorinsky, 2001). Different disciplines have their own accepted norms for what constitutes a good argument during discourse. For example, in mathematics it is demonstration of a logical proof (Lehrer, Randle, and Sancilio, 1989; Kazami, 1999), in science it is often modeling to explain a

process (White and Frederiksen, 1998), and in English language arts it is the appropriate citation of primary and secondary texts (Lee, 2001). Such norms for each discipline need to be acquired to meet teacher expectations.

Participation frameworks or structures (Erickson, 1992; Goffman, 1981) influence who can say what and to whom, which also affects the nature of discourse in classroom activities. A common participation framework for discourse in the classroom is the initiate-respond-evaluate/feedback (IRE/F) script (Lemke, 1990; Mehan, 1979), which is a framework in which the teacher asks questions that can be answered briefly and quickly by the students. The teacher then simply evaluates the response as correct or incorrect or provides feedback to students regarding the accuracy and completeness of their answers (Mehan, 1979; O'Connor and Michaels, 1996). The social configuration of a classroom is also important to classroom discourse because it too can constrain the types of activities and communication that take place. For example, a study of a seventh-grade mathematics classroom found that after whole-class discussions, forming small-scale problem-solving groups afforded opportunities for peer-to-peer teaching (Enyedy, 2003). This is a common type of informal learning that occurs in classrooms and is largely outside a teacher's control, which may have an eventual impact on the rate and attainment of AEL for some ELL students. Opportunities to learn or missed opportunities to learn, as these cases may be, could impact AEL acquisition along with academic achievement.

Assessing Academic English Language: The Next Steps

It is critical that we have a greater understanding of the nature of language—its complexity and usage in academic settings—if we are to effectively teach and assess ELL students. The necessary next steps, particularly for the improved assessment of AEL, are documented in the ensuing chapters. The current chapter has reviewed the language and discourse literature to help determine what we know already of AEL and what aspects we must research further. From this review we learn a lot about AEL, but we also learn that it will be necessary to create a more complete picture both across the grade spans and as regards the entire school curriculum. We learn in addition that we must render this and all new information into a

usable form for educational applications. Particularly in the case of our main focus, AEL assessment, this will need to be in the form of detailed specifications about the lexical, syntactic, and discourse demands at different grade levels and in different content areas.

The next chapter will document the first empirical studies that went into a *needs analysis* to show the mismatch between the language skills that have traditionally appeared on ELD assessments and the language demands that are truly expected in school. Chapter 3 details the next critical step—government policy to hold states accountable for the development of English language proficiency and to guide state efforts in the creation of assessments better aligned to the language demands of school. The remaining steps described in chapters 4 to 7 involve the development of a conceptual framework within which to conduct empirical research of AEL, devising state-of-the-art test development procedures, and extending the AEL construct into classroom-based assessment practices. Always an iterative process, the final step in this test development process, and the subject of chapter 8, is to ask: What new steps must we then still take?

Notes

1. This brief review is designed to provide an illustration of the range of test development efforts in the K-12 arena and is not intended to be exhaustive of all the tests that have been recently developed or are currently under development. At the time of writing, other well-known test development initiatives include the New York State English as a Second Language Achievement Test (NYSESLAT, Harcourt Assessment, 2005), commissioned by the New York State Education Department; the Test of Emerging Academic English (TEAE, MetriTech, 2001); and the Test of Emerging Academic Listening and Speaking (TEALS, Lidget Green and Language Learning Solutions, in preparation), developed by the State of Minnesota; the Iowa Tests of English Language Learning (ITELL, University of Iowa, in preparation), developed by the State of Iowa Department of Education and designed in part to correspond to the content areas represented on the Iowa Test of Basic Skills (Riverside Publishing, 2001); and the Illinois Measure of Annual Growth in English (IMAGE, Illinois State Board of Education, Division of Standards and Assessment, 2005). This final test was first created to assess the English reading and writing skills of ELL students until they are sufficiently proficient in English to meaningfully take the Illinois aca-

demic achievement tests. (For further details, see also www.nysed.gov, education.state.mn.us, www.state.ia.us, and www.ccsd15.net.)

2. These are roughly analogous to tier 2 and tier 3 words in the taxonomy devised by Beck, McKown, and Kucan (2002) for instructional purposes.

3. For examples of these syntactic structures, see table 5.2 in Bailey et al., chapter 5.

References

Abedi, J., Courtney, M., and Leon, S. (2001). *Language accommodations for large-scale assessment in science* (Final Deliverable to OERI, Contract No. R305B960002). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Abedi, J., and Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education*, 14(3), 219–234.

Abedi, J., Lord, C., Hofstetter, C., and Baker, E. (2000). Impact of accommodation strategies on English language learners' test performance. *Educational Measurement: Issues and Practice*, 19(3), 16–26.

Abedi, J., Lord, C., and Plummer, J. (1997). *Final report of language background as a variable in NAEP mathematics performance* (CSE Tech. Rep. No. 429). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Allen, R. V. (1976). *Language experience activities.* Boston: Houghton Mifflin.

August, D., and Hakuta, K. (1997). *Improving schooling for language-minority children: A research agenda.* Washington DC: National Academy Press.

Bailey A. L. (in press). From Lambie to Lambaste: The conceptualization, operationalization, and use of academic language in the assessment of ELL students. In K. Rolstad (Ed.), *Rethinking school language.* Mahwah, NJ: LEA.

Bailey, A. L., and Butler, F. A. (2002/3). *An evidentiary framework for operationalizing academic language for broad application to K-12 education: A design document* (CSE Tech. Rep. No. 611). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Bailey, A. L., and Butler, F. A. (2004). Ethical considerations in the assessment of the language and content knowledge of English language learners, K-12. *Language Assessment Quarterly*, 1, 177–193.

Bailey, A. L., and Butler, F. A. (2005, April). *ELD standards linkage and test alignment under Title III.* Paper presented at the American Educational Research Association Annual Conference, Montreal, Canada.

Bailey, A. L., Butler, F. A., LaFramenta, C., and Ong, C. (2001/4). *Towards the characterization of academic language in upper elementary science*

classrooms (CSE Tech. Rep. No. 621). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Bailey, A. L., and Luoma, S. (2003, July). *Defining the construct "academic language" through empirical research.* Paper presented at the Language Testing Research Conference, Reading, UK.

Bailey, A. L., Luoma, S., and Cho, Y. (2005). *Pre-K-K IDEA proficiency test.* Brea, CA: Ballard and Tighe.

Ballard and Tighe Publishers. (2005). *IPT/PEM (IDEA proficiency test/Pearson Educational measurement) Title III testing solution.* Brea, CA: Author.

Bazerman, C. (2003, May). *What does it mean to think, read, and write analytically?* Berkeley: University of California Literacy Consortium.

Beck, I. L., McKown, M. G., and Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction.* New York: Guilford.

Bever, T. (1970). The cognitive basis for linguistic structure. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 279–353). New York: John Wiley.

Botel, M., and Granowsky, A. (1974). A formula for measuring syntactic complexity: A directional effort. *Elementary English*, 1, 513–516.

Butler, F. A., and Stevens, R. (1997). *Accommodation strategies for English language learners on large-scale assessments: Student characteristics and other considerations* (CSE Tech. Rep. No. 448). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

California State Board of Education. (1999). *English-language development standards for California public schools: Kindergarten through grade twelve.* Sacramento: California Department of Education.

Castellon-Wellington, M. (1999). *The impact of preference for accommodations: The performance of English learners on large-scale academic achievement tests* (CSE Tech Rep. No. 524). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Cazden, C. B. (1988). Environmental assistance revisited: Variation and functional equivalence. In F. Kessel (Ed.), *The development of language and language researchers: Essays in honor of Roger Brown* (pp. 281–297). Hillsdale, NJ: Lawrence Erlbaum.

Cazden, C. (2001). *Classroom discourse: The language of teaching and learning* (2nd ed.). Portsmouth, NH: Heinemann.

CCSSO (2005). *The English Language Development Assessment.* Retrieved December 1, 2005, from http://www.ccsso.org/projects/SCASS/Projects/Assessing_Limited_English_Proficient_Students/

Celce-Murcia, M., and Larsen-Freeman, D. (1983). *The grammar book: An ESL/EFL teacher's course.* Rowley, MA: Newbury House.

Chamot, A. U. (2005). The Cognitive Academic Language Learning Approach (CALLA): An update. In P. Richard-Amato and M. A. Snow (Eds.), *Academic Success for English Language Learners* (pp. 87–102). White Plains, NY: Longman.

Chamot, A. U., and O'Malley, J. M. (1994). The CALLA handbook: Implementing the cognitive academic language learning approach. Reading, MA: Addison-Wesley.

Cobb, P., Stephan, M., McClain, K., and Gravemeijer, K. (2001). Participating in classroom mathematical practices. *Journal of the Learning Sciences, 10*, 113–163.

Cocking, R. R., and McHale, S. (1981). A comparative study of the use of pictures and objects in assessing children's receptive and productive language. *Journal of Child Language, 8*, 1–13.

Collier, V. (1987). Age and rate of acquisition of second language for academic purposes. *TESOL Quarterly, 21*(4), 617–641.

Collier, V. (1989). How long? A synthesis of research on academic achievement in a second language. *TESOL Quarterly, 23*(3), 509–531.

Collier, V. (1995). *Acquiring a second language for school: Directions in Language and Education*, 1:4. Washington, DC: National Clearinghouse for Bilingual Education.

Collier, V., and Thomas, W. (1989, Fall). How quickly can immigrants become proficient in school English? *Journal of Educational Issues of Language Minority Students, 5*, 26–38.

Corson, D. (1997). The learning and use of academic English words. *Language Learning, 47*(4), 671–718.

CTB/McGraw-Hill. (2004). *California English language development test: Form D*. Monterey, CA: Author.

Cummins, J. (1980). The construct of proficiency in bilingual education. In J. E. Alatis (Ed.), *Georgetown University Round Table on Languages and Linguistics: Current issues in bilingual education, 1980* (pp. 81–103). Washington, DC: Georgetown University.

Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. In California State Department of Education (Ed.), *Schooling and language minority students: A theoretical framework* (pp. 3-49). Los Angeles: National Dissemination and Assessment Center.

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

Cunningham, J. W., and Moore, D. W. (1993). The contribution of understanding academic vocabulary to answering comprehension questions. *Journal of Reading Behaviors, 25*, 171–180.

De Avila, E. (1997). Setting expected gains for non and limited English proficient students. *NCBE Resource Collection Series No.8*. Washington, DC: National Clearinghouse for Bilingual Education.

De Avila, E. A., and Duncan, S. E. (1987, 1988, 1989, 1990). *Language assessment scales, oral administration manual, English: Forms 2c and 2d*. Monterey, CA: CTB MacMillan McGraw-Hill.

de Villiers, J., and de Villiers, P. (1973). Development of the use of word order in comprehension. *Journal of Psychological Research, 2*, 331–341.

Dewey, J. (1897). My pedagogic creed. *School Journal, 54*, 77–88.

Dickinson, D. K., McCabe, A., Anastasopoulos, L., Peisner-Feinberg, E. S., and Poe, M. D. (2003). The comprehensive language approach to early literacy: The interrelationships among vocabulary, phonological sensitivity, and print knowledge among preschool-aged children. *Journal of Educational Psychology, 95*(3), 465–481.

Douglas, D. (2000). *Assessing language for specific purposes*. Cambridge: Cambridge University Press.

Duncan, S. E., and De Avila, E. A. (1990). *Language assessment scales (LAS) reading and writing English: Forms 1a and 3a*. Monterey, CA: CTB/McGraw-Hill.

Enyedy, N. (2003). Knowledge construction and collective practice: At the intersection of learning, talk, and social configurations in a computer-mediated mathematics classroom. *Journal of the Learning Sciences, 12*(3), 361–407.

Erickson, F. (1987). Transformation and school success: The politics and culture of educational achievement. *Anthropology and Education Quarterly, 18*(4), 335–356.

Erickson, F. (1992). Ethnographic microanalysis of interaction. In M. D. LeCompte, W. L. Millroy, and J. Preissle (Eds.), *The handbook of qualitative research in education* (pp. 201-225). Burlington, MA: Academic Press.

Fraser, C., Bellugi, U., and Brown, R. (1963). Control of grammar in imitation, comprehension, and production. *Journal of Verbal Learning and Verbal Behavior, 2*, 121–135.

Garcia, G. R. (1991). Factors influencing the English reading test performances of Spanish-speaking Hispanic children. *Reading Research Quarterly, 26*, 371–391.

Gee, J. P. (1996). *Social linguistics and literacies: Ideology in discourse*. London: Falmer Press.

Gibbons, P. (1998). Classroom talk and the learning of new registers in a second language. *Language and Education, 12*(2), 99–118.

Goffman, E. (1981). *Forms of talk*. Philadelphia: University of Pennsylvania Press.

Greeno, J. G., and Hall, R. P. (1997). Practicing representation: Learning with and about representational forms. *Phi Delta Kappan*, 361–367.

Gumperz, J. J., and Cook-Gumperz, J. (1980). *Beyond ethnography: Some uses of sociolinguistics for understanding classroom environments*. Los Angeles: National Dissemination and Assessment Center, California State University, Los Angeles.

Gutiérrez, K. D. (1994). How talk, context, and script shape contexts for learning: A cross-case comparison of journal sharing. *Linguistics and Education, 5,* 335–365.

Gutiérrez, K. (1995). Unpackaging academic discourse. *Discourse Processes, 19*(1), 21–37.

Hakuta, K., Butler, Y., and Witt, D. (2000). *How long does it take English learners to attain proficiency?* Santa Barbara, CA: UC Linguistic Minority Research Institute.

Haladyna, T. M., and Downing, S. M. (2004). Construct-irrelevant variance in high-stakes testing. *Educational Measurement: Issues and Practice, 23,* 17–27.

Halliday, M. A. K., and Martin, J. R. (1993). *Writing science: Literacy and discursive power.* Pittsburgh: University of Pittsburgh Press.

Harcourt Assessment Inc. (2005). *New York State English as a Second Language Achievement Test (NYSESLAT).* San Antonio, TX: Author.

Harcourt Educational Measurement (2003). *Stanford English language proficiency test (Stanford ELP).* San Antonio, TX: Author.

Hicks, D. (1994). Individual and social meanings in the classroom: Narrative discourse as a boundary phenomenon. *Journal of Narrative and Life History, 4*(3), 215–240.

Hunt, K. W. (1965). *Grammatical structures written at three grade levels* (Research Report No. 3). Urbana IL: National Council of Teachers of English.

Hunt, K. W. (1977). Early blooming and late blooming syntactic structures. In C. R. Copper and L. Odell (Eds.), *Evaluating writing: Describing, measuring, judging.* Urbana, IL: National Council of Teachers of English.

Hymes, D. (1974). *Foundations in sociolinguistics: An ethnographic approach.* Philadelphia: University of Pennsylvania Press.

Illinois State Board of Education, Division of Standards and Assessment. (2005). *Illinois measure of annual growth in English (IMAGE).* Springfield, IL: Author.

Johns, A. M. (1997). *Text, role, and context: Developing academic literacies.* Cambridge: Cambridge University Press.

Kazami, E. (1999, Spring). Mathematical discourse that promotes conceptual understanding. *Connections.* Los Angeles: UC Regents.

Kemper, S., Jackson, J. O., Cheung, H., and Anagnopoulous, C. A. (1993). Enhancing older adults' reading comprehension. *Discourse Processes, 16,* 405-428.

Kern, R. (2000). *Literacy and language teaching.* Oxford: Oxford University Press.

Kopriva, R. (2000). *Ensuring accuracy in testing for English language learners.* Washington, DC: Council of Chief State School Officers.

Labov, W. (1972). *Language in the inner city: Studies in the black English vernacular.* Philadelphia: University of Pennsylvania Press.

LaCelle-Peterson, M., and Rivera, C. (1994). Is it real for all kids? A framework for equitable assessment policies for English language learners. *Harvard Educational Review, 64* (1), 55–75.

Lampert, M., and Ball, D. (1998). *Teaching, multimedia, and mathematics: Investigations of real practice.* New York: Teachers College Press.

Larson-Freeman, D. (2000). Second language acquisition and applied linguistics. *Annual Review of Applied Linguistics, 20,* 165–181.

Lave, J., and Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge: Cambridge University Press.

Lee, O. (2001). Culture and language in science education: What do we know and what do we need to know? *Journal of Research in Science Teaching, 38*(5), 499–501.

Lehrer, R., Randle, L., and Sancilio, L. (1989). Learning preproof geometry with LOGO. *Cognition and Instruction, 6*(2), 159–184.

Lehrer, R., Schauble, L., Carpenter, S., and Penner, D. (2000). The interrelated development of inscriptions and conceptual understanding. In P. Cobb, E. Yackel, and K. McClain (Eds.), *Symbolizing and communicating in mathematics classrooms: Perspectives on discourse, tools, and instructional design* (pp. 325–360). Mahwah, NJ: Lawrence Erlbaum.

Lemke, J. (1990). *Talking science: Language, learning, and values.* Norwood, NJ: Ablex. (ERIC Document Reproduction Service No. Ed362379)

Lidget Green and Language Learning Solutions. (In preparation). *Test of emerging academic listening and speaking (TEALS).* Eugene, OR: Author.

Lord, C. (2002). Are subordinate clauses more difficult? In J. Bybee and M. Noonan (Eds.), *Complex sentences in grammar and discourse: Essays in honor of Sandra A. Thompson.* Philadelphia: John Benjamins.

Luoma, S., Cho, Y., and Buck, G. (2004, March). *From needs analysis to test development in a test for ELLs.* Paper presented at the Language Testing Research Colloquium, Temecula, CA.

McLaughlin, B., August, D., and Snow, C. (2000). *Vocabulary knowledge and reading comprehension in English language learners* (Final Performance Report to OERI, Award No. R306F60077–97). Washington, DC: U.S. Department of Education.

Mehan, H. (1979). *Learning lessons: Social organization in the classroom.* Cambridge, MA: Harvard University Press.

Menyuk, P. (1995). Language development and education. *Journal of Education, 177*(1), 39–62.

MetriTech Inc. (2001). *Test of emerging academic English (TEAE).* Champaign, IL: Author.

Nation, I. S. P. (2001). *Learning vocabulary in another language.* Cambridge: Cambridge University Press.

Nation, I. S. P., and Coxhead, A. (2001). The specialised vocabulary of English for academic purposes. In J. Flowerdew and M. Peacock (Eds.),

Research perspectives on English for academic purposes (pp. 252–267). Cambridge: Cambridge University Press.

National Educational Goals Panel. (n.d.). *Reconsidering children's early development and learning: Toward common views and vocabulary.* Retrieved June 20, 2004, from http://www.negp.gov/Reports/child-ea.htm.

Nippold, M. A. (1995). Language norms in school-age children and adolescents: An introduction. *Language, Speech, and Hearing Services in Schools, 26,* 307–308.

No Child Left Behind Act. (2001). Pub. L. No. 107-110, 115 Stat. 1425, December 13.

O'Connor, M. C., and Michaels, S. (1996). Shifting participant frameworks: Orchestrating thinking practices in group discussion. In D. Hicks (Ed.), *Discourse, learning, and schooling* (pp. 63–103). New York: Cambridge University Press.

Olson, L. (2002, December). States scramble to rewrite language-proficiency exams. *Education Week.* Available from http://www.edweek.com/ew/ew_printstory.cfm?slug=14lep.h22.

Palinscar, A., and Brown, A. (1984). Reciprocal teaching of comprehension-fostering and comprehension monitoring activities. *Cognition and Instruction, 1*(2), 117–175.

Pauley, A., and Syder, F. H. (1983). Natural selection in syntax: Notes on adaptive variation and change in vernacular and literary grammar. *Journal of Pragmatics, 7,* 551–579.

Philips, S. U. (1972). Participant structures and communicative competence: Warm Springs children in community and classroom. In C. B. Cazden, V. P. John, and D. Hymes, (Eds.), *Functions of language in the classroom* (pp. 370–394). New York: Teachers College Press.

Pimm, D. (1987). *Speaking mathematically: Communication in mathematics classrooms.* London: Routledge and Kegan Paul.

Pimm, D. (1995). *Symbols and meanings in school mathematics.* London: Routledge.

Reppen, R. (2001). Register variation in student and adult speech and writing. In S. Conrad and D. Biber (Eds.), *Variation in English: Multi-dimensional studies* (pp. 187–199). Harlow, Essex, UK: Pearson Education.

Rivera, C., and Stansfield, C. W. (2001, April). *The effects of linguistic simplification of science test items on performance of limited English proficient and monolingual English-speaking students.* Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

Rivera, C., Stansfield, C., Scialdone, L., and Sharkey, M. (2000). *An analysis of state policies for the inclusion and accommodation of English language learners in state assessment programs during 1998–1999.* Arlington, VA: George Washington University, Center for Equity and Excellence in Education.

Riverside Publishing. (2001). *Iowa tests of basic skills (ITBS).* Itasca, IL: Author.

Roth, W. M. (2001). Situating cognition. *Journal of the Learning Sciences, 10,* 27–61.

Roth, W. M., and Bowen, G. M. (1995). Knowing and interacting: A study of culture, practices, and resources in a Grade 8 open-inquiry science classroom guided by a cognitive apprenticeship metaphor. *Cognition and Instruction, 13*(1), 73–128.

Sato, E., Lagunoff, R., Worth, P., Bailey, A. L., and Butler, F. A. (2005). *ELD standards linkage and test alignment under Title III: A pilot study of the CELDT and the California ELD and content standards.* Deliverable to California Department of Education.

Scarcella, R. (2003). *Academic English: A conceptual framework* (Tech. Rep. No. 2003-1). Santa Barbara UC Linguistic Minority Research Institute.

Scarcella, R., and Zimmerman, C. (1998). Academic words and gender: ESL student performance on a test of academic lexicon. *Studies in Second Language Acquisition, 20,* 27–49.

Schleppegrell, M. J. (2001). Linguistic features of the language of schooling. *Linguistics and Education, 12*(4), 431–459.

Schleppegrell, M. J. (2004). *Language of schooling: A functional linguistics perspective.* Mahwah, NJ: Lawrence Erlbaum.

Short, D. J. (1993). *Integrating language and culture in middle school American history classes* (Report to OERI). Washington, DC: Center for Applied Linguistics and the National Center for Research on Cultural Diversity and Second Language Learning.

Short, D. (1994). Expanding middle school horizons: Integrating language, culture, and social studies. *TESOL Quarterly, 28*(3), 581–608.

Smagorinsky, P. (2001). If meaning is constructed, what is it made from? Toward a cultural theory of reading. *Review of Educational Research, 17*(1), 133–169.

Snow, C. (1990). The development of definitional skill. *Journal of Child Language, 3,* 697–710.

Snow, C. (1991). Diverse conversational contexts for the acquisition of various language skills. In J. Miller (Ed.), *Research on child language disorders* (pp. 105–124). Austin, TX: Pro-Ed.

Spanos, G., Rhodes, N. C., Dale, T. C., and Crandall, J. (1988). Linguistic features of mathematical problem solving: Insights and applications. In R. R. Cocking and J. P. Mestre (Eds.), *Linguistic and cultural influences on learning mathematics* (pp. 221–240). Hillsdale, NJ: Erlbaum.

Stack, J. (2002, September). *California English language development test (CELDT): Language proficiency and academic achievement.* Paper presented at the annual conference of the National Center for the Study of Evaluation, Standards, and Student Testing, UCLA, Los Angeles.

Stauffer, R. (1980). *The language-experience approach to the teaching of reading.* New York: Harper Row.

Stevens, R. A., Butler, F. A., and Castellon-Wellington, M. (2000). *Academic language and content assessment: Measuring the progress of ELLs* (Final Deliverable to OERI, Contract No. R305B960002). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

TESOL. (1997). *ESL standards for pre-K-12 students.* Alexandria, VA: Author.

University of Iowa. (In preparation). *Iowa tests of English language learning (ITELL).* Iowa City: Author.

U.S. Department of Education. (2005). *FY 2002–2004 biennial report to Congress on the implementation of NCLB, Title III, the State Formula Grant Program.* Washington, D.C. Retrieved March 31, 2005, from http://www.ncela.gwu.edu/oela/biennial05/index.htm.

Wang, M. D. (1970). The role of syntactic complexity as a determiner of comprehensibility. *Journal of Verbal Learning and Verbal Behavior, 9,* 398–404.

White, B. Y., and Frederiksen, J. R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction, 16*(1), 3–118.

WIDA Consortium. (2005). Retrieved January 31, 2005, from http://www.wida.us.

CHAPTER 2

ELLs and Standardized Assessments: The Interaction between Language Proficiency and Performance on Standardized Tests

FRANCES A. BUTLER, ROBIN STEVENS, AND MARTHA CASTELLON

Within the context of accountability at the local and state levels, the growing need during the past decade to assess both the English language proficiency of English language learners (ELLs) and their understanding of subject matter across content areas has created challenges for educators. In this chapter, three critical issues regarding the assessment of ELLs are presented and discussed—first, the heterogeneity of ELLs, that is, the notion that ELLs are not all alike and that awareness and documentation of their differing needs should be a part of any decision-making process; second, the mismatch between the language assessed on English language proficiency tests and the language used on content tests; and third, the importance of operationalizing the language of school (academic English) for test development purposes and for broad educational applications. Data from two studies (Butler and Castellon-Wellington, 2000/5; Stevens, Butler, and Castellon-Wellington, 2000) that examine student concurrent performance on a test of English proficiency and tests of academic achievement inform the discussion. These studies were the first to empirically demonstrate the interplay between language and content assessments, validating researchers' and

Abbreviations

AEL academic English language
AELP Academic English Language Proficiency
AERA American Educational Research Association
ALADIN Academic Language Assessment and Development of Individual Needs
ALEC Academic Language Exposure Checklist
AMAO Annual Measurable Achievement Objective
ANOVA analysis of variance
APA American Psychological Association
AYP adequate yearly progress
BICS basic interpersonal communication skills
CAHSEE California High School Exit Examination
CAL Center for Applied Linguistics
CALLA Cognitive Academic Language Learning Approach
CALP cognitive academic language proficiency
CAPA California Alternate Performance Assessment
CELDT California English Language Development Test
CHILDES Child Language Data Exchange System
CI confidence interval
CLAN Computer Language Analysis
CRESST National Center for Research on Evaluation, Standards, and Student Testing
CST California Standards Test
EFL English as a foreign language
EL English learner
ELA English language arts
ELD English language development
ELDA English Language Development Assessment

ELL English language learner
ELP English language proficiency
EO English-only
ESEA Elementary and Secondary Education Act
ESL English as a second language
FEP fluent English proficient
HLM Hierarchical Linear Modeling
IASA Improving America's School Act
IFEP initial fluent English proficient
IMAGE Illinois Measure of Annual Growth in English
IRE/F initiate-respond-evaluate/feedback
IPT IDEA Proficiency Test
ITBS Iowa Test of Basic Skills
ITELL Iowa Tests of English Language Learning
LAS Language Assessment Scales
LEA local educational agency
LEP limited English proficient
NCE normal curve equivalent
NCELA National Clearinghouse for English Language Acquisiti
NCLB No Child left Behind Act of 2001
NCME National Council on Measurement in Education
NLLIA National Languages and Literacy Institute of Australia
NRC National Research Council
NYSESLAT New York State English as a Second Language Achieve ment Test
OELA Office of English Language Acquisition
OTL opportunity to learn
RE reverse engineering
RFEP redesignated (or reclassified) fluent English proficient
SDAIE Specially Designed Academic Instruction in English
SIOP Sheltered Instruction Observation Protocol
SLA second language acquisition
TEAE Test of Emerging Academic English
TEALS Test of Emerging Academic Listening and Speaking
TESOL Teachers of English to Speakers of Other Languages
WIDA Wisconsin, Delaware, and Arkansas (Assessment Consortiu
ZPD zone of proximal development