Focus on Meaning Structure Matters for English Learners: Exploring Opportunity to Learn Academic Language

Zenaida Aguirre-Munoz
College of Education,
Texas Tech University

Abstract

Past research has demonstrated the importance of opportunity to learn (OTL) variables in explaining student achievement. However few OTL studies have included English learners (ELs). This exploratory study used a mixed-method approach to investigate the relationship between learning opportunities of middle school ELs and their writing assessment performance. The primary OTL indicator explored in this study was academic language instruction and its' relationship to ELs' assessment performance. Quantitative results suggest that academic language instruction is an important OTL indicator for ELs; it was the only classroom level indicator consistently related to writing performance after controlling for classroom and school effects. Implications are discussed and suggestions for future research are offered.

Keywords: English language learners, at-risk students, adolescence, assessment, accountability, pedagogy, writing performance

Focus on Meaning Structure Matters for English Learners: Exploring Opportunity to Learn Academic Language

Substantial empirical evidence has demonstrated the importance of opportunity to learn (OTL) variables in explaining student achievement (e.g., Authors, 2005; Brophy & Good, 1986; Schmidt, Cogan, & McKnight, 2010; Stevenson & Stigler, 1992). OTL information can serve to document a school's provision of educational opportunities and provide detailed explanatory information regarding student achievement (Authors, 2008; Porter, 1991). However, most of this research has not examined directly the

appropriateness of applying OTL indicators, investigated predominately with majority group student samples, to linguistically diverse student populations. From the perspective of OTL, the population of English learners (ELs) poses important and complex instructional challenges that demand direct attention (First Author, 2010). The purpose of the study was to: collect descriptive information regarding the relationship between academic language instruction and EL achievement; examine other OTL indicators relevant for ELs (e.g., sheltered instruction); and use what is learned to inform research targeting effective learning environments for ELs.

Opportunity to Learn and ELs

Differences in OTL have repeatedly been associated with background characteristics such as ethnicity, gender, and (more recently) language background (see Heartal, Moss, Pullen, & Gee, 2008). For example, Guiton and Oakes (1995) found that classes predominantly composed of White and Asian students had higher levels on all of their indicators of teacher quality (teacher experience, education, and assignment) than mixed or predominantly minority classes. That is, as the minority composition in classes increased, sustained OTL dropped precipitously. These researchers also found that regardless of students' initial achievement level, those who were placed in lower level courses obtained smaller gains over time than students of comparable achievement who were placed in higher level courses. Other studies reveal that ELs are consistently tracked into courses that limit their OTL (e.g., Callahan, 2005), sometimes under the guise of differentiated instruction (First Author, 2003). These findings underscore the discrepancies in OTL that arise from tracking practices that can lead to discrepancies in learning outcomes.

The national response to these disparities has been to focus on performance outcomes (Shepard, 2009). Yet, more recent OTL research suggests that an emphasis on performance outcomes does not guarantee increased equity in the distribution of student learning, particularly for ELs (e.g., Authors, 2008; Lee & Wong, 2004). For example, Authors (2008) investigated the impact of the introduction of a standards-based, writing performance measure on teacher practice and student outcomes. Although teachers increased their standards-based instruction and student writing mean scores improved overall, the learning opportunities that resulted from increased exposure were unequally distributed as indicated by the increase in the achievement gap based on language background. Authors (2008) found that instruction was not individualized to students' specific linguistic needs and may have contributed to the disparities in achievement between ELs and non-ELs. This study demonstrated that an increase in content exposure alone may not create equitable OTL; ELs did not appear to be able to take advantage of such an increase. The results of these studies also raises questions about the effectiveness of learning environments that do not provide instructional experiences that simultaneously address ELs' linguistic, cultural, and content learning needs. Given these trends, it is reasonable to argue that increasing curricular access may, at a minimum, require linguistic accommodations during instruction in order to increase ELs' opportunity to engage academic tasks productively.

Academic language instruction. Increasingly researchers direct attention to the lack of instructional focus on academic language in explaining ELs' achievement disparities (e.g., First Author, 2008; Gee, 2008). Academic language refers to the language of school-based tasks and activities whose linguistic features contrasts with the language students encounter outside of school (Cummins, 1979, 1984). To describe academic language, we draw on systemic functional linguistics (SFL) (Halliday, 1975). From an SFL perspective meaning is (re) constructed by analyzing and discussing the connections between form and function in academic tasks. SFL examines how linguistic meaning structures are used to (a) represent experience and involves the representation of ideas, activities, and surrounding circumstances; (b) negotiate relationships and conveys to the reader (or the listener) opinions and attitudes; and (c) organize information into a coherent and cohesive text for a given genre and context. Academic language is therefore described here in terms of the lexical and grammatical choices that are made to create different types of meanings in ways that are acceptable in school settings.

Corpus-based studies in applied linguistics demonstrate the difficulty ELs have in attaining command of academic language and suggest that exposure to focused instruction, from an SFL perspective, leads to improved awareness and command of academic language in various content areas (e.g., Gebhard, Harman, & Seger, 2007; Gibbons, 2003; Schleppegrell, 2003). In addition, Schleppegrell, (2004) demonstrated that unlike conversational language, academic language is not learned naturally. Therefore, ELs' limited knowledge of English may make it more challenging for them to gain from undifferentiated exposure to academic language (Authors, 2008). Thus, we argue that a focus on linguistic meaning structures may provide differentiation to AL instruction for ELs.

Linguistic and cultural factors. This body of work has demonstrated the need to expand the discussion of OTL to include linguistic and cultural factors that may explain the persistent gap in achievement that is observed even after ELs achieve English proficiency. Therefore, we argue that the strategies schools and teachers use to implement the intended curriculum ought to be contextualized to the instructional needs of the EL population. Although general strategies appear to benefit all students (i.e., scaffolding, tapping into prior knowledge etc.) the manner in which these strategies are delivered may need to be fundamentally different for ELs than native English speaking students. Unfortunately, few empirical studies have examined effective instructional strategies and practices for ELs, outside of basic literacy development at the elementary school level. To address this research gap, this study also investigated OTL indicators consistent with those proposed by First Author (2010) in an OTL model intended to better reflect the instructional needs of ELs. Included in that model are two general domains (content and teaching quality) that describe the characteristics of instruction contextualized to ELs as well as modifiers of instructional characteristics intended to address additional adjustments ELs may need to benefit from increased OTL (August & Shanahan, 2006).

Curriculum & teaching quality. Curriculum quality indicators include exposure to core content topics, content-based academic language as well as opportunities to develop higher-order cognitive functions. Teaching quality refers to teachers' understanding of the content in relation to what is known to be effective for ELs (Darling-Hammond, 2000). That is, teacher's ability to deliver content to ELs that results in appropriate curriculum access and development of their knowledge systems. Curriculum access includes strategies that have documented evidence in providing ELs with increased access to the curriculum, such as scaffolded instruction (e.g., Walqui, 2006), sheltered content instruction (e.g., Echevarria, Short, & Powers 2006), and explicit instruction (e.g., Greenleaf & Freedman, 1993; Linan-Thompson & Vaughn, 2007). These instructional activities foster development of deep understanding of the content and language, including opportunities for extended discourse, group work, and feedback to students through informal or formal classroom assessments.

Modifiers of instructional characteristics. An important subset of teaching quality, are three sets of additional factors—modifiers of instructional characteristics—that impact the kinds of affordances students receive in a given classroom. These factors help explain the extent to which students can take advantage of the opportunities provided in the learning context. These factors include student characteristics (e.g., EL status), teacher characteristics (e.g., knowledge of linguistic features of AL), and course-specific resources (e.g, use of realia, first language dictionaries etc.).

With this model in mind, the specific research questions investigated were:

1. What student, classroom, and school-level factors are significantly related to writing performance?
2. After controlling for student- and school-level factors, to what extent does academic writing performance vary within classrooms as a function of academic language instruction? Is there a differential impact of AL instruction between ELs and non-ELs?
3. After controlling for student- and school-level factors, to what extent does academic writing performance vary within classrooms as a function of other OTL factors? Does the relationship between AL and performance change when other OTL indicators are examined?

Method

Participants

Teachers. A sample of 32 language arts teachers from three urban middle schools in Southern California participated in the study. Therefore, 21 out of 32 teachers were selected by district and school administrative personnel to be trained on instructional strategies to incorporate academic language instruction in classrooms. The remaining eleven teachers who participated in the study were not trained in AL instruction. Randomization was not possible due to policy constraints and priorities in the districts involved in the study. The trained group was over-sampled due to potential attrition as well as the potential for low implementation by trained teachers. Since this was an exploratory study and not an efficacy study,

training was limited in scope and follow-up training was minimal. Although a study limitation, randomization is not necessary in exploratory studies because the intent was not to establish cause-and-effect relationships. The intent is to begin to gather empirical information about the relationship between AL instruction and EL assessment scores to inform future studies.

The total years of teaching experience for all teachers ranged from 1 to 27 years with an average of 10 years. In general, trained and untrained teachers were very similar in terms of their teaching experience. Untrained teachers did have a slightly higher average number of years teaching sheltered English than the trained teachers (5.33 and 3.27 for comparison and trained teachers respectively). Although we only recruited language arts teachers, the majority of the teachers did not major in English as undergraduates. Twenty-four out of thirty-two teachers held teaching credentials.

Students. A total of 1,646 middle school students enrolled in language arts classes completed the Language Arts Performance Assignment (LAPA) at the end of the spring semester. Students with California English Language Development Test (CELDT) scores and a designated English Language Development (ELD) level were classified as ELs in this study. In general, the background characteristics of students from trained and untrained teachers were very similar. The proportion of EL students was slightly higher in the trained group (57%) as compared to untrained teachers (43%). This difference reflected the districts' commitment to serve the greatest amount of ELs as possible. Ninety percent of the student sample is identified as Hispanic; 60% of the sample was designated as ELs at the time of the study.

Procedures and Instruments

Teacher Training on Academic Language Instruction

The teacher training has been described elsewhere (see First Author, 2008), therefore, we only briefly describe it here. The training focused on SFL-based concepts corresponding to writing skills. The goal of the five-day training was to move teachers from a traditional approach to grammar instruction to one focused on meaning construction wherein linguistic elements are made transparent for ELs. The first two days targeted linguistic meaning structures that correspond to response to literature tasks in teacher-friendly terms. The remaining time focused on analysis of student writing and the collaborative development of AL lessons. Evaluation of the teacher training revealed that the training was effective in building teacher knowledge of the use of SFL in identifying areas of instructional support for ELs, and in providing feedback to students on their written academic discourse skills. Post-test scores on a teacher performance assessment were significantly higher than pre-test scores (p's $< .05$). Trained teachers were also more likely to focus instruction on clarifying meaning than untrained teachers (see First Author, 2008).

Teacher Opportunity to Learn Survey

Teachers completed a teacher survey intended to capture critical aspects of OTL indicators targeted in this study. The survey targeted two curriculum quality domains: (a) academic language coverage, and (b) English language arts (ELA) content coverage; two teaching quality domains: (a) curriculum access; (b)

knowledge development; and two variables in the modifiers of instructional practice domain: (a) teacher expertise, and (b) teacher training. Teacher reported curriculum quality and teaching quality on 6-point scales. Table 1 presents the constructs targeted by the survey items and a brief description.

Table 1 Survey Constructs and Definitions

Construct	Definition	Scale
Curriculum Quality Domain		
Academic Language Coverage	Amount of instruction on grammatical features of school-based writing genres such as "verbs choices that signal analysis of a character or situation"	0-5
English Language Arts Coverage	Amount of time spent on learning or doing activities related to literary analysis such as "summarizing the plot of novels" and "writing about heroic qualities of characters"	0-5
Teaching Quality Domain		
Curriculum Access	Amount of time spent using sheltered instructional strategies such as links to background knowledge and scaffolding	0-5
Knowledge Development	Amount of time spent on extended discourse and providing feedback to students	0-5
Modifiers of Instructional Practice Domain		
Credential Status	Whether or not teacher holds a credential	0=No 1=Yes
Teacher Training	Number of college level English Language Arts courses	Count
Post Graduate Degree	Whether or not teachers completed a graduate degree	0=No 1=Yes

The reliability of the items was evaluated using the alpha coefficient. Alpha coefficients for the five constructs (excluding the years of teaching experience) ranged from 0.86 to 0.95. These alphas provide evidence of strong to high reliability across the five constructs. Confirmatory factor analysis (CFA) was also conducted to test whether the items are sufficiently representing the specified constructs. The results of the CFA suggest that the items in general appear to be adequately measuring the proposed constructs. The good model fit indices (CFI's > .90) and high factor loading all indicate that the items have high internal consistency and factorial validity of the constructs (Authors, 2011). Although reliability was high and the

CFA indicated that the constructs had some degree of validity, we found evidence of a teacher background effect on the interpretation of academic knowledge coverage items as wells as a social desirability effect (Authors, 2011). Therefore, caution should be taken when interpreting the results of the academic language coverage construct based on the survey items.

Qualitative Data on Classroom Practice

In addition to administering the teacher surveys, we observed classrooms of participating teachers at two different time points and conducted detailed interviews regarding teacher practice to gather OTL data on two constructs not included in previous OTL studies: academic language coverage and curriculum access strategies. Teachers were observed for two consecutive lessons. After the observations, teachers were interviewed to discuss the planning for the observed lesson, their judgments on their ability to achieve the lesson objectives, and how typical the lesson was in comparison to other lessons delivered throughout the year. To achieve high inter-rater reliability on the observation protocol, an observer calibration session was conducted prior to site visits that involved the use of video-taped lessons, discussions around key constructs, and practice observations at one of the pilot-testing sites. Percent exact score agreement on the seven items ranged from 76% to 95% and kappa coefficients ranged from .79 to .93 indicating strong inter-rater reliability.

AL implementation. AL implementation refers to the amount and quality of sustained exposure to meaning-based AL instruction. Both teacher interview and observation data were used to make a global judgment regarding the quality of academic language instruction on a four-point scale, ranging from 0 to 3. This score was used in the analysis to capture both quantity and quality. While the interview data served as the primary data source for this dimension, two items on the observation protocol were used to verify the accuracy of these reports of instruction and scores were lowered when observations did not coincide with teacher descriptions of their practice. Two raters made judgments on this dimension; one was blind to which group (trained or untrained) the teacher belonged and one was not. Kappa coefficients for this dimension was .80, suggesting they were consistent in their judgments of AL implementation.

Curriculum access indicators. Three sets of items targeted curriculum access practices based on the Sheltered Instruction Observation Protocol (Echavarria et. al, 2006). Three items targeted teachers' expectations, defined in terms of learning objectives. Five items targeted sheltered instructional strategies, which include provision of comprehensible input; and scaffolding techniques. Four items targeted student engagement activities, such as targeted practice and application of content. Observers rated teacher practices on a 5-point, Likert-scale, such as, from 0 representing "no use of scaffolding techniques" to 4 representing "consistent use of scaffolding techniques throughout lesson." Thus, judgments pertaining to sheltered instruction, teacher expectations, and engagement strategies were made globally for an entire lesson.

Student Assessment

The outcome measures for the study were four scores on a curriculum-embedded performance assessment designed to assess student understanding and skills in English language arts. This Language Arts Performance Assignment (LAPA) was modeled after previous work and has undergone validation studies (Authors, 2005; 2008). In this assessment, students were asked to select a literary piece that contains a heroic character and describe the qualities of that character in writing, citing detailed information from the literary work. Thus, students were expected to analyze the story beyond the surface features of the plot and support all assertions about the character with accurate and supporting citations. Students were given time to complete the stages of the writing process (5-10 hours of class time over 1-2 weeks).

Scoring Student Work

A two-day training session for scoring student essays was conducted with thirteen raters (seven researchers and six middle school teachers).

Holistic dimension. Following the review of the writing task and rubric, the raters read and discussed the anchor papers that represented the performance levels of the rubric. The discussion centered on qualities of performance described in the rubric which included descriptions of the character, quality of the references from the text, overall organization, and extent of mechanical errors. Once anchor papers were discussed, and before raters began scoring student work, the raters practiced applying the scoring criteria by individually scoring sets of six "practice" papers. After scoring the practice papers individually, the raters discussed their qualities in reference to the rubric and anchor papers. Any major discrepancies were discussed thoroughly. This process continued until exact score agreement on a set of practice papers was at least 80%.

Academic language dimensions. Analytic scoring of the AL dimensions of the LAPA assessment targeted the use of the linguistic choices essential for successful characterization. They included noun phrases, linguistic structure used to present circumstances (circumstances), and linguistic structures to refer to the character (character references). Training on these AL dimensions followed the same procedure as the holistic training described above. Once rater exact score agreement on the set of practice papers reached 80% the raters began to score student responses. Kappa coefficients for these dimensions on 30 common papers were .87, .78, 74, and 88 for the holistic, noun phrases, circumstances and character reference scores respectively. These coefficients suggest adequate to strong reliability.

Analysis

Responses to the OTL survey, interview data, and classroom observation data were analyzed in concert with student performance results using eleven different two-level ordinal logistic hierarchical linear models (ordinal logistic HLM). HLM models provide a systematic way to investigate how teacher-level OTL variables influence student-level outcomes and whether these variables have any differential effects on

EL performance after adjusting for student-level variables. Further, both teacher- and student-level factors can be examined simultaneously. Given that our primary research question is related to the relationship between AL instruction and ELs' writing performance, we included AL instruction implementation in all of the models and examined this along with one of the other 11 OTL variables in any given model. Due to sample size restrictions at the teacher level, we examined the impact of the 11 OTL variables separately, with academic language implementation (the 12th OTL variable) being the common OTL variable in each model.

Student-level findings (described below) revealed systematic differences in student performance by school. Therefore, in all the models, we included schools as a covariate in order to control for initial differences across the three schools. This systematic difference was also found to be confounded by the proportion of ELs in a given classroom. Therefore, also included in the models is the proportion of ELs in classrooms to control for this effect. Thus, for each of the four LAPA scores in each of the HLM models, we looked at the effect of (a) school differences, (b) proportion of ELs, (c) AL implementation, and (d) one additional OTL variable. The OTL variable names and descriptions are presented in Table 2. Descriptive information is provided in Table 3.

Table 2

Description of Student and Teacher Level OTL Indicators Used in HLM Analyses

Indicator	Description
Student Level Indicators	
Gender	Student gender
ELL Status	Whether or not the student is an EL or a non-EL
Grade 7	Whether or not the student is in seventh grade.
Grade 8	Whether or not the student is in eighth grade
Hispanic	Whether or not the student is Hispanic
Holistic	Holistic score on the LAPA (1-4)
Noun Phrases	LAPA Score for Noun phrases (1-3)
Circumstances	LAPA Score for Circumstances (1-3)
Character References	LAPA Score for Character references (1-3)
Teacher Level OTL Indicators from Survey Measures	
Expert	Teacher content expertise (1-6)
ELA_CC	ELA content coverage (1-6)
AL_CC	Academic language reports (1-6)
SI_SR	Sheltered Instruction (1-6)
ASSES	Knowledge Development: Feedback and assessment (1-6)

Cred	Whether or not teacher hold credentials
NuC	Number of ELA courses taken (0-16)
Grad	Whether or not teacher has graduate studies
Teacher Level OTL Indicators from Qualitative Data	
ALIm	Academic language implementation (0-3)
SI_O	Sheltered Instruction implementation (0-3)
Expect	Clear content expectations (0-4)
Engage	Student engagement activities (0-4)
School Level Indicators	
Los Niños	Whether or not student attends Los Niños School
Casi	Whether or not student attends Casi Middle School

Note: Numbers in parentheses represent the range of possible scores.

Table 3

Descriptive Information for LAPA scores and Teacher (OTL) Variables

OTL Variables	N	Mean	SD
LAPA Scores			
Holistic	1,606	2.19	0.54
Noun Phrases	1,606	1.49	0.33
Circumstances	1,606	1.63	0.32
Character References	1,606	1.67	0.25
Teacher Variables			
Curriculum Quality			
1. AL implementation	30	1.27	1.23
2. ELA content coverage (ELA_CC)	32	4.68	1.09
3. AL Content Coverage (AL_CC)	32	3.23	0.88
Teaching Quality			
4. Clear content expectations (Expect)	32	2.25	0.97
5. Sheltered instruction self-report (SI_SR)	32	4.28	0.94
6. Engagement Strategies (Engage)	32	1.87	0.98
7. Sheltered instruction observed (SI_O)	30	1.53	0.90
8. Feedback & assessment (Asses)	32	3.78	1.17
Modifiers of Instructional Practice			
9. Teacher expertise (Expert)	32	4.29	0.92
10. Completed graduate studies (Grad)	32	0.66	0.48
11. Number of ELA courses (NuC)	32	5.22	4.89
12. Credential status (Cred)	32	0.75	0.44

Note: teacher variables 2, 3, 5, 8-12 are self reported data; the remaining variables were obtained from qualitative data (interview and observation data); variable labels appear in parentheses.

Results and Analysis

Relationship between Student, Classroom, and School-Level Factors and Writing Performance

Gender & Ethnicity. Gender was significantly associated with the holistic, circumstances and character references scores. Overall, female students out-performed their male counterparts. This result is consistent with literature regarding the performance gap between males and females on language arts content. Authors, 2008, for example, found a significant gap between EL males and non-EL females increased as the amount of (undifferentiated) exposure to writing instruction increased. We also found that Hispanic students performed significantly lower on the LAPA holistic score compared to other students, which is also consistent with other studies (e.g., Abedi et al., 2002).

Proportion of ELs. The proportion of ELs (indicated in the tables as Proportion EL) in the classroom was negatively associated with the noun phrases and circumstances scores (see Tables 4 & 5). Although individual student EL status was not a significant factor, classrooms with higher proportions of ELs had significantly lower performance for the noun phrases and circumstances scores. Students in classrooms with higher proportions of ELs received lower means on these dimensions compared to students in classrooms with lower proportions of ELs. Since classrooms with larger proportions of ELs were also found to have lower implementation of AL instruction than classrooms with smaller proportions of ELs, the lower performance associated with higher proportions of EL suggests that EL status may be an important indicator affecting teacher expectations of student performance, beyond that of grade level affiliation.

Table 4

HLM Results for Holistic Score & Noun Phrases Score

	Asses	Cred	ELA_CC	AL_CC	SI_SR	Expert	Grad	NuC	Engage	Expect	SI_O
	Holistic Score										
Common intercept[a]											
Mean	-	-	-2.65*	-2.16*	-1.76	-	-	-	-	-1.60*	-
Los Niños[c]	1.29*	1.47*	1.27*	1.26*	1.17*	1.29*	1.28*	1.22*	0.87	0.83	1.03
Casi[d]	1.82*	1.93*	1.81*	1.84*	1.77*	1.83*	1.84*	1.78*	1.49	1.25	1.50
OTL effect[e]	**-0.02**	**-0.60**	**0.02**	**0.17**	**0.22**	**-0.04**	**0.12**	**0.04**	**0.40**	**0.56****	**0.34**
AL	**0.53***	**0.58***	**0.51***	**0.56****	**0.59***	**0.55***	**0.52***	**0.53***	**0.51****	**0.43***	**0.55***
Proportion	-0.75	-0.86	-0.76	-0.83	-0.97	-0.77	-0.73	-0.65	-0.77	-0.85	-0.65
Gender difference (male)[j]	- 0.31*	- 0.31*	-0.31*	-0.30*	- 0.31*	- 0.31*	- 0.31*	- 0.31*	-0.30*	-0.32*	- 0.31*
Hispanic[k]	- 0.59*	- 0.60*	-0.59*	-0.59*	- 0.59*	- 0.59*	- 0.59*	- 0.60*	-0.59*	-0.62*	- 0.60*
Threshold(2)[s]	2.32**	2.31**	2.32**	2.32**	2.31**	2.32**	2.32**	2.31**	2.31**	2.32**	2.31**
Threshold(3)[t]	4.48**	4.48**	4.48**	4.48**	4.48**	4.48**	4.49**	4.48**	4.49**	4.49**	4.48**
	Noun Phrases										
Common intercept[a]											
Mean	-	-	-3.07**	-2.46**	-	-	-	-	-	-	-
Los Niños[c]	0.61	0.53	0.65	0.56	0.56	0.58	0.56	0.56	0.45	0.25	0.59
Casi[d]	1.10	1.09	1.11	1.12	1.08	1.15	1.12	1.06	1.04	0.80	1.16
OTL effect[e]	**-0.13**	**0.13**	**-0.12**	**0.03**	**-0.07**	**-0.06**	**-0.10**	**-0.00**	**0.14**	**0.35**	**-0.04**
AL	**0.31***	**0.34***	**0.37***	**0.36****	**0.34***	**0.37***	**0.31***	**0.35***	**0.33***	**0.31**	**0.35***
Proportion	-0.91	-0.96	-1.03**	-1.01**	-0.94	-	-0.90	-0.98	-	-	-
Threshold(2)[s]	2.20**	2.20**	2.20**	2.20**	2.20**	2.19**	2.20**	2.20**	2.19**	2.21**	2.20**
Threshold(3)[t]							4.49**	4.48**	4.49**	4.49**	4.48**

Note: School names are fictitious; * p ≤ .05; ** p ≤. .01.

Table 5

HLM Results for Circumstances and Character References Scores

	Asses	Cred	ELA_CC	AL_CC	SI_SR	Expert	Grad	NuC	Engage	Expect	SI_O
	Circumstances										
Common intercept[a]											
Mean	-	-	-2.62**	-2.40**	-	-	-	-	-	-	-
Los Niños[c]	0.64*	0.62	0.64	0.62	0.66*	0.63*	0.63*	0.61*	0.48	0.56	0.62
Casi[d]	1.60*	1.61*	1.61**	1.61**	1.62*	1.61*	1.62*	1.63*	1.51**	1.52**	1.60**
OTL effect[e]	**-0.03**	**0.02**	**-0.01**	**0.06**	**-0.07**	**0.02**	**0.11**	**0.03**	**0.14**	**0.08**	**0.01**
AL	**0.41***	**0.41***	**0.42****	**0.42****	**0.40***	**0.41***	**0.40***	**0.41***	**0.40****	**0.40****	**0.425**
Proportion	-	-	-0.89*	-0.92*	-	-	-	-.81*	-0.87*	-0.90*	-0.89*
Gender difference (male)[j]	0.33*	0.33*	-0.33*	-0.32*	0.32*	0.32*	0.32*	0.32*	-0.33*	-0.33*	-0.32*
Threshold(2)[s]	3.00*	3.00*	3.00**	3.00**	3.01**	3.00**	3.00**	3.00**	3.00**	3.00**	3.01**
Threshold(3)[t]							4.49**	4.48**	4.49**	4.49**	4.48**
	Character References										
Common intercept[a]											
Mean	-	-	-2.25**	-2.01*	-	-	-	-	-	-1.58*	-
Los Niños[c]	0.46	0.50	0.51	0.42	0.47	0.41	0.39	0.42	0.27	0.33	0.51
Casi[d]	0.89	0.95	0.93	0.88	0.89	0.88	0.88	0.92	0.78	0.78	1.01
OTL effect[e]	**-0.13**	**-0.22**	**-0.09**	**-0.05**	**-0.18**	**0.06**	**0.12**	**0.02**	**0.18**	**0.15**	**-0.16**
AL	**0.28**	**0.33***	**0.32***	**0.31***	**0.29***	**0.30***	**0.30**	**0.31**	**0.230***	**0.28**	**0.31***
Proportion	0.19	0.08	0.11	0.13	0.25	0.13	0.13	0.19	0.14	0.11	0.05
Gender difference (male)[j]	0.34*	0.34*	-0.34*	-0.33*	0.34*	0.34*	0.34*	0.34*	-0.34*	-0.34*	-0.34*
Threshold(2)[s]	2.90*	2.89*	2.89**	2.89**	2.90**	2.89**	2.89**	2.89**	2.89**	2.89**	2.90**

Note: School names are fictitious; * $p \leq .05$; ** $p \leq .01$.

Grade-level effects. Although the LAPA was administered at three different grade levels (Tables 4 & 5), there seems to be no systematic differences in student performance across the different grade levels (β's ranged from -0.24 to 0.60 and p's > .05). These statistics are not presented in the tables as due to space constraints. This finding is likely attributed to both the nature of the student population targeted by this study (ELs) and the overlap in the language arts curriculum at the middle school in terms of literary analysis. Proficiency scores for students at these grade levels are based on the same English proficiency assessment (CELDT) that is vertically aligned to document growth overtime. Therefore, one would expect that sixth grade students classified as having intermediate English proficiency would be similar to seventh and eighth grade students designated with the same level of proficiency.

School effect. As mentioned previously, a systematic school effect (indicated on the tables as Los Niños and Casi) was found for the holistic and circumstances scores. For both measures, students enrolled in Casi and Los Niños Middle Schools performed consistently higher than the students in Wood Middle School (WMS). At the time of the study WMS was among the lowest performing in the state and qualitative data revealed that the school climate was the least conducive to instructional change (First Author, 2008). Thus, school level trends may also reflect the cumulative effects of low teacher expectations as well as inadequate support provided to teachers for implementing instructional change.

Relationship Between Writing Performance and Academic Language Instruction

Tables 4 & 5 also show that implementation of AL instruction was positively and consistently associated with LAPA scores. This relationship is noted in the tables as AL implementation and is in bold for easy reference. Across all four LAPA scores, the students in classes with teachers who had high AL implementation had significantly higher performance on the LAPA than the students in the classrooms with low AL implementation. As indicated in the table, this relationship was significant even after taking into account each of the other OTL indicators (these are identified in the column heads). Further, Figure 2, presents the probability (based on log-odds ratios) of receiving the highest score on any of the scoring dimensions (4 for holistic, 3 for all others). The graph illustrates that as the LAPA increases the level and quality of AL implementation increases. When no AL instruction was provided to ELs, the probability of receiving a high score on any of the dimensions was about 25%. In contrast, when high quality AL instruction was provided to ELs, the probability of receiving the highest score increased to just over 60%. In other words, the probability of receiving the highest scores more than doubled when ELs received high quality AL instruction. It appears that ELs who receive high quality AL instruction may be better positioned to take advantage of the learning opportunities their teachers provided.

Further, we found that AL learning opportunity equally benefited both ELs and non-ELs. That is, the distribution of scores was consistent between EL and non-ELs. No statistically significant effect was found for any of the analyses (β's ranged from 0.25 to 0.32 and p's > .05; these are not included in the tables due to space constraints). Unlike the Authors, 2008 study, we did not find an increase in the performance gap

as a result of increased OTL for ELs. This finding suggests that AL instruction may be necessary to reduce achievement disparities between ELs and non-ELs. This claim is further supported by the consistency in which AL implementation was found to be the most important OTL variable associated with student performance on all four LAPA scores (see Tables 4 & 5). For all eleven HLM analyses, AL implementation was significantly related to writing performance, even after controlling for student, classroom, and school level factors (e.g., gender, EL-proportion etc).

Relationship Between Writing Performance and Other OTL Indicators

An unexpected finding was that while all of the other OTL variables were significantly related to the outcomes measures, all but one was significant after taking into account AL implementation. Examining the rows labeled as "OTL effect" in Tables 4 and 5, one finds only a few significant statistics. The OTL effect represents the effect of the OTL variables identified in the column heads after taking into account the AL implementation effect (row is in bold for easy reference). The one outcome measure we found to be significant after the effect of the AL implementation was clearly communicated expectations. Teachers who provided clear learning expectations seemed to also have a positive impact on performance on the holistic score. This relationship was not found significant for the other three dimensions. It is possible that a focus on the meaning structures of language makes expectations clearer for students and consequently explicit statements on the objective of the lesson may not add to students' understanding of the lesson. However, the inconsistency with which this OTL indicator was significantly associated with LAPA scores makes drawing strong conclusions nebulous at this point. Additional research is needed to determine if the inconsistency was due to limitations of this study (e.g., conducting separate HLM analyses) or to other factors in the learning environment.

Although all of the OTL variables were significantly associated with LAPA scores, none were significantly associated with student performance on the LAPA after taking into account the effect of AL implementation, school, and proportion of ELs in classrooms (all p's > .05). Given the limitations of this study, it is premature to conclude that the other OTL indicators are not important for ELs. Additional research is needed to determine whether or not the lack of significance of their contribution to student performance is the result of the importance of AL instruction or an artifact of the limitations of the study.

Limitations of the Study

A major limitation of the study was the teacher sample size which prohibited detecting moderate and small OTL variable effects. While a sample of 32 is sufficient for detecting large effects using 4 variables (2 controls and 2 OTL indicators), a minimum of 53 is needed for detecting moderate and small effects. The sample size also prohibited running the analysis with all the variables simultaneously. Such a procedure requires a sample size of over 100 to yield more robust information regarding the relative importance of the remaining OTL indicators. These constraints notwithstanding, the study does provide initial empirical support for the argument that AL instruction appears to be important in writing performance

and may be more important than other types of opportunities ELs are provided. Another major limitation of the study was how ELs were grouped; either EL or non-EL. Such a grouping does not shed light on a potential differential effect of AL instruction on students with more or less English proficiency. One could argue that certain aspects of AL instruction may not be appropriate for ELs with beginning proficiency, for example, because they do not have the necessary vocabulary knowledge to be able to take advantage of opportunities focused on abstract vocabulary or complex syntax.

Discussion

The most important finding from this study was the consistent, significant, and positive relationship between AL implementation and student writing performance. This pattern supports the argument that meaning-based AL instruction may be important for ELs to fully benefit from standards-aligned, assessment-driven reform. Further, the linear relationship between AL implementation and EL performance is consistent with past research on the importance of fidelity to the instructional approach (e.g., Allen, 2007; Echevarria, Richards-Tutor, Chinn, & Ratlef, 2011;Goldenberg, 2004). As the frequency and quality of the AL implementation increased, the level of EL performance increased. That is, the more teachers followed the model for AL instruction, the greater the impact on student writing performance. Thus, a systematic focus on the linguistic meaning structures relevant to targeted school-based tasks may provide ELs with tools to cope with increased content instruction. Without such linguistic supports, ELs may not benefit from outcome-based reform efforts, even if such reforms are aimed at improving educational opportunity. However, additional research is needed to draw conclusions about the causal link between AL instruction and EL assessment performance as well as for understanding its impact relative to other OTL indicators and across content areas.

The finding that the proportion of ELs was negatively associated with teacher practice and outcomes is cause for concern. In light of the low levels of sheltered instruction teachers reported coupled with the lack of a significant grade level effect, we have suggested that teacher expectations may be a possible explanation for this finding. One would expect higher levels of sheltered instruction techniques in classrooms with high proportions of ELs to provide access to the content. Sheltered instruction techniques are specifically designed for ELs with low English proficiency yet teachers reported lower levels of curriculum access strategies than expected and they were not found to be significantly related to student performance. Therefore, it is reasonable to raise the question of teacher expectations as a plausible source of this trend. A recent study of academic language opportunities (Martinez, Bailey, Kerr, Huan, & Beauregard, 2010) reported consistent trends for science teachers. In the Martinez et. al.,(2010) study, teachers reported differences in foci of evaluation and reported less emphasis on developing ELL students' scientific vocabulary and writing skills compared to non-ELLs.

Moreover, our qualitative findings (reported in First Author, 2008) also suggest that misconceptions about EL instructional strategies may have reduced teachers' motivation to alter their practices. All the teachers in this study taught at schools identified for program improvement by the state. Therefore, the need to show that ELs' instructional needs were met may have been important for these teachers. This sentiment is consistent with past research findings revealing teacher concerns about deviating from standard curriculums when funding is based on test performance. In Gebhard et. al. (2007), for example, teachers tended to adopt a "more behavioral… conception of grammar and language teaching" (p. 429). This conception of language teaching tends not to be focused on meaning construction or language learning for authentic purposes. Therefore, these teachers may have perceived authentic meaning construction to be at odds with the perceived pressure to make gains on state assessments. Thus, research in teacher training should also address potential teacher misconceptions and test whether addressing misconceptions improves the level of implementation of instruction focused on meaning construction.

Despite these concerns, what is clear is that the provision of appropriate instructional support for ELs is a complex issue that demands closer attention to multiple factors in the instructional context. To be informative for EL achievement, OTL models and research should be expansive and move beyond content exposure and focus attention on the antecedents of EL outcomes in relation to linguistic and cultural dimensions of schooling.

References

Abedi, J. (2002). Standardized achievement tests and English language learners: Psychometric Issues. *Educational Assessment, 8,* 231-257.

Allen, R. (2007). The essentials of science, grades 7-12: Effective curriculum, instruction, and assessment. Alexandria, VA: Association for Supervision and Curriculum Development.

August, D., & Shanahan, T. (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on language-minority children and youth.* Mahwah, New Jersey: Lawrence Erlbaum

Authors. (2005). *Educational Assessment.*

Authors. (2008). *Journal of Latinos and Education.*

Authors. (2011). Paper submitted to *Educational Assessment*.

Brophy J., & Good, T. L. (1986). Teacher behavior and student achievement. In M. C. Wittrock (Ed.), *Handbook on research on teaching* (pp. 328-375). New York: MacMillan.

California English/Language Arts Committee (1999). English-Language Arts Content Standards for California Public Schools (Kindergarten Through Grade Twelve). Sacramento: California Department of Education.

Callahan, R. (2005). Tracking and high school English language learners: Limiting opportunity to learn. *American Educational Research Journal*, *42*, 305–328.

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research, 49*(2), 222-251.

Cummins, J. (1984). *Bilingualism and special education: Issues in assessment and pedagogy*. Clevedon, England: Multilingual Matters.

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives, 8*, 1-50.

Echevarria, J., Richards-Tutor, C., Chinn, V. P., & Ratleff, P. A. (2011). Did they get it? The role of fidelity in teaching English learners. *Journal of Adolescent & Adult Literacy, 54(6),* 426-434.

Echevarria, J., Short, D., & Powers, K. (2006). School reform and standards-based education: A model for English-language learners. *Journal of Educational Research 9*, 195-210.

First Author. (2008). *Bilingual Research Journal*, *31,* 295-322.

First Author. (2010). *Journal of Education for Students Placed at Risk.*

Gebhard, M., Harman, R. & Seger, W. (2007). Reclaiming Recess: Learning the Language of Persuasion. *Language Arts* 84.5, 419–430.

Gee, J. P. (2008). A Sociocultural perspective on opportunity to learn. In P.A. Moss, D. C. Pullen, J. P. Gee, E. H. Haertel, & L. J. Young (Eds.). *Assessment, equity, and opportunity to learn,* (pp. 1-16). New York: Cambridge University Press.

Gibbons, P. (2003). Mediating language learning: Teacher interactions with ESL students in a content-based classroom. *TESOL*, *37*, 247-273.

Goldenberg, C. (2008). Teaching English language learners: What the research does—and does not—say. *American Educator, 32(2),* 8-23, 42-44.

Greenleaf, C., & Freedman, S. W. (1993). Linking classroom discourse and classroom content: Following the trail of intellectual work in a writing lesson. *Discourse Processes, 16*, 465-505.

Guiton, G., & Oakes, J. (1995). Opportunity to learn and conceptions of educational equality. *Educational Evaluation and Policy Analysis, 17*(3), 323-336.

Haertal, E. H., Moss, P. A., Pullen, D. C., & Gee, J. P. (2008). Introduction. In P.A. Moss, D. C. Pullen, J. P. Gee, E. H. Haertel, & L. J. Young (Eds.), *Assessment, equity, and opportunity to learn* (pp. 1-16). New York, NY: Cambridge University Press.

Halliday, M. A. K. (1975*). Learning how to mean: Explorations in the development of language.* London: Edward Arnold.

Lee, C. D. (2008). Cultural modelling as opportunity to learn: Making problem solving explicit in culturally robust classrooms and implications for assessment. In P.A. Moss, D. C. Pullen, J. P. Gee, E. H.

Haertel, & L. J. Young (Eds.), *Assessment, equity, and opportunity to learn* (pp. 136-169). New York, NY: Cambridge University Press.

Linan-Thompson, S., & Vaughn, S. (2007). *Research Based Methods of Reading Instruction for English Language Learners*. Alexandria, VA: ASCD.

Martinez, J. F., Bailey, A. L., Kerr, D., Huang, B., Beauregard, S. (2010). *Measuring opportunity to learn and academic language exposure for English language learners in elementary science classrooms.* (CRESST Report 767). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Porter, A. C. (1991). Creating a System of School Process Indicators. *Educational Evaluation and Policy Analysis, 13(1)*, 13-29.

Schleppegrell, M. J. (2003). *Grammar for writing: Academic language and the ELD standards.* UCLMRI Final Report.

Schleppegrell, M. J., Achugar, M., & Oteiza, T. (2004). The grammar of history: Enhancing content-based instruction through a functional focus on language. *TESOL Quarterly, 38,* 67–93.

Stevenson, H., & Stigler, J. (1992). *The learning gap: Why are schools are failing and what we can learn from Japanese and Chinese education.* New York: Summit Books.

Walqui, A. (2006). Scaffolding instruction for English language learners: A conceptual framework. *International Journal of Bilingual Education and Bilingualism, 9,* 159-180.

Figure 1. Probability of Receiving the Highest Score on the LAPA Scores as a Function of Academic Language Implementation. The same trend is found for each LAPA score.