Issues in Measuring Opportunity to Learn Academic Language: Social Desirability and Teacher Education Effects

Zenaida Aguirre-Munoz[1,*], Christy Kim Boscardin[2] & Rong Chang[3]

[1]Psychological Health & Learning Sciences Department, University of Houston, Houston, Texas 77204, USA

[2]School of Medicine, University of California, San Francisco, USA

[3]Advancing Community Engagement & Service Institute, University of Houston, Houston, Texas 77204, USA

*Corresponding author: Psychological Health & Learning Sciences Department, University of Houston, Houston, Texas 77204, USA. E-mail: zaguirre-munoz@uh.edu

Abstract

Many assert the importance of academic language (AL) instruction in raising achievement for English learners (EL). The increased academic communication demands of the Common Core State Standards has resulted in greater attention to learning opportunities provided to students in meeting disciplinary learning. Yet, few studies have tested operationalizations of AL to investigate opportunity to learn (OTL) for English language learners. This paper discusses the issues in the measurement of AL instruction using survey instruments in bilingual contexts. Although results suggest our initial attempt to develop a survey scale show items are reliable and exhibits some degree of content validity, latent class analysis revealed that teachers interpreted the items differently depending on their level of education. Responses also appear to be confounded by social desirability evident in the lowest performing classrooms. Implications for the measurement of OTL academic language are discussed.

Keywords: Opportunity to learn, Academic Language, English learners, Instructional indicators

1. Introduction

The growing diversity of English Learners (ELs) increases the complexity of challenges for teachers in mainstream and bilingual contexts across the nation. In the context of the No Child Left Behind Act (NCLB), the need to raise the quality of instruction for ELs is further heightened because high stakes assessments define student learning opportunities (Aguirre-Muñoz & Amabisca, 2010; Gee, 2008). Compounding this complexity is the tendency for states to include ELs in their accountability systems as early as one year after enrolling in US public schools. In this context, the instruction of linguistic skills needed for mainstream class participation, referred to as academic language (Bailey, 2007; Schleppegrell, 2004a), is particularly concerning. Despite the general consensus that academic English proficiency is paramount for the educational success of ELs (August & Shanahan, 2006; Cummins, 2014; Ucelli, et al., 2015), to date few published empirical studies have examined directly the relationship between academic language instruction and EL achievement. This study was part of a previously reported investigation (Aguirre-Muñoz et al., 2006; Aguirre-Muñoz, 2014) examining an opportunity to learn (OTL) framework that incorporated instructional processes and content opportunities thought to be relevant for explaining EL achievement. Given the increased academic communication demands of the Common Core State Standards and attention to learning opportunities provided to students in meeting disciplinary learning, we believe the current study can offer insight in the measurement of OTL disciplinary language learning.

The complete OTL framework for EL outcomes has been described elsewhere (Aguirre-Muñoz et al., 2006; Aguirre-Muñoz & Amabisca, 2010; Aguirre-Muñoz, 2014). The focus here uses an unpublished subset of the data from the previous work to highlight issues in the measurement of opportunity to learn discipline-specific academic language. We found potential confounding effects related to a commonly used method for collecting OTL information–teacher self-reports of content coverage and exposure. While these effects are well known measurement phenomena, these effects may be intensified in contexts where large numbers of ELs are underperforming and there is significant pressure to improve test scores.

1.1 Past OTL Research

OTL variables are important in explaining students' test scores (e.g., Abedi & Herman, 2010; Boscardin, Aguirre-Muñoz, Stoker, Kim, & Kim, 2005; Aguirre-Muñoz, 2010; Aguirre-Muñoz & Amabisca, 2010; Monte-Sano & De La Paz, 2012; Stevenson & Stigler, 1992). For example, students are more likely to respond to an item correctly if they had the opportunity to learn the tested concepts and skills (Muthén, Kao, & Burstein, 1991). In addition, OTL variables (content coverage, content exposure, content emphasis, and quality of instructional delivery) have been shown to be significant predictors of both written and hands-on test scores and the effects of OTL varied by test format (Wang, 1998). In the Wang (1998) study, content exposure was the most significant predictor of students' written test scores, whereas quality of instructional delivery was the most significant predictor of the hands-on test scores.

Moreover, the impact of school capacity and social support on academic achievement can be traced to test-driven accountability (Lee & Smith, 1999; Newmann, King, & Rigdon, 1997). For example, research found smaller achievement gaps among racial and socioeconomic groups in states with a more equitable distribution of school resources and classroom OTL indicators (Wong & Lee, 1998). Student background characteristics such as language background, ethnicity, and gender have repeatedly been associated with levels of OTL. For example, Abedi and Lord (2001) found that students' language proficiency was positively associated with performance on the National Assessment of Educational Progress in mathematics. Abedi & Herman, (2010) found that lower OTL reported by middle school ELs, is related to lower performance on content assessments and found that as the proportion of ELs in classrooms increased, the level of OTL dropped precipitously. Likewise, Guiton and Oakes (1995) found that classes predominantly composed of White and Asian students reported higher levels of teacher quality (teacher experience, education, and assignment) than mixed or predominantly minority classes. Correlations between new content coverage in five topic areas suggested that significantly more fraction and ratio subtopics were "introduced" rather than covered in depth as minority composition of the classes increased. These researchers also found that regardless of students' initial achievement level, those who were placed in lower level courses showed smaller gains over time than students of comparable achievement who were placed in higher level courses. Subsequent studies have found similar results, particularly for ELs (e.g., Abedi & Herman, 2010; Dabach, 2014, 2015; Muller, Rieglo-Crumb, Schiller, Schiller, Wilkinson, & Frank, 2010).

A narrow focus on performance outcomes does not promise improved equity in the distribution of student learning. For example, Aguirre-Muñoz & Boscardin (2008) found an increase in the writing performance gap between ELs and non-ELs as the amount of content exposure to writing instruction increased. 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. These researchers found that instruction was not individualized to students' specific linguistic needs and may have contributed to the achievement gaps. Given these trends, we have argued that in order for increases in content exposure to yield desired outcomes for ELs, linguistic accommodations that scaffold the discipline-specific academic language may be necessary.

Increasingly, achievement inequalities are attributed to the lack of instructional focus on academic language (e.g., Aguirre-Muñoz & Gregory, 2019; Aguirre-Muñoz, Park, Amabisca, & Boscardin, 2008; Pando & Aguirre-Muñoz, 2019; Cummins, 2014; Gee, 2008; Ucelli et al., 2015). Applied linguistics research frequently demonstrates the difficulty ELs have in achieving control of academic language and suggests that exposure to focused instruction in this area leads to improved learning of academic language in various content areas (Gebhard, Harman, & Seger, 2007; Gibbons, 2002; Schleppegrell, 2004a; Schleppegrell, Achugar, & Oteiza, 2004; Spycher, 2007). Recent design-based research has demonstrated that students as young as second grade do benefit from OTL discipline-specific academic language instruction (e.g., Aguirre-Muñoz & Gregory, 2019; Haneda, 2014; Moore & Schleppegrell,

2014). These findings support the need for integrating language factors, academic language OTL in particular, that are relevant for ELs into OTL models. OTL models that do not contextualize learning opportunities for various groups of students do not account for antecedents of EL learning and thus provide limited guidance for developing reform policies that adequately address their instructional needs (Cummins, 2014).

1.2 Importance of Academic Language as an OTL Construct

Research in educational linguistics demonstrates that proficiency in conversational English is necessary but insufficient for academic success (Bailey, 2007; Cummins, 1979; Shanahan & Shanahan, 2008). Cummins (1979), an influential bilingual education scholar who first alerted the education community of the role of academic language in EL achievement maintained that language of school-based tasks should be considered from at least two dimensions: the level of contextualization and cognitive demand (mental behaviors required by school activities). Within these dimensions, Cummins described academic registers of schooling as context-reduced and cognitively demanding and everyday language as context-embedded and cognitively undemanding. Cummins demonstrated that acquiring proficiency in context-reduced language that involves cognitively demanding tasks requires a more time than acquiring context-embedded language involving cognitively undemanding tasks. Although Cummins' initial definition was highly criticized for various theoretical and empirical reasons, his conceptualization continues to inform academic language research and scholarship (Haneda, 2014).

More recent studies in this area have demonstrated that although academic language in subject areas appears to be important for creating meaning in school-based tasks (Bailey 2007; Gee, 2003, 2008), it does not develop naturally in language development (Brisk, 2015; Schleppegrell, 2004b). Even students, who have had the luxury of participating in bilingual or dual language programs (programs that utilize the child's home language as the primary instructional tool for content instruction), cannot rely solely on positive first language transfer as originally argued by bilingual education scholars (Gebhard et al., 2007). Moreover, as students advance in grades at school, the linguistic expectations increase sharply (Bailey, 2007; Harklau, 2002). Further, Uccelli, and her colleagues (Uccelli et al., 2015) found that academic language knowledge proficiency can predict reading compression skills in upper elementary and middle school students. Thus, it is reasonable to argue that OTL academic language can have positive effects in achievement, particularly for ELs. Operationalizing a construct targeting OTL academic language is therefore an essential step in the development of an OTL model that is contextualized to the needs of ELs. In this study, we examine the technical quality of a measurement scale deigned to address specific academic language content coverage items for language arts instruction. Although we found technical flaws in the scale, latent class analyses point to potential sources of these issues and qualitative data provide additional evidence of these potential sources of error.

1.3 Theoretical Basis for Operationalization of Academic Language

Genre-based literacy programs informs our approach to academic language (e.g., Christie, 1999; Martin & Rose, 2005; Rose & Martin, 2012) and is grounded in Systemic Functional

Linguistics (SFL) Theory (Halliday, 1975, 1994). Within this framework, language is viewed as a system of language resources for creating meanings serving specific purposes in the social context where it is used. Genre-based pedagogy advocates explicit teaching of academic genres that students are expected to master in school-based tasks to be successful in school settings (Brisk, 2015). It specifically involves teaching of linguistic features found in academic genres. In this paper, we use the term *genre* to refer to types of texts (e.g., exposition, narration, explanation, etc.) and the term *register* to refer to a set of lexical and grammatical resources that characterize a genre (Halliday & Hassan, 1976).

In operationalizing academic language, we draw on the SFL notions of field, tenor, and mode, which construe ideational, interpersonal, and textual meanings, respectively. These three language metafunctions offer specific targets for framing academic language instruction (Aguirre-Muñoz, 2014; Aguirre-Muñoz et al., 2008; Brisk, 2015; Moore & Schleppegrell, 2014) and for operationalizing it as an OTL construct (Aguirre-Muñoz et al., 2006). Field refers to how ideas (or content) are expressed through noun groups, verbal groups, and adverbial expressions. Tenor refers to the way the writer (or the speaker) conveys to the reader (or the listener) a stance toward the text being created. Such interpersonal meanings are rendered through modal verbs/adverbs (e.g., may, can, would, perhaps, etc.), evaluation-laden lexical choices, as well as constructions such as "I think that…" or "I feel that…" Mode refers to the way that language is delivered (e.g., written or spoken, formal or informal, etc.), rendering textual meanings. Textual meanings are realized through various cohesive devices including connectors (e.g., however, therefore, etc.) as well as clause-combining strategies. Academic language is thus described in terms of the lexical and grammatical choices that are made to create ideational, interpersonal, and textual meanings in ways that are acceptable in school settings. (See Schleppegrell 2004 for comprehensive descriptions of these elements.) Table 1 presents examples of these dimensions and depicts important distinctions between academic language and everyday language in the context of persuasive texts; the targeted genre in the larger study.

Utilizing this theoretical framework as the starting point to operationalizing the construct in the OTL context addresses the limitations of narrow definitions found in past research. Narrow definitions define academic language as knowledge of academic words and may also include knowledge of morphology or word parts (Uccelli et al., 2015). Although word knowledge has been consistently shown to be predictive of literacy skills, vocabulary interventions have not resulted in consistent and robust effects for improving literacy skills (Deshler, Palinscar, Biancarosa, & Nair, 2007; Proctor et al., 2011). These research patterns indicate that academic language instruction focused narrowly on vocabulary knowledge may not provide exposure to additional key academic language skills that support development of a broader set of academic language skills necessary to benefit from vocabulary intervention. Uccelli et al. (2015) have begun to apply a broader operationalization of academic language to develop and test the construct they call Core Academic Language Skills (CALS). Their work shows promise in the utility of using a broader definition of academic language skills to predict reading comprehension in upper elementary and middle school students. However, their approach contrasts from ours in that, they target "language forms and functions that

co-occur with oral and written school learning tasks *across disciplines*" (Uccelli et al., 2015 p. 1079, emphasis added).

Table 1. Comparison of Academic Language to Everyday Language

Task/Activity Expectation	Grammatical Features	
Display knowledge (field)	*Academic Language*	*Everyday Language*
	(a) Complex noun groups with specialized, technical, and abstract vocabulary	
	Recent brain study *data* indicate a clear and positive link between *physical activity* and *brain function*.	Students need to run around skip and play. It's good for them. It's not fair that students cannot have recess!!!!!
	(b) Verbs that enable clause-internal reasoning with nouns, verbs, and prepositions instead of conjunctions	
	Children *engaged* in moderately vigorous playground activity were less fidgety and more attentive after recess.	Recess makes children active *and* it makes them pay attention after recess.
Be authoritative (tenor)	(a) Declarative mood and modal verbs to accomplish "reasoned" judgments	
	The increase in engaged activity during instruction *should* produce improved achievement.	Why is recess so important? It makes kids smarter.
	(b) Implicit evaluation construction	
	The evidence clearly demonstrates the advantages of recess in learning.	I think that this shows that recess will make students learn better…
Structure text in expected ways (mode)	(a) Clause-combining strategies of condensation and embedding (subordinate clauses)	
	The widely held belief in the adult world is that workers are more satisfied, alert, and attentive *if given sufficient breaks.* (1 clause; embedding)	A lot of adults think that giving breaks to workers makes them happier and they also think it makes them pay attention more and be more prepared to do their work. (2 clauses; no embedding)
	(b) Lexically dense clauses through the use of abstract nouns that express a whole clause of information (grammatical metaphor/nominalization)	
	The widely held belief in the adult world…	Different clauses are used to express ideas or to describe

Note: Each contextual feature described has multiple strategies for accomplishing linguistic expectations; Theme is not included in the mode dimension due to space constraints, the use of extended text to illustrate the difference between academic and everyday language is needed; Some of the academic language examples are adapted from Gebhard et al. (2007).

Our focus is on discipline-specific academic language for writing. Consistent with Moore & Schleppegrell, (2014), we recognize that the specific features of academic language vary by task, subject matter, and grade level. To be useful instructionally, scaffolding strategies for academic language use should reflect specific contexts of use. One way this specificity can be addressed in language arts is to target linguistic features that are relevant for genres common in schools. In this study, language for persuasion (also referred to as argumentation) in English language arts was selected because the writing assessment task targeted in the larger study asked students to identify a heroic character in a fictional story and present a case (an argument) for depicting the character as heroic. Thus, the task requires the use of lexical and grammatical resources needed for explanation and argumentation as well as (character) description. This type of character analysis is a common and valued type of writing genre for district and state assessments (Bailey, 2007; Christie & Derewianka, 2008; Schleppegrell 2004a).

The argument of this line of research is that OTL school-based content involves opportunities to increase ELs' ability to make more strategic linguistic choices; choices that are sensitive to their immediate and disciplinary-specific context by exploring the way language functions to: (a) convey meaning or ideas (field), (b) enact relationships among participants (tenor), and (c) reflect the expected medium or mode of communication (mode). Based on this model for addressing ELs' OTL academic language, this study explored: (1) how academic language instruction can be captured in a survey instrument by examining evidence of reliability and construct validity; (2) response patterns of middle school teachers and the factors that might be accounting for these response patterns; and (3) the degree of correspondence between survey responses and other (qualitative) sources of OTL data. The following questions guide this study:

1. To what extent does the internal consistency of the items provide evidence of reliability and construct validity?

2. What levels of OTL academic language are provided to middle school ELs? Do trained and untrained teachers report different levels of OTL academic language?

3. What factors account for teacher OTL response patterns?

4. To what extent do teachers' self-reports of OTL correspond to researcher observations of teacher practice and teachers' qualitative descriptions of instructional opportunities?

2. Method and Data Source

A mixed methods research approach was employed to address the research questions. We examined the instructional practice data from three sources: (a) the teacher OTL survey responses, (b) classroom observations, and (c) teacher interviews. The OTL survey measured the quantity of specific instructional practices aligned with OTL variables; the observation and interview data provided additional information on the quality of those practices.

2.1 Sample

A sample of 32 language arts teachers from three urban middle schools in Southern California participated in the study (19, 9, and 4 teachers from Wood, Los Ni ños, and Casi Middle Schools respectively; school names are pseudonyms). We did not expect teachers to have relevant background experiences in academic language as defined by this study. Therefore, 21 of the teachers received professional training in academic language (summarized below) in the winter and additional follow-up training in the spring. The training group was over-sampled due to anticipated attrition and potential for low implementation of training content.

The majority of teachers participating in the study were recruited from the largest of the three schools, Wood Middle School. The total years of teaching experience at the time of the study ranged from 1 to 27 years with an average of 10 years for both trained and untrained teachers (see Table 2). In general, trained and untrained teachers were very similar in terms of their teaching experience. Differences in the average number of years teaching sheltered English language arts was not statistically significant ($M = 2.91$, $SD = 1.58$ and $M = 2.22$, $SD = 1.63$ for untrained and trained teachers respective), $p > .05$. Although we only recruited teachers who taught sheltered English language arts, the majority of the teachers did not major in English or language arts as undergraduates. Sheltered instruction is an approach to teaching ELs which integrates language and content instruction with a focus on improving English language proficiency (Echevarria, Short, & Powers, 2006). Twenty-four teachers (75%) held teaching credentials. Only 20% ($N = 4$) of the trained teachers and 33% ($N = 4$) of the untrained teachers did not hold teaching credentials at the time of the original study. Seven out of 32 teachers indicated that they either already had an advanced degree or were currently enrolled in a graduate program.

Table 2. Background Information for Comparison and Trained Teachers

Variable Category	Comparison			Trained		
	N	*Mean*	*SD*	*N*	*Mean*	*SD*
Total years of teaching	12	10.04	7.34	20	9.55	8.58
Years at this school	12	4.63	4.04	20	4.50	4.12
Years of teaching English language arts	10	5.05	5.35	18	7.78	6.46
Years of teaching sheltered English	9	5.33	5.32	15	2.27	2.25
Exact No. English/LA courses – und	9	4.11	2.80	17	3.65	3.24
Exact No. English/LA courses – grad	12	2.25	2.80	14	2.93	3.05
Number of courses - sheltered/SDAIE	11	2.91	1.58	18	2.22	1.63
Number of courses – ESL	12	2.08	1.88	17	2.29	1.61

Note: SDAIE = Specially Designed Academic Instruction in English; ESL = English-as-a-Second Language.

The missing entries were omitted in the comparison of teaching background.

2.2 Teacher Training

The intent of the larger study was to determine the relationship between academic language instructional opportunities and EL writing outcomes in the context of a broader OTL model. At the time of the OTL study, our conception of academic language informed by SFL (instruction

beyond engagement with academic words and word parts) was a new construct in teacher training. Therefore, in order to determine the impact of this newly defined construct on OTL we needed to create the conditions where academic language instruction aligned to this definition would be observed in enough classrooms to be able to measure impact. To this end, we provided some teachers with a five-day training in academic language instruction. The teacher training has been described elsewhere (see Aguirre-Muñoz et al., 2006; Aguirre-Muñoz et al., 2008b), therefore, only a brief description is included. The training was narrowly focused on developing teacher knowledge of the SFL metalanguage so trained teachers would be able to provide concrete assistance to student learning how to write about their evaluations of characters in stories. Thus, the training provided teachers with engaging content related to specific concepts corresponding to SFL language metafunctions (field, tenor, mode) for writing instruction. This focus "enables teachers to foreground meaning while also being explicit about language forms" (Moore & Schleppegrell, 2014, pg. 93).

Four modules were used to target academic language content, instructional models, and provide model lessons. On each of the first four days of training, a different module was introduced. Within each module, the presentation of the material utilized several whole- and small-group activities. The first two modules targeted concepts related to the SFL approach, highlighting the three contextual features of language: field, tenor, and mode in teacher-friendly terms. The third module addressed instructional strategies that emphasized co-construction of knowledge through classroom talk. Module four, presented during the fourth day of the training was dedicated to the analysis of student writing, and the collaborative development of lessons using SFL concepts. On the fifth day of the training session, teachers collaboratively-developed lessons and completed post-workshop instruments designed to measure the impact of the training on improving teacher knowledge and application of training content. The training evaluation showed that it was effective in building teacher knowledge of SFL metafunctions, identifying specific areas of instructional support for ELs, and providing feedback to students on their written academic discourse skills. Post-test scores on a performance assessment were significantly higher than pre-test scores (p-values for each linguistic dimension was lower than .05) (See Aguirre-Muñoz et al., 2008).

2.3 Instrumentation

2.3.1 OTL Scale

All participants completed a teacher survey intended to capture OTL coverage and exposure to academic language instruction, the two most common constructs in the OTL literature (Abedi & Herman, 2010). We confined OTL data to teacher survey responses and did not include student responses due to past research indicating a high mismatch between teacher and EL responses (Abedi & Herman, 2010). Further, we used topics that were not part of typical language arts instruction for ELs during the year the data were collected; we did not have the resources to train students to reliably complete the survey. We used research observation and interview data for cross-validation of survey responses. Teachers rated the average frequency (exposure) in which they provided explicit instruction in seven academic language topics (coverage) in lessons targeting writing development. Each topic represented a linguistic feature

related to one of the three language SFL metafunctions: field, tenor, or mode. Items targeting frequency of coverage included: "verb choices that signal analysis of a character or situation" (tenor) and "grammatical structures that build cohesion at the sentence level" (mode). Teachers responded on a six-point scale, ranging from "Never" to "2 or more times per day."

2.3.2 Teacher Observation Protocol

To obtain additional information about classroom practice, classroom instruction was observed twice: two and three months after the teacher training. For each round of observations, teachers were observed for two consecutive lessons. Two sets of items on the observation protocol focused on students' level of exposure to linguistic features relevant to written characterizations. The first item asked observers to rate globally the degree to which classroom activities supported the application of language concepts. Observers rated teacher's delivery of language related activities on a scale from "highly evident," meaning the teacher provided frequent opportunities (exposure) for students to apply language knowledge in the classroom to "not evident," where the teacher provided no activities for students to apply language knowledge in the classroom. The second item set included six linguistic features that correspond more directly to academic language informed by SFL. For each of the six features, observers rated the extent to which specific linguistic features were addressed in the lesson from "a great extent," to "barely covered" (coverage). These linguistic features included (a) use of various clauses and phrases in theme position (SFL syntactic elements that corresponds to complete subject) to create sentence variety (mode); (b) use of adjectives, verbs, and adverbs to describe events (field); (c) use of various verb types to provide textual interpretations (tenor); (d) use of connectors and expanded noun phrases to build cohesion (mode); (e) use of model verbs that frame the writer's point of view (tenor); and (f) use of grammatical structures to generate an impersonal context (mode).

To obtain a more complete picture of classroom processes, observation field notes were taken to contextualize information gleaned from the analysis of the structured classroom observation and teacher interview data (described below). To ensure solid inter-rater reliability on the observation protocol, the research team participated in an observer calibration session 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.

2.3.3 Teacher Interview Protocol

A semi-structured interview protocol were conducted when interviewing participating teachers. Teacher responses to particular items from the interview protocol were analyzed in line with our operationalization of academic language. These included broad questions such as, "How do you prepare ELs to process academic language?" as well as more specific questions such as, "How do you integrate grammar instruction with literary analysis or writing development?" For trained teachers, additional questions about the linguistic features addressed in the teacher training were asked, such as, "Which activities/strategies have you used from the institute so far?" Researchers also reviewed teacher responses to questions pertaining to the observed

lessons, which often led to a discussion of typical classroom practices, including how (if at all) they included academic language instruction in their practice. Answers were then coded for evidence of exposure to and coverage of specific grammatical features or functions, including: "grammatical structures that build cohesion," "vocabulary that reveals interpretation," "long noun phrases to increase sentence variety," and "overall essay cohesion and organization."

Each teacher was interviewed following each of the two classroom observations. The audio-taped interviews were then transcribed and coded according to the SFL concepts. To achieve coding consistency, each of two researchers independently coded a set of interview transcripts, reviewed each other's preliminary codes, and in consultation with each other refined them further. This process continued until a 90% exact score agreement rate was reached and Kappa coefficients on 3 coded transcripts was at least .75 for each of the codes pertaining to academic language. Trends in observation, survey, and teacher interview data were compared and served as additional validity evidence of the academic language OTL survey scale.

3. Results and Analysis

3.1 Internal Consistency of Academic Language Items (Research Question 1)

We examined teacher response patterns as evidence of reliability and construct validity. The Cronbach alpha obtained for the academic language coverage items was .89 and thus provides evidence of strong reliability for this scale. In addition, construct validity of the survey items was examined. We conducted confirmatory factor analysis (CFA) as it is traditionally used to test whether the items are sufficiently representing the pre-specified constructs or content domains. The robust maximum likelihood (ML) estimation and the fixed-factor scaling method (McArdle & McDonald, 1984) were used for better interpretation of item coefficients. The

latent factor analysis demonstrated good model fit with fit indices $\chi^2 = 2.25$, $p < .001$ and CFI =

0.92. As shown in Table 3, all the factor loadings to the latent construct were significant ($p < .001$), suggesting that all the items were adequately measuring the proposed construct. The relatively high and consistent loading estimates also indicated that the indicators have great consistency and predictability.

3.2 Teacher Reported Levels of OTL Academic Language (Research Question 2)

As presented in Table 3, most teachers reported frequently engaging students in instruction on general language grammatical structures (item a). Sixty-one percent of teachers reported engaging students in this type of instruction at least twice per week. In terms of specific grammatical features (items b-g), teachers reported devoting the most attention to vocabulary focused on presenting clear descriptions (field) followed by cohesion strategies (mode). Respectively, 50.0% and 43.8% of teachers reported focusing on these areas at least two times per week. The area teachers reported spending the least amount of time was grammatical structures that generate impersonal tone where 16.1% of teachers reported focusing on this area at least twice per week and 51.6% of teachers reported devoting instructional attention to this topic "less than once per week" or "never."

Table 3. Confirmatory Factor Analysis & Teacher Self-Reports on Academic Language Coverage Items

| Linguistic feature | | Percentage of Respondents | | | | | CFA Results | | | |
	Never	Less than once a week	Once a week	2-4 times a week	Once a day	2 or more times a day	Esti-mates	S.E.	Est./ S.E.	p
a. English language grammatical structures	3.23	3.23	32.26	32.26	22.58	6.45	1.00	0.00	0.00	.000
b. Long noun phrases to increase sentence variety in a piece of writing (field)	12.50	25.00	28.13	28.13	3.13	3.13	1.05	0.19	5.58	.000
c. Vocabulary to describe of characters or situations (field)	0.00	12.50	37.50	28.13	15.63	6.25	1.17	0.21	5.55	.000
d. Verb choices that signal analysis of a character or situation (tenor)	6.67	16.67	36.67	26.67	13.33	0.00	0.52	0.15	3.53	.000
e. Grammatical structures that build cohesion at the sentence level (mode)	3.13	25.00	28.13	34.38	6.25	3.13	1.18	0.17	7.06	.000
f. Grammatical structures that signal point of view (tenor)	3.13	31.25	34.38	21.88	3.13	6.25	1.60	0.28	5.70	.000
g. Grammatical structures that generate an impersonal tone (tenor)	9.68	41.94	32.26	12.90	0.00	3.23	1.61	0.30	4.70	.000

Note: $\alpha = .89$; n = 28.

Response patterns of the items targeting linguistic features most germane to the written genre of characterization also reveal lower levels of academic language instruction than expected for sheltered content courses. Recall that sheltered content courses require an explicit and significant focus on content-based language learning. The percentage of teachers who reported engaging students in instruction on specific linguistic features with high frequency (at least once per day) ranged from 3.2% to 21.8%. Further, the percentage of teachers who reported focusing instruction on specific features less than once per week ranged from 12.5% to 51.6%. When these responses were combined with teachers reporting coverage at least once per week, the range was 37.5% to 83.88%. Given that ELs need significant amounts of time actively engaging English, particularly written modes (Feng & Powers, 2005; Ferris, 1997; Harklau, 1994; Hakuta, Butler, & Witt, 2000; Sasaki & Hirose, 1996), direct exposure to academic language at either of these levels is arguably not an optimal amount of time if the goal is to raise student achievement in English language arts. Finally, as indicated previously, instructional

focus was highest for "vocabulary to describe characters or situations," with 50.0% of teachers reporting frequent (at least two times per week) instruction in this area. This is not surprising given that most teacher training presents academic language as a narrow construct (Aguirre-Muñoz & Gregory, 2019; Pando & Aguirre-Muñoz, 2019). These findings suggest that overall most teachers reported relatively low levels of engagement in linguistic features that comprise academic language needed to produce a written character analysis or persuasive text.

3.3 Factors Contributing to OTL Responses (Research Question 3)

To further explore the validity of the survey items, we compared the response patterns of the trained and untrained teachers. Since the training process data (see Aguirre-Muñoz et al., 2008) showed the training was effective in developing applied teacher knowledge of academic language, we expected trained teachers to report significantly higher levels of OTL than untrained teachers. The survey instrument did not capture differences in self-reports of OTL academic language between teachers trained in SFL and the untrained group. No such statistically significant differences were found ($p > .05$) on the survey instrument. This finding called for further examination of the survey items.

To understand teacher response patterns to the academic language items on the teacher survey instrument, we conducted Latent Class Analysis (LCA). LCA allows us to examine potential explanations for the lack of statistical differences in reporting patterns between trained and untrained teachers as well as possible explanations for discrepancies between the survey and qualitative data (described below). To select the number of classes for the final LCA model, a series of models with increasing class numbers was fit using items targeting the academic language OTL construct. The Bayesian information criterion (BIC) and Lo-Mendell-Rubin-Likelihood Ratio Test (LMR-LRT) was applied to determine a preferred number of classes. With a significant p-value in LMR-LRT and higher level of BIC index, a K-class model is statistically superior than (k-1) class. Table 4 presents a summary of the results for the measurement model. Although LMR-LRT was not statistically significant (p = .28), the BIC for the 2-class model was significantly smaller than the BIC for the 1-class model. A series of simulation studies have proved (Nylund, Asparouhov, & Muthén, 2007) that BIC is a better indicator for class enumeration compared to LMR-LRT. Therefore, the 2-class model was more sensitive for us to understand teacher response pattern.

Table 4 also presents the estimated mean values for each of the items within each of the latent classes. These class-specific mean values are useful for understanding the character of the classes. The classes are similarly defined for all items: Class 1 represents those teachers that responded low in all the items, and Class 2 represents those teachers that answered relatively high in all the items. To help describe these two different sets of responders, a post-hoc investigation of correlates related to the two classes was conducted based on modal class assignment. We opted for post-hoc analysis rather than including the correlates in the LCA models due to the small sample size. This analysis can be argued as less conservative because it does not account for uncertainty of class membership. For this reason, results are generally treated as a more descriptive and exploratory technique. However, our final model indicated a

very high estimated classification precision. This level of classification makes changes to inferences regarding the possible correlates unlikely. There is also an ease in interpretability not present when including such correlates simultaneously within the LCA framework.

To further examine the pattern of responses found in the two classes and explain class membership, several different background variables were investigated. Post-hoc analysis (Table 4), indicated teachers with less education consistently responded differently than teachers with higher education levels. The total number of English or language arts content courses that the teacher took in his or her undergraduate and graduate studies (number of courses) was the only significant variable related to class membership in the academic language coverage construct ($p = .03$). This finding suggests that for academic language coverage, level of education, operationalized as the number of English/language arts courses taken, seems to be the best predictor for how a teacher responds to these items. That is, whether a teacher would respond high or low on this set of items, regardless of whether or not they were trained to provide academic language instruction. In other words, teachers who were in Class 1 (characterized as teachers reporting low OTL academic language), were also those teachers with lower levels of education in the discipline. Further, teachers who were in Class 2 (those reporting high OTL academic language) were also teachers who reported higher levels of education in the discipline. Thus, more educated teachers may have been compelled to report higher levels of academic language coverage and exposure because they may have believed they should be providing more exposure to academic language. Teachers with less education, on the other hand, may have not been sure about what level of academic language coverage might be appropriate or desirable to report. The untrained teachers reported higher levels of course specific education than the trained teachers and tended to report higher levels of academic language instruction, which was not substantiated by the qualitative data discussed next. Qualitative results (described below) support the possibility that a social desirability effect may have contributed to observed response patterns.

Upon further examination of the items on the survey, we noted that these items needed explicit examples to ensure teachers from different educational backgrounds and experience would understand the items similarly. That is, in addition to a potential social desirability effect, the language and design of the items could have also contributed to differences in item interpretations. Teachers with less education, particularly those who were not trained, would be less familiar with the grammatical features targeted by the items and may have interpreted the items differently. In addition, trained teachers with low education levels may have needed more training on academic language instruction to be able to report it accurately on a survey instrument. The format and language of the items coupled with possible teacher education differences may have contributed to the lack of statistical difference between these two groups of teachers.

Table 4. LCA Model Results & Summary: Academic Language Coverage

Academic language items	Estimated means	S.E.	Est./S.E.	Estimated means	S.E.	Est./S.E.
	Latent Class 1			*Latent Class 2*		
a. English language grammatical structures	3.71	0.29	12.78	4.31	0.27	16.18
b. Long noun phrases to increase sentence variety in a piece of writing	2.53	0.28	9.11	3.49	0.33	10.60
c. Vocabulary to describe characters or situations	2.88	0.17	16.94	4.65	0.22	21.63
d. Verb choices that signal analysis of a character or situation	2.47	0.20	12.48	4.06	0.24	17.12
e. Grammatical structures that build cohesion at the sentence level	2.80	0.30	9.22	3.93	0.22	18.23
f. Grammatical structures that signal point of view	2.44	0.18	13.59	3.95	0.33	12.15
g. Grammatical structures that generate an impersonal tone	2.00	0.14	14.11	3.31	0.31	10.72
LCA Model Summary						
	2 Classes			1 Class		
BIC	594.68			607.78		
LMRT-LRT p-value	0.28					
Entropy	0.92					
Classes						
CI	14 (50%)					
C2	14 (50%)					

Including "Number of Courses" to Predict Class Membership Summary					
	Construct	Class count	"No. Courses" log-odds	Estimated	LMR-LRT Test (p-value)
Academic Language Coverage	0.96	C1: 20	-.23*	C1: 3.95	.034
		C2: 8	(0.10)	C2: 9.18	

Note: C = Class; No. = Number.

3.4 Qualitative Evidence of OTL Academic Language (Research Question 4)

3.4.1 Teacher Interviews

In contrast to the high survey reports of academic language instruction, analysis of teachers' interview comments revealed that the untrained teachers rarely discussed explicit instruction of grammatical and lexical resources targeted by the training. Lack of support provided to ELs for generating academic texts was expressed by most untrained teachers and is illustrated in the following quote.

> *Regularly, I would just give them the prompt and say, "Do it." Then I would have them do the draft, and then have them edit it and have them rewrite it. That takes like 3 or 4 days.*

Low evidence of support in developing academic language proficiency was most evident in the descriptions of writing instruction of untrained teachers. Their responses indicated they tended

to focus simply on content and ideas, and on a broad and superficial level of writing instruction (e.g., an overall essay structure). Although these teachers attempted to provide instruction in writing, their attempts did not target the linguistic features appropriate for the targeted academic genre. Nor did their instruction provide the level of specificity ELs may require to gain control of academic language. Thus, in addition to the effect of teacher education level on survey response items described above, these teachers also appeared to have a more traditional conception of grammar instruction. The interview questions and prompts attempted to target grammar from an SFL lens, yet these teachers only discussed lessons that entailed traditional grammar instruction. Examples of this kind of instruction included worksheets that highlighted grammatical forms (e.g., tense, noun-verb agreement, etc.) detached from the functions of language use and a lack of opportunity to apply concepts to students' own writing.

Teachers trained in SFL (71%) on the other hand described prewriting activities that helped students develop ideas. Their focus on elaboration and academic language allowed them to provide instruction that met grade level expectations during the first draft and revision phases. Trained teachers were much more likely to describe academic language instruction that targeted directly the genre and was more specific to the language learning needs of the ELs. The quotes below illustrates the types of instruction these teachers reportedly provided.

> *I'm breaking it down more… like the participants [nouns] and the verbs…, whereas before [my instruction] wasn't really specific…Now I ask, "How does this verb fit into what you're trying to say and the thought.* [T1]

> *After they were done with it, I pointed out what kind of verbs they were using in those paragraphs. They're like, "Oh, actually, we're using attitudinal verbs and we're using attributive verbs." I said, "Exactly, and … you can actually label those now and see what kind of purpose they [have] in writing."* [T2]

Thus, interview data suggests that, as expected, teachers trained in the SFL approach reported more coverage of academic language (compared to the untrained teachers) as well as more specific grammatical structures (exposure) consistent with the academic language as defined by this study.

3.4.2 Classroom Observations

Consistent with the overall frequencies of the survey data, 60% of teachers (trained and untrained) were observed to provide activities for students to apply language knowledge in the classroom. Since the teachers who participated in this study were sheltered content teachers, we expected all teachers to engage students in language application activities in all lessons. However, much of this application, particularly in the untrained group, tended to be focused on traditional grammar and vocabulary instruction. That is, grammar instruction was worksheet driven, did not involve analysis of model text or grammatical structures that would direct attention to the use of academic language to clarify meaning, and did not provide students with opportunities to apply concepts to their own writing. Revision lessons for many teachers tended to be focused on editing as opposed to clarification of meaning. The amount of

instruction in specific linguistic features of academic language relevant for persuasive texts, as recorded in the observation protocol, was lower than teachers' self-reports. About half (53%) of all teachers (15 of the trained group and 2 of the untrained group) were observed to engage students in academic language instruction, compared to 68.78% of teachers who reported regular (at least once per week) instruction in this area. Further, approximately one third of trained teachers (8 of 21) across both observations were rated in the observation protocol as providing high quality instruction (adequate coverage and exposure) in academic language instruction; one third of trained teachers (7 of 21) were rated as providing moderate quality instruction addressing academic language. The remaining teachers in the trained group were not observed to provide academic language instruction corresponding to the training targets at either of the two time points. Yet, these teachers reported very high levels of OTL in the survey instrument indicating a possible social desirability effect. As expected, none of the teachers in the untrained group were observed to engage students in discipline-specific academic language instruction. However, two (18.18%) did provide instruction in literary analysis germane to character analysis.

Lastly, the number of trained teachers who provided instruction on the linguistic features that comprise tenor (such as the presentation of implicit opinion and grammatical structures generating an impersonal context) was lowest across both time points and was lower than their survey-reports of time spent on these topics, as compared to the other two language dimensions. Only one to two teachers (per observation time points) were observed to provide instruction in principles pertaining to tenor.

Despite the issues with the survey instrument, the available data showed that teachers trained in academic language from an SFL perspective were more likely than untrained teachers to develop students' (a) grammatical sensitivity; (b) lexical knowledge in context; and (c) revision skills directed at two contextual features: field and mode (Aguirre-Muñoz et al., 2008). The field contextual variable (expanded noun phrases and verb choices) was the area most addressed during instruction, followed by internal structure or mode (conjunctions and transitions and clause linking strategies). Only a couple of teachers provided instruction in tenor (modal verbs and third-person references). The lower rates of coverage of linguistic features pertaining to tenor may have been attributed to limitations of the training, the lack of ongoing support of teachers after the initial training, as well as the complexity of this concept. Of the three contextual features, tenor may require deeper knowledge of English grammatical structures than the sampled teachers held during the course of this study. More research is needed to determine the potential sources of the low rates of coverage in this area.

4. Summary

We found the survey items targeting academic language to be, on the surface, reliable and exhibit some degree of construct validity in the preliminary investigations. Teachers reported lower than expected rates of exposure to general and specific grammatical features of academic language that would be expected for sheltered content classes. However, the validity of the instrument came into question when we compared teacher survey responses between trained

and untrained teachers. Trained teachers did not report higher levels of coverage or exposure to academic language as we expected given the success of the training to increase teacher knowledge of discipline specific academic language and in teachers' ability to develop lessons and provide feedback to students (Aguirre-Muñoz, 2008). The lack of statistical difference in academic language instruction between trained and untrained teachers based on the survey responses prompted us to examine what accounted for this lack of significance among the two groups. The latent class analysis revealed that teachers interpreted the items differently depending on their level of education as measured by the number of undergraduate discipline-specific courses they reported taking. Since differences were found between the trained and untrained teachers in their reports of the number of courses taken, their background knowledge may have impacted the way they interpreted the items. This finding suggests that the survey should be augmented to include more explicit examples of what is meant by each linguistic feature to ensure all teachers interpret the items similarly.

Additional concerns were apparent when comparing survey self-reports with classroom observation and interview data. Although most teachers reported frequent engagement in the application of language knowledge, analysis of observation data revealed instruction for at least half of the teachers sampled was traditional in nature and did not support ELs' specific linguistic needs. This instructional preference is problematic in that a primary principle of an SFL approach to supporting ELs in academic writing is that it should not be done in a pedantic manner (Moore & Schleppegrell, 2014). Consistent with this general trend, 60% of the observed instruction was centered on general grammatical structures. However, specific grammatical elements that comprise academic language (as operationalized here) occurred in just over half of the classrooms visited. Unlike what was reported in the survey, none of the untrained teachers were observed to provide this kind of instruction to students.

Despite the possible interpretation issue with the language of the survey described in the results section, qualitative findings demonstrated that, as expected, trained teachers were more likely than untrained teachers to engage in academic language instruction beyond superficial applications of grammar such as worksheet exercises. Although the quality of instruction varied, two thirds of trained teachers compared to none of the untrained teachers engaged students in meaning-based, grammar instruction that focused on discipline-specific academic language as defined here. These teachers attempted to reveal the importance of linguistic patterns in realizing school-based functions (character description, analysis, etc.) and clarify meaning. Another important finding was half of all teachers did not provide academic language instruction to ELs. This finding is concerning considering all of these classrooms are sheltered content courses. If schools and districts with significant proportions of ELs are to show adequate yearly progress, ELs likely ought to be engaged in meaningful academic language instruction more consistently (August & Shanahan, 2006; Johns, 1997). Despite expert recommendations about increasing academic language instruction, few investigations have examined the relationship and impact of such instruction on student achievement gains. One related line of research conducted by Ucelli et al. (2015) presents evidence of a positive and significant relationship between academic language knowledge and reading comprehension. In addition, previously we reported on the impact of OTL academic language instruction by

utilizing the interview and observation data to identify levels of OTL and found a significant and positive relationship between OTL academic language and performance on writing assessments targeting persuasive texts (Aguirre-Munoz, 2014).

In addition, the interview and observation data revealed that trained teachers provided higher levels of OTL in this area. Trained teachers provided the most instructional attention to the field contextual variable to support written descriptions followed by the mode contextual variable to support within paragraph organization. Tenor was only addressed by two teachers, which suggests that teachers may need additional training on ways to support ELs' written arguments and implicit communication of personal stance. This finding was consistent across three types of data sources (survey, observation ratings, and interview themes). Although the survey data is difficult to interpret given the interpretation issues, it is notable that this pattern was consistent with the other data sources. Qualitative data indicated an adequate degree of correspondence between observations and interview data related to the overall levels of academic language instruction provided to students and in the relative differences in OTL across the language metafunctions.

Taken together, these findings call for continued refinement of the survey instrument to reduce teachers' inflated reports of academic language coverage particularly for teachers with lower education levels and those not trained in SFL. Teachers' self-reports of academic language coverage may have been influenced by at least three factors. First, teachers' general awareness that their practices should be consistent with ELs' linguistic needs may have compelled them to report higher levels of coverage than they actually implemented, particularly for those who had higher proportions of ELs in their classrooms. Moreover, the need to show that ELs' instructional needs were met may have been important for these teachers because they all taught at schools identified for program improvement by the state department of education, therefore. This interpretation is consistent with past research findings (e.g., Gebhard et al., 2007) 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 narrow view of grammar and language teaching. This perception of language teaching is not generally focused on meaning construction or language learning for authentic purposes. Therefore, these teachers may have perceived our approach focused on authentic meaning construction to be at odds with the perceived pressure to make gains on state assessments.

Second, teachers may have not fully understood the survey items. The latent class analyses suggests that teacher lacked subject matter understanding associated language arts, particularly what distinguishes various genres. Teachers' misunderstanding of academic language is not surprising since, at the time of study, pre-service training did not provide teachers with exposure to the linguistic structures that differentiate academic language from informal language or instructional strategies to target development of academic language in this way (Wong-Fillmore & Snow, 2000). Hill and Ball (2004) also found the terms used in measures of teacher knowledge and practice are often highly open to interpretation as well as susceptible to social desirability of responses. The combination of background knowledge and social desirability appeared to have impacted teacher responses to the OTL academic language survey items. Refinement in the language of the items is needed to address the interpretability

of the item responses. Additionally, investigations should be conducted to determine if training in responding to survey items targeting academic language can improve response accuracy from teachers representing a wide range of educational background and teaching experience.

5. Conclusion

An important aim in OTL research is the need to improve relevant instructional antecedents for student achievement (Arnold-Berkovits, Kurz, & Reddy, 2018; Kurz, Elliott, Kettler, & Yel, 2014). An assumption scholars in this line of research make, is that understanding the impact of instructional inputs on student achievement will improve test score interpretation. Past OTL research has established a range of OTL indices that can improve the validity of test score interpretations (Kurz, 2011). However, much of the prior work has not targeted an important instructional dimension emphasized in the current standards-based reform and that which is also considered paramount to addressing ELs' needs. An assumption undergirding current large-scale assessment and accountability systems is that all participating students have the opportunity to learn what they are expected to know and on which they are tested (Elliot, 2015). Yet, this assumption has not been fully examined with respect to ELs. We have argued that discipline-specific academic language is important to integrate in OTL models that provide information instrumental in understanding antecedents to academic outcomes of ELs. Although our initial attempt to develop a survey scale failed to yield valid inferences about exposure to and coverage of academic language, others (e.g., Abedi & Herman, 2010; Porter, 2002) have shown survey instruments can be valid instruments for capturing quantity and curricular depth. Still others show that instructional logs can also provide technically sound information about teachers' instructional practices (Kurz, 2011; Kurz et al., 2014). To shed light on specific targets for instructional improvement of ELs, however, additional studies addressing teachers' interpretations of disciplinary-specific academic language as well as clearer operationalization of academic language indicators are needed to adequately account for the contexts they encounter in schools. Consistent with Elliot (2015), we believe "high-quality instruction…and academic growth cannot rest on policy intent…but must be contextualized by… instruction and growth that occurs …as a function of their exceptionality and background" (pg. 63). Continued efforts in the operationalization of disciplinary-specific academic language would provide such contextualization for ELs.

6. Acknowledgement and Sponsoring Information

References

Abedi, J., & Herman, J. (2010, March). Assessing English Language Learners' Opportunity to Learn Mathematics: Issues and Limitations. *Teachers College Record 112*(3), 723-746.

Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education, 14*(3), 219-234. https://doi.org/10.1207/S15324818AME1403_2

Arnold-Berkovits, I., Kurz, A., & Reddy, L. A. (2018). Teacher log of students' opportunity to learn and classroom observation: an initial investigation of convergence. *Educational Assessment, Evaluation, and Accountability, 31*(1), 97-119. https://doi.org/10.1007/s11092-018-9288-2

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 Associates, Inc. https://doi.org/10.4324/9780203937600

Aguirre-Muñoz, Z. (2010, Fall). Assessment accommodations for English language learners. *AccELLerate!, 31,* 14-16.

Aguirre-Muñoz, Z. (2014). Focus on meaning structures matters for English learners: Exploring opportunity to learn academic language**.** *International Journal of Liberal Arts and Sciences, 2*(6), 1-20.

Aguirre-Muñoz, Z., Boscardin, C. K., Jones, B., Park, J. E., Chinen, M., Shin, H. S., Lee, J., Amabisca, A. A., & Brenner, A. (2006). *Consequences and validity of performance assessment for English Language learners: Integrating academic language and ELL instructional needs into opportunity to learn measures.* CSE/ CRESST Technical Report No. 678, Los Angeles.

Aguirre-Muñoz, Z., & Boscardin, C. K. (2008). Opportunity to learn and English learner achievement: Is increased content exposure beneficial? *Journal of Latinos and Education*, *7*(3), 186-205. https://doi.org/10.1080/15348430802100089

Aguirre-Muñoz, Z., Park, J. E., Amabisca, A. A., & Boscardin, C. K., (2008). Developing teacher capacity for serving ELLs writing instructional needs: A case for systemic functional linguistics. *Bilingual Research Journal, 31,* 295-322. https://doi.org/10.1080/15235880802640755

Aguirre-Muñoz, Z., & Amabisca, A. A. (2010). Defining opportunity to learn for English language learners: Linguistic and cultural dimensions of ELLs' instructional contexts. *Journal for the Education of Students Placed at Risk, 15*(3), 259-278. https://doi.org/10.1080/10824669.2010.495691

Aguirre-Muñoz, Z., & Gregory. M. O. (2019). Concept-based teaching in dual language science classrooms: Using oral language routines to develop scientific descriptions and arguments. In P. Spycher and E. Haynes (Eds.), *Culturally and Linguistically Diverse Learners and STEAM: Teachers and Researchers Working in Partnership to Build a*

Better Path Forward (pp 45-75). Charlotte, NC: Information Age Publishing.

Aguirre-Muñoz, Z., & Pando, M. (2017). Knowing and Teaching Elementary Math to Bilingual Students: Examining the Role of Teaching Self Efficacy on Content Knowledge. *Journal of Bilingual Education Research and Instruction (JBERI), 19*(1), 126-145.

Pando, M., & Aguirre-Muñoz, Z. (2019). Engaging English language learners in model-based science instruction. In L. C. de Oliveira, K. Obenchain, R. Kenney, & A. Oliveira (Eds.), *Teaching the Content Areas to English Language Learners in Secondary Schools: English Language Arts, Mathematics, Science, and Social Studies* (pp. 215-230). Switzerland: Springer.

Bailey, A. L. (2007). *The language demands of school: Putting academic English to the test.* New Haven, CT: Yale University Press.

Boscardin, C. K., Aguirre-Muñoz, Z., Stoker, G., Kim, J., & Kim, M. (2005). Relationship between opportunity to learn and student performance on English and Algebra assessments. *Educational Assessment Journal, 10*(4), 147-172. https://doi.org/10.1207/s15326977ea1004_1

Boyle, B., Lamprianou, I., & Boyle T. (2005). A longitudinal study of teacher change: What makes professional development effective?, Report of the second year of the study. *Journal of School Effectiveness and School Improvement*, *16*(1), 1-27. https://doi.org/10.1080/09243450500114819

Brisk, M. E. (2015). *Engaging students in academic literacies: Genre-based pedagogy for K–5 classrooms.* New York, NY: Routledge. https://doi.org/10.4324/9781317816164

Brophy, J., & Good, T. L. (1986). Teacher behavior and student achievement. In M. C.Wittorck (Ed.), *Handbook of research on teaching* (3rd ed., pp. 328–375). New York: Macmillan.

Christie, F. (1999, Winter). Genre theory and ESL teaching: A systemic functional perspective. *TESOL Quarterly*, *33*(4), 759-764. https://doi.org/10.2307/3587889

Christie, F., & Derewianka, B. (2008). *School discourse: Learning to write across the years of schooling.* London: Continuum.

Common Core State Standards Initiative. (2010). *English language arts standards.* Retrieved from http://www.corestandards.org/

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research, 49*(2), 222-251. https://doi.org/10.3102/00346543049002222

Cummins, J. (2014). Beyond language: Academic communication and student success. *Linguistics and Education, 26,* 145-154. https://doi.org/10.1016/j.linged.2014.01.006

Dabach, D. B. (2014). "I am not a shelter!": Stigma and social boundaries in teachers'

accounts of students' experience in separate "sheltered" English learner classrooms. *Journal of Education for Students Placed at Risk (JESPAR)*, *19*, 98-124. https://doi.org/10.1080/10824669.2014.954044

Dabach, D. B. (2015). Teacher placement into immigrant English learner classrooms: Limiting access in comprehensive high schools. *American Educational Research Journal, 52*(2), 243-274. https://doi.org/10.3102/0002831215574725

Deshler, D. D., Palincsar, A. S., Biancarosa, G., & Nair, M. (2007). *Informed choices for struggling adolescent readers: A research-based guide to instructional programs and practice*. NewYork: International Reading Association.

Echevarria, J., Short, D., & Powers, K. (2006). School reform and standards-based education: A model for English-language learners. *Journal of Educational Research, 9*(4), 195-210. https://doi.org/10.3200/JOER.99.4.195-211

Elliot, S. N. (2015). Measuring opportunity to learn and achievement growth: Key research issues with implications for the effective education of all students. *Remedial and Special Education, 36*(1), 58-64. https://doi.org/10.1177/0741932514551282

Ferris, D. (1997). The influence of teacher commentary on student revision. *TESOL Quarterly, 31,* 315-339. https://doi.org/10.2307/3588049

Feng, S., & Powers, K. (2005). The short- and long-term effect of explicit grammar instruction on fifth graders' writing. *Reading Improvement, 42*(2), 67-72.

Gebhard, M., Chen, I., & Britton, B. (2014). Miss, nominalization is a nominalization: English language learners' use of SFL metalanguage and their literacy practices. *Linguistics and Education, 26,* 106-125. https://doi.org/10.1016/j.linged.2014.01.003

Gebhard, M., Harman, R., & Seger, W. (2007). Reclaiming Recess: Learning the Language of Persuasion. *Language Arts, 84*(5), 419-430.

Gee, J. P. (2003). Opportunity to learn: A language-based perspective on assessment. *Assessment in Education, 10*(1), 27-46. https://doi.org/10.1080/09695940301696

Gee, J. P. (2008). A Sociocultural perspective on opportunity to learn. In P. Moss, D. Pullin, J. Gee, E. Haertel, & L. Young (Eds.), *Assessment, Equity, and Opportunity to Learn* (Learning in Doing: Social, Cognitive and Computational Perspectives, pp. 76-108). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511802157.006

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. (2002). *Scaffolding language, scaffolding learning: Teaching second language learners in the mainstream classroom.* Portsmouth, NH: Heinemann.

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

https://doi.org/10.3102/01623737017003323

Halliday, M. A. K. (1975*). Learning how to mean: Explorations in the development of language.* London: Edward Arnold. https://doi.org/10.1016/B978-0-12-443701-2.50025-1

Halliday, M. A. K. (1994). *An introduction to functional grammar* (2nd ed.). London: Edward Arnold.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.

Haneda, M. (2014). From academic language to academic communication: Building on English learners' resources. *Linguistics and Education, 26,* 126-135. https://doi.org/10.1016/j.linged.2014.01.004

Harklau, L. (2002). The role of writing in classroom second language acquisition. *Journal of Second Language Writing, 11*, 329-350.

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

Hill, H. C., & Ball, D. L. (2004). Learning mathematics for teaching: Results from California's Mathematics Professional Development Institutes. *Journal of Research in Mathematics Education, 35,* 330-351. https://doi.org/10.2307/30034819

Johns, A. (1997). *Text, role, and context.* New York: Cambridge University Press. https://doi.org/10.1017/CBO9781139524650

Kurz, A. (2011). Access to what should be taught and will be tested: Students' opportunity to learn the intended curriculum. In S. N. Elliott, R. J. Beddow, & A. Kurz, (Eds.), *Handbook of accessible achievement tests for all students: Bridging the gaps between research, practice, and policy* (pp. 99-129). New York, NY: Springer. https://doi.org/10.1007/978-1-4419-9356-4_6

Kurz, A., Elliott, S. N., Kettler, R. J., & Yel, N. (2014). Assessing students' opportunity to learn the intended curriculum using online teacher log: Initial validity evidence. *Educational Assessment, 19,* 159-184. https://doi.org/10.1080/10627197.2014.934606

Lee, J., & Wong, K. K., (2004). The impact of accountability on racial and socioeconomic equity: considering both school resources and achievement outcomes. *American Education Research Journal, 41*(4), 797-832. https://doi.org/10.3102/00028312041004797

Lee, V. E., & Smith, J. B. (1999). Social support and achievement for young adolescents in Chicago: The role of school academic press. *American Educational Research Journal, 36*, 907-945. https://doi.org/10.3102/00028312036004907

Martin, J. R., & Rose, D. (2005). Designing literacy pedagogy: Scaffolding asymmetries. *Continuing discourse on language*, 251-280.

McArdle, J. J., & McDonald, R. P. (1984). Some algebraic properties of the Reticular Action Model for moment structures. *British Journal of Mathematical and Statistical Psychology*, *37*(2), 234-51. https://doi.org/10.1111/j.2044-8317.1984.tb00802.x

McDonnell, L. M., Burstein, L., Ormseth, T., Catterall, J., & Moody, D. (1990). *Discovering what schools really teach: Designing improved coursework indicators* (Report No. R3886). Santa Monica, CA: RAND Corporation. https://doi.org/10.1037/e670602011-001

Monte-Sano, C., & De La Paz, S. (2012). Using writing tasks to elicit adolescents' historical reasoning. *Journal of Literacy Research, 44*(3), 273-299. https://doi.org/10.1177/1086296X12450445

Moore, J., & Schleppegrell, M. (2014). Using a functional linguistics metalanguage to support academic language development in the English Language Arts. *Linguistics and Education, 26,* 92-105. https://doi.org/10.1016/j.linged.2014.01.002

Muller, C., Riegle-Crumb, C., Schiller, K. S., Wilkinson, L., & Frank, K. A. (2010). Race and academic achievement in racially diverse high schools: Opportunity and stratification. *Teachers College record (1970)*, *112*, 1038.

Muthén, B. O., Kao, C. F., & Burstein, L. (1991). Instructionally sensitive psychometrics: Application of a new IRT-based detection technique to mathematics achievement test items. *Journal of Educational Measurement, 28,* 1-22. https://doi.org/10.1111/j.1745-3984.1991.tb00340.x

Newmann, F. M., King, M. B., & Rigdon, M. (1997). Accountability and school performance: Implications from restructuring schools. *Harvard Educational Review, 67*(1), 41-74. https://doi.org/10.17763/haer.67.1.14141916116656q6

Nylund, K. L., Asparouhov, T., & Muthén, B. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal*, *14,* 535-569. https://doi.org/10.1080/10705510701575396

Porter, A.C. (2002). Measuring the content of instruction: Uses in research and practice. *Educational Researcher, 31*(7), 3-14. https://doi.org/10.3102/0013189X031007003

Proctor, P., Dalton, B., Uccelli, B. G., Mo, E., Snow, C. E., & Neugebauer, S. (2011). Improving comprehension online: Effects of deep vocabulary instruction with bilingual and monolingual fifth graders. *Reading and Writing, 24*, 517-544. https://doi.org/10.1007/s11145-009-9218-2

Rose, D., & Martin, J. R. (2012). *Learning to Write, Reading to Learn: Genre, Knowledge and Pedagogy in the Sydney School*. Sheffield, UK: Equinox Publishing.

Sasaki, M., & Hirose, K. (1996). Explanatory variables in EFL students' expository writing. *Language Learning, 46,* 137-174. https://doi.org/10.1111/j.1467-1770.1996.tb00643.x

Schleppegrell, M. J. (2004a). Teaching academic writing to English learners. *University of California Linguistic Minority Research Institute (UCLMRI) Newsletter*, *13*(2), 1-2.

Schleppegrell, M. J. (2004b). *The language of schooling: A functional linguistics perspective.* Mahwah, NJ: Erlbaum. https://doi.org/10.4324/9781410610317

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. https://doi.org/10.2307/3588259

Shanahan, T., & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: Rethinking content-area literacy. *Harvard educational review*, *78*, 40-59. https://doi.org/10.17763/haer.78.1.v62444321p602101

Spycher, P. (2007). Academic writing of adolescent English learners: Learning to use 'although.' *Journal of Second Language Writing, 16*, 238-254. https://doi.org/10.1016/j.jslw.2007.07.001

Stevens, R. A., Butler, F. A., & Castellon-Wellington, M. (2000). *Academic language and content assessment: Measuring the progress of English language learners (ELLs)* (CSE Technical Report No. 552). Los Angeles: University of California, National Center for Research on Evaluation, Standards, Student Testing.

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.

Uccelli, P., Barr, C. D. , Dobbs, C. L., Galloway, E. P., Meneses, A., & Sanchez, E. (2015). Core academic language skills: An expanded operational construct and a novel instrument to chart school-relevant language proficiency in preadolescent and adolescent learners *Applied Psycholinguistics 36,* 1077-1109. https://doi.org/10.1017/S014271641400006X

Wang, J. (1998). Opportunity to learn: The impacts and policy implications. *Educational Evaluation and Policy Analysis, 20*(3), 137-156. https://doi.org/10.3102/01623737020003137

Wong-Fillmore, L., & Snow, C. E. (2000). *What teachers need to know about language.* ERIC Clearinghouse on Languages and Linguistics Special Report. US Department of Education, Office of Educational Research and Improvement.

Wong, K. K., & Lee, J. (1998). Interstate variation in the achievement gap among racial and social groups: Considering the effects of school resources and classroom practices. In K. K. Wong (Ed.), *Advances in educational policy* (Vol. 4, pp. 119-144). Greenwich, CT: JAI Press.

Copyright Disclaimer