Read Writ (2014) 27:1327–1352
DOI 10.1007/s11145-013-9489-5

Signaling organization and stance: academic language use in middle grade persuasive writing

Christina L. Dobbs

Published online: 22 October 2013

Abstract Effective academic writing is accessible to readers because writers follow shared conventions for organization and signal their stance on particular topics; however, few specifics are known about how middle graders might develop knowledge of and use these academic language forms and functions to signal their organization and stance in persuasive essays. This study examined how differences in organization and stance marker use was related to writing quality in 664 persuasive essays written by 176 sixth, seventh, and eighth grade students. Essays were collected in the context of a supplemental vocabulary program, transcribed and analyzed for length by researchers, scored for overall writing quality by a team of teachers, and then coded for markers of organization and stance by researchers. Multilevel modeling results reveal that two specific emergent organizational marker types (evidence markers and code glosses) have statistically significant negative relationships to quality, and the variety of stance markers used is a positive predictor of quality when an interaction with length is included in the model. Findings give insight into the ways students are using organization and stance markers and point toward these language forms and functions as potentially pedagogically-relevant and worth assessing.

Keywords Persuasive essay · Adolescent literacy · Academic language · Writing quality · Organization · Stance

Introduction

Writing is a major mechanism by which students are expected to demonstrate their academic skills, reflecting their capacity to read, research, evaluate, and synthesize

C. L. Dobbs (✉)
Boston University, 2 Silber Way, Rm 241, Boston, MA 02215, USA
e-mail: cdobbs@bu.edu

information (Graham & Perin, 2007a). Given the Common Core State Standards' emphasis on writing (2010) and its importance in college and job success (ACT, 2005; National Commission on Writing, 2004), US students' widespread struggles with academic writing across grades and demographic categories are concerning (Ferretti, MacArthur, & Dowdy, 2000; Graham, 1990; Graham, MacArthur, Schwartz, & Page-Voth, 1992; Salahu-Din, Persky, & Miller, 2008; Troia & Graham, 2002). Despite the clear importance of writing, historically more attention and discussion have focused on reading skill, with fewer studies devoted to writing (Coker & Lewis, 2008; Graham & Perin, 2007a). Researchers have more to learn about the development of specific writing skills, especially with underprivileged student groups (Chall & Jacobs, 1983; Graham & Perin, 2007a, b).

More precise information about student writing development could be learned by in-depth study of particular features, as writing is typically scored with holistic rubrics that provide a more global score (Llosa, Beck, & Zhao, 2011). This study focuses on two central aspects of persuasive writing, middle school students' use of organizational and stance markers. It is worthwhile to study how these students utilize these particular academic language (AL) features in their writing and the relationships between uses of these markers and writing quality. In order to better understand how particular writing features are deployed, more fine-grained ways of analyzing student writing must be developed and employed. Additionally, this study explores the writing across topics of a group of middle graders from low socioeconomic-status backgrounds, a group that has historically underperformed on writing tasks and whose use of particular features is worth exploring in depth.

The persuasive writing task

The persuasive writing task is an ideal context in which to conduct an examination of AL features used by middle graders, as it is a task commonly employed by middle grade teachers. Middle school students[1] are often asked to write persuasive essays, requiring them to use evidence to support their claims and persuade an audience. A persuasive written argument involves strong understanding of the subject matter and the demonstration of perspective-taking and critical thinking skills (Gorman, Purves & Degenhart, 1988; Toulmin, 2003). The task can also come with AL demands, as in the context of this study, and ultimately, academic persuasive writing requires skills similar (though developmentally less sophisticated) to later argumentative writing (Dombek & Herndon, 2004; Hillocks, 2002).

The ability to make a claim and support it with relevant evidence, skills developed when writing persuasive essays in the middle grades, is receiving increasing attention in the new standards developed by the Common Core State

[1] Prior to the middle grades, students often have little exposure to non-narrative writing tasks until a sudden shift in the secondary grades (Applebee, Langer, & Mullis, 1986; Pritchard & Honeycutt, 2006). Though a variety of genres require some elements of persuasion, in the middle grades, the persuasive essay seems to be a sort of mini-genre in its own right, with many resources for practitioners focusing units on the writing of a persuasive task (See Robb, 2012 for a sample guide).

Standards Initiative (2010). Additionally, many writing measures at the middle grades have used persuasive tasks as the genre for assessing secondary writing proficiency (College Board, 2011; Salahu-Din et al., 2008). For instance, in Massachusetts, where the data for this study were collected, middle school writing proficiency has been assessed using a persuasive writing prompt[2] (Massachusetts Comprehensive Assessment System, 2010).

In the studies that have focused on persuasive writing skill, few have analyzed proficiency with particular linguistic or grammatical skills, but factors such as use of cohesive devices (Crowhurst, 1987), use of rhetorical devices (Connor, 1990), topic knowledge (Chesky and Hiebert 2001; Langer, 1984), syntactic complexity and lexical diversity (Beers & Nagy, 2007; McNamara, Crossley, & McCarthy, 2010), and disciplinary reasoning (De La Paz, 2005) have been found to contribute to higher quality writing. One additional study of the academic persuasive writing of high schoolers found that organizational markers, markers which signify sequences of ideas in a piece, and some markers of stance, conceptualized as linguistic markers that convey attitudes toward propositions, contributed to writing quality beyond the contribution of lexical and syntactic variables (Uccelli, Dobbs, & Scott, 2013). This study extends those findings and attempts to fill a gap in the research by closely analyzing the language resources middle school students from low socioeconomic-status backgrounds marshal as they write to persuade.

Developing academic language skills

Writing persuasively in academic tasks can require the use of a particular register of language—academic language. AL is a register of language[3] typically associated with school and academic contexts, and it is often conceptualized as a question of vocabulary.[4] Academic language consists of a variety of features including vocabulary, in addition to an authoritative detached stance (e.g., passive voice), complex sentence structure (e.g., embedded clauses), different types of vocabulary (e.g., lower-frequency, domain-specific vocabulary), stepwise logical argumentation, and markers that indicate the writer's rationale (Bailey, 2007; Schleppegrell, 2001; Snow & Uccelli, 2009). These features of AL span a variety of levels, from the lexical to the rhetorical and require the flexibility to determine when a particular form is most appropriate (Nir & Berman, 2010; Ravid & Tolchinsky, 2002). The ability to package ideas in an academic fashion that demonstrates writers' expertise about the task (Hyland & Tse, 2007; Scarcella, 2003; Schleppegrell, 2001; Snow & Uccelli, 2009), and these skills are refined and integrated over time, with middle graders beginning to learn to effectively deploy these features.

[2] The holistic writing rubric used for the purpose of assessing writing quality is the rubric used to measure writing quality on the MCAS exam.

[3] For the purposes of this study, we will focus exclusively on written language to the extent possible, though there are certainly important oral forms of the AL register that students also develop over the course of the secondary grades.

[4] See Bar-Ilan and Berman (2007), Corson (1997), Nagy and Townsend (2012), for detailed discussion of academic vocabulary.

This study adopts a sociocultural and pragmatics-based view of language to consider how language develops as students expand the contexts in which they participate and thus expand their need for new language resources (Snow & Uccelli, 2009) by considering the task itself and the larger academic context (Bereiter & Scardamalia, 1987; Blum-Kulka, 2008; Flood & Lapp, 2000; Gee, 2008; Hayes, 1996, 2000; Kellogg, 2008; Nystrand, 2006; Prior, 2006; Pritchard & Honeycutt, 2006). This work is also drawn from prior work in functional linguistics approaches (Berman & Ravid, 2009; Halliday & Matthiesen, 2004; Schleppegrell, 2004).

A strategy to assess two key domains of academic language

The constellation of academic language skills developed during the middle grades have been theorized as core academic language skills (CALS), which can be defined as "knowledge, deployment, and awareness of a set of prevalent language forms and functions that co-occur with learning-related tasks across content areas and across modalities at school" (Uccelli, Barr, Dobbs, Phillips Galloway, Meneses, & Sánchez, under review, p. 24). CALS includes several domains, including organizational strategy and appropriate stance, which are the focus of this study. An analytic strategy for these two domains was developed and modified from prior work in metadiscourse and from developmental linguistics studies on propositional attitudes.

Signaling organization and stance

In academic writing, students must convey an organized sequence of ideas to a reader, and one way they do so is by using organizational discourse markers to orient the reader and to signal relationships between individual clauses, as well as larger units of text such as paragraphs (Givón, 1992; Van de Kopple 1985). Research has demonstrated how effective deployment of these markers explicitly demonstrates coherence in a piece. This study draws from the theory of organization introduced by Hyland (2005), who terms these markers *interactive metadiscourse*.[5] This study focuses on the following markers:

- Transition markers: mark relations between clauses (e.g., *but, or, therefore*[6])
- Frame markers: refer to sequences or stages (e.g., *first, second, next*)
- Goal markers[7]: refer to the overall organization of the text (e.g., *teen smoking is wrong for three reasons*)
- Conclusion markers: mark the conclusion of a piece (e.g., *in sum, in conclusion*)

[5] The original metadiscourse framework was developed to understand how a mature writer communicated with readers using explicit linguistic markers (Hyland, 2005; Van de Kopple, 1985).

[6] Samples listed in parentheses throughout this section are drawn from the dataset.

[7] Hyland's list of markers contains several categories, not all of which were found to be relevant in studying the writing of novice AL users in pilot coding. Hyland (2005) treats goal markers and conclusion markers as types of frame markers. In the middle school writing studied here, it seemed that goal markers of overall organization, conclusion markers, and general frame markers operated differently, so they were treated as separate categories.

- Code glosses: mark examples, paraphrases, or meanings to elaborate (e.g., *for example, namely*)
- Evidentials: refer to sources of information, including text (e.g., *in the passage, according to scientists*)

Additionally, this study looked at how novice academic writers marked elements of their stance to indicate their perspective. Conveying stance in academic tasks often involves being assertive about one's reasoning while being epistemically cautious. The term stance typically refers to the ways that a writer demonstrates his or her attitudes toward the propositions in an argument using a variety of tools; in the context of this study, we will conceptualize a portion of this larger construct—the attitudinal markers of stance that indicate perspective as delineated by Berman and colleagues[8] (Berman, Ragnarsdóttir, & Strömqvist, 2002; Berman & Ravid, 2009; Reilly, Zamora, & McGivern, 2005). Prior research in the secondary grades by the author and others (Dobbs, under review; Uccelli, Dobbs, & Scott, 2013) demonstrated that these elements in particular were used in interesting and relevant ways by developing AL users and that these uses had important relationships with quality. The work of Berman and colleagues on propositional attitudes offered a useful frame in considering how novice writers convey stance (Reilly, Baruch, Jisa, & Berman, 2002), and this study used two categories they developed that were salient to middle grade academic writing[9]:

- Deontic markers: markers which convey a categorical attitude or judgmental attitude toward a claim (e.g., *people must, it is wrong*)
- Epistemic markers: markers which convey the writer's attitude about the certainty or possibility of a claim (e.g., *it is true, it might be*).

In addition to these facets of writing, some from the metadiscourse frame were used to more fully capture the ways middle graders were explicitly conveying perspective through stance. Hyland (2005) uses the term *interactional metadiscourse* to refer to the markers of these attitudes, and the following types of stance markers were coded:

- Hedges: show a cautious attitude (e.g., *likely, potentially*)
- Boosters: demonstrate a clear attitude (e.g., *definitely, absolutely*)
- Engagement markers: explicitly reference the reader or attempt to engage the reader (e.g., the inclusive *we, you*)

Though hedges and boosters are similar to epistemic and deontic markers, in coding novice academic writing, these categories of booster and hedge were

[8] This work is drawn from earlier work on the written and spoken language distinctions laid out by Chafe (1994).

[9] The notion of stance is much more complex than the small number of explicit features explored for the purposes of this study, and it often includes elements of detachment, as well as uses of modal expressions that are not included in the deontic and epistemic types mentioned here. Pilot study of the writing of middle graders found deontic markers and epistemic markers to be particularly frequently used and relevant in persuasive tasks, though not an exhaustive means of coding for stance.

reserved only for adverbial expressions of certainty or caution (For a full explication of the coding scheme, see "Appendix").

Prior to this study, it seemed clear that these markers had the potential to improve scorers' impressions of student writing quality. Studies of metadiscourse, which have typically examined texts by older adolescents and adults, find that organizational and stance elements contribute to the clarity and overall quality of the text (Hyland, 1998, 2002a, b, 2006, 2007, 2008). It is important to note, however, that the texts often considered the most persuasive contain moderate levels of both types of metadiscourse markers, a finding in keeping with our own study of high school writers (Uccelli, Dobbs, & Scott, 2013; Dafouz-Milne, 2008). However, it was less clear how these markers, and stance markers in particular, would be associated with quality in persuasive essays in the writing of younger students.

A small number of studies have analyzed metadiscourse use by young writers beyond the study of high school writing mentioned previously. These studies include an initial pilot study (Dobbs, under review), which was used to develop the coding scheme, and a second study of 65 essays (Author, under review), which found that seventh grade students deployed organizational and stance markers with varying degrees of effectiveness in a small sample of persuasive drafts. The most interesting substantive finding from these early studies was that, even though the essays were not very high quality overall, students were using these markers in their persuasive drafts.

It is a basic principle of language development that form and function development are intertwined, classically stated as "new forms first express old functions, and new functions are first expressed by old forms" (Slobin, 1973, p. 184). These initial pilot studies revealed examples of students using both advanced, more formal forms in emergent ways and overly informal language to express complex ideas. For instance, complex transition forms such as *whereas* and *however* were present in the data but often used to connect sentences whose meanings were not contrastive in nature. Students also used informal forms, such as text abbreviations (e.g., *LOL*) or other types of informal language (e.g., saying a proposition is '*mad right*') to convey stance. These findings suggest that a fine-grained approach is promising in revealing the incremental nature of developing marker use, rarely captured in the holistic assessment of adolescents' writing.

Research questions

The current study assesses the dimensions of organization and stance language used by a sample of middle schoolers, writing across a variety of persuasive topics in an unusually large corpus of data. The study addresses the following questions:

1. How often do students use academic language markers of organization and stance in their persuasive essays on a variety of topics?
2. Controlling for essay length and grade level, is the use of organizational or stance markers associated with overall writing quality in middle graders' persuasive essays across a variety of topics?

Current study methods

Setting and curriculum context

Data in this study were collected in 2009–2010 from a large urban middle school located in the Northeast. The school implemented a supplemental academic vocabulary program called Word Generation (2010), a program developed and supported by the Strategic Education Research Partnership and that has been shown effective at building vocabulary (Snow, Lawrence, & White, 2009).

Word Generation is a 24-week vocabulary program, with the aim of supporting students in acquiring high-utility academic vocabulary words (Word Generation, 2010). Each weekly unit introduced students to general academic vocabulary words from the Academic Word List (Coxhead, 1998) in the context of a high-interest, controversial topic. Instruction was delivered by teachers across the four major content areas. The program's design allows for cross-sectional exploration of sixth, seventh and eighth grade writings about the same topic, because students were writing frequently and the same topic was taught across levels simultaneously. In the final weekly activity, students drafted short essays about their positions on each topic in approximately 12–15 minutes, wherein they were encouraged to use academic language.

Participants

The school's population consists of approximately 550 students, who are diverse over a number of demographic dimensions, though the population was primarily low socioeconomic status. The curriculum materials consisted of two workbooks, each containing 12 units, and all available workbooks were collected from the site. Students' essays were included in the data corpus for this study if both workbooks were collected and demographic information was available. The essay corpus consisted of essays written at 5 time points over the first 15 weeks of the program. Students were retained in the sample for analysis if they contributed at least 2 possible essays from the units being analyzed.

The result was a participant sample consisting of 67 sixth graders, 60 seventh graders, and 40 eighth graders, for a total of 167 students (see Table 1 for details about sample demographics). The participants in this sample were generally reflective of the overall school population and were gender-balanced and linguistically and ethnically diverse. They also spanned the achievement spectrum in their performance on the English/language arts portion of the Massachusetts state test. The ethnic breakdown of the sample was 54 % African American, 29 % Latino, 8 % White, and 8 % Asian American.[10] The majority of students in the sample did not receive services for English language learners (72 %), while 15 % were classified as limited English proficient (15 %) and 13 % as formerly limited English proficient (13 %), meaning they were still likely becoming fully fluent in English. The vast majority of students in the sample (81 %) qualified for free or reduced lunch, a rough proxy for low socioeconomic status.

[10] A single student chose the 'Other' category for race/ethnicity.

Data corpus

The total corpus contained 664 essays (76,493 words) written by 167 students. The essays that make up this corpus were collected at weeks 1, 5, 8, 12, and 15 of the curriculum in order to assess student marker use across an array of topics, and the five different topics for these weeks were from a variety of domains (see Table 2 for a full list of topics, essay prompts, and taught academic words). Topics included the purpose of school, censorship of hip hop music, causes of global warming, junk food in schools, and responsibility for teen smoking, an array about which students were likely to have varying prior knowledge.[11] Qualitative reports from teachers and data from workbooks showed high levels of participation during those weeks.

The final data corpus consisted of 284 essays written by sixth graders, 222 by seventh graders, and 158 by eighth graders. The greatest number of essays was collected during Week 8 (global warming) and the fewest at Week 15 (teen smoking). Each collection time point varied in the number of essays collected, due to a variety of reasons that might include student absences, field trips, or choices not to write essays on particular days (see Table 3 for a full description of essay distribution totals).

Data transcription and analysis

Originally each essay was handwritten on lined pages in the workbooks. Essays were transcribed into Microsoft Word to facilitate the quality scoring process and to avoid bias due to handwriting differences. Then to facilitate automated linguistic analysis, essays were transcribed using CHAT (Codes for the Human Analysis of Transcripts) conventions in the CLAN program (MacWhinney, 2011). Following the transcription process, the following measures were used to analyze the data:

Length, lexical diversity, and syntactic complexity

Basic length and lexical diversity measures for each essay were generated using CLAN, including number of words (Tokens), variety of words (Types), and ratio of types to tokens (Type–Token Ratios). As a more useful means of measuring length, essays were divided into clauses by the research team. The clausing process utilized a simple definition of a clause laid out by Berman and Slobin (1994); they defined a clause as "a unified predicate dealing with a single situation" (p. 660). Clause demarcation was governed by the same process used in the author's prior study of high school writing (2013), and all clause coding was verified by at least two other researchers. Then CLAN was used to generate length as the number of clauses per essay, the measure ultimately used for analyses based on early descriptive analyses and guidance from previous studies. After clauses were defined for each essay, syntactic complexity was estimated. It was defined as average number of words per clause.

[11] It is important to note that though several of the essay prompts do not begin with modal verbs, students had extensive exposure to the variety of perspectives that could be taken when discussing these topics through the curricular materials taught earlier in the week. Particular arguments in response to these questions had been explored in a variety of ways over the course of the instruction.

Table 1 Demographic information for students in sample (n = 167)

	Number of students (%)
Grade	
6th grade	66 (40)
7th grade	60 (36)
8th grade	40 (24)
Gender	
Female	81 (49)
Male	86 (51)
Ethnicity	
African American	89 (54)
Asian American	14 (8)
Latino	49 (29)
White	14 (8)
Other	1 (>1)
English proficiency status[a]	
Receives no ELL services	121 (72)
Formerly limited english proficiency	21 (13)
Limited english proficiency	25 (15)
Socioeconomic status	
Receives free or reduced lunch	136 (81)
Does not receive free/reduced lunch	31 (19)
MCAS achievement level	
Advanced	8 (5)
Proficient	93 (56)
Needs improvement	51 (30)
Warning	15 (9)

[a] These are designations laid out by the district. It is likely that many students who did not receive ELL services at the time of the study were bilingual or multilingual, and it is unknown how many received ELL services at some point over the course of their school careers prior to the years surrounding this study time frame

Organizational markers

Essays were coded for organizational markers using a coding scheme described previously and drawn from prior work by Hyland (2005) and adapted to be more developmentally appropriate and attuned to markers typically used by middle graders, and including goal markers, frame markers, conclusion markers, code glosses, evidentials, and transition markers.

Essays were coded by a team of three, who underwent a training session and fairly extensive coding practice using essays not analyzed as a part of this data corpus. Then coders went through formative reliability scoring to establish interrater reliability. Coders scored 20 % of the essays with high levels of reliability, yielding a Cohen's kappa statistic of .95.

Stance markers

Markers of stance were coded using a similar procedure. The coding scheme described earlier was drawn from work by Hyland (2005) and Reilly et al. (2002).

Table 2 Topics/essay prompts selected for the data corpus

Week No.	No. of essays	Essay prompt	AWL words taught
1	138	What is the purpose of school?	Analyze, factor, function, interpret, structure
5	145	Should hip hop music be censored?	Considerable, contribute, demonstrate, sufficient, valid
8	150	Who is responsible for stopping global warming?	Attribute, cycle, hypothesis, project, statistics
12	139	Should junk food be sold in schools?	Acknowledge, incidence, incorporate, initiative, transport
15	92	Who is responsible for the problem of teen smoking?	Accumulation, contradict, exhibit, inevitable, manipulate

Table 3 Number of essays in the data corpus by grade and topic/time (n = 664)

Grade	No. of students	Time 1: Week 1	Time 2: Week 5	Time 3: Week 8	Time 4: Week 12	Time 5: Week 15	Total
6th	67	55	62	62	64	41	284
7th	60	48	53	52	37	32	222
8th	40	35	30	36	38	19	158
Totals	167	138	145	150	139	92	664

Coded markers included deontic modality markers, epistemic modality markers, boosters, hedges, and engagement markers. The three coders went through similar processes to train and practice this coding, and interrater reliability coding yielded a Cohen's kappa statistic of .92.

Writing quality

In order to understand markers in relationship to the overall essay, essays were scored for overall quality using the rubric from the MCAS.[12] This rubric is designed and used to holistically measure quality on a state assessment of achievement. The MCAS rubric is divided into two analytic trait scores, for topic/idea development (TISCORE) and conventions (CSCORE). A third score, designed to represent the global quality score, was created by the researcher by summing these two scores to create a total quality score (TOTSCORE).

Essays were scored by a pair of scorers who had experience at the middle school level. To ensure that their judgment of quality was not biased, scorers did not have

[12] The MCAS assessment is administered in Massachusetts to assess student achievement in a variety of subject areas. The writing assessment asks students to write in response to a prompt, and the style of writing varies across grade levels. At the middle school level, students are typically asked to write in a persuasive or expository style. Regardless of the grade level or the type of task assigned, the MCAS rubric remains the same, and is the one used in this study. Essays are scored holistically, with students receiving a single number score to encompass their achievement in topic/idea development and a single score to capture their knowledge of conventions. No more specific feedback is given.

knowledge of the other study components; they underwent training in scoring and practiced extensively using essays not included in the dataset. Using a procedure quite similar to the MCAS scoring, scorers had to have exact or adjacent score agreement; adjacent agreement was allowed in order to maximize score variability. Essays with more than a 2-point score differential were sent to a third scorer, whose score was doubled. A total of four essays were sent to a third reader, constituting less than 1 % of the sample.

Results

Descriptive and correlational analyses

First, a series of descriptive analyses were conducted. On the holistic quality measures, essays had an average topic/idea score of 6.40 out of a 12-point scale, an average conventions score of 4.98 of a possible 8, and an average total score of 11.38 of a possible 20 (see Table 4 for descriptive statistics describing quality measures). Quality scores were highly correlated across the quality measures (see Table 5 for correlational data for the variables of interest in this study). Analyses were always conducted using each of the three possible outcome measures, and relationships of interest are reported here.

The students who scored highest were the 66 sixth graders, regardless of the measure being used, a phenomenon worth exploring in more depth in the discussion. Their average on the topic/idea development rubric was 6.87, with seventh and eighth graders performing lower on average (means = 6.07 and 6.03, respectively). The pattern held on the conventions rubric, though the scores were closer on a rubric with a smaller possible range. Finally, on the total score measure the sixth grade students' essays were highest in quality, and they had the smallest variability in scores (mean = 12.01, $SD = 2.30$).

The sixth graders in this sample also wrote the longest essays, whether clauses or tokens are used as the measure of length (see Table 6 for measures of length). If we consider average essay length by topic, the essays written about teen smoking, from time point 5, are the longest with the largest standard deviation (mean tokens = 125.12, $SD = 55.94$), and the essays written about the purpose of school, Time 1, are the shortest (mean tokens = 109.62, $SD = 50.34$).

Additional analyses were conducted, exploiting the longitudinal nature of the dataset, to determine whether there was a systematic trend in essay quality over time, due to either topic differences that influenced overall quality or to an overall trend in improvement over the 15 program weeks.[13] This led to the conclusion that

[13] First, to assess differences between means for the five essay time points, pairwise comparisons of means (using Tukey's honestly significantly different test) were conducted. Only two topics were identified as having significantly different means, topics 3 (global warming) and 5 (teen smoking). Other pairs showed no statistically significant mean differences in overall writing quality. Next, an investigation was conducted to discern whether a longitudinal trend existed in the data that would need to be accounted for; preliminary longitudinal analyses were conducted to assess whether there were systematic differences in how well topic predicted overall quality. The subsequent fitted model showed that topic/time point was

Table 4 Descriptive statistics for measures of overall essay quality (n = 664)

	M (SD)	Possible range	Min.	Max.	6th grade *M* (*SD*)	7th grade *M* (*SD*)	8th grade *M* (*SD*)
Topic and idea score (TISCORE)	6.40 (1.76)	2–12	2	11	6.87 (1.70)	6.07 (1.71)	6.03 (1.75)
Conventions score (CSCORE)	4.98 (1.21)	2–8	2	8	5.13 (1.14)	4.81 (1.34)	4.92 (1.11)
Total score (TOTSCORE)	11.38 (2.44)	4–20	5	19	12.01 (2.30)	10.87 (2.50)	10.94 (2.39)

there was no systematic difference over time in using topic to predict quality. Consequently, the multilevel models generated were not treated as longitudinal. Subsequent analyses were conducted using multilevel approaches that treated the data as series of essays nested within individual students.

RQ1: Organizational and stance marker frequency

In answer to research question one, the first analysis conducted was to determine the frequency of markers used. First, a total of 4,684 organizational markers appear in the data corpus. Over half of those markers were transition markers, simple, non-temporal connectives between clauses. The other five types of markers were used far less frequently, with goal markers, markers that express the overall intent of a piece, (n = 166) and conclusion markers (n = 285) appearing least often (see Table 7).

A total of 645 of the 664 essays in the data corpus contained at least one instance of an organizational marker. The number of organizational markers per essay ranged from zero to 31.[14] Essays averaged a use of seven organizational markers per essay ($SD = 4.43$). Of the six types of coded organizational markers, essays averaged at 2.8 types per essay ($SD = 1.39$). A total of 27 essays contained all six marker types, and 19 essays contained no organizational markers at all (see Table 8).

Stance markers were used slightly more frequently than organizational markers, with a total of 4,802 markers appearing in the entire data corpus. Engagement markers were most frequent with 1,385 instances across the corpus and a maximum of 28 uses of engagement markers in a single essay. Both deontic and epistemic modality markers were frequently occurring, with a total of 1,305 deontic and 1,079

Footnote 13 continued
not a significant predictor of quality and would explain less than 1 % of the variability in the outcome, and there was no systematic trajectory of overall writing quality, in fact, trajectories of writing quality varied widely by student. Attempts at modeling using additional shapes, including nonlinear forms or logarithmic transformations, were not successful at predicting quality.

[14] In the essay with 31 organizational markers, almost all were transition markers.

Table 5 Pairwise correlations between writing quality, length, organizational markers, stance markers, and vocabulary measures

	1	2	3	4	5	6	7	8	9
1. Total quality score	1								
2. Topic/idea development score	.8842***	1							
3. Conventions score	.7345***	.3325***	1						
4. Length in clauses	.4548***	.6124***	.0285	1					
5. Syntactic complexity	−.0180	−.0302	.0076	−.2432***	1				
6. Organizational tokens	.3281***	.4999***	−.0637	.7440***	.7356***	1			
7. Organizational types	.1850***	.3217***	−.0931	.4937***	.4899***	.6795***	1		
8. Stance tokens	.2349***	.3738***	−.0688	.6419***	.6221***	.5181***	.2308***	1	
9. Stance types	.2341***	.3568***	−.0457	.5425***	.5268***	.3922***	.2141***	.6499***	1

* $p < .05$, ** $p < .01$, *** $p < .0001$

Table 6 Descriptive statistics for measures of length by grade level (n = 664)

Grade	*n*	Clauses			Tokens		
		M (*SD*)	Min.	Max.	*M* (*SD*)	Min.	Max.
6th	284	19.95 (9.26)	1	55	125.12 (55.14)	24	382
7th	221	17.52 (8.39)	2	48	114.02 (51.11)	16	290
8th	159	15.15 (6.47)	1	38	99.11 (37.92)	10	245
Total	664	17.99 (8.58)	1	55	115.20 (51.11)	10	382

epistemic markers. There were a total of 785 boosters in the corpus, and hedges were by far the least frequently used, with only 248 instances across the corpus.

Essays ranged from zero stance markers to 39 total stance markers.[15] A total of 47 essays contained all five types of coded stance markers. Essays used an average of 7.2 stance markers ($SD = 4.86$). Essays used a variety of the possible stance marker, with an average of 2.90 of the possible 5 ($SD = 1.16$).

RQ2: Relationships between markers and quality

Multilevel models were built to understand the relationships between individual marker types, various constellations of markers (e.g., total number of organizational markers, total number of types of stance markers used, etc.), and holistic essay quality scores. These models are two-level, with essays on several topics nested within students. However, prior to introducing organizational or stance markers to models, a basic model with control predictors was built. Predictors included grade level, a student-level predictor, as well as length in clauses and syntactic complexity, essay-level predictors. See Table 9 for a taxonomy of fitted models using these controls. All three control predictors were, as expected, statistically significant, with essay-level predictors length ($\beta = .12$; $p < .001$) and syntactic complexity ($\beta = .09$; $p = .001$) associated with a positive increase in total essay quality. The student-level control, grade, had a small negative association with total quality scores ($\beta = -.26$; $p = .03$); given the higher performance overall of sixth grade students, this result was unsurprising.

Next, individual organizational markers were added to the control model to determine the effect of the frequency of individual markers on quality. Evidentials were the first marker to be modeled as a statistically significant negative predictor of quality ($\beta = -.15$; $p = .03$) (see Table 10 for model details for models including individual markers). This relationship was evident when using the overall quality measure and the one for topic/idea development; evidentials were not a statistically significant predictor when the conventions rubric was used to measure quality. The second type of organizational marker with a statistically significant negative relationship to overall quality was the use of code glosses, or explicit markers of

[15] Nine essays contained no stance markers, and the essay with 39 markers primarily contained engagement markers.

Table 7 Frequency of occurrence of coded organizational and stance markers across data corpus (n = 664)

Type of marker	Total coded instances	No. of essays containing at least 1 instance	Min, max uses in a single essay
Organizational markers			
Goal markers	166	153	0, 4
Frame markers	776	345	0, 8
Evidentials	394	207	0, 5
Code glosses	457	295	0, 7
Transition markers	2,606	603	0, 25
Conclusion markers	285	256	0, 2
Total organizational markers	4,684	645	0, 31
Stance markers			
Deontic modality markers	1,305	474	0, 11
Epistemic modality markers	1,079	497	0, 10
Boosters	785	403	0, 9
Hedges	248	183	0, 4
Engagement markers	1,385	377	0, 28
Total stance markers	4,802	655	0, 39
Total markers	9,486	663	0, 69

examples or paraphrases ($\beta = -.23$; $p = .005$); these relationships were similar using the topic/idea development score and the conventions score.[16]

Finally, constellations of organizational markers were used as question predictors. The frequency of organizational markers in an essay and the diversity of organizational markers were both tested in relationship to the quality outcome measures. Neither frequency nor diversity of organizational markers was a significant predictor of quality on any of the rubrics.[17]

When stance markers were used as question predictors, no individual marker was a statistically significant predictor of quality on any of the outcome rubrics. The

[16] Because of the high performance of sixth graders in the sample as compared to seventh and eighth graders, analyses were also run on only the sixth grade essays and then the seventh and eighth grade essays. In all cases, the directionality of relationships remained the same for both organizational and stance markers. However, in the sixth grade sample, evidentials were no longer statistically significant negative predictors of quality. Code glosses approached statistical significance as predictors of quality in the sixth grade sample, but the relationships between individual markers and quality were less clear in the sixth grade sample. Relationships between stance markers and quality remained the same no matter the sample, as well as relationships between constellations of markers and quality.

[17] Then, given the overwhelming frequency of transition markers in the organizational category, transition markers were separated from the other types of organizational markers and the frequency and diversity of markers retested, excluding transition markers. The frequency of organizational markers, not including transition markers, and the diversity of organizational markers, also excluding transition markers, approached significance at the .10 level, with very small negative parameter estimates. However, given the high potential for error in these estimates, it is safe to assume that overall frequency and diversity of organizational marker use were not significant quality predictors.

Table 8 Descriptive statistics for constellations of organizational and stance markers (n = 664)

	M (*SD*)	Min.	Max.
Organizational markers			
Frequency of markers	7.00 (4.43)	0	31
Diversity of markers	2.80 (1.39)	0	6
Stance markers			
Frequency of markers	7.20 (4.86)	0	39
Diversity of markers	2.90 (1.16)	0	5

frequency of stance markers was not a statistically significant predictor of quality. However, when both variety of stance markers and the interaction between length and variety of stance markers were included in the model, statistically significant results were obtained. The parameter results were as follows: variety of stance markers—$\beta = .30$; $p < .001$; interaction between length and variety of stance markers—$\beta = -.02$; $p < .001$ (see Table 11 for model details). Substantively, this means that having more types of stance markers in an essay is associated with a higher quality score, but this effect is dampened as essays get longer.

Discussion

Prior to exploring the study results more fully, it is important to note and understand the implications of the high performance by sixth graders on several different metrics in this study. One could typically expect longer and higher quality essays from students as they progress through the grades, a finding not borne out in this study. Though we can not be certain why this happened, several theories are possible in this particular context. First, it is possible that the sixth grade teachers might have implemented the program differently than other teachers or devoted slightly more time to it. Or perhaps, sixth graders, who had not participated in the program in previous years, could have been more engaged in the program overall, while older-grade students were experiencing a sort of fatigue having participated in the program for multiple years. Though we do not know specifics of what happened here, it seems likely that the high sixth grade performance was a result of instructional environment somehow, rather than a result of developmental differences.

Substantively, these analyses reveal some important points about how low-SES middle grade students marshal their language resources to write persuasively, and placing these findings alongside findings about students from higher-SES backgrounds could reveal interesting patterns for this particular populations. These short essays collected on a variety of topics reveal interesting patterns of organization and stance. First, these essays show clearly that students make many attempts at using these markers to organize their essays and assert their stance toward their arguments. Only one essay in the corpus did not have a single organizational or stance marker, meaning that all of these 167 writers, regardless of of their overall writing skill, were using organizational and stance markers as they made their cases.

Table 9 Taxonomy of fitted multilevel models describing the relationship between overall essay quality and grade, length, and syntactic complexity (n = 664)

	Multilevel model			
	*M*1	*M*2	*M*3	*M*4
Fixed effects				
Intercept	11.26***	15.68***	11.86***	10.97***
Grade		−.65***	.37*	−.26*
Length in clauses			.11***	.12***
Syntactic complexity (words per clause)				.09**
Random effects				
σ_u^2	2.90	2.66	1.99	1.97
σ_ε^2	3.12	3.11	2.70	2.65
Goodness of fit				
−2LL	2,894.21	2,882.21	2,769.34	2,757.79

* $p < .05$; ** $p < .01$; *** $p < .001$

Table 10 Taxonomy of fitted multilevel models describing the relationship between overall quality and relevant individual organizational markers, controlling for grade, length, and syntactic complexity (n = 664)

	Multilevel model *M*1	Multilevel model *M*2
Fixed effects		
Intercept	11.26***	10.99***
Grade	−.39*	−.37*
Length in clauses	.12***	.13***
Syntactic complexity (words per clause)	.09**	.10**
Evidentials	−.15*	
Code glosses		−.23**
Random effects		
σ_u^2	1.96	1.92
σ_ε^2	2.63	2.63
Goodness of fit		
−2LL	2,753.28	2,749.90

* $p < .05$; ** $p < .01$; *** $p < .001$

Despite this overwhelming evidence of attempts at using markers, the number of markers used in an essay was not a statistically significant predictor of quality, but the small-to-moderate, statistically significant correlations between organizational tokens and quality or stance tokens and quality show that students wrote across the rubric scale for overall quality using an array of organizational and stance marker tools. The mere presence of these markers does not necessarily mean that an essay is judged as being higher quality overall. These findings from middle school samples do not show what earlier studies found in high school writing, that a variety of organizational marker tools or particular types of stance markers were predictive of

Table 11 Taxonomy of fitted multilevel models describing the relationship between quality of topic/idea development and variety of stance markers (types), controlling for grade, length, and syntactic complexity ($n = 664$)

	Multilevel model		
	*M*1	*M*2	*M*3
Fixed effects			
Intercept	10.97***	4.92***	3.93***
Grade	−.26*	−.21*	−.20*
Length in clauses	.12***	.12***	.18***
Syntactic complexity (words per clause)	.09**	.08***	.09***
Variety of stance markers		.03	.30***
Length * variety of stance markers			−.02***
Random effects			
σ_u^2	1.97	.53	.51
σ_e^2	2.65	1.33	1.30
Goodness of fit			
−2LL	2,757.79	2,230.29	2,213.51

* $p < .05$; ** $p < .01$; *** $p < .001$

overall quality (Uccelli, Dobbs, & Scott, 2013), which is perhaps due to difference in task attributes,[18] age, or developmental stage. In these data, organization and stance markers behaved somewhat differently, so trends from each category will be discussed separately and in more detail.

Organization

Students tended to use some organizational markers at much higher rates than others, and not necessarily in ways that contribute to higher quality writing overall. For example, students used an overwhelming number of transition markers (2,606) throughout the essays and regardless of topic. Many of these uses were simple sentence-level connectives, such as *but*, *so*, and *yet*, used to connect clauses to each other, and this had implications for sentence length. In some cases, this overuse of transition markers led to essays filled with run-on sentences. For example, Student 181 wrote the following sentence using many simple connectives to create a run-on:

> One reason is that they should sell juck food & nutrition too because sometimes if you eat too much juck you get sick of eating it, but then agian if you eat to much nutrition you get sick too, but like alot of people start eating both they will have both sugar & also nutrition.

[18] The study of high school writing focused on a similar, but more abstract, writing task designed to model the SAT writing task, which relies more heavily on specific argumentation (See Uccelli, Dobbs, & Scott, 2013). Additionally, this study found that organizational markers were used more frequently in essays written by high school students (mean = 10.63; *SD* = 6.04). Stance markers were used much less frequently in the high school samples (mean = 2.84; *SD* = 2.49). It is important to note that the essays written for the high school study were slightly longer than those collected from middle graders.

These examples of students using many transition markers were quite common in the dataset, and the numbers would have been even higher if the simplest connective *and* had been counted. Other essays contained slightly more advanced uses of connectives, such as *whereas* and *however*.

Other organizational markers were less common, appearing in about half or less than half of the essays. A total of 394 essays contained frame markers, which ranged from simple (e.g., *first, next, then*) to more involved (e.g., *the first reason, another reason*). Other markers appeared less frequently, and goal markers, the least frequent, appeared in exactly one quarter of the essays, 166. Students are trying to explicitly signal their organization scheme to the readers of their persuasive essays, but they are using the variety of organizational tools at varying frequencies.

However, merely containing those markers of organization is not enough to ensure that an essay receives a higher quality score, as none of the markers is associated with a systematic gain in essay quality on any of the rubric measures. No constellation of organizational markers, either frequency or diversity of markers, was significantly associated with essay quality in this sample. It is likely that to see a high quality score, organization must not only be marked, but the markers must be used well. Given the important principle discussed earlier about forms and functions, it would seem that our coding scheme uncovered a variety of uses of particular forms that were developing, that is they did not convey the appropriate functions in entirely conventional ways.

In fact, when some markers are employed poorly, they can be associated with lower quality scores, which is true in the case of code glosses and evidential uses. Many students attempted to explicitly mark examples in their essays using code glosses, but many of these uses were not employed with much finesse. Often there was a mismatch between the reason provided to support an argument and the example called on to support it. For instance, Student 100 writes the following reason and supporting code gloss:

> People should buy better cars and more efficient electrical appliances in order to slow down or stop global warming. For example if we could drive less and use hybrid cars, walk, ride a bike and etc.

The mismatch here between the assertion that people should buy better cars and the example of means by which people might avoid driving is confusing for the reader and results in a muddled argument. Many of the code glosses in the corpus are similar to this one, examples that do not necessarily match up with reasons. The small but statistically significant negative association between quality scores and code glosses is likely a reflection of these mismatches and misuses. Students seem to need support in refining their work to reflect well-matched reasons and glosses.

When modeling using evidentials, the models revealed a small negative association with each of the quality measures. There were two basic types of evidence used by students—references to general parties of people such as 'scientists' or 'experts' and references to weekly curricular passages. Students consistently referred back to the weekly passage as a source of evidence, a trend seen in over half of the evidentials coded, and these sometimes demonstrated a lack of facility at integrating sources. Students seemed to have the inclination to quote

and otherwise paraphrase, an important understanding in the context of this genre, but they seemed to struggle to artfully integrate their chosen evidence. The following were evidentials used in the essays:

> On the 2nd paragraph, last sentence, it says that this earth is done if we keep ignoring it like we are now.
> Tobacco companies take part of this blame because they manipulate teens with advertisements, said exactly on ¶ 3.
> The symptoms are coughing, low energy, poor circulation, and respiratory difficulties. That information was in paragraph 1.

These uses were often supportive of the broader argument, but they were not integrated smoothly. Despite this, it seems as though it would not take immense revision to help students integrate quotes more smoothly into essays, and potentially reverse the negative association between evidential use and quality.

The other organization markers coded, goal markers, frame markers, and conclusion markers, were not statistically significant predictors of writing quality. Overall, qualitatively they showed a wide array of uses from more novice to more sophisticated. For instance, many writers used a traditional *in conclusion* to close their essays, but others used a more colloquial *that's why I think …* to close the piece. Though the more colloquial uses found here are not conventional AL uses, they do reflect a basic awareness of how essay endings are organized, despite a lack of a precise academic form. These novice uses point to clear opportunities for teachers to provide instruction so students learn more advanced forms of AL to signal organization. This explicit instruction could improve the students' usages of organizational markers.

Stance

The ways that students used coded stance markers varied somewhat from the ways that they used organizational markers. Many instances of epistemic and deontic modality markers were found, as well as high numbers of engagement markers. Students again showed that they were aware that stance should be signaled in trying to persuade a reader. In addition, more emphatic markers (deontic markers and boosters) appeared more frequently than more measured markers (epistemic markers and hedges). It is possible this is an indication of the age and/or developmental phase of the student writers studied.

Prior research showed that deontic modality markers appeared earlier in development than epistemic ones (Reilly et al., 2002), and initially, hypotheses were tested about whether epistemic markers would appear more in the essays of older writers or epistemic markers might show some association with higher quality since they are later-appearing. None of these hypotheses was borne out in the data, however, and neither epistemic nor deontic markers were statistically significant predictors of quality on any rubric. Essays with epistemic and deontic markers scored at both ends of the quality range, and students across the grades used both frequently. In fact, no individual markers were statistically significant predictors of quality when considering stance.

As in the case of organization, the frequency across types of stance markers were tested as predictors of quality, and the number of types was not a statistically significant predictor. However, when we looked at the diversity of stance markers, a different story was revealed. When the interaction term between length and variety of markers is included in the model, the addition of more types of stance markers is associated with higher quality, though that effect is dampened as essays get longer. This could be cautiously interpreted a few ways. In many of the essays in the corpus, students do not provide more than two reasons to support their overarching positions, so perhaps repeatedly using stance markers becomes redundant. It is also possible that longer essays simply contained an overwhelming number of markers; the essay with the most stance markers contained 39 markers, possibly an overwhelming number to read. These essays containing such large numbers of stance markers were often full of redundancies or extensive uses of second-person engagement markers, which were often unclear in terms of reference. These essays at the high end of the stance marker use quantity spectrum are potentially driving the interaction in the results, indicating a threshold at which additional stance markers decrease quality rather than influencing it positively. More exploration of stance might further illuminate the relationship between stance marker use and writing quality.

Overall, students tended to make emphatic arguments in these essays, expressing their ideas in absolute terms. For example, one student argued the following about hip hop music:

> Rap music is a freedom of speech, is an important american right, so rap lyrics should never be censored no matter what they may say.

These strong opinions appeared throughout the corpus, and it seems clear that asserting strong opinions is a trend for these students. However, asserting these strong opinions does not necessarily lend itself to higher quality essays overall, and essays where stance is marked in more measured ways received both high and low quality scores. Having a variety of stance tools at one's disposal can improve an essay's quality, if employed effectively.

Conclusion

A clear takeaway from this study is that students were overwhelmingly attempting to use markers to explicitly signal organization and stance. This reflects a base of developing knowledge that we must find ways to build upon to ensure that students grow and expand their skills as academic writers. This study highlights the incremental nature of language learning, something we know to be true from prior work about vocabulary learning (Nagy & Scott, 2000) extended to language learning more broadly. As students expand their repertoires of AL forms, they might use forms in ways that are not fully conventional at first, but as they come to understand a form and its function more fully, we might see writing that is clearer and more precise.

However, this study, while informative about how students use many typical AL markers, has several important limitations that could be explored in future research. First, the coding scheme employed here looked for the presence of a range of academic language forms that middle school students used to signal organization and stance in their drafts. However, it is possible that there were instances where students knew certain functions were necessary but did not mark them with accompanying forms in a way that was discernible to coders. We also did not sort the coded markers by the range of more colloquial to academic types found. Perhaps the use of simpler forms, such as *but* or *so* are not associated with quality in the same way that more sophisticated markers, *therefore* or *in sum* are, and this is certainly an important direction to explore. Additionally because this was a secondary analysis of essays, little is known about the specific classroom environments in which these data were gathered. It is not clear whether students remained in the same classrooms throughout the intervention or why they might have written or not written in particular weeks. We are unable to determine what teachers did to encourage students to write essays in particular ways. We are also unable to determine whether students from higher-SES backgrounds would have performed similarly to the sample of students studied here, a question certainly worth exploring more.

Future research would do well to closely track the instruction that students receive to better understand whether certain contexts lend themselves to more attempts at AL use. This research clearly establishes that essays written across the quality spectrum overwhelmingly make attempts to use particular markers, but when the AL structures are not well integrated into the content being discussed, we see less-clear relationships with the overall quality of the piece. Perhaps more targeted instruction could help students refine the tools they use to demonstrate their organization and stance. Or it is possible that students simply need more time and support in revising their essays to better integrate emerging forms and add new forms, possibly in a workshop- or conference-type context.

Though this study offers no particular guidance about what type of instructional strategies might help develop these practices, the findings here do point toward particular areas that might be considered when choosing pedagogical approaches that will foster effective AL use. It is possible that helping students organize their texts and explicitly mark that organization or helping them express their stance toward propositions clearly could help them use these types of AL markers more effectively. The findings of this study show the large number of attempts students are making at using academic language. The findings about relationships to quality, though somewhat modest, reveal the value in closely assessing student writing. This type of close study of writing using a key set of useful organizational and stance forms, as well as other marker types not explored here, has potential for revealing even more specific and relevant information to help guide teachers' instruction and assessment of students' language skills.

Acknowledgments The research reported here was supported by the Institute of Education Sciences, US Department of Education, through Grant R305A090555 to Harvard University and Grant R305F100026 to the Strategic Education Research Partnership as part of the Reading for Understanding Research Initiative. The opinions expressed are those of the authors and do not represent views of the Institute or the US Department of Education. The author wishes to thank Catherine E. Snow, Paola

Uccelli, and Nonie Lesaux for their invaluable insights into this work, and Emily Phillips Galloway, Jessica Scott, Larinda Dobbs, and Kyle Dobbs for their work in transcribing, coding, and providing thoughtful feedback.

Appendix

See Table 12.

Table 12 Coding manual and sample markers adapted from Uccelli, et. al. (under review) study and Uccelli, Dobbs, & Scott (2013)

Type of marker	Sample marker from data
Organizational markers	
Code glosses: expressions that express an example, definition, or paraphrase	*In other words*, we need to justify our rights to debate in schools
Evidentials: markers that acknowledge the source of a claim	*In the text about renting a pet*, people should make it legal to own a pet because some animals are very expensive
Conclusions markers: markers that express summary or conclusion	*In conclusion*, the family is not important and the school is
Goals: markers that express the goal or aim of a section of text	In this open response, *I explained my position on the question*
Transition markers: connectives including sentence-level connections and markers of macrostructural discourse (excluding temporal connectives)	A debate in one class a week would be helpful *because* the students would be more informed in the work…
Frame markers: markers that indicate overall scheme, text structure, and text boundaries	*To some people* renting a pet can be a good thing…
Stance markers	
Hedges: markers that delimit or soften a claim or statement	… but that is *just* my perspective
Boosters: markers that intensify or strengthen a claim or statement	If they saw how debates helped students they would *definitely* change their minds
Engagement markers: markers which address a reader directly or otherwise make a relationship with a reader	Whereas, *you* will get to keep it for a few days *you* want it to have fun
Deontic modality markers: expressions of attitude toward propositions that demonstrate degrees of prohibition or obligation	Schools *should* only be concerned with student learning
Epistemic modality markers: expressions of attitude toward propositions that demonstrate degree or certainty or degree of knowledge as well as hedges that convey an epistemic function	*I think* you should be able to rent it out…

References

ACT. (2005). *Crisis at the core: Preparing all students for college and work.* Retrieved from http://www.act.org/path/policy/pdf/crisis_report.pdf.

Applebee, A., Langer, J., & Mullis, I. (1986). *The writing report card: Writing achievement in American schools*. Princeton, NJ: Educational Testing Services.

Author. (2009, December). *Increasing academic vocabulary and written academic language skills: Exploring students' writing skills in the Word Generation program*. Paper presented at the meeting of the National Reading Conference, Albuquerque, NM.

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

Bar-Ilan, L., & Berman, R. A. (2007). Developing register differentiation: The Latinate–Germanic divide in English. *Linguistics, 45*, 1–35.

Beers, S. F., & Nagy, W. E. (2007). Writing development in four genres from grades three to seven: Syntactic complexity and genre differentiation. *Reading and Writing: An Interdisciplinary Journal, 24*, 183–202.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum.

Berman, R. A., Ragnarsdóttir, H., & Stömqvist, S. (2002). Discourse stance: Written and spoken language. *Written Language and Literacy, 5*, 255–289.

Berman, R. A., & Ravid, D. (2009). Becoming a literate language user: Oral and written text construction across adolescence. In D. R. Olsen & N. Torrance (Eds.), *Cambridge handbook of literacy* (pp. 92–111). Cambridge, UK: Cambridge University Press.

Berman, R. A., & Slobin, D. I. (1994). *Relating events in narrative: A crosslinguistic developmental study*. Hillsdale, NJ: Erlbaum.

Blum-Kulka, S. (2008). Language, communication and literacy: Major steps in the development of literate discourse. In P. Klein & Y. Yablon (Eds.), *From research to practice in early education* (pp. 117–154). Jerusalem, Israel: Israeli Academy of Science (translated from Hebrew).

Chafe, W. L. (1994). *Discourse, consciousness, and time: The flow of language in speech and writing*. Chicago, IL: University of Chicago Press.

Chall, J. S., & Jacobs, V. A. (1983). Writing and reading in the elementary grades: Developmental trends among low SES children. *Language Arts, 60*, 617–626.

Chesky, J., & Hiebert, E. H. (2001). The effects of prior knowledge and audience on high school students' writing. *Journal of Educational Research, 80*, 304–313.

College Board. (2011). *About the tests.* Retrieved from http://sat.collegeboard.org/about-tests.

Coker, D., & Lewis, W. E. (2008). Beyond *Writing Next*: A discussion of writing research and instructional uncertainty. *Harvard Educational Review, 78*, 231–251.

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

Connor, U. (1990). Linguistic/rhetorical measures for international persuasive student writing. *Research in the Teaching of English, 24*, 67–87.

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

Coxhead, A. (1998). *An academic word list*. Wellington, NZ: Victoria University of Wellington.

Crowhurst, M. (1987). Cohesion in argument and narration at three grade levels. *Research in the Teaching of English, 21*, 185–201.

Dafouz-Milne, E. (2008). The pragmatic role of textual and interpersonal metadiscourse markers in the construction and attainment of persuasion: A cross-linguistic study of newspaper discourse. *Journal of Pragmatics, 40*, 95–113.

De La Paz, S. (2005). Effects of historical reasoning instruction and writing strategy mastery in culturally and academically diverse middle school classrooms. *Journal of Educational Psychology, 97*, 139–156.

Dobbs, C. L. (under review). Learning to be convincing: Metadiscourse and the academic writing of middle graders.

Dombek, K., & Herndon, S. (2004). *Critical passages: Teaching the transition to college composition*. New York, NY: Teachers College Press.

Ferretti, R. P., MacArthur, C. A., & Dowdy, N. S. (2000). The effects of an elaborated goal on the persuasive writing of students with learning disabilities and their normally achieving peers. *Journal of Educational Psychology, 92*, 694–702.

Flood, J., & Lapp, D. (2000). Teaching writing in urban schools: Cognitive processes, curriculum resources, and the missing links—management and grouping. In R. Indrisano & J. R. Squire (Eds.), *Perspectives on writing: Research, theory, and practice* (pp. 187–213). Newark, NJ: International Reading Association.

Gee, J. P. (2008). *Social linguistics and literacies: Ideology in discourses*. New York, NY: Routledge.
Givón, T. (1992). The grammar of referential coherence as mental processing instructions. *Linguistics, 30*, 5–56.
Gorman, T. P., Purves, A. C., & Degenhart, R. E. (Eds.). (1988). *The IEA study of written composition* (1st ed.). Oxford, UK: Pergamon Press.
Graham, S. (1990). The role of production factors in learning disabled students' compositions. *Journal of Educational Psychology, 82*, 781–793.
Graham, S., MacArthur, C., Schwartz, S., & Page-Voth, V. (1992). Improving the compositions of students with learning disabilities using a strategy involving product and process goal setting. *Exceptional Children, 58*, 322–334.
Graham, S., & Perin, D. (2007a). *Writing next: Effective strategies to improve writing of adolescents in middle and high school*. Washington, DC: Alliance for Excellent Education.
Graham, S., & Perin, D. (2007b). What we know, what we still need to know: Teaching adolescents to write. *Scientific Studies of Reading, 11*, 313–335.
Halliday, M. A. K., & Matthiesen, C. M. I. M. (2004). *Introduction to functional grammar* (3rd ed.). London, UK: Hodder Education.
Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 1–27). Mahwah, NJ: Lawrence Erlbaum.
Hayes, J. R. (2000). A new framework for understanding cognition and affect in writing. In R. Indrisano & J. R. Squire (Eds.), *Perspectives on writing: Research, theory, and practice* (pp. 6–44). Newark: International Reading Association.
Hillocks, G. (2002). *The testing trap: How state writing assessments control learning*. New York, NY: Teachers College Press.
Hyland, K. (1998). Persuasion and context: The pragmatics of academic metadiscourse. *Journal of Pragmatics, 30*, 437–455.
Hyland, K. (2002a). Authority and invisibility: Authorial identity in academic writing. *Journal of Pragmatics, 34*, 1091–1112.
Hyland, K. (2002b). Options of identity in academic writing. *ELT Journal, 56*, 351–358.
Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. London, UK: Continuum.
Hyland, K. (2006). Representing readers in writing: Student and expert practices. *Linguistics and Education, 16*, 363–377.
Hyland, K. (2007). Applying a gloss: Exemplifying and reformulating in academic discourse. *Applied Linguistics, 28*, 266–285.
Hyland, K. (2008). Persuasion, interaction, and the construction of knowledge: Representing self and others in research writing. *International Journal of English Studies, 8*(2), 1–23.
Hyland, K., & Tse, P. (2007). Is there an "academic vocabulary"? *TESOL Quarterly, 41*, 235–253.
Kellogg, R. T. (2008). Training writing skills: A cognitive developmental perspective. *Journal of Writing Research, 1*, 1–26.
Langer, J. A. (1984). The effects of available information on responses to school writing tasks. *Research in the Teaching of English, 18*, 27–44.
Llosa, L., Beck, S. W., & Zhao, C. G. (2011). An investigation of academic writing in secondary schools to inform the development of diagnostic classroom assessments. *Assessing Writing, 16*, 256–273.
MacWhinney, B. (2011). *The CHILDES project: Tools for analyzing talk*. Pittsburgh, PA: Carnegie Mellon University.
Massachusetts Comprehensive Assessment System. (2010). *Massachusetts Comprehensive Assessment System*. Retrieved from http://www.doe.mass.edu/mcas/.
McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic features of writing quality. *Written Communication, 27*, 57–86.
Nagy, W. E., & Scott, J. A. (2000). Vocabulary processes. In M. L. Kamil, P. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 269–284). Mahwah, NJ: Lawrence Erlbaum.
Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly, 47*, 91–108.
National Commission on Writing. (2004). *Writing: A ticket to work…or a ticket out: A survey of business leaders.* Retrieved from http://www.writingcommission.org/report/html.
Nir, B., & Berman, R. A. (2010). Complex syntax as a window on contrastive rhetoric. *Journal of Pragmatics, 42*, 744–765.

Nystrand, M. (2006). The social and historical context for writing research. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 11–27). New York: Guilford.

Prior, P. (2006). A sociocultural theory of writing. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 54–66). New York, NY: Guilford.

Pritchard, R. J., & Honeycutt, R. L. (2006). The process approach to writing instruction: Examining its effectiveness. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 275–290). New York, NY: Guilford.

Ravid, D., & Tolchinsky, L. (2002). Developing linguistic literacy: A comprehensive model. *Journal of Child Language, 29*, 417–447.

Reilly, J. S., Baruch, E., Jisa, J., & Berman, R. A. (2002). Propositional attitudes in written and spoken language. *Written Language & Literacy, 5*, 183–218.

Reilly, J., Zamora, A., & McGivern, R. F. (2005). Acquiring perspective in English: The development of stance. *Journal of Pragmatics, 37*, 185–208.

Robb, L. (2012). *Smart writing: Practical units for teaching middle school writers*. Portsmouth, NH: Heinemann.

Salahu-Din, D., Persky, H., & Miller, J. (2008). *The Nation's Report Card: Writing 2007* (NCES 2008-468). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Scarcella, R. (2003). *Academic English: A conceptual framework (Technical Report 2003-1)*. Irvine, CA: The University of California Linguistic Minority Research Institute.

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

Schleppegrell, M. (2004). *The language of schooling*. New York, NY: Lawrence Erlbaum.

Slobin, D. I. (1973). Cognitive prerequisites for the development of grammar. In E. A. Ferguson & D. I. Slobin (Eds.), *Studies of child language development* (pp. 175–208). New York, NY: Holt, Rinehart, and Winston.

Snow, C. E., Lawrence, J. F., & White, C. (2009). Generating knowledge of academic language among urban middle school students. *Journal of Research on Educational Effectiveness, 2*, 325–344.

Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. In N. Torrance & D. R. Olson (Eds.), *The Cambridge handbook of literacy* (pp. 112–133). Cambridge, UK: Cambridge University Press.

Toulmin, S. E. (2003). *The Uses of Argument, Updated Edition.* Cambridge, UK: Cambridge University Press.

Troia, G. A., & Graham, S. (2002). The effectiveness of a highly explicit, teacher-directed strategy instruction routine: Changing the writing performance of students with learning disabilities. *Journal of Learning Disabilities, 35*, 290–305.

Uccelli, P., Barr, C., Dobbs, C. L., Phillips Galloway, E., Meneses, A., & Sánchez, E. (under review). Cross-disciplinary academic language skills: Developmental trends and individual variability during the adolescent years.

Uccelli, P., Dobbs, C. L., & Scott, J. (2013). Mastering academic language: Organization and stance in the persuasive writing of high school students. *Written Communication, 30*, 36–62.

Van de Kopple, W. J. (1985). Some exploratory discourse on metadiscourse. *College Composition and Communication, 36*, 82–93.

Word Generation. (2010). *Word Generation: Middle school literacy development using academic language*. Retrieved from http://www.wordgeneration.org/.