Title
Comparing Elementary Students' Explanatory Language across Oral and Written Modes

Journal
The Elementary School Journal, 122(3)

ISSN
0013-5984

Authors
Blackstock-Bernstein, Anne
Woodbridge, Amy
Bailey, Alison

Publication Date
2022-03-01

DOI
10.1086/718077

Peer reviewed

Comparing Elementary Students' Explanatory Language Across Oral and Written Modes

Anne Blackstock-Bernstein

Amy Woodbridge

Alison Bailey

University of California, Los Angeles

Students receive instruction to develop their competencies in writing throughout elementary school, and their written language skills are frequently assessed (Cutler & Graham, 2008). Traditionally, less attention has been given to the instruction and assessment of students' oral language abilities unless students are acquiring English as a new language or experience difficulties with language development (Bailey, 2010). However, instructional and assessment ramifications of the college and career ready standards for K-12 education have meant that students are increasingly required to use both oral and written language—especially explanations—to demonstrate their conceptual knowledge in the content areas (e.g., the Common Core State Standards of the National Governors Association Center for Best Practices & Council of Chief State School Officers [NGA & CCSSO], 2010; and the Next Generation Science Standards, NGSS Lead States, 2013). For example, mathematics standards demand that elementary students provide explanations of concepts and procedures (e.g., multiplication with fractions) and communicate their mathematical reasoning (NGA & CCSSO, 2010; see Bailey & Heritage, 2014, for a more comprehensive discussion of language demands in content standards). Educators who primarily assess students' maturing language abilities in only one mode (e.g., writing) cannot necessarily assume a similar level of language development in the other mode without an understanding of how these modes relate to one another.

Despite the utility of explanations in the classroom, most previous comparisons of oral and written language have examined children's narratives (Kim et al., 2014; Kim et al., 2015; Olinghouse & Leaird, 2009) or corpora of adult speech and writing (Biber et al., 2011; Chafe, 1982). Important differences between narrative and expository language, as well as between adult and child language, prevent us from generalizing these studies' findings to elementary students' explanatory uses of oral and written language. The current study examines whether

previous studies' more general findings about the differences between oral and written language—such as greater vocabulary diversity in writing—hold true for children's explanations specifically. We also investigate whether important demographic and contextual factors such as children's gender, English learner (EL) status, or the task at hand moderate the differences between their oral and written explanations. The additional understanding gleaned from this study will support teachers in their contingent formative assessment of students' language by highlighting areas that need additional support and providing clarity as to which modes of language allow students the opportunity to fully express their content knowledge.

The Relationship between Oral and Written Language

Early conversational skills can provide a basis for later academic skills, including writing (Snow & Uccelli, 2009). In order to produce a written text, an author needs to generate ideas, translate them into language, and transcribe them. A range of oral language skills are important for this process, and several studies have found that children's performance on standardized oral language assessments is related to their writing quality. As young as kindergarten, children's oral vocabulary and grammatical knowledge relate to their early writing skills (Kim et al., 2011). First grade children's oral vocabulary knowledge and grammatical skills are related to their word choice and quality of written narratives (Coker, 2006; Kim et al., 2014). Third graders' oral grammatical knowledge is also related to their written compositional quality (Olinghouse, 2008), although Berninger and Abbott (2010) found that oral expressive skills are predictive of written expressive skills for third graders but not for fifth graders. While oral and written language share some of the same processes, they are distinct systems that have independent but interconnected developmental trajectories (Berninger, 2000; Berninger et al., 2006).

However, direct comparisons between the two domains remain scarce, particularly in regards to early writing development and for populations of linguistically diverse learners (Kim et al., 2011; Silverman et al., 2015). Most of the above studies measure children's oral language skills using standardized assessments rather than collecting samples of child language to examine parallel features between the two modes. In one of the few direct comparisons of oral and written language with both adults and children, Danielewicz (1984) examined the language of two eight-year-olds, two twelve-year-olds, and two adults. She found that by eight years of age, children are beginning to show small differences between their speech and writing, particularly at the sentence level, but these differences are smaller than the differences found between adult speech and writing. Children begin integrating more complex syntactic and discursive structures into their writing during elementary school (Berninger et al., 1996). Individual children, however, may have varying degrees of skill in the three levels of language—word, sentence, and discourse—and skill in one level does not necessarily predict skill in another.

Surprisingly few studies have built from these early findings, and we emphasize the need for new empirical research. In addition, many studies use "planned" written language (e.g., academic papers) as their corpus. Because both spoken and written language can undergo various degrees of planning (Danielewicz, 1984), the type of language children produce in school may differ from the language studied by previous researchers. For example, children may produce informal written responses to a prompt as an "exit ticket" before the end of a lesson. In the following sections, we review previous findings on differences between oral and written language at the word, sentence, and discourse levels. Few of the existing studies are specific to children's oral language and literacy, further highlighting the need for additional research.

Vocabulary-level Differences

Generally, adults' written vocabulary reflects a higher level of sophistication than their oral vocabulary, with a greater diversity of words and more instances of nominalization (e.g., Drieman, 1962; Chafe, 1982). Adult written language also has a higher "lexical density," or proportion of content words (Danielewicz, 1984; Halliday, 1994). However, Danielewicz finds that eight- and twelve-year-old children demonstrate no real difference in lexical density from their spoken to written language, supporting her aforementioned claim that the structures of children's written language adhere more closely to oral language through middle childhood.

Sentence-level Differences

Previous findings on sentence-level differences between oral and written language yield mixed results. Chafe's (1982) seminal study on the differences between modes finds that adult writing contains a higher proportion of serial phrases, complement clauses, and relative clauses than adult speech. Danielewicz (1984) also finds that written sentences contain more dependent clauses than spoken sentences, and that these differences show up in children as young as eight years of age, although they are not as pronounced in children as they are in adults. Similarly, O'Donnell et al.'s (1967) study of third, fifth, and seventh graders found that their written narratives contained more frequent use of a range of subordinate clauses than their spoken narratives. More recent studies, however, find that spoken language contains a higher proportion of clauses (Biber et al., 2011; Halliday, 1994). When studying corpora of adult conversation and adult writing, many researchers conclude that clausal subordination, the most common measure of syntactic complexity, is more common in spoken language (Biber, 1992; Biber et al., 2002; Biber et al., 2011).

Discourse-level Differences

Most language that students encounter in the classroom takes place as discourse: for the expressive language that is the focus of this paper, either as multi-sentence written passages or as multi-sentence spoken monologues or conversations. Discourse-level features of language go beyond the boundaries of a single sentence, requiring coherence and cohesion among sentences and ideas. Coherence involves the logical or accurate presentation of information and ideas in a sequential, chronological, or causal order. Cohesion occurs when one element of the discourse or text is tied to another, as in the use of pronouns to refer to previously introduced nouns (Halliday & Hasan, 1976). Cohesive ties can refer to elements within the same sentence or across sentence boundaries to elements found elsewhere in the discourse or text. Textual cohesion reduces repetition, making language more parsimonious.

For decades, theorists have suggested that written language is more cohesive than spoken language (Halliday, 1977; Pellegrini & Galda, 1986). According to Halliday, oral language should rely less on grammatical cohesion than written language, in part because it typically occurs when the speaker and listener share physical proximity and can refer to objects in the immediate context. There is limited empirical evidence, however, to support this theory. In one of the few studies to investigate oral versus written cohesion, Pellegrini et al. (1984) found that third graders' writing contained more grammatical cohesion (i.e., use of pronouns, ellipsis, and substitution) than their speech. The fifth graders in their sample showed no significant differences between their oral and written cohesion. In terms of coherence, Chafe (1982) found that among adults at least, writing included more sequencing phrases (e.g., words like "then," "next," "finally") than their oral language.

Studying Elementary Students' Oral and Written Explanations

While many previous investigations of the relationship between children's oral and written language have focused on only one level of language, such as vocabulary (Kim et al., 2015; Olinghouse & Leaird, 2009), the current study simultaneously investigates differences at the word, sentence, and discourse levels (i.e., topic vocabulary, sentence structure, and coherence/cohesion). We can therefore explore whether differences between the modes cut across the three levels of language. To understand the unique impact of mode, it is also important to consider student characteristics that may play a role in students' oral and written explanations, such as their grade level, gender, and EL status.

Student Characteristics

Students develop both oral and written language skills as they progress through the elementary school grades. Oral and written language development are necessarily on different trajectories, however, given that oral language begins developing in infancy, but most children do not encounter writing instruction until school age (Shanahan, 2006). (This is, of course, assuming that the child's language of schooling is the same as the language he or she acquired at home; we will discuss the progression of oral and written language for English learners below.) Therefore, we control for grade level when examining the difference between the two modes, because we expect the progression of students' oral and written language production to vary at different stages of their elementary education. Along with the receptive language skills of listening and reading, the expressive language skills of speaking and writing develop in "overlapping, parallel waves" (Berninger, 2000, p. 66). Oral and written expressive language appear to have a bidirectional influence on each other in third grade, although this same effect does not appear in fifth grade (Berninger & Abbott, 2010). The increased cognitive and physical

demands of writing (e.g., Hayes & Flower, 1980; Singer & Bashir, 2004) likely impact its progression in the elementary years in a way that differs from the progression of oral language.

We also consider EL status among our participants. We expect EL students' development by grade to differ from their English proficient counterparts based on the fact that, as stated above, they did not acquire spoken English in early childhood to the extent that their peers did, which may put their oral and written language on different trajectories from those of English proficient students. Based on this, we have reason to control for the effect of EL status on children's oral and written language progressions.

Finally, we know that gender differences exist in language, although these differences vary based on the mode of language (e.g., reading, writing, speaking, or listening) being assessed (Taylor et al., 2007). Studies utilizing large-scale data do show that girls outperform boys on tests of writing, and in orthographic and compositional fluency (e.g., number of letters/words produced within a time limit; Berninger et al., 1996; Taylor et al., 2007). It is more difficult to know exactly how the groups differ in oral language, which is not frequently formally assessed in children other than with English learners or students with learning or developmental disabilities (Bailey, 2010; Taylor et al., 2007). However, findings from available standardized assessments of English learners provide some understanding of how the oral and written modes differ by gender. In California, where data on EL language test scores are disaggregated by gender, elementary-aged girls score higher than boys in oral language by a small margin and score higher than boys in written language by a larger margin (Taylor et al., 2007). Because we do not know whether these same patterns hold true for English proficient students, we do not fully understand how the relationship between oral and written language differs by gender.

In sum, while existing literature demonstrates an effect of oral language on children's later written language skills, we know very little about the concurrent progression of both modes of language in children. Evidence suggests that, unlike adult vocabulary patterns, children's written vocabulary maintains the lexical density of their oral vocabulary (Danielewicz, 1984). Current literature regarding the sophistication of sentence structure in oral and written language is mixed: while a number of studies assert that oral language contains a higher percentage of subordinate clauses than written language (e.g., Biber et al., 2011; Halliday, 1994), previous studies indicate this may not be true of children's language (Danielewicz, 1984; O'Donnell et al., 1967). When examining discourse-level differences, we find some evidence to suggest that written language contains more devices of coherence and cohesion (Chafe, 1982; Pelligrini et al., 1984), although findings are sparse. In addition, because language development continues across the grades and may differ by gender and EL status, there is reason to investigate whether the relationship between the two modes differs by these characteristics as well.

Systemic Functional Linguistics

Our approach in this study is guided by a systemic functional linguistics (SFL) framework. Originated by Halliday (1985), systemic functional linguistics views language as a resource through which to make meaning in a particular context. The goal of a functional approach to language in education is to provide students with the linguistic resources to communicate in a way that is appropriate for the context (Schleppegrell & Go, 2007; Spycher, 2009). SFL provides a basis for studies connecting linguistics and literacy due to its focus on the co-occurrence of lexical and syntactic forms and the bi-directional relationship between language and social context (Halliday, 1985; Schleppegrell, 2004).

A number of the studies reviewed above consider the accuracy of written conventions, such as spelling, punctuation, and handwriting, in their outcomes (e.g., Kim et al., 2014; Olinghouse, 2008). Based on an SFL framework, we argue that descriptive measures of children's writing should focus on communication and meaning-making during the phases in which they are still developing skills and knowledge. Snow and Uccelli (2009) advocate for a focus on the practical application of the framework that provides insight to the skills and instructional techniques necessary for students' development of academic language. The current study examines linguistic aspects of children's oral and written explanations across tasks in order to understand how teachers might support their development of the disciplinary language of the academic content standards. Findings can support teachers in the use of contingent formative assessment in the classroom by offering details about which aspects of oral and written language may differ, making them candidate targets of embedded assessment to guide instruction.

The Current Study

Children in the middle and upper grades of elementary school are at an important stage of their writing development, typically progressing from learning the mechanics of writing to learning to write for multiple purposes (NGA & CCSSO, 2010). Previous studies of children's performances on oral and written language tasks have generally used standardized assessments of oral language to predict features of written language, including mechanical conventions, and most typically examine outcomes in narrative or descriptive writing. Much of the existing body of literature relies on theoretical assumptions or limited empirical data, and the few studies to directly compare oral and written language are decades old; new educational demands and changing student populations call for new studies to build upon these early findings. In the current study, we elicited oral and written explanations from third through sixth grade students to

investigate differences between the modes and to examine and compare student language use in two tasks: academic (i.e., mathematics) and non-academic (i.e., a personal routine). We have adopted Beals' (1993) description of explanation as an interactional exchange in which the speaker (writer) assumes or is informed that she/he knows something that the addressee needs to know. Students were asked to provide explanations to a hypothetical student. The student had to take account of the limited point of view of the hypothetical student and reflect this understanding linguistically in their attempts to be fully explanatory in both their oral and written explanations. We addressed two research questions:

1. Do students' oral explanations differ from their written explanations in terms of their vocabulary, sentence, and discourse structures, when controlling for task (academic vs. non-academic), grade level, English learner (EL) status, and gender?
2. Do differences between modes (oral vs. written) vary by task, EL status, or gender?[1]

Anticipated Results

Based on previous literature, we hypothesized that students' vocabulary sophistication would not differ significantly between oral and written modes at this age. It was more difficult to hypothesize findings regarding sentence structure due to conflicting previous findings; however, more recent literature suggests that adults' oral syntax is more complex overall than their written syntax. Furthermore, the fact that our participants were children, whose written language was still developing, supported our hypothesis that we would see a higher level of sentence sophistication in children's oral explanations than their written explanations at this age level. Finally, we expected that on the discourse level, students' written explanations would show higher levels of coherence and cohesion than their oral explanations.

Our second research question was more exploratory, although we expected to see some differences. For example, we would logically expect that students are more accustomed to speaking about personal routines than writing about them, and we would anticipate a result reflecting this familiarity. Because most English proficient students have been using spoken English longer than they have been writing, we expect to see larger differences between their oral and written explanations, compared with English learners, many of whom have been developing their oral and written English language in a schooling context simultaneously. Finally, based on our knowledge that girls have been found to outperform boys in written assessments (e.g., Taylor et al., 2011), we predict that while both boys and girls will demonstrate higher performance in their oral explanations than their written explanations, we will see a smaller difference between the modes for girls than for boys.

Method

Setting

School Sites

Students attended one of four elementary schools: two public, one public charter, and one laboratory school at a public university, all located in a metropolitan area of the western United States. Schools were selected to attain socioeconomic status (SES) and racial/ethnic diversity within the sample. At the three public schools, the percentage of students eligible for free or reduced-price lunch ranged from 58% to 99%. The laboratory school was not eligible for the national school lunch program, so free or reduced-price lunch data was not available as a proxy for SES.

Students

Students (n = 128) were drawn from a larger research project (Bailey & Heritage, 2014). Following institutional review board procedures, recruitment letters were distributed to parents of kindergarten through sixth grade students in one to ten classrooms at each school, and all consented students were enrolled in the larger study (n = 324). Students were eligible for the current study (n = 189) if they were in third through sixth grade and completed both tasks (explanations of a personal routine and a math activity) in both modes (oral and written language). Sixty-one of the third through sixth grade students did not complete all tasks required for the current study and were therefore excluded from analyses.[2]

Our analytic sample (n = 128) comprised third (n = 35), fourth (n = 20), fifth (n = 46), and sixth (n = 27) grade students (52% female; mean age = 10.1, SD = 1.09) and was representative of their overall school populations in terms of race/ethnicity: 46% of the students were Latinx, 18% were White, 17% were multi-racial/multi-ethnic, 10% were Asian, 6% were Black, and 2% were unspecified. For privacy reasons, the schools did not provide information about individual students' socioeconomic status. The age range was 7 years; 11 months to 12 years; 1 month. Thirteen students (10%) had received a diagnosis that made them eligible for special education services; no students in the sample participated in gifted programs. Of the 128 students, 25.8% (n = 33) were classified as EL students based on two different designation systems. In the public schools, the district determined EL classification based on students' performances on the state English language proficiency assessment (California Department of Education [CDE], 2012). Of these students, one student was at the Beginning level of overall proficiency, four were Early Intermediate, 12 were Intermediate, 11 were Early Advanced, and one was Advanced (CDE, 2013).[3] At the laboratory school, EL designations (n = 4) were based on a combination of home language survey, teacher report, and student performance on the

Stanford Achievement Test Series, Tenth Edition (SAT10) Reading, Language, and Listening strands (Pearson Assessments, 2011).

Procedures

Standardized prompts were used to elicit students' explanations and were worded similarly for the oral and written modes. Prior research has found that some differences between oral and written language can be attributed to the additional time that is often available for outlining, drafting, and revising written text (e.g., Pellegrini et al., 1984), so in our study, children produced unplanned explanations in both spoken and written modes. Unplanned writing allows for more direct comparison between the two modes, and it aligns with many written tasks that students are likely to encounter even at the end of elementary school (e.g., class exit tickets).

Students provided their explanations during one-on-one sessions with a trained researcher in a quiet space near their classroom. The two tasks (i.e., a non-academic personal routine explanation task and a mathematics activity explanation task) were completed in separate sessions and counterbalanced to account for participant fatigue. Eliciting student explanations in these two tasks enabled us to examine potential differences between students' language use for academic and non-academic topics. These topics differed by the degree of abstraction and conceptualization required; the personal routine (teeth cleaning) being a quotidian and fairly concrete topic to explain and the mathematics activity involving an unanticipated task to complete and a chosen mathematical operation to explain.

Within each task session, students produced an oral explanation followed by a written explanation. This ordering was designed to reflect authentic classroom practices of trying out language and ideas in speech before writing them down (Calkins & Mermelstein, 2003). Oral explanations were audio recorded and transcribed verbatim. During transcription, we parsed the

speech stream into words and sentences based on standard conventions, and this parsing was then verified by a second trained transcriber.

Personal Routine Explanations

Students produced explanations about a familiar routine—teeth cleaning—which provided students at all grade levels with similar expertise and a similar capacity to be fully explanatory. In the oral mode, the student was first asked to explain how they cleaned their teeth and then why they cleaned their teeth. Following these introductory explanations, the student was asked to provide an explanation for a naïve listener, based on the following prompt: "Pretend you're talking to your friend and he/she does not know how to clean his/her teeth. When you're ready, tell him/her how to do it and why he/she should do it."[4] Each student was also asked to produce a written explanation ("Write a note to a friend who doesn't know how to clean his/her teeth. Write down how to do it and why he/she should do it."). See Figure 1 for a sample explanation.

Mathematics Activity Explanations

This task is based on the cardinality standards of the CCSSO Math Standards (NGA Center & CCSSO, 2010). This task was developed because these standards begin in kindergarten with students' counting one-by-one to find a total number of items. The authenticity of the content of this task is part of the validity argument for the progressions (see Bailey & Heritage, 2014), and students are free to choose how to determine a total number of items so that the task is age appropriate (e.g., older students may choose to group items and use multiplication of the number of groups and group size to complete the task; Bailey et al., 2015). Depending on the grade level, the student was given 50 or 100 colorful cubes and asked to find the total number of cubes. In the oral mode, the student was next asked to explain the procedure they used and their

rationale for using it. They were then given the following prompt: "Pretend you are talking to a classmate who has never done this activity. When you're ready, tell him/her how to use the cubes to find out how many there are, and why using the cubes this way helps him/her.*"* In the written mode, the student was asked to answer a similar prompt by writing a note to a classmate ("Write a note to a friend who has never done this activity. Write down how to use the cubes to find out how many there are and why using the cubes this way helps him/her.").

For both tasks and in both modes, if the student did not provide either the procedural portion of the explanation (i.e., how) or the justification portion (i.e., why), they were reminded to do so by the researcher. After responding to the prompt, the researcher asked, "Anything else?" to give students an opportunity to elaborate on their responses.

Measures

As part of the larger research project, researchers created and validated learning progressions for language features at the word, sentence, and discourse levels in the explanation tasks using a modified bookmark method for rank ordering sophistication (e.g., Mitzel et al., 2001), establishing high levels of interrater reliability between teams of raters (average Krippendorff's alphas of .761 and .829 for oral and written progressions respectively; Bailey & Heritage, 2014; 2019). In the current study, key progressions of interest were sophistication of topic vocabulary, sophistication of sentence structure, and coherence/cohesion. For each of these features, coders determined an explanation's "best fit" placement on a four-point progression: Not Evident, Emerging, Developing, and Controlled by sets of performance descriptors. In general, more precise, appropriate, and varied use of each linguistic feature was evidence of the increasing complexity of student explanations (for a detailed description of our coding decisions, please see Bailey & Heritage, 2014). Following Valdés' (2005) argument that

the focus should be on students' communication rather than correctness while acquiring a new language, our learning progression does not evaluate accuracy/correctness until the highest placement (Controlled). In an explanation that is Controlled for any of the three features, errors should not interfere with listener or reader comprehension, otherwise the explanation is considered to be still at the Developing phase. Expressed as Cohen's kappa, which takes account of chance agreement, inter-rater reliability was considered substantial (.79, .69, and .74 for topic vocabulary, sentence structure, and coherence/cohesion, respectively; Landis & Koch, 1977).

Prior studies of the progressions support a validity argument for internal structure and external relations to other variables such as state ELA test scores (Bailey et al., 2015), including an IRT model that revealed sequencing (Bailey & Heritage, 2015) and other studies of order of emergence effects (Bailey, 2017). Additionally, a recent study confirming concurrent validity with the TOEFL Junior® Speaking Tests was conducted with a subsequent sample of English as a foreign language learners (Huang et al., 2020).

Topic Vocabulary

At the word level, we evaluated students' use of the vocabulary words necessary to convey meaning in each task. The topic vocabulary feature measures the sophistication of students' vocabulary use and progresses from the use of a small set of core topic words at the Emerging level to a more encompassing topic-specific lexicon and precise, low-frequency topic vocabulary (e.g., "cavity," "multiplication"); explanations placed at the Controlled level use a variety of precise words. A list of topic vocabulary words was determined for each task through analysis of the vocabulary used by students in their explanations.

Sentence Structure

At the sentence level, we evaluated students' use of a variety of sentence structures. The sentence structure feature progresses from the use of sentence fragments to the use of complex sentences with varied dependent clause structures (e.g., relative, complement, adverbial). As with topic vocabulary, varied usage indicates increasing complexity; explanations placed at the Developing level demonstrated use of at least one subordinate clause, and those with a variety of complex clause structures were placed at the Controlled level.

While we parsed students' oral explanations into sentences during transcription, we decided to maintain students' own punctuation choices in their written explanations. We were concerned that using a standardized parsing system would not necessarily reflect the students' intended meaning. In order to demonstrate control of sentence structure, students needed to accurately signal the end of their sentences in some way, such as punctuation, formatting, and/or capitalization. In other words, students' explanations could not be placed at the Controlled level if their sentence boundaries were unclear, as unclear boundaries could interfere with reader comprehension. For example, in the example shown in Figure 1, the student used both complement clauses ("why you should always brush your teeth") and adverbial clauses ("so it won't get yellow"). While this variety of complex clause structures could signal a Controlled placement, the student's explanation was placed at the Developing level, because the sentence boundaries were not made clear.

Coherence/Cohesion

At the discourse level, we examined students' use of specific strategies that help the listener/reader understand their explanations. Our measure of discourse examines coherence and cohesion, which are essential to both oral and written language. Coherence evaluates students' use of temporal discourse connectors (e.g., "first," "next," "then") and conjunctions to logically

sequence procedural steps. Cohesion is maintained by the accurate application of referential ties across and within sentences, such as students' use of cohesive devices (e.g., pronouns, ellipses, substitution) to create a parsimonious explanation. For example, the following excerpt from an oral explanation contains both pronominal reference and substitution to create cohesion: "…give each two blocks together. Give them a bubble space from other ones." The student uses both a pronoun ("them") and substitution ("ones") to refer to "blocks," thus tying the two sentences together and minimizing repetition. In ellipsis, the student omits words that grammatically do not need to be repeated, such as a verb or noun appearing in a prior clause. For example, this student omits "I clean" at two points in her explanation: "And then I clean my tongue and […] the upper side and […] my side teeth."). Due to minimal sequencing and limited use of cohesive devices, explanations at the Emerging level may require a lot of effort from the listener/reader to understand, whereas Controlled explanations, which are logically sequenced and use accurately tied cohesive devices, require little effort to understand. In the oral mode, deictic references reliant on a shared context in the "here and now" (e.g. phrases such as "that one" without anaphoric or cataphoric reference to a full noun) were not considered accurately tied cohesive devices. These deictic uses were not sufficiently effective communication for placement at the Controlled level, because the task demands were overt about the need to be fully explanatory (verbally or in print) for a naïve listener. Moreover, gestures in place of speech were not taken into account to interpret meaning so that both the oral and written modes were parallel.

Analytic Plan

We first examined how students' explanations were distributed along each progression: topic vocabulary, sentence structure, and coherence/cohesion. Fisher's exact tests of the

frequency of placements at each level revealed whether the distributions differed by each of the control variables (i.e., gender, EL status, grade, task).

We used ordered logistic regression to look for differences between the oral and written modes for our three outcome variables—topic vocabulary, sentence structure, and coherence/cohesion—each of which was placed on a four-level ordinal scale (i.e., Not Evident, Emerging, Developing, Controlled). The unit of analysis was the explanation, because we were interested in differences between oral and written explanations, rather than differences between children. Each student completed an oral and written explanation for both the personal routine and mathematics tasks, resulting in 512 explanations that were nested within the 128 students in our sample. We therefore used multilevel ordered logistic models, because they estimate more precise standard errors when observations are clustered, such as when multiple explanations are produced by one student (Raudenbush & Bryk, 2002). For each outcome variable, a likelihood-ratio test compared nested models (with and without a random effect) and indicated that there was enough variability between students to justify the use of two-level (i.e., explanations nested within students) mixed-effects ordered logistic regression instead of standard ordered logistic regression ($p < .001$).[5]

For each outcome, we started by building a model that included explanation characteristics as Level-1 fixed effects; this model, Model 1, examined the effect of the mode (i.e., written or oral) and task (i.e., mathematics or personal routine) on each language feature. Next, we added children's characteristics (i.e., gender, EL status, grade level) as control variables in Model 2. This allowed us to determine whether there were statistically significant differences between students' oral and written explanations after controlling for variance due to task, gender, EL status, and grade. Finally, we added interaction terms in Model 3 to examine

whether the extent of any differences between oral and written explanations varied by task, gender, or EL status. Unless otherwise noted, likelihood ratio tests indicated that each model was a better fit than the previous one.

Results

Descriptive and Preliminary Analyses

Overall, students' explanations were most often placed at the Developing phase of the progression in terms of topic vocabulary and sentence structure, and at the Emerging phase in terms of coherence/cohesion (see Table 1). Fisher's exact tests examined whether the distribution of explanations along the progression differed by each of the control variables. They revealed that girls' and boys' explanations were significantly different from one another, with girls' explanations placed higher than boys' on the progressions for topic vocabulary ($p = .002$), sentence structure ($p = .001$), and coherence/cohesion ($p = .001$). We also found significant differences based on EL status for all three language features: English proficient students' explanations were placed higher than ELs' explanations on the topic vocabulary ($p < .001$), sentence structure ($p = .001$), and coherence/cohesion ($p = .005$) progressions. Both sentence structure and coherence/cohesion placements differed by grade ($p < .001$ and $p = .001$, respectively), such that older students' explanations were placed higher on the progressions; there was not a statistically significant grade difference for topic vocabulary ($p = .094$).[6] Finally, explanations differed significantly based on task for some features; students' personal routine explanations were placed higher on the vocabulary progression than their mathematics explanations ($p = .038$), whereas for sentence structure, their mathematics explanations were placed higher than their personal routine explanations ($p = .006$). There was no difference in the distributions in terms of coherence/cohesion ($p = .272$).

Oral versus Written Topic Vocabulary

Analyses indicated that, across both tasks, sophistication of topic vocabulary was similar between students' oral and written explanations. Our first model (Model 1 in Table 2) controlled only for task and found no significant difference in the odds of a student's oral or written explanation being placed at a higher level on the topic vocabulary progression (e.g., placed at Developing or Controlled instead of Emerging). Similarly, in Model 2, while controlling for task, grade, gender, and EL status, we found no significant difference between students' oral and written explanations (Odds ratio = 1.11, p = .563; see Table 2).

There were several interaction effects between mode and children's characteristics in Model 3, as shown in Table 2. One interaction—mode with gender—revealed that girls' written explanation vocabulary was placed higher than their oral explanation vocabulary, but boys' written explanation vocabulary was placed lower than their oral (see Figure 2). Another interaction—mode with EL status—found that English proficient students' written explanations were placed slightly lower on the vocabulary progression than their oral explanations, but EL students' written explanations were placed higher than their oral explanations for vocabulary, and were nearly comparable to the written explanation vocabulary of English proficient students (see Figure 3). There were no statistically significant interactions between mode and task for this language feature.

Oral versus Written Sentence Structure

Models 1 and 2 for sentence structure revealed that, while controlling for task, grade, gender, and EL status, there were differences in sentence structure between students' oral and written explanations. As shown in Model 2 in Table 3, oral explanations had 1.69 times the odds of being placed at a higher level on the sentence structure progression compared to written

explanations (p = .005). This finding suggests that students have greater control of sentence structure in their oral explanations than in their written explanations. A likelihood ratio test indicated that Model 2 fit the data better than Model 3, because there were no significant interaction effects.

Oral versus Written Coherence/Cohesion

Like sentence structure, there were differences between the oral and written coherence/cohesion of students' explanations. Oral explanations had 1.71 times the odds of being placed at a higher level on the coherence/cohesion progression compared to written explanations (p = .003; see Model 2 in Table 4). This suggests that students demonstrate more control of coherence/cohesion in their oral explanations than in their writing. Interaction analyses in Model 3 suggested that the difference between students' oral and written coherence/cohesion may depend in part on the task at hand. There was a significant interaction between task and mode (Odds ratio = .432, p = .018), suggesting that there is less of an effect of mode on math explanations than on explanations about a personal routine. As Figure 4 shows, students were likely to demonstrate similar coherence/cohesion in their oral and written math explanations, but for personal routine explanations, their oral coherence/cohesion was likely to be placed higher on the progression than their written coherence/cohesion. A summary table of all main effects and interaction effects is presented in Table 5.

Discussion

The results of this study contribute to the understanding of how explanatory language in the intermediate and upper elementary grades differs between oral and written modes. While we found that students' oral explanations were more sophisticated than their written explanations in terms of sentence structure and coherence/cohesion, our findings reveal that contextual factors

(e.g., academic versus non-academic tasks) may play a key role in understanding differences between the two modes. These findings provide a basis for teachers to identify content areas and features of language in which students may continue to require additional support, especially in writing. This is particularly important when considering formal assessment of content knowledge, which has traditionally been conducted through written, rather than oral, language.

Descriptive Findings

We first confirmed that student performance was consistent with prior studies of oral and written language development. Indeed, based on our initial descriptive analyses, girls' explanations were placed higher on all three progressions (topic vocabulary, sentence structure, and coherence/cohesion) than those produced by boys, and the responses of English proficient students were more likely to be placed higher than those of their peers who were acquiring English as a new language (e.g., Berninger, et al., 1996; Taylor et al., 2007). Students in higher grades were placed higher on the sentence structure and coherence/cohesion progressions than those in lower grades.

In both oral and written explanations, students' coherence/cohesion was less sophisticated than their topic vocabulary and sentence structure. The ELA Common Core State Standards, to which many US states currently adhere, introduce the use of "linking words and phrases"—a key element of our coherence/cohesion progression—by third grade (NGA & CCSSO, 2010). For comparison, the use of "precise and domain-specific vocabulary"—similar to our topic vocabulary progression—does not become part of the ELA standards until fourth grade (NGA Center & CCSSO, 2010). We would therefore have expected to see students develop control of written coherence and cohesion earlier than control of written sophistication of topic vocabulary, but this was not the case. Areas such as vocabulary may be considered by

many teachers the "low-hanging fruit" in terms of language development, which may explain its higher placement than elements at the discourse level. In addition, our data underscore the concerns raised by Bailey and Heritage (2014), who argue that current standards "do not provide a comprehensive, empirically validated detailed description of how language develops" (p. 501). The findings of this study can help guide future revisions of educational standards towards data-driven, rather than theoretical, progressions of language learning.

When comparing the language of explanations in academic versus non-academic tasks, we saw significant differences in the sophistication of children's vocabulary and sentence structure, but not in their coherence and cohesion. Children's vocabulary in their explanations of personal routines was more frequently placed at the Controlled level of the progression than the vocabulary used in their math explanations. This may be due to the fact that children's increased comfort with the everyday activity of cleaning their teeth results in an increased familiarity with a wider range of topic vocabulary. In terms of the sophistication of students' sentence structure, however, more math explanations than personal routine explanations were placed at the higher levels of the progression. This finding is not surprising when considering the ideas being expressed in a mathematical task and the complex relationships among them. Children used adverbial and relative clauses (e.g., "if you have this number of cubes"; "the cubes that you've counted") more frequently than they did when explaining how to clean their teeth. We were also not especially surprised to find that placements of coherence and cohesion did not differ based on the task at hand. The measurement of coherence, in particular, relied on the frequency of sequencing words and transitions, and is closely tied to the genre of procedural explanations. Because both prompts were about procedures, there is no compelling reason they should have differed in this regard.

Effects of Mode of Production

Vocabulary

Our analyses revealed no significant difference in the sophistication of topic vocabulary in students' oral and written explanations. These findings are consistent with Danielewicz's (1984) findings that the lexical density of children's oral language is similar to that of their written language even through age 12. However, we found significant interactions with both gender and EL status. Not only were girls' explanations more likely to receive a higher topic vocabulary placement than boys' in both oral and written modes, but the relationship between the two modes was reversed for boys and girls: girls' written vocabulary was placed higher than their oral vocabulary, while boys' oral vocabulary was placed higher than their written vocabulary. Our findings are supported by analyses of large-scale assessment results that find that boys' writing is generally less sophisticated than girls' throughout their school careers (Reynolds et al., 2015; Taylor et al., 2007). Because adult language demonstrates a higher level of vocabulary sophistication in written language than in oral language (e.g., Chafe, 1982; Drieman, 1962; Halliday, 1994), it is possible that the girls in our study more closely approximate the adult-like pattern in terms of their language development. Boys are also found to struggle with compositional fluency (Berninger et al., 1996; Berninger & Fuller, 1992), which may help explain the relationship between their oral and written topic vocabulary use. From an instructional standpoint, it is essential that all children continue to develop sophisticated topic vocabulary particularly in their writing, as the lack of a present interlocutor with whom one shares common ground can demand an additional level of explicitness or lexical precision (Bailey, 2020; Galloway & Uccelli, 2015).

We also found an interaction between mode and EL status: overall, English proficient students' topic vocabulary was more sophisticated in their oral explanations than in their written explanations, whereas EL students used more sophisticated topic vocabulary in their written explanations than in their oral explanations. In fact, although the topic vocabulary placements of explanations given by EL students were lower overall, their written explanations were placed nearly as high as the written explanations of English proficient students. This interaction may be due in part to a performance effect, an effect that is documented in foreign language instruction literature (e.g., Cheng et al., 1999). A written task may alleviate some of the potential anxiety that students—particularly those still acquiring English—may face when producing oral explanations for another individual. In addition, written tasks afford the student additional time to review or monitor as they compose (Flower & Hayes, 1981), factors that may positively impact EL students more so than students who are already English proficient. These findings demonstrate that EL students are capable of producing topic vocabulary that is almost as sophisticated as their English proficient peers. However, they serve as a reminder that EL students may need additional support in their expressive oral vocabulary, given that they are frequently assessed in this area.

Sentence Structure

Students demonstrated more sophistication of sentence structure in their oral explanations than in their written explanations. Our finding adds to the body of evidence demonstrating a similar relationship in adult language (e.g., Biber et al., 2011; Halliday, 1994), and contradicting previous studies of children's language (Danielewicz, 1984; O'Donnell et al., 1967). We found no statistically significant interactions between mode and task, EL status, or gender in terms of students' sentence structure.

In order to demonstrate understanding of material in many content areas, students are often asked to explain abstractions such as relationships between ideas (NGA & CCSSO, 2010), which may demand the use of complex clauses. The participants in our study demonstrated less control of complex clause structures in their written language than in oral language. Given this finding, we encourage educators to examine the complexity of the ideas and actions of the task at hand when determining how to assess content knowledge through students' explanations. Children in mid-elementary grades may be better able to convey their content knowledge orally if that knowledge requires them to produce complex sentences.

Coherence/Cohesion

The findings also add to the current knowledge of students' use of discourse features in their oral and written language. Theoretically, we expect written language to demand more devices of coherence and cohesion than oral language, because the reader is dependent on the writer to provide linguistic guidance, given the lack of shared physical context (Halliday, 1977; Pellegrini & Galda, 1986). An initial review of our findings suggests the opposite: in general, elementary students' oral explanations were placed higher on the coherence and cohesion progression than their written explanations. However, we observed an interaction between task and mode showing that students' personal routine explanations demonstrated more control of coherence and cohesion markers in oral language as opposed to written language, whereas academic explanations showed little difference between modes. In other words, any higher-level control of coherence/cohesion in oral language was being driven by students' non-academic explanations. This interaction is illustrated in Figure 4. This finding suggests that elementary students may need focused instruction in crafting cohesive, coherent explanations.

Limitations and Future Directions

A strength of this research is its focus on explanation as a genre that has high utility across content areas in elementary school, in contrast to most previous studies that focused on children's narrative skills. The results of this study, however, cannot be generalized to that genre. Future studies should include both explanations and narratives. Researchers can also consider extending the results of this study to explanations in other content areas, such as science, given SFL postulates disciplinary-specific language characteristics. In addition, this study's correlational design prevents us from making causal inferences about the relationships between task and mode. As previously noted, students were given prompts in the oral mode first, then written, in order to reflect authentic classroom language use (Calkins & Mermelstein, 2003). We recognize that this introduces a potential rehearsal effect and we might expect to see an advantage in written explanations. However, children's oral explanations were generally placed higher on the progressions, suggesting that differences between the two modes may in fact be even stronger than our data show.

One limitation in this analysis is our coding's reliance on frequency rather than proportions in determining students' levels on the progressions. This may have given an advantage to oral language, since children's written stamina may be lower, resulting in shorter written explanations. We attempted to minimize this advantage by requiring the coding of explanations to take a child's repertoire into account, rather than relying on repeated tokens of a single type. That is, placement at the Developing and Controlled levels required an increasing variety of forms (e.g., range of lexical types, sentence structures, cohesive devices) in an attempt to balance the coding qualitatively rather than coding simply on raw frequency of one form.

We investigated third through sixth graders because they are in the midst of developing their writing skills. We recommend that future studies include older students in order to better

understand how the relationship between oral and written language may continue to change over time and at what point children's language reaches the "adult-like" patterns that have been established in previous literature. A longitudinal design would be ideal for examining the development of such patterns. In addition, while we do see development across grade levels even within this age range, a larger sample would allow for investigation of more fine-grained mode and task interactions within each grade. Analyses using EL students' English proficiency levels, rather than simply their status, would also allow for additional fine-grained examination.

Finally, this study was intended to look at specific features of elementary-aged children's language. However, we realize this is an incomplete picture, as effective explanations must also reflect accurate content knowledge. We suggest future studies examine correlations between ratings of students' language and content knowledge to better understand how students' language reflects their knowledge.

Implications

Much of the existing literature on children's oral and written language has focused on the beginning stages of written language development (Kim et al., 2014). This study builds on existing knowledge and expands our understanding of how children's oral language progresses as their writing skills become more developed. Analyzing children's oral and written language separately in terms of their topic vocabulary, sentence structure, and coherence/cohesion of their discourse highlights the fact that simply because children are showing controlled use of language within one feature does not necessarily mean they have control within another feature. When identifying instructional needs, teachers can more accurately monitor specific features that may still be developing. In addition, teachers cannot assume that a child who demonstrates control in his oral language abilities has a similar level of control in his written language, and vice versa.

Literacy, including writing skills, is assessed more frequently than speaking, particularly in high-stakes summative assessments (Bailey, 2010). However, elementary students' oral language is still developing and requires structured guidance. This may be particularly true for students who are acquiring English as a new language. This group placed higher overall on the vocabulary progression for their writing than for their oral explanations, with written explanations being nearly at the level of those of their English proficient peers. These findings for writing abilities are unexpected, and as previously noted, may be partially explained by performance anxiety during assessment of oral language. However, these findings may also reflect literature indicating that many EL students receive less direct instruction in oral language than in written language (August & Shanahan, 2006). Assuming this is indeed true, the strength of EL students' written vocabulary in the current study suggests that the vocabulary of EL students can and does progress when supported academically. The findings also call attention to the need for continued support of EL students' oral language development, as this population of students is assessed annually in both modes of expressive language and must demonstrate proficiency in both for redesignation purposes.

Based on our findings, we also encourage teachers to consider allowing students to produce oral explanations when they are formatively assessing students' content knowledge. Our elementary-aged participants demonstrated more sentence complexity in their oral explanations, suggesting that they may initially be better able to represent complex ideas and relationships in speech. We recommend that when appropriate, teachers begin with this approach, later providing students with writing instruction to help them transfer their knowledge to the page.

Children are also expected to use temporal words and linking phrases in first and second grade, respectively (NGA & CCSSO, 2010), but our findings demonstrate that even older

students may need additional guidance in producing academic explanations that use such devices to be coherent and cohesive. We recommend teachers provide additional instruction at this level in both oral and written language in order to support students.

Because our results suggest that boys demonstrate greater control of oral vocabulary than written vocabulary, we offer the following example of how teachers might apply these findings: a fifth-grade teacher, conducting formative assessment in a pair-share activity during math instruction, notes a boy's use of sophisticated topic vocabulary in his oral explanation of a problem-solving strategy. Later, turning to his writing about a similar problem, she may not see the same level of vocabulary being used. Because she had previously attended closely to his oral vocabulary, the teacher would recognize that the student had sufficient content knowledge and would be attuned to the fact that he had control of topic vocabulary in one mode, and that he was ready to further progress his written vocabulary with instruction.

Prior researchers working in the SFL tradition argue that teachers can use SFL to analyze the language of students acquiring English as a new language by breaking down utterances in order to isolate the "grammatical constituents" of meaning making (e.g., Schleppegrell & Go, 2007). Once teachers recognize the meaning their students have succeeded in communicating, they can then isolate areas of potential linguistic growth that will allow students to express themselves in a way that is valued and accepted in an academic context. While we agree that this potential applies to emergent bilingual students, we believe it extends to all students. Our study provides an empirically derived understanding of the level of children's abilities as they make progress in their language and literacy learning, and it highlights the heterogeneous nature of these progressions in different language and literacy contexts students encounter and among students who differ by gender, language background, and grade.

References

August, D., & Shanahan, T. (Eds.). (2006). Executive summary: Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth. Lawrence Erlbaum Associates Publishers.

Bailey, A. L. (2010). Assessment in schools: Oracy. In James, M. (Section Ed.), *International encyclopedia of education* (3rd ed., pp. 285–292). Elsevier.

Bailey, A. L. (2017). Progressions of a new language: Characterizing explanation development for assessment with young language learners. *Annual Review of Applied Linguistics*, *37*, 241.

Bailey, A. L. (2020). The discourse of explicitness: Mathematics explanatory talk and self-assessment by Spanish-speaking emergent bilingual students in elementary classrooms. *Theory Into Practice*, *59*(1), 64-74.

Bailey, A. L., Blackstock-Bernstein, A., & Heritage, M. (2015). At the intersection of mathematics and language: Examining mathematical strategies and explanations by grade and English learner status. *The Journal of Mathematical Behavior*, *40*, 6-28.

Bailey, A. L., & Carroll, P. E. (2015). Assessment of English language learners in the era of new academic content standards. *Review of Research in Education*, *39*(1), 253-294.

Bailey, A. L. & Heritage, M. (2014). The role of language learning progressions in improved instruction and assessment of English language learners. *TESOL Quarterly*, *48*(3), 480-506.

Bailey, A. L. & Heritage, M. (2015, July). *Empirically-Derived Language Learning Progressions in Assessment for Learning.* International Conference on Learning, Madrid, Spain.

Bailey, A. L., & Heritage, M. (2019). *Progressing Students' Language Day by Day*. Corwin.

Beals, D. E. (1993). Explanatory talk in low-income families' mealtime conversations. *Applied Psycholinguistics*, *14*(4), 489–513.

Berninger, V. W. (2000). Development of language by hand and its connections with language by ear, mouth, and eye. *Topics in Language Disorders*, *20*(4), 65–84.

Berninger, V. W., & Abbott, R. D. (2010). Listening comprehension, oral expression, reading comprehension, and written expression: Related yet unique language systems in grades 1, 3, 5, and 7. *Journal of Educational Psychology*, *102*(3), 635–651.

Berninger, V. W., Abbott, R. D., Jones, J., Wolf, B. J., Gould, L., Anderson-Youngstrom, M., Shimada, S., & Apel, K. (2006). Early development of language by hand: Composing, reading, listening, and speaking connections; three letter-writing modes; and fast mapping in spelling. *Developmental Neuropsychology*, *29*(1), 61–92.

Berninger, V. W., & Fuller, F. (1992). Gender differences in orthographic, verbal, and compositional fluency: Implications for assessing writing disabilities in primary grade children. *Journal of School Psychology, 30*(4), 363-382.

Berninger, V. W., Fuller, F., & Whitaker, D. (1996). A process model of writing development across the life span. *Educational Psychology Review*, *8*(3), 193–218.

Biber, D. (1992). On the complexity of discourse complexity: A multidimensional analysis. *Discourse Processes*, *15*(2), 133–163.

Biber, D., Conrad, S., Reppen, R., Byrd, P., & Helt, M. (2002). Speaking and writing in the university: A multidimensional comparison. *TESOL Quarterly*, *36*(1), 9–48.

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly, 45*(1), 5–35.

California Department of Education (2012). *California English Language Development Test.* 2012-2013 edition. California Department of Education.

California Department of Education (2013). *Technical Report for the California English Language Development Test (CELDT) 2012−13 Edition.*

Calkins, L., & Mermelstein, L. (2003). *Launching the Writing Workshop*. Heinemann.

Chafe, W. L. (1982). Integration and involvement in speaking, writing, and oral literature. In D. Tannen (Ed.), *Spoken and written language: Exploring orality and literacy* (pp. 35–53). Ablex Publishing Corporation.

Cheng, Y., Horwitz, E. K. & Schallert, D. L. (1999). Language anxiety: Differentiating writing and speaking components. *Language Learning, 49*(3), 417-446.

Huang, B. H., Bailey, A. L., Sass, D. A., & Shawn Chang, Y. H. (2020). An investigation of the validity of a speaking assessment for adolescent English language learners. *Language Testing*.

Coker, D. (2006). Impact of first-grade factors on the growth and outcomes of urban schoolchildren's primary-grade writing. *Journal of Educational Psychology, 98*(3), 471-488.

Cutler, L., & Graham, S. (2008). Primary grade writing instruction: A national survey. *Journal of Educational Psychology, 100*(4), 907–919.

Danielewicz, J. M. (1984). The interaction between text and context: A study of how adults and children use spoken and written language in four contexts. In A. D. Pellegrini & T. D.

Yawkey (Eds.), *The development of oral and written language in social contexts* (pp. 243–260). Ablex Publishing Corporation.

Drieman, G. H. (1962). Differences between written and spoken language: An exploratory study. *Acta Psychologica, 20*, 78-100.

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, *32*(4), 365–387.

Galloway, E. P., & Uccelli, P. (2015). Modeling the relationship between lexico-grammatical and discourse organization skills in middle grade writers: Insights into later productive language skills that support academic writing. *Reading and Writing, 28*(6), 797-828.

Halliday, M. A. K. (1977). Text as semantic choice in social contexts. In *Grammars and descriptions: Studies in text theory and text analysis*. W. de Gruyter.

Halliday, M. A. K. (1985). *An introduction to functional grammar*. E. Arnold.

Halliday, M. A. K. (1994). Spoken and written modes of meaning. In D. Graddol & O. Boyd-Barrett (Eds.), *Media texts, authors and readers: A reader* (pp. 51–73). Multilingual Matters.

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

Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3–30). Erlbaum Associates.

Kim, Y.-S., Otaiba, S. A., Puranik, C., Folsom, J. S., Greulich, L., & Wagner, R. K. (2011). Componential skills of beginning writing: An exploratory study. *Learning and Individual Differences, 21*(5), 517-525.

Kim, Y.-S., Otaiba, S. A., Folsom, J. S., Greulich, L., & Puranik, C. (2014). Evaluating the dimensionality of first-grade written composition. *Journal of Speech, Language, and Hearing Research*, *57*(1), 199–211.

Kim, Y.-S., Puranik, C., & Otaiba, S. A. (2015). Developmental trajectories of writing skills in first grade: Examining the effects of SES and language and/or speech impairments. *The Elementary School Journal*, *115*(4), 593–613.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159–174.

Mitzel, H.C., Lewis, D. M., Patz, R. J., & Green, D. R. (2001). The bookmark procedure: Psychological perspectives. In G. J. Cizek (Ed.), Setting performance standards: Concepts, methods, and perspectives (pp. 249-281). Lawrence Earlbaum Associates.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards.* National Governors Association Center for Best Practices & Council of Chief State School Officers.

NGSS Lead States. (2013). *Next Generation Science Standards: For States, By States*. National Academies Press.

O'Donnell, R. C., Griffin, W. J., & Norris, R. C. (1967). *Syntax of Kindergarten and Elementary School Children: A Transformational Analysis*. National Council of Teachers of English.

Olinghouse, N. G. (2008). Student- and instruction-level predictors of narrative writing in third-grade students. *Reading and Writing*, *21*(1), 3–26.

Olinghouse, N. G., & Leaird, J. T. (2009). The relationship between measures of vocabulary and narrative writing quality in second- and fourth-grade students. *Reading and Writing*, *22*(5), 545–565.

Pearson Assessments. (2011). *Stanford 10 information packet.* http://www.pearsonassessments.com/hai/Images/PDF/Webinar/Stanford_Testing Info Packet1272011.pdf

Pellegrini, A. D., & Galda, L. (1986). The role of theory in oral and written language curricula. *The Elementary School Journal, 87*(2), 201–208.

Pellegrini, A. D., Galda, L., & Rubin, D. L. (1984). Context in text: The development of oral and written language in two genres. *Child Development, 55*(4), 1549–1555.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd edition). Sage.

Reynolds, M. R., Scheiber, C., Hajovsky, D. B., Schwartz, B., & Kaufman, A. S. (2015). Gender differences in academic achievement: Is writing an exception to the gender similarities hypothesis? *The Journal of Genetic Psychology, 176*(4), 211-234.

Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective*. Lawrence Erlbaum Associates.

Schleppegrell, M. J., & Go, A. L. (2007). Analyzing the writing of English learners: A functional approach. *Language Arts, 84*(6), 529–538.

Shanahan, T. (2006). Relations among oral language, reading, and writing development. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 171–183). Guilford Press.

Silverman, R. D., Coker, D., Proctor, C. P., Harring, J., Piantedosi, K. W., & Hartranft, A. M. (2015). *The Elementary School Journal, 116*(1), 103-125.

Singer, B. D., & Bashir, A. S. (2004). Developmental variation in writing composition skills. In C. A. Stone, E. R. Silliman, B. J. Ehren, & K. Apel (Eds.), *Handbook of language and literacy: Development and disorders* (pp. 559–582). Guilford Press.

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

Spycher, P. (2009). Learning academic language through science in two linguistically diverse kindergarten classes. *The Elementary School Journal*, *109*(4), 359–379.

Taylor, A., Bailey, A., Cooper, P., Dwyer, C. A., Kramarae, C., & Lieb, B. (2007). Gender equity in communication skills. In S. S. Klein, B. Richardson, D. A. Grayson, L. H. Fox, C. Kramarae, D. S. Pollard, & C. A. Dwyer (Eds.), *Handbook for achieving gender equity through education* (pp. 281-304). Lawrence Erlbaum Associates.

Valdés, G. (2005). Bilingualism, heritage language learners, and SLA research: Opportunities lost or seized? *The Modern Language Journal, 89(3)*, 410-426.

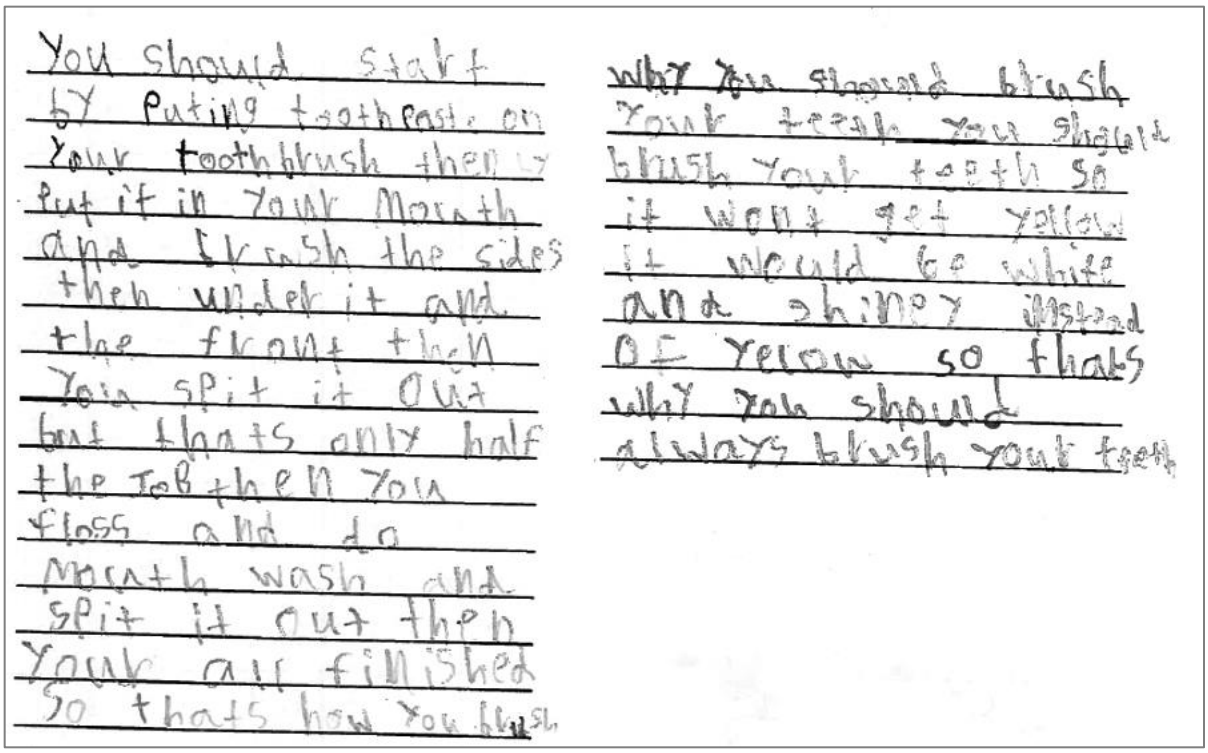

Figure 1. Sample written explanation.

Transcript: You should start by putting toothpaste on your toothbrush then put it in your mouth and brush the sides then under it and the front then you spit it out but that's only half the job then you floss and do mouth wash and spit it out then your all finished so thats how you brush why you should brush your teeth you should brush your teeth so it wont get yellow it would be white and shiney instead of yelow so thats why you should always brush your teeth

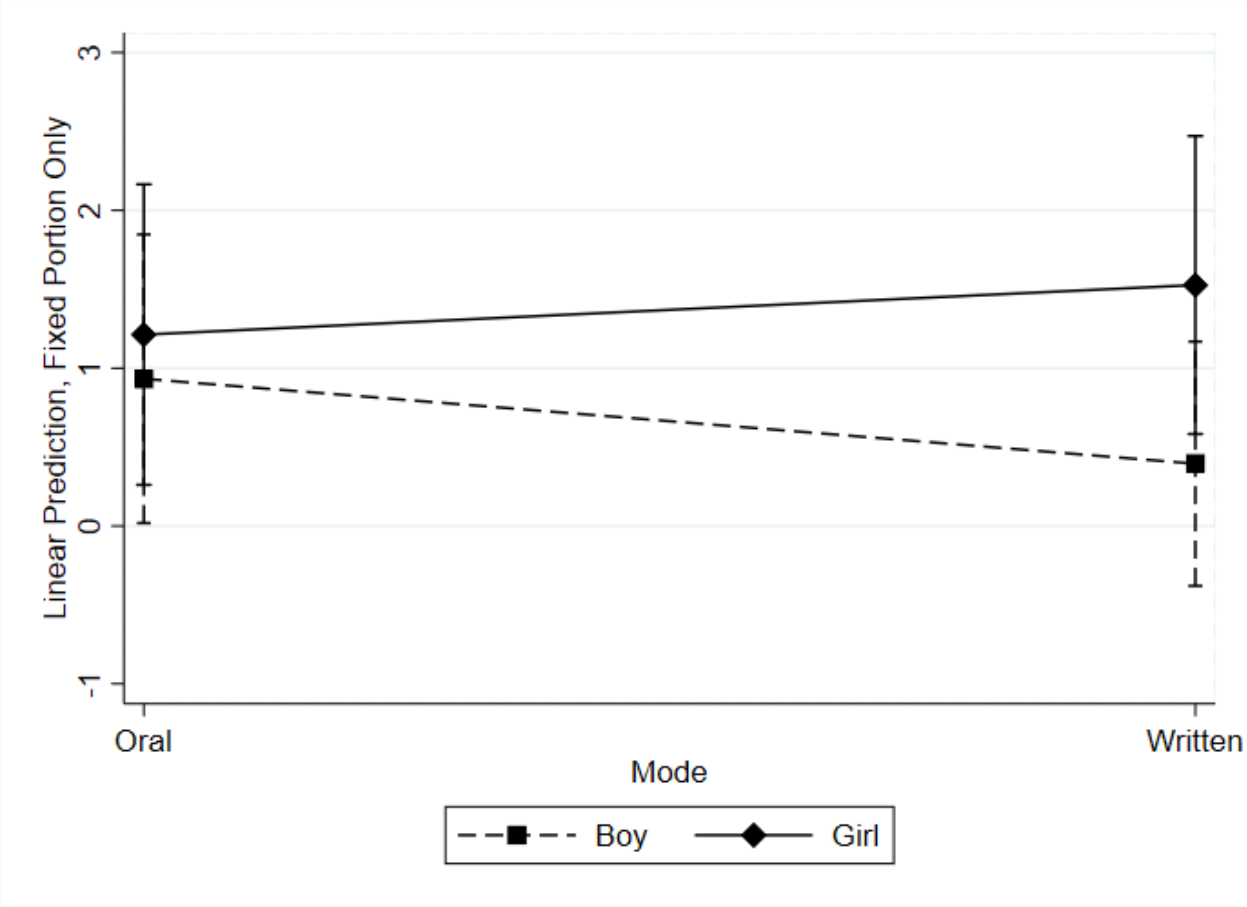

Figure 2. Predictive margins of mode and gender for topic vocabulary

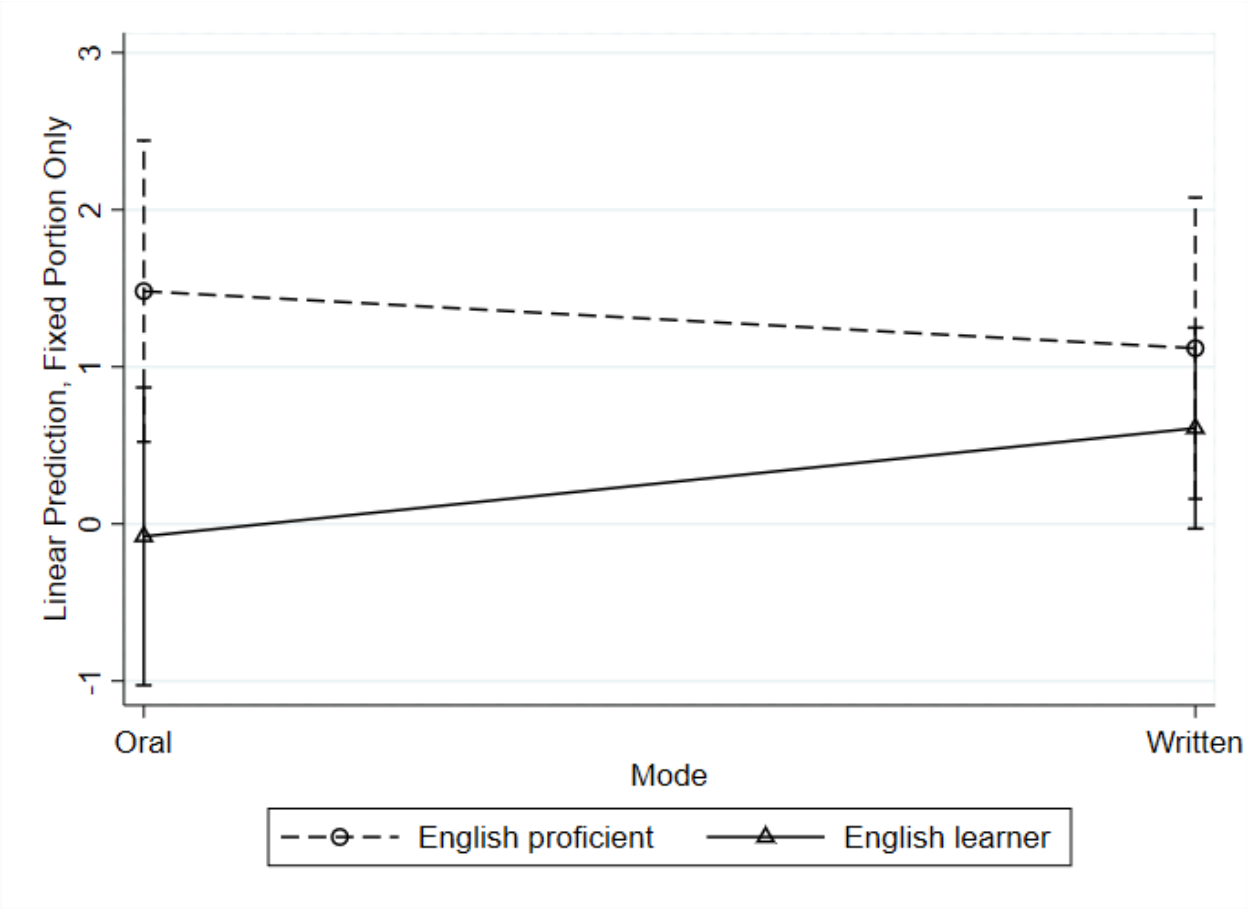

Figure 3. Predictive margins of mode and EL status for topic vocabulary

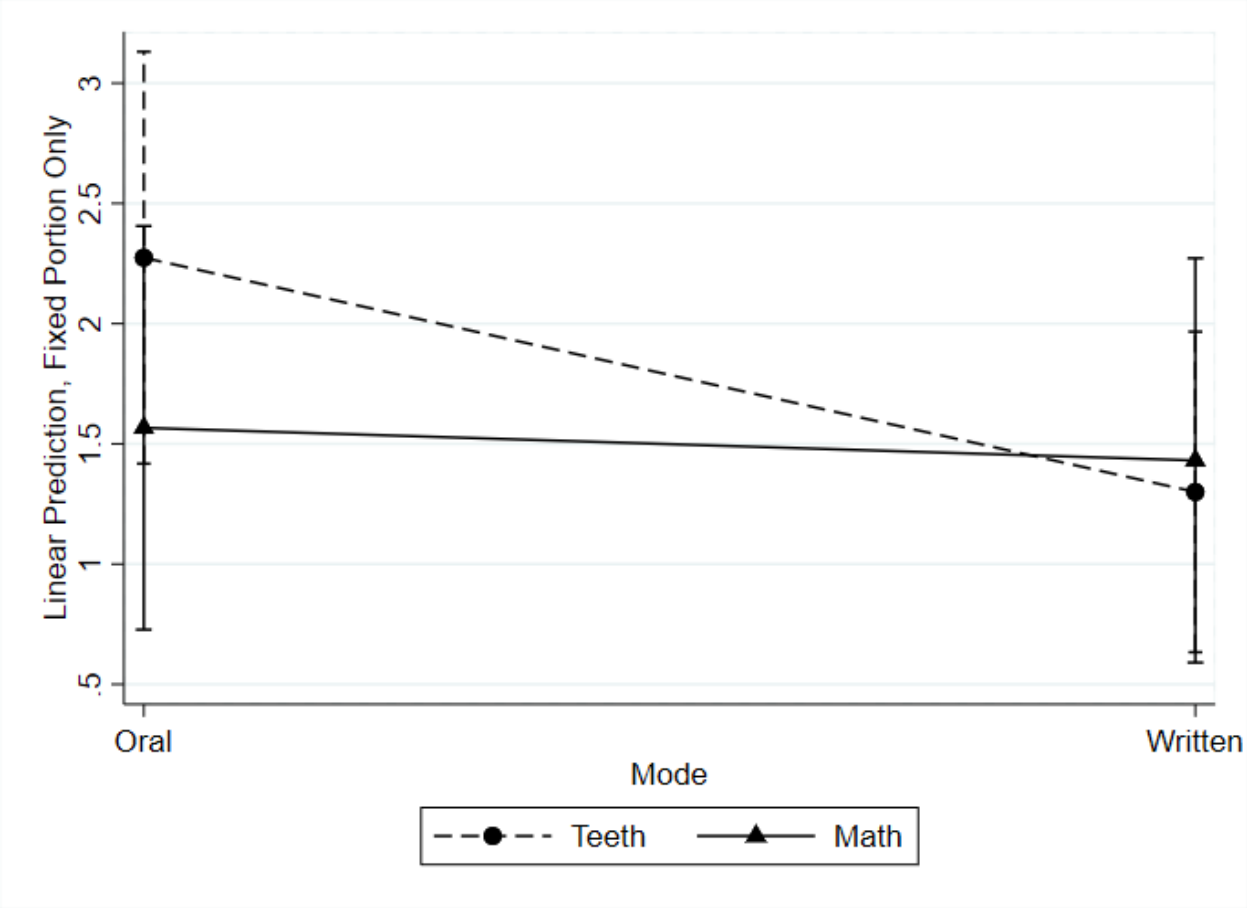

Figure 4. Predictive margins of mode and task for coherence/cohesion

Table 1
Frequency of Language Feature Placements by Mode and Task

	Mode		Task	
	Oral (%)	Written (%)	Non-academic (%)	Academic (%)
Topic vocabulary:				
Not evident	0.4	2.3	1.2	1.6
Emerging	16.4	12.9	10.9	18.4
Developing	53.1	58.6	55.5	56.2
Controlled	30.1	26.2	32.4	23.8
Sentence structure:				
Not evident	0.4	1.9	2.0	0.4
Emerging	5.9	7.4	8.2	5.1
Developing	48.8	55.9	56.6	48.0
Controlled	44.9	34.8	33.2	46.5
Coherence/Cohesion:				
Not evident	5.1	14.9	9.8	10.2
Emerging	55.9	54.3	52.7	57.4
Developing	27.3	19.9	23.4	23.8
Controlled	11.7	10.9	14.1	8.6

Note. N = 256 explanations for each mode/language feature or task/language feature combination.

Table 2
Results of Models for Topic Vocabulary

	Model 1	Model 2	Model 3
	Odds Ratio	Odds Ratio	Odds Ratio
Level-1 fixed effects (explanation):			
Mode			
Oral	1.12 (.20)	1.11 (.20)	1.25 (.56)
Task			
Math	.56*** (.10)	.56*** (.10)	.75 (.19)
Level-2 fixed effects (student):			
Gender			
Female		1.99** (.49)	3.15*** (.98)
EL Status			
Non-EL		2.77** (.96)	1.84 (.75)
Grade			
4th		.62 (.31)	.62 (.32)
5th		1.19 (.45)	1.19 (.46)
6th		2.05 (1.09)	2.08 (1.12)
Interaction effects:			
Oral x math			.55 (.20)
Oral x non-EL			2.35* (.98)
Oral x girl			.41* (.15)
Variance component:			
Student	.86 (.30)	.66 (.27)	.73 (.28)

Note. Standard errors in parentheses. In all models: explanations, $n = 512$; students, $n = 128$; classrooms, $n = 12$
* $p < .05$.
** $p < .01$.
*** $p < .001$.

Table 3
Results of Models for Sentence Structure

	Model 1	Model 2	Model 3
	Odds Ratio	Odds Ratio	Odds Ratio
Level-1 fixed effects (explanation):			
Mode			
Oral	1.69** (.31)	1.69** (.32)	3.86** (1.80)
Task			
Math	2.00*** (.38)	2.01*** (.38)	2.61*** (.69)
Level-2 fixed effects (student):			
Gender			
Female		1.48 (.37)	2.03* (.64)
EL Status			
Non-EL		1.91* (.59)	2.27* (.86)
Grade			
4th		2.02 (.83)	2.04 (.84)
5th		2.63** (.83)	2.67** (.84)
6th		3.74*** (1.44)	3.76*** (1.47)
Interaction effects:			
Oral x math			.60 (.22)
Oral x non-EL			.73 (.31)
Oral x girl			.54 (.20)
Variance component:			
Student	.81 (.31)	.72 (.28)	.75 (.29)

Note. Standard errors in parentheses. In all models: explanations, $n = 512$; students, $n = 128$; classrooms, $n = 12$
* $p < .05$.
** $p < .01$.
*** $p < .001$.

Table 4
Results of Models for Coherence/Cohesion

	Model 1	Model 2	Model 3
	Odds Ratio	Odds Ratio	Odds Ratio
Level-1 fixed effects (explanation):			
Mode			
Oral	1.72** (.31)	1.71** (.31)	3.82** (1.70)
Task			
Math	0.73 (.13)	0.73 (.13)	1.14 (.29)
Level-2 fixed effects (student):			
Gender			
Female		2.10** (.51)	2.86** (.89)
EL Status			
Non-EL		1.82* (.55)	1.94 (.72)
Grade			
4th		1.01 (.40)	0.99 (.40)
5th		1.20 (.36)	1.21 (.37)
6th		2.40* (.88)	2.43* (.90)
Interaction effects:			
Oral x math			.43* (.15)
Oral x non-EL			.91 (.37)
Oral x girl			.57 (.20)
Variance component:			
Student	.87 (.29)	.72 (.26)	.75 (.26)

Note. Standard errors in parentheses. In all models: explanations, $n = 512$; students, $n = 128$; classrooms, $n = 12$
* $p < .05$.
** $p < .01$.

Table 5
Summary of Main Effects and Interaction Effects

	Topic Vocabulary	Sentence Structure	Coherence/Cohesion
Main Effect			
Mode	No difference between modes	Oral higher than written	Oral higher than written
Interaction Effects			
Gender:			
Girls	Written higher than oral	-	-
Boys	Oral higher than written	-	-
EL Status:			
English Proficient	Oral higher than written	-	-
English Learner	Written higher than oral	-	-
Task:			
Academic	-	-	No difference between modes
Non-academic	-	-	Oral higher than written

Note. A dash (-) indicates no significant interaction effect.

Notes

[1] We did not investigate interactions between mode and grade level, because there is no substantively meaningful reference group for comparisons of the four grade levels, and the addition of three extra parameters would have made our models less parsimonious.

[2] The incomplete data for these 61 third through sixth graders was a result of data collection scheduling and prioritizing data collection with a longitudinal subsample. Few student absences accounted for incomplete data and no students withdrew from the study. These 61 students did not differ from the analytic sample in terms of gender, race/ethnicity, or EL status.

[3] This student was yet to be redesignated by the district as Fluent English proficient, which is a decision that involves more than ELP test performance alone.

[4] The pronoun used to describe the friend (he/she) was matched to the student's gender.

[5] For each of our three outcome variables, we found that these two-level models provided a better fit over three-level models that included classroom (i.e., explanations nested within students, nested within classrooms) or school (i.e., explanations nested within students, nested within schools). This is because the variation between individual students was much larger than the variation between classrooms or schools.

[6] The Fisher's exact test for grade was conducted using a 2-level grade variable that combined third/fourth graders in one group and fifth/sixth graders in another; this streamlined computation and enabled us to determine whether there were differences in the placement of explanations produced by students in intermediate versus upper elementary grades.

This work was supported by funding to Alison Bailey from the WIDA Consortium at the Wisconsin Center for Educational Research and the WIDA ASSETS Enhanced Assessment Grant from the U.S. Department of Education. However, the content does not necessarily represent the policy of the U.S. Department of Education and you should not assume endorsement by the federal government. Alison Bailey also acknowledges serving as a consultant and advisory board member for WIDA projects. We thank the students, teachers and administrators who participated in the DLLP project, as well as colleagues and graduate students who assisted, particularly Despina Goral, Eve Ryan, Sandy Chang, and Margaret Heritage. We also thank the editors and anonymous reviewers for their helpful comments and suggestions.