

Language Sample Analysis of Writing in Children and Adolescents

Assessment and Intervention Contributions

Cheryl M. Scott

Competent writing eludes many school children and adolescents, especially those with developmental language and reading disorders. Language sample analysis (LSA) of writing is a powerful but underused assessment tool that can also inform instruction. Similar to LSA of oral language, writing has been analyzed at word, sentence, and text levels using measures that include lexical diversity, sentence length, clause density, grammaticality, productivity, and text organization and content. Unique observations of writing include spelling and literate semantic and syntactic features. This article offers a review of analytic writing measures in the LSA literature from perspectives that include developmental change, language ability differences, relation to quality ratings, practical utility, and effects of genre and task. Writing samples from two 12-year-old students, with and without a language disorder, are used to illustrate application of these measures and suggest potential instruction targets. **Key words:** *adolescents, assessment, children, genre, grammar, language disorder, language sample analysis, struggling writers, writing*

FROM EARLY elementary through the secondary school years, writing holds a prominent place in school curricula. This is evident when examining the considerable detail devoted to grade-specific writing standards adopted across states (Common Core State Standards Initiative, 2010), as well as the place of writing as a separate assessment, alongside reading, in the National Assessment of Educational Progress (NAEP), administered by the U.S. Department of Education. Amidst the backdrop of its importance, however, is

the reality that many school children struggle to write well. In a population-based birth cohort study of written language disorders (WLD), epidemiologists established cumulative incidence rates that varied from 6.9% to 14.7% depending on information available in school records (Katusic, Colligan, Weaver, & Barbaresi, 2009). A considerable number of children and adolescents write below an expected standard according to results of the writing assessment conducted as part of the NAEP. On the 2011 assessment (the most recent results available), slightly over half of the nation's 8th- and 12th-grade students wrote at a basic level, meaning they had not fully mastered essential writing skills for their respective grades. Only a quarter were proficient writers (National Center for Education Statistics, 2012).¹

Author Affiliation: *Department of Communication Sciences and Disorders, Rush University, Chicago, Illinois.*

The author has indicated that she has no financial and no nonfinancial relationships to disclose.

Corresponding Author: *Cheryl M. Scott, PhD, Department of Communication Sciences and Disorders, Rush University, 600 S. Paulina St, Chicago, IL 60612 (cheryl_m_scott@rush.edu).*

DOI: 10.1097/TLD.0000000000000213

¹For a description of basic, proficient, and advanced writing levels at 4th, 8th, and 12th grades, see the NAEP

In the course of explaining criteria used to identify WLD from school records, Katusic et al. (2009) noted that “currently there are no universally accepted tests, assessment batteries, or standards for identifying children with WLD” (p. 1311). Along the same lines, Troia (2009) wrote that it is difficult to determine how well a student writes because of the myriad of factors that can shape a final product (e.g., task, topic, background knowledge, motivation, instruction). And, the evaluation of any one student’s writing can be further complicated by the high probability that other developmental language-based disorders (reading, oral language) or comorbid conditions may compete for attention.² Rarely are writing problems seen in isolation, nor are they the only factor impacting a student’s academic achievement (see *Diagnostic and Statistical Manual*, 5th edition, American Psychiatric Association, 2013). Perhaps it should not come as a surprise when surveys report that many speech-language pathologists and classroom teachers feel unprepared to address writing difficulties of their students (Fallon & Katz, 2011; Graham, Cappizi, Harris, Hebert, & Morphy, 2014; Pavelko, Owens, Ireland, & Hahs-Vaughn, 2016).

In this tutorial, I argue that language sample analysis (LSA) of writing offers promise for narrowing the gap between need and current clinical and educational practice. Although the LSA literature on child and adolescent writing is not as extensive as that on spoken language, there is by now enough research to help practitioners determine whether a student’s writing meets peer-referenced standards for age and genre and to suggest targets for individualized intervention. The first section provides a framework for LSA based on a developmental model of

the writing process and two critical skills needed for successful school writing—facility with genre and written sentence form. In the second section, I use a multidimensional model of language to explore commonly used LSA measures at word, sentence, and text levels. Developmental and language ability differences, relation to quality ratings, effects of genre and task, and practical utility also are addressed. The third section illustrates application of these measures in expository writing samples of two 12-year-old students, one with typical language development (TD) and the other with specific language impairment (SLI), and suggests intervention targets that follow from the analysis. In concluding remarks, contributions of LSA are summarized and future directions are suggested. When considering educational and clinical applications of the information in this article, the student I have in mind is one who lags behind age/grade peers on writing assignments. Some of these students may have an individualized educational program under the classification of specific learning disability, speech-language impairment, or other disabilities, but others may not. The term *struggling writer* is often used in articles and books for this broad, high-incidence group of school-aged students.

THE PLACE OF LSA IN A DEVELOPMENTAL FRAMEWORK

Developmental models of writing

Writing is one of the most complex things students do. The cognitive and linguistic requirements are substantial. When Hayes and Flower (1980) asked adult writers to think aloud about what was going on in their minds as they wrote, the cognitive processes involved became more transparent. The planning, generating, and revising stages of writing looked less sequential and more integrated and recursive. Later, Berninger and her colleagues proposed two models of writing appropriate for children; one they called the simple view of writing, and a later

Writing Achievement Levels, retrieved from <https://nces.ed.gov/nationsreportcard/writing/achieve.aspx>

²The incidence of WLD in children with ADHD was more than 50% (Yoshimasu et al., 2011) in the same birth cohort cited above (Katusik et al., 2009).

iteration the not-so-simple view (see reviews of both models in Berninger, Garcia, & Abbott, 2002). In the simple view, a transcription component (handwriting, keyboarding, and spelling) and an executive function component (self-regulation of attention and strategy use for planning, reviewing, and revising) feed into the component of text generation (also called ideation or translation) where words, sentences, and texts are produced. The not-so-simple view advocates for the central role of memory processes, both long- and short-term, as they interact with transcription, as well as an (expanded) executive function component, leading to text generation. Work concerned with further specifying the nature of developmental writing components continues; for example, the work of Kim and Schatschneider (2017) investigates those aspects of text generation (e.g., foundational oral language skills, inference) that best account for quality ratings.

Where does LSA fit within a developmental model of writing? By definition, LSA examines the product of the text generation component—the words, sentences, and text that are committed to paper (or screen) to complete a writing prompt or assignment. The status of a child's spelling and handwriting (output of the transcription component) can be assessed as well. Although executive function and memory processes are less directly accessible, LSA of written text can point in the direction of underlying processes that could be impacting writing. For instance, grammatical errors in long, complex sentences could point to the lack of long-term memory templates for complex sentences or to a lack of attention control mechanisms needed for ongoing monitoring or revision. Processes hypothesized to explain observations in written text could then be assessed more directly. Recently, several researchers have evaluated planning and revision directly in written language samples (e.g., instances of crossing out words or phrases indicating revision by Troia, Shen, & Brandon, 2019; number of ideas generated in planning outlines by Koutsoftas, 2016). In these studies, initial

directions to writers increased the likelihood that planning or revision would occur (e.g., “take time to plan” . . .). However, most studies reviewed later have not addressed planning or revising in directions to participants and do not report on these writing processes.

Genre considerations

My focus in this tutorial is on academic writing, from early elementary through secondary school, rather than texting, messaging, e-mailing, and posting on social media communication platforms used primarily during off-school hours.³ Genre refers to the broad type of text and current state and national curriculum standards have been written to address the three broad genres of narrative, expository, and persuasive writing. The Common Core State Standards expect second-grade children to compose text fairly independently in all three genres and recommend introducing the three purposes of writing in kindergarten (cf., Common Core State Standards Initiative, 2010). State and national assessments of writing (e.g., the NAEP assessment) use prompts geared to the same three genres. Likewise, LSA studies of child or adolescent writing reviewed in the next section have used narrative, expository, or persuasive writing prompts, so there is a good match between LSA research and school-sanctioned writing tasks. Several studies have compared two genres to explore structural differences, such as asking whether narratives or expository texts differ in terms of sentence complexity at particular ages.

It is important to consider effects of genre in LSA. For any given student, proficiency in one type of writing may not “hold” for another. One explanation follows from research on the development of genre. Narratives, based on experiences or themes often familiar to young children and with elements

³Although children's messaging/writing via social media is of great interest, there are less data on this type of writing.

arranged chronologically, are commonly the first texts children write independently, often by the end of first grade (Sulzby, 1996). Also, narratives are often emphasized in early elementary curricula (Scott, 2012). Expository texts, on the contrary, are logically based and often deal with unfamiliar topics. Anecdotally, I remember my own daughters as second and third graders writing multipage fiction stories with all components characteristic of narratives (setting, initiating event, etc.), while at the same time their attempts at expository writing were much shorter, lacked overall macrostructure (text organization schema elements such as introductions or conclusions), and read more like a list of facts they knew about a topic. Evidence suggests that students first gain proficiency in narrative writing, followed by expository writing, and then persuasive writing (as reviewed by Scott, 1994, 2012). Language sample analysis investigations generally reflect the same developmental sequence in that researchers have used narrative prompts when focusing on younger writers, while introducing expository or persuasive prompts for older writers. Making decisions about genre skills for a particular writer is complicated by the fact that the developmental course of writing is a long one extending throughout the school-age years. As shown later, however, general guidelines for how well a piece of writing conforms to genre expectations can be gleaned from LSA.

Syntax considerations

When attempting to specify the nature of the text generation component of writing within a developmental model, Kim and Schatschneider (2017) noted that the ability to formulate words and sentences when writing is based on oral language; therefore, they included oral language measures of vocabulary and grammar in their model. Although it goes without saying that we write the same language we speak, there are important syntactic differences in the two modalities and children's writing begins to reflect such differences by mid-elementary years (Kroll, 1981). Writing, devoid of intona-

tion and listener feedback found in speaking, communicates theme and focus by placing the most important information at the end of a sentence and moving background information to the front. Two examples include (a) the fronting of adverbial clauses to sentence initial position (e.g., *After making her way to the podium*, the candidate spoke at some length about her views on health care), and (b) cleft constructions (e.g., *I walked through the shelter and there in the last cage of the last row was the most adorable puppy*). Scott and Balthazar (2010) discuss other grammatical features found more often in written text, including long and complex noun phrases, multi-clausal subordination/embedding (the packing of several clauses that relate logically into a single sentence), passive voice, and nominalization (turning verbs into nouns (e.g., *evaporate* > *evaporation*). Perera (1984) and Scott (1988a,b; 2012) have summarized a developmental progression for many of these syntactic features in children's writing. In LSA research to date, these types of structural features characteristic of writing have been neglected, even in older age groups, but I will suggest that a more nuanced analysis should consider them. Perera stated that even a single occurrence of such a structure is important and has an impact on quality (1984, p. 248). This is particularly true for younger children, starting at age 9 or 10; by later middle school years, frequencies of complex syntactic structures should increase.

LSA MEASURES AND THEIR INTERPRETATION

In this section, I use a language-level framework to consider LSA measures at word, sentence, and text levels. This approach reflects the fact that there are many features that describe a piece of writing (e.g., complexity and accuracy at both word and sentence levels) and contribute to writing quality. Studies reviewed here have used analytic or quantitative methods, where word, sentence, and

text features are counted and quantified in some manner. Although research is ongoing, there is increasing evidence validating the use of analytic methods within levels of language framework for evaluating writing quality and distinguishing age and language ability groups of school-aged children (Troia et al., 2019). This way of looking at writing samples contrasts with holistic methods, or qualitative methods, where a sample is rated on a small number of traits using rubrics such as the *6 + 1 Traits of Writing* in which organization, word choice, sentence fluency, conventions, and so forth, are evaluated (Culham, 2003).

There are scores of possible things that could be identified and counted in a writing sample. Measures discussed later have been chosen because they are common across studies and reflect productivity, grammaticality, and complexity features shown to be predictive of writing quality (Troia et al., 2019). To assist clinicians in the interpretation of these measures, when there is sufficient evidence, I address the following questions: (1) Is this measure sensitive to developmental changes (age/grade differences)? (2) Is this measure sensitive to group membership classification as either TD or developmental language disorders including SLI or language learning disorders/learning disabilities?⁴ (3) How well does this measure predict quality ratings? (4) Overall, does this measure have practical utility for educators and clinicians?

Written words

By far, the most common word-level measure included in LSA studies of writing to date

is lexical diversity, coded as the number of different words (NDW). A skilled writer has the ability to draw on a large NDW rather than reusing the same ones over and over. This measure is calculated automatically by the Systematic Analysis of Language Transcripts (SALT) computer program—an analysis tool used often in both oral and written language LSA research (Miller & Iglesias, 2016). Although studies of NDW that include participants across four or more grades usually find a significant main effect, it is less common to find significant change in 1- or 2-year comparisons (Nelson & Van Meter, 2007; Wood, Bustamante, Schatschneider, & Hart, 2019). So, comparing a second grader with a fifth grader, one is likely to find a significant difference in NDW but much less likely in a comparison of a fourth-grade writer with a fifth-grade writer. On a cautionary note, when comparing two developmental groups (or individuals), the sample length should be the same. This is because sample length as measured in total number of words (TNW) shows robust growth with age and NDW naturally rises along with sample length. It is therefore important to note whether a researcher has controlled for sample length by using truncated samples (most frequently 50- or 100-word samples) when comparing age or ability groups.

Comparisons of language ability groups on NDW are mixed. Some studies have not reported significant results (Scott & Windsor, 2000, using 50-word samples) whereas others have. A recent meta-analysis of the writing of students with learning disabilities (LD) did find a significant effect size (-0.89) for vocabulary (defined as diversity and accuracy) across 10 studies comparing LD and TD students (Graham, Collins, & Rigby-Wills, 2017). As one would expect, NDW was moderately correlated with a standardized test of reading vocabulary in upper elementary grades in a study by Wood et al. (2019)—a finding that speaks to lexical diversity as a cross-modality core linguistic trait. Nonsignificant correlations at lower grade levels in the same study, however, are indicative of an often-repeated

⁴Most research exploring writing skills in language ability groups have compared children or adolescents with typical language development (TD) and those meeting inclusion criteria for specific language impairment (SLI) or language learning disability (LLD). In SLI and LLD, language is disordered in spite of broadly normal functioning in other cognitive domains and no other neurodevelopmental or medical diagnoses (e.g., ASD, FX). When discussing results from these studies, I adhere to terminology used by study authors.

finding of the effect of age in both spoken and written LSA research—a significant finding at one age is not significant at another.

In terms of clinical or educational utility, NDW has limitations. For one, it is difficult to see what a significant difference in a group study signals clinically. To illustrate, Koutsoftas and Gray (2012) reported means of 36 for typically developing students (TD) versus 33 for language learning disability (LLD)—a significant difference, but the clinical meaning of that difference is not immediately evident. Although the literature supports slow increases with age and there is some support for language ability differences, interpretation for an individual student is problematic. One can imagine that genre and topic could affect NDW in major ways. For example, narratives contain proportionally higher percentages of pronouns that repeat, compared with science topics. This and other genre-specific content and structural features could easily impact NDW. Younger children, for whom spelling can be a slow and arduous process, write less, using a more constrained vocabulary, perhaps even consciously using words they have more confidence spelling correctly. Unless a clinician evaluating an individual child consulted comparison values derived from the same topic, task, and sample length, interpretation would be suspect.

Given these issues with NDW, practitioners are encouraged to examine a piece of writing for specific types of words known to reflect developmentally higher level of vocabulary skills. Nippold (1998) has advanced the notion of a “literate lexicon” that includes words like adverbs of magnitude and likelihood (*probably, somewhat, extremely*) and metacognitive verbs (*remember, decide, conclude*). Words with derivational affixes (e.g., *generous, unfaithful*) are further examples as are later-developing subordinate conjunctions (*whenever, although*) used to connect clauses, and adverbial conjuncts (*further, however, in conclusion*) that connect two sentences. In their study of persuasive writing in 11-, 17-, and 24-year-old individuals

with typical language skills, Nippold, Ward-Loneragan, and Fanning (2005) found developmental increases in the use of adverbial conjuncts, abstract nouns, and metaverbs. For instance, 24-year-olds used three times the number of abstract nouns as 11-year-olds. In a study of adolescent narrative writing, Sun and Nippold (2012) counted instances of abstract nouns and metacognitive verbs and found a significant difference for age at 11, 14, and 17 years. Not only did the older students use these types of words with greater frequency, they used them with increased diversity (different exemplars). Another way of noting lexical strengths in writing samples is to look for long words (mean number of syllables or letters/word). Because of an inverse relationship between word length and word frequency (Zipf, 1932), word length is a proxy for lower frequency words. Troia et al. (2019) included two measures indicative of the use of longer, lower frequency words in their study of narrative writing of fourth- through sixth-grade students but found little change in mean syllables/word (1.21, 1.23, 1.22 in fourth, fifth, sixth) or in a word frequency metric across the three grades. In an analysis of five expository writing samples (three students with LLD, and two with TD) from the Scott and Windsor database (2000), I identified low-frequency (higher level) vocabulary words in each sample (Scott, 2009, p. 369). In the three LLD samples, there were a total of 40 words (e.g., *predators, moisture, adapted*), but in the two TD samples, there were 75 such words (*occasional, competition, mainly*). It is highly likely that genre and task influence a writer’s use of longer, lower frequency words.

Spelling, transparent in handwritten samples (or in computer samples with spell-checking disabled), also should be examined in a writing sample. Several LSA studies include information about spelling performance, calculated as the percentage of correctly spelled words divided by total words, or the reverse, the percentage of misspelled words (Coker, Ritchey, Uribe-Zarain, & Jennings, 2018; Dockrell, Lindsay,

Connelly, & Mackie, 2007; Koutsoftas & Gray, 2012; Nelson & Van Meter, 2007; Puranik, Lombardino, & Altman, 2008; Troia et al., 2019). These studies have shown that younger children with typical language (first and second grades) spell 80%–84% of their words correctly and this percentage increases to around 95% by the fifth grade. Studies that include ability comparisons of spelling have found that children with language disorders are about 10% less accurate than their TD peers (Koutsoftas & Gray, 2012; Nelson & Van Meter, 2007). Besides proportion of correct spellings, further analysis of types of errors is seldom reported in LSA studies but could be undertaken by clinicians looking for intervention direction. For example, consider the word *trapped* spelled as *trap* by one child versus *trappt* by another. Both misspellings count equally in a rate calculation. However, in the first instance, the speller misses any representation, even phonological, of the obligatory past tense inflectional morpheme *-ed*, whereas in the second case, the morpheme is represented phonologically as is the consonant doubling rule, indicating greater phonological and orthographic knowledge.

Written sentences

Complexity

Sentences carry a heavy burden in writing. They must follow multiple grammatical rules to convey simple and complex ideas and their chronological and logical relationships. Also, sentences must indicate how information in any one sentence relates to preceding information and previews upcoming information. It is helpful to remember that the average adult informational written sentence is 22 words (Francis, Kucera, & Ackers, 1982)—a considerable expanse of words to coordinate for structural and semantic purposes. And, it is not uncommon to find sentences of 30 or more words in adult reading material. To become fully “linguistically literate,” developing writers learn to encode complex ideas in various registers, genres, and content areas

using increasingly complex sentences (Ravid & Tolchinsky, 2002).

The two most common measures of overall sentence complexity calculated in LSA studies of writing are (a) sentence length, usually in words rather than morphemes, and (b) the extent to which sentences are simple (only one clause) or complex (more than one clause). Most research studies have measured sentence length as mean length of T-unit (MLTU). The T-unit is a standard way of segmenting a sample into sentence-like units based on clear structural features and used in many studies of writing since first proposed by Hunt (1965). Of note, T-units may or may not match a child’s uses of punctuation. The extent to which a writer uses multi-clause sentences is captured by the measure of clause density (CD), defined as the number of clauses (main and subordinate) in a text divided by the total number of T-units. In some studies, this same measure is referred to as the subordination index (SI). To calculate either, one looks at each sentence separately and marks the number of clauses, then adds these for all sentences (T-units) in the sample and divides by the number of sentences. A CD of 1.10 would indicate that the writer’s sentences are usually one-clause (simple), but a CD of 1.90 (close to 2.0) communicates that a writer routinely constructs sentences with two clauses (along with some single-clause and some 2+ clause sentences).

Table 1 shows MLTU and CD means from LSA studies of writing arranged in rows from first through 11th grades. Some investigations were restricted to just one grade/age, whereas others span several age groups. Within and across studies, we can see developmental change in MLTU over the full span of school years such that first- and second-grade children write sentences less than eight words, but by sixth grade, they are writing 10-word sentences, and in high school well more than 10 words. Like NDW, however, for shorter periods of time such as between adjacent grades, or even spans of 2 or 3 years, progress is not consistent. The same grade/age trends hold for CD. Children in first

Table 1. Mean values for sentence complexity measures of length (MLTU) and CD in studies of language sample analysis of writing arranged by grade

Grade	MLTU		CD	
	TD	SLI or LLD	TD	SLI or LLD
1 (Coker et al., 2018)	7.96 (N) 6.16 (E)		1.13 (N) 1.08 (E)	
1 (Nelson & Van Meter, 2007)	5.88 N	5.83 M		
2 (Nelson & Van Meter, 2007)	6.70 N	6.23 N		
2 (Hall-Mills & Apel, 2015)	7.51 (N) 7.58 (E)		1.40 (N) 1.25 (E)	
3 (Hall-Mills & Apel, 2015)	8.49 (N) 8.58 (E)		1.43 (N) 1.58 (E)	
3 (Nelson & Van Meter, 2007)	7.66 N	6.50 N		
3 (Puranik et al., 2008)	9.6 E		1.78 E	
4 (Puranik et al., 2008)	10.5 E		1.77 E	
4 (Hall-Mills & Apel, 2015)	7.98 (N) 8.33 (E)		1.46 (N) 1.61 (E)	
4 (Nelson & Van Meter, 2007)	8.23 N	7.13 N		
4-5 (Koutsoftas & Gray, 2012)			1.47 (N) 1.69 (E)	1.28 (N) 1.55 (E)
5 (Nippold & Sun, 2010)	12.33 E			
5 (Sun & Nippold, 2012)	9.14 N		1.50 N	
5 (Puranik et al., 2008)	10.5 E		1.83 E	
5 (Nelson & Van Meter, 2007)	8.46 N	7.23 N		
5-6 (Scott & Windsor, 2000)	10.3 (N) 11.4 (E)	9.1 (N) 9.7 (E)	1.9 (N) 1.79 (E)	1.75 (N) 1.66 (E)
6 (Nippold et al., 2005)	11.29 P		1.63 P	
6 (Puranik et al., 2008)	10.3 E		1.82 E	
7-8 (Beers & Nagy, 2009)	11.0 (N) 15.0 (P)		1.5 (N) 2.0 (P)	
8 (Nippold & Sun, 2010)	14.53 E			
8 (Sun & Nippold, 2012)	11.19 N		1.71 N	
9 (Brimo & Hall-Mills, 2019)			2.3 (E) 2.58 (P)	
11 (Sun & Nippold, 2012)	11.27 N		1.63 N	
11-12 (Nippold et al., 2005)	13.48 P		1.67 P	

Note. Troia et al. (2019) provide measures of sentence complexity (words/sentence and percent sentences with subordination) for a large-N study of fourth-, fifth-, and sixth-grade writers, but these are not included in Table 1 because of differences in how these measures were calculated. CD = clause density; E = expository; LLD = language learning disability/disorder; MLTU = mean length of T-unit; N = narrative; P = persuasive; SLI = specific language impairment; TD = typical language development.

grade rarely write multclause sentences, but by fourth grade and beyond, such sentences are becoming more common. Language ability (see studies by Nelson & Van Meter, 2007; Koutsoftas & Gray, 2012; and Scott & Windsor, 2000, in Table 1) impacts both MLTU and CD in the expected direction with children from SLI and LLD groups writing shorter sentences with less clausal subordination than their TD peers. This agrees with findings in a meta-analysis of studies comparing students with LD and their TD peers where sentence fluency, a measure that included sentence length, showed a significant effect size (Graham et al., 2017).

By far, the most frequently sampled genre in LSA of writing is the narrative, particularly at younger ages. Four studies in Table 1 (see studies by Coker et al., 2018; Hall-Mills & Apel, 2015; Koutsoftas & Gray, 2012; and Scott & Windsor, 2000, in Table 1) compared sentence complexity in narrative and expository writing. In lower grades, narratives, common in younger children's school writing, exceed expository in global measures of sentence complexity (Hall-Mills & Apel, 2015). Then, by the late elementary years as children are exposed to more expository texts and presumably have more practice writing in the expository genre, these samples post higher length and clausal complexity values than narratives. In older students, sentences in persuasive writing are more complex than in both narrative and expository (Beers & Nagy, 2009; Brimo & Hall-Mills, 2019).

For an assessment of any one student's writing sample(s) for clinical purposes, and assuming an "apples to apples" comparison of genre and task, clinicians can use the values in Table 1 to gauge "ballpark" comparisons with average MLTU and CD complexity values, keeping in mind that standard deviations behind these averages are often large. Sentence length, by itself, is not a measure that leads directly to any meaningful intervention goal; rather, an increase in sentence length would be a by-product of most structural and semantic goals that add complexity of ideas and nuance to a text (see Scott & Balthazar,

2013, for a list of examples of structures that would increase sentence length).

Clause density translates more directly to intervention goals because a low value points directly to the goal of increasing the writer's ability to combine clauses in sentences. Moreover, increasing clausal subordination is a manageable target because there are 3 major categories of subordinate clauses (adverbial, object complements/nominal, and relative clauses) and developmental patterns are fairly well established (Perera, 1984; Scott, 1988a,b). A low CD value should prompt finer-grained analysis of the types of subordination that are underused. Several studies that include details about the types of subordinate clauses used in writing provide guidance for such an analysis. We know that adverbial clauses (e.g., My friend got in trouble *because she came home late*) and object complement clauses (e.g., I think *there should be less homework in 4th grade*) are used earlier and more often in children's writing than relative clauses (e.g., The candidate *who wins the primary* goes on to the general election; Scott, 2003). In an analysis of persuasive writing (Nippold, Ward-Lonergan, & Fanning, 2005), 11-year-olds used adverbial clauses (found in 19% of sentences) and nominal clauses (36% of sentences) at frequencies similar to 17-year-olds. Relative clauses occurred less often for both groups, but usage was significantly higher for the older writers. When it comes to types of subordination, variety is a good thing. Two studies have reported significant differences between TD and LLD groups when comparisons are made of the extent to which writers *combine* various types of subordinate clauses within sentences (Gillam & Johnston, 1992; Scott, 2003). It goes without saying that increases in CD (using more subordinate clauses per sentence) would simultaneously increase MLTU.

Grammaticality

After MLTU and CD, the third most quantified sentence variable in children's writing is overall grammaticality (also called accuracy). Grammaticality has usually been measured as

either an error rate per T-unit or percentage of correct T-units. Proportion of errors in narratives for TD students in middle to late elementary years has been reported as 0.07 (Koutsoftas & Gray, 2012), 0.22 (Hall-Mills & Apel, 2015), and 0.11 (Scott & Windsor, 2000), with higher rates for expository (0.09, 0.29, and 0.15 in the same studies, respectively). For TD students in the same grades, we can expect 78%–82% of all T-units to be error-free (Hall-Mills & Apel, 2015; Puranik et al., 2008).

Grammaticality is an important observation because, unlike MLTU and CD, it consistently distinguishes children with TD from those with developmental language disorders and does so across grades and genres. Error percentages for students with language disorders are considerably higher, approximately three times as high. In several studies with group comparisons that have included language age (LA) matches in addition to chronological age (CA) matches, children with lower language ability make more grammatical errors than *both* CA- and LA-matched children, where LA groups are typically 2 or more years younger than their peers with language disorders (Mackie & Dockrell, 2004; Scott & Windsor, 2000; Windsor, Scott, & Street, 2000). This contrasts with other sentence-level measures where LA matches perform similarly to those children with language disorders.

Several studies have delved into specific morphosyntax errors in children's writing. As reported by Green et al. (2003), inflectional morphology was widely used and mostly accurate in the writing of TD students (88%–94%), depending on the morpheme, by the end of fourth grade. For children with LLD, the picture is very different. Windsor, Scott, and Street (2000) compared CA-, LLD-, and LA-matched groups of middle- to late-elementary-aged students on error proportions for verb finiteness markers (regular past tense, third person singular present tense, copula and auxiliary BE) and noun markers (regular plural, possessive, and articles). Regular past tense markers were omitted

once in every four obligatory contexts (26%) by LLD students, and regular plurals at a rate of 12% by the same students. Comparable rates for CA peers were 1.5% and 2.6% and for LA matches were 3.3% and 5%, respectively. Students with LLD were more accurate in spoken language samples than in written samples. The stark group and modality differences in morphosyntax error rates underscore the centrality of morphosyntax both as a key diagnostic index feature in these children and its fragile representation under the additional stresses imposed by writing when compared with speaking.

Written text

A review of the writing LSA literature reveals common use of two types of quantitative text-level measures. One is length of the text, commonly referred to as productivity and typically measured as TNW or, in some studies, number of utterances (T-units; some studies report both measures). Although productivity has been included with other microstructural measures as a word or sentence characteristic (cf. Hall-Mills & Apel, 2015), I view productivity as a text-level (macrostructure) trait due to the fact that text length is difficult to separate from the overall success of a piece of writing. A text that is too short to accomplish a particular purpose (e.g., telling a good story or offering a sufficient explanation of an event or phenomenon) is not a good text. Furthermore, in most cases, productivity will vary directly with the second common way of quantifying text-level traits—the extent to which a text includes all of the organizational components of good narrative, expository, or persuasive text, which several studies report as a way of quantifying text-level content and organization. For example, Nippold et al. (2005) compared persuasive writing in three age groups (11, 17, and 24 years) by counting the number of reasons offered to support a position; Koutsoftas and Gray (2012) counted the number of complete episodes in narrative writing in their comparison of TD and LLD 11-year-olds.

Compared with quantitative measures at word and sentence levels discussed previously, productivity is a more consistent index of developmental change, even when comparing 1- or 2-year increments (Hall-Mills & Apel, 2015; Koutsoftas & Gray, 2012). Wood et al. (2019) reported a threefold increase in productivity for TD students between first grade and fifth grade on written responses to their narrative prompt. A similarly robust productivity increase was reported by Hall-Mills and Apel (2015) with a twofold increase from second grade to third grade and a threefold increase between second grade and fourth grade in both narrative and expository samples. Nine-year-olds wrote only 60% as many words as 11-year-olds in both narrative and expository summaries in a research study by Scott and Windsor (2000). Nippold et al. (2005) showed a long-lasting but slower rate of change for older TD individuals in persuasive productivity, with 24-year-olds writing twice as much as 11-year-olds.

Productivity also is a robust sign of language ability difference. Narratives written by TD students in the study by Nelson and Van Meter (2007) averaged 34 and 171 words in the first and fifth grades, respectively, but only 24 and 91 words for students with special needs in the same grades. Productivity of the TD group exceeded that of the SLI group for narrative writing but not for expository writing in the Koutsoftas and Gray (2012) investigation, whereas both genres were shorter for LLD children in the research by Scott and Windsor (2000). In two studies that included LA as well as CA matches, children with SLI (Mackie & Dockrell, 2004) and LLD (Scott & Windsor, 2000) were significantly less productive than CA peers at 11 years of age but similar to LA matches who were 2 years younger. One interpretation is that children with language-based literacy problems continued to struggle with writing, even after 2 years of additional instruction and experience.

There are both gender and genre caveats when measuring productivity. Several studies have reported that girls wrote more than

same-age boys on narratives (cf. Fey, Catts, Proctor-Williams, Tomblin, & Zhang, 2004). This might be expected, given the popular notion that girls are acculturated to be more interested in stories. However, some research has shown that they also exceed boys in expository productivity. To illustrate, on an expository prompt, fifth-grade girls wrote an average of 176 words compared with 112 for boys, and they wrote as much as eighth-grade boys, as reported by Nippold and Sun (2010). Genre effects on productivity have been mixed. In a study comparing narrative and expository writing, children's narrative texts were longer than expository texts when the task was one of summarizing videos relating either a story or information (Scott & Windsor, 2000), but there were no genre differences when writing in response to narrative versus expository prompts reported by Hall-Mills and Apel (2015). Productivity findings summarized here underscore the difficulty of generalizing findings from one study to the next or from research to a particular clinical case when prompts and tasks differ.

Genre implies a particular set and organization of text components. For example, narratives typically begin with a setting, then proceed to a problem, followed by the protagonist's reaction to the problem and formation of a plan to deal with the problem and so forth. Persuasive texts state a point of view or opinion on a topic, frequently a controversial one, and proceed to give reasons supporting the opinion. More seasoned persuasive writers anticipate counterarguments that a reader might think of and address these as well. Expository texts have common subtypes that include description, procedure, problem-solution, cause-effect, enumerative, and compare-contrast (Nippold & Scott, 2010), each with its own organization (see also the article by Lundine in this issue). Text-level analysis of writing is therefore specific to the genre, subtype, and topic. The typical process for quantifying how well a writing sample meets expectations involves constructing an organizational template for an expected response and then evaluating the

presence/absence of components that match that template. For example, Scott and Jennings (2004) constructed a template of topics covered in the audio portion of an informational video that 11-year-old participants were asked to summarize. The children's written summaries were then matched against that template. Group results showed that, compared with CA peers, children with LLD wrote summaries that (a) were half the length (TNW), (b) addressed fewer topics, (c) addressed topics less completely, and (d) contained fewer topic generalizations. Although this example might seem tedious for clinical purposes, the basic procedure would be similar—that of analyzing a piece of writing against an expected organizational template specific to the task, content, and genre. Some researchers have used more holistic analyses of text structure. For example, Hall-Mills and Apel (2015) assigned a number to writing samples based on the sum of ratings of organization, text structure, and cohesion, each rated on a 4-point scale; Nelson and Van Meter (2007) rated narratives according to a developmental maturity rubric from 1 to 6.

Markers of text cohesion also are an important text-level observation. The student who enumerates reasons in a persuasive piece with the connectives *first*, *second*, and so forth, and ends with *in conclusion* demonstrates knowledge about text cohesion. Likewise, the student who uses ellipsis or substitution to tie successive sentences together shows cohesion skill (e.g., *She failed to take any notes during the interview. That was a serious mistake*; in the second sentence, the pronoun *that* substitutes for the entire clause forming the first sentence).

Analytic measures as predictors of writing quality

A recent analysis of relationships between analytic measures and narrative writing quality was undertaken by Troia et al. (2019). In a study of 362 students in grades 4 through 6, the researchers measured 17 writing variables spanning word-, sentence-, and text-level features; narrative quality was an average rating across five traits (conventions,

sentence fluency, word choice, organization, and ideas). Using hierarchical regression analysis, word-level predictors of writing quality included accuracy (a measure of spelling and capitalization), lexical diversity, and content word frequency. Sentence-level predictors were percent grammatical sentences and mean words/sentence, and text-level predictors included total words (productivity) and process use (indicators of planning or revising). The three top predictors were total words, word accuracy, and percent grammatical sentences. Thus, not only is writing productivity a robust indicator of age and ability but also a strong predictor of quality judgments by trained raters.

Results from other studies linking sentence complexity and quality according to genre effect are of interest. Results with middle school children indicate that words *per clause* related positively to quality ratings for expository essays but not narrative papers; CD, on the contrary, related positively only to narrative ratings (Beers & Nagy, 2009). To explain the discrepancy, the authors observed that clauses in expository writing are prone to be “packed” with information in the form of prepositional phrases, attributive adjectives, and long and complex noun phrases—all structures that would increase internal clause length. They also noted that multi-clause sentences in expository essays were often “formulaic” structures like *I think X because Y*. These are structures that raise the CD value but may not have impressed raters. Findings like these underscore the importance of going beyond global measures like MLTU and CD to uncover underlying structural reasons behind the quantitative numbers. There is no question that in LSA certain genres and topics “set up” writers to use particular structures.

LSA MEASURES ILLUSTRATED IN TWO WRITTEN SAMPLES

Two writing samples are shown in Tables 2 and 3 to illustrate calculation of word-, sentence-, and text-level measures discussed previously. Both 12-year-old students wrote a

Table 2. Expository uncorrected writing sample for Matt (aged 12 years)

Strength in Your Dreams
<p>1. This <i>episode</i> of NOVA is about how sleep works, and what it does for your brain.</p> <p>2. I learned that <i>the silly sounding part of the brain, the Hippocampus</i>, is a place where <i>thoughts are stored, used, and/or strengthened</i>.</p> <p>New paragraph</p> <p>3. It started out when a scientist put some fruit flies into a machine like ferris wheel during an earthquake.</p> <p>4. She made it start to spin and left it spin over night.</p> <p>5. When she came back in the morning, she compared the <i>spun flies</i> with the not spun ones.</p> <p>6. She was comparing fruit flies because they need sleep as well, much like us humans.</p> <p>7. They might need sleep for the same reason as us.</p> <p>New paragraph</p> <p>8. The <i>flies that hadn't been spun</i> were <i>acting normally</i>.</p> <p>9. They were moving around and looking to cause <i>mischief while</i> the other flies were not moving.</p> <p>10. <i>Most</i> seemed to be dead.</p> <p>11. They were alive, just sleeping.</p> <p>12. When <i>they were kept up</i> all night, they needed sleep a lot like we do when we miss a night of rest.</p> <p>New paragraph</p> <p>13. I also learned that if you are learning to do something, if you get a full night of rest, you will be able to come back and do it better the next day.</p> <p>14. <i>This</i> is because during the night your <i>brain reviews</i> all that you learned that day, and increases your ability to do it by 20%.</p> <p>New paragraph</p> <p>15. So fluff up your pillow, and get ready for a good night's sleep.</p> <p><i>Sentence complexity measures:</i> NDW (130) = 96; MLTU = 15.8; CD = 3.1; Grammatical T-units = 100%; TNW = 238</p> <p><i>Observations (transcript line numbers referenced in parentheses):</i></p> <ul style="list-style-type: none"> • Higher level vocabulary: <i>episode</i> (1); <i>compared</i> (5); <i>normally</i> (8); <i>mischief</i> (9) • Metacognitive content: <i>thoughts are stored, used, and/or strengthened</i> (2); <i>the brain reviews</i> (14) • Literate long noun phrases: <i>the silly sounding part of the brain</i> (2); <i>a place where ... strengthened</i> (2) • Literate appositive construction: <i>the Hippocampus</i> (2) • Literate nominalization: verb <i>spin</i> in (4) becomes <i>spun flies</i> in (5) • Passive voice: <i>flies that hadn't been spun</i> (8); • Different types of subordination in one sentence: (14) • Center-embedded relative clause: <i>that hadn't been spun</i> (8) • Adverbial fronting: <i>when they were kept up all night</i> (12) • Ellipsis as a cohesion device: <i>most</i> (10) • Substitution as a cohesion device: <i>this</i> (14)

summary of an age-appropriate 12-min NOVA video on the topic of the importance of sleep for memory function. The two samples illustrate differences based on language ability. The first summary was written by a 12-year-old with TD (Matt, Table 2) and the second by

a same-age peer (John, Table 3), who met SLI criteria for participation in a treatment study that targeted complex sentences (Balthazar & Scott, 2018). The tables show consecutively numbered T-units with preserved (i.e., uncorrected) spelling and punctuation.

Table 3. Expository uncorrected writing sample for John (aged 12 years)

<p>1. Flies sleep all morning and work all night.</p> <p>2. They say sleeping can you <i>remember</i> thing like what you ate.</p> <p>3. They say you can't get a enough sleep.</p> <p>4. That you can work a lot faster after you haved a nap or a good night sleep.</p> <p>5. <i>If you go to bed at the same time you well rested.</i></p> <p>6. You should least get up to 8 hours of sleep a day</p> <p>7. They say if you get up to 8 hours of sleep you will have a lot <i>energy</i> as faster</p> <p>8. You gain <i>stergnth</i> by energy</p> <p>9. You need sleep because it helps</p> <p>10. And when you sleep you get energy when help</p> <p>11. You need sleep because it helps work the brain</p> <p>12. and it gives it power.</p> <p>13. Sleep is the one thing you needed because it helps</p> <p><i>Sentence complexity measures:</i> NDW (130) = 56; MLTU = 10.7; CD = 1.82; grammatical T-units = 46%; TNW = 130</p> <p><i>Observations (transcript line numbers referenced in parentheses):</i></p> <ul style="list-style-type: none"> • Repeated content and wording • Clause density: Subordinate conjunctions restricted to early developing <i>if, when, because</i> • Higher level vocabulary: <i>well-rested</i> (5); <i>energy</i> (8); <i>strength</i> (8) • Metacognitive content: <i>remember</i> (2); <i>helps work the brain</i> (11); <i>and it gives it power</i> (12) • Literate placement of adverbial clause: <i>if you go to bed at the same time</i> (5); <i>if you get up to 8 hours of sleep</i> (7)
--

Note. CD = clause density; MLTU = mean length of T-unit; NDW = number of different words; TNW = total number of words.

The measures were calculated by hand by the author. Clinicians who have used the computer tool SALT (Miller & Iglesias, 2016) for LSA of spoken language could construct a transcript for these written samples and the program would automatically calculate TNW, NDW, and MLTU in words. Systematic Analysis of Language Transcripts also calculates an SI (same as CD) but, based on their knowledge of complex sentence structure, clinicians still need to determine the number of clauses for each T-unit and hand-code each T-unit accordingly. If a handwritten sample is entered into a computer file, most word processors provide a word count that facilitates calculation of MLTU. Clinicians can also explore online word diversity calculators, but be aware that some of these count only lexical words (nouns, verbs, adjectives, adverbs) whereas NDW as calculated in SALT and discussed here is based on all words written. On-

line tools (mentioned later) also are available to analyze word frequency in a text. For example, a program will identify which words in a text fall below the 2,000 most frequent words in terms of word frequency.

Comparisons of word, sentence, and productivity values in Tables 2 and 3 indicate substantial differences. At the word level, NDW was calculated for samples of 130 words. Because John's sample was only 130 words in total, the first 130 words of Matt's sample were used to equate length, as previously discussed. Matt's sample contained 96 different words and 100% correct spelling compared with 56 different words and 96% correct spelling for John. For both samples, I note vocabulary that I consider to be lower frequency, "literate" words including vocabulary and phrases considered to be metacognitive in nature. These are listed as observations under each sample in the tables

(the T-unit numbers shown in parentheses indicate where the words can be found for full context). These observations document Matt's higher level of literate vocabulary skill.

Sentence-level complexity differences, quantified as MLTU and CD, also are substantial. Although John writes sentences that are reasonably long and complex when compared with values in Table 1, it would be a mistake not to look at individual sentences behind those numbers. His multiclausal sentences are developmentally immature for a 12-year-old. Adverbial clauses are restricted to early developing structures with subordinate conjunctions *because*, *if*, and *when*, and there are no relative clauses. In terms of grammatical accuracy, a strong determinant of quality ratings (Troia et al., 2019), Matt's sentences are completely grammatical (100%) with correct punctuation whereas John's sentences contain numerous issues including word omissions, odd word pairings, and morphosyntactic difficulties, resulting in only 46% grammatically correct sentences; punctuation is inconsistent. Tables 2 and 3 list additional observations about literate sentence structures that, while not completely absent for John, are more extensive in Matt's sample.

At the text level, Matt's sample is considerably longer (238 words) than John's 130 words. Matt shows expository text structure acumen by providing a title, beginning with an overall topic generalization statement, dividing content into thematically based paragraphs, and concluding with a "catchy" statement. John's summary reads more like a list of unrelated points, which he repeats several times. For content, he seems to draw on his general background knowledge about sleep generally being a good thing rather than specific content from the video he watched.

The question then becomes how this LSA analysis could assist those working with John on his writing. Perhaps the most significant contribution is the specificity LSA can bring to intervention planning. John can use help at all levels of his writing—at the word level using more diverse and higher level vocabulary, and at the sentence level working on

complexity, particularly use of developmentally higher levels of clause subordination, and correcting grammar issues. At the text level, John needs work on expository text structure, perhaps using graphic organizers to plan in advance of writing, and generating sufficient content. Instructional targets at each level can be singled out for attention in a variety of decontextualized exercises (e.g., sentence-combining exercises to encourage clausal subordination) but with the caveat that clinicians look for ways to transfer skills to real writing tasks, hopefully in the same teaching session (Balthazar & Scott, 2017; Berninger et al., 2009). With these many skills in need of attention, clinicians will need to prioritize and decide what to address first. My own recommendation for John would be the production of grammatically correct sentences and generating enough content to adequately address a writing assignment.

Part of John's intervention might include building recognition of his own writing targets in materials he is currently reading. He could be taught to recognize "interesting" words in a reading passage and discuss their characteristics (e.g., *evaporation*—this is a word for a process in nature, and it ends with the common suffix *-tion*) and then encouraged to use words like this in an appropriate writing task. Overall, he needs explicit instruction that addresses weaknesses identified by LSA. He needs an organization scheme for thinking about his writing at the word, sentence, and text levels and labels for these traits. Language sample analysis gives clinicians and teachers a scheme and relevant terminology to start to build language skills that support better writing.

CONCLUSIONS AND FUTURE DIRECTIONS

One of the goals of this tutorial was to review research on LSA of student writing to see whether these studies provide benchmark data useful to professionals assessing students in their classrooms or clinics. Toward this end, studies were organized by grade and values of commonly used measures

at word, sentence, and text levels were reported and evaluated for their ability to distinguish developmental and language ability differences. Summing across data from several studies provides some guidelines for professionals assessing individual students or groups of students. Attention to these LSA features was supported by research demonstrating their relationship to writing quality. I suggested additional characteristics to look for that could provide more nuanced observation of the types of literate words and sentences used by the writer. The same measures were illustrated in language sample analyses of two students.

Looking to the future, I urge a perspective on LSA within a context of what actually transpires in a student's classroom. To my knowledge, we do not yet know how well LSA results like those shown in Table 1 would predict findings on common types of writing assignments found in the classroom. By the middle elementary grades and beyond, teachers ask students to write summaries of their current reading as well as reports that require integration of information across several sources. These types of assignments seem like very different tasks than writing to a prompt used in most research studies. Future research should explore resulting differences, if any. It also is important to consider results of an LSA against students' classroom experiences with writing. What amount of classroom time is devoted to writing, and of what type? Writing develops very much within a context of supportive instruction, which, multiple surveys show, can vary a great deal from classroom to classroom (as reviewed in Graham, 2019). The written products of any one student are best evaluated in relation to a local standard of classmates, classrooms, and culture.

In the future, connections between writing and reading and oral language should be emphasized as they impact both assessment and intervention. We know that writing problems are rarely isolated and I advocated for attention to reading when discussing possible intervention targets for John. Children with

SLI identified in the early elementary years are at high risk for persistent writing difficulties down the road (Dockrell et al., 2007; Fey et al., 2004). Both reading comprehension and writing rely heavily on metacognitive processes of language construction and integration. Studies have shown that instruction directed at writing can improve reading comprehension and vice versa (Caccamise, 2011; Graham & Hebert, 2011). In a recent article calling for changes in the way writing is taught, Graham (2019) argued for enhanced teacher knowledge of the connections between writing, reading, and language generally as well as the unique skills required for writing—all topics emphasized in this tutorial. Language sample analysis of writing, using a multidimensional framework based on language at the word, sentence, and text levels, fits well with Graham's argument for change.

It may seem daunting to follow the analytic approach advocated here to count and quantify multiple measures at varied levels of language in the analysis of written language samples. There are ways to ease the time burden of LSA by the use of computers. Although the SALT (Miller & Iglesias, 2016) computer program is more commonly used in analyses of spoken language, it is increasingly used for written LSA (see Nelson, 2018, for a tutorial on applications of SALT for writing samples). Troia et al. (2019) used Coh-Metrix for several automatically calculated measures in their study. Coh-Metrix is an online platform for text analysis that includes more than 100 variables related to text difficulty, structure, and cohesion (Graesser, McNamara, & Kulikowich, 2011; www.cohmetrix.com). Researchers of school-age writing predict increasing use of online text analysis systems such as Coh-Metrix based on advances in computational linguistics and natural language processing (G. Troia, personal communication, October 24, 2019). Hopefully, information in this article, along with the development of automation, will result in the increased use of LSA by teachers and clinicians.

REFERENCES

- American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: Author.
- Balthazar, C., & Scott, C. (2017). Complex sentence intervention. In R. J. McCauley, M. E. Fey, & R. B. Gillam (Eds.), *Treatment of language disorders in children* (2 ed., pp. 349-388). Baltimore, MD: Brookes.
- Balthazar, C. E., & Scott, C. M. (2018). Targeting complex sentences in older school children with specific language impairment: Results from an early phase treatment study. *Journal of Speech, Language, and Hearing Research*, *61*, 713-728.
- Beers, S. F., & Nagy, W. E. (2009). Syntactic complexity as a predictor of adolescent writing quality? Which measures? Which genre? *Reading and Writing: An Interdisciplinary Journal*, *22*, 185-200.
- Berninger, V. W., Garcia, N. P., & Abbott, R. D. (2009). Multiple processes that matter in writing instruction and assessment. In G. A. Troia (Ed.), *Instruction and assessment for struggling writers: Evidenced-based practices* (pp. 15-50). New York, NY: Guilford.
- Brimo, D., & Hall-Mills, S. (2019). Adolescents' production of complex syntax in spoken and written expository and persuasive genres. *Clinical Linguistics & Phonetics*, *33*, 237-255.
- Caccamise, D. (2011). Improved reading comprehension by writing. *Perspectives on Language Learning and Education*, *18*, 27-31.
- Coker, D., Ritchey, K. D., Uribe-Zarain, X., & Jennings, A. S. (2018). An analysis of first-grade writing profiles and their relationship to compositional quality. *Journal of Learning Disabilities*, *51*, 336-350.
- Common Core State Standards Initiative. (2010). *Common core state standards for English language arts*. Retrieved from www.corestandards.org
- Culham, R. (2003). *6 + 1 traits of writing: The complete guide grades 3 and up*. New York, NY: Scholastic.
- Dockrell, J. E., Lindsay, G., Connelly, V., & Mackie, C. (2007). Constraints in the production of written text in children with specific language impairments. *Exceptional Children*, *73*, 147-164.
- Fallon, K. A., & Katz, L. A. (2011). Providing written language services in schools: The time is now. *Language, Speech, and Hearing Services in Schools*, *42*, 3-17.
- Fey, M. E., Catts, H. W., Proctor-Williams, K., Tomblin, J. B., & Zhang, X. (2004). Oral and written composition skills of children with language impairment. *Journal of Speech, Language, and Hearing Research*, *47*, 1301-1318.
- Francis, W. N., Kucera, H., & Ackers, A. W. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Chicago, IL: Houghton Mifflin Harcourt.
- Gillam, R., & Johnston, J. (1992). Spoken and written language relationships in language/learning-impaired and normally achieving school-age children. *Journal of Speech and Hearing Research*, *35*, 1303-1315.
- Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix: Providing multilevel analyses of text characteristics. *Educational Researcher*, *40*(5), 223-234.
- Graham, S. (2019). Changing how writing is taught. *Review of Research in Education*, *43*, 277-303.
- Graham, S., Cappizi, A., Harris, K. R., Hebert, M., & Morphy, P. (2014). Teaching writing to middle school students: A national survey. *Reading & Writing: An Interdisciplinary Journal*, *27*, 1015-1042.
- Graham, S., Collins, A. A., & Rigby-Wills, H. (2017). Writing characteristics of students with learning disabilities and typically achieving peers: A meta-analysis. *Exceptional Children*, *83*, 2017.
- Graham, S., & Hebert, M. (2011). Writing to read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review*, *81*, 710-744.
- Green, L., McCutchen, D., Schwiebert, C., Quinlan, T., Eva-Wood, A., & Juelis, J. (2003). Morphological development in children's writing. *Journal of Educational Psychology*, *95*, 752-761.
- Hall-Mills, S., & Apel, K. (2015). Linguistic feature development across grades and genre in elementary writing. *Language, Speech, and Hearing Services in Schools*, *46*, 242-255.
- 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). Hillsdale, NJ: Erlbaum.
- Hunt, K. (1965). *Grammatical structures written at three grade levels* (Research Rep. No. 3). Champaign, IL: National Council of Teachers of English.
- Katusic, S. K., Colligan, R. C., Weaver, A. L., & Barbaresi, W. J. (2009). The forgotten learning disability: Epidemiology of written-language disorder in a population-based birth cohort (1976-1982), Rochester, Minnesota. *Pediatrics*, *123*, 1306-1313.
- Kim, Y. G., & Schatschneider, C. (2017). Expanding the developmental models of writing: A direct and indirect effects model of developmental writing (DIEW). *Journal of Educational Psychology*, *109*, 35-50.
- Koutsoftas, A. D. (2016). Writing process products in intermediate-grade children with and without language-based learning disabilities. *Journal of Speech, Language, and Hearing Research*, *59*, 1471-1483.
- Koutsoftas, A. D., & Gray, S. (2012). Comparison of narrative and expository writing students with and without language-learning disabilities. *Language, Speech, and Hearing Services in Schools*, *43*, 395-409.
- Kroll, B. (1981). Developmental relationships between speaking and writing. In B. Kroll & R. Vann (Eds.), *Exploring speaking-writing relationships: Connections*

- and contrasts (pp. 32-54). Urbana, IL: National Council of Teachers of English.
- Mackie, C., & Dockrell, J. E. (2004). The nature of written language deficits in children with SLI. *Journal of Speech, Language, and Hearing Research*, 47, 1469-1483.
- Miller, J., & Iglesias, A. (2016). *Systematic Analysis of Language Transcripts (SALT): Research Version 16 [Computer software]*. Middleton, WI: SALT Software, LLC.
- National Center for Education Statistics. (2012). *The nations' report card: Writing 2011 (NCES 2012-470)*. Washington, DC: Institute of Educational Sciences, U.S. Department of Education.
- Nelson, N. (2018). How to code written language samples for SALT analysis. *Perspectives of Language Learning and Education*, 3, 45-55.
- Nelson, N. W., & Van Meter, A. M. (2007). Measuring written language ability in narrative samples. *Reading and Writing Quarterly*, 23, 287-309.
- Nippold, M. A. (1998). *Later language development* (2nd ed.). Austin, TX: Pro-Ed.
- Nippold, M. A., & Scott, C. M. (2010). *Expository discourse in children, adolescents, and adults: Development and disorders*. New York, NY: Psychology Press.
- Nippold, M. A., & Sun, L. (2010). Expository writing in children and adolescents: A classroom assessment tool. *Perspectives on Language Learning and Education*, 17, 100-107.
- Nippold, M. A., Ward-Loneragan, J., & Fanning, J. L. (2005). Persuasive writing in children, adolescents, and adults: A study of syntactic, semantic, and pragmatic development. *Language, Speech, and Hearing Services in Schools*, 36, 125-138.
- Pavelko, S. L., Owens, R. E., Ireland, M., & Hahs-Vaughn, D. L. (2016). Use of language sample analysis by school-based SLPs: Results of a nationwide survey. *Language, Speech, and Hearing Services in Schools*, 47, 246-258.
- Perera, K. (1984). *Children's writing and reading: Analyzing classroom language*. London: Blackwell.
- Puranik, C. S., Lombardino, L. J., & Altmann, L. J. P. (2008). Assessing the microstructure of written language using a retelling paradigm. *American Journal of Speech-Language Pathology*, 17, 107-120.
- Ravid, D., & Tolchinsky, L. (2002). Developing linguistic literacy: A comprehensive model. *Journal of Child Language*, 29, 417-447.
- Scott, C. (1988a). Spoken and written syntax. In M. Nippold (Ed.), *Later language development: Ages 9 through 19* (pp. 45-95). San Diego, CA: College Hill Press.
- Scott, C. (1988b). Producing complex sentences. *Topics in Language Disorders*, 8(2), 44-66.
- Scott, C. (1994). A discourse continuum for school-age students: Impact of modality and genre. In G. Wallach & K. Butler (Eds.), *Language learning disabilities in school-age children and adolescents: Some underlying principles and applications* (2nd ed., pp. 219-252). Columbus, OH: Macmillan/Merrill.
- Scott, C. (2003, June). *Literacy as variety: An analysis of clausal connectivity in spoken and written language of children with language learning disabilities*. Paper presented at the annual Symposium on Research in Child Language Disorders, Madison, WI.
- Scott, C. (2009). Language-based assessment of written expression. In G. A. Troia (Ed.), *Instruction and assessment for struggling writers: Evidence-based practices* (pp. 358-385). New York, NY: Guilford.
- Scott, C. (2012). Learning to write. In A. Kamhi & H. Catts (Eds.), *Language and reading disabilities* (3rd ed., pp. 244-268). Boston, MA: Pearson.
- Scott, C., & Balthazar, C. (2010). The grammar of information: Challenges for older students with language impairments. *Topics in Language Disorders*, 30(4), 288-307.
- Scott, C., & Balthazar, C. (2013). The role of complex sentence knowledge in children with reading and writing difficulties. *Perspectives on Language and Literacy*, 39(2), 16-24.
- Scott, C., & Jennings, M. (2004, November). *Expository discourse in children with LLD: Text level analysis*. Poster session presented at the annual meeting of the American Speech Language Hearing Association, Philadelphia, PA.
- Scott, C., & Windsor, J. (2000). General language performance measures in spoken and written narrative and expository discourse in school-age children with language learning disabilities. *Journal of Speech, Language, and Hearing Research*, 43, 324-339.
- Sulzby, E. (1996). Roles of oral and written language as children approach conventional literacy. In C. Pontecorvo, M. Orsolini, B. Burge, & L. B. Resnick (Eds.), *Children's early text construction* (pp. 25-46). Mahwah, NJ: Erlbaum.
- Sun, L., & Nippold, M. A. (2012). Narrative writing in children and adolescents: Examining the literate lexicon. *Language, Speech, and Hearing Services in Schools*, 43, 2-13.
- Troia, G. A. (2009). *Instruction and assessment for struggling writers: Evidence-based practices*. New York, NY: Guilford.
- Troia, G. A., Shen, M., & Brandon, D. L. (2019). Multi-dimensional levels of language writing measures in grades 4 to 6. *Written Communication*, 36(2), 231-266.
- Windsor, J., Scott, C., & Street, C. (2000). Verb and noun morphology in the spoken and written language of children with language learning disabilities. *Journal of Speech, Language, and Hearing Research*, 43, 1322-1336.
- Wood, C. L., Bustamante, K. N., Schatschneider, C., & Hart, S. (2019). Relationship between children's

lexical diversity in written narratives and performance on a standardized reading vocabulary measure. *Assessment for Effective Intervention*, 44, 173-183.

Yoshimasu, K., Barbaresi, W. J., Colligan, R., Killian, J. M., Voigt, R. G., Weaver, A. L., et al. (2011). Written-

language disorder among children with and without ADHD in a population-based birth cohort. *Pediatrics*, 128, 605-612.

Zipf, G. (1932). *Selected studies of the principle of relative frequency in language*. Cambridge, MA: Harvard University Press.