Hammill Institute on Disabilities

Communication Disorders Quarterly
2018, Vol. 40(1) 28–39

DOI: 10.1177/1525740118763063
cdq.sagepub.com
SAGE

English Narrative Language Growth Across the School Year: Young Spanish–English Dual Language Learners

Carla Wood, PhD[1], Mary Claire Wofford, MS[1], Clariebelle Gabas, MS[1], and Yaacov Petscher, PhD[1]

Abstract

This study aimed to describe the narrative retell performance of dual language learners (DLLs) in the fall and spring of the school year and examine predictive relationships. Participants included 74 DLLs in kindergarten and first grade from low socioeconomic backgrounds. Microstructural measures included number of different words (NDW), words per minute (WPM), and verb accuracy. Macrostructural measures included number of total story elements and number of different types of story elements. Path analysis models were used to test the relations among variables. Findings indicated that narrative measures were sensitive to developmental differences across the school year. Fall NDW performance in narrative retells was moderately related to both spring NDW and the total number of macrostructural elements in the spring. Spring WPM was uniquely predicted by fall WPM. Authors concluded that narrative retells are sensitive to developmental differences across a school year for DLLs. Findings support the use of narrative retell measures as a promising tool to examine and describe English language growth of young DLLs within a school year.

Keywords

dual language learner, language development, narratives, progress monitoring, Spanish–English speakers

Given the changing demographic characteristics of schools in the United States, there is an increasing demand for language assessments suitable for progress monitoring of dual language learners (DLLs) to support teachers and related personnel in meeting the educational needs of students from linguistically diverse backgrounds. In essence, DLLs can refer to a widely diverse group of language learners, though they generally include simultaneous bilinguals and second language learners (Paradis, Genesee, & Crago, 2011). As a result, there is a growing demand for less biased assessment options and more knowledge relating to expected rates of language growth for progress monitoring language development of students who are from linguistic minority backgrounds, particularly Spanish–English speaking DLLs. Although language norms are well established for monolinguals (Paradis et al., 2011), assessment and progress monitoring of DLLs' language skills may yield different results given that Spanish–English speaking children in the United States have unique language input and experience (Hoff et al., 2012) resulting in fundamentally different oral language and growth rates than monolinguals. Differences in the opportunity to learn words and develop oral language skills are influenced by differences in socioeconomic, cultural, and linguistic backgrounds, which warrant consideration and further research on language performance of DLLs on assessments (Bandel, Atkins-Burnett, Castro, Wulsin, & Putnam, 2012). More information on typical school year growth is needed for progress monitoring and instructional planning for the provision of appropriate supports responsive to the needs of DLLs.

One approach to assessing language growth during the school year is through language sampling. The use of language sampling is a well-established "best practice" in the assessment of children's expressive language skills and is preferred practice for child language assessment (American Speech-Language-Hearing Association [ASHA], 2004). The use of language sampling through oral narratives is one option for gathering authentic assessment information on language development for children from culturally and linguistically diverse backgrounds (Cheatham & Jimenez-Silva, 2011). Among proposed strengths, it is suggested that narrative retells are less culturally biased than formal standardized

[1]Florida State University, Tallahassee, USA

Corresponding Author:
Carla Wood, Florida State University, 201 W. Bloxham, Tallahassee, FL 32306-1200, USA.
Email: carla.wood@cci.fsu.edu

tests (Miller & Iglesias, 2010). Language samples, such as narrative retells, have been applauded as forms of authentic assessment and may task the integration and execution of multiple language skills simultaneously (Gutiérrez-Clellen, 2002; Justice et al., 2006; Westby, Van Dongen, & Maggart, 1989). The use of narrative retells has been regarded as a useful progress monitoring tool because of its relative ease of administration, tendency toward less bias than standardized tests, and usefulness for differentiating children with and without language learning disorders (Squires et al., 2014). Given the suggested strengths and promising sensitivity of narrative measures, oral narrative retell measures hold promise as a useful tool for educational personnel, particularly if retell measures are sensitive to incremental changes in the language skill development of DLLs over the course of a school year. However, there remains a paucity of research on expected change in narrative retell performance of Spanish–English DLLs, particularly in relation to other types of language measures such as standardized measures of vocabulary and grammar.

Theoretical Motivation

The theoretical motivation to examine the rate of change in oral narrative retells for DLLs is grounded within a constructivist framework. Following this theoretical model, input frequency is thought to be important, along with the underlying language processing mechanisms to support growth in language acquisition (Ambridge & Lieven, 2011; Cameron-Faulkner, Lieven, & Tomasello, 2003). Placing emphasis on the role of environmental input, children's development of rich oral narrative retells is thought to be facilitated by the frequent exposure to storytelling in their environments. Accordingly, over the course of a school year, as children are exposed to examples of stories with initiating events, attempts, and consequences as well as rich vocabulary, children's stories would be expected to contain more lexical diversity and story components at the end of the school year. In addition, as children from linguistic minority backgrounds gain increasing exposure to English over a school year, they would be expected to show improved overall English language skills resulting in gains in verb accuracy, sentence length, lexical diversity, and words per minute (WPM).

Use of Narratives to Measure Language Growth

Given the potential utility of narrative retells as one component of language assessments, we reviewed the available literature on studies that examined the effect of time on macrostructural or microstructural elements of narratives. In language sample analysis, microstructure refers to the lexical, syntactical, and morphological features of a language. Macrostructure, by contrast, describes the overall organization of the discourse and the way concepts and ideas relate to each other. Typical narrative development is characterized by micro- and macrostructural changes as the child's oral language skills mature and become more sophisticated (Brooks & Kempe, 2014).

Effect of Time

Burgeoning studies have begun to examine school year growth in narrative production on measures of lexical diversity (e.g., number of different words or NDW), story components, and utterance length (e.g., mean length of utterance [MLU]). In one study of narrative production growth (Bitetti & Hammer, 2016), findings included that NDW tended to grow an average of 5.74 words every year from the beginning of preschool to the end of first grade, beginning on average at 43.20 words and achieving on average 83.38 words at the end of first grade. Growth was also evidenced in utterance length with MLU-morpheme score increasing 0.49 points each year, averaging 4.32 morphemes at the beginning of first grade and 7.75 morphemes at the end of first grade. Results also indicated that participants grew by 1.02 elements on average on the Narrative Scoring Scheme, an index of macrostructure that characterizes the complexity of content within a narrative (Heilmann, Miller, Nockerts, & Dunaway, 2010). Participants increased on average from 8.83 to 15.97 points.

Other studies have considered the interplay between languages for Spanish–English DLLs specifically. Uccelli and Paéz (2007) examined the narrative skills and vocabulary development of 24 children from low socioeconomic status (SES) backgrounds in kindergarten and first grade. In this study, children produced narratives in English and Spanish in response to a set of pictures. Narrative samples were collected in kindergarten and first grade in addition to expressive vocabulary standardized measures. Findings showed moderate associations between vocabulary and narrative skills within one language only. The authors reported that NDW was more sensitive to developmental change for English narrative productivity than total number of words (TNW). Furthermore, children's Spanish story structure in kindergarten predicted their English narrative quality in first grade (Uccelli & Paéz, 2007).

Interaction of Time and Language Ability

Few studies have included children with below average or low levels of language development (e.g., Rezzonico et al., 2015; Squires et al., 2014). In one such study, Squires and her colleagues examined microstructure and macrostructure scores in Spanish–English speaking bilingual children including a group who were typically developing and a group with language impairment. A battery of standardized measures (e.g., semantics, morphosyntax, narrative skills) was used to obtain descriptive information on children's language abilities in kindergarten and first grade. Overall, bilingual children with language impairment performed significantly lower than their typically developing peers in kindergarten. Results of the study substantiated the use of narrative retells as progress monitoring tools for bilinguals of varying abilities and to differentiate language delays from disorders. Based on examination of narratives of 42 bilingual children (21 with language impairments and 21 typically developing), significant growth in both micro- and macrostructure was observed for children who were typically developing but not for children with language impairments (Squires et al., 2014).

By contrast, a significant effect of time on micro- and macrostructural measures was reported in a longitudinal narrative study that included monolingual and bilingual children with and without specific language impairment (SLI; Rezzonico et al., 2015). The study spanned a 6-month period and included 52- to 58-month-old bilingual children from a variety of linguistic backgrounds (17 different languages including Spanish). Microstructure measures included lexical diversity, sentence length (average of the five longest sentences), verb accuracy, and the use of first mentions. Macrostructure was measured by the amount of information in the retell based on the inclusion of key events and key words. Within-subjects analyses revealed that both typically developing children and children with SLI demonstrated significant growth in micro- and macrostructural measures regardless of their language status (i.e., monolingual or bilingual). In addition, the authors found significant group differences between monolingual and bilingual children in verb accuracy, adding further support for the need for additional study on rates of narrative growth.

Given the mixed results of extant research, it remains unclear whether initial language performance influences the amount of growth on narrative retells. As such, there is a need for additional study on rates of narrative growth in DLLs, and initial English oral language knowledge warrants further consideration as an influencing factor on expected school year growth in English narrative skills.

Influencing Factors

In addition to initial English language knowledge, several factors may influence the expected school year growth in English narrative skills for DLLs. Among potential influencing factors, existing literature on monolinguals' narrative performance suggests that story effects should be considered. In other words, children's performance across multiple time points may be influenced by the story itself. Story effects could include their interest and familiarity with the story, syntactic complexity and imageability of the story (Graesser et al., 2014), the concreteness of the words in the story (McNamara, Graesser, McCarthy, & Cai, 2014), and the cohesiveness of the story influenced by the use of connectives and overlapping words and ideas in neighboring sentences (Graesser et al., 2014; McNamara et al., 2014).

Although burgeoning studies suggest narrative retell measures are sensitive to developmental growth for bilinguals, fewer studies have examined concurrent and predictive relationships between measures. One study (Bitetti & Hammer, 2016) found that book reading in the home contributed significantly to growth in macrostructural elements, but not microstructural elements. Such studies are limited by the heterogeneity in DLL's initial English language skills, which is often not taken into account. In another study, Uccelli and Paéz (2007) reported that English vocabulary performance predicted bilinguals' narrative performance. A similar relationship between vocabulary skills and narrative performance was noted in Terry, Mills, Bingham, Mansour, and Marencin (2013), although this study did not include DLLs.

Research Aims

Although an increasing number of studies have examined narratives, additional studies are needed to broaden our knowledge base of expected development of narrative performance of DLLs. The need to expand the knowledge base is particularly important because DLLs vary greatly in their initial English vocabulary knowledge at elementary school entry (Jackson, Schatschneider, & Leacox, 2014). Given that NDW and macrostructural measures of narrative retell have shown promise for progress monitoring language development of DLLs, additional research is warranted to add to our understanding of typical expected development relative to initial English vocabulary skills. Moreover, longitudinal research is needed to improve our understanding of the relationships between narrative retell measures across the school year. In response, the current study was designed to address the research questions:

Research Question 1: What is the relation between fall and spring English microstructural and macrostructural aspects of narrative retells for young DLLs when controlling for initial vocabulary and form effect of passage?

Research Question 2: What factors contribute to narrative retell outcomes at the end of the school year for DLLs in kindergarten and first grade?

Method

Data for the current project were collected as part of a package of assessment measures administered in a larger development grant funded by the Institute of Education Sciences, U.S. Department of Education. The study procedures were

Table 1. Participants' Performance on Language and Literacy Assessments.

Assessment	Description	*n*	*M*	*SD*
PTONI	Measures nonverbal reasoning abilities	74	97.55	20.26
TVIP	Measures receptive vocabulary in Spanish	50	85.30	19.15
WRMT-III—Phonological awareness	Measures phonological awareness skills such as first and last sound matching, rhyme production, blending	20	93.40	17.85
WRMT-III—Letter identification	Measures ability to identify letters	20	98.35	10.99
WRMT-III—RAN	Measures RAN ability for numbers, letters, colors, and objects	61	93.69	17.85
BESA—Sentence repetition	Measures ability to repeat sentences containing increasingly complex forms in English and Spanish	74	23.07	6.91

Source. Peabody Picture Vocabulary Test–Fourth Edition (Dunn & Dunn, 2007)..
Note. PTONI = *Primary Test of Nonverbal Intelligence*; TVIP = *Test de Vocabulario en Imagenes Peabody* (Dunn, Lugo, Padilla, & Dunn, 1986). WRMT-III = *Woodcock Reading Mastery Tests–Third Edition*; RAN = rapid automatic naming; BESA = Bilingual English–Spanish Assessment (Peña, Gutiérrez-Clellen, Iglesias, Goldstein, & Bedore, 2013).

reviewed and approved by the university's committee on research involving human subjects (HSC No. 2016.18265). The current project used extant data from 1 year of the funded project with the four participating elementary schools that included narrative data in the assessment battery. Due to time demands, not all of the partnering schools included narrative retells in the assessment battery at two time points. The larger project did not explicitly teach narratives or include retell activities between test points.

Participants

The sample for this study included children from four schools who had narrative retell samples in both fall and spring time points, collected in September and May of the school year. Of the 74 participating children, 36 were girls and 38 were boys. All attended kindergarten (n = 29) or first grade (n = 45). Eligibility requirements included parents report that Spanish was spoken in the home. Exclusionary criteria included sensory impairments, identified disability, or an inaudible audio sample on either assessment time point due to noise in adjacent spaces of the school setting. A nonverbal intelligence test (*Primary Test of Nonverbal Intelligence* [PTONI]) was administered to ensure participants were within 1.5 standard deviations of the mean. The investigators administered assessments of language and emerging literacy performance in the fall to further describe participants (see Table 1). Performance on standardized assessments was included to describe the participants' language and literacy skills. Scores on language and literacy assessments were not used for inclusionary or exclusionary decisions. No one was excluded based on his or her performance.

Bilingual research assistants conducted phone interviews with the children's parents to gather information on the families' cultural and linguistic backgrounds. Parents were asked to report on children's language use at home. This was done by asking parents to indicate whether the child spoke some English, English and Spanish equally, or primarily English or Spanish (see Table 2). The largest percentage of families (71%) reported speaking primarily Spanish-only in the home. The majority of the parents reported their highest education to be a high school diploma or less (68% of mothers and 80% of fathers). The most commonly reported countries of origin of the participating parents included Mexico and Guatemala. Based on school and family reports, approximately 93% received free lunch and the remaining 3% qualified for reduced lunch. Refer to Table 2 for more detailed family demographic information of participants.

Materials

Animated narratives. The participants viewed one of four randomly assigned short animated movies on a computer. Each movie was approximately 2½ min in length and exhibited a simple story line following a narrative scheme that included characters (up to three), an initiating event, attempt and outcome. The animated narratives included voice-over story recordings. The stories were similar to each other in terms of length and semantic and syntactic complexity (Diehm, Wood, Messier, & Callender, submitted). To ensure relatively adequate opportunities for microstructure models, each narrative included at least three instances of each microstructural element of the Narrative Assessment Protocol (NAP; Pence, Justice, & Gosse, 2007) such as regular and irregular past tense verbs, prepositional phrases, and complex sentences.

Procedures

The investigators and bilingual research assistants administered narrative retell measures in September and May of the

Table 2. Demographic Characteristics of Participants and Families.

Demographic Characteristic	*n*	% of respondents
Characteristics of participants		
Primary language of child	59	
Spanish	27	37
Spanish, some English	4	6
Both languages equally	19	32
English, some Spanish	3	5
Primarily English	5	8.5
Free/reduced lunch	56	
Free	52	93
Reduced	4	7
Characteristics of families		
Country of origin	58	
Mexico	25	43
El Salvador	4	7
Honduras	2	3
Guatemala	14	24
The United States	4	7
Puerto Rico	5	8
Other	5	8
Primary household language	58	
Spanish	41	71
English	8	14
Both languages equally	9	12
Mother's education	52	
Beyond high school	2	4
High school/GED	16	31
Some high school	9	17
Less than high school	20	37
Mother's occupation	54	
Professional	0	0
Skilled labor	3	6
Unskilled labor	14	26
Unemployed	37	69
Father's education	52	
Beyond high school	2	4
High school/GED	16	31
Some high school	9	17
Less than high school	25	48.5
Father's occupation	49	
Professional	0	0
Skilled labor	14	29
Unskilled labor	32	65
Unemployed	3	6

Note. GED = General Education Diploma.

academic school year. Children viewed the same animated story with the same recorded narrative at both time points. Research assistants in a speech-language pathology program listened to the audio files and transcribed the samples following traditional procedures in accordance with conventions established for *Systematic Analysis of Language Transcripts* (SALT; Miller & Iglesias, 2010). In the case that an audio file was inaudible due to background noise (e.g., competing speakers in close proximity, reduced signal to noise ratio), it was not transcribed for inclusion in the final sample. Children were not discouraged from using Spanish; however, only one participant code-switched between English and Spanish in his narrative retell, likely because the narrative was presented in English and the school setting involved English-only instruction. For the child who produced Spanish in the narrative retell, six of 14 utterances contained Spanish and 20 out of 78 words were in Spanish. Regarding content, the child using Spanish asked clarification questions in Spanish, and the examiner prompted the child to try to use English. Utterances delivered in Spanish did not affect the macrostructural score. Because 99% of the retells were produced in English only, narratives retells in English became the focus of the current study. Interrater agreement was determined by dividing the instances of agreement at the word level by the total number of opportunities; percentage agreement between research assistants was calculated to be 83.31% on a randomly selected 20% of transcripts.

Microstructure coding. Microstructural measures included NDW, number of total words (NTW), verb accuracy, and WPM. NDW and NTW were derived by generating standard measures reports in SALT. From the transcribed retells, research assistants calculated verb accuracy microstructure variables using methods employed in two previous studies to allow for comparison (Hipfner-Boucher et al., 2014; Rezzonico et al., 2015). For verb accuracy, the number of correct verbs was divided by the number of total verbs in the narrative. WPM is a measure intended to represent the ease or difficulty with which a child formulates and produces an oral language task (Rojas & Iglesias, 2013). Children who have difficulty assembling and/or expressing their ideas would be expected to have low WPM. It has been used in Spanish–English language sample analysis as a measure of oral fluency (Heilmann, Miller, Nockerts, & Dunaway, 2010; Miller, Andriacchi, & Nockerts, 2015; Price, Hendricks, & Cook, 2010), and has reported to positively correlate with age (Miller & Heilmann, 2004). WPM was calculated using SALT by dividing the NTW by the time elapsed from the examiner's first prompt to the child's last utterance. Some words were not included in total word count, such as reformulated and repeated words; these were coded in parentheses applying SALT conventions.

Macrostructure coding. Macrostructural measures included an index of the quantity of elements which we will refer to as number of total story elements regardless of the type of element (e.g., plot, setting, characters). In addition, we included an index to quantify the different types of story elements such as plot, characters, and reactions, which we will refer to as the number of different types of macrostructural codes. The narrative transcriptions were analyzed and coded for eight storytelling components adapted from previous research (Schachter & Craig, 2013; Spencer & Peterson, 2012): character, setting, plot, initiating event/problem, reactions/emotions, attempt, consequences, and ending. The rubric and set of specific procedures were developed to classify cases of unclear or inadequate narrative statements. One of the authors trained a team of three undergraduate research assistants on the coding procedures, and provided them with opportunities for practice and thorough feedback. All narrative samples were divided into three sets. Each set was independently coded by two coders. Agreement ranged from 76% to 97% depending on macrostructural aspect. When discrepancies occurred, a third independent coder reviewed and rescored the transcript and discussed the contrasts with the coders in an attempt to refine the rubric to clarify difficult differentiations (e.g., consequence vs. ending).

Analyses

Path analysis models were used to test the relations between the identified independent and dependent variables for each research question. A key role for the path analysis models is the ability to evaluate the extent to which a hypothesized model provides good fit to the data, change portions of a model, and re-evaluate structural relations among key variables in the model. Two key considerations drove the initial specification of the models for each research question. First, it was important to account for students' baseline vocabulary, as measured by the PPVT, because DLLs vary greatly in their initial English vocabulary knowledge upon elementary school entry (Jackson et al., 2014), and previous studies suggest narrative performance is correlated with and predicted by vocabulary skills (Terry et al., 2013). Second, students were exposed to multiple stories that varied in number of initiating events and attempts. To adjust for the possibility of form effects influencing score results (e.g., Petscher & Kim, 2011), story sequence was included as a categorical covariate. Model fit for the path analysis was evaluated according to the comparative fit and Tucker–Lewis indexes (CFI and TLI), as well as the root mean square error of approximation (RMSEA). CFI and TLI values of at least .90 and RMSEA values less than .10 provide evidences of model with acceptable fit. Chi-square difference analyses were used for nested models to test for statistical improvement in the incremental fit due to model re-specification. All path analysis models were estimated using Mplus 7.1 software (Muthén & Muthén, 2004) using 1,000 bootstrapped samples due to the small sample size.

Results

Descriptive Statistics

To describe narrative retell performance of DLLs, we first report descriptive statistics at the beginning and end of the school year. Means and standard deviations for narrative measures of the full sample are provided in Table 3. On average, children included 6.8 more NDW at the end of the school year when compared with narrative retells in the fall (M_{Fall} = 30.93, SD_{Fall} = 17.9; M_{Spring} = 37.73, SD_{Spring} = 16.4). WPM increased by 14.5 on average (M_{Fall} = 59.07, SD_{Fall} = 32.1; M_{Spring} = 73.56, SD_{Spring} = 28.1). Participants showed an increase in the total number of verbs and verb accuracy across the school year. Total verbs used on average increased from 18.46 (SD = 12.7) in fall to 19.16 in spring (SD = 9.3). Children demonstrated a difference of 13.6% accuracy between the fall and spring averages (i.e., 74% accuracy on verbs on average in the fall and 88% accuracy in the spring). For macrostructural aspects, the total number of components increased by 2.6 at the end of the school year (M_{Fall} = 7.20, SD_{Fall} = 4.7; M_{Spring} = 9.77, SD_{Spring} = 4.9), and children included approximately one additional type of component in the spring when compared with their fall retell (M_{Fall} = 4.32, SD_{Fall} = 2.1; M_{Spring} = 5.19, SD_{Spring} = 1.6). Average frequency of story components inclusion is displayed by individual components in Figure 1. Although grade-level differences were not a specific research question in the current project, average inclusion of story components by grade is depicted in Figure 2. Increases in children's use of attempts and consequences were noted when comparing fall and spring macrostructure performance.

Path Analysis Models

The initial specification of the path analysis model for Research Question 1 included spring NDW and the macrostructure total score as outcomes with the covariates (i.e., fall PPVT and passage form) and the fall auto-regressor predictor in the model. This initial specification of the path analysis provided poor fit to the model, $\chi^2(2) = 16.41$, CFI = .90, TLI = .43, RMSEA = .312 (95% confidence interval [CI] = [.184, .460], p close < .001). Inspection of the results suggested that the model could be improved by including fall NDW as a predictor of spring macrostructure total. The revised model was significantly improved ($\Delta\chi^2 = 16.28$, $\Delta df = 1$, $p < .001$) with acceptable fit to the data, $\chi^2(1) = 0.13$, CFI = 1.00, TLI = 1.00, RMSEA = .000 (95% CI = [.000, .221], p close = .741). Standardized coefficients from the revised model (included in Figure 3) show that fall NDW performance was moderately related to both spring NDW (.64, $p < .001$) and spring macrostructure total (.53, $p < .001$) outcomes. No other covariates were statistically significant in the model; however, the inclusion of all predictors resulted in 45% and 47% of the variance in spring NDW and macrostructure total explained, respectively.

The initial model for Research Question 2 included NDW and macrostructure total scores in the spring as outcomes as well as the number of different macro codes, WPM, and TNW. Similar to the initial model for the first

Table 3. Descriptive Statistics for Overall Sample.

	Fall narrative retell				Spring narrative retell			
Measures	*M*	*SD*	Minimum	Maximum	*M*	*SD*	Minimum	Maximum
PPVT Raw	80.58	22.8	16.0	122.0	94.86	20.6	50.0	143.0
PPVT SS	85.31	12.2	46	111.0	88.77	11.7	65	121.0
NDW	30.93	17.9	0.0	95.0	37.73	16.4	7.0	79.0
WPM	59.07	32.1	4.3	125.8	73.56	28.1	18.9	154.3
Total macro	7.20	4.7	0.0	17.0	9.77	4.9	2.0	23.0
Macro codes	4.32	2.1	0.0	8.0	5.19	1.6	1.0	8.0
Total verbs	18.46	12.7	0	52	19.16	9.3	4	51
Verb accuracy	74.07	25.1	0	100	87.63	14.8	31.3	100

Note. n = 74 for all measures, except posttest PPVT n = 73. PPVT reflects raw and standard scores (SS) on the *Peabody Picture Vocabulary Test* (Dunn & Dunn, 2007); NDW reflects number of different words; WPM reflects words per minute; Total macro reflects the total number of macrostructure elements coded; Macro codes reflects the number of different types of macrostructural components present out of eight total elements (character, setting, plot, initiating event, emotions/reactions, attempt, consequence, ending); Verb accuracy refers to the overall percentage accuracy for verb usage in the narrative retell.

research question, fall PPVT and story effects were included as covariates, as well as the fall auto-regressor for each outcome. This model resulted in acceptable fit, $\chi^2(20) = 32.79$, CFI = .97, TLI = .92, RMSEA = .093 (95% CI = [.024, .148], p close = .115); however, model fit indices suggested that improvements could be made by adding fall NDW as a predictor of both spring macrostructure totals and macro codes final model, $\Delta\chi^2 = 12.89$, $\Delta df = 2$, $p = .002$). Results for the revised model, $\chi^2(18) = 19.90$, CFI = .99, TLI = .99, RMSEA = .038 (95% CI = [.000, .113], p close = .540) are displayed in Figure 4. For NDW, the results pointed to fall NDW as the only significant predictor (.51, $p < .001$) in the presence of the other covariates in the model, with 41% of the variance explained. Macrostructure total showed that both fall NDW and PPVT were unique predictors of spring performance (.38 and .22, respectively, $p < .05$; 41% of the variance explained). Spring TNW was best explained by prior TNW (.44, $p < .001$; 29% of the variance explained); spring WPM was uniquely explained only by fall WPM (.57, $p < .001$; 40% of the variance explained); and spring macro codes was explained uniquely by fall passages (.19, $p < .01$), spring passages (.27, $p < .001$), and fall NDW (.34, $p < .001$) with 52% of the total variance explained.

Discussion

Key Findings

The purpose of this study was to describe the narrative retell performance of DLLs at the beginning and end of the school year and examine predictive relationships between English microstructural and macrostructural aspects of oral narrative retells for DLLs. Across the school year, children's average narrative performance increased in lexical diversity, rate or WPM, and accuracy. They demonstrated 6.8 new words in NDW, rate increases of 14.5 WPM on average, and average gains of 13.6% accuracy in verb use. Finally, children also produced more story components as demonstrated by 2.6 new macrostructural components, and included at least one different type of macrostructural component on average. Fall NDW performance in narrative retells was moderately related to both spring NDW and spring macrostructure total. In other words, children's English lexical diversity in the fall predicted their spring narrative retell performance in terms of lexical diversity and the total number of story elements included. The total number of macrostructural elements in the spring was predicted by both fall NDW and PPVT performance, suggesting initial vocabulary was an important predictor of narrative retell performance at the end of the school year.

Taken together, the relationships provide insights on the constructs of oral narrative retell measures. Findings substantiate that vocabulary is an important variable in assessing and monitoring development of oral narrative retells. Lexical diversity (NDW) significantly predicted end of school year performance in measures of narrative retells with the exception of WPM. Spring WPM was uniquely predicted by fall WPM which may suggest that growth in WPM across the school year does not rely on vocabulary development, but rather reflects a different language component. In future studies, it would be interesting to consider other measures of language performance in Spanish and English to further explore factors that predict the developmental trajectory of oral narrative retells.

Comparison With Literature

Lexical diversity. The current results for NDW are most similar to those reported by Bitetti and Hammer (2016). The current finding that NDW grew by 6.6 in the school year

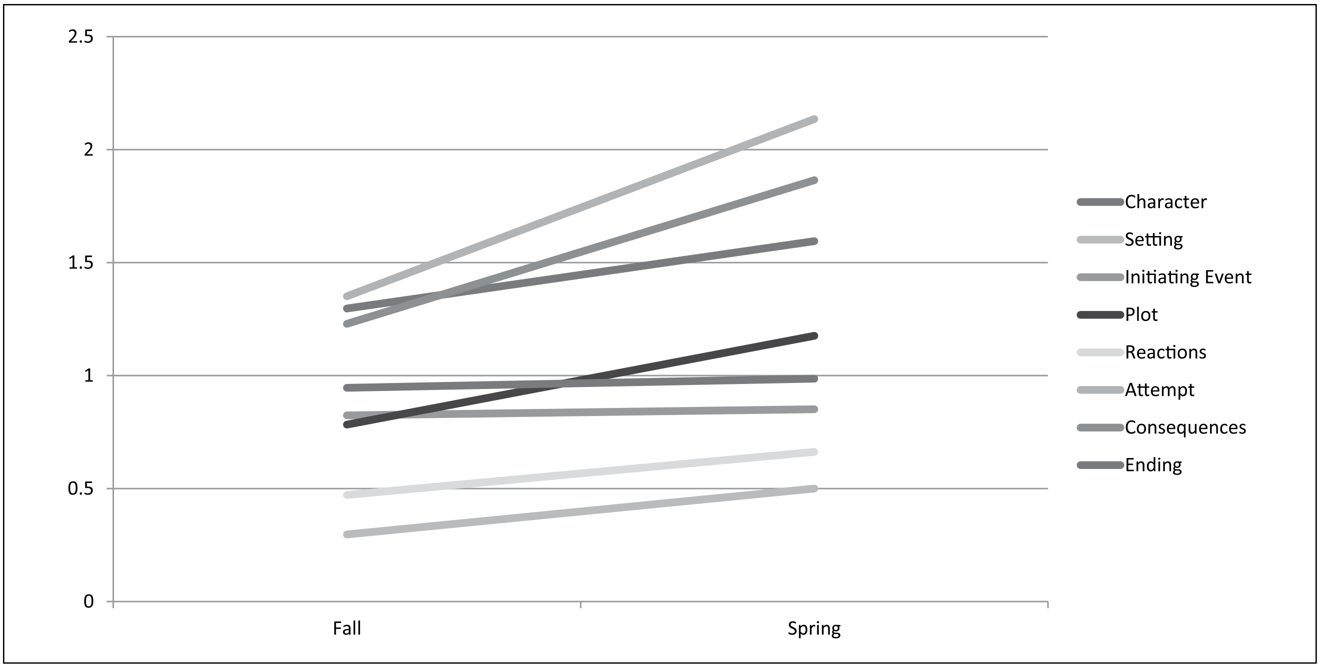

Figure 1. Children's inclusion of macrostructural components in oral narrative retells at two time points across academic year.

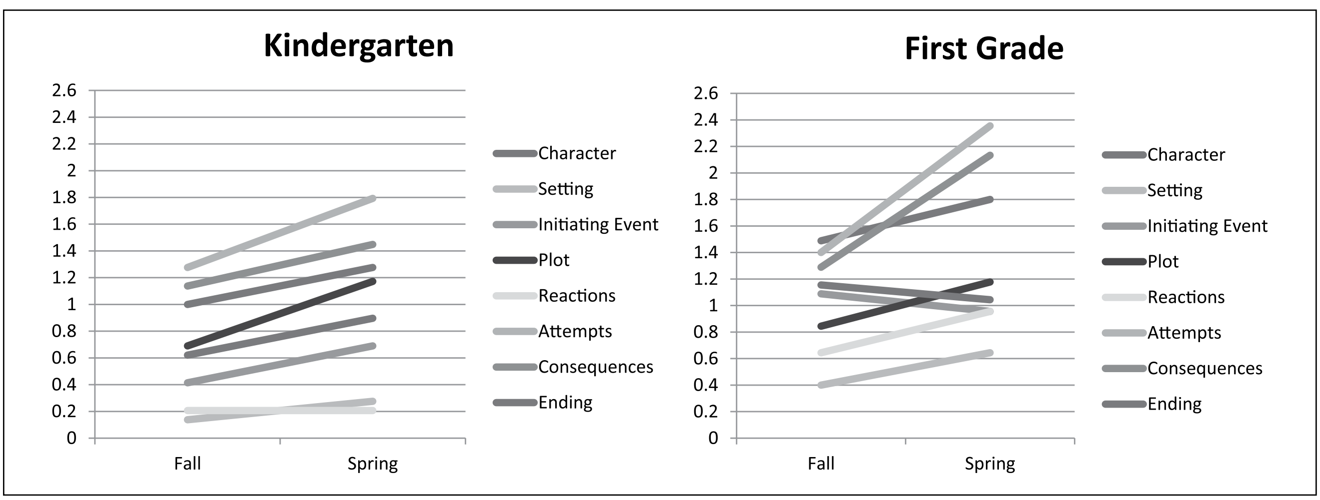

Figure 2. Kindergarten and first-grade children's inclusion of macrostructural components in oral narrative retells at two time points across academic year.

was similar to the average of 5.74 NDW growth reported in the Bitetti and Hammer study. This may be explained, in part, by the similarities in the participant characteristics of the studies. Comparing the two studies, the participants in the Bitetti and Hammer study showed similar (below average) initial Spanish and English receptive vocabulary skills based on the PPVT and *Test de Vocabulario en Imagenes Peabody* (TVIP) measures. In contrast, Terry et al. (2013) did not find significant growth in NDW from fall to spring; however, the participants were monolingual dialect users with average initial receptive vocabulary scores on the PPVT compared with expectations for their age.

Macrostructural components. Developmental change in children's macrostructural components is difficult to compare across existing studies due to the differences in measures used between studies. In general, most studies have shown developmental change in macrostructural measures across the school year, which is substantiated by the current findings. The current findings extend our understanding of the

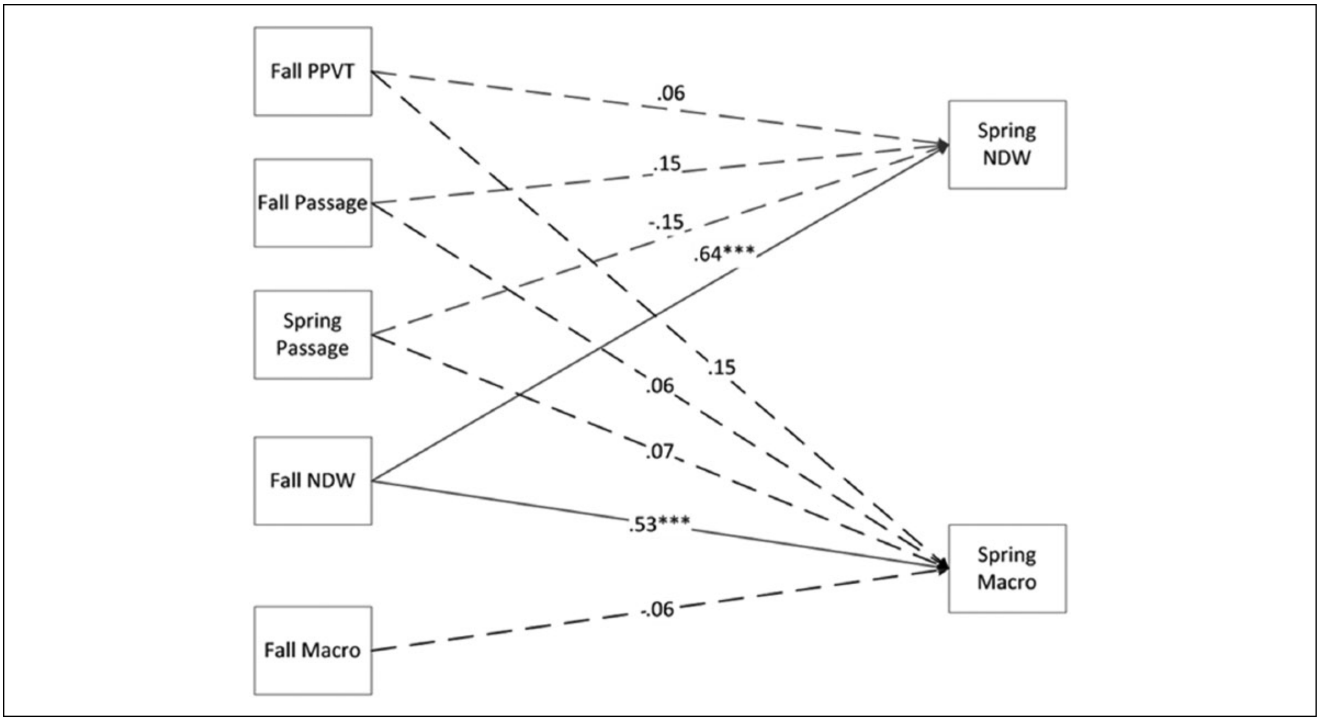

Figure 3. Standardized path coefficients for Research Question 1.
Note. PPVT = *Peabody Picture Vocabulary Test–Fourth Edition*; NDW = number of different words; passage = assigned story (1 of 4); macro = total macrostructure.

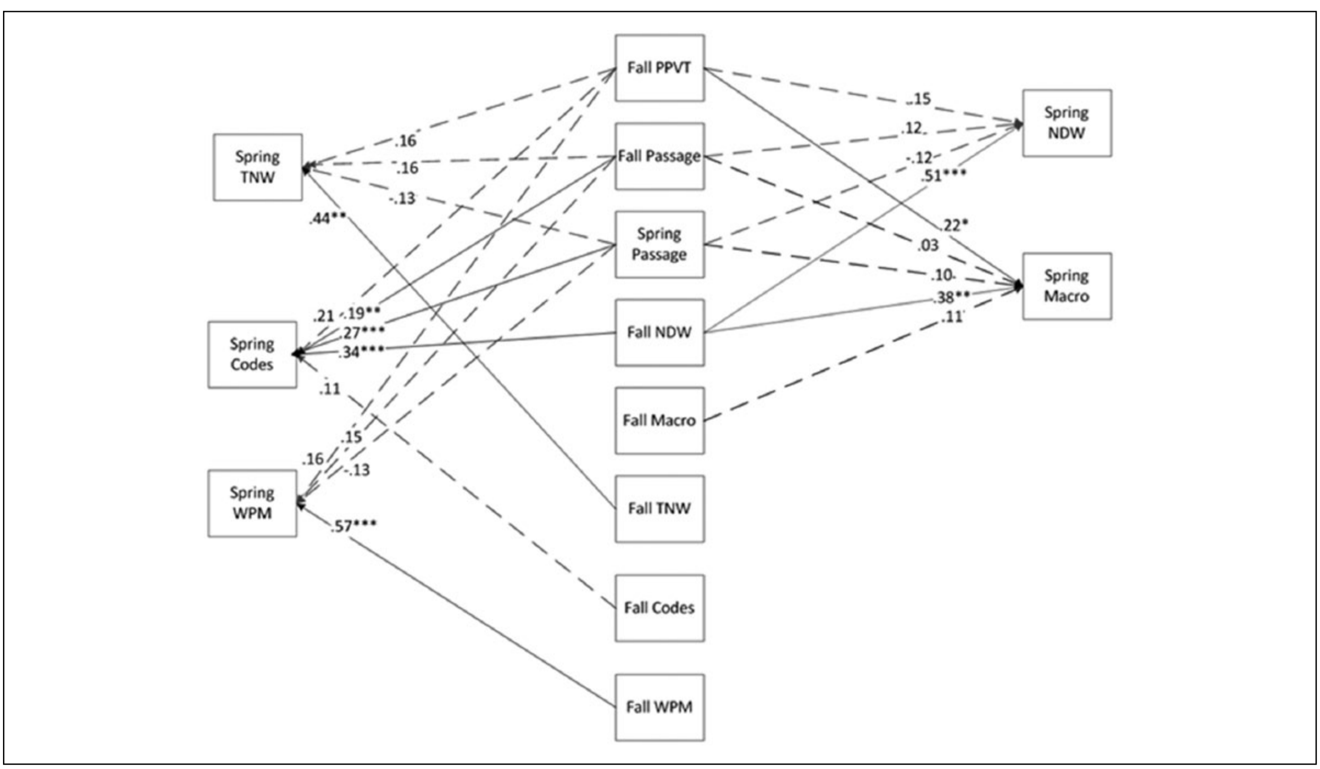

Figure 4. Standardized coefficients for Research Question 2.
Note. TNW = total number of words; PPVT = *Peabody Picture Vocabulary Test–Fourth Edition*; NDW = number of different words; codes = number of different macro codes; passage = assigned story (1 of 4); macro = total macrostructure; WPM = words per minute.

relationship between initial vocabulary skills (fall NDW) and spring macrostructural performance. The ability to tell a more complete story was associated with having good receptive vocabulary skills and higher lexical diversity. With regard to specific components of storytelling, it was interesting that children's spring macrostructural components in retelling the same stories included more attempts and consequences than in the fall. One potential explanation is that consequences have lower imageability and are more abstract or conceptually complex than story components such as characters. This finding is consistent with developmental expectations described by McCabe and Peterson (1991) suggesting that children generally provide all narrative structures, notably, including a description of the resolution, by 6 or 7 years of age.

Limitations

The use of only one sample at each time point was a noted limitation in addition to the transcription reliability rate of 83%. Given a single narrative retell task, it cannot be assumed that similar results would be observed using other types of prompts or narrative tasks. Similarly, it should be noted that narrative retells were conducted in English only, which limits information about the interplay between languages for DLLs in the study. It would be best practice to sample in each language (Peña & Halle, 2011); however, in this study, only one collected sample at each time point was available. In a future study, it would be interesting to examine differences in other types of samples (e.g., personal narratives, picture description, and wordless books) and their use as progress monitoring tools for young DLLs.

The current study involved a convenience sample with regard to the proportion of English and Spanish spoken at home and the socioeconomic backgrounds of families whose children participated. As such, it cannot be assumed that the current findings would generalize to DLLs from backgrounds unlike those in the current study. It should be noted that some families reported that both Spanish and English were spoken approximately equally at home. The lack of precision in quantifying the percentage of use of each language at home is a weakness of the current study, because we recognize that parents' use of language at home is expected to impact rate of growth in each language.

Implications

Despite limitations of the study, the findings substantiate the sensitivity of narrative retell measures to developmental changes in children's performance across the school year. The developmental changes support the usefulness of narrative retells for progress monitoring the language development of DLLs, particularly. The current findings also highlight specific components that appear to be malleable across the school year. The identification of components that are expected to show growth across the school year may be particularly useful for teachers and related personnel in progress monitoring and program planning. Furthermore, the important role of initial vocabulary skills on narrative performance adds to our understanding of children's outcomes and predicted performance across the school year.

Increasingly critical for educators is greater recognition of the typical effects of second language acquisition and diverse home language resources on academic performance. In light of this, the current study contributed to the extant literature base by targeting DLLs from low SES backgrounds, while the majority of studies have historically focused on middle-class English-monolingual speakers. Use of novel methodological elements, including stimuli (e.g., animated movies), analyses (e.g., path modeling), and narrative measures (e.g., diversity of macrostructural codes) allows researchers to broaden the knowledge base of oral narrative retell performance with the aim of improving language assessment for DLLs.

Authors' Note

The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Acknowledgment

The authors are especially grateful to partnering schools, participating families, and support from research assistants who were essential to project completion.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A130460 to Florida State University.

ORCID iDs

Carla Wood https://orcid.org/0000-0002-0551-0638
Mary Claire Wofford https://orcid.org/0000-0001-6243-7214

References

Ambridge, B., & Lieven, E. V. (2011). *Child language acquisition: Contrasting theoretical approaches*. New York, NY: Cambridge University Press. doi:10.1017/CBO9780511975073

American Speech-Language-Hearing Association. (2004). *Preferred practice patterns for the profession of speech-language pathology* (Preferred practice patterns). doi:10.1044/policy.pp2004-00191

Bandel, E., Atkins-Burnett, S., Castro, D. C., Wulsin, C. S., & Putnam, M. (2012). *Examining the use of language and literacy assessments with young dual language learners*. Chapel Hill: Center for Early Care and Education Research—Dual Language Learners, Frank Porter Graham Child Development Institute, The University of North Carolina. Retrieved from http://fpg.unc.edu/sites/fpg.unc.edu/files/resources/reports-and-policy-briefs/FPG_CECER-DLL_AssessingDLLs.pdf

Bitetti, D., & Hammer, C. S. (2016). The home literacy environment and the English narrative development of Spanish–English bilingual children. *Journal of Speech, Language, and Hearing Research*, *59*, 1159–1171. doi:10.1044/2016_jslhr-l-15-0064

Brooks, P. J., & Kempe, V. (2014). *Narrative development: Encyclopedia of language development*. Thousand Oaks, CA: SAGE.

Cameron-Faulkner, T., Lieven, E., & Tomasello, M. (2003). A construction based analysis of child directed speech. *Cognitive Science*, *27*, 843–873. doi:10.1016/j.cogsci.2003.06.001

Cheatham, G. A., & Jimenez-Silva, M. (2011). What makes a good story? Supporting oral narratives of young children from culturally and linguistically diverse backgrounds. *Childhood Education*, *84*, 4261–4268. doi:10.1080/00094056.2011.10523188

Diehm, E., Wood, C., Messier, J., & Callender, M. (In-revision). Technology enhanced narrative assessment: Effect of presentation modality on preschoolers' narrative retell.

Dunn, L., Lugo, S., Padilla, R., & Dunn, L. (1986). Test de Vocabulario en Imagenes Peabody. [Peabody Picture Vocabulary Test]. Circle Pines, MN: American Guidance Service.

Dunn, L. M., & Dunn, L. M. (2007). Peabody Picture Vocabulary Test- Fourth. Circle Pines, MN: American Guidance Service.

Graesser, A. C., McNamara, D. S., Cai, Z., Conley, M., Li, H., & Pennebaker, J. (2014). Coh-Metrix measures text characteristics at multiple levels of language and discourse. *The Elementary School Journal*, *115*, 210–229. doi:10.1086/678293

Gutiérrez-Clellen, V. F. (2002). Narratives in two languages: Assessing performance of bilingual children. *Linguistics and Education*, *13*, 175–197.

Heilmann, J., Miller, J. F., Nockerts, A., & Dunaway, C. (2010). Properties of the narrative scoring scheme using narrative retells in young school-age children. *American Journal of Speech-Language Pathology*, *19*, 154–166.

Hipfner-Boucher, K., Milburn, T., Weitzman, E., Greenberg, J., Pelletier, J., & Girolametto, L. (2014). Relationships between preschools' oral language and phonological awareness. *First Language*, *34*, 178–197. doi:10.1177/0142723714525945

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, *39*, 1–27. doi:10.1017/S0305000910000759

Jackson, W. C., Schatschneider, C., & Leacox, L. (2014). Longitudinal analysis of receptive vocabulary growth in young Spanish-English speaking children from migrant families. *Language, Speech, and Hearing Services in Schools*, *45*, 40–51. doi:10.1044/2013_LSHSS-12-0104

Justice, L. M., Bowles, R. P., Kaderavek, J. N., Ukrainetz, T. A., Eisenberg, S. L., & Gillam, R. B. (2006). The index of narrative microstructure: A clinical tool for analyzing school-age children's narrative performances. *American Journal of Speech-Language Pathology*, *15*, 177–191.

McCabe, A., & Peterson, C. (Eds.). (1991). *Developing narrative structure*. Hillsdale, NJ: Psychology Press. doi:10.1037/032742

McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge, MA: Cambridge University Press. doi:10.1017/cbo9780511894664.006

Miller, J. F., Andriacchi, K., & Nockerts, A. (2015). *Assessing language production using SALT software: A clinician's guide to language sample analysis*. Middleton, WI: Systematic Analysis of Language Transcripts Software.

Miller, J. F., & Heilmann, J. (2004, October). *Bilingual language project update*. Paper presented at the Department of Communication Disorders Colloquium, Madison, WI.

Miller, J. F., & Iglesias, A. (2010). Systematic Analysis of Language Transcripts (SALT), English & Spanish (Version XX) [Computer software]. Madison: Language Analysis Laboratory, Waisman Center, University of Wisconsin–Madison.

Muthén, L. K., & Muthén, B. O. (2004). *Mplus user's guide* (7th ed.). Los Angeles, CA: Author.

Paradis, J., Genesee, F., & Crago, M. B. (2011). *Dual language development & disorders*. Baltimore, MD: Paul H. Brookes.

Peña, E. D., Gutiérrez-Clellen, V. F., Iglesias, A., Goldstein, B. A., & Bedore, L. M. (2013). *Bilingual English–Spanish Assessment (BESA)*. San Diego, CA: AR-Clinical Publications.

Peña, E. D., & Halle, T. G. (2011). Assessing preschool dual language learners: Traveling a multiforked road. *Child Development Perspectives*, *5*, 28–32. doi:10.1111/j.1750-8606.2010.00143.x

Pence, K., Justice, L. M., & Gosse, C. (2007). *Narrative assessment protocol*. Columbus: Preschool Language & Literacy Lab, The Ohio State University. doi:10.1037/t53458-000

Petscher, Y. & Kim, Y. S. (2011). The utility and accuracy of oral reading fluency score types in predicting reading comprehension. *Journal of School Psychology, 49*, 107-129. doi:10.1016/j.jsp.2010.09.004

Price, L. H., Hendricks, S., & Cook, C. (2010). Incorporating computerized language sample analysis into clinical practice. *Language, Speech, and Hearing Services in Schools, 41*, 206-222.

Rezzonico, S., Chen, X., Cleave, P. L., Greenberg, J., Hipfner-Boucher, K., Johnson, C. J., . . . Girolametto, L. (2015). Oral narratives in monolingual and bilingual preschoolers with SLI. *International Journal of Language & Communication Disorders*, *50*, 830–841. doi:10.1111/1460-6984.12179

Rojas, R., & Iglesias, A. (2013). The language growth of Spanish-speaking English language learners. *Child Development*, *84*, 630–646. doi:10.1111/j.1467-8624.2012.01871.x

Schachter, R. E., & Craig, H. K. (2013). Students' production of narrative and AAE features during an emergent literacy task. *Language, Speech, and Hearing Services in Schools*, *44*, 227–238.

Spencer & Peterson. (2012). *Test of Narrative Retell–Preschool (TNR-P)*. Available from http://www.languagedynamics-group.com/

Squires, M. K. E., Lugo-Neris, M. J., Peña, E. D., Bedore, L. M., Bohman, T. M., & Gillam, R. B. (2014). Story retelling by bilingual children with language impairments and typically developing controls. *International Journal of Language & Communication Disorders*, *49*, 60–74. doi:10.1111/1460-6984.12044

Terry, N. P., Mills, M. T., Bingham, G. E., Mansour, S., & Marencin, N. (2013). Oral narrative performance of African American prekindergartners who speak nonmainstream American English. *Language, Speech, and Hearing Services in Schools*, *44*, 291–205.

Uccelli, P., & Paéz, M. M. (2007). Narrative and vocabulary development of bilingual children from kindergarten to first grade: Developmental changes and associations among English and Spanish skills. *Language, Speech, and Hearing Services in Schools*, *38*, 225–236.

Westby, C. E., Van Dongen, R., & Maggart, Z. (1989). Assessing narrative competence. *Seminars in Speech and Language*, *10*, 63–76.