

Assessing Young Children's Oral Language: Recommendations for Classroom Practice and Policy

Alesia Malec

University of Toronto

Shelley Stagg Peterson

University of Toronto

Heba Elshereif

University of Toronto

Abstract

A systematic review of research on oral language assessments for four-to-eight-year-old children was undertaken to support a six-year action research project aimed toward co-creating classroom oral language assessment tools with teachers in northern rural and Indigenous Canadian communities. Through an extensive screening process, 10 studies were assessed as highly rated and identified for inclusion in the final review. Narrative, vocabulary, and syntax assessments were the most common assessment types found in the final review. Assessment practices in all studies in the final review

involved gathering language samples in one-on-one adult-directed contexts. The systematic review also revealed that a preponderance of the research on young children's oral language assessment has been published in speech-language pathology and language testing journals. Although educational researchers recognize the importance of oral language to children's literacy and learning, there is a paucity of research on oral language assessment conducted by educational researchers and published in educational research journals. Implications to policy and classroom practice include recommendations for increased research collaboration between speech-language pathology researchers and literacy researchers along with input from early childhood educators to develop oral language assessment instruments that support children's oral language in classroom settings.

Keywords: oral language, systematic review, early childhood education, young children, classroom assessment

L'évaluation de la langue orale des jeunes enfants : des recommandations pour les pratiques pédagogiques et pour les politiques

Résumé

Une revue systématique des évaluations de la langue orale chez les enfants de quatre à huit ans a été entreprise en vue d'étayer un projet de recherche-action visant à co-crée des outils d'évaluation du langage parlé en classe avec des enseignants dans des communautés autochtones et dans des communautés rurales dans le nord du Canada. À la suite d'un processus de dépistage exhaustif, dix études ont reçu une excellente cote et été retenues pour la synthèse finale. Les types d'évaluations les plus courantes, telles que répertoriées dans la synthèse finale, portaient sur les compétences narratives, le vocabulaire et la syntaxe. Les pratiques d'évaluation dans toutes les études retenues dans la synthèse finale reposaient sur la collecte d'échantillons de langage dans des échanges interpersonnels dirigés par un adulte. Il est ressorti de la revue systématique que les recherches

sur l'évaluation de la langue orale chez les jeunes enfants sont surtout publiées dans des revues axées sur l'orthophonie et l'évaluation des connaissances linguistiques. Bien que les chercheurs en éducation reconnaissent l'importance du langage parlé pour la littératie et l'apprentissage durant l'enfance, il y a peu de recherches sur l'évaluation de la langue orale menées par des chercheurs en éducation et publiées dans des revues centrées sur la recherche en éducation. Sur le plan des politiques et des pratiques pédagogiques, il y aurait donc lieu que se multiplient des recherches concertées entre les chercheurs en orthophonie et les chercheurs en littératie et qu'on fasse appel aux éducateurs de la petite enfance pour mettre au point des outils d'évaluation du langage parlé afin de favoriser le développement de la langue orale chez les enfants en classe.

Mots-clés : langue orale, revue systématique, éducation de la petite enfance, jeunes enfants, évaluation en classe

Acknowledgements

We wish to thank Dr. Laureen McIntyre for her guidance in setting up the parameters for the systematic review. We also gratefully acknowledge two anonymous reviewers for their generous and invaluable feedback. Finally, we thank the Social Sciences and Humanities Research Council for funding this research through a Partnership Grant.

Introduction

In this article we draw on the results of a systematic review of research evaluating young children's oral language assessment to identify principles and practices for classrooms and make recommendations for policy development. We begin with the research-based and curriculum-supported assumption that oral language is foundational to literacy and to all learning. Children's oral vocabulary, storytelling performance, and phonological awareness at school entry are predictors of later reading comprehension and decoding (Dickinson & Porche, 2010; Resnick & Snow, 2009). In addition to using oral language to express ideas and intentions in communication with others, children use language to organize and reflect on experience (Alexander, 2006; Barnes, 1992; Mercer & Littleton, 2007; Wells, 1999; Vygotsky, 1978). Putting words together to communicate with others helps children to clarify and deepen understanding. Oral language is foundational to social, as well as conceptual, learning. Children are socialized into the cultural practices of their communities through their everyday interactions with others (Cekaite, Blum-Kulka, Grøver, & Teubal, 2014). Children's choice of words, sounds and structures, and the ways in which they use these features of language to achieve their own purposes, are some of the indicators of their social learning.

The importance of oral language to literacy and all learning is described in many language arts curricula in Canada (e.g., Alberta Education, 2000; Manitoba Education, 2011; Ontario Ministry of Education, 2006; Saskatchewan Ministry of Education, 2010) and around the world (Department for Education, 2013; Ministère de l'éducation nationale, de l'enseignement supérieur et de la recherche, 2016). In Ontario, for example, the *Language* curriculum document (Ontario Ministry of Education, 2006) contains an assertion that "oral communication skills are fundamental to the development of literacy and essential for thinking and learning" (p. 9). As such, supporting children's oral language, in part through carrying out ongoing, observational formative assessments, should be an integral part of classroom practice (Fisher & Frey, 2007). Student learning is enhanced when teachers gather and analyze information about students' learning throughout the school year, using it to guide their teaching (Bennett, 2011; Black & Wiliam, 2009; Hon-drich, Hertel, Adl-Amini, & Kleime, 2015).

Given the importance of oral language to literacy and all learning in school and beyond, there is a need for educators and educational researchers to devote time and

resources to understanding effective ways of assessing and supporting children's oral language in classrooms. This is the premise underlying a six-year action research project aimed toward co-creating culturally relevant (Ladson-Billings, 1995) classroom oral language assessment and instructional tools and approaches with teachers of four-to eight-year-old children in northern rural and Indigenous Canadian communities. Participants are Indigenous and non-Indigenous kindergarten and Grade 1 teachers and their students in rural northern communities in four provinces. In an ongoing cycle of data collection, analysis, and implementation, teachers and researchers video-record children's play interactions. Based on collaborative analysis of the recordings, we plan together pedagogical and assessment approaches that teachers then implement to support children's language learning and development.

To inform the action research, we conducted a systematic review of research on oral language assessment published between 1980 and 2015. We identified research-supported characteristics of oral language assessment practices for developing assessment tools and practices appropriate for primary classrooms whose makeup differs from a European-heritage urban/suburban middle-class demographic. The question guiding the systematic review was: What practices and tools are effective for assessing young children's oral language? We focused on the following aspects of the research on oral language assessment because they have greatest relevance to classroom practice, as identified by teachers participating in our previous interview research (Peterson, McIntyre, & Forsyth, 2016):

1. features of children's oral language that are assessed in the research studies;
2. practices for gathering and analyzing oral language of young children; and
3. considerations for classroom assessment of young children's oral language.

Following a description of our systematic review process, we discuss the oral language assessment features, practices, and considerations that are found in the identified research studies in terms of their potential to inform classroom practice and policy decisions. We draw on the literature in the literacy education field and our ongoing action research to support our recommendations.

Methods

In 2015, a literacy researcher/teacher educator and a language researcher/speech language pathologist developed the systematic review protocol to guide a team of five graduate student researchers through the subsequent search and screening process. All phases of the review were guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group, 2009), which detail the essential elements of a review through a comprehensive checklist and flow diagram (see prisma-statement.org for further information). The two lead researchers agreed on a definition of oral language as consisting of three major components: language form (e.g., phonology, morphology, syntax), semantics, and pragmatics/communicative competence (Owens, 2012). Children demonstrate communicative competence when they use language appropriately in various social contexts (Hymes, 1966). Oral language is viewed as both “a code whereby ideas about the world are represented through a conventional system of arbitrary signals for communication” (Bloom & Lahey, 1978) and as an expression of “culture and [ways] to express intentions in congruence with the culture” (Bruner, 1986, p. 65). Thus, in addition to learning vocabulary, speech sounds, and ways of putting words and phrases together, children learn possible roles, relationships, perspectives, meanings, and modes of meaning available within their social worlds as they learn language (Genishi & Dyson, 2009; Vygotsky, 1978).

With this sociolinguistic view of language (Bruner, 1986; Vygotsky, 1978) as the conceptual framework for our systematic review, we selected three online bibliographic databases to conduct the search: ERIC, ProQuest Education, and PsycINFO. In the initial phase, two graduate student researchers searched the targeted databases using the core search terms: *oral language assessment* and *oral communication assessment*. During the second electronic search phase, five graduate student researchers searched the same electronic databases using a core search term and a sub-term. Examples of sub-terms include language knowledge, expressive language, language structure, lexical development, emerging language, and functional language, among others. For example, a student would search for *oral language assessment* together with *expressive language*. A third non-electronic phase consisted of physical searches of five peer-reviewed journals related to early childhood education and oral language development, targeting journals that the initial search phase had shown were likely to contain oral language assessment-focused

research. One graduate student researcher conducted physical searches in the *American Journal of Speech Language Pathology*; *Language, Speech and Hearing Services in Schools*; *Journal of Research in Childhood Education*; *Communication Disorders Quarterly*; and *Applied Psycholinguistics* following the completion of the electronic searches.

Screening: Applying Inclusion/Exclusion Criteria

Guided by the goal of informing classroom teachers in contexts such as the northern Canadian rural and Indigenous classrooms involved in our action research, our configurative systematic review (Oliver, Dickson, & Newman, 2012) included research that evaluated the effectiveness of particular assessment tools and approaches across a range of cultural contexts. Primary research, published in English from 1980 to 2015, and conducted with participants aged four to eight years of age, was targeted for inclusion (Gough & Thomas, 2012) in the review. The decision was made to begin the search in 1980 to cover a sufficient span of time to capture the types of oral language assessments that may be used with young children, while acknowledging that searching further into the past was unlikely to be productive to our review. Participant age ranges were permitted to overlap with our targeted age range (e.g., a study with participants from ages two years to six years would be included because the participant ages crossed over with the target age range for the review). A list of key word descriptors related to oral language assessment was also applied as an inclusion criterion (e.g., language use, emergent language, developing language, speech development, expressive language, receptive language, etc.).

The review protocol and PRISMA work-flow diagram (Moher et al., 2009) were used to screen over 30,000 articles found in the initial search of the two core search terms: *oral language assessment* and *oral communication assessment*. The language researcher/speech pathologist trained two graduate student researchers on the systematic review process and these graduate student researchers trained an additional three graduate student researchers. The team of five graduate student researchers exported search lists from the three databases, which included article titles, abstracts, publication dates, and keywords, using each core search term (e.g., *oral language assessment*) plus sub-term (e.g., *expressive language*), into Excel files. Following removal of duplicate entries, titles and abstracts were reviewed in Excel and inclusion criteria were applied (primary research, participant age, publication date, publication language, key words). Where

insufficient information was available in the title and abstract (e.g., research participants' ages), the full article was consulted for the relevant information. In this screening process, we identified approximately 1,200 articles for potential inclusion in the review. The number of articles meeting the existing inclusion criteria called for a honing of the criteria which would better represent the intent of the systematic review's guiding question—to gain a more comprehensive understanding of effective children's oral language assessment tools and practices.

In a team meeting attended by the five graduate student researchers and the literacy researcher/teacher educator, two additional inclusion criteria were generated. Graduate student researchers had noted that much of the identified research in the initial search phase targeted non-typically developing children (e.g., hearing-impaired children, children with special identified needs), bilingual children, or children learning an additional language. They also identified a group of studies that used oral language assessments to assess the effectiveness of interventions or were designed to determine potential relationships between results of oral language and literacy (primarily reading) assessment results. These research studies did not directly address our research question, which focused specifically on practices and tools that would inform classroom assessment of young children's oral language. As a result, during the next iteration of the search process (Brunton, Stansfield, & Thomas, 2012), we excluded them in the next round of screening. We reasoned that studies primarily interested in the relationship between reading and oral language development, while important to literacy development, fell outside the scope of our research question in that these studies had a primary focus on reading and not oral language.

The second phase of screening, conducted by the five graduate student researchers, resulted in 201 non-excluded studies. Two graduate student researchers subsequently read the resulting 201 studies in their entirety to verify the presence of the inclusion criteria (primary research, participant age, publication date, publication language, key words) and relevancy to the research question. These two graduate student researchers then met to discuss their independent assessments of the 201 articles using the inclusion and relevance criteria. They agreed upon 33 studies as meeting all inclusion criteria and being relevant to the research question.

Final Screening Process

The final set of 33 research studies was evaluated by applying a research quality and research relevance assessment (Gough, 2007). Research quality referred to evaluations of validity (internal, external, and construct) and research relevance related to fidelity to the review's research question, including whether language form, content, and use had been addressed within the study. The literacy researcher/teacher educator also reviewed the final set of 33 research studies for quality and research relevance following the graduate students' assessments, which provided an additional layer of validation to the graduate student researchers' assessments. All three researchers agreed on the final assessments for the last set of 33 studies. Ten of the 33 studies resulted in a high overall rating, 18 studies achieved a medium rating, and five studies received a low overall rating. A study receiving a high rating represented high relevancy to the research question, high levels of internal, external, and construct validity, and addressed at least two of the three language elements targeted (form, content or meaning, and use). Only highly rated studies were included in the final synthesis. Table 1 provides detailed information about the language features assessed, participant information, and the journals in which each of the 10 identified studies were published.

Table 1. High-rated studies used in discussion

Citation	Journal	Participant Age	Methodology	Target Demographic	# of Participants	Language Features Assessed
Bradfield, Besner, Wackerle-Hollman, Albano, Rodriguez, & McConnell (2014)	Assessment for Effective Intervention	Pilot – 36 to 71 months Main – 48 to 71 months	6 language assessments measured expressive and receptive vocabulary; correlations calculated across all assessments	Preschool children	Pilot – 55 Main – 756	Vocabulary
Craig & Washington (2002)	American Journal of Speech-Language Pathology	47 to 78 months	Spontaneous language samples scored for linguistic features (e.g., utterance length, complex syntax); scores analyzed using ANOVA and regression modeling	African American children speaking African American English; Low and Middle income SES	100	Vocabulary
Craig, Washington, & Thompson (2005)	American Journal of Speech-Language Pathology	Grade 1–5	Picture description task and comprehension tasks analyzed for linguistic features; results analyzed using MANOVA	African American children speaking African American English; Low and Middle income SES	295	Vocabulary
Heilmann, Miller, & Nockerts (2010)	Language Testing	60 to 84 months	Narrative retell task scored with Narrative Scoring Scheme (NSS); scores analyzed through descriptive statistics, including skewness, kurtosis and examiner judgement	Typically developing children	129	Narrative features
Hipfner-Boucher, Milburn, Weitzman, Greenberg, Pelletier, & Girolametto (2014)	First Language	46 to 71 months	Assessment battery of 10 tests included 2 narrative retell tasks and 1 story generation task; scores analyzed through correlation, linear regression modeling	English speaking children	89	Narrative features
Patton Terry, Mills, Bingham, Mansour, & Marencin (2013)	Language, Speech, and Hearing Services in Schools	55 to 71 months	Narrative elicitations scored using NSS and analyzed for syntax structures; results correlated across tasks	African American children speaking African American English	146	Syntax Narrative performance

Citation	Journal	Participant Age	Methodology	Target Demographic	# of Participants	Language Features Assessed
Qi, Kaiser, Milan, & Hancock (2006)	Language, Speech & Hearing Services in Schools	36 to 54 months	4 tests administered to establish convergent validity; scores analyzed through correlation and regression analysis	African American and European American children in same community; Low income SES	482	Vocabulary
Qi & Marley (2011)	Topics in Early Childhood Special Education	34 to 62 months	Battery of tests to validate Preschool Language Scale for target population; results analyzed through MANOVA, regression analysis, item response theory	Hispanic American and European American children	440	Vocabulary
Smith, Myers-Jennings, & Coleman (2000)	Communication Disorders Quarterly	36 to 71 months	4 tests administered to target population; scores analyzed using ANOVA, descriptive statistics	Rural children African American and Caucasian middle income SES	160	Communicative competence
Schneider & Vis Dubé (2005)	American Journal of Speech-Language Pathology	62 to 82 months	3 story retell tasks scored using Systematic Analysis of Language Transcripts; results analyzed through ANOVA, <i>t</i> tests	Typically developing English speaking children	44	Narrative features; Syntax

The final 10 highly rated studies represented a wide range of assessment types and target participants. To mine the studies for their insights into oral language assessment, the same two graduate student researchers applied an inductive coding process (Glaser, 1978) to the studies to determine patterns in features of the children's oral language that were assessed, the assessment practices employed, and considerations for classroom assessment of young children's oral language.

Before discussing the results of our systematic review, we present limitations to our search process that must be considered when applying our findings to classroom practice and policy development. In the extensive process of having five research assistants screening the large number of studies initially identified (over 30,000), it is possible that studies qualifying for full review may have been missed. The length of time taken to conduct the review means that studies published since the review began have not been included and we acknowledge

that there may be more recent research that would be relevant to our research question and meet our inclusion criteria. Additionally, we acknowledge that the systematic review was undertaken to inform our own action research project with kindergarten and Grade 1 teachers and children in northern Canadian rural and Indigenous communities. Our research goals and sociocultural theoretical lens have influenced our choice of databases and inclusion criteria, as well as our evaluation of the 201 studies that were identified in the second screening. Because children in our action research study speak and represent diverse English dialects, we focused on the literature within these domains. The wealth of research on bilingual children's language learning and development indicates that an additional review focusing on this body of the literature is warranted. We must also acknowledge that the primary or sole language of all the researchers involved in the review is English, thus limiting the inclusion of articles to those published in English while understanding that we live and conduct research in a country (Canada) with two official languages and where a multitude of languages are in use.

One observation of note is that the oral language assessment research identified in our search has not been published to a great degree in early childhood education and literacy journals, nor in classroom-based assessment journals. We discuss this finding later in our implications for policy section, as well as a discussion of implications for classroom practice that arise from our systematic review of the literature. First we present the results of the review.

Results

As shown in Table 1, narrative comprehension, syntax, and vocabulary were the most commonly assessed oral language features in the final set of research studies identified in our systematic review process. Communicative competence was assessed as one of a number of language features in one of the identified studies. These features demonstrate direct correspondence with our initial three-pronged definition of oral language: language form (e.g., vocabulary and syntax), semantics (e.g., narrative comprehension), and pragmatics or communicative competence. We use the assessed features of oral language—narrative, syntax, vocabulary, and communicative competence—as organizational

categories for reporting the results of the review, followed by a brief discussion of oral language assessment practices.

Narrative Assessments

Researchers in four of the 10 research studies identified in our systematic review (Heilmann, Miller, & Nockerts, 2010; Hipfner-Boucher et al., 2014; Patton Terry et al., 2013; Schneider & Vis Dubé, 2005) gathered language samples by inviting children to narrate a story while looking at wordless picture books. These studies' practices and assumptions align with social constructivist views of language and learning to some degree. The use of narrative, for example, fits with Bruner's (1990) argument that narrative structures provide frames that help children both construct meaning and communicate their understanding of that meaning. Children's oral narrative retellings of stories from wordless picture books, comprised of sentences tied together by themes and narrative structures, provide a more comprehensive picture of their language and thinking, than do standardized tests that evaluate limited sets of skills that have been stripped of their communicative context (Heilmann, Miller, & Nockerts, 2010). Researchers examining children's oral narratives recognize that a number of factors will influence children's retellings of stories, including whether and how many times they have previously heard the story (Martinez & Roser, 1985), their familiarity with the activities that characters undertake in story events (Hudson, Shapiro, McCabe, & Peterson, 1991), the complexity of the story (Heilmann, Miller, & Nockerts, 2010), and the adult input (directions and expectations) when eliciting the narrative retelling (DeTemple, Wu, & Snow, 1991).

Researchers conducting two of the narrative-focused studies (Hipfner-Boucher et al., 2014; Schneider & Vis Dubé, 2005) examined episodic structure, number of characters identified, initiating event, internal response, and the setting, attempt, and consequence of story events in children's narratives to examine individual structures within the narratives (Schneider & Vis Dubé, 2005) or potential relationships between narratives and phonological awareness (Hipfner-Boucher et al., 2014). Schneider and Vis Dubé (2005) found that children in Grade 2 produced significantly more story grammar elements than children in kindergarten when stories were presented orally (either with or without pictures) by a researcher for retelling by the child in the assessment. The picture-only presentation did not result in a significant difference in the number of story of elements

produced by the children based on age/grade level. Hipfner-Boucher and colleagues (2014) combined narrative assessment with a battery of standardized tests and found strong correlations between narrative story structure and phonological awareness. Their results support emphasizing the importance of assessing children on narrative abilities as well as vocabulary skills and phonological awareness.

Heilmann, Miller, and Nockerts (2010) assessed the appropriateness of four scoring schemes for analyzing 129 five- to seven-year-old children's oral narratives. They used the Narrative Scoring Scheme (NSS) (developed by the Language Analysis Lab at the University of Wisconsin-Madison) to examine the macrostructure of children's narratives (e.g., use of story grammar—providing information about the introduction, character development and mental states and intentions, conflict/resolution, and conclusions, in addition to sequencing events accurately and providing transitions between events). The other three assessments were (1) a procedure focusing on plot and theme (Reilly, Losh, Bellugi, & Wulfeck, 2004), which assesses setting, instantiation (providing concrete examples), episodes, and resolution; (2) an ordinal adaptation of Applebee's Narrative Maturity Scale (Manhardt & Rescorla, 2002), which assesses narratives holistically, placing them on a 5-point scale ranging from a "Heap" of events with no central theme, to a "True narrative" with well-developed events and central theme or moral; and (3) an ordinal adaptation of Stein's story levels (Pearce, McCormack, & James, 2003), in which an 11-point continuum ranges from an isolated description of random characters and actions to interactive episodes taking multiple perspectives of multiple characters and showing how their goals mutually influence each other (Heilmann, Miller, & Nockerts, 2010, pp. 621–622). The NSS achieved a wider, but normal, distribution of scores (compared to the other measures) and no child achieved "mastery" (p. 612) (a score over 90%) on the measure. Results from each of the other three measures found over one-third of participants scoring over 90%, indicating a ceiling effect for the other measures. The NSS includes examiner judgement, and thus does not strictly rely on numerical outcomes on the discrete elements of assessment and "tap[s] into the perceptual elements of the narrative process" (p. 615).

Patton Terry and colleagues' (2013) narrative assessment research examined relationships between participating children's use of their spoken African American English dialect, narrative macrostructure, and story comprehension. They developed performance descriptions of pre-kindergarten African American children's narrative performance from

the fall to the spring of one school year. They found that their participants performed within age-appropriate expectations on three measures used in the study (Narrative Assessment Protocol, as noted in Pence, Justice, & Gosse, 2007; Narrative Scoring Scheme, as noted in Heilmann, Miller, & Nockerts, 2010; Heilmann, Miller, Nockerts, & Dunaway, 2010; High Point Analysis, as noted in McCabe, Bliss, Barra, & Bennett, 2008) and, that as children's narratives became more complex throughout the school year, the scores on two of the assessments significantly improved. These analysis schemes provide an abundance of narrative features and ways to assess young children's narrative comprehension for classroom use.

Viewed through a sociolinguistic lens, the oral narrative assessments in our identified studies overlook a critical component of storytelling, however, as they do not take into account cultural differences in narrative styles. As Gutierrez-Clellen and Quinn (1993) explain, "Storytelling is a social event governed by cultural norms and values. These extralinguistic rules dictate appropriate narrative behavior" (p. 4). In classroom practice, it is important to consider the ways in which narratives are structured according to cultural expectations within the children's communities when developing schemes for assessing children's oral narratives (Melzi, Schick, & Kennedy, 2011; Michaels, 1981; Ochs & Capps, 2001; Peltier, 2010; Piquemal, 2003). The notion of narratives having a story grammar consisting of a setting, problem or initiating event, attempts to solve the problem, and eventual problem resolution is not a universal view of how narratives are structured. Assuming that children's narrative retellings will conform to the dominant narrative structures creates a biased assessment context favouring children who have grown up hearing and reading narratives with these structures.

Syntax Assessments

Researchers who have published their work in literacy journals have found that children's oral narratives are useful sources of information about children's story comprehension (e.g., Morrow, 1990; van Kraayenoord & Paris, 1996). The studies identified in our systematic review, published predominantly in speech-language pathology and language testing journals, expand the assessment focus when children's oral narratives are the data sources. They show how the narratives can provide information about children's use of syntax, a language feature that correlates with children's overall language growth (Craig

& Washington, 1994), as well as their later reading achievement (Craig, Connor, & Washington, 2003).

In one of the identified studies (Patton Terry et al., 2013), children's narrative samples were used to examine language productivity (e.g., number of words and number of utterances), language complexity (e.g., use of cohesive devices, such as conjunctions, clause density, and number of complex structures in each utterance), and accuracy (e.g., number of grammatically correct utterances and adequacy of cohesive devices used in the retelling). Children's use of syntactic structures in their oral language was an assessed feature of oral language in four of the 10 identified studies in the final review (Craig & Washington, 2002; Craig, Washington, & Thompson, 2005; Patton Terry et al., 2013; Smith, Myers-Jennings, & Coleman, 2000).

The results of three of the 10 identified studies underscore the need for teachers to consider local English dialects when assessing the syntax of children's oral language (Craig & Washington, 2002; Craig et al., 2005; Patton Terry et al., 2013). The impetus for these studies was research showing that "many typically developing African American students fall below published normative means on standardized tests that are based primarily on mainstream students" (Craig et al., 2005, p. 121), as well as research showing that children from higher socio-economic communities have performed better than children from lower socio-economic communities (Craig & Washington, 2002). Qi and Marley (2011) explain in their study that the norming procedures, selection of test items, and the interaction patterns used to gather language samples contribute to findings of cultural bias in standardized tests. Many oral language assessments have been normed with children using Mainstream American English (MAE). The test items require that children have experience with typical European American heritage middle-income families' lifestyles, and are familiar with interaction patterns that require labelling or pointing to objects. As such, the performance on standardized oral language assessments of children who speak English dialects other than MAE may reflect language and cultural experiences that differ from children used to norm the assessments, rather than true language delays.

When the syntax of African American English was recognized in the analysis of children's free-play utterances and narrative retellings, Craig and Washington (2002) found that African American children performed within age-appropriate expectations. These findings contrast with results of standardized tests that have been normed using

Mainstream American English showing consistent patterns of underachievement of African American children (Qi et al., 2006; Qi & Marley, 2011; Smith et al., 2000). It is apparent that designing linguistically and culturally appropriate assessments should involve consultation with community members about communicative practices and dialect within the community. Research studies in our identified set of 10 (e.g., Smith et al., 2000) provide a model for considering dialectal differences when scoring children's responses to prompts. Smith and colleagues (2000) found that two tests (*Test of Language Development-2 Primary* [TOLD-2P] and *Test of Pragmatic Skills*) showed little to no difference between their participants' results and the normed group, indicating these tests may be culturally valid for a wide range of children. However, the *Test of Auditory Comprehension of Language—Revised* (TACL-R) scores were significantly lower for the participant children compared to the normed group, which may indicate a potential for bias. Scores on the *Patterned Elicitation Syntax Test* (PEST) were also significantly lower for the participant group compared to the normed group. However, once scores were adjusted for dialectal differences, the number of participants “classified as language disordered was reduced by more than two thirds” (Smith et al., 2000, p. 105).

Vocabulary Assessments

Given the predominance of vocabulary assessments in the oral language assessment research identified in our systematic review (five of the final 10), together with the abundance of research showing correlations between vocabulary and reading comprehension (Dickinson & Tabors, 2001) and between vocabulary and children's abilities to comprehend what others say and to communicate with others (Marchman & Fernald, 2008), it is clear that vocabulary should be a component of classroom oral language assessment practices.

The identified studies that used the *Peabody Picture Vocabulary Test* (PPVT-3 or 4) (Dunn & Dunn, 2007) and the *Preschool Language Scale* (PLS-3 or 4) (Zimmerman, Steiner, & Pond, 2002) included those that attempted to establish a criterion or validation measure for correlational purposes when testing and validating a newly designed vocabulary assessment (*Picture Naming* task in Bradfield et al., 2014) or other forms of assessments (Craig et al., 2005; Patton Terry et al., 2013), or to examine the validity of standardized tests across varied demographic populations (Qi et al., 2006; Qi & Marley,

2011). Other vocabulary assessments that involve pointing to pictures representing target vocabulary, such as *Test of Auditory Comprehension of Language-Revised* (TACL-R; Carrow-Woolfolk, 1985) and the *Test of Language Development-2 Primary* (TOLD-2; Shulman, 1986) were used by Smith and colleagues (2000) to determine the effect of linguistic variation in English on rural preschool children's language. The *Test of Preschool Early Literacy* (TOPEL; Lonigan, Wagner, Torgesen, & Rashotte, 2007) was used alongside the PPVT-4 in Patton Terry and colleagues' (2013) study of correlations between narrative performance, oral expressive and receptive vocabulary, spoken dialect use, and story comprehension. The actual tests that were used and the administration processes carried out in these vocabulary-oriented studies are not as readily applicable to classroom settings as the narrative-oriented studies. For example, in the identified studies, an adult, usually a researcher or speech language pathologist, assessed individual children's receptive language by asking them to point to pictures of objects and actions that the adult named. Children's expressive language was assessed when they named objects and actions in pictures. However, given that vocabulary is best learned when children hear and use words in meaningful contexts (Neuman & Dwyer, 2009) and when children are interactive and responsive in social situations, rather than passive listeners of others' language (Harris, Michnick Golinkoff, & Hirsh-Pasek, 2011), it follows that classroom assessments should also take place in meaningful interactive contexts. Inviting children to name or point to pictures of objects and actions removes their use of vocabulary from actual communicative experiences where there are social purposes for the language.

In two of the studies, researchers targeting vocabulary assessment chose alternative instruments or methods to administer their assessments. Bradfield and colleagues (2014) went to great lengths to create lists of words familiar to children of the target age when developing oral vocabulary assessments by consulting three different word-list sources (e.g., Biemiller, 2009; Dale & Fenson, 1996; Hart & Risley, 1999) and eight published preschool curricula to develop an appropriate word list for their Picture Naming assessment. Only one of the identified 10 research studies (Craig & Washington, 2002) involved the recording of spontaneous language samples in 15–20 minute segments while children played in pairs with action figures, dolls, and a toy school house. Craig and Washington assessed children's expressive vocabulary by calculating the number of different words that children expressed, average clause length, and measures of syntactic complexity in the recorded segments.

Communicative Competence Assessment

Fundamental to oral language assessment is gathering information about children's communicative competence; information that shows how children achieve social intentions across a wide range of situations encountered in daily life (Genishi & Dyson, 2009; Ochs & Capps, 2001). Children's communicative intent and use of language to achieve social purposes comprised a small piece of one of the final set of 10 studies. In this study (Smith et al., 2000), the researchers engaged children in four dyadic-conversations. They asked questions to elicit 10 categories of communicative intentions (e.g., greeting, requesting, informing, rejecting, reasoning, closing the conversation) using the *Test of Pragmatic Skills* (Shulman, 1986). Smith and colleagues (2000) explain that adult-child conversations used to assess children's communicative intent are meant to provide a "more natural conversational context" (p. 103).

While it is true that children are conversational participants in this assessment and that the language samples are more authentic than those of the standardized vocabulary assessments, the flow of conversation is somewhat contrived in an effort to elicit particular communicative intentions that align with the assessment protocol. Furthermore, this rather restrictive communicative context provides limited information about children's use of language to achieve social intentions, an assessment purpose aligned with the sociolinguistic theoretical approach (Halliday, 1975; Vygotsky, 1978) underpinning our systematic review.

Assessment Practices

As an extension of the communicative competence findings, a discussion of the assessment practices enlisted by researchers in the final 10 studies of the review requires attention. The most frequently used assessment practice involved adults asking an individual child to respond to oral language input related to pictures. This practice was evident in all 10 of the studies included in the final synthesis. For example, in the administration of the PPVT-4 for receptive vocabulary and TOLD-2 Primary test of semantic, morphological, and syntactical receptive language skills, an adult presents pictures of four common objects and actions, asking the child to point to one of them. In a similar manner, assessments of children's expressive vocabulary, grammatical skills, and spoken dialect use (e.g., PPVT-4; Expressive One Word Picture Vocabulary Test—III, Picture Naming Test;

TOLD-2; TACL-R) ask children to name objects and actions in pictures or repeat stimuli provided for pictures. Tests of oral comprehension (Craig & Washington, 2002; Craig et al., 2005) involved an adult asking “Wh” questions about activity pictures and asking children to point to pictures showing the actions in sentences with active and passive constructions that are spoken aloud by the adult.

Children's narrative language samples were elicited from the wordless picture book, *Frog, Where Are You?* (Mayer, 1969) in three studies (Heilmann, Miller, & Nockerts, 2010; Hipfner-Boucher et al., 2014; Patton Terry et al., 2013). The adult examiner read the story aloud to the child, inviting the child to follow along by looking at the pictures, and then asked the child to retell the story. Schneider and Vis Dubé (2005) presented three stories that they had developed from the pictures of one book (*Oops* by Mercer Mayer, 1977) to typically developing, English-speaking kindergarten and Grade 2 children in their study. The stories were then presented orally, orally with pictures, or with pictures only to each child by a researcher, and the child retold the story to a different researcher.

Adult-child conversations used to assess children's communicative intent are meant to provide a “more natural conversational context” (Smith et al., 2000, p. 103), though it is the adults that select conversation topics and direct the flow of conversation in an effort to elicit particular communicative intentions in children's language.

Summary

Our review of oral language assessments for four-to eight-year-old children revealed a diverse body of assessments administered for an array of testing purposes. The use of narrative assessments focused attention on the story grammar knowledge of young children (i.e., narrative macrostructures) (Heilmann, Miller, & Nockerts, 2010; Hipfner-Boucher et al., 2014; Patton Terry et al., 2013; Schneider & Vis Dubé, 2005), syntactic structures (i.e., narrative microstructures) (Craig & Washington, 2002; Craig et al., 2005; Patton Terry et al., 2013), and in one instance, dialect (Craig & Washington, 2002). Vocabulary assessments, such as the PPVT and the PLS, were used, not only to directly assess oral language (Qi et al., 2006; Qi & Marley, 2011), but also as validation instruments for other vocabulary assessments (Bradfield et al., 2014), for narrative assessments

(Patton Terry et al., 2013), and for others addressing syntax (Craig et al., 2005). Less diversity was observed in the types of assessment practices of the identified studies for review. Nine of the 10 studies involved adult-directed assessment practices conducted by researchers or speech language pathologists. An acknowledgment of the importance in studying underrepresented populations was clear, as six of the 10 studies specifically targeted underrepresented ethnic, language, or socio-economic communities (Craig et al., 2002; Craig & Washington, 2005; Patton Terry et al., 2013; Qi et al., 2006; Qi & Marley, 2011; Smith et al., 2000). These studies demonstrate some of the limitations of administering standardized tests without considering a child's local dialect and/or socio-economic context.

Our systematic review of research has revealed that researchers in the speech-language pathology field are far more likely to take up research questions related to oral language assessment, than are educational researchers. Despite educational researchers' recognition of the importance of oral language to children's learning and literacy, there is a paucity of classroom-based educational research and resources for teachers on the topic of oral language assessment. As shown through the results of the systematic review, the oral language assessment research findings have much to offer classroom teachers in terms of the content of oral language assessment (e.g., vocabulary, narrative structures, syntactic structures, and communicative competence), including a recognition of the syntax of children speaking diverse English dialects. The criterion-referenced assessments conducted in the identified research studies (e.g., analysis of children's narrative retellings and of their spontaneous free-play interactions) provide the needed flexibility to capture the complexities of children's language across linguistic and cultural contexts. Given that the research findings of studies highlighted in our systematic review address research purposes specific to speech-language pathology and special education, however, it is important to integrate the findings with those of research having a classroom focus. The integration of the two fields can enrich teachers' classroom practices to support children's oral language development.

Implications

Our systematic review of oral language assessment research provides implications for policy development, as well as classroom practice, including that arising from our action research. Given that most of the 201 studies that were relevant to our search and all of the 10 articles on the final list have been published in journals of speech language pathology, communication disorders, language or psychoeducational testing, special education, and assessment for intervention, it is apparent that the fields of speech language pathology and special education have much to offer classroom teachers in their assessment of young children's oral language.

For Policy

Many structures are in place that separate researchers and practitioners in the two fields, making it difficult for each to inform the other. For example, the preparation of speech-language pathologists often takes place in science faculties and speech-language assessment research is generally funded from science research envelopes, whereas teacher preparation takes place in social sciences faculties and classroom assessment research funds generally come from social science-oriented sources. In our Canadian province, funding for speech-language services provided for children comes from either the Ministry of Health, Ministry of Community and Social Services, or the Ministry of Education, depending on the nature of individual children's speech and language needs, and on their age (Ontario Ministry of Education, 2016). Indeed, it is only through activities related to our action research project (e.g., carrying out this systematic review and being contacted by speech-language pathologists working in the Indigenous communities that are part of our project to collaborate in designing dynamic assessment tools appropriate for Indigenous children) that we literacy researchers/teacher-educators have become aware of the healthy overlap, in terms of the goals of the two fields, and the considerable expertise and knowledge that each field can offer the other. To our knowledge, the collaborative research in which speech-language pathologists and we literacy researchers are engaging is rather unique.

An immediate implication for policy, thus, is to create forums for teachers, literacy educators, and speech language pathologists to collaborate and learn from each

other in order to support all children's oral language in classroom and speech-language service settings. Additionally, both fields would benefit from collaborative research that integrates research purposes relevant to the work of speech-language pathologists and classroom teachers, and draws on the expertise of each. The fundamental nature of oral language to children's learning and literacy should be recognized in policies guiding teacher education programs and classroom assessment, as well as the provision of speech and language services in schools.

For Practice

Our experience in working with kindergarten and Grade 1 teachers and with speech-language pathologists as part of our action research shows that both groups of professionals, who work directly with young children, recognize the importance of oral language and have observed the need for a focus on supporting oral language in classroom instruction and in research. The findings of the systematic review show that oral language assessment research using children's oral narratives has direct application to classroom practice, both in the methods used to gather language samples and in the features of children's language that are assessed. However, through interviews with 36 primary teachers, early childhood educators, and consultants in northern communities in four Canadian provinces, our action research project has, so far, highlighted the need for a framework to guide classroom assessment of children's oral language. Participants in our action research project describe being underprepared to assess and support children's oral language because of a lack of attention on oral language in teacher education programs, and a dearth of resources available to teachers (Peterson, McIntyre, & Forsyth, 2016). Teachers participating in our action research tell us that it is not practical for teachers in classrooms of 20 or more children to assess children's use of syntax to the degree that researchers do in a one-time, one-on-one gathering of language samples. Additionally, classroom language assessments are most helpful to teachers when they provide information about children's response to as wide a range as possible of communicative demands that are placed on them in their everyday lives (Barnes, 1992; Halliday, 1975).

Together with our participant teachers, we are drawing on the oral language assessment research identified in our systematic review to create a manageable classroom assessment scheme that will provide useful information about children's language to

inform teaching practice. Specifically, we are working on approaches to assessing children's language growth in terms of the length of children's utterances, as well as their use of conjunctions and clauses, in interactions while engaged in dramatic and construction play. Sociodramatic play is recommended as a context for supporting (Nicolopoulou, McDowell, & Brockmeyer, 2006) and assessing children's expressive vocabulary, as we are finding, where our participant children have used language for 35 specific purposes in their play. In sociodramatic play, they most frequently used language for the purpose of imagining, and in construction play, they used language most frequently for learning and to satisfy their own needs. Participating teachers have co-created a classroom assessment tool (*Assessing Children's Use of Language—ACUL*) to record and analyze individual children's use of language in play and other contexts where children interact with others in the classroom (Portier & Peterson, 2017).

The professional learning that we all have gained through drawing on each other's knowledge and experience shows us, on a small scale, the benefits of collaboration between two fields that have traditionally been separate, yet have much in common as we draw on assessment information to support young children's oral language.

References

- Alberta Education. (2000). *English language arts K-9*. Edmonton, AB: Alberta Education. Retrieved from <http://education.alberta.ca/media/450519/elak-9.pdf>
- Alexander, R. (2006). *Towards dialogic teaching*. York, England: Dialogos.
- Barnes, D. (1992). *From communication to curriculum* (2nd ed.). Portsmouth, NH: Boynton/Cook Publishers.
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 18(1), 5–25.
- Biemiller, A. (2009). *Words worth teaching*. DeSoto, TX: SRA/McGraw-Hill.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(5), 5–31.
- Bloom, L., & Lahey, M. (1978). *Language development and language disorders*. New York, NY: John Wiley & Sons.
- Bradfield, T. A., Besner, A. C., Wackerle-Hollman, A. K., Albano, A. D., Rodriguez, M. C., & McConnell, S. R. (2014). Redefining individual growth and development indicators: Oral language. *Assessment for Effective Intervention*, 39(4), 233–244.
- Bruner, J. S. (1986). *Actual minds, possible worlds*. Cambridge, MA: Harvard University Press.
- Bruner, J. S. (1990). *Acts of meaning*. Cambridge, MA: Harvard University Press.
- Brunton, G., Stansfield, C., & Thomas, J. (2012). Finding relevant studies. In D. Gough, O. Sandy, & J. Thomas (Eds.), *An introduction to systematic reviews* (pp. 107–134). London, England: Sage Publications.
- Carrow-Woolfolk, E. (1985). *Test for Auditory Comprehension of Language—Revised*. Allen, TX: DLM Teaching Resources.
- Cekaite, A., Blum-Kulka, S., Grøver, V., & Teubal, E. (2014). Children's peer talk and learning: Uniting discursive, social and cultural facets of peer interactions: Editors' introduction. In A. Cekaite, S. Blum-Kulka, V. Grover, & E. Teubal (Eds.), *Children's peer talk: Learning from each other* (pp. 3–19). Cambridge, England: Cambridge University Press.

- Craig, H. K., & Washington, J. A. (1994). The complex syntax skills of poor, urban, African American preschoolers at school entry. *Language, Speech, and Hearing Services in Schools, 25*(3), 181–190.
- Craig, H. K., & Washington, J. A. (2002). Oral language expectations for African American preschoolers and kindergartners. *American Journal of Speech-Language Pathology, 10*(1), 59–70.
- Craig, H. K., Connor, C. M., & Washington, J. A. (2003). Early positive predictors of later reading comprehension for African American students: A preliminary investigation. *Language, Speech, and Hearing Services in Schools, 34*(1), 31–43.
- Craig, H. K., Washington, J. A., & Thompson, C. A. (2005). Oral language expectations for African American children in grades 1 through 5. *American Journal of Speech-Language Pathology, 14*(2), 109–130.
- Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavioral Research Methods, Instruments, & Computers, 28*(1), 125–127.
- Department for Education. (2013). *English programmes of study: Key stages 1 and 2 National curriculum in England*. London, England: Author.
- DeTemple, J. M., Wu, H. F., & Snow, C. E. (1991). Papa Pig just left for Pigtown: Children's oral and written picture descriptions under varying instructions. *Discourse Processes, 14*(4), 469–495.
- Dickinson, D. K., & Porche, M. V. (2010). Relation between language experiences in preschool classrooms and children's kindergarten and fourth-grade language and reading abilities. *Child Development, 82*(3), 870–886.
- Dickinson, D. K., & Tabors, P. O. (2001). *Beginning literacy with language*. Baltimore, MD: Paul H. Brookes.
- Dunn, L., & Dunn, L. (2007). *The Peabody picture vocabulary test* (4th ed.). Bloomington, MN: Pearson.
- Fisher, D., & Frey, N. (2007). *Checking for understanding*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Genishi, C., & Dyson, A. H. (2009). *Children, language, and literacy: Diverse learners in diverse times*. New York, NY: Teachers College Press.

- Glaser, B. G. (1978). *Theoretical sensitivity*. Mill Valley, CA: Sociology Press.
- Gough, D. (2007). Weight of evidence: A framework for the appraisal of the quality and relevance of evidence. *Research Papers in Education, 22*(2), 213–228.
- Gough, D., & Thomas, J. (2012). Commonality and diversity in reviews. In D. Gough, O. Sandy, & J. Thomas (Eds.), *An introduction to systematic reviews* (pp. 35–65). London, England: Sage Publications.
- Gutierrez-Clellen, V. F., & Quinn, R. (1993). Assessing narratives of children from diverse cultural/linguistic groups. *Language, Speech, and Hearing in Schools, 24*, 2–9.
- Halliday, M. A. K. (1975). *Learning how to mean: Explorations in the functions of language*. London, England: Edward Arnold.
- Harris, J., Michnick Golinkoff, R., & Hirsh-Pasek, K. (2011). Lessons from the crib for the classroom: How children really learn vocabulary. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research Volume 3* (pp. 49–65). New York, NY: Guilford Press.
- Hart, B., & Risley, T. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: P. H. Brookes.
- Heilmann, J., Miller, J. F., & Nockerts, A. (2010). Sensitivity of narrative organization measures using narrative retells produced by young school-age children. *Language Testing, 27*(4), 603–626.
- 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 preschoolers' oral language and phonological awareness. *First Language, 34*(2), 178–197.
- Hondrich, A. L., Hertel, S., Adl-Amini, K., & Kleime, E. (2015). Implementing curriculum-embedded formative assessment in primary school science classrooms. *Assessment in Education: Principles, Policy & Practice, 23*(3), 353–376.

- Hudson, J. A., Shapiro, L. R., McCabe, A., & Peterson, C. (1991). From knowing to telling: The development of children's scripts, stories, and personal narratives. In A. McCabe & C. Peterson (Eds.), *Developing narrative structure* (pp. 89–136). Hillsdale, NJ: Lawrence Erlbaum.
- Hymes, D. (1966). Two types of linguistic relativity. In W. Bright (Ed.), *Sociolinguistics* (pp. 114–158). The Hague, the Netherlands: Mouton.
- Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal*, 33(3), 465–491.
- Lonigan, C. J., Wagner, R. K., Torgeson, J. K., & Rashotte, C. A. (2007). *Test of preschool early literacy*. Austin, TX: Pro-Ed.
- Manhardt, J., & Rescorla, L. (2002). Oral narrative skills of late talkers at ages 8 and 9. *Applied Psycholinguistics*, 23(1), 1–21.
- Manitoba Education. (2011). *Early returns: Manitoba's early learning and child care curriculum framework for preschool centres and nursery schools*. Winnipeg, MB: Manitoba Education. Retrieved from https://www.gov.mb.ca/fs/childcare/resources/pubs/early_returns.pdf
- Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, 11(3), F9–F16.
- Martinez, M., & Roser, N. (1985). Read it again: The value of repeated readings during storytime. *The Reading Teacher*, 38(8), 782–786.
- Mayer, M. (1969). *Frog, where are you?* New York, NY: Dial Press.
- McCabe, A., Bliss, L., Barra, G., & Bennett, M. (2008). Comparison of personal versus fictional narratives of children with language impairment. *American Journal of Speech-Language Pathology*, 17, 194–206.
- Melzi, G., Schick, A. R., & Kennedy, J. L. (2011). Narrative elaboration and participation: Two dimensions of maternal elicitation style. *Child Development*, 82(4), 1282–1296.
- Mercer, N., & Littleton, N. (2007). *Dialogue and the development of children's thinking: A sociocultural approach*. Abingdon, England: Routledge.

- Michaels, S. (1981). 'Sharing time': Children's narrative styles and differential access to literacy. *Language in Society*, 10(3), 423–442.
- Ministère de l'éducation nationale, de l'enseignement supérieur et de la recherche. (2016). Les programmes de l'école élémentaire. *Présentation des programmes et des horaires à l'école élémentaire*. Retrieved from http://www.education.gouv.fr/cid38/presentation-des-programmes-et-des-horaires-a-l-ecole-elementaire.html#Cycle_2_-_cycle_des_apprentissages_fondamentaux_CP_CE1_et_CE2
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic review and meta-analyses: The PRISMA statement. *PLoS Med*, 6(7): e1000097. doi:10.1371/journal.pmed.1000097
- Morrow, L. M. (1990). Assessing children's understanding of story through their construction and reconstruction of narrative. In L. M. Morrow & J. K. Smith (Eds.), *Assessment for instruction in early literacy* (pp. 110–134). Englewood Cliffs, NJ: Prentice-Hall.
- Neuman, S. B., & Dwyer, J. (2009). Missing in action: Vocabulary instruction in pre-K. *The Reading Teacher*, 62(5), 384–392.
- Nicolopoulou, A., McDowell, J., & Brockmeyer, C. (2006). Narrative play and emergent literacy: Storytelling and story-acting meet journal writing. In D. Singer, R. Golinkoff, & K. Hirsh-Pasek (Eds.), *Play=Learning*. New York, NY: Oxford University Press.
- Ochs, E., & Capps, L. (2001). *Living narrative: Creating lives in everyday storytelling*. Cambridge, MA: Harvard University Press.
- Oliver, S., Dickson, K., & Newman, M. (2012). Getting started with a review. In D. Gough, O. Sandy, & J. Thomas (Eds.), *An introduction to systematic reviews* (pp. 66–82). London, England: Sage Publications.
- Ontario Ministry of Education. (2006). *The Ontario curriculum grades 1-8 language*. Toronto, ON: Author.
- Ontario Ministry of Education. (2016). *Interministerial guidelines for the provision of speech and language services (as applicable to the Education Act)*. Retrieved from <http://www.edu.gov.on.ca/extra/eng/ppm/guide.html>

- Owens, R. E. (2012). *Language development: An introduction* (8th ed.). Upper Saddle River, NJ: Allyn & Bacon.
- Patton Terry, N., 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*(3), 291–305.
- Pearce, W., McCormack, P., & James, D. (2003). Exploring the boundaries of SLI: Findings from morphosyntactic and story grammar analyses. *Clinical Linguistics & Phonetics, 17*(4-5), 325–334.
- Peltier, S. (2010). *Valuing children's storytelling from an Anishinaabe orality perspective* (Unpublished master's thesis). North Bay, ON: Nipissing University. Retrieved from <http://www.ecdip.org/docs/pdf/Sharla's%20Final%20Revision%20without%20signatures.pdf>
- Pence, K., Justice, L. M., & Gosse, C. (2007). *Narrative assessment protocol*. Columbus, OH: Preschool Language and Literacy Lab, The Ohio State University.
- Peterson, S. S., McIntyre, L., & Forsyth, D. (2016). Supporting young children's oral language and writing development: Teachers' and early childhood educators' goals and practices. *Australasian Journal of Early Childhood, 41*(3), 11–19.
- Piquemal, N. (2003). From Native North American oral traditions to Western literacy: Storytelling in education. *Alberta Journal of Educational Research, 49*(2), 113–122.
- Portier, C., & Peterson, S. S. (2017). Rural northern Canadian teachers' discoveries about young children's oral language. *Language and Literacy, 19*(2), 109–126.
- Qi, C. H., Kaiser, A. P., Milan, S., & Hancock, T. (2006). Language performance of low-income African American and European American preschool children on the PPVT-III. *Language, Speech & Hearing Services in Schools, 37*(1), 5–16.
- Qi, C. H. & Marley, S. C. (2011). Validity study of the Preschool Language Scale–4 with English-speaking Hispanic and European American children in Head Start programs. *Topics in Early Childhood Special Education, 31*(2), 89–98.

- Reilly, J., Losh, M., Bellugi, U., & Wulfeck, B. (2004). Frog, where are you? Narratives in children with specific language impairment, early focal brain injury and Williams Syndrome. *Brain & Language*, 88(2), 229–247.
- Resnick, L. B., & Snow, C. E. (2009). *Speaking and listening for preschool through third grade*. Newark, DE: International Reading Association.
- Saskatchewan Ministry of Education. (2010). *Saskatchewan curriculum: Kindergarten*. Retrieved from https://www.edonline.sk.ca/bbcswebdav/library/curricula/English/Kindergarten_K_2010.pdf
- Schneider, P., & Vis Dubé, R. (2005). Story presentation effects on children's retell content. *American Journal of Speech-Language Pathology*, 14(1), 52–60.
- Shulman, B. (1986). *Test of pragmatic skills—Revised*. Tucson, AZ: Communication Skill Builders.
- Smith, T., Myers-Jennings, C., & Coleman, T. (2000). Assessment of language skills in rural preschool children. *Communication Disorders Quarterly*, 21(2), 98–103.
- van Kraayenoord, C. E., & Paris, S. G. (1996). Story construction from a picture book: An assessment activity for young learners. *Early Childhood Research Quarterly*, 11(1), 41–61.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wells, G. (1999). *Dialogic inquiry: Towards a sociocultural practice and theory of education*. Cambridge, England: Cambridge University Press.
- Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2002). *Preschool language scale—4*. San Antonio, TX: Pearson.