

**Critical Review:**  
**How does performance on narrative retell compare between English Language Learners with and without language impairment?\***

Sunny Shaffner  
M.Cl.Sc. SLP Candidate  
University of Western Ontario: School of Communication Sciences and Disorders

This critical review compares the performance of English Language Learners (ELLs) with language impairment (LI) to English Language Learners with typical language (TL) on narrative retell. The five studies examined whether there is a difference between the two groups on narrative retell and evaluate elements of micro- and macrostructure to determine which differences are present. All studies employed a nonrandomized design, with two between group studies and two mixed studies. Overall, the research provides compelling levels of evidence that ELLs with language impairment perform significantly lower on narrative retell compared to ELLs without language impairment. More research is needed to determine which elements of micro- and macrostructure differ between the two groups. This research may have clinical implications in the use of narrative retell in language assessment of ELLs.

### *Introduction*

Within the Canadian multilingual and multicultural context, there is an increasing need for research to guide evidence-based language assessment of English Language Learners (ELLs) from diverse first language backgrounds. ELL language assessment issues are complex. Bilingual children, including ELLs, have been over-identified as having language impairment on standardized tests normed for monolingual children. At the same time, a “wait and see” approach prevents early intervention and it may take up to five years for the language impairment to be recognized (Paradis, Schneider, & Duncan, 2013). The diversity of first languages spoken by ELLs increases the complexity of assessment, as it increases the difficulty of providing comprehensive assessment of the children’s dual language abilities (Paradis et al., 2013). In addition to English and French, 120 languages are spoken by children in Canada (Paradis et al., 2013). Therefore, it is essential to develop assessment measures to reliably identify language impairment through assessment of second language ability.

Language sample analysis (LSA) has been proposed as the gold standard for the assessment of language impairment in children (Jacobson & Walden, 2013), especially for the assessment of linguistically diverse children. Narratives are often used as a language-sampling task because they can evaluate children’s knowledge of vocabulary, grammatical structures, and story structure (Squires et al., 2014). Narratives are also more naturalistic than standardized tests. It has been suggested that narrative retell places additional language processing demands compared with spontaneous narrative tasks, since narrative retell requires the use of particular vocabulary and expanded use of specific

grammatical forms (Squires, 2014). Narrative retell tasks include processing demands of attention and memory, which are skills underlying the difficulties learning language for children with language impairment (Peña, 2014). Narrative retell may be used to examine macrostructure (e.g., story grammar) or microstructure (e.g., cohesion devices, mental and linguistic verbs, and elaborated noun phrases).

Narrative retell has the potential to compare the performance of ELLs to their ELL peers. Research regarding ELL norms for narrative retell in typically developing children and children with language impairment is in the initial stages. The first step in determining the clinical use of narrative retell to assess language status in ELLs is to determine whether there is a significant difference in performance between children with and without language impairment, and to determine which specific outcome measures are different.

### *Objectives*

The primary objective of this paper is to critically evaluate existing literature comparing performance on narrative retell between English Language Learners with and without language impairment. The secondary objectives of this review are to explore the clinical implications and to make recommendations for future research.

### *Methods*

#### Search Strategy

Computerized databases including PubMed, CINAHL, SCOPUS, and Google Scholar were searched. The following keywords were used: [(narrative) OR (story

*\*This paper was created as a required assignment for the CSD9639 Evidence Based Practice for Clinicians course at Western. While it has been evaluated by course instructors for elements of accuracy and style, it has not undergone formal peer-review.*

retell)) AND ((bilingual) OR (English as a Second Language) OR (English Language Learners) OR (multilingual)) AND ((language impairment) OR (language disorder) OR (language delay)]. Reference lists of retrieved articles were reviewed to find additional relevant articles.

#### Selection Criteria

Selected studies for inclusion in this review were required to compare performance on narrative tasks between ELLs with language impairment and typical language. Studies included preschool and elementary school age children. Both sequential and simultaneous bilingual children were included. Studies using a narrative retell task were included, but studies using personal narratives were excluded.

#### Data Collection

Results of the literature search yielded four articles that met selection criteria. All articles were nonrandomized; two were between group and two were mixed designs.

### ***Results***

Two of the studies reviewed employed a nonrandomized between subjects design, while two studies used a nonrandomized mixed design. In all studies, the comparison of narrative performance was evaluated between the LI and TL groups, which was appropriate to the research question. Randomization was not possible since one group must include ELLs with language impairment and the other group must include ELLs with typical language.

**Jacobson and Walden (2013)** compared ELLs with LI (n=22) and ELLs with TL (n=26) on a narrative retell task using wordless picture books. All children were sequential Spanish-English bilinguals. This study also explored whether lexical diversity and word or morpheme omission could predict language status. However, the present critical review will focus on the comparison between groups. The story retell was completed in both languages, and transcribed and coded using the Systematic Analysis of Language Transcripts (SALT) coding conventions. Appropriate independent t-test comparisons revealed statistically significant differences between groups on the number of word and morpheme omissions, but no difference for lexical diversity. The results held true for narrative retell in both English and Spanish.

One strength of this study was that it specifically evaluated omission errors and lexical diversity through narrative retell, in order to determine which specific elements of story retell are different between groups. The measures and procedures were described clearly.

Importantly, the authors made a distinction between language learning ability and language proficiency, and focused on poor language learning ability as an indicator of language impairment. Highly reliable assessments were used to classify participants as having LI or TL initially. Inter-rater reliability for coding of transcripts was high and was described in detail. Finally, a strength of the use of SALT coding conventions is that it could be used to analyze narratives for any language written using the Roman alphabet.

A limitation to this study is the lack of blinding of the clinicians or researchers. This study included only Spanish-English speakers, who have been more researched than any other bilingual population (Paradis, 2013), and whose results cannot be generalized to children with linguistically diverse backgrounds.

Overall, this study provides compelling evidence of a difference between ELLs with LI and TL on word and morpheme omissions in a narrative retell task.

**Paradis, Schneider, and Duncan (2013)** examined a combination of language measures to distinguish between LI and TL in ELLs from diverse language backgrounds. The present critical review will focus on the group comparison on the story retell task. The performance of sequential bilingual ELLs with TL (n=152) and LI (n=26) on narrative macrostructure (story grammar) was compared using the Edmonton Narrative Norms Instrument (ENNI). Story retells were completed in English. Story retells were recorded and then transcribed using the CHAT system. Appropriate independent samples t-tests indicated that the ELL TL group scored significantly higher on the ENNI compared to the LI group. The effect size between groups was medium. The authors concluded that narrative retells could be valuable to include as part of a bilingual language assessment.

A strong rationale was provided to include narrative skills within the larger test battery given in this study, and to examine narrative macrostructure specifically. One strength of this study was the diversity of the first languages of participants, which increases generalizability. ELLs with TL and with LI were matched according to tense marking or non-tense marking first languages, as well as the length of their exposure to English. Importantly, these variables were based on previous research of bilingual children. There was high inter-rater reliability reported for coding transcripts.

This study provides compelling evidence of a difference in story grammar on narrative retell between ELLs with TL and ELLs with LI.

**Peña, Gillam, and Bedore (2014)** explored the identification accuracy of dynamic assessment of narrative ability by comparing ELLs with LI (n=18) and TL (n=18). Of relevance to the present review is the group comparison of narrative ability. Children completed a story retell task with wordless picture books before and after mediated learning sessions took place. Stories were transcribed using SALT, and scored on 10 qualitative items that resulted in a total story score (e.g., story components, story ideas, episode structure). An appropriate ANOVA was conducted, and pairwise comparisons indicated that children in the LI group scored significantly lower than the TL group on the total story score. ELLs with LI produced a greater proportion of ungrammatical utterances compared to the TL group. No significant differences were found between groups for number of different words, total number of words, mean length of utterance, and number of main verbs.

One strength of this study is the strong evidence based classification of children into the LI and TL groups. All examiners were blind to children's language ability. Elements of both microstructure and macrostructure were included.

Only Spanish-English bilinguals were included, despite the author explicitly stating the need for language assessment of bilingual children from diverse language backgrounds.

Overall, this study provides compelling evidence of a difference between ELLs with LI and ELLs with TL on narrative retells, specifically in the areas of macrostructure and proportion of grammatical utterances.

**Squires et al. (2014)** conducted a longitudinal study to determine whether ELLs with TL (n=21) and with LI (n=21) present similar gains from kindergarten to first grade in the macro and microstructure of story retell. A mixed design was employed. However, the present critical review will focus only on the comparison between groups. Children were asked to retell wordless picture books during kindergarten and first-grade years in both Spanish and English. Macro and microstructure were coded with an adapted version of Monitoring Indicators of Scholarly Language (MISL). Appropriate tests of simple main effects using the estimated marginal means procedure in SPSS revealed that the TL group earned significantly higher macrostructure scores than the children in the LI group at kindergarten and at first grade. An ANOVA revealed that the TL group performed significantly higher on microstructure than the LI group at both time points.

ELLs were identified as LI or TL according to best practices cited in previous research, with agreement between three SLPs who independently evaluated the children. Participants were carefully matched based on factors (e.g., language input and output in both languages, and results of the Universal Non-Verbal Intelligence Test). The inter-rater reliability was high for scoring story retells. The authors included scoring rubrics for the story retells, and described the study in enough detail that it could be easily replicated.

A limitation of this study is that only Spanish-English bilinguals were included in the study, which limits the generalizability of the results.

Overall, this study provides compelling evidence of a difference between groups in narrative retell on both micro- and macrostructure.

### *Discussion*

The primary objective of the present critical review was to evaluate existing literature comparing performance on narrative retell between English Language Learners with and without language impairment. Overall, the level of evidence is compelling that there is a significant difference between these groups.

The overall findings of the four reviewed studies are in agreement that there is a difference in performance on narrative retell between ELLs with LI and TL. Each study is different in the outcome measures it evaluates, but taken together the studies can be seen as initial evidence of differences between the two groups.

Three of the four studies revealed differences in microstructure elements between the TL and LI groups (Jacobson and Walden, 2013; Peña, 2014; Squires, 2014). Specifically, the LI group had more word and morpheme omissions (Jacobson and Walden, 2013) and a higher proportion of grammatical utterances (Peña, 2014). Three of the four studies indicated differences in macrostructure elements between groups on elements such as story components, story ideas, and episode structure (Peña, 2014; Paradis, 2013; Squires, 2014).

The comparison between the LI and TL groups on narrative retell was also informative to determine which elements were not significantly different between groups. Jacobson and Walden (2013) indicated that there was no difference between groups on lexical diversity, and indicated that lexical diversity gave an indication of language proficiency but not language impairment. Peña (2014) indicated no significant differences were found between groups for number of different words, total number of words, mean length of utterance, and number of main verbs.

Overall, the reviewed studies shared several strengths. The classification of language status as language impairment or typical language was strong in all studies, and was based on previous literature in this area. Studies ensured that speech language pathologists involved in identifying language status were bilingual and were familiar with bilingual language assessment constraints. Reliability between SLP ratings of language status was very high.

The four studies were described in detail and could be replicated. Narrative story retell was completed using wordless picture books. The study by Squires et al. (2014) included a scoring rubric and a detailed description of coding. Inter-rater reliability was high for coding story retells, and was a strength across studies.

One limitation to the generalizability of evidence from the studies reviewed is that three of the four studies included only sequential Spanish-English bilingual children. Spanish-English bilinguals have been researched more than any other bilingual population (Paradis, 2013). There is a recognized need to conduct research on children from linguistically diverse backgrounds, particularly within the multilingual context of Canada (Paradis, 2013).

One limitation to the present critical review was the heterogeneity of outcome measures present in each study. Although the use of narrative retell was consistent across studies, some studies chose to examine only microstructure elements or only macrostructure elements. In addition, some studies described elements of micro- and macrostructure in detail, while others did not.

It should be noted that the studies included in this review were all published during the last five years, and provide initial evidence that there is a difference between LI and TL groups on narrative retell in both micro- and macrostructure. More research is needed to determine which elements are most indicative of a difference between groups. The importance of performance on microstructure compared to macrostructure should be determined. Since a focus of recent research on language assessment of bilingual children is with children from diverse language backgrounds, these children should be represented in research. It is important to step away from a focus on only Spanish-English bilinguals as a next step. Future research may aim to study the reliability and validity of narrative retell as an assessment tool for ELLs, and may

include a monolingual control group in addition to the bilingual LI and TL groups.

### ***Clinical Implications and Conclusion***

Evidence of a difference in narrative retell between LI and TL groups is the first step toward using narrative retell in assessment of ELLs in order to determine language status.

Further predictive studies as well as norms for ELLs are needed. It has been suggested by Paradis (2013) that narratives may currently be valuable in combination with other measures as part of bilingual language assessment. Narrative retell may be used as an informal measure to add to the clinician's understanding of language ability. Additionally, the evidence from the reviewed studies supports the idea of language assessment in a child's second language.

In conclusion, the four studies reviewed provided compelling evidence that there is a difference between LI and TL bilingual children on performance on narrative tasks. Future research is needed to identify the most important outcome measures, to increase the linguistic diversity of participants studied, and to study the diagnostic reliability and validity of narrative retell.

### ***References***

- Jacobson, P., & Walden, P. (2013). Lexical diversity and omission errors as predictors of language ability in the narratives of sequential Spanish-English bilinguals: A cross language comparison. *American Journal of Speech-Language Pathology*, 22, 554-565.
- Paradis, J., Schneider, P., & Duncan, T. (2013). Discriminating children with language impairment among English language learners from diverse first-language backgrounds. *Journal of Speech, Language, and Hearing Research*, 56, 971-981.
- Peña, E., Gillam, R., & Bedore, L. (2014). Dynamic assessment of narrative ability in English accurately identifies language impairment in English language learners. *Journal of Speech, Language, and Hearing Research*, 57, 2208-2220.
- Squires, K., Lugo-Neris, M., Peña, E., Bedore, L., Bohman, T., & Gillam, R. (2014). Story retelling by bilingual children with language impairments and typically developing controls. *International Journal of Language and Communication Disorders*, 49 (1), 60-74.

