A Comparison of Mercer Mayer’s Frog Books in the Narrative Assessment of Cantonese-English Bilinguals

Laura G. Peterson

*University of Colorado Boulder, lape8064@colorado.edu*

Follow this and additional works at: [https://scholar.colorado.edu/honr_theses](https://scholar.colorado.edu/honr_theses)

Part of the Speech and Hearing Science Commons, and the Speech Pathology and Audiology Commons

**Recommended Citation**


[https://scholar.colorado.edu/honr_theses/2055](https://scholar.colorado.edu/honr_theses/2055)

This Thesis is brought to you for free and open access by Honors Program at CU Scholar. It has been accepted for inclusion in Undergraduate Honors Theses by an authorized administrator of CU Scholar. For more information, please contact cuscholaradmin@colorado.edu.
A Comparison of Mercer Mayer’s *Frog* Books in the Narrative Assessment of Cantonese-English Bilinguals

Laura Peterson

October 30th, 2019

Undergraduate Honors Thesis

University of Colorado, Boulder

Department of Speech, Language, Hearing Sciences

Committee Members:

Pui Fong Kan, Ph.D., CCC-SLP, Speech Language and Hearing Sciences

Kathryn Arehart, Ph.D., Speech Language and Hearing Sciences

Chase Raymond, Ph.D, Ph.D, Linguistics
Acknowledgements

First, I would like to thank Dr. Pui Fong Kan for her continual support and advising during the thesis writing process. I truly appreciate your guidance as both a professor and mentor. I would not have had such a wonderful experience pursuing undergraduate research without you as my advisor.

I would also like to thank Dr. Kathryn Arehart and Dr. Chase Raymond for reading my paper, asking thought-provoking questions, and supporting my research endeavors. The time spent engaging with my ideas and work is deeply appreciated.

I would like to thank the lab members of the Child Language Learning Lab at CU who helped to collect and transcribe the data that is used in this thesis. Their time and hard work was essential to the final product seen here.

I would also like to thank the Undergraduate Research Opportunities Program at CU for providing funding and support for this research. It is truly wonderful to be so well supported by the University of Colorado as an undergraduate student.

Lastly, I would like to thank my incredible support network, Maxwell Watrous, Megan Pielke and Samantha Bartolo, for reading and re-reading versions of the paper and providing emotional support through this process. I love you all and truly appreciate you. I couldn’t have done this without you.
Abstract

The purpose of this study is to compare two commonly used wordless picture books (*Frog On His Own* and *Frog, Where Are You?*) in the context of narrative sample analysis for bilingual assessment. Thirty-one sequential Cantonese-English bilingual language learners produced narrative retells of each book in English, which were then transcribed and analyzed using SALT 18. Samples were analyzed for multiple measures of language development and narrative microstructure including morpheme accuracy, mean length of utterance, number of different words, and number of total words. Results from the study show no significant differences between the books in terms of mean length of utterance (MLU), number of different words (NDW), or number of total words (NTW). Significant differences were observed between the two books in the number of opportunities provided to produce regular past tense markers (-ed) and articles, and in morpheme accuracy for the possessive –s. The number of omissions for the past tense –ed was also significantly higher in *Frog On His Own*. The production of other grammatical morphemes was not significantly different between the books. The findings from this study suggest that performance on the narrative retell task may vary by book with respect to the production and accuracy of specific morphemes, although no significant differences were observed for general measures (MLU, NDW, NTW). Clinical implications for narrative assessment with Cantonese English bilingual populations are discussed.
Introduction

The effective assessment of bilingual populations is a growing concern within the field of speech language pathology. In the United States, it is estimated that 20.7% of individuals, ages five and older, speak a language other than English in the home (U.S. Bureau of Census, 2015). Moreover, it is estimated that over 50% the global population is bilingual (Grosjean, 2012). Comparatively, demographics provided by the American Speech Language Hearing Association (ASHA) indicate that only 6% of speech language pathologists (SLPs) are bilingual. Given these figures, it is conceivable that a practicing monolingual SLP will find her/himself assessing and possibly treating a bilingual child (and likely many bilingual children) during his/her career.

When this speech language pathologist evaluates a bilingual child for the first time, it is essential that they are equipped with evidence-based assessment tools to provide an accurate assessment and diagnosis of that child. Without effective assessment, bilinguals are at risk of being over or under diagnosed (Paradis, 2010), which in turn can result in a loss of time and resources for the child who is over diagnosed and result in poor language outcomes for the child who is under diagnosed. Thus, effective assessment for bilinguals is of critical importance.

To account for growing culturally and linguistically diverse populations, research in the field has focused on identifying new assessment methods that accurately predict language impairment in bilingual children with both sensitivity and specificity (Gutierrez-Clellen & Peña, 2001; Gillam & Peña, 2004; Petersen, Chantrongthip, Ukrainetz, Spencer, & Steeve, 2017; Bedore & Peña, 2008, Laing & Kamhi, 2003). In the comprehensive assessment of bilingual children, clinicians and researchers often use narrative retell tasks as a way of eliciting robust language samples in both languages. The samples can be analyzed in multiple linguistic domains, including semantics, morphology, and syntax. The measures are resistant to cultural
and linguistic biases typically found in standardized tests (Rojas & Iglesias, 2009). This narrative sample analysis is useful for clinicians and researchers alike, as it provides an accessible, flexible and culturally sensitive way of assessing bilingual language learners. Current research investigating language sample analysis in bilingual populations finds that the narrative task is sensitive to developmental changes (Muñoz, Gillam Peña & Gulley-Faehnle, 2003), predictive of later literacy (Miller, Heilmann, Nockerts, Iglesias, Fabiano, & Francis, 2006; Uchikoshi, Lu & Liu, 2017) and promising for distinguishing typically developing children from children who may have developmental language disorders (Squires, Lugo-Neris, Peña, Bedore, Boham & Gillam, 2014).

In both clinical and research-based settings, Mercer Mayer’s *Frog* Books, a popular series of wordless picture books, are frequently used for the narrative retell task. Given their widespread popularity in the field of language assessment, more research is needed to examine the possible effects of book choice on the narrative samples produced by bilingual language learners. The purpose of the present study is to comparatively analyze two of these *Frog* books to investigate their interchangeability in assessment settings.

**Language Development in Cantonese-English Sequential Bilinguals**

Dual language learners (DLLs) represent a diverse body of individuals—some learn two languages from birth, others learn a second language in early childhood, while still others learn a second language late into their adolescent and adult years. Moreover, dual language learners vary greatly in the languages that they speak, their exposure to those languages, and the domains where their languages are spoken (Pearson, 2007). Given the diversity of dual language learners, bilingual language development is closely tied with the experiences of the individual learner. Current understanding of bilingual language development rejects the notion that a dual language
learner is the sum of two monolingual speakers and instead focuses on the interaction and coexistence of their languages to create a “different but complete language system” (Grosjean, 2010, 75). In this framework, the two or more languages spoken by a bilingual interact with each other during language development and throughout the lifespan.

The present study focuses explicitly on sequential Cantonese-English bilinguals and their morphosyntactic and semantic development. The participants in this study learned Cantonese at home as their first language (L1) and were beginning to learn English as their second language (L2) in the preschool setting. There is limited research documenting the exact nature of language development in this specific population, however there is evidence detailing the use of verbs and tense-aspect markers in these developing bilinguals.

Current research looking at morphological development suggests that Cantonese-English sequential bilinguals may have later acquisition of verbal tense markers when compared to age-matched monolinguals. Rezzonico (2017) analyzed the morphological development of tense marking and the verb ‘be’ in forty-seven sequential Cantonese-English bilingual preschoolers, ages 4-5. Participants were measured for tense marking and ‘be’ accuracy through narrative samples and through elicitation tasks. Results from the study indicated that participants had low accuracy rates for tense morphology (third singular -s, past tense –ed) with an emergent mastery of the verb ‘be’. The participants in the study scored lower than their typically developing, age-matched monolingual peers on the same tasks. These results suggest that Cantonese-English sequential bilinguals may acquire these morphological features at different rates than monolingual peers, for whom ‘be’ and tense agreement tend to arise at the same time. The results from this study corroborate similar findings from other studies investigating tense and verb-marking development in Cantonese-English bilinguals (Nicoladis, 2012). The differences in the
morphological structure of Cantonese and English, namely that Cantonese does not require bound tense-aspect markers on verbs, could provide one explanation for the pattern of development seen in these children.

**Wordless Picture Books and Language Sample Analysis**

Language sample analysis (LSA) is an assessment tool used by researchers and clinicians to analyze the various structural components of a child’s language. One method of eliciting these language samples is through a narrative retell task. In the narrative retell task, children are told a scripted story accompanied by a wordless picture book and asked to retell that story using the picture book as a guide. The story retell task, compared to a simple story telling task, is developmentally appropriate for preschoolers who may have difficulties generating a story on their own. The retell task allows the examiner to provide the initial story structure which can be beneficial when testing young learners. Each narrative sample is then recorded, transcribed, and analyzed by the clinician or researcher. These samples can be analyzed for both macrostructure and microstructure. Macrostructure refers to the larger structural elements of narrative storytelling, including story grammar, sequencing, events, and characters. Alternatively, microstructure refers to specific elements of morphology and syntax including mean length of utterance, number of different words, tense agreement etc. While both domains are of interest, the present research focuses only on microstructure. LSA and narrative retell tasks have been shown to be an accurate indicator of language development in monolingual populations (Bishop & Edmundson, 1987; Paul & Smith, 1993) and are becoming increasingly more common in the assessment of bilingual populations (Rojas & Iglesias, 2009).

Language sample analysis is a beneficial and effective way of assessing bilingual language development that provides an alternative to typical standardized testing. Standardized
tests of language are often normed on populations that are neither culturally nor linguistically diverse. Thus, it is inappropriate to interpret the scores of a child from a culturally or linguistically diverse background in the context of these norms. Cleave et al (2010) compared monolingual English speaking children with developmental language disorders (DLD) to bilingual children with DLD. When administered a standardized test, the monolingual children did significantly better than their bilingual counterparts. However, when assessed using LSA and story retell, there were no significant differences in the performances of the two groups. This study highlights the potential biases of standardized tests, while simultaneously supporting the use of LSA as a less biased method for assessing bilingual children.

While LSA is beneficial in that it is culturally, linguistically, and developmentally appropriate for emergent bilinguals, measures obtained therein have also been shown to significantly predict both language development and reading skills for bilinguals. Miller, Heilmann, Nockerts, Iglesias, Fabiano, & Francis (2006) studied 1,531 Spanish-English bilingual children and their language skills. The study involved eliciting a narrative language sample from the story book *Frog, Where Are You?* and testing the children on reading comprehension in both Spanish and English using the Woodcock Passage Comprehension test. Miller found that performance on the narrative retell task predicted reading comprehension both within and across languages. This predictive power highlights the applicability of LSA in assessing language development and reading outcomes for bilingual children.

The results from Miller et al were further supported and extended by Uchikoshi, Lu & Liu (2017) who analyzed both Spanish-English bilinguals and Cantonese-English bilinguals. In this longitudinal study, they found that narrative quality, as measured in the first grade, was predictive of English reading comprehension as measured a year later in the second grade. The
researchers found this to be true for narratives produced in English as well as narratives produced in the participants’ L1, either Cantonese or Spanish. These results are especially important as this study expands current research relating to Cantonese-English bilinguals, a lesser studied population in the field of speech, language, hearing sciences.

In addition to predicting reading outcomes for bilingual children, LSA can also be used to accurately measure the language development of bilingual children. Muñoz, Gillam, Peña & Gulley-Faehnle (2003) studied Spanish-English bilingual youth in two groups separated by age. They were then asked to produce a narrative sample, again using the wordless picture book *Frog, Where Are You?* The researchers then analyzed these language samples for multiple language measures, including productivity measures (total number of words and total different words), syntactic accuracy, and story grammar. They found that for two of three measures (story grammar and syntactic accuracy) scores improved significantly with age, suggesting that these measures, when produced in the context of narrative samples, are a reliable indicator of language development.

Perhaps more important than measuring language development, evidence suggests that LSA might also be used to differentiate typically developing children from those with language impairment. Squires, Lugo-Neris, Peña, Bedore, Boham & Gillam (2014) analyzed narrative samples in Spanish-English bilingual children who are typically developing (TD) and who have primary language impairment (PLI). These two groups produced narrative samples in kindergarten and again in first grade to track their language output over time. The researchers found that the TD group did significantly better in measures of macro and microstructure in both languages at both time intervals. Additionally, they found that the TD group showed significantly greater improvement over time in their macrostructure scores when compared to the
PLI group. The PLI group did improve in both measures between assessments, but their scores measured in first grade were lower than those of the TD group measured in kindergarten. The findings from this study indicate that LSA and the narrative retell task is sensitive to the differences between TD and PLI groups and that this analysis technique accurately tracks language development over time.

**Measures of Microstructure**

Because narrative samples are incredible robust and can be analyzed on multiple levels, it is essential to determine which measures are most indicative of language ability when analyzing such samples. Due to the limited scope of this project, this thesis will focus explicitly on the microstructure of the language samples, though measures of macrostructure are also frequently and successfully used in the context of LSA.

Several measures of language ability can be analyzed from a single narrative sample. The current research focuses on three such measures: mean length of utterance (MLU), the total number of words (NTW), and the number of different words (NDW). MLU is a productivity measure that indicates the average length (in either words or morphemes) of each intelligible utterance produced by a child. It is well established in the field that MLU increases with age in monolingual populations (Brown, 1973), making it a useful indicator of language development over time. Previously cited studies, including Miller et al. (2006) and Muñoz et al. (2003), indicate that this age-driven increase in MLU is also present in language samples from Spanish-English bilingual populations. In a study comparing narrative sample measures with standardized language scores in Spanish-English bilinguals, Bedore, Peña & Gillam (2010) found that MLU in English was significantly correlated with standardized scores, indicating that it could be useful in differentiating bilingual children with language impairment from TD peers. Similarly, in a
comparison of typically developing and language delayed Spanish-English bilinguals, the language delayed group presented with lower MLU scores than their TD counterparts (Simon-Cereijido & Gutierrez-Clellen, 2009). In Cantonese-speaking monolinguals, MLU has been shown to increase with age (Tse, Kwong, Chan, Li, 2002; Klee, Stokes, Wong, Fletcher, Gavin, 2004) and, when combined with age and D (a measure of linguistic diversity) was shown to accurately differentiate speakers with language impairment from typically developing peers (Klee et al., 2004).

Measures of lexical productivity (the number of total words) and lexical diversity (the number of different words) are also frequently analyzed by researchers and clinicians in language sample analysis. Lexical diversity is a more commonly used index of language development, as it measures the breadth of a child’s lexicon rather than their capacity to produce a certain number of words in an utterance. Lexical diversity has been shown to increase with age in Spanish-English bilingual populations (Muñoz et al., 2003; Ucelli & Paéz, 2007), and has been shown to be lower in language delayed populations when compared to TD bilingual peers (Simon-Cereijido & Gutierrez-Clellen, 2009). However, other studies have indicated that linguistic diversity is not necessarily a significant predictor of language impairment. Jacobson & Walden (2013) considered the use of lexical diversity in differentiating language impaired Spanish-English bilinguals from TD peers. Results from the study did not support NDW as a successful predictor of impairment. Other measures of lexical diversity have been proposed, but also show limited predictive power. In the of Cantonese monolinguals, D (a different measure of lexical diversity) was not found to be a significant predictor of specific language impairment on its own, however when combined with MLU and age was more predictive of language impairment (Klee et al., 2002). Despite the presence of conflicting evidence in the literature,
lexical diversity is still a commonly identified measure in studies looking at narrative production in bilingual populations, making it an important measure to consider in the context of the present study.

A less frequently measured index of language development is morpheme accuracy, as determined by morpheme omission. While many studies consider the overall grammaticality of a narrative (i.e., how many utterances are grammatically correct), fewer studies investigate the morphemic omission patterns present in narrative samples. Although a less pervasive measure in the literature, research does indicate that studying omission patterns can be a valuable predictor of language impairment. Jacobson & Walden (2013) looked at both measures of lexical diversity and morpheme omissions in bilingual Spanish-English populations to determine their sensitivity in predicting primarily language impairment (PLI) in bilingual children. In this study, participants from two groups (TD and PLI) were asked to retell *Frog, Where Are You?* in English and *A Boy, a Dog, and a Frog* in Spanish. Researchers looked at measures of lexical diversity (including NDW) and also looked at omissions of bound grammatical morphemes and single words. They found that NDW was not a significant predictor of language impairment in either language, but that the number of omissions was a statistically significant predictor of language impairment in both languages. These results highlight the potential of using morpheme and word omission scores to assess language impairment in bilingual populations.

In monolingual English language development, research has shown that the order and timeline of morpheme acquisition is relatively fixed. Brown (1973) was the first to study and describe the typical order of morpheme acquisition in monolingual English-speaking children. These morphemes, known as Brown’s morphemes (described in Table 1), are often analyzed as an indicator of typical or atypical language development in monolingual children. Some
morphemes are acquired earlier in development (-ing, in, on) while other morphemes are acquired in later stages of development (contractible copula, contractible auxiliary). While bilingual morphosyntactic development does not parallel morphological development in monolingual English-speaking children (Paradis, 2005), Brown’s morphemes are still of interest as they capture nominal, verbal, and prepositional areas of early morphosyntactic development. Restrepo & Kruth (2000) compared the language development in two Spanish-English bilinguals, one who was typically developing and one who had a language impairment (LI). Among other measures, they looked at the development of Brown’s morphemes and found that the child with LI had greater difficulty with definite articles, the uncontractible auxiliary, third person singular and irregular third person. Research by Guttieréz-Clellen & Simon-Cereijido (2007) also attempted to accurately identify bilinguals with language impairment by using the E-MST, an assessment that analyzes grammatical ability. Within this test, several of Brown’s morphemes are included (past tense –ed, third person singular –s, plural –s, possessive –s, articles, auxiliary forms and copula forms). Results from this study showed that the test was relatively sensitive to language impairment, but not sufficiently specific in ruling out those who do not have LI. Despite this research conducted with Spanish-English populations, it remains unclear if children who learn two languages with distinct typology, such as Cantonese and English, have similar morphological developmental patterns.

Table 1

*Brown’s Morphemes*

<table>
<thead>
<tr>
<th>Morpheme</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present progressive -ing</td>
<td>He is running <strong>ing</strong></td>
</tr>
<tr>
<td>In</td>
<td>He jumps <strong>in</strong> the pool.</td>
</tr>
<tr>
<td>On</td>
<td>The frog is <strong>on</strong> the log.</td>
</tr>
</tbody>
</table>
Peterson

Plural –s
Irregular Past Tense
Possessive –’s
Uncontracted Copula
Articles
Past Tense –ed
Third singular –s
Third person irregular
Uncontracted auxiliary
Contractible copula
Contractible Auxiliary

The frogs are happy.
The frogs went to the pond.
That is the frog’s log.
He is happy.
The frog is happy.
The frog jumped on the log.
The frog jumps.
The boy has a frog.
The frog is jumping.
The frog’s happy.
He’s jumping.

Language and Story Effects on Language Samples in Bilingual Children

Several studies have sought to gauge the effect of language on narrative sample outcomes for bilingual populations. For example, Rezzonico, Goldberg, Mak, Yap, Milburn, Balletti & Girolametto (2016) analyzed the narrative macro and microstructure of 47 bilingual Cantonese-English speaking children ages four and five utilizing the Mercer Mayer wordless picture book Frog, Where Are You? (Mayer, 1969). The purpose of this study was to determine if the language of assessment (Cantonese or English) affected the language measures that were being tested. The results of this study found that the microstructure of the narratives (including MLU and vocabulary diversity measures) was similar cross-linguistically and that there were no statistically significant differences between the two languages. In analyzing the macrostructure, the researchers found that participants scored higher for their story grammar scores in English than they did in Cantonese. These results support the use of wordless picture books as a linguistically appropriate assessment of bilingual Cantonese-English speaking children.

Similarly, Fiestas & Peña (2004) also analyzed narrative language samples in Spanish-English bilingual children between the ages of four and seven. The narrative retell procedure in
this study was again based on the book *Frog, Where Are You?* (Mayer, 1969). The researchers in this study looked at the story complexity and productivity of language samples produced in each language. The results of this study showed similarities in measures of productivity (MLU and lexical diversity) across both languages, and similar levels of story complexity in both languages. The only significant difference observed was the inclusion of specific narrative elements in the retells, with Spanish having more initiating events and English narratives containing more consequences. Despite this difference, this study further supports the use of a wordless picture book task in analyzing and assessing language ability in bilingual populations, specifically in analyzing microstructure.

While there have been several studies looking at the effect of language on narrative samples, comparatively few have considered the effect of the wordless picture book itself on the samples. Currently, the only study looking at the effect of book choice on language measures was conducted by Heilmann, Rojas, Iglesias & Miller (2016). This study compared five different *Frog* books with Spanish-English bilingual participants. The five books compared were *A Boy, a Dog, and A Frog* (Mayer, 1967), *A Boy, A Dog, A Frog, and a Friend* (Mayer, 1971), *Frog Goes to Dinner* (Mayer, 1974), *Frog On His Own* (Mayer, 1975), and *One Frog Too Many* (Mayer, 1975). Participants were assigned to tell one of the five stories in both Spanish and English, and their narrative samples were compared for differences in mean length of utterance, number of total utterances, number of different words, and narrative story structure. Heilmann et al. found no clinically significant differences between the books, though there were statistically significant differences in the number of different words (NDW) used in two of the five stories. These differences were attributed to structural differences between the stories, including the number of episodes and characters in the two books. Even with differences in NDW, the results of this
study support the use of these five books interchangeably in a clinical setting for assessing language in Spanish-English bilingual children.

The current thesis seeks to build on the research conducted by Heilmann et al. (2016) by considering a different bilingual population, namely Cantonese-English bilinguals, and by comparing two different frog books: *Frog On His Own* (Mayer, 1975) and *Frog, Where Are You?* (Mayer, 1969). Through a comparative analysis of the language samples elicited from these commonly used wordless picture books, this thesis seeks to further the findings of the Heilmann study and support the clinical use of these picture books in assessing Cantonese-English bilingual children.

**The Present Study**

The purpose of the present study is to analyze and compare the language samples in English elicited from two commonly used wordless picture books, *Frog On His Own* (Mayer, 1975) and *Frog, Where Are You?* (Mayer, 1969). It is worthwhile to note that this study only considers the English narrative samples produced by the participants, not the narrative samples produced in Cantonese. The reason for this decision was two-fold. First, the participants in this study were learning English as their second language and already had a strong developmental foundation in Cantonese. Thus, the development of their second language is of more concern as it is emergent rather than established. Second, the present study looks extensively at the morphological features (including Brown’s morphemes) present in the language samples. Given that these features are unique to English, we only analyzed the English language samples.

Through an analysis of the language samples produced using the two books, this thesis will attempt to answer the following questions:

1. Does morpheme accuracy vary by book?
2. Are language measures (MLU, NTW, NDW) similar between the books?

3. Do both books provide similar opportunities to produce specific morphemes?

In addressing the first research question, I will analyze omission patterns and accuracy rates to determine morpheme distribution patterns by book. In answering the second question, I will analyze several measures of narrative microstructure including mean length of utterance (MLU) and measures of lexical diversity including the number of total words (NTW) as well as the number of different words (NDW). To assess the final research question, I will analyze the occurrences, opportunities, and omission patterns for morphemes as produced in the language samples. Finally, I will consider two individual case studies from the task to highlight the potential differences that might exist between the same individual’s two performances.

It is expected that language measures will be similar between the books (consistent with the findings in Heilmann et al. (2016)), that there will be no significant differences in morpheme accuracy, and that opportunities to produce specific morphemes will be similar between both books. This study will test the reliability of using these books interchangeably to elicit language samples and will provide further evidence supporting the use of the *Frog* books in assessing Cantonese-English bilingual children using narrative retell tasks and language sample analysis.

**Methods**

**Participants**

This sample consisted of 31 bilingual Cantonese-English speaking participants (Mean age = 4;8, SD = 0;10), recruited from a Head Start program in San Francisco, CA. These participants were sequential bilinguals and were exposed to Cantonese as their L1 from birth at home and started to learn English as their L2 in preschool settings. According to parent and teacher reports,
these children had stronger Cantonese skills at the time of testing. Cantonese was the primary language used at home. Both Cantonese and English were used during classroom activities. Parents and teachers had no concerns about these children’s language development. Each participant produced two narrative retells in both languages: one for *Frog, Where Are You?* and one for *Frog on His Own*. The samples were previously collected by Dr. Pui Fong Kan in 2011 as a part of a larger study.

**Story Retell Procedure**

The participants were tested using a story retell task. Each child was initially read the story by the examiner, using a developmentally and culturally appropriate script. The child was then asked to retell the story to the examiner, using the picture book as a guide. The order in which the stories were told was random. Children were tested in both English and Cantonese, and the order of the language was counterbalanced. For the purposes of this thesis, only the English samples were analyzed. The narrative samples produced by each child were recorded, transcribed, and then analyzed using the Systematic Analysis of Language Transcript (SALT) Version 18.

**Sample Selection**

The samples were chosen based on the following criteria: 1) that the participant produced a language sample for both books, 2) that the majority of the participant’s language sample was produced in English, 3) that the language sample included at least one instance of morphosyntactic structure (i.e., that the sample contained more than just one word utterances). Initially there were over one hundred language samples, which were reduced to just thirty-one after applying these criteria.
Transcription

Each story retell audio recording was transcribed according to SALT transcription conventions. The procedure for coding in SALT requires that one first identify utterance boundaries, defined as the end of independent clauses. Once utterance boundaries are established, one must then account for all bound grammatical morphemes. Morpheme boundaries are coded with a ‘/’ symbol, followed by the morpheme. In the case of a final -s, which can serve many purposes, specific morpheme codes are used (‘3S’ for third singular -s, ‘s’ for plural -s, and ‘z’ for possessive -s). Once morpheme boundaries are coded, all repetitions and filler words (um, uh etc.) are mazed, or excluded from analysis. Additionally, phrases which the child understands as one lexical entry (e.g., uh_oh or frog_where_are_you) are consolidated with underscores and are considered to be one word for the purpose of coding. Finally, all unintelligible words or morphemes are coded using the symbol ‘XXX’. Utterances which contain any unintelligible sounds or words are not counted in the analysis, even if the remainder of the utterance is intelligible. A sample transcription is shown below:

Utterance: “The frog um walked by the ponds and he uhh liked he liked the water.”

SALT Transcription: The frog (um) walk/ed by the pond/s

And he (uh) lik/ed (he liked) the water.

Once transcribed, the language samples were analyzed using SALT, a language analysis software specifically designed to provide quantitative data for a given language sample. The software provides an analysis of the sample, which includes the number of utterances, the number of morphemes, mean length of utterance, and lexical distributions within the sample.

For each language sample, I analyzed mean length of utterance (MLU), number of total words (NTW) and number of different words (NDW). Beyond these measures, I also counted the
presence of Brown’s morphemes in each sample and omissions of those same morphemes. As mentioned before, only complete, intelligible utterances were considered in these measures, following SALT guidelines for sample analysis.

**Brown’s Morphemes**

The morphemes analyzed in this study are the fourteen morphemes, first described by Brown (1973), that are acquired sequentially in monolingual English language development. These morphemes are described previously in Table 1. In order to account for the fact that several of Brown’s morphemes are free morphemes (e.g., in and on), we analyzed both the morphemic output (to account for bound morphemes) and the lexical distributions (to account for free morphemes) generated by the SALT software. Brown’s morphemes were counted using the software output and crosschecked with the original transcriptions.

All instances of the verb ‘to be’ were analyzed individually in the transcript to determine their function (copula or auxiliary). The verb ‘to be’ is considered a copula when it is the only verb in a sentence (e.g., he is happy) and considered an auxiliary when it is secondary to another verb (e.g., he is running). An uncontracted copula or auxiliary occurs when the verb ‘to be’ occurs in its full form, that is, not contracted (e.g., he is happy). The opposite, the contractible auxiliary and copula, occurs when the verb ‘to be’ is contracted with the subject (e.g., he’s happy or he’s running). For irregular past tense forms, any past tense verb that does not end in –ed was counted (e.g., saw, ate, said, went). The irregular third person verbs are “to have” (has) and “to do” (does). As mentioned earlier, only Brown’s morphemes which appeared in fully intelligible utterances were included in the final counts.
Opportunities & Omissions

In order to determine a measure of accuracy, any omission of Brown’s morphemes were recorded. For the purpose of analysis, an omission is defined as an obligatory context where the morpheme was omitted, or not present in the language sample. An obligatory context is a context in which a morpheme would necessarily appear in typical adult language for Standard American English. For example, if the participant produced the sentence “Frog go to the park”, the third singular (3S) morpheme is omitted on the word ‘go’ because it is obligatory in this context as determined by the conventions of Standard American English. In this example, we would count one omission of the 3S morpheme. Omissions were included in the original transcriptions (denoted with an * in SALT) and were crosschecked at the time of coding to ensure the accuracy and reliability of the transcriptions.

In addition to coding omissions, we were also interested in the number of opportunities that each book provided to produce specific morphemes, to determine the potential similarities and differences therein. An opportunity was coded as either: 1) the correct presence of a morpheme in an obligatory context (the correctly used morphemes as they appeared in the language samples), or 2) the omission of a morpheme in an obligatory context (where a morpheme should have appeared, but did not). All instances of correct morphemes and morpheme omissions (across all participants) were summed to arrive at the total number of opportunities to produce a given morpheme in each book.

Other Errors

With respect to morphemic errors, if the child produced the wrong morpheme instead of the correct one (i.e., getted instead of the irregular past tense ‘got’), then the regular past tense morpheme (-ed) was not counted and an omission was counted for the irregular past tense
morpheme ‘got’. If instead, an extra morpheme appeared that was incorrect but did not replace a
different morpheme, this morpheme was simply not counted in the total number of correct
morphemes. For example, if the child produced the utterance “they eat/3S”, the 3S morpheme is
an error but does not replace a required morpheme, so the 3S was not included in total counts.

**Measures of Productivity and Lexical Diversity**

In addition to the presence and omission of Brown morphemes, we were also concerned
with the participants’ mean length of utterance (both in words and in morphemes), total number
of words, and number of different words for each sample. The MLU is generated by averaging
the number of morphemes (including free morphemes and bound grammatical morphemes)
produced in each utterance for the entire language sample. Unintelligible or abandoned
utterances are not included in this average. The total number of words (NTW) is a measure
which counts all words that appear in the transcription, including repeated words, only excluding
those words which appear in unintelligible or abandoned utterances. The number of different
words (NDW) is a word count which counts each word only when it first appears, and then does
not count any further repetitions of the same word. The SALT software was used to calculate
MLU, NTW and NDW values.

**Results**

**Does Morpheme Accuracy Vary by Book?**

Tables 2a and 2b respectively provide a statistical analysis of morpheme accuracy in each
book. The percentages in this table reflect the number of correct uses of a morpheme divided by
the total number of opportunities available to use that morpheme across all language samples for
each book. The range of morpheme accuracy for *Frog On His Own* is 76.58%, with the
morpheme ‘on’ representing the highest level of accuracy (97.22%) and the third singular ‘-s’
representing the lowest (20.64%). Comparatively, the range for *Frog Where Are You* was 89.6%, with the highest morpheme accuracy being the contractible auxiliary (100%, though with a relatively low sample size of six) followed by the articles (98.18%), and the lowest again being the third singular ‘-s’ (8.58%). As shown in table 2b, there are almost no significant differences in levels of accuracy by morpheme between the two books. The only exception is the possessive ‘-s’ morpheme which shows higher levels of accuracy in *Frog On His Own* than it does in *Frog, Where Are You?* (p=0.01). For the 3rd irregular morpheme, the contractible copula, and the contractible auxiliary, the sample sizes were too small to provide conclusive statistical analysis as to their significance.

Table 2a

*Morpheme accuracy across both books*

<table>
<thead>
<tr>
<th>Morpheme</th>
<th>Own (N)</th>
<th>Where (N)</th>
<th>Own Mean (%)</th>
<th>Own SD (%)</th>
<th>Where Mean (%)</th>
<th>Where SD (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ing</td>
<td>27</td>
<td>23</td>
<td>82.10</td>
<td>28.91</td>
<td>71.86</td>
<td>32.67</td>
</tr>
<tr>
<td>In</td>
<td>20</td>
<td>17</td>
<td>91.50</td>
<td>24.55</td>
<td>94.16</td>
<td>24.09</td>
</tr>
<tr>
<td>On</td>
<td>18</td>
<td>13</td>
<td>97.22</td>
<td>11.79</td>
<td>84.62</td>
<td>37.55</td>
</tr>
<tr>
<td>Plural -s</td>
<td>15</td>
<td>20</td>
<td>93.33</td>
<td>17.59</td>
<td>80.00</td>
<td>34.45</td>
</tr>
<tr>
<td>Irregular Past</td>
<td>20</td>
<td>17</td>
<td>89.17</td>
<td>27.72</td>
<td>95.24</td>
<td>11.29</td>
</tr>
<tr>
<td>Possessive -s</td>
<td>9</td>
<td>5</td>
<td>50.00</td>
<td>50.00</td>
<td>56.67</td>
<td>43.46</td>
</tr>
<tr>
<td>Uncontracted Copula</td>
<td>18</td>
<td>24</td>
<td>62.04</td>
<td>44.21</td>
<td>65.14</td>
<td>43.05</td>
</tr>
<tr>
<td>Articles</td>
<td>30</td>
<td>31</td>
<td>93.85</td>
<td>18.36</td>
<td>98.18</td>
<td>5.04</td>
</tr>
<tr>
<td>Past Tense -ed</td>
<td>19</td>
<td>8</td>
<td>73.25</td>
<td>34.75</td>
<td>92.50</td>
<td>21.21</td>
</tr>
<tr>
<td>Third singular -s</td>
<td>29</td>
<td>26</td>
<td>20.64</td>
<td>30.73</td>
<td>8.58</td>
<td>11.90</td>
</tr>
<tr>
<td>Third Irregular</td>
<td>9</td>
<td>4</td>
<td>44.44</td>
<td>52.70</td>
<td>25.00</td>
<td>50.00</td>
</tr>
<tr>
<td>Uncontracted Auxiliary</td>
<td>23</td>
<td>18</td>
<td>44.06</td>
<td>42.98</td>
<td>34.72</td>
<td>44.67</td>
</tr>
<tr>
<td>Contractible Copula</td>
<td>11</td>
<td>13</td>
<td>90.91</td>
<td>30.15</td>
<td>89.74</td>
<td>28.50</td>
</tr>
<tr>
<td>Contractible Auxiliary</td>
<td>6</td>
<td>6</td>
<td>88.33</td>
<td>20.41</td>
<td>100.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table 2b

*Morpheme accuracy across both books cont.*
### Morphological Omissions across Books

Table 3 and Figure 1 provide an analysis of the omission patterns by morpheme in each book. The values presented in Table 3 reflect the total number of morpheme omissions (in an obligatory context) across all participants for each morpheme in each book, as well as a statistical analysis of the omission patterns between books. The omission patterns were largely similar between books, although the past tense marker (-ed) was omitted more frequently in *Frog On His Own* (n=15, p=.033). There were no other significant differences between the two book with respect to omissions. Notably, for both books, we see the highest concentration of omissions with the third person singular ‘-s’, followed distantly by the uncontracted auxiliary. Notably, across both books, omissions tend to appear more frequently with verbal morphemes (third singular –s, uncontractible copula, uncontractible auxiliary) when compared to nominal morphemes such as plural –s, possessive -s or articles.
Table 3

*Number of Morpheme Omissions by Book*

<table>
<thead>
<tr>
<th>Morpheme</th>
<th>Own (n)</th>
<th>Where (n)</th>
<th>DF</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ing</td>
<td>12</td>
<td>18</td>
<td>1</td>
<td>.280</td>
</tr>
<tr>
<td>In</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>.325</td>
</tr>
<tr>
<td>On</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>.572</td>
</tr>
<tr>
<td>Plural -s</td>
<td>2</td>
<td>7</td>
<td>1</td>
<td>.096</td>
</tr>
<tr>
<td>Irregular Past</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>.712</td>
</tr>
<tr>
<td>Possessive -s</td>
<td>8</td>
<td>4</td>
<td>1</td>
<td>.354</td>
</tr>
<tr>
<td>Uncontracted Copula</td>
<td>20</td>
<td>19</td>
<td>1</td>
<td>.813</td>
</tr>
<tr>
<td>Articles</td>
<td>16</td>
<td>9</td>
<td>1</td>
<td>.374</td>
</tr>
<tr>
<td>Past Tense -ed</td>
<td>15</td>
<td>3</td>
<td>1</td>
<td>.033*</td>
</tr>
<tr>
<td>Third singular -s</td>
<td>219</td>
<td>174</td>
<td>1</td>
<td>.538</td>
</tr>
<tr>
<td>Third Irregular</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>.264</td>
</tr>
<tr>
<td>Uncontracted Auxiliary</td>
<td>25</td>
<td>25</td>
<td>1</td>
<td>.899</td>
</tr>
<tr>
<td>Contractible Copula</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>.712</td>
</tr>
<tr>
<td>Contractible Auxiliary</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>.184</td>
</tr>
</tbody>
</table>

*p<.05

Figure 1

*Number of Omissions by Morpheme and Book*

![Omissions by Morpheme and Book](image-url)
Are Language Measures (MLU, NTW, NDW) Similar Between the Books?

One of the primary research questions guiding this study is the comparability of the language outcomes between books. Table 4 provides a statistical analysis of the average mean length of utterance and lexical diversity for each book. MLU is, on average, higher for narrative retellings of *Frog On His Own* (4.59) than for *Frog, Where Are You?* (4.14). However, this difference between the books is not statistically significant (p=.069). Similarly, NTW and NDW were both higher on average in *Frog On His Own* (119.84 and 47.68, respectively) than in *Frog, Where Are You?* (99.65 and 41.32, respectively). Again, the differences in NTW and NDW were not found to be statistically significant (p=.128; p=.069).

Table 4

*The mean length of utterance (MLU) & lexical diversity (NTW and NDW) for both books*

<table>
<thead>
<tr>
<th></th>
<th>Own Mean</th>
<th>Own SD</th>
<th>Where Mean</th>
<th>Where SD</th>
<th>F</th>
<th>DF</th>
<th>DF 2</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLU</td>
<td>4.59</td>
<td>1.50</td>
<td>4.14</td>
<td>1.44</td>
<td>3.56</td>
<td>1</td>
<td>30</td>
<td>0.069</td>
</tr>
<tr>
<td>NTW</td>
<td>119.84</td>
<td>79.92</td>
<td>99.65</td>
<td>54.48</td>
<td>2.44</td>
<td>1</td>
<td>30</td>
<td>0.128</td>
</tr>
<tr>
<td>NDW</td>
<td>47.68</td>
<td>21.6</td>
<td>41.32</td>
<td>16.89</td>
<td>3.57</td>
<td>1</td>
<td>30</td>
<td>0.069</td>
</tr>
</tbody>
</table>

Do both books provide similar opportunities to produce specific morphemes?

Also of great consideration in this study is the distribution of opportunities across the books. The number of opportunities indicates the level of similarity between the stories with respect to morphosyntactic complexity. Table 5, represented visually in Figure 2, shows the distribution of the morpheme opportunities in each book. The total number of opportunities was calculated by summing the occurrences of each morpheme (those that appeared in obligatory contexts in the samples) and the omissions of each morpheme (those that did not appear, but should have appeared, in obligatory contexts in the samples) across all participants. From Figure 2, it is apparent that there is a high number of opportunities in both books to utilize articles and
the third singular ‘-s’. With the exception of plural s, the uncontracted copula, and the contractible copula, *Frog On His Own* provided more opportunities to produce all other morphemes than *Frog, Where Are You?* These findings are consistent with higher MLU scores for *Frog On His Own*.

A statistical analysis of the number of opportunities present in each book is also outlined in Table 4. For the majority of the morphemes, there were no statistically significant differences between the two books. However, the number of opportunities for articles (n=605,  p=.048) and for the past tense marker (–ed) (n=53,  p=.013) were significantly higher in *Frog On His Own*. Cohen’s effect size value for articles (d=0.42) and for past tense –ed (d=0.16) indicate a low practical significance for both morphemes.

Table 5

*Number of Morpheme Opportunities by Book*

<table>
<thead>
<tr>
<th>Morpheme</th>
<th>Own (n)</th>
<th>Where (n)</th>
<th>Own Mean (SD)</th>
<th>Where Mean (SD)</th>
<th>DF</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ing</td>
<td>89</td>
<td>69</td>
<td>2.87 (2.33)</td>
<td>2.23 (1.99)</td>
<td>1</td>
<td>.077</td>
</tr>
<tr>
<td>In</td>
<td>48</td>
<td>45</td>
<td>1.55 (1.67)</td>
<td>1.45 (2.03)</td>
<td>1</td>
<td>.822</td>
</tr>
<tr>
<td>On</td>
<td>38</td>
<td>25</td>
<td>1.23 (1.5)</td>
<td>0.81 (1.38)</td>
<td>1</td>
<td>.196</td>
</tr>
<tr>
<td>Plural -s</td>
<td>31</td>
<td>45</td>
<td>1.00 (1.39)</td>
<td>1.45 (1.52)</td>
<td>1</td>
<td>.194</td>
</tr>
<tr>
<td>Irregular Past</td>
<td>48</td>
<td>44</td>
<td>1.55 (2.28)</td>
<td>1.42 (1.72)</td>
<td>1</td>
<td>.804</td>
</tr>
<tr>
<td>Possessive -s</td>
<td>13</td>
<td>8</td>
<td>.042 (0.72)</td>
<td>0.26 (0.68)</td>
<td>1</td>
<td>.305</td>
</tr>
<tr>
<td>Uncontracted Copula</td>
<td>51</td>
<td>64</td>
<td>1.65 (2.09)</td>
<td>2.06 (2.18)</td>
<td>1</td>
<td>.405</td>
</tr>
<tr>
<td>Articles</td>
<td>605</td>
<td>437</td>
<td>19.52 (13.84)</td>
<td>14.10 (11.75)</td>
<td>1</td>
<td>.048*</td>
</tr>
<tr>
<td>Past Tense -ed</td>
<td>53</td>
<td>20</td>
<td>1.71 (1.81)</td>
<td>0.65 (1.68)</td>
<td>1</td>
<td>.013*</td>
</tr>
<tr>
<td>Third singular -s</td>
<td>256</td>
<td>189</td>
<td>7.61 (5.14)</td>
<td>6.10 (4.86)</td>
<td>1</td>
<td>.249</td>
</tr>
<tr>
<td>Third Irregular</td>
<td>10</td>
<td>4</td>
<td>0.97 (3.59)</td>
<td>0.13 (0.34)</td>
<td>1</td>
<td>.208</td>
</tr>
<tr>
<td>Uncontracted Auxiliary</td>
<td>44</td>
<td>40</td>
<td>1.42 (1.26)</td>
<td>1.29 (1.35)</td>
<td>1</td>
<td>.677</td>
</tr>
<tr>
<td>Contractible Copula</td>
<td>17</td>
<td>17</td>
<td>0.55 (0.81)</td>
<td>0.55 (0.77)</td>
<td>1</td>
<td>1.00</td>
</tr>
<tr>
<td>Contractible Auxiliary</td>
<td>16</td>
<td>14</td>
<td>0.52 (1.15)</td>
<td>0.45 (1.12)</td>
<td>1</td>
<td>.745</td>
</tr>
</tbody>
</table>

*p<.05
While analyzing these books for larger statistical differences and similarities is important, I also wanted to highlight the importance of how an individual performance might differ when given the different stories. Speech language therapy is primarily focused on improving outcomes for individuals, making individual performance relevant to this thesis. In order to look at these individuals in an unbiased way, I randomly selected two participants and analyzed the differences in their results for each book for their MLU, NTW and NDW scores. Table 6 shows the results for Participant 1, and Table 7 the results for Participant 2.
Table 6

*Individual Productivity Scores for Participant One*

<table>
<thead>
<tr>
<th></th>
<th><em>Frog On His Own</em></th>
<th><em>Frog, Where Are You?</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>MLU</td>
<td>4.56</td>
<td>5.33</td>
</tr>
<tr>
<td>NTW</td>
<td>47</td>
<td>140</td>
</tr>
<tr>
<td>NDW</td>
<td>23</td>
<td>56</td>
</tr>
</tbody>
</table>

Table 7

*Individual Productivity Scores for Participant Two*

<table>
<thead>
<tr>
<th></th>
<th><em>Frog On His Own</em></th>
<th><em>Frog, Where Are You?</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>MLU</td>
<td>5.81</td>
<td>3.54</td>
</tr>
<tr>
<td>NTW</td>
<td>166</td>
<td>132</td>
</tr>
<tr>
<td>NDW</td>
<td>69</td>
<td>53</td>
</tr>
</tbody>
</table>

As described in Table 6, participant one had higher scores in all measures for *Frog, Where Are You*. Perhaps most notable are the differences in lexical diversity between these two samples, with over three times as many words being used in their *Frog, Where Are You* sample (NTW=140) when compared to their *Frog On His Own* sample (NTW=47). Similarly, the different words used in *Frog, Where Are You* (NDW=56) are almost twice that of those used in *Frog On His Own* (NDW=23). In all measures, it appears that participant one performed better when tested with the story *Frog, Where Are You*.

Contrastingly, participant two scored higher in all measures for *Frog On His Own*. Most strikingly in this example is that their MLU score is much higher for *Frog On His Own* (MLU=5.81) when compared to *Frog, Where Are You* (MLU=3.54). Measures of lexical diversity are closer in this case study but are still higher in *Frog On His Own* (NTW=166,
NDW=69) than they are in *Frog, Where Are You?* (NTW=132, NDW=53). For all measures, participant two performed better when tested with *Frog On His Own*.

**Discussion**

The present study provides a comparative analysis of two Mercer Mayer *Frog* Books to determine their interchangeability in a clinical- or research-based narrative assessment. The participants were 31 typically developing Cantonese-English sequential bilingual preschool children who produced narrative retells for *Frog On His Own* and *Frog, Where Are You?* The language samples were analyzed for measures of language productivity, morpheme distribution, and morpheme accuracy. The results from this analysis show limited differences between the books, including higher accuracy for the possessive -s morpheme, increased opportunities for the production of the regular past tense (-ed) and articles (the/a), and increased omissions of the regular past tense marker (-ed) in *Frog On His Own*. Despite moderate differences in the morphemic production across books, the findings of this study support the use of either book in clinical- or research-based narrative assessment of Cantonese-English bilingual children.

**Morpheme Accuracy across Books**

An analysis of morpheme accuracy revealed almost no significant differences between the books. These findings are consistent with our initial prediction that the participants would have similar language samples in both books. The results of the present study indicate that these accuracy measures are relatively stable when tested across books. Previous research indicates that morpheme accuracy is an indicator of language development (Jacobson & Walden, 2013). The results from the present study suggest that children demonstrate similar morpheme accuracy when tested twice with different books, lending support to the stability of the measure.
Despite the typological differences of Cantonese and English, there still remains the possibility of syntactic and semantic transfer between the languages. Language transfer is the process by which a syntactic feature or lexical meaning in one language influences the production of a similar or related linguistic item in the second language. Language transfer is well documented in bilingual populations (Scheele, Leseman, Mayo 2010; Melby-Lervag 2011), with the dominant first language often influencing the emergent second language. Although Cantonese and English share neither a typological nor genetic classification, there is still evidence that transfer occurs between the two languages (Yip & Matthews, 2000). In the present study, the difference in the accurate production of the possessive –s morpheme could result from linguistic transfer.

The present study found that the possessive -s morpheme was used with higher rates of accuracy in *Frog On His Own* than in *Frog, Where Are You?* While both books provide a similar number of opportunities to produce the morpheme, children in this study were able to produce the morpheme with higher rates of accuracy in *Frog On His Own*. This difference is possibly the result of a semantic-morphological transfer between Cantonese and English. In one scene in *Frog On His Own*, there is a baby in a stroller. The Cantonese word for stroller is a compound word that roughly translated into English is a “baby-car”. Instead of producing the English lexical entry “stroller”, some children in the study instead produced the Cantonese influenced phrase “baby’s car”. It is possible that the semantic understanding of a stroller in Cantonese enabled the children to correctly produce the possessive –s morpheme on the word baby. As no such similar context existed in *Frog, Where Are You?* this linguistic interaction could be responsible for the increased accuracy rates for the morpheme in *Frog On His Own*. 
In this study, omission patterns were also documented as an indicator of inaccurate morpheme production. The highest instances of omissions occurred for the third person singular \(-s\) and the uncontractible copula (‘be’). These results are consistent with the findings of Rezzonico et al. (2017) which revealed that sequential Cantonese-English bilinguals had low accuracy scores when producing morphological tense markers, and only an emergent mastery of the verb ‘be’. These findings, as well as our own, make sense in the context of the structural differences that exist between Cantonese and English. The participants in the present study were beginning to learn English and therefore were only just encountering bound tense and subject markings. This typological difference between Cantonese and English likely contributed to the high rates of omissions for the third person singular \(-s\).

**Productivity Measures across Books**

Analysis of the narrative samples revealed that productivity measures (MLU, NTW, NDW) do not vary significantly by book. These results are consistent with our prediction that the language samples would be structurally similar across books and are also supported by the findings of Heilmann et al (2016). While not significant, scores were, on average, higher for all measures in *Frog On His Own*. MLU scores for *Frog On His Own* were on the threshold of being significantly higher than those for *Frog, Where Are You?*, \(p=0.056\). A possible explanation for these elevated MLU scores could lie in the structure of the story itself. Perhaps the story contains more actions or more characters, resulting in elevated MLU scores for those narratives. Further analysis of the book components (characters, actions, objects, sequences etc.) could provide information that would illuminate the nature of this difference. This difference, while interesting, does not impact the efficacy of these books in clinical or research based assessment.
The individual case studies highlighted the differences that can emerge in language samples even when controlling for the participant. While the differences between the books are not necessarily statistically significant when analyzed across all participants, individuals can be tested twice with different language sample outcomes. We see that for both randomly selected participants, each performed much better on one narrative sample than on the other. Participant One scored higher on all measures of productivity (MLU, NTW, NDW) for *Frog On His Own* whereas Participant Two scored higher on all these same measures for *Frog, Where Are You?* Notably, Participant One had a very low score for NTW (47) for *Frog On His Own*. After reanalyzing their language sample, I have concluded that this figure is the result of many utterances containing unintelligible words. The presence of unintelligible words excludes those utterances from analysis, resulting in a relatively low NTW score. For the entire language sample, the NTW score was much higher at 125, which is comparable to their NTW score for *Frog Where Are You?* (NTW = 140). Even with closer analysis of the language samples, it remains unclear why participants did better on one book compared to the other. It is possible that motivation during testing or other external factors could have contributed to the difference in performances. Regardless, these results stress the importance of using a variety of testing measures when analyzing individuals. As of yet, it cannot be said that there is a single gold standard approach for accurately identifying bilingual children with developmental language disorders, so the use of multiple assessment measures is highly recommended in clinical settings.

**Morpheme Distribution across Books**

Between the books, the type and number of morphemes that appear are fairly consistent. However, Figure 2 shows that there are more opportunities to produce articles and the third singular ‘s’ when compared to other morphemes across both books. Higher instances of the third
singular morpheme can be explained because the story is comprised of many simple actions performed by one character. Given this narrative structure, many actions in the retell have a third person singular subject (he, she, it, or a singular noun like ‘frog’). It should be noted that the majority of opportunities to produce the third person singular morpheme were omissions as opposed to correct productions (85.5% for *Frog On His Own* and 92.1% for *Frog Where Are You*?). Contrastingly, the high presence of articles were primarily correct productions with only 2.6% of *Frog On His Own* productions and 2.1% of *Frog, Where Are You?* productions being omissions.

Statistical analysis of the number of opportunities across books revealed a significantly higher number of opportunities to produce articles and the past tense marker (-ed) in *Frog On His Own*. The effect size of these differences, however, is relatively small. The scripts are initially told in past tense for both books, so the original tense of narration does not impact the retells disproportionately in one book or the other. The elevated presence of the past tense marker in *Frog On His Own* indicates that the structure of the story may be more conducive to past tense narration. The story *Frog, Where Are You?* follows a boy as he actively searches for his pet frog, a narrative structure which may implicitly call for the use of the present tense. Comparatively, *Frog On His Own* is a story about a frog and his mischievous day in the park, which is easily told in either tense. The exact nature of the difference remains unclear, but the possibility for tense difference between the stories should be noted in clinical assessment.

An analysis of the differences in morpheme omissions between the books revealed a significantly higher number of omissions for the past tense marker (-ed) in *Frog On His Own*. This finding is consistent with the elevated number of opportunities in *Frog On His Own* to use the same marking. If a child was narrating in the past tense and struggled to produce regular
verbal endings, this could lead to greater instances of omissions for the past tense marker. Additional studies evaluating the tense choice and tense production of narrative assessments could prove beneficial in illuminating the nature of this difference.

The elevated number of articles suggests that *Frog On His Own* may have more nouns than *Frog, Where Are You?* Analyzing the scripts for both books, *Frog On His Own* has 90 articles (72 definite, 18 indefinite) whereas *Frog, Where Are You?* has 97 articles (77 definite, 20 indefinite). These numbers suggest that the stories are roughly comparable with respect to article distribution, with slightly higher numbers for *Frog, Where Are You?* An analysis of the core nouns in each story reveals that the nouns in *Frog On His Own* are more immediately recognizable and common in a preschool lexicon (a flower, a bumblebee, a boat, a baby, a bottle) when compared to the core nouns in *Frog, Where Are You?* (a jar, a beehive, an owl, a deer, a cliff). Children may produce more objects (and therefore more articles) if they have higher familiarity with the nouns used in the *Frog On His Own*. While seemingly innocuous in the larger context of language sample analysis, elevated article opportunities could result in higher MLUs and NTWs, impacting a child’s overall performance on the task.

**Conclusion**

This study provides a comparison of the popular Mercer Mayer *Frog* books as a clinical tool for Cantonese-English bilingual assessment. The books are compared across multiple measures of narrative microstructure including MLU, lexical diversity, morpheme accuracy, and morpheme distribution. Measures of productivity (MLU and NTW) are similar between both books, as is lexical diversity (NDW). Patterns of morpheme accuracy were also found to be similar across the books, with the only exception being the possessive -s morpheme, which was
used with significantly higher accuracy rates in *Frog On His Own*. Overall, the number of opportunities to produce Brown’s morphemes were similar across the books, with higher numbers of opportunities to produce articles and regular past tense marker –ed in *Frog On His Own*. Omission patterns revealed a high number of omissions for the third singular –s when compared to other morphemes, and *Frog On His Own* showed significantly higher instances of omission for the past tense marker (-ed) when compared to *Frog, Where Are You?* Results are consistent with current research supporting the interchangeability of the books in the *Frog* series (Heilmann et al., 2016).

This study is limited in that it focuses explicitly on the microstructure of narrative language samples. Further research is needed to address the similarities and differences between the books with respect to macrostructure. Additionally, this study is limited in that it focuses explicitly on the L2 (English) of the participants. Future studies could consider the interaction of Cantonese and English in the narrative task by assessing samples produced in both languages. These studies could provide further support for using LSA to assess language development for sequential Cantonese-English bilingual populations.

**Clinical Implications**

Individual case studies reveal that, while the books have limited significant differences, individual narrative samples can vary. Given this variability, utilizing multiple methods of assessment in conjunction with narrative production is recommended. Current research suggests that dynamic assessment following a test-teach-retest model utilizing narrative retells could be effective in identifying bilingual children with language impairment (Petersen et al., 2017). Future studies are needed to examine whether dynamic assessment technique could enhance
bilinguals’ morpheme and narrative production and provide a more accurate and holistic diagnosis of their language development.

The results from the present study suggest that *Frog On His Own* may offer more opportunities to produce certain morphemes (articles and –ed), may allow for more accurate productions of the possessive –s, and may have higher omissions of the past tense marker –ed. Despite these limited differences at the level of individual morphemes, children produced similar language samples regardless of book when assessing MLU and vocabulary measures. The findings of this study support the use of both books, though clinicians should understand the possibility for elevated scores when using *Frog On His Own*. The results from this study, combined with results from Rezzonico et al. (2016), provide evidence for using narrative *Frog* book samples to assess language in Cantonese-English bilingual children in both clinical and research settings. Taken together with studies assessing Spanish-English bilinguals (Fiestas & Peña, 2004; Heilmann et al., 2016; Miller et al., 2006) this study contributes to the body of evidence in support of narrative language sampling as an assessment tool for bilingual populations.
References


