Testing Local: Small-Scale Language Sample Databases for ESL Assessment

Évaluation Locale. Petites Bases de Données d'échantillons Linguistiques Pour l'évaluation de l'anglais Langue Seconde.

Kate Ballem Chase
Judith R. Johnston

KEY WORDS

LANGUAGE ASSESSMENT

ESL

LANGUAGE SAMPLE ANALYSIS

NARRATIVE LANGUAGE

SCHOOL-AGED CHILDREN

Abstract

Purpose: This study describes the development of a small scale, local language sample database of children who were learning English as a second language. Goals were to develop a clinically useful, cost efficient means for comparing second language learning children to an appropriate peer group, and to determine whether a localized approach to database development could provide a practical solution to English Second Language (ESL) assessment.

Method: Narrative language samples were gathered from 18 typically developing 7 year olds learning English as a second language. These children spoke either Mandarin or Cantonese as their first language, and had entered kindergarten with little or no knowledge of English. Samples were compared on a full range of linguistic variables to samples collected from age-matched monolingual English speakers, drawn from a Systematic Analysis of Language Transcripts database.

Results: ESL children performed similarly to monolingual children on measures of lexical diversity, syntax, and language processing, but made more morphological errors.

Conclusion: The pattern of strengths and weaknesses observed in the ESL children differentiates them from monolingual English speakers and also from children with learning impairments, and thus demonstrates the potential usefulness of language sample databases in the assessment of this population.

Abrégé

But : Cette étude décrit le développement d'une petite base de données d'échantillons locaux de langage d'enfants qui apprenaient l'anglais comme langue seconde. Les buts étaient de développer un moyen économique et cliniquement utile pour comparer un groupe d'enfants apprenant une langue seconde à un groupe de pairs approprié, et de déterminer si le développement d'une base de données locale pourrait offrir une solution pratique à l'évaluation de l'anglais langue seconde.

Méthodologie : Des échantillons de langage narratif ont été recueillis auprès de 18 enfants de sept ans au développement typique qui apprenaient l'anglais comme lange seconde. Ces enfants parlaient le mandarin ou le cantonais comme langue maternelle et étaient entrés à la maternelle avec peu ou pas de connaissances de l'anglais. Les échantillons ont été comparés selon des variables linguistiques à des échantillons recueillis auprès de locuteurs unilingues anglophones du même âge, tirés d'une base de données d'analyse systématique de transcriptions linguistiques.

Résultats : Les enfants ayant l'anglais comme langue seconde ont obtenu un niveau de réussite égal à celui des enfants unilingues sur les mesures de diversité lexicale, de syntaxe et de traitement du langage, mais ils ont fait plus d'erreurs morphologiques.

Conclusion : Le tableau regroupant les forces et des faiblesses observées chez les enfants apprenant l'anglais comme langue seconde se différencie de celui des enfants anglophones unilingues ainsi que de celui des enfants ayant des troubles d'apprentissage. Les résultats illustrent donc l'utilité potentielle des bases de données d'échantillons linguistiques dans l'évaluation de cette population.

Kate Ballem Chase, D. Phil.
Clinical Assistant Professor
School of Audiology and Speech Sciences,
University of British Columbia,
Vancouver, BC
Canada

Judith R. Johnston, Ph.D.
Professor Emerita,
University of British Columbia,
Vancouver, BC
Canada

Providing speech-language services to multilingual children and their families is a challenging task that is becoming an everyday occurrence for many speech-language pathologists (S-LPs). In urban areas, large numbers of children are entering the school system with little or no exposure to English. In Vancouver, British Columbia, for example, the school board reported that in 2009 over 60% of the students spoke a language other than English at home, and listed over 126 home-languages (Vancouver School Board, 2011). Facts such as these help to explain the growing attention being paid to the practice of speech-language pathology among culturally and linguistically diverse populations. In this paper we focus in particular on the assessment of language abilities in ESL children, and propose a new approach to this task.

The Challenge of ESL Assessment

A primary source of difficulty in the identification of language disorder in children who are learning a second language is the similarities seen in the language characteristics of second language learners and children with specific language impairment (SLI). This similarity has been identified in languages of diverse typology including Swedish, Quebec French, and English. Hakansson and Nettelbladt's (1993) longitudinal study compared the development of subject verb inversion in two Swedish second language learners (L2), two Swedish children with SLI (SLI) and two typically developing children (L1) matched for MLU. The L1 children used varied word order patterns from the outset of their grammatical development. In contrast, the L2 and SLI children shared a clear pattern of progression through three stages, beginning with a stage in which there was exclusive and uniform use of subject verb object (SVO) word order. In regard to word order, the Swedish children who were second language learners resembled the children with SLI more than they resembled the typical first language learners. Studies of morphological development report similar findings. Paradis and Crago's (2000) study of Quebec French second language learners and Quebec French children with SLI found that both groups of children used significantly fewer finite verb constructions than monolingual children of the same age. Paradis (2005) later tested a group of 24 typically developing ESL children with less than two years of exposure to English, and found that 87.5% fell within the clinical range on morphological measures included in the Test of Early Grammatical Impairment (Rice & Wexler, 2001). Finally, in a study of English as L2, Gutierrez-Clellan, Simon-Cereijido, and Wagner (2008) found comparable levels of finite verb accuracy in typically developing ESL children, monolingual children with language impairment, and bilingual children with language impairment.

The similarities observed in SLI and L2 populations can make it difficult to interpret any language limitations observed in a second language learner, which can ultimately result in the over-identification of SLI. Alternatively, clinicians aware of similarities between second language learners and children with SLI may fail to diagnose language disorder by attributing a child's difficulty to their second language status. This dilemma has received significant attention in recent years, and a variety of solutions have been proposed: the translation of standardized tests into various home-languages, the use of interpreters (Laing & Kamhi, 2003), and the creation of alternate norms for bilingual children (Oller & Eilers, 2002).

Lack of Feasibility in Current Proposals

Educators and researchers in the field of bilingual assessment agree that bilingual children should be assessed in both of their languages (Bedore & Pena, 2008; Oller & Eilers, 2002; Kohnert, 2010; Miller et al, 2006). Indeed, assessment in L1 has been enshrined in the Individuals with Disabilities Education Act Amendments of 1997 in the United States (US) in an effort to avoid under- or over-diagnosis of language disorder in culturally and linguistically diverse (CLD) populations. This requirement rests on the reasonable assumption that assessment in both L1 and L2 will provide a truer picture of a child's language capabilities than assessment in L2 alone. However, with the exception of a few European languages, tools for assessment in languages other than English remain quite limited, and culturally and linguistically appropriate translations of English assessment tools are very difficult to create (Bedore & Pena, 2008). Furthermore, even if such tools existed it is not clear who would use them. A recent survey of S-LPs in British Columbia (Simmons & Small, 2011) showed that only 24 (18%) of the 136 clinicians included in the study were bilingual. Matching the few bilingual S-LPs to the many ESL children needing assessment is virtually impossible, especially in cities such as Vancouver that are linguistically diverse and contain a number of language preserving communities. Interpreters can assist in assessment sessions, but require significant training to be effective (Langdon & Quintanar-Sarellana, 2003). It is unlikely that the resources necessary to train and employ this additional group of professionals would be available in many jurisdictions.

Bilingual Norms for L2

Faced with the difficulties of assessing ESL children in L1, researchers have looked for alternate approaches. One possibility would be to compare a given ESL child's knowledge and use of English to the knowledge and

use of English seen in a normative group of bilingual children. There would be challenges in the development of the database since a number of factors such as nature of L1, length of exposure to L2, age at acquisition of L2, and home language environment are known to influence L2 learning (Oller and Eilers, 2002). Some of these factors would need to be specified as inclusion criteria; all of them would require careful description. Further research would be needed to determine the points at which heterogeneity compromised validity. However, once the normative database is created, this approach seems quite feasible, less costly, and well within the capabilities of most S-LPs. Substantial normative work has been done with Spanish-English bilingual children; Oller and Eilers' "Miami Project" gathered data on the performance of nearly one thousand bilingual children on standardized tests of English and Spanish vocabulary and literacy skills, and also collected data on socioeconomic status, home language environment, and school language environment. In the US, Pena and colleagues (Pena, Gutierrez-Clellen, Iglesias, Goldstein & Bedore, in development), have developed an experimental standardized assessment for Spanish-English bilingual children called the Bilingual English Spanish Assessment (BESA). Miller et al. (2006) have developed a database of Spanish and English narrative language samples from over 1500 Spanish speaking ESL children. These large-scale projects in the US provide rich sources of data, in both English and Spanish, for the assessment of Spanish speaking children who are learning English at school. As well as showing the importance of normative groups of bilingual children, they show the feasibility and potential value of assessments conducted in L2 and provide the tools to conduct them – but only for Spanish-English bilingual children.

The unfortunate fact is that despite these demonstrations of best practice, many S-LPs will not have appropriate normative data for bilingual children on their caseload or the ability to assess all children in their home language. For these clinicians, the current literature suggests a multi-pronged approach to assessment, including extensive child and family histories, classroom observation, peer-based comparisons, and dynamic assessment approaches (Bedore & Pena, 2008; Kohnert, 2010). Once again these practices are time consuming, costly and thus unlikely to be available to the number of children needing them. Clinicians remain in need of assessment solutions that are sensitive to cultural and linguistic differences yet feasible within the available resources.

Localized and Focused Language Sample Databases

Discussions of language assessment with ESL children have tended to focus on the appropriateness and availability of standardized tests. The work of Miller and colleagues (Miller et al, 2006) is a noteworthy exception to this trend. The Systematic Analysis of Language Transcripts (SALT) software (Miller & Iglesias, 2012) uses naturalistic samples of spontaneous speech (e.g. conversation, narrative, expository) as the basis for language assessment. The program compares the language of one child to the language of a normative peer group, providing standard scores for a number of language characteristics in a variety of domains (e.g., semantic, syntactic, lexical). Use of spontaneous language avoids the cultural specificity of test items and the need for "test-taking" skill. The most important feature of the SALT program may be that it allows individual users to develop their own normative databases for use with children who are learning languages other than English or who represent some particular population. Thus far this capability has led to the development of databases of language samples from children learning Spanish (now provided with SALT), Quebec French (Thordardottir, 2005), and Turkish (Acarlar & Johnston, 2006).

SALT's ability to work with the user's own databases would seem to provide exactly the tools needed to differentiate ESL and SLI learners. A reference database of English (L2) language samples could be collected from typically developing second language learners. This database could then form a standard of comparison for children who seem to be having difficulty learning English. This evaluation could function as an initial "screening" in which the assessment question would be: Does this child speak English as well as other ESL children at the same age and from similar language and cultural groups? If a child was found to compare poorly to the pertinent reference database, his/her language abilities could be more fully examined, including, where possible, an assessment in L1. To our knowledge, the only normative SALT database containing language samples from bilingual children is the one described by Miller et al. (2006), consisting of Spanish and English narrative language samples from over 1500 Spanish speaking ESL children. This database is now available for download with SALT (Miller & Iglesias, 2012), and provides rich data for the assessment of Spanish speaking children who are learning English in school.

SALT software invites clinical researchers to develop additional large-scale SALT databases that target various groups of second language learners (e.g. Spanish ESL children in California; Chinese ESL children in Vancouver, Canada; Punjabi ESL children in the United Kingdom). However, the cost of such a project would be considerable and perhaps, as in the US, could be justified only for the largest groups. However, as an alternative to large-N databases spanning a range of ages, one could

build small-scale databases that focused on children of a particular age or experience. Large or small, these databases could be used by clinicians to evaluate a bilingual child's English (L2) proficiency by comparison to other ESL children with similar cultural backgrounds and learning opportunities. This evaluation would help to identify children who warranted an in-depth assessment.

The Current Study

This paper describes the development of a small-scale database containing language samples collected from Grade 2 Chinese ESL learners. We used the SALT program to determine whether the language of the bilingual children in this database showed consistent areas of strength and weakness when compared to that of monolingual age peers and if so, whether the identified profile was also likely to distinguish ESL and SLI learners. If so, small-N normative databases that were focused on particular points of developmental might provide a feasible solution to the immediate clinical needs of S-LPs serving multilingual caseloads.

Methods

Participants

Participants in this study consisted of 18 children enrolled in the first semester of Grade 2 in English-speaking public schools in the Greater Vancouver area. The mean age of these children was 7;2 years with a range from 6;10 to 7;9. Local schools use a three level framework to describe the language proficiency of ESL children at school entry. Table 1 lists the criteria for each of the three levels of English proficiency. To be included in this study, participants needed to speak Mandarin or Cantonese as their first language and to have been judged by school personnel to speak English at Levels 1 or 2 when they entered kindergarten. They were also required, in the teacher's opinion, to be progressing in their current classrooms as expected given their ESL status, and could not have participated in any intensive English programs outside the normal classroom environment. Information pertinent to the selection of participants was obtained from school personnel and from parents via a written questionnaire, in both English and Chinese, regarding language knowledge and use in the home (see appendix). Table 2 summarizes the information that parents provided.

All of the children were reported to have learned Cantonese or Mandarin as their first language. Two thirds of the families reported speaking only Chinese at home, with the remaining third using both Chinese and English. The mean age of initial exposure to English reported by parents was 3;10 years, indicating that most of the children did have some pre-kindergarten exposure to English. Despite this apparent early exposure, English proficiency at kindergarten entry was reported to be quite limited. Twelve of the children were judged to know only a few words (Level 1) at kindergarten entry, and 4 were judged to be able to use only short phrases (Level 2). The remaining 2 children were judged to be more advanced by their parent, but at Level 2 by their teacher and school S-LP. An explanation for this otherwise surprising outcome may lie in the fact that first exposure occurred primarily in the context of preschool or daycare programs. There are wide

Table 1. Three-level classification of early English proficiency.

Level 1: Emerging	Level 2: Beginning	Level 3: Developing
• Very limited understanding. • Speaks in isolated words or short phrases. • Often silent. • Repeats, uses body language to be understood. • Limited pronunciation of English sounds	• Has difficulty following what is said even when slowed. • He sitant in everyday conversation. • Understands/uses simple concrete words, phrases, sentences. • Often silent. Requires long wait before answering questions. • Needs extensive support with content language. • Repeats, uses body language to be understood. • Often hard to understand.	• Usually understands and engages in conversations with peers. • Begins to participate in classroom discourse. • Uses varied vocabulary, sentence structure. • Requires less waiting prior to answering questions. • Needs support with content language. • Uses/requires repetition and rephrasing of new material. • Begins to self-correct. • Occasionally hard to understand.

Table 2. Parent responses to language use survey regarding use of Chinese languages (Cantonese or Mandarin) and English.

Question:	Parent Responses:		
	Chinese only:	Chinese/ English:	English only: 0
Language used by parents:	12	6	0
Language used by child:	13	4	1
Child's L1:	18	0	0
Mean age of initial exposure (SD):	3;10 (0;11)		
Location of initial L2 exposure:	Preschool/Daycare	14	
	Other:	4	
Level of English ability at kindergarten entry:	"Spoke only a few words":	12	
	"Could put a few words together":	4	
	"Could talk about most topics but made lots of mistakes...":	2	

differences in patterns of language use in Vancouver preschools and daycares. Those designated as "English" may still have a significant number of Cantonese/ Mandarin speaking teachers and peers. The English immersion expected in these schools is frequently absent. Children attending different preschools for the same number of years could have widely differing exposure to English. Based on the low ratings of proficiency at kindergarten entry, it seems improbable that any of the participants were early coordinate bilinguals.

Our inclusion criteria address at least five of the factors known to influence the learning of a second language: age, L1, home culture, context of L2 learning, and length of exposure to L2. This study was not, however, designed to confirm these influences or to delineate the factors in detail. Our more practical purposes led us to prioritize variables that were currently available and feasible in a school service setting. For example, in the absence of appropriate standardized tests we relied on the ability of teachers to describe the language level of the bilingual children as they entered kindergarten and to identify children in Grade 2 whose English was advancing at a satisfactory rate. Given the ethnic profile of the school district, all teachers would have considerable experience observing and interacting with children from the Chinese community, including many children who entered kindergarten with little or no English.

Our inclusion criteria and selection process also did not guarantee a representative sample of children who enter kindergarten with English proficiency at Levels 1 or 2. Teachers might not have referred children unless they were confident about the normalcy of his/her progress in learning L2. If our database were intended for use as a full diagnostic instrument, this potential bias towards more competent speakers would be problematic. However, our database was intended only to identify children for whom further assessment seems warranted and a tendency toward over-identification, if present, could be viewed as suitably cautious.

We decided to focus this analysis exclusively on Chinese children in grade 2 for several reasons. First, the use of a smaller, but focused database required fewer resources. If the outcomes of the analysis were to indicate that this assessment approach is viable, similar databases would be economically feasible in most school settings. Second, although we would prefer to identify children with learning problems at an earlier age, the initial months of Grade 2 had been identified as an important educational decision point by the S-LPs who participated in this study. Although the development of English proficiencies that can fully support classroom learning requires some four years of school experience (Ramirez, 1998, as cited in Johnston, 2006), after two years of English schooling, many Grade 2 bilingual children have considerable English language and literacy capabilities. Those who do not are beginning to be referred for S-LP services. Finally, including only children from a single culture group reduced the need to identify cultural differences that could lead to differences in English proficiency, e.g., differences in family support or expectations for academic achievement.

Potential participants identified by their school-based S-LP were sent home with an explanatory letter, a consent form, and a questionnaire regarding the child's developmental history (see appendix). Children whose parents provided consent were then interviewed individually by the school-based S-LP to ensure that the child met inclusion criteria.

A comparison group of monolingual English speaking children was drawn from an existing database of samples gathered during the standardization of Gillam and Pearson's (2004) Test of Narrative Language (TNL). Eighteen age-matched monolingual English children were randomly selected from the database. The mean age of the monolingual group was 7;2, with a range of 6;10 to 7;9. This comparison group was included to in order to determine whether and in what ways the C-ESL group could be expected to differ from monolingual English speakers, i.e. to identify the 'normal' limitations of a C-ESL child at this point in development.

Procedure

Narrative Task

A substantial body of literature attests to the value of narrative in language assessment. Narrative draws on a child`s social, linguistic, world, and conceptual knowledge as well as knowledge specific to storytelling (Johnston, 2008). Narratives require more intensive planning and coordination of cognitive processes than conversational speech (Hadley, 1998), and they tend to elicit longer utterances (Miller & Leadholm, 1992). These features contribute to the usefulness of narratives in the assessment process. Indeed, Masterson and Kamhi (1991) suggest that narrative discourse may reveal areas of difficulty where conversational data do not. This is particularly important where second language learners are concerned, as their competence in L2 may be adequate for conversation after a short exposure but be inadequate for more complex discourse forms (Ramirez, 1998, as cited in Johnston, 2006). A secondary, but important consideration in our decision to use a narrative task was the goal of reducing cultural bias in the assessment process. Although narrative traditions vary across cultures, Fiestas and Pena (2004) found that Spanish-English bilingual children were able to produce narratives of equivalent complexity in both Spanish and English on a picture-book task. Cleave, Girolametto, Chen and Johnson (2010) compared the performance of monolingual and bilingual children with SLI on a narrative task, and found similar levels of performance in the two groups. In short, there is growing consensus that narratives provide a supportive context in which to elicit maximum linguistic complexity while minimizing the effects of cultural bias (Bedore, Pena, Gillam & Ho, 2010; Cleave et al, 2010; Rojas & Iglesias, 2009).

Each child in this study was asked to produce two narratives. The children were informed that the S-LP was conducting a project to learn more about how children tell stories. S-LPs then asked the children to produce two stories, each based on a different picture stimulus taken from the Test of Narrative Language (Gillam & Pearson, 2004). The first stimulus was a sequence of five pictures showing a boy waking up late and his ensuing difficulties in getting to school on time. The second stimulus was a single picture showing two children watching an alien spaceship landing in a park. The children were reminded that stories should have a beginning, a middle, and an end, and were encouraged to make them as long as possible. The samples were audiotaped and returned to the University of British Columbia for transcription and analysis by the authors. The monolingual comparison group samples had been gathered using identical stimuli during the standardization of the Test of Narrative Language (Gillam & Pearson, 2004). It should be noted that the entire Test of Narrative Language was administered to the monolingual comparison group, while only the two sections described above were used in the current study.

Transcription

Language samples were transcribed by trained graduate students using SALT coding conventions. In addition to the standard coding of inflectional

morphemes, mazes, and abandoned utterances, the transcribers coded any words that were omitted or incorrectly selected. This additional coding was also carried out on the monolingual language samples. Utterances were divided into C-units (communication units), which consist of an independent clause and its modifiers (Miller & Leadholm, 1992), as this is the unit of analysis used in the language samples of the TNL database available with SALT (SALT Software, 2010).

The process of *consensus transcription* developed in the SALT lab (Heilmann, Miller, Nockerts, & Andriacchi, 2006) was used to ensure the reliability of the transcriptions. Each sample was transcribed by one coder and then checked by a second, who listened to the sample, reviewed the original transcript, and noted any disagreements. These disagreements were resolved through discussion between the two transcribers. The initial level of agreement between transcribers was calculated to determine the accuracy of the transcriptions on a number of aspects of the transcript: disagreements on utterance boundaries, words, and morphemes were each considered separately. Agreement levels were above 95% for all measures, on all transcripts.

Analysis

SALT software (Miller & Iglesias, 2003-2007) was used to compare the ESL and monolingual groups on a number of measures across the domains of vocabulary, morphology, syntax, and language processing (see Table 2). In each domain we measured aspects of expressive language that have been considered the hallmarks of specific language impairment in children learning English as a first language. If, as suggested by recent studies (Hakansson & Nettelbladt, 1993; Paradis & Crago, 2000; Paradis, 2005; Paradis 2010), these measures also reveal areas of difficulty for our second language learners, there would be little support for using L2 Language Sample Analysis as a screening tool to differentiate between the two groups. If, on the other hand, the performance of the ESL children on some or all of these variables fails to indicate learning difficulties, it would suggest the viability of this approach for distinguishing between language impaired children and second language learners. Since this project was also intended to inform current practice patterns, we selected variables that were both clinically relevant and readily accessible. All of the measures could be calculated using the basic operations of the SALT Program and most were included in the standard set of SALT analyses. These measures are outlined in Table 3 and discussed in the following sections.

Vocabulary:

NDW. The number of different word (NDW) roots in a language sample is one of a variety of measures of lexical diversity, which include the traditional type-token ratio (TTR, Templin, 1957), as well as more complex measures such as "D" (Richards & Malvern, 1997). Significant debate has occurred over the usefulness of various measures and their potential confounds with syntax and sample size (Hewitt, Hammer, Yont & Tomblin, 2005; Richards & Malvern, 1997; Watkins, Kelly, Harbers & Hollis, 1995). Nevertheless, NDW has been shown to differentiate typical and language disordered populations when the total number of words in a sample is controlled (Hewitt et al., 2005). NDW-100 has also been found to distinguish typical and language disordered populations in monolingual Cantonese pre-schoolers (Klee, Stokes, Wong, Fletcher & Gavin, 2004). Thus, we opted to use NDW calculated in 100 word samples (NDW-100) as our measure of lexical diversity.

Content word errors and omissions. In addition to lexical diversity, we also investigated the number and rate of incorrectly selected content words ("word errors") and omitted content words ("word omissions"). Word errors were coded when a child used the wrong word for the context. For example, a word error was judged to occur when a child said, "I saw TV" instead of "I watched TV." Word omissions were judged to occur when a word was omitted, as in "I *VERB to the park," where the necessary verb was omitted. Note that only word errors and omissions that expressed primary propositional content were included in this category. Errors or omissions of grammaticized words such as articles, pronouns and prepositions were coded separately.

Morphology:

English-speaking children with specific language impairment have difficulty with grammatical morphemes, including both affixes and unbound functors (Johnston & Schery, 1976; Rice & Wexler, 1996; Bedore & Leonard, 1998). Children who are learning English as a second language have also been shown to make errors in their use of grammatical morphemes (Paradis, 2005). We calculated the number of types of bound morphemes used in the samples, and the number of omissions of bound morphemes and unbound functors. Finally, we investigated the number of word errors that involved grammaticized forms such as articles, pronouns and prepositions. For example, in "Him went to the park", the case of the pronoun is incorrect, but the propositional content is not lost. Errors involving these closed-class words were coded as morphosyntactic word errors.

Table 3. Measures of language ability used to compare samples.

Domain	Measure	Description
Vocabulary	NDW	Number of different words used in the transcript.
	NTW	Total number of words used in the transcript.
	Word Errors (Content)	Number of content words used incorrectly.
	Word Omissions (Content)	Number of content words omitted.
Morphology	Types	Types of bound morphemes used.
	Rate omitted	Rate of omission of obligatory bound morphemes
	Word Errors (Morphosyntactic)	Number of morphosyntactically based word errors.
	Word Omissions (Morphosyntactic)	Number of morphosyntactically based word omissions.
Syntax	MLUw	MLU, calculated in words
	Conj/Wh-comps (Types)	Types of conjunctions and complementizing wh-words used in non-initial positions.
	Subordination Index	Ratio of total number of clauses to the total number of C-units.
Processing	Mazes (Number)	Total number of mazes (linguistic non-fluencies such as false starts, revisions, pauses, and repetitions)
	Mazes (Rate)	Percentage of utterances with mazes.

Syntax:

MLU. Mean length of utterance (MLU) has a long history of use as an index of syntactic complexity (McCarthy, 1943) and general level of language proficiency. Hewitt et al (2005) examined MLU in kindergarteners and found that MLU was lower for children with SLI as compared to typically developing peers. Additionally, Klee et al.'s (2004) study of monolingual Cantonese pre-schoolers found lower values for MLU in children with SLI. MLU has been shown to increase with age throughout the school years (Miller & Chapman, 1981; Miller & Leadholm, 1992). Despite this correlation, it has been argued that the validity of MLU as a measure of syntactic ability is lower in older children, as the length of children's utterances becomes relatively less dependent on syntactic ability and relatively more dependent on discourse factors (Johnston, Miller, Curtiss, & Tallal, 1993). Johnston (2001) introduced an alternate calculation of MLU2 that serves to reduce the effect of discourse context on MLU by eliminating exact repetitions and responses to questions. Due to the narrative nature of these transcripts, there were no opportunities for these discourse factors to affect the calculation of MLU, and so this procedure was not necessary. To prevent a confound between morphological and syntactic abilities, mean length of communication units was calculated in words, rather than morphemes.

Use of Conjunctions and Wh-words. The use of conjunctions (e.g., *He cried because the bus left*) and complementizing wh-words (e.g. *I know where my shoes are*) was investigated as a second measure of the syntactic complexity. Complementizing wh-words were defined as wh-words used in non-initial positions. The number of different types for each of these categories was counted. This syntactic measure was included to determine whether the utterance length measure reflected utterances with propositional complexity as well as expansions of the primary verb or noun phrase (Johnston & Kamhi, 1984).

Subordination Index. A final analysis of syntactic complexity was conducted using the subordination index (SI), which is a ratio of the total number of clauses divided by the total number of C-units. The coding required for the SI is described in detail in the documentation available with SALT (SALT Software, 2010). Codes for SI are now included in the standard databases provided with SALT. While this improves clinical accessibility for SI, it remains more time consuming to code for this than to calculate other measures of syntactic complexity (e.g. MLU, conjunctions/wh-words). We nevertheless included SLI in our measures of SI to ensure that we accurately described the syntactic complexity of the ESL language samples, and to validate any findings based on the simpler measures used.

Language Processing:

Mazes were used as an index of processing difficulties with the narrative task. Mazes have been defined as linguistic non-fluencies that do not form part of the intended utterance (Loban, 1976). Mazes include repetitions, revisions, false starts, and abandoned utterances. They occur in the utterances of typical children and adults, as well as in children with language disorders. Mazes are linked to the processing required for planning and producing sentences (Rispoli & Hadley, 2001), and their frequency tends to increase when children are using more complex or newly learned language patterns (Miller & Leadholm, 1992; MacLachlan & Chapman, 1988). They also occur more frequently in the language of children who have known difficulty with language processing; as in SLI (Bond & Schneider, 2005; Thordardottir & Ellis Weismer, 2002).

Results

The narrative language samples of the Chinese-ESL (C-ESL) and monolingual English (ML) groups were compared on the variables described above, and the reliability of any group difference was evaluated using one-tailed t-tests. Tests were one-tailed as we were expecting the ML group to outperform the C-ESL children. Direct comparisons of the raw scores from various measures were not made as they would be uninterpretable due to differences in scale. We considered standardizing the scores then using repeated measures ANOVA with 'families' of variables, e.g., syntactic, lexical, etc. This design would increase the likelihood of a main effect of Group but it also assumes a dimensional relationship among family members that research has yet to demonstrate. With these considerations in mind a series of t-tests seemed most appropriate. However, due to the increased risk of Type 1 error with multiple comparisons, alpha levels were set family-wise at 0.05. The alpha level for individual tests was .017 or .013 depending upon the number of tests conducted within the family. Cohen's d was used to calculate effect size for all measures. Table 4 provides an overview of all comparisons conducted.

Sample Size

Given that both groups of children were completing the same constrained narrative task, we chose not to control for differences in the length of the language samples, but treated length as a dependent variable. If group differences did emerge, statistical controls could be used where pertinent in the analysis of other variables. The C-ESL group had a mean of 26.7 utterances per sample (SD = 12.6), while the monolingual control group had a mean of 25.8 utterances per sample (SD = 10.8). This difference was not significant, $p > 0.4$: the two groups produced narrative samples of a similar length.

Vocabulary

Analysis of NDW-100 revealed highly comparable values for the C-ESL children (M = 47.4, SD = 5.4) and the ML children (M = 50.1, SD = 8.2). There was no significant difference between the two groups on this measure, t(34)=1.15, $p > 0.13$, and the effect size was small, d = 0.39. Overall, the C-ESL and ML children showed similar levels of lexical diversity. Analysis of errors in lexical selection showed that these errors were infrequent in both groups, with $M_{\text{C-ESL}}$= 2.1 ($SD_{\text{C-ESL}}$=2.3) and $M_{\text{ML=}}$1.0 ($SD_{\text{ML=1.6}}$). This difference did not achieve significance, $p > 0.09$, however there was a medium effect size, d = 0.55. It is possible that in longer language samples or with a larger sample size, this measure would reveal differences. Omissions of content words were not analyzed as there were so few in the entire database: only 2 instances in the C-ESL samples and only 1 instance in the ML samples. Overall, our analysis of vocabulary measures suggests that the C-ESL children in this study had an adequate range of vocabulary for this narrative task after two years of English schooling.

Morphology

Analysis of the different types of morphemes used by the two groups of children suggested that both groups used a similar variety of bound morphemes ($M_{C\text{-}ESL}$ = 3.89, SD = 1.60, M_{ML} = 4.17, SD = 1.69, $t(34)$=0.51, p > 0.30, d = 0.19). Children in the C-ESL group omitted a mean of 17.6% of bound morphemes (SD = 21.6%), compared with just 4.10% of bound morphemes in the ML group (SD = 12.2%). This difference was significant ($t(34)$ = 2.31, p = 0.016), and the effect size was medium-large, d = 0.77. The large values for standard deviation suggest this was a measure on which the performance of the C-ESL group varied considerably. Analysis of errors in grammaticized words showed that C-ESL children made significantly more word errors ($t(34)$ = 4.73, p < 0.0001, d = 1.57) than did ML children ($M_{C\text{-}ESL}$=4.83, $SD_{C\text{-}ESL}$=3.49, M_{ML}=0.83, SD_{ML}=0.86). Taken together, these data seem to suggest that the C-ESL group are continuing to struggle with the morphosyntactic system of English in their Grade 2 year.

Syntax

Table 4 shows the mean values for MLU in each group. Although the ML (M_{ML}= 7.67, SD_{ML}=1.55) group had a slightly higher MLU than the C-ESL ($M_{C\text{-}ESL}$=7.30, $SD_{C\text{-}ESL}$=1.24), group, there was no significant difference in utterance length between the two groups, $t(34)$ = 0.79, p > 0.79, d = 0.56. Children in the ML and C-ESL groups also had similar variety in the different types of conjunctions and complementing wh-words ($M_{C\text{-}ESL}$=3.89, $SD_{C\text{-}ESL}$=1.60, M_{ML}=4.17, SD_{ML}=1.69, $t(34)$=0.51, p > 0.12, d = 0.17). Finally, there were no differences between the two groups on the subordination index, t(34) = 0.05, p > 0.90. In summary, none of the syntactic measures used in this study showed group differences. This suggests that the syntactic complexity evident in the oral narratives of the C-ESL children was comparable to that seen in monolingual English speaking children of the same age.

Language Processing

There were no statistically significant differences between the two groups for total number of mazes in the sample ($M_{C\text{-}ESL}$ = 12.8, SD = 9.87; M_{ML} = 10.78, SD = 8.59; $t(34)$ = 0.64, p > 0.25, d = 0.22). There was also no significant difference in the rate of maze occurrence between the two groups (p > 0.95). As measured by mazes, it seems that this oral narrative task did not

Table 4. Comparisons between Chinese ESL (C-ESL) and monolingual (ML) samples on selected language measures. Significance was determined using a family-wise alpha of 0.05. Statistically significant differences noted with an asterisks.

Language Domain	Measure	Mean C-ESL (SD)	Mean ML (SD)	t(34)	d
Sample Size	**Number of Utterances**	26.7 (12.6)	25.7 (10.8)	0.22	0.09
Vocabulary	**NDW-100**	47.4 (5.39)	50.1 (8.21)	1.15	0.39
	Word Errors (Content)	2.11 (2.27)	1.00 (1.61)	1.70	0.56
	Word Omissions (Content)	not analyzed			
Morphology	**Morphemes (Types)**	4.28 (1.23)	4.50 (1.04)	0.59	0.19
	% Morphemes Omitted	17.6 (21.6)	4.10 (12.2)	2.31*	0.77
	Word Errors (Morphosyntactic)	4.83 (3.49)	0.83 (0.86)	4.73*	1.57
Syntax	**MLUw**	7.30 (1.24)	7.67 (1.55)	0.79	0.26
	Conj/Wh-comps (Types)	3.89 (1.60)	4.17 (1.69)	0.51	0.17
	Subordination Index	1.33 (0.27)	1.33 (0.22)	0.05	0.0
Language Processing	**Number of Mazes**	12.8 (9.87)	10.8 (8.59)	0.65	0.22
	Rate of Mazes per Utterance	0.46 (0.26)	0.40 (0.24)	0.81	0.24

cause more processing difficulties for the C-ESL children in this study than for their monolingual peers.

Discussion

Comparing the language profiles of Grade 2 monolingual, C-ESL and SLI children

This project was undertaken with the goal of improving the diagnosis of language disorder in learners of English as a second language. To this end, we investigated the viability of an assessment strategy that would begin with language sample analysis of narratives in L2, using a database of language samples from typically developing ESL children in the local community as a reference point. This strategy would only work if ESL children presented a profile of language strengths and weaknesses that differed from the profile seen in children with SLI and also from children who are monolingual speakers of English. The literature primarily points to similarities in the language patterns of ESL and SLI children (Gutierrez-Clellen et al, 2008; Hakansson & Nettelbladt, 1993; Paradis & Crago, 2000; Paradis, 2005), but studies comparing these two groups have focused primarily on grammatical morphology. Our study compared the three groups on a broader set of language variables.

Consider first the language measures taken individually. The C-ESL learners in this study demonstrated some difficulty in their use of grammatical morphology compared to the monolingual children, with higher rates of omissions and errors in their use of bound morphemes and unbound functors. This finding is consistent with earlier reports (Paradis & Crago, 2000; Paradis, 2005; Guitierrez-Clellen et al, 2008), and indicates that analysis of ESL children's morphological abilities alone would not differentiate between typical and atypical second language learners. It also points to an important function of bilingual L2 reference databases: to identify those errors that are commonly found in the language of bilingual children in grade 2 and should not be treated as evidence of language learning difficulties.

The fact that our second language learners had levels of lexical diversity (NDW, number of different words) similar to those of monolingual children may seem to run counter to evidence that second language learners have smaller vocabularies in L2 than do monolingual speakers of that same language (Genesee, Paradis & Crago, 2004). A measure of lexical diversity, however, does not directly evaluate vocabulary size. Instead, it evaluates the ability to use a variety of words within a communicative task - an ability that seems to imply a lower cost for lexical processing and greater attention to differences in meaning. School-aged English speaking children with language disorders and Cantonese speaking preschool children with SLI have been shown to score lower than their age peers on this measure (Hewitt et al., 2005; Klee et al., 2004). The current data suggest that while ESL children may know fewer words than their monolingual peers, they use what they know in a more mature fashion than is seen in children with SLI.

Values for MLU have been shown to increase throughout the school years, and MLU has long been used in the identification of language disorder (Miller & Leadholm, 1992). For example, Hewitt et al. (2005) recently found lower values of MLU in a group of Kindergarten and Grade 1 students with language disorder as compared to a group of typically developing peers. Klee et al.'s (2004) study of Cantonese speaking pre-schoolers with SLI likewise showed reduced MLU in the language-disordered group. In contrast, the typically developing C-ESL students sampled in the current study had values of MLU that were similar to those of the monolingual children. Lexical indicators of syntactic complexity such as conjunctions and the calculated indices of clausal complexity were in accord with MLU.

Complex syntax is generally motivated by complex ideas. It makes sense that since children in the C-ESL group were normal learners, they would know from experience that language is capable of indicating pragmatic focus and expressing abstract relationships between ideas and/or events. They would thus attempt these same functions in L2, creating language that was syntactically complex, though not always grammatical.

Finally, children with language disorder have been reported to have higher rates of maze occurrence than age-matched peers (MacLachlan & Chapman, 1988; Bond & Schneider, 2005; Thordardottir & Ellis Weismer, 2002). The C-ESL children in this study did not show this pattern; their rates of maze occurrence were at the same low level seen in monolingual age peers. This pattern of group differences suggests that the presence or absence of mazes may be another way to distinguish between children learning English as L2 and children with language learning disorders. However, since interpretation of maze data is best done with knowledge of a given child's familiarity with specific lexical and syntactic forms as well as detailed syntactic analyses, clinical decisions should probably not be based on frequency of mazes alone.

We turn next to consider our findings on the various language measures taken as sets. When compared with monolingual peers, there was evidence that children in our C-ESL group were more likely to omit or make errors in their use of bound morphemes and unbound

functors. However, there were no reliable group differences on measures of lexical diversity (NDW-100), syntactic complexity (MLU, use of conjunctions/wh-words, SI), or language processing (maze rates). All of our measures were chosen because prior research had shown them to be areas of characteristic difficulty for children with SLI (Hewitt et al., 2005; Johnston & Kamhi, 1984; Johnston & Schery, 1976; Klee, et al, 2004; Paradis, 2005; Bond & Schneider, 2005; Thordardottir & Ellis Weismer, 2002). The fact that the ESL group did show significant difficulty with two of the morphological measures but resembled native L1 speakers in the remaining variables, suggests that it should be possible to distinguish the language profile of a typically developing ESL child from the profile of a bilingual child with SLI. Although we would have liked to further test this conclusion with a comparison group of Chinese SLI children learning English in a primary school setting, an extended search for such children in Vancouver, Canada and in Hong Kong was unsuccessful.

Typical "norms" are drawn from large, carefully identified samples, not language samples collected from 18 children who are "doing just fine". Without denying that a larger number of participants would improve our database, two factors do seem to mitigate this concern. First, our normative database focuses on a very narrow developmental window, i.e., the first three months of Grade 2. If we think of our normative sample as a cell within a larger database, and each school year had four cells with 18 samples, the entire normative sample would consist of 72 children at each grade level, a more readily acceptable size. Secondly, our normative data were collected in the home communities of the children who will be evaluated with them. This design feature is rare and virtually guarantees a level of pertinence that is usually achieved through increases in the size of the normative database.

The individual variability seen in the performance of the Chinese ESL children on each of the measures described here may also raise concern. However, such variability is reported in many studies of second language learners (Genesee, Paradis, & Crago, 2004), and similar variability exists within language-disordered populations (Hewitt et al., 2005). The degree to which this variability reflects sampling decisions or is inherent in the learning process is unknown. However, when local SALT L2 databases are available, it becomes possible to investigate a child's performance on a range of measures with reference to the same normative group for all areas of language proficiency. This, in effect, controls for sampling differences and allows us to focus on differences in individual variability among the different language domains.

To summarize the argument thus far, our data suggest that there are a number of areas of language in which C-ESL children, after two years of schooling in English, can be as proficient as their monolingual peers – at least when telling stories in English. Two such measures, MLU and NDW, would seem to be particularly useful in the early phases of assessment; they are comprehensive, independent of topic, reliable across samples of varying length, and included among the standard SALT variables. In clinical practice, children whose stories were comparable in length and lexical diversity to those told by children in a local ESL SALT database would be judged to have low priority for in-depth assessment despite morphological errors. In contrast, the SLI literature indicates that language disordered children, including those who are bilingual, are likely to have difficulty in lexical diversity, syntax, and language processing in addition to grammatical morphology. If a child's L2 English usage fell below the levels evident in the local ESL reference database in several areas, that child could be scheduled for additional assessment.

These preliminary practice guidelines will need confirmation in further research, as studies of language impairment and bilingualism to date have not typically looked at the full profile of language measures simultaneously, nor have there been many studies of bilingual children with language impairment. Also, as discussed earlier, the relatively small number of participants and our selection process make our conclusions somewhat less certain. It is important to remember, however, that the goal of this project was not to provide normative data on the English competencies of second grade children from Chinese speaking homes. We intended only to provide empirical support for further research on assessment strategies. Findings reported here do seem to indicate that use of local reference databases of L2 narratives in the early stages of language assessment is a strategy that merits further investigation.

Current research does not yet support the creation and use of a single database of L2 samples elicited from children who have learned a variety of first languages, nor does it support using an L2 database from one language, cultural or socioeconomic group to evaluate the same L2 spoken by children from some other group. Several lines of evidence suggest caution in generalizing from one language or cultural group to another. Johnston and Wong (2002) and Simmons and Johnston (2007), for example, identified significant cultural differences in the verbal interaction patterns of parents with young children, and in the beliefs of parents about their role in language learning. These differences in parental beliefs and practices could well affect the course of language learning although the

necessary research has yet to be done. The nature of the child's first language is known to influence the timing and sequence in which its grammatical forms are learned (Aksu-Koc, 1998; Slobin, 1973). For example, in a recent comparison study of 11 languages, grammatical morphology was learned earliest in languages with the richest morphological systems (Xanthos, et al, 2011). In the area of child L2 acquisition, recent research has confirmed that transfer effects do play some role in children's acquisition of English morphosyntax (Paradis, 2011; Zdorenko & Paradis, 2011). For example, Zdorenko and Paradis (2011) compared article use in L2 English children learning from L1 languages with and without articles. Children from all backgrounds overused the definite article, a developmental trend that has long been identified in L1 acquisition of English (e.g., Brown, 1973). Additionally, however, Zdorenko and Paradis found that children from no-article L1s (Cantonese/Mandarin) frequently omitted articles, while such omissions were very unusual in the language of children from article-including L1s (Spanish, Hindi/Urdu/Punjabi). Similarly, Paradis (2011) found that verb tense and agreement marking of a child's L1 was one predictor of L2 children's performance on the Test of Grammatical Impairment (Rice & Wexler, 2001). Overall, current research suggests that it would be advisable to minimize language, cultural and social differences between the children being assessed and the children in the reference databases.

The Feasibility of "Testing Local"

If confirmed, the patterns of performance reported in this study would emphasize the importance of assessment methods that allow children to demonstrate their language strengths as well as weaknesses. Language sampling provides just such a tool. With relatively little assessment time and minimal coding, we were able to obtain a snapshot of children's performance in a variety of language domains and the resultant profiles served to distinguish the groups. Narrative samples as short as about 25 utterances require little time to collect and transcribe but have been shown to yield the same values as longer samples for general measures such as MLU, NDW and rate of maze occurrence (Heilmann, Nockerts & Miller, 2010). A 1997 survey of S-LPs in the US indicated that some 83% of them were already using language sample analysis in their work with children Kemp & Klee, 1997). Data regarding use of SALT are not available, but our teaching experience indicates that the basic analyses necessary to implement the assessment strategy presented here can be learned in 30-60 minutes by a computer literate S-LP. The skill needed to create a database might require another hour or two, but would not be required of all users. Interpretation of the SALT profile for a given bilingual child requires only the same body of knowledge that all practice with children with language disorders requires. In short, the local SALT database approach thus seems to provide a practical and feasible means by which to identify those ESL children who require more extensive assessment.

Conclusion

Findings from this study point to the potential value of an assessment strategy that begins by comparing a given ESL child's English narratives to those of typically developing ESL children in a local normative database. The bilingual children in our local reference database displayed a profile of language abilities that showed many similarities to monolingual children, despite a distinct difficulty in the domain of morphology. This contrasts with the profile of SLI children found in the literature, in which children with SLI have been shown to demonstrate difficulty in morphology, but also in vocabulary, syntax, and language processing. Assessments of the language learning ability of ESL children have been limited by cultural bias (e.g. standardized tasks), lack of standardization (e.g., observational data), or extreme logistical challenges (e.g. assessment in child's first language). The use of computer assisted language sample analysis could substantially reduce these limitations. A local normative database such as the one described here could be created and a profile for ESL learners from that particular cultural, linguistic and economic group could be determined. Armed with this information, clinicians could implement a step-by-step process to decision making in the assessment of an ESL child. We are suggesting here that the initial step in this process would be the use of the local normative database to assess the child's performance on a short, narrative language sample in L2. If the child were to perform poorly relative to ESL peers, further investigation of the child's language learning ability in L1 or L2 would be warranted. If the child performed well relative to ESL peers, in-depth assessment at that time would not be required.

Further work will be required to assess the sensitivity and specificity of this assessment strategy, to identify the most discriminating variables, and to establish the predictive validity of language sample data with ESL learners. The findings reported here would seem to indicate the value of continuing this line of investigation. Small-scale databases representing the language proficiencies of children from specific language and cultural communities would seem to be an appropriate and practical option for improving our standard of care among culturally and linguistically diverse communities.

Acknowledgements

This report as well as the major portion of the data analysis and project management were submitted by the first author in fulfilment of the requirements of the M.Sc. degree under the supervision of the second author. Funding for this project was provided by the Community Care Foundation. We thank them for this support, and also thank the families, teachers and graduate students who contributed ideas, information and transcription. We especially thank the Speech Language Pathologists in the Vancouver School Board and the Surrey School Board who made this project possible by identification of the participating children and collection of the language samples.

References

Acarlar, F. & Johnston, J. (2006). Computer-based analysis of Turkish child language: Clinical and research applications. *Journal of Multilingual Communication Disorders, 4*, 78-94.

Aksu-Koc, A. (1998) The role of input versus universal predispositions in the emergence of tense-aspect morphology: evidence from Turkish. *First Language, 54*, 255-280.

Bedore, L.M., & Leonard, L.B. (1998). Specific language impairment and grammatical morphology: A discriminant function analysis. *Journal of Speech, Language and Hearing Research, 41*, 1185-1192.

Bedore, L.M., & Pena, E.D. (2008). Assessment of bilingual children for identification of language impairment: Current findings and implications for practice. *International Journal of Bilingual Education and Bilingualism, 11*, 1-26.

Bedore, L.M., Pena, E.D., Gillam, R.B., & Ho, T.H. (2010). Language sample measures and language ability in Spanish-English bilingual kindergarteners. *Journal of Communication Disorders, 43*, 498-510.

Bond, S. & Schneider, P. (2005). Maze usage by typically developing and impaired children in a story context. *Poster presented at the convention of the American Speech and Hearing Association,* San Diego, CA.

Brown, R. (1973). *A first language: The early stages.* London: George Allen & Unwin.

Cleave, P.L., Giralometto, L.E., Chen, X., & Johnson, C.J. (2010). Narrative abilities in monolingual and dual language learning children with specific language impairment. *Journal of Communication Disorders, 43*, 511-522.

Fiestas, C.E., & Pena, E.D. (2004). Narrative discourse in bilingual children: Language and task effects. *Language, Speech, and Hearing Services in Schools, 35*, 155-168.

Genesee, F., Paradis, J., & Crago, M.B. (2004). *Dual language development and disorders: A handbook on bilingualism and second language learning.* Baltimore, MD: Paul H. Brookes Publishing Co.

Gillam, R. & Pearson, N. (2004). *The Test of Narrative Language.* Austin, TX: Pro-Ed.

Gutierrez-Clellan, V.F., Simon-Cereijido, G., & Wagner, C. (2008). Bilingual children with language impairment: A comparison with monolinguals and second language learners. *Applied Psycholinguistics, 29*, 3-19.

Hadley, P.A. (1998). Language sampling protocol for eliciting text-level discourse. *Language, Speech, and Hearing Services in Schools, 29*, 132-147.

Hakansson, G., & Nettelbladt, U. (1993). Developmental sequences in L1 (normal and impaired) and L2 acquisition of Swedish syntax. *International Journal of Applied Linguistics, 3*, 131-157.

Heilmann, J., Miller, J.F., Nockerts, A., & Andriacchi, K. (2006, June). Agreement and reliability of language sample analysis revisited. *Paper presented to the Symposium on Research and Child Language Disorders,* Madison, Wisconsin.

Heilmann, J., Nockerts, A. & Miller, J. (2010) Language sampling: Does the length of the transcript matter? *Language, speech and hearing services in schools, 41,* 393-404.

Hewitt, L.E., Hammer, C.S., Yont, K.M., & Tomblin, J.B. (2005). Language sampling for kindergarten children with and without SLI: Mean length of utterance, IPSYN, and NDW. *Journal of Communication Disorders, 38,* 197-213.

Johnston, J.R. (2001). An alternate MLU calculation: Magnitude and variability of effects. *Journal of Speech, Language, and Hearing Research, 44*, 156-164.

Johnston, J.R. (2006). *Thinking about child language: Research to practice.* Eau Claire, WI: Thinking Publications.

Johnston, J.R. (2008). Narratives: Twenty-five years later. *Topics in Language Disorders, 28*, 93-98.

Johnston, J.R. & Kamhi, A. (1984). The same can be less: Syntactic and semantic aspects of the utterances of language impaired children. *Merrill-Palmer Quarterly, 30*, 65-86.

Johnston, J.R., Miller, J.F., Curtiss, S., & Tallal, P. (1993). Conversations with children who are language impaired: Asking questions. *Journal of Speech and Hearing Research, 36*, 973-978.

Johnston, J.R. & Schery, T. (1976). The use of grammatical morphemes by children with communication disorders. In D.M. Morehead & A.E. Morehead (Eds.), *Normal and deficient child language* (pp. 239-258). Baltimore, MD: University Park Press.

Johnston, J.R. & Wong, M.Y. (2002). Cultural differences in beliefs and practices concerning talk to children. *Journal of Speech, Language and Hearing Research, 45*, 916-926.

Kemp, K. & Klee, T. (1997) Child language sampling practices: Results of a survey of speech language pathologists in the United States. *Child Language Teaching and Therapy, 13,* 161.

Klee, T., Stokes, S.F., Wong, A.M.Y., Fletcher, P., & Gavin, W.J. (2004). Utterance length and lexical diversity in Cantonese-speaking children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research, 47*(6), 1396-1411.

Kohnert, K. (2010). Bilingual children with primary language impairment: Issues, evidence and implications for clinical actions. *Journal of Communication Disorders, 43*, 456-473.

Laing, S.P., & Kamhi, A. (2003). Alternative assessment of language and literacy in culturally and linguistically diverse populations. *Language, Speech, and Hearing Services in Schools, 34*, 44-55.

Langdon, H.W. & Quintanar-Sarallana, R. (2003). Roles and responsibilities of the interpreter in interactions with speech-language pathologists, parents, and students. *Seminars in Speech and Language, 24*(3), 235-244.

Loban, W. (1976). *Language development: Kindergarten through grade twelve* (Research Report 18). Urbana, IL: National Council of Teachers of English.

MacLachlan, B.G., & Chapman, R.S. (1988). Communication breakdowns in normal and language learning-disabled children's conversation and narration. *Journal of Speech and Hearing Disorders, 53*, 2-7.

Masterson, J.J. & Kamhi, A.G. (1991). The effects of sampling conditions on sentence production in normal, reading-disabled, and language-learning-disabled children. *Journal of Speech and Hearing Research, 34*, 549-558.

McCarthy, D. (1943). Language development in the preschool child. In R.G. Barker, J.S. Kounin, H.F. Wright (Eds), *Child behavior and development: A course of representative studies* (pp. 107-128). New York, NY: McGraw-Hill.

Miller, J.F. & Chapman, R.S. (1981). The relation between age and mean length of utterance in morphemes. *Journal of Speech and Hearing Research, 24*, 154-161.

Miller, J.F., Heilmann, J., Nockerts, A., Iglesias, A., Fabiano, L., & Francis, D.J. (2006). Oral language and reading in bilingual children. *Learning Disabilities Research and Practice, 21*(1), 30-43.

Miller, J.F. & Iglesias, A. (2012). Systematic Analysis of Language Transcripts (SALT, Version 9) [Computer Software]. Madison, WI: Language Analysis Laboratory, University of Wisconsin.

Miller, J.F. & Leadholm, B. (1992). *Language sample analysis guide: The Wisconsin guide for identification and description of language impairment in children.* Wisconsin Department of Public Instruction.

Oller, D.K., & Eilers, R.E. (2002). *Language and literacy in bilingual children.* Clevedon: Multilingual Matters.

Paradis, J. (2005). Grammatical morphology in children learning English as a second language: Implications of similarities with specific language impairment. *Language, Speech and Hearing Services in Schools, 36*, 172-187.

Paradis, J. (2010). The interface between bilingual development and specific language impairment. *Applied Psycholinguistics, 31*, 227-252.

Paradis, J. (2011). Individual differences in child English second language acquisition: Comparing child-Internal and child-external factors. *Linguistic Approaches to Bilingualism, 1*(3), 213-237.

Paradis, J. & Crago, M. (2000). Tense and temporality: Similarities and differences between language-impaired and second-language children. *Journal of Speech, Language and Hearing Research, 43*(4), 834-848.

Pena, E.D., Gutierrez-Clellen, V., Iglesias, A., Goldstein, B., & Bedore, L.M. Bilingual English-Spanish Assessment (BESA). Manuscript in preparation.

Richards, B.J. & Malvern, D.D. (1997). *Quantifying lexical diversity in the study of language development.* Reading, U.K.: The University of Reading, New Bulmershe Papers.

Rice, M. & Wexler, K. (1996). Toward tense as a clinical marker of SLI in English-speaking children. *Journal of Speech, Language and Hearing Research, 39*, 1239-1257.

Rice, M. & Wexler, K. (2001). Test of Early Grammatical Impairment. New York: The Psychological Corporation.

Rispoli, M., & Hadley, P. (2001). The leading edge: The significance of sentence disruptions in the development of grammar. *Journal of Speech, Language, and Hearing Research, 44*, 1131-1143.

Rojas, R., & Iglesia, A. (2009). Making a case for language sampling: Assessment and intervention with (Spanish–English) second-language learners. *The ASHA Leader, 14*, 10- 13.

SALT Software (2010). Gillam Narrative Tasks Reference Database. Retrieved from http://www.saltsoftware.com/salt/downloads/referencedatabases.cfm.

Slobin, D. (1973) Cognitive prerequisites for the acquisition of grammar. In C. Ferguson & D. Slobin (Eds.), *Studies of child language development* (pp. 175-208). New York: Holt, Reinhardt & Winston.

Simmons, N. & Johnston, J.R. (2007). Cross-cultural differences in beliefs and practices that affect the language spoken to children: Mothers with Indian and Western heritage. *International Journal of Language and Communication Disorder, 42*(4), 445-465.

Simmons, N. & Small, J. A.. *A survey of speech-language pathology service to bilingual clients in British Columbia.* Manuscript in preparation.

Templin, M.C. (1957). *Certain language skills in children; Their development and inter-relationships*. Minneapolis, MN: University of Minnesota Press.

Thordardottir, E.T. (2005). Early lexical and syntactic development in Quebec French and English: Implications for cross-linguistic and bilingual assessment. *International Journal of Communication Disorders, 40*(3), 243-278.

Thordardottir E.T., & Ellis Weismer, S. (2002). Content mazes and filled pauses in the spontaneous speech of school-aged children with specific language impairment. *Brain and Cognition, 48*, 587-592.

Vancouver School Board (2011). About VSB. Retrieved from http://www.vsb.bc.ca/about-vsb.

Watkins, R.V., Kelly, D.J., Harbers, H.M., & Hollis, W. (1995). Measuring children's lexical diversity: Differentiating typical and impaired language learners. *Journal of Speech and Hearing Research, 38*(6), 1349-1355.

Xanthos, A., Laaha, S., Gillis, S., Stephany, U., Aksu-Koc, A., Christofidou, A.,... Dressler, W. (2011). On the rule of morphological richness in the early development of noun and verb inflection. *First Language, 31*(4), 461-479.

Zdorenk, T. & Paradis, J. (2011). Articles in child L2 English: When L1 and L2 acquisition meet at the interface. *First Language*, 1-25. DOI: 10.1177/0142723710396797

Authors' Note

Correspondence concerning this article should be addressed to Kate Ballem Chase, UBC School of Audiology and Speech Sciences, 2177 Wesbrook Mall, Vancouver, BC, V6T 1Z3. Email: kdballem@mail.ubc.ca.

Received date: November 4, 2011

Accepted date: March 18, 2012

Appendix

Language Use at Home:

Please answer the following questions to give us a better idea of your Grade 2 child's language experience. (You may write your answers in English or in your native language, whichever you prefer).

Birth date of Grade 2 Child: _____

1. What language do you use when you talk to your Grade 2 child at home?
2. If you have older children, what language do they use when talking to your Grade 2 child at home?
3. What language did your child learn first?
4. What language does your child usually use when he/she talks to you at home?
5. How well does your child speak his/her first (native, home) language?

   ___ Was later and slower than most children in learning to talk.

   ___ Just as well as most children of the same age.

   ___ Better than other children of the same age.

6. At what age did your Grade 2 child start to learn English?

   Where did he/she first learn English?

   ___ Preschool/Daycare ___ Babysitter ___ Other

7. How much English did your child speak when he/she first went to school in kindergarten in Canada?

   ___ Only a few words

   ___ Could put a few words together to say simple things about play activities, food, or TV programs.

   ___ Could talk about most topics, but made lots of mistakes and didn't always know the words he/she needed.

   ___ Knew a lot of English and used everyday.