

EXPLORING THE COMPLEX SYNTAX OF SCHOOL-AGED CHILDREN: THE
EFFECTS OF AGE, DISCOURSE AND BILINGUALISM

by

Peter T. Cahill

Submitted in partial fulfilment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
March, 2018

© Copyright by Peter T. Cahill, 2018

TABLE OF CONTENTS

LIST OF TABLES.....	iii
LIST OF FIGURES	iv
ABSTRACT.....	v
LIST OF ABBREVIATIONS USED	vi
ACKNOWLEDGEMENTS	vii
CHAPTER 1 INTRODUCTION.....	1
CHAPTER 2 REVIEW OF LITERATURE	3
2.1 Complex syntax.....	3
2.2 Notes on multilingualism.....	5
2.3 Eliciting complex syntax	8
2.4 Measuring complex syntax	12
2.5 Disordered complex syntax.....	15
2.5 Preliminary conclusions	25
CHAPTER 3 METHODS	28
3.1 Participants	28
3.2 Materials and procedures	28
3.3 Reliability	31
CHAPTER 4 RESULTS.....	32
4.1 Status variables	32
4.2 Experimental variables	34
4.2.1 Surface measures.....	35
4.2.2 Mean deployment of MOVE	37
4.2.3 Distribution of MOVE complexity.....	41
4.2.4 Error patterns.....	42
4.2.5 Strategies for managing complexity	43
CHAPTER 5 DISCUSSION	45
5.1 Effects of age	45
5.2 Effects of discourse.....	48
5.3 Effects of multilingualism.....	49
5.3.1 Mazing and errors	50
5.3.2 Multilingualism and syntax	52
5.4 Comparing measures of syntax	54
CHAPTER 6 CONCLUSION	57
6.1 Summary of findings	57
6.2 Limitations and future directions	58
BIBLIOGRAPHY	60
APPENDIX A CODING GUIDELINES FOR CLAUSAL DENSITY.....	69
APPENDIX B CODING GUIDELINES FOR MOVE.....	73

LIST OF TABLES

Table 1	Participant characteristics: group size, age, expressive vocabulary	32
Table 2	Bilingual participant characteristics: Language exposure and proficiency	32
Table 3	Surface measures – means and standard deviations	35
Table 4	Mean deployment of MOVE – means and standard deviations	40
Table 5	Mazing and errors patterns – means and standard deviations.....	43
Table 6	List of codes by clause type	69
Table 7	Examples of coordination coding guidelines.....	70
Table 8	Examples of let clause coding guidelines	71
Table 9	Examples of infinitive vs. non-finite clause coding guidelines	71
Table 10	Examples of progressive participial clause coding guidelines	71
Table 11	Examples of perfective participial clause coding guidelines	72
Table 12	Examples of quoted discourse coding guidelines	72
Table 13	Codes for movement.....	73
Table 14	Examples of argument movement coding.....	77

LIST OF FIGURES

Figure 1	Overall percentages of clause types for all participants by discourse type... 36
Figure 2	Ratio per group of children who used adjunct movement at least once in the sample 38
Figure 3	Ratio per group of children who used argument movement at least once in the sample 38
Figure 4	Ratio per group of children who used MOVE at least once in the sample (including argument and adjunct movement) 39

ABSTRACT

Purpose: The aim of the study was to evaluate established and novel metrics of syntactic development in school-age children, to compare elicitation tasks, and to evaluate the effects of bilingualism on measures of syntactic complexity.

Method: The participants were 48 children recruited from the province of Nova Scotia in four groups: monolingual and bilingual French-English children 7 and 12 years of age. The groups were compared using established and novel language sample measures on English language samples.

Results: Important syntactic development continues during the early school-age years. Established and novel metrics generally showed sensitivity to age and discourse type, but not bilingualism. The measures validated interest in the application of theoretical syntax to language assessment. Clinical and theoretical implications for syntax in developmental language disorders are discussed.

LIST OF ABBREVIATIONS USED

CD	Clausal density
DLD	Developmental language disorder
EVIP	Échelle de vocabulaire en images Peabody
LSA	Language sample analysis
MDM	Mean deployment of wh-movement
MLU	Mean length of utterance
MLU _m	Mean length of utterance in morphemes
MLU _w	Mean length of utterance in words
PPVT	Peabody Picture Vocabulary Test
SALT	Systematic Analysis of Language Transcripts
SLI	Specific language impairment
TD	Typically developing
TNL	Test of Narrative Language

ACKNOWLEDGEMENTS

I would like to express my overwhelming gratitude to Dr. Patricia Cleave for the countless hours of guidance and assistance she provided during this project, and for her dedication to cultivating an outstanding next cohort of clinician-researchers.

My thanks to Dr. Elissa Asp for the many hours spent discussing syntactic theory and its application to this project. I learned more in those discussions than in years of reading about syntax.

Thank you as well to Dr. Elizabeth Kay-Raining Bird for serving as the loyal opposition and for teaching my classmates and I that the amygdala is a critical processing centre for clinical thinking.

Finally, I would like to acknowledge the contributions of Dr. Ann Sutton, who provided many suggestions for this project in its early development.

My thanks to my colleagues and classmates for their emotional and intellectual support during this journey. I want to thank my colleagues in the Child Language Lab, Bonita Squires and Eleanor Campbell, for their contributions during this process. I would also like to thank my School of Human Communication Disorders colleague, Merlin.¹

Finally, I would like to thank those who supported me financially during this project. This work was supported by the Canadian Institutes of Health Research – Institute of Genetics [funding reference number ISU-146055] and the Social Sciences and Humanities Research Council of Canada Joseph Armand Bombardier Canada Graduate Scholarship-Master's. The Killam Pre-Doctoral Fellowships and Nova Graduate Scholarships also provided generous support for this project.

¹ This was revealed to me in a dream by courtesy of Crystal Primeau.

CHAPTER 1 INTRODUCTION

Complex syntax is known to be an area of difficulty for children with a language disorder (Penke, 2015). However, the exact presentation of these difficulties remains inadequately described (Scott & Koonce, 2014; Scott & Stokes, 1995). Additionally, researchers and clinicians are poorly served by the metrics currently available to measure and describe a child's syntactic abilities and performance. These measures are broad-brush and sometimes poorly sensitive to changes in the early school age years (Nippold, Hesketh, Duthie, & Mansfield, 2005; Nippold, Mansfield, Billow, & Tomblin, 2008). Recent scholarship has identified underlying areas of possible difficulty for children with language disorders, providing new possibilities for assessment and intervention (see Frizelle & Fletcher, 2014; Levy & Friedmann, 2009 for examples). These studies have predominately utilized laboratory tasks and have focused on clinical populations. To increase confidence in these findings, as well as sketch a clearer picture of age-appropriate expectations and candidacy for assessments and interventions, this study seeks to describe the syntactic abilities of typically developing mono- and bilingual children in an ecologically valid task – language sample analysis. By observing how children without a language disorder utilize complex syntax in comparatively natural contexts, the goal is to inform clinicians and researchers regarding what can be expected from school-aged children, as well as the most appropriate ways to measure a child's syntax. This study is interested in the following general research questions.

1. What differences exist between surface form measures and measures of underlying syntactic operations?

2. How do children's syntactic competencies develop during the school-age years when measured with different metrics?
3. To what extent does children's use of syntax differ according to task demands?
4. What effect does learning two languages have on measurements of children's syntactic performance?
5. What, if any, strategies do children use to manage complexity in their spoken production?

CHAPTER 2 REVIEW OF LITERATURE

2.1 Complex syntax

As children acquire language, they must learn how to begin to combine words into longer and longer sentences to express more complex thoughts about their world, eventually reaching an adult-like grammar (Turnbull & Justice, 2012). A sentence consists of units called constituents arranged in a hierarchical order, which underlies the eventual temporal ordering of words and morphemes (Carnie, 2007). In this way, all sentences can be thought of as complex and rule-governed, so it is necessary to define complex syntax and the scope of this study. What complexity must a sentence demonstrate to be considered an example of complex syntax?

Complex syntax can be defined as sentences formed by multiple clauses (a sentence element containing one verb phrase (Turnbull & Justice, 2012)) joined together by embedding or coordination (Paul & Norbury, 2012). The present study further restricts the definition by Paul and Norbury (2012), as that definition includes coordinated multi-clausal sentences (sentences consisting of two equally weighted clauses, usually connected by ‘and’ or ‘but’). Coordination is the earliest developing method to combine two clauses into a single sentence (Paul, 1981), as well as the simplest (de Villiers & de Villiers, 1985). Due to the simplicity and early emergence of coordination, the current study focuses on sentences with subordinate clauses. A final approach used in some studies (such as Salameh, Håkansson, & Nettelbladt, 2004; van der Lely, Jones, & Marshall, 2011; van der Lely & Marshall, 2011), defines complex sentences as those containing non-local operations. Non-local operations refer to syntactic phenomena which are best explained through the hierarchical relationships between constituents, and

are very poorly explained by relationships between constituents immediately adjacent in the temporal spell-out of a sentence. For example, in a sentence such as “Who are you going to the concert with?” children must learn that there is a relationship between the “who” and the “with” despite being separated by six intervening words. As children acquire language, their grammar expands in its ability to process these non-local operations across larger syntactic distances. This study defines complex syntax as both multiclausal utterances that contain at least one case of subordination and/or a significant non-local operation, such as wh-movement.

Complex syntax appears in at least rudimentary form early in the language acquisition process, although researchers disagree on the extent of children’s mastery of these constructions by the end of the preschool years (de Villiers & de Villiers, 1985). Paul (1981) found that, by the age of 4, the children in her sample had acquired verbal complement clauses, finite embedded wh-clauses, multiple embedding, and infinitive clauses that share a subject with the matrix clause. The children were still acquiring gerundial, unmarked infinitive, wh-infinitive, and relative clauses. Relative clauses have been found to occur, albeit sometimes in reduced forms, as early as 2 years of age (de Villiers & de Villiers, 1985; Diessel & Tomasello, 2005). Despite an early ability to produce relative clauses, children’s performance on receptive tasks with relative clauses does not reach adult-like performance by 6 years of age (de Villiers & de Villiers, 1985). Over several studies, Diessel and Tomasello (2005) found that by three years of age children can produce well-formed relatives, although only those subordinated to copular clauses and elliptically isolated nouns. As children approach the age of 5, they become more and more likely to subordinate a relative clause within more complex matrix

clauses (Diessel & Tomasello, 2005). The picture that these data paint suggests that, as children enter the school age years, they are becoming adept at a wide variety of complex syntactic structures, but that there remains significant room for growth as they continue to build upon their language systems. Using surface measures, a large, well-powered language sample study demonstrated that syntactic development continues from 8 years until middle adulthood (Nippold et al., 2005). Not present in the literature are studies focused on syntactic development in the transitional period of the early school years. The goal of the present study is to better describe the syntactic development that takes place in the early school-age years (7 to 12 years of age) to assist clinicians and researchers in their understanding of the role of syntax in the transition from learning language to “language for learning” (Paul & Norbury, 2012).

2.2 Notes on multilingualism

The proportion of individuals using more than one language daily is increasing in Canada. Bilingualism in English and French reached a historic high point of 17.9% in the 2016 census (Statistics Canada, 2017a). Those who use a language other than English, French, First Nations, Métis, or Inuit languages (e.g. Tagalog, Mandarin, Spanish, Urdu, Punjabi, Hindi, etc...) at home have also increased to 20.1% of the population (Statistics Canada, 2017c). Over 19% of the total population uses at least two languages at home on a daily basis (Statistics Canada, 2017b). Speakers of non-official languages remain heavily concentrated in the largest metropolitan areas of the country (Statistics Canada, 2012, 2017b), with 1.8 million Toronto residents using primarily a non-official language at home in 2011 (Statistics Canada, 2012). Despite considerable variation in the prevalence of bilingualism in communities across Canada, significant bilingual

populations are present in all provinces and territories (Statistics Canada, 2017a, 2017b, 2017c). Additionally, retention of mother languages by many First Nations, Métis, and Inuit communities exceeds 70% (Statistics Canada, 2017b). Using these figures as a rough guide, Canadian clinicians can expect about 2 out of every 5 children to be exposed to two or more languages, although this will vary greatly by community. In keeping with the terminology used by Statistics Canada (2017b) to capture the wide variety of language abilities and experiences of Canadians, the terms *multilingualism*, *bilingualism*, and *dual-language learning* will be used interchangeably, although many subgroups and distinctions can be made within and across these terms (see de Jong, 2008).

Language assessment for multilingual children remains a challenge for clinicians. Standardized assessment tools developed for the language are often not accessible, and translations and adaptations frequently do not meet psychometric standards (Bedore & Peña, 2008). Standardized testing in English only has also been found to have poor specificity for multilingual children (Cleave, Girolametto, Chen, & Johnson, 2008; Paradis, 2016; Peets & Bialystok, 2015). Both over- and under-identification of multilingual children in preschool and beyond has been repeated findings in the literature (Bedore & Peña, 2008; Paradis, Genesee, & Crago, 2011; Salameh, Nettelbladt, Hakansson, & Gullberg, 2002). Paradis (2016) has argued that standardized assessments with monolingual norms are inappropriate for dual language learners even after years of exposure to English. Bedore and Peña (2008) point out the double bind in which clinicians find themselves without appropriate assessment tools for these children. For a child whose profile is ambiguous between a language disorder and multilingual language development, clinicians can err on the side of caution, over-diagnose, and provide

services where not necessary. Alternatively, they can delay diagnosis or discharge children, failing to provide services to children in need. Both situations are problematic, and improved assessment tools for multilingual language development could add to clinicians' toolboxes for investigating complex cases.

The assessment of morphology and syntax in multilingual language development is also complex. In English, clinical markers for DLD and the language development profiles of sequential bilinguals have been found to overlap (Paradis, 2010). Clinical markers for DLD also vary cross-linguistically, and some established areas of difficulty in English are not impaired in other languages and vice-versa (Leonard, 2014). Bedore and Peña (2008) recommended paying close attention to clinical markers in both languages and Paradis (2016) also noted that specific aspects of the inflectional system of English appear particularly vulnerable in multilingual children with DLD. However, in both of these overviews, the discussion focuses predominantly on morphosyntax, and not complex syntax. To date, no studies have been found regarding the assessment of complex syntax for multilingual children.

Strategies for improving assessment in multilingual contexts include assessing in all languages and/or using modified assessment techniques in the clinician's language (Paradis, 2016). The latter is likely more feasible in many clinical contexts. Suggestions for modified assessment techniques include testing language structures known to be extremely difficult for children with DLD (Bedore & Peña, 2008) and focusing on language-general, rather than language-specific features (Paradis et al., 2011). Language-general features are those structures or operations of a language system which are more likely to be shared among languages whereas language-specific features are those parts

which are more likely to be idiosyncratic to the language. An example relevant to the present study is the language-specificity of morphosyntax in English and French compared to the more language-general nature of relative clause constructions. English and French not only differ in the types of inflectional morphology that are licenced and required, but the inflectional morphology that is challenging for children with DLD also differs between these two languages (Crago, Paradis, & Menn, 2009). In contrast, French and English are both head-initial languages and share five out of six relativizable positions (Keenan & Comrie, 1977). The structures of relative clauses in English and French are far more similar, and therefore better candidates for consideration as language-general features, than are the inflectional systems of two languages, which are more language specific. To provide some guidance as to where to begin the search for assessment techniques for complex syntax in multilingual children, these suggestions from the literature can be combined. Researchers should attempt to locate structures and/or operations which are highly difficult for children with DLD and are also more likely to be language-general.

2.3 Eliciting complex syntax

Language sample analysis (LSA) involves presenting a communicative situation to a child (typically conversation with the examiner about the child's interests or recent experiences), recording the interaction, and subsequently transcribing and analysing the child's language production. LSA has the benefit of providing data which are relatively naturalistic when compared to controlled experimental tasks. A sample is assumed to be at least partially representative of the child's language use in similar, naturalistic contexts. For example, a conversational sample would represent the child's ability to use

language in dialogue, whereas a narrative sample would demonstrate the child's abilities to create and effectively communicate a monologue which recounts a specific story or experience.

This methodology has several advantages for researchers and clinicians. It can go beyond experimental tasks and show whether children do in fact use these structures in naturalistic speech, and if they do, with what frequency and degree of mastery. LSA is also recognized as a clinical technique for speech-language pathologists to assess, select appropriate intervention goals, and evaluate outcomes of intervention (Hesketh, 2004; Paul & Norbury, 2012). Using LSA to explore the development of syntax can demonstrate the ecological validity of the structures of interest, as well as provide information regarding typical performance on these structures. Performance by TD children on language samples can then be used as a control against which clinical samples can be compared. For example, SALT currently offers some normative databases for clinicians and researchers (Miller & Iglesias, 2012), and other databases are in development. However, complex syntax has not been a focus in the development of these databases, reducing their utility as a clinical tool in this specific domain of language. Additionally, language sample analysis has been found to be a promising assessment tool for multilingual children (Ooi & Wong, 2012), and to show greater specificity for these children compared to standardized norms-based measures (Cleave et al., 2008; Peets & Bialystok, 2015).

LSA also has important pitfalls. It is difficult to control and isolate linguistic variables. Also, there is the risk that the child will simply not produce targets of interest at all, despite having the underlying competence to do so. Whereas inflectional morphology

has obligatory contexts as determined by the syntactic frame of any given sentence and will have at least some exemplars in a language sample of nominal length, complex syntax has no such obligatory contexts and is produced as a matter of syntactic choice rather than obligation (Scott & Stokes, 1995). Marinellie (2004) utilized 100 utterance conversational samples from children between 10 and 11 years of age and noted the conspicuous sparseness of complex syntactic constructions. She went as far as to directly question the content validity of conversational samples for syntactic analysis. However, a different elicitation task may increase the chances of a child using complex syntax. The increasing complexity of language content (the intended message) is thought to drive increases in the complexity of language form, including syntax (Balthazar & Scott, 2015; Nippold, 2016). Therefore, other discourse genres with more complex content may prove useful for eliciting more complex syntax.

Two options for increasing the odds of eliciting complex constructions include narrative discourse, where the child is asked to produce or retell a story, or expository discourse, where the child is asked to communicate factual or technical information to the examiner. Narratives have been shown to have great clinical utility to the assessment of grammar, both for diagnostic (Cleave et al., 2008; Guo & Schneider, 2016) and generalization purposes (Hesketh, 2004). Narrative samples have been shown to elicit greater syntactic complexity than conversational samples (Nippold et al., 2014). Expository discourse is also a strong candidate for eliciting syntactic complexity. A large cross-sectional study of English-speaking monolingual children, teens, and adults from 8 through 40 years of age showed that syntactic complexity was significantly greater for an expository task than a conversational one across all ages (Nippold et al., 2005). The same

task was also used to compare the syntactic production of typically developing children to those with DLD and languages disorders with differentiating conditions, showing that the task was sensitive to these differences (Nippold et al., 2008). All three groups in this study demonstrated greater syntactic complexity on the expository task than a conversational sample (Nippold et al., 2008). Children with language disorders increased their production of complex syntax when presented with an expository task, presumably to match the increased cognitive complexity of language content, although they were not as successful as their typically developing peers in doing so (Nippold et al., 2008). Thus, the evidence suggests that both narrative and expository discourse are more appropriate for the assessment of complex syntax than conversation.

It is unclear how narrative and expository discourse compare in the elicitation of complex syntax. Using written samples, researchers found that argumentation (which can be considered a type of expository discourse) elicited longer T-units than narrative tasks based on the same visual stimuli (Crowhurst, 1980; Crowhurst & Piche, 1979). In contrast, Nippold and colleagues (2015) had children retell a fable (a narrative task), and then complete a critical thinking task regarding the fable they had just retold (an expository task). The narratives yielded significantly higher MLU than did the critical thinking tasks, although clausal density did not differ significantly between the discourse types. However, the narrative was a retell task, which has been shown to provide scaffolding or priming for narrative productions (Lofranco, Peña, & Bedore, 2006). To date, the literature does not provide a clear picture as to which discourse genre is the most promising for eliciting complex syntax. This question is explored in the present study.

Finally, narrative and expository discourse texts have the advantage of closely

imitating the types of language use required of children in school environments (Nippold, 2016; Peets & Bialystok, 2015; Petersen & Petersen, 2016). Therefore, narratives and expository tasks are promising as measuring of functional language use for children in the language for learning stage. These elicitation tasks are promising as methods to evaluate how language deficits may impact performance on school-based tasks and affect overall school success, which are well-documented sequelae of paediatric language disorders (Catts, Fey, Tomblin, & Zhang, 2002; Johnson, Beitchman, & Brownlie, 2010).

2.4 Measuring complex syntax

There are several, widely-used surface measures of syntactic complexity used to analyze language sample data. The most widely recognized measure is likely mean length of utterance (MLU). There are numerous variations of this metric, including mean length of T-unit, mean length of C-unit, and mean length of utterance, and each of these can be measured in words, morphemes, or syllables. This metric has been suggested to be the most reliable measure, showing the most consistent gains from 8 to 44 years of age (Nippold et al., 2005) and differentiating between DLD and TD (Nippold et al., 2008). However, there are several critical issues with this particular measure. First, MLU has limited face validity as a measure of syntactic complexity. It is a measure of productivity and not complexity. A high MLU suggests that the child's sentences are complex, but only the length of the utterance is measured, and its complexity is simply inferred. For example, see the following two utterances.

(1) My brother did this really crazy thing yesterday.

(2) What my brother did yesterday was really crazy.

Both of these sentences have an MLU of 8 words; however, they are not equally

complex. The first sentence is simple as it consists of a single clause. The second sentence is not only a complex sentence with two clauses joined by subordination, but it also features an object-gapped, headless relative clause. MLU is unable to recognize the greater complexity of the second sentence. A second pitfall of MLU is that it is not particularly useful for treatment planning. It can suggest no other treatment targets beyond expanding the length of the utterance. MLU tells clinicians nothing about the types or relative presence of various constructions in the child's current linguistic repertoire, nor which may be emerging or stimulable. For these reasons, MLU appears to be a poor gold standard.

Clausal density (CD) is a frequently used metric that yields a mean number of clauses per utterance. Clausal density has also been shown to demonstrate gradual, age-related changes from 8 to 44 years (Nippold et al., 2005); however, CD dipped in this study around thirteen years of age. While this dip has only been found in a single study, this is one of the most sufficiently powered LSA studies to date. As such, this result is somewhat concerning regarding the validity of CD. Further, mean length of T-unit and relative clause use were found to show more robust age related gains in this study, suggesting that CD is not as sensitive as these other measures (Nippold et al., 2005). In addition to these studies of English-speaking children, a study with French-speaking children calculated several different versions of CD (Tuller, Henry, Sizaret, & Barthez, 2012). Results followed similar patterns across these measures, which distinguished between French-speaking children with a language impairment (10;11 to 15;7, with a mean of 12;6) and typically developing children who were on average three years younger (Tuller et al., 2012). These studies suggest that clausal density may be a useful

metric, but that more research is required to determine which method of calculating it yields the most robust results. Additionally, CD does not suggest any particular treatment targets beyond expanding the use of any type of subordination.

Another approach is to count subordinate clause types based on their role in the matrix clause. Subordinate clauses that replace noun phrases are called nominal phrases; those that replace adverbs are adverbial phrases, and those that modify noun phrases are called relative clauses. An advantage of this approach is that it is straightforward and does not require an advanced understanding of syntactic theory. However, this approach does not account for the varied complexities and forms that these clauses can display. Additionally, only relative clause use has been shown to be sensitive to age when this approach is used (Nippold et al., 2005). Relative clause use was observed to decrease from 8 years to dip at 13 (the early school years), before subsequently increasing into middle adulthood (Nippold et al., 2005), although this may have been an artifact. Additionally, only relative clause use has been found to be sensitive to the presence of a language disorder when using this system (Nippold et al., 2008). This system does not appear to be promising for assessing a child's syntactic system, although results from this system emphasize the potential of relative clause structures.

A more exhaustive clause profiling system has been used to classify all English subordinate clauses into fourteen categories based on the clauses' specific structures (Arndt & Schuele, 2013). This system was used for a case study of a child with developmental language disorder (Schuele & Dykes, 2005). While this provides much more detail than counting clauses based on their role in the matrix clause, it is a labour-intensive system. Further, it is dependent on exemplars of specific structures appearing in

the language sample. As such, more finely tuned metrics that can be calculated based on a variety of clause types may be more desirable. In other words, a measure which can be used to “average” syntactic complexity across a wide variety of specific embedded structures would be preferable because of its general nature and decreased dependence on the presence of relatively infrequent structures.

It is clear that best practices in the measurement of complex syntax are far from established and that more work is needed to evaluate the most reliable, valid, and efficient ways to measure a child’s syntactic abilities. The evidence for expressive syntactic interventions is currently mixed (Law, 2004). The lack of truly valid and rigorous methods to measure syntax may contribute in part to the poor quality of the evidence for these interventions. This justifies the creation and testing of new measures of syntax, based either on developmental literature, or on research into grammatical impairments in paediatric language disorders. The current study takes the latter approach.

2.5 Disordered complex syntax

Some children have difficulties with the acquisition of language which do not resolve overtime without intervention. These difficulties can be associated with a larger disorder complex, such as Down syndrome or autism spectrum disorder; however, some children demonstrate difficulties with learning language without any known aetiological factor. This is known as developmental language disorder (DLD) (Bishop et al., 2017), formerly called specific language impairment or primary language impairment, among other terms (see Bishop, 2014; and Reilly et al., 2014). Pre-schoolers’ difficulties with expressive morphosyntax in English are perhaps the most well-studied area of developmental language disorder (Leonard, 2014a) and are one of the hallmarks of this

disorder in the literature (Cummings, 2013). It is generally recognized that, as these children enter the school years, these classic symptoms typically resolve, but that residual difficulties with language form remain (Scott & Koonce, 2014; Scott & Stokes, 1995). Complex syntactic constructions are thought to be an area of difficulty (Fletcher, 2009; Scott & Koonce, 2014; Scott & Stokes, 1995). While childhood language disorders associated with conditions such as Down syndrome, Williams syndrome, and autism spectrum disorder also demonstrate syntactic difficulties of interest to researchers and clinicians (Clahsen, 2008; Fortunato-Tavares et al., 2012, 2015), scholarly attention on disordered syntax has focused on children with DLD, presumably due to hypothesized deficits in the neurological architecture of the language system. Despite this scholarly attention, the aetiology remains a mystery (Bishop, 2014; Reilly et al., 2014). Although the present study only includes typically developing children, the literature on syntactic difficulties in DLD is reviewed here. This is because analyses were developed from what is known about disordered syntax in order to restrict the number of analyses and increase the likelihood of clinically useful findings. If a structure which is thought to be vulnerable to DLD can be shown to be mastered or emerging in typically developing children, the structure may prove useful to clinicians as a clinical marker. Difficulties demonstrated beyond what is found in typically developing children increase confidence that the errors are not age appropriate and are rather indicative of a language disorder.

Marinellie (2004) profiled syntax within conversational language samples and found that children with DLD used significantly fewer adverbial, relative and coordinate clauses, as well as multiple embeddings, than typically developing (TD) children. Schuele and Dykes (2005) explored the emergence of complex syntax by collecting

language samples from a child diagnosed with DLD over five years. Using their exhaustive clause profiling system, they determined that syntax emerges in the same general pattern as would be expected from typical development, but with a delay of a few years. The least observed structures across all samples and so presumably latest developing structures were non-finite wh-clauses, non-subject relative clauses, and participial clauses. Three errors patterns that are not reported in the typically developing literature were present: omission of obligatory infinitival *to*, omission of obligatory relative markers and pronouns, and omission of embedded wh-pronouns. Another study compared DLD and TD children on a handful of syntactic measures, and only found a significant difference between the groups for mean length of T-unit, and then only in expository discourse (Nippold et al., 2008). Tuller et al. (2012) observed that French-speaking adolescents with DLD produced less syntactically dense samples, and that their rates of subordination were lower than typically developing children. The children with DLD did not appear to avoid any particular subordination construction, instead producing fewer examples of all types of subordination. However, the DLD group showed strong preferences for certain sub-types of these constructions. When producing relatives, DLD children were much more likely to opt for copular or existential matrix clauses without significant semantic value. Only half of these children ever attempted a relative clause embedded within a matrix clause other than a copular or existential clause. This pattern mirrors the relative clause use patterns of early versus late preschool children in English (see Diessel & Tomasello, 2005). Children with DLD also produced more morphosyntactic errors in embedded clauses, although this finding only trended towards significance.

Van der Lely and colleagues have demonstrated that, at least for children with a language impairment characterized by significant grammatical difficulties, DLD causes difficulties in several areas of non-linear syntactic processing, including: anaphor (van der Lely & Stollwerck, 1997), passive sentences (Marinis & Saddy, 2013) and wh-questions (Marinis & van der Lely, 2007; van der Lely & Battell, 2003; van der Lely et al., 2011). The evidence for difficulties with wh-questions is particularly strong, as it has been identified by other studies in English (de Villiers, de Villiers, & Roeper, 2011; Deevy & Leonard, 2004), French (Jakubowicz, 2011; Stavrakaki, Chrysomallis, & Petraki, 2011), Italian (Cipriani, Bottari, Chilosi, & Pfanner, 1998), Greek (Stavrakaki, 2006; Stavrakaki et al., 2011), Hebrew (Friedmann & Novogrodsky, 2011), Swedish (Hansson & Nettelbladt, 2006), German (Hamann, Penner, & Lindner, 1998), and Cantonese (Wong, Leonard, Fletcher, & Stokes, 2004). The error patterns observed in these studies vary considerably across languages, and even across participants. However, errors seem particularly likely whenever wh-questions are involved. Studies have also identified relative clauses as an area of difficulty in English (Frizelle & Fletcher, 2014; Marinellie, 2004; Schuele & Nicholls, 2000), French (Tuller et al., 2012), Italian (Cipriani et al., 1998), Hebrew (Novogrodsky & Friedmann, 2006), Danish (Jensen de López, Sundahl Olsen, & Chondrogianni, 2014), Swedish (Håkansson & Hansson, 2000), and Dutch (Zwitserlood, Wijnen, van Weerdenburg, & Verhoeven, 2015). Children with DLD also demonstrate difficulties with topic-comment structures in Mandarin Chinese (Yu, 2016). A parsimonious explanation would account for difficulties with these various structures across such a diversity of languages. One such explanation is that children with DLD have difficulty with structures where elements are embedded, and then displaced to

another location in the syntactic tree – MOVE according to the minimalist program (Corrêa & Augusto, 2011). According to the system developed to explain agrammatic aphasia in English (Thompson & Shapiro, 2005), children with DLD have difficulties when attempting wh-movement structures (also called A'-movement). If MOVE were the locus of these difficulties, this would provide an explanation for these difficulties in every language mentioned above, with the exception of Cantonese which remains unexplained (Wong et al., 2004). Although research does suggest that children also have difficulties with NP-movement (Jensen de López et al., 2014; Marinis & Saddy, 2013), there are relatively fewer studies exploring this type of movement, so the evidence is less clear. Given the evidence available, the present study is predominately interested in exploring wh-movement.

Studies also find that, although they perform better than children with DLD, typically developing children continue to have some difficulties with wh-movement operations during the early school years. For example, van der Lely and Battell (2003) found trace location and wh-word effects even in typically developing control groups. Jensen de López and colleagues (2014) found that typically developing children were less accurate for relative clauses where the trace was in the object position of the embedded clause rather than in the subject position. Frizelle and Fletcher (2014) found that performance by DLD and TD groups generally worsened as the trace location moved deeper into the phrase structure of the embedded clause. Diessel and Tomasello (2005) observed a similar pattern of results in both English and German on experimental tasks with typically developing children, despite how these two languages diverge in how they structure relative clauses. Even fluent adult speakers of English have been found to

produce object-gapped relatives more slowly than subject-gapped ones, and to maze more frequently when attempting the construction (Scontras, Badecker, Shank, Lim, & Fedorenko, 2014), showing that these processes are psycholinguistically complex even for mature grammars. Additionally, Frizelle and Fletcher (2014) found that the presence of significant semantic information in the matrix clause affects accuracy in the construction of relative clauses by TD children, which has also been found for children with DLD (Frizelle & Fletcher, 2014; Tuller et al., 2012). In the case of typical development, the clearest findings appear to be that the location of the trace in the syntactic tree and the semantic content of the matrix clause can impact the accuracy of wh-movement production. Because these factors affect the performance of typically developing children and fluent adult speakers, they are likely even more difficult for children with DLD.

There are many theoretical explanations of wh-movement difficulties for children with DLD. These include: treating obligatory movement operations as optional (van der Lely & Battell, 2003), difficulties with the acquisition of complementizers and subordinate clause structure (Håkansson & Hansson, 2000), difficulties with the assignment of thematic roles (Novogrodsky & Friedmann, 2006), difficulties with intervening candidates for dependencies (Jensen de López et al., 2014), difficulties with feature checking (de Villiers et al., 2011), processing demands dependent upon the number of MERGE and MOVE operations required (Jakubowicz, 2011), and processing demands combined with structural syntactic features as well as frequency of input (Frizelle & Fletcher, 2014). As these theoretical explanations are focused on disordered syntax, not all provide clear predictions regarding the course of non-impaired relative

clause development. However, three theories provide predictions which may be extended to the acquisition of wh-movement in typical development.

According to the account proposed by Jensen de López, Sundahl Olsen, and Chondrogianni (2014), Danish-speaking TD children will demonstrate sensitivity to intervening candidates (see examples 3 and 4) for a long-distance dependency, and may attempt several strategies to avoid these situations, specifically by passivizing object relatives to “smuggle” the theme noun phrase past the intervening agent noun phrase before wh-movement occurs. This account is bolstered by the finding that adult English speakers also attempt passivization as a frequent strategy in lieu of producing true object relatives (Scontras et al., 2014). This account would be confirmed by a higher proportion of younger children producing passivized object relatives than the older cohort, who would be more likely to simply produce an un-passivized object relative. It should be noted that this account would also be consistent with processing accounts of developmental language disorder, as it would minimize the working memory requirements by reducing the temporal distance between the moved constituent and the trace(s).

Presence (object relative fragment) versus absence (passivized relative fragment) of intervening candidate noun phrase (Jensen de López et al., 2014, p. 61)

(3) The girl_{*t*} that the grandmother_{*intervening candidate*} kissed *t*.

(4) The girl_{*t*} that *t* is kissed by the grandmother.

According to the account provided by Jakubowicz (2011) called the Derivational Complexity Hypothesis, the multiple derivations required to avoid an object relative with

passivization would predict that the younger cohort would be less likely to produce passivized relatives, opting instead for simple object relatives when compared to the older children. A single, long-distance movement (example 5) would be derivationally less complex than transforming the embedded clause by A-movement, and then performing a second, interclausal A'-movement to finally place the moved constituent into the matrix clause (example 6). This account appears to be more consistent with domain-specific explanations of developmental language disorder, although it does also include computational complexity concerns. According to this perspective, the processing capacities may be overwhelmed by an attempt to compensate for a weaker underlying linguistic system. As such, children would avoid multiple instances of movement, and would not use a combination of A-movement and A'-movement to turn an object relative into a subject relative, especially in younger age ranges. Presumably, children would either simply produce the object relative, or avoid producing the structure at all, perhaps by producing several simple sentences without embedding to communicate the same semantic information. Passivization has been found to be a strategy used by English-speaking adults (Scontras et al., 2014), suggesting that some factor besides the co-occurrence of A- and A'-movement is at work.

Single trace object relative fragment with one trace/movement versus passivized relative fragment with two traces/movements

(5) The girl_{*t*} that the grandmother kissed *t*.

(6) The girl_{*t*} that *t* is kissed *t* by the grandmother.


Finally, according to the general conclusions of Frizelle and Fletcher (2014), the

syntactic position of the trace, as well as the presence of semantic content beyond a simple copular construction (called relatives versus clefts in the acquired disorder literature, such as Thompson & Shapiro, 2006) would affect a child's ability to produce the construction. This account would predict that younger children produce a higher number of clefts when compared to relatives. The role of the syntactic location of the trace is somewhat more complex. Frizelle and Fletcher (2014) discuss the possibility that the accuracy observed for subject relatives is influenced by the canonical order of these sentences. Ignoring the complementizer would allow children to reconstruct the appropriate thematic relations between the agent and theme without needing to process any type of movement (see examples 7 and 8 below, as opposed to the non-canonical order of example 9). However, there are substantial reasons to suspect that it is not canonical word order that causes this subject-relative advantage. If canonical word order underlies this production advantage, presumably children would show an object-relative advantage in head-final languages, such as Mandarin Chinese, as object relatives follow canonical word order and allow the reconstruction of appropriate thematic relations without processing the wh-movement in these languages (see examples 10-12 below). The opposite has been observed, as even in Mandarin Chinese children appear to use and master subject-relatives before object-relatives (Hsu, Hermon, & Zukowski, 2009). This provides cross-linguistic evidence that distance in the syntactic tree can provide an underlying explanation for this phenomenon. Combining these findings, we predict that the distance moved in the syntactic tree by a wh-moved constituent will show age-related increases, with subject gaps being more prevalent in samples from younger children, and object and oblique gaps becoming increasingly common in samples with age. This

account includes both domain-general and domain-specific concerns, and would be consistent with processing capacities being unable to fully compensate for difficulties with more complex syntactic constructions.

(7) Plain sentence – The boy saw a squirrel.


(8) Subject gap - The boy_{*t*} that *t* saw a squirrel.



(9) Object gap - A squirrel_{*t*} that the boy saw *t*.



(10) Plain sentence - 那个 男孩 看见 一只 松鼠。


Na-ge nanhai kanjian yi-zhi songshu.


That-CL boy saw a-CL squirrel.


(11) Subject gap- *t* 看见 一只 松鼠 的 男孩 *t*_o.


t kanjianyi-zhi songshu de nanhai_{*t*}.


t saw a-CL squirrel DE boy_{*t*}.


(12) Object gap - 那个 男孩 看见 *t* 的 松鼠 *t*_o.


Na-ge nanhai kanjian *t* de songshu_{*t*}.


That-CL boy saw *t* DE squirrel_{*t*}.


All of these accounts have been developed and tested in experimental laboratory probes. While these provide rigorous controls to answer theoretical questions, there are also drawbacks to generalizability and comparison across studies. The present study seeks to address this gap in the literature and contribute to the debate by studying wh-movement via language sample analysis. Additionally, given the well documented

difficulties that children with DLD have with wh-movement, the emerging mastery of the operation by typically developing children, as well as the presence of wh-movement in several types of syntactic constructions (including wh-questions and relative clauses), we hypothesize that measuring wh-movement in language samples may provide a more sensitive language-general measure of syntactic development than surface structure measures, such as MLU and clausal density. If so, it would also be more useful for treatment planning than established metrics, as it may help determine candidacy for treatment options, such as discourse-based syntactic intervention (To, Lui, Li, & Lam, 2015), MetaTaal (Zwitserlood et al., 2015), traditional metalinguistic syntactic treatments (To et al., 2015), or the paediatric adaptation of the Treatment of Underlying Forms (Levy & Friedmann, 2009) which was originally developed for acquired language disorders (Thompson, 2008; Thompson & Shapiro, 2005).

2.5 Preliminary conclusions

The present study is interested in the development of complex syntax in the early school years, and how multilingualism and task effects may affect measurement. One goal of the study is to determine the method of measuring syntactic development that demonstrates the most robust age-related gains, and appears to be maximally language-general or an area of extreme difficulty for children with DLD. Another goal is to determine whether either discourse type yields consistently greater syntax. It is hoped that a combination of these factors may inform best practice in the assessment of complex syntax in research and/or clinical practice. The research questions are revisited below, and preliminary hypotheses provided.

- 1. What differences exist between surface form measures and measures of underlying*

syntactic operations?

We predict that measures of underlying form will be more sensitive to age-related changes in syntactic complexity and less likely to demonstrate effects of multilingualism than surface measures.

- 2. How do children's syntactic competencies develop during the school-age years when measured with different metrics?*

We expect measures based on movement to be the most robust to age-related differences and least sensitive to language status effects. Syntactic depth is predicted as the most critical variable affecting children's performance on these sentences, as evidenced by greater frequency of object and oblique gaps in samples from 12-year-old children. Further, we predict that clausal density calculated with all clause types will be more robust to age-related gains than clausal density based on only one or a few types of clauses. Finally, MLU is predicted to be the least sensitive measure.

- 3. To what extent does children's use of syntax differ according to task demands?*

Expository discourse is predicted to elicit greater syntactic complexity as measured by all metrics.

- 4. What effect does learning two languages have on measurements of children's syntactic performance?*

Due to the scarcity of research on complex syntactic development in multilingual children, it is difficult to make specific predictions. The English syntax of monolingual English-speaking and bilingual English-French-speaking children will be compared and

contrasted based on both surface measures and underlying forms metrics. Possible divergences between these measures will also be explored.

5. *What, if any, strategies do children use to manage complexity in their spoken production?*

We predict that younger children will be more likely to use forms such as passive subject relatives as opposed to object relatives to reduce the syntactic complexity of these structures. Additionally, measures of performance (mazing and errors) are predicted to increase in tandem with syntactic complexity.

CHAPTER 3 METHODS

3.1 Participants

Forty-Eight English monolingual and English-French bilingual children aged 7-8 and 11-12 years were recruited in Nova Scotia. Twelve children were recruited for the monolingual and bilingual 7-year-old groups; 13 for the bilingual 12-year-old group, and 11 for the monolingual 12-year-old group. All monolingual children attended school in English and spoke English at home. All bilingual children attended school in French and were exposed to French and English at home, although to varying degrees. A parent questionnaire was used to inquire about language use and exposure. Participants were recruited through their schools, and had no known communication or learning difficulties, as established by parental report.

3.2 Materials and procedures

The children participated in language assessment and sampling sessions with members of the Dalhousie University Child Language Lab from the School of Communication Sciences and Disorders as part of a larger protocol. All examiners were trained in the administration of the tasks. The monolingual children participated in only one session; the bilingual children participated twice, with one session in each language. The order of administration of the languages was counterbalanced. Each session consisted of four parts: a vocabulary test, two narrative tasks, an expository task and 10 minutes of conversation. All sessions began with the vocabulary test, and the order of the remaining tasks was counterbalanced. As even monolingual individuals in Canada are exposed to French, the French receptive vocabulary skills of the monolingual English-speaking children were assessed at the end of sessions. In this study, only English

language narrative and expository samples were analyzed.

The Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4) (Dunn & Dunn, 2007), and its *Québécoise* adaption *Échelle de vocabulaire en images Peabody* (EVIP) (Dunn, Thériault-Whalen, & Dunn, 1993) were used to assess receptive vocabulary. The narrative task included an adaptation of the Test of Narrative Language (TNL) (Gillam & Pearson, 2004) and story stems. TNL involves the elicitation of narratives in three tasks. For this study, the data from only one task was used. In this task, the children were shown a complex picture and the examiner told a story about the picture. The children were then shown another complex picture that is designed to elicit a similar story, and they were asked to tell a story about the new picture. This tell-retell technique is thought to provide scaffolding and priming to increase the children's performance on the task (Lofranco et al., 2006). In the story stem tasks, the examiner began a story and then prompted the child to complete the story in whatever way seems best. The TNL was in revision at the time, and the revised unpublished version was used. With the permission of the authors and the publisher, the research team developed an English version of the Spanish adaptation of the TNL (Gillam, Peña, & Bedore, n.d.). The English narratives developed from the Spanish version were tested for equivalence with the English TNL. Both English versions were then translated into French. Two story stem prompts were used, each of which were paired with a version of the TNL. The story stems always followed the TNL story retell task.

A modification of the favourite game or sport task described in Nippold, Kesketh, Duthie, and Mansfield (2005) was used for the expository samples. The children were asked to explain their favourite game or sport, and were given time to prepare. The

modification involved providing paper with prompts for the many parts of the question (the point of the game, how to score the game, when the game ends... etc), and the child was allowed to take notes on the paper, if desired. A French translation of the instructions and prompt sheet were developed. In the second session, bilingual children were prompted to describe a game or sport other than the one they described in the first session.

The children's English language samples were audio-recorded and later transcribed by trained graduate students in speech-language pathology following SALT conventions (Miller & Iglesias, 2012). Of relevance to the current study, these conventions provided the template for decision making regarding utterance segmentation and the presence of mazing and errors. Transcripts were further coded for both surface measures and underlying syntactic structures separately. Clauses were identified and coded according to their verbal morphology into the categories of: finite, marked infinitive, unmarked infinitive, progressive participial, and perfective participial clauses. Occurrences of wh-movement were identified and coded as to whether they involved movement of an argument or an adjunct out of the subordinated clause. Cases of argument movement were further coded according to: syntactic depth of the co-indexed trace (6 levels) and the semantic content of the matrix clause (2 levels). Additionally, subject and direct-object gaps were coded for transitivity (subject) and animacy (object). This yielded 15 possible codes for instances of MOVE. Due to too few exemplars in many categories, several were collapsed prior to statistical analyses. The syntactic depth code was collapsed from 6 levels to 3 (subject, object, oblique). The transitivity and animacy distinction was also ignored. This consolidation yielded 7 codes. For guidelines

and examples of both novel coding systems, see appendices A and B.

3.3 Reliability

The audio files were transcribed according to SALT conventions (Miller & Iglesias, 2012) by trained graduate students. The first author reviewed all of the transcripts twice, once listening to the audio to check the transcription and a second time to ensure that SALT conventions were followed for words, utterance segmentation, mazing, and errors.

To evaluate coding reliability of clausal density, 17% of samples were re-coded by a trained doctoral student (two samples per age group and per language group). 93% agreement in total presence and type of clause was found between the two coders. Movement was also re-coded in 17% of the samples, yielding 83% exact agreement of presence and type in expository samples and 64% in narratives. An inspection of disagreements revealed that 31% of disagreements involved whether embedded *when* and *where* clauses featured movement or if these were conjunctive adverbs generated *in-situ*. Based on discussion and a review of the literature, the coding guidelines for *when* and *where* clauses were revised. The remaining 83% of the samples were reviewed by both coders using the new coding guidelines. Agreement for these remaining samples before discussion was 96%. After consensus was reached, agreement was 100%.

CHAPTER 4 RESULTS

4.1 Status variables

Status variables were tested to confirm that the participants met the assumptions of the study design. Scores for receptive vocabulary measures for the participants can be found in Table 1. For the bilingual participants, parental report of language exposure and language proficiency in all four modalities can be found in Table 2.

Table 1 Participant characteristics: group size, age, expressive vocabulary

Group ¹	n	Age ²	PPVT-4 SS ³	EVIP SS ⁴	EVIP AE ⁵
M7	12	91.08 (4.64)	108.00 (11.05)		25.83 (1.75)
M12	11	140.64 (8.61)	104.91 (9.26)		30.30 (5.60)
B7	12	94.00 (5.49)	113.75 (19.82)	108.75 (15.43)	102.17 (19.82)
B12	13	138.23 (6.91)	112.38 (15.95)	105.08 (13.68)	164.00 (67.91)

Table 2 Bilingual participant characteristics: Language exposure and proficiency

Group ¹	n	Proportional English exposure	Percent English lifetime exposure	Percent French exposure	Speak English ⁶	Understand English ⁶	Read English ⁶	Write English ⁶
B7	13	40.75 (23.05)	85.71 (34.99)	96.50 (9.31)	3.77 (0.93)	4.15 (0.56)	3.85 (0.80)	4.38 (0.87)
B12	12	39.41 (24.93)	78.21 (38.57)	95.82 (9.35)	3.92 (1.24)	4.00 (1.04)	5.00 (1.28)	5.67 (1.07)

Key for Tables 1 and 2

¹Group designations

M7 – English monolingual 7 and 8 year olds

M12 – English monolingual 11 and 12 year olds

B7 – French-English bilingual 7 and 8 year olds

B12 – French-English bilingual 11 and 12 year olds

²Chronological age in months

³PPVT-4 SS – PPVT-4 standard score

⁴EVIP SS – EVIP standard score

⁵EVIP AE – EVIP age-equivalent score in months

⁶Proficiency in English relative to French on a 7-point scale with 1 = only English, 4 = English and French equivalent, 7 = only French

A two-way ANOVA with age in months as the dependent variable found a significant main effect for age group ($F(1,44) = 614.819, p = .000, \eta_p^2 = .933$), while the main effect of language status was not significant ($F(1, 44) = .018, p = .893, \eta_p^2 = .000$), nor was the interaction effect ($F(1,44) = 1.980, p = .166, \eta_p^2 = .043$). A two-way ANOVA with PPVT-4 standard scores as the dependent variable failed to find significant main or interaction effects ($p > .12$). These findings confirm that the groups differed according to age, but that they did not differ significantly in their English language abilities, as measured by a test of semantic development.

A final two-way ANOVA with the same fixed factors and the age-equivalent EVIP scores as the dependent variable found significant main effects for both age ($F(1,43) = 9.172, p = .004, \eta_p^2 = .176$) and language status ($F(1,43) = 92.049, p = .000, \eta_p^2 = .682$). The interaction effect was also significant ($F(1,43) = 6.867, p = .012, \eta_p^2 = .138$). Post-hoc independent t-tests were used to compare each group along the factors of age and language status. Significantly greater age-equivalent scores were found for the 12-year-olds, both for monolingual ($t(20) = -2.625, p = .016$) and bilingual participants ($t(23) = -3.033, p = .006$). Significantly greater scores were also found for bilingual participants when compared to monolingual participants in both the younger ($t(22) = 13.292, p = .000$) and older groups ($t(21) = 6.176, p = .000$). The bilingual group clearly demonstrated a significantly larger receptive French vocabulary than the monolingual group. This score increased with age in the bilingual group as expected. The significant difference in age-equivalent scores in the monolingual group was not expected, as this suggests that there has been a growth in French receptive vocabulary in the monolingual participants as well. However, there are many cognates between English

and French (such as *ambulance*) on the test, so increasing English vocabulary might also improve scores on the French adaptation of the test. Nova Scotian curriculum also includes core French as a mandatory subject. Additionally, the 12-year-old monolingual children achieved a mean age-equivalent score of 2.4 years of age, demonstrating that this statistically significant growth in French receptive vocabulary has limited clinical significance. These children have very limited French competence as measured by this test of receptive vocabulary. These findings increase confidence that there were not significant unplanned differences among the groups, and that the between-subjects factors of age and language status differed according to the intended study design.

4.2 Experimental variables

Statistical analyses were performed to explore the English syntax of school-age children and evaluate the impacts of age, language status, and discourse type. Age and language status (monolingual versus bilingual) were set as the between groups factors, and the within-subjects factor was defined as discourse type (expository versus narrative tasks). For all analyses, the narrative tasks adapted from the TNL (Gillam & Pearson, 2004), as well as the story stem tasks, are included as a single narrative task to increase the amount of data. As some structures included in these analyses are relatively infrequent in spontaneous speech, the narrative samples were combined to increase the chances of these structures appearing in the sample. Combining the two narrative tasks into a single sample is appropriate as only microstructural measures were calculated for the present study. The expository samples consisted of the favourite-game-or-sport task only.

4.2.1 Surface measures

Means and standard deviations for surface measures (mean length of utterance in morphemes and five calculations of clausal density) can be found in Table 3 separated by group and discourse type.

Table 3 Surface measures – means and standard deviations

Group ¹	Expository						Narrative					
	MLU	CD _a	CD _b	CD _c	CD _d	CD _e	MLU	CD _a	CD _b	CD _c	CD _d	CD _e
M7	9.36	1.30	1.63	1.63	1.68	1.68	8.24	1.19	1.33	1.36	1.37	1.43
<i>SD</i>	<i>1.93</i>	<i>0.28</i>	<i>0.41</i>	<i>0.41</i>	<i>0.41</i>	<i>0.41</i>	<i>1.49</i>	<i>0.13</i>	<i>0.17</i>	<i>0.18</i>	<i>0.17</i>	<i>0.16</i>
B7	8.05	1.14	1.46	1.46	1.49	1.49	7.53	1.11	1.21	1.25	1.23	1.29
<i>SD</i>	<i>2.49</i>	<i>0.48</i>	<i>0.65</i>	<i>0.65</i>	<i>0.65</i>	<i>0.65</i>	<i>1.32</i>	<i>0.24</i>	<i>0.25</i>	<i>0.26</i>	<i>0.27</i>	<i>0.26</i>
M12	10.92	1.41	1.76	1.78	1.80	1.82	9.47	1.41	1.63	1.66	1.71	1.77
<i>SD</i>	<i>2.32</i>	<i>0.37</i>	<i>0.44</i>	<i>0.45</i>	<i>0.48</i>	<i>0.49</i>	<i>1.39</i>	<i>0.30</i>	<i>0.40</i>	<i>0.39</i>	<i>0.41</i>	<i>0.39</i>
B12	10.26	1.40	1.66	1.66	1.69	1.70	9.28	1.37	1.52	1.55	1.55	1.59
<i>SD</i>	<i>1.91</i>	<i>0.23</i>	<i>0.25</i>	<i>0.25</i>	<i>0.26</i>	<i>0.26</i>	<i>1.74</i>	<i>0.24</i>	<i>0.24</i>	<i>0.24</i>	<i>0.25</i>	<i>0.25</i>

Key for Table 3

¹Group designations

M7 – English monolingual 7 and 8 year olds

M12 – English monolingual 11 and 12 year olds

B7 – French-English bilingual 7 and 8 year olds

B12 – French-English bilingual 11 and 12 year olds

MLU – mean length of utterance in morphemes

CD_a – clausal density with finite clauses only

CD_b – clausal density with finite and marked infinitive clauses

CD_c – clausal density with finite, marked infinitive, and unmarked infinitive clauses

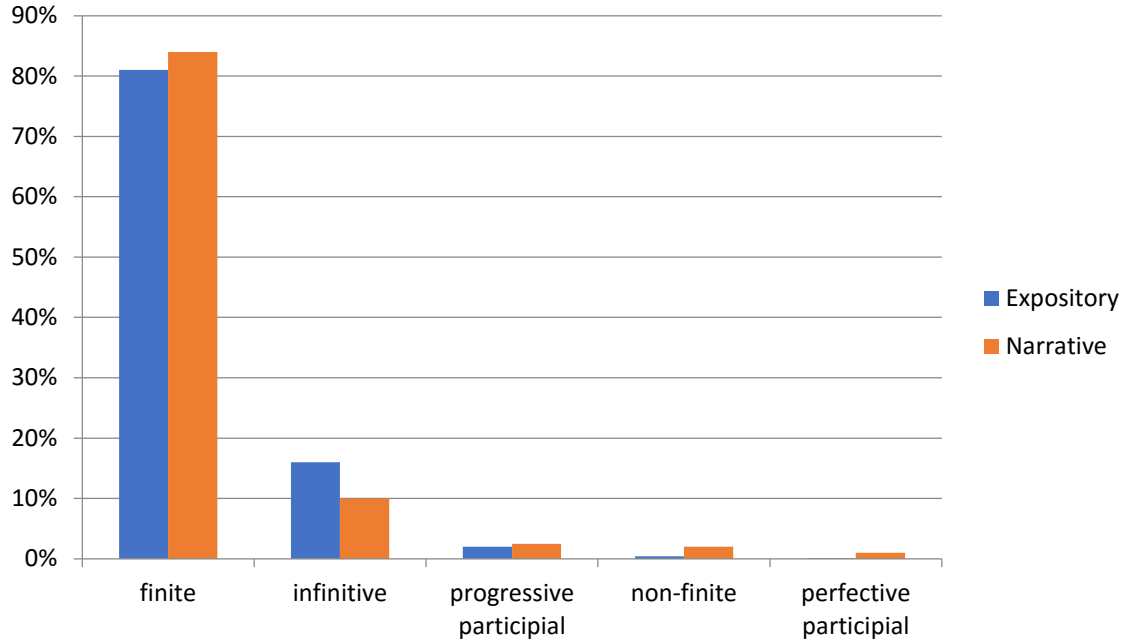
CD_d – clausal density with finite, marked infinitive, and progressive participial clauses

CD_e – clausal density with all clause types

The vast majority of clauses produced by all participants were finite, with marked infinitive clauses comprising most of the remaining clauses in the data set. Combined, these clause types make up 95% of clauses produced in the study. Unmarked infinitives, as well as progressive and perfective participial clauses made up less than 5% of all clauses produced. The overall percentage of clause types from all participants divided by

discourse can be found in Figure 1.

Figure 1 Overall percentages of clause types for all participants by discourse type



To evaluate the influence of age, language status, and discourse type on surface measures of syntactic complexity, a series of 2x2x2 mixed model ANOVAs were run. The first series of analyses were performed using the following as the dependent variables: mean length of utterance in morphemes (MLU_m); clausal density calculated only with finite clauses (CD_a); clausal density calculated with both finite and overtly marked infinitive clauses (CD_b); clausal density calculated with finite, marked infinitive, and unmarked infinitive clauses (CD_c); clausal density calculated with finite, marked infinitive, and progressive participial clauses (CD_d); and clausal density calculated with finite, marked infinitive, unmarked infinitive, progressive participial clauses, and perfective participial clauses (CD_e). Results are as follows. For MLU_m , main effects were found for discourse with the expository samples having higher MLU ($F(1,43) =$

10.577, $p = .002$, $\eta_p^2 = .197$) and age group with the 12 year olds producing longer utterances ($F(1,43) = 14.279$, $p = .000$, $\eta_p^2 = .249$). The main effect of language group and all interaction effects were not significant ($p > .1$). For CD_a, only a main effect for age group reached significance, with a greater clausal density in samples by the 12 year olds ($F(1,43) = 9.308$, $p = .004$, $\eta_p^2 = .178$). All other effects were not significant ($p > .2$). All other versions of clausal density found the same patterns of significance, with main effects for discourse (expository > narrative) and age group (12 year olds > 7 year olds): for CD_b, discourse ($F(1,43) = 10.314$, $p = .003$, $\eta_p^2 = .193$), age ($F(1,43) = 6.789$, $p = .013$, $\eta_p^2 = .136$); CD_c, discourse ($F(1,43) = 7.581$, $p = .009$, $\eta_p^2 = .150$), age ($F(1,43) = 7.164$, $p = .010$, $\eta_p^2 = .143$); for CD_d, discourse ($F(1,43) = 9.702$, $p = .003$, $\eta_p^2 = .184$), age ($F(1,43) = 7.294$, $p = .010$, $\eta_p^2 = .145$); for CD_e, discourse ($F(1,43) = 5.628$, $p = .022$, $\eta_p^2 = .177$), age ($F(1,43) = 7.254$, $p = .010$, $\eta_p^2 = .144$). All other effects were non-significant ($p > .1$), with the exception of CD_e, where a trend towards significance was observed for language group ($F(1,43) = 2.872$, $p = .097$, $\eta_p^2 = .063$), with better performance by the monolingual children.

4.2.2 Mean deployment of MOVE

Most of the children were found to use movement at least once in their samples. The following figures show the proportions of children within a group to use at least one exemplar of adjunct (Figure 2), argument (Figure 3), or any type of movement (Figure 4). There were more instances of argument movement compared to adjunct movement across the data set.

Figure 2 Proportion per group of children who used adjunct movement at least once in the sample

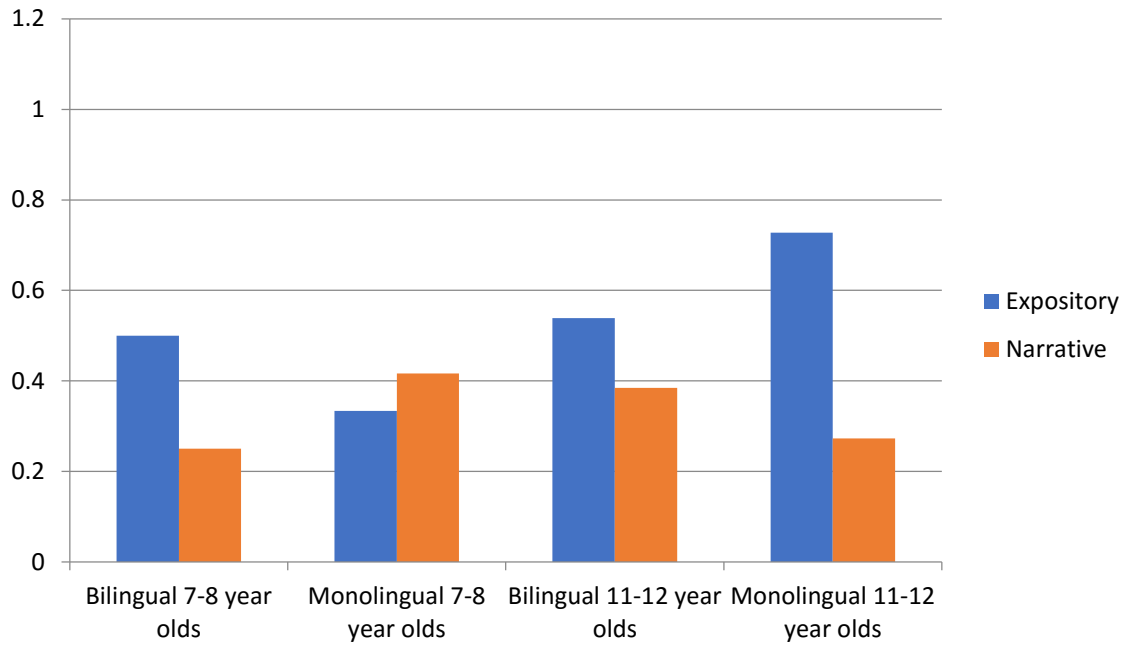


Figure 3 Proportion per group of children who used argument movement at least once in the sample

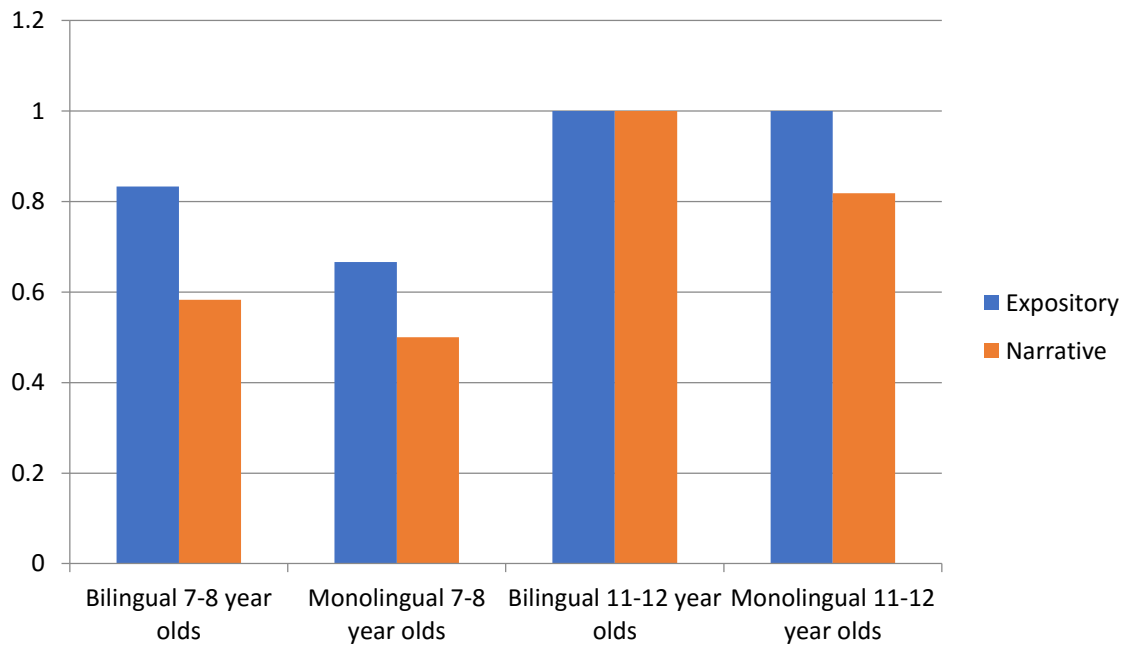
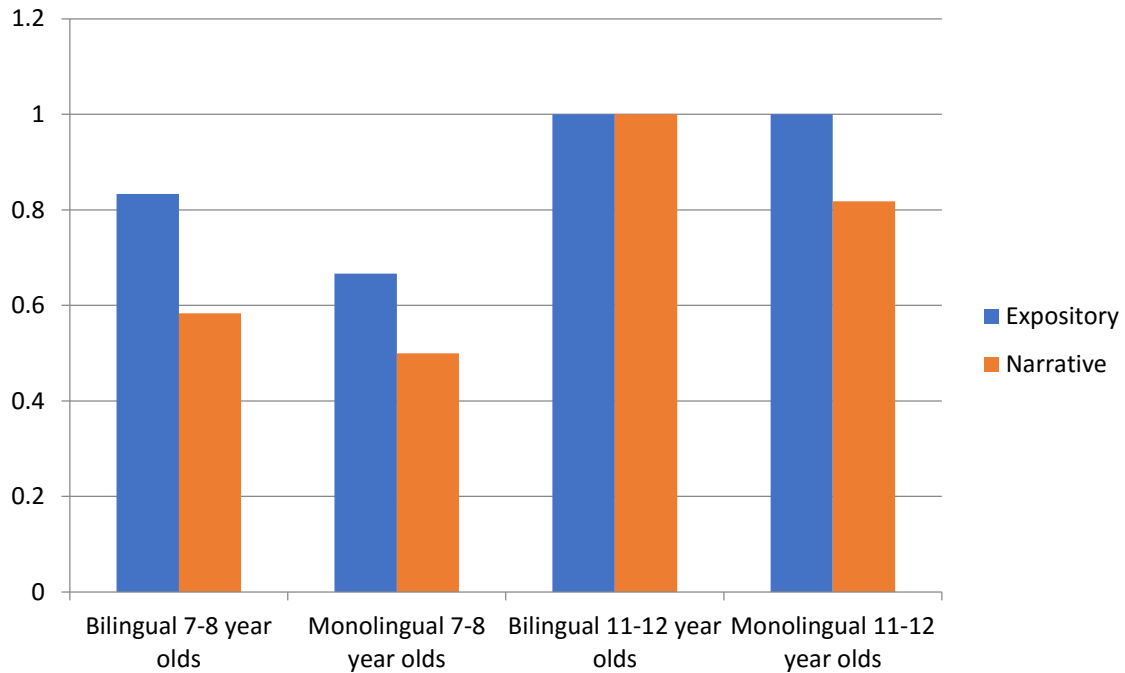


Figure 4 Proportion per group of children who used MOVE at least once in the sample (including argument and adjunct movement)



The means and standard deviations of the use of adjunct, argument and combined movement instances over the total number of utterances can be found in Table 4. To evaluate whether the average use of MOVE varied with the independent variables, another series of 3-way ANOVAs were run with the same between and within-subjects factors that were used for the surface measures. The first set of analyses explored whether average use of MOVE showed significant differences among the groups. These metrics are referred to as mean deployment of MOVE (MDM). MDM_a was calculated by finding the average number of MOVEs per utterance where the moved constituent was an adjunct in the subordinate clause. There was a significant main effect for discourse type ($F(1, 43) = 13.119, p = .001, \eta_p^2 = .247$), with greater presence of moved adjuncts in expository discourse, with no main effects for age ($F(1, 43) = .083, p = .776, \eta_p^2 = .002$) or language status ($F(1, 43) = 3.568, p = .066, \eta_p^2 = .077$). MDM_b was calculated as the average

number of moved constituents per utterance where the argument of the verb in the subordinate clause was moved. This yielded main effects for discourse type ($F(1,43) = 14.313, p = .000, \eta_p^2 = .250$) where argument movement was more frequent in expository discourse than in narration, and for age ($F(1, 43) = 8.302, p = .006, \eta_p^2 = .162$) with the 12 year olds using more argument movement than the 7 year olds. Mean deployment of move was also calculated combining argument and adjunct movement (MDM_c). The results mirrored argument movement, with significant main effects for discourse ($F(1, 43) = 21.482, p = .000, \eta_p^2 = .333$) and age group ($F(1, 43) = 4.579, p = .038, \eta_p^2 = .096$). Average use of any type of movement was greater in expository discourse when compared to narrative discourse, and movement of an argument in the subordinate clause was more frequent in samples by the 12 year olds compared to the 7 year olds. All other main effects and interaction effects for these three metrics were not significant ($p > .1$).

Table 4 Mean deployment of MOVE – means and standard deviations

Group ¹	Expository			Narrative		
	Adjunct ²	Argument ³	Combined ⁴	Adjunct ²	Argument ³	Combined ⁴
M7	0.0525	0.0760	0.1285	0.0191	0.0315	0.0506
<i>SD</i>	<i>0.0506</i>	<i>0.0919</i>	<i>0.1293</i>	<i>0.0276</i>	<i>0.0357</i>	<i>0.0477</i>
B7	0.0284	0.0996	0.1280	0.0135	0.0558	0.0692
<i>SD</i>	<i>0.0373</i>	<i>0.0852</i>	<i>0.0985</i>	<i>0.0249</i>	<i>0.0672</i>	<i>0.0745</i>
M12	0.0630	0.1269	0.1899	0.0167	0.0662	0.0828
<i>SD</i>	<i>0.0690</i>	<i>0.0702</i>	<i>0.1281</i>	<i>0.0297</i>	<i>0.0581</i>	<i>0.0792</i>
B12	0.0315	0.1449	0.1764	0.0123	0.0716	0.0839
<i>SD</i>	<i>0.0336</i>	<i>0.0650</i>	<i>0.0697</i>	<i>0.0239</i>	<i>0.0426</i>	<i>0.0503</i>

Key for Table 4

¹Group designations

E7 – English monolingual 7 and 8 year olds

E12 – English monolingual 11 and 12 year olds

B7 – French-English bilingual 7 and 8 year olds

B12 – French-English bilingual 11 and 12 year olds

²Moved constituent is an adjunct of the verb in the relative clause

³Moved constituent is an argument of the verb in the relative clause

⁴Moved constituent is either an adjunct or argument of the verb in the relative clause

4.2.3 Distribution of MOVE complexity

To evaluate whether the distribution of complexity within MOVE utterances differed with any of the independent variables, a series of χ^2 and McNemar tests were run. These analyses were chosen as there appeared to be too few exemplars of MOVE to yield a reliable weighted mean of complexity. Instead, the presence/absence of varying degrees of complexity were explored. These tests examined the location of the gap/trace in the subordinate clause and the semantic complexity of the matrix clause. As language group had not been shown to significantly affect the average use of argument movement in samples, this distinction was ignored in these analyses. Instead, groups were collapsed across language group and discourse type in the χ^2 tests to test the effect of age only. To compare whether the complexity of MOVE differed between discourse types, McNemar tests ignoring age and language group were conducted. Bonferroni corrections were used to correct for multiple tests for both the χ^2 and the McNemar tests. For syntactic position, the corrected p-value was (three levels, $.05/3 = .016$) $p = .016$ and for semantic complexity of the matrix clause, the corrected value was (two levels, $.05/2 = .025$) $p = .025$.

For the effect of age, subject ($\chi^2 (1, N = 48) = 6.701, p = .010, V = .374$) and object ($\chi^2 (1, N = 48) = 7.111, p = .008, V = .385$) MOVE were found to significantly differ in their distribution, with the older children's samples being more likely to use these structures. Oblique MOVE did not differ significantly ($\chi^2 (1, N = 48) = 4.148, p = .042, V = .294$), likely due to the few exemplars of oblique movement across the data set. The presence of true relatives was significantly higher in the 12-year-old group ($\chi^2 (1,$

$N = 48$) = 9.600, $p = .002$, $V = .447$); however, cleft structures did not differ in their presence across age groups ($\chi^2(1, N = 48) = .137$, $p = .712$, $V = .053$).

For the effect of discourse type, true relatives were more likely to be present in expository discourse ($p = .006$, $V = .481$). Clefts did not differ significantly in their presence/absence across discourse type ($p = .150$, $V = .296$). The presence of subject ($p = .019$, $V = .225$), object ($p = .052$, $V = .114$), and oblique movement ($p = .359$, $V = .081$) did not significantly differ between discourse types, although with a less conservative correction, subject movement would have been more likely to occur in expository samples.

4.2.4 Error patterns

Samples were also analyzed for evidence of performance difficulties, which were measured by frequency per utterance. Particularly, the rates of mazing and errors were of interest. Means and standard deviations for these measures can be found in Table 5.

Three-way mixed ANOVAs were used to explore mazing and error patterns. The average frequency of mazing of any type per utterance showed main effects for discourse type ($F(1,43) = 31.563$, $p = .000$, $\eta_p^2 = .423$) and language group ($F(1,43) = 7.728$, $p = .008$, $\eta_p^2 = .152$), with greater frequency of mazing in expository discourse and in samples by bilingual children. All other main and interaction effects were not significant ($p > .3$). To maximize power, morphological, lexical, and syntactic errors were all collapsed into a single error category. The main effects of age ($F(1,43) = 4.725$, $p = .035$, $\eta_p^2 = .099$) and language status ($F(1,43) = 11.819$, $p = .001$, $\eta_p^2 = .216$) were significant; however, the two-way interaction between age and language status was also significant ($F(1,43) = 5.052$, $p = .030$, $\eta_p^2 = .105$). To follow up on this interaction effect, two 2-way mixed

ANOVAs were run, one with age as the between-subjects variable, and one with language status. Only the main effect of language group was significant ($F(1,43) = 9.042$, $p = .004$, $\eta_p^2 = .167$) with a greater frequency of errors by the bilingual children, whereas discourse and the interaction effect were not significant ($p > .18$). The within-subjects variable (discourse type) and its interaction effects were not significant ($p > .2$). The analysis of age group did not yield any significant main or interaction effects ($p > .08$). Originally, an inferential analysis had been planned to compare mazing and error rates in the context of MOVE to the overall rate of mazing and errors to evaluate whether the increased complexity of these constructions would lead to decreased performance. Few exemplars and floor effects made it inappropriate to carry out these analyses.

Table 5 Mazing and errors patterns – means and standard deviations

Group ¹	Expository		Narrative	
	Mazes/utterance	Errors/utterance	Mazes/utterance	Errors/utterance
M7	0.5108	0.0238	0.3308	0.0716
<i>SD</i>	<i>0.2536</i>	<i>0.0323</i>	<i>0.1557</i>	<i>0.0876</i>
B7	0.7073	0.1469	0.4682	0.1593
<i>SD</i>	<i>0.2274</i>	<i>0.1395</i>	<i>0.2525</i>	<i>0.1129</i>
M12	0.4927	0.0509	0.2945	0.0475
<i>SD</i>	<i>0.1995</i>	<i>0.0392</i>	<i>0.1921</i>	<i>0.0642</i>
B12	0.6562	0.0812	0.5192	0.0651
<i>SD</i>	<i>0.2901</i>	<i>0.0643</i>	<i>0.3495</i>	<i>0.0655</i>

Key for Table 5

¹Group designations

M7 – English monolingual 7 and 8 year olds

M12 – English monolingual 11 and 12 year olds


B7 – French-English bilingual 7 and 8 year olds


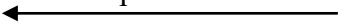
B12 – French-English bilingual 11 and 12 year olds


4.2.5 Strategies for managing complexity

Based on our interpretation of the data from Frizelle and Fletcher (2014), it was hypothesized that children would be sensitive to the syntactic depth of gap location and use strategies to avoid movement across longer distances. One strategy is using passives

to “smuggle” noun phrases into the subject position of the embedded clause (Jensen de López et al., 2014), allowing a shorter interclausal movement operation to occur. We hypothesized that children would use this strategy in our samples, especially in the younger age group. There is little evidence for this in the data. None of the occurrences of argument movement in narrative samples could be considered passive versions of object gaps. Only 4 instances (3.3%) could be considered examples of passivization in expository sample subject MOVEs. Additionally, the frequency of double extraction, where a constituent is moved at least twice (see examples 13 and 14), was greater in both discourse types (8.4% for expository and 1.1% in narrative) than was the frequency of passivization. One child specifically produced the transitive verb “smell” instead of the intransitive version, even though the transitive version requires an object MOVE instead of a subject MOVE (example 15). Using passive constructions within embedded clauses to reduce the syntactic distance of movement operations as a strategy to reduce complexity was not supported by the data.

(13) C *the mistake_t* that (many) I have seen *t* make people make *t* is they will

 put it too close to the enemy.

(14) C the other player ((randomly because the first time you have no idea))
 randomly start/3s and put/3s four peg/s of *whichever colour_t* they think *t*

 the other person/z code is *t*.


(15) C and then the *something good_t* he smell/ed *t* was a bear.


CHAPTER 5 DISCUSSION

This study uses English language samples from monolingual English and bilingual English/French school-aged children to explore the development of syntax from 7 to 12 years old, focusing on how age, type of discourse, and bilingualism affect the measurement of syntactic development. It also compares measures of syntax to explore best-practices in assessment for both research and clinical contexts.

5.1 Effects of age

Slow, age related gains were observed for all metrics of syntactic complexity, which is consistent with the literature on the development of school-aged syntax (Nippold, 2016; Scott, 1988). The 12-year olds produced more complex syntax as measured by all metrics than the 7-year olds, and their group performance is separated by moderate effect sizes, although the effect size of MLU_m ($\eta_p^2 = .249$) was notably larger than for other metrics. Although the groups differed significantly, the difference between the means in absolute terms was small – fewer than two morphemes across between age groups and about 0.2 for CD measures.

The proportion of children using argument movement at least once in their samples within each group increased with age, from about half of the younger children to almost all of the older children. The older children were also much more likely to attempt both subject and object movement structures than the younger cohort. The presence of oblique movement, which is hypothesized to be the most complex structural configuration included in this study, did not differ significantly with age, although this may be due to a small number of exemplars in all groups. Adjunct movement did not display age-related changes, suggesting that these types of movement do not follow the

same developmental trajectory. Attempting at least one relative structure also increased with age, whereas cleft structures were not observed to differ between the samples from older and younger children. These results support a general growth in the children's willingness to use MOVE during the early school years. The difference in the use of relative structures as opposed to clefts also differed more significantly with larger effect sizes between the age groups than did the position of the trace in the subordinate clause, suggesting that the semantic load of the matrix clause is more difficult for younger school age children than is a long-distance MOVE. However, no particular structure was found to appear only in the older children. This suggests that there are not specific "milestones" in the acquisition of complex movement structures in this age-range. These results support a generally increasing willingness to attempt movement, and to do so with more complex structures as the children age.

These results were group trends, and there was substantial individual variation, especially within the 7-year-old group, with some children producing no instances of movement, and other producing relative clauses with oblique extraction, as well as instances of movement across several clauses. While the likelihood of children attempting the more complex constructions did increase with age, the younger children produced some of the most complex structures included in this study. This suggests that some children have already acquired the necessary syntactic competence to produce these structures as they enter school, but others are significantly less likely to attempt complex constructions when compared to their 12-year-old counterparts.

Because this study used a language sampling methodology and did not require children to use any particular syntactic structures, the absence of certain structures in the

samples of some children cannot be interpreted as evidence that these children cannot produce the target structures. The lack of exemplars may be due to absence of the structure in the grammar, but it also may be due to the optionality of the task. More tightly controlled elicitation tasks may be able to clarify whether expecting highly-developed proficiency with movement is a truly age-appropriate expectation for very young school-aged children, or whether this study has captured the varied timing of the emergence of using complex MOVE, with some 7 year children having a certain facility with the operation that others do not. The present methodology cannot distinguish between these explanations, and the results were consistent with both hypotheses.

All children in the older age groups demonstrated use of movement at least once in their samples, although two of the monolingual children did so only in their expository samples. The older children used movement significantly more frequently per utterance than did the younger children, and they used more exemplars of subject and object movement, as well as true relative structures. This means that, by this age, clinicians and researchers can expect children to have a well-developed mastery of movement. Research and clinical assessment tasks that manipulate the length of movement and the semantic load of the matrix clause appear to be age-appropriate based on these results. Importantly, these results indicate that, by this age, syntactic competence should include highly complicated structures such as object relatives, confirming these complex structures as age-appropriate goals for intervention (see Levy & Friedmann, 2009 for an example of such an intervention). The children's use of movement appeared to differ from English-speaking adult production only in the conspicuously infrequent use of passivized relatives, a common alternative to object relatives for English-speaking adults (Scontras

et al., 2014), and a strategy found in children speaking the closely related language of Danish (Jensen de López et al., 2014). The fact that these children were not observed to use this strategy frequently could be due to the complexity of processing two separate movement operations, consistent with the Derivational Complexity Hypothesis (Jakubowicz, 2011). An alternative explanation is that passivization is a feature of an academic register of English which children learn in adolescence. The data presented in this study cannot comment on these explanations as the study design did not require children to produce any particular syntactic targets; however, the conspicuously low frequency of this strategy contradicted our original hypotheses and is an area for future study.

5.2 Effects of discourse

The effects of discourse type were consistently less likely to be due to chance and demonstrated larger effect sizes than did age effects. This pattern was true for both surface and underlying measures, with the exception of MLU_m . This suggests that the type of discourse encourages differences in syntactic complexity greater than a 4 to 5-year age difference. In the present study, expository discourse was consistently measured as more complex than were narratives. This also confirms previous findings that syntactic complexity can vary widely depending on stimulus type, and that close attention must be paid to contextual influences when studying and analyzing the syntactic production of children in language sample tasks (Nettelbladt, Hansson, & Nilholm, 2001). This study also parallels findings from studies of written discourse which found that differences in the writing task yielded more significant differences than those between grade-levels (Crowhurst, 1980; Crowhurst & Piche, 1979). This demonstrates the importance of

carefully selecting a comparison database for language sample analysis, as a mismatch in language sample task would provide inappropriate assessment results. Careful control of elicitation procedures is recommended for both clinical and research assessments.

These findings support expository discourse as the better elicitation method for complex syntax. However, they do not support expository discourse as the single best choice for eliciting language in the school-age years. The selection of discourse genre will vary depending on the clinical question of interest. Narratives are a frequent area of difficulty for children with DLD (Paul & Norbury, 2012), Poor performance on narratives has been suggested as an important diagnostic feature of the disorder (Bishop et al., 2016). Further, narratives are important to school success (Nippold, 2016) and narrative intervention may be effective for improving expository as well as narrative discourse, both of which are critical for complex language tasks (Petersen & Petersen, 2016). Additionally, conversational discourse, although less ideal for the assessment of syntax (Nippold et al., 2005), has been suggested as the most appropriate discourse genre for the assessment of derivational morphology (Squires, Kay-Raining Bird, Cahill, & Cleave, in revision). In summary, clinicians and researchers should carefully choose elicitation tasks according to the relevant clinical and theoretical hypotheses under investigation, and consider collecting samples in more than one discourse genre.

5.3 Effects of multilingualism

The present study explored the effects of multilingual development on measures of syntactic development in English samples only. All measures of syntactic complexity, including MLU, CD, and metrics based on movement, pointed to age-related gains and the greater demands of expository discourse. None of these showed significant

differences between the mono- and bilingual groups. This divergence between syntactic performance and complexity suggests that these bilingual children are developing the general competencies necessary to produce the complexity and breadth of structures that their monolingual peers are attempting. These results indicate that acquiring two languages does not have a negative effect of syntactic complexity in the school-age years. While multilingualism did not appear to have any effect on syntactic competence, mazing and error patterns suggested that multilingual children may lag behind their monolingual peers in syntactic facility.

5.3.1 Mazing and errors

The bilingual children mazed significantly more often than did the monolingual children. This pattern held across both age group and discourse type. Higher rates of mazing in bilingual children is problematic for clinicians as mazing has been suggested as a potential clinical marker of DLD (Leadholm & Miller, 1995). This has been validated by studies showing that children with DLD maze more often than both their TD peers (Thordardottir & Weismer, 2002) and their peers with a phonological impairment (Nettelbladt & Hansson, 1999). Because bilingual children and children with DLD may both demonstrate increased rates of mazing, clinicians must be cautious to compare bilingual children with a suspected language impairment only to bilingual norms on this measure.

However, this finding must also be interpreted with caution, as it contributes to the growing evidence that mazing measures require further investigation and development. A well-powered study with a similar methodology compared bilingual Spanish-English American children to their functionally monolingual English- and

Spanish-speaking peers and did not find any significant differences in overall rate of mazing or in overall mazing type distribution (Bedore, Fiestas, Peña, & Nagy, 2006). The discrepancy may be due to different methods of calculation, as the Spanish-English study used proportion of utterances with mazes and the present study used mazes per utterance. The opposing findings of the present study and the Spanish-English study, despite their similarities, support the need for further investigation of mazing measures. Other researchers have also suggested that current mazing measures are insufficiently well-developed. Different types of mazing may be indicative of different processing patterns, and not all types of mazing may be true clinical markers for a language disorder (Thordardottir & Weismer, 2002). Additionally, significant variation in mazing frequency has been noted in both typical development and in the presence of DLD (Bedore et al., 2006; Lofranco et al., 2006; Nettelbladt & Hansson, 1999), including in the present study. For these reasons, it is suggested that mazing metrics require further refinement.

The frequency of errors also was significantly higher in the bilingual group. Errors consisted predominately of omitted or incorrect inflectional morphology or the incorrect use of function words. Additionally, the interaction effect between age and language status demonstrated that error frequency followed different age-related patterns in these two groups. For the bilingual children, error frequency decreases from the 7-year-old group to the 12-year-old group. In the monolingual group, error frequency did not differ with age. This suggests that mastery of morphosyntax with occasional residual errors is an appropriate expectation for monolingual children as they enter school. However, this is not an appropriate assumption for at least these bilingual children, as these early school age children are still mastering the inflectional morphology of English

in their dual-language development. Error rates were found to converge in the older groups. These findings help to contextualize the results of language assessment for these children, as unusual frequency and persistence of morphosyntactic errors and mazing can be clinical markers for DLD (Leadholm & Miller, 1995; Leonard, 2014a). As such, assessment of early school-aged children who are learning two or more languages should be careful not to overestimate the clinical significance of high error rates.

5.3.2 Multilingualism and syntax

In summary, the effects of bilingual language acquisition on syntax appeared to be minimal. Language group was found to exert a significant effect only for mazing and errors (measures of syntactic facility). This contrasts starkly with the consistent age and discourse effects that were found for other measures, which were designed to measure syntactic complexity and competence. For these reasons, measures of syntactic complexity appear to be specific by not measuring language status, and as such are promising as a method of assessment which may be less likely to confuse patterns of dual-language development and language disorders, at least for the English skills of English/French bilinguals. These findings in syntax parallel other studies which have found that language sample analysis is a promising assessment technique for multilingual children (Cleave et al., 2008; Peets & Bialystok, 2015). Given how these findings support previous research using more controlled tasks such as sentence repetition (Frizelle & Fletcher, 2014), this study also provides some suggestion that these tasks may also be blind to language status in the domain of syntactic competence, which would be useful for clinicians working with multilingual populations.

Mazing and error patterns followed divergent patterns in this study. Mazing, by

varying directly with increases in discourse complexity, appeared to be related to the processing demands required by the utterances that the children were producing. Mazing was also consistently higher in samples by bilingual children. Errors, on the other hand, did not vary with discourse complexity, suggesting that in this age group (at least for monolingual children), these types of errors have reached a floor-effect in the age ranges studied here where errors occur occasionally in speech but remain uncommon. The older bilingual children showed a similar pattern to both monolingual groups, while the younger bilingual children produced a significantly higher rate of errors than all other groups. It is unclear as to why mazing and error patterns not only demonstrated different patterns of effects than did measures of complexity, but also differed from each other. This dissociation between mazing and morphological, lexical, and syntactic errors suggests that different psycholinguistic processes underlie the two types of errors. Future studies should investigate this matter further.

A word of caution regarding interpreting these results – all participants, including those attending school in French, live in the predominantly English-speaking sociolinguistic environment of Nova Scotia. In their review of previous literature comparing morphosyntactic development in monolingual and bilingual children, Paradis, Genesee, and Crago (2011) argue that the sociolinguistic context of bilingual children may affect whether they remain on par with their monolingual peers. Given the sociolinguistic status of English in Nova Scotia, and the fact that this study only evaluated these children's English language capabilities, the lack of significant syntactic differences between the groups is unsurprising. However, these findings may not be generalizable to other language contact situations with different patterns of language use

and ideology, or even to the French abilities of the bilingual participants.

5.4 Comparing measures of syntax

Overall, similar patterns were observed irrespective of whether surface measures or novel, theoretically-driven measures were used. Main effects of age and discourse type were consistently found, with the latter being less likely to be due to chance and demonstrating larger effects sizes. The difference in syntactic complexity between narrative and expository discourse were consistently greater than a 5-year age gap with both types of measures. As statistical results do not provide significant reasons to prefer any measure, other factors must be considered.

MLU continues to be the gold-standard for measuring age-related gains in syntactic complexity even in the early school years. This result parallels findings validating MLU as a measure of syntactic development across three decades of development in English (Nippold et al., 2005) and in the preschool and early school years in French (Mimeau, Plourde, Ouellet, & Dionne, 2015). MLU has the additional benefit of being easy to calculate and of being already well-known by clinicians. However, it is a poor treatment planning measure at this stage of language, as the only treatment goal that the measure can recommend is expanding utterances with any type of word/morpheme. While this may be an appropriate and clinically-feasible goal for children in early language stages, the language of early school age children is more complex and requires different goals.

CD was also found to be appropriate for measuring syntactic development, as most versions of the measure were sensitive to both age- and discourse-related increases in complexity. This result parallels similar validation of CD as a measure of syntactic

development for school-aged, French-speaking children (Mimeau et al., 2015). The single calculation of CD that was not sensitive to discourse differences was CD calculated with exclusively finite clauses. This suggests that counting finite clauses alone misses some increased complexity and is therefore a less sensitive method of calculation. The lack of sensitivity may be why CD has been shown to be less sensitive than MLU in studies using only finite clauses (e.g. Nippold et al., 2005). Clauses bearing aspect but not tense were found to not be sufficiently present to justify the time spent coding. Results from this study suggest that the most efficient methods for calculating CD with appropriate sensitivity involve counting marked infinitive clauses and possibly unmarked infinitive clauses in addition to finite clauses. As omission of ‘to’ in infinitive constructions has been suggested as a clinical marker for DLD (Arndt & Schuele, 2013; Schuele & Dykes, 2005), and so including marked infinitive clauses for the calculation of CD has the added benefit of drawing the clinician’s attention to potential instances of these constructions and their obligatory contexts. However, like MLU, CD provides little guidance for treatment planning.

The measures based on quantifying movement demonstrated patterns of results that paralleled MLU and clausal density. Age-related growth in the frequency of MOVE use (specifically overall and argument movement) was observed between grades 1 and 6; expository texts elicited significantly more frequent use of movement than did narrative texts. However, these measures are more time consuming to calculate than MLU_m or CD, and require significantly more metalinguistic knowledge on the part of the evaluator. For these reasons, measures based on MOVE in language samples do not appear to be clinically feasible.

Significant increases in the presence of subject and object MOVE gap and true relative structures were also observed. Importantly, all of the children in the older group produced at least one exemplar of movement. This provides strong evidence that use of movement, and especially the complexity of the overall embedded multiclausal sentence, is a process that children are continuing to master during the early school age years, and that children can be expected to have good grasp of these structures by the age of 12. The developmental patterns of movement seen between these age groups, although too optional and time-intensive to feasibly measure in language sample analysis, may provide an area of language sufficiently complex and language-general to be appropriate for the assessment of syntax in both mono- and multilingual children. Sentence repetition tasks have recently been shown to be promising assessment techniques (Conti-Ramsden, Botting, & Faragher, 2001; Everitt, Hannaford, & Conti-Ramsden, 2013; Frizelle & Fletcher, 2014; Riches, Loucas, Baird, Charman, & Simonoff, 2010). A sentence repetition task which takes into account the developmental pattern of movement is suggested for future studies.

CHAPTER 6 CONCLUSION

6.1 Summary of findings

This study explored the measurement of syntactic development along three key parameters: age-related gains, task demand-related differences, and the possible impact of dual language learning on complex syntax. In particular, an approach to the measurement of syntax based on theoretical underlying operations was proposed as potentially more valid and/or efficient than surface measures. This prediction was only partially supported by the data. Similar to surface measures, movement-based measures showed sensitivity to age-related increases in syntactic complexity and demonstrated that expository discourse encourages more complex syntax on average than does narrative discourse. Both types of measures were not sensitive to language status, further adding to the body of evidence that language sample tasks hold promise as appropriate assessment techniques for multilingual children (Cleave et al., 2008; Peets & Bialystok, 2015). However, movement-based measures require more time and greater metalinguistic knowledge to calculate, making them clinically unfeasible. Instead, these results can be used in future research to develop and/or modify more feasible tasks, such as sentence repetition. The movement of constituents from argument positions and the complexity of the matrix clause appeared to be critical factors in the development of movement in this age group. Mean length of utterance and clausal density (when calculated with a sufficient diversity of clause types) continue to be useful for the measurement of syntactic development through the early school years.

Mazing and errors patterns followed unexpected patterns, and appeared to disassociate from each other. Mazing increased in expository samples and those samples

by bilingual children. It did not show significant age-related differences. Errors patterns, on the other hand, appeared to demonstrate floor effects for all groups except for the bilingual 7-8 year olds. The divergent findings from these two measures of language performance strongly suggest that they represent and measure different psychological factors in speech production. This divergence should be explored in future studies.

6.2 Limitations and future directions

As a language sampling analysis study, relatively little control was exerted on linguistic microstructure. Therefore, the findings in this study can only comment on what children choose to say, and not on what they are capable of saying, during this age range. The group design also limits the discussion to general age-related trends. Longitudinal studies could offer much more detailed descriptions of the developmental process. The sample size was also small, increasing the likelihood of failing to find significant differences between groups. Although the bilingual group varied in their exposure to English and French, they all were acquiring the same language-pair and lived within a small number of communities in a single Canadian province. Generalizability of these findings to multilingual children learning other language pairs or within communities with different language practices and culture may be limited. Additionally, this study only examined the impact that learning French has on English syntax, and did not explore the impact that learning English has on French syntax.

This study highlights several opportunities for clarification and further exploration by future research. The present study suggests that movement-based assessment tasks show promise as valid measures for syntactic development in both mono- and multilingual children. Incorporating these findings into other assessment tasks

may yield more efficient means for clinicians and researchers to measure children's syntactic development. The exploration of complex syntax in multilingual children should also be explored in language pairs which differ in more substantial ways in non-local operations (e.g. English and the Chinese languages, etc...). These comparisons may yield further insights into the dual language acquisition process, crosslinguistic transfer, and the language-generalness of complex syntax. Further, the unexpected divergence in measures of language performance (mazing and errors) should be explored in further studies with greater power, greater experimental control, and a more detailed coding system for these features of discourse production. The complex relationship of these language performance measures with dual language development should also be further explored.

BIBLIOGRAPHY

- Arndt, K. B., & Schuele, C. M. (2013). Multiclausal utterances aren't just for big kids: A framework for analysis of complex syntax production in spoken language of preschool- and early school-age children. *Topics in Language Disorders, 33*(2), 125–139. <https://doi.org/10.1097/TLD.0b013e31828f9ee8>
- Balthazar, C., & Scott, C. (2015). Syntax in school-age assessment and intervention. In T. Ukrainetz (Ed.), *School-age language intervention: Evidence-based practices* (pp. 281–333). Austin, TX: PRO-ED, Inc.
- Bedore, L. M., Fiestas, C. E., Peña, E. D., & Nagy, V. J. (2006). Cross-language comparisons of maze use in Spanish and English in functionally monolingual and bilingual children. *Bilingualism: Language and Cognition, 9*(3), 233–247. <https://doi.org/10.1017/S1366728906002604>
- Bedore, L. M., & Peña, 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), 1–29. <https://doi.org/10.2167/beb392.0>
- Bishop, D. V. M. (2014). Ten questions about terminology for children with unexplained language problems. *International Journal of Language and Communication Disorders, 49*(4), 381–415. <https://doi.org/10.1111/1460-6984.12101>
- Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., Adams, C., Archibald, L., ... House, A. (2017). Phase 2 of CATALISE: A multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology and Psychiatry, 58*(10), 1068–1080. <https://doi.org/10.1111/jcpp.12721>
- Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., Adams, C., Archibald, L., ... Whitehouse, A. (2016). CATALISE: A multinational and multidisciplinary Delphi consensus study. Identifying language impairments in children. *PLoS ONE, 11*(7), 1–26. <https://doi.org/10.1371/journal.pone.0158753>
- Carnie, A. (2007). *Syntax: A generative introduction* (2nd ed.). Malden, MA: Blackwell Publishing Ltd.
- Catts, H. W., Fey, M. E., Tomblin, J. B., & Zhang, X. (2002). A longitudinal investigation of reading outcomes in children with language impairments. *Journal of Speech, Language, and Hearing Research, 45*(6), 1142–1157. [https://doi.org/10.1044/1092-4388\(2002/093\)](https://doi.org/10.1044/1092-4388(2002/093))
- Cipriani, P., Bottari, P., Chilosi, A., & Pfanner, L. (1998). A longitudinal perspective on the study of specific language impairment: The long term follow-up of an Italian child. *International Journal of Language & Communication Disorders, 33*(3), 245–280.

- Clahsen, H. (2008). Chomskyan syntactic theory and language disorders. In M. J. Ball, M. R. Perkins, N. Müller, & S. Howard (Eds.), *The handbook of clinical linguistics* (pp. 165–183). Malden, MA: Blackwell Publishing Ltd.
- Cleave, P., Girolametto, L., Chen, X., & Johnson, C. (2008). Narrative abilities in monolingual and dual language learning children with specific language impairment. *Journal of Communication Disorders*, *43*(6), 511–22. <https://doi.org/10.1016/j.jcomdis.2010.05.005>
- Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology and Psychiatry*, *42*(6), 741–748. <https://doi.org/10.1111/1469-7610.00770>
- Corrêa, L. M. S., & Augusto, M. R. A. (2011). Possible loci of SLI from a both linguistic and psycholinguistic perspective. *Lingua*, *121*(3), 476–486. <https://doi.org/10.1016/j.lingua.2010.10.011>
- Crago, M., Paradis, J., & Menn, L. (2009). Cross-linguistic perspectives on the syntax and semantics of language disorders. In M. J. Ball, M. R. Perkins, N. Müller, & S. Howard (Eds.), *The handbook of clinical linguistics* (pp. 275–289). Malden, MA: Blackwell Publishing Ltd. <https://doi.org/10.1002/9781444301007.ch17>
- Crowhurst, M. (1980). Syntactic complexity in narration and argument at three grade levels. *Canadian Journal of Education*, *5*(1), 6–13.
- Crowhurst, M., & Piche, G. L. (1979). Audience and mode of discourse effects on syntactic complexity in writing at two grade levels. *Research in the Teaching of English*, *13*(2), 101–109.
- Cummings, L. (2013). *Communication disorders*. Basingstoke, Hampshire, U.K.: Palgrave Macmillan.
- de Jong, J. (2008). Bilingualism and language impairment. In M. J. Ball, M. R. Perkins, N. Müller, & S. Howard (Eds.), *The handbook of clinical linguistics* (pp. 261–274). Malden, MA: Blackwell Publishing Ltd. <https://doi.org/10.1002/9781444301007.ch16>
- de Villiers, J. G., & de Villiers, P. A. (1985). The acquisition of English. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition volume 1: The data* (pp. 27–139). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- de Villiers, J. G., de Villiers, P. A., & Roeper, T. (2011). Wh-questions: Moving beyond the first phase. *Lingua*, *121*(3), 352–366. <https://doi.org/10.1016/j.lingua.2010.10.003>
- Deevy, P., & Leonard, L. B. (2004). The comprehension of wh-questions in children with specific language impairment. *Journal of Speech, Language & Hearing Research*, *47*(4), 802–815. [https://doi.org/10.1044/1092-4388\(2004/060\)](https://doi.org/10.1044/1092-4388(2004/060))

- Diessel, H., & Tomasello, M. (2005). A new look at the acquisition of relative clauses. *Language*, 81(4), 882–906. <https://doi.org/10.1353/lan.2005.0169>
- Dunn, L., & Dunn, L. (2007). *Peabody Picture Vocabulary Test* (Fourth). Minneapolis, MN: Pearson Assessments.
- Dunn, L., Thériault-Whalen, C., & Dunn, L. (1993). *Échelle de vocabulaire en images Peabody* (Adaptation). Toronto, ON: Psycan.
- Everitt, A., Hannaford, P., & Conti-Ramsden, G. (2013). Markers for persistent specific expressive language delay in 3-4-year-olds. *International Journal of Language and Communication Disorders*, 48(5), 534–553. <https://doi.org/10.1111/1460-6984.12028>
- Fletcher, P. (2009). Syntax in child language disorders. In R. G. Schwartz (Ed.), *Handbook of child language disorders* (pp. 388–405). New York: Psychology Press.
- Fortunato-Tavares, T. M., Andrade, C. R. F., Befi-Lopes, D., Hestvik, A., Epstein, B., Tornyova, L., & Schwartz, R. G. (2012). Syntactic structural assignment in children with specific language impairment, autism and Down syndrome. *Journal of Speech, Language & Hearing Research*, 55(August), 1097–1111.
- Fortunato-Tavares, T. M., Andrade, C. R. F., Befi-Lopes, D., Limongi, S. O., Fernandes, F. D. M., & Schwartz, R. G. (2015). Syntactic comprehension and working memory in children with specific language impairment, autism or Down syndrome. *Clinical Linguistics & Phonetics*, 29(7), 499–522. <https://doi.org/10.3109/02699206.2015.1027831>
- Friedmann, N., & Novogrodsky, R. (2011). Which questions are most difficult to understand? The comprehension of Wh questions in three subtypes of SLI. *Lingua*, 121(3), 367–382. <https://doi.org/10.1016/j.lingua.2010.10.004>
- Frizelle, P., & Fletcher, P. (2014). Profiling relative clause constructions in children with specific language impairment. *Clinical Linguistics & Phonetics*, 28(6), 437–49. <https://doi.org/10.3109/02699206.2014.882991>
- Gillam, R., & Pearson, N. (2004). *Test of Narrative Language*. (Pro-Ed., Ed.). Austin, TX.
- Gillam, R., Peña, E. D., & Bedore, L. M. (n.d.). *Spanish version of the Test of Narrative Language*.
- Guo, L.-Y., & Schneider, P. (2016). Differentiating school-aged children with and without language impairment using tense and grammaticality measures from a narrative task. *Journal of Speech, Language, and Hearing Research*, 59(2), 317–329. <https://doi.org/10.1044/2015>
- Håkansson, G., & Hansson, K. (2000). Comprehension and production of relative

- clauses: a comparison between Swedish impaired and unimpaired children. *Journal of Child Language*, 27(2), 313–33. <https://doi.org/10.1017/S0305000900004128>
- Hamann, C., Penner, Z., & Lindner, K. (1998). German impaired grammar: The clause structure revisited. *Language Acquisition: A Journal of Developmental Linguistics*, 7(2–4), 193–245. <https://doi.org/10.1207/s15327817la0702-4>
- Hansson, K., & Nettelbladt, U. (2006). Wh-questions in Swedish children with SLI. *Advances in Speech-Language Pathology*, 8(4), 376–383. <https://doi.org/10.1080/14417040600880722>
- Hesketh, A. (2004). Grammatical performance of children with language disorder on structured elicitation and narrative tasks. *Clinical Linguistics and Phonetics*, 18(3), 161–182. <https://doi.org/10.1080/02699200310001659061>
- Hsu, C. C. N., Hermon, G., & Zukowski, A. (2009). Young children’s production of head-final relative clauses: Elicited production data from Chinese children. *Journal of East Asian Linguistics*, 18(4), 323–360. <https://doi.org/10.1007/s10831-009-9047-y>
- Jakubowicz, C. (2011). Measuring derivational complexity: New evidence from typically developing and SLI learners of L1 French. *Lingua*, 121(3), 339–351. <https://doi.org/10.1016/j.lingua.2010.10.006>
- Jensen de López, K., Sundahl Olsen, L., & Chondrogianni, V. (2014). Annoying Danish relatives: Comprehension and production of relative clauses by Danish children with and without SLI. *Journal of Child Language*, 41(1), 51–83. <https://doi.org/10.1017/S0305000912000517>
- Johnson, C. J., Beitchman, J. H., & Brownlie, E. B. (2010). Twenty-year follow-up of children with and without speech-language impairments: Family, educational, occupational, and quality of life outcomes. *American Journal of Speech-Language Pathology*, 19(1), 51–65. [https://doi.org/10.1044/1058-0360\(2009/08-0083\)](https://doi.org/10.1044/1058-0360(2009/08-0083))
- Keenan, E. L., & Comrie, B. (1977). Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry*, 8(1), 63–99. <https://doi.org/10.2307/4177973>
- Law, J. (2004). The implications of different approaches to evaluating intervention: Evidence from the study of language delay/disorder. *Folia Phoniatica et Logopaedica*, 56(4), 199–219. <https://doi.org/10.1159/000078343>
- Leadholm, B. J., & Miller, J. F. (1995). *Language sample analysis: The Wisconsin guide*. Madison, WI: Wisconsin Department of Public Instruction. Retrieved from <http://files.eric.ed.gov/fulltext/ED371528.pdf>
- Leonard, L. B. (2014a). *Children with specific language impairment* (2nd ed.). Cambridge, MA: MIT Press.
- Leonard, L. B. (2014b). Specific language impairment across languages. *Child*

- Development Perspectives*, 8(1), 1–5. <https://doi.org/10.1111/cdep.12053>
- Levy, H., & Friedmann, N. (2009). Treatment of syntactic movement in syntactic SLI: A case study. *First Language*, 29(1), 15–49. <https://doi.org/10.1177/0142723708097815>
- Lofranco, L. A. L., Peña, E. D., & Bedore, L. M. (2006). English language narratives of Filipino children. *Language, Speech, and Hearing Services in Schools*, 37(1), 28–38. [https://doi.org/10.1044/0161-1461\(2006/004\)](https://doi.org/10.1044/0161-1461(2006/004))
- Marinellie, S. A. (2004). Complex syntax used by school-age children with specific language impairment (SLI) in child-adult conversation. *Journal of Communication Disorders*, 37(6), 517–33. <https://doi.org/10.1016/j.jcomdis.2004.03.005>
- Marinis, T., & Saddy, D. (2013). Parsing the passive: Comparing children with specific language impairment to sequential bilingual children. *Language Acquisition*, 20(2), 155–179. <https://doi.org/10.1080/10489223.2013.766743>
- Marinis, T., & van der Lely, H. (2007). On-line processing of wh-questions in children with G-SLI and typically developing children. *International Journal of Language & Communication Disorders*, 42(5), 557–582. <https://doi.org/10.1080/13682820601058190>
- Miller, J., & Iglesias, A. (2012). *Systematic Analysis of Language Transcripts (SALT)* (Research V). Middletown, WI: SALT Software, LLC.
- Mimeau, C., Plourde, V., Ouellet, A. A., & Dionne, G. (2015). Comparison of measures of morphosyntactic complexity in French-speaking school-aged children. *First Language*, 35(2), 163–181. <https://doi.org/10.1177/0142723715577320>
- Nettelbladt, U., & Hansson, K. (1999). Mazes in Swedish pre-school children with specific language impairment. *Clinical Linguistics & Phonetics*, 13(6), 483–497. <https://doi.org/10.1080/026992099298997>
- Nettelbladt, U., Hansson, K., & Nilholm, C. (2001). Why ask questions? Contextual effects on grammatical structure in the language production of children with specific language impairment. *Child Language Teaching and Therapy*, 17(2), 89–106. <https://doi.org/10.1191/026565901678225637>
- Nippold, M. A. (2016). *Later Language Development* (4th ed.). Austin, TX: PRO-ED, Inc.
- Nippold, M. A., Frantz-Kaspar, M. W., Cramond, P. M., Kirk, C., Hayward-Mayhew, C., & MacKinnon, M. (2014). Conversational and narrative speaking in adolescents: Examining the use of complex syntax. *Journal of Speech, Language, and Hearing Research*, 57(3), 876–86. [https://doi.org/10.1044/1092-4388\(2013/13-0097\)](https://doi.org/10.1044/1092-4388(2013/13-0097))
- Nippold, M. A., Frantz-Kaspar, M. W., Cramond, P. M., Kirk, C., Hayward-Mayhew, C.,

- & MacKinnon, M. (2015). Critical thinking about fables: Examining language production and comprehension in adolescents. *Journal of Speech, Language, and Hearing Research, 58*(2), 325–35. <https://doi.org/10.1044/2015>
- Nippold, M. A., Hesketh, L., Duthie, J., & Mansfield, T. (2005). Conversational versus expository discourse: A study of syntactic development in children, adolescents, and adults. *Journal of Speech, Language, and Hearing Research, 48*(5), 1048–1064. [https://doi.org/10.1044/1092-4388\(2005/073\)](https://doi.org/10.1044/1092-4388(2005/073))
- Nippold, M. A., Mansfield, T. C., Billow, J. L., & Tomblin, J. B. (2008). Expository discourse in adolescents with language impairments: Examining syntactic development. *American Journal of Speech-Language Pathology, 17*(4), 356–366. [https://doi.org/10.1044/1058-0360\(2008/07-0049\)](https://doi.org/10.1044/1058-0360(2008/07-0049))
- Novogrodsky, R., & Friedmann, N. (2006). The production of relative clauses in syntactic SLI: A window to the nature of the impairment. *Advances in Speech-Language Pathology, 8*(4), 364–375. <https://doi.org/10.1080/14417040600919496>
- Ooi, C. C. W., & Wong, A. M.-Y. (2012). Assessing bilingual Chinese-English young children in Malaysia using language sample measures. *International Journal of Speech-Language Pathology, 14*(6), 499–508. <https://doi.org/10.3109/17549507.2012.712159>
- Paradis, J. (2010). The interface between bilingual development and specific language impairment. *Applied Psycholinguistics, 31*(2), 3–28. <https://doi.org/10.1017/S0142716409990373>
- Paradis, J. (2016). The development of English as a second language with and without specific language impairment: Clinical implications. *Journal of Speech, Language, and Hearing Research, 59*(1), 171–182. <https://doi.org/10.1044/2015>
- Paradis, J., Genesee, F., & Crago, M. B. (2011). *Dual language development and disorders: A handbook on bilingualism and second language learning* (2nd ed.). Baltimore, MD: Paul H. Brooks Publishing Co., Inc.
- Paul, R. (1981). Analyzing complex sentence development. In *Assessing language production in children: Experimental procedures* (pp. 36–40, 67–71). Baltimore, MD: University Park Press.
- Paul, R., & Norbury, C. F. (2012). *Language disorders from infancy through adolescence: Listening, speaking, reading writing, and communicating* (4th ed.). St. Louis, MO: Elsevier Mosby.
- Peets, K. F., & Bialystok, E. (2015). Academic discourse: Dissociating standardized and conversational measures of language proficiency in bilingual kindergarteners. *Applied Psycholinguistics, 36*(2), 437–461. <https://doi.org/10.1017/S0142716413000301>

- Penke, M. (2015). Syntax and language disorders. In T. Kiss & A. Alexiadou (Eds.), *Syntax - theory and analysis* (Vol. 3, pp. 1833–1874). Berlin: De Gruyter Mouton.
- Petersen, D. B., & Petersen, J. W. (2016). Re-conceptualizing expository language as narration. *Perspectives of the ASHA Special Interest Groups: SIG 1, 1*(3), 109–117. <https://doi.org/10.1044/persp1.SIG1.109>
- Reilly, S., Tomblin, B., Law, J., McKean, C., Mensah, F. K., Morgan, A., ... Wake, M. (2014). Specific language impairment: A convenient label for whom? *International Journal of Language and Communication Disorders, 49*(4), 416–451. <https://doi.org/10.1111/1460-6984.12102>
- Riches, N. G., Loucas, T., Baird, G., Charman, T., & Simonoff, E. (2010). Sentence repetition in adolescents with specific language impairments and autism: An investigation of complex syntax. *International Journal of Language and Communication Disorders, 45*(1), 47–60. <https://doi.org/10.3109/13682820802647676>
- Salameh, E. K., Håkansson, G., & Nettelbladt, U. (2004). Developmental perspectives on bilingual Swedish-Arabic children with and without language impairment: A longitudinal study. *International Journal of Language and Communication Disorders, 39*(1), 65–91. <https://doi.org/10.1080/13682820310001595628>
- Salameh, E. K., Nettelbladt, U., Hakansson, G., & Gullberg, B. (2002). Language impairment in Swedish bilingual children: A comparison between bilingual and monolingual children in Malmö. *Acta Paediatrica, 91*(2), 229–234. <https://doi.org/10.1080/080352502317285261>
- Schuele, C. M., & Dykes, J. (2005). Complex syntax acquisition: A longitudinal case study of a child with specific language impairment. *Clinical Linguistics & Phonetics, 19*(4), 295–318. <https://doi.org/10.1080/02699200410001703709>
- Schuele, C. M., & Nicholls, L. (2000). Relative clauses: Evidence of continued linguistic vulnerability in children with specific language impairment. *Clinical Linguistics & Phonetics, 14*(8), 563–585. <https://doi.org/10.1080/026992000750048116>
- Scontras, G., Badecker, W., Shank, L., Lim, E., & Fedorenko, E. (2014). Syntactic complexity effects in sentence production. *Cognitive Science, 39*(3), 559–583. <https://doi.org/10.1111/cogs.12168>
- Scott, C. M. (1988). Producing complex sentences. *Topics in Language Disorders, 8*(2), 44–62.
- Scott, C. M., & Koonce, N. M. (2014). Syntactic contributions to literacy learning. In C. A. Stone, E. R. Silliman, B. J. Ehren, & G. P. Wallach (Eds.), *Handbook of language and literacy: Development and disorders* (2nd ed., pp. 283–301). New York, NY: Guilford Press.

- Scott, C. M., & Stokes, S. L. (1995). Measures of syntax in school-age children and adolescents. *Language, Speech, and Hearing Services in Schools*.
<https://doi.org/10.1044/0161-1461.2604.309>
- Squires, B., Kay-Raining Bird, E., Cahill, P., & Cleave, P. (n.d.). Undating and quarterponies: Identifying derived words and compounds in child language samples. In A. Tsedryk & C. Doe (Eds.), *Words: Description, Acquisition and Pedagogy*.
- Statistics Canada. (2012). *Linguistic characteristics of Canadians. Statistics Canada Catalogue no. 98-314-X2011001*. Ottawa. Retrieved from
<http://www12.statcan.gc.ca/census-recensement/2011/as-sa/98-314-x/98-314-x2011001-eng.pdf>
- Statistics Canada. (2017a). Census in Brief: English–French bilingualism reaches new heights. Retrieved from <http://www12.statcan.gc.ca/census-recensement/2016/as-sa/98-200-x/2016009/98-200-x2016009-eng.cfm>
- Statistics Canada. (2017b). Census in Brief: Linguistic diversity and multilingualism in Canadian homes. Retrieved from <http://www12.statcan.gc.ca/census-recensement/2016/as-sa/98-200-x/2016010/98-200-x2016010-eng.pdf>
- Statistics Canada. (2017c). Immigrant languages in Canada. <https://doi.org/98-314-X2011003>
- Stavrakaki, S. (2006). Developmental perspectives on specific language impairment: Evidence from the production of wh-questions by Greek SLI children over time. *Advances in Speech-Language Pathology*, 8(4), 384–396.
<https://doi.org/10.1080/14417040600880714>
- Stavrakaki, S., Chrysomallis, M.-A., & Petraki, E. (2011). Subject-verb agreement, object clitics and wh-questions in bilingual French-Greek SLI: The case study of a French-Greek-speaking child with SLI. *Clinical Linguistics & Phonetics*, 25(5), 339–367.
<https://doi.org/10.3109/02699206.2010.538954>
- Thompson, C. K. (2008). Treatment of syntactic and morphologic deficits in agrammatic aphasia: Treatment of underlying forms. In R. Chapey (Ed.), *Language intervention strategies in aphasia and related neurogenic communication disorders* (5th ed., pp. 735–755). Philadelphia, PA: Lippincott Williams & Wilkins.
- Thompson, C. K., & Shapiro, L. (2005). Treating agrammatic aphasia within a linguistic framework: Treatment of underlying forms. *Aphasiology*, 19(10–11), 1021–1036.
<https://doi.org/10.1080/02687030544000227>
- Thordardottir, E. T., & Weismer, S. E. (2002). Content mazes and filled pauses in narrative language samples of children with specific language impairment. *Brain and Cognition*, 48(2–3), 587–592. <https://doi.org/10.1006/brcg.2001.1422>
- To, C. K. S., Lui, H. M., Li, X. X., & Lam, G. Y. H. (2015). Randomized controlled trial

- of two syntactic treatment procedures with Cantonese-speaking, school-age children. *Journal of Speech, Language, and Hearing Research*, 58, 1258–1272. <https://doi.org/10.1044/2015>
- Tuller, L., Henry, C., Sizaret, E., & Barthez, M.-A. (2012). Specific language impairment at adolescence: Avoiding complexity. *Applied Psycholinguistics*, 33(1), 161–184. <https://doi.org/10.1017/S0142716411000312>
- Turnbull, K. L. P., & Justice, L. M. (2012). *Language development: From theory to practice* (2nd ed.). Upper Saddle River, NJ: Pearson Education, Inc.
- van der Lely, H., & Battell, J. (2003). WH-movement in children with grammatical SLI: A test of the RDDR hypothesis. *Language*, 79(1), 153–181. <https://doi.org/10.2307/4489388>
- van der Lely, H., Jones, M., & Marshall, C. R. (2011). Who did Buzz see someone? Grammaticality judgement of wh-questions in typically developing children and children with Grammatical-SLI. *Lingua*, 121(3), 408–422. <https://doi.org/10.1016/j.lingua.2010.10.007>
- van der Lely, H., & Marshall, C. (2011). Grammatical-specific language impairment: A window onto domain specificity. In *The handbook of psycholinguistic and cognitive processes* (pp. 401–418). New York, NY: Taylor & Francis Group.
- van der Lely, H., & Stollwerck, L. (1997). Binding theory and grammatical specific language impairment in children. *Cognition*, 62(3), 245–290. [https://doi.org/10.1016/S0010-0277\(96\)00783-4](https://doi.org/10.1016/S0010-0277(96)00783-4)
- Wong, A., Leonard, L. B., Fletcher, P., & Stokes, S. F. (2004). Questions without movement: A study of Cantonese-speaking children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research : JSLHR*, 47(6), 1440–1453. [https://doi.org/10.1044/1092-4388\(2004\)107](https://doi.org/10.1044/1092-4388(2004)107)
- Yu, H. (2016). The grammar impairment of Mandarin Chinese SLI children : Evidence from topic-comment. *Journal of Language Teaching and Research*, 7(2), 299–306.
- Zwitserslood, R., Wijnen, F., van Weerdenburg, M., & Verhoeven, L. (2015). “MetaTaal”: Enhancing complex syntax in children with specific language impairment-a metalinguistic and multimodal approach. *International Journal of Language & Communication Disorders*, 50(3), 273–297. <https://doi.org/10.1111/1460-6984.12131>

APPENDIX A CODING GUIDELINES FOR CLAUSAL DENSITY

Calculating clausal density

There is no consensus on the best method to calculate clausal density. Typically, it has involved counting only finite clauses, although infinitive clauses are also included on occasion. It is not clear why other types of clauses, such as gerundial clauses or non-finite clauses that are not marked with a clear infinitive, are excluded from the calculations. This study adopts a system which is inspired by the framework which Arndt and Schuele (2013) have utilized to examine the emergence of complex syntax. The clausal types and their corresponding codes can be found in the following table. They are ordered in according to their hypothesized rarity of use in spontaneous speech by school-age children.

Table 6 List of codes by clause type

Clause type	Code
Finite clause	[^c]
Infinitive clause	[^ci]
Non-finite clause	[^cn]
Progressive participial clause	[^cg]
Perfective participial clause	[^cp]

The matrix clause of an utterance is usually (although not exclusively) a finite clause. These clauses are marked for tense, including simple tense which predominately communicates temporal information (e.g. he walks, he walked, we will walk, etc...) and complex tenses that communicate both temporal and aspectual information (e.g. he is walking, we will have walked when..., etc...) This is complicated by the fact that in English tense is sometimes marked through a zero morpheme (e.g. I walk), although it is usually identifiable by surrounding contexts, such as the actor of the verb bearing nominal case (if it is a pronoun). For example “I am walking[^c]” versus “He

saw [^c] me walking [^cg]”. Native speakers with adequate metalinguistic awareness are typically highly accurate in identifying finite clauses.

There are cases where the matrix clause of an utterance is not finite. This typically occurs in elliptical responses. For example, the examiner may ask “What did you see?”, to which the child responds “him walking [^cg]”. This would be coded as a gerundial clause, which is also serving as the matrix clause for the utterance.

In this system, coordination is considered to extend an utterance regardless of the number of times that it is utilized in an utterance. If the subject is maintained (elided) across coordinated clauses, these are treated as a single utterance. Coordinated clauses with different or repeated subjects are treated as separate utterances. However, verbs which are coordinated without intervening constituents are treated as a single clause, as it is possible that the coordination is occurring at the lexical rather than clausal level.

Table 7 Examples of coordination coding guidelines

Coordination	
Multiple utterances	He kick/ed it [^c]. And then he hit it [^c]. And then he threw it [^c]
Multiple finite clauses	He kick/ed it [^c] and then hit it [^c] and then threw it [^c].
Single finite clause	He kicked and hit and threw it [^c].

All subordinated clauses are retained as part of the utterance of the matrix clauses. Following the framework by Arndt and Schuele, “let’s” is considered the matrix clause and subsequent clauses are coded as non-finite.

Table 8 Examples of let clause coding guidelines

Subordination with <i>let</i>	
Let's	Let's [^c] go to the store [^cn].
Let	Let him [^c] go to the store [^cn].

Subordinated clauses are coded as infinitive only if they are explicitly marked with 'to,' or the presence of 'to' is considered an obligatory context to render the utterance grammatical, even if the child had omitted the word. Clauses where the infinitival 'to' would be grammatically incorrect are coded as non-finite. The verbal element in a non-finite clause must not bear tense, nor any aspectual morphology. Unmarked verbs in a serial verb construction are coded as non-finite.

Table 9 Examples of infinitive vs. non-finite clause coding guidelines

Infinitive versus non-finite clauses	
Infinitive, <i>to</i> not omitted	I want [^c] to go to the store [^ci].
Infinitive, <i>to</i> as obligatory context	He went [^c] *to see the penguin/s [^ci].
Non-finite, no errors	I saw her [^c] go to the park [^cn]
Non-finite, with incorrect <i>to</i>	I saw her [^c] to[ew] go to the park [^cn]
Serial verb construction	Let's [^c] go [^cn] see her [^cn].

Progressive participial clauses are coded where a verb ends in the progressive suffix “-ing” and the following constituents stem directly from the verbal root. If there are no subsequent syntactically/semantically related constituents, the verb is considered a gerund or a post-modifier and is not coded as a clause.

Table 10 Examples of progressive participial clause coding guidelines

Progressive participial clauses	
Post-modifiers with <i>-ing</i> (not coded)	I saw her standing [^c].
Gerunds (not coded)	Eating is awesome [^c].
Gerundial clause	I saw her [^c] standing in line [^cg]
“	Eating kimchi [^cg] is awesome [^c]

Perfective participial clauses follow the same criteria as perfective participial clauses, but instead of “-ing” involved verbal elements with past-participle morphology (e.g. “-ed”).

Table 11 Examples of perfective participial clause coding guidelines

Perfective participial clauses	
Post-modifier with –ed (not coded)	It look/ed like a city destroyed [^c].
Participle clause	It look/ed like a city [^c] destroyed by a fire [^cp]

In quoted discourse, only the first matrix clause and all of its subordinate clauses are included within the same utterance as the clause reporting the dialogue. All other matrix clauses are coded as separate utterances.

Table 12 Examples of quoted discourse coding guidelines

Quoted discourse	
Single, complex quoted utterance	He said [^c] “that is [^c] because he had already hit the ball [^c] he had been pitched out of the park [^c].”
Multiple quoted utterances	He said [^c] “look [^c]”. “There it is [^c]”.

APPENDIX B CODING GUIDELINES FOR MOVE

Table 13 Codes for movement

<i>Pneumonic</i>	gap location	pseudo (cleft)	true (relative)
<i>Move other</i>	adjunct	[mr]	[mr]
<i>Move Subject</i>	subject	[msp]	[mst]
<i>Move Object</i>	object	[mop]	[mot]
<i>Move oblique</i>	oblique	[mqp]	[mqt]

The first decision made when coding is whether something has moved.

[mr]

Once move have been identified, the next step is to identify the location of the gap. If the moved constituent does not have a canonical argument position, then it is considered non-canonical and coded as [mr]. Some examples include:

*I don't know **how** to do that [mr].* "How" has likely moved, as evidenced by the fact that it cannot be further moved around within this new construction. However, the various replacements for "how", such as prepositions phrases, adverbs, etc... can move around substantially in the clause. Since it is difficult to propose a specific location out of which the "how" is being moved, we code as [mr].

Special cases: When and where

Ignore when and where clauses if they are adverbial clauses.

When I went to the store, I saw something strange.

I will meet you where you told me earlier.

Do code when and where if:

- a. They are wh-questions
- b. They head required complement clauses

I don't know where he went.

I don't know when she will go.

- c. They head embedded infinitive clauses

That is where to go.

That is when to go.].

[msX]

If the moved item is gapped in the subject position in the subordinated clause, it is coded as [msX]. Some examples:

*I don't know **who** went there [mst].* The "who" is the subject of the embedded clause and has been moved to fill in for object of "know" as well.

***The guy** that worked there [mst] called me.* "The guy" is both the subject of the matrix and embedded clause. It presumably has been moved from the subject of the embedded clause to the subject of the whole utterance.

[moX]

If the moved item is gapped in the direct object position in the subordinated clause, it is coded as [moX]. Some examples include:

*It's **the guy** that the horse bucked [mop].* "The guy" here is likely generated as the object of the embedded clause and moved to be the complement of the copula in the matrix clause.

The guy, who I had called [mot], called her back instead. "The guy" is generated as the object of the embedded clause and is now the subject of the matrix clause as well.

[mqX]

If the moved item is in the indirect object position, the object of a prepositional phrase (that has not been moved in its entirety as a constituent), or the possessor of a possessive construction are coded as oblique. In the data set, this category consists almost exclusively of objects of prepositions. Some examples include:

That's the car I got in [mqp]. "The car" is presumably generated as the object of the preposition "in" and is moved to also be the complement of the copula in the matrix clause.

The girl he talked to [mqt] called me back. "The girl" generates within the prepositional phrases headed by "to" and is moved to subject position in the matrix clause.

[p] versus [t]

The final decision for coding is whether each clause with move is a "true" move clause or a "pseudo" move. In contrast to the gapping decision, which is made based on the syntax of embedded clause, the true/pseudo distinction is based on the semantic load of the matrix clause in which the subordinate clause is directly embedded. Utterances are coded as pseudo [p] if they:

1. are produced as an elliptical response without a matrix clause

E you don't know>

C what to do [mop].

2. are embedded only within dialogue

C then he said "what are you doing [mop]"?

3. are directly embedded within a copular clause

C That's the guy who I saw [mop].

Clauses are coded as true [t] if they do not fall into these categories. In other words, it is coded as true if the clause that immediately governs the clause from which an item has been moved carries its own semantic weight. Below is a somewhat complicated example.

C I don't know what it is [mot] he said [mop]. The first clause is coded as true because it is embedded within a clause that provides semantic meaning beyond an existential copula. The second clause is coded as pseudo because the clause in which it is immediately embedded carries no semantic weight besides the copula.

NOTE: The [mr] code does not receive the true/pseudo distinction.

NOTE: These codes can be applied to incomplete, abandoned, interrupted, and even partially unintelligible utterances as long as sufficient information is present to make a proper syntactic judgement about the location of the gap and the semantic load of the matrix clause.

Examples of the argument movement can be found in Table 14 (following page).

Table 14 Examples of argument movement coding
 (adapted from Frizelle & Fletcher (2014, p. 259))

	<i>Cleft</i>	<i>Relative</i>
<i>Subject</i>	There is the sheep that <i>[msp]</i> drank the water this morning.	Eddie met the girl who <i>[mst]</i> broke the window last week.
<i>Object</i>	There is the boy that Emma helped <i>[mop]</i> in the kitchen.	The boy rode the horse that Anne put <i>[mot]</i> in the field.
<i>Oblique</i>	There is the tree that the car crashed into <i>[mqp]</i> last night.	Anne painted the picture that the girl looked at <i>[mqt]</i> today.