Capturing Complex Syntax Development:
Comparing Late Talkers to Children with Typical Language

By

Laurel Jean Teller

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

In

Hearing and Speech Sciences

May 8, 2020

Nashville, Tennessee

Approved:

C. Melanie Schuele, Ph.D.

James Bodfish, Ph.D.

Stephen Camarata, Ph.D.

Ann Kaiser, Ph.D.

Duane Watson, Ph.D.

ACKNOWLEDGEMENTS

I wish to thank the many people who have helped in the completion of this dissertation. First, to Dr. Schuele for her careful feedback and for providing me with the opportunity to complete my doctoral studies. Second, to my committee members, Dr. Camarata, Dr. Kaiser, Dr. Bodfish, and Dr. Watson, for their conversations, time, and feedback. Third, to my labmates and friends who I am so glad travelled along beside me and to my family for their love and support. Finally, to the many people who worked carefully to provide the language transcripts that are freely available on CHILDES.

Funding for the Ellis Weismer Corpus was provided by National Institute on Deafness and Other Communication Disorders Grant 5 R01 DC03731, “Linguistic Processing in Specific Language Delay” (S. Ellis Weismer, PI, J. L. Evans & R. S. Chapman, Co-PIs) and support from the National Institute of Child Health and Human Development P30 core grant awarded to the Waisman Center. The completion of this dissertation study was supported by a United States Department of Education Preparation of Leadership Personnel grant (H325D140087; PI: Schuele). The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institute on Deafness and Other Communication Disorders, the National Institute of Child Health and Human Development, or the United States Department of Education.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

“Late talkers” is a term that refers to a group of children who demonstrate late language emergence despite otherwise typical development (e.g., Klee, Carson, Gavin, Hall, Kent, & Reece, 1998; Paul, Looney, & Dahm, 1991). Children frequently are classified as late talkers based on parent report of delays in expressive vocabulary and/or lack of word combinations at 24 months. Zubrick, Taylor, Rice and Slegers (2007) concluded that late language emergence is a relatively common developmental concern. In their population-based study, 19% of the 24-month-old children demonstrated late language emergence based on parent report of the presence or absence of word combinations.

By kindergarten most late talkers demonstrate performance in the typical range on norm-referenced expressive language tests (e.g., Girolametto, Wiigs, Smyth, Weitzman, & Pearce, 2001; Paul, 1996). Group means on language outcomes continue to be within the average range into adolescence albeit statistically lower as compared to group means for peers with typical language (Rescorla, 2009). However, some late talkers fail to achieve scores within the average range. Unfortunately, predicting who will move within the average range and who will not has proven to be difficult (e.g., Dollaghan, 2013). Determining prognosis has consequences for family stress and financial decisions about whether or not to proceed with intervention (e.g., Paul, 1996).

Tense marking is an area of consistent deficit for children with specific language impairment (SLI), a group of children who have long-term language impairment. Thus, tense marking is an important area to evaluate in determining which late talkers will achieve language scores within the average range. In a follow-up to their population-based study, Rice, Taylor, and Zubrick (2008) found that late talkers demonstrated poorer performance on assessments that tapped the bound tense morphemes, third person -s and past tense -ed, as compared to same-aged peers (*d* = 0.50 – 1.93). Prominent deficits in third person -s and past tense -ed are unsurprising given that deficits in tense morphemes have been reported in other studies of late talkers as well as children diagnosed with SLI (e.g., Rescorla & Turner, 2015; Rice & Wexler, 1996).

The development of dependent clauses – complex syntax – represents another important but understudied aspect of preschool grammatical development (Barako Arndt & Schuele, 2013). There is evidence that late talkers and children with SLI demonstrate deficits in complex syntax development (e.g., Paul & Alforde, 1993; Morton, Delgado, & Schuele, 2019). The relevance of complex syntax to persistent language deficits promotes the need for a greater understanding of the longitudinal trajectory of complex syntax development in late talkers. The aim of this dissertation was to characterize the complex syntax development of late talkers across four time points from 30 months to 66 months as compared to same-aged peers with typical language.

Model of Complex Syntax Development

Barako Arndt and Schuele (2013) proposed a multi-dimensional model of complex syntax development that emphasizes the syntactic and the lexical aspects of complex syntax development. Broadly, there are four categories of complex syntax – infinitival clauses (e.g., *I want to go*), complement clauses (e.g., *I think I can go*), subordinate conjunction clauses (e.g., *I'll go because I need a walk*), and relative clauses (e.g., *That's the park that I went to yesterday*). Each category is further subdivided into multiple complex syntax types (e.g., Barako Arndt & Schuele, 2013; Diessel, 2004). As children produce new complex syntax types they demonstrate what is termed vertical growth and as they produce a wider diversity of lexical items within the structures, they demonstrate what is termed horizontal growth (see Figure 1).

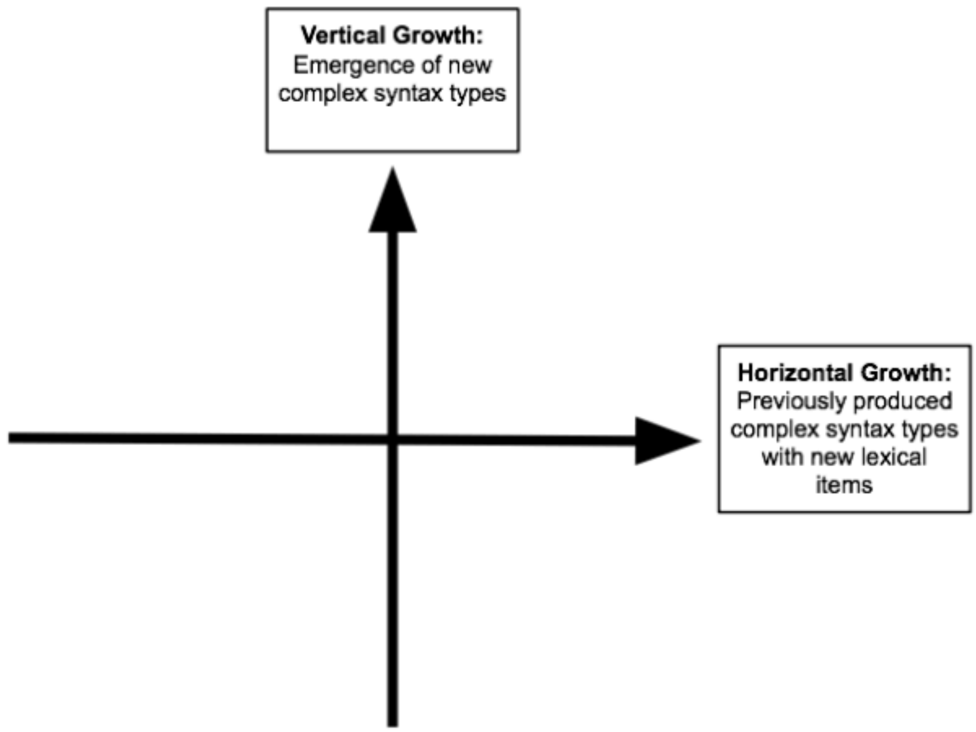

Figure 1. Multi-dimensional model of complex syntax development.

Three of the complex syntax categories – infinitival clauses, complement clauses, and subordinate conjunction clauses – are reliant on the production of specific lexical items. We term these lexical items complex syntax vocabulary. Complex syntax vocabulary comprises subordinate conjunctions (e.g., *because, when, if*) and verbs that subcategorize for infinitives (e.g., *want, need*) and/or complement clauses (e.g., *think, know*). As a group these verbs are termed complement clause verbs. Greater diversity in complex syntax vocabulary indicates greater horizontal growth within the related complex syntax types or categories. For example, a child demonstrates horizontal growth when she uses subordinate conjunction clauses not only with *because*, but also with *until* and *after*. In contrast, relative clauses are not reliant on the production of specific lexical items. We might expect late talkers, who typically demonstrate initial delays in expressive vocabulary, to have delays in vertical and horizontal development and thus in the syntactic aspects and the lexical aspects of complex syntax development.

Complex Syntax Research in Typical and Atypical Populations

Late talkers. The extant literature provides preliminary evidence that late talkers demonstrate delays in vertical and horizontal complex syntax development relative to same-aged peers with typical language. These delays appear to include syntactic and lexical aspects of complex syntax. With respect to vertical development Lee and Rescorla (2002) reported that three-year-old late talkers produced fewer exemplars of four types of complex syntax (conjoined sentences, non-catenative infinitives, propositional complements, wh-clauses) relative to same-aged peers with typical language as measured on the Index of Productive Syntax (IPSyn) Sentence Structure subscale (Scarborough, 1990). Lee and Rescorla (2002) further subdivided these late talkers into late bloomers (based on mean length of utterance) and those with continuing delay at three years of age. The late talkers with continuing delay produced fewer tokens across all four types of complex syntax as compared to same-aged peers; whereas the late bloomers produced fewer tokens of only two types (propositional complements, non-catenative infinitives).

Within a sample of children overlapping with Lee and Rescorla (2002), Rescorla and Turner (2015) found that five-year-old late talkers with continuing delay demonstrated a lower mean score on the Sentence Structure subscale as compared to same-aged peers (d = 1.47). In contrast, children classified as late bloomers at five years of age did not demonstrate a lower mean score as compared to same-aged peers (d = 0.41). Between-group differences for specific complex syntax types were not compared in this study.

With respect to horizontal development, Lee and Rescorla (2002) found that three-year-old late talkers (late bloomers + late talkers with continuing delay) produced fewer cognitive state words (d = 1.47) as compared to same-aged peers. The subset of late bloomers produced fewer cognitive state words (d = 1.09) as compared to same-aged peers. Late bloomers and late talkers with continuing delay did not differ from each other in the number of cognitive state words produced. The majority of the cognitive state words were complement clause verbs (88% in the late talker group, 98% in the same-aged peers) and thus were potentially produced with complement clauses. Children's production of cognitive state words correlated with propositional complements on the IPSyn (r = .74). Not surprisingly given the differences in the number of cognitive state words, at each age late talkers produced fewer propositional complements (e.g., *She didn't know that I wasn't here*) as compared to same-aged peers.

In a follow-up study Lee and Rescorla (2008) found that at three, four, and five years of age late talkers produced a lower percentage of utterances that included cognitive state words as compared to same-aged peers with effect sizes of approximately d = 2.92, 4.75, and 3.46 respectively.[1] The production of propositional complements and cognitive state words were correlated at each time point (rs = .75 – .79). At five years of age, 55% of the late talkers produced propositional complements as compared to 100% of the same-aged peers. Lee and Rescorla (2008) was limited to propositional complements and thus did not provide a full characterization of complex syntax development in late talkers nor did it differentiate between late talkers with continuing delay and late bloomers. Further investigation is needed to more fully characterize the syntactic and lexical aspects of complex syntax development in late talkers. Individual differences within the late talker group must be considered.

Children with specific language impairment. Multiple research teams have reported that complex syntax production in children with SLI, another linguistically vulnerable population, differs from same-aged and MLU-matched peers (e.g., Barako Arndt & Schuele, 2012; Marinellie, 2004; Owen & Leonard, 2006). Schuele and Dykes (2005) represents the first

---

[1] Information to calculate effect sizes was ascertained using WebPlotDigitizer 4.2 https://apps.automeris.io/wpd/ and thus are approximated.

comprehensive, longitudinal analysis of complex syntax development in one child with SLI from three to seven years. Delays were evident in complex syntax quantity, emergence and diversity of complex syntax types, and grammatical accuracy as compared to extant descriptions of complex syntax development in same-aged peers (e.g., Tyack & Gottsleben, 1986; Diessel, 2004). Other studies provide further evidence of between-group differences in grammatical accuracy and complex syntax quantity in children with SLI as compared to children with typical language (Barako Arndt & Schuele, 2012; Morton et al., 2019).

Present Study

Schuele and Dykes (2005) proposed a framework of four areas in which to measure complex syntax delays in children with SLI: (a) complex syntax quantity, (b) complex syntax type diversity, (c) emergence of complex syntax types, and (d) grammatical accuracy. In this dissertation we focused on the first three areas to compare the complex syntax development of late talkers to that of same-aged peers with typical language. We asked five research questions. The first two are falsifiable research questions. The last three are descriptive research questions.

1. Within each language group is there growth in complex syntax quantity and diversity at 12-month intervals?
2. Are there between-group differences in the complex syntax quantity and diversity at each measurement time point?
3. What percent of late talkers have complex syntax proportion and diversity no more than one SD below the mean of same-aged peers at each time point?
4. What proportion of participants produced exemplars from each complex syntax category at each time point measured cumulatively, that is inclusive of previous time points?

5. What is the proportional composition of tokens from each complex syntax category relative to the sum of complex syntax tokens in the four categories at each time point?

CHAPTER II

METHOD

The study methods were approved by the Vanderbilt University Institutional Review Board.

Ellis Weismer Corpus and Participant Criterion

This study involved a secondary analysis of extant language sample data from the Ellis Weismer Corpus in the Child Language Data Exchange System (CHILDES) collected between 1999 and 2004 (e.g., Heilmann, Ellis Weismer, Evans, & Hollar, 2005; Moyle, Ellis Weismer, Evans, & Lindstrom, 2007). The Ellis Weismer Corpus comprises language transcripts and audio files from annual lab-based, adult-child interactions at child age of 30, 42, 54, and 66 months. Audio files and language transcripts are not available for all children at each time point. At 30 and 42 months, children engaged in examiner-child and parent-child play-based interactions. At 54 months, children engaged in an examiner-child play-based interaction and an examiner-child structured interview. At 66 months, children engaged in an examiner-child structured interview.

The Ellis Weismer Corpus includes data from 80 children with typical language (NL) and 57 late talkers (LT). Group assignment was based on child score at 24 months on the productive vocabulary portion of the MacArthur Communicative Development Inventories: Words and Sentences (CDI; Fenson et al., 1993). Children in the NL group received a score at or above the 20^{th} percentile and children in the LT group received a score at or below the 10^{th} percentile. Per parent report each child in the LT group was typically developing except in the domain of language with no report of hearing loss or cognitive impairment. All children were monolingual English speakers. Specific demographic information for each child is not available via CHILDES. However, likely children are predominantly white and from middle-class backgrounds (Ellis Weismer, 2007; Heilmann et al., 2005).

At 30 months children met several eligibility criteria: (a) typical overall development as screened by the Denver II (Frankenburg et al., 1990), (b) typical oral and speech motor abilities as measured by Robbins and Klee's (1987) clinical assessment tool, and (c) typical hearing as screened by distortion product otoacoustic emissions using a Biologic OAE screener (2000, 3000, 4000, and 5000 Hz in at least one ear). Receptive language level was not a component of the eligibility criteria and therefore, was left free to vary.

Participants

Children in the Ellis Weismer Corpus who had available audio files for experimenter-child play-based interactions at 30, 42, and 54 months were selected to participate in the current study. The NL participant group comprised 16 children (75% male) and the LT participant group comprised 26 children (73% male). The percent of boys in the LT group was similar to other studies of late talkers (e.g., Paul & Smith, 1993). See Table 1 for mean age by participant group at each time. Using a Mann-Whitney *U* test, between-group differences for age were non-significant at each timepoint (ps > .05).

Table 1

Age by Time and Group

	Participant Group			
	NL Group		LT Group	
Child Age	Mean	SD	Mean	SD
30 months	30.44	0.63	30.15	0.46
42 months	42.63	0.81	42.42	0.76
54 months	54.13	0.50	54.16	0.62
66 months	66.27	0.70	66.29	0.46

Note. NL group, n = 16 and LT group, n = 26 with the exception of the NL group at 66 months (n = 15), the LT group at 54 months (n = 25), and the LT group at 66 months (n = 24). Participants with missing data at a given time point were excluded from analyses at that time point.

Norm-referenced assessments. Ellis Weismer Corpus participants completed norm-referenced standardized testing at each time. Data for norm-referenced assessments were provided for a subset of participants in the current study (Ellis Weismer, personal communication, 10/24/2018). At 30 months scores were provided for 63% (10/16) of the NL group and 50% (13/26) of the LT group: (a) raw scores for CDI productive vocabulary, (b) percentiles for CDI productive vocabulary, (c) raw scores for CDI complexity, and (d) Total Score on the Preschool Language Scale-3 (PLS-3; Zimmerman, Steiner, & Pond, 1992). At 66 months the Listening Quotient and Speaking Quotient from the Test of Language Development-3 Primary (3rd Edition; TOLD-P:3; Newcomer & Hammill, 1997) were provided for 53% (8/15) of the NL group[2] and 42% (10/24) of the LT group.

30-month assessments. At 30 months we compared the MLU of participants in each group for whom we had norm-referenced assessments (subset) with those for whom we did not (non-subset). The mean MLU for the NL subset was 3.28 (SD = 0.69) and the mean MLU for

[2] There was an 80% overlap between participants in the NL group who had available scores at 30 months and 66 months and a 77% overlap between participants in the LT group at each time point.

the NL non-subset was 3.12 (SD = 0.61). The mean MLU for the LT subset was 1.91 (SD = 0.63) and the mean MLU for the LT non-subset was 2.33 (SD = 0.68). There was no between-group difference for the NL subset as compared to the NL non-subset, $t(14) = 0.45$, $p = .656$, two-tailed nor for the LT subset as compared to the LT non-subset, $t(24) = -1.64$, $p = .114$, two-tailed. Thus, we report the subset of norm-referenced assessments as tentatively representative of the study sample.

At 30 months participants in the NL subset and in the LT subset had a mean percentile on the CDI productive vocabulary of 42.40 (SD = 25.91) and 13.38 (SD = 15.10), respectively, with a large effect size between groups ($d = 1.42$). Additionally, at 30 months participants in the NL subset and in the LT subset had a mean PLS-3 Total Score standard score of 110.60 (SD = 17.82) and 94.54 (SD = 14.40) respectively, with a large effect size between groups ($d = 1.01$). Only 31% (4/13) of the LT subset group had PLS-3 Total Score standard scores below average, that is less than 85.

66-month assessments. At 66 months we compared the MLU of participants in each group for whom we had norm-referenced assessments (subset) with those for whom we did not (non-subset). At 66 months the mean MLU for the NL subset was 6.24 (SD = 1.77) and the mean MLU for the NL non-subset was 6.41 (SD = 0.68). The mean for the LT subset was 5.98 (SD = 1.05) and the mean MLU for the non-subset was 6.49 (SD = 1.20). There was no between-group difference for the NL subset as compared to the NL non-subset, t(9.27) = -0.27, p = .797, two-tailed, equal variances not assumed, nor for the LT subset as compared to the LT non-subset, t(22) = -1.08, p = .292, two-tailed. Thus, we report scores on the TOLD-P:3 at 66 months as tentatively representative of the study sample.

At 66 months participants in the NL subset and in the LT subset had a mean TOLD-P:3 Listening Quotient of 122.13 (SD = 9.75) and 120.10 (SD = 11.32) respectively, with a small effect size between groups ($d = 0.19$). In contrast, participants in the NL subset and in the LT subset had a mean TOLD-P:3 Speaking Quotient of 116.13 (SD = 8.77) and 103.00 (SD = 8.12) respectively, with a large effect size between groups (*d* = 1.56). The standard scores for all children were above a standard score of 85.

Transcript Analysis Set

The analysis set from which we derived dependent variables at each time point was defined as the first 50 complete and intelligible utterances in the language sample transcript. We chose to include only complete and intelligible utterances to mirror analysis sets that are commonly used in the calculation of MLU (e.g., Eisenberg, Fersko, & Lundgren, 2001; cf. Schuele & Dykes, 2005). Utterances for the analysis set were drawn primarily from the examiner-child play interaction transcripts at 30, 42, and 54 months and solely from the examiner-child interview transcripts at 66 months. If 50 complete and intelligible utterances were not available in the examiner-child play interaction transcripts at 30, 42, and 54 months, the analysis set was completed with utterances from the parent-child play interaction transcript at 30 and 42 months and from the examiner-child interview transcript at 54 months (see Table 2).

Table 2

Proportion of Participants with Utterances Drawn from Two Interaction Contexts

	Participant Group	
Time	NL Group	LT Group
30 months	0.25	0.62
42 months	0.13	0.27
54 months	0.19	0.24

Note. NL group, n = 16 and LT group, n = 26 with the exception of the NL group at 66 months (n = 15), the LT group at 54 months (n = 25), and the LT group at 66 months (n = 24).

Several children who were selected to participate did not produce a 50-utterance analysis set across available transcripts. At 30 months four participants in the LT group and at 54 months one participant in the LT group did not produce a 50-utterance analysis set inclusive of examiner- and parent-interaction transcripts. As a result, variables were calculated from the shortened analysis set (Range = 41 – 47 utterances). At 54 months one child who met selection criteria produced only 12 utterances in the examiner-child play interaction and the audio for the 54-month examiner-child interview was not available; variables for this participant were not calculated at this time point (i.e., missing data). At 66 months three participants (NL, n = 1; LT, n = 2) did not have audio files, and therefore their 66-months data were considered missing data.

Language Transcript Revisions

Language transcripts for each participant were imported from the Computerized Language Analysis (CLAN) format in CHILDES to the Systematic Analysis of Language Transcripts (SALT) format to allow for analysis in SALT (Miller & Iglesias, 2012).[3] We revised the archival transcripts in three ways to allow for the analyses of interest. First, the principal investigator coded the utterances for complex syntax. We followed conventions from Schuele (2009a) as well as study-specific conventions. Second, the principal investigator aligned the written transcripts with Schuele (2009b) and study-specific conventions: (a) slash bound morphemes (e.g., cat/s), (b) remove filler words, sound effects not in response to questions, and nonverbal productions from the analysis set, (c) correct orthographic transcription errors (e.g., word misspellings), (d) apply study-specific utterance segmentation rules (see next paragraph), and (e) separate multi-word sequences formatted as single words in the archival transcripts (e.g., a_lot_of = a lot of, have_to = have to, but not compound words, e.g., bath_tub). Third, the

[3] At 30 months one participant who was selected to participate had an audio file of the examiner-child play interaction, but no written transcript. The principal investigator transcribed this 50-utterance analysis set in accordance with the orthographic transcription conventions.

principal investigator listened to each audio file to make additional transcript revisions including revisions of (a) partial words and unintelligible words, (b) utterance boundaries, and (c) transcription of reduced infinitives and full infinitives (e.g., gonna vs. going to).

Utterance boundaries in this study were primarily defined by syntactic boundaries, that is we maintained the syntactic integrity of utterances, and to a lesser degree intonational contour (e.g., Miller & Chapman, 1981; Schuele, 2009b). As a result, utterance boundary revisions in the transcripts included segmentation of independent clauses combined with *and* (e.g., C *I ran and I jumped*. was revised to C *I ran*. C *and I jumped*.). Combined imperatives (e.g., C *Run around the block and jump on the trampoline*) and phrases with the same implied subject (C *She plays games*. C *watches movies*. was revised to C *She plays games *and watches movies*) were not segmented.

Dependent clauses were combined with the immediately prior utterance if the two utterances were topically related (e.g., C *I want to go*. C *Because I'm hungry*. was revised to C *I want to go because I'm hungry*). Likewise, sentence-initial noun phrases (e.g., C *Aunt Mary, she's the best*) and variants of *yes* and *no* (e.g., *Yes, I want to go*) that were topically-related formed one utterance. Constituents that could be syntactically combined as well as tag questions were combined to form one utterance (e.g., C *The dog walks.* C *around the block* was revised to C *The dog walks around the block;* C *pickles*. C *and tomatoes* was revised to C *pickles and tomatoes*). Single words that were repeated across sequential utterances were removed from the analysis set (e.g., C *I bounce the ball*. C *Ball.* was revised to C *I bounce the ball (ball)*).

Complex Syntax Coding

The principal investigator coded all child utterances that included one or more dependent clauses as complex syntax utterances. Some researchers (e.g., Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002) have coded only complete sentences with dependent clauses for complex syntax; in contrast we coded any utterance with a dependent clause as complex syntax whether or not the utterance was a complete sentence (e.g., Schuele & Dykes, 2005). Including children's non-sentence dependent clauses in the analysis set may be particularly important for capturing children's earliest productions of complex syntax. For example, a child may respond to a why question with a dependent clause starting with *because* (e.g., M *Why did you eat the cookie?* C *because I'm hungry*). Each complex syntax token was coded as an *attempt* in complex syntax production regardless of grammaticality based on the gloss. Non-exemplars of complex syntax included attention getters (e.g., *see, I can do it*), repeated verbs (e.g., *go, go, go*), sequential verbs (e.g., *go put it away*), reduced infinitives (i.e., *gonna, gotta, hafta, wanna*), tag questions (e.g., *I'm going, aren't you*?), and passives (e.g., *They got smashed*). Each complex syntax token was coded for complex syntax type (e.g., Schuele, 2009a). See Table 3 for a description of each complex syntax type and an example of each type from the participant transcripts.

Table 3

Complex Syntax Types, Descriptions, and Examples

Complex Syntax Type	Description	Study Transcript Examples
Coordinate Conjunction Clause	Independent clause joined to a dependent clause with a coordinate conjunction	I'm gonna stand up and do that.
Subordinate Conjunction Clause	Clauses joined with subordinate conjunction or single dependent clause headed with subordinate conjunction	And there's a steering wheel so it can back it up.
Marked Infinitive	Nonfinite verb clause that requires obligatory infinitival TO	She always wants to play that.
Unmarked Infinitive	Nonfinite verb clause that does not require infinitival TO	And I let him do lots of things.
Full Propositional Complement Clause	Finite clause that is complement of mental state verbs (e.g., *think*) or verbs of communication (e.g., *tell*); complementizer may be optional or required depending on verb	My mom said it will be some milk and a graham_cracker.
WH Finite Clause	Finite WH clause - *when, what, where, how,* or *why* - that is complement of mental state verb (e.g., *think*) or verb of communication (e.g., *tell*)	I don't remember what Mr_Potato_Head talks like.
WH Nonfinite Clause	Nonfinite WH clause - *when, what, where, how,* or *why* - that is complement of mental state verb (e.g., *think*) or verb of communication (e.g., *tell*)	I learned how to swim with my head up.
Subject Relative Clause	Finite clause that modifies a noun phrase; gap in the relative clause is in the subject position; obligatory relative pronoun or relative marker *that*	Some tractors that don't weigh a lot, they'll go really slow.
Other Relative Clause	Includes all other finite clauses that modify a noun phrase, (a) e.g., gap in the relative clause in object or indirect object position, optional relative pronoun or relative marker *that,* and (b) adjunct relative clause, with relative *where* or *why*	The guy you're holding is the baby's dad.
Nominal Relative Clause	Finite clause in which the modified noun phrase and the relative pronoun are coalesced (i.e., no overt noun phrase)	This is where I save people.
Participle Clause	Nonfinite clause that includes a past participle or a present participle	You can go swimming in that lake.
Other	Any dependent clause that does not fit into one of the above categories	It's like they're pulling wagons.

Note. Complex syntax types and descriptions reproduced from Schuele, Teller, Kaiser, & Camarata, in preparation; adapted from Barako Arndt & Schuele (2013).

Additionally, each complex syntax token was coded for the presence of errors in the syntactic or semantic production of each dependent clause (see Table 4). Complex syntax error codes were not created *a priori*, but rather derived through a review of the complex syntax tokens and the existing literature. The following errors have been reported in the literature: omission of infinitival TO, of obligatory relative clause markers, and of WH pronouns and semantic and syntactic errors in the production of subordinate conjunctions (e.g., Barako Arndt & Schuele, 2012; Schuele & Dykes, 2005; Schuele & Tolbert, 2001; Weiler & Schuele, 2014). Errors in tense and agreement were not coded as errors in the production of the dependent clauses.

Table 4

Complex Syntax Error Descriptions and Examples

Complex Syntax Type	Description of Observed Errors	Study Transcript Examples
Coordinate Conjunction Clause	Missing conjunction	I wanna do this *and put this in the toys.
Subordinate Conjunction Clause	Omitted or misused conjunction	*If you squeeze his feet, his hands will come up (intonation + context demonstrate need for subordinate conjunction); for the babies can sleep (for used instead of so)
Marked Infinitive	Error in word order for infinitive clause; Omitted infinitival TO	Baby needs *to sleep by them; no I need to this come out
Unmarked Infinitive	NA	No exemplars
Full Propositional Complement Clause	Omitted complement clause verb; inappropriate complementizer	No how about we *pretend it's a firestation.
WH Finite Clause	Omitted or inappropriate WH word	Look *what I found.
WH Nonfinite Clause	Omitted or inappropriate WH word or TO;	I don't know how *to do it.
Subject Relative Clause	Omitted relativizer	
Other Relative Clause	Inappropriate relativizer	That's playdough when you squeeze and it comes out.
Nominal Relative Clause	Omitted or inappropriate production of relative pronoun	We thought what is that? Gloss = We thought about what that was.
Participle Clause	Agrammatical production of participle (e.g., go for going)	Go biking ride
Other	Agrammatical complex syntax utterance that does not fit one of the above categories	And that goes right eat here.

Note. * = omitted word.

Derivation of Dependent Variables

MLU in morphemes and number of different words (NDW) were derived in SALT as descriptive language sample measures. MLU captures growth in syntactic complexity and has been found to correlate with complex syntax variables (Paul, 1981). NDW captures growth in lexical diversity, which is relevant to delays in productive vocabulary in late talkers. These variables allow for a descriptive comparison of growth in more frequently-used language sample measures alongside growth in the complex syntax variables.

Three variables were derived to capture complex syntax quantity: (a) complex syntax proportion, (b) number of complex syntax tokens, and (c) complex syntax density. Complex syntax proportion was defined as the proportion of 50 utterances with at least one complex syntax token. Number of complex syntax tokens was defined as the raw frequency of complex syntax tokens across 50 utterances. An utterance could contain more than one complex syntax token. Complex syntax density was defined as the sum of complex syntax tokens divided by the number of utterances that contained at least one complex syntax token (variable minimum = 1.0; i.e., one complex syntax token per utterance). Complex syntax density was derived only for participants whose analysis set included at least one complex syntax token. See Table 5 for the proportion of participants who produced at least one complex syntax token at each time point.

Table 5

Proportion of Participants Who Produced at Least One Complex Syntax Token by Time and Group

	Participant Group	
Time	NL Group	LT Group
30 months	0.81	0.31
42 months	0.81	0.92
54 months	0.94	1.00
66 months	1.00	0.96

Note. NL group, n = 16 and LT group, n = 26 with the exception of the NL group at 66 months (n = 15), the LT group at 54 months (n = 25), and the LT group at 66 months (n = 24).

One variable was derived to capture complex syntax diversity. Complex syntax type diversity was defined as the number of different complex syntax types (max = 11; other excluded) in the 50-utterance analysis set and was a measure of the variety of dependent clauses that each child produced. Complex syntax type diversity was a measure of vertical complex syntax development.

Exploratory variables. Quantity variables were derived for each of four complex syntax categories by combining complex syntax types that have similar linguistic properties (Diessel, 2004; Schuele, Teller, Kaiser, & Camarata, in preparation). The infinitival clause composite was defined as the summed frequency of marked infinitive tokens plus unmarked infinitive tokens. The complement clause composite was defined as the summed frequency of full propositional complement tokens plus WH finite clause tokens plus WH nonfinite clause tokens. The combined clause composite was defined as the summed frequency of coordinate conjunction clause tokens plus subordinate conjunction clause tokens. The relative clause composite was defined as the summed frequency of subject relative clause tokens plus other relative clause tokens plus nominal relative clause tokens.

Some types of complex syntax are reliant on the production of specific lexical items. For example, subordinate conjunctions are produced as part of subordinate conjunction clauses and complement clause verbs subcategorize for infinitival clauses and/or complement clauses. Three complex syntax vocabulary variables were derived from a subset of utterances in the transcript for each participant at each time: (a) complement clause verb diversity with marked infinitives, (b) complement clause verb diversity with complement clauses and (c) subordinate conjunction diversity. Complex syntax vocabulary was a measure of horizontal development. Complement clause verb diversity with marked infinitives was defined as the number of different complement clause verbs that were produced with marked infinitives (excludes copula BE as main verb, used to, BE + supposed to). Complement clause verb diversity with complement clauses was defined as the number of different complement clause verbs (CCV) in the subset of utterances coded as including complement clauses – full propositional complements, WH finite clauses, and WH nonfinite clauses. Subordinate conjunction diversity was defined as the number of different subordinate conjunctions in the subset of utterances coded as including subordinate conjunction clauses. See Table 3 for definitions of complex syntax types that are reliant on complex syntax vocabulary.

Two exploratory variables were derived to describe the frequency of errors in complex syntax tokens: proportion of errored productions per complex syntax token and proportion of errored productions in each complex syntax category. Participants who did not produce any complex syntax tokens at a given time point were not included in this analysis. For proportion of errored productions per complex syntax token we divided the sum of errors in complex syntax production by the total number of complex syntax tokens for each child at each time point. We then calculated the mean proportion of tokens in error within each complex syntax category in each group at each time point. To derive the proportion of errored productions in each complex syntax category we tallied a raw count of complex syntax tokens in error within each complex syntax category for each participant at each time point. We then calculated the proportion of tokens in error from each complex syntax category relative to all the tokens in each complex syntax category for each participant at each time point. Finally, we calculated the mean proportion of tokens in error within each complex syntax category in each group at each time point.

Reliability

Training. The principal investigator trained a research assistant who had previous orthographic transcription and complex syntax coding experience to check orthographic transcription and complex syntax coding. The research assistant completed approximately 10 hours of training, which involved reading lab coding manuals (Schuele, 2009a; Schuele, 2009b) and a study-specific manual supplement as well as practice checks of orthographic transcription and complex syntax coding.

For training purposes the principal investigator revised eight non-study transcripts from the Ellis Weismer Corpus according to orthographic transcription conventions and coded each transcript for complex syntax. The research assistant and the principal investigator reviewed the orthographic transcription and complex syntax coding for one non-study practice transcript together. Subsequently, the research assistant independently checked orthographic transcription and complex syntax coding on seven additional non-study practice transcripts. Disagreements with the principal investigator's transcription and complex syntax coding were discussed with the principal investigator after the research assistant completed each transcript.

Orthographic transcription reliability. The research assistant checked orthographic transcription for 31% (n = 5) of the NL group transcripts and 23% (n = 6) of the LT group transcripts at each time point. After the principal investigator listened to and aligned a set of two to five transcripts, with a few exceptions, according to study-specific rules, the research assistant randomly selected one transcript from the set to check for orthographic transcription. The research assistant was blind to child group and child age. After each transcript was checked, the principal investigator and research assistant discussed any transcription disagreements for training purposes.

Reliability was calculated for morpheme-by-morpheme orthographic transcription and utterance segmentation by comparing the research assistant's checked transcript to the principal investigator's original transcript. Reliability for morpheme-by-morpheme orthographic transcription was calculated by summing the total number of words + bound morphemes – disagreements and dividing this sum by the sum of total number of words + bound morphemes. Total number of words and bound morphemes was derived from the principal investigator's transcript using SALT. Average reliability across checked transcripts at each time point ranged from 96 – 99% for morpheme-by-morpheme orthographic transcription. Reliability for utterance segmentation was calculated by subtracting utterance segmentation disagreements from 50 and then dividing that sum by 50. Average reliability across checked transcripts at each time point ranged from 99 – 100% for utterance segmentation.

Transcript revisions by consensus. The research assistant reviewed each printed transcript and marked disagreements in slashing bound morphemes, spelling, and mazes. The principal investigator reviewed the disagreements and consensus was achieved on these additional changes to the transcripts.

Complex syntax coding consensus. Likewise, complex syntax coding was reviewed and final coding was achieved via consensus as has been typically completed in the extant literature (e.g., Dunn Davison, Schuele, Fisher, Dickinson, & Combs, 2020). The research assistant reviewed each printed transcript and marked complex syntax coding disagreements. The principal investigator and research assistant discussed disagreements. Unresolved differences in complex syntax coding were discussed with the dissertation committee chair to establish consensus. The principal investigator made the agreed-upon coding changes to each transcript to create a final transcript for analysis.

Data Analysis Plan

Prior to data analysis we assessed the potential influence of individual data points on the results and the distribution of each dependent variable. No data points had a substantial influence on the regression model for any of the dependent variables with time and group as predictor variables (Cook's D < 1). Based on the Shapiro-Wilk test MLU and NDW were normally distributed at each time point (*p*s > .05). In contrast, none of the complex syntax variables were consistently normally distributed (*p*s < .05). Thus, we employed parametric statistics for MLU and NDW and non-parametric statistics for each complex syntax variable.

Statistical analyses were completed within IBM SPSS (Version 25). A repeated measures ANOVA was employed to assess main effects as well as the time by group interaction for MLU and NDW. A Friedman test (i.e., non-parametric alternative of repeated measures ANOVA) was employed to assess the main effect of time for each complex syntax variable in each group; the interaction between time and group cannot be assessed using a Friedman test. Listwise deletion was applied to missing data (NL, n = 3; LT, n = 1) and thus, the repeated measures ANOVAs and the Friedman tests represent a subset of the participants in the study; these analyses included 94% (15/16) of the NL group and 88% (23/26) of the LT group for each dependent variable except for complex syntax density, which included 56% (9/16) of the NL group and 27% (7/26) of the LT group. These analyses for complex syntax density included only participants who produced at least one complex syntax token at each time.

Follow-up analyses were employed to determine growth between adjacent time points and to compare between-group performance at each time point. All follow-up analyses were one-tailed with the expectation that performance would increase at annual intervals and that the NL group would demonstrate greater performance as compared to the LT group at each time point. Significance was set at $p < .05$ for all statistical tests. We did not correct for familywise error due to the exploratory nature of this study. Effect sizes for follow-up analyses were interpreted with values of .20 to .49 considered to be small, values of .50 to .79 considered to be moderate, and values of 0.80 or greater considered to be large (Cohen, 1988).

CHAPTER III

RESULTS

Description of Statistics

We first report descriptive statistics for MLU and NDW and for each complex syntax variable. We included median and range for each variable that was assessed non-parametrically. Additionally, we report correlations between all MLU, NDW, complex syntax variables, and complex syntax vocabulary variables. We next report results for the repeated measures ANOVAs and follow-up analyses for MLU and NDW to describe performance on these common language sample measures for each group. We then present results for each research question. Finally, we conclude the results section with several exploratory analyses to further describe between-group differences.

Table 6

MLU and NDW by Time and Group

	NL Group		LT Group		
Time	Mean	SD	Mean	SD	*d*
MLU					
30 months	3.22	0.65	2.12	0.68	1.64***
42 months	4.19	0.69	3.88	0.63	0.47
54 months	4.86	0.70	4.72	0.70	0.20
66 months	6.32	1.33	6.28	1.15	0.03
NDW					
30 months	63.06	10.99	44.65	11.78	1.60***
42 months	81.81	11.69	78.35	11.77	0.29
54 months	97.88	13.68	89.80	14.47	0.57*
66 months	124.40	21.19	120.54	19.42	0.19

Note. *$p < .05$; **$p < .01$; ***$p < .001$ NL group, n = 16 and LT group, n = 26 with the exception of the NL group at 66 months (n = 15), the LT group at 54 months (n = 25), and the LT group at 66 months (n = 24).

Table 7

Descriptive Statistics for Complex Syntax Variables by Time and Group

	NL Group				LT Group				
Time	Mean	SD	Median	Range	Mean	SD	Median	Range	*d*
				Complex Syntax Proportion					
30 months	0.03	0.03	0.02	0.00 – 0.08	0.01	0.02	0.00	0.00 – 0.08	0.83**
42 months	0.10	0.07	0.10	0.00 – 0.24	0.08	0.06	0.06	0.00 – 0.24	0.31
54 months	0.17	0.07	0.18	0.00 – 0.26	0.14	0.07	0.14	0.04 – 0.28	0.43
66 months	0.25	0.11	0.24	0.08 – 0.44	0.22	0.11	0.22	0.00 – 0.42	0.27
				Number of Complex Syntax Tokens					
30 months	1.69	1.49	1.00	0 – 5	0.50	0.95	0.00	0 – 4	1.01**
42 months	5.81	4.65	5.00	0 – 15	4.12	3.49	3.00	0 – 14	0.43
54 months	9.63	4.18	10.00	0 – 17	7.92	4.32	7.00	2 – 16	0.40
66 months	17.27	8.95	18.00	4 – 35	14.58	8.65	14.00	0 – 32	0.31
				Complex Syntax Density[a]					
30 months	1.02	0.07	1.00	1.00 – 1.25	1.00	0.00	1.00	1.00 – 1.00	0.36
42 months	1.11	0.20	1.00	1.00 – 1.67	1.05	0.11	1.00	1.00 – 1.42	0.41
54 months	1.15	0.13	1.17	1.00 – 1.33	1.15	0.16	1.09	1.00 – 1.50	0.00
66 months	1.32	0.23	1.32	1.00 – 1.73	1.27	0.24	1.22	1.00 – 1.82	0.21
				Complex Syntax Type Diversity					
30 months	1.31	1.08	1.00	0 – 3	0.38	0.64	0.00	0 – 2	1.12**
42 months	2.75	1.81	3.00	0 – 5	2.23	1.24	2.00	0 – 5	0.35
54 months	5.06	2.05	5.00	0 – 8	4.12	1.59	4.00	2 – 7	0.53*
66 months	6.20	1.86	6.00	2 – 9	4.88	2.01	5.00	0 – 8	0.67*

Note. **p* < .05; ***p* < .01; ****p* < .001; Participants for all variables excluding complex syntax density NL group, n = 16 and LT group, n = 26 with the exception of the NL group at 66 months (n = 15), the LT group at 54 months (n = 25), and the LT group at 66 months (n = 24). [a]Participants for Complex Syntax Density at 30 months: NL group, n = 13; LT group, n = 8; at 42 months, NL group, n = 13; LT group, n = 24; at 54 months: NL group, n = 15; LT group, n = 25; at 66 months: NL group, n = 15; LT group, n = 23.

Correlations. To characterize the associations between dependent variables, we calculated correlations at each time point across groups using Pearson's *r* (see Table 8 – 11). All dependent variables were correlated at each time (*p*s < .01) except complex syntax density, which was inconsistently correlated with the other variables.

Table 8

Correlations at 30 Months

Variables	1.	2.	3.	4.	5.	6.
1. MLU		.86***	.60***	.59***	.14	.66***
2. NDW			.57***	.55***	.01	.62***
3. CSP				.99***	.43	.93***
4. # of CS Tokens					.56**	.93***
5. CS Density						.43
6. CS Type Diversity						--

Note. *p < .05, **p < .01, ***p < .001. Participants, n = 42 except for CS Density, n = 21. CSP = Complex syntax proportion; # of CS Tokens = Number of complex syntax tokens; CS Density = Complex syntax density; CS Type Diversity = Complex syntax type diversity.

Table 9

Correlations at 42 Months

Variables	1.	2.	3.	4.	5.	6.
1. MLU		.65***	.65***	.62***	.27	.58***
2. NDW			.50**	.56***	.51**	.56***
3. CSP				.97***	.44**	.85***
4. # of CS Tokens					.66***	.85***
5. CS Density						.50***
6. CS Type Diversity						--

Note. *p < .05, **p < .01, ***p < .001. Participants, n = 42 except for CS Density, n = 21. CSP = Complex syntax proportion; CS Density = Complex syntax density; CS Type Diversity = Complex syntax type diversity.

Table 10

Correlations at 54 Months

Variables	1.	2.	3.	4.	5.	6.
1. MLU		.78***	.66***	.71***	.30	.53***
2. NDW			.50***	.53***	.26	.41**
3. CSP				.96***	.02	.81***
4. # of CS Tokens					.28	.81***
5. CS Density						.15
6. CS Type Diversity						--

Note. *p < .05, **p < .01, ***p < .001. Participants, n = 40 except for CS Density, n = 41. CSP = Complex syntax proportion; # of CS Tokens = Number of complex syntax tokens; CS Density = Complex syntax density; CS Type Diversity = Complex syntax type diversity.

Table 11

Correlations at 66 Months

Variables	1.	2.	3.	4.	5.	6.
1. MLU		.85***	.83***	.83***	.56***	.65***
2. NDW			.75***	.71***	.44**	.60***
3. CSP				.96***	.54***	.83***
4. # of CS Tokens					.75***	.82***
5. CS Density						.60***
6. CS Type Diversity						--

Note. *p < .05, **p < .01, ***p < .001. Participants, n = 39 except for CS Density, n = 38. CSP = Complex syntax proportion; # of CS Tokens = Number of complex syntax tokens; CS Density = Complex syntax density; CS Type Diversity = Complex syntax type diversity.

MLU. The repeated measures ANOVA showed a main effect for time for MLU, $F(2.21) = 141.36$, $p < .001$, $\eta_p^2 = .80$ and for group, $F(1) = 5.07$, $p < .05$, $\eta_p^2 = .12$. The assumption of sphericity was not met based on Mauchly's Test of Sphericity (*ps* < .05). Thus, the Huynh-Feldt correction was applied to degrees of freedom. The group by time interaction was significant, $F(2.21) = 4.08$, $p < .05$, $\eta_p^2 = .10$ indicating differential growth over time between groups (see Figure 2).

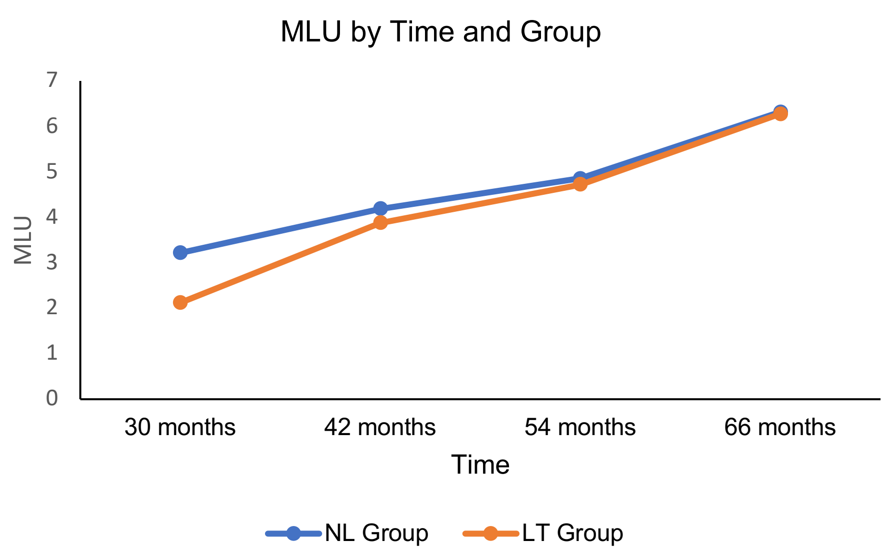

Figure 2. MLU by time and group.

Planned follow-up one-tailed paired t-tests were conducted to assess growth between adjacent time points (see Table 12 for results). Between 30 and 42 months all of the participants in each group were included in analyses. Due to missing data, between 42 and 54 months 96% (25/26) of the LT group were included and between 54 and 66 months 94% (15/16) of the NL group and 88% (23/26) of the LT group were included. There was significant growth between adjacent time points in each group with the largest effect size in each group between 30 months and 42 months.

Table 12

Effect Sizes for MLU and NDW Between Time Points by Group

	NL Group	LT Group
Time	*d*	*d*
MLU		
30 months vs. 42 months	1.24***	2.44***
42 months vs. 54 months	0.69**	1.51***
54 months vs. 66 months	0.94**	1.41***
NDW		
30 months vs. 42 months	1.52***	2.74***
42 months vs. 54 months	0.79**	0.73***
54 months vs. 66 months	1.13***	2.09***

Note. $*p < .05$, $**p < .01$, $***p < .001$. NL group, n = 16 and LT group, n = 26 with the exception of the NL group at 54 vs. 66 months (n = 15), the LT group at 42 vs. 54 months (n = 25), and the LT group at 54 vs. 66 months (n = 23).

Planned follow-up one-tailed independent sample t-tests were conducted to assess between-group differences at each time point (see Table 6). All of the participants in each group were included in analyses at 30 and 42 months. Due to missing data, at 54 months 96% (25/26) of the LT group were included and at 66 months 94% (15/16) of the NL group and 92% (24/26) of the LT group were included. There was a significant between-group difference only at 30 months, $t(40) = 5.18$, $p < .001$, $d = 1.64$, one tailed. There was significant growth between adjacent time points in each group with the largest effect size in each group between 30 months and 42 months.

NDW. The repeated measures ANOVA showed a main effect for time for NDW, $F(2.51) = 177.16$, $p < .001$, $\eta_p^2 = .83$ and for group, $F(1) = 6.14$, $p < .05$, $\eta_p^2 = .15$. The interaction was non-significant. Again, the assumption of sphericity was not met based on Mauchly's Test of Sphericity (*ps* < .05). Thus, the Huynh-Feldt correction was applied to degrees of freedom.

Follow-up analyses included the same participants as were included for MLU. Planned follow-up one-tailed paired t-tests were conducted to assess growth between adjacent time points (see Table 12 for results). Again, there was significant growth between adjacent time points in each group with the greatest growth between 30 months and 42 months. Planned follow-up one-tailed independent sample t-tests were conducted to assess between-group differences at each time point (see Table 6). There was a significant between-group difference at 30 months, $t(40) = 5.04$, $p < .001$, $d = 1.60$, one tailed and at 54 months, $t(39) = 1.78$, $p < .05$, $d = 0.57$. The NL group produced a greater NDW at each time point as compared to the LT group.

Longitudinal Development and Between-Group Differences

We now address each research question in turn. Our first research question was: Within each language group is there growth in complex syntax quantity and diversity at 12-month intervals? Our second research question was: Are there between-group differences in the complex syntax quantity and diversity at each measurement time point? To address these research questions, we explored longitudinal growth and between-group differences in each complex syntax variable at 12-month intervals. One Friedman test was completed for each group for each of the following variables: (a) complex syntax proportion, (b) number of complex syntax tokens, (c) complex syntax density, and (d) complex syntax type diversity.

Follow-up analyses. Again due to missing data, follow-up analyses included varying numbers of participants. For differences between time points for each complex syntax variable (with some exceptions for complex syntax density) between 42 and 54 months 96% (25/26) of the LT group were included and between 54 and 66 months 94% (15/16) of the NL group and 88% (23/26) of the LT group were included. Complex syntax density was only calculated at each time point for participants who produced at least one complex syntax token. Thus, complex syntax density was calculated for a different number of participants as compared to other complex syntax variables. For complex syntax density between 30 and 42 months, 63% (10/16) of the NL group but only 31% (8/26) of the LT group were included, between 42 and 54 months 75% (12/16) of the NL group and 88% (23/26) of the LT group were included, and between 54 and 66 months 88% (14/16) of the NL group and 85% (22/26) of the LT group were included.

For between-group differences all participants in each group were included at 30 and 42 months. For each complex syntax variable (with some exceptions for complex syntax density) at 54 months 96% (25/26) of the LT group were included and at 66 months 94% (15/16) of the NL group and 92% (24/26) of the LT group were included. For complex syntax density at 30 months 81% (13/16) of the NL group but only 31% (8/26) of the LT group were included. At 42 months 81% (13/16) of the NL group and 92% (24/26) of the LT group were included, at 54 months 94% (15/16) of the NL group and 96% (25/26) of the LT group, and at 66 months 94% (15/16) of the NL group and 88% (23/26) of the LT group were included.

Table 13

Effect Sizes for Complex Syntax Variables Between Time Points by Group

Time	NL Group *d*	LT Group *d*
Complex Syntax Proportion		
30 months vs. 42 months	0.91**	1.22***
42 months vs. 54 months	0.72**	0.72***
54 months vs. 66 months	0.74**	0.72**
Number of Complex Syntax Tokens		
30 months vs. 42 months	0.81**	1.14***
42 months vs. 54 months	0.67**	0.72***
54 months vs. 66 months	0.85**	0.76***
Complex Syntax Density[a]		
30 months vs. 42 months	0.17	0.77
42 months vs. 54 months	0.28	0.45*
54 months vs. 66 months	0.67*	0.37*
Complex Syntax Type Diversity		
30 months vs. 42 months	0.66**	1.65***
42 months vs. 54 months	0.80**	0.89***
54 months vs. 66 months	0.36	0.32

Note. *$p < .05$, **$p < .01$, ***$p < .001$. NL group, n = 16 and LT group, n = 26 with the exception of the NL group at 54 vs. 66 months (n = 15), the LT group at 42 vs. 54 months (n = 25), and the LT group at 54 vs. 66 months (n = 23). [a]Participants for Complex Syntax Density at 30 vs. 42 months: NL group, n = 10; LT group, n = 8; at 42 months vs. 54 months, NL group, n = 12; LT group, n = 23; at 54 vs. 66 months: NL group, n = 14; LT group, n = 22.

Complex syntax proportion. The Friedman test for complex syntax proportion was significant for time for the NL group, $F_{r(3)} = 30.87$, $p < .001$ and for the LT group, $F_{r(3)} = 52.80$, $p < .001$. Planned follow-up Wilcoxon tests (i.e., non-parametric alternative of paired t-tests) were conducted to assess growth between adjacent time points (see Table 13 for results). There was significant growth between adjacent time points in each group with the largest effect size in each group between 30 months and 42 months.

Planned follow-up one-tailed Mann-Whitney *U* tests (i.e., non-parametric alternative for independent t-tests) were conducted to assess between-group differences at each time point (see Table 7). Using the Mann-Whitney *U* test, there was a significant between-group difference only at 30 months (U = 99, n_{NL} = 16, n_{LT} = 26, $p < .01$, d = 0.83). The NL group produced a greater complex syntax proportion at 30 months as compared to the LT group.

Number of complex syntax tokens. The Friedman test for number of complex syntax tokens was significant for time for the NL group, $F_{r(3)} = 30.43$, $p < .001$ and for the LT group, $F_{r(3)} = 53.19$, $p < .001$. Planned follow-up paired Wilcoxon tests were conducted to assess growth between adjacent time points (see Table 13 for results). There was significant growth between adjacent time points in each group.

Planned follow-up one-tailed Mann-Whitney *U* Tests were conducted to assess between-group differences at each time point (see Table 7). Using the Mann-Whitney *U* test, there was a significant between-group difference only at 30 months (U = 95, n_{NL} = 16, n_{LT} = 26, $p < .001$, d = 1.01). At 30 months the NL group produced a greater number of complex syntax tokens as compared to the LT group.

Complex syntax density. The Friedman test for complex syntax density was significant for time for the NL group $F_{r(3)} = 13.88$, $p < .01$ and for the LT group, $F_{r(3)} = 8.76$, $p < .05$. Planned follow-up Wilcoxon tests were conducted to assess growth between adjacent time points (see Table 13 for results). There was significant growth between 54 months and 66 months in the NL group and between 42 months and 54 months and 54 months and 66 months in the LT group. Planned follow-up one-tailed Mann-Whitney *U* Tests were conducted to assess between-group differences at each time point (see Table 6). The effect size between 30 and 42 months in the

LT group suggests that the analysis at that time point may have been underpowered to detect an effect. Using the Mann-Whitney *U* test, between-group differences were not significant at any time point (*ps* > .05).

Complex syntax type diversity. The Friedman test for complex syntax type diversity was significant for time for the NL group, $F_{r(3)} = 31.01$, $p < .001$ and for the LT group, $F_{r(3)} = 53.20$, $p < .001$. Planned follow-up Wilcoxon tests were conducted to assess growth between adjacent time points (see Table 13 for results). There was significant growth between 30 months and 42 months and between 42 months and 54 months in each group.

Planned follow-up Mann-Whitney *U* Tests were conducted to assess between-group differences at each time point (see Table 7). Using the Mann-Whitney *U* test, there was a between-group difference at 30 months ($U = 96$, $n_{NL} = 16$, $n_{LT} = 26$, $p < .001$, $d = 1.12$), 54 months ($U = 133.5$, $n_{NL} = 16$, $n_{LT} = 25$, $p < .05$, $d = 0.53$), and 66 months ($U = 115$, $n_{NL} = 15$, $n_{LT} = 24$, $p < .05$, $d = 0.67$). At each time point, the NL group produced about one more complex syntax type as compared to the LT group.

In summary, for our first and second research questions each variable showed fairly consistent growth over time except complex syntax density for which the number of participants between time points varied widely. Between-group differences were most prominent at 30 months with additional between-group differences for complex syntax type diversity at 54 and 66 months.

Individual Differences in Complex Syntax Development

Our third research question was: What percent of late talkers have complex syntax proportion and diversity no more than one SD below the mean of same-aged peers (i.e., average range) at each time point? To address this research question, we first calculated the mean and standard deviation for the NL group for one quantity variable – complex syntax proportion – and one diversity variable – complex syntax type diversity (see Table 14 and 15). We then determined the proportion of participants in the LT group who performed within the following ranges relative to the NL group mean at each time point: (a) greater than or equal to -1.00 SD, (b) less than -1.00 SD to greater than or equal to -1.50 SD, (c) less than -1.50 to greater than or equal to -2.00 SD, and (d) less than -2.00 SD below the NL group mean. Proportions greater than 80% approximate the number of NL participants in the average range.

Table 14

Proportion of the LT Group Producing Complex Syntax Proportion Relative to the NL Group Mean

Time	≥ -1.00 SD*	< -1.00 to ≥ -1.50 SD	< -1.50 to ≥ -2.00 SD	< -2.00 SD
30 months	0.31	0.69	NA	NA
42 months	0.77	0.23	NA	NA
54 months	0.68	0.04	0.28	0.00
66 months	0.79	0.08	0.08	0.04

Note. *Average range. NA because range is below 0. NL group, n = 16 and LT group, n = 26 with the exception of the NL group at 66 months (n = 15), the LT group at 54 months (n = 25), and the LT group at 66 months (n = 24).

Table 15

Proportion of the LT Group Producing Complex Syntax Type Diversity Relative to the NL Group Mean

Time	≥ -1.00 SD*	< -1.00 to ≥ -1.50 SD	< -1.50 to ≥ -2.00 SD	< -2.00 SD
30 months	0.31	0.69	NA	NA
42 months	0.92	0.08	0.00	NA
54 months	0.60	0.40	0.00	0.00
66 months	0.58	0.13	0.21	0.08

Note. *Average range. NA because range is below 0. NL group, n = 16 and LT group, n = 26 with the exception of the NL group at 66 months (n = 15), the LT group at 54 months (n = 25), and the LT group at 66 months (n = 24).

There was a dramatic increase in the proportion of the LT group who performed within the average range for each variable between 30 months and 42 months. For complex syntax proportion, the proportion the LT group who performed in the average range dipped at 54 months but at 66 months returned to a similar proportion of participants who performed within the average range as was demonstrated at 42 months. In contrast, for complex syntax type diversity the proportion of the LT group who performed in the average range dipped at 54 months, but did not increase again at 66 months like complex syntax proportion did.

Emergence of Complex Syntax Categories

The fourth research question was: What proportion of participants produce exemplars from each complex syntax category at each time point measured cumulatively inclusive of previous time points? We explored emergence of complex syntax categories by measuring the cumulative proportion of participants in each group who produced at least one exemplar from each complex syntax category by each time point (see Table 16). For example, a child who produced a complement clause at 54 months, but not at 66 months received credit for a complement clause at 54 months as well as at 66 months. This cumulative measure was deemed appropriate due to the brevity of the language transcript analysis set.

Table 16

Cumulative Attempts of Complex Syntax Categories by Time and Group

	Complex Syntax Category							
	NL	LT	NL	LT	NL	LT	NL	LT
Time	Infinitival Clauses		Complement Clauses		Combined Clauses		Relative Clauses	
30 months	0.50	0.19	0.06	0.00	0.25	0.19	0.25	0.00
42 months	0.88	0.81	0.63	0.42	0.63	0.46	0.38	0.15
54 months	1.00	0.92	0.94	0.80	0.94	0.80	0.75	0.64
66 months	1.00	1.00	1.00	0.92	1.00	1.00	1.00	0.83

Note. NL group, n = 16 and LT group, n = 26 with the exception of the NL group at 66 months (n = 15), the LT group at 54 months (n = 25), and the LT group at 66 months (n = 24).

Infinitival clauses. At 30 months the proportion of participants in the NL group (.50) who produced infinitival clauses was substantially greater than the proportion of participants in the LT group (.19). By 42 months over .80 of the participants within each group produced infinitival clauses. The increase in participants who produced infinitival clauses was dramatic between 30 months and 42 months. The dramatic increase was particularly evident for the LT group (.81), which almost caught up to the NL group (.88) by 42 months. By 54 months virtually all of the participants in each group had produced at least one exemplar of an infinitival clause.

Complement clauses. At 30 months very few of the participants in the NL group or the LT group produced complement clauses, a proportion of .06 and .00 respectively. By 42 months the majority (.63) of the participants in the NL group produced complement clauses, whereas the majority of the participants in the LT group did not (.42). Majority was defined as greater than .50. By 54 months at least .80 of the participants within each group produced complement clauses, but the LT group (.80) lagged behind the NL group (.94). By 66 months virtually all of the participants in each group had produced at least one exemplar of a complement clause.

Combined clauses. At 30 months the proportion of participants in the NL group (.25) who produced combined clauses was similar to the proportion of participants in the LT group (.19). Like complement clause production, by 42 months the majority (.63) of the participants in the NL group produced combined clauses, whereas the majority of the participants in the LT group did not (.46). By 54 months at least .80 of the participants across both groups produced combined clauses but again the LT group (.80) lagged behind the NL group (.94). By 66 months all of the participants in each group had produced at least one exemplar of a combined clause.

Relative clauses. At 30 months the proportion of participants in the NL group (.25) who produced relative clauses was substantially greater than the proportion of participants in the LT group (.00). By 42 months the proportion of participants who produced relative clauses in the NL group and in the LT group increased to.38 and .15 respectively. At 54 months the majority (.75) of the participants in the NL group and in the LT group (.64) produced relative clauses. By 66 months all of the participants in the NL group and over .80 of the participants in the LT group produced at least one exemplar of a relative clause.

Composition of Complex Syntax Categories

The fifth research question was: What is the proportional composition of tokens from each complex syntax category relative to total complex syntax tokens at each time point? To address this research question, we calculated the proportion of the tokens in each complex syntax category relative to the sum of the complex syntax tokens in the four complex syntax categories at each time point (see Table 17). Participants who did not produce any complex syntax tokens at a given time point were excluded from the analysis (see Table 4). Again, majority was defined as a proportion greater than .50.

At 30 months complex syntax tokens represented each complex syntax category in the NL group, whereas the LT group produced complex syntax only from the infinitival clause and combined clause categories. At 30 months infinitival clauses comprised the majority (.60) of the complex syntax tokens produced by the NL group and infinitival clauses (.50) and combined clauses (.50) comprised the majority of the complex syntax tokens produced by the LT group. The low number of complex syntax tokens produced by the LT group likely influenced the proportion at 30 months.

Table 17

Proportion of Total Complex Syntax Tokens by Complex Syntax Category by Time and Group

			Complex Syntax Category							
Time	Number of Complex Syntax Tokens Across All Participants in Group		Infinitival Clauses		Complement Clauses		Combined Clauses		Relative Clauses	
	NL	LT	NL	LT	NL	LT	NL	LT	NL	LT
30 months	24	13	0.60	0.50	0.04	0.00	0.14	0.50	0.23	0.00
42 months	87	102	0.41	0.62	0.35	0.22	0.23	0.12	0.01	0.05
54 months	148	191	0.23	0.24	0.37	0.32	0.27	0.28	0.12	0.15
66 months	240	325	0.25	0.32	0.13	0.11	0.40	0.44	0.23	0.14

Note. 30 months: NL group, n = 13; LT group, n = 8; 43 months, NL group, n = 13; LT group, n = 24; 54 months: NL group, n = 15; LT group, n = 25; 66 months: NL group, n = 15; LT group, n = 23.

At 42 months the proportion of infinitival clauses (.41) and relative clauses (.01) produced by the NL group decreased alongside an increase in complement clauses (.35) and combined clauses (.23). In contrast, although complex syntax tokens produced by the LT group at 42 months comprised tokens from each complex syntax category, infinitival clauses represented the majority (.62) of complex syntax tokens. The increase in proportion of complement clauses was particularly notable in each group between 30 months and 42 months.

At 54 months the proportion of complex syntax tokens from each category was similar in each group representing the majority of the complex syntax tokens produced by participants in the NL group (.37) and the LT group (.32). At 66 months the proportion of complement clauses in each group decreased as compared to 54 months with a notable increase in combined clauses and relative clauses for the NL group and in combined clauses for the LT group. At 66 months the proportion of complex syntax tokens from each category was similar across groups except for relative clauses, which were more prominent in the NL group.

Exploratory Analyses

We performed several exploratory analyses to further describe between-group differences.

Complex syntax categories. To compare between-group differences for complex syntax categories we conducted a series of one-tailed Mann-Whitney *U* tests for the frequency of exemplars within each complex syntax category at each time point. There were no between-group differences for complement clauses nor combined clauses (*p*s > .05; see Table 18). At 30 months there was a significant between-group difference for infinitival clauses ($U = 140$, $n_1 = 16$, $n_2 = 26$, $p < .05$, $d = .74$) and relative clauses ($U = 156$, $n_1 = 16$, $n_2 = 26$, $p < .05$, $d = .91$). The NL group produced more infinitival clauses and relative clauses as compared to the LT group. Additionally, at 66 months there was a significant between-group difference for relative clauses

($U = 97$, $n_1 = 15$, $n_2 = 24$, $p < .01$, $d = .76$). The NL group produced more relative clauses as compared to the LT group.

Table 18

Complex Syntax Categories by Time and Group

	NL Group				LT Group				
Time	Mean	SD	Median	Range	Mean	SD	Median	Range	*d*
					Infinitival Clauses				
30 months	0.88	1.09	0.50	0 – 3	0.27	0.60	0.00	0 – 2	0.74*
42 months	2.19	2.29	1.50	0 – 7	1.92	1.77	1.50	0 – 7	0.14
54 months	1.88	1.31	2.00	0 – 5	1.72	1.46	2.00	0 – 5	0.11
66 months	3.87	2.67	4.00	0 – 9	4.42	3.36	4.50	0 – 13	0.18
					Complement Clauses				
30 months	0.06	0.25	0.00	0 – 1	0.00	0.00	0.00	0 – 0	0.39
42 months	1.88	2.09	1.00	0 – 6	1.08	1.90	0.00	0 – 8	0.41
54 months	3.56	2.34	3.50	0 – 9	2.84	2.90	2.00	0 – 11	0.27
66 months	2.40	2.64	2.00	0 – 10	1.33	1.52	1.00	0 – 6	0.53
					Combined Clauses				
30 months	0.31	0.60	0.00	0 – 2	0.23	0.51	0.00	0 – 2	0.15
42 months	1.25	1.57	1.00	0 – 5	0.69	1.09	0.00	0 – 5	0.43
54 months	2.63	2.22	2.50	0 – 7	1.88	1.72	1.00	0 – 6	0.39
66 months	6.33	3.85	5.00	2 – 13	5.92	4.44	5.50	0 – 17	0.10
					Relative Clauses				
30 months	0.25	0.45	0.00	0 – 1	0.00	0.00	0.00	0 – 0	0.91*
42 months	0.13	0.34	0.00	0 – 1	0.23	0.59	0.00	0 – 2	0.20
54 months	1.19	1.38	1.00	0 – 4	1.20	1.29	1.00	0 – 4	0.01
66 months	3.40	1.99	4.00	0 – 8	1.88	2.01	1.00	0 – 8	0.76*

Note. *$p > .05$. NL group, n = 16 and LT group, n = 26 with the exception of the NL group at 66 months, (n = 15), the LT group at 54 months, (n = 25), and the LT group at 66 months, (n = 24).

Complex syntax by type. In addition to our cumulative measure of emergence by complex syntax category, we calculated the proportion of participants in each group who used each complex syntax type at each time point non-cumulatively (see Table 19). Marked infinitives were produced by the greatest proportion of participants at all time points except at 66 months when subordinate conjunction clauses were produced by the greatest proportion of participants. Coordinate conjunction clauses, subordinate conjunction clauses, marked infinitives, full propositional complements, and other relative clauses were used by more than .50 of the participants in both groups at least at one time point. In addition, more than .50 of the NL participants produced WH finite complements, subject relative clauses, and participle clauses at one time point, whereas the LT participants did not do so at any time point. The greatest difference in percent of participants from each group who used a given type were for marked infinitives at 30 months, coordinate conjunction clauses at 42 months, WH nonfinite clauses at 54 months, and subject relative clauses at 66 months.

Table 19

Proportion of Participants Who Produced at Least One Exemplar of Each Type at Each Time Point

NL Group	CC	SC	SI	UIC	FPC	WFC	WNFC	SRC	RC	NRC	PC	OTHER
30 months	0.00	0.25	0.50	0.06	0.06	0.00	0.00	0.00	0.06	0.19	0.19	0.00
42 months	0.38	0.38	0.69	0.06	0.56	0.19	0.19	0.00	0.06	0.06	0.19	0.13
54 months	0.56	0.63	0.88	0.13	0.81	0.69	0.38	0.25	0.25	0.31	0.25	0.06
66 months	0.67	1.00	0.93	0.13	0.53	0.33	0.27	0.67	0.73	0.40	0.53	0.27
LT Group												
30 months	0.15	0.04	0.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
42 months	0.12	0.35	0.81	0.15	0.35	0.15	0.04	0.08	0.00	0.08	0.12	0.04
54 months	0.36	0.64	0.72	0.24	0.64	0.48	0.08	0.16	0.32	0.32	0.20	0.04
66 months	0.67	0.88	0.79	0.25	0.54	0.21	0.13	0.08	0.58	0.29	0.46	0.29

Note. CC = coordinate conjunction clauses; SC = subordinate conjunction clause; SI = marked infinitive; UIC = unmarked infinitive; FPC = full propositional complement; WFC = WH finite complement; WNFC = WH nonfinite complement; SRC = subject relative clause; RC = other relative clause; NRC = nominal relative clause.

Complex syntax categories by MLU. MLU is a measure of utterance length. Utterance length increases as children produce a greater number of bound morphemes and a greater number of clauses within an utterance (e.g., Paul, 1981; see also Olson & Masur 2019). We explored emergence of complex syntax categories relative to MLU range (see Table 20).

Various complex syntax types may lengthen utterances at various MLU ranges. Furthermore, late talkers may production different complex syntax types as compared to same-aged peers at different MLU ranges.

First, we sorted all language samples by MLU in increments of 0.50 from 0.00 – 6.99 and then in increments of 1.0 from 7.00 – 7.99 and 8.00 – 8.99. The last two increments included very few transcripts. Second, we calculated the proportion of transcripts in each group that included at least one exemplar from each complex syntax category within each MLU increment. Each complex syntax category was initially represented in transcripts from each group when MLU reached 3.50 – 3.99. At least 80% of the transcripts from each group included exemplars from each complex syntax category when MLU reached 6.00 – 6.49 in each group.

In general infinitival clauses were the most frequently occurring complex syntax category at MLU ranges lower than 4.00. Combined clauses and to some extent complement clauses became more prominent in the LT group at a lower MLU range than in the NL group. In contrast, relative clauses were more prominent in the NL group at a lower MLU range than in the LT group. The LT group generally had a higher mean age at a lower MLU. Thus, the participants in the LT group may have had greater experience with complex syntax vocabulary at a lower MLU as compared to the NL group. After an MLU of 5.00 differences were less evident for the number of transcripts from each group that contained a give complex syntax category. Our findings suggest that the LT group produced slightly more complex syntax at a slightly lower MLU, but between-group differences quickly disappeared.

Table 20

Proportion of Transcripts that Included Complex Syntax Categories by Group

MLU	Mean Age NL	Mean Age LT	Number of transcripts NL	Number of transcripts LT	Infinitival NL	Infinitival LT	Complement NL	Complement LT	Combined NL	Combined LT	Relative NL	Relative LT
1.00 – 1.49	--	30	0	7	--	0.00	--	0.00	--	0.00	--	0.00
1.50 – 1.99	30	30	1	4	0.00	0.00	0.00	0.00	0.00	0.25	0.00	0.00
2.00 – 2.49	31	31	1	8	0.00	0.13	0.00	0.00	0.00	0.13	0.00	0.00
2.50 – 2.99	34	30	4	7	0.50	0.29	0.00	0.00	0.00	0.43	0.25	0.00
3.00 – 3.49	36	43	6	6	0.67	0.83	0.00	0.50	0.33	0.00	0.00	0.00
3.50 – 3.99	40	46	12	14	0.58	0.86	0.17	0.43	0.42	0.50	0.08	0.36
4.00 – 4.49	43	48	6	12	0.50	0.75	0.83	0.58	0.67	0.58	0.50	0.25
4.50 – 4.99	50	52	12	14	0.75	0.57	0.92	0.50	0.83	0.86	0.50	0.29
5.00 – 5.49	56	58	7	7	1.00	0.86	0.71	0.71	0.86	0.86	0.71	0.86
5.50 – 5.99	54	62	3	6	1.00	0.67	1.00	0.50	1.00	1.00	0.33	0.83
6.00 – 6.49	64	66	5	5	1.00	1.00	1.00	0.80	1.00	0.80	1.00	0.80
6.50 – 6.99	67	66	2	6	1.00	1.00	0.50	0.67	1.00	1.00	1.00	0.83
7.00 – 7.99	67	66	2	3	1.00	1.00	1.00	0.67	1.00	1.00	1.00	0.67
8.00 – 8.99	66	67	2	2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Complex syntax vocabulary. We completed a series of Mann-Whitney *U* tests to assess between-group differences for complement clause verb diversity and subordinate conjunction diversity at each time point (see Table 21). We chose to employ non-parametric statistics for complex syntax vocabulary variables because these variables were not consistently normally distributed across time points based on the Shapiro-Wilk test (*p*s > .05).

Table 21

Complex Syntax Vocabulary by Time and Group

	NL Group				LT Group				
Time	Mean	SD	Median	Range	Mean	SD	Median	Range	*d*
				Subordinate Conjunction Diversity					
30 months	0.19	0.40	0.00	0 – 1	0.00	0.00	0.00	0 – 0	0.78*
42 months	0.38	0.72	0.00	0 – 2	0.38	0.70	0.00	0 – 3	0.00
54 months	1.25	1.29	1.00	0 – 4	1.00	1.04	1.00	0 – 4	0.22
66 months	2.33	1.35	2.00	1 – 6	1.92	1.18	2.00	0 – 4	0.33
				Complement Clause Verb Diversity with Complement Clauses					
30 months	0.19	0.54	0.00	0 – 2	0.00	0.00	0.00	0 – 0	0.57
42 months	1.13	1.20	1.00	0 – 4	0.58	0.81	0.00	0 – 3	0.56
54 months	2.31	1.58	2.00	0 – 5	1.60	1.32	2.00	0 – 5	0.51
66 months	1.40	0.99	2.00	0 – 3	1.13	1.23	1.00	0 – 5	0.24
				Complement Clause Verb Diversity with Marked Infinitives					
30 months	0.56	0.63	0.00	0 – 1	0.12	0.33	0.00	0 – 2	0.94*
42 months	0.88	0.96	1.00	0 – 2	0.96	0.77	1.00	0 – 3	0.09
54 months	1.19	0.66	1.00	0 – 5	0.92	1.12	1.00	0 – 2	0.28
66 months	2.13	1.81	2.00	0 – 4	2.04	1.40	2.00	0 – 7	0.06

Note. *$p < .05$; NL Group, n = 16 and LT Group, n = 26 with the exception of NL group at 66 months, (n = 15); LT group at 54 months, (n = 25); LT group at 66 months, (n = 24).

There was a significant between-group difference for subordinate conjunction diversity at 30 months ($U = 169$, $n_{NL} = 16$, $n_{LT} = 26$, $p < .05$). The NL group produced a greater subordinate conjunction clause diversity as compared to the LT group. *Because* (38%), *when* (21%), *if* (19%), and *so* (13%) comprised 91% of the subordinate conjunctions summed across time and group. All between-group differences for complement clause verb diversity with complement clauses were non-significant (*p*s > .05). *Think* (32%) and *know* (28%) comprised 60% of the complement clause verbs with complement clauses summed across time and group. There was a significant between-group difference for complement clause verb diversity with marked infinitives at 30 months ($U = 140.5$, $n_{NL} = 16$, $n_{LT} = 26$, $p < .05$). The NL group produced a greater complement clause verb diversity with marked infinitives as compared to the LT group.

Get (20%), *have* (18%), *need* (14%), and *want* (12%) comprised 64% of the complement clause verbs with marked infinitives summed across time and group.

Late bloomers and late talkers with continuing delay. To further examine between-group differences we divided the LT group into two groups at 66 months based on the NL group mean for MLU: (a) late bloomers were defined as participants who performed no more than one standard deviation below the NL group mean (n = 21) and (b) late talkers with continuing delay were defined as participants who performed more than one standard deviation below the NL group mean (n = 3). We calculated between-groups effect sizes for MLU, complex syntax proportion, and complex syntax type diversity for each group of late talkers as compared to the NL group using the means and standard deviations from each group (see Table 22). Effect sizes suggest that late bloomers differ widely from late talkers with continuing delay and more closely resemble the NL group than do late talkers with continuing delay.

Table 22

Effect Sizes Between Late Talker Groups and NL Group at 66 Months

	Late Talker Group	
Variable	Late Bloomers (n = 21)	Late Talkers with Continuing Delay (n = 3)
MLU	0.21	1.53
Complex Syntax Proportion	0.10	1.81
Complex Syntax Type Diversity	0.51	2.28

Error analyses. Following the framework proposed by Schuele and Dykes (2005) we completed an error analysis to explore the grammatical accuracy of complex syntax tokens produced by participants in each group.

Proportion of total complex syntax tokens in error. Using a Mann Whitney U Test, there were no between-group difference for number of errors per complex syntax token at any time point (ps > .05, one-tailed; see Table 23). As would be expected, the mean number of errors per complex syntax token across groups was greatest at 30 months. The errors produced by the NL group dramatically decreased by 42 months. In contrast, the errors produced by the LT group did not dramatically decrease until 54 months. Participants who did not produce any complex syntax tokens at a given time point were not included in this analysis.

Table 23

Number of Errors Per Complex Syntax Token by Time and Group

	NL Group			LT Group			
Child Age	Mean	SD	Median	Mean	SD	Median	*d*
30 months	0.43	0.46	0.40	0.34	0.44	0.13	0.20
42 months	0.12	0.15	0.10	0.30	0.37	0.13	0.57
54 months	0.10	0.10	0.09	0.09	0.20	0.00	0.06
66 months	0.06	0.07	0.04	0.09	0.15	0.04	0.24

Note. 30 months NL group, n = 13 and LT group, n = 8; 42 months NL group, n = 13, LT group, n = 24, 54 months, NL group, n = 15, LT group, n = 25, 66 months, NL group, n = 15, LT, n = 23.

Proportion of errored productions by complex syntax category. At 30 months the most common error in the NL group was in the production of infinitival clauses and the most common error in the LT group was in the production of infinitival clauses and combined clauses (see Table 24). At 42 months the most common error in the NL group was in the production of relative clauses and the most common error in the LT group was again in the production of infinitival clauses.

Table 24

Proportion of Errored Productions Relative to Total Tokens in Each Complex Syntax Category

	Complex Syntax Category							
	NL	LT	NL	LT	NL	LT	NL	LT
Time	Infinitival		Complement		Combined		Relative	
30 months	.58	.40	.00	NA	.25	.40	.50	NA
42 months	.15	.38	.12	.11	.32	.21	.50	.13
54 months	.09	.13	.00	.03	.12	.05	.22	.11
66 months	.11	.04	.02	.02	.07	.14	.06	.06

Note. NA = no exemplars produced by LT group participants at 30 months

At 54 months the most common error in the NL group was in the production of relative clauses and the most common error in the LT group was in the production of infinitival clauses with combined clauses and relative clauses trailing close behind. At 66 months the most common error in the NL group was in the production of infinitival clauses and the most common error in the LT group was in the production of combined clauses. Participants who did not produce any complex syntax tokens in a given complex syntax category at a given time point were not included in this analysis.

Chapter IV

DISCUSSION

In this study we characterized the complex syntax development in a group of late talkers (LT Group) as compared to children with typical language (NL group) from 30 months to 66 months. We measured complex syntax development longitudinally in each group using a comprehensive set of variables. Then, following the framework proposed by Schuele and Dykes (2005), we compared the groups on measures of complex syntax quantity and diversity, emergence of complex syntax categories, and grammatical accuracy. Additionally, we compared the groups on measures of complex syntax vocabulary.

Grammatical development is a common area of long-term delay in children with late language emergence (e.g., Rice et al., 2008). Tense morphemes and complex syntax production are particular areas of concern for late talkers (e.g., Paul & Alforde, 1993; Rescorla & Turner, 2015). Longitudinal studies of complex syntax development in late talkers are important to fully characterize grammatical skills in this population. This study is the first to comprehensively evaluate complex syntax development in late talkers as compared to same-aged peers.

Longitudinal Development

As expected, variables measuring complex syntax quantity and diversity, with a few exceptions, increased between each adjacent 12-month interval in each group. Furthermore, complex syntax variables were generally correlated at all time points. Thus, the variables did not capture fully unique aspects of complex syntax development. Complex syntax proportion was consistently sensitive to change in each group between each adjacent time point. Despite substantial individual variability and differences in coding, the mean complex syntax proportion in the NL group was similar to proportion metrics reported in previous studies of children with typical language (see Table 25).

Diessel (2004) reported that between 24 months and 36 months the mean proportion of utterances with at least two verbs was .04, similar to the mean complex syntax proportion of .03 in the NL group at 30 months. Diessel (2004) reported that between 36 months and 48 months the mean proportion of utterances with at least two verbs was .12 similar to the mean complex syntax proportion of .10 in the NL group at 42 months. Similarities between Diessel (2004) and the current study are notable because Diessel measured complex syntax development in a much smaller group of children (n = 5), but within longer samples as compared to the current study. Jackson and Roberts (2001) reported a complex syntax proportion similar to the NL group and the LT group with a mean complex syntax proportion of .06 at 36 months and .12 at 48 months. The mean complex syntax proportion of children with SLI in previous studies (Delgado, Morton, & Schuele, 2018; Schuele & Dykes, 2005) lags behind the mean complex syntax proportion in the LT group.

Table 25

Cross-Study Comparison of Metrics of Complex Syntax Proportion

Citation	Age in Years	Complex syntax proportion (NL Group)	Complex syntax proportion (LT Group/SLI)
Current study	2;6	.03	.01
	3;6	.10	.08
	4;6	.17	.14
	5;6	.25	.22
Diessel, 2004	< 2;0	.00	NA
	2 – 3	.04	
	3 – 4	.12	
Jackson & Roberts, 2001	3	.06	NA
	4	.11	
Schuele & Dykes, 2005	3;3 – 3;9	NA	.008 – .02[a]
	4;0 – 4;8		.04 – .18[a]
	5;3 – 5;9		.15 – .18[a]
Delgado, Morton, & Schuele, 2018	5	.20	.13[a]

Note. [a]The data from these studies represents children with SLI.

Unlike complex syntax proportion, number of complex syntax tokens is unconstrained by proportion (see Craig & Washington, 1994). This variable captures variation that is due to multiple dependent clauses within a single utterance. The number of complex syntax tokens patterned after growth in complex syntax proportion in that it was consistently sensitive to change between each adjacent time point in each group. Like number of complex syntax tokens, complex syntax density is sensitive to multiple embeddings. In contrast however, complex syntax density was not sensitive to change between the first two time points in either group. It is important to note that this variable was only calculated for participants who produced at least one complex syntax token at a given time point and thus does not represent all participants at each time point. This variable increased more gradually as compared to other complex syntax variables.

Complex syntax type diversity captures the variety of complex syntax types that a child produces. Measures that have captured complex syntax type diversity previously have been found to be sensitive measures of preschool language development (e.g., Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010; Jackson & Roberts, 2001; Scarborough, 1990). Complex syntax type diversity was sensitive to change between 30 months and 42 months and between 42 months and 54 months in each group. The lack of sensitivity to change between 54 months and 66 months in each group may be constrained by the lack of opportunities to produce a variety of structures within a 50-utterance analysis set.

As expected, MLU and NDW were sensitive to change in each group between each adjacent time point. Unlike all complex syntax variables, MLU and NDW were consistently normally distributed at each time point. Complex syntax variables, except for complex syntax density, were consistently intercorrelated with each other and with MLU and NDW. As utterance length increased, children produced a greater diversity of words and complex syntax types as well as more complex syntax tokens.

Despite variability in the age of acquisition as well as differences in complex syntax coding, the order of acquisition in which greater than 50% of the participants in each group produced exemplars from each complex syntax category was similar to previous studies (Bloom, Lahey, Hood, Lifter, & Fiess, 1980; Limber, 1973). Infinitival clauses emerged followed by complement clauses which were in turn followed by relative clauses in each group. The emergence of combined clauses generally coincided with the emergence of complement clauses. Similar to previous studies, the diversity of complex syntax increased greatly in each group between an MLU of 3.50 and 5.49 (Schuele & Dykes, 2005; Tyack & Gottsleben, 1986).

In summary, the quantity and diversity of complex syntax increased from 30 to 66 months in late talkers and same-aged NL peers. Complex syntax categories followed a similar albeit somewhat delayed developmental trajectory in late talkers as compared to same-aged peers. The proportion of participants in each group who had produced at least one exemplar from each complex syntax category was roughly equivalent by 66 months.

Complex syntax increased alongside established language sample measures – MLU and NDW. Specific properties of each variable are presented in Table 26.

Table 26

Properties of Dependent Variables

Variable	Sensitive to change at annual intervals	Normally distributed	Correlation with other complex syntax variables	Number of between-group differences (max = 4)
MLU	Yes	Yes	Yes	1
NDW	Yes	Yes	Yes	2
Complex syntax proportion	Yes	Inconsistent	Yes	1
Number of complex syntax tokens	Yes	Inconsistent	Yes	1
Complex syntax density	Inconsistent	Inconsistent	Inconsistent	0
Complex syntax type diversity	Inconsistent	Inconsistent	Yes	3

Between-group Differences

At 30 months between-group differences were widespread. The LT group produced a lower complex syntax quantity and diversity of complex syntax types, fewer infinitival clauses, fewer relative clauses, and less subordinate conjunction diversity as compared to the NL group. Between-group differences disappeared by 42 months. Measures that captured individual differences showed a slightly different story. The proportion of the LT group participants who produced each complex syntax category lagged behind the NL group participants from 30 months to 54 months. However, by 66 months the proportion of children who produced each complex syntax category were essentially equivalent except for a small lag in relative clauses. Likewise by 66 months, the composition of complex syntax category tokens was generally similar across groups, but relative clauses made up a smaller proportion of the complex syntax tokens produced by the LT group participants as compared to the NL group participants. In general, the patterns of between-group differences for complex syntax quantity, complex syntax vocabulary and the emergence and composition of complex syntax categories suggest that the LT group demonstrated an initial delay as compared to the NL group followed by a period of catch up. Although most of the participants achieved performance within the average range, a few did not.

In contrast, between-group differences for complex syntax type diversity were more long-lasting, evident not only at 30 months, but also at 54 months and 66 months. Individual differences showed that more participants in the LT group performed below the average range at 66 months for complex syntax type diversity than for complex syntax proportion, MLU, and NDW. Relatedly, the IPSyn Total Score, which accounts for specific grammatical structures previously has been found to be more sensitive to between-group differences as compared to MLU at five years of age (Rescorla & Turner, 2015).

The fairly consistent between-group differences in complex syntax type diversity supported an exploratory analysis of between-group differences in complex syntax categories.

As previously stated, at 30 months the LT group participants produced fewer infinitival clauses and relative clauses as compared to the NL group participants. At 66 months the LT group produced fewer relative clauses as compared to the NL group. Furthermore, at 66 months 67% of the NL group, but only 8% of the LT group, produced subject relative clauses.

As mentioned in the introduction, infinitival clauses, complement clauses, and subordinate conjunction clauses are reliant on the production of specific lexical items; whereas, relative clauses are not. Longer-lasting between-group differences for relative clauses suggests that the syntactic aspects of complex syntax development may be particularly challenging for late talkers even after they catch up on variables that tap the lexical aspects of complex syntax development. Prior studies have found morphosyntax to be particularly challenging for late talkers as compared to vocabulary skills (e.g., Rescorla et al., 1997; Rescorla, Mirak, & Singh, 2000; Rice et al., 2008).

Illusory recovery. Initial between-group differences for complex syntax type diversity and relative clauses that resurfaced are potentially suggestive of illusory recovery (Scarborough & Dobrich, 1990; Rescorla & Turner, 2015). Illusory recovery is a phenomenon in which between-group differences disappear for a time and then resurface later. Illusory recovery has been demonstrated previously at 42 months when children with typical language plateau on grammatical measures followed by an acceleration at 60 months (Scarborough & Dobrich, 1990). Interestingly, this timeline is similar to what was found in the current study. Illusory recovery was further evidenced in the current study by the high proportion of LT group participants who performed within the average range for complex syntax type diversity at 42 months with a subsequent decrease in the proportion of participants who performed within the average range at 54 months and 66 months.

There are two potential reasons for illusory recovery in complex syntax development. First, illusory recovery may occur because a task is too easy and thus does not differentiate groups. The conversation-based interaction at 66 months as compared to predominantly play-based interactions at prior time points was likely more linguistically challenging for the participants and thus, may have led to a reemergence of between-group differences.

Second, we speculate that illusory recovery may be explained at least in part by the initial reliance in complex syntax development on high frequency lexical items produced in formulaic phrases such as *I think* (Diessel, 2004). Although late talkers demonstrate initial expressive vocabulary delays, they may catch up more quickly in producing formulaic utterances as compared to later, novel utterances that require greater syntactic skill. As children with typical language move on to these later utterances, late talkers who are relying on formulaic utterances may appear to catch up. However, as children with typical language expand their vertical development to include more complex syntax types that do not rely on frequent lexical items, late talkers may be outpaced again.

Relation of Current Findings to the Rescorla Corpus

Between-group differences for complex syntax variables, MLU, and NDW do not appear to be as pervasive or long-lasting in the current study as compared to between-group differences found in the Rescorla sample (e.g., Rescorla, Dahlsgaard, & Roberts, 2000). Consideration of the participant characteristics in our study that led to greater catch up as compared to the Rescorla sample may illuminate future areas of investigation and potential malleable factors for intervention.

Rescorla and colleagues completed a longitudinal analysis of late talkers as compared to same-aged peers for children from 24 months to 17 years. Intake age ranged from 24 to 31 months (e.g., Rescorla et al., 1997; Rescorla, 2009). Late talkers in this corpus produced a lower MLU and IPSyn total score at three and four years as compared to same-aged peers as well as fewer cognitive state words, the majority of which were complement clause verbs, from three to five years of age (Rescorla et al., 2000a; Rescorla & Turner, 2015). Differences in findings between the Rescorla corpus and the current study may be driven by a number of sample characteristics that illuminate the following factors to consider in future studies – intake age, late talker eligibility criterion, individual differences in complex syntax development, and language sample length.

First, all participants in the Ellis Weismer corpus entered the study at 24 months; whereas, the children in the Rescorla corpus entered the study between 24 months and 31 months. Children who are identified as late talkers when they are older have had more time to catch up but have not and thus, may be less likely to catch up as compared to children who are identified earlier (e.g., Rescorla & Schwartz, 1990).

Second, the Rescorla corpus appears to include a greater number of late talkers who did not catch up as compared to our sample. Rescorla and Turner (2015) separated late talkers into two groups: (a) children with continuing language delays and (b) late bloomers. Late bloomers were defined as children who had an MLU within 1 SD of the group mean of the same-aged peers. At five years, children with continuing delay represented 19% of their sample. A comparable 15% of the current study sample met the criterion for late bloomers at four-and-a-half years, but at five-and-a-half years only 8% of our sample met the criterion. At five years of age Rescorla and Turner (2015) found between-group differences on the IPSyn Sentence Structure subscale, which taps several types of complex syntax, for late talkers with continuing delay as compared to same-aged peers, but not for late bloomers as compared to same-aged peers. Thus, having a greater proportion of participants with continuing language delays likely led to longer-lasting between-group differences in their corpus as compared to the current study sample. Grammatical measures appear to be particularly important to determine catch up or lack thereof in late talkers (e.g., Rice et al., 2008).

Third, Rescorla used a 100-utterance analysis set as compared to our 50-utterance analysis set. Our shorter analysis set likely led to greater variability in the complex syntax variables and thereby may have obscured between-group differences. When Moyle and colleagues (2007) completed an analysis with a subset of children from the Ellis Weismer corpus, they found between-group differences for MLU and NDW in 50-utterance language transcripts only at 30 months and 42 months but not at 54 months and 66 months as compared to consistent between-group differences on norm-referenced assessments. They proposed that the lack of between-group differences at 54 months and 66 months were due to the substantial variability for MLU and NDW within each group. Gavin and Giles (1996) reported that preschoolers need samples of 100 utterances to attain reliability of at least .80 for MLU. Likewise, we expect that samples of at least 100 utterances are needed to obtain stable estimates of complex syntax variables. Longer samples are likely to be particularly important for sampling complex syntax types and complex syntax vocabulary that are relatively rare as well as for establishing sufficient power to detect between-group differences (Marinellie, 2004).

Fourth, Lee and Rescorla (2008) demonstrated consistent between-group differences in the percent of utterances containing cognitive state words from three to five years. The current study did not reveal between-group differences in the diversity of complement clause verbs or in the frequency of complement clauses at any time point. There are three potential reasons for the differences in findings between studies. First, Lee and Rescorla (2008) included cognitive state words that were not complement clause verbs. Second, the short samples in the current study included only a small number complement clause verbs and thus, may have not allowed us to detect effects. Third, the smaller between-group differences in the current study may simply reflect that the LT group had greater facility with complement clause verbs as compared to Rescorla's sample. This latter point supports the possibility of less persistent deficits in the lexical aspects of complex syntax development in some late talkers. Future research may explore whether children who demonstrate deficits in the lexical and syntactic aspects of complex syntax development have greater overall language delays as compared to children who only demonstrate deficits in the syntactic aspects of complex syntax development.

Limitations

Prior to discussing implications and future directions, we acknowledge three limitations in the current study. First, as previously mentioned the language samples were brief and thus, the complex syntax variables were likely subject to variability. Second, due to the brevity of the language samples in some instances we drew utterances from a secondary transcript to complete the analysis set. This decision may have affected results particularly at 54 months in which we drew utterances from play-based and interview-based samples. There is some evidence that complex syntax types vary by context (Klein, Moses, & Jean-Baptiste, 2010). However, we drew utterances from the secondary transcripts for a similar proportion of participants across groups at 54 months and the between-group difference for the average number of utterances drawn from the secondary transcript was non-significant at that time point (U = 169, n_{NL} = 16, n_{LT} = 25, p = .822, d = 0.13). Third, limited information was available to characterize the participant sample. Receptive language, SES level, and family history of late language emergence are each relevant predictors of language outcomes in late talkers and thus, are needed to more fully explain between-group differences (Fisher, 2017; Zubrick et al., 2007).

Clinical Implications

As demonstrated in past studies, complex syntax is an age-appropriate language skill to be targeted in the preschool years (e.g., Barako Arndt & Schuele, 2013). Our findings from the NL group provide a set of general benchmarks for complex syntax development (see Table 27). It is important to note that based on the subset of children for whom we have norm-referenced standardized testing the NL group demonstrates strong language skills and thus these benchmarks may overestimate the population average. Previous studies have suggested that adult input should be slightly more advanced than children’s level of production (e.g., Nelson, Denninger, Bonvillian, Kaplan, & Baker, 1984; see also Rowe et al., 2012). Thus, alongside the developmental benchmarks we provide preliminary suggestions for enhancing input to preschoolers.

Table 27

Complex Syntax Developmental Benchmarks and Suggestions for Enhancing Input to Preschoolers

Age	Complex Syntax Proportion	Complex Syntax Type Diversity	Typical Benchmarks	Suggestions for Enhancing Input
2½	0.03	1.31	Complex syntax is mainly comprised of infinitival clauses.	Model complement clause verbs in play.
3½	0.10	2.75	Children increase the frequency with which they produce complement clauses. The majority of complex syntax is produced with complement clause verbs in infinitival clauses and complement clauses. Watch out for children who are not producing infinitival clauses.	Model complement clause verbs and conjunctions while reading books, explaining how to make arts and crafts, and playing with toys (see Schuele & Dykes, 2005; Teller & Schuele, 2019). Ask and model answers to *why* questions to increase subordinate conjunction clauses in the input (Diessel, 2004). Model indirect why questions (e.g., *I wonder why . . .*).
4½	0.17	5.06	Children produce a wide variety of complex syntax types in play. Complex syntax types will vary by activity context. Children may produce more complement clauses in play and more relative clauses in conversation.	Model diverse complement clause verbs to add diversity in play. Model relative clauses and subordinate conjunctions in conversation to get children ready for kindergarten.
5½	0.25	6.20	Children produce a wide variety of complex syntax types in conversation.	Model relative clauses and diverse subordinate conjunctions in conversation.

Factors in Individual Differences

Evidence for between-group differences in complex syntax development supports future study of the mechanisms of difference. Although the current study is exploratory and thus not positioned to make strong arguments regarding mechanisms driving individual differences, we suggest three endogenous and one environmental factor that likely contribute to individual differences in complex syntax development: (a) family history and genetics, (b) language processing speed, (c) receptive language abilities, and (d) adult input.

Endogenous and environmental factors in individual differences. First, the literature supports a genetic or familial component for individual differences in late talkers (Reilly et al., 2010; Zubrick et al., 2007; cf. Fisher, 2017). Children who have a family history of late language emergence are at greater risk for late language emergence. Second, individual differences in the language processing speed predict vocabulary and syntactic growth (e.g., Peter, Durrant, Jessop, Bidgood, Pine, & Rowland, 2019). Late talkers who have more efficient word recognition skills have more accelerated vocabulary growth as compared to late talkers with less efficient word recognition skills (Fernald & Marchman, 2012).

Third, receptive language abilities are predictive of late talker outcome (Fisher, 2017). Children with delays in receptive and expressive language have poorer outcomes as compared to children with expressive delays alone. For example, Thal, Marchman, and Tomblin (2013) found that a similar proportion of children with delays in expressive language produced complex sentences at 36 months as did the children with typical language. In contrast, a smaller proportion of children with delayed production and comprehension produced complex sentences. In the Ellis Weismer corpus, receptive language was left free to vary. However, the subset of children for whom we have norm-referenced assessments at 66 months suggests that receptive language may be similar between groups. Thus, individual differences in receptive language may have been a relatively small contributor to individual differences in our study.

Fourth, individual differences may be related to frequency of complex syntax in the adult input. Frequency of complex syntax in adult input is related to the age of acquisition of complex syntax vocabulary. For example, *think* and *know* are the most commonly occurring complement clauses verbs in maternal input and are the earliest and most frequent complement clause verbs in child production (e.g., Diessel, 2004). Although late talkers likely require more complex syntax input for development (see Riches, Tomasello, & Conti-Ramsden, 2005), late talkers receive less complex syntax input as compared to children with typical language (d = 1.08; Schuele et al., in preparation). If late talkers produce less complex syntax, adults may in turn respond with less complex syntax potentially reducing learning opportunities for complex syntax vocabulary and structures (see Camarata & Yoder, 2002).

Future Directions

This study demonstrates that individual differences in complex syntax development are an important area of study in child language development. Comparisons between our findings and previous studies illuminate considerations for future research. First, researchers should make systematic choices about participant characteristics in order to control for factors that might reduce the interpretability of study results. Important characteristics include a consistent intake age and eligibility criteria that specifies expressive and receptive language level. When completing prospective studies, researchers should use an adequate sample size preferably with population-based samples that represents a range of SES levels (Reilly et al., 2010; Zubrick et al., 2007). Carefully designed large-scale studies are needed to describe similarities and differences should they exist between children with typical language, late bloomers, late talkers with continuing language delays, and children with SLI. Such studies may uncover different mechanisms for language delays in children who resolve as compared to those who do not resolve (see Bavin & Bretherton, 2013, p. 17). Large-scale studies will allow for exploration of complex syntax development using more sophisticated statistical techniques such as growth curve modeling. Growth curve modeling allows researchers to account for individual differences in initial performance as well as in rate of growth over time and thus to measure complex syntax development more precisely. Measurement at 6-month intervals will likely further illuminate small changes in growth over time that were not captured in the current study.

Second, researchers should investigate how to improve the temporal stability of complex syntax variables to improve replicability and interpretability of findings. We recommend that complex syntax variables be investigated in language sample utterances with at least 100 utterances with preference to language samples of at least 200 utterances (Rice et al., 2008). The stability of the complex syntax variables may be improved through careful control of stimuli, conversational topic, and elicitation techniques (Heilmann, DeBrock, & Riley-Tillman, 2013). A structured conversational protocol, such as that found in Hadley (1998), has been found to elicit a variety of complex syntax types (Barako Arndt & Schuele, 2013). A productivity criterion for complex syntax type diversity may allow for greater stability and better observation of subtle changes in the development within complex syntax types (e.g., Eisenberg, 1997; Hadley & Short, 2005). Furthermore, the combination of complex syntax type diversity and MLU – a well-established, normally-distributed corollary to other complex syntax measures – may be explored to determine whether a more limited set of variables is sufficient to capture complex syntax development. The relation between findings from complex syntax variables and other assessment measures that are particularly sensitive to language impairment, such as the Test of Early Grammatical Impairment (Rice & Wexler, 2001) and sentence imitation, should be explored.

Third, researchers should further explore errors in complex syntax production in late talkers as compared to same-aged peers following the framework proposed by Schuele and Dykes (2005). Tense-marking morphemes (e.g., - 3s and past tense -ed) as well as errors in complex syntax markers (e.g., infinitival TO; relative clauses markers) have been particularly useful in identifying children with SLI (e.g., Barako Arndt & Schuele, 2012; Goffman & Leonard, 2000; Schuele & Tolbert, 2001). Thus, between-group differences in complex syntax errors may be useful as a marker of catch up or lack thereof in late talkers. Fourth, researchers should continue to explore mechanisms of individual differences in complex syntax development. Exploring the longitudinal relation between adult input and child complex syntax production will likely further our understanding of complex syntax development.

Conclusion

We demonstrated that a group of late talkers initially lagged behind typical children in the production of complex syntax. Many of the children subsequently caught up to same-aged peers with the greatest growth between 30 months and 42 months. By 66 months most children in both groups produced exemplars from each of the major complex syntax categories. Complex syntax type diversity and relative clause frequency demonstrated the greatest evidence of between-group differences across time points. This study links previous studies of complex syntax development to future studies that will investigate mechanisms of individual differences in complex syntax development and refine complex syntax measurement. Our findings behoove researchers and clinicians (a) to investigate factors that lead to variability in complex syntax development and (b) to investigate complex syntax as an important preschool intervention target to prepare for children with linguistic vulnerabilities for school-age language demands.

References

Barako Arndt, K., & Schuele, C. M. (2012). Production of infinitival complements by children with specific language impairment. *Clinical Linguistics & Phonetics, 26*, 1-17. doi:10.3109/02699206.2011.584137

Barako Arndt, K., & 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*, 125-129. doi:D - EJ1013676

Bavin, E. L., & Bretherton, L. (2013). The early language in Victoria study: Late Talkers, Predictors, & Outcomes. In L. A. Rescorla, & P. S. Dale (Eds.), *Late talkers: Language development, intervention, and outcomes*. Brooks Publishing Company.

Bloom, L., Lahey, M., Hood, L., Lifter, K., & Fiess, K. (1980). Complex sentences: Acquisition of syntactic connectives and the semantic relations they encode. *Journal of Child Language, 7*, 235-261. doi:10.1017/S0305000900002610

Camarata, S., & Yoder, P. (2002). Language transactions during development and intervention: Theoretical implications for developmental neuroscience. *International Journal of Developmental Neuroscience, 20*, 459-465. doi:10.1016/S0736-5748(02)00044-8

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Craig, H. K., & Washington, J. A. (1994). The complex syntax skills of poor, urban, African-American preschoolers at school entry. *Language, Speech, and Hearing Services in Schools, 25*, 181-190. doi:10.1044/0161-1461.2503.181

Delgado, A., Morton, I., & Schuele, C. M. (2018, July). Complex syntax production in 5-year-old children with specific language impairment (SLI). Poster presented at the Poster Session for the School for Science and Math at Vanderbilt, Nashville, TN.

Diessel, H. (2004). *The acquisition of complex sentences*. Cambridge University Press.

Dollaghan, C. (2013). Late talker as a clinical category: A critical evaluation. In L. A. Rescorla, & P. S. Dale (Eds.), *Late talkers: Language development, intervention, and outcomes* (pp. 91-112). Brooks Publishing Company.

Dunn Davison, M., Schuele, C. M., Fisher, J., Dickinson, D. & Combs, S. (2020). *An initial exploration of complex syntax input in preschool classrooms of children from low-SES families.* Unpublished manuscript.

Eisenberg, S. (1997). Investigating children's language: A comparison of conversational sampling and elicited production. *Journal of Psycholinguistic Research, 26*, 519-538.

Eisenberg, S. L., Fersko, T. M., & Lundgren, C. (2001). The use of MLU for identifying language impairment in preschool children. *American Journal of Speech-Language Pathology, 10*, 323-342. doi:10.1044/1058-0360(2001/028)

Ellis Weismer, S. (2007). Typical talkers, late talkers, and children with specific language impairment: A language endowment spectrum? In R. Paul (Ed.), *Language disorders from a developmental perspective: Essays in honor of Robin S. Chapman. New directions in communication disorders research* (pp. 83-101). Psychology Press, Taylor & Francis Group.

Fenson, L., Resznick, S., Thal, D., Bates, E., Hartung, J., Pethick, S., & Reilly, J. (1993). *The MacArthur Communicative Developmental Inventory*. Singular Publishing Group.

Fisher, E. L. (2017). A systematic review and meta-analysis of predictors of expressive-language outcomes among late talkers. *Journal of Speech, Language, and Hearing Research, 60*, 2935-2948. doi:10.1044/2017_JSLHR-L-16-0310

Fernald, A., & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Development, 83*, 203-222. doi:10.1111/j.1467-8624.2011.01692.x

Frankenburg, W., Dodds, J., Archer, P., Bresnick, B., Maschka, P., Edelman, N., & Shapiro, H. (1990). *Denver Screening Test II*. Denver Developmental Materials.

Gavin, W. J., & Giles, L. (1996). Sample size effects on temporal reliability of language sample measures of preschool children. *Journal of Speech, Language, and Hearing Research, 39*, 1258-1262. doi:10.1044/jshr.3906.1258

Girolametto, L., Wiigs, M., Smyth, R., Weitzman, E., & Pearce, P. S. (2001). Children with a history of expressive vocabulary delay. *American Journal of Speech-Language Pathology, 10*, 358-369. doi:10.1044/1058-0360(2001/030)

Goffman, L., & Leonard, J. (2000). Growth of language skills in preschool children with specific language impairment: Implications for assessment and intervention. *American Journal of Speech-Language Pathology, 9*, 151-161. doi:10.1044/1058-0360.0902.151

Hadley, P. A. (1998). Language sampling protocols for eliciting text-level discourse. *Language, Speech, and Hearing Services in Schools, 29*, 132-147. doi:10.1044/0161-1461.2903.132

Hadley, P. A., & Short, H. (2005). The onset of tense marking in children at risk for specific language impairment. *Journal of Speech, Language, and Hearing Research, 48*, 1344-1362. doi:10.1044/1092-4388(2005/094)

Heilmann, J., DeBrock, L., & Riley-Tillman, T. C. (2013). Stability of measures from children's interviews: The effects of time, sample length, and topic. *American Journal of Speech-Language Pathology, 22*, 463-475. doi:10.1044/1058-0360(2012/11-0035)

Heilmann, J., Weismer, S. E., Evans, J., & Hollar, C. (2005). Utility of the MacArthur—Bates Communicative Development Inventory in identifying language abilities of late-talking and typically developing toddlers. *American Journal of Speech-Language Pathology, 14*, 40-51. doi:10.1044/1058-0360(2005/006)

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children’s language growth. *Cognitive Psychology, 61*, 343-365. doi:10.1016/j.cogpsych.2010.08.002

Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology, 45*, 337-374. doi:10.1016/S0010-0285(02)00500-5

IBM Corp. Released 2017. IBM SPSS Statistics for MacIntosh, Version 25.0. IBM Corp.

Jackson, S. C., & Roberts, J. E. (2001). Complex syntax production of African American preschoolers. *Journal of Speech, Language, and Hearing Research, 44*, 1083-1096. doi:10.1044/1092-4388(2001/086)

Klee, T., Carson, D. K., Gavin, W. J., Hall, L., Kent, A., & Reece, S. (1998). Concurrent and predictive validity of an early language screening program. *Journal of Speech, Language, and Hearing Research, 41*, 627-641. doi:10.1044/jslhr.4103.627

Klein, H. B., Moses, N., & Jean-Baptiste, R. (2010). Influence of context on the production of complex sentences by typically developing children. *Language, Speech, and Hearing Services in Schools, 41*, 289-302. doi:10.1044/0161-1461(2009/08-0080)

Lee, E. C., & Rescorla, L. (2002). The use of psychological state terms by late talkers at age 3. *Applied Psycholinguistics, 23*, 623-641. doi:10.1017/S014271640200406X

Lee, E. C., & Rescorla, L. (2008). The use of psychological state words by late talkers at ages 3, 4, and 5 years. *Applied Psycholinguistics, 29*, 21-39. doi:10.1017/S0142716408080028

Lee, L. L. (1974). *Developmental sentence analysis: A grammatical assessment procedure for speech and language clinicians*. Chicago, IL: Northwestern University Press.

Limber, J. (1973). The genesis of complex sentences. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 169-185). Academic Press.

Manhardt, J., & Rescorla, L. (2002). Oral narrative skills of late talkers at ages 8 and 9. *Applied Psycholinguistics, 23*, 1-21. doi:10.1017/S0142716402000012

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*, 517-533. doi:10.1016/j.jcomdis.2004.03.005

Miller, J. F., & Chapman, R. S. (1981). The relation between age and mean length of utterance in morphemes. *Journal of Speech, Language, and Hearing Research, 24*, 154-161. doi:10.1044/jshr.2402.154

Miller, J., & Iglesias, A. (2012). Systematic Analysis of Language Transcripts (SALT), Student Version 2012 [Computer Software]. Middleton, WI: SALT Software, LLC.

Morton, I., Delgado, A., & Schuele, C. M. (2019). *The production of complex syntax in five-year-old children with specific language impairment*. Paper presented at the Symposium on Research in Child Language Disorders, Madison, WI.

Moyle, M. J., Ellis Weismer, S., Evans, J. L., & Lindstrom, M. J. (2007). Longitudinal relationships between lexical and grammatical development in typical and late-talking children. *Journal of Speech, Language, and Hearing Research, 50*, 508-528. doi:10.1044/1092-4388(2007/035)

Nelson, K. E., Denninger, M. M., Bonvillian, J. D., Kaplan, B. J., & Baker, N. D. (1984). Maternal input adjustments and non-adjustments as related to children's linguistic advances and to language acquisition theories. In A. D. Pelligrini (Ed.), *The development of oral and written language in social contexts* (Vol. 13, pp. 31-56). Ablex Publishing.

Newcomer, P., & Hammill, D. (1997). *Test of Language Development-3 (Primary)*. Austin, TX: Pro-Ed.

Olson, J., & Masur, E. F. (2019). Developmental changes in the frequency and complexity of mothers' internal state utterances across the second year. *First Language, 39*, 462–476. doi:10.1177/0142723719850001

Owen, A. J., & Leonard, L. B. (2006). The production of finite and nonfinite complement clauses by children with specific language impairment and their typically developing peers. *Journal of Speech, Language, and Hearing Research, 49*, 548-571. doi:10.1044/10902-4388(2006/040)

Paul, R. (1981). Analyzing complex sentence development. In J. Miller (Ed.), *Assessing language production in children: Experimental procedures* (pp. 36-40). Baltimore, MA: University Park Press.

Paul, R., & Alforde, S. (1993). Grammatical morpheme acquisition in 4-year-olds with normal, impaired, and late-developing language. *Journal of Speech, Language, and Hearing Research, 36*, 1271-1275. doi:10.1044/jshr.3606.1271

Paul, R., Looney, S. S., & Dahm, P. S. (1991). Communication and socialization skills at ages 2 and 3 in "late-talking" young children. *Journal of Speech, Language, and Hearing Research, 34*, 858-865. doi:10.1044/jshr.3404.858

Paul, R., & Smith, R. L. (1993). Narrative skills in 4-year-olds with normal, impaired, and late-developing language. *Journal of Speech, Language, and Hearing Research, 36*, 592-598. doi:10.1044/jshr.3603.592

Paul, R. (1996). Clinical implications of the natural history of slow expressive language development. *American Journal of Speech-Language Pathology, 5*, 5-21. doi:10.1044/1058-0360.0502.05

Peter, M. S., Durrant, S., Jessop, A., Bidgood, A., Pine, J. M., & Rowland, C. F. (2019). Does speed of processing or vocabulary size predict later language growth in toddlers? *Cognitive Psychology, 115*, 101238. doi:10.1016/j.cogpsych.2019.101238

Reilly, S., Wake, M., Ukoumunne, O. C., Bavin, E., Prior, M., Cini, E., . . . Bretherton, L. (2010). Predicting language outcomes at 4 years of age: Findings from Early Language in Victoria Study. *Pediatrics, 126*, e1530-e1537. doi:10.1542/peds.2010-0254

Rescorla, L. (2002). Language and reading outcomes to age 9 in late-talking toddlers. *Journal of Speech, Language, and Hearing Research, 45*, 360-371. doi:10.1044/1092-4388(2002/028)

Rescorla, L. (2005). Age 13 language and reading outcomes in late-talking toddlers. *Journal of Speech, Language, and Hearing Research, 48*, 459-472. doi:10.1044/1092-4388(2005/031)

Rescorla, L. (2009). Age 17 language and reading outcomes in late-talking toddlers: Support for a dimensional perspective on language delay. *Journal of Speech, Language, and Hearing Research, 52*, 16–30. doi:10.1044/1092-4388(2008/07-0171)

Rescorla, L., Dahlsgaard, K., & Roberts, J. (2000a). Late-talking toddlers: MLU and IPSyn outcomes at 3;0 and 4;0. *Journal of Child Language, 27*, 643-664. doi:10.1017/S0305000900004232

Rescorla, L., Mirak, J., & Singh, L. (2000). Vocabulary growth in late talkers: lexical development from 2;0 to 3;0. *Journal of Child Language, 27*, 293-311. doi:10.1017/S030500090000413X

Rescorla, L., Roberts, J., & Dahlsgaard, K. (1997). Late talkers at 2: Outcome at age 3. *Journal of Speech, Language, and Hearing Research, 40*, 556-566. doi:10.1044/jslhr.4003.556

Rescorla, L., & Schwartz, E. (1990). Outcome of toddlers with specific expressive language delay. *Applied Psycholinguistics, 11*, 393-407. doi:10.1017/S0142716400009644

Rescorla, L., & Turner, H. L. (2015). Morphology and syntax in late talkers at age 5. *Journal of Speech, Language, and Hearing Research, 58*, 434-444. doi:10.1044/2015_JSLHR-L-14-0042

Reynell, J. (1977). *Reynell Developmental Language Scales*. NFER.

Rice, M. L., Taylor, C. L., & Zubrick, S. R. (2008). Language outcomes of 7-year-old children with or without a history of late language emergence at 24 months. *Journal of Speech, Language, and Hearing Research, 51*, 394–407 doi:10.1044/1092-4388(2008/029)

Rice, M. L., & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech, Language, and Hearing Research, 39*, 1239-1257. doi:10.1044/jshr.3906.1239

Rice, M., & Wexler, K. (2001). *Rice Wexler Test of Early Grammatical Impairment*. The University of Kansas.

Riches, N., Tomasello, M. & Conti-Ramsden, G. (2005). Verb learning in children with SLI: Frequency and spacing effects. *Journal of Speech, Language, and Hearing Research, 48,* 1397-1411. doi:10.1044/1092-4388(2005/097)

Robbins, J., & Klee, T. (1987). Clinical assessment of oropharyngeal motor development in young children. *Journal of Speech and Hearing Disorders, 52*, 271-277. doi:10.1044/jshd.5203.271

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development, 83*, 1762-1774. doi:10.1111/j.1467-8624.2012.01805.x

Scarborough, H. S. (1990). Index of productive syntax. *Applied Psycholinguistics, 11*, 1-22. doi:10.1017/S0142716400008262

Scarborough, H. S., & Dobrich, W. (1990). Development of children with early language delay. *Journal of Speech, Language, and Hearing Research, 33*, 70-83. doi:10.1044/jshr.3301.70

Schuele, C. M., & Dykes, J. C. (2005). Complex syntax acquisition: A longitudinal case study of a child with specific language impairment. *Clinical Linguistics & Phonetics, 19*, 295-318. doi:10.1080/02699200410001703709

Schuele, C. M. (2009a). *Complex syntax coding manual*. Unpublished manuscript. Child Language and Literacy Lab, Department of Hearing and Speech Sciences. Nashville, Tennessee.

Schuele, C. M. (2009b). *Transcription and basic coding manual*. Unpublished manuscript. Child Language and Literacy Lab, Department of Hearing and Speech Sciences. Nashville, Tennessee.

Schuele, C. M., Teller, L., Kaiser, A., & Camarata, S. M. (2020). *Maternal complex syntax input to four-year-old boys: Does child history of Language Impairment make a difference?* Unpublished manuscript.

Schuele, C. M., & Tolbert, L. (2001). Omissions of obligatory relative markers in children with specific language impairment. *Clinical Linguistics & Phonetics, 15*, 257-274. doi:10.1080/02699200010017805

Teller, L., & Schuele, C. M. (2019, November). *Simple strategies to increase preschool teacher complex syntax in play interactions.* Poster presented at the Annual Convention of the American Speech-Language-Hearing Association, Orlando, FL.

Thal, D. J., Marchman, V. A., & Tomblin, J. B. (2013). Late-talker toddlers: Characterization and prediction of continued delay. In L. A. Rescorla, & P. S. Dale (Eds.), *Late talkers: Language development, intervention, and outcomes* (pp. 83-101). Paul H. Brooks Publishing Company.

Tyack, D. L., & Gottsleben, R. H. (1986). Acquisition of complex sentences. *Language, Speech, and Hearing Services in Schools, 17*, 160-174. doi:10.1044/0161-1461.1703.160

Weiler, B. K., & Schuele, C. M. (2014). Joining clauses with subordinate conjunctions: One type of complex syntax. *Perspectives on Language Learning and Education, 21*, 182-191. doi:10.1044/lle21.4.182

Zimmerman, I., Steiner, V., & Pond, R. (1992). *Preschool Language Scale-3.* Psychological Corporation.

Zubrick, S. R., Taylor, C. L., Rice, M. L., & Slegers, D. W. (2007). Late language emergence at 24 months: An epidemiological study of prevalence, predictors, and covariates. *Journal of Speech, Language, and Hearing Research, 50*, 1562–1592. doi:10.1044/1092-4388(2007/106)