Proceedings of The National Conference On Undergraduate Research (NCUR) 2018 University of Central Oklahoma Edmond, Oklahoma April 5-7, 2018

Syntactic Complexity of Narratives Produced by Typically Developing Children Ages 4-7

Madeline Peterson, Alison Barlow, Natalie Green Communicative Disorders Utah State University Old Main Hill Logan, UT 84322 USA

Faculty Mentor: Sandra Gillam

Abstract

Children with Language Disorders (LD) tend to demonstrate significant difficulty using multiple clause sentences that make up complex syntax. More research is needed to determine how children who are typically developing use complex syntax so that we have a better standard for determining when children may be experiencing difficulty developing competent syntax abilities. Our current study's goal was to extend findings from previous research in children who are typically developing by using narratives as the context for eliciting and examining the use of complex syntax and by including younger children. We studied the syntactic complexity of narratives of 260 children who were typically developing between the ages of 4 and 7. Children were given 3 narrative tasks which included retelling a story, composing a story from sequenced pictures, and producing a story from a single picture prompt. The narratives were transcribed and coded for the use of complex sentences by research assistants who were at least 90% reliable using the *Systematic Analysis of Language Transcripts*⁶. The syntactic complexity of these narratives was analyzed by coding for Subordination Index (SI). Our results indicated that SI scores do differ across age levels for each story context, including retell, sequenced pictures, and single scene prompt. This means that older children typically use more complex syntax in their narratives. We also found that the SI scores did not vary significantly across the 3 narrative tasks with age collapsed which may indicate that SI is a stable measure across context.

Keywords: Childhood Language, Syntactic Complexity, Narratives

1. Introduction

Children with Developmental Language Disorders (DLD) tend to demonstrate significant difficulty using multiple clause sentences that make up complex syntax. Given the importance of complex syntax comprehension and production during the school-age years, it is important that we have a firm understanding of how it develops in children who are typically developing so we may better identify and treat problems with complex syntax proficiency in children with DLD.

Prior research that has investigated complex syntax use by children who are typically developing and those with a history of late language emergence (LLE) has shown that the context in which language is elicited plays an important role in whether or not complex syntax is used^{1,5,7,8}. For example, Norbury and Bishop⁹ compared the narrative skills of children who are typically developing to children with the following three diagnoses: specific language impairment (SLI), pragmatic language impairment (PLI), and high functioning autism (HFA). The participants included 17 children with SLI, 21 children with PLI, 12 children with HFA, and 18 children who were developing typically between the ages of 6 to 10 years. The children were asked to compose a narrative from the wordless picture book *Frog, Where are You?* The results indicated that the children from all three groups showed no statistical significance

in the length of their stories between the groups. However, the children with SLI and HFA demonstrated less syntactic complexity and a higher number of tense marking errors in their narratives compared to the children who are typically developing. In general, narrative language sampling has been proposed as the best method to characterize expressive language across a wide range of ages as well as disorder types (i.e., Autism, Language Disorder)¹. For this reason, narrative language sampling was used in the current study to examine complex syntax in children who are typically developing ranging in age from 4 to 7.

2. Methodology

2.1 Participants

The participants included 4-year-olds (n = 42), 5-year-olds (n = 57), 6-year-olds (n = 66), and 7-year-olds (n = 94) whose narrative abilities were assessed through the Test of Narrative Language³. They were typically developing, were not receiving special education services, and did not demonstrate hearing, vision, neurological or other medical conditions. The data for the children who participated in this study were recruited at various sites across the United States including Texas and Utah.

2.2 Procedures

The TNL consists of various narrative comprehension and production tasks. The format of the TNL is as follows. The child listens to a story about a boy, a girl, and their mother who go to McDonalds after school. After answering comprehension questions, the child is asked to retell the story as similarly as possible. Next, the child listens to a story that follows a series of sequenced pictures about a girl who breaks her project on her way to school. Following the comprehension questions, the child is requested to generate a story based on a novel series of sequenced pictures about a boy who is late for school. Finally, the child listens to a story based on a single picture scene about two children who find a dragon. The child is then asked to formulate a story about a novel picture scene which depicts two children and their encounter with an alien family. This study focused on the narrative production tasks. Thus, children were asked to produce stories in three different contexts; a story retell after a clinician model, a story generation based on a novel picture scene after the presentation of a model, and a story generation based on a single picture prompt after a model was given. The narratives of the children who participated in the normative sample for the TNL were analyzed for syntactic complexity using the subordination index (described in the outcome measures section).

3. Outcome Measures

The narratives were transcribed and coded using the *Systematic Analysis of Language Transcripts (SALT)*⁶. SALT is a language sample transcription and coding system that is commonly used in speech language pathology research. SALT prescribes how samples are transcribed, segmented and coded. Five research assistants who were at 90%+ reliable in using the transcription, segmentation and coding rules analyzed all of the data used in this project. The syntactic complexity of the narratives elicited from the participants was also coded using the Subordination Index (SI) scheme used in the SALT program. SI is a ratio of the number independent clauses to the number of C-units (i.e. independent main clauses and phrases/clauses subordinated to it) in each transcript. Each utterance was coded according to whether the subordination index was X (unintelligible), 0 (not a complete sentence), 1 (contains one independent clause), 2 (contains 2 independent/dependent clauses), 3 (contains 3 independent/dependent clauses) or 4 (contains 4 independent/dependent clauses). No participant demonstrated the use of more than 4 independent/dependent clauses in their narrative(s).

3.1 Reliability

Before coding SI using the data in the study, reliability was established by having five research assistants score 30 transcripts that were compared to a gold standard transcript. Two of the RAs were blind to the purpose of the study. When the RAs were at 90% or higher in coding SI, they began coding the transcripts for the study. The transcripts were scored by the RAs who met weekly to resolve any discrepancies between their independently coded transcripts.

Percentage of agreement for coding SI, point by point, between primary and secondary assistants (blind to the purpose of the study) was 90%.

4. Results

Table 1.1 shows the means and standard deviations for the SI scores of the story retell, sequenced pictures, and the single scene picture prompt from the TNL.

A one-way, multivariate analysis of variance (MANOVA) was conducted to determine the effect of age on SI scores for the McDonald's stories (story retell), Late For School (LFS; sequenced pictures, and Aliens (single scene prompt), and overall performance on the Test of Narrative Language Index (TNLAI raw score). Significant differences were found among the age groups on the dependent measures, Wilks's $\Lambda = .852$, p < .001. The multivariate n² based on Wilks's Λ was strong, n²=.052. Analyses of variances (ANOVA) on the dependent variables were conducted as follow-up tests to the MANOVA. Using the Tukey method, each ANOVA was tested at the .05 level.

The ANOVA for the SI scores for the Aliens story was significant, F(3, 255) = 7.32, p = .001, $n^2 = .079$, as was the ANOVA for the LFS story, F(3, 255) = 5.63, p = .001, $n^2 = .062$, and the ANOVA for the McDonald's story, F(3, 255) = 4.75, p = .003, $n^2 = .053$. The ANOVA for the TNLAI raw score was also significant, F(3, 220) = 42.89, p = .001, $n^2 = .369$.

Post hoc analyses to the univariate ANOVA for the scores revealed that the six- and seven-year-old children earned significantly higher SI scores on the Aliens story than did the four- and five-year-old children. There was no significant difference in the SI Aliens story scores for the six- and seven-year-old children. For the LFS story, the seven-year-old children earned higher SI scores than the four- and five-year-old children. The five and six-year-old children earned higher SI scores than the four-year-old children. For the McDonald's story, the six and seven-year-old children earned higher SI scores than the four and five-year-old children. There was no significant difference in the SI scores for the four and five-year-old children.

Age (years)	McDonald's(SD)	Late for School(SD)	Aliens(SD)
4 (n=42)	0.95(.25)	0.92(.36)	0.85(.49)
5 (n=57)	0.93(.52)	1.04(.16)	0.92(.36)
6 (n=67)	1.08(.37)	1.06(.30)	1.11(.16)
7 (n=93)	1.15(.37)	1.13(.29)	1.08(.37)

Table 1.1. Mean Subordination index scores by age for McDonald's (story retell), Late for School (sequenced pictures), and Aliens (single scene).

Post hoc analyses to the univariate ANOVA for the scores revealed that the seven-year-old children earned significantly higher raw scores on the TNLAI than the all of the other age groups. The six-year-old children performed significantly higher than the four- and five-year-old children. The five-year-old children performed significantly higher than the four-year-old children.

A one-way, multivariate analysis of variance (MANOVA) was conducted to determine the effect of age on the total number of utterances used for the Alien (single scene prompt), LFS (sequenced pictures), and McDonald's stories (story retell). Significant differences were found among the age groups on the dependent measures, Wilks's $\Lambda = .922$, p <.014. The multivariate n² based on Wilks's Λ was strong, n² = .027. Analyses of variances (ANOVA) on the dependent variables were conducted as follow-up tests to the MANOVA. Using the Tukey method, each ANOVA was tested at the .05 level. The ANOVA for the Total number of utterances for the Alien story was not significant; however, the ANOVA for the LFS story, F(3, 255) = 4.00, p = .008, n² = .045, and the ANOVA for the McDonald's story, F(3, 255) = 5.42, p = .001, n² = .06 were significant.

Post hoc analyses to the univariate ANOVA for the scores revealed that the seven-year-old children used more utterances than the four-year-old children in the LFS story and used more utterances than both the four- and five-year-old children in the McDonald's story.



Figure 1. Total number of utterances by age group for each story context.

A one-way, multivariate analysis of variance (MANOVA) was conducted to determine the effect of age on the total number of different words used for the Aliens (single scene prompt), LFS (sequenced pictures), and McDonald's stories (story retell). Significant differences were found among the age groups on the dependent measures, Wilks's $\Lambda = .738$, p <.001. The multivariate n-based on Wilks's Λ was strong, n²=.096. Analyses of variances (ANOVA) on the dependent variables were conducted as follow-up tests to the MANOVA. Using the Tukey method, each ANOVA was tested at the .05 level. The ANOVA for the Number of Different Words used for the Aliens story, the LFS story and McDonald's story were significant (p = <.05).

Post hoc analyses to the univariate ANOVA for the scores revealed that the six- and seven-year-old children used more total number of different words (TNDW) in their Aliens story than did the four- and five-year-old children. The five-year-old children used more TNDW in their Aliens story than the four-year-old children. For LFS, the seven-year-old children used more TNDW in their story than any of the other age groups. The six-year old children used more TNDW than the four-year-old children. No other differences were significant for the LFS story. For the McDonald's story, the seven-year-old children used more TNDW than the four- and five-year-old children. The four- and five-year-old children. The four- and five-year-old children. The four- and five-year-old children are groups, followed by the six-year-old children, who used more TNDW than the four- and five-year-old children. The four- and five-year-old children did not differ on NDW used in the McDonald's story.



Figure 2. Number of different words by age group for each story context.

Paired samples t-tests showed that there were no significant differences in SI scores between the three story contexts. Correlation coefficients were computed among the SI scores, total number of utterances (TNU) used, TNDW used and the TNLAI raw scores. These correlations and their corresponding *p*-values are shown in Table 1.2. The results of the correlational analyses show that the correlations between the SI scores and measures of language productivity (TNU) and language diversity (TNDW) were significant.

	ATotNu mUt.T	Total Produc tion Raw Score	AND W.A	SIA	LFSTot NumUt .T	LFSND W.A	SILFS	McDTot NumUt.T	McDN DW.A	Mc DSI
ATotNumU t.T	1									
Total Production Raw Score	.367**	1								
ANDW.A	.838**	.610**	1							
SIA	.222**	.203**	.424**	1						
LFSTotNu mUt.T	.445**	.429**	.503**	0.117	1					

Table 1.2. Correlations for subordination index scores, total number of utterances used, total number of different words used, and production raw scores.

LFSNDW. A	.487**	.623**	.629**	.208**	.898**	1				
SILFS	.138*	.320**	.274**	.304**	.278**	.432**	1			
McDTotNu mUt.T	.329**	.475**	.389**	0.067	.382**	.410**	.127*	1		
McDNDW. A	.361**	.754**	.553**	.209**	.398**	.539**	.270**	.798**	1	
McDSI	.127*	.393**	.226**	0.08	.131*	.229**	.299**	.231**	.443**	1

Key:

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Note: ATotNumUt.T = Total number of Utterances in Aliens Story; Total Production Raw Score = Total Production Score across all stories; ANDW.A - Number of Different Words in Aliens Story; SIA = Subordination Index for Aliens Story; LFSTotNumUt.T = Total number of utterances in Late For School Story; LFSNDW.A = Number of Different Words in Late For School Story; SILFS = Subordination Index for Late For School Story; McDNDW.A = Number of Different Words in McDonald's; McDSI = Subordination Index for McDonald's Story

5. Discussion

The current study addressed the question: do SI scores differ across age levels for each story context, including retell, sequenced pictures, and single scene prompt? The results affirmed that six- and seven-year-old children earned significantly higher SI scores on the Aliens story than did the four- and five-year-old children. For the LFS story, the seven-year-old children earned higher SI scores than the four- and five-year-old children, while the five- and six-year-old children earned higher SI scores than the four- and five-year-old children; while the six- and seven-year-old children earned higher SI scores than the four- and five-year-old children; however, the six- and seven-year-old children earned higher SI scores than the four- and five-year-old children; however, there was no significant difference in the SI scores between the four- and five-year-old children. These findings are consistent with Tyack and Gottsleben's¹⁰ findings that age is a significant predictor of syntactic complexity.

Another goal of this study was to determine whether SI scores differed across story context with age-level collapsed. We found that there was no significant difference in the SI scores across the 3 story contexts. This may suggest that SI is a stable measure across various contexts. In addition, our research concluded that SI scores correlated with measures of language productivity (TNU) and diversity (TNDW). SI scores for the Aliens and LFS stories had small correlations with the TNU and TNDW. The McDonald's story had a small to moderate correlation with the TNU and TNDW. SI scores for the Aliens and LFS stories had small correlations with the TNLAI. SI scores for the McDonald's story had a moderate correlation with the TNLAI.

6. Clinical Implications

Greater narrative proficiency scores may be correlated with more syntactically complex narratives. This finding is consistent with previous research findings that children who can tell better stories use more complex sentences in their stories. For example, Zanchi et al.¹¹ analyzed the syntactic complexity of narratives elicited from a wordless picture book produced by children attending kindergarten (T1) and one year post kindergarten attendance (T2). They found that at T1 the children who demonstrated more narrative competence (number of acting characters and number of events told) used more syntactically complex utterances in their narratives. These findings are also consistent with Israelsen and Gillam's⁴ findings when examining the syntactic complexity of narratives elicited through story retells

from children with ASD. They found that after three phases of narrative intervention, the children who told better stories demonstrated more syntactically complex narratives compared to their narratives prior intervention. This could mean that if we teach children to tell better stories, that we may also expect increases in the complexity of their language.

In the current study, higher subordination index scores were associated with longer stories that contained a greater variety of words. A higher subordination index score typically indicates a more complex story, which usually indicates that the story was longer and contained more complex syntax. The increase in syntactic complexity was also associated with the use of more diverse vocabulary in their stories.

One last observation was that children tended to use more diverse vocabulary when they were allowed to create their own stories as opposed to retelling stories they heard. For example, all of the students in this study used more diverse vocabulary (not using the same words over and over) in the Aliens story when they were asked to create their own story as compared to the McDonald's story that they were asked to retell a story they heard. The four- and five-year-old children were also observed to use more diverse vocabulary in the story that required them to create a narrative from sequenced pictures (LFS) as compared to the McDonald's story (a retell).

Future research should investigate potential differences in SI scores for older children across these varied narrative contexts to see if the patterns observed in the younger children are maintained over time.

7. Acknowledgements

I want to especially thank Samantha Winward, Graduate Student Mentor for her guidance.

8. References

1. Channell, M. M., Loveall, S. J., Conners, F. A., Harvey, D. J., & Abbeduto, L. (2018). Narrative Language Sampling in Typical Development: Implications for Clinical Trials. *Am J Speech Lang Pathol*, 27(1), 123-135. doi: 10.1044/2017_AJSLP-17-0046.

2. Domsch, C., Richels, C., Saldana, M., Coleman, C., Wimberly, C., & Maxwell, L. (2012). Narrative skill and syntactic complexity in school-age children with and without late language emergence. International Journal of Language & Communication Disorders, 47(2), 197-207. doi:10.1111/j.1460-6984.2011.00095.x

3. Gillam, R., & Pearson, N. (2004). Test of Narrative Language. Austin, TX: PRO-Ed Inc.

4. Israelsen, M., Gillam, S. L. (2016) The Relationship between Narrative Proficiency and Syntactic Complexity of Story Retells Elicited from Children with ASD Spectrum Disorders (ASD). Autism Open Access 6: 181. doi:10.4172/2165-7890.1000181

5. Kapantzoglou, M., Fergadiotis, G., & Restrepo, M. A. (2017). Language Sample Analysis and Elicitation Technique Effects in Bilingual Children With and Without Language Impairment. *J Speech Lang Hear Res*, 60(10), 2852-2864. doi: 10.1044/2017_JSLHR-L-16-0335.

6. Miller, J. & Iglesias, A. (2010). Systematic Analysis of Language Transcripts (SALT), Research Version 2010 [Computer Software], SALT Software, LLC.

7. Nippold, M. A., Frantz-Kaspar, M. W., Cramond, P. M., Kirk, C., Hayward-Mayhew, C., & MacKinnon, M. (2014). Conversational and Narrative Speaking in Adolescents: Examining the Use of Complex Syntax. *J Speech Lang Hear Res*, 57(3), 876-886. doi: 10.1044/1092-4388(2013/13-0097).

8. Nippold, M. A., Frantz-Kaspar, M. W., & Vigeland, L. M. (2017). Spoken Language Production in Young Adults: Examining Syntactic Complexity. *J Speech Lang Hear Res*, 60(5), 1339-1347. doi: 10.1044/2016_JSLHR-L-16-0124.

9. Norbury, C. F., & Bishop, D. M. (2003). Narrative Skills of Children with Communication Impairments. *International Journal Of Language & Communication Disorders*, *38*(3), 287-313.

10. Tyack, D. L., & Gottsleben, R. H. (1986). Acquisition of Complex Sentences. Language, Speech, And Hearing Services In Schools, 17(3), 160-74.

11. Zanchi, P., Zampini, L., Fasolo, M., & D'Odorico, L. (2016). Syntax and Prosody in Narratives: A Study of Preschool Children. *First Language*, *36*(2), 124-139.