2012-12-14

Noun Clauses in Clinical Child Language Samples

Christine Beate Scoville
Brigham Young University - Provo

Noun Clauses in Clinical Child Language Samples

Christine Scoville

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Ron W. Channell, Chair
Shawn L. Nissen
Martin Fujiki

Department of Communication Disorders

Brigham Young University

December 2012

ABSTRACT

Noun Clauses in Clinical Child Language Samples
Christine Scoville

Department of Communication Disorders, BYU
Master of Science

Noun clauses are grammatical constructions that are of relevance both to typical language development and impaired language development. These clauses have been part of published techniques for the clinical analysis of language samples, and computer software for the automated analysis of clinical language samples has attempted to identify noun clauses, with limited success. The present study examined the development and clinical use of noun clauses as well as the automated identification of these clauses.

Two sets of language samples were examined. One set consisted of 10 children with specific language impairment (SLI) whose age ranged from 7;6 to 11;1 (years;months), 10 peers matched for language development equivalence, and 10 peers matched for chronological age. The second set of samples were from 30 children considered to be typically developing, who ranged in age from 2;6 to 7;11.

Language sample utterances were manually coded for the presence of noun clauses (including *wh- noun clauses*, *that-* noun clauses, and gerunds.) Samples were then automatically tagged using software. Results were tabulated and compared for accuracy. ANCOVA revealed that differences in the frequencies of WH-infinitive noun clauses and gerunds were significant between the matched groups. "Zero that clauses" (*that*-noun clauses containing no subordinator *that*) and gerunds were significantly correlated with age. Kappa levels revealed that agreement between manual and automated coding was high on WH-infinitive clauses, gerunds, and finite wh-noun clauses.

Key words: child language, language development, noun clauses, grammatical analysis

ACKNOWLEDGEMENTS

First and foremost I would like to thank my advisor, Dr. Channell, for being available continually throughout the process of this project, offering direction, encouragement, and humor.

My husband, David offered encouragement and practical help in terms of formatting of this manuscript. Thank you, David.

Thank you also to Dr. Nissen and Dr. Fujiki, who offered valuable input and perspective on this project and for supporting me in this process.

Thank you to everyone else who has offered encouragement and support. You are appreciated; I could not have done this alone!

Table of Contents

List of Tables

List of Appendixes

Description of Structure and Content

The body of this thesis is written as a manuscript suitable for submission to a peer-reviewed journal in speech-language pathology. An annotated bibliography is presented in Appendix A.

Introduction

Noun clauses are grammatical constructions that are of relevance both to typical language development (O'Grady, 1997) and impaired language development (e.g., Schuele & Tolbert, 2001). Perhaps because of this relevance, published techniques for the clinical analysis of language samples generally include noun clauses, and computer software for the automated analysis of clinical language samples has attempted to identify noun clauses, though with limited success. The present study aims to extend our knowledge regarding the development and clinical use of noun clauses as well as the automated identification of these clauses.

Several syntactic constructions can function as noun clauses in English. Noun clauses are created when an entire clause is placed in the grammatical position of a sentence that a pronoun might otherwise occupy. For example, in the sentence *I know why you like peas,* the clause *why you like peas* is considered a noun clause because it could be replaced with a pronoun such as *it*. Noun clauses are a type of a subordinate clause or embedded clause because they cannot stand alone and fit syntactically within a matrix clause (a clause that contains the main verb of the sentence). Relative clauses and adverbial clauses are other examples of subordinate clauses (Diessel, 2004). Noun clauses fall into two classes: finite and nonfinite (Diessel, 2004). In finite types, the verb is marked for tense and number. An example of a finite noun clause is *Peter promised that he would come*. Nonfinite noun clauses use the same verb regardless of tense or number; for example *Sue wants Peter to leave*. Finite noun clauses can be divided into three forms: *wh*-noun clauses (WNC), *that*-noun clauses (TNC), and *if*-noun clauses (INC). For the purposes of this study TNCs will be separated into TNCt (containing the subordinator *that)* and TNCz ("zero" noun clauses, not containing the subordinator *that*.) Nonfinite noun clauses include varieties such as *to*-infinitives and *gerunds.* Typically, noun clauses are paired with a matrix clause, which contains the main verb of the sentence. These matrix clause verbs are most often mental state verbs such as *think*, *believe*, and *assume*, but may also be use- and communication verbs such as *see, say,* and *tell* (Diessel, 2004; Owen & Leonard, 2006).

Children as young as two to two and a half years begin to develop complex language, which typically begins with noun clauses (Diessel, 2004; Scott, 1988). In English speakers, earliest noun clauses emerge in the form of a subject-verb-object (SVO) sentence structure (Limber, 1973; Scott, 1988), with noun clauses emerging attached to formulaic attention getters around two years of age (Diessel, 2004), or as nonfinite clauses introduced with *wanna* and *gonna,* also known as catenatives (Bloom, Tackeff, & Lahey, 1991; Limber, 1973; Owen & Leonard, 2006; Tyack & Gottsleben, 1986). Full constructions of the *to-infinitives* WNCs, and TNCs are often next to emerge (Limber, 1973). As the child's language continues to mature, other complex grammatical constructions such as relative and adverbial clauses begin to emerge (Scott, 1988).

Several studies have examined aspects of the typical or impaired development of noun clauses in individuals with language impairment (LI). Schuele and Dykes (2005) demonstrated that children with delayed language displayed decreased levels of syntactic complexity and omissions of complex syntactic structures such as infinitival *to*, relative markers and pronouns, and *wh*-pronouns. Despite delays in acquisition of complex syntax, children with LI tend to acquire language constructs in the same order as typically developing children (Schuele & Dykes, 2005). However, children with language delay tend to omit many important grammatical markers such as infinitival *to* in nonfinite noun clauses (e.g. *I want play outside* instead of *I want to play outside*; Owen & Leonard, 2006; Schuele & Dykes, 2005). Factors such as the delayed acquisition of complex syntax and omission of grammatical markers can be important diagnostic indicators for children with LI. Much has not yet been studied, however. The frequency of occurrence of noun clauses in language samples from individuals with LI versus typical language development has been examined for finite clauses (Nippold, Mansfield, Billow, & Tomblin, 2008) but not for non-finite noun clauses. Examples of these non-finite clauses are gerunds, or –*ing* clauses (e.g. *watching television* keeps them out of mischief) and wh-infinitives (e.g. I know *when to go*).

Published techniques for the clinical, grammatical analysis of children's spontaneous language often involve the study of noun clauses. Important information regarding the level of complexity of language of individuals on a clinician's caseload can be obtained by determining the appearance and frequency of noun clauses in a child's language sample, especially if the individual is identified as having LI. Because of the important role that noun clauses play in the development of child language, their identification and isolation has long been an object of clinical language sampling. Lee (1974), as part of the Developmental Sentence Scoring (DSS) procedure, allocated points to *wh*-subordinating pronouns in the *personal pronouns* category, to *that*-subordinators in the *conjunctions* category, and to gerunds in the *secondary verb* category. Paul (1981) included noun clauses among the structures to be identified in conjunction with the analysis of complex utterances. Crystal, Garman, and Fletcher (1989) included noun clauses at stage five of seven when developing the Language Assessment, Remediation and Screening Procedure (LARSP). As part of Index of Productive Syntax (IPSyn), Scarborough (1990) included noun clauses as one of the 56 syntactic/morphological items to be scored. Noun clauses specifically have not yet been studied in a developmental framework, however.

Limitations on speech-language clinician time (Long, 2001) and training (Long, 1996) have prompted attempts to automate the analysis of child language samples for information on constructions such as noun clauses. However, as yet, the accuracy of this software for the analysis of noun clauses has not been specified. Long and Channell's (2001) study yielded only 12% accuracy in agreement between manual coding and automated analysis on LARSP's subordinate clauses level of analysis using CP (Long & Channell, 2001), the level which includes noun clauses. Channell's (2003) study of automated DSS analysis yielded a point-by-point agreement average of 40% between automated and manual analysis on the categories that include noun clauses. Clearly, improvement in automated recognition of structures such as noun clauses is necessary for any clinical or research application.

Improvement in the accuracy of automated parsing might be possible by using recently developed and available tools such as the Stanford parser (Klein & Manning, 2003). The Stanford parser uses a probabilistic context-free grammar model to create an analysis of syntactic structure. However, this parser does not directly label constituents such as noun clauses, which would be useful in speech-language pathology for comparing children's utterance productions to developmental data. If the output of the Stanford parser could be interpreted by another program which identified constructions such as noun clauses in this output, perhaps substantive improvement in accuracy could be obtained, and thus the clinical utility of automated analysis of clinical child language samples would be enhanced.

Thus further attention to the development and clinical use of noun clauses is warranted, as is the collection of data regarding the accuracy of the automated identification of noun clauses. The present study compared the use of several varieties of noun clauses in the samples of typically developing children and children with language impairment. It also examined the accuracy with which these noun clause varieties could be identified in children's samples using computer software.

Method

Language Samples

The language samples used in the present study were previously collected by other researchers for purposes that differed from those of the present study.

Reno Samples. In 1990, Fujiki, Brinton, and Sonnenberg collected language samples from 30 children as part of a sample for a longer study of conversational repairs with children with LI. These samples will be referred to hereafter as the Reno samples.

The sampling groups consisted of 10 children with LI, 10 peers matched for language age (LA), and 10 peers matched for chronological age (CA). Each group included five males and five females. Children with LI ranged in chronological age (CA) from 7;6 to 11;1, CA-matched children ranged from 7;6 to 11;2, and LA-matched children ranged in age from 5;6 to 8;4.

The children with LI did not show any signs of intellectual impairment and were receiving speech and language services at the time the samples were taken. These children were diagnosed as LI by demonstrating significant deficits in both expressive and receptive language. Significant receptive delays were determined through a score outside of one standard deviation from the mean on two or more of the following tests: Peabody Picture Vocabulary Test-Revised (Dunn & Dunn, 1981), the Test for Auditory Comprehension of Language-Revised (Carrow-Woolfolk, 1985), select subtests of the Test of Language Development-Primary (TOLD-P; Newcomer & Hammill, 1982), and the Clinical Evaluation of Language Functions Screening Test (CELF-S; Semel & Wiig, 1980).

Expressive delays were determined through subtests taken from the TOLD-P and CELF-S and the Clinical Evaluation of Language Functions-Diagnostic Battery (Semel-Mintz & Wiig, 1982). All children with LI had been receiving language therapy since first grade and were receiving therapy at the time that the samples were collected, in addition to being seen by a learning disabilities specialist, primarily for the issue of communication disorders.

Children diagnosed with LI scored below normal limits in in both expressive- and receptive language skills, but scored within normal limits on nonverbal intelligence. LA-matched children were matched to the children with LI by scores from the Utah Test of Language Development (Mecham, Jex, & Jones, 1967). Children matched for CA were matched within four months to a child with LI and attended the same elementary school.

Language samples were collected through a 30 minute conversation with an adult examiner. Several games and toys were used in the collection process, including Viewmasters, the *Guess Who* game, Transformer toys, and a magic kit. Topics were introduced by the examiner (e.g., movies, Christmas vacation), and 200 to 600 utterances were collected from each child.

Provo Samples. The Provo samples were collected by Barber (1989), Chamberlain (1989), and Taylor (1989) as part of three separate thesis studies. The sampled children were all judged to be typically developing, and each child passed a hearing screening. The children ranged in age from 2;6 to 7;11 years. All children lived in a student family housing complex at Brigham Young University in Provo, Utah. Three children from each six-month age interval were randomly selected from a pool of volunteers. Each child provided a language sample of at least 200 child utterances, during which generally only the child and examiner were present. The first ten minutes of each sample were not transcribed, so as to obtain language that was as representative of the child's abilities as possible.

Procedure

Any noun clauses occurring in the language samples were manually coded. Noun clauses were divided into six subcategories for tabulation: (a) finite *wh*-noun clauses, whether serving as the verb complement or as the grammatical subject (sometimes called nominal or headless relative clauses), (b) wh-alternative noun clauses (sometimes called *yes/no noun clauses* or wh-noun clauses starting with *whether* or *if*), (c) *that- noun clauses* (sometimes called full propositional complement clauses) subdivided as to whether the clause started with the complementizer *that* or not (called *zero-that* complement clauses), (d) gerund clauses, and (e) wh-infinitive noun clauses.

Interrater reliability of the manually coded files was calculated by having a second clinician independently code the structures of interest in 20% of the samples. The number of classification agreements was divided by the total number of classification judgements. Using this formula, interrater reliability was found to be 93%.

Following manual analysis, each sample was prepared for automated analysis using a utility program which removed details such as speaker codes, utterances not produced by the target child, parenthetical material, and manual codes. The prepared files were then grammatically analyzed by the Stanford parser (Klein & Manning, 2003). The Stanford parser is a probabilistic context-free grammar parser which uses grammatical data extracted from training corpora to isolate the grammatical constituents of sentences. The Stanford parser output was then analyzed for utterances containing targeted varieties of noun clauses using software which was written as part of the current study and called *cxs*. The cxs software finds patterns in the Stanford parser output such as *(SBAR (WHNP (WP dominated by a VP node* which suggests, in this case, that the utterance contains a wh-noun clause. The output from the cxs program's analysis was then compared to the manual coding of the various noun clause varieties in each utterance of the sample.

Data Analysis. The data from comparing the manual and automated analysis of each noun clause type in each child utterance were assigned to four categories: (a) utterances identified as containing a noun clause only by manual analysis, (b) utterances identified as containing a noun clause only by the software, (c) utterances identified as containing a noun clause by both manual analysis and software analysis, and (d) utterances identified as containing no noun clauses both by manual analysis and software analysis.

Cohen's Kappa levels were calculated for each group of participants to quantify manual to computer agreement while controlling for the possibility of chance agreement. A alpha level of $p < .05$ was used for all statistical comparisons.

Results

Reno Samples

Table 1 shows the frequency of occurrence of each type of noun clause structure for each child in the three Reno groups. It may be seen in Table 1 that children varied greatly in terms of the number of utterances and number of occurrences of each noun clause structure produced. It can also be seen that wh-noun clauses and zero-that noun clauses were the varieties of noun clauses most commonly produced by these children. Most of the gerunds were produced by the chronological age matched children, and two children in the group with LI actually produced the highest number (6) of wh-alternative noun clauses, though the majority of children in that group produced no wh-alternative noun clauses at all.

Table 1

Descriptive Statistics for the Reno Samples, including age, number of utterances, and frequency of each noun clause variety

Child	Age	N Utt.	WNC	INC	TNCt	TNCz	gerund	WH-inf
RLI 1	111	188	1	0	0	0	2	0
RLI 2	90	377	4	0	2	5	0	0
RLI 3	111	123	0	0	0	0	0	0
RLI 4	104	251	3	0	1	18	2	0
RLI 5	104	392	29	6	0	14	2	2
RLI 6	113	301	7	6	1	2	2	1
RLI 7	119	533	15	2	1	8	4	7
RLI 8	133	401	7	0	0	6	2	1
RLI 9	104	198	2	0	0	2	0	0
RLI 10	109	190	0	0	0	1	1	0
RLA 1	91	269	16	0	0	6	0	1
RLA 2	88	181	1	0	0	2	0	0
RLA 3	95	261	9	0	0	2	1	0
RLA 4	66	145	2	0	0	1	1	0
RLA 5	82	219	8	1	0	2	0	2
RLA 6	100	426	10	0	4	11	3	7
RLA 7	69	274	6	2	0	3	1	2
RLA 8	77	259	4	0	1	2	0	1
RLA 9	83	446	25	3	0	7	2	5
RLA 10	84	318	5	0	2	0	1	1
RCA 1	90	375	6	1	0	9	3	1
RCA 2	108	321	2	1	5	6	4	1
RCA 3	106	360	13	0	5	13	3	1
RCA 4	100	405	5	0	0	7	4	0
RCA 5	122	266	1	0	1	7	6	0
RCA 6	110	424	16	0	2	4	1	1
RCA 7	106	307	6	1	1	3	8	3
RCA 8	104	370	9	1	0	3	3	0
RCA 9	132	263	6	0	0	3	19	0
RCA 10	110	288	9	0	1	2	0	1

Table 2 shows a summary of the descriptive statistics for the Reno samples, organized by the type of noun clause structure. It may be seen in Table 2 that the standard deviations were often larger than the means, suggesting that the mean was generally not a reliable indication of group performance. A one-way ANOVA was used to compare the frequencies of noun phrase types among the three groups. Groups differed significantly on only one type of noun clause: gerunds, with F (2, 29) = 4.9; p = .015. A posthoc Neuman-Kuels analysis showed the RCA group to differ from the other two groups, which did not differ from each other.

Because the RCA group had significantly larger utterance numbers per sample, a higher number of noun clauses might be due to the larger number of utterances. Accordingly, an ANCOVA was performed to compare the three groups while controlling for sample length. With this analysis, the difference in the frequency of WH-infinitives, with F (2, 26) = 4.4; p = .023, as well as the frequency of gerunds, with F (2, 26) = 4.8; p = .017 was significant between groups.

Table 2

Summary Statistics for the Reno Samples: Means and Standard deviations for each group and grammatical structure

Group		WNC	INC	TNCt	TNCz	gerund	WH-inf
RLI							
	M	6.8	1.4	0.5	5.6	1.5	1.1
	SD	9.0	2.5	0.7	6.2	1.3	2.1
RLA							
	M	8.6	0.6	0.7	3.6	0.9	1.9
	SD	7.2	1.1	1.3	3.4	1.0	2.3
RCA							
	M	7.3	0.4	1.5	5.7	5.1	0.8
	SD	4.6	0.5	2.0	3.4	5.4	0.9

Provo Samples

Table 3 shows the frequency of each type of noun clause for children in the Provo group.

Table 3

Descriptive Statistics for the Provo Samples, including age, number of utterances, and frequency of each noun clause variety

Child	Age	N Utt.	WNC	INC	TNCt	TNCz	gerund	WH-inf
P1	30	190	0	0	0	0	1	0
P2	30	223	1	0	0	0	0	3
P3	33	194	1	0	0	0	0	1
P4	35	222	2	0	0	3	0	0
P5	37	233	1	0	0	2	0	0
P6	39	221	0	0	0	0	0	0
P7	45	238	3	0	0	1	2	0
P8	45	270	3	1	0	1	1	1
P9	46	206	3	0	0	0	0	0
P10	53	218	9	0	0	13	2	1
P11	56	215	4	1	0	5	1	0
P12	59	217	13	0	1	5	2	1
P13	59	260	3	0	1	0	6	3
P14	62	200	0	0	0	5	1	0
P15	62	216	5	2	0	0	1	0
P16	64	237	9	0	0	0	3	0
P17	65	226	2	0	0	2	0	1
P18	65	283	10	0	0	8	4	0
P19	66	230	4	0	0	4	3	0
P20	68	217	2	0	0	2	1	2
P21	69	378	11	0	1	1	7	0
P22	72	226	6	0	0	3	2	1
P23	75	250	5	1	0	7	5	0
P24	77	328	0	0	0	0	2	6
P25	79	221	3	0	2	2	6	0
P26	79	226	8	0	0	5	0	1
P27	84	259	4	0	0	4	2	5
P28	91	222	5	1	0	5	5	0
P29	94	301	6	2	1	8	12	1
P30	95	313	3	0	4	13	3	2

It may be seen in Table 3 that children varied greatly in terms of the number of utterances and in the number of occurrences of each noun clause structure produced. Though the samples are arranged by child age, no clear pattern of age-related increase in frequency of noun clause types is generally apparent. Because the older children generally produced longer samples, partial correlations were used to examine the relationship between age and frequency of noun clauses of each type while controlling for the number of utterances. These partial correlations are presented in Table 4.

Table 4

Partial correlations between grammatical structures and age (with df =27; 2-tailed)

	WNC	INC	TNCa	TNCz	gerund	WH-inf
Correlation	0.250	0.277	0.305	0.471	0.462	0.094
Significance	0.191	0.146	0.108	0.010	0.012	0.628

It may be seen in Table 4 that the frequencies of zero-*that*-clauses (TNCz) and of gerunds were significantly correlated with age. Other correlations did not reach statistical significance.

As a child's mean length of utterance (MLU) is often a better predictor of syntactic complexity than is the child's age, the relationship between a child's MLU and the frequencies of the varieties of noun clauses was also examined. Using partial correlations to control for differences in the length of samples, the obtained values are presented in Table 5. It can be seen in Table 5 that the frequencies of wh-noun clauses, zero-*that*-clauses, and of gerunds were significantly correlated with MLU.

Table 5

Partial correlations between grammatical structures and MLU (with df =27; 2-tailed)

	WNC	INC	TNCa	TNCz	gerund	WH-inf
Correlation	0.381	0.256	0.144	0.601	0.441	-0.087
Significance	0.042	0.180	0.455	0.010	0.017	0.653

Accuracy of Automated Analysis

The level of accuracy for the automated analyses of noun clause structures was assessed using the Kappa statistic, which relates the number of agreements between automated and manual analysis of both presence and absence of an item to the number of misses and false positives. Many researchers use the guidelines for Kappa interpretation published by Landis and Koch (1977) which rate Kappas from .61 to .81 as *substantial* and .82 to 1.00 as *almost perfect* (Boslaugh & Watters, 2008). Kappa levels are presented in Table 6.

Table 6

Kappa levels for Reno and Provo groups for each noun clause variety

	WNC	INC	TNCa	TNCz	gerund	WH-inf
RLI	0.771	0.823	0.259	0.570	0.290	1.000
RLA	0.723	0.587	0.398	0.547	0.298	0.914
RCA	0.599	0.615	0.409	0.538	0.655	0.875
Provo	0.685	0.363	0.390	0.625	0.558	0.928

In Table 6 it may be seen that Kappa levels ranged from 0.259 to 1.000 in size; the mean of these Kappa values was .556.

Another indicator of the accuracy of automated recognition of structures such as noun clauses may be obtained by examination of the rates of sensitivity and specificity. In the present context, sensitivity refers to the likelihood that the software would identify a noun clause that had been identified by the manual analysis, and specificity refers to the likelihood that the software would not identify a noun clause in an utterance that had not been identified by manual analysis. The percentage rates for specificity and sensitivity of the automated analysis are presented in Table 7. It may be seen in Table 7 that the levels of both sensitivity and specificity were high, averaging 93%.

Table 7

Sensitivity and Specificity percentage rates for the automated analysis of each noun clause variety

	WNC	INC	TNCt	TNCz	gerund	WH-inf
RLI						
Sensitivity	90	100	60	71	87	100
Specificity	99	100	100	99	98	100
RLA						
Sensitivity	86	83	86	78	89	84
Specificity	99	100	99	99	99	100
RCA						
Sensitivity	67	100	87	67	88	88
Specificity	99	99	99	99	99	100
Provo						
Sensitivity	75	100	90	85	71	90
Specificity	99	100	100	99	99	100

Discussion

The present study examined the frequencies of several varieties of noun clauses in two sets of clinical language samples and also assessed the accuracy of software for automated recognition of these noun clauses. Noun clause frequencies in samples of children with LI were compared with those of children who were similar either in language test scores or in chronological age. Frequencies of most varieties of noun clauses did not differ significantly among groups, perhaps because of the large variation among children of each group; only gerunds and wh-infinitive noun clauses differed significantly among groups. The frequencies of various noun clause varieties were correlated with the ages of a second group of children, but only the frequencies of zero-*that*-clauses and of gerunds were significantly correlated with age. The accuracy of automated recognition of noun clauses was high but imperfect both in sensitivity (identifying correctly when a noun clause was present in an utterance) and specificity (not falsely concluding that a noun clause was present).

The findings of the present study regarding differences between children with LI and typically developing children both corroborate and extend the findings of Nippold et al. (2007), who found no differences in frequency of finite noun clauses (as well as other finite embedded constructions) between groups of children with and without LI. The present study did find differences in the use of gerunds and wh-infinitive noun clauses between groups, two non-finite grammatical constructions which had not been assessed by Nippold et al. However, the patterns of group performance differed for these two constructions: The language-similar children used wh-infinitive noun clauses more frequently than did the chronological-age-matched children, whereas the age-matched children used gerunds much more frequently than did the other two groups. Nevertheless, the examination of additional grammatical structures in the present study suggests that certain structures might help distinguish children with language impairment from those without impairment. Further research into the development and use of non-finite noun clause constructions such as gerunds and wh-infinitive clauses might be of interest and perhaps clinical importance.

No previous studies had examined the age-related changes in frequency of noun clauses in clinical language samples. The present study's findings that older children tended to use gerunds and zero-that-noun clauses more frequently, even when controlling for sample length, is of interest. It should be noted that the constructions studied were quite sparse and infrequent, even though the developmental language samples studied all exceeded 190 child utterances, averaging 242 utterances per sample. Perhaps this is due to the conversational nature of the samples, and expository samples such as those of Nippold et al. (2007) would yield more data and allow additional relationships to be seen.

Though presented as an assessment technique and not a study of language development and impairment per se, the findings of Lee (1975) have relevance to the current findings. Lee examined a large variety of syntactic constructions in the development of Developmental Sentence Scoring, including both finite and non-finite forms. Those constructions which most clearly differentiated younger children (ages 2;6 - 3;0) from older children (ages 6;0 - 7;0) were assigned a higher point value on a 1 to 8 point scale. Utterances in a child's sample were then awarded a number of points based on the syntactic structures they contained, and the average number of points per utterances was called the Developmental Sentence Score. In the resulting scale that Lee developed, wh-infinitive noun clauses and gerunds were both given 8 points, the highest point value. Thus the present study corroborates Lee's claim that these two grammatical structures were of clinical and developmental importance, a claim which has not been further examined in the intervening decades.

An issue related to the use of frequency of these structures to attempt to separate groups and to show abge-related trends concerns the focus on frequency data itself. In looking at patterns of language development and language impairment, Bloom and Lahey (1978) proposed the use of criteria of productivity and emergence, rather than the use of raw frequency data, to evaluate language development. If a child used a construction 2 or 3 times in a sample, that construction was judged as emerging, and if used four or more times, the construction was judged as productive. Frequency beyond that point was more a function of conversational topic rather than linguistic development. Thus a future task might be to re-analyze the present data on the basis of emergence and productivity rather than frequency, either to distinguish among groups or to study the emergence of both finite and non-finite grammatical constructions in relation to age.

It might also be of interest to look at the occurrence of multiple embeddings used in an utterance rather than just the frequencies of specific types of embedding, as adult language is marked by multiple embeddings per utterance. Once a child has mastered construction and production of a grammatical structure, further development may be marked by the use of multiple embedded structures in an utterance rather than mere increases in the frequency of use of that grammatical structure. Relatedly, Diessel (2004) found that early uses of embedding were restricted to particular grammatical contexts, and thus examination of the variety of different grammatical contexts for embedding rather than the frequency of use of an embedded construction in any context would be of clinical and developmental interest.

As can be observed in the descriptive data of both the Reno and Provo samples (Tables 1 and 2), there was great variability even among children within the same group, making a significant difference between groups difficult to find. Likewise, those in the RLI group had been matched to RLA children based on language test score similarity. However, in terms of the noun clause structures of interest, performance between these two groups of participants was often quite different. This may be because the participants in the two groups had been matched based on norm-referenced test scores, and these scores were largely based on grammatical features other than the embedded clause constructions of interest in the present study. Embedding may be a rare item of these tests, and thus a prediction of grammatical similarity based on test scores is not corroborated by these childrens' spontaneous language performance.

Finally, in relation to the accuracy of automated identification of noun clauses in children's conversational language samples, previous evaluations of software for automated analysis of clinical language samples did not specifically give data on the accuracy of noun clause identification, but their accuracy on items or levels containing noun clauses as well as other embedded structures was quite low. Thus the findings of high levels of sensitivity and specificity for identification of noun clause varieties in utterances in children's clinical language samples suggests value in the further refinement of the approach to automated analysis used in the present study. However, the fluctuation in the Kappa levels for the analysis of the different varieties of noun clauses (ranging from 0.26 to 1.00 in size with a mean of .56) suggests that further improvement in the automated identification of these constructions is necessary prior to uncorrected clinical or research use of this software-based analysis. Some of these disagreements between manual and automated analysis may be due to mistagging by the Stanford parser; others may be due to misinterpretation by the cxs software of the output from the Stanford parser. This issue should be addressed in future applications and studies of the parser.

Nevertheless, the present study has contributed to greater understanding of the profile of noun clause use between children with LI and typically developing children, of the age- and MLU-related frequencies of noun clause use, and of the currently obtainable levels of accuracy for the automated identification of noun clauses.

References

Barber, M. (1989). Children's repetition of sentences previously produced spontaneously. Unpublished Master's Thesis, Brigham Young University.

Bloom, L., & Lahey, M. (1978). *Language development and language disorders*. New York: Wiley.

Bloom, L., Tackeff, J., & Lahey, M. (1991). Infinitive complements with *to*. In L. Bloom (Ed.), *Language development from two to three* (pp. 290-309). Cambridge, UK: Cambridge University Press.

Boslaugh, S., & Watters, P. A. (2006). *Statistics in a nutshell: A desktop quick reference.* Sebastopol, CA: O'Reilly.

Carrow-Woolfolk, E. (1985). *Test for Auditory Comprehension of Language (Rev. ed.)*. Allen, TX: DLM Teaching Resources.

Channell, R. W. (2003). Automated Developmental Sentence Scoring using Computerized Profiling software. *American Journal of Speech-Language Pathology, 10,* 180-188.

Crystal, D., Garman, M., & Fletcher P. (1989). *The grammatical analysis of language disability: A procedure for assessment and remediation* (2nd ed.). London: Cole and Whurr.

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

Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test* (Rev. Ed.). Circle Pines, MN: American Guidance Service.

Fujiki, M., Brinton, B., & Sonnenberg, E. A. (1990). Repair of overlapping speech in the conversations of specifically language-impaired and normally developing children. *Applied Psycholinguistics, 11,* 201-215.

Klein, D., & Manning, C. D., (2003). Accurate unlexicalized parsing. *Proceedings of the 41st Meeting of the Association for Computational Linguistics, 423-430.*

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159-174.

Lee, L. (1974). *Developmental sentence analysis*. Evanston: Northwestern University.

Limber, J. (1973). The genesis of complex sentences. In T. E. Moore (ed.) *Cognitive development and the acquisition of language*. Retrieved April 9, 2012, from http://pubpages.unh.edu/~jel/JLimber/Genesis_complex_sentences.pdf

Long, S. H. (1996). Why Johnny (or Joanne) can't parse. *American Journal of Speech Language Pathology, 5,* 35-42.

Long, S. H. (2001). About time: A comparison of computerised and manual procedures for grammatical and phonological analysis. *Clinical Linguistics and Phonetics, 15*, 399-426.

Long, S. H., & Channell, R. W. (2001). Accuracy of four language analysis procedures performed automatically. *American Journal of Speech-Language Pathology, 10*, 180-188.

Mecham, M. J., Jex, J. L., & Jones, J. D. (1967). *Utah Test of Language Development.* Salt Lake City, UT: Communication Research Associates.

Newcomer, P. L., & Hammill, D. D. (1982). *Test of language development-Primary.* Austin, TX: Pro-Ed.

Nippold, M. A. (2007). Later language development: School-age children, adolescents, and young adults. Austin, TX: Pro-Ed.

Nippold, M. A., Mansfield, T. C., Billow, J. L., & Tomblin, J. B. (2008). Expository discourse in adolescents with language impairments: Examining syntactic development [Electronic Version]. *American Journal of Speech Language Pathology*, *17*, 356-366.

O'Grady, W. (1997). *Syntactic development.* Chicago, IL: The University of Chicago Press.

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.

Paul, R. (1981). Analyzing complex sentence development. In J. F. Miller (Ed.), *Assessing language production in children: Experimental procedures* (pp. 36-40). Needham Heights, MA: Allyn & Bacon.

Scarborough, H. S. (1990). The index of productive syntax. *Applied Psycholinguistics, 11,* 1-22.

Schuele, C. M., & Dykes, J. C. (2005). Complex syntax acquisition: A longitudinal case study of a child with specific language impairment [Electronic Version]. *Clinical Linguistics & Phonetics, 19* (4), 295-318.

Schuele, C. & Tolbert, L. (2001) Omissions of obligatory relative markers in children with specific language impairment. *Clinical Linguistics and Phonetics* 15, 257-74.

Scott, C. M. (1988). Producing complex sentences. *Topics in Language Disorders, 8*(2), 44-62.

Semel, E. M., & Wiig, E. H. (1980) *Clinical Evaluation of Language Functions Screening Test.* Columbus, OH: Charles E. Merrill.

Semel-Mintz, E. M., & Wiig, E. H. (1982). *Clinical Evaluation of Language Functions Diagnostic Battery*. Columbus, OH: Charles E. Merrill.

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

Appendix A: Annotated Bibliography

Bloom, L., & Lahey, M. (1978). *Language development and language disorders*. New York: Wiley.

This book is an overview of both child language development and child language impairment. Of central interest to the present study is a discussion as to what clinical frequency data from naturalistic, spontaneous language samples means in terms of drawing conclusions regarding a child's competence for grammatical and semantic constructions. The notions of a criterion for emergence and for the productive use of a grammatical construction are presented and justified.

Bloom, L., Tackeff, J., & Lahey, M. (1991). Infinitive complements with *to*. In L. Bloom (Ed.), *Language development from two to three* (pp. 290-309). Cambridge, UK: Cambridge University Press.

This chapter presents in-depth information about the grammatical construction *to* and its use as a verb infinitive marker as opposed to the use of *to* as a preposition indicating motion toward something. Children learn to with the *verb + to* construction in the context of a small number of verbs that are subcategorized for taking infinitival complements. Semantically, the notion of "motion toward something" was similar both for verbs and the use of *to* as a preposition.

Channell, R. W. (2003). Automated Developmental Sentence Scoring Using Computerized Profiling Software. *American Journal of Speech-Language Pathology, 12,* 369-375.

Developmental sentence scoring (DSS) is a widely used measure for the quantification of expressive syntax development. Accuracy of fully automated DSS analysis using Computerized

Profiling (CP) software was evaluated using samples from 48 school-age children (28 with language impairment). Manual and automated analyses were compared, yielding a point-by-point agreement of 78% and correlation of *r* =.97. Per-category agreements ranged from 0% to 98% and agreement levels on samples from children with language impairment were lower by about 2%.

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

This examination of the development of complex sentences in early child speech is based on the hypotheses that the development of complex sentences progresses gradually from simple nonembedded sentences to multiple clause constructions, and that these early complex sentences become increasingly abstract. The progression to increasingly complex syntax was found to be influenced by the following determining factors: the frequency of complex syntactic construction in the child's ambient language, the complexity of these emerging constructions, the communicative functions of these complex sentences, and the social-cognitive development of the child. The hypotheses were tested using data from 5 English-speaking children ages 1;8 to 5;1, taken from the CHILDES database. Multiple-clause utterances were identified and analyzed, and findings indicate that the earliest multiple-clause utterances occurred around age two, which increased steadily in frequency until age 4;0. The earliest multiple-clause constructions were sentences including nonfinite subordinate clauses, followed by finite subordinate clauses. Multiple-clause utterances using two finite clauses were overall more frequent than multiple-clause utterances including a finite and a non-finite clause. Most infinitival and participial constructions were complement clauses. Finite complement clauses were the most frequent, while finite relative clauses were much less frequent. Conjoined clauses, or clauses consisting of either a finite adverbial clauses or co-ordinate clause, were also very frequent in the corpus. The author discusses a framework of development, that children's early complex sentences are organized around a formulaic matrix clause, a particular conjunction, or some other lexical expressions providing a frame for the rest of the utterance. This study shows that the development of complex sentences is characterized by both the steady increase of syntactic complexity, as well as becoming increasingly more abstract as the concrete constructions of early child are linked to abstract constructional schemas.

Limber, J. (1973). The genesis of complex sentences. In T. E. Moore (ed.) *Cognitive development and the acquisition of language*. Retrieved April 9, 2012, from http://pubpages.unh.edu/~jel/JLimber/Genesis_complex_sentences.pdf

In this book chapter, John Limber reports on and discusses his observations of the spontaneous production of complex sentences by English-speaking children under age 3. The subjects were a number of children in the Boston area between the ages of 1;6 and 3;0 who had been participating in a longitudinal developmental study of early language acquisition the past year and a half. The authors describe the stages of syntax acquisition in the terms of precomplex constructions, and complex constructions. Precomplex constructions included simple names and predicates, referential pronouns, and wh-questions. Complex constructions included complements (the earliest complex constructions were object complements (or nominals), Wh-Clause Constructions, and conjunctions. Limber concludes that (a) by age 3, children are able to generate syntactically complex names and descriptions: complements and relatives, allowing them to linguistically individuate a wide variety of abstract and concrete entities, (b) children's utterances display the basic structural features of English, with the exception of those aspects of syntax not present in the child's repertoire during this developmental period, (c) the major developments of complex syntax in the third year progress as follows: simple N-V-N sequences, expansions or substitutions of N-V-N sequence for noun phrases, and finally the conjoining of sentences.

Long, S. H., & Channell, R. W. (2001). Accuracy of four language analysis procedures performed automatically. *American Journal of Speech-Language Pathology, 10*, 180-188.

Computerized profiling uses probabilistic parsing algorithms to automatically identify grammatical word classes. It produces four types of clinically useful analyses: MLU, LARSP, IPSyn, and DSS. Accuracy of these measures was assessed on 69 language samples from individuals with typical language, speech impairment, and language impairment, ranging in age from 2 years 6 months to 7 years 10 months. Results from manual and automated coding were compared; correlation between these values was comparable to published data on interrater reliability of these procedures.

Nippold, M. A. (1993). Developmental markers in adolescent language: Syntax, semantics, and pragmatics. *Language, Speech, and Hearing Services in Schools, 24,* 21-28.

Marilyn Nippold discusses normal development during adolescence in the areas of syntax, semantics, and pragmatics in a comprehensive literature review. Syntax is discussed in the aspects of sentence length, subordination and cohesive devices. Nippold concludes that, although adolescents tend to have acquired a great deal of syntactic maturity by late adolescence, it should not be assumed that all aspects of syntax will be mastered by then or that speech will be free of syntactic errors. She points out that even well educated adults may have difficulty with low-frequency syntactic structures.

Nippold, M. A. (2007). Later language development: School-age children, adolescents, and young adults. Austin, TX: Pro-Ed.

According to Nippold, syntactic complexity in adolescents varies widely across genres such as conversation, narrative, persuasive, and expository, and should be studied in a variety of genres. She discusses, in general, the various ways in which sentence length increases over time, as adolescents mature. The increased use of subordinate clauses is one way in which sentence length increases in adolescents. These first appear in the speech of preschool children, but their frequency of use gradually increases in both spoken and written language throughout childhood and adolescence and into adulthood. Nominal clauses tend to be more frequent in spoken language, while adverbial clauses predominate in written discourse. Greater sentence length does not always imply greater sentence complexity, as certain types of verb phrases such as participial phrases, infinitive phrases, and gerund phrases, can suggest even greater syntactic sophistication when used in place of dependent clauses, because of their conciseness of expression. However, use of subordinate clauses is an important syntactic attainment.

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.

The authors explored the production of finite and nonfinite complement clauses in children with Language Impairment, ages 5;1–8;0, as compared to typically developing age- and vocabulary-matched children. The study included two experiments; in experiment one, simple finite and nonfinite complement clauses were elicited from the children through puppet show enactments, while in experiment two, finite and nonfinite complement clauses that required an additional argument were elicited from the children. The results were such that children in all three groups were more accurate in their use of nonfinite complement clauses than finite complement clauses, but the children with SLI were less proficient than both comparison groups. The SLI group was more likely than the typically developing groups to omit finiteness markers, the nonfinite particle *to*, arguments in finite complement clauses, and the optional complementizer *that*. The authors concluded that current theories of SLI need to be extended or altered to account for these results.

Paul, R. (1981). Analyzing complex sentence development. In J. F. Miller (Ed.), *Assessing language production in children: Experimental procedures* (pp. 36-40). Needham Heights, MA: Allyn & Bacon.

Rhea Paul outlines in this section of a developmental language textbook some milestones in development of complex sentence production. These are based on an analysis of previously obtained transcripts from 59 children between the ages of 2;5 and 6;11, engaged in 15-minute free-play sessions with their mothers, videotaped and transcribed by another researcher, who calculated MLU. The children were placed in five groups based on their MLU, and each transcript was analyzed for various forms of embedding, conjoining, and individual conjunctions. Proportions of subjects within each MLU group that used at least one form were calculated and analyzed for the presence of various forms of embedding, conjoining, and individual conjunctions. Upon calculating the proportions of each MLU grouping that used each form at least once, to derive stage placements for complex sentence constructions. Regression analysis on the data showed that MLU was a good predictor of complex sentence use than either age or cognitive level.

Rosenberg, S., & Abbeduto L. (1987). Indicators of linguistic competences in the peer group conversational behavior of mildly retarded adults. ***Applied Psycholinguistics, 8,*** **19-32.**

The authors provide an ordinal scaling system for classification of language complexity. They examined conversational language samples from adults classified with mild mental retardation for indications of mature linguistic competence, or specifically, grammatical morpheme and complex sentence use. The findings confirmed the authors' expectation that "the eventual level of mastery of these aspects of linguistic competence in adults classified with mild metal retardation is relatively high." The authors found no relationship between frequency of occurrence of individual complex sentence structures and a presumed order of acquisition. However, subjects' ability to combine complex sentences did "appear to be related to the presumed order of acquisition." The authors also found a negative correlation between relative frequency of complex sentence usage and estimated conversational communicative competence.

Scarborough, H. S. (1990). The index of productive syntax. ***Applied Psycholinguistics, 11,*** **1-22.**

The Index of Productive Syntax is a method of describing the grammatical complexity of preschool-age language, developed by Hollis S. Scarborough in 1990. Occurrences of 56 syntactic and morphological forms were counted and a total score and sub scores for noun phrases, verb phrases, questions/ negations, and sentence structures were obtained. Good content validity, strong age differentiation, and a strong correlation with MLU were demonstrated. Evidence for predictive validity for individual differences assessed by the IPSyn was also obtained. Limitations of the instrument are that IPSyn does not provide detailed diagnostic information about a child's mastery of particular syntactic or morphological rules; it can only serve as a rough indication of which aspects of performance should be analyzed in depth by other means. The scores published obtained through the first implementation of the measure are not norms, and when comparing individuals or groups, identical criteria for defining language samples must have been applied for comparisons between groups to be valid. IPSyn uses a types–based approach to evaluate grammatical types, rather than tokens and appears to be a reliable and potentially valid method of evaluating the syntactic and morphological complexity of preschool language corpora, and may be most useful if modified according to the needs of the researcher.

Schuele, C. M., & Dykes, J. C. (2005). Complex syntax acquisition: A longitudinal case study of a child with specific language impairment [Electronic Version]. ***Clinical Linguistics & Phonetics, 19*** **(4), 295-318.**

In order to better characterize the linguistic deficiencies evident in those with SLI, this article presents a longitudinal case study of a child with SLI from ages 3 to 7, examining the development of complex syntax during these years. The authors analyzed twelve conversational language samples for emergence of complex syntax types, proportional use of complex syntax, and complex syntax production errors. They found that the earliest emerging complex syntax types were catenatives and *let's* clauses, which are not always classified as complex syntax, as well as simple infinitives. These forms constituted 88% of complex syntax tokens through age 4;8 (MLU 3.12). At age 5;9 (MLU 4;27), a diverse range of complex syntax types was produced consistently, including wh-clausal complements, relative clauses, and full propositional clauses. Production errors noted on complex syntax included omissions of infinitival to, omissions of wh-pronouns in wh-clausal complements, omissions of obligatory relative markers, and an omission of a complementizer. These errors persisted into the latest language sample taken at age 7;10 (MLU 5.46).

Scott, C. M. (1988). Producing complex sentences. *Topics in Language Disorders, 8*(2), 44-62.

In this book chapter, Scott expounds on the aspects that characterize complex sentences and provides a structural framework for describing language complexity. Scott establishes that embedding and subordinating are one important aspect of complex syntax. The conditions under which the language is produced affect complexity. Scott discusses clausal complexity in the context of coordination and subordination, both of which are ways in which clauses to sentences. Nonclausal complexity is also discussed in terms of conjuncts and disjuncts, which introduce complexity by signaling meaning relations between sentences. Scott then discusses a developmental order of emergence of various complex syntactic structures.

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

The authors analyzed complex syntax in 110 children with typical language development, ages 1;8 to 4;9. The data analysis revealed a direct relationship between chronological age, mean length of utterance, and percent of complex sentences. Analysis of the complex sentences in each sample revealed subcategories for complexity type, which had a distinct order of acquisition. Furthermore, subcategories of certain types of complex sentences often did not appear until after other types of embedding.