Analysis of oral narratives of children who stutter and their fluent peers: Kindergarten through second grade

AMIT BAJAJ

Emerson College, Boston, MA, USA

(Received 24 June 2006; accepted 18 October 2006)

Abstract
Measures of language sample length (in c-units) and morphological, syntactic, and narrative abilities were obtained from oral narrative transcripts of 22 children who stutter and 22 children who do not stutter; participants attended kindergarten, first, and second grades. A two-way MANOVA yielded significant main effects for grade, with significant differences on some measures evidenced between participants in kindergarten and second grades. No significant differences between groups or group-grade interaction effects on the measures were obtained. Grade-wise comparisons (through t-tests) indicated that the performance of children who stutter did not differ significantly from their typically fluent peers on all dependent measures; however, kindergarten children who stutter obtained the most discrepant (lower) scores than their grade-matched fluent peers on the Narrative Scoring Scheme measure, with group differences approaching statistical significance on this measure. The findings suggest that children who do and do not stutter evidence similar expressive language abilities, even as subgroups of children who stutter may lag behind their grade-matched fluent peers in particular language domains.

Keywords: Oral narratives, stuttering, children, expressive language

Introduction
Research on the relationship between stuttering and language now has a 75-year history, much of which has involved children who stutter. Central to this research is the question of whether children who stutter possess the same language abilities as typically fluent children or whether they have language deficiencies that predispose them to stuttering. Per the demands-capacities model (see Adams, 1990; Starkweather & Gottwald, 1990, for elaboration), a leading etiological explanation for stuttering, it is worth considering whether children who stutter have inherently reduced language capacities so they exhibit aberrant speech motor responses (i.e. stuttering) to even typical language demands.

The position that the speech motor system may be rendered unstable due to language deficiencies is bolstered by research evidence demonstrating that premotor processes, such as language formulation, have systematic effects on the speech motor control system. For example, physiological studies indicate that when both fluent and dysfluent adults produce
longer and grammatically more complex sentences, their values on speech kinematic indices also rise (Kleinow & Smith, 2000; Maner, Smith, & Grayson, 2000). Results from several studies with children also have demonstrated that utterance length and grammatical complexity are positively associated with stuttering (see Bernstein Ratner, 1995, for review). Similar outcomes are reported in other studies that have examined the effect of syntactic complexity on speech fluency among children with typical language and those with language impairments (Colburn & Mysak, 1982a, b; Hall, Yamashita, & Aram, 1993).

While most research literature has supported a positive relationship between syntactic complexity and speech fluency, the status of language abilities among children who stutter remains unclear. Empirical outcomes have been divergent in that, relative to typically fluent children, children who stutter have evidenced language abilities that are less developed (Byrd & Cooper, 1989; Silverman & Bernstein Ratner, 2002), comparable (Nippold, Schwartz, & Jescheniak, 1991; Howell, Davis, & Au-Yeung, 2003) or exceeded developmental expectations (Watkins, Yairi, & Ambrose, 1999). Watkins and Johnson (2004) submitted that significant methodological differences between studies and failure to incorporate control over key variables that impact language development account for the current ambiguity in this area. Along with reports of considerable heterogeneity in language skills among children who stutter, there is ample evidence that some children who stutter have additional phonological and language disorders (see Nippold, 2004, for discussion). While estimates of such concomitant communication disorders vary widely (e.g. Blood & Seider, 1981; Arndt & Healey, 2001), the robustness of this finding attests that children who evidence communication disorders concomitant with stuttering constitute a subgroup in the population.

Recent forays in assessing language performance among children who stutter signal a shift in research direction. It has been suggested that language abilities among children who stutter should be studied using tasks that incorporate subtle language measures and target individuals’ performance in specific domains of language competence (Bernstein Ratner, 1997, for discussion; Watkins et al. 1999, for research suggestions). Recent studies have used such tasks as non-word repetition (Hakim & Bernstein Ratner, 2004; Anderson, Wagovich, & Hall, 2006); metalinguistic tasks, such as phoneme reversal and grammatical judgments (Bajaj, Hodson, & Schommer-Aikins, 2004); and experiments utilizing priming methodologies to examine language encoding abilities among children who stutter (Melnick, Conture, & Ohde, 2003; Anderson & Conture, 2004; Pellowski & Conture, 2005). Studies using such tasks and experiments, which are commonly used to examine working memory abilities, have generally reported significant group differences, with typically fluent children outperforming children who stutter on some selected measures (see Melnick et al., 2003, and Bajaj et al., 2004, for partial exceptions). Such results are noteworthy because the subtle language measures used in these studies allow positing a potential link between working memory and stuttering.

Connected speech samples have served as other means for obtaining a wide array of language measures upon which the performance of children who stutter can be examined. In this context, there is considerable value in eliciting connected speech samples through narrative tasks, given that children’s narratives serve as indices of their abilities in multiple developmental domains, such as comprehension of ideas, expressive language abilities, literacy skills, and overall intellectual and emotional states (Engel, 1995).

Oral narratives generated by children who stutter are of interest in the current study for several reasons. First, narratives provide a rich mix of data on grammatical complexity and content organization that is not readily obtained from structured language tasks in
standardized tests or other connected speech samples. Performance on a variety of measures can be examined through two kinds of narrative analyses: microstructural and macrostructural (Hughes, McGillivray, & Schmidek, 1997). Microstructural analysis includes the study of cohesive devices and other linguistic aspects, such as tense markers, vocabulary, and sentence complexity. Macrostructural analyses relate to the ways in which story content is organized in language output, such as the use of story grammar components in the narrative task. The description and distinctions between the two types of analyses is maintained throughout this manuscript.

The second reason to examine narrative abilities of children who stutter is that narrative analyses have been found to be clinically useful for several populations. For example, narrative skills evidenced through story retelling tasks are considered robust predictors of persistent language disorders among children (e.g. Bishop & Edmundson, 1987; Paul & Smith, 1993). Additionally, there are several available resources to assess the level of narrative skills among typical and language-disordered school-age children (e.g. Strong, 1998; Justice et al., 2006) or those from culturally-linguistically diverse backgrounds (e.g. Gutierrez-Clellen & Quinn, 1993; Lofranco, Pena, & Bedore, 2006). The development of narrative analysis methodologies and their use in assessing the language difficulties of language-disordered children supports the rationale for using narrative analyses to examine the language skills of children who stutter.

Third, various kinds of narratives (e.g. accounts, recounts, event casts and stories; see Heath, 1986) appear naturally in most children’s communicative environments; they are likely to hear them when interacting with other people or produce them spontaneously. Therefore, ‘narratives have ecological validity’ (Hughes et al., 1997, p. 7) unlike standardized language tasks that tend to be more contrived in nature. From a demands-capacities perspective, narratives offer the opportunity to explore how children who stutter formulate and use language in real communicative situations. In the process, narrative tasks enable researchers to study how everyday language demands might challenge speech motor capacities, while making it possible to analyse children’s performance on a wide range of language measures.

A limited number of studies on children who stutter have used narrative tasks. Two studies involving such children have utilized narrative samples to obtain speech data that was used for comparing the number and types of disfluencies between groups (Trautman, Healey, & Norris, 2001; Boscolo, Bernstein Ratner, & Rescorla, 2002). Some studies, reviewed here because of their relevance to the current research questions, have addressed differences between fluent children and children who stutter on various language measures derived from narrative tasks.

An initial study by Nippold et al. (1991) examined story length, syntactic complexity, story grammar components, and story comprehension in a story retelling task and reported no significant group differences (10 children who do and 10 who do not stutter) on any of the measures. Weiss and Zebrowski (1994), in a second investigation, compared story retelling abilities of eight children who stutter and eight fluent peers under two conditions: story retelling to familiar listeners and naïve listeners. Narrative transcripts were analysed for story grammar elements, c-units, and maze behaviour. Although no significant group differences were obtained on these measures, the authors reported qualitative differences between groups on the length and detail included in the stories; the stuttering group generated stories that were shorter and less detailed than their fluent peers when retelling them to naïve listeners. The authors suggested that this qualitative outcome may signal differences in the pragmatic abilities of the groups.
Scott, Healey, and Norris (1995) also examined narrative skills of six children who stutter and six fluent children under two conditions: story retelling and story generation. Narrative samples under both conditions were analysed for t-units, attempted story grammar episodes, and grammatical cohesion indices (various conjunctions and pronoun references). In addition, stuttering frequencies of children in the stuttering group were compared across the two narrative elicitation tasks. Although no significant group differences were obtained on the language measures, significantly greater stuttering counts were obtained from the stuttering group during the story retell condition.

The aforementioned studies on narrative abilities of children who stutter differ from each other in language elicitation procedures, participant characteristics, and some measures of language abilities. Further, all such studies have the following limitations: small samples, which reduce the ability to detect subtle disturbances and limit the external validity of results (Maxwell & Satake, 2005); few target-cohort group matching criteria, raising questions about the comparability of participant groups; and unaccounted age or grade differences that potentially relate to the narrative abilities of the participant groups. Given such limitations in methodology, findings pertaining to the narrative abilities of children who stutter remain in need of further scrutiny.

The purpose of the current study is to draw upon narrative methodology to examine the expressive language abilities of children who stutter, using measures and analyses that are different from previous studies on this topic. Thus, the current study differs from previous studies on this topic in a number of ways. First, in the current study, both overall group differences and group differences per grade-level were examined. The sample size (44 participants) was adequate for segmenting the participants in three grade-based groups so that grade-wise differences in the groups’ performance could be analysed.

Analysing the language performance of children who stutter at different grade levels is pertinent, given evidence that school experiences influence language learning among children (e.g. Wadsworth, 1986) and certain curricular experiences, shaped by teacher-child interactions, impact children’s discourse development (Dickinson, 1991). For example, children’s literacy experiences are noted to influence their narrative styles (Michaels, 1981), such as the extent of elaboration included in their personal narratives (Minami & McCabe, 1991). From a demands-capacities perspective, it is worth probing whether elementary school children who stutter in lower and higher grades demonstrate different language abilities relative to their grade-matched fluent peers in a demanding language formulation task.

The second way in which the present study differs from previous work is that three of the four dependent measures used in this study have not been examined in previous studies on the narrative performance of children who stutter. They include percentage accuracy of tense marking, subordination index, and composite scores obtained on the Narrative Scoring Scheme (all dependent measures are described under Method). All such measures relate to the adequacy of underlying language competence and story organization among participants. The number of c-units, the fourth dependent measure, was included in the current study to verify the results of previous studies with a larger sample. That said, it should be noted that other story grammar elements in the narrative productions of children who stutter have been examined previously (see Nippold et al., 1991; Weiss & Zebrowski, 1994; Scott et al., 1995). Also, overall measures of grammatical complexity and tense-related morphological indices have been examined from conversational samples of children who stutter (Watkins et al., 1999)
Finally, the elicitation procedures used here, story generation from a wordless picture book, have not been utilized in previous studies on narrative abilities of children who stutter. It is widely reported that stimulus properties and elicitation procedures impact the depth and breadth of language structures that are obtained; for example, utterance length varies according to the types of samples elicited (O’Donnell, Griffin, & Norris, 1967; Loban, 1976; Klecan-Aker & Lopez, 1985). Among narrative elicitation techniques, story retellings are considered less challenging than story generation procedures; this is reported in the added length and sophistication in story grammar elements that are evidenced in story generation, in contrast to story retelling tasks (Meritt & Liles, 1987). Wordless picture books have been found to be appropriate for eliciting narratives that are long, informative, and complex, relative to other stimuli, such as single-scene descriptions (Pearce, 2003).

The following research questions are addressed in this study:

1. Does the performance of children who stutter on language sample length (in c-units) and morphological, syntactic, and story grammar measures differ significantly from their typically fluent cohorts? Additionally, is there a group (children who do and do not stutter) and grade (kindergarten, first, and second) interaction for the dependent measures?

2. Does group performance differ significantly when group comparisons on all dependent measures are made at each grade level?

**Method**

**Participants**

Forty-four Caucasian boys between the ages 5;10 and 8;10 (years;months), participated in this study, 22 of whom were children who stutter (CWS) and 22 were children who do not stutter (CWNS). There were 10 children in kindergarten (age range 5;10 to 7;6), 24 in first grade (age range 6;8 to 8;3), and 10 in second grade (age range 7;8 to 8;10), with an equal number of CWS and CWNS in each grade. Table I presents the age ranges and average age of participants in each group according to their grade levels. Speech-language pathologists (SLPs) from 19 Midwestern elementary schools in the USA referred the children on their caseload for this study. The SLPs consulted the classroom teachers to refer grade-matched typically fluent children as well. All data were collected during the second semester of the school year. By teacher and school SLP reports, participants selected for this study met the following criteria: (a) no histories of organic anomalies, neurological conditions or hearing deficits (this was confirmed during testing), (b) performance at or above grade level in reading, and (c) English as their native language.

<table>
<thead>
<tr>
<th>Grade</th>
<th>CWS Age Range</th>
<th>CWS Average Age</th>
<th>CWNS Age Range</th>
<th>CWNS Average Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kindergarten</td>
<td>5;10–7;6</td>
<td>6;3</td>
<td>6;0–6;9</td>
<td>6;4</td>
</tr>
<tr>
<td>First</td>
<td>6;8–8;3</td>
<td>7;3</td>
<td>6;8–7;11</td>
<td>7;1</td>
</tr>
<tr>
<td>Second</td>
<td>7;8–8;9</td>
<td>8;4</td>
<td>7;9–8;10</td>
<td>8;3</td>
</tr>
</tbody>
</table>

Table I. Age ranges and average age of CWS and CWNS in kindergarten, first, and second grades.
Children who stutter (CWS)

Individualized Education Plans, which documented treatment services for stuttering, were available for 21 of the 22 CWS; the remaining one participant was referred before his treatment plan was completed, but was confirmed by the examiner to present stuttered speech. Based on the Stuttering Severity Instrument-Revised (Riley, 1986), four participants were rated as “very mild”, nine were rated at “mild”, six as “moderate”, and three as “severe”. Such severity ratings are provided only to better characterize the sample, as they are based on one speech sample per participant. They should be interpreted judiciously because stuttering behaviour varies across persons and situations (see Bloodstein, 1995).

Of the 22 CWS, four received therapy for concomitant speech-sound disorders, one was being monitored for speech-sound problems, two received concomitant services for both speech-sound and language problems, and one had a history of speech-sound problems that had normalized. CWS with such concomitant concerns were included as participants for two reasons. First, given that multiple studies provide evidence that some CWS tend to evidence speech-sound disorders and language difficulties (see Nippold, 2004), including such subgroups in the participant pool better represents the overall population of CWS than constituting “pure” samples where such subgroups are excluded. Second, SLP reports indicated that in each of the cases with concomitant speech or language problems, stuttering was the major concern due to its severity, perceived impact on functioning or parental concerns and that the concomitant speech-sound or language concerns were minor in comparison.

Children who do not stutter (CWNS)

One grade-matched peer was selected by the teacher and SLP from the same classrooms as a CWS. This cohort selection criterion was met for 21 of the 22 CWNS; for one kindergarten level CWS who was home schooled, a CWNS peer from an area kindergarten class was selected. Drawing grade-matched pairs from the same classrooms permitted some degree of control over classroom experiences that may have affected the children’s performance on the measures examined in the narrative tasks. In selecting a CWNS peer, teachers and SLPs were asked to select a child whose academic performance was comparable to that of the CWS, whose chronological age was within six months of the CWS, and who presented no speech, language or academic problems.

It should be noted that participant selection criteria provided a reasonable basis for matching the groups on speech and language skills so that the presence of stuttering could serve as a focal basis for group comparison. First, although eight out of 22 CWS evidenced current or past articulation or language concerns, the referring SLPs had adjudged that such concerns did not interfere with their academic performance. In other words, these children were performing at or above grade levels like their classroom-matched typically fluent peers selected for this study. Recall that the rationale for including a subgroup of CWS with concomitant speech and language concerns was to constitute the sample that would adequately represent the population of CWS where such subgroups are well reported. Second, given that all participants were performing at academic levels adjudged as adequate or better, both groups likely demonstrated a range of language based abilities that varied within normal limits; an equal number of participants in both groups (22 per group) and at each grade level further bolsters this likelihood.
Data collection procedures

Data for this study were collected as part of a larger investigation on metalinguistic skills of children who stutter, published elsewhere (Bajaj et al., 2004; Bajaj, Hodson, & Westby, 2005). In the overall testing sequence, the narrative task reported in the current study was administered early; story telling samples were elicited immediately after engaging participants in a short rapport-building conversation and before other tasks were administered.

A wordless picture book, *Frog Goes to Dinner* (Mayer, 1974), was used to obtain story telling samples. The visual stimuli in this book are line-drawn picture sequences that present the adventures of an errant frog (a boy’s pet) who hides in the boy’s pocket and arrives at a restaurant. The book depicts the frog’s misadventures (it gets into a band member’s saxophone, into someone’s food, etc.) until it is finally apprehended and returned to its rightful owner.

The narrative task was administered by the examiner in the following manner. The examiner, who sat next to each participant, gave him the wordless picture book to browse. While the participant was browsing the book, the examiner told him that wordless picture books contain no words but they do have a real story to tell. The story can be told by examining the details of the pictures on each page and talking about them in a way that makes careful connections with preceding pages. Following this introduction, the examiner narrated a “model” story unfolding on the first five pages of the book. The script of the model story had been memorized and its production rehearsed by the examiner so that each participant was provided the model in the same manner.

After providing the participant with the model story, the participant was told that it was his turn to tell the story from the beginning until the end of the book. Once participants began telling the story, the examiner was silent except for (a) occasional verbally-limited responses (“yes”, “that’s right”) to encourage participants’ verbal output, and (b) prompting participants to elaborate (“tell me more”, “what else is going on here?”), when they generated minimal stories.

The story telling samples were audio recorded. Participants’ audio recordings were transcribed by a graduate student to render nearly verbatim texts of their stories; filled pauses and repaired utterances were included in the transcription, but inflectional patterns and stuttered speech (such as, repetitions: f-f-f-frog) were excluded. Another graduate student cross-checked the transcripts to establish an exact recording-transcript match. Finally, the author spot checked the accuracy of 22 (of 44) transcripts and achieved 100% inter-rater reliability with the graduate student who provided the most recent corrections. Ultimately, 44 story telling transcripts constituted the verbal data for this study.

Descriptions of dependent measures

Microstructural measures.

Overall length of sample: Number of communication (c)-units. The overall narrative sample length was measured by segmenting the language sample according to c-units and counting the total number of c-units in the sample. C-units (Loban, 1976) are either main/independent clauses (e.g. “John ate an apple”) or main clauses with their modifiers (e.g. “John ate an apple that was rotten”). C-unit segmentation and analysis is a common procedure for examining language transcripts and has been utilized in previous studies on narrative abilities of CWS (e.g. Nippold et al., 1991; Weiss & Zebrowski, 1994).
Morphological: Percentage accuracy of tense marking. This measure relates to the accuracy of finite-tense marking and the use of auxiliary and copula verbs. To summarize, English verbs are marked for tense through the use of bound morphemes, such as, the third person singular -3s (e.g. he jumps, he eats) or past tense –ed (e.g. he jumped); past tense is marked also with irregular forms for some verbs (e.g. he ate). Together with the use of auxiliary and copula forms of BE and DO verbs (e.g. he is jumping/he does jump and he is sad, respectively), which follow similar generative rules, tense marking accuracy has been studied as an index of morphological development (Dethorne, Johnson, & Loeb, 2005) that is incorporated in the grammatical repertoire of most typically developing children by the time their language conforms to adult models (Rice, Wexler, & Hershberger, 1998). Accordingly, its development has been examined as a clinical marker for specific language impairment (SLI) among children (Rice, Haney, & Wexler, 1998; Rice & Wexler, 1996; Leonard, Miller, & Gerber, 1999; Pearce, McCormack, & James, 2003), given that “a clinical marker would be one that would show variation across children where no variation is expected” (Rice & Wexler, 1996, p. 1240) once language abilities approximate adult standards. The children in the current sample were not diagnosed as having SLI, and they were in the age range (5;10 to 8;10) in which little variation is expected in the use of tenses. Tense marking was, therefore, utilized as a sensitivity measure for examining whether subtle differences exist in the morphological abilities of CWS and CWNS.

Following the conventions suggested by Rice and Wexler (1996), four tense and verb forms were included in the tense marking measure: past tense regular (-ed); third person singular (henceforth, -3s); and auxiliary and copula forms of BE and DO verbs. To compute this measure for each transcript, the total number of these tense and verb forms used correctly was added. That total number of correct uses was divided by the number of obligatory contexts in which those forms appeared to yield a percentage accuracy of tense marking.

It should be noted that while it is possible that dialectal differences can influence tense marking performance, selecting participants from a limited Midwestern region in United States diminished the potential influence of dialect on tense marking ability in the current sample. Further, it is improbable that articulation concerns among some CWS would have impacted tense marking performance to a significant extent. Two CWS had a history of articulation concerns that had normalized at the time of data collection, and SLPs had adjudged the current articulation difficulties among the remaining six CWS as ‘mild’ in relation to their stuttering. Also, the relatively minor impact of such difficulties on the intelligibility of the six CWS was noted during data collection.

Syntactic: Subordination Index. Subordination Index (SI) relates to clause density in language samples, that is, the extent to which subordinate (or dependent) clauses are included in the utterances/sentences. It has been used as an overall measure of syntactic complexity in the language samples obtained from various speakers, including English-speaking children (Klecan-Aker & Lopez, 1985) and Spanish-speaking preschoolers (Gutierrez-Clellen & Hofstetter, 1994). Here, the SI was computed for each transcript by first identifying the number of clauses (main and other types) embedded in each c-unit, then summing the number of clauses across c-units to yield a total number, and finally, dividing that total number by the number of c-units in the transcript.

Macrostructural measure: Narrative Scoring Scheme (NSS) scores. Attributed to Miller and Heilmann (2004) and based on story grammar components developed by Stein and Glenn
(1979, 1982), the NSS is a recently developed measure of the extent to which each of following seven story grammar components is incorporated in the narrative: Introduction, Character Development, Mental States, Referencing, Conflict/Resolution, Cohesion, and Conclusion. A summary description of these components is provided in Appendix A.

In the NSS scoring system, performance on each story grammar component is assigned a value between 0 and 5. "Proficient" use of the category receives 5 points, "emerging/inconsistent" use receives 3 points, and 1 point is assigned to "immature/minimal use". Zero points are assigned to "poor performance", which includes incomplete or unintelligible utterances, incorrect story productions, story productions in a non-target language, inability to perform the task, or imitations of the NSS target components. Intermediate performance, between 5 and 3 and 3 and 1, is assigned 4 and 2 points, respectively. A narrative transcript, provided in Appendix B, includes the NSS scores for individual story grammar categories as well as a composite NSS score, which is the sum of individual category scores. NSS composite scores were used in the data analysis for this study.

Data analysis and reliability

The 44 oral narrative transcripts were converted to text files, which were coded using conventions of the Systematic Analysis of Language Transcripts (SALT v. 9; Miller & Chapman, 1984–2006), a computerized language analysis program. SALT encoding of the transcripts was performed in the Language Analysis Laboratory at the University of Wisconsin-Madison. The laboratory has been involved in the collection, transcription, and analysis of language samples since 1983 and has generated over 3000 SALT-format language transcripts in the past 3 years. A team of four personnel, each with at least 2 years of experience in transcription, converted the orthographic transcripts to a SALT format for this project.

The author assigned an alphanumeric identification number to each participant. This number did not disclose the group (CWS or CWNS) to which the participants belonged, so that the SALT coding personnel remained blind to participants’ group status, thereby reducing the chance of potential coding biases.

Per the transcription and coding protocols observed at the Language Analysis Laboratory, one person computed the SI scores; therefore, no inter-rater reliability was computed for this measure. Given the level of training and experience among the personnel at the laboratory, there were sufficient reasons to expect that objectivity and accuracy were maintained in the SI scoring task. The same applies to segmenting the transcripts into C-units.

NSS scoring was carried out independently by two laboratory personnel: an initial scorer and a second scorer. After both sets of scores were obtained, the initial scorer checked both scorers’ tallies for all NSS scores. When inter-scorer differences of less than one point were discovered, the initial scorer’s counts of the story grammar scores superseded those of the second scorer. For inter-scorer discrepancies of greater than one point on any story grammar category, the two scorers caucused to review the details of the scoring rubric and reached an agreement on the final score.

After all transcripts were coded in the SALT format, the tense marking measure was computed manually by the author and a colleague with considerable research experience in child language disorders. The examiners first reached agreement on the type of tense marking forms to be identified and the procedure for estimating them from bound
morpheme tables and word root tables, generated as part of the SALT standard outputs. Both examiners then analysed independently the free and bound morpheme summaries to compute the percentage accuracy of obligatory use of tense markers by all participants; the second examiner was blind to the group status of participants. A 98% inter-examiner agreement was obtained on the percentage accuracy of tense marking.

Results

The research questions posed in this study were examined statistically using the Statistical Package for the Social Sciences (version 12.0) program. Findings relating to each question are presented next.

1. Does the performance of children who stutter on language sample length (in c-units) and morphological, syntactic, and story grammar measures differ from their typically fluent cohorts? Additionally, is there a group (children who do and do not stutter) and grade (kindergarten to second grade) interaction for any of the dependent measures?

Table II summarizes the group means and standard deviations for all dependent measures for each group. Homogeneity of variances for all dependent measures, obtained through the Levene’s Test, confirmed the appropriateness of using parametric statistical techniques for analysing the data. The footnote for Table II shows that Levene’s test F values for all dependent variables were greater than .05.

A two-way multiple analysis of variance (MANOVA) was computed to examine main effects for group and grade, and interaction effects between group and grade for all dependent variables. Table III presents the MANOVA results. The results suggested no significant group differences (i.e. no main effects for group) on any of the dependent measures (c-units: F=.31, p>.05; percentage accuracy of tense marking: F=1.60, p>.05; SI: F=2.81, p>.05; NSS scores: F=.58, p>.05).

Main effects for grade were obtained for c-units (F=3.54, p<.05) and SI (F=5.49, p<.05) only, suggesting a significant difference between the grades (kindergarten through second) on the participants’ performance on the number of C-units and SI scores. The significant main effects required using a post-hoc test to locate the differences in performance between specific grades. Tukey’s post-hoc test indicated that significant grade differences for both c-units and SI scores were found between kindergarten and second grades (both p<.05), a predictable developmental difference in language sample length and

<table>
<thead>
<tr>
<th>Dependent Measures</th>
<th>CWS Group</th>
<th>CWNS Group</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td></td>
</tr>
<tr>
<td>Microstructural Measures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-Units</td>
<td>36.50</td>
<td>13.85</td>
<td>33.00</td>
</tr>
<tr>
<td>Percentage Accuracy of Tense Marking</td>
<td>93.09</td>
<td>10.78</td>
<td>97.64</td>
</tr>
<tr>
<td>Subordination Index (SI)</td>
<td>1.20</td>
<td>.15</td>
<td>1.26</td>
</tr>
<tr>
<td>Macrostructural Measure: Narrative Scoring Scheme (NSS)</td>
<td>16.59</td>
<td>4.14</td>
<td>17.27</td>
</tr>
</tbody>
</table>

Note: Levene’s Test of Equality of Error Variances F values: C-units=0.96; Percentage Accuracy of Tense Marking=1.03; SI=.92; NSS=2.01. p>.05 for all measures.
syntactic complexity between lower and higher elementary grade children. As Table III shows, there was no significant interaction effect between grades and groups on any of the dependent measures.

2. Does group performance differ significantly when group comparisons on all dependent measures are made at each grade level?

Grade-wise comparisons were made to examine whether group performance differed significantly at each grade level. Grade-wise group comparisons were made because fluent peers were selected by grade, not age (mean age difference between the pairs was six months). Also, a case can be made for the role of curricular influences on the development of discourse production (e.g. Michaels, 1981; Dickinson, 1991; Minami & McCabe, 1991), justifying selecting grade-wise comparisons as the unit of analysis.

When participants were categorized into their respective grades—kindergarten, first, second—the sample size per grade was reduced. To note, there were 10 participants attending kindergarten, 24 attending first grade and 10 in second grades. There were equal numbers of participants in the CWS and CWNS group at each grade level. In order to examine group differences on the four dependent variables, four t-tests were run per grade. It has been noted that conducting multiple t-tests to discern group differences increases the probability of a Type I error, i.e. obtaining statistically significant results by chance; the proposed solution is to revise the alpha ($\alpha$) value following Sidak’s or Bonferroni’s adjustment procedures (Sankoh, Huque, & Dubey, 1997). Accordingly, the statistical significance or alpha level was adjusted from its customary value of .05 to .0125 using the Bonferroni adjustment statistic (see Sankoh et al., 1997, for procedure). Tables IV, V, and VI present means, standard deviations, and t-test results for group differences for kindergarten, first, and second grades, respectively. Levene’s test of equality of variances indicated homogeneity of variances for all dependent measures for all grade levels. Results indicate that the groups did not differ significantly on any dependent measure at each grade level; however, as noted in Table IV, most non-discrepant group performance was evidenced in NSS scores in kindergarten, where NSS scores obtained by CWS in

### Table III. Summary of two-way multiple analysis of variance (MANOVA) with group and grade main effects and group-grade interaction effects.

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-units</td>
<td>1</td>
<td>.31</td>
<td>.58</td>
</tr>
<tr>
<td>Percentage Tense Accuracy</td>
<td>1</td>
<td>1.60</td>
<td>.21</td>
</tr>
<tr>
<td>Subordination Index</td>
<td>1</td>
<td>2.81</td>
<td>.10</td>
</tr>
<tr>
<td>Narrative Scoring Scheme</td>
<td>1</td>
<td>.58</td>
<td>.45</td>
</tr>
<tr>
<td>Grade</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-units</td>
<td>2</td>
<td>3.54</td>
<td>.04*</td>
</tr>
<tr>
<td>Percentage Tense Accuracy</td>
<td>2</td>
<td>.64</td>
<td>.54</td>
</tr>
<tr>
<td>Subordination Index</td>
<td>2</td>
<td>5.49</td>
<td>.01*</td>
</tr>
<tr>
<td>Narrative Scoring Scheme</td>
<td>2</td>
<td>1.97</td>
<td>.15</td>
</tr>
<tr>
<td>Group* Grade</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-units</td>
<td>2</td>
<td>.83</td>
<td>.45</td>
</tr>
<tr>
<td>Percentage Tense Accuracy</td>
<td>2</td>
<td>.76</td>
<td>.48</td>
</tr>
<tr>
<td>Subordination Index</td>
<td>2</td>
<td>.62</td>
<td>.54</td>
</tr>
<tr>
<td>Narrative Scoring Scheme</td>
<td>2</td>
<td>1.24</td>
<td>.30</td>
</tr>
</tbody>
</table>

$z = .05$; *$p < .05$.

Note: Tukey’s post-hoc test yielded the following significant $p$ values ($z = .05$) between kindergarten and second grades: c-units $p = .03$; Subordination Index $p = .02$. 
Kindergarten (M=13.80, SD=1.92) were considerably lower (t(8)=-2.74, p=.03) than those obtained by their fluent peers in kindergarten (M=18.20, SD=3.03).

**Discussion**

Research on language abilities of CWS indicates that although some studies have reported significant differences between CWS and CWNS on syntactic measures obtained from spontaneous language samples, this finding has not been universally supported (see
Bernstein Ratner, 1997, for discussion). Current results converge with findings from other studies that have reported adequate performance of CWS on language measures drawn from narrative tasks. A significant grade-based (kindergarten and second grades) difference in participants’ performance on two measures—c-units and SI scores—was expected, given that these measures are indexed to developmental changes in language abilities across grade levels (see Scott, 1988) and that children in second grade were, on average, about 2 years older than those in kindergarten (see Table I).

A grade-wise analysis of performance on the dependent measures provides additional insights into the differences in groups’ performance. It should be noted that there were sufficient reasons to expect parity between the groups at each of the grade levels, given the following aspects of this study’s methodology: (a) CWS and CWNS pairs were selected from the same classrooms at each grade level (barring one exception, where a homeschooled CWS was matched with his age-matched fluent peer in kindergarten); (b) all data were collected during a 3-month period, when participants were in their final semesters at each grade level; and (c) all participants were deemed average to above-average performers in reading and academic domains at each grade level.

It is noteworthy, therefore, that the most discrepant performance among the four dependent measures was evidenced among Kindergarten CWS on the NSS measure, with group differences approaching statistical significance ($p = .03$; Table IV) as CWNS outperformed CWS on this measure. Although this finding is preliminary because it is based on comparisons between 10 participants (five CWS and five CWNS), it provides evidence that differences in language abilities of CWS and CWNS may manifest among particular grades. It is possible that CWS in lower grades, such as kindergarten, are less likely to avail curricular opportunities for discourse development than their grade-matched fluent peers. The lack of significant group differences on the NSS scores between the same numbers of participants in the second grade attests to this possibility.

It is worth pondering whether, grammatical ability notwithstanding, the responses of kindergarten CWS that manifested in lower scores on the NSS measure reflect a limited willingness to speak. Overall, though non-significant, group differences in c-units for this grade level (CWS mean = 24.0; CWNS mean = 29.4) support this possibility. To the extent that limitations in verbal output are suggestive of speakers’ reluctance to respond, the performance of kindergarten CWS allows speculation about their self-perceptions as communicators. Studies on communicative attitudes of children who stutter have reported that such children tend to evidence negative attitudes towards their own communication (e.g. Vanryckeghem & Brutten, 1996; Vanryckeghem, Hylebos, Brutten, & Peleman, 2001). Indeed, the current participant set demonstrated a divergence in the groups’ responses concerning positive and negative speaker attributions and self-perceptions as communicators (Bajaj et al., 2005). Whether kindergarten CWS felt constrained in elaborating their narrative responses due to their speech impairment remains unknown; however, readers are alerted to the possibility that competencies and psychological dispositions may interact significantly to affect children’s responses.

Although no significant group difference among first graders on the percentage accuracy of tense marking was obtained, statistical significance was most closely approached on this measure ($p = .09$; Table V) in contrast to other measures. CWNS outperformed CWS, prompting a scrutiny of individual scores of CWS on this measure. Observation of individual percentage accuracy scores from CWS participants revealed that three participants in this group obtained tense marking accuracies that were under
90% (i.e. individual accuracies were 50%, 89%, 89%), where 90% accuracy in obligatory contexts represents a significant benchmark for acquisition of particular grammatical features (see Ingram, 1989, for discussion). All CWNS obtained accuracies of 90% and above.

It is noteworthy that the SLPs and teachers, who referred participants for this study, used their judgment to refer participants who were performing at or above grade level on reading and academic skills. Where speech or language concerns were present, the children were selected only when those concerns were adjudged minor in comparison to stuttering. It is noteworthy, therefore, that the subgroup of participants who did not meet the benchmark of 90% accuracy for tense marking emerged in the CWS group only. The subgroup was more likely to emerge in relatively expanded participant set, so it is not surprising that the outliers belonged to the first grade where the highest number of CWS (12 participants) were included.

Overall results of this study suggest that although within-group variations in the performance of CWS can be striking, such children are likely to demonstrate expressive language abilities comparable to their typically fluent peers. Current results and findings from previous studies on narrative abilities, where related measures were used, collectively support the view that “although stuttering may interact in theoretically and clinically interesting ways with language tasks requirements, the condition of stuttering cannot clearly be linked to obvious underlying deficiencies in language knowledge or to evidently atypical patterns of language performance” (Bernstein Ratner, 1997, p. 103). Although current results add to the mounting evidence supporting this view, the presence of negative outcomes implies the possibility that pursuit of different language measures may yield different results. Also, additional investigations or replication efforts should consider the limitations of the current study. For example, a stronger analysis would be possible if equal numbers of participants were selected at each grade level. Although the current study included equal numbers of participants in both the CWS and CWNS groups at each grade level, there were discrepancies in the total number of participants between the grades (10 in kindergarten, 24 in first grade, and 10 in second grades).

In summary, the findings from this study support the view that CWS may not differ significantly from typically fluent children in their language performance even during linguistically demanding tasks; however, notable discrepancies may manifest among children who do and do not stutter in early elementary grades, such as kindergarten. Although current findings await further scrutiny through different language measures obtained from larger and more varied samples, the results have implications for the demands-capacities model of stuttering etiology. While one must remain vigilant to the possibility that global or subtle deficiencies in multiple domains may interact with environmental variables to put some children at risk for stuttering, there is limited support that language domains pose special challenges for all children who exhibit stuttering.

Acknowledgements

I thank the participants for their time and patience. Thanks also to Karen Andriacchi and the personnel at the Language Analysis Laboratory, University of Wisconsin-Madison for their SALT encoding services. Finally, thanks to Daniel Kempler for supporting this research project and Dorothy Aram for help in coding tense markers and reviewing this manuscript.
References


Appendix A

Descriptions of story characteristics that feature in the Narrative Scoring Scheme (NSS), reprinted from the SALT manual.

The NSS consists of the following seven story characteristics:

Introduction: Scores are determined by the presence, absence, and qualitative depiction of character and setting components.

Character Development: Scores are based on the acknowledgement of characters and their significance throughout the story.

Mental States: Narratives are evaluated based on the vocabulary used to convey character emotions and thought processes. The frequency as well as the diversity of mental state words is considered. For example, if a story provides frequent opportunities to verbalize anger themes and a child marks each of these “mad,” he/she will not receive as high of a score as a child that explains one opportunity using “mad,” another using “angry,” another using “upset,” and so on. Mental state words can be either adjectives (sad, happy, scared) or active cognitive-state words (believe, know, remember).

Referencing: Scores are given according to the consistent and accurate use of antecedents and clarifiers throughout the story.

Conflict/Resolution: Conflict/Resolution scores are based on the presence/absence of conflicts and resolutions required to express the story as well as how thoroughly each is described.

Cohesion: The sequencing of, details given to, and transitions between each event are examined.

Conclusion: A score is based on the conclusion of the final event as well as the wrap-up of the entire story.

There was this boy. And his three pets were sad because the boy was leaving them for the first time. And then, before he left, he petted his dog and his turtle bye. But then the frog jumped into his pocket. And then the boy said, "Bye" to his pets. And his pets looked really sad. But the frog was so happy because he got to go with him. Then the guy parked the car. And then they went into a restaurant. While they were looking at the menus, the frog jumped (out of his pocket) out of the boy's pocket and landed in a trumpet. When the man was playing the trumpet, his face started to turn red. And then (he tried he saw) he saw what was up in his trumpet. Then the frog fell out of the trumpet and landed on the guy's face. (And then they all) and then they all fell (in a trash fight) in a wastebasket. And then they started to have a trash fight. And then the guy that was playing the trumpet, he broke the drum that he was starting to play. And then the frog jumped onto someone's food. The lady and the man were gonna eat the food. (But then) but then the lady took a bite and saw the frog under the food. Then the frog jumped (out) off the table and landed in a guy's cup. And then the guy that was playing the trumpet met the lady (who was) which the frog jumped in her food. And then the guy saw the frog in his cup. And he was really, really mad. And then the guy tried to catch the frog because he did not see him. And then a lady saw the frog in there and then just fell to the ground.
C And then a waiter walked by [SI-1].
C And he was carrying the frog by his legs [SI-1].
C And then the boy got his frog back from the guy [SI-1].
C And then they got out of the restaurant [SI-1].
C And then they went home [SI-1].
C And everyone was mad at the boy [SI-1].
C Then (the mom) the dad sent the boy to his room [SI-1].
C And then he played even more [SI-1].
+ Introduction: 2
+ CharacterDev: 2
+ MentalStates: 3
+ Referencing: 2
+ ConflictRes: 2
+ Cohesion: 2
+ Conclusion: 2

+ NSS: 15
+ SI: 1.44