The Effects of Discourse Genre on English Language Complexity in School-Age Deaf Students

Brenda Schick
University of Colorado at Boulder

This study explored the effects of discourse genre on the complexity of English language produced by school-age deaf students. Subjects were 13 profoundly deaf students, ages 7;1 to 14;8, who had been educated using only an English-based sign system in public schools. Language samples were elicited in three conditions: a traditional spontaneous interaction, a narrative elicitation using pictures, and an interview in which the examiner asked cognitively and linguistically challenging questions. The results show that students produced the most complex English language during the interview, as measured by the use of embedded clauses, conjunctions, and modals. In contrast to previous studies, utterance length in this one did not vary as a function of sample type, likely because segmentation of the samples was based on the T-unit, not sentences. The results have implications for language sampling with school-age students, both in terms of efficacy and ecological validity.

It is well established that deaf students frequently have difficulty attaining levels of English competence comparable to those of their hearing peers (Bornstein, Saulnier, & Hamilton, 1980; Geers, Moog, & Schick, 1984; Kretschmer & Kretschmer, 1978; Quigley, Power, & Steinkamp, 1977). Deaf students' difficulties with English grammar have significant impact on their reading skills and academic achievement, which also tend to be below that of their hearing peers (Geers & Moog, 1989; Moores, 1987; Schildroth & Karchmer, 1986). For example, a large cross-sectional study on deaf adolescents educated in both oral and signing environments identified English language ability, including vocabulary, syntax, and discourse skills, as a primary factor associated with achieving reading skills commensurate with those of normal-hearing students (Geers & Moog, 1989; Moores & Sweet, 1990). Currently, the majority of deaf students are educated in public school settings (Moores, 1987) and, consequently, educators and clinicians are often required to monitor and evaluate the extent to which a deaf student is making satisfactory progress. To do so, they need evaluation tools that can assess English language skills in students who are producing complex English language.

This need to assess English language in older students, hearing and deaf, has posed special difficulties for professionals. Research on spoken language development and clinical intervention for language delay have traditionally emphasized language produced by younger children. However, significant growth in language, particularly in areas of discourse, pragmatics, and nonliteral language, as well as some advanced syntactic structures, occur throughout the preadolescent and adolescent years (Nippold, 1988). Evaluating language skills in older children poses special problems because changes in language behavior in this population become apparent only when "sophisticated linguistic phenomena are analyzed" (Nippold, 1988, p. 3). In addition, for deaf students, it is important to contin-
Continuously assess the extent to which knowledge of English is a barrier to academic success, particularly since teachers and subject areas differ in their reliance on linguistic skills (e.g., reading versus math).

This is particularly important because the role of language in education shifts as a student progresses through the grades. For younger children, education focuses on developing an ability to express oneself, as well as learning content material (Cazden, 1972). As students become older, the emphasis shifts from learning new uses of language to using language to learn. This places school-age deaf students who have yet to master advanced English language structures at a disadvantage, since the classroom does not support their language learning during content subjects as well as when they were younger.

Language and Academic Success

The difficulties assessing English language skills in school-age deaf students is exacerbated by the fact that language, as it is used in conversation, is different from language used in classroom discourse. Any consideration of a deaf student's English language skills must be viewed in terms of functionality in the classroom. Cummins (1984) has proposed a model of linguistic proficiency that provides a useful means of conceptualizing the relationship between cognitive and academic functioning. He argues that language proficiency should be addressed in terms of the degree of contextual support and cognitive involvement required in any given linguistic task, as illustrated in Figure 1. Contextual support may range from context-embedded to context-reduced communication, represented on the horizontal continuum. In context-embedded interactions (quadrants A and B), the participants can negotiate meaning and the language is supported by paralinguistic and situational cues. In contrast, in context-reduced communication (quadrants C and D), language may assume a greater role in conveying the message and there is an added pressure for the linguistic message to be elaborated precisely and explicitly. Cummins (1991) has more recently described this distinction as conversational-versus academic-language proficiency.

In addition, as represented on the vertical continuum, the degree of cognitive involvement will influence language proficiency. Numerous language routines are automatized and require little active cognitive involvement (quadrants B and C). At the other end of the continuum are communication tasks that require active cognitive involvement (quadrants D and C). Cummins suggests that persuading another individual is an example of a situation that often requires a person to stretch linguistic resources to achieve a communicative goal. While a student might have language skills adequate for context-embedded, cognitively unchallenging tasks, classroom lessons requiring a high degree of cognitive involvement, with little environment support, might pose difficulty. There is some evidence that contextualized and decontextualized language skills are relatively independent of each other in bilingual students (Cummins, 1991; Snow, Cancino, De Temple, & Schley, 1991).

For students who continue to be at risk for difficulties with the English language, such as school-age deaf students, the issue is not whether they are able to use a given language structure, but whether they can use it to learn new information and whether a particular language structure is productive in the classroom learning environment.

Assessing Language Competence

To determine linguistic competency, a student's language skills must be sampled in some way, typically accomplished by either using tests or unstructured free-

Figure 1 Model of the interaction of language with cognitive complexity (Cummins, 1984).
play. Language tests make elicitation more reliable and efficient in that they attempt to obligate the use of certain structures. However, most standardized tests evaluate aspects of language acquired relatively early. Not many standardized tests evaluate later aspects of language development, such as the use and understanding of advanced connectives (e.g., although, likewise) and advanced syntactic constructions (Nippold, 1988) or the use of embedded clauses. In addition, it is difficult to obligate relatively rare structures. Often there are alternative ways to state complex ideas that may not be as linguistically sophisticated but are not incorrect. Obligating complex syntax often involves highly structured activities, which present additional problems. However, given the low frequency of occurrence of some complex structures, tests provide a reliable method of determining competence.

While standardized language tests might make elicitation more reliable, they do not focus on use of language or always provide a representative picture of a student's language behavior. The goal of most language tests is to determine whether the student is competent in the use of a certain language structure, without the confound of a cognitively difficult task. In fact, in most language tests, especially for younger children, cognitive complexity is considered an unwanted interference. To reduce this confound, tests are generally context-dependent, frequently using pictures or toys for stimulus materials. In essence, competence is tested in Quadrant A of Cummins's model. While this provides some information, it is likely inadequate to predict academic success.

Some tests designed for older students and adolescents elicit language in a decontextualized context. However, Nippold (1988) argues that many of these tests are not theoretically motivated, have psychometric problems, or, most importantly, do not truly assess aspects of language that develop at older ages. Rather, they assess features of language production typically present in older students, even those with difficulties in language. More recently, assessment tools have focused on language appropriate for older students, such as the Test of Adolescent Language (TOAL-2; Hammill, Brown, Larsen, & Wiederholt, 1987), the Test of Language Competence (TLC; Wiig & Secord, 1989), and the Test of Language Development (TOLD-2; Hammill & Newcomer, 1988). These tests often evaluate the use of ambiguity, figurative expressions, and superordinate terms in vocabulary, all of which require metalinguistic skills.

These areas can be problematic for the deaf. Nonliteral use of English language is typically more difficult for deaf students than syntactic structure (Conley, 1976; Iran-Nejad, Ortony, & Rittenhouse, 1981; Yoshinaga-Itano & Downey, 1986; see, however, Marschark & West, 1985). That is, a deaf student who has difficulty with these areas of English may still have relatively intact syntax. In addition, tests for older students often assess syntactic skills by means of reading and writing, which are also difficult for deaf students (Moores, 1987). However, the main difficulty with depending on tests designed for hearing students to assess English language skills in deaf students is that the tests assume a certain relationship between overall language skills and the individual measures, such as vocabulary skills and metalinguistic knowledge. These relationships may not be the same in deaf and hearing students. Schick and Moeller (1992) found evidence for this interpretation in a study of the English language skills of school-age deaf students. The ability to use morphological affixes and articles did not predict their overall ability to produce complex language, even though the misuse of morphological affixes is quite predictive of language impairment in hearing students.

In addition, issues related directly to deafness and sign language sometimes make test items less effective when used with deaf students than with hearing students. For example, a TOAL-2 subtest requires students to identify pictures that represent two possible meanings of a single word, such as haste as in sewing and cooking. However, the signs that represent these homonyms do not share the same phonological form. So hearing students would have an advantage in that they would have the additional phonological clue not available to the deaf student. Finally, these tests often do not directly test the grammatical structure of English, and many deaf students continue to have difficulty in sentence construction even in upper grade levels. Educators often need a broad sample of syntactic skills even when more formal tests are administered.

Despite these limitations, tests often provide information that can be compared across students more reli-
ably than spontaneous samples of language use. Even with deaf students, formal tests provide some information regarding the student’s potential ability to perform in a mainstream setting. However, given that formal tests results are likely to be below average for deaf students, educators should also obtain a reasonable sample of the student’s expressive language, a recommendation often made for hearing students (Bloom & Lahey, 1978; Lund & Duchan, 1988; Miller, 1981).

Clinicians and educators typically obtain a spontaneous language sample in a naturalistic context by observing a child’s language during free-play interactions. Bloom and Lahey (1978) describe language samples as low-structured observations in which the adult is as unobtrusive as possible. However, it is not clear whether this strategy of eliciting a language sample is also optimal for older students.

There are a few reports of alternative means of obtaining language samples for clinical purposes (Dollaghan, Campbell, & Tomlin, 1990; Evans & Craig, 1991). In order to investigate the relative effectiveness of language sampling techniques, Evans and Craig (1991) compared language obtained during a free-play interaction, during which the child determined the activity and the adult followed the child’s lead, with language obtained during an interview, during which the child was asked about family, school, and free-time activities. For 8- and 9-year-old children with specific language impairment, they found that an interview was a reliable and valid means of eliciting language. The children used similar language in both free-play and in the interview condition, but the majority of behaviors occurred significantly more often during the interview. In addition, the children were more responsive and produced more overall language during the interview. They concluded that the interview was a more efficient means of collecting a representative language sample.

Language Sampling and Language Variability

Clearly, language obtained during sampling can vary as a function of extrinsic factors, such as those related to the environment, the interactant, or the topic of conversation. Environmental variables that affect spoken language production include the use of props or pictures (Atkins & Cartwright, 1982; Brown & Ruder, 1985; Dollaghan, Campbell, & Tomlin, 1990; Longhurst & Grubb, 1974; Longhurst & File, 1977; Wren, 1985); the conversational partner (Olswang & Carpenter, 1978; Scott & Taylor, 1978; Wilkinson, Hiebert, & Rembold, 1981); the type of question asked (Blank, 1975; Evans & Craig, 1991); the extent of listener knowledge (Masterson & Kambhi, 1991; Skarakis & Greenfield, 1982); and topic (Cazden, 1970; Masterson & Kambhi, 1991; MacLachlan & Chapman, 1988; Scott, 1988a).

However, there are several problems with generalizing results from previous studies to the task of eliciting language in older students. First, the vast majority of research completed on variables affecting language production has focused on young children, often in the earliest stages of language development. Second, researchers have typically used MLU, type/token ratios, and number of utterances as language variables to evaluate which sampling technique is the “best” (see, however, Evans & Craig, 1991; Masterson & Kambhi, 1991; Scott, 1988b). Those sampling techniques that resulted in longer utterances or more utterances were considered better than others; length and complexity are often considered synonymous. Third, sampling techniques have mostly emphasized which technique provides the most language, without considering whether the language sample reflects classroom needs. Ecological validity and academic requirements on language are important when evaluating language in older students. Fourth, to some extent, spontaneous language sampling makes comparisons across students more difficult. When a student controls the topic, comparing the language produced by one student with that produced by another becomes more difficult. Each student decides which propositions are discussed, and the syntactic structures used, which results in problems of replicability.

Research on written language in older students provides some information on how best to elicit complex language at advanced levels in that there is some evidence that cognitively challenging tasks elicit more complex written language in school-age students. Perron (1977a, 1977b) found that mode of discourse was a significant factor in influencing syntactic complexity in the written language of students in third, fourth, and fifth grades. He found that students used more com-
plex syntax in their writing, as measured by clausal length as well as the use of subordination, when they were asked to write an argumentation, as opposed to an exposition, narration, or description. Blair and Crump (1984) found similar results for written language of language-disordered students in the sixth, eighth, and tenth grades, indicating that genre might be a highly relevant variable in language sampling, particularly with older students.

This study investigates English language complexity in two traditional language sampling techniques in comparison with a sampling technique that, to use Cummins's model, is both context-free and cognitively challenging. The first two techniques include a spontaneous interview similar to that described by Evans and Craig (1991) and a narration using complex picture prompts. The sampling technique of interest to this investigation uses cognitively challenging questions to elicit a student's opinion. Of particular interest was whether a sampling technique that better represents academic use of language would elicit a language sample comparable to a free-speech sample, and whether the added cognitive complexity of the sampling condition would induce more errors in language production. And because students have been shown to use more complex written language when the task entails it (Perron, 1977a, 1977b), I hypothesized that the cognitively challenging condition would elicit more complex language structures.

To examine this hypothesis, I examined several measures of English language complexity, including the use of subordinate clauses, modals, and diversity of conjunction use. Studies have shown that even when deaf students use complex sentences, they tend to rely on a subset of those used by hearing individuals (Berenst, 1988; Quigley, Smith, & Wilbur, 1974; Quigley, Wilbur, & Montenelli, 1976). Conjunctions were analyzed because of their important role in structuring discourse since they create connections across discourse and thus reflect a more literate form of language (Snow & Patton, 1993; Wallach & Miller, 1988) and greater cognitive complexity. Modals indicate metalinguistic ability and skills in social cognition because they often communicate the degree of belief about the truth of a statement. Coates (1988) found that development of modals continues into the school years, even until age 12 years, because, she argues, the modal system is cognitively quite complex. Studies have shown that deaf students have difficulty using modals (Wilbur, Goodhart, & Fuller, 1989). And given the difficulty that some deaf students have with complex sentences and discourse construction (see Weiss & Johnson, 1993), I analyzed use of conjunctions to provide some measure sensitive to the use of complex language and discourse complexity. For conjunctions and modals, I used a broad measure of use, which was defined as the number used per utterance. Similarly, a measure of diversity was used, defined as the number of unique uses. A measure of utterance length and a type/token ratio were also included to allow comparison with previous studies on elicitation procedures and to gain a broader picture of the effects of sampling condition on use of English language.

In addition, previous research has shown that errors in language production may increase with syntactic complexity (MacWhinney, 1978; Panagos, Quine, & Klich, 1979). If that relationship were present here, it might help professionals better understand classroom performance. Thus, measures of English language errors typical in deaf students' expressive language were included.

The data used for this study were taken from an investigation on the learnability of English sign systems in deaf students, age 7–15 years (Schick & Moeller, 1992). There has been considerable controversy concerning whether visual forms of English (e.g., Signing Exact English, Signed English) are learnable as a native language. Some researchers (Gee & Goodhart, 1985; Supalla, 1991) have argued that a visual representation of English, because it depends on an auditory language, conflicts with the manner in which the eye processes language, thus making a visual representation of English theoretically unlearnable. All previous evidence has been negative, in that previous studies have examined children in environments with questionable support of language learning, regardless of the language form issue. The goal in the Schick and Moeller study was to determine whether students who had been educated using a visual form of English and had been in environments supportive of language learning would learn English or would continue to demonstrate the types of problems with English found
in previous research. Because deaf students typically have difficulty with complex English structures, a sampling protocol was developed to elicit language in a variety of settings, with the goal of obtaining a broad picture of the students' skills with relatively rare structures. Schick and Moeller found that for measures of language complexity, the deaf students were comparable to hearing students who were slightly older. The students had very good skills in global aspects of English, such as word order, use of sentence embedding, and frequency of types of embedding. For measures of language errors in English, they continued to have difficulty using closed-class morphology, including inflectional morphemes, auxiliaries and copulas, and articles. This study examines these data for the three different sampling conditions used in the protocol.

Method

Subjects. Subject selection was based on the following criteria: (1) profoundly deaf; better-ear pure frequency average (PFA) at 500, 1000, and 2000 Hz greater than 90 dB (ANSI, 1969); (2) chronological age between 7;0 and 15;11; (3) no additional disabilities of educational significance; (4) scores within the normal range on measures of general intelligence (≥85), as defined by a performance scale; and (5) enrolled continuously in the same educational program since three years of age. Parents of all subjects who met the criteria were sent a letter describing the project goals and procedure. All students whose parents responded to the mailing were included in the study; producing a total of 13 subjects; all students were Caucasian, 8 were girls, 5 were boys. Students were told that the goal of the project was to investigate the English skills in deaf children, and all subjects were paid for their time. Table 1 shows subject information, which includes age, hearing level, scores on a Performance IQ test from the WISC-R or the Hiskey, percentile scores in reading, and standard scores on the Peabody Picture Vocabulary Test, as well as grade level at testing.

All students attended two schools strongly committed to providing a complete visual English representation using an English sign system. Both programs used Signing Exact English (SEE II; Gustason, Pfetzinger, & Zawolkow, 1980). There are large differences among the artificially created systems; SEE II attempts to represent English literally and purports to follow a strict one sign for one English free morpheme or "word" criterion (although it does not always do so).

Table 1  Subject information for the deaf subjects including age, pure tone average (PTA, the Better Ear Speech Frequency Average at 500, 1000, and 2000 Hz) and performance IQ (PIQ), reading percentile (hearing norms), and standard scores from the Peabody Picture Vocabulary Test (PPVT), and grade level at time of testing

<table>
<thead>
<tr>
<th>Subject</th>
<th>Age</th>
<th>PTA</th>
<th>PIQ</th>
<th>Grade level</th>
<th>Reading percentile</th>
<th>PPVT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7;1</td>
<td>111</td>
<td>120</td>
<td>1st</td>
<td>98*</td>
<td>111</td>
</tr>
<tr>
<td>2</td>
<td>7;6</td>
<td>105</td>
<td>121</td>
<td>1st</td>
<td>80*</td>
<td>125</td>
</tr>
<tr>
<td>3</td>
<td>8;8</td>
<td>106</td>
<td>132</td>
<td>3rd</td>
<td>62*</td>
<td>83</td>
</tr>
<tr>
<td>4</td>
<td>9;0</td>
<td>105</td>
<td>119</td>
<td>3rd</td>
<td>71*</td>
<td>71</td>
</tr>
<tr>
<td>5</td>
<td>9;7</td>
<td>105</td>
<td>130</td>
<td>4th</td>
<td>40*</td>
<td>99</td>
</tr>
<tr>
<td>6</td>
<td>9;11</td>
<td>105</td>
<td>120</td>
<td>4th</td>
<td>32*</td>
<td>86</td>
</tr>
<tr>
<td>7</td>
<td>10;11</td>
<td>100</td>
<td>109</td>
<td>4th</td>
<td>40*</td>
<td>72</td>
</tr>
<tr>
<td>8</td>
<td>11;3</td>
<td>100</td>
<td>130</td>
<td>5th</td>
<td>66*</td>
<td>99</td>
</tr>
<tr>
<td>9</td>
<td>11;10</td>
<td>106</td>
<td>130</td>
<td>5th</td>
<td>50*</td>
<td>NA</td>
</tr>
<tr>
<td>10</td>
<td>12;3</td>
<td>95</td>
<td>NA</td>
<td>5th</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>11</td>
<td>12;4</td>
<td>93</td>
<td>127</td>
<td>5th</td>
<td>20*</td>
<td>77</td>
</tr>
<tr>
<td>12</td>
<td>14;4</td>
<td>100</td>
<td>111</td>
<td>9th</td>
<td>53*</td>
<td>88</td>
</tr>
<tr>
<td>13</td>
<td>14;8</td>
<td>123</td>
<td>115</td>
<td>9th</td>
<td>63*</td>
<td>76</td>
</tr>
</tbody>
</table>

* Stanford Achievement Test.
* Woodcock Johnson.
* California Achievement Test.
This means that the signing includes English word ordering, all function words, marking for complement and relative clauses, signs for English contractions (e.g., can't is signed can + n't), and a substantial number of inflectional and derivational morphological affixes (~70; e.g., -ed, pre-).

Procedure. The same examiner served as the conversational partner in all conditions. The examiner is skilled in the evaluation of deaf and hard-of-hearing students and fluent in SEE II. All of the subjects had known the examiner for many years due to annual intensive evaluations of their English language skills and academic achievement. Thus, the students did not need any encouragement to talk with the examiner.

Language samples were collected in three different conditions: a traditional spontaneous language sample (SPON), a narration elicited with a picture story (NARR), and a sample from the Language Proficiency Interview (LPI), an interview technique that involves asking complex questions. The spontaneous language sample was elicited by discussing topics of general interest with the examiner asking general open-ended questions, avoiding difficult topics. Most of the spontaneous language samples involved what the students had been doing since their last contact, such as school, common friends, dating, summer vacations, and so forth.

Second, subjects told a story using a picture story book that has no print but is complex in terms of narrative structure (Mayer, 1967). The story has 25 pictures and involves numerous complicated event structures. There is ample opportunity for students to discuss internal states, causations, contingencies, and dependencies. The story was not varied across age group because it had potential for complex narrative structure and there were difficulties matching stories in terms of structure.

The third sample was collected using the LPI (Caccamise, Newell, & Mitchell-Caccamise, 1983; Newell, Caccamise, Boardman, & Holcomb, 1983), an interview technique originally devised by the Foreign Service Institute to assess second language competence. Its focus is on how people use language to discuss topics that vary in cognitive complexity and abstraction. It has been adapted by Caccamise and his colleagues to obtain a broad measure of communicative competence in sign language. Particularly, language skills are assessed by an examiner who chooses cognitively challenging topics and asks linguistically complex questions. The goal is to ask questions that require a person to support a personal opinion with observations and learned information and to ask the question without reducing linguistic complexity. Questions that require a great deal of prior knowledge may be used, but if the student does not know the information, the examiner goes on to another question. The examiner also encourages the student to discuss the topic fully by asking additional questions to clarify and expand the topic. Naturally, what is a challenging question for a high school student and a junior high school student may differ, but since the goal is linguistic and cognitive complexity, the examiner is free to develop questions relevant to specific age ranges and students. The Appendix presents examples of questions for high school, junior high, and elementary students.

In addition to choosing cognitively challenging topics, the examiner purposely manipulates the syntax and lexicon to increase the demands of the question. For example, an examiner might initially ask a student, “Should the federal government pass a law to prohibit smoking in public areas? Why do you feel that way?” If the student has difficulty with the question, it may be rephrased as, “Do you feel that people should smoke in restaurants? Should we have a law to stop people from smoking? Why do you feel that way?” The goal is to see how the student can cope with the combination of complex topic, syntax, and vocabulary. In this study, the LPI was not scored, as Caccamise et al. do; rather, the subjects’ utterances during this interview were treated as a language sample.

Transcription and scoring. The samples from each of the three sampling conditions were first transcribed by a research assistant, a mother of a deaf child (who was not in the study) who is fluent in SEE II. The examiner or I checked the samples (I am a hearing native signer of American Sign Language [ASL], having grown up with deaf parents who used ASL, and am reasonably fluent with various English sign systems). Each transcription contained both the spoken and signed productions noting any omissions of obligatory words or
morphemes in either spoken or signed production. Imitative utterances were excluded from the transcript, because they did not reflect linguistic productivity, as were single-word responses to the examiner's question, because they were typically a response to a question. All sign transcriptions followed the rules of SEE II, which requires all function words, auxiliaries, and prepositions to be present. For example, if the student used the ASL sign DON'T-KNOW (a single sign), credit was given for know in the transcript, but not for I, do, or n't. Even though the student may have been semantically appropriate, we wanted to investigate structural competence in English. It should be noted that the students rarely used ASL signs. Signs that were unambiguous references even though they were not in SEE II were accepted. For example, the fingerspelled loan sign form for car was acceptable.

For each sampling condition, a combination transcript was created containing whichever production, voice or sign, was the subject's best production. That is, if a subject included an article in either speech or sign, but not necessarily in both, it was included in the best transcript. This investigation will focus on only the best productions.

The language samples were segmented into T-units, or terminable units (Hunt, 1970). The T-unit consists of a main clause and all of its subordinate clauses. By this definition, all compound sentences are considered two or more T-units (see Scott, 1988b, for a description of this linguistic unit). Language samples ranged in length: the spontaneous sample tended to be the longest (52-226; mean $M = 115$; standard deviation $SD = 50$), followed by the LPI (33-130; $M = 69$; $SD = 27$), with the narrative yielding the fewest utterances (31-120; $M = 58$; $SD = 22$).

Data coding and analysis. All transcriptions were coded using a computer program designed to facilitate structure coding and analysis, Systematic Analysis of Language Transcripts (SALT; Miller & Chapman, 1983). All utterances were coded, using SALT, with a taxonomy that focused on the English language complexity and types of errors produced by the subjects, discussed later. SALT allows coding on a sentence by sentence basis, where the coder views the actual sentence and makes a judgment. SALT then inserts a user-defined code, such as AUXD, for the deletion of an auxiliary. In other cases, SALT allows a researcher to define a closed set of words, such as conjunctions. SALT will search for all the sentences that contain a word from the set and will insert a user-defined code into each sentence, such as CONJC for the use of a clausal conjunction. Frequency per T-unit for each code was calculated using SALT and exported to a spreadsheet for further analysis.

Reliability. Reliability was calculated for transcription and coding for 6 of the 39 transcripts (15%). Intertranscriber item-by-item reliability between us, the examiner and me, ranged between .89 to .96 with a mean of .93. All disputes were resolved before coding the transcripts. Intercoder reliability between the graduate assistants and me was calculated for several categories: T-unit segmentation, number of dependent clauses, and the absence of obligatory sentential arguments. Reliability was computed as the ratio of the total number of agreements to the total number within each category. Agreement for T-unit segmentation was 97%; for number of dependent clauses, 95%; and for missing arguments, 91%.

Measures of English language structure. Of primary interest was the elicitation of relatively rare syntactic structures in the English language of deaf students, such as the use of subordinate clauses, clauses, and modals and auxiliaries. To determine the extent to which the subjects used embedded sentences, sentences were coded for the number of dependent clauses used within each T-unit. The average ratio of the number of dependent clauses to T-units (D/T) was computed for each sample for each student (see Hunt, 1970; Scott, 1988b). This measure shows the extent to which subjects used embedding; a high D/T ratio represents a high use of complex sentences with subordinate clauses. An example of computing the D/T ratio appears below. Main clause verbs are shown in bold and subordinate clause verbs are underlined. For this sample, there are 11 clauses in the 4 T-units, resulting in a subordination index of 2.75:

1. He works pretty hard without giving up because his father taught him how not to give up and quit.
2. But sometimes they say solve it yourself if it's an easy one.

3. Because I want to be the first deaf kid to go in space.

4. I saw that last year.

For both conjunctions and modals, SALT was used to determine whether a particular type of sample elicited a greater diversity of these elements. First, each sentence was coded according to whether it had a conjunction or modal or whether one was obligatory but missing. The diversity of conjunction use was calculated as the ratio of the number of unique conjunctions used divided by the total number of conjunctions used. This can be thought of as a type/token ratio (TTR) of conjunction use, with a low number reflecting reliance on a few conjunctions and a higher number representing a greater number of different conjunctions. Similarly, modal use was calculated as the number of unique modals divided by the total number of modal use per utterance.

In order to measure length of utterance, a mean length in words (MLW) was calculated. An MLU was not used because it is not sensitive to language growth at this language level (Miller & Chapman, 1981). In addition, Schick and Moeller (1992) found a lack of correlation between the correct use of morphemes and measures of syntactic complexity and age for school-age deaf students. To obtain a measure of lexical diversity, a TTR was calculated (Miller, 1981; Templin, 1957).

For the measures of English language complexity, the hypothesis was that the LPI would elicit higher levels of complexity, followed by the SPON sample, with the NARR sample eliciting the least complex language (LPI > SPON > NARR). This hypothesis was tested using a repeated measures analysis of variance (MANOVA) with sampling condition as the repeated measure, and D/T, conjunction diversity, unique modal use, total modal use, and TTR as the dependent variables. MLW was tested separately, using the same statistical design, but with a null hypothesis, that sample types would not vary in the length of T-units.

Measures of language errors in English. An area of particular difficulty for deaf students is the correct use of English argument structure (Geers, Moog, & Schick, 1984). That is, deaf students tend to have missing subjects, objects, and other obligatory syntactic arguments. The use of English argument structure was assessed by coding the absence of an obligatory argument of the verb, which included subjects, objects, indirect objects, prepositional phrases, and obligatory verb complements. In order to determine whether an argument was missing, or whether the student had made an appropriate phrasal comment, which is permitted in some contexts in English, the student's or examiner's previous utterance was examined. If the missing argument was not appropriate, given the structural context, it was coded as missing. Otherwise, it was considered correct. Examples of argument errors appear below, with omitted arguments underlined.

1. And I also had another good friend named Dalia and there were five boy/s who picked on me all month. They drove me nut/s.
2. I'd like a dumb waiter in it in the kitchen. You use it to pull up food for party/s.
3. And then one time I went to my room and I couldn't get it open and then I look at the number. They had taken my key out of my jacket and trade it.

Articles, auxiliaries, and copulas were coded as present or obligatory and missing. An error rate was determined by computing the percentage of contexts that were missing these elements. All inflectional morphemes were coded according to SALT format and analyzed using predefined programs in SALT, rather than codes we developed.

The measures of language errors in English were analyzed using a repeated measures MANOVA, with sampling condition as the repeated measure (LPI, SPON, NARR) and argument errors, morpheme errors, article errors, and auxiliary/copula errors as dependent variables.

Results

The results for the measures of English language complexity are shown in Figures 2–6. As can be seen, for
each measure, the D/T ratio (Figure 2), conjunction diversity (Figure 3), unique modal use (Figure 4), total modal use (Figure 5), and TTR (Figure 6), the LPI elicited higher scores than either the SPON or the NARR sample. A repeated measures MANOVA for the hypothesis that scores would be higher on the LPI followed by the SPON condition, followed by the NARR condition on the set of variables confirmed that, overall, the differences were significant ($F([12] = 69.57, p < .0001)$, as summarized in Table 2. Follow-up tests showed that these differences were results from higher scores for the LPI ($p < .001$), for the group of variables, than for SPON, and SPON scores were greater than NARR ($p < .05$).

Follow-up univariate tests showed that a similar pattern of results obtained across all dependent measures, also shown in Table 2. That is, for each variable, there were significant differences across sampling conditions (D/T: $F(2,24) = 13.42, p < .0001$; conjunction diversity: $F(2,24) = 4.86, p < .0001$; unique modal...
Figure 4  Shown for each sample type is the number of unique modals per total utterances.

Figure 5  Shown for each sample type is the total number of modals per total utterances.

use: $F[2,24] = 8.82, p < .0001$; total modal use: $F[2,24] = 12.35, p < .0001$; TTR: $F[2,24] = 9.36, p < .0001$). Follow-up tests revealed that four of these five measures followed the predicted relationship for all three conditions in that scores obtained during the LPI were higher than those obtained during SPON for D/T ratio ($p < .001$), for conjunction diversity ($p < .05$), for unique modal use ($p < .001$), and for total modal use ($p < .05$). For TTR, the differences between the LPI and the SPON condition were not significant. Differences between the SPON and NARR condition were not significant for the measures of D/T ratio, conjunction diversity, and modal diversity. However, there were significant differences in the use of total modals and TTR, with the SPON condition eliciting a higher measure than the NARR condition.

For the measurement of MLW, the effect was somewhat different, as can be seen in Figure 7. A re-
repeated measures ANOVA confirmed found no significant differences in the length of utterances, after samples had been segmented into T-units. No sample condition resulted in students producing longer T-units.

Unlike the measures of language complexity, for the four measures of language errors (argument, morpheme, article, or auxiliary/copula errors), the three sampling conditions did not elicit differences. A repeated measures MANOVA revealed no significant differences for the set of four measures of errors, for the main effect of sampling condition ($F[2,11] = .95$, $p < .40$). This means that no sampling condition caused the student to produce more errors, even though for some students, the error rates were high across all samples. Table 3 shows the rate of errors for these language structures and illustrates that students made errors in all conditions and that the error rate
Table 3  Percentage of errors for argument errors, bound morphemes, articles, and auxiliaries per T-unit for each condition

<table>
<thead>
<tr>
<th>Condition</th>
<th>Argument errors</th>
<th>Bound morpheme errors</th>
<th>Article errors</th>
<th>Auxiliary errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>NARR</td>
<td>.13</td>
<td>.24</td>
<td>.02</td>
<td>.15</td>
</tr>
<tr>
<td>SD</td>
<td>.02</td>
<td>.19</td>
<td>.01</td>
<td>.12</td>
</tr>
<tr>
<td>Range</td>
<td>0-.42</td>
<td>0-.19</td>
<td>0-.01</td>
<td>0-.12</td>
</tr>
<tr>
<td>SPON</td>
<td>.21</td>
<td>.27</td>
<td>.025</td>
<td>.18</td>
</tr>
<tr>
<td>SD</td>
<td>.01</td>
<td>.20</td>
<td>.12</td>
<td>.11</td>
</tr>
<tr>
<td>Range</td>
<td>.03-.41</td>
<td>0-.69</td>
<td>0-.035</td>
<td>0-.04-.56</td>
</tr>
<tr>
<td>LPI</td>
<td>.23</td>
<td>.22</td>
<td>.03</td>
<td>.17</td>
</tr>
<tr>
<td>SD</td>
<td>.02</td>
<td>.14</td>
<td>.015</td>
<td>.13</td>
</tr>
<tr>
<td>Range</td>
<td>0-.51</td>
<td>0-.48</td>
<td>0-.04</td>
<td>0-.06-.50</td>
</tr>
</tbody>
</table>

was not so high that there was no potential for further increase in errors.

To determine whether the differences in sampling condition were related to differences in comprehension of language, Pearson correlations were calculated between the students' D/T ratios and their PPVT scores and their reading standard scores. There were no significant correlations by condition for either PPVT (LPI: \( r = .155, p = .631 \); SPON: \( r = -.177, p = .583 \); NARR: \( r = -.014, p = .966 \)).

Discussion

The results show that language produced by deaf students using the LPI interview technique is generally more complex than that produced during a more informal spontaneous interview or elicited using a complex picture story. During the LPI, the students produced more embedded clauses as shown by the higher proportion of dependent clauses to T-units (D/T ratio), a greater diversity of modals, a greater number of modals overall, and a greater diversity of conjunctions. For the D/T ratio, conjunction and modal diversity, there were no differences between the SPON and NARR condition. The differences in conjunction diversity would indicate that the higher D/T ratio is not due to a certain type of complex sentences being used more often; rather, students were using a greater diversity of complex sentences in these conditions.

In contrast, sampling condition appeared to have no effect on the length of utterance, after utterances were segmented into T-units. This finding contradicts previous research (MacLachlan & Chapman, 1988), which found that narratives elicited longer communi-
cation units than conversations. This likely reflects the fact that narrations tend to elicit a greater number of compound structures, in which a child might string together several independent clauses into what appears to be one sentence when written on paper. Segmenting the utterances into T-units removes this confound, in that there is no subjective judgment concerning the starting and ending point of a sentence.

Given that the reading scores and vocabulary scores of these students were relatively good, these data likely reflect differences in sampling methods rather than differences in comprehension of English. The lack of correlations between the use of complex sentences in the three sampling conditions and reading and vocabulary scores indicates that differences in results were not due to differences in comprehension.

While an increase in cognitive complexity did result in more complex linguistic structures, an increase in the number of language errors was not observed. Students did not make more errors using inflectional morphemes, auxiliary or copula verbs, or articles during any single condition. However, there were high rates of errors in using these linguistic structures, so the potential for an increase in errors was present.

The results from this study appear to support a model in which cognitive complexity appears to increase the opportunity for linguistic complexity (Cazden, 1970). To convey the interrelationship of ideas necessary in higher levels of abstraction, one must use more complex languages structures to represent propositional complexity. In a sense, language function dictates linguistic complexity in that discussing complex ideas will entail using embedded sentences that indicate causality, conditionals, implicatures, and so on. Complex ideas often require a speaker to show the relationship between two propositions, which require conjunctions. In addition, complex ideas will entail the use of modals in order to modify predicates beyond simple tensing. Those conditions that result in more routinized and casual language also result in less complex language. This relationship was confirmed in this study.

These results might be due in part to the differences in register in the examiner's questions. When discussing topics of a personal nature (e.g., school, home, family), an informal register is appropriate in both the questions and the responses. However, when the topics involve opinions about international events, political issues, roles and responsibilities, a more formal register is appropriate. Likely the shifts in register in the examiner's questions obligated a similar shift in the student's register.

It is interesting to note that these results are similar to those found by Perron (1977a, 1977b) for written language, reported previously. That is, mode of discourse significantly influences syntactic complexity. This finding is encouraging since with school-age students, the relationship between verbal and written language skills is of primary importance. This study suggests that an interview technique such as the LPI could provide a better indication of whether the student's language skills were likely to impede success in literacy.

It is also clear that cognitive complexity did not interfere with the student's ability to produce language. Although these students were not producing perfect English (see Schick & Moeller, 1992), in general their language was not less accurate when the task was more cognitively complex. Students did not make more errors in the use of morphemes, auxiliaries/copulas, or articles during the LPI than in less demanding language tasks, despite the high rate of errors on these linguistic elements in general, which would indicate a potential for differences in error rates.

Following the model of language proficiency proposed by Cummins (1984), the language obtained during the LPI condition provides a better indication of how students use their language skills to interact and learn in an academic environment. In a sense, the LPI represents an ecologically valid means of sampling and evaluating school discourse. Not only is the language obtained during the LPI more complex, which provides more information about the use of rare structures, but the examiner is able to see language used without the support of context. In addition, because the topic is essentially dictated with the asking of a direct question with follow-up questions, it is possible to see how well students can maintain a discussion on a single issue, how well they can support their point, and whether they can use the narrative structure of argumentation and debate. Because the examiner's questions are not routinized language, the student is more likely to have comprehension breakdowns, which also provide more information about the student's potential
to succeed in a classroom. For example, some of the students in this study switched topics to their own choosing when they did not comprehend a question. In essence, their solution to a difficult language task was to produce a monologue and make it difficult for the examiner to interrupt. In other instances, it was difficult to change topics with the student; questions had to be repeated and restructured in order for them to respond appropriately. This might indicate that the students would have difficulty with topic shifts in the classroom, even though their language production appears good. And some students did not indicate that they did not understand the question, but when the examiner continued to request a response, it became clear that the student had not understood. By using an elicitation technique in which the examiner’s participation is as important as the student’s, aspects of language that involve more than sentence structure can be observed.

The LPI also allows similar questions to be used with different students so that the examiner can more easily compare responses across individuals. When a student controls the topic, as with more traditional language sampling, it is difficult to compare his or her ability to participate in the classroom environment with that of other students. Using similar questions across students allows an educator to compare a student’s performance with that of his or her peers.

In addition, with a technique in which the examiner controls the topics, social class and personality differences among students may not affect the resulting sample as much. Lawton (1968, in Cazden, 1970) reported that in open, unstructured discussion, middle-class British boys (12–15 years of age) used more abstract arguments and complex language than working-class boys, who used more concrete, narrative/descriptive language, clichés, and anecdotes. When highly directive, abstract questions were asked, social class differences were much smaller. Similar results are reported in Williams and Naremore (1969). This would indicate that working-class children might demonstrate higher levels of language functioning if they were asked challenging questions relevant to their lives. A very positive aspect of the LPI is that questions may be developed for the hot issues that affect students in a given school, thus increasing relevance and ecological validity.

These results also have implications for evaluating the best, in the sense of most representative, sampling technique. Previous measures of utterance length would indicate that a sampling technique that elicits a narrative would be the most effective and efficient. In this study, once subjective segmentation and concatenation of independent clauses were controlled, there were no differences in length of utterance among the three sampling conditions. However, measures of language complexity show that an interview technique would be the most efficient means of obtaining information about language complexity. In the past, many studies that investigated sampling conditions focused on utterance length (not T-unit length) and lexical diversity. If the results here were predicated solely on these types of measures, the conclusions would be very different. It is important that, for older students, measures of complexity reflect the use of sentence embedding, modals, and conjunctions, and not overall length.

In conclusion, the results suggest that discourse genre significantly affects language complexity in school-age deaf children. Because these students performed similarly to older hearing students, similar results likely would be found with hearing school-age students. In addition, using a discourse genre that more accurately reflects classroom discourse might provide better information about classroom performance, particularly given similar results in the written language of hearing school-age students (Perron, 1977a, 1977b).

Appendix

Sample Questions for the Language Proficiency Interview

High School Students

1. Do you believe that the Bill Cosby show provides a good role model for American families? Why or why not? How does your family compare with theirs?

2. Would you support a law to ban the sale of handguns? What should we do to reduce the use of weapons?

3. If a student gets low grades in school, should they be allowed to participate in sports? What can schools do to encourage students to obtain better grades when they are having difficulty?
Junior High School Students

1. What could the President do to solve the problems of homeless people today?

2. If NASA announced a new program that would allow a deaf child to go into space, would you volunteer? Why or why not?

3. If you were to meet some children from the Soviet Union, what would you tell them about the United States?

Elementary School Students

1. How do you feel about violence on television? Should parents control which programs their children watch?

2. Should people smoke in restaurants and other public places? Do you feel the government should pass a law to prevent people from smoking?

3. Do you have responsibilities around your house? Do you feel that children should have to help with housework? Why or why not?

Notes

1. It is equally important to understand what type of ASL skills the deaf student might have. The goal is to find out the student's underlying language competence, not just English competence.

2. Performance scores were obtained from the student's file, which contained a mix of how the scores were reported. Some were reported in age levels and some in standard scores. In addition, students were assessed at varying points in their educational program since testing was completed only to verify scores within the normal range. All tests were administered by licensed and trained testers who were experienced in evaluating students who are deaf and hard of hearing.

3. The examiner was a SEE II signer with 12 years of signing experience. She received an Advanced rating in an evaluation given by the SEE Center.

4. The LPI uses questions at various levels of complexity to determine the highest level of cognitive complexity the individual can maintain an interaction. In this study, all the questions used would be considered the highest level in the LPI.

5. More specifically, Salfile was used, which allows easy interface between SALT I and other programs, such as a spreadsheet. Using this utility, a user can select variables of interest, such as MLW or use of word roots, and output the data directly to a text file. Since multiple files can be processed as a single batch, analysis is fast, once transcripts are properly entered.

6. Follow-up tests were conducted using Systat, a software package for data analysis. In Systat, these tests are simply termed "comparison studies."

References


