Grammaticality Judgments of an Extended Optional Infinitive Grammar: Evidence From English-Speaking Children With Specific Language Impairment

Mabel L. Rice
University of Kansas
Lawrence

Kenneth Wexler
Massachusetts Institute of Technology
Cambridge

Sean M. Redmond
University of Utah
Salt Lake City

This study reports on the outcomes of an investigation designed to evaluate competing accounts of the nature of the grammatical limitations of children with specific language impairment (SLI) with a new comprehension measure involving well-formedness judgments. It is a follow-up to the longitudinal study of Rice, Wexler, and Hershberger (1993), which reported on the production of grammatical morphemes by young children with SLI and 2 control groups of children, one at equivalent levels of mean length of utterance at the outset of the study, the other of equivalent age. In this investigation, we report on grammaticality judgment measures collected from the same 3 groups of children over a period of 2 years for 5 times of measurement. It is the first longitudinal study of grammaticality judgments of children with SLI. The findings show that children's grammatical judgments parallel their productions: Children with SLI can make fine-tuned grammatical judgments to reject morphosyntactic errors they are unlikely to commit, whereas they accept morphosyntactic errors that they are likely to produce. The findings support the extended optional infinitive (EOI) account of morphosyntactic limitation based in underlying grammatical representations and do not support accounts of input processing deficits or production constraints.

KEY WORDS: grammaticality judgments, specific language impairment, extended optional infinitives, children's language impairments, optional infinitives

Three generalizations are now well established about the grammar of young English-speaking children with specific language impairment (SLI). One is that some, but not all, of their morphosyntax is likely to be less mature than that of younger children at equivalent levels of mean length of utterance. For the affected morphemes, a group of children with SLI, on average, has a lower percentage correct for morpheme use than that of a younger language control group. The second generalization is that English SLI grammar differs from the adult grammar in that forms are likely to be dropped in contexts where they should appear. An important corollary is that errors of form choice are infrequent in young children's utterances. The third generalization is that this pattern of morphological limitations can persist for years beyond the time when unaffected children show a similar pattern.
Multiple explanations have been proposed for these grammatical limitations. Four particular models are of interest here. The approaches differ in the putative locus of the primary source of the grammatical impairments. Two approaches focus on the underlying grammatical representations: One of these, an extended optional infinitive account, highlights the role of deficits in tense marking; the second highlights a misunderstanding of subject-verb agreement. Two alternative approaches assume the grammar is intact and other factors account for non-adultlike morphosyntax. One possibility is the influence of production constraints. A fourth account focuses on limited input processing abilities. There is ongoing evaluation of the relative merits of the competing accounts, as scholars continue to weigh the evidence and related theoretical arguments.

To date, a potentially significant kind of evidence has been unavailable. Current evidence consists of production data, either from children’s spontaneous utterances or their responses to experimental probes designed to elicit target morphemes. A certain kind of comprehension evidence is much needed in order to evaluate important and contrasting predictions that can be drawn from the competing accounts. Although general tests of comprehension of morphology exist, they are unsuitable because they do not evaluate a child’s interpretation of a missing form (i.e., whether affected children regard an utterance of the kind they are likely to produce, such as “Patsy happy,” as allowable). Instead, available comprehension tests evaluate the semantic components of morphology, as in past versus present tense, where a child is asked to “show me mother baked the cake” versus “show me mother is baking the cake.” A child’s ability to point to the appropriate picture for either of these items does not tell us if the child is likely to regard dropped forms or wrong form choice as allowable alternatives to the adult grammar.

This study reports on the outcomes of an investigation designed to evaluate competing accounts of the source of the grammatical limitations of children with SLI, with a new comprehension measure involving well-formedness judgments. This investigation of children’s grammaticality judgment abilities is a follow-up to the longitudinal study of Rice, Wexler, and Hershberger (1998), which reported on the production of grammatical morphemes by young children with SLI and two control groups of children, all of whom were followed for 3 years. In this investigation, we report on grammaticality judgment measures collected from the same three groups of children, over a period of 2 years, for a total of five times of measurement. This is the first longitudinal evidence of this sort that we know of. In the following sections, we briefly review key empirical evidence; then we describe four alternative models of morpheme dropping.

We specify predictions for each model for performance on a grammaticality judgment task. We conclude that the affected children, as well as the younger children, make judgments that mirror their morphological productions in ways predicted by the extended optional infinitive model, thus placing the limitations primarily in underlying grammatical representations of a certain kind. Furthermore, the predicted judgments show the expected transition over time for the affected and younger control children, such that the earlier immature grammar gradually approximates the adult grammar.

**Empirical Evidence**

There is now a large body of evidence attesting to the fact that young English-speaking children with SLI are likely to drop morphemes, especially certain verbal morphemes (for reviews, see Bishop, 1997; Leonard, 1998; Rice, 1996). Relevant to this study is the now well-replicated finding that regular past tense -ed is likely to be omitted. Key studies include those by Bishop (1994); Leonard, Bortolini, Caselli, McGregor, and Sabbadini (1992); Marchman and Ellis Weismer (1994); Oetting and Horrobin (1997); Rice and Wexler (1996); and Rice, Wexler, and Cleave (1995). Another well-established finding is that third-person singular -s is likely to be dropped from obligatory contexts (Bishop, 1994; Leonard et al., 1992; Rice & Oetting, 1993; Rice & Wexler, 1996; Rice et al., 1995). Less widely studied are copula and auxiliary BE, which are known to be omitted from obligatory contexts (Cleave & Rice, 1997; Hadley & Rice, 1996; Ingram, 1972; Rice et al., 1995, 1998; Rice & Wexler, 1996). Finally, auxiliary DO is omitted by children with SLI (Rice et al., 1995, 1998; Rice & Wexler, 1996).

Following their theoretical model, described below, Rice et al. (1998) predicted that this set of morphemes (i.e., past tense -ed, third-person singular -s, BE, and DO), would individually and collectively differentiate affected children from younger control children over

---

1 The EOI model has recently been updated to an agreement-tense omission model (ATOM, Wexler, Schütze, & Rice, in press). Because the predictions for this study are not affected, and in the interest of continuity with Rice, Wexler, and Hershberger (1998), the EOI label will be used here. Wexler (in press-a) provides a model of the OI stage which derives the ATOM and its effects from a syntactic model involving the “unique checking constraint.” Again, since the predictions for English development are not affected, we will not discuss this model here.

2 Linguistic theorists differentiate between grammatical representations, such as features of grammatical tense (TNS) and subject-verb agreement (AGRE), and computational processes, such as operations involving feature checking and movement of sentence constituents. Scholars are presently working out this distinction in adult grammars. The distinction is not relevant to this investigation, where we use “grammatical representations” as a general term encompassing grammatical features as well as the computational processes within the grammar that act upon these features.
time, and would show a protracted period of acquisition. They carried out the first detailed longitudinal study of children with SLI and two control groups of children for this set of morphemes. These predictions were upheld; the SLI group, on average, performed at lower levels of percentage use in obligatory contexts than the younger control children throughout the period examined, from age 5 to 8 years for the SLI group and from 3 to 6 years for the younger control children. Furthermore, hierarchical linear modeling of growth curves revealed that change in the acquisition of these morphemes was not predicted by mother’s education, the children’s initial receptive vocabulary level, or the children’s performance on nonverbal intelligence measures. The children’s initial mean length of utterance was a weak predictor, accounting for less than 1% of the observed variation. These findings are consistent with the possibility that grammatical change in this domain is driven by maturational forces relatively specific to morphosyntax.

Of equal interpretive importance is the replicated evidence showing what is not affected in the productions of affected children. Progressive -ing, for example, as in “Patsy is talking,” is likely to appear in obligatory contexts, indicating that the problem is not one that equally affects all verbal morphology (Rice & Wexler, 1996). Furthermore, errors of overt form usage, as in “Patsy are talking” are very unlikely (Rice et al., 1995, 1998; Rice & Wexler, 1996; van der Lely, in press). Also, plural -s, a morpheme with great surface similarity to third-person singular -s, is known to be relatively unaffected (Bishop, 1994; Rice & Wexler, 1996).

Four Alternative Interpretations of the Observed Morpheme Dropping

Extended Optional Infinitive Model

As Wexler (1994) showed, in the first study which argued that there was a cross-linguistic “optional infinitive” phenomenon, observations of young children speaking German (Clahsen, 1982; Miller, 1979; Mills, 1985; Poeppel & Wexler, 1993; Weissborn, 1990; Weverink, 1989) French (Pierce, 1992), and a number of other languages (Wexler, 1994), revealed that the youngsters sometimes used infinitival forms of verbs where finite forms were expected. This was apparent because in these languages infinitives appear with audible morphemes that clearly distinguish them from finite forms of a given verb. For example, in German, bringen [bracht] is an infinitival form of the verb “to bring,” whereas bringen [bracht] is a finite form; in French, manger [manger] is an infinitival form of the verb “to eat,” whereas mange [manger] is a finite form. This period of children’s use of infinitival verbs where finite verbs are expected became known as an “optional infinitive” (or, alternatively, “root infinitive”) period of grammatical development. Wexler (1994, 1996) hypothesized that such a period was also operative in English-speaking children’s grammatical development, where the bare stem form of the verbs would appear as infinitives in contexts where finite forms were required. This could be seen in contexts requiring the use of third-person singular -s or past tense -ed. Wexler’s model also predicts omission of auxiliary and copula BE and auxiliary DO. The surface grammatical manifestations of this period are attributed to an immature underlying grammatical representation, in which features of grammatical tense marking (TNS), which are obligatory in the adult grammar, are treated as optional by children. These features, in turn, influence the presence of third-person singular -s, -ed, BE, and DO, in that these forms can be omitted if tense features are optionally deleted in a clause. Equally important is the prediction of the optional infinitive (OI) model—that if the tense feature is marked in a given clause, the surface forms will appear as expected. That is, a wrong form choice, such as “Patsy are talking” should not appear because the underlying grammatical representations will not allow such a clause to be produced. If such errors appear, they should be infrequent, attributable to performance effects outside the grammar.

The hypothesis for children with SLI is that an OI period will be evident, similar to that of younger unaffected children, but it will be extended for a longer time period, hence an “extended” optional infinitive (EOI) period. Under this account, the observed dropping of surface morphemes, whether they be affixes on lexical stems, such as third-person singular -s or regular past tense -ed, or free-standing morphemes such as BE and DO, is attributable to an underlying immaturity of grammatical representations of the grammatical feature TNS, similar to that of younger unaffected children, but for a much more protracted period of time. This model was upheld by cross-sectional comparisons of children with SLI and younger control children (e.g., Rice et al., 1995; Rice & Wexler, 1996) and in a longitudinal study that documented the very protracted acquisition by affected children (Rice et al., 1998). A further prediction upheld was that morphemes of similar surface properties but outside the domain of TNS, such as plural -s, did not show the same level of affectedness (cf. Rice et al., 1998; Rice & Wexler, 1996). Note that for the EOI model, children’s productions and understanding of morphemes are expected to show parallels because they both tap into the underlying grammatical representations.

---

3The prediction of correct form choice for subject/verb agreement assumes that the relevant alternative forms are available for selection (i.e., have been entered into a child’s lexicon). For the morphological forms studied here, alternative form use was evident.
Furthermore, although some (not all) verbal morphemes (i.e., those that mark TNS) may optionally appear in the surface structures, they should not show misaligned person/number agreement with the subject of the clause.

**Deficit in Subject-Verb AGR**

An alternative view of possible deficits in underlying grammatical representations posits that the limitation is not attributable to TNS marking, but instead is located in subject-verb agreement (Clahsen, 1991; Clahsen, Bartke, & Gollner, 1997). Under this model, contrary to the predictions of the EOI account, errors such as “Patay are walking” are to be expected because person and number inflections on verbs are alleged to be severely impaired in children with SLI. Essentially, the model posits that children do not know that if the verb appears in the clause it must agree with the person/number features on the subject. Production evidence in support of the model is available from German (Clahsen, 1991), although counter evidence from German-speaking children with SLI is also reported (Rice, Noll, & Grimm, 1998). Clahsen et al. (1997) challenge the conclusions of investigators of SLI in English-speaking children with an analysis of a sample of older children with SLI collected by van der Lely (in press), but Clahsen et al.’s conclusions are countered by an analysis of the same data by van der Lely. These contradictions show that production evidence for the model is mixed. Note that under the subject-verb agreement model, production and grammatical judgments should be aligned, as in the EOI. Unlike the EOI model, children should regard misaligned subject/verb person/number features as allowable.

**Production Constraint**

Another line of explanation has been put forward for the dropped surface morphemes of English that posit an underlying deficit in production capacities. Bishop (1994), drawing on Levelt’s (1989) model of the processes involved in speech production, suggests that the pattern of intermittent omission of grammatical forms seen in English-speaking children with SLI is attributable to demands on the sentence formulator, such that omissions are associated with utterances characterized by greater processing demands, such as multiple phrases, unfavorable serial position, and phonological complexity (number of syllables). Bishop’s (1994) analysis of spontaneous utterances yields, as she concludes, less than compelling evidence for this account. She calls for experimental evidence. Of relevance here is the underlying assumption of the model: “SLI children do have underlying competence in that they understand the grammatical function of morphological markers, but fail to apply their knowledge consistently because of limitations on their processing capacity” (Bishop, 1994, p. 508, italics inserted). We take this to mean that, under this account, the underlying grammatical representation is intact and the same as that for unaffected children. The problem is located instead in an over-taxed production system that breaks down under the pressures of immediate spoken clauses, presumably because some components cannot be retrieved, accessed, or pronounced as readily as others, and it is these components that get dropped. Within this account, although production may be faulty, grammatical judgments, as an index of underlying competence, should be unaffected because the underlying grammar is unaffected. This logic is preceded by that of Leonard, McGregor, and Allen (1992), who state: “If differences between children with SLI and MLU controls were found to be smaller in comprehension than in production of these forms, then production factors would be implicated” (p. 1082, italics inserted). Note that this position is not differentiated from that of the EOI model on the expectation that processing factors such as utterance length could influence children’s utterances, or their grammaticality judgments. The EOI model does not rule out such influences. The differentiation of the two models hinges on whether or not the underlying grammar/competence is affected.

**Limited Input Processing Mechanisms**

A well-known perspective on the condition of SLI argues that the children’s morphological deficits are linked to a problem of input processing. A number of proposals have appeared within this framework. Two underlying assumptions unify them. As with the production constraint model above, a simplifying assumption is that the underlying grammar is intact. In contrast to the production constraint model, models of limited input processing argue that the problem of affected children is due to a protracted intake of relevant features of the input they receive. The temporal processing deficit model of Tallal et al. (1996) is a well-known model of this type, hypothesizing a deficit in the processing of brief speech stimuli and the processing of rapidly changing speech stimuli that accounts for the grammatical deficits of children with SLI. Tallal et al. (1996) also conclude that such a deficit accounts for children’s apparent gain in general language comprehension scores following an intensive intervention program in which practice in stimuli detection is provided. Rice (in press) points out that such an input processing limitation would pose big problems for a child’s grammar, with a much broader scope of likely errors in children’s utterances than are actually observed. Of relevance here are the predictions for judgments of omitted versus nonagreed finite forms, or for omitted forms that do not mark finiteness. These predictions are more fully specified in a model with similarities and differences from that of the temporal processing deficit model. Leonard
and his colleagues (Leonard, 1988, 1996; Leonard, Eyer, Bedore, & Grela, 1997) proposed a surface hypothesis, in which the acoustic properties of the grammatical markers combined with the processing demands of paradigm building are thought to be central factors in the tendency of affected children to drop these morphemes.

With regard to production data, limitations of input processing models have been noted. Rice and Oetting (1993) point out that, although plurals are predicted to be affected for preschool children with SLI relative to younger language controls, this prediction is not met, as also found by Bishop (1994), Rice and Wexler (1996), and Rice et al. (1998). As noted above, a number of investigators find that children with SLI are unlikely to make errors of form choice with surface morphology, which Leonard (1998) acknowledges could happen under a surface deficit account. Cleave and Rice (1997) report that counter to expectations, brief contractible forms of BE show higher percentages of use in obligatory contexts than do more acoustically salient uncontractible forms of BE for children with SLI. Presumably, the contracted forms of BE, which appear frequently in conversational speech, would make it difficult for a child to learn which form of BE goes with which subject. Finally, Wexler, Schutze, and Rice (in press) argue that pronominal case assignment is linked to TNS/AGR marking, a grammatical contingency outside the framework of input processing deficit models. In English, children show systematic substitution of non-nominative case marking (e.g., him, in contexts where nominative case, he, is required, even though he is the more frequent form in the adult grammar); they do not make the opposite substitution (i.e., he for him). Most importantly, the likelihood of nominative case is linked to agreement marking on the verb. These facts are not captured, collectively, by a limited input processing account.

The relevant point here is that the interpretation of production evidence does not support input processing models. Very little relevant evidence is available for children’s understanding of the morphemes. Fellbaum, Miller, Curtiss, and Tallal (1995) examined children’s understanding of regular past tense -ed, third-person singular present tense -s, plural -s, and possessive -s, as well as comparative -er, as in “Which girl is happier?” They compared sixty-seven 6-year-old children with SLI who had been identified in preschool and twenty-seven 6-year-old unaffected control children. The task required pointing to pictures that contrasted semantic cues, such as pointing to one sheep eating as an indication of the comprehension of “sheep eats” as contrasted to multiple sheep eating for “sheep eat.” They found significant group differences for -s, third-person singular -s, past tense -ed, and possessive -s, differences they attribute to the relatively brief duration of the morphemes. They contrast this with pronoun performance, which did not differentiate the two groups, and this they regard as attributable to the longer duration of pronoun forms. Although they do not make specific predictions for grammaticality judgments, they take this reasoning to suggest that children with input processing limitations should be unlikely to detect/differentiate dropped forms as ungrammatical and may well have faulty choice of forms of BE because these often appear as brief, unstressed, contracted forms.

**Grammaticality Judgments**

**Children With SLI**

Because we were interested in children’s interpretations of clauses with certain dropped elements (e.g., “Patsy walking” instead of “Patsy is walking” or “Patsy walks” instead of “Patsy walk”), we turned to a grammaticality judgment task, asking children for well-formedness judgments. Kamhi and Catts (1986) utilized a similar task with children with SLI, as one of a battery of measures of phonological awareness thought to be a factor in reading impairments. In that study of 12 affected children, ages 6;11–9;1, the task consisted of 16 sentences, 12 of which had morphological errors, such as “They throwing the stick.” No further description of the items is provided. Of interest here is the finding that the children in the SLI group made more errors than their age peers on this task. This was regarded as an unsurprising finding, given earlier studies showing that children with SLI performed below their age peers on metalinguistic tasks and were thought to have poor metalinguistic skills in general (cf. Kamhi & Koenig, 1985). This is of relevance here because it will be important to know if poor performance on a grammaticality judgment task is attributable to a limited general metalinguistic competence versus discrete areas of non-adultlike judgments.

Van der Lely and Ullman (1996) report grammaticality judgment outcomes for 12 children with “grammatical” SLI, ages 9;3–12;10, compared to language-ability-matched unaffected children. On a task in which unmarked verb forms were presented (i.e., “look,” where “looked” was required), the SLI group was more likely to accept the unmarked forms as grammatical than their younger language control children. At the same time, the children with SLI did correctly reject irregularized verb forms, such as “lecked” for “looked.”

In a study of tag questions, Weckerly, Contreras, Wulfeck, Bellugi, and Reilly (1997) found lower levels of grammaticality judgment performance by affected school-aged children than by age-matched controls for tag questions (e.g., “A needle never hurts Mary, does it?”). They explored agreement of the tag auxiliary verb form with a subject pronoun, tense, and polarity. They
conclude that the affected children's lower performance on dimensions of polarity and long distance agreement are attributable to limited processing abilities, not qualitative deficits in grammatical knowledge. Although processing factors could well be involved, it is not possible to rule out whether the complexity of the target sentences also tapped into underlying grammatical interpretations that complicated the children's judgments. What is needed is information from simple clausal constructions that allow for examination of subject-verb agreement, for example, without the additional grammatical complexities of tag questions.

Unaffected Younger Control Children

We have argued (Rice et al., 1995, 1996; Rice & Wexler, 1996; Wexler et al., in press) that SLI is a stage of EOI (extended optional infinitives), "extended because there is an OI (optional infinitive) stage in normal children in many languages, ending in English around 3;6-4;0 (cf. Wexler, 1994, 1996). The work on the OI stage in normal children has thus had to deal with fairly young children. Therefore the methods of investigation have been mostly limited to production data, since comprehension experiments on children younger than 3 years of age are quite difficult and typically elicit truth-value judgments involving semantic reference (e.g., identifying who performed an action; cf. Gordon, 1996; McDaniel & Cairns, 1996). The one comprehension study of the OI stage in English of which we are aware is Schoonenberger, Pierce, Wexler, and Wijnen (1995), which showed that non-finite root forms were interpreted by children as ongoing activity rather than as "modal." Grammaticality judgment evidence pertinent to the OI stage is not available for unaffected children. This study adds valuable first information on the grammaticality judgments in this domain for 4-year-old control children. Further, the SLI group can be considered as a model for even younger children at lower levels of obligatory TNS marking. Eliciting grammaticality judgments from a 6-year-old child with SLI is to have the 6-year-old's ability to form metalinguistic judgments working on the grammar of a 2- or 3-year-old (with respect to morphosyntactic abilities). In this way, the findings here bear on our understanding of the comprehension of younger unaffected children.

Predictions for Children's Performance on the Grammaticality Judgment Task in an OI or EOI Stage

The intent for this investigation was to study the grammar of children with SLI who are known to show an EOI stage. Although a grammaticality judgment (GJ) task was used, this was not a study only of whether children with SLI can judge grammaticality, though even that would be interesting. We predicted that they have a certain grammar, distinct from the adult grammar, and we were testing that grammar and comparing it to other ungrammatical grammars. Thus, we were interested in formulating items that were grammatical within the adult grammar and contrasting them with (a) deviations from the adult grammar that are allowed within an OI grammar and (b) deviations that are not allowed. We predicted that the children with SLI would differentiate between (a) and (b), would not differ from their controls in their judgments of (b), but would differ from controls in judgments of (a).

Predictions are laid out in the following way. First, we lay out precise predictions for an EOI grammar in the form of children's expected responses to clauses at the levels of individual sentences, their performance relative to the control groups, and the longitudinal trends. Following that section, we summarize the predictions for the EOI model, contrasted with the three alternative models and the default possibility that children with SLI lack the ability to make metalinguistic judgments.

Allowable Deviations From the Adult Grammar

Prediction 1: Children in an OI/EOI grammar will accept finiteness-marked items as grammatical. They will judge "he is running" and "he looks big" as grammatical.

Prediction 2: Children in an OI/EOI grammar will also accept OI items as grammatical (i.e., "he running away" and "he look happy" will be judged as acceptable).

Prediction 3: Children with an adult grammar will judge OI items as ungrammatical (i.e., "he running away" and "he look happy" will be judged as not acceptable).

Non-Allowable Deviations From the Adult Grammar

Prediction 4: Children in an OI/EOI grammar will judge non-OI ungrammaticality, or bad agreement (BA), to be not acceptable. They will know that "he is run" and "he are running" are both ungrammatical.

Prediction 5: Children in an OI/EOI grammar will differentiate between non-TNS items and items that show BA. They will be more likely to accept the former and reject the latter.

Predictions About the Performance of the SLI Group Relative to the Control Groups

Prediction 6: Given their extended period of OI, children with SLI will be more likely to judge OI sentences as grammatical than their younger OI controls. That is,
the difficulty with grammatical judgments of this sort is not fully predicted by general language growth indexed by utterance length. Further, they will vary greatly from their age peers who will know the adult grammar and judge accordingly (Prediction 3).

Prediction 7: At the same time, children with SLI will know that non-OI ungrammaticality, or BA, is ungrammatical. They will not differ from their controls on Prediction 4.

Predictions About Longitudinal Trends

Prediction 8: During the ages of 3–4 years, when the OI stage resolves in unaffected children, this group will show increasingly adultlike grammaticality judgments. This is because we expect, as their underlying grammar evolves, that their judgments will change accordingly.

Prediction 9: During the ages of 5–7 years, when the EOI stage is apparent in children with SLI, this group will show a protracted progression toward the adult grammar. This is to say that Prediction 6, relative to group differences, should persist during this time period.

Predictions About Group Performance

The predictions for group performance are laid out in Table 1, according to the theoretical model and the experimental grammaticality judgment conditions (i.e., the type of grammar4 and grammar contrasts thought to be involved). These are (a) a grammar that allows tense marking to be dropped (optional infinitives); (b) a comparison of the OI/EOI grammar to a grammar that allows for a verb form to appear that is not correctly marked for person/number of the subject (BA); and (c) a comparison to a grammar that allows for an omitted non-tense marking form, such as the present progressive -ing, as in “he is running” (Dropped -ing; DI). In the first line of the table is the predicted outcome if the children with SLI show a general metalinguistic deficit, relative to their age peers (i.e., lower performance in each of three possible grammatical contrasts). The second line shows the group predictions generated by an EOI model in which the children with SLI are at lower levels than their age-matched controls and their language-matched controls for an OI grammar, but perform better on a BA or DI grammar, thereby reducing differences with the controls. In the third line, the AGR deficit account differs from the EOI model in the prediction that the SLI group will perform equally poorly in judgments of an OI as well as a BA grammar, thereby falling below control groups on both measures. In the fourth line, the production constraint model contrasts with the EOI model in the prediction that the children with SLI will have intact understanding of the morphological functions tested, with no group differences. Finally, the input processing deficit contrasts with the previous models in the prediction that all three kinds of non-adultlike usage of verbal morphemes should be at reduced levels of performance for the SLI group, relative to the controls.

<table>
<thead>
<tr>
<th>Model</th>
<th>Optional Infinitive</th>
<th>Bad Agreement</th>
<th>Dropped -ing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metalinguistic deficit</td>
<td>SI &lt; AC</td>
<td>SI &lt; AC</td>
<td>SI &lt; AC</td>
</tr>
<tr>
<td>EOI</td>
<td>SI &lt; AC</td>
<td>SI = AC</td>
<td>SI = AC</td>
</tr>
<tr>
<td>SLI &lt; LC</td>
<td>SI = LC</td>
<td>SI = LC</td>
<td>SI = LC</td>
</tr>
<tr>
<td>AGR deficit</td>
<td>SI &lt; AC</td>
<td>SI &lt; AC</td>
<td>SI = AC</td>
</tr>
<tr>
<td>Production constraint</td>
<td>SI = AC</td>
<td>SI = AC</td>
<td>SI = LC</td>
</tr>
<tr>
<td>Input processing deficit</td>
<td>SI &lt; AC</td>
<td>SI &lt; AC</td>
<td>SI &lt; AC</td>
</tr>
<tr>
<td>SLI &lt; LC</td>
<td>SI &lt; LC</td>
<td>SI &lt; LC</td>
<td>SI &lt; LC</td>
</tr>
</tbody>
</table>

*Here we use the term "grammar" as a short-hand way to describe the deviations from the adult grammar that are expected to be accepted, such as "she is running" (as optional infinitive grammar) versus those deviations expected not to be accepted, such as "she are running" (BA) and "she is run" (DI). It is important to keep in mind that in all cases it is expected that children will accept adult utterances. Although one can think of the deviations from the adult grammar as "experimental conditions" evaluated here, this terminology does not capture the comparison with the adult grammar that is essential.

Table 1. Predictions of group differences according to theoretical model and experimental grammatical judgment conditions (grammars).

Expectations for Growth Curve Modeling

Finally, it is possible to investigate the shape of the growth curve outcomes for the grammaticality judgment measures, along with the role of the predictors evaluated in Rice et al. (1998). There is no previous report of evidence of this sort. We adopt the strongest assumption, that the children's productions and grammaticality judgments both derive from the same underlying base. Even though the two measures are not the same, and grammaticality judgment tasks undoubtedly tap into areas of linguistic, cognitive, and social functioning not necessarily the same as those addressed in children's productions, we nevertheless expect that the same general patterns of growth should hold for grammaticality judgment outcomes as well as production outcomes, given that the judgment tasks are designed to closely approximate children's actual utterances. Further, we expect that the relation to the four predictors evaluated in the production data (i.e., mother's education, children's nonverbal intelligence, receptive vocabulary, and mean length of utterance) should replicate for the grammaticality judgment outcomes.

Rice et al. | Children's Grammaticality Judgments of an EOI Grammar | 949
Method

Participants

The participants in this study were drawn from the children who participated in the longitudinal study of Rice et al. (1998). Three groups of children participated: a group of children initially identified with specific language impairment (henceforth, the SLI group), who met conventional inclusionary and exclusionary criteria for receptive/expressive language impairment when they were recruited into the study in the year before school entry; a second group of younger control children at equivalent levels of mean length of utterance (MLU) at initial recruitment (henceforth, LC for language controls); and a third control group of age equivalent children (AC, for age controls). The inclusionary criteria for entry into the SLI group were (a) one standard deviation or more below age mean levels on the Peabody Picture Vocabulary Test–Revised (PPVT–R; Dunn & Dunn, 1981); and (b) mean length of utterance one standard deviation or more below age expectations (Leadholm & Miller, 1993). All but 2 children were one standard deviation or more below the mean on the 5-subtest language quotient of the Test of Language Development–Primary (TOLD–P;2; Newcomer & Hammill, 1988). These two children were at low-normal range and met all other criteria. The children were screened for phonological competencies. They were not selected for grammatical morpheme performance, which was treated as a dependent variable. Nonverbal intelligence was within normal limits, or above, on the Columbia Mental Maturity Scale (CMMS; Burgemeister, Blum, & Lorge, 1972).

The first time of measurement in this study was in the third round of data collection for the longitudinal study. The SLI group consisted of 21 children, mean age of 6.0, with the following means and standard deviations (in parentheses) on language measures and nonverbal intelligence testing: PPVT standard score, 83 (12); TOLD–P,2, Spoken Language Quotient 5 subtests, 82 (07); and MLU-morpheme, 4.17 (.55); CMMS, 93 (08). For the LC group (N = 19, mean age of 4.1): PPVT–R mean standard score, 106 (12); TOLD–P,2, 107 (09); MLU, 4.18 (.63); CMMS, 110 (09). For the AC group: N = 21, mean age of 6.0; PPVT–R, 115 (09); TOLD–P,2, 111 (08); MLU, 4.6 (.79); CMMS, 115 (14).

At the end of the study, one child had dropped from the SLI group, one from the LC group, and one from the AC group. It can be seen that the SLI group, on average, continued to show lower performance on conventional language measures at the first time of measurement for this study, a relative deficit that persisted throughout subsequent rounds of measurement, although nonverbal intelligence measures remained within the normal range. As noted earlier, the SLI group lagged behind the control groups throughout the study on a composite measure of grammatical tense marking, computed as the arithmetic mean across 7 measures of elicitation and spontaneous performance on the following morphemes: third-person singular present tense -s, past -ed, BE, and DO. At the outset of this study, that is, during Round 3 of the data collection on these children, the group means were as follows: SLI, 67% (19); LC, 88% (10); and AC, 95% (06).

Grammaticality Judgment Task

The task format was that of a story description task in which the children observed as the examiner acted out a simple story with small toy objects. Two robot toys were introduced as people from outer space, and the child’s task was to tell the examiner if their language was “good” or “not so good.” The examiner pretended to speak for the robots and asked the child after each target clause to make a judgment. Our pilot testing established that this format was interesting to the children and eliminated spurious truth-value judgments. The online story description put the event frame, the reference frame, and the speech frame in alignment (i.e., the statements were about immediate events) to enhance interpretability of judgments. This is important because we wanted to ensure as much as possible that the children were interpreting the items as intended. For example, an item such as “he is hiding” was spoken as the examiner showed a toy bear hiding behind a box; likewise, “he is cough” was spoken as the toy bear was coughing, complete with the examiner uttering a pretend cough; “he is jump” as the toy bear was demonstrated jumping.

Experimental items were preceded by introductory items in which feedback was given to establish the kinds of judgments of interest. The 15 practice items were regular plurals (present or omitted) and –ing progressive (present or omitted). The -ing progressive was also included in the experimental items as a check for attention and as a control for expected correct judgments. No feedback was given during the experimental items. Items could be repeated if the child had not given a judgment and appeared to be distracted or requested repetition.

Examiners were trained to administer the task in a way to minimize prosodic cues for the different item types. For example, they were trained not to highlight in any way the wrong form choice items of the BA items, but to deliver them in a prosodic contour similar to the items with agreement. Examiner presentation during data collection was monitored by a lab supervisor and by audiotaped recordings. Any deviations were quickly noted and corrected. Further checks were made on the
reliability of examiners' coding of children's responses on the tasks. This was carried out by having multiple examiners view the same video recording of a child's performance. Pairwise comparisons across three raters found 100% agreement on the child's responses.

The actual items are listed in Table 2. The notation used in the table for items ungrammatical in the adult grammar is the standard asterisk (*). An asterisk followed by a plus sign (+) indicates the adult ungrammatical items that are predicted to be acceptable to children in an OI/EOI grammar; all others with an asterisk are predicted to be unacceptable. All items are simple declarative sentences similar to the prototypical clauses attempted by young children with SLI. A total of 35 items was given, broken out as follows: 10 items each in the three grammatical categories of adult grammar, OI grammar, and BA grammar, equally distributed across lexical affixes and forms of *BE*; an additional five items, with dropped progressive -ing, were included as a control for possible attentional drift during the testing. The items are intermixed in the actual protocol, following the story line. The maximum clause length was six words; the minimum was two words, with the majority in the three- to five-word range.

**Data Collection**

The grammaticality judgment task was included in the protocols of the longitudinal study, beginning at the third round of measurement (cf. Rice et al., 1995). Data collection was every 6 months, once in the fall and once in the spring. A total of five rounds of data was collected. The control items for the DI grammar were added at the second time of testing. Because of ceiling effects and limited time, the grammaticality judgment task was dropped from the protocol for the AC group at the final time of measurement.

**Dependent Variable**

The analyses followed the logic of signal detection theory. A child's response was "yes/good" or "no/not so

<table>
<thead>
<tr>
<th>Table 2. Experimental items.</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEX</td>
</tr>
<tr>
<td>Adult utterance items</td>
</tr>
<tr>
<td>Wow, he looks big</td>
</tr>
<tr>
<td>He landed on the box</td>
</tr>
<tr>
<td>Maybe he drinks milk</td>
</tr>
<tr>
<td>Maybe he busted his head</td>
</tr>
<tr>
<td>Maybe he loves you</td>
</tr>
<tr>
<td>Optional infinites</td>
</tr>
<tr>
<td>*+He eats toast</td>
</tr>
<tr>
<td>*+Now the bear want a drink</td>
</tr>
<tr>
<td>*+Maybe he like Pepsi better</td>
</tr>
<tr>
<td>*+Maybe he need a Band-Aid</td>
</tr>
<tr>
<td>*He look happy now</td>
</tr>
<tr>
<td>Bad agreement</td>
</tr>
<tr>
<td>*I likes toast</td>
</tr>
<tr>
<td>*You jumps on the box</td>
</tr>
<tr>
<td>*I drinks milk</td>
</tr>
<tr>
<td>*I wants Pepsi</td>
</tr>
<tr>
<td>*You needs to help him</td>
</tr>
</tbody>
</table>

\*+Dropped -ing*  

*He is cough*  
*He is jump*  
*The bear is look for something*  
*He is cry*  
*He is smile*

Note. * = ungrammatical in the adult grammar. *+ = ungrammatical in the adult grammar but predicted to be acceptable to children with an OI/EOI grammar.  
*This set of items was added at the second time of testing.*
good." For illustration, let us assume a contrast between the adult grammar (AG) items and the optional infinitive (OI) items. If the child was presented with items from the AG (e.g., "he runs," "he is running") and said "yes," this was a "hit." If the items were from the OI grammar (e.g., "he run," "he running") and the child said "yes," this was a "false alarm." If a child said "no" to an AG item, this was a "miss," but if a child said "no" to an OI item, this was a "correct rejection.

Other investigators report that children have a bias to say "yes" in their responses, presumably a prosocial bias toward acceptance (cf. Bishop, 1997, p. 135; McDaniel & Cairns, 1996, p. 252). Note that this bias works against detection of the predicted rejections of a BA grammar. Following other investigators (e.g., Weckerly et al., 1997; Wulfek, 1993), we computed an adjusted measure of sensitivity, A', which can be interpreted as the proportion of correct responses attainable in a two-alternative, forced-choice procedure (Green, 1964; Grier, 1971). In other words, it is as if the children were asked which of the two alternatives of a given contrast were grammatical. This was calculated by determining the values x, the proportion of false alarms, and y, the proportion of hits. In effect, this calculation adjusts for a child's bias toward accepting items. Following the formula taken from Linebarger, Schwartz, and Saffran (1983), A' = 0.5 + (y - x) / (1 + y - x) / 4y (1 - x). Perfect discrimination (i.e., following AG) yields an A' of 1.00. "Yes" responses to both adult grammatical and OI items (i.e., a full OI grammar that accepts both OI and AG items) yields an A' of .50, as does 50% "yes" responses to each of the options. A tendency to say "no" to all items can yield A' values of less than .50.

Results

Illustration of Raw Data

Table 3 provides an illustration of the data, laid out as percentages of hits, misses, false alarms, and correct rejections, for the third time of measurement for each group and for each grammar contrast. The third time of measurement was chosen because it is midway, where there is interesting variation in the groups and the conditions, and it shows representative patterns of responding. It is clear that the responses match the overall expectations from the theoretical perspective of an EOI grammar, but what is needed is a summary statistical representation to capture the probabilities across the cells, adjusting for the tendency for young children to default to "false alarms" (i.e., false "yes") responses. The following results are reported in terms of A values. The same patterns hold for percentage correct outcomes, although the evidence is not as succinctly and accurately summarized as with the A measure.

Predictions About Judgment of OI Grammar Over Time (Predictions 1, 2, 3, 6, and 8)

It was expected that children in an OI period would judge items as grammatical that are grammatical in the adult grammar (henceforth, AG; i.e., "he is running" and "he runs"). At the same time, they would also accept items that are grammatical in an OI grammar but not in the adult grammar (i.e., "he running" and "he run"). To evaluate this prediction, children's performance on the AG items was compared with that of the OI items.

The mean A' values for each of the groups for the

<table>
<thead>
<tr>
<th>OI Grammar</th>
<th>+/-AGR</th>
<th>Drop -ing Grammar</th>
</tr>
</thead>
<tbody>
<tr>
<td>he runs/he run</td>
<td>he is running/he is run</td>
<td></td>
</tr>
<tr>
<td>Adult grammar</td>
<td>OI</td>
<td>+/-AGR</td>
</tr>
<tr>
<td>L Group</td>
<td>Yes</td>
<td>68%</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>18%</td>
</tr>
<tr>
<td>3N Group</td>
<td>Yes</td>
<td>85%</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>8%</td>
</tr>
<tr>
<td>5N Group</td>
<td>Yes</td>
<td>95%</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>4%</td>
</tr>
</tbody>
</table>
adult versus OI grammar are shown in Figure 1. This figure reports A’ at five times of measurement. It is clear that by 6 years of age, unaffected children knew this grammatical distinction very well; the AC group at the first time of testing performed above .90, and they continued at that level throughout. On the other hand, 4-year-old children at the end of an OI period were likely to accept the OI items as grammatical, but the group showed sharp progress toward the adult grammar during the time period tested (i.e., they moved from A’ = .65 to about .90 during the time interval sampled). The SLI group, in comparison, performed below their younger language-matched controls, a generalization that became increasingly strong over subsequent times of measurement. That is, although the SLI group did improve over time, their progress toward the adult grammar did not keep pace with that of the younger comparison group. For example, the group average hovered around .65–.70 for ages 6–7 years.

The findings shown in Figure 1 address Predictions 1, 2, 3, 6, 8, and 9. That is, it was expected that A’ for AG/OI would be low for children in an OI/EOI stage (i.e., the SLI and LC groups), but at adult levels for children in an adult grammar (i.e., the AC group; Predictions 1, 2, 3). As shown in Table 1, it was also expected that the SLI group would perform below the two control groups, most interestingly, the younger language control group (the LC group). Finally, it was expected that the children’s performance would show progress toward the adult grammar over the times of measurement (Prediction 8), but the SLI group would show more protracted change (Prediction 9). Recall from Table 1 that this prediction differentiates the EOI model from the production constraint model. To examine overall Group × Time measurement effects, a MANOVA with group as a between-subjects factor and time of measurement as a within-subjects factor was carried out, with a significant main effect for group \(F(2, 54) = 15.6, p < .001\) and time \(F(3, 52) = 15.6, p < .001\) and a Group × Time interaction \(F(6, 104) = 2.4, p < .052\). To examine the interaction of Group × Time, at each time, a one-way ANOVA of group was carried out. As expected, at each time of measurement there was a significant group effect. Of greater interest is the performance of the SLI group as compared to the AC and LC groups, which was examined by carrying out a series of one-way ANOVAs with two groups. These are summarized in Table 4, which shows F values, degrees of freedom, p values, eta square effect size estimates, and power estimates expressed as Cohen’s (1988) f values (where effect sizes of .10 are regarded as “small,” .25 as “medium,” and .40 as “large”) at each round of data collection. At each time, the SLI group had significantly lower OI A’ values than the AC group, an expected finding that, as shown in Table 4, had a consistently large effect size. In the comparison of the SLI and younger LC group at Time 1, the SLI group = LC. At Times 2, 3, and 4, the SLI group lagged behind the LC group, with significant differences. At the final time of measurement, the SLI group has closed the gap to a nonsignificant difference. The effect sizes are lower than the SLI/AC contrast, as expected, but

<table>
<thead>
<tr>
<th>Time</th>
<th>F</th>
<th>Sig</th>
<th>Eta sq (R)</th>
<th>f</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>56.42</td>
<td>.000</td>
<td>.59</td>
<td>1.21</td>
<td>1.00</td>
</tr>
<tr>
<td>2</td>
<td>37.66</td>
<td>.000</td>
<td>.49</td>
<td>.98</td>
<td>1.00</td>
</tr>
<tr>
<td>3</td>
<td>32.51</td>
<td>.000</td>
<td>.45</td>
<td>.91</td>
<td>1.00</td>
</tr>
<tr>
<td>4</td>
<td>23.18</td>
<td>.000</td>
<td>.38</td>
<td>.78</td>
<td>1.00</td>
</tr>
<tr>
<td>5</td>
<td>.15</td>
<td>.230</td>
<td>.04</td>
<td>.20</td>
<td>.22</td>
</tr>
</tbody>
</table>

Table 4. Group comparisons for OI grammar at each round: obtained F values, significance, effect size, and power estimates.

Because the AC group was not assessed at the final time of measurement, the full Group × Time analysis was carried out through the fourth time of measurement. In the following analyses within round, all rounds of data were examined for the SLI/LC comparison.

Verb type is not considered in the analyses reported here. In preliminary analyses, a main effect for verb type appeared in which performance on the BE items is more accurate than on the lexical items. This difference does not interact with group status or with time of measurement. Because it does not contribute to the predictions specified here or further our understanding of how the groups differ, the analyses of grammatical differences reported here are collapsed across the two verb types.
still, as shown in Table 4, they are large for Times 2, 3, and 4, when the SLI group was 6.5 to 7.5 years of age and the younger children were 4.5 and 5.5.

**Predictions About BA Grammar: Predictions 5 and 7**

With regard to the comparison of the OI grammar (OI) and the bad agreement (BA) items (Predictions 5 and 7), our prediction was that the children with an OI/EOI grammar (i.e., the LC and SLI groups) would be better at detecting ungrammaticality (i.e., deviations from the adult grammar) for items in which subject-verb agreement is faulty (i.e., BA), such as “he are mad,” than they would be for items in an OI grammar, such as “he brown.” Recall that, as shown in Table 1, the OI/EOI model is differentiated from the BA and the input processing deficit models for this prediction. These comparisons (collapsed across verb type) are shown in Figure 2 (for the SLI group), Figure 3 (LC), and Figure 4 (AC), in which A’ for OI (as reported in Figure 1) is now graphed on the same figure with A’ for BA. (The third line, for DI, will be discussed in the following section.) To test these group differences, a MANOVA was run with group as a between-subjects variable and grammar (OI/BA) and time as within-subjects variables. As expected, there was a main effect for group [F(2, 54) = 25.17, p < .001] and for grammar [F(1, 54) = 60.78, p < .001], a Group × Grammar interaction [F(2, 54) = 9.69, p < .001], and a main effect for time [F(3, 52) = 27.14, p < .001].

To address the predicted better performance of a BA grammar over an OI grammar for each group at each time of testing, a series of two-tailed t tests were carried out. These are reported in Table 5, where it can be seen that, for the SLI group, their judgments of BA items were more sensitive than for the OI items for four consecutive rounds, as predicted. For the younger group, the expected difference was evident for three of the five times of measurement. At the first round, when they were 3.5–4.5 years old, the differential sensitivity was not detected. For unknown reasons, at the third round, the mean levels of performance moved closer together. The older children, whose sensitivity for both kinds of items was higher than the other two groups (throughout, near ceiling levels), also showed a statistically significant advantage for the BA grammar at the last two rounds of testing. The first and third round differences between the OI grammar and the BA grammar may be a reflection of the underlying “yes” bias that works against detection of a BA grammar even at generally high levels of performance, such that in the few times when inattention or other factors are operative, children default to “yes.” Overall, the differentiated levels of A’ for the OI/BA distinction show that the SLI group, as well as the control children, can detect BA, thereby
revealing that they do draw upon underlying representations of person/number features and of the requirement that these match for subjects and their predicates. This is at the same time that the children with SLI, as well as the controls, treat sentences with dropped TNS as grammatical. The Group × Grammar interaction was attributable to a greater difference between the OI and BA A' values for the SLI group relative to the controls. This is illustrated in Figure 5 for Time 3, where it is clear that the SLI group, as expected, is much more likely to detect BA violations of the adult grammar than OI violations.

Predictions About Acceptable Ungrammaticality (OI) and Unacceptable Ungrammaticality (BA/DI)

Next we examined Prediction 4 and compared it to Prediction 5 (i.e., compared three adult-ungrammatical contexts, “he is run” [DI] vs. “he running away” [OI] vs. “he are mad” [BA]). The prediction was that children would know that DI was incorrect and BA was incorrect, but that OI was acceptable. Recall from Table 1 that this prediction differentiates the EOI model from the input processing deficit model.

The A' values are graphed for each group for each time of measurement in Figures 2, 3, and 4, where we now focus on the DI line as well as BA and OI. It is clear from the figures that DI and BA ungrammaticality, as predicted, are very likely to be detected. The comparison of the DI grammar with OI and BA is reported in Tables 6 and 7, respectively, for each group at each time of measurement. As predicted, the SLI group consistently shows greater sensitivity to a dropped present progressive -ing than to an OI omission, as is true for the control groups as well, with the exception of the youngest children (i.e., the first round of testing with the LC group). On the other hand, as predicted, the DI/BA sensitivity is equivalent at each round for each of the groups. This is clear evidence that each group of children can detect subtle grammatical violations when they are not allowed in their underlying grammatical representations.

Growth Curve Modeling

The same logic of the growth curve modeling carried out by Rice et al. (1998) for the children's production data was followed here. As before, we modeled the growth of the 3N and the SLI group, where sufficient variation over time existed for modeling. For each of the three grammars, we addressed these questions: (a) Do children grow linearly in their grammaticality judgments? (b) Do children grow nonlinearly in their grammaticality judgments? (c) Are there individual differences in the rate and type of growth? (d) Are there group differences (between the SLI and 3N children) in the rate and type of growth? (e) Are there individual and group differences in growth after covarying out individual differences at the outset due to nonverbal intelligence (CMMS), comprehension vocabulary (PPVT-R), mother's education, or MLU? and (f) How much variance is accounted for by the initial predictors? Following Rice et al. (1998), the covariate data were drawn from the children's status at initial identification, approximately one year prior to the first time of measurement for grammaticality judgments.

Table 6. OI/DI grammar differences for each group at each time of measurement.

<table>
<thead>
<tr>
<th>Time</th>
<th>OI</th>
<th></th>
<th>DI</th>
<th></th>
<th>SLI</th>
<th></th>
<th>LC</th>
<th></th>
<th>AC</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NA</td>
<td></td>
<td>NA</td>
<td></td>
<td>NA</td>
<td></td>
<td>NA</td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-3.33</td>
<td>.003</td>
<td>-4.86</td>
<td>.641</td>
<td>NA</td>
<td></td>
<td>NA</td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-5.84</td>
<td>.000</td>
<td>-4.26</td>
<td>.000</td>
<td>-2.76</td>
<td>.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-4.05</td>
<td>.001</td>
<td>-2.61</td>
<td>.018</td>
<td>-2.17</td>
<td>.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-4.24</td>
<td>.000</td>
<td>-2.51</td>
<td>.022</td>
<td>NA</td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7. BA/DI grammar differences for each group at each time of measurement.

<table>
<thead>
<tr>
<th>Time</th>
<th>OI</th>
<th></th>
<th>DI</th>
<th></th>
<th>SLI</th>
<th></th>
<th>LC</th>
<th></th>
<th>AC</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NA</td>
<td></td>
<td>NA</td>
<td></td>
<td>NA</td>
<td></td>
<td>NA</td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-.67</td>
<td>.40</td>
<td>.71</td>
<td>.49</td>
<td>-7.6</td>
<td>.46</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-1.22</td>
<td>.24</td>
<td>-1.93</td>
<td>.07</td>
<td>NA</td>
<td></td>
<td>NA</td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-2.27</td>
<td>.79</td>
<td>-5.5</td>
<td>.62</td>
<td>NA</td>
<td></td>
<td>NA</td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-2.28</td>
<td>.78</td>
<td>-1.17</td>
<td>.26</td>
<td>NA</td>
<td></td>
<td>NA</td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Modeling was carried out by hierarchical linear modeling (HLM) procedures (Bryk & Raudenbush, 1992; Bryk, Raudenbush, & Congdon, 1994) for each of the three A’ measures, EOI/OI, BA, and DI. HLM is a mixed-model analysis where predictor covariates are considered as fixed variables and the children’s growth over time (both linear and nonlinear) are random effects. As random effects, a different linear and nonlinear regression coefficient is obtained for each child. It is these regression coefficients that provide information concerning the rate of change of each child. Further, variability among children’s regression coefficients provides information concerning individual differences in rate of change.

As in the production analysis, four nested models were evaluated: Model 1 examined whether the four covariates (mother’s education, child CMMS, PPVT–R, MLU) were significantly related to judgments of a grammar over time; Model 2 added to Model 1 a linear growth term in order to determine whether, on the average, children (regardless of group membership) changed linearly across time; Model 3 added to Model 2 a nonlinear (quadratic) growth term in order to determine whether, on average, children (regardless of group membership) changed nonlinearly over time; Model 4 added to Model 3 two interaction terms. One term represented the interaction between group membership and linear growth; the second, the interaction between group membership and nonlinear growth. The significance of either term would indicate that the groups grew differently. For example, perhaps the 3N group experienced nonlinear growth, whereas nonlinear growth was absent for the SLI group.

The outcomes for the OI and BA grammars are summarized in Table 8, which reports the coefficients and standard errors (in parentheses) for each term in the four models. What is striking in the outcomes is that findings for Models 1, 2 and 3 are the same for the grammaticality judgments of OI and BA grammars, and they replicate the findings for the production outcome data of Rice et al. (1996; i.e., the only significant predictor is the MLU value at the outset of the study, and both linear and nonlinear components are evident in the growth curve). As was the case for the production outcomes, mother’s education, child nonverbal intelligence, and child receptive vocabulary do not predict growth in

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th></th>
<th></th>
<th>Model 2</th>
<th></th>
<th></th>
<th>Model 3</th>
<th></th>
<th></th>
<th>Model 4</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OI</td>
<td>BA</td>
<td></td>
<td>OI</td>
<td>BA</td>
<td></td>
<td>OI</td>
<td>BA</td>
<td></td>
<td>OI</td>
<td>BA</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>.80**</td>
<td>.73**</td>
<td></td>
<td>( .28)</td>
<td>( .23)</td>
<td></td>
<td>(.069)</td>
<td>(.06)</td>
<td></td>
<td>(.00)</td>
<td>(.00)</td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td>-.10</td>
<td>-.08</td>
<td></td>
<td>(.059)</td>
<td>(.06)</td>
<td></td>
<td>(.00)</td>
<td>(.00)</td>
<td></td>
<td>(.00)</td>
<td>(.00)</td>
<td></td>
</tr>
<tr>
<td>Intelligence</td>
<td>.001</td>
<td>.001</td>
<td></td>
<td>(.00)</td>
<td>(.00)</td>
<td></td>
<td>(.00)</td>
<td>(.00)</td>
<td></td>
<td>(.00)</td>
<td>(.00)</td>
<td></td>
</tr>
<tr>
<td>PPVT</td>
<td>.00</td>
<td>.00</td>
<td></td>
<td>(.00)</td>
<td>(.00)</td>
<td></td>
<td>(.00)</td>
<td>(.00)</td>
<td></td>
<td>(.00)</td>
<td>(.00)</td>
<td></td>
</tr>
<tr>
<td>MLU</td>
<td>.082*</td>
<td>.09**</td>
<td></td>
<td>(.03)</td>
<td>(.03)</td>
<td></td>
<td>(.02)</td>
<td>(.01)</td>
<td></td>
<td>(.01)</td>
<td>(.01)</td>
<td></td>
</tr>
<tr>
<td>Mother education</td>
<td>.005</td>
<td>.01</td>
<td></td>
<td>(.03)</td>
<td>(.03)</td>
<td></td>
<td>(.00)</td>
<td>(.00)</td>
<td></td>
<td>(.00)</td>
<td>(.00)</td>
<td></td>
</tr>
<tr>
<td>Linear</td>
<td></td>
<td>.05***</td>
<td>.06***</td>
<td></td>
<td>(.01)</td>
<td>(.01)</td>
<td></td>
<td>(.00)</td>
<td>(.00)</td>
<td></td>
<td>(.00)</td>
<td>(.00)</td>
</tr>
<tr>
<td>Quadratic</td>
<td></td>
<td>-.02***</td>
<td>-.02***</td>
<td></td>
<td>(.00)</td>
<td>(.00)</td>
<td></td>
<td>(.00)</td>
<td>(.00)</td>
<td></td>
<td>(.00)</td>
<td>(.00)</td>
</tr>
<tr>
<td>Linear x Group</td>
<td></td>
<td></td>
<td></td>
<td>-.03</td>
<td>-.11</td>
<td></td>
<td>(.08)</td>
<td>(.10)</td>
<td></td>
<td>(.00)</td>
<td>(.00)</td>
<td></td>
</tr>
<tr>
<td>Quadratic x Group</td>
<td></td>
<td></td>
<td></td>
<td>.003</td>
<td>.01</td>
<td></td>
<td>(.00)</td>
<td>(.00)</td>
<td></td>
<td>(.00)</td>
<td>(.00)</td>
<td></td>
</tr>
<tr>
<td>Proportion reduction in variance</td>
<td>.25</td>
<td>.52</td>
<td>.92</td>
<td>.88</td>
<td>.99</td>
<td>.99</td>
<td>.99</td>
<td>.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .05. **p < .01. ***p < .001.
grammaticality judgments. The question of whether the two groups show the same growth curves yields a finding of no group differences for either the linear or quadratic terms for either the OI grammar or the BA grammar. Recall that the production data outcomes showed no difference in the groups' growth curve (cf. Rice et al., 1998). As before, the full model accounted for almost all the individual variance (i.e., .99 for both OI and BA grammars).

The values of A for the dropped -ing grammar are not included in Table 8 because the lack of variability made it impossible to estimate values for Model 3 (quadratic) and Model 4 (group interactions). The values for Model 1 showed that MLU was the only significant predictor, and for Model 2 showed that there is linear change over time for two groups, as is evident in Figures 2 and 3.

Discussion

Let us return to Table 1 and begin with the possibility that children with SLI have limited metalinguistic abilities in general (Kamhi & Catts, 1988; Kamhi & Koenig, 1985). Recall that the comparison group for this conclusion was an age-matched group. The findings reported here were arguably not inconsistent with that conclusion, in that for many of the grammaticality judgments observed, the SLI group performed below that of their age peers. At the same time, the more specific findings suggest that the conclusion is misleading in some important ways. For the DI grammar, and even for the BA grammar, by 7 years of age the performance of the SLI group is, on average, at a high level, at an A' of .90 or above, which would be too high to be characterized as "lacking" metalinguistic abilities for relatively fine-grained grammatical judgments. So although the SLI group continues to be statistically lower than the AC group, it is at the same time at relatively high levels of performance. Furthermore, the performance of children with SLI compared to the younger LC group is very similar on the BA and DI contrasts, indicating that, once general language level was controlled, the apparent deficit in metalinguistic ability was much diminished. Finally, the grammatical sensitivity of the children with SLI is differentially sensitive, such that OI items are much less likely to be judged ungrammatical than BA or DI items. Thus, our conclusion is that metalinguistic judgments of affected children depend upon the grammatical structure assessed and control for overall language levels. It would be too strong to conclude that children with SLI lack the ability to make metalinguistic judgments across the board, as part of a general deficit in an ability to reflect on language properties, independent of language level and the degree of affectedness of particular structures. Instead, although there was some unexplained variation in their performance, the SLI group demonstrated they could differentiate between grammatically allowable deviations from the adult grammar and deviations not consistent with an immature OI grammar on these very finely tuned grammatical judgments.

In our interpretation, the control comparisons of BA and DI involve incorrect form choice resulting from lack of subject-verb agreement or incorrect omission of a progressive -ing affix. We note that in our production evidence, reported in Rice et al. (1996), we find very few errors of this kind and, more generally, very few errors of wrong forms or inserted morphemes. We must consider some alternative interpretations of the children's detection of BA and DI grammatical errors, and possible inadvertent confounds in the construction of the experimental items.

One possibility is an inadvertent semantic confound, such that the children's detection of the DI items is attributable to their interpretation of "he is cough" as unacceptable because it could be construed as intending "cough" as a predicate noun, which children would be likely to reject on semantic grounds, instead of a rejected DI from the verb. First, recall that these children showed that they could readily detect omission of regular plural -s from nouns, a detection so likely that it was used as a training context. So we know they can detect omitted morphemes. Several additional observations mitigate against the possible semantic confound interpretation of rejected DI. One is that as the items were administered, children saw the relevant action demonstrated. It is very unlikely they would construe "cough" to mean "a cough" as the examiner was demonstrating coughing. Second, not all the five items of this type lend themselves to this possibility, and it seems even less plausible for "the bear is look for something." Inspection of item performance shows minimal variation across items. A related alternative explanation is that children could regard the DI items as unacceptable because of a possible substitution of "is" for a modal such as "can" or "will." Again, the semantic referencing of the item administration rules against this interpretation; it is obvious that the bear is in the act of coughing, not just capable of it ("can") or going to ("will").

Another possible alternative explanation is that the SLI group's rejection of BA items is because there is an overt error of form choice, an "added" illicit form, in contrast to the DI items which involve an omitted obligatory element. One way to examine this possibility is to
compare items of the type “he are spitting it out/he am hurt” with “he is cough.” The former involve a possible “added” form, whereas the latter, we argue, involve a dropped morpheme. We calculated for each five-item set the percentage correct rejections and carried out a series of t test comparisons of the two different item sets for each round, for the SLI group and the LC controls. Not one of these comparisons, for either group, reached statistical significance. Moreover, children with SLI judge the omission of plural -s as deviant, as we have already pointed out. Thus the idea that it is omitted material that is judged acceptable by the SLI group could not explain the results for plural -s either. Our conclusion, although the matter is not conclusively settled by these early findings, is that it is most likely that the children’s rejections of incorrect items are not primarily attributable to “added” versus “dropped” elements.

Let us now return to Table 1 for the primary four models under analysis in this study. As predicted by the EOI model, children with SLI and their younger LC comparison group were, on average, more likely to accept predictably allowable deviations from the adult grammar (such as “he running away” and “he look happy”) than non-allowable deviations from the adult grammar (such as “he is run” or “he are running”), a difference that persisted over repeated times of measurement. At the same time, as noted above, the children with SLI were sensitive to the non-allowable deviations (i.e., they were likely to reject “he is run” or “he are running”). So it is not the case that they were simply likely to accept all utterances, regardless of grammatical configuration. Furthermore, the differentiation between the two grammars remained throughout the period sampled, indicating that, while the general performance levels improved, the OI grammar remained as an option. The trailing OI grammar was what clearly set the affected group apart from the control children, as was evident in the pattern of effect sizes obtained in the group comparisons, where it was clear that the strongest effects for group differences were for the OI grammar, over time. Furthermore, these are persistent findings, evident across different rounds of testing.

The remarkable similarity of the growth curve modeling for the grammaticality judgment outcomes reported here, as compared to the earlier modeling for the EOI production data, adds further support to the expectation that the grammaticality judgments draw upon the underlying grammatical representations of an EOI grammar. For both sets of models, MLU is a significant predictor of growth, but mother’s education, child nonverbal intelligence, and child receptive vocabulary do not predict growth, suggesting that growth in this part of the grammar is independent of the other predictors. Unlike the BA grammar judgments and the production outcomes, the judgments of an OI grammar showed differences from the LC control group, attributable to a prolonged period of little change in judgments before growth accelerated. This suggests that the bias to accept OI violations of the adult grammar may be especially slow to be relinquished by affected children, but because this seems to rest on the performance of the SLI group at one time of measurement, it should be replicated in future studies before strong conclusions can be made.

The deficit in subject-verb agreement account of SLI predicted that affected children would perform equally poorly on the OI and the BA grammatical variations from the adult grammar. With one important exception, that did not prove to be the case. Except for the first time of measurement (where the obtained t values approach but do not meet the conventional .05 level of probability), the SLI group showed greater sensitivity to the BA than the OI items and consistently showed this pattern over the five rounds of assessment. This finding would be difficult to explain if the grammars of children with SLI allowed for deficits in subject-verb agreement that would allow nonagreeing verb forms to appear. However, maybe it is the case that at the younger ages, sensitivity to BA violations does not exceed sensitivity to OI violations. Support for this possibility is found in the fact that the younger LC group also did not show greater sensitivity to BA violations at Time 1. On the other hand, they show a somewhat mixed picture on this comparison over time, again dropping to near significant levels at Time 3. Consideration of the AC group, whose performance is above .90 throughout on both grammars, reveals that they also sometimes (Times 3 and 4), but not always, show a difference in favor of the BA grammar. Thus, the interpretive issue, considering all three groups, is whether we assume that children have a fluctuating enhanced sensitivity to BA violations. This would pose a large challenge to theoretical accounts. Our interpretation is that, for all the children, as expected, detection of BA violations of the adult grammar are more likely than detection of OI violations, a pattern clearly evident in the group means, but the tendency of children to default to a “yes” response, even when this tendency is adjusted arithmetically, makes it difficult to detect at levels of statistical significance. Parsimony seems to call for the assumption that the advantage for the BA grammar is evident throughout the different times of measurement, but is sometimes just below the level of statistical detection.

The production constraint model generates the expectation that the children’s underlying grammars are intact and therefore the children should be sensitive to OI and BA grammatical variations. This clearly was not the case. Instead, the children with SLI were quite accepting of OI items, although they were, at the same
time, able to show that they could detect grammatical variations not allowed in an OI grammar. Furthermore, this pattern was extended over multiple times of measurement, so the apparent differences were not a one-time manifestation detectable only at a certain time of measurement. Of course, as noted by Bishop (1994), putative production constraints could interact with grammatical complexity, such that, in production, more complex utterances would be more likely to show omitted elements. But this interpretation does not apply here for two reasons. One reason is that interpretation, not production, is tapped in the grammaticality judgment tasks. As noted earlier, the prediction is that interpretation will be unaffected. Second, the items for the different grammars are arguably of the same complexity, at least as indexed by utterance length and the fact that simple declarative clausal structures are used throughout. Furthermore, BA violations ("I am going") are probably more "complex" than OI violations ("he going"), since in the former, features of the subject must be "matched" against features of the verb in order to judge the item's well-formedness, whereas in the latter only an obligatory feature of the sentence must be detected as present or missing. Thus, a "complexity" theory would predict exactly the opposite result, that BA violations would be more difficult to judge than OI violations.

The limited input processing mechanisms model also lacks an explanation of how the affected children could be more sensitive to the BA items than to the OI items, where both depend on very finely tuned grammatical distinctions evident in small, unstressed surface forms of relatively short duration (e.g., "he is a bear" vs. "I am coming back" or "running away"). One possibility is that if a limited input processing mechanism caused a child to drop morphemes because they were not encountered in input (i.e., a child "skipped over" them), then such a limitation would apply to the small grammatical units involved in the BA grammar, and the two grammars would be equally likely to be accepted or rejected, although this was not the case. Alternatively, as suggested by Leonard (1998), an input limitation could cause affected children to have difficulty building morphological paradigms, in which the subject/verb agreement demands are misinterpreted, such that they could assume that "am," for example, could appear with third-person singular subjects. That would also lead to the prediction that children's judgments of BA should be at the same level of sensitivity as the OI grammar, which was not the case. We can consider another assumption of the surface account, that the affected children's underlying grammars are the same as those of unaffected younger children, in which case we would expect to see similar levels of sensitivity to the OI items in the SLI and LC groups. Instead, the findings revealed that the SLI group's performance was significantly below that of the LC group for 3 of the 5 rounds for the OI grammar.

Does this mean that there is no possible role for processing influences on the grammaticality judgments of children with SLI? Surely not, and that is not our claim. Our point is that the omissions in surface structure of the very particular kind we have studied here, that is, the omission of tense-marked English forms, does not yield to a simple limitation of input processing of the sort described in current accounts, given the pattern of BA detection, nor does it yield to a production constraint notion of processing, as noted above. Wexler (in press-a, in press-b) discusses the difficulty of accounting for the OI stage with a "processing deficit" account, and those arguments carry over to the grammaticality judgment data. This certainly does not exhaust the realm of processing factors. As noted earlier, in this study, children with SLI performed at lower levels than their age-matched controls, a difference that could be attributable to an interaction of the underlying grammatical limitations with attentional processes that are mediated by grammatical representations. For more complex grammatical constructions, processing factors involving memory may be implicated. Further, it is likely that if the input conditions are degraded by noise or other processing complications, performance on grammaticality judgment items will be affected.

All things considered, our conclusion is that children with SLI, and younger unaffected children, are likely to interpret simple clausal structures, under simple input conditions, according to the same principles that guide their productions (i.e., that they are willing to accept clauses with deletion of grammatical tense, as in an OI grammar) but are more sensitive to clauses with morphological variations not allowed by an OI grammar. It remains for future studies to reveal additional details to clarify the robustness and range of this conclusion. Of clinical import, the strong effect sizes for comparison with their age peers on the OI grammar point toward a comprehension index, as well as production measures, of an EOI period as a possible clinical identifier of children with SLI. We need not defer possible clinical applications as we investigate further the interpretive nuances and empirical facts associated with the grammaticality judgments of children with SLI.

Acknowledgments

This study was supported by National Institute on Deafness and Other Communicative Disorders Award RO1 DCO1803 to Mabel L. Rice and Kenneth Wexler. We express special appreciation to Karla R. Haney, Chien J. Wang, Laura Smith, and Carol Schedak for data collection and to Esther Lerner for data management and assistance with data analysis. We thank Scott Hierhoffer for the growth curve analyses. We also greatly appreciate the assistance of...
Patsy Woods. We are especially grateful for the support of the parents, who agreed to share their children for the purpose of investigation, and to the many teachers, school administrators, and preschool/day care providers who assisted us in the study. Although their contributions are worthy and each special in their own way, it is not possible to list names here because this list now numbers over 100 attendance centers. Preliminary reports of this investigation were presented at the 22nd Annual Boston University Conference on Language Development, November 1997 and the 57th Annual Convention of the American Speech-Language-Hearing Association, November 1997.

References


Received April 20, 1998
Accepted February 8, 1999
Contact author: Mabel L. Rice, Child Language Doctoral Program, University of Kansas, 1932 Dole, Lawrence, KS 66045. Email: mabel@ukans.edu