# Development of copula and auxiliary BE in children with Specific Language Impairment and younger unaffected controls*

KATE L. JOSEPH
LUDOVICA SERRATRICE  } *University of Manchester*
GINA CONTI-RAMSDEN

ABSTRACT

Children with Specific Language Impairment (SLI) have well-known difficulties with the obligatory use of verbal tense-marking morphology which they use optionally for a longer period of time than typically developing children. Copula BE and auxiliary BE are two tense markers that have been shown to be problematic for children with SLI, but their status as independent allomorphs is somewhat controversial in the literature. In the present study we argue for copula BE and auxiliary BE as two separate tense markers showing different developmental curves both for children with SLI and for younger unaffected controls. It is suggested that copular versus auxiliary status, morphological form, and the frequency of the lexical construction containing BE all affect children's provision of copular and auxiliary forms. Implications for identifying variables constraining optionality are discussed.

INTRODUCTION

Children with Specific Language Impairment (SLI) have long been

---

known to experience extraordinary difficulties with verbal tense-marking morphology, so much so that some researchers have argued for the use of tense as a clinical marker of SLI (Rice & Wexler 1996). According to Rice, Wexler and colleagues (Rice & Wexler 1996, Rice, Wexler & Cleave 1995, Rice, Wexler & Hershberger 1998), the linguistic profile of children with SLI can be characterized as an Extended Optional Infinitive (EOI) period. Similarly to younger, typically developing children, children with SLI go through a phase in which they use tense markers appropriately but only optionally so. For example, rather than supplying an inflected finite verb form such as in (1a) and (2a) they will produce (1b) and (2b):

(1a)  She walked home.
(1b)  *She walk home.
(2a)  Thomas likes ice cream.
(2b)  *Thomas like ice cream.

In the case of copula and auxiliary BE, the non-finite equivalent is the omission of the tense marker altogether:

(3a)  I am tired.            (fulfilled copular context)
(3b)  *I tired.             (unfulfilled copular context)
(4a)  I am going home.        (fulfilled auxiliary context)
(4b)  *I going home.         (unfulfilled auxiliary context)

Agreement errors are relatively rare, a fact that Rice, Wexler and colleagues interpret as an indication that, although children with SLI do not know that finite forms are obligatory in main clauses, they do know how to use finite forms appropriately when required. However, unlike in the course of typical development, in the case of children with SLI this phase of optional omission of finiteness lasts for a protracted period of time, hence the definition of Extended Optional Infinitive period. In the EOI framework, the prediction is that in principle all tense markers will be equally affected by the phenomenon of optionality. Because they all spell out abstract temporal properties, there is no *a priori* reason why some tense markers should be more vulnerable to omission than others. Recently Rice *et al.* (1998) have advocated the notion of Composite Tense according to which the proportions of use in obligatory contexts of third person singular –*s* inflections, regular past tense –*ed* inflections, auxiliary DO, copula and auxiliary BE forms are collapsed to form one single measure of children's proficiency in the marking of tense. In particular, copula and auxiliary BE are treated as one and the same morpheme, as previously reported in Rice *et al.* (1995) and Rice & Wexler (1996).

This approach contrasts markedly with numerous other accounts in the literature where copula and auxiliary BE are treated as two separate tense markers, despite their allomorphy. In his seminal study of three typically developing children Brown (1973: 269) was the first to note that 'the copula-auxiliary distinction is a significant one'. A number of studies investigating the grammatical proficiency of children with SLI have indeed considered production in obligatory contexts for copula and auxiliary BE separately. Leonard (1995) reported that, both for the children with SLI and the typical MLU controls in his study, the proportion of copular forms in obligatory contexts was higher than the corresponding proportion for auxiliary BE. Moreover, the MLU controls outperformed the children with SLI on both measures. Hadley & Rice (1996) found the order of acquisition of copula and auxiliary BE to be the same for a group of children with SLI and a group of younger controls: in both sets of children copula BE emerged before auxiliary BE. In a cross-sectional study on the effect of contractibility on the production of copula and auxiliary BE, Cleave & Rice (1997) reported that contractible forms were supplied significantly more accurately than uncontractible forms, both for the children with SLI and the younger typical controls. Analyses of both spontaneous and probe data provided further support for the finding that children with SLI supply the copula less accurately than younger controls (Leonard, Eyer, Bedore & Grela 1997).

In sum, those studies that have kept copula and auxiliary BE separate have consistently reported that the copula both emerges earlier and is produced more accurately than auxiliary BE in obligatory contexts, and that on both measures children with SLI perform worse than younger typically developing children. None of these studies, however, have so far attempted to investigate in detail the extent to which different person/number combinations contribute to children's use of the two tense markers, or whether copula and auxiliary BE are more likely to be produced in certain constructions than others. We believe that by exploring patterns in the use and omission of copula and auxiliary BE we can begin to shed some light on the nature of the optionality of tense marking in development.

In a top-down theoretical approach as that underlying the EOI, the expression of an abstract syntactic feature such as tense must be uniformly realized, regardless of its morphological form or the lexical context in which it is used. Contrary to the theoretical stance of the EOI model, the bottom-up constructivist approach adopted here is based on the assumption that children's use of inflected forms is constrained by a number of extra-syntactic factors other than the realization of an

abstract tense feature. We believe that factors such as frequency and morphological and lexical specificity play a non-trivial role in predicting the pattern of optionality of different tense-marking morphemes.

Previous work has also revealed differences between the acquisition of the morphological forms of BE. Brown (1973) notes that for the three children in his study the contracted copula and auxiliary BE –s and –z allomorphs were generally supplied more accurately than either the –m or –r morphemes. Hadley & Rice (1996) also report that for the children with SLI they studied, third person singular contexts dominated in the use of copula and auxiliary BE. The predominance of third person singular forms seems to hold cross-linguistically; Pizzuto & Caselli (1992) report a similar finding for use of the copula by three typically developing Italian-speaking children. This asymmetry in the use of copula and auxiliary BE forms indicates that, although children's production is initially inconsistent, it is not random.

*Factors affecting the provision of copula and auxiliary BE*

The consistent finding that certain forms are more likely to be used than others – namely third person singular forms of BE are supplied more accurately in obligatory contexts than any other form – has a number of possible explanations. Firstly, it has been suggested that the third person inflection can be considered the most neutral or unmarked form for person reference (Lyons 1977). Secondly, Western middle-class parents tend to engage in frequent naming games with their children, both producing and eliciting a large number of utterances containing the copula of the type 'What's that?', 'It's a dog', and so on. Finally, adults and children playing together commonly provide a running commentary of their actions that typically results in a large number of auxiliary BE forms in third person contexts once again, e.g., 'What's the doggie doing?', 'Is he eating his lunch?', 'Look, he's chasing the cat'.[1]

A number of different studies have put forward possible explanatory variables for the differences in provision both between copula and auxiliary BE and within each morpheme. These factors include whether the context is nominative or predicative, verb argument structure complexity and the specific lexical construction in which BE appears.

With respect to patterns of copular use and omission, a recent proposal by Becker (2000) identified clear-cut differences between the

[1] We believe this to be particularly true of situations in which parents are aware that their task is to try to elicit as many utterances as possible from their children.

production of copular forms in nominal predicative and locative predicative contexts in four typically developing English-speaking children. Becker's findings show that for all four children the copula was significantly more likely to be omitted in locative predicative contexts than in nominal predicative contexts. The children were thus more likely to omit the copula in a locative context and produce (5b) instead of (5a), rather than in a nominal context using (6b) instead of (6a):

(5a)  My pen is down there.
(5b)  *My pen down there.          (Peter, file 6)
(6a)  I'm a big boy.
(6b)  *I a big boy.                (Adam, file 10)

Although the pattern is quite striking there is reason to question the reliability of the identification of the locative contexts in Becker's analysis. While in (6b) it is reasonable to assume that there is a missing first person copula inasmuch as 'big boy' is in a clear predicative relation with respect to the subject 'I', in (5b) the interpretation of the telegraphic utterance is not so straightforward. While it is possible that the child is describing a state of affairs (e.g., the pen is in a box), it is also possible that he is expressing the result of an action that he is carrying out (e.g., he is putting the pen in a box). The example below from a child with SLI in the current study illustrates this possibility.

(7)    CHI:  that on roof.          (Harry, 3:10.16)
       (Child is putting a piece on the toy house)

       MOT: put it on the roof you mean?

Especially in the initial stages of acquisition when children's utterances are characterized by frequent ellipsis, it is an extremely arduous task to try to identify which elements are missing. One possible reason why the locative contexts in Becker's study contained such a low proportion of copular forms is that in fact some of these were not proper copular contexts at all, but rather stranded resultative small clauses.[2]

Another attempt to identify a consistent pattern of use and omission of tense markers is a recent study on auxiliary omission by children with SLI and MLU controls by Grela & Leonard (2000). In this carefully controlled experimental investigation with English-speaking

---

[2]  See Lebeaux (2000: 59–60) for the proposal that telegraphic utterances containing a noun and a preposition (e.g., 'boot off', 'him out') are to be analysed as resultative small clauses as in 'I want boot off', 'I want him out'.

children with SLI and MLU controls, the authors proposed that one variable predicting patterns of auxiliary BE use in progressive constructions was the complexity of the verb's argument structure. Even after utterance length had been accounted for, both the children with SLI and the younger MLU controls in the study were more likely to omit an auxiliary with ditransitive verbs than with simple transitive or intransitive verbs, and the children with SLI also performed marginally worse than the MLU controls. However, there were no significant differences between the children's provision of the auxiliary BE in the transitive and intransitive constructions. This finding is especially relevant because it enables the researcher to formulate more sophisticated predictions as to when a child is more or less likely to omit a certain tense marker.

Another factor that might contribute to an explanation of children's optional use of copula and auxiliary BE is the interaction between the use of finite forms and that of specific lexical constructions. Construction is defined here as a linear string where in addition to the copula at least one other pre-copular or post-copular element is fixed (see Method section for a more comprehensive definition of the four constructions of interest in the present study). Kuczaj (1985/6) found that his two sons used uncontracted copula BE only in specific constructions during the early stages of acquisition. For example, Abe used his first forms of copula BE only in declarative contexts such as 'these/those are + x', whereas Ben initially used the copula exclusively in sentence-final position, e.g., 'Here it is'. Thus, Abe's apparently high provision of copula 'are' can be explained by the observation that early on he would supply this form only in contexts beginning with 'those' or 'these'. If the use of such verb forms is tied to specific lexical constructions, one would expect to find significant differences in the pattern of omission of verb forms as a function of lexical constructions. For example, it could be the case that a child produces copular forms in 80% of obligatory contexts, but that she only ever does so with 'What's that?' and 'That's mine'. This would clearly be an extreme case of lexical specificity as the expected pattern of copular use could be entirely predicted by the identification of these two constructions. Clearly, children's language becomes less lexically-specific over time; however, we believe that the repeated use of low-level lexically specific schemas plays a major role in the early stages of acquisition (Pine, Lieven & Rowland 1998). In the case of children with SLI, their poorer distributional learning abilities (Conti-Ramsden & Jones 1997) could lead to a greater reliance on a small number of lexically-specific frames.

*Aims and objectives of this paper*

In the present paper we use longitudinal data for three preschool children with SLI and a group of eleven typically developing MLUw (Mean Length of Utterance in words) controls to study the development of copula and auxiliary BE forms. Our main aim is to identify patterns in the omission of copula and auxiliary BE that allow us to constrain the optionality of tense marking in a predictable manner.

Firstly, we investigate whether the two morphemes are produced at different levels in obligatory contexts. In line with previous studies that have kept the two morphemes separate (Hadley & Rice 1996, Leonard 1995), we predict that copula BE will be produced more often in obligatory contexts than auxiliary BE for both groups of children.

. Secondly, we examine children's mastery of the morphological paradigm of copula and auxiliary BE by breaking down their use of forms into the six person and number combinations. Maternal speech data are also analysed in order to assess to what extent children's production reflects a distributional bias in the input.

Finally, we investigate whether the provision of copula and auxiliary BE varies as a function of lexical constructions. Four constructions were identified by the pre-copular or pre-auxiliary element featuring: noun, personal pronoun, other pronouns (e.g., demonstrative, existential and indefinite pronouns) and *wh*-question words. We predict that not all subjects will be equally likely to be used with copula and auxiliary BE. In particular, in the case of the copula we expect a large proportion of third person singular subjects because of the naming bias we know to exist in early adult-child interaction. A related prediction is that semi-formulaic frames such as 'What's that?' will frequently feature in children's production, especially in the initial stages of acquisition when they rely more heavily on a small repertoire of stock phrases.

METHOD

*Participants*

Approximately 10 months of spontaneous data were analysed for two boys and one girl with SLI, aged 3;1.9, 3;5.0 and 4;0.9 at the start of the study and 3;10.22, 4;5.2 and 4;8.30 at the end of the study. The children were recruited from speech and language therapists in the north-west of England on the basis of receptive and expressive language delay in the absence of any additional problems. That is, none of the children showed any evidence of hearing impairment or frank

neurological damage. They showed normal levels of social interaction as measured by the Autistic Screening Questionnaire (Berument, Rutter, Lord, Pickles & Bailey, 1999) and they all scored above 90 on the Leiter Performance Scale (1979). Receptive language was measured using the Reynell Developmental Scales and all children scored below the 16th centile, providing evidence of severe difficulties with comprehension (Edwards, Fletcher, Garman, Hughes, Letts & Sinka 1997). In addition, the children were reported by speech therapists as being at the early stages of multiword speech, demonstrating severe expressive problems. The children were visited at home twice a month where they were audiorecorded for an hour-long informal play session with their mothers. The investigator was sometimes present as a participant observer.

The 11 typically developing children in the control group were taken from the Manchester corpus (Theakston, Lieven, Pine & Rowland 2001). The children ranged in age from 1;8.22 at the beginning of the study to 2;4.21 at the end of the study. They were visited twice every three weeks at home where one hour of mother-child interaction was audiorecorded.

Only a subset of the data available from the Manchester corpus was analysed for the purposes of the present paper. With the exclusion of one child, Ruth, who produced very few finite forms throughout the period of data collection and therefore provided little relevant data for this analysis, the remaining 11 children were matched with the children with SLI for MLUw (Mean Length of Utterance in words) starting at 1.8. Three developmental stages were identified at the following MLUw levels: stage 1 (MLUw 1.8), stage 2 (MLUw 2.4), stage 3 (MLUw 2.9).

For each child the first session that came closest to MLUw 1.8 was selected, and six more sessions were included afterwards at intervals of approximately five weeks. A total of 77 hour-long sessions were therefore analysed for the MLUw controls. Summary statistics for the three children with SLI and the eleven MLUw controls are provided in Table 1.

*Speech corpora*

All recordings were orthographically transcribed in CHAT format as described in the CHILDES manual (MacWhinney 1995). Complete morphological tagging was automatically created using the MOR and POST programs on CHILDES. The criteria for inclusion for children's utterances were ones that there were: (a) fully intelligible; (b) used spontaneously (i.e., were neither self repetitions nor imitations); and

TABLE 1. *Summary statistics*

| Group | Child | Age range | MLUw range | No. sessions analysed | No. utterances | No. utterances analysed |
|---|---|---|---|---|---|---|
| SLI | Bonnie | 4;0.09–4;8.30 | 2.2–3.4 | 14 | 6810 | 490 |
| | Harry | 3;5.0–4;5.2 | 1.7–3.3 | 24 | 8117 | 375 |
| | Nathan | 3;1.9–3;10.22 | 1.6–3.6 | 22 | 12337 | 779 |
| MLUw controls | Anne | 1;10.07–2;7.01 | 1.5–2.8 | 7 | 4218 | 132 |
| | Aran | 2;0.2–2;8.12 | 1.7–3.4 | 7 | 3724 | 248 |
| | Becky | 2;2.15–2;10.25 | 1.7–3.1 | 7 | 4681 | 235 |
| | Carl | 1;8.22–2;5.8 | 1.9–3.1 | 7 | 5568 | 312 |
| | Dominic | 2;1.11–2;9.26 | 1.6–2.7 | 7 | 4535 | 216 |
| | Gail | 1;11.27–2;8.6 | 1.6–3.0 | 7 | 3580 | 176 |
| | Joel | 1;11.22–2;9.13 | 1.5–2.8 | 7 | 3071 | 223 |
| | John | 1;11.15–2;7.24 | 1.8–2.8 | 7 | 3123 | 70 |
| | Liz | 2;0.07–2;8.14 | 1.6–3.5 | 7 | 3713 | 165 |
| | Nicole | 2;4.21–3;0.10 | 1.6–3.0 | 7 | 3772 | 144 |
| | Warren | 1;10.6–2;6.23 | 1.9–3.8 | 7 | 3522 | 227 |

(c) not strings from songs or nursery rhymes. Imitations and self-repetitions were defined as the exact repetition of a multiword utterance in the immediately preceding line. For the mother, the criteria for inclusion were utterances that were (a) fully intelligible, and (b) not strings taken directly from songs or read directly from books.

The corpora were used to search for strings containing copula and auxiliary BE. In addition, searches were conducted for strings containing obligatory contexts for both forms of BE. For copula BE, it was necessary to search manually through strings of utterances that should have contained a copular form. Thus an obligatory context for the copula was defined as an utterance in which the copula was omitted. For auxiliary BE, it was possible to search for all utterances containing progressive participles. It should be noted that, in line with previous studies, only aspectual auxiliary BE (rather than passive auxiliary BE) was investigated in this study. Thus, an obligatory context for auxiliary BE was defined as an utterance containing a progressive participle without an accompanying auxiliary.[3]

In addition, detailed lexical searches were conducted to identify both the specific constructions containing copula and auxiliary BE forms and the constructions that constitute obligatory contexts for these forms. Four constructions were identified by the pre-copular or pre-auxiliary element featuring: noun (e.g., 'my kitty is nice', 'Tom and Jane are leaving'), personal pronoun (e.g., 'you are naughty', 'she's laughing'), other pronouns (demonstrative, existential and indefinite pronouns, e.g., 'that is blue', 'there is a dog', 'someone's hiding') and wh-question words (e.g., 'what's that noise?', 'where's he going?').

In order to compare mother and child distributions of BE forms, a composite mother group was formed. This was done by pooling independent samples of seven sessions from the 11 mothers of the MLUw controls (i.e., these samples did not overlap with the samples of child speech data analysed). The decision was made to exclude the mothers of the children with SLI from the overall mother group as there has been no research into the possible similarities and differences in the lexical, morphological and syntactic characteristics of the language

---

[3] Although the omission of auxiliary BE in progressive contexts is simple to identify, Stan Kuczaj has pointed out to us that the child may not always produce the progressive form with the correct –ing inflection (i.e., she might instead produce a bare stem form in a progressive context). Thus, the number of auxiliary BE progressive contexts reported could be an underestimation of the actual number that were produced by the children.

produced by mothers of typically developing children and mothers of children with SLI.

RESULTS

To begin, it should be noted that copula BE was used far more often than auxiliary BE by all the children in the study. For the MLUw controls as a group, copula BE accounted for 86.3% of all BE forms. The corresponding percentages for the children with SLI were 90.6% for Bonnie, 90.1% for Harry and 83.6% for Nathan. The predominance of copula BE held over the three stages, decreasing only slightly at stage 3 for both the MLUw controls and the three children with SLI. It should also be noted that the mothers of the MLUw controls used copula BE much more frequently than auxiliary BE, with the copula accounting for over 70% of all BE forms for the mother group. The higher frequency of copula BE had consequences for the results to follow, in that certain analyses could not be conducted for auxiliary BE due to an insufficient number of tokens.

*Provision of copula and auxiliary BE in obligatory contexts*

The proportions of copula and auxiliary BE forms that were supplied in obligatory contexts by the MLUw controls over time were compared. Table 2 presents the mean number of child utterances per session, with the numbers of obligatory contexts for copula and auxiliary BE and the percentage of contexts that were finite for the MLUw group and the three children with SLI.[4]

The most common way of reporting the proportions of forms supplied in obligatory contexts in the literature is in terms of percentages. However, we believe that it is also important to take into account the absolute frequencies with which the relevant forms are produced and omitted. The impact of the frequency of fulfilled and unfulfilled contexts for a given verb form should not be easily dismissed by only considering proportional measures. A measure of absolute frequency is important in itself as an indicator of how centrally, or peripherally, a given form features in the acquisition process.

---

[4] The sessions for the children with SLI were sometimes slightly longer than one hour, and this may partly explain why the children with SLI produced higher mean numbers of tokens than the MLUw controls. Note also that the length of sessions for Harry were reduced to around 40 minutes at stage 3, and therefore fewer data were collected for the child in this stage.

TABLE 2. *Mean number of utterances per session, mean number of copula and auxiliary BE contexts per session and percentage of finite contexts*

| Group | Child | Stage | Mean no. utterances per session | Copula BE | | Auxiliary BE | |
|---|---|---|---|---|---|---|---|
| | | | | No. contexts (tokens) | Proportion of finite tokens (%) | No. contexts (tokens) | Proportion of finite tokens (%) |
| SLI | Bonnie | 1 | 488.6 | 300 | 76.3 | 36 | 44.4 |
| | | 2 | 502.5 | 186 | 85.5 | 42 | 50.0 |
| | | 3 | 471.4 | 346 | 90.5 | 83 | 44.6 |
| | Harry | 1 | 424.5 | 158 | 43.7 | 19 | 15.8 |
| | | 2 | 333.1 | 301 | 64.5 | 79 | 27.8 |
| | | 3 | 227.5 | 85 | 74.1 | 27 | 40.7 |
| | Nathan | 1 | 636.3 | 219 | 26.1 | 64 | 1.6 |
| | | 2 | 477.3 | 401 | 67.3 | 162 | 38.3 |
| | | 3 | 580.7 | 566 | 86.8 | 190 | 51.1 |
| MLUw controls | Group mean | 1 | 571.6 | 187.2 | 32.6 | 16.0 | 10.3 |
| | | 2 | 579.8 | 323.2 | 68.6 | 76.8 | 44.6 |
| | | 3 | 590.6 | 238.6 | 84.8 | 62.8 | 46.1 |
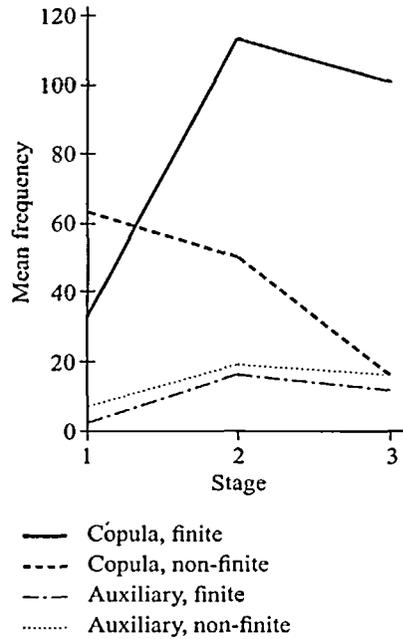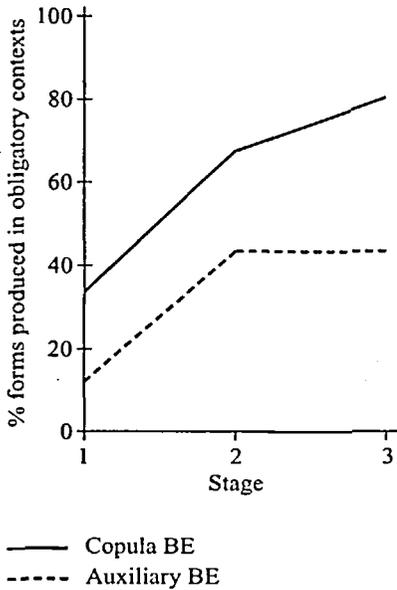
Fig. 1a.  Percentages of copula and auxiliary BE forms used in obligatory contexts by MLUw controls

Fig. 1b.  Mean frequency of finite and non-finite forms of copula and auxiliary BE used by MLUw controls

For example Hoff-Ginsberg (1992) argues that more consideration should be give to measures of total frequency in addition to proportional measures in the study of child-directed speech.

It is for this reason that provision of copula and auxiliary BE is presented not only with respect to percentages in obligatory contexts in Fig. 1a but also with respect to the absolute frequencies of fulfilled finite and unfulfilled non-finite contexts in Fig. 1b.[5] Examples of fulfilled and unfulfilled copula and auxiliary BE contexts are given in (3) and (4).

---

[5] The decline in the overall number of copula contexts between time 2 and time 3 is most likely related to the fact that some of the children in the group started using a number of other verb constructions such as modal constructions.

Figure 1a shows that the MLUw controls were always better at producing copula BE in obligatory contexts than they were at supplying auxiliary BE where required. However, interpretation of the graph based on the percentage data would suggest there to be little difference between the rates at which provision of copula and auxiliary BE change over time. If Fig. 1b is considered, it is clear that for copula BE the number of fulfilled and unfulfilled contexts is larger than for auxiliary BE. There are also differences in the rates at which the provision of the two morphemes change over time. The simultaneous presentation of Figs 1a and 1b, which are based on the same data set, raises an important methodological issue in the assessment of grammatical morphology. That is, it can be misleading to present percentages of forms produced in obligatory contexts, especially when the percentages are based on relatively small numbers of tokens, as is the case here for auxiliary BE. Furthermore, it is somewhat problematic to compare percentages that are obtained from very different numbers of tokens. Henceforth, percentages in obligatory contexts will be presented in the text in line with previous studies; however, all figures will plot absolute frequency data in order to provide a more informative picture of the longitudinal development of the two morphemes.

Statistical analyses were conducted to evaluate the patterns observed in the provision of copula and auxiliary BE by the MLUw controls. It should be noted that only the theoretically interesting results will be reported here. For example, where a significant interaction was found between morpheme and stage, simply denoting children's increased use of both morphemes and unfulfilled contexts for morphemes over time, this was not reported. Analyses were conducted first for the MLUw controls as a group and then individually for the three children with SLI.

Log transformations were carried out as the frequency distribution of the data for the MLUw control group was highly positively skewed. A three-way, repeated measures ANOVA was conducted with morpheme (copula BE, auxiliary BE), finiteness (finite, non-finite) and stage (1, 2, 3) as factors. A significant third-order interaction was found between morpheme, stage and finiteness $(F(2,20) = 8.0, p < 0.01)$, indicating that copula and auxiliary BE were produced with significantly different levels of finiteness over time. It should be noted that the significant differences between the two morphemes held at stage 2 $(F(1,10) = 19.1, p < 0.001)$ and stage 3 $(F(1,10) = 39.6, p < 0.001)$, but that at stage 1 there was no significant difference between the proportions of finite copula and auxiliary BE forms, probably because of the low levels of performance on both morphemes.

In order to investigate further the third-order interaction between morpheme, stage and finiteness, two separate, two-way ANOVAs were conducted on copula and auxiliary BE to analyse the relationship between stage and finiteness within the two BE paradigms. For copula BE, there was a significant interaction between stage and finiteness ($F(2,20) = 10.5$, $p < 0.001$), reflecting the increase in the number of finite forms together with the decrease in the number of non-finite forms from stage 1 to 2. It is clear that the main advances made by the MLUw controls occur between stages 1 and 2 (MLUw 1.8–2.4) and that their provision of finite copular forms changes less considerably between stages 2 and 3 (MLUw 2.4–2.9).

The results for the ANOVA on auxiliary BE highlight the different developmental trajectories of the two BE allomorphs. For auxiliary BE, there was only a significant main effect of stage ($F(2,20) = 15.3$, $p < 0.001$), which is a function of the overall increase in auxiliary BE contexts from stage 1 to 2. It is important to note that the lack of a significant interaction between stage and finiteness signals that the proportions of finite and non-finite auxiliary BE contexts remained relatively stable over time, and the lack of a main effect for finiteness shows that there was little difference between the frequencies with which finite and non-finite forms in auxiliary BE contexts were produced. This pattern is illustrated in Fig. 1b by the flat, parallel lines denoting finite and non-finite contexts for auxiliary BE. Thus, the MLUw controls seemed to be supplying auxiliary BE in around 50% of obligatory contexts. Whereas they show clear improvement in their provision of copula BE in obligatory contexts over time, they demonstrate little evidence of knowledge that auxiliary BE is required with progressive forms, even at stage 3. Corresponding analyses of the provision of copula and auxiliary BE in obligatory contexts were carried out individually for the three children with SLI. Figures 2, 3 and 4 plot the frequencies of finite and non-finite contexts for copula and auxiliary BE over time for each child.[6]

Visual inspection of Figs 2, 3 and 4 clearly shows that the combined number of fulfilled and unfulfilled contexts for copula BE is far greater than for auxiliary BE. In addition, while the three children with SLI were still supplying auxiliary BE in 40.0–51.1% of obligatory contexts at stage 3, they all showed a clear improvement in their provision of

---

[6] For Harry the smaller number of copula and auxiliary contexts is to be ascribed to his preference for modal constructions at stage 3.
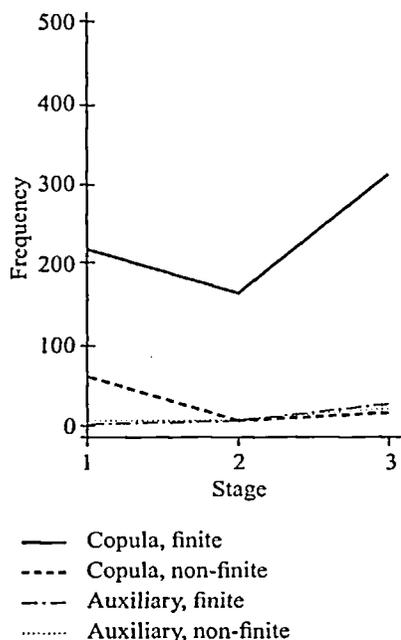
Fig. 2. Frequency of finite and non-finite forms of copula and auxiliary BE used by Bonnie (SLI)
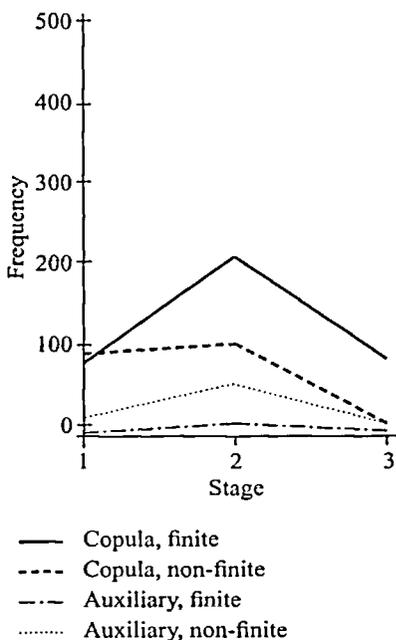
Fig. 3. Frequency of finite and non-finite forms of copula and auxiliary BE used by Harry (SLI)

copula BE, reaching between 74.0% and 90.0% in obligatory contexts by the end of the study.

Hierarchical log linear models were used to analyse the data for each child with SLI separately. For Nathan there was significant third-order interaction between morpheme, stage and finiteness ($\chi^2 = 13.8$, $p < 0.01$). Thus, Nathan showed significant variation in his marking of finiteness for copula and auxiliary BE over time. In order to identify the source of this variation, separate hierarchical log linear analyses were conducted for Nathan's use of copula and auxiliary BE. Significant interactions were found between finiteness and stage for both copula BE ($\chi^2 = 268.5$, $p < 0.001$) and auxiliary BE ($\chi^2 = 65.1$, $p < 0.001$). These results reflect the fact that Nathan improved in his marking of finiteness on both copula and auxiliary BE over time. It should be noted
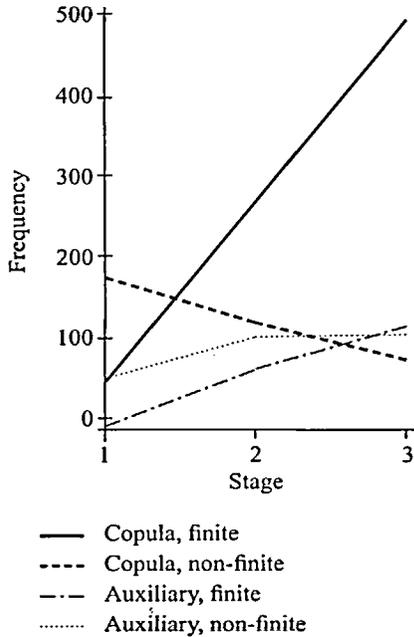
Fig. 4. Frequency of finite and non-finite forms of
copula and auxiliary BE used by Nathan (SLI)

that that there was a significant stage-finiteness interaction for auxiliary
BE for Nathan but not for the MLUw controls. At first glance, it might
appear that Nathan improved on his provision of auxiliary BE in
obligatory contexts over and above the MLUw controls. However, on
closer inspection of the data, the significant interaction between stage
and finiteness for Nathan is likely to be a function of his very low
percentage of auxiliary BE forms in progressive contexts at stage 1
(1.6%), as he was still marking finiteness on auxiliary BE in only 51.0%
of obligatory contexts by stage 3. Thus the pattern shown by Nathan is
broadly similar to that of the MLUw controls, although the controls do
not show any significant change in their provision of auxiliary BE over
time (from 10.3% at stage 1 to 46.1% at stage 3) due to the fact that
they started off with a higher level of provision in stage 1.

There were no significant third-order interactions for either Bonnie or Harry, suggesting that their provision of copula and auxiliary BE in obligatory contexts did not differ significantly over time. The data for both children, however, yielded significant interactions between morpheme and finiteness (Bonnie: $\chi^2$ = 109.0, $p$ < 0.001; Harry: $\chi^2$ = 50.2, $p$ < 0.001). This interaction signals that copula and auxiliary BE are produced with differing levels of finiteness, the copula being produced more reliably in obligatory contexts than auxiliary by both Bonnie and Harry.

*Morphological analysis of BE*

Given the clear advantage of copula BE over auxiliary BE for both the MLUw controls and the children with SLI, the use of BE forms within the copula and auxiliary paradigms was investigated.[7] Tables 3 and 4 present the distribution of agreement inflections on the present tense copula and auxiliary BE. It is clear that both for the children with SLI and the MLUw controls, the impressive performance on the copula, in both its contracted and uncontracted forms, is almost exclusively accounted for by knowledge of the third person singular form (see Table 3). The MLUw controls produced a mean of 37.1 copular forms per one-hour session (range 7.7–70.7). Similarly, the children with SLI used a mean of 35.2 copular forms per one-hour session (range 13.7–49.9). The mean proportion of third person singular copular forms for the MLUw controls was 90.0%, (range 68.0–98.3%). The figures are comparable for the children with SLI, whose mean use of third person singular copular forms was 91.4%, (range 86.3–95.1%). The two groups of children in the current study produced copula BE with both similar frequency and similar distribution of morphological forms.

The data in Table 3 are collapsed over time. However, it is also important to consider the emergence of the individual contracted and uncontracted forms of copula BE over the three stages. With respect to uncontracted forms, Bonnie and Harry showed a strong bias towards 'is' at stages 1 and 2, although more 'are' forms were produced by the end of data collection. Whereas Harry used 'am' a handful of times during each stage, Bonnie never used this morpheme. Nathan used

---

[7] Although the amount of data for the MLUw controls and the children with SLI is roughly equivalent, we are aware that, due to the smaller number of children in the SLI group, the role played by individual variation in this group may be more important.

TABLE 3. *Percentage of present tense copula BE forms for three children with SLI and eleven MLU controls*

| Group | Child | No. copula BE forms | Relative percentage frequency | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | 1 sg. | 2 sg. | 3 sg. | 1 pl. | 2 pl. | 3 pl. |
| SLI | Bonnie | 699 | 0.4 | 1.8 | 95.1 | 1.3 | 0.0 | 1.3 |
| | Harry | 329 | 5.5 | 0.9 | 86.3 | 0.0 | 0.0 | 7.3 |
| | Nathan | 913 | 2.2 | 0.9 | 92.8 | 0.1 | 0.0 | 4.0 |
| | Group mean | 647.0 | 2.7 | 1.2. | 91.4 | 0.5 | 0.0 | 4.2 |
| MLUw controls | Anne | 320 | 0.9 | 3.1 | 92.5 | 0.0 | 0.0 | 3.4 |
| | Aran | 234 | 0.4 | 0.9 | 98.3 | 0.4 | 0.0 | 0.0 |
| | Becky | 454 | 2.0 | 5.5 | 83.0 | 3.1 | 0.0 | 6.4 |
| | Carl | 413 | 0.0 | 5.9 | 93.2 | 0.0 | 0.0 | 0.9 |
| | Dominic | 151 | 3.9 | 4.6 | 88.0 | 0.0 | 0.0 | 3.3 |
| | Gail | 495 | 0.8 | 0.4 | 97.6 | 0.4 | 0.0 | 0.8 |
| | Joel | 97 | 2.0 | 15.5 | 68.0 | 13.4 | 0.0 | 1.0 |
| | John | 176 | 0.0 | 1.1 | 98.3 | 0.0 | 0.0 | 0.6 |
| | Liz | 272 | 2.9 | 8.8 | 84.2 | 0.4 | 0.0 | 3.7 |
| | Nicole | 54 | 0.0 | 3.7 | 90.7 | 0.0 | 0.0 | 5.5 |
| | Warren | 198 | 0.0 | 1.0 | 96.5 | 0.0 | 0.0 | 2.5 |
| | Group mean | 260.4 | 1.2 | 4.6 | 90.0 | 1.6 | 0.0 | 2.6 |

TABLE 4. *Percentage of present tense auxiliary BE forms for three children with SLI and eleven MLU controls*

| Group | Child | No. auxiliary BE forms | Relative percentage frequency | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | 1 sg. | 2 sg. | 3 sg. | 1 pl. | 2 pl. | 3 pl. |
| SLI | Bonnie | 65 | 9.2 | 3.0 | 81.5 | 3.0 | 0.0 | 3.0 |
| | Harry | 36 | 33.3 | 8.3 | 50.0 | 5.5 | 0.0 | 2.7 |
| | Nathan | 158 | 25.9 | 2.5 | 69.6 | 0.6 | 0.0 | 3.4 |
| | Group mean | 86.3 | 22.8 | 4.6 | 67.0 | 3.0 | 0.0 | 3.0 |
| MLUw controls | Anne | 37 | 16.2 | 8.1 | 70.2 | 2.7 | 0.0 | 2.7 |
| | Aran | 20 | 0.0 | 0.0 | 100.0 | 0.0 | 0.0 | 0.0 |
| | Becky | 86 | 34.8 | 9.3 | 52.3 | 2.3 | 0.0 | 1.1 |
| | Carl | 120 | 4.2 | 1.6 | 93.3 | 0.0 | 0.0 | 0.0 |
| | Dominic | 9 | 33.3 | 11.1 | 55.5 | 0.0 | 0.0 | 0.0 |
| | Gail | 28 | 14.2 | 3.6 | 64.3 | 17.8 | 0.0 | 0.0 |
| | Joel | 45 | 24.4 | 15.5 | 55.5 | 2.2 | 0.0 | 2.2 |
| | John | 18 | 0.0 | 0.0 | 100.0 | 0.0 | 0.0 | 0.0 |
| | Liz | 31 | 48.3 | 9.6 | 38.7 | 0.0 | 0.0 | 3.2 |
| | Nicole | 8 | 12.5 | 25.0 | 75.0 | 0.0 | 0.0 | 0.0 |
| | Warren | 24 | 0.0 | 0.0 | 100.0 | 0.0 | 0.0 | 0.0 |
| | Group mean | 38.7 | 17.1 | 7.6 | 73.1 | 2.3 | 0.0 | 0.8 |

'am', 'are' and 'is' at all three stages, but 'is' accounted for a consistently high proportion of the data (over 80%) throughout. As a group, the MLUw control children used no 'am' forms at stage 1 and only a handful of 'am' forms at the latter two stages. 'Are' and 'is' forms were equally frequent at stage 1, but at stages 2 and 3 'is' was produced more often than 'are'. Thus, both the children with and without SLI used little or no 'am' forms early on, and exhibited a strong preference for 'is' forms over time.

With respect to contracted forms of copula BE, there was less development over time: contracted third singular forms remained the main contracted form used by all the children with SLI throughout, with only a few contracted first and second person singular forms being produced. Similarly, the vast majority of the contracted copular forms produced by the MLUw controls were third singular, accounting for over 80% of all contracted forms at all three stages.

Table 4 shows that, on average per session, both the children with SLI and the MLUw controls produced considerably fewer auxiliary BE forms than copular forms. The MLUw controls used a mean number of 5.2 auxiliary BE forms per session, (range 1.1–17.1) and for the children with SLI the equivalent figure was 4.3 (range 1.5–7.5). Like the copula, auxiliary BE use consisted largely of third person singular forms, with a mean proportion of 73.1% for the MLUw controls. However, the range was considerably wider than that reported for the copula, between 38.7% and 100.0%. For the children with SLI the mean proportion of third person singular auxiliary BE forms was 67.0%, with a range between 50.0% and 81.5%. For two of the children with SLI (Harry and Nathan), and for four of the MLUw controls (Becky, Dominic, Joel and Liz), the distribution of auxiliary forms was not so heavily skewed towards the use of third person singular forms, and a sizeable proportion of auxiliary BE forms were instead found in first person singular contexts.

The children in both groups also showed interesting developmental patterns for the different morphological forms of auxiliary BE. Bonnie and Harry used very few uncontracted forms of auxiliary BE, and 'is' was the most frequent form at each stage. Nathan used more uncontracted forms of auxiliary BE, producing 'am' and 'is' with similar frequencies, although there was little sign of development over time. The MLUw control group used very few uncontracted auxiliary BE forms at stage 1, increasing mainly in their use of second and third singular forms at stages 2 and 3. All three children with SLI used very few contracted forms of auxiliary BE, with Bonnie using mainly third singular forms, and Harry and Nathan producing a combination of first and third

TABLE 5. *Relative percentage frequencies of contracted and uncontracted copula and auxiliary BE forms in the combined corpus of eleven mothers' speech*

|                          |              | 1 sg. | 2 sg. | 3 sg. | 1 pl. | 2 pl. | 3 pl. |
|--------------------------|--------------|-------|-------|-------|-------|-------|-------|
| Copula (N = 13174)       | Contracted   | 1.7   | 2.8   | 57.4  | 0.1   | 0.0   | 1.2   |
|                          | Uncontracted | 0.3   | 3.6   | 27.0  | 1.3   | 0.0   | 4.6   |
| Auxiliary (N = 5372)     | Contracted   | 5.8   | 12.4  | 29.4  | 2.3   | 0.0   | 2.0   |
|                          | Uncontracted | 0.9   | 25.6  | 14.3  | 2.9   | 0.0   | 4.4   |

singular forms. The MLUw controls used few contracted forms of auxiliary BE at stage 1, increasing primarily in their use of first and third singular contracted forms over time.

*Distributional bias in maternal input*

Given the asymmetrical distributions of morphological forms for both copula and auxiliary BE in the children's speech, the corresponding distributions in the mothers' speech data were considered. Although we are not arguing for a simple input-output relationship to account for children's use of copula and auxiliary BE forms, if a similar distribution of morphological forms is found in the input that children are likely to hear, then the distribution of these forms might at least be partly determined by their distribution in the mothers' speech.[8] Table 5 presents the distribution of morphological forms of the copula and auxiliary BE in the maternal input. The mothers show a highly skewed distribution of copular forms, in which over half of all copular tokens are contracted third person singular forms and over 80% of all forms

---

[8] Note that by pooling together all the mothers of the MLUw controls to form one mother group we have obtained a representative sample of the statistics of British English addressed to children in the early stages of multiword utterances. For each MLUw child, his or her own mother only contributes one-eleventh of the whole input data, and in the case of the children with SLI there is no overlap at all; therefore we have minimized the likelihood that similarities might be due to context-specific effects of correlating input and output from the same recordings.

consist of contracted and uncontracted third singular forms. Third person singular contracted forms were also the most frequent for the auxiliary BE. However, there was more variation in the distribution of inflected forms, and second person singular constituted over one-third of all auxiliary BE forms in the maternal input.

Correlational analyses were performed to assess the extent to which children's use of copula and auxiliary BE mirrors the high-frequency BE forms in the maternal input. Although there are only three finite surface forms of BE (am, are, is), the data were divided with respect to the six person/number combinations. The distinction was made between contracted and uncontracted forms; therefore a total of twelve BE forms (six contracted and six uncontracted) formed the basis for the separate copula and auxiliary BE correlations.

Again, due to the highly positively skewed distributions of the frequency data, log transformations were conducted. A Pearson's pairwise correlation between the log frequencies of copula BE forms for the MLUw group and mother group yielded $r = 0.93$ (df $=11, p < 0.001$). The corresponding correlation between the SLI group and the mother group was also significant: $r = 0.83$ (df $= 11, p < 0.001$). The analyses for copula BE were also repeated after removing the figures for contracted and uncontracted third person singular forms, which accounted for around 90% of all copular tokens in both the children's and the mothers' productions. In this case, too, a significant correlation was obtained for the MLUw group and the mother group, $r = 0.90$ (df $= 9$, $p < 0.001$), and the SLI group and the mother group, $r = 0.70$ (df $= 9, p < 0.05$). This shows that the correlation is not carried by the highly frequent third person singular copular form. Significant correlations were also obtained for the log frequencies of contracted and uncontracted auxiliary BE forms, both between the MLUw group and the mother group, $r = 0.75$ (df $= 11, p < 0.01$), and between the SLI group and the mother group, $r = 0.64$ (df $= 11, p < 0.05$).

The correlations between the mother and child distributions of copula and auxiliary BE forms point to the highly skewed distributions of these forms in English and to the strong bias towards third person singular copular forms. However, it is important to note that even when the highly frequent third person singular forms were removed from the copula analysis, the relationship between the distributions remained, with a fairly fixed frequency hierarchy from frequent second person singular forms down to second person plural forms that were rarely produced by either mothers or children. The mother-child correlations for the MLUw controls and the children with SLI were of similar strength, suggesting that the morphological distribution of BE forms in

the input language accounted for a considerable proportion of the variance in the children's distribution of morphological forms of the copula (86% for the MLUw controls and 69% for the children with SLI), and a slightly lower, but still substantial amount of the variance in the children's distribution of auxiliary BE forms (56% for the MLUw controls and 41% for the children with SLI).

## Construction-based analysis of BE

Data analyses were carried out on all constructions of the type X + (copula) + Y, where X is either a subject (nominal or pronominal) or a *wh*-word, and Y is a nominal or adjectival predicate. All actual occurrences of the copula were counted as instances of finite fulfilled contexts and all constructions where the copula was omitted were counted as instances of non-finite unfulfilled contexts. All the copular constructions listed were then divided into four construction types (noun, personal pronoun, other pronoun, *wh*-construction) for the purpose of further analyses (see Method section). The same criteria were applied to the categorization of progressive constructions with or without the auxiliary BE.

Patterns of provision and omission of copula and auxiliary BE forms in these four constructions were analysed in order to ascertain whether there was a relationship between finiteness and construction over time. For copula BE, a three-way, repeated measures ANOVA was conducted on the data for the MLUw controls, with construction (noun, personal pronoun, other pronoun and *wh*-word), finiteness (finite, non-finite) and stage (1, 2, 3) as factors. A significant third-order interaction was found between construction, finiteness and stage ($F(6,60) = 4.3$, $p < 0.001$), indicating that the constructions were produced with different levels of finiteness over the three stages.

It is clear that the children's provision of copula BE in *wh*-constructions outstripped their provision of copula BE in the other three constructions, especially at stage 1. Consequently, when the *wh*-construction was removed from the analysis, there was no longer a third-order interaction between construction, stage and finiteness. That is, finiteness marking did not differ with respect to the personal pronoun, other pronoun and noun constructions over time. There were, however, differences between the provision of the copula in the noun, personal pronoun and other pronouns constructions when the three stages were collapsed, as indicated by the significant interaction between construction and finiteness ($F(2,10) = 13.4$, $p < 0.001$). The copula was produced least often in obligatory contexts in noun constructions, followed by the other pronoun and the personal pronoun constructions.

An ANOVA was also conducted on the MLUw controls' provision of auxiliary BE for the same four constructions. In contrast to the results for copula BE, there was no significant third-order interaction between construction, finiteness and stage, indicating that the constructions did not differ with respect to the production of auxiliary BE in obligatory contexts over time. There was, however, a significant interaction between construction and finiteness ($F(3,30) = 3.8$, $p < 0.05$), indicating that the four constructions contained different proportions of auxiliary BE in obligatory contexts. The pattern was the same as that for the copular constructions, where auxiliary BE forms are most likely to be supplied in obligatory contexts in *wh*-constructions, followed by the personal pronoun, other pronoun and noun constructions.

The use of copula and auxiliary BE in the four common constructions was also investigated for the children with SLI using hierarchical log linear models for each child. For Bonnie's use of copula BE in the four constructions there was a significant interaction between construction and finiteness ($\chi^2 = 19.6$, $p < 0.001$). This denotes the different levels of finiteness marking on copula BE in the four constructions. The data for the four constructions containing copula BE for Harry and Nathan yielded third-order interactions between construction, finiteness and stage (Harry: $\chi^2 = 35.1$, $p < 0.001$; Nathan: $\chi^2 = 36.5$, $p < 0.001$). These results provide evidence of the significant differences in the provision of copula BE in the four constructions over time. Nathan improved on his production of copula BE in the noun, other pronoun and *wh*-constructions from stage 1 to 2. Harry showed a main improvement on copulas in the noun construction from stage 2 to 3, whereas his performance on personal pronoun and *wh*-constructions was more stable over time.

The results for auxiliary BE show less differentiation between the four constructions. However, it should be noted that these results are based on a smaller number of tokens than the results for copula BE. For Bonnie, there was an interaction between construction and stage ($\chi^2 = 17.8$, $p < 0.01$), denoting her increased use of the personal pronoun construction over time. For Harry, there was an interaction between construction and finiteness ($\chi^2 = 26.7$, $p < 0.001$), denoting his increased use of finite auxiliary BE forms in all except for *wh*-constructions from stage 1 to 2. For Nathan there was a significant interaction between construction and finiteness ($\chi^2 = 18.1$, $p < 0.001$), indicating the different levels of finiteness marking on auxiliary BE for the four constructions. Note that Nathan performed poorly on the provision of auxiliary BE in *wh*-constructions (13.8%) in comparison with his provision of copula BE in *wh*-constructions (65.5%).

*Lexical analysis of BE*

Having established that BE forms were mainly produced in third person singular form and were likely to appear in certain constructions more often than others, it was necessary to consider the lexical diversity of the constructions in which BE was used. That is, to what extent can children's knowledge of BE be attributed to a handful of lexical items in the pre-copular position? The combined frequency of the three top-frequency frames containing the copula for the MLUw group was calculated as a percentage of all utterances containing copula BE forms. At all three stages, these high-frequency frames were 'it's + x', 'what's + x', and 'that's + x', which accounted for 44.7% of all constructions containing a copula. There was, however, individual variation among the three top-frequency frames used by the MLUw controls. 'It's + x' was the only top-frequency frame produced by all the children in the MLUw group. 'That's + x' and 'what's + x' were both used highly frequently by seven of the eleven children.

The same three frames were also used most frequently by Bonnie, accounting for 53.1% of all her copular productions. It is important to note that Nathan began to produce contracted copular forms only by stage 3. For stages 1 and 2, Nathan's high-frequency copular forms were 'it is + x', 'that is + x' and 'what is + x'. At stages 1 and 2, these uncontracted forms accounted for 50.9% of Nathan's productions containing copula BE. At stage 3, the corresponding contracted forms were used repeatedly, accounting for 71.4% of Nathan's copular use at this stage. Harry used a mixture of contracted and uncontracted copular forms frequently: 'it's + x', 'what's + x' and 'what is + x' accounting for 43.3% of all copular forms over time. Thus, both for the MLUw controls and the children with SLI, three high-frequency contracted copular forms accounted for a important proportion of all copular use. In addition to the previous finding of morphological specificity (i.e., the predominance of the third person singular form), this analysis emphasizes the lexical specificity in early use of copula BE.

The lexical analysis was further extended to investigate the reasons behind the different levels of finiteness for the four copula BE constructions. The appearance of a small number of lexical items in pre-copular position can be used to explain the pattern of provision of copular BE in personal pronoun, other pronoun and *wh*-constructions at stage 1. For the MLUw controls, recall the considerable proportion of fulfilled copula BE contexts in *wh*-constructions at stage 1 (68.0%) compared with personal pronoun (40.2%) and other pronoun (15.0%)

constructions. If the lexical form with the highest token frequency in each of the constructions is considered, it appears that use of copula BE in *wh*-constructions at stage 1 was accounted for by 'what's' 64.7% of the time. It is suggested that the higher percentage production of *wh*-constructions in obligatory contexts is a function of the children's use of the high-frequency 'what's + x' construction.

Within the personal pronoun construction, 'it's + x' accounted for 51.3% of all copula BE uses, which corresponds approximately to the provision of copula BE in this construction at just below 50%. Again, the distribution of lexical items in a construction is reflected in the children's ability to produce the required BE form in obligatory contexts. It should be noted that this explanation has mainly been used for stage 1, where there were clear differences between the provision of the copula. Analyses of lexically-specific constructions may be less applicable to the later stages of development. Unfortunately the numbers of individual lexical constructions containing auxiliary BE do not provide a reliable basis for an equivalent analysis to be conducted.

The relationship between a handful of lexical items within the personal pronoun, other pronoun and *wh*-constructions and provision of copula BE was investigated for each of the children with SLI. Like the typically developing children, Harry relied on the use of a single high-frequency frame in *wh*-constructions at stage 1. Specifically, 'what's + x' was the most frequent *wh*-construction (85.7% of all *wh*-constructions), which corresponds to the high percentage of copula BE in obligatory contexts in this construction at stage 1 (97.7%). Bonnie was fairly accurate in her provision of copula BE in all four constructions over time (ranging from 67.4% to 98.2% in obligatory contexts). We suggest that her impressive performance is related to the repeated use of high-frequency pre-copular lexical items. Like the controls, 'it's + x', 'that's + x' and 'what's + x' accounted for between 48.4% and 62.9% of Bonnie's productions in each individual construction. Unlike the MLUw controls and the other two children with SLI who all used mostly contracted copular forms, Nathan showed an unusual transition from favouring uncontracted copular frames at stages 1 and 2 to producing a majority of contracted copular frames at stage 3. For Nathan at stage 1, the construction with the largest number of finite copular forms was the personal pronoun construction (50.0%), in which 'it is + x' accounted for 80% of all tokens in this construction. At stage 3, Nathan's provision of copula BE was high for all three constructions (ranging from 69.6% to 84.8%), which corresponds to the large proportion of tokens in each construction containing the top frequency lexical item (56.1% to 81.4%).

It is interesting to note that there were similarities between the mother and child frequency distributions of the four constructions containing copula and auxiliary BE.[9] The noun construction made up less than 7% of all copular constructions used by the mothers, MLUw controls and children with SLI alike. The *wh*-construction was used with moderate frequency (20–34% of all copular constructions) by the mothers and both groups of children, with the other pronoun construction being used slightly more often (33–36%). Finally, the personal pronoun construction accounted for between 25% and 40% of all mother and child copular constructions. Thus, in the inventory of copular expressions used by mothers and by the two groups of children, constructions with a pronoun featured prominently, while constructions with a subject NP were comparatively rare.

There were also similar patterns for mother and child use of constructions containing auxiliary BE. The construction used least frequently was the other pronoun construction (2–4% of all auxiliary BE constructions). For all three groups, the personal pronoun construction was the most frequent, appearing in 46–59% of all auxiliary BE utterances.

DISCUSSION

Data from children with SLI and younger MLUw controls were used to investigate patterns in the provision of copula and auxiliary BE. First, we assessed whether copula and auxiliary BE are produced with similar levels of provision in obligatory contexts. Second, we explored the possibility that optionality is lexically constrained by the entrenchment of low-level schemas.

*Provision of copula and auxiliary BE*
In the EOI model proposed by Rice, Wexler and colleagues the two BE morphemes are collapsed into one single measure of composite tense together with the third person singular -*s* inflection, the regular past

---

[9] Mothers' use of the copula was classified here in terms of the different constructions in which it appeared, regardless of whether the construction would be appropriate in the adult language. Oshima-Takane & Derat (1996) report the tendency of mothers to use proper names in what would normally be considered pronominal contexts in the early stages of the child's language acquisition. Similar findings have been reported for mothers of children with language impairment (Conti-Ramsden 1989).

tense -ed inflection and auxiliary DO. Specifically, copula and auxiliary BE are treated as a unique BE morpheme, and thus any differences that might exist between them are obliterated. Other studies that have kept the two tense markers separate have consistently reported significant differences in the provision of copula and auxiliary BE, with the copula both emerging earlier and being produced more accurately in obligatory contexts, for children with SLI as well as for MLUw controls. Although it is common practice to report performance in obligatory contexts as percentages, the difference between Fig. 1a (reporting percentages) and Fig. 1b (reporting absolute frequencies of fulfilled and unfulfilled contexts) shows that the conversion of frequency data into percentages can paint a considerably different picture of the developmental trajectory of forms.

The results of this study clearly show that these two morphemes indeed develop differently. Significant differences were found in the production in obligatory contexts of copula and auxiliary BE in both sets of children. We believe that a number of factors account for this difference. Firstly, we reported a frequency effect in the distribution of these two morphemes in the maternal input. Mothers, too, use more copular than auxiliary forms in the speech to their children. As a consequence children have more opportunities to hear a copular form than an auxiliary form and are therefore more likely to learn the former than the latter.

A second finding of this study is that both the children with SLI and the MLUw controls displayed an impressive command of copula BE with near ceiling performances, at least by the end of the period of observation. Further investigation of the use of different copula BE forms however revealed that both sets of children used an overwhelming majority of third person singular forms: their mastery of copula BE forms equated with mastery of this one morpheme. This result is in line with findings from a number of other studies reporting a predominance of third person singular verb forms (Hadley & Rice 1996). For auxiliary BE, too, third person singular forms tended to be the most frequent overall; however, for two of the children with SLI and four of the MLUw controls, a consistent proportion of first person singular auxiliary forms were reported. This observed variation is likely to reflect individual children's inclination to talk about themselves while performing certain actions rather than comment on the actions of third person referents.

In addition to the unmarked status of third person singular referents (Lyons 1977), we propose that the frequency distribution of morphological forms in the maternal input accounts for a large proportion of

the variance in the children's use. There were significant positive correlations between the children's and the mothers' use of contracted and uncontracted copular and auxiliary forms. Specifically, correlations were found between the children with SLI and the mother group, as well as between the MLUw controls and the mother group. The generally high values of the correlation coefficients indicate that frequency alone accounts for a sizeable proportion of the variance. Positive correlations between the production of copular forms by both sets of children and by the mother group were also obtained, even after third person singular forms were removed from the analysis, showing that an input effect is also found for the use of other person/number combinations. We are not arguing here for a simple input-output relationship to account for children's use of copula and auxiliary BE forms. We simply point out the fact that, when predicting patterns of use and omission in children's production, the higher frequency with which they hear some forms over others must be a factor. The greater the exposure to a given form, the greater the chances that that form will become entrenched and used by the child.

*The role of lexical specificity*

The final prediction concerning the role of lexical specificity was also upheld by the results on the use of copula and auxiliary BE in four different constructions. For the MLUw controls, a significant difference was reported in the provision of copula BE forms over time. This interaction was accounted for by the frequent use of copula BE forms in *wh*-constructions, especially at stage 1. Once the *wh*-construction was removed from the analysis, there was no longer a significant difference in the use of the copula in the remaining three constructions over time. However, overall differences remained when time was removed as a factor: finite forms of the copula were least likely to be found in the noun construction.

Similar results were obtained for the three children with SLI where statistical evidence indicated that the provision of the copula varied across the four constructions. There were also significant differences over time for two of the children with SLI (Harry and Nathan). As for the MLUw controls, provision of copula BE differed across constructions, suggesting that children's use of this morpheme was more entrenched in certain constructions than in others. There was also anecdotal evidence to support the view that copula BE is particularly entrenched in certain frames. For example, two of the MLUw controls used apparently unanalysed contracted copular forms in phrases with uncontracted copular forms, such as 'that's isn't faster' (Becky, 2;6.19) and 'where's is the baby?' (Anne, 2;5.2). However, in the sessions analysed, such

utterances were produced infrequently (approximately 0.6% of all finite copular forms for the MLUw controls).

In the case of the auxiliary BE, for the MLUw controls the difference in provision across the four constructions only held overall and not for the three different stages. A similar pattern was found for two of the three children with SLI (Harry and Nathan).

*Optionality is not random*

The finding that both for the children with SLI and the MLUw controls provision of copula and auxiliary BE was influenced by the construction type in which the morpheme appeared corroborates the notion that optionality is not a random, across-the-board phenomenon. We have in fact identified one possible factor that is likely to affect children's use of these two morphemes. Both in the case of copula and auxiliary BE we observed the same general pattern: the finite verb form was most likely to occur in *wh*-constructions and least likely to appear in noun constructions.

The first issue to resolve when addressing this kind of asymmetry is whether children's behaviour is in any way different from that of adult speakers. If children's use is indeed more restricted than adults', one can conclude that it is in some way deficient. If, on the contrary, the same trend is found in child-directed adult speech, then children's asymmetrical use of a given form may well mirror what they hear in the input. In the case of the frequency distribution of copula and auxiliary BE, the latter hypothesis seems to be correct. We suggest that the poorer performance of both groups of children in the use of copula and auxiliary BE in noun constructions is most likely to be explained by the frequency with which such constructions appear in the maternal input: less than one-tenth of mothers' copular constructions were noun constructions. In other words, the children were significantly more likely to hear copular constructions, such as 'that's nice' or 'he's nice', than 'the doggie's nice'.

It should also be noted that the noun construction is a much more open-ended schema than any of the other constructions. In principle there are no restrictions on the nouns that can feature in pre-copular position; any number of mass and count, singular and plural nouns are admitted. The open-endedness of the noun construction schema has two main implications: on the one hand it allows maximum productivity (Bybee 1995). On the other hand, the number of exemplars that the child will have to accumulate to reach a sufficiently abstract and schematic representation is larger than for other constructions. The likely consequence is therefore that it will take children longer to

become proficient in the use of the copula with nouns than with other subjects or with *wh*-words. Unlike in the noun construction, in *wh*- and pronoun constructions the elements that appear in the pre-copular slot belong to a closed class. This means that only a finite number of items are available to fill this slot. It will take a relatively short time before children have accumulated a large enough number of examples to form low-level schemas of the type 'it's + x' or 'what is + x?'. Additionally, the accumulation of a variety of these low-level schemas will contribute to the formation of a more abstract schematization of the construction. For example, in the case of the *wh*-construction, repeated exposure and use of instances of 'what's + x?' will lead to the entrenchment of this lexically specific schema, in which the pre-copular slot is occupied by 'what' and the post-copular slot by a variety of noun phrases, e.g., 'what's that thing?', 'what's this toy for?'. A number of other lexically specific *wh*-constructions might also be learnt at the same time, e.g., 'who's that?', 'who's this for?', 'where's that?', 'where's the dog?' The gradual accumulation of a range of these lexically specific *wh*-constructions will ultimately lead to the abstraction of a more general schema in which not only 'what', but a number of other *wh*-words will appear in sentence-initial position followed by a copular form.

For children with SLI this process of lexical learning and schematization is likely to take longer, partly due to a combination of factors such as poorer distributional learning abilities (Conti-Ramsden & Jones 1997), difficulties with establishing stable phonological representations (Gathercole & Baddeley 1990) and heightened sensitivity to item-level phonological properties interfering with generalization mechanisms (Marchman, Wulfeck & Ellis Weismer 1999). It is possible that excessive reliance on a small number of highly frequent, well-entrenched lexically specific schemas will be an obstacle to further abstraction into more general schemas.

*Learning to use copula BE in frames*

A more detailed investigation of the *wh*-words and pronouns used by the two groups of children in the various copula and auxiliary BE constructions revealed an interesting pattern. The MLUw controls' relatively good performance at stage 1 in copula *wh*-constructions was largely accounted for by use of the *wh*-frame 'what's'; two-thirds of all *wh*-constructions did in fact contain either 'what's that?' or 'what's that + n?' At stages 2 and 3 the proportion of *wh*-constructions with 'what' diminished steadily to just over a third by the end of the period of observation. This trend clearly shows that with time the control children became less reliant on specific lexical frames and started using

a larger variety of elements in pre-copular position. For the children with SLI the pattern was similar for Bonnie, but not for the other two children. As for the MLUw controls, Bonnie's consistent use of copula BE forms was mainly due to her repeated use of the same three constructions as the control children. However, Harry and Nathan both produced high-frequency frames containing an uncontracted copula. It is possible that the uncontracted copular forms are likely to be more salient than contracted forms, especially when they appear in sentence-final position where duration might also be longer (e.g., 'that's not yours!', 'it is!'). It has been suggested that children with SLI are more sensitive than typically developing children to the phonological salience of forms in the input language and that less salient items will be more difficult to process for affected children (Leonard *et al.* 1997).

The findings on the use of copula and auxiliary BE forms in different constructions confirmed our initial prediction concerning the role of lexical specificity in children's acquisition. Although both the children with SLI and the MLUw matches seemed to control the use of the two tense markers in a range of constructions, their performance differed considerably depending on the element featuring in pre-copular position. A pattern began to emerge in the optional provision of copula and auxiliary BE: provision was at least partially dependent on the choice of subject or sentence-initial *wh*-word. If this is the case, children's omission of the copula is not an entirely random phenomenon, and optionality can be constrained into a more systematic and predictable pattern. The finding that the copula was provided most accurately in *wh*-constructions and omitted most often in noun constructions accords with the principles of item-based acquisition (Rowland & Pine 2000). In the case of *wh*-constructions the successful performance of both sets of children was principally due to their repeated use of one single 'what's + x' frame. Learning one instance of a contractible copula and related *wh*-word ensured the provision of finiteness in a considerable proportion of obligatory contexts. In the case of the noun construction a similar one-type strategy was obviously not possible. Children hear a large variety of noun types being used in copular constructions and therefore no one-to-one relationship can be easily established whereby the acquisition of one noun with a contracted or uncontracted copular form can be used repeatedly to guarantee the successful provision of finiteness in this construction.

CONCLUSION

Identifying patterns of omission and possible predictors is important to

try to tease apart the underlying causes of the extended period of optionality in children with and without SLI. In a model like the EOI, optionality of finiteness marking is ascribed to children's immature knowledge of the obligatoriness of tense and, although children can and do use tense markers appropriately, initially they do not mark tense in 100% of obligatory contexts as the target grammar requires. Although the EOI model offers a maturational account of the optionality of finiteness marking, it does not address the issue of whether any constraints can be identified to predict when a tense marker is more or less likely to be overtly realized.

We propose that a constructivist approach, which takes into account extra-syntactic factors such as frequency effects (Marchman *et al.* 1999, Oetting & Horohov 1997) and item-based learning (Conti-Ramsden & Jones 1997, Pine *et al.* 1998), offers a more psychologically plausible account of typical and impaired language acquisition. As shown in the present paper, such an approach has the potential for identifying constraints on optionality and for predicting the likelihood that a target morpheme will be omitted or used appropriately.

We believe that studies which aim to identify constraints on optionality are particularly important because they expose systematic patterns in an otherwise seemingly random phenomenon. The maturational account on which the EOI rests does not make any specific predictions as to what shapes optionality over time, and we think that future research should be addressing the issue of constrained optionality in more precise terms (cf. Grela & Leonard 2000).

The present study has made a contribution to the understanding of the nature of optionality in the early grammars of children with and without SLI. Three main variables were identified as predictors of omission of the BE morpheme: auxiliary vs. copular status, type of construction, and nature of the lexical item in pre-copular position. That is, BE is more likely to be omitted if it is an auxiliary form, if it is in a noun construction, and if the pre-copular element with which it appears is low-frequency. Future research should address in greater detail, possibly with larger numbers of children in controlled experimental conditions, which other variables constrain the phenomenon of optional finiteness.

REFERENCES

Becker, M. (2000). The acquisition of the English copula. *Proceedings of the 24th Annual Boston University Conference on Language Development* (Sommerville, MA: Cascadilla Press), 104–15.

Berument, S. K., Rutter, M., Lord, C., Pickles, A. & Bailey, A. (1999). Autistic screening questionnaire: diagnostic validity. *British Journal of Psychiatry*, 175, 444–51.

Brown, R. (1973). *A First Language: The Early Stages* (Cambridge, MA: Harvard University Press).

Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, 10, 425–55.

Cleave, P. L. & Rice, M. (1997). An examination of the morpheme BE in children with specific language impairment: the role of contractibility and grammatical form class. *Journal of Speech Language and Hearing Research*, 40, 480–92.

Conti-Ramsden, G. (1989). Proper name usage: mother-child interactions with language-impaired and non-language-impaired children. *First Language*, 9, 271–84.

Conti-Ramsden, G. & Jones, M. (1997). Verb use in specific language impairment. *Journal of Speech and Hearing Research*, 40, 1298–313.

Edwards, S., Fletcher, P., Garman, M., Hughes, A., Letts, C. & Sinka, I. (1997). *The Reynell Developmental Language Scales III: The University of Reading Edition* (Windsor: NFER-Nelson).

Gathercole, S. & Baddeley, A. (1990). Phonological memory deficits in language-disordered children: is there a causal connection?. *Journal of Memory and Language*, 29, 336–60.

Grela, B. G. & Leonard, L. B. (2000). The influence of argument structure complexity on the use of auxiliary verbs by children with SLI. *Journal of Speech and Hearing Research*, 43, 1115–25.

Hadley, P. & Rice, M. (1996). Emergent uses of BE and DO: evidence from children with specific language impairment. *Language Acquisition*, 5, 209–43.

Hoff-Ginsberg, E. (1992). How should frequency in input be used?. *First Language*, 12, 233–44.

Kuczaj, S. (1985/6). General developmental patterns and individual differences in the acquisition of copula and auxiliary *be* forms. *First Language*, 6, 111–17.

Lebeaux, D. (2000). *Language Acquisition and the Form of the Grammar* (Amsterdam: John Benjamins).

*Leiter International Performance Scale* (Chicago: Stoelting, 1979).

Leonard, L. B. (1995). Functional categories in the grammars of children with specific language impairment. *Journal of Speech and Hearing Research*, 38, 1270–83.

Leonard, L., Eyer, J., Bedore, L. & Grela, B. (1997). Three accounts of the grammatical morpheme: difficulties of English-speaking children with Specific Language Impairment. *Journal of Speech, Language and Hearing Research*, 40, 741–53.

Lyons, J. (1977). *Semantics* (Cambridge: Cambridge University Press).

MacWhinney, B. (1995). *The Child Language Data Exchange System* (Hillsdale, NJ: Lawrence Erlbaum Associates).

Marchman, V., Wulfeck, B. & Ellis Weismer, S. (1999). Morphological productivity in children with normal language and SLI: a study of the English past tense. *Journal of Speech, Language and Hearing Research*, 42, 206–19.

Oetting, J. & Horohov, J. (1997). Past tense marking by children with and without Specific Language Impairment. *Journal of Speech Language and Hearing Research*, 40, 62–74.

Oshima-Takane, Y. & Derat, L. (1996). Nominal and pronominal reference in maternal speech during the later stages of language acquisition: a longitudinal study. *First Language*, 16, 319–38.

Pine, J. M., Lieven, E. V. M. & Rowland, C. F. (1998). Comparing different models of the development of the English verb category. *Linguistics*, 36, 807–30.

Pizzuto, E. & Caselli, M. C. (1992). The acquisition of Italian morphology: implications for models of language development. *Journal of Child Language*, 19, 491–557.

Rice, M. L. & Wexler, K. (1996). A phenotype of Specific Language Impairment: Extended Optional Infinitives. In M. L. Rice (ed.), *Towards a Genetics of Language* (Hillsdale, NJ: Lawrence Erlbaum Associates), 215–37.

Rice, M. L., Wexler, K. & Cleave, P. L. (1995). Specific Language Impairment as a period of extended optional infinitive. *Journal of Speech and Hearing Research*, **38**, 850–63.

Rice, M. L., Wexler, K. & Hershberger, S. (1998). Tense over time: the longitudinal course of tense acquisition in children with Specific Language Impairment. *Journal of Speech, Language and Hearing Research*, **41**, 1412–31.

Rowland, C. & Pine, J. (2000). Subject-auxiliary inversion errors and wh-question acquisition: 'what children do know?'. *Journal of Child Language*, **27**, 157–81.

Theakston, A., Lieven, E., Pine, J. & Rowland, C. (2001). The role of performance limitations in the acquisition of verb-argument structure. *Journal of Child Language*, **28**, 127–52.