J. Child Lang. **32** (2005), 269–289. © 2005 Cambridge University Press
doi:10.1017/S0305000905006860 Printed in the United Kingdom

Testing the Agreement/Tense Omission Model: why the data on children's use of non-nominative 3psg subjects count against the ATOM*

JULIAN M. PINE, CAROLINE F. ROWLAND
University of Liverpool

ELENA V. M. LIEVEN
Max Planck Institute for Evolutionary Anthropology

AND

ANNA L. THEAKSTON
University of Manchester

(*Received* 16 *October* 2000. *Revised* 20 *September* 2004)

ABSTRACT

One of the most influential recent accounts of pronoun case-marking errors in young children's speech is Schütze & Wexler's (1996) Agreement/Tense Omission Model (ATOM). The ATOM predicts that the rate of agreeing verbs with non-nominative subjects will be so low that such errors can be reasonably disregarded as noise in the data. The present study tests this prediction on data from 12 children between the ages of 1;8.22 and 3;0.10. This is done, first, by identifying children who produced a reasonably large number of non-nominative 3psg subjects; second, by estimating the expected rate of agreeing verbs with masculine and feminine non-nominative subjects in these children's speech; and, third, by examining the actual rate at which agreeing verb forms occurred with non-nominative subjects in those areas of the data in which the expected error rate was significantly greater than 10%. The results show, first, that only three of the children produced enough non-nominative subjects to allow a reasonable test of the ATOM to be made; second, that for all three of these children, the only area of the data in which the expected frequency of

[*] We would like to thank all the families that took part in the research reported here. We would also like to thank Melissa Bowerman and two anonymous reviewers for their constructive comments on an earlier version of this paper. The research was funded by the Economic and Social Research Council, Grant Numbers: R000236393 and R000237911. Address for correspondence: Julian M. Pine, School of Psychology, University of Liverpool, L69 7ZA, UK. e-mail: Julian.Pine@Liverpool.ac.uk

agreeing verbs with non-nominative subjects was significantly greater than 10% was their use of feminine case-marked subjects; and third, that for all three of these children, the rate of agreeing verbs with non-nominative feminine subjects was over 30%. These results raise serious doubts about the claim that children's use of non-nominative subjects can be explained in terms of AGR optionality, and suggest the need for a model of pronoun case-marking error that can explain why some children produce agreeing verb forms with non-nominative subjects as often as they do.

INTRODUCTION

It is a well-documented fact that English-speaking children sometimes use non-nominative pronouns in subject position, producing utterances such as 'me go', 'my going', 'him gone' and 'her went' (Gruber, 1967; Bloom, 1970; Huxley, 1970; Webster & Ingram, 1972; Brown, 1973; Suppes, 1974; Chiat, 1981; Budwig, 1989; Loeb & Leonard, 1991; Rispoli, 1994, 1998*a*, *b*, 1999; Vainikka, 1994; Schütze & Wexler, 1996; Schütze, 1997; Radford, 1998; Wexler, 1998; Wexler, Schütze & Rice, 1998). One of the most influential recent accounts of this phenomenon is provided by Schütze & Wexler's (1996) Agreement/Tense Omission Model (henceforth ATOM).

According to the ATOM, by the time young children begin to produce multi-word speech, they have already correctly set all the basic inflectional and clause structure parameters of their language. However, there is a stage of development (the optional infinitive stage) during which the abstract features of Tense and Agreement may be underspecified in the underlying representation of the sentence. This results in children sometimes using nonfinite verb forms in contexts in which finite verb forms would be obligatory in adult speech, and producing errors such as those in 1, 2 and 3:

(1) That go in there
(2) Daddy going to the shops
(3) Mummy gone to work

Moreover, since Agreement is assumed to assign Nominative case, it also results in children using non-nominative pronouns such as 'me', 'my', 'him' and 'her' in contexts in which nominative pronouns are required, and producing case-marking errors such as those in 4, 5, 6 and 7:

(4) My go in there
(5) Me went home
(6) Him going to the shops
(7) Her gone to work

These errors are assumed to reflect the absence of AGR from the underlying representation of the sentence and, according to the ATOM, can

TABLE 1. *Summary of the patterns of AGR and TNS omission and NOM and Non-NOM subject use predicted by the ATOM*

AGR/TNS	NOM	Examples
Predicted to occur		
+AGR/+TNS	+NOM	I'm going, He goes, She's gone, She went
+AGR/−TNS	+NOM	I going, He go, She gone
−AGR/+TNS	−NOM	Me going, Him go, Her gone, Her went
Predicted not to occur		
+AGR/+TNS	−NOM	*Me am going, *Him goes, *Her's gone

be distinguished from errors such as those in 8, 9 and 10, which are also predicted by the ATOM, but are assumed to reflect the absence of TNS.

(8) I going to the shops
(9) He gone to work
(10) She gone to work

In this latter case, the assumption is that the child uses an uninflected verb form with a nominative subject because, although Agreement is present in the underlying representation of the sentence, the verb itself is not marked for agreement because of the absence of Tense.

The main strength of the ATOM is that it provides a unified account of patterns of AGR/TNS omission and patterns of pronoun case-marking error in young children's speech. Thus, according to the ATOM, children produce nominative subjects with verbs that are inflected for tense and agreement when TNS and AGR are present in the underlying representation of the sentence; they produce nominative subjects with verbs that are not inflected for tense or agreement when AGR is present but hidden because TNS is absent; and they produce non-nominative subjects with verbs that are not inflected for agreement, but may be inflected for tense, when AGR is absent (see Table 1 for a summary of the pattern of case, tense and agreement marking that the ATOM predicts). However, although the ATOM is able to explain a number of features of children's early multi-word speech, its empirical status as a model of pronoun case-marking error is problematic for a number of reasons.

The first problem is that, because of the rather complex relation between the underlying knowledge assumed by the ATOM and the way that this knowledge is assumed to manifest itself in the child's speech, there has been a certain amount of confusion in the literature about how the ATOM should be operationalized and tested. Thus, the ATOM has often been treated as if all it predicted were significant associations between correct subject case and agreement marking in children's speech, or significant differences in the rate at which children use agreeing and non-agreeing verbs with non-nominative subjects. For example, Wexler *et al.* (1998) report significant group differences in the proportion of agreeing and non-agreeing verbs with non-nominative subjects, and significant associations between correct subject case and agreement marking in data pooled across children, and interpret these results as providing strong support for the ATOM. In fact, however, although the effects reported by Wexler *et al.* are consistent with the ATOM, they do not map directly onto the ATOM's central prediction, which is that children will not use agreeing verb forms with non-nominative subjects – or that the rate at which agreeing verb forms occur with non-nominative subjects will always be so low that such errors can be reasonably disregarded as noise (Schütze, 2001).

As Schütze himself acknowledges, this prediction is not equivalent to the prediction of a significant difference in the rate at which children use agreeing or non-agreeing verbs with non-nominative subjects or a significant association between correct subject case and agreement marking. This is because there is no reason why one should not find children with low rates of agreeing verbs with non-nominative subjects who fail to show such effects (because they also have low rates of non-agreeing verbs with non-nominative subjects), or children with high rates of agreeing verbs with non-nominative subjects who do show such effects (because they have even higher rates of non-agreeing verbs with non-nominative subjects). It follows that the critical test of the ATOM is not whether children show a significant association between correct subject case and agreement marking, but whether children produce agreeing verbs with non-nominative subjects at rates that are too high to be reasonably disregarded as noise.

The most straightforward way of testing this prediction is by setting an upper limit on the rate of agreeing verb forms with non-nominative subjects that one would be prepared to disregard as noise and determining whether the rate at which such errors actually occur is significantly greater than this figure. To our knowledge, the ATOM has never been explicitly tested in this way (though Schütze & Wexler, 1996, do use precisely the same logic to reject a simple Tense Omission model of case-marking error).[1] The present study therefore represents an attempt to perform such a test by explicitly focusing on the claim that the rate of agreeing verb forms with non-nominative subjects in children's speech is so low that it can be reasonably disregarded as noise in the data. This will be done by setting the upper limit on the rate of agreeing verbs with non-nominative subjects that one would

[1] Schütze & Wexler (1996) point out that, although there is a significant association between correct subject case and tense marking in their data, a simple Tense Omission model of case-marking error cannot be correct since the rate at which past tense forms occur with non-nominative subjects 'ranges from 4% to as high as 30% in our data' (Schütze & Wexler, 1996: 676).

be prepared to disregard as noise at 10% and determining whether the rate at which agreeing verb forms actually occur with non-nominative subjects is significantly greater than this figure. If the rate of agreeing verbs with non-nominative subjects is not found to be significantly greater than 10%, then the data will be viewed as consistent with the ATOM. If the rate of agreeing verbs with non-nominative subjects is found to be significantly greater than 10%, then the data will be taken as evidence against the ATOM.

This brings us to a second problem with the ATOM, which is that, once one has established what the ATOM really predicts, it becomes clear that the range of situations in which this prediction can be properly tested is actually rather limited. That is to say, since the central prediction of the ATOM is a prediction about the absence of agreeing verb forms with non-nominative subjects, and since the ATOM makes no predictions about the rate at which non-nominative subjects will actually occur in any particular dataset, the ATOM clearly capitalizes on the fact that in many datasets the rate at which agreeing verb forms with non-nominative subjects would be expected to occur by chance may actually be very low (i.e. on the fact that in many datasets the low rate at which agreeing verb forms occur with non-nominative subjects may be a straightforward consequence of the low rate at which any verb forms occur with non-nominative subjects). The implication is that the ATOM can only be properly tested on data in which the rate at which agreeing verbs with non-nominative subjects would be expected to occur by chance is reasonably high (i.e. high enough to disconfirm the theory if it were wrong). Thus, if one assumes that the ATOM can only be falsified by the discovery of children who produce agreeing verbs with non-nominative subjects at rates significantly greater than 10%, then it follows that the ATOM can only be properly tested on datasets in which the rate at which such errors would be expected to occur by chance is also significantly greater than 10% (i.e. greater than 10% and based on a high enough number of agreeing verb forms to be considered statistically reliable).

In fact, however, there are very few datasets in the literature that meet this criterion. Thus, Schütze & Wexler (1996) and Schütze (1997) report data from three children (Nina and Peter for 1psg and Nina and Sarah for 3psg). These data are reproduced in Tables 2 and 3, together with measures of the frequency with which one would expect the various combinations to occur by chance given the distribution of nominative and non-nominative subjects and agreeing versus non-agreeing verb forms in the data. The observed frequencies are taken from Schütze (1997). The expected frequencies were calculated using the formula: Expected frequency = Row Total × Column Total/Grand Total as in a standard Chi-square analysis. Also presented are measures of the rate at which agreeing verb forms occur with non-nominative subjects and estimates of the rate at which such errors

TABLE 2. *Distribution of NOM/Non-NOM 1psg subjects with/without Agreeing verb forms for Peter and Nina (expected values in parantheses)*

	I	Me/My
Peter		
Agreeing	136 (122·68)	0 (13·32)
Non-Agreeing	564 (577·32)	76 (62·68)
Expected rate of Agreeing verbs with Non-Nominative subjects = 9·8%		
Observed rate of Agreeing verbs with Non-Nominative subjects = 0%		
Nina		
Agreeing	15 (13·02)	0 (1·98)
Non-agreeing	888 (889·98)	137 (135·02)
Expected rate of Agreeing verbs with Non-Nominative subjects = 13·2%		
Observed rate of Agreeing verbs with Non-Nominative subjects = 0%		

TABLE 3. *Distribution of NOM/Non-NOM 3psg subjects with/without Agreeing verb forms for Sarah and Nina (expected values in parentheses)*

Sarah	She	Her
Agreeing	11 (8·34)	0 (2·66)
Non-Agreeing	33 (35·66)	14 (11·34)
Expected rate of Agreeing verbs with Non-Nominative subjects = 24·2%		
Observed rate of Agreeing verbs with Non-Nominative subjects = 0%		
Nina	He/She	Him/Her
Agreeing	213 (160·79)	10 (62·21)
Non-agreeing	185 (237·21)	144 (91·79)
Expected rate of Agreeing verbs with Non-Nominative subjects = 27·9%*		
Observed rate of Agreeing verbs with Non-Nominative subjects = 4·5%		

* Significantly greater than 10% by Binomial test at $p < 0{\cdot}05$.

would be expected to occur by chance. The observed error rates were calculated using the formula: Observed rate of Agreeing verb forms with Non-Nominative subjects = Observed frequency of Agreeing verb forms with Non-Nominative subjects/Observed frequency of Agreeing verb forms with Nominative and Non-Nominative subjects. The expected error rates were calculated using the formula: Expected rate of Agreeing verb forms with Non-Nominative subjects = Expected frequency of Agreeing verb forms with Non-Nominative subjects/Observed frequency of Agreeing verb forms with Nominative and Non-Nominative subjects.

It can be seen from Tables 2 and 3 that the only dataset in which the expected rate of agreeing verbs with non-nominative subjects is significantly greater than 10% is Nina's 3psg dataset (where the expected error rate is 27·9% compared with an observed error rate of 4·5%). Nina's 3psg dataset is thus the only dataset reported in the literature to date in which the expected

TABLE 4. *Distribution of NOM/Non-NOM 3psg subjects with/without Agreeing verb forms for Nina broken down by gender (expected values in parentheses)*

	He	Him	She	Her
Agreeing	207 (203·11)	3 (6·89)	6 (1·25)	7 (11·75)
Non-Agreeing	176 (179·89)	10 (6·11)	9 (13·75)	134 (129·25)
	Expected rate of Agreeing verbs with Non-NOM masculine subjects = 3·3%		Expected rate of Agreeing verbs with Non-NOM feminine subjects = 90·4%*	
	Observed rate of Agreeing verbs with Non-NOM masculine subjects = 1·4%		Observed rate of Agreeing verbs with Non-NOM feminine subjects = 53·8%*	

* Significantly greater than 10% by Binomial Test at $p < 0.05$.

error rate is high enough to allow the ATOM to be properly tested.[2] The implication is that, although Nina's data do provide some limited support for the ATOM, the ATOM derives much of its credibility as a model of pronoun case-marking error from the fact that, in most of the datasets currently available, the rate at which agreeing verbs would be expected to occur with non-nominative subjects by chance is not high enough to allow a reasonable test of the ATOM to be made.

This brings us to a third problem with the ATOM, which is that even the idea that Nina's 3psg data provide any real support for the ATOM relies very heavily on the way in which the observed and expected error rates are computed. Thus, Table 4 presents Nina's 3psg data broken down by gender. It can be seen from these data that the rates at which Nina produces non-nominative masculine subjects and the rate at which she produces non-nominative feminine subjects are very different, and hence that the overall expected error rate reported above hides very different expected error rates for the masculine and feminine parts of the system.[3] Thus, although the rate at which agreeing verbs would be expected to occur with feminine non-nominative subjects is very high (90·4%), the rate at which agreeing verbs

[2] In fact there are also 4 children described in Rispoli (1999) who meet this criterion. These children produced agreeing verbs with non-nominative subjects at rates of 57·1, 41·2, 30·0 and 27·8%, respectively. However, Schütze (2001) argues that these datasets cannot be used to test the ATOM since the ATOM only makes predictions about the use of non-nominative subjects in children who also have the relevant nominative pronouns in their productive lexical inventories, and it is unclear whether this was the case for these children.

[3] The observed frequencies presented in Table 4 are also taken from Schütze (1997), who presents detailed breakdowns of Nina's data by gender and verb type. However, for some reason, Schütze (1997) ignores the high rate of agreeing verbs with non-nominative feminine subjects in Nina's data, and only considers the overall rate of agreeing verbs with non-nominative subjects when evaluating the fit between the data and the model.

would be expected to occur with masculine subjects is much lower (i.e. only 3·3%). The implication is that the only area of the data in which the expected error rate is high enough to allow a reasonable test of the ATOM to be made is Nina's use of feminine case-marked subjects. However, the observed rate of agreeing verbs with non-nominative feminine subject is 53·8% – a figure that, although significantly lower than one would expect by chance ($p<0.001$ by Fisher's Exact), is also significantly higher than 10% ($p<0.001$, by a Binominal test). Thus, although it is clear that there is an association between correct nominative case and agreement marking in Nina's data, it is also clear that Nina produces agreeing verbs with non-nominative feminine subjects at a rate that is too high to be reasonably disregarded as noise in the data.

One obvious way of interpreting this pattern of results is to argue that although Nina's grammar does allow agreeing verb forms with non-nominative subjects, Nina produces significantly fewer of these combinations than one would expect by chance because her use of agreeing verb forms with nominative subjects is boosted by her use of unanalysed subject + copula and subject + auxiliary contractions such as 'he's' and 'she's' (Rispoli, 1999). Indeed, in a recent study Wilson (2003) has shown that there are significant asymmetries in Nina's provision of copula BE and auxiliary BE with pronominal and lexical subjects that suggest that this may be exactly what is going on in her data. The critical point, however, is that Nina's use of agreeing verbs with non-nominative feminine subjects is not consistent with the ATOM's central prediction. The implication is that Nina's data provide much less support for the ATOM than has tended to be assumed in the past. An additional implication is that, if one is able to isolate areas of particular children's data in which the expected rate of agreeing verbs with non-nominative subjects is reasonably high, it is possible to find rates of agreeing verbs with non-nominative subjects that are much higher than the ATOM would predict.

To summarize, a detailed consideration of the ATOM and the data on which it is based raises a number of questions about the empirical status of the ATOM as a model of pronoun case marking error. In view of these questions, the aim of the present study is to take a fresh look at the ATOM by testing its central prediction on 12 longitudinal datasets (the Manchester corpus) that have recently been made available by our research group through the CHILDES system (Theakston, Lieven, Pine & Rowland, 2001). The study represents an attempt to assess the claim that the rate at which children produce agreeing verbs with non-nominative subjects is so low that such errors can be reasonably disregarded as noise in the data (Schütze, 2001). This will be done, first, by identifying children who produce a reasonably large number of non-nominative 3psg subjects; second, by estimating the expected rate of agreeing verbs with non-nominative masculine and feminine subjects in these children's data; and, third, by examining the actual rates at which agreeing verb forms occur with non-nominative subjects in those areas of the data in which the expected error rate is significantly greater than 10%. If the error rates in these areas of the data are not found to be significantly greater than 10%, then they will be viewed as consistent with the ATOM. If the error rates in these areas of the data are found to be significantly greater than 10%, then they will be taken as evidence against the ATOM.

METHOD

Participants

Participants in the study were 12 first-born children (6 boys and 6 girls) from predominantly middle-class backgrounds. They were recruited through advertisements in local newspapers, nurseries and doctors' surgeries. All of the children were from monolingual English-speaking families and all were cared for primarily by their mothers. At the beginning of the study the children's ages ranged from 1;8.22 to 2;0.25 and their MLUs from 1·06 to 2·22 morphemes. At the end of the study the children's ages ranged from 2; 8.15 to 3;0.10 and their MLUs from 2·85 to 4·12 morphemes.

Procedure

Each child was audio-recorded at home for an hour twice every three weeks over a twelve-month period. Each recording session consisted of two thirty-minute recordings of the child at play with his or her mother. The first of these was a recording of the child playing with his or her own toys. The second was a recording of the child playing with toys provided by the investigator. In order to reduce the artificiality of the situation, the investigator behaved as a participant observer throughout. However, an attempt was made to 'leave the floor' to the child and his or her mother as much as possible.

Children's speech corpora

All audio-recordings were orthographically transcribed into a computerized database using the CHAT system from the CHILDES project (MacWhinney, 2000). The resulting transcripts have since been made available through CHILDES as the Manchester Corpus (Theakston *et al.*, 2001). These transcripts were used to build up a corpus of utterances for each of the 12 children. The criteria for inclusion of utterances in the children's corpora were that the utterances were (a) fully intelligible; (b) complete (i.e. not interrupted utterances or false starts); (c) used spontaneously by the child (i.e. neither imitations nor immediate self-repetitions); and (d) neither strings of numbers nor fragments of songs or nursery rhymes.

Coding scheme

Each of the children's corpora was searched for utterances in which the 3psg pronouns 'he', 'him', 'she' or 'her' occurred in subject position. This included multiple instances of the same utterance type provided these were not immediate self-repetitions (see above). It also included utterances such as 'He a good boy' and 'Her ready' where the relevant pronoun could be interpreted as the subject of a missing copula. All pronominal subjects were then classified as either nominative ('he' or 'she') or non-nominative ('him', or 'her') and all verb forms (including missing copulas) were classified as either agreeing or non-agreeing. The category of agreeing verb forms included all instances of unambiguously agreeing verb forms (i.e. lexical verbs marked for 3psg, and contracted and non-contracted forms of unambiguously agreeing copulas and auxiliaries). The category of non-agreeing verb forms included all instances of verbs forms that were not unambiguously agreeing (i.e. modals, past tense forms, ambiguous forms of the auxiliaries 'have' and 'do', uninflected lexical verbs, progressive and perfect participles without auxiliary 'be' or auxiliary 'have', missing copulas and infinitival forms of auxiliary 'be' and copula 'be'). The resulting datasets were used to identify children who produced more than a handful of non-nominative 3psg subjects and to draw up a contingency table for each of these children showing the observed frequencies of nominative and non-nominative 3psg subjects with agreeing and non-agreeing verb forms. These contingency tables were then used to derive the following measures:

1. *The expected frequencies of nominative and non-nominative 3psg subjects with agreeing and non-agreeing verb forms*. These measures were derived by applying the formula: Expected frequency = Row Total × Column Total/Grand Total. This procedure is equivalent to that used in a standard Chi-square analysis and provides an estimate of the frequency with which one would expect to observe each possible combination of subject case and agreement marking if there were no relation between subject case and agreement marking in the data.

2. *The observed rate of agreeing verb forms with 3psg non-nominative subjects*. This measure was derived using the formula: Observed rate of Agreeing verb forms with Non-Nominative subjects = Observed frequency of Agreeing verb forms with Non-Nominative subjects/Observed frequency of Agreeing verb forms with Nominative and Non-Nominative subjects.

3. *The expected rate of agreeing verb forms with 3psg non-nominative subjects*. This measure was derived using the formula: Expected rate of Agreeing verb forms with Non-Nominative subjects = Expected frequency

TABLE 5. *Descriptive Statistics on the 12 children's Age and MLU ranges, the size of their multi-word speech corpora, and the total number of 3psg nominative contexts and 3psg non-nominative subjects in these corpora*

	Age range	MLU range	Total multi-word utterances	Total 3psg NOM contexts	Total 3psg Non-NOM subjects
Anne	1;10–2;9	1·61–3·46	9510	309	10
Aran	1;11–2;11	1·41–3·84	8966	371	2
Becky	2;0–3;0	1·46–3·24	10 144	511	18
Carl	1;9–2;9	2·17–3·93	13 867	936	0
Dominic	1;11–2;11	1·20–2·85	8686	81	0
Gail	1;12–2;11	1·76–3·42	7971	223	19
Joel	1;11–2;10	1·33–3·32	8174	400	1
John	1;11–2;11	2·22–2·93	5681	30	0
Liz	1;11–2;11	1·35–4·12	7420	139	0
Nicole	2;1–3;0	1·06–3·26	5588	12	2
Ruth	2;0–3;0	1·41–3·35	7698	21	1
Warren	1;10–2;10	2·01–4·12	9515	226	0

of Agreeing verb forms with Non-Nominative subjects/Observed frequency of Agreeing verb forms with Nominative and Non-Nominative subjects.

To ensure that children's use of agreeing verbs with non-nominative subjects did not simply reflect ignorance of the appropriate nominative forms, utterances were only counted from the point at which the relevant nominative forms appeared in the data. To ensure that the expected frequencies of agreeing verbs with non-nominative subjects were not distorted by developmental changes in the children's use of non-nominative subjects and/or developmental changes in their use of agreeing verb forms, all of the measures were based only on data from the period during which the child was producing both the relevant non-nominative subjects and the relevant agreeing verb forms with nominative subjects. Thus, all of the above measures were based on the period during which the child was producing both utterances including non-nominative subjects (e.g. 'Him big' or 'Her going') and utterances including nominative 3psg subjects with unambiguously agreeing verb forms (e.g. 'He wants a drink', 'She's a good girl' or 'He's coming'). In practice, imposing these restrictions had very little effect on the overall pattern of results.

RESULTS

Table 5 presents data on the size of the children's multi-word speech corpora together with the children's age and MLU ranges and the total number of 3psg nominative and non-nominative subjects in their corpora. It can be seen from these data that although all of the 12 children produced at

TABLE 6. *Distribution of NOM/Non-NOM 3psg subjects with/without Agreeing verb forms for Anne, Becky and Gail* (*expected values in parentheses*)

	NOM	Non-NOM
Anne		
Agreeing	141 (139·48)	5 (6·52)
Non-agreeing	73 (74·52)	5 (3·48)
Expected rate of Agreeing verbs with Non-Nominative subjects = 4·5%		
Observed rate of Agreeing verbs with Non-Nominative subjects = 3·4%		
Becky		
Agreeing	239 (241·38)	16 (13·62)
Non-agreeing	80 (77·62)	2 (4·38)
Expected rate of Agreeing verbs with Non-Nominative subjects = 5·3%		
Observed rate of Agreeing verbs with Non-Nominative subjects = 6·3%		
Gail		
Agreeing	146 (138·32)	13 (20·68)
Non-agreeing	48 (55·68)	16 (8·32)
Expected rate of Agreeing verbs with Non-Nominative subjects = 13·0%		
Observed rate of Agreeing verbs with Non-Nominative subjects = 8·2%		

least some utterances including 3psg nominative contexts, only 3 of the 12 children (Anne, Becky and Gail) produced more than 2 non-nominative 3psg subjects during the course of the study. Since the ATOM can only be tested on children who produce a reasonably large number of non-nominative subjects, all subsequent analyses focus specifically on these three children.

Data on the distribution of nominative and non-nominative 3psg subjects with and without unambiguously agreeing verbs in Anne, Becky and Gail's data are presented in Table 6. This table also includes estimates of the expected frequency of agreeing verbs with nominative and non-nominative subjects and estimates of the rate at which agreeing verbs would be expected to occur with non-nominative as opposed to nominative subjects if there were no relation between case and agreement marking in the data. It can be seen from these data that, consistent with the predictions of the ATOM, for all three of the children the rate at which agreeing verbs occur with non-nominative subjects was relatively low (i.e. less than 10%). However, it can also be seen that for none of the children was the expected rate of agreeing verbs with non-nominative subjects significantly greater than 10%, and hence high enough to allow a reasonable test of the ATOM to be made.

Table 7 presents the same data broken down by gender. It is clear from Table 7 that all three of the children produced non-nominative feminine subjects at much higher rates than non-nominative masculine subjects and hence that for all three of the children the expected error rates reported in

TABLE 7. *Distribution of NOM/Non-NOM 3psg subjects with/without Agreeing verb forms for Anne, Becky and Gail broken down by gender (expected values in parentheses)*

	He	Him	She	Her
Anne				
Agreeing	133 (131·97)	1 (2·03)	8 (8·77)	4 (3·23)
Non-Agreeing	62 (63·03)	2 (0·97)	11 (10·23)	3 (3·77)
	Expected rate of Agreeing verbs with Non-NOM masculine subjects = 1·5%		Expected rate of Agreeing verbs with Non-NOM feminine subjects = 26·9%*	
	Observed rate of Agreeing verbs with Non-NOM masculine subjects = 0·7%		Observed rate of Agreeing verbs with Non-NOM feminine subjects = 33·3%*	
Becky				
Agreeing	213 (212·09)	3 (3·91)	26 (30·69)	13 (8·31)
Non-Agreeing	58 (58·91)	2 (1·09)	22 (17·31)	0 (4·69)
	Expected rate of Agreeing verbs with Non-NOM masculine subjects = 1·8%		Expected rate of Agreeing verbs with Non-NOM feminine subjects = 21·3%*	
	Observed rate of Agreeing verbs with Non-NOM masculine subjects = 1·4%		Observed rate of Agreeing verbs with Non-NOM feminine subjects = 33·3%*	
Gail				
Agreeing	132 (128·73)	4 (7·27)	14 (10·86)	9 (12·14)
Non-Agreeing	45 (48·27)	6 (2·73)	3 (6·14)	10 (6·86)
	Expected rate of Agreeing verbs with Non-NOM masculine subjects = 5·3%		Expected rate of Agreeing verbs with Non-NOM feminine subjects = 52·8%*	
	Observed rate of Agreeing verbs with Non-NOM masculine subjects = 2·9%		Observed rate of Agreeing verbs with Non-NOM feminine subjects = 39·1%*	

* Significantly greater than 10% by Binomial test at $p<0\cdot05$.

Table 6 hide very different expected error rates in the masculine and feminine parts of the system. Thus, none of the three children had an expected rate of agreeing verbs with non-nominative masculine subjects that was significantly greater than 10%, but all of the three children had expected rates of agreeing verbs with non-nominative feminine subjects that were significantly greater than 10%. It is also clear that all three of the children produced agreeing verbs with non-nominative feminine subjects at rates significantly greater than 10%. Thus Anne produced agreeing verbs with non-nominative feminine subjects at a rate of 4/12 = 33·3%, (which is significantly greater than 10% at $p<0\cdot05$ by a Binomial test); Becky produced agreeing verbs with non-nominative feminine subjects at a rate of 13/39 = 33·3% (which is significantly greater than 10% at $p<0\cdot0001$ by a

TABLE 8. *All instances of unambiguously Agreeing verb forms with Non-NOM 3psg subjects in Anne, Becky and Gail's data*

Anne	Him doesn't And her has Probably her's a baby A big girl now, her is I think her was crying for me
Becky	Where does him go? Her is gonna make a dinner Her is gonna make it Her's got bin, haven't they, Mum? Her's being lovely Her's sixteen, Mum Her's fifteen and sixteen and nineteen Her's finished lunch now Her hasn't got some nighties, has she? Her isn't Her is not here. Her's can take this one to home Him's eating you, crocodile Him's ready to have in a bath How old is her? Her's cross
Gail	Her's go on this Her's go in desk Her's want to do some cooking Him's go in Him's going to bed now Him's want to be a monster Her's going to party. Her's going to party, Mummy Her's going to a party Her's getting cold Him's going to sleep Her's got a tie thing Her's not got any clothes on

Binomial test); and Gail produced agreeing verbs with non-nominative feminine subjects at a rate of $9/23 = 39.1\%$ (which is significantly greater than 10% at $p < 0.001$ by a Binomial test). These results show that the low overall rate of agreeing verbs with non-nominative subjects cannot be taken at face value, and that when one does isolate areas of the children's data in which the expected rate of agreeing verbs with non-nominative subjects is reasonably high, one finds rates of agreeing verbs with non-nominative subjects that are much higher than the ATOM would predict.

Table 8 presents all the instances of agreeing verbs with non-nominative 3psg subjects observed in the three children's speech. These data include a relatively wide range of different non-nominative subject + agreeing verb

TABLE 9. *Distribution of NOM/Non-NOM 3psg feminine subjects with non-contracted Agreeing verb forms for Anne, Becky and Gail*

	She	Her	Rate (%)
Anne			
Non-contracted	1	3	75·0*
Becky			
Non-contracted	13	6	31·6*
Gail			
Non-contracted	3	0	0

* Significantly greater than 10% by Binomial test at $p < 0{\cdot}05$.

combinations in a number of different sentential contexts and are hence consistent with the view that, during the period of development covered by this study, the use of agreeing verbs with non-nominative subjects was a real grammatical possibility for these children. However, the data presented in Table 8 also raise two possible objections to this interpretation of the results.

The first is that many of the utterances presented in Table 8 involve contracted her+copula or her+auxiliary combinations that could, in principle, be viewed as instances of 'hers'+non-agreeing verb combinations (i.e. cases where the child concerned is incorrectly substituting a genitive pronoun for a nominative pronoun). This raises the possibility that the relatively high rate of agreeing verbs with non-nominative subjects in the feminine part of the system may be a consequence of the misclassification of 'hers for she' errors as 'her's for she's' errors.

One obvious way of controlling for this possibility is by focusing specifically on the rate of non-contracted agreeing verb forms with non-nominative feminine subjects and determining whether this rate is significantly greater than 10%. The results of such an analysis are presented in Table 9 from which it can be seen that although Gail did not produce any non-contracted agreeing verbs with non-nominative feminine subjects, both Anne and Becky produced such errors at rates significantly greater than 10%. Thus, Anne produced non-contracted agreeing verbs with non-nominative feminine subjects at a rate of $3/4 = 75\%$ (which is significantly greater than 10% at $p < 0{\cdot}01$ by a Binomial test), and Becky produced non-contracted agreeing verbs with non-nominative feminine subjects at a rate of $6/19 = 33{\cdot}3\%$ (which is significantly greater than 10% at $p < 0{\cdot}01$ by a Binomial test). The implication is that the high rates of agreeing verbs with non-nominative feminine subjects in these children's speech cannot be explained away as genitive for nominative errors.

A second possible objection is that although there are several instances of agreeing copulas and auxiliaries with non-nominative subjects there are no

TABLE 10. *Distribution of NOM/Non-NOM 3psg subjects with/without Agreeing verb forms by gender and verb-type for Abe (expected values in parentheses)*

	He	Him	She	Her
Agreeing Lexical verbs	17 (16·92)	0 (0·08)	1 (0·67)	3 (3·33)
Agreeing copulas and auxiliaries	91 (90·56)	0 (0·44)	4 (2·83)	13 (14·17)
Non-Agreeing verbs	99 (99·52)	1 (0·48)	5 (6·50)	34 (32·50)
	Expected rate of Agreeing verbs with Non-NOM masculine subjects = 0·5%		Expected rate of Agreeing verbs with Non-NOM feminine subjects = 83·3%*	
	Observed rate of Agreeing lexical verbs with Non-NOM masculine subjects = 0%		Observed rate of Agreeing lexical verbs with Non-NOM feminine subjects = 75·0%*	
	Observed rate of Agreeing cop/aux with Non-NOM masculine subjects = 0%		Observed rate of Agreeing cop/aux with Non-NOM feminine subjects = 76·5%*	

* Significantly greater than 10% by Binomial test at $p < 0·05$.

instances of agreeing lexical verbs with non-nominative subjects. This raises the possibility that there may be something special about agreeing function verbs that interacts with the child's grammar to allow non-nominative subjects with agreeing copulas and auxiliaries, but not with agreeing lexical verbs. It is difficult to assess the validity of this objection on Anne, Becky and Gail's data since agreeing lexical verbs were so rare in all of these children's speech that the rate at which agreeing lexical verbs would be expected to occur with non-nominative subjects was never significantly higher than 10%. This issue was therefore investigated by searching the CHILDES database for children who produced a reasonably large number of agreeing lexical verbs and a reasonably high proportion of non-nominative 3psg subjects during the same developmental period.

The only child we were able to find who met this criterion was Abe (Kuczaj, 1976), who produced 21 agreeing lexical verbs with 3psg case-marked subjects, and 51/268 non-nominative 3psg subjects during a 9 month period between 2;5·10 and 3;2·7. A detailed breakdown of Abe's data is presented in Table 10, and a list of all Abe's agreeing verbs with non-nominative subjects is presented in Table 11.

It is clear from Table 10 that the overall pattern of Abe's data is remarkably similar to that reported for Anne, Becky and Gail. Thus, like Anne Becky and Gail, Abe has a much higher expected rate of agreeing verbs with non-nominative feminine than with non-nominative masculine subjects (83·3% vs. 0·5%); and, like Anne Becky and Gail, Abe produces agreeing verbs with non-nominative feminine subjects at a much higher rate

TABLE 11. *All instances of Agreeing verb forms with Non-NOM 3psg subjects in Abe's data*

Lexical verbs
Sometimes her barks nice (sometimes her don'ts)
Her has a tummy ache
Because her gets real angry at people that jump on the bed and break my stuff
Non-contracted copulas and auxiliaries
Her is pretty fine and I like her
Her was a baby
Three weeks old and her was little tiny
Sometimes her don'ts
Sometimes her is
Is her a little girl?
And this girl's dirty too so her has to take a bath
Contracted copulas and auxiliaries
Why we are daddies and her's girl?
Her's in the bathroom teeteeing
Sometimes her's not friendly
Her's going to put me in bed and read to me
Her's already clean

than the ATOM would predict (16/21 = 76·2%, which is significantly greater than 10% at $p < 0{\cdot}00001$ by a Binomial test). However, one important difference between Abe's data and Anne, Becky and Gail's data is that, because of the very high proportion of non-nominative feminine subjects, and the relatively high number of agreeing lexical verbs, Abe also has an expected rate of agreeing lexical verbs with non-nominative feminine subjects that is significantly greater than 10%. This means that Abe's data allow a reasonable test of the hypothesis that there is an important difference between errors involving agreeing function verbs and errors involving agreeing lexical verbs. In fact, however, Abe's data provide very little support for this hypothesis. Thus, Abe produces agreeing lexical verbs with non-nominative feminine subjects at a rate of 3/4 = 75·0%, which is very similar to the rate at which he produced agreeing function verbs with non-nominative feminine subjects (13/17 = 76·5%), and is significantly greater than 10% ($p < 0{\cdot}01$ by a Binomial test). This finding suggests that the absence of agreeing lexical verbs with non-nominative subjects in Anne, Becky and Gail's data is not a theoretically interesting fact, but a straightforward consequence of the low frequency of agreeing lexical verbs in their speech. More importantly, it shows that when one focuses on datasets in which the expected error rate is reasonably high, one finds rates of both agreeing function verbs and agreeing lexical verbs with non-nominative subjects that are significantly higher than the ATOM would predict.

DISCUSSION

The aim of the present study was to take a fresh look at the ATOM by testing its central prediction that the rate of agreeing verbs with non-nominative subjects in English-speaking children's speech is so low that such errors can be reasonably disregarded as noise in the data. The results count against the ATOM in three ways.

First, they show that it is surprisingly difficult to find children, either in the Manchester corpus or in the CHILDES database as a whole, who produce enough 3psg non-nominative subjects to allow a reasonable test of the ATOM to be made. The implication is that the low rates of agreeing verbs with 3psg non-nominative subjects in most of the datasets currently available reflect the low rates of non-nominative subjects in these datasets and should not be taken as evidence in favour of the ATOM.

Second, they show that even for those children who do produce a reasonably large number of non-nominative 3psg subjects, overall expected error rates tend to hide very different expected error rates for agreeing verbs with masculine non-nominative subjects and agreeing verbs with feminine non-nominative subjects. Thus, in all of the datasets analysed in the present study, the only area of the data in which the expected rate of agreeing verbs with non-nominative subjects was high enough to allow the ATOM to be properly tested was the use of feminine case-marked pronouns. The implication is that the kind of low overall error rates reported in previous research should not be taken at face value since they collapse together potentially informative error rates in the feminine part of the system and trivially low error rates in the masculine part of the system.

Third, they show that when one does focus on areas of the data in which the expected rate of agreeing verbs with non-nominative subjects is reasonably high, one finds that the observed error rate is also reasonably high, and much higher than the ATOM would predict. Interestingly, this is true not only for the children whose data were analysed in the present study, but also for Nina (the child whose data are generally taken to provide the strongest support for the ATOM). The implication is that, when analysed appropriately, the data on children's use of non-nominative 3psg subjects not only fail to support the ATOM, they actually count directly against it.

These results raise serious doubts about the claim that children's use of non-nominative subjects can be explained in terms of Agreement optionality, and suggest that previous support for this view reflects the failure to operationalize the ATOM correctly. More importantly, they suggest the need for a model of case-marking error that can not only explain why children make pronoun case-marking errors, but also why some children produce agreeing verbs with non-nominative subjects as often as they do. One possible explanation is that pronoun case-marking errors reflect intermediate stages in a paradigm-building process (Rispoli, 1994, 1998*a*, *b*, 1999). According to this view, although children know the grammatical features of Person, Number and Case, they have still to fully learn the word-specific paradigms of the personal pronouns, and produce non-nominative subjects with both agreeing and non-agreeing forms when the grammatical features of Person, Number and Case fail to converge on the correct form of the relevant pronoun.

Another possible explanation is that pronoun case-marking errors reflect the absence of abstract knowledge of Case and Agreement. According to this view, to the extent that young children have knowledge of Case and Agreement, this knowledge is still largely embedded in lexically-specific constructions, and children produce pronoun case-marking errors with both agreeing and non-agreeing verbs when they attempt to use case-marked and/or agreeing verb forms more productively. This possibility is broadly consistent with the results of a recent study by Wilson (2003) which shows that young children's use of inflection tends to pattern around a relatively small set of high frequency pronominal subjects, and is hence much more lexically-specific than would be predicted by most current generativist accounts.

In addition to their theoretical implications, however, the results of the present study also have a number of methodological implications for the field. First, they underline the dangers of arguing for syntactic knowledge on the basis of the low frequency of particular kinds of errors in children's speech. Such arguments are very common in the literature. However, they are impossible to evaluate in the absence of information about the rate at which such errors would be expected to occur by chance even if such knowledge were absent. This is because the frequency with which particular kinds of grammatical errors occur in children's speech is critically dependent on the frequency with which instances of the relevant grammatical features occur in the speech of the children concerned. Errors that seem to be notable by their absence may therefore simply be errors that are extremely unlikely to occur given the frequency with which particular grammatical features occur in the data. The implication is that the low frequency of particular kinds of errors should only be considered of theoretical interest if it can be shown that the frequency with which such errors would be expected to occur by chance is reasonably high.

Second, they suggest that, even when one has allowed for the distribution of the relevant grammatical features in children's speech, low overall error rates can still be rather misleading. This is partly because any rote-learning on the part of the child will have the effect of artificially increasing the frequency with which children produce adultlike combinations and hence of artificially inflating the expected error rate (Rispoli, 1999; Wilson, 2003). However, it is also because the different components of particular grammatical systems are rarely evenly distributed in children's speech. Thus, low error rates in high frequency parts of the system can easily hide theoretically interesting pockets of error in low frequency parts of the system (Rubino & Pine, 1998). The implication is that even error rates that are significantly lower than one would expect by chance should be treated with some caution, since they are often open to a variety of possible interpretations.

Finally, they provide a good illustration of the way in which performing more detailed analyses of children's errors can lead one to rather different conclusions about the theoretical significance of these errors than one might otherwise have reached. Thus, the relatively high rates of agreeing verbs with nominative subjects found in the present study not only count against the ATOM, but also suggest that, once one takes into account the rate at which such errors would be expected to occur by chance, the occurrence of such errors is a surprisingly robust phenomenon. This raises the possibility that rates of error in other aspects of children's production may also reflect peculiarities in the distribution of particular grammatical features and particular lexical forms, and hence that the commonly held assumption that children's early speech is essentially free of errors of commission may actually be false. Of course, whether or not this turns out to be the case is an empirical question. What is clear from the results of the present study, however, is that it is a question that is much more difficult to answer than has tended to be assumed in much of the previous literature. Certainly, dismissing low frequency errors as noise on theoretical grounds would appear to be a dangerous strategy, and one that, we would argue, is more likely to hold back than to advance our understanding of the language acquisition process.

REFERENCES

Bloom, L. (1970). *Language development: form and function in emerging grammars*. Cambridge, MA: MIT Press.

Brown, R. (1973). *A first language: the early stages*. Cambridge, MA: Harvard University Press.

Budwig, N. (1989). The linguistic marking of agentivity and control in child language. *Journal of Child Language* **16**, 263–84.

Chiat, S. (1981). Context-specificity and generalisation in the acquisition of pronominal distinctions. *Journal of Child Language* **8**, 75–91.

Gruber, J. (1967). Topicalization in child language. *Foundations of Language* **3**, 37–65.

Huxley, R. (1970). The development of the correct use of subject personal pronouns in two children. In G. Flores d'Arcais & W. Levelt (eds), *Advances in Psycholinguistics*. Amsterdam: North Holland.

Kuczaj, S. (1976). -ing, -s and -ed: a study of the acquisition of certain verb inflections. Unpublished doctoral dissertation, University of Minnesota.

Loeb, D. F. & Leonard, L. B. (1991). Subject case marking and verb morphology in normally developing and specifically language-impaired children. *Journal of Speech and Hearing Research* **34**, 340–6.

MacWhinney, B. (2000). *The CHILDES project: tools for analyzing talk* (*3rd edition*). Mahwah, NJ: Erlbaum.

Radford, A. (1998). Genitive subjects in child English. *Lingua* **106**, 113–31.

Rispoli, M. (1994). Pronoun case overextension and paradigm building. *Journal of Child Language* **21**, 157–72.

Rispoli, M. (1998*a*). Me or My: two different patterns of pronoun case errors. *Journal of Speech, Language and Hearing Research* **41**, 385–93.

Rispoli, M. (1998*b*). Patterns of pronoun case error. *Journal of Child Language* **25**, 533–54.

Rispoli, M. (1999). Case and agreement in English language development. *Journal of Child Language* **26**, 357–72.

Rubino, R. B. & Pine, J. M. (1998). Subject–verb agreement in Brazilian Portuguese: what low error rates hide. *Journal of Child Language* **25**, 35–59.

Schütze, C. T. (1997). *INFL in child and adult language: agreement case and licensing*. Doctoral Dissertation, Massachusetts Institute of Technology.

Schütze, C. T. (2001). Productive inventory and case/agreement contingencies: a methodological note on Rispoli (1999). *Journal of Child Language* **28**, 507–15.

Schütze, C. T. & Wexler, K. (1996). Subject case licensing and English root infinitives. In A. Stringfellow, D. Cahma-Amitay, E. Hughes & A. Zukowski (eds), *Proceedings of the 20th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.

Suppes, P. (1974). The semantics of children's language. *American Psychologist* **29**, 103–14.

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

Vainikka, A. (1994). Case in the development of English syntax. *Language Acquisition* **3**, 257–325.

Webster, B. & Ingram, D. (1972). The comprehension and production of the anaphoric pronouns he, she, him, her in normal and linguistically deviant children. *Papers and Reports on Child Language Development* **4**, 55–73.

Wexler, K. (1998). Very early parameter setting and the unique checking constraint: a new explanation of the optional infinitive stage. *Lingua* **106**, 23–79.

Wexler, K., Schütze, C. T. & Rice, M. (1998). Subject case in children with SLI and unaffected normal controls: evidence for the Agr/Tns omission model. *Language Acquisition* **7**, 317–44.

Wilson, S. (2003). Lexically specific constructions in the acquisition of inflection in English. *Journal of Child Language* **30**, 75–115.