Can input explain children's me-for-I errors?

Journal of Child Language / Volume 36 / Issue 05 / November 2009, pp 1091 - 1114
DOI: 10.1017/S0305000909009350, Published online: 16 February 2009

J. Child Lang. **36** (2009), 1091–1114. © 2009 Cambridge University Press
doi:10.1017/S0305000909009350 Printed in the United Kingdom

Can input explain children's me-for-I errors?*

MINNA KIRJAVAINEN AND ANNA THEAKSTON
University of Manchester, UK

AND

ELENA LIEVEN
Max Planck Institute for Evolutionary Anthropology, Leipzig

(*Received* 12 *January* 2007. *Revised* 24 *July* 2008. *First published online* 16 *February* 2009)

ABSTRACT

English-speaking children make pronoun case errors producing utterances where accusative pronouns are used in nominative contexts (*me do it*). We investigate whether complex utterances in the input (*Let me do it*) might explain the origin of these errors. Longitudinal naturalistic data from seventeen English-speaking two- to four-year-olds was searched for 1psg accusative-for-nominative case errors and for all 1psg preverbal pronominal contexts. Their caregivers' data was also searched for 1psg preverbal pronominal contexts. The data show that the children's proportional use of me-for-I errors correlated with their caregivers' proportional use of *me* in 1psg preverbal contexts. Furthermore, the verbs that children produced in me-error utterances appeared in complex sentences containing *me* in the input more often than verbs that did not appear in me-for-I errors in the children's speech. These findings are discussed in the context of current explanations for children's case marking errors.

INTRODUCTION

It is well known that many English-speaking two- to three-year-old children make pronoun case errors producing utterances such as *me do it*, *her going*, *him was crying* where accusative (ACC: all abbreviations used in

[*] We would like to thank Brian and his family, and Fraser and his family, for their time and patience, the team of research assistants who collected and transcribed the data, and Jeannine Goh for supervising the dense database project. Thanks also to Julian Pine for helpful discussions about the data and to the editors and two anonymous reviewers for helpful suggestions. This research was funded by a Max Planck PhD studentship to the first author. Address for correspondence: Minna Kirjavainen, School of Psychological Sciences, University of Manchester, Oxford Road, Manchester, M13 9PL, UK. e-mail: minna.kirjavainen@manchester.ac.uk

the current paper are defined in footnote[1]) pronouns are used in nominative (NOM) contexts, and utterances such as *my found it* where genitive (GEN) pronouns are used in nominative contexts (Budwig, 1989; Pine, Rowland, Lieven & Theakston, 2005; Rispoli, 1998; Schütze & Wexler, 1996). These errors cannot be explained by children being confused as to which pronoun to use in different contexts, as errors where a NOM pronoun is used in a non-NOM context are rare (e.g. *She hit I*) (Schütze & Wexler, 1996; but see Rispoli, 1998, for some counter-examples). Moreover, children often produce both correct and incorrect pronominal forms at the same stage in development, so a lack of lexical knowledge cannot account for these errors. Errors are concentrated in a small set of persons and numbers – 1psg (*me/my*), and 3psg FEM (*her*) errors are widely reported (e.g. Budwig, 1989; Pine, Rowland, Lieven & Theakston, 2005; Rispoli, 1998, 1999; Schütze & Wexler, 1996; Vainikka, 1994) whereas errors with other pronouns have been observed less frequently (Rispoli, 2005). Children are not uniform in their pronominal error production either. Some children produce many errors, others virtually none, although these differences could in part reflect sampling problems. Some children who make errors do so with a range of pronominal forms, others with only one particular pronoun type (Pine *et al.*, 2005).

Explanations for the pattern of case errors observed in children's speech have been suggested by researchers from a range of theoretical perspectives. According to the Agreement/Tense Omission Model (ATOM) (Wexler, 1998), case errors occur when children fail to check both tense (TNS) and agreement (AGR) in their utterances. As AGR is thought to licence NOM case, failing to check AGR leads to the production of utterances where ACC or GEN pronouns may appear in NOM contexts (e.g. *me do it*). According to the model, children should not produce errors where non-NOM pronominal forms are produced alongside verbs that are overtly marked for AGR (e.g. *her cries*, *me is happy*), as the presence of AGR on the verb should licence the use of NOM pronouns in subject position (Schütze & Wexler, 1996; see also Wexler, Schütze & Rice, 1998, for an analysis on typically developing and SLI children; although see Pine *et al.*, 2005, for counter-arguments and data).

Although the ATOM can account for some of the linguistic patterns observed, some research has indicated that in fact not all of the ATOM's predictions are correct (Ambridge & Pine, 2006; Pine, Joseph & Conti-Ramsden, 2004; Pine *et al.*, 2005). In addition, the ATOM does not address

[1] NOM = nominative case, ACC = accusative case, GEN = genitive case, 1psg = first person singular forms, 3psg = third person singular forms, FEM = feminine forms, V = verb, TNS = tense, AGR = agreement, SLI = specific language impairment, RI = root infinitive, INFL = development of finiteness, MASC = masculine forms.

the question of why some children produce many pronoun errors while others produce virtually none, why some pronominal forms are more susceptible to errors than others, or why different children make errors with different forms (Rispoli, 2005).

Rispoli's (1994, 1998, 1999, 2005) PARADIGM BUILDING model holds that pronoun case errors occur when the retrieval of a given pronoun fails for one reason or another. Importantly, Rispoli has offered explanations for some aspects of the pronominal error pattern that have been previously overlooked, for example, the presence of both me-for-I and my-for-I errors and the variable pronoun error rate across children. Rispoli (1998) explains the presence of both me-for-I and my-for-I errors partly in terms of the relative frequency of use of different pronominal forms. He categorized a group of twelve children as 'me-children' (9/12 who predominantly overextended *me*) or 'my-children' (2/12 who predominantly overextended *my* – one child had equal *me* and *my* error rates), providing evidence that most children make most of their 1psg errors with a single pronominal form. Rispoli then calculated a me-error rate and the overall production of *me* for each of his participants and found that at a certain point in development, me-for-I errors were highly positively correlated with the correct production of *me* in the children's speech.

Rispoli (2005) further explains differential error rates across children by a combination of lexical retrieval problems and the development of finiteness (INFL). He suggests that children who produce few finite verb forms, but attempt to produce a diverse range of pronouns, are likely to produce many errors as their underdeveloped case feature in INFL fails to determine the correct pronoun form from a choice of many. Children who do not attempt to produce many pronoun forms are less likely to make errors because only a fairly limited activation from the case feature in INFL is needed to produce the correct pronominal form from a small selection.

Input

Some researchers have reported a close link between various aspects of the input and the acquisition patterns observed in children, (e.g. Freudenthal, Pine, Aguado-Orea & Gobet, 2007; Huttenlocher, Vasilyeva, Cymerman & Levine, 2001; Naigles & Hoff-Ginsberg, 1998; Rowland & Pine, 2000; Theakston, Lieven, Pine & Rowland, 2001, 2002, 2004, 2005; Theakston, Lieven & Tomasello, 2003). However, many researchers have claimed that pronoun case errors could not be explained by the speech children hear (e.g. Hoekstra & Hyams, 1998; Wexler, 2003; Wexler *et al.*, 1998), because NOM pronouns are more frequent in the input than non-NOM pronouns (with the possible exception of 3psg FEM *her*), yet errors typically involve the extension of non-NOM pronouns into NOM contexts (Schütze, 1997, based on the Brown corpus (Francis & Kucera, 1982) and analyses of three children's naturalistic data on CHILDES). In addition, GEN forms do not appear immediately preceding verbs without intervening material in the input, yet these errors (e.g. *my do it*) occur in children's speech.

Budwig (Budwig, 1996; Budwig & Wiley, 1995) has suggested that input factors may play a part in pronoun errors, in particular the distribution in maternal speech of *me, my, mine, Mummy* and the child's own name in relation to sentence position and the semantics with which children use particular verbs. Tomasello (2000, 2003) has also suggested that ACC-for-NOM errors could be explained by input – not in the form of simple sentences or overall frequency of particular pronouns – but by omitting the beginning of complex sentences (e.g. *Let me do that*, See *her opening it*). This proposal is situated within the broader usage-based, constructivist approach to language acquisition, which suggests that children's early utterances (grammatical as well as ungrammatical) are largely (but not solely) a reflection of the speech they have been exposed to. If a particular word or construction appears regularly in the input, children are more likely to learn and start using that word or construction in their own speech (e.g. Rowland & Pine, 2000; Theakston *et al.*, 2001, 2002). Within this approach, children are expected to extract lexically specific chunks from complex but relatively frequent utterances in the input, as well as learning shorter utterances as a whole. Thus, errors where a NOM pronoun is erroneously replaced with an ACC pronoun could be due to children hearing both I + verb (e.g. *I do that every day*) and me + verb (e.g. *Let me do it*) sequences, which could result in children having two competing constructions for a given verb (e.g. *I/me + do*) when referring to themselves. Children are also, under this view, abstracting from this lexically specific material from the outset. Abstraction is thought to take place more rapidly if a large number of different types occur within a given slot, for example *me + V*, where *V* is instantiated by a number of different verbs. This means that although the earliest stages of learning might be tied to specific lexical items, children will form links between items, resulting in more abstract schemas, although the scope of this abstraction changes with development. Abstract schemas allow children to produce combinations not previously heard in the input, for example *me want that*. A defining feature of the usage-based approach is, then, that the categories and generalizations of the linguistic system are built up rather than pre-given.

Wexler (2003) asserts that the source of ACC-for-NOM errors could not be complex sentence input, as to make ACC-for-NOM errors, children would have to ignore the fact that they never hear the ACC pronoun in sentence-initial positions. Note, however, that children do hear the form *her* in utterance-initial position, albeit performing a different syntactic role (GEN) (e.g. *Her drink is over there*).

There is some empirical support for the claim that some of the children's errors may reflect the learning of sequences of words within a complex sentence in the input. First, Theakston *et al.* (2003) elicited novel verbs from children aged 2;6 to 3;0 in 3psg contexts. The verbs had previously been modelled for the children in either declaratives only (*it tams*), questions only (*will it tam?*) or in both declaratives and questions. The children's production of 3psg vs. unmarked verb forms in 3psg contexts was related to the sentence context in which novel verbs were previously modelled. Children who heard verbs modelled in declaratives produced 3psg verb forms in their speech, whereas children who heard novel verbs modelled in questions showed much lower use of 3psg verb marking in 3psg contexts, suggesting that they had extracted the form *it tam* from questions. Second, Freudenthal *et al.* (2007) conducted a computational modelling study which showed that errors where a root verb form is used in finite contexts (root infinitive (RI) errors) could be explained by gradual learning and generalization from the input. The model had an utterance-final and an utterance-initial bias for learning sequences of words, together with a capacity for generalization based on co-occurrences between adjacent words. Child-directed speech from English, Dutch, German and Spanish was fed into the model. The model's ability to combine utterance-initial subjects with utterance-final non-finite verbs generated novel, subject–verb sequences with RIs mirroring the different proportions of RI errors observed in the speech of children learning the four languages. Higher error levels were observed for German and Dutch because these are V2 languages in which utterances with complex verb phrases place the non-finite verb form at the end of the utterance. Thus the suggestion from these studies is that children extract sections of more complex utterances from the input and that this sometimes results in errors (e.g. *Me do it* from *Let me do it; He go there* by combining utterance-initial *He* with *go there* from utterances such as *It can go there*).

The present paper reports an input driven analysis of the pronominal case errors observed in children's early speech. More precisely, the aim of the present analyses was to investigate whether 1psg ACC-for-NOM pronoun errors (e.g. *Me do it*) could be explained by the input children hear. Study 1 tackles the question as to why different children exhibit different error rates, and investigates whether this can be explained by the different proportions of me+V sequences in complex sentence structures (e.g. *Let me do it*, *Did you see me doing it?*) in the children's caregivers' speech. Study 2 examines me-for-I errors in a more detailed manner, to see whether children have had input of the specific me+V sequences that they produce in their me-for-I error utterances.

STUDY 1

The aim of Study 1 was to determine whether the difference in me-for-I error rates in children could be explained by the rate at which their caregivers produce me+V sequences.

METHOD

Corpora

A corpus analysis was conducted on data from seventeen monolingual English-speaking children. These children were: (1) the twelve children in the Manchester corpus (Anne, Aran, Becky, Carl, Dominic, Gail, Joel, John, Liz, Nicole, Ruth and Warren, recorded interacting with their mothers for two separate hours in every three-week period between the ages of 2;0 and 3;0; Theakston, Lieven, Pine & Rowland, 2001; available on CHILDES, MacWhinney, 2000); (2) two children (Brian & Fraser) recorded in Manchester, UK, interacting primarily with their mothers in their home environments by researchers at the Max Planck Child Study Centre, University of Manchester (Brian: five separate hours in every week period between the ages of 2;0 and 3;2, then five hours during a one-week period in every month until 5;0; Fraser: five separate hours in every week period for six weeks from 2;0 and 3;0, between these periods for five separate hours in one week of every month); (3) three children (Abe, Nina, Peter) taken from the CHILDES database (MacWhinney, 2000) (Abe, 210 recordings collected between 2;4 and 5;0 while interacting with his parents (Kuczaj, 1976); Nina, 59 recording sessions, collected between 1;11 and 3;11 while interacting mostly with her mother (Suppes, 1974); Peter, 20 recordings of between three and four-and-a-half hours duration collected between 1;4 and 2;10 while interacting with researchers and his mother (Bloom, Hood & Lightbown, 1974)).

Determining the error-files

To ensure that the children were at roughly the same developmental period, the first recording in which each child's MLU was $\geqslant 2$ was established using the MLU program in CLAN (MacWhinney, 2000). The first MLU $\geqslant 2$ file and nine files following it were searched for me+V errors in each child's speech.

If a child reached MLU $\geqslant 2$ before file six, resulting in fewer than five files preceding the MLU $\geqslant 2$ file for the parents' input analysis (explained below), the child's errors were analyzed in files 6–15 and files 1–5 were used as the input files. The children this affected were: Aran (who reached MLU2 in file 4), Carl (who reached MLU2 in file 3), and Abe (who had reached MLU3 in his first file). MLU tends to fluctuate across recordings, which meant that Carl had an MLU of less than 2 in file 6, his first analysis file.

After determining the first file where each child reached MLU $\geqslant 2$, or file 6 for those who reached MLU $\geqslant 2$ earlier, an additional search was conducted to find each child's first utterances in which the NOM and ACC 1psg pronouns occurred in the correct sentence structure, i.e. the child produced the NOM pronoun *I* in sentence subject position and the ACC pronoun *me* in a V+me context (in an object or ditransitive construction). This was to establish that the child had at least some knowledge of the correct usage of the two pronoun cases and therefore that any pronoun case errors were not due to the child not knowing the relevant lexical forms. Single-word utterances were not taken as evidence that the child could use case marked pronouns in their correct syntactic position. The utterance *excuse me*[2] was also ignored in the V+me search as this sequence was regarded as a frozen phrase and did not provide clear evidence that the child could use the 1psg ACC pronoun correctly. All children apart from Brian, Nina, Peter and Warren had produced NOM and ACC pronouns in correct sentence structures before MLU $\geqslant 2$. For these four children, the analysis files consisted of the first ten files immediately following the child's first correct use of both NOM and ACC 1psg pronouns. For Brian, this also means that his first analysis file had MLU <2 due to fluctuations in MLU. Table 1 shows the error-file details for each child.

Me error-search in the children's speech

In the ten error files, all the me+V errors were extracted from the transcripts. Only me+V sequences were included, and not for example me+adjective (*Me happy*) sequences.[3] In addition, errors that were preceded by an inaudible sequence or a schwa were excluded (e.g. *xxx me do it*)

[2] Two children's target sentences were affected by this decision (Brian and Warren). The first verb+me sequence these two children produced was *excuse me*. The first non-formulaic verb+me sequence following this frozen phrase was taken as the first correct ACC pronoun use.

[3] In the current study, we were interested in whether the non-finite complement clause construction was related to me-for-I errors. However, we also conducted an analysis including me+adjective errors and input sequences. Only me+adj sequences in which the adjective was not premodifying a noun were included in the analysis as these appeared to be me-for-my errors (e.g. *There me big tower fall down* [Ruth, file 18b]). Me+adj errors were only produced by Ruth ($n=12$), Abe ($n=1$) and Fraser ($n=1$). Only Joel and Abe's mothers produced me+adj sequences – one each (*you really make me happy Abe* (Abe's mother file 2), *Drives me mad* (Joel's mother, file 12b). If these errors and input are included in the analysis with the me+verb sequences, the correlation remains significant ($r_s=0{\cdot}512$, $n=17$, $p=0{\cdot}036$).

TABLE 1. *Error files for each child, MLU and age for the error analysis*

Child	Files	MLU in 1st file	MLU in last file	Age
Abe	6–15	4·607	5·021	2;5.16–2;6.14
Anne	11–20	2·270	2·461	2;1.20–2;4.14
Aran	6–16	2·134	2·440	2;1.07–2;4.20
Becky	10–19	2·128	2·715	2;3.06–2;6.19
Brian	118–127	1·900	1·833	2;5.28–2;6.12
Carl	6–15	1·753	2·289	1;10.19–2;1.16
Dominic	12–21	2·029	2·006	2;2.16–2;5.22
Fraser	11–20	2·005	1·503	2;0.15–2;0.27
Gail	7–16	2·064	2·371	2;2.05–2;4.28
Joel	12–21	2·113	2·398	2;3.04–2;6.12
John	19–28	2·158	2·557	2;5.20–2;8.21
Liz	11–20	2·065	2·729	2;2.23–2;5.22
Nicole	16–25	2·049	2·126	2;6.11–2;9.09
Nina	9–18	2·096	3·085	2;1.06–2;3.18
Peter	7–16	3·018	3·444	2;1.0–2;7.13
Ruth	18–27	2·091	2·131	2;5.06–2;8.21
Warren	10–19	2·501	3·051	2;1.14–2;4.13

as it would be difficult to determine whether this was an error or whether the child was trying to produce a grammatical utterance (e.g. *Let me do it*). No intervening words were allowed in the error-utterances apart from the adverb *just* (e.g. *Me just do it*) and negation *no/not* (e.g. *Me no do it*).

A search for all other sentence subject (not object) 1psg pronoun + verb contexts was also conducted. Hence, all I + V and my + V sequences were extracted from the child's speech. I + V utterances included instances of *I* with or without a contracted auxiliary or main verb form (i.e. *I, I'm, I've, I'll*). Null forms (e.g. *go there*) were not included in the analysis. Each child's proportional use of me + V errors as a function of all of his/her sentence subject 1psg pronoun + V utterances (I + V, me + V, my + V) was then calculated.

Many child language corpus analyses only analyze spontaneous child utterances, while excluding imitations. However, in this study, me + V errors that were produced as imitations of prior adult utterances were included (if a caregiver said *Let me do that for you*, and the child's response was *Me do that*, this would be included as an erroneous use of the ACC form *me*). These utterances were included because we hypothesize that children's errors originate in their use of short sequences taken from complex utterances in the input. However, as there were only five errors in total that followed adult utterances of this type occurring in the speech of three different children, direct imitations alone do not account for the pattern of results found.

Search for me + V input

The five files preceding the child's error files were searched for me + V sequences in the input. Five files were selected because this provides a sizable input corpus while ensuring that the data for most children could be taken from the first file where MLU reached 2 without overlapping with this input sample. For all the children apart from Abe, the mother's speech was searched for these sequences because the mothers were most often the parent present during the recording sessions.[4] As both Abe's mother and father regularly interacted with him during the recordings, the speech of both was included in Abe's input analysis.

All complex sentences containing a non-finite complement clause that contained a me + V sequence were extracted from the caregivers' data. Consistent with the analysis of the children's data, no intervening words were allowed apart from the adverb *just* (e.g. *Let me just do this and then we'll go out*) and negators, although in practice no negators occurred within me + V sequences in the input. An additional search was also conducted in which all other me + V sequences occurring within a single utterance in the input were extracted (me + V sequences that were divided by a sentence boundary were excluded, e.g. *I thought you gave it to me. Have you seen it?*). These additional me + V sequences consisted of utterances containing tag questions (*You gave it to me, didn't you?, n* = 13, in six children's input), ungrammatical utterances and utterances that are grammatical only when used in incredulity constructions (**Me don't +//, *Me put it away? n* = 2, in two children's input), and a complex sentence containing a finite complement (*Show me is there any dark left*, *n* = 1) (see Appendix for the utterances and the children whose input they were found in). The status of these utterances as me + V sequences was less clear due to deviant intonation contour (tags, and possibly the finite complement clause), and adult error. Therefore, these me + V sequences were grouped together and a second analysis was run where both clear me + V utterances and more ambiguous me + V sequences were taken as input for me + V sequences in the children's speech.

To determine the caregivers' proportional use of me + V sequences as a function of all 1psg pronoun + V contexts, all I + V sequences were also extracted from the caregivers' speech, including uses of *I* with a clitic auxiliary or main verb form.

[4] In Peter's case, the recordings usually consisted of interaction between the child, mother and researcher(s) and therefore his data was analysed in two ways: (1) so that only the mother's data, and (2) so that all adult data were included in the input analysis. The results are the same regardless of whether only the mother or all adult data were included.

Three control analyses

(*1*) *The proportional use of me in the input*. It is possible that there is a relation between the me-error rate observed in children and the overall frequency of use of the 1psg ACC pronoun *me* in the input, rather than any particular influence of the presence of me+V sequences in the input. Therefore, the total number of ACC and NOM 1psg pronouns in the input files examined above for each child was determined and the proportion of 1psg ACC pronouns over 1psg ACC+1psg NOM pronouns calculated, including clitic forms as above.

If it is the overall pronoun rates rather than the proportion of me+V sequences in the input that result in errors, we should find a correlation between the children's me-for-I error rates and the proportion of 1psg ACC pronouns in the input.

(*2+3*) *Complexity of input*. It is possible that children whose caregivers produce more complex language input in general (i.e. longer and more complex utterances) find it more difficult to acquire the sentence subject grammar than children whose caregivers produce less complex language. This could explain the me-for-I errors rather than me+V sequences in the input. To investigate this, we conducted two analyses. First, we determined the caregivers' MLU in the input files and ran a correlation analysis comparing the children's error rates to these MLU scores. If input complexity in terms of utterance length causes me-for-I errors, we should find that the mothers' MLUs correlate with their children's me-for-I error rates.

Second, it could be that non-finite complement clause input results in children producing not only me-for-I errors, but also other error types. That is, this input may be difficult to organize in general and this may be evidenced not only as a correlation between me+V input and me-for-I errors but also as a correlation between me+V input and other error types. Thus, we determined the children's 3psg nominative (MASC and FEM) pronoun+verb-form error rates (i.e. lack of agreement and/or tense on the main verb [e.g. *S/he know it*], auxiliary verb [e.g. *S/he don't*],[5] and omitted auxiliaries [e.g. *S/he stuck, S/he going*]), in the same ten files included in the main analysis by dividing the number of errors by the total number of utterances with s/he+verb sequences (erroneous or correct). Only declarative sentences were included in the analysis. Self- and interlocutor repetitions were excluded. These error proportions were then compared to the me+V input proportions. If the non-finite complement clause input was affecting children's language, making it more prone to all types of

[5] No erroneous pronoun+modal combinations were produced in the search files by any of the children.

TABLE 2. *The number of ACC, NOM and GEN 1psg pronoun + verb sequences in the ten error files for Study 1*

CHILD	me + verb	I + verb	my + verb
Abe	3	164	0
Anne	18	275	2
Aran	1	653	1
Becky	3	691	2
Brian	0	5	0
Carl	1	139	5
Dominic	0	523	3
Fraser	0	68	1
Gail	3	229	13
Joel	4	345	2
John	0	229	0
Liz	4	601	5
Nicole	2	132	3
Nina	4	458	45
Peter	30	1332	22
Ruth	211	124	9
Warren	5	219	13

errors – not just me-for-I errors – the me + V input should correlate with the 3psg pronoun + verb form errors.

RESULTS

Me-for-I errors were relatively infrequent in the ten hours of data for many of the children – some children did not produce any errors during the age range included in the analysis (see Table 2 for the number of me + V, I + V and my + V sequences). The data violated the assumptions for parametric tests, and therefore rank-order correlations were conducted. The children's proportion of me-errors was compared to their caregivers' proportion of me + V sequences.

In the first analysis, only non-finite complement clauses containing me + V sequences were included as input (i.e. other types of me + V sequences were excluded). A Spearman's rho (two-tailed) test showed a significant correlation between the proportion of me + V errors the children produced and the proportional use of me + V sequences in the input ($r_s = 0{\cdot}554$; $n = 17$; $p = 0{\cdot}021$).[6] In the second analysis, non-finite complement clauses as well as the more ambiguous me + V sequences (tags,

[6] Spearman's rho test is less reliable with ties with small sample sizes, and as there was a four-way tie in the children's rank due to four children not making any me-errors, permutation tests (of 1000 permutations) were also run on the data. The permutation test results were the same as the Spearman's rho results in all analyses in Study 1.

ungrammatical utterances and the complex sentence) were treated as input. A Spearman's rho (two-tailed) test showed a significant correlation between the proportion of me-errors the children produced and their caregivers' production of me+V sequences ($r_s = 0.523$; $n = 17$; $p = 0.031$).

The control analyses showed that: (1) the overall proportion of 1psg ACC over 1psg NOM+ACC pronouns in the caregivers' speech was not correlated with the proportion of me-errors the children produced ($r_s = 0.31$; $n = 17$; $p = 0.229$[7]); (2) the children's error rates did not correlate with the caregivers' MLU ($r_s = 0.211$; $n = 17$; $p = 0.417$); and (3) we also found no correlation between the proportion of children's 3psg pronoun+verb errors and me+V sequences in the input, regardless of whether we only included the (more common) MASC pronoun (only me+V input included: $r_s = 0.215$; $n = 17$; $p = 0.408$; less clear me+V input also included: $r_s = 0.230$; $n = 17$; $p = 0.374$) or both MASC and FEM pronouns (only me+V input included: $r_s = -0.029$; $n = 17$; $p = 0.911$; less clear me+V input also included: $r_s = -0.012$; $n = 17$; $p = 0.963$).

SUMMARY

A significant correlation was found between the proportion of me+V errors children make and the proportion of me+V sequences in the adult speech they hear. This was the case whether only non-finite complement clauses were included in the input analysis (e.g. *Let me do it*), or all me+V sequences, including tags and the like. There was no significant correlation between the overall proportional use of *me* in the input and children's erroneous use of *me* in preverbal position, showing that it is the specific sequences involving *me* in preverbal position in the input rather than an overall tendency to produce the form *me* in the input that is related to children's me+V error rates. Also, the control analyses showed that the errors could not be explained by the general complexity of the caregivers' language. Hence, it seems that children learn me+V sequences from the input. However, Study 1 tells us little about the detailed relation between specific me+V input and children's me+V errors. If the input directly contributes to me+V errors, we would expect to find that children produce me-errors with the same verbs that they have heard produced in me+V sequences in the input. We examine this possibility in Study 2.

[7] The control correlation results are also non-significant if the number of 1psg ACC pronouns was taken as a proportion of the 1psg ACC+1psg NOM pronouns, excluding NOM forms occurring with cliticized auxiliary and main verb forms.

STUDY 2

The aim of Study 2 was to see whether the verbs children used in their me+V errors were the same verbs that they had previously heard being used in me+V sequences in their caregivers' speech.

METHOD

A corpus analysis was conducted on the data of the same seventeen children as in Study 1. The analysis proceeded in three stages.

Error search

We created a list of all the verb types for individual children that were produced in me+V errors by extracting all utterances containing a me+V error that occurred after the child had already produced at least one NOM and one ACC 1psg pronoun in the correct sentence structure. Verb types could be either a lexical verb or auxiliary, in contracted or freestanding form (although in practice no contracted forms were found in combination with *me* in the children's speech). If the pronoun was followed by an auxiliary+verb sequence, only the first verb type (in this case the auxiliary form) was documented. The adverb *just* and the negators *no* and *not* were ignored for the purposes of verb type classification. Only the first error for each verb form of each verb type was included in the analysis. The same verb was included more than once only if it appeared in several different forms (e.g. *me go, me going, me gone*). Error sequences which were preceded by a schwa or an inaudible word/sequence were excluded from the analysis due to the difficulty of determining whether the utterance was a true me-error or whether the child was trying to produce a more complex sentence structure in which the 1psg ACC pronoun use is grammatical. Utterances with an omitted main verb (*me happy*) were also excluded.

All available data for the Manchester corpus children and Fraser, Nina and Peter was searched for errors. Abe's data was searched between 2;4.24 and 3;5.06 (the first 100 files; no errors occurred after file 28), and Brian's data was searched between 2;0.12 and 3;11.06 (the first 322 files; no errors occurred after file 305).

Counterpart search

A baseline was needed for how often any random set of verbs are preceded by me in the input. Therefore, a counterpart-utterance search was conducted for each child separately. For this search, all I+V-type sequences produced by the child were extracted from the same files in which the child's me+V errors had occurred. These sequences could include the adverb *just* or the negators *no* or *not*, although these forms were ignored for the purposes of verb type classification. An I+V counterpart for each me+V error sentence was then randomly chosen by selecting every third I+V sequence, where possible from the same file in which the error occurred. If there were no I+V sequences in that file, or if all I+V-type sequences in a particular error file already appeared in the counterpart list for that child, a counterpart was randomly selected from the previous file. The criterion for an I+V sequence was the same as in the me-error search except that I+V sequences that contained a contracted auxiliary verb form were excluded. This was because the 1psg ACC pronoun does not take clitics (*me'll, me've*) in adult language and no me+clitic errors appeared in the children's data. Including I+clitic verbs might create a set of control verbs of which a large proportion could not appear with me in the input, thus biasing the results.

The error and counterpart searches resulted in two lists of verbs for each child: a list of the verb types appearing in their me+V errors and a list of randomly selected verb types appearing in I+V sequences in their speech. The verbs in the two lists were not mutually exclusive; if a child produced an error *me do it*, and one of the random verb type counterparts for that child was *I do that*, the verb *do* would be included in both lists. For example, Joel's error and counterpart sentences and verb types were:

Errors (me+V)		Counterparts (I+V)	
me do it	(do)	I found	(found)
me fetch it	(fetch)	I just turn around	(turn)
me sing	(sing)	I fetch it	(fetch)
me have some more please	(have)	I don't know	(don't)
me go there	(go)	I want to sit there	(want)
me read it	(read)	I didn't put it in, did I?	(didn't)

Input search

All the files prior to each target sentence (me+V errors and I+V counterparts), as well as the previous discourse in the file in which the error/counterpart occurred were then searched for me+V input for each of the verb types listed for an individual child. Thus, if the target error sentence was *Me do it*, we looked for the sequence *me do* in the input. The same procedure was followed for verbs in the counterpart list. Thus, if the counterpart was *I have it*, we searched for the sequence *me have*; if the counterpart was *I don't know*, we searched for *me don't*.

We found in Study 1 that me+V input does not only appear in non-finite complement clauses (e.g. *Let me do it*), but that several caregivers also produced utterances where me was followed by a tag (e.g. *That's for me, is it?*). The intonation of these me+tag sequences is different to that of

TABLE 3. *Number of verb types and tokens produced in me+V errors; children who produced 4+ error types in Study 2*

Child	No. of verb types	No. of verb tokens
Abe	6	6
Anne	16	24
Becky	4	4
Brian	13	17
Dominic	7	7
Fraser	5	5
Gail	7	8
Joel	6	9
Liz	8	12
Nicole	13	17
Nina	9	11
Peter	17	34
Ruth	93	532
Warren	8	8

utterances containing non-finite complement clauses, so two input analyses were conducted.

(1) Only complex sentences containing non-finite complement clauses (e.g. *Let me do it*) were included.
(2) Any me+V sequence in the adult speech was seen as input. These included tags (e.g. *You showed me, didn't you*) ($n=6$), coordinated structures (*you and me can play*) ($n=1$) and adult errors ($n=1$). However, me+V sequences that were divided by a sentence boundary were not included (e.g. *I thought you gave it to me. Have you seen it?*).

The input search was a binary analysis, in which the target sentences were coded as either having or not having had input of me+V. We then calculated the proportion of verb types appearing in the me-error-lists and the proportion of verb types appearing in the counterpart-verb lists that had been preceded by input of me+V for each child separately.

RESULTS

The error and counterpart search

Three of the seventeen children did not produce more than one or two errors, and were hence excluded from the analysis. These children were Aran, Carl and John. The remaining fourteen children produced between 4 and 93 verb types in me-for-I errors. Table 3 shows the number of verb types produced in me+V errors for each of the 14 children included in this analysis.

TABLE 4. *Percentage of verb types preceded by me + V input (non-finite complement clauses)*

	% me + V error verb types	% I + V counterpart verb types
Abe	17	0
Anne	31	19
Becky	50	0
Brian	62	23
Dominic	57	14
Fraser	40	0
Gail	29	0
Joel	33	0
Liz	25	13
Nichole	31	23
Nina	22	11
Peter	41	29
Ruth	12	20
Warren	0	0
Mean	32	11

As can be seen in Table 3, Ruth produced 93 me-for-I verb type errors. She only produced 64 I + V types in all of her data. As the analysis was done by comparing the proportion of verb types that occurred in the input in me + V sequences, rather than exclude 29 of her me-error verb types, all of her me + V and I + V verb types were included, even though these two lists did not have the same number of targets. For all the other children the number of verb types in the me + V and I + V lists was the same.

Input search

The search for input of me + V sequences for the verb types listed in the error and counterpart lists revealed that for most of the fourteen children the verbs that were produced in me + V error sentences were more likely to have been preceded by me + V sequences in the input than the verbs that appeared in the counterpart (I + V) lists. Paired samples *t*-tests showed that this difference was statistically significant both when only utterances containing non-finite complement clause me + V sequences were included in the input sample ($t(13)=4{\cdot}519$, $p=0{\cdot}001$) (see Table 4) and when other me + V sequences such as tags were also included ($t(13)=4{\cdot}035$, $p=0{\cdot}001$) (see Table 5). As a further check to ensure that the auxiliary forms, for example *can't*, *don't, did*, found in the counterpart list were not biasing the results (as there were more auxiliary forms in the counterpart list than in the me-errors list), the proportion of verb forms found in me + V sequences in the input was calculated excluding verb types that constituted auxiliary

TABLE 5. *Percentage of verb types preceded by me + V input (non-finite complement clauses + other me + V sequences)*

	% me + V error verb types	% I + V counterpart verb types
Abe	17	0
Anne	38	25
Becky	50	0
Brian	62	46
Dominic	57	14
Fraser	40	20
Gail	29	0
Joel	33	0
Liz	25	13
Nichole	31	31
Nina	22	11
Peter	47	35
Ruth	12	20
Warren	0	0
Mean	33	15

verb forms. The results mirrored those reported above such that verb types produced in me-errors by the children were more likely to be found in me + V sequences in the input than verb types produced in I + V sequences by the children ($t(13) = 4{\cdot}23$, $p = 0{\cdot}001$).

SUMMARY

Study 2 shows that there is a relation between the specific verbs produced in complex utterances containing me + V sequences in the input and the specific verbs that appear in children's me + V errors. This suggests that the input may provide the source of these errors in children's speech.

OVERALL DISCUSSION

This paper aimed to investigate whether patterns of language use in the input, specifically the use of 1psg ACC pronoun + V sequences in complex utterances such as *Let me do it*, might provide an explanation for children's 1psg ACC pronoun + V errors, e.g. *me do it*. Study 1 showed that there is a significant correlation between the relative proportional use of me + V sequences in the input as a function of all 1psg + V contexts and the proportion of me + V errors in children's speech as a function of all 1psg + V contexts. Study 2 showed that the specific verbs that appear in me + V errors are also more likely to appear in me + V combinations in the input than verbs that appear in I + V combinations in the children's speech (even though these groups of verbs were not mutually exclusive). These findings suggest that complex utterances in the input that model me+V word order may provide an important source of English-speaking children's me-for-I errors. This suggestion, based on our corpus analysis, is supported by Theakston *et al.*'s (2003) experimental study which showed that children's subject–verb agreement errors were related to subject+non-finite verb sequences they heard in questions (*Will it tam?* vs. *it tams*), and also by Theakston & Lieven's (2008) finding, based on both experimental and naturalistic data, that auxiliary omission in declaratives (*He dancing*) is related to hearing questions where the auxiliary appears preverbally, rather than between the subject and verb (*Is he dancing?*) in the input.

The literature was not clear as to what the ideal period for the error analysis would be. Wexler (2003: 15) suggested that the period during which optional marking of tense and agreement can be observed extends from birth to age 3;0, and all of our children's data for Study 1 was well within that period (see Table 1). Although we aimed to study the children at roughly the same developmental period (MLU $\geqslant 2$), the sampling density and length of recordings used with different children were not identical. This meant that some children's data were analysed over longer developmental periods than others. Nevertheless, for most of the children, me-for-I case errors were spread out across the files, so that although not all of the ten files necessarily had errors in them, errors were not concentrated in only a small number of files either. The fact that a significant correlation was observed between the children's error rates and the me+V input could be taken to indicate that the effect of the input is not restricted to a very narrow age or MLU period, but rather that the input seems to be associated with error rates at least until age 2;9.

Prior to the present study, only a few suggestions have been made as to why different children produce different pronoun error rates. Rispoli (1998) suggested that children's me-error rates were associated with their correct use of *me* in ACC contexts, such that children who made me-for-I errors were more likely to use *me* correctly in ACC contexts than children who made my-for-I errors. However, Rispoli calculated me-error rates as a proportion of 1psg errors rather than as a proportion of overall 1psg use. Thus, his analysis does not differentiate between children who produce more errors overall, but between those who tend to produce *me* errors vs. those who produce *my* errors, irrespective of their overall error rate. The present study provides an explanation for different overall error rates across children, without making any assumptions about the relation between *me* and *my* errors.

There are, however, a number of issues that require some elaboration. Even though the usage-based approach assumes that children's early utterances are largely a reflection of the speech they have been exposed to, it does not claim that children will rigidly stick to only producing exactly what they have heard. Instead, children are assumed to extract more abstract constructions from the language they hear and produce on the basis of the detailed type and token frequencies of items within a construction, resulting in productive language use. This can be illustrated by looking more closely at Ruth's data. She is the only child for whom a greater proportion of the verbs found in I+V structures in her speech are found in me+V sequences in the input, in comparison with the verbs she produced in me+V errors. However, that was not the only difference between her and the other children. Ruth produced 211 me+V error tokens in the initial set of 10 error files (Study 1), and 93 me+V-type errors (532 tokens) in total (Study 2), very many more than any of the other children, and the number of me+V types (and tokens) outnumbered the number of I+V types (and tokens) in her speech. These differences are not simply a result of the differential sampling resulting from different length recordings or different contexts, as the same length of recordings generated from the same play contexts are analysed for the other eleven children from the Manchester corpus who produced far fewer error tokens. Nor are age or MLU likely predictors: Abe's, Brian's and John's files for Study 1 begin at roughly the same age, yet they produce fewer error tokens, while Dominic's, Fraser's, Gail's and Liz's files begin at roughly the same MLU, yet they produce far fewer error tokens than Ruth (see Tables 1 and 2). Ruth therefore appeared to be operating with a highly productive schema for the erroneous use of the 1psg ACC pronoun, even though she knew the correct NOM form. This is likely to have made her errors less closely related to the patterning of me+V in the input, although the input could still provide the origin for her erroneous use if she initially learned specific me+V sequences from the input, and only later developed a more abstract me+V schema.

In fact, many of the children show some degree of productivity in their use of me+V sequences. Eleven of the fourteen children in Study 2 produced at least one me+finite verb form error, and finite verb forms are much less likely than non-finite forms to be produced in me+V sequences in the language addressed to children. However, the rate at which finite verb forms were observed was lower than the use of non-finite forms. Only 3 of the 18 most commonly produced verb types were verbs that the children are unlikely to have heard in a me+V sequence in the input in any form (tags included) (see Table 6 for the most commonly produced me-error verbs in the children's speech).

Thus, although children use the 1psg ACC pronoun creatively, producing errors with verbs such as *got, want* and *gonna*, a large majority of the verbs produced in me+V errors are those that children are likely to have heard produced in me+V sequences in the input. It is worth noting,

TABLE 6. *Verbs produced at least by three of the fourteen children in me + V errors*

Me + V	No. of children producing me + V error (n = 14)
do	10
have	10
go	8
get	7
*got	6
*want	5
*gonna	4
make	4
put	4
read	4
see	4
eat	3
find	3
going	3
like	3
take	3
wash	3
can't	3

Note: The asterisk denotes a verb form which is not likely to appear in the input in me + verb sequences.

therefore, that the usage-based approach makes a similar prediction to the ATOM (Schütze & Wexler, 1996), which claims that verbs marked for person and hence checked for agreement should not co-occur frequently with non-NOM subjects. An input-driven approach would predict a notable absence of these forms due to the infrequent use of these combinations in the input. However, the usage-based approach can explain why these errors ARE found in the speech of some children (Ambridge & Pine, 2006; Pine *et al.*, 2005) by appealing to: (a) input of these forms in combinations of the ACC pronoun and a tag (*She's found me, hasn't she*) or multiverb utterances with finite verb forms such as (*Don't hit me cries the little bear*); and (b) the abstraction of a me + V construction, which can result in children using the ACC pronoun with verbs they have not heard produced with the ACC pronoun. Exactly which constructions will be abstracted by individual children, and their degree of productivity, is expected to be closely related to the type and token frequencies of a range of linguistic forms in the input (Bybee, 1998) and the alternative forms available to a child at a given stage in development (Rispoli, 2005).

In the present studies only unambiguous me + V sentence subject errors were included in the analyses. However, other types of me-errors could potentially also be explained with reference to the language children hear and the development of lexically based schemas. For instance, errors such as *Me too burn* (Abe, file 29 age 2;8.8) and *That's a fishing pole like me got* (Abe, file 14, age 2;6.10) may be explained, not by complex sentence input, but by input and/or children's use of other lexical combinations that contain the 1psg ACC pronoun. Children may produce the above errors because the sequences *me too* and *like me* are frequent in the input and are well known to them. They consequently get activated when children want to express a meaning related to the meaning of the frequently heard/produced sequence. Exploring these avenues as additional explanations for me-errors is an important next step in developing a comprehensive account of ACC-for-NOM errors from a usage-based perspective.

The usage-based view differs from many syntax-orientated views in that single-factor models of acquisition are generally not assumed to explain development. A wide range of factors including the child's existing knowledge of language, the distributional properties of the input, perceptual salience, the child's understanding of pragmatics and semantics, and the child's communicative goals are thought to contribute to the state of the child's linguistic system at any given point in development. For this reason, children's errors can be seen as deriving from a number of different sources. As far as pronominal case errors are concerned, a single mechanism is unlikely to explain all of the observed errors, and clearly a straightforward input-driven account cannot explain why many children produce GEN-for-NOM (i.e. my-for-I) errors, as these combinations are not found in the input. However, a recent study of the acquisition of negation may cast light on this problem.

Cameron-Faulkner, Lieven & Theakston (2007) examined the relation between Brian's use of negators in neg+V(root form) constructions and the use of negation in the input. In the early stages of negator use, Brian produced a large number of ungrammatical utterances based around the frame no+V (*no move, no reach*) that were not obviously modelled on patterns in the input. Next he moved on to produce not+V sequences, but again these were not straightforwardly predicted by patterns of use in the input. Finally, he began to produce the auxn't+V constructions that were found in the input (*don't V, can't V*). However, detailed analysis showed that there was a relation between the use of negators in the input and Brian's use, but it was mediated by the current level of his linguistic knowledge. Thus, at the earliest stages he took the most frequent negator overall in the input (*no*) and combined it productively with a variety of verbs. Next he moved on to the most frequent negator in multiword utterances in the input (*not*) and combined this productively. Finally, he acquired the more complex AUX*n't* forms, but the order in which he did so was based on their relative frequency of use within given pragmatic functions in the input.

With respect to pronominal case errors, it is possible that different children acquire different 1psg pronominal forms (*I/me/my*) at different stages in development, perhaps reflecting the use of proper names rather than 1psg NOM pronouns in the input, a proclivity to talk about ownership and/or different frequencies of use of these forms in the input. Such differences in early acquisition may lead to different patterns of case error reflecting what the child is trying to express through language, and the linguistic forms available to them at that stage (Budwig, 1996; Pine, pc; Rispoli, 2005). Future research should investigate potential usage-based explanations for GEN-for-NOM pronoun errors; our framework suggests that a number of interacting processes working in parallel will provide the basis of an explanation.

GENERAL CONCLUSION

It might be possible to explain children's production of me-errors by the ATOM (Wexler, 1998, 2003; Schütze & Wexler, 1996) or Rispoli's (1998) paradigm building theory and the development of INFL. However, as shown by the present studies, children's me + V errors can also be explained by the input they hear; children seem to make me-for-I errors because their caregivers produce utterances such as *Let me do it* and *Did you see me doing it*. Input was found not only to give a plausible explanation for me-for-I errors, it also provides an explanation for different me-error rates observed in different children. The present paper therefore contributes to the growing body of research which suggests that many features of child language, including some of the errors they make, are linked to children's linguistic experience (Freudenthal *et al.*, 2007; Rowland & Pine, 2000; Theakston & Lieven, 2008; Theakston *et al.*, 2003).

REFERENCES

Ambridge, B. & Pine, J. (2006). Testing the Agreement/Tense Omission Model using an elicited imitation paradigm. *Journal of Child Language* **33**, 879–98.

Bloom, L., Hood, L. & Lightbown, P. (1974). Imitation in language development: If, when and why. *Cognitive Psychology* **6**, 380–420.

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

Budwig, N. (1996). What influences children's patterning of forms and functions in early child language? In D. Slobin, J. Gerhardt, A. Kyratzis & G. Jiansheng (eds), *Social interaction, social context and language: Essays in honour of Susan Ervin-Tripp*, 143–56. Mahwah, NJ: Lawrence Erlbaum.

Budwig, N. & Wiley, A. (1995). What language reveals about children's categories of personhood. In L. Sperry & P. Smiley (eds), *Exploring young children's concepts of self and other through conversation*, 21–32. San Francisco, CA: Jossey-Bass Publishers.

Bybee, J. (1998). The emergent lexicon. *Papers from the Regional Meetings, Chicago Linguistics Society* **2**, 421–35.

Cameron-Faulkner, T., Lieven, E. & Theakston, A. (2007). What part of *no* do children not understand? A usage-based account of multiword negation. *Journal of Child Language* **34**(2), 251–82.

Francis, N. & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. New York: Houghton Mifflin.

Freudenthal, D., Pine, J. M., Aguado-Orea, J. & Gobet, F. (2007). Modelling the developmental patterning of finiteness marking in English, Dutch, German and Spanish using MOSAIC. *Cognitive Science* **31**(2), 311–41.

Hoekstra, T. & Hyams, N. (1998). Aspects of root infinitives. *Lingua* **106**, 81–112.

Huttenlocher, J., Vasilyeva, M., Cymerman, E. & Levine, S. (2001). Language input and child syntax. *Cognitive Psychology* **45**, 337–74.

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

MacWhinney, B. (2000). *The CHILDES project*. Hillsdale, NJ: Erlbaum.

Naigles, L. & Hoff-Ginsberg, E. (1998). Why are some verbs learned before others? *Journal of Child Language* **25**, 95–120.

Pine, J. M., Joseph, K. L. & Conti-Ramsden, G. (2004). Do data from children with specific language impairment support the agreement/tense omission model? *Journal of Speech, Language, and Hearing Research* **47**, 913–23.

Pine, J. M., Rowland, C. F., Lieven, E. V. M. & Theakston, A. L. (2005). Testing the Agreement/Tense Omission Model: Why the data on children's use of non-nominative 3psg subjects counts against the ATOM. *Journal of Child Language* **32**, 269–89.

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

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

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

Rispoli, M. (2005). When children reach beyond their grasp: Why some children make pronoun case errors and others don't. *Journal of Child Language* **32**, 93–116.

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

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

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, Vol. 2*, 670–81. Somerville, MA: Cascadilla Press.

Suppes, P. (1974). The semantics of children's language. *American Psychologist* **29**, 103–114.

Theakston, A. & Lieven, E. V. M. (2008). The influence of discourse context on children's provision of auxiliary BE. *Journal of Child Language* **35**(1), 129–58.

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

Theakston, A. L., Lieven, E. V. M., Pine, J. M. & Rowland, C. F. (2002). Going, going, gone: the acquisition of the verb 'Go'. *Journal of Child Language* **29**, 783–811.

Theakston, A. L., Lieven, E. V. M., Pine, J. M. & Rowland, C. F. (2004). Semantic generality, input frequency and the acquisition of syntax. *Journal of Child Language* **31**, 61–99.

Theakston, A. L., Lieven, E. V. M., Pine, J. M. & Rowland, C. F. (2005). The acquisition of auxiliary syntax: BE and HAVE. *Cognitive Linguistics* **16**, 247–77.

Theakston, A. L., Lieven, E. V. M. & Tomasello, M. (2003). The role of the input in the acquisition of the third person singular verbs in English. *Journal of Speech, Language, and Hearing Research* **46**, 863–83.

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition* **74**, 209–253.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition.* Cambridge, MA: Harvard University Press.
Vainikka, A. (1994). Case in the development of English syntax. *Language Acquisition* **3**, 257–325.
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. (2003). Lenneberg's dream: Learning, normal language development, and specific language impairment. In Y. Levy & J. Schaeffer (eds), *Language competence across populations: Towards a definition of specific language impairment*, 11–61. Mahwah, NJ: Erlbaum Associates.
Wexler, K., Schütze, C. T. & Rice, M. (1998). Subject case in children with SLI and unaffected controls: Evidence for the Agr/Tns Omission Model. *Language Acquisition* **7**, 317–44.

APPENDIX

STUDY 1

Less clear input of me + V (These utterances were included in the additional analysis)

(1) *Me + tag sequences in the* ***input***

Anne: That's a blanket for me, is it?
You showed me, didn't you?
That's what you're telling me, isn't you?
It's no good getting all cross with me, is it?
Becky: It's a crown for me, is it?
That's for me, is it?
She's going to bounce on me, is she?
You did that this morning with me, didn't you?
Brian: Oh it's going to see me, is it?
Gail: That's me, isn't it?
Joel: … but you won't tell me, will you?
You're kidding me, aren't you?
Liz: He's the one that doesn't like me, isn't he?

(2) *Other complex sentence* ***input*** *with me + V sequences*

Ruth: Show me is there any dark left.

(3) *Ungrammatical* ***adult*** *utterances*

Carl: *Me put it in?
Liz: Me don't +//…