

Why Do Children Say “Me do it”?

Rick Kazman
Department of Philosophy and
Software Engineering Institute
Carnegie Mellon University
Pittsburgh, PA 15213
rnk@sei.cmu.edu

Abstract

A common feature of early speech is that children use case marking incorrectly. Several researchers have proposed that the child’s mistakes are limited to the misuse of nominative case, and are corrected once the child acquires verbal morphology. In this paper I will show that this characterization of the problem is incorrect: children misuse *all* case forms, not just nominative case. In addition, I will show that the child’s use of case is related to the acquisition of *nominal* morphology, not verbal.

Case marking can be better understood as a result of the child learning the *productive agreement processes* of his language. This characterization accounts for the acquisition of case and the “waffling” which children exhibit, and does so within a unified theory of lexical and syntactic acquisition.

Introduction

Many language researchers have noted a salient characteristic of early speech:¹ children appear to assign case roles incorrectly. They say things like: *Me go home*, *Her too cold*, and *I knock he over*. Several researchers have attempted to explain why children have difficulty marking case (Radford, 1990), [Aldridge, 1989]. Their explanations center around the connection between the child’s acquisition of verbal morphology and the assignment of nominative case. They claim that the child cannot assign nominative case until he has first gained knowledge of the language’s verbal morphology, paralleling the link between these two phenomena proposed in Government Binding theory [Chomsky, 1981]. I will show that this characterization of child language is incorrect—children do not simply lack the ability to assign nominative case, they lack the ability to assign *all* syntactic case—and I will present a model of language acquisition which relates the child’s acquisition of case to the acquisition of *nominal* morphology, and does so within a unified framework of the acquisition of the lexicon and syntax.

¹The period of acquisition which I will be concerned with is roughly ages 18 to 36 months. During this time, most of the major components of language—word order, argument structure, agreement and case marking—are acquired.

The computational model of the acquisition of morphology and its relationship to syntax which I will be adopting has been discussed elsewhere ([Kazman, 1990], [Kazman, 1991]). In this paper, I will show that by extending the notion of *agreement* in this model to include case marking, the acquisition of case—both the time course of acquisition and the mistakes which children commonly make—is explained. In this way the child’s acquisition of case is accounted for by positing a minimal change to a language acquisition system which already demonstrably accounts for much of the child’s early syntax and inflectional morphology.

In addition, this proposal will account for the waffling stage through which children seem to pass on their way to the acquisition of a particular facet of the adult language. Waffling refers to the tendency in the child’s speech, at this time, to use lexical items and grammatical constructions intermittently, and seemingly randomly.

Consider, for example, the following data taken from the speech of Adam, at age 2;10:²

Adam: Look my telephone.
You don’t like me tow truck?
Mine expresso.
I making noise.
I’m swimming.

In these utterances—all of which were spoken in a period of approximately 20 minutes—Adam sometimes uses the correct possessive *my*, and at other times uses *mine* or *me* as a possessive. Similarly, he sometimes uses the auxiliary verb *be* with progressive forms of verbs, and other times does not (compare *I’m swimming* and *I making noise*).

The acquisition procedure proposed here can account for the waffling in both case marking and auxiliary usage because they are really different manifestations of the same thing—the child’s incomplete understanding of the agreement processes in his language.

²All of the data used in this paper was taken from the CHILDES database [MacWhinney and Snow, 1985]. The data for Adam was taken from [Brown, 1973], the data for Peter was taken from [Bloom, 1973], the data for Shem came from [Clark, 1978], and the data for Nina was taken from [Suppes, 1973]. Some further examples from a number of children will be quoted from [Radford, 1990]. The age notation that I will be using throughout this paper is years;months.

The Acquisition Model

The Induction of the Lexicon

The acquisition model which I am proposing is based upon the assumption that the basic principles of the grammar of a language can be determined by an analysis of the morphology of the target language, and by an examination of the interface between morphology and syntax in this language. The analysis of the lexicon and the development of rules of grammar appear to be significantly different, although related, classes of problems. Accordingly, different learning strategies are applied to them: I am advocating an inductive approach to the acquisition of the lexicon and a deductive approach to the acquisition of syntax.

The lexicon is constructed inductively, through a pseudo-statistical ranking procedure which postulates hypotheses about lexical items based upon their input frequency and other properties such as their phonological salience. The main purpose of the lexical acquisition procedure is to discover the root word forms and the productive affixes of a language. Once the inductive procedure has discovered the form and meaning of the productive affixes, the *structure* of the lexicon can be said to be acquired. Certainly, there is a great deal more to lexical acquisition than this, but once the structure of the lexicon is learned, much of lexical acquisition is reduced to a relatively rote process of adding new entries which conform to previously learned paradigms. That is, once the procedure has learned *bites*, *walks*, *throws*, *wants*, *gives*, etc., learning *brings* or *takes* involves only the application of the rule that the child already knows, i.e. add *-s* to the verb root.

The lexical acquisition procedure proposes hypotheses for word roots and affixes. These hypotheses are ranked, and their ranks are continually updated based upon how well they account for the model's input. In this way, the most productive morphological rules will be highly reinforced by the input, whereas less productive rules will either be pruned or simply listed in the lexicon as irregular forms. Initially, the ranks for all affixes will be low. This corresponds to the period in which words are typically produced by the child uninflected. As the model receives more input, the ranks of the productive affixes increase relatively quickly. During this stage, waffling and overgeneralization³ occur. Overgeneralization occurs because the child learns about the productive affixes more quickly than he learns about irregular affixes (since productive affixes are reinforced by many different words within the language) and, in this model, roots and affixes are combined exhaustively and then unproductive combinations are pruned by the ranking procedure.

Waffling occurs because the child possesses roots and affixes which have a rank which has not yet reached certainty. The highly productive inflectional affixes encode agreement information on their host category—for instance, the verbal

³Overgeneralization refers to the tendency of children to over-apply productive affixes, producing novel forms such as *foots*, *ated*, *goed* and so on. This subject is treated more fully elsewhere ([Marcus *et al.*, 1990], [Kazman, 1990])

-ed affix in English encodes verbal past tense. This piece of information about *-ed* is not innate however. It represents an inductive generalization by the child based upon exposure to a large number of verbs. The rank of an affix is a quantification of the child's belief that that word undergoes a *productive agreement process*. The model of language generation being assumed here assumes that all semantically valid forms of a word activated in parallel (following [Simpson, 1984] and [Carlson and Tannenhaus, 1988]), and the child chooses stochastically among them, based upon their normalized ranks. Assume, for example, that the child has a root, *swim*, and two affixes: the *-ed* affix, with a rank of 0.25, and the *i→a* ablaut, with a rank of 0.50. These ranks are a statistical measure of the child's degree of belief in these affixes (or agreement process) and so the child will, at this stage, produce *swimmed* 1/3 of the time and *swam* 2/3. The result of this ranking procedure, to the external observer, will be a waffling in the child's morphological productions, agreement processes and case marking.

The Deduction of Syntax

The deductive portion of the acquisition model is logically dependent upon the results produced by the inductive portion. The knowledge-base for making deductions about the language is the lexicon. The lexicon, when properly structured, contains much detailed information about the agreement properties of the categories of a language.

Intuitively, this dependence upon the inductive lexical acquisition process can be seen as a way of buffering a deductive model of syntax from the enormous amount of variation and inconsistency in language. Once the inductive analysis has provided a clear model of the lexicon, then the deductive portion of the model may safely make inferences based upon this information.

The deductive portion of the model is structured as a rule-based system, like a production system—a system of if-then rules tied to an associated database of facts, where any rule may be activated when its preconditions are met in the database.⁴ To give a concrete example, we might postulate the following rules for the deductive model of syntax:

IF a category X exhibits regular agreement
THEN hypothesize an *Agr* node dominating XP

IF a category X has an *Agr* node and there is
a function word which subcategorizes for X
THEN analyze this function word in XP's *Agr* node

These rules provide a means by which the target grammar can be gradually refined, based upon an analysis of grammatical relations revealed in the lexicon. In the above example, these rules determine if and when two categories

⁴A general introduction to production systems can be found in [Davis and King, 1984]. Production systems have been used in the past to model human cognitive processes. Newell, for instance, used it to model human performance in a memory scanning task [Newell, 1973]. Furthermore, rule-based systems form the backbone of most expert systems.

| Morpheme | Order of Acquisition | | |
|---------------------|----------------------|-------|-------------|
| | Brown | Babel | De Villiers |
| Present Progressive | 1 | 1 | 2 |
| on | 2 | 2 | 2 |
| in | 2 | 3 | 4 |
| Plural | 4 | 4 | 2 |
| Possessive | 5 | 6 | 6 |
| Articles | 6 | 5 | 5 |
| Past Regular | 7 | 7 | 7 |
| 3rd Person Regular | 8 | 8 | 7 |

Figure 1: Brown vs. *Babel* vs. De Villiers

should be constrained to agree in the syntax. Rules like the above must be included in any serious model of grammar, since cross-linguistically different languages exhibit different agreement facts, and developmentally, children begin their linguistic careers with little or no knowledge of agreement and gradually develop this knowledge over time.

This model has been implemented as a computer program, written in C, called *Babel*. *Babel* has been shown to correctly simulate the acquisition of English and Polish—two historically unrelated languages. To be more precise, this model of lexical and syntactic acquisition makes certain predictions about the rate and order of acquisition of productive affixes and syntactic agreement phenomena. The results of simulating acquisition with this model⁵ are shown in figure 1, compared with the order of acquisition of the same morphemes by children, as documented in Brown's seminal work *A First Language: The Early Stages* [Brown, 1973], and in De Villiers & De Villiers *The Acquisition of English* [de Villiers and de Villiers, 1985]. The Pearson product moment coefficient of correlation⁶ was calculated for the three sets of results. The order of acquisition produced by *Babel* is correlated at a level of 0.99 with Brown's results, and 0.96 with the De Villiers' results.

In addition, the predictions which the model makes for Polish have been shown to conform to the order of acquisition presented in *The Acquisition of Polish* [Smoczyńska, 1985]. Furthermore, these results also hold for Dutch, French and Hebrew, as I have discussed elsewhere ([Kazman, 1990]), although I have not, as yet run computer simulations of the acquisition of these languages.

Case marking

Several researchers have suggested that case marking is missing in early child language, and is acquired at the same time as function words like determiners and auxiliaries (see [Radford, 1990], [Kazman, 1988], [Aldridge, 1989], [Guilfoyle and Noonan, 1988]). Different theories have been

⁵These results were based upon an input sample of over 44,000 words of transcribed spoken adult English taken from the CHILDES database.

⁶The Pearson product moment coefficient of correlation is a scaleless measure of correlation between two variables. Its value varies from -1 (completely unrelated) to +1 (completely related).

proposed to account for this, but they largely rely on case being a property of a phrasal category, Infl, which is a projection of a verb's inflectional information. One problem with these theories is that they force acquisition to rely on an innate knowledge of putative linguistic particulars (such as the notion that nominative case is assigned by a tensed Infl), or on the maturation of some linguistic property, such as Case marking. This amounts to little more than stipulation.

For instance, Radford and Aldridge suggest that children acquire nominative case marking as a result of the creation of Infl. Nominative case is assigned, they claim, once the child has created an Infl node in his syntactic representations. In order to justify this claim, they must not only show that nominative case is not operational in the child's grammar before Infl has been acquired, but that accusative case is operational (since accusative case is assigned by a verb to its direct object—a construction which the children have already acquired). Radford provides evidence suggesting that this is the case. He first provides examples of children using pronouns marked for accusative case:

- Paula: Paula put *them*. Throw *them* in.
Daniel: Mum, Wayne hit *me*. Want kick *him*.
Put *him* bed.
Hayley: Help *me* out. [Radford, 1990, p. 133]

He then provides examples of children using accusative pronouns in positions which would call for nominative ones:

- Angharad: *Me* have biscuit.
Neil: *Him* on there.
Hayley: *Her* do that. [Radford, 1990, pp. 133-4]

On closer inspection, however, it does not appear that this analysis can be maintained. If Radford and Aldridge are correct in assuming that the child can use accusative case (because this is assigned by a verb) before being able to use nominative case (which is assigned by Infl), then they would predict that accusative case should be much more productive than nominative case in the child's spontaneous speech. Radford clarifies this point as follows: "Our central claim here is that children at the lexical stage in their categorial development have not *mastered* the conditions under which nominative case marking takes place in adult English . . . However, the hypothesis that children have not *mastered* nominative case assignment should not be taken as implying that we never find any examples of (what might appear to be) nominative forms in the speech of children at this stage. . . . It is certainly true to say that we find the sporadic occurrence of nominative pronouns in the transcripts of some children who otherwise appear to be at the lexical stage in their categorial development." [Radford, 1990, pp. 134-5] To paraphrase, we might find some examples of nominative case marking in the child's speech, but they should be infrequent and unproductive; simply the result of the child's use of rote-learned phrases such as *Off he goes*.

However, it appears that Radford and Aldridge's assumption can not be substantiated. The frequencies of pronoun usage for Nina and Adam are given in figure 2. These are counts of the number of utterances in which the children use a given pronoun, taken from the samples before Infl has

been acquired by the children (as indicated by metrics such as the child's use of auxiliary verbs and productive verbal affixes).

| Pronoun | Adam | Pronoun | Nina |
|---------|------|---------|------|
| I | 719 | I | 154 |
| me | 494 | you | 92 |
| you | 452 | him | 62 |
| he | 50 | her | 49 |
| him | 46 | me | 41 |
| them | 25 | he | 39 |
| we | 14 | them | 4 |
| her | 10 | they | 4 |
| she | 8 | she | 3 |
| us | 6 | we | 3 |
| they | 3 | us | 0 |

Figure 2: Adam and Nina's Pronoun Usage

If the Radford/Aldridge hypothesis were true, we would expect to see significantly higher usage of accusative pronouns than nominative. However, there does not appear to be any tendency to use accusative pronouns more frequently. In fact, just the opposite effect appears to be the case (although only slightly): the most productive pronoun, *I*, is used significantly more often by both children than its accusative counterpart, *me*.

Radford and Aldridge also claim that children use accusative pronouns in cases where nominative ones are called for, thus indicating that nominative case marking has not been acquired. However, another logical possibility may also be considered: that children simply have *no* notion of case marking at this stage, and so they use their pronouns fairly randomly. If this hypothesis were true, what you would expect to see in the child data is a wide variety of misuses of pronouns, indicating that no case marking of any type has been acquired. So, in addition to the examples above which show children using accusative pronouns in nominative contexts, you would expect to find nominative pronouns in accusative contexts and genitive pronouns in nominative contexts. This is exactly the case, as shown in the following representative examples:

Adam: Let he walk.
See he walk like dat.
Her sing it.
I going knock they down.

Nina: My close it.
My get my car.
Hold him eyes.
Him can't see.

Shem: You heard he climbing in the doggie.
Me no have juice.
Him go meow.

As a consequence of examples such as these, the Radford/Aldridge hypothesis—that children have mastered only

accusative case before developing Infl, and master nominative case only after Infl is acquired—cannot be maintained. Another hypothesis can be suggested given the theoretical framework I have described: case is first acquired by the child as another feature of lexical items. Whereas inflectional affixes are an audible form of agreement indicating some intrinsic feature of a word (such as number or tense) case is an audible form of agreement indicating an *extrinsic* feature of a word: its function in the sentence.

In effect, this is another application of the notion of “bootstrapping”, proposed by [Macnamara, 1982] and [Pinker, 1984]. The child starts out with a simplistic but natural notion of case—that it is another agreement feature of a word—and uses this to “lift himself by the bootstraps” into the more complex adult notion of case. The hypothesis that case is first acquired as a type of agreement not only accounts for the acquisition facts of case, but does so by positing no new mechanisms, and by claiming no reliance on either innate knowledge or maturation.

This proposal carries some empirical consequences with it. It suggests that: 1) we should see the same waffling in the acquisition of case that we see in the acquisition of morphology; 2) case marking should not be productive in the stage before inflectional morphology has been acquired; 3) case marking should be acquired at the same rate as inflectional morphology. These predictions appear to be substantiated. First, we do see waffling in both case marking and inflectional morphology, as pointed out earlier. Second, case marking *is* acquired at the same time as inflectional morphology. The table given in figure 3 shows Peter's usage of plurals, pronouns and genitive 's. The numbers of plurals, genitive 's and pronouns appear to be increasing over time⁷ in lock step. In fact, the Pearson coefficient of correlation for plurals and pronouns is 0.87, for pronouns and 's is 0.83 and for plural and 's is 0.96.

| Age | Plural | Pronouns | 's |
|------|--------|----------|----|
| 1;9 | 6 | 1 | 1 |
| 1;10 | 20 | 1 | 2 |
| 1;11 | 27 | 17 | 2 |
| 2;0 | 28 | 33 | 5 |
| 2;1 | 16 | 50 | 5 |
| 2;2 | 18 | 80 | 11 |
| 2;3 | 93 | 219 | 32 |
| 2;4 | 21 | 137 | 19 |
| 2;5 | 83 | 224 | 26 |
| 2;6 | 36 | 169 | 14 |

Figure 3: Peter's Acquisition of Nominal Elements

What we can infer from this high level of correlation

⁷The numbers do not monotonically increase, in part, because of the different number of utterances collected for Peter each month. For instance, at age 2;3 3300 utterances were recorded for Peter, compared with less than 1600 for 2;4. The other part of the variation is attributable to the stochastic nature of the child's speech.

is that case marking, as evidenced in pronouns and in the genitive marker 's, is acquired along with agreement, as evidenced in plurals. In a theory such as the current one, where case marking and agreement are simply different manifestations of the same process, this correlation is predicted.

Finally, the third prediction—that the *rates* of acquisition of case marking and inflection are related—can also be substantiated. In a language like English, where inflection and case markings are relatively rare and typically phonologically reduced, the acquisition of inflection and case is relatively late. Most children do not achieve anything like adult mastery until about 3 years of age. In Polish, on the other hand, which has abundant, phonologically salient inflection and case marking, the acquisition of inflection and case is precocious. Smoczyńska reports: “as far as the acquisition of specific case endings is concerned, most of them are used correctly from the very moment of emergence of a given category” and “the acquisition of gender is precocious in Polish children. Most of the children have acquired this distinction before the age of 2” [Smoczyńska, 1985, pp. 618, 645].

Conclusions

This proposal explains why children use case marking inconsistently in the early stages. Since case marking is simply another type of agreement, the child, in the stage before he has identified the productive agreement processes of his language, will have several competing hypotheses for the structure of nouns in his language, each of which is weighted by its rank, as determined by the lexical acquisition procedure. Thus, if a pronoun is associated with two possible forms, one correct and one incorrect, each of which have a rank of 0.5 then we would expect that the child will use the correct form about 50% of the time and the incorrect form about 50% of the time—this will appear, to the listener, to be waffling. As the child learns more about his language, the rank of correct forms will near 1, the child will use case and agreement with adult proficiency, and waffling will disappear. The waffling stage will be the same time that the child will be overgeneralizing morphology and use other types of agreement inconsistently. For instance, during this time the child will produce utterances such as “a shoes” (where the number of the determiner does not agree with the number of the noun), “that’s feet” (where the demonstrative and copula don’t agree in number with the noun) and “sit on the lap” (where the determiner and noun don’t agree in definiteness).

So, why do children say *Me do it?* For the same reason that they say *He want some:* because they have not learned nominal and verbal morphology. In particular, they have not learned that a verb agrees with its subject, and they have not learned that a pronoun “agrees” with its position in the sentence. This is also the reason why children produce utterances like *I making noise.* The auxiliary verb *am* is missing from this sentence because the child is not enforcing verbal agreement in his syntax.

The unified lexical and syntactic acquisition procedures presented here not only learn the productive agreement processes of the target language, but do so in a way which is

consistent with the acquisition of the lexicon, as recorded in longitudinal studies: words are first learned unanalyzed, then some words are produced incorrectly and gradually, over time, an adult-like knowledge of the language’s syntax and morphology, including case marking, is learned.

References

- [Aldridge, 1989] Michelle Aldridge. *The Acquisition of Infl.* PhD thesis, University College of North Wales, Bangor, 1989.
- [Bloom, 1973] Lois Bloom. *One Word at a Time.* Mouton, The Hague, 1973.
- [Brown, 1973] Roger Brown. *A First Language: The Early Stages.* Harvard University Press, Cambridge, Mass., 1973.
- [Carlson and Tannenhaus, 1988] Greg N. Carlson and Michael K. Tannenhaus. Thematic roles and language comprehension. In Wendy Wilkins, editor, *Syntax and Semantics, Volume 21: Thematic Relations.* Academic Press, San Diego, 1988.
- [Chomsky, 1981] Noam Chomsky. *Lectures on Government and Binding.* Foris, Dordrecht, Holland, 1981.
- [Clark, 1978] Eve Clark. Discovering what words can do. In D. Farkas, W. M. Jacobsen, and K. W. Todrys, editors, *Papers from the parasession on the lexicon*, pages 34–57. Chicago Linguistic Society, Chicago, 1978.
- [Davis and King, 1984] Randall Davis and Jonathan J. King. The origin of rule based systems in AI. In Bruce G. Buchanan and Edward H. Shortliffe, editors, *Rule Based Expert Systems.* Addison Wesley, Reading, Mass., 1984.
- [de Villiers and de Villiers, 1985] Jill G. de Villiers and Peter A. de Villiers. The Acquisition of English. In Dan I. Slobin, editor, *The Cross-linguistic Study of Language Acquisition, Volume 1: The Data*, pages 27–139. Lawrence Erlbaum, Hillsdale, New Jersey, 1985.
- [Guilfoyle and Noonan, 1988] Eithne Guilfoyle and Maire Noonan. Functional categories and language acquisition. Presented to the Boston University Conference on Language Acquisition, 1988.
- [Kazman, 1988] Rick Kazman. Null subjects and the acquisition of Case and Infl. Presented to the 13th Boston University Conference on Language Development, 1988.
- [Kazman, 1990] Rick Kazman. The Genesis of Functional Categories. Presented to the 15th Boston University Conference on Language Development, 1990.
- [Kazman, 1991] Rick Kazman. On building a model of grammar from information in the lexicon. Working Notes: AAAI Spring Symposium Series, Workshop on Machine Learning of Natural Language and Ontology, 1991.
- [Macnamara, 1982] John Macnamara. *Names for Things, A Study of Human Learning.* Bradford Books/MIT Press, Cambridge, Mass., 1982.

- [MacWhinney and Snow, 1985] Brian MacWhinney and Catherine Snow. The child language data exchange system. *Journal of Computational Linguistics*, 12:271–296, 1985.
- [Marcus *et al.*, 1990] Gary F. Marcus, Michael Ullman, Steven Pinker, Michelle Hollander, T. John Rosen, and Fei Xu. Overregularization. Technical Report Center for Cognitive Science Occasional Paper #41, Massachusetts Institute of Technology, 1990.
- [Newell, 1973] Alan Newell. Production systems: Models of control structures. In W. G. Chase, editor, *Visual Information Processing*, pages 463–526. Academic Press, New York, 1973.
- [Pinker, 1984] Stephen Pinker. *Language Learnability and Language Development*. Harvard University Press, Cambridge, Mass., 1984.
- [Radford, 1990] A. Radford. *Syntactic Theory and the Acquisition of English syntax*. Blackwell, Oxford, UK, 1990.
- [Simpson, 1984] Greg B. Simpson. Lexical ambiguity and its role in models of word recognition. *Psychological Bulletin*, 96(2):316–340, 1984.
- [Smoczyńska, 1985] Magdalena Smoczyńska. The Acquisition of Polish. In Dan I. Slobin, editor, *The Cross-linguistic Study of Language Acquisition, Volume 1: The Data*, pages 595–686. Lawrence Erlbaum, Hillsdale, New Jersey, 1985.
- [Suppes, 1973] Patrick Suppes. The semantics of children's language. *American Psychologist*, pages 103–114, February 1973.