J. Child Lang. **26** (1999), 749–755. Printed in the United Kingdom
© 1999 Cambridge University Press

NOTE

Different rates of pronoun case error: comments on Rispoli (1998)*

CARSON T. SCHÜTZE
UCLA

(*Received* 14 *October* 1998. *Revised* 18 *March* 1999)

ABSTRACT

Rispoli (1998) presents data to motivate a model of pronoun case errors in child English. His data consist of relative rates of occurrence of errors involving particular forms in the study of twenty-seven English children between the ages of 2;6 and 4;0. I show that his claim that overextensions of *he* and *him* are antagonistic cannot be maintained. I argue that his explanation for why *her* subjects are more frequent than other errors is insufficient, and suggest an account in terms of relative input frequencies. Finally, I demonstrate that the fundamental assumption underlying Rispoli's model is untenable, and that his findings are not counterevidence for developmental syntax models such as that of Schütze (1997).

INTRODUCTION

In a recent paper, Rispoli (1998) sets out to further document and explain the common observation that English children's case errors occur at different rates for different pronouns. Specifically, he focuses on the relationship between rates of *he* overextension and *him*-for-*he* errors, and the relationship between rates of *her*-for-*she* versus *him*-for-*he* and *them*-for-*they* substitution. He proposes an account that is intended to describe patterns of relative frequency of these errors found in the study of 27 English children between the ages of 2;6 and 4;0. Rispoli's proposal can be construed as an attempt to explain errors not explained by developmental syntax accounts (such as that of Schütze, 1997) while at the same time explaining differential rates of subject errors that do fall under such accounts. Unfortunately, Rispoli does not successfully treat either phenomenon. He also presents an inaccurate view of Schütze's work on this topic (Schütze & Wexler, 1996; Schütze, 1997). In

[*] This work was supported by a UCLA Academic Senate grant. Address for correspondence: Carson T. Schütze, Department of Linguistics, University of California, Los Angeles, Box 951543, Los Angeles, CA 90095–1543, U.S.A. e-mail: cschutze@ucla.edu. fax: (310) 206–5743.

https://doi.org/10.1017/S0305000999003992 Published online by Cambridge University Press

this note I elucidate these specific problems; see Schütze (1999) for a more general critique of Rispoli's model.

OVEREXTENSIONS OF *he* AND 'ANTAGONISM'

In introducing the topic of *he* overextensions, Rispoli (1998) cites Schütze (1997) among those who have 'focused exclusively on subject case errors' (p. 534). This is incorrect. In that work, Schütze tabulated pronoun uses in all syntactic environments, including object positions. Out of a total of 583 object environments in three large longitudinal transcripts from CHILDES (MacWhinney & Snow, 1990) – Nina (Suppes, 1973; 1;11·24–2;5·28), Peter (Bloom, 1970; 1;11·7–2;5·0), and Sarah (Brown, 1973; 2;8·25–3;1·24) – he found a total of zero overextensions of nominative forms, and only three errors altogether, all by Nina (an error rate of 0·5 %). In contrast, over the same files, there were 2434 subject pronoun environments, with 381 errors (15·7 %) – overextensions of objective and genitive forms. Given the overwhelming asymmetry in the frequency of the two error types, with errors in subject position roughly 30 times more frequent than errors in object position, Schütze contends that they should not be treated as the same phenomenon. In particular, subject errors are argued to result largely from the child's syntactic representations, whereas the much rarer object errors are taken to reflect performance factors.[1] Rispoli presents no arguments against this view.

Rispoli suggests that, descriptively speaking, *him*-for-*he* and *he*-for-*him*/*his* overextensions are 'antagonistic,' by which he means, 'Two error patterns seem to be mutually inhibitory' (p. 537). That is, children who make one kind of error are claimed to be less likely to make the other kind of error. This claim is difficult to assess, because Rispoli does not report the number

[1] I can do no more than sketch the general properties of this model here. (For full details see Schütze (1997); for suggestions about how the model fits into the bigger picture of a theory of morphosyntactic development, see Wexler (1998).) The central claim is that children's syntax differs from that of adults only in highly circumscribed respects, such that a child who has mastered the relevant lexical entries (e.g. those for pronouns) will produce nonadult patterns of case marking that are argued still to conform to Universal Grammar. In particular, the child learning English is limited to the case-assigning properties of the functional elements tense and agreement found in adult English, but may combine these elements in ways not seen in adult main clause declaratives. Specifically, unlike the adult grammar, the child's allows agreement to be omitted from such clauses, but when this happens the rest of the grammar still functions as it would for adults: in the absence of agreement, nominative case cannot be assigned to the subject, so it must be spelled out in the default case of the language, accusative. (This is not the only possibility – see note 5.) In this way, subject case 'errors' result from an adultlike lexicon coupled with a slightly non-adultlike syntax; this system will not yield any OBJECT errors because they too would show up as default case forms, which happen to be indistinguishable from the (correct) accusative forms in English. When this whole system (a valid grammar that happens to be not quite the grammar of English) is hooked up to an (imperfect) performance system, we expect to observe the basic asymmetry of errors just described along with a small amount of 'noise' distributed randomly across it.

https://doi.org/10.1017/S0305000999003992 Published online by Cambridge University Press

or the rate of *he* overextensions produced by the children in this study. Rather, he reports only whether or not a child made one or more such errors versus none at all. The calculation that is meant to provide evidence for antagonism shows only that children who produced at least one *he* overextension had a mean rate of *him* overextension lower than that of the children who produced no *he* overextensions. But if the actual rate of *him* overextensions is relevant, why is the actual rate of *he* overextensions not relevant? If recording sessions had continued for a longer period of time, all of the children might eventually have produced a *he* overextension, at which point Rispoli's measure of antagonism could not be calculated at all.

Fortunately, in earlier work Rispoli (1994) does report quantitative data on *he* overextensions for a different group of 12 children recorded by Betty Hart (Hart, 1991; Hart & Risley, 1995), so we can test the claim of antagonism more precisely. This is done in Table 1, in which children are listed in order

TABLE 1. *Percentage rates of pronoun overextension for 12 children (from Rispoli, 1994), ordered by rate of* he *overextension*

Participant number	Percent *he* overextensions	Percent *him*-for-*he* overextensions
3	0	4
4	1	7
6	2	3
8	2	9
7	3	0
1	4	2
Mean	2	4·17
9	5	0
10	5	11
2	8	3
11	8	38
5	14	15
12	14	20
Mean	9	14·5

of increasing rate of *he* overextension. According to the antagonism hypothesis, the children with higher rates of *he* overextension should have lower rates of *him* overextension. However, in Table 1 just the opposite is true: the mean rate of *him* overextensions is much HIGHER for the children with higher *he* overextension rates than for those with lower rates.[2] This can be

[2] Although it is not evident from Rispoli's (1994) Table 5, whence some of these figures are drawn, it is noted in the text that there were no *his*-for-*he* overextensions, hence all the 'oblique overextensions' were uses of *him*. As a result, the two sets of figures in Table 1 correspond exactly to the two error types analysed by Rispoli (1998).

https://doi.org/10.1017/S0305000999003992 Published online by Cambridge University Press

seen by comparing the mean *him* overextension rate for the first six children, 4·17%, with the mean rate of the other six children, 14·5%. Thus, regardless of what the actual rates of *he* overextension may be for the children discussed by Rispoli (1998), Rispoli's earlier study shows that the pattern predicted by antagonism is not consistently found for English children. This casts serious doubt on a theory such as Rispoli's that predicts antagonism.

THE *her* ERROR RATE AND THE DOUBLE-CELL EFFECT

Like many others who have studied English pronoun case errors, Rispoli (1998) finds that *her* subjects are proportionately much more frequent than other subject errors: in this instance, 49% for *her* vs. 11% for *him* and 12% for *them*. More importantly, *her* subjects are much more frequent than the combined rate of *me* and *my* subject errors. The latter were not studied here, but in Rispoli (1994) we find *me*+*my* subject errors totalling 10% while *her* errors occurred at a rate of 47%. For Nina (Suppes, 1973), Schütze (1997) reports 90% *her* subjects versus 13% *me* and *my* subjects. Rispoli (1998) claims that his model accounts for this difference by virtue of the 'double cell effect,' the fact that *her* occupies both the genitive and the objective cells of the third person singular feminine pronoun paradigm. However, no account is provided of how this prediction actually follows from the model. What might follow is that *her* subjects are more frequent than *him* subjects or *me* subjects or *my* subjects, but the fact that *her* occupies two cells would seem to predict that it should arise as an error with about the same frequency as the forms from the two corresponding cells in the first person singular paradigm. That is, the *me*+*my* error rate and the *her* error rate should be about the same. The above figures show, however, that *her* errors are much more frequent than this. What is required, then, is a mutually reinforcing effect, rather than just an additive one, resulting from the double cell situation. Rispoli does not provide a formal specification for 'retrieval strength,' the central concept in his model, but nothing in his informal description suggests that it would have this property.

Indeed, another factor might provide a better explanation for the size of these differences than the pronoun paradigm, namely, the relative frequencies of the various forms in the input, as proposed in Schütze (1997). There I observed that a corresponding pattern is found in a well-known corpus of approximately one million words of English text (Francis & Kučera, 1982): the ratio of occurrences of *I* to those of *me* and *my* combined is approximately 2 to 1, whereas the ratio of *she* to *her* occurrences is approximately 0·94 to 1. That is, in the first person singular, nominatives greatly outnumber non-nominatives, while in the third person singular feminine, non-nominatives slightly outnumber nominatives. One would obviously like to verify that this is also true of speech to children, but the difference is certainly suggestive. If

https://doi.org/10.1017/S0305000999003992 Published online by Cambridge University Press

a child's rate of overextension of a given form in subject position is proportional to the relative frequency with which she hears that form, then the observed higher rates of *her* overextensions, vs. *me* and *my* combined, would follow. While I am not proposing a specific model here, the suggested frequency effect would fit with the view that the English child effectively has a free choice of pronoun case in subject position (but not in other positions). Choosing whether to say *she* vs. *her* is then influenced by the relative number of occurrences of *she* and *her* that the child has heard.[3] This approach is also consistent with Rispoli's (1998) *him* error data: in the corpus, the ratio of *he* to *him* occurrences is approximately 3·7 to 1. Comparing this to the ratio for *her* given above, we see that *him* has proportionally about one quarter the relative frequency of *her*, and Rispoli found that *him* subjects were proportionally about one quarter as frequent as *her* subjects.[4]

WHERE IS THE NOISE?

Rispoli's account relies crucially on imperfect lexical retrieval to model the existence of rare types of case errors. In this respect, his approach differs fundamentally from that of many other researchers. In particular, it should be noted that in stating his own position, he incorrectly attributes to Schütze & Wexler the view that 'during the early multiword phase of syntactic development, the influence of the grammatical feature case in the retrieval process is weak or intermittent (Schütze & Wexler, 1996; Schütze, 1997)' (p. 547). However, in the former work there is an explicit statement to the contrary: 'Our account does not rely on any kind of random choice of case from the lexicon' (Schütze & Wexler, 1996, p. 678). It is therefore necessary to consider what Schütze & Wexler's approach has to say about infrequent types of errors.

Rispoli cites a single utterance, *my taked it off*, as a counterexample to Schütze's hypothesis (which predicts that genitive subjects should not co-occur with past tense marking).[5] Then he remarks that such counterexamples

[3] This proposal is not part of Schütze's (1997) model, which addresses a different question, but it is compatible with it. Schütze's model embodies claims about the syntactic representations that underlie different types of child utterances, in particular, utterances with nominative versus non-nominative subjects, but it says nothing about the basis for choosing among the different possible representations, only that a certain range of them is available to the child. The suggestion in the text is that lexical frequencies could be an important factor in determining the rate at which different syntactic options are chosen.

[4] This calculation ignores the genitive form *his*, because as Rispoli has noted, it either does not occur as a subject error or else goes undetected.

[5] This prediction resulted from the hypothesis that children's genitive subjects have the same kind of source as genitive subjects in adult English, which appear in gerunds (e.g. *my reading the book*). Since gerunds are not marked for tense or agreement contrasts, children's *my* subjects were predicted to occur only in utterances lacking tense or agreement marking. See Schütze (1997) for fuller discussion; as is noted there, this specific prediction about genitives and tense marking derives from assumptions sup-

https://doi.org/10.1017/S0305000999003992 Published online by Cambridge University Press

'are attributed to ''noise'' in the data (Schütze, 1997: 236), although the source of that noise is never fully addressed' (p. 535). The latter assertion is certainly true at some level, but the same can be said for any other theory of human behaviour – there are no COMPLETE theories of any facet thereof. Schütze (1997, § 3·3) did, however, suggest several possible sources for this noise, including production errors (e.g. slips of the tongue) on the part of the children, and mishearings and transcription errors on the part of the investigators. This proposal is therefore quite different from Rispoli's, because Rispoli claims that the word retrieval component of the child's language system differs SYSTEMATICALLY from that of an adult's, inasmuch as case features have little or no effect on the retrieval of pronoun forms: 'Case features do not regularly activate any output units' (p. 548). In contrast, Schütze & Wexler argue that the child's lexical retrieval system is just like the adult system, but that the words that a child actually produces are sometimes affected in NONSYSTEMATIC ways by production errors, which would in turn be expected to exhibit the same qualities as production errors that adults make in spontaneous speech.

The evidence for the claim that case DOES play a consistently systematic role in word retrieval is the virtual absence of case errors outside subject position. If Rispoli's view were correct, and case features really had a minimal effect on lexical retrieval, the distribution of pronouns in children's speech should be virtually identical in all sentence positions, i.e. subject, object, and genitive. But Rispoli's own data in his previous study show nothing of the kind. For example, according to Rispoli's (1994) tables 4 and 5, for first person singular pronouns, subject position contained 90% nominative forms, while object and genitive positions contained 0% nominatives; for third person singular masculine, subject position contained 95% nominative forms, object and genitive positions only 9% nominatives; and so forth. Clearly, the dominant factor in the choice of pronoun form that a child will produce is the case that is appropriate to its position in the sentence.

CONCLUSIONS

There is much that remains puzzling about English children's pronoun case usage; no one has a worked-out theory that explains all the facts. Nonetheless, I have argued here that Rispoli's particular proposal does not advance our understanding of this phenomenon. On the two specific points it is intended to address, it proves to be empirically inadequate. At the more general level of investigative strategy, Rispoli concludes by saying that in his approach 'no

plemental to those that underlie the account of case errors in general. Thus, even if it should turn out that this one prediction is incorrect, the other results of this model would not be affected.

https://doi.org/10.1017/S0305000999003992 Published online by Cambridge University Press

error is treated as inconvenient noise' (p. 30). But in trying to explain everything at once, one is doomed to explain nothing. From the point of view of any one mental system, certain isolated instances of human behaviour surely ARE noise, that is, the result of some nonsystematic occurrence in one of the other mental systems with which it interacts. Progress in understanding these systems can come only from a foundation based in what the individual components do systematically.

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.

Francis, W. N. & Kučera, H. (1982). *Frequency analysis of English usage : lexicon and grammar*. Boston: Houghton Mifflin.

Hart, B. (1991). Input frequency and children's first words. *First Language* **11**, 289–300.

Hart, B. & Risley, T. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore: Brookes.

MacWhinney, B. & Snow, C. (1990). The Child Language Data Exchange System: an update. *Journal of Child Language* **17**, 457–72.

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

Rispoli, M. (1998). Patterns of pronoun case error. *Journal of Child Language* **25**, 533–84.

Schütze, C. T. (1997). *INFL in child and adult language : agreement, case and licensing*. Ph.D. dissertation, MIT, Cambridge, MA. [Distributed by MIT Working Papers in Linguistics.]

Schütze, C. T. (1999). Explaining patterns of pronoun case error: response to Rispoli. *Journal of Speech, Language, and Hearing Research* **42**.

Schütze, C. T., & Wexler, K. (1996). Subject case licensing and English root infinitives. In A. Stringfellow, D. Cahana-Amitay, E. Hughes & A. Zukowski (eds), Proceedings of the 20th Annual Boston University Conference on Language Development, volume 2 (670–81). Somerville, MA: Cascadilla Press.

Suppes, P. (1973). The semantics of children's language. *American Psychologist* **88**, 103–14.

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

https://doi.org/10.1017/S0305000999003992 Published online by Cambridge University Press