Learnability and the Acquisition of Auxiliaries

by

Karin J. Stromswold

A.B. Psychology and Social Relations, Harvard University
(1983)

Submitted to the Department of Brain and Cognitive Sciences in
Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy
in Cognitive Science

at the

Massachusetts Institute of Technology
September, 1990

Signature of Author _____
Department of Brain and Cognitive Sciences
September 1990

Certified by _____
Steven Pinker
Professor, Department of Brain and Cognitive Sciences

Accepted by _____
Emilio Bizzi
Chairman, Department of Brain and Cognitive Sciences

Abstract

This thesis examines how children acquire the English auxiliary system. The behavior of auxiliaries is extremely intricate and, at times, seemingly arbitrary. Therefore, auxiliaries pose a number of difficult learning problems for the child. Perhaps, the most difficult is how the child distinguishes between auxiliaries and their homophonous main verb counterparts. Before the child has mastered the grammar of English, how does the child distinguish auxiliary *be*, *do*, and *have* from main verb *be*, *do*, and *have*? How does the child master the grammar, if she has yet to distinguish auxiliary *be*, *do*, and *have* from main verb *be*, *do*, and *have*?.

Spontaneous speech transcripts containing over 50,000 uses of auxiliaries or homophonous main verbs and over 40,000 questions were analyzed to determine how children acquire the auxiliary system. In addition, two judgment experiments were conducted in which children judged the grammaticality of sentences with inversion errors and tensing errors. The result of these analyses and experiments suggest that from a very early age children distinguish auxiliaries from main verbs and treat the two kinds of verbs differently. They apply a productive learning strategy with main verbs and a fairly conservative strategy with auxiliary verbs. Because children adopt a fairly conservative strategy for dealing with auxiliaries, they make very few of the many types of auxiliary errors that could potentially be made. What errors they do make seem to involve inversion and *do*-support. From a linguistic standpoint and a learnability standpoint, children's problems with subject-auxiliary inversion and *do*-support are quite understandable.

Acknowledgments

I would like to express my heart-felt thanks to my advisor, Steven Pinker, for all of his advice and support during college, medical school and graduate school. Nine years ago I took an undergraduate course in research methods in language acquisition from Steve Pinker. I have vivid memories of him helping me solder some piece of experimental apparatus. Since then, Steve has helped me wrestle with matters ranging from details of statistical analysis to the overall issues of language, cognition and learning. It was he who convinced me to pursue graduate training in cognitive science. For that alone, I will always be grateful.

I would also express my gratitude to the kindness and advice that my other committee members have shown me. I am grateful to Sue Carey for helping me keep the big picture in mind. As I wrote sections of this, I could hear Sue Carey's voice saying "so? how does all this auxiliary business relate to the rest of cognitive development?" I am grateful to Ken Wexler for directing me to the relevant linguistic literature, and for making me try to see my developmental work in light of recent developments in linguistic theory. I can never thank Jane Grimshaw enough for her calm, clear way of giving advise and criticism. In particular, I thank her for reading and giving me comments on portions of an early version of this thesis as she literally was packing for vacation.

I would also like to thank the following people who have given me insightful comments on portions of this work: Nuel Belnap, Paul Bloom, Kay Bock, Noam Chomsky, Robin Clark, Leon Eisenberg, Jacob Feldman, David Finkelstein, Jerry Kagan, Brian MacWhinney, David Pesetsky, Molly Potter, Tom Roeper, and Dan Slobin. I would also like to give special thanks to Roger Brown who taught a class during my first year of college which convinced me to major in psychology instead of chemistry.

Much of the work represented in this thesis would have been much more onerous if not for the assistance given to me by three MIT undergraduate research assistants, Marie Coppola, William Snyder and Elsa Tsao. I would especially like to thank Marie who regularly went above and beyond the call of UROP-dom during the year she worked with me.

I am grateful to the McArthur Foundation for making it possible for me to pursue a Ph.D. in addition to an M.D.. In addition to the support I received from the MacArthur Foundation, I have also been generously supported by a National Science Foundation Graduate Fellowship and a Sherman Fairchild Pre-doctoral Fellowship.

Lastly, I would like to thank my friends and family. I thank my family for not accepting the limitations that others would have placed on me. I thank my friends for standing by me even as I neglected them,

particularly during this past year when I balanced the demands of medical school and graduate school. A finer group of friends I couldn't ask for. I especially would like to thank Jacob Feldman. If not for the food he brought me, I surely would have starved to death on or about August 12, 1990.

I am grateful to all of the above people. I only wish this thesis were as good as the support they gave me.

Table of Contents

Abstract	2
Acknowledgments	3
Table of Contents	5
List of Tables	6
List of Figures	7
Chapter 1: Introduction	11
Chapter 2: Linguistic and Learnability Issues of Auxiliaries	22
Chapter 3: Auxiliary Errors in Spontaneous Speech	43
Chapter 4: First Uses of Auxiliaries and their Homophonous Main Verb Counterparts	78
Chapter 5: Subject-Auxiliary Inversion	109
Chapter 6: Tense and Auxiliaries: Difficulties with *Do*-support	216
Chapter 7: Conclusion	260
References	267
Tables	277
Figures	310

List of Tables

Table 3.1	Transcripts Used in Error Analyses	277
Table 3.2	Overregularization Errors	278
Table 3.3	Auxiliary and Main Verb Overregularization Baselines	279
Table 3.4	Double-tensed Declarative Utterances	281
Table 3.5	Auxiliary/Inflection Mismatches	282
Table 4.1	Transcripts Used in Age of First Use Analyses	283
Table 4.2	Number of Missing Data per Child	284
Table 4.3	Number of Missing Data per Sentence Modality	285
Table 4.4	Number of Missing Data per Auxiliary and Copula Form	286
Table 5.1	Interaction Among *Wh*-placement, Inversion and Prosody	287
Table 5.2	Comparison of Prior Subject-Auxiliary Inversion Studies	288
Table 5.3	Number of Scoreable Questions per Child	289
Table 5.4	Overall Frequency of Inversion Errors	290
Table 5.5	Children's Inversion Rates for *Yes/No* and *Wh*-Questions	291
Table 5.6	Inversion Rates for Each Child and Each Question Type	292
Table 5.7	Overall Inversion Rates for Each Auxiliary Type	293
Table 5.8	Inversion Rates for Each Auxiliary in Each Question Type	294
Table 5.9	Overall Inversion Rates for Negated and Non-Negated Auxiliaries	295
Table 5.10	Inverted Embedded Questions	296
Table 5.11	Inversion Rates for Matrix and Embedded *Wh*-Question	298
Table 5.12	Practice Questions Used in Judgment Experiments	299
Table 6.1	Transcripts Used in Tensing Analyses	300
Table 6.2	Double-Tensed Questions with More than One Auxiliary	301
Table 6.3	Double-Tensed Questions with One Auxiliary	303
Table 6.4	Double-Tensed Questions with Irregular Past Tense Verbs	306
Table 6.5	"Good" Examples of Double-Tensed Questions	308
Table 6.6	Percent of Auxiliary-less Questions with Tensed Main Verbs	309

List of Figures

Figure 4.1 Original, Unreplaced Data: Interaction between Auxiliary/Main Verb and Be/Do/Have 310
Figure 4.2 Replaced Data: Interaction between Auxiliary/Main Verb and Be/Do/Have 311
Figure 4.3 Interaction between Tense and Be/Do/Have (All Data) 312
Figure 4.4 Interaction between Tense and Be/Do/Have (Auxiliary-Only Data) 313
Figure 4.5 Interaction between Tense and Be/Do/Have (Main Verb-Only Data) 314
Figure 4.6 Interaction between Be/Do/Have/Modal and Sentence Type 315
Figure 4.7 Interaction between Be&Have/Do/Modal and Sentence Type 316
Figure 4.8 Interaction between Modal/NonModal and Sentence Type 317
Figure 4.9 Interaction between Auxiliary/Copula *Be* and Tense 318
Figure 4.10 Interaction between Sentence Type and Tense (Aux & Copula Data) 319
Figure 4.11 Interaction between Sentence Type and Tense (Copula-Only Data) 320
Figure 4.12 Interaction between Sentence Type and Tense (Auxiliary-Only Data) 321
Figure 5.1 Pollock/Chomsky V-Raising Account of Inversion 322
Figure 5.2 Developmental Trend: Adam 323
Figure 5.3 Developmental Trend: Allison 324
Figure 5.4 Developmental Trend: April 325
Figure 5.5 Developmental Trend: Eve 326
Figure 5.6 Developmental Trend: Mark 327
Figure 5.7 Developmental Trend: Naomi 328
Figure 5.8 Developmental Trend: Nathan 329
Figure 5.9 Developmental Trend: Nina 330
Figure 5.10 Developmental Trend: Peter 331
Figure 5.11 Developmental Trend: Ross 332
Figure 5.12 Developmental Trend: Sarah 333
Figure 5.13 Developmental Trend: Shem 334
Figure 5.14 Overall Inversion Rate for Each Type of Question 335
Figure 5.15 Individual Children's Inversion Rates for *Yes/No* Questions 336
Figure 5.16 Overall Developmental Trend: *Yes/No* Questions 337
Figure 5.17 Individual Children's Inversion Rates for *Who* Questions 338
Figure 5.18 Individual Children's Inversion Rates for *How* Questions 339
Figure 5.19 Individual Children's Inversion Rates for *Where* Questions 340

List of Figures (continued)

Figure 5.20 Individual Children's Inversion Rates for *What* Questions 341
Figure 5.21 Individual Children's Inversion Rates for *Which* Questions 342
Figure 5.22 Individual Children's Inversion Rates for *When* Questions 343
Figure 5.23 Individual Children's Inversion Rates for *Why* Questions 344
Figure 5.24 Overall Developmental Trend: *Who* Questions 345
Figure 5.25 Overall Developmental Trend: *How* Questions 346
Figure 5.26 Overall Developmental Trend: *Where* Questions 347
Figure 5.27 Overall Developmental Trend: *What* Questions 348
Figure 5.28 Overall Developmental Trend: *Which* Questions 349
Figure 5.29 Overall Developmental Trend: *When* Questions 350
Figure 5.30 Overall Developmental Trend: *Why* Questions 351
Figure 5.31 Overall Inversion Rates for Forms of *Be* 352
Figure 5.32 Overall Inversion Rates for Forms of Auxiliary *Do* 353
Figure 5.33 Overall Inversion Rates for Forms of Auxiliary *Have* 354
Figure 5.34 Overall Inversion Rates for Modal Auxiliaries 355
Figure 5.35 Overall Inversion Rates for *Be*, *Do*, *Have*, and Modal Auxiliaries 356
Figure 5.36 Overall Inversion Rates for Negated and Non-Negated Auxiliaries 357
Figure 5.37 Inversion Rates for Negated and Non-Negated Auxiliaries: Adam 358
Figure 5.38 Inversion Rates for Negated and Non-Negated Auxiliaries: Allison 359
Figure 5.39 Inversion Rates for Negated and Non-Negated Auxiliaries: April 360
Figure 5.40 Inversion Rates for Negated and Non-Negated Auxiliaries: Eve 361
Figure 5.41 Inversion Rates for Negated and Non-Negated Auxiliaries: Mark 362
Figure 5.42 Inversion Rates for Negated and Non-Negated Auxiliaries: Naomi 363
Figure 5.43 Inversion Rates for Negated and Non-Negated Auxiliaries: Nathan 364
Figure 5.44 Inversion Rates for Negated and Non-Negated Auxiliaries: Nina 365
Figure 5.45 Inversion Rates for Negated and Non-Negated Auxiliaries: Peter 366
Figure 5.46 Inversion Rates for Negated and Non-Negated Auxiliaries: Ross 367

List of Figures (continued)

Figure 5.47 Inversion Rates for Negated and Non-Negated Auxiliaries: Sarah 368
Figure 5.48 Inversion Rates for Negated and Non-Negated Auxiliaries: Shem 369
Figure 5.49 Percent of *Wh*-Questions Judged Grammatical 370
Figure 5.50 Interaction Between MLU and Matrix/Embedded Question 371
Figure 5.51 Interaction Between Inversion and Clause Type 372
Figure 5.52 Interaction Among MLU, Inversion, and Matrix/Embedded Question (All *Wh*-Questions) 373
Figure 5.53 Interaction Among MLU, Inversion, and Matrix/Embedded Question (All *Wh*-Questions except *How come* Questions) 374
Figure 5.54 Interaction Among MLU, Inversion, and Matrix/Embedded Question (Only *How come* Questions) 375
Figure 5.55 Argument vs. Adjunct *Wh*- Questions 376
Figure 5.56 Argument vs. Adjunct *Where* Questions 377
Figure 6.1 Aux-less *Yes/No* Questions with Tensed Main Verbs: All Children 378
Figure 6.2 Aux-less *Yes/No* Questions with Tensed Main Verbs: Adam 379
Figure 6.3 Aux-less *Yes/No* Questions with Tensed Main Verbs: April 380
Figure 6.4 Aux-less *Yes/No* Questions with Tensed Main Verbs: Eve 381
Figure 6.5 Aux-less *Yes/No* Questions with Tensed Main Verbs: Mark 382
Figure 6.6 Aux-less *Yes/No* Questions with Tensed Main Verbs: Naomi 383
Figure 6.7 Aux-less *Yes/No* Questions with Tensed Main Verbs: Nathan 384
Figure 6.8 Aux-less *Yes/No* Questions with Tensed Main Verbs: Nina 385
Figure 6.9 Aux-less *Yes/No* Questions with Tensed Main Verbs: Peter 386
Figure 6.10 Aux-less *Yes/No* Questions with Tensed Main Verbs: Ross 387
Figure 6.11 Aux-less *Yes/No* Questions with Tensed Main Verbs: Sarah 388
Figure 6.12 Aux-less *Yes/No* Questions with Tensed Main Verbs: Shem 389
Figure 6.13 Aux-less *Wh*-Questions with Tensed Main Verbs: All Children 390
Figure 6.14 Aux-less *How* Questions with Tensed Main Verbs: All Children 391
Figure 6.15 Aux-less *What* Questions with Tensed Main Verbs: All Children 392
Figure 6.16 Aux-less *When* Questions with Tensed Main Verbs: All Children 393
Figure 6.17 Aux-less *Where* Questions with Tensed Main Verbs: All Children 394
Figure 6.18 Aux-less *Which* Questions with Tensed Main Verbs: All Children 395

List of Figures (continued)

Figure 6.19 Aux-less *Why* Questions with Tensed Main Verbs: All Children 396
Figure 6.20 Interaction between +/-Auxiliary and +/- Tensed Main Verb 397
Figure 6.21 Grammaticality Judgments & the Number of Tense Markers 398
Figure 6.22 Interaction between MLU and +/-Auxiliary 399
Figure 6.23 Interaction between *Wh*-word and +/- Auxiliary 400
Figure 6.24 Interaction between *Wh*-word and +/- Tensed Main Verb 401

Chapter 1

Introduction

This thesis examines how children acquire auxiliary verbs.[1] The acquisition of the auxiliary system is of interest for a number of reasons. First, auxiliaries are necessary to speak in a syntactically, semantically and pragmatically acceptable way. Auxiliary verbs serve many important functions in English. Syntactically, they are needed to form questions and to negate. On a pragmatic/discourse level, they are used to indicate politeness or to indicate that one is uttering a moral, hypothetical, or tentative statement.

In addition, auxiliaries are of interest because increasingly they are becoming part of the core grammar of prominent linguistic theories. For example, in current Government-Binding theory (Chomsky, 1986a), auxiliaries have been incorporated into the X-bar framework. In various lexicalist theories, descriptions of the auxiliary system are viewed as among the great success stories for non-transformational systems. To the extent that different linguistic systems predict that different types of acquisitional errors will or will not occur, acquisitional data can be a valuable differential test of these systems. For example, Government-Binding (GB) theory distinguishes modals from other auxiliaries whereas Lexical Functional Grammar (LFG) and Generalized Phrase Structure Grammar (GPSG) do not. According to GB, we might expect different patterns of

[1] Unless I specify otherwise, I will use the term 'auxiliary' to refer generally to auxiliary *be*, the perfect or aspectual *have*, auxiliary *do*, and the modal auxiliaries.

acquisition for modals and the other auxiliaries. LFG and GPSG would predict no difference in acquisition.

The acquisition of auxiliaries is particularly interesting for a number of reasons. Children say their first auxiliaries at a young age but complete mastery of all of the auxiliaries takes a considerable amount of time. Hence, by tracing the acquisition of auxiliaries, one can see whether (and how) children's acquisition of language changes over time. Second, auxiliaries are one of the few aspects of language for which it has been demonstrated that acquisition is directly affected by input (Newport, Gleitman, and Gleitman, 1978). The fact that the acquisition of the auxiliary system is so protracted and seems to be influenced by input suggests that the auxiliary system may pose particularly difficult learning problems for the child. As we shall see, the behavior of auxiliaries is very complex and seemingly arbitrary at times. It is no surprise, therefore, that children find it difficult to master.

Chapter 2 outlines the intricate linguistic behavior of the auxiliary verbs. The semantic, syntactic and lexical properties of auxiliaries are discussed and the similarities and differences between auxiliary verbs and main verbs are explored. In particular, the similarities and differences between auxiliary and copula *be* are explored. Chapter 2 also outlines some of the approaches that have been taken to explain the properties and complex behavior of the auxiliaries. The "affix-hopping" approach of Chomsky (1957) which relies on phrase structure rules and an affix-hopping transformation is discussed as are more recent approaches such as those that assume the behavior of auxiliaries are the result of semantic and selection restrictions of auxiliaries (Emonds, 1976, 1985; McCawley, 1988; Pullum and Wilson, 1977).

Having outlined the behavior of the auxiliaries and some of the linguistic approaches that have been taken to explain this behavior, chapter 2 then discusses the problems associated with learning the auxiliary system. In order to learn the complexities of the auxiliary system, children must distinguish auxiliary verbs from main verbs. They must also learn the behavior of each auxiliaries. There are many approaches children could take to this second problem. At one extreme children could apply what they know about the behavior of main verbs to auxiliaries, and vice versa. The problem with this approach is that children will make errors from which recovery is impossible without negative evidence. At the other extreme, children could be completely conservative, making no generalizations or assumptions whatsoever about the behavior of auxiliaries. This approach is untenable because, even if children correctly identified the auxiliaries (no mean feat), they still would have to consider 2^n possible combinations of auxiliaries (where n = number of auxiliaries). In addition to these two extreme possibilities, there are a number of intermediate possibilities in which children make limited assumptions or generalizations. Chapter 2 concludes with a list of seven plausible approaches children could take in the task of acquiring the auxiliary system. Each approach makes certain predictions about the relative order of acquisition of the auxiliaries and what errors children will make.

As is discussed in chapter 2, the inflections, combinations and orders of auxiliaries which are acceptable are quite restricted. Chapter 3 provides an overview of the inflection, combination, and order errors 14 children made in over 55,000 spontaneous uses of auxiliaries. Overall, children made very few errors. There were no inflection errors (i.e., there were no errors that involved either adding an unacceptable inflection or using an

unacceptable form of an acceptable inflection). With the exception of double-tensing errors involving *do*-support, there were very few examples of children using illicit combinations of auxiliaries or illicit combinations of auxiliaries and inflections. Lastly, with the exception of subject-auxiliary inversion errors, children made essentially no order errors.

The *do*-support and inversion errors that children did make were quite restricted in type. Children's *do*-support errors were of two sorts. First, there were examples which contained *did* or *didn't* plus an irregular past tense verb. The structure of these examples suggest that many double-tensing errors are the result of difficulties involving lexical access of irregular past tense verbs. The second type of *do*-support errors were errors in which children unnecessarily invoked *do*-support in utterances which contained copula *be* (e.g., **does it be around it?*and **this doesn't be straight*). The structure of these examples is consistent with the children having treated copula *be* as a main verb. Children's inversion errors were also restricted. Children essentially never inverted a verb which is not invertible and no child ever consistently failed to invert a verb which should invert.

The paucity of most types of auxiliary errors suggests that either the children made no generalizations about auxiliaries or they made only limited generalizations within the *be*, *do*, and *have* subtypes of auxiliaries. The fact that *do*-support errors and inversion errors were common suggests that these phenomena are different from most aspects of the auxiliaries. In chapters 5 and 6, I argue that inversion errors and *do*-support errors both represent difficulties involving verb-raising (or inversion).

Chapter 4 discusses the results of analyses of patterns of acquisition of the various auxiliaries and their homophonous main verbs. The logic behind these analyses is described in the following example: Let's imagine we believe that auxiliary verbs and main verbs are distinct and that we also believe that there is a auxiliary *have* category which is distinct from the auxiliary *do* category. If this was true of children's grammars, we might expect children to generalize what they know about the behavior of auxiliary *do* to auxiliary *does,* and auxiliary *did.* We would predict the children would not generalize what they know about the behavior of main verb *do* to auxiliary *do*, and vice versa. Furthermore, we would predict that children would not generalize what they know about auxiliary *do* to auxiliary *have*, and vice versa. Hence, by looking at the pattern of acquisition of the various auxiliaries--which auxiliaries come in at the same time and which auxiliaries begin to invert or negate at the same time--we may be able to uncover signs of subtypes within which children generalize.

Notice that while finding acquisitional clusters among the auxiliaries or homophonous main verbs suggests that children generalize within the clusters, failing to find clusters does not indicate that children do not generalize. As is discussed in chapter 4, there are many reasons why clustering might not be evident despite generalizations. However, if there is clustering, this suggests that there is generalization.

The goodness of eight different verb groupings were examined using analyses of variance. These analyses revealed that there was a significant difference in the patterns of acquisition for auxiliaries and homophonous main verbs. Despite the extensive similarities between copula and auxiliary *be*, there was a significantly different pattern of acquisition for the two types of *be*. In addition, there was evidence of difference between the

different auxiliaries. Analyses of variance were performed to determine which grouping of auxiliaries was best. Of the models tested, the best was one which grouped auxiliaries into forms of auxiliary *be* and *have*, forms of auxiliary *do*, and the modal auxiliaries. Another notable finding was that grouping the auxiliaries into two subtypes--modal auxiliaries and nonmodal auxiliaries--explained less of the variance than grouping the auxiliaries into 18 single-member groups. This suggests that there were acquisitionally significant differences among auxiliary *be*, *do*, and *have*. In summary, the results of the patterns of acquisition analyses discussed in chapter 4 were consistent with the error analyses discussed in chapter 3.

Chapter 5 examines inversion errors which are remarkable if only because they are not as remarkably rare as most other types of auxiliary errors. In addition to being fairly frequent, inversion errors are important because they have been used to claim that children have qualitatively different grammars than adults. A number of findings are reviewed in chapter 5. First, while inversion errors did occur, and while they are an error which every child made, no child passed through a stage during which he or she habitually failed to invert. Second, no child either habitually failed to invert an invertible auxiliary or inverted a verb which does not invert in English. Thus, the children appear to have attached the +INVERT feature to all and only those verbs which invert. Thus, when the children failed to invert it was because either they failed to recognize a structural context which requires inversion or something about a particular structure made inversion less favorable.

Contrary to claims by Bellugi (1971) that children go through a stage during which they invert *yes/no* questions but not *wh*-questions, I found no difference in inversion for *yes/no* questions and *wh*-questions.

Hence, there was no evidence of children have qualitatively different grammars from adults as Bellugi claims. While there was no difference in inversion rates for *yes/no* and *wh*-questions, there was a significant difference in the inversion rates for the different types of *wh*-questions. Children inverted more frequently in argument questions than adjunct questions. This difference was apparent even for argument and adjunct *where* questions and argument and adjunct *how* questions. Thus, there appears to be some structural difference between argument and adjunct questions that makes inversion more felicitous for argument questions than adjunct questions. If having a clausal scope element in a certain position triggers inversion, one possible explanation for the difference in inversion rates for argument and adjunct questions would be that when the children did not invert adjunct questions either the adjunct *wh*-words inherently did not have clausal scope or they were not in the correct position to trigger inversion.

While the children inverted all and only those verbs which invert in English, they inverted some auxiliaries significantly more often than others. Inversion rates were greatest for the auxiliary forms of *do*, followed by auxiliary *be*, *have*, and the modals. The forms of copula *be* were inverted the least. While this pattern of results may reflect structural differences such as those described in the recent work of Chomsky (1989) and Pollock (1989), this pattern of results is exactly what one might predict for learnability reasons. Children inverted *do* the most because the only reason *do* was present was to invert. If children were not intending to invert, they would have asked auxiliary-less questions (e.g., *who she knows?* for *who does she know?*) Auxiliary *be*, *have*, and the modals were all inverted about 90% of the time because this was the children's true

inversion rate. The children may have inverted copula *be* less because they were somewhat confused by the fact that a main verb could invert.

In addition to inverting some auxiliaries more than others, the children inverted non-negated auxiliaries significantly more often than they inverted negated auxiliaries. This effect was most pronounced for the modals. It is possible that the reason the children inverted negated auxiliaries less often is that they might have found it difficult to simultaneously invoke two operations that have clause level scope. Alternatively, the children might have inverted negated auxiliaries less frequently because of a structural difference between utterances with negated and non-negated auxiliaries. This type of structural difference is better handled by a theory such as GB than by a lexicalist theory. Within the principles-and-parameters approach, this difference might have to do with the relationship between negation and barrier-hood or with the ability to delete a negated trace at Logical Form.

Chapter 5 also discusses errors of inverting in *how come* questions and embedded questions. Given that only one child ever inverted in *how come* questions, and this child did it in non-productively, inversion errors in *how come* questions are probably just speech errors. Inverted embedded questions were more common. Overall, about 10% of the embedded questions than contained auxiliaries were inverted. Furthermore, eight of the children made this error at least once. Four possible explanations for the existence of inverted embedded questions are considered. In addition to reporting the results of transcript analyses, chapter 5 describes an experiment which was designed to elicit children's grammaticality judgments for inverted and non-inverted matrix and embedded questions

The results of this experiment suggest that children's grammars are essentially identical to adult grammars.

Chapter 6 examines children's knowledge of tense and auxiliaries. As discussed in chapter 3, the children occasionally asked questions that were marked twice for tense. A careful examination of these errors suggest that they were quite rare and that those that were present were probably the result of difficulties involving *do*-support and lexical access of irregular past tense verbs. The second kind of error examined in chapter 6 involves questions that lack obligatory auxiliaries. An examination of the auxiliary-less questions in the transcripts indicates that they were not simply the result of a production constraint or of the children not knowing that matrix sentences must be tensed. These auxiliary-less questions appear to have been the result of failure to invoke *do*-support for inversion. Thus, auxiliary-less questions can be thought of as a type of inversion error.

In addition to describing the results of these transcript analyses, chapter 6 describes the result of a judgment experiment in which children judged the grammaticality of questions that had no tense markers, questions that had only a tensed auxiliary, questions that had only a tensed main verb, and questions than had both a tensed auxiliary and a tensed main verb.

Children judged questions that had either one of the two tense markers (e.g., *what does Kermit eat?* and **what Kermit eats?*) to be equally good and significantly better than questions that had neither tense marker (e.g., **what Kermit eat?*) or both tense markers (e.g., **what does Kermit eats?*). This suggests that the double-tensed questions were simply production errors, but that the auxiliary-less questions were the result of children not realizing that tense inflection must be carried by an auxiliary

in matrix questions. The results of the transcript analyses and the judgment experiment suggest that while children know a great deal about tensing, they do not realize the necessity of *do*-support in questions.

Chapter 7 brings together the results of the previous chapters and presents a picture of how children accomplish the task of learning the auxiliary system. The results discussed suggest that children approach the acquisition of the auxiliary verbs in a fairly conservative fashion, while they use a productive strategy in the acquisition of main verbs. They do not produce forms of auxiliaries that they have never heard (i.e., they do not use general rule to create new auxiliaries). However, once they hear a particular auxiliary, they generalize across tense, number,and person within the *Be*, *Do* and *Have* subtypes.

The most salient fact about the acquisition of the auxiliary system is that mastery of the auxiliary system occurs early and is essentially error-free. Given that most of the possible generalizations would lead to unrecoverable errors, unconstrained generalization is not possible. The combinatorics of the auxiliary system are such that an extremely conservative approach involving no generalizations or assumptions about the auxiliaries is untenable.

Thus, children must make whatever generalizations are possible. In large part, Universal Grammar determines what these possible generalizations are. Some of these generalizations concern the universality of auxiliaries and the universal properties of auxiliaries. Others generalizations involve other universal aspects of grammar which interact with the auxiliaries (e.g., the universal aspects of tense). Children quickly master the properties of the auxiliary system which are reflections of Universal Grammar. The one clear example of failure to master an aspect

of the auxiliary system, mainly *do*-support, represents an aspect of the auxiliary system which is particular to English.

Chapter 2

Linguistic and Learnability Issues of Auxiliaries

The auxiliary system has received considerable attention from linguists and language acquisition researchers. One of the reasons for this attention is that the behavior of auxiliaries is very complicated; while certain syntactic and semantic traits are associated with many auxiliaries, some auxiliaries do not exhibit these traits. Linguists have tried to address the question of why the auxiliary system is so complicated or, less ambitiously, how its behavior can be described. Researchers studying language acquisition have attempted to explain how children simultaneously learn the generalizations and restrictions on auxiliary order, combination, inflection, and placement (see Baker, 1981; Pinker, 1984). This thesis focuses on how children acquire the auxiliary system.

Auxiliary verbs include progressive and passive *be* (*am*, *are*, *be*, *been*, *is*, *was*, *were*), perfect *have* (*has*, *had*, *have*), auxiliary *do* (e.g., *do*, *does*, *did*), and the modals (*can*, *could*, *may*, *might*, *must*, *shall*, *should*, *will*, and *would*). Auxiliaries are similar to main verbs in a number of ways. Like main verbs, they lack all of the privileges associated with nouns (e.g., following *the* or *a*), adjectives (e.g., preceding a noun), and adverbs (e.g., taking the inflection *-ly*). In addition, auxiliaries have a number of main verb privileges. For example, most auxiliaries mark subject-verb agreement and tense. In addition, as is the case with main verbs, auxiliary verbs can follow a subject noun phrase.

However, auxiliaries differ from main verbs in a number of ways. In questions, auxiliaries can appear before the subject, whereas main verbs generally cannot[1] (e.g., *is she going?* but not **goes she?*). The negation marker *not* can appear after an auxiliary but not after a main verb (e.g., *she will not go* but not **she goes not*).[2] A main verb can be the sole verb in a non-ellipitical sentence, whereas an auxiliary cannot (e.g., *she likes fish* but not **she must*). Several auxiliaries can appear in the same sentence, whereas usually only one main verb can appear in a sentential clause (e.g., *she must have been expecting guests* but not **she wan: expecting guests*).[3] Lastly, with the exception of passive *be*, main verbs can take the progressive inflection *-ing*,[4] whereas auxiliaries cannot (e.g., **musting*).

The child learning the English auxiliary system could adopt various strategies. At one extreme, she could acquire the auxiliary system productively, generalizing what she knows about the behavior of main verbs to auxiliaries and vice versa. If she does, however, she will make many mistakes. Consider, for example, what would happen if she generalized the behavior of *kick* to *must*. If she did, she might notice that in the sentence *she kicked the ball*, the past tense inflection *-ed* is attached to *kick* and *kick* precedes the noun phrase *the ball*. Having noticed these things about the linguistic behavior of *kick*, if she generalized this behavior

---

1 There are two exceptions. In all modern dialects of English the copula *be* inverts. In addition, in some British dialects possessive *have* can invert. I will return to these points later.

2 This is generally true if we ignore tense-less clauses like *she tried not to fall asleep* and the copula *be* (and in U.K. dialects possessive *have*).

3 Certain verbs like *go* can be followed by other verbs (e.g., *I am going to go get dinner*). Another group of verbs can be followed by infinitival verbs (e.g., *I am trying to finish*).

4 Some types of verbs such as stative verbs may seem somewhat marked with the progressive inflection, but in the proper context, it is possible to add the progressive inflection to all main verbs.

to *must* she would say things like **she musted the ball* to mean something like "she had to have the ball".

At the other extreme, the child could adopt a conservative strategy and refuse to make any generalizations about the behavior of auxiliary from another auxiliary. She would learn the entire auxiliary system by rote, producing only those constructions which had been positively attested to in the input. If she adopted this strategy, she would make no errors. However, her acquisition of the auxiliary system would be extremely slow and protracted since she would never transfer what she knows about one auxiliary to another auxiliary. For example, such a child would is some sense consider each of the 20! (2.43×10^{18}) possible combinations and orders of auxiliaries equally likely to occur. Hence, the child would be waiting for confirmation from the input on the acceptability of over 2.43×10^{18} strings of auxiliaries. In order to determine which auxiliaries invert in which settings, the child might have to consider over 13,000 possibilities.[5] Given that children master most of the details of the auxiliary system by age 5, it is impossible that the child makes no generalizations or assumptions about the behavior of auxiliaries.

In this chapter, I argue that the auxiliary system is so complicated that, if a child were to make most of the logically possible generalizations from one auxiliary to another, she would almost certainly make errors. In chapter 3, I present the results of a search of an large corpus of children's speech which indicate that inversion errors are the only type of auxiliary error that children make with any appreciable frequency. I argue that the paucity of most types of auxiliary errors suggests that children acquire

---

[5] Please see chapter 5 for an explanation of how this figure was determined.

most aspects of the auxiliary system conservatively. Given that it is essentially impossible to acquire the auxiliary system in 5 years without making generalizations or assumptions, it is clear that something must be guiding the children to make the correct generalizations and assumptions. I assert that this something is Universal Grammar.

1. Phenomena of the Auxiliary System

Types of auxiliaries.

As mentioned above, there are five basic kinds of auxiliaries. The first subtype is the progressive *be* (e.g., the *is* in *she is eating*). The second subtype is the passive *be* (e.g., the *was* in *he was killed*). Not all instances of *be* are auxiliaries. For example, the *is* in *she is happy* is a copula, not an auxiliary. In some ways copula *be* acts very much like an auxiliary. It can negate (e.g., *she isn't happy* and *she is not happy*) and it can invert (e.g., *is she happy?*). On the other hand, copula *be* can appear as the sole verb in a non-elliptical sentence. Copula *be* can directly precede a noun phrase or adjectival phrase (e.g., *she is a good friend*). Likewise the action pro-verb *being* in *she was being naive* is not an auxiliary.

The third subtype of auxiliary is the perfect *have* (e.g., the *have* in *they have been eating*). Again not all instances of *have* are auxiliaries. The *have* in the sentence *they have food* is a main verb denoting possession. With the exception of certain British dialects, possessive *have* can neither invert (e.g., **have they food?*) nor negated (e.g., **they haven't food* and **they have not food*). There is also the *have* which takes an infinitival verb complement (e.g., the *have* in *they have to go*). Like perfect *have*,

this *have* cannot be the sole verb in a non-elliptical sentences (e.g., **they have to*). In addition like perfect *have*, it must precede a verb phrase and never a noun phrase or adjectival phrase (e.g., *they have to go* but not **they have to food*). Perfect *have* and the *have* in *have to* differ in that the *have* in *have to* can neither invert (e.g., **have they to eat?*) nor negate (e.g., **they have not to eat* and **they haven't to eat*).

The fourth subtype of auxiliary is the auxiliary *do* (e.g., the *do* in the sentence *I do not eat*). Auxiliary *do* must be distinguished from the proverb *do* found in sentences like *I do this*. Pro-verb *do* can appear as the sole verb in a non-elliptical sentence, and it can precede a noun phrase and take the progressive inflection *-ing* (e.g., *she is doing the dishes*). Like all main verbs, proverb *do* cannot invert (e.g., **does she the dishes*) or negate (e.g., **she does not the dishes*).

The fifth subtype of auxiliaries are the modal auxiliaries. The modals include *can*, *could*, *will*, *would*, *shall*, *should*, *may*, *might*, *must*. Modals *can* and *will* must be distinguished from their homophonic main verb counterparts found in the sentences *he canned peas* and *he willed her to come*.

In addition to these true auxiliaries, there are a group of verbs that exhibit some isolated auxiliary-like traits. I will refer to these verbs as restricted modals. Examples of restricted modals include the verbs *need*, *better*, *ought* and *dare*. So, for example, one can say *you better go*, *you better not go*, *you better be going*, *you better have been going*, etc.. Notice, however, that each restricted modal lacks certain behaviors which most auxiliaries exhibit. So, for example, restricted modal *need* can only appear when negated (i.e., one cannot say **you need go*). Restricted modal *ought* and *dare* can only precede a bare infinitival verb if a *not* is placed

between them and the verb (e.g., *I ought not go* and *I ought to go* but not **I ought go*). Only some of restricted modals can appear before the subject (e.g., **better you go?*, *?need I wait?*, and *dare I leave you in such a state*). Some restricted modals cannot have an attached negation marker (e.g., **I bettern't go*, **you daren't go*, but *you needn't go*, *you oughtn't have said that*).

There are also main verbs like *gonna* (= *going to*), *gotta* (= *got to*), *hafta* (= *have to*), and *wanna* (= *want to*) which some authors refer to as "semi-auxiliaries" because they have similar semantic meanings to the auxiliaries and can precede an infinitival "to verb" complement. These verbs seem to have none of the other properties of auxiliaries. It is possible, however, that children might be confused about the status of these elements.

Restriction on auxiliaries

Using Chomsky's (1957) notation, the restrictions on the auxiliary system can be summarized as follows:

AUX --->(Modal) (*have -en*) (progressive *be -ing*) (passive *be -en)*

Any or all of the auxiliaries are optional, but if they are present, they must occur in the above order. In addition, each auxiliary requires that the verb that follows it be of a certain form. Modals require that the verb that follows be an infinitival form (e.g., *eat*), perfect *have* requires that the verb that follows be a perfect participle (e.g., *eaten*) progressive *be* requires that the verb that follows be a progressive participle (e.g., *eating*),

and passive *be* requires that the main verb be a passive participle (e.g., *eaten*). In addition, the first verbal element must be tensed in a matrix clause. Lastly, questions and negative statements are formed by inverting or negating the first auxiliary. If no auxiliary is present, *do*-support is required.

For ease of exposition, I will divide the restrictions on the behavior of auxiliaries into 4 types--inflectional restrictions, combination restrictions, order restrictions, and inversion restrictions-- and discuss each type of restriction in turn.

Inflectional restrictions. With the exception of modal auxiliaries, all auxiliaries agree with their subjects (subject-verb agreement). So, for example, one says *she is going* not **she are going*. A second inflectional restriction is that the only auxiliary that can take the progressive *-ing* inflection is the passive *be* (for example, the *being* in *he is being traded to the Mets*). One can't say **musting*, **being* (if *be* is a progressive auxiliary), **doing* (if *do* is auxiliary *do*) or **having* (if *have* is perfect *have*). Thus, the *being*, *doing*, and *having*, are perfectly acceptable forms in the lexicon but only when *be*, *do*, and *have* are main verbs. In addition, most linguists argue that a third inflectional restriction is that the modal auxiliaries do no inflect for tense (e.g., **canned*, **mayed*, **shalled*, etc.).[6]

Combination restrictions. In general, most of the possible combinations of auxiliaries are not acceptable. Progressive *be* can only

[6]However see Schacter (1983) and Pullum and Wilson (1977) for arguments that *could* is the irregular past tense form of *can*, etc..

appear with verbs that have progressive inflections. Thus, for example, *she is eating* is acceptable but **she is eat*, **she is eats*, **she is eaten*, where *is* is progressive *be* and not passive *be* are not. Passive *be* can only appear with a past participle verb form. For this reason, sentences such as *the food was eaten* and *the food was being eaten* and *the food has been eaten* are acceptable, whereas the sentences **the food was eating*, **the food was eat*, and **the food was ate* are not. Modals can only precede infinitival verb forms. Thus, for example, *she may eat* and *she may have eaten* are acceptable, while **she may eats*, **she may eating*, **she may ate*, **she may eaten* are not.

Even within subtypes of auxiliaries, some combinations are acceptable whereas others are not. For example, the combination of modal + aspectual *have* is acceptable for all modals except *can*. Sentences like **I can have eaten*, **I can have been eating*, **I can have been kicked*, and **I can have been hungry* are unacceptable even though these same sentences are grammatical with every other modal. Attaching the contracted negation marker *-n't* is only acceptable for some members of auxiliary subtypes. The contracted form of *not* (*n't*) can attach onto all forms of copula and progressive and passive *be* except *am* (i.e., one can say *isn't*, *aren't*, *wasn't*, *weren't*, but not **amn't*). Among the forms of the auxiliary *do*, only *don't* is irregular (i.e., one says *didn't* and *doesn't* but not **do-n't*). Among the 9 modals, 4 do not have regular contracted forms (**willn't*, **shalln't*, **mayn't* and **can-n't*).

Order restrictions. The subtypes of auxiliaries are strictly ordered as follows:

<MODAL> <HAVE> < PROGRESSIVE BE> <PASSIVE BE>

One can best appreciate this ordering in sentences which contain all of the subtypes of auxiliaries. Consider, for example, the sentence, *he must have been being tortured the entire time he was in jail.* Notice that if one switches the order of any of the auxiliaries, the sentence is ungrammatical (e.g., **he must have being been tortured* , etc.). A second order restriction is that if there is a tense or subject-verb agreement inflection, it must appear on the first auxiliary. A third restriction is that, if the content of the entire sentence is to be negated, the negation marker must follow the first auxiliary.[7]

Inversion Restrictions. Lastly, there are restrictions which determine whether the order of an subject and an auxiliary appear in standard position or whether the order is inverted. In questions, the left-most auxiliary and the subject invert in matrix questions but not in embedded questions or in *how come* questions.[8] So, for example, *when can he go?* is a grammatical matrix question, whereas, **when he can go?* is not. In embedded questions and *how come* questions the auxiliary and subject do not invert. Thus, *I wonder if he can go* and *how come he can go?* are grammatical, whereas the inverted **I wonder if can he go* and **how come can he go?* are ungrammatical.

---

[7] One can say, for example, *I could have not been late*, but semantically what seems to be negated isn't the entire utterance, but rather the "been late" portion.

[8] The auxiliary can also occur before the subject in exclamatives (e.g., *can she cook!*) and negative polarity sentences (e.g., *Never have I seen such a mess*). Because these forms are very rare in early speech, they will not be discussed.

Things are not that simple, however. The first complication is that some non-auxiliaries can invert. In all modern dialects of English, the copula inverts. So, for example, one says *is she happy?* even though *is* is not strictly an auxiliary. If the inversion rule were simply that only the first auxiliary can invert, we would expect *do*-support to be necessary in sentences that contain copula *be*. In other words, we would expect **does she be happy* to be the grammatical way to question the sentence *she is happy*. Additionally, in United Kingdom dialects of English, the main verb *have* can invert (e.g., *have you a match?*).[9] A third complication is that many restricted modals cannot invert (e.g., **Better you go?*).

2. Linguistic Accounts of the Above Phenomena

Linguistic theories have differed radically in their treatment of auxiliary verbs. In Chomsky (1957), auxiliaries were treated as members of the category Aux which was distinct from the main verb category Verb. Ross (1969) introduced the idea of treating auxiliaries as main verbs, the so-called "Main Verb Analysis". Since then, hybrid positions have been developed in which auxiliaries are considered members of a sub-category Auxiliary which is part of a super-category Verb. This sub-category-within-a-super-category position has often been instantiated by having all and only the auxiliaries carry a +Aux feature (see Gazdar, Pullum and Sag, 1982, LaPointe, 1980, and Pullum and Wilson, 1977 for a description).

---

[9] It probably isn't the case that U.K.-English possessive *have* can invert because it is a full-fledged auxiliary. Unlike a true auxiliary, it can appear as the sole verb in a sentence (e.g., *they have apples*). Furthermore, sentences in which possessive *have* precede an uncontracted *not* sound stilted at best (e.g., **?I have not apples*).

A number of different approaches have been taken to explain the combination, order and placement restrictions of auxiliaries which were described above. In Chomsky's (1957) "affix-hopping" account, Aux was expanded as:

AUX -->Tense (modal) (*have -en*) (*be -ing*) (*be -en*)

and an affix-hopping transformation adjoined affixes to the verb to the right of the affix, with *do*-support being invoked if no such verb existed. Gazdar, Pullum and Sag (1982) argued that this "affix-hopping" account is at odds with transformational theory both as practiced in 1957 and as practiced in 1982.

Since then, a number of different approaches have been taken to explain the restrictions in the behavior of the auxiliary system. One general approach has been to assume that auxiliary verbs are complement-taking verbs. In such an account restrictions in the behavior of auxiliaries are reflections of restrictions in the verb complements each type of auxiliary can take. These restrictions in turn are the combined result of defective morphology of auxiliary verbs (see Gazdar, Pullum and Sag, 1982; LaPointe 1980; Pullum and Wilson 1977; and Schacter, 1983) combined with semantic restrictions on possible combinations of meaning (see McCawley, 1988). Thus, for semantic reasons, each auxiliary requires that the following verb be of a certain form. Each auxiliary can only occur in the position that it occurs in because each auxiliary lacks certain morphological forms. For example, the modals cannot follow any auxiliary because they do not have infinitival, perfect participle, progressive participle or passive participle forms. The perfect *have* can

follow modals because it has an infinitival form. Perfect *have* cannot follow progressive or passive *be* because it does not have a progressive or passive participle form. The fact that certain combinations are or seem unacceptable (e.g., *she can have eaten*), is due to the semantics of those combinations.

Thus, the behavior of each type of auxiliary verb is restricted and these restrictions have been fleshed out in different ways by different linguistic theories. At one extreme there is the approach taken by GPSG which essentially uses feature co-occurrence rules and phrase-structure rules and a detailed lexicon to list those inflections, order, and combinations of auxiliaries which are acceptable. Other theories have argued that the restrictions on auxiliaries are the result of selection restrictions based on the defective morphological nature of the auxiliaries. Others have argued that semantic restrictions are essential in determining which combinations and orders are acceptable.

3. Learnability and the acquisition of auxiliaries

Baker (1981) and Pinker (1984) had a slightly different agenda than that of other linguistics when they set out to describe the behavior of the auxiliaries. Whereas most of the linguistic theories of auxiliaries have merely been content to adequately describe the behavior of the auxiliaries within a one particular linguistic framework or another, Baker and Pinker wanted a description of the auxiliary system which would be learnable to a child.

In order to acquire the intricacies of the auxiliary system, children must determine what words are auxiliaries. To the extent that auxiliaries

are universal to all languages and to the extent that the features auxiliaries exhibit across languages are similar, children could be born knowing much of what they need to know in order to distinguish auxiliary verbs from main verbs. Steele et al. (1981) identify the following universal properties of AUX :

> AUX = a set of language-particular properties Aux_L such that:
>
> a. Aux_L is a constituent
> b. Aux_L occurs in first, second or final position
> c. For most choices of L, Aux_L may attach to some adjacent element
> d. Aux_L contains a specified, i.e. fixed and small, class of elements
> e. These elements occur in a fixed order
> f. Aux_L must include elements marking tense or modality
> g. Aux_L may include, as well, elements indicating subject marking, subject agreement, question, evidential, emphasis, aspect, object marking, object agreement, and negation.
> h. Insofar as these notational types can be identified across languages, their relative order does not follow from any general principles(s)
>
> (from Steele et al., 1981. p. 155-156)

Children could use properties (b), (f), and (g) to identify which elements are auxiliaries. Knowledge of property (d) might help children by making them cognizant that some categories are closed, i.e., that a productive strategy is not always appropriate. Knowledge of properties (c) and (e) would be useful in that they constrain the possible grammars the child has to consider. While these universal properties by no means identify all and only auxiliaries, they help limit the number of elements which must be considered as potential auxiliaries. As children acquire more knowledge of the syntax of auxiliaries, this will help delimit further which elements are auxiliaries. Conversely, as children acquire more knowledge of the

semantics of particular auxiliaries, this will help in the acquisition of the syntax of the auxiliaries. Thus, semantic and syntactic bootstrapping working in concert is necessary, using the universal aspects of auxiliaries as a starting point.

Let us assume that some combination of semantic and syntactic bootstrapping allows the child to correctly identify the auxiliary verbs. How do children acquire knowledge of the behavior of auxiliaries? Baker and Pinker argued that the behavior of auxiliaries is best explained by assuming that auxiliaries have defective paradigms which restrict the possible combinations, inflections, and orders of auxiliaries. They made the critical assumption that children would not use knowledge about the behavior of main verbs or other auxiliaries to fill out the paradigms for an auxiliaries. Most importantly, children would not not generalize what they know about main verb *be*, *do*, or *have* to auxiliary *be*, *do*, or *have*, despite the identical appearance of the two forms. Both Baker and Pinker assumed that each auxiliary had different paradigms for declarative and non-declarative contexts. Given this strategy of no generalization, children would have to hear each of the acceptable 10,240 auxiliary combinations in order to use it.[10]

There are generalizations that can safely be made. If children never create a form they haven't heard, then children could safely generalize within the subtypes of *be*, *do*, *have*, and the modals. If children did this, then they would only have to learn 64 possible combinations.[11] The critical assumption that this potential learning strategy shares with Baker

---

[10] The 10,240 combinations is calculated as follows: 10 modal choices x 4 *have* choices x 8 progressive *be* choices x 8 passive *be* choices x 4 modalities.

[11] The number of combinations is calculated as follows: 2^4 x 4 modalities.

and Pinker's is that children never invent forms. Therefore, children won't invent forms like **mayen* or**maying*. If we accept the arguments of Zwicky and Pullum (1983) that the contracted negation marker, *-n't*, is an inflectional affix and not a clitic, this assumption will suffice to prevent children from using forms like **mayn't*, **willn't*, etc.. Given that there is a safe generalizations that can be made, it behooves the children to make them.

4. Possible Generalizations

The intricate nature of the phenomena described in section 1 and the lack of consensus on how to account for this behavior (see section 2) suggest that the behavior of auxiliaries is extremely complicated. The auxiliary system seems to be filled with linguistic landmines waiting for the child if she makes the wrong generalizations. Before examining the errors children make, let's consider a few of the generalizations a child learning English might make and the errors that would result from these generalizations.

The child could generalize what she knows about main verbs to auxiliary verbs, and vice versa. This generalization might be tempting for three reasons. First, while auxiliaries tend to convey information about tense, mood and aspect, there is considerable overlap between the meanings conveyed by main verbs and auxiliary verbs. This is evident by a number of main verbs which are nearly synonymous with auxiliary verbs. Consider, for example, the following main verb/auxiliary pairs: *able/can*,

seems/is, *exists/is*, *have to/must*, *going to/will*, etc.. Second, as discussed above, main verbs and auxiliary verbs share certain syntactic privileges. For example, both types of verbs follow subject noun phrases and neither type exhibits any of the syntactic behavior associated with nouns, adjectives, etc. Third, as discussed above, there is a remarkable degree of lexical overlap between main verbs and auxiliary verbs. Of the 20 auxiliary forms, only *could*, *may*, *might*, *must*, *shall*, *should*, and *would* do not have main verb counterparts.

Based on these semantic, syntactic and lexical similarities between main verbs and auxiliary verbs, the child could conclude that there was no difference between main verbs and auxiliary verbs. If she did this, she would make many errors. One type of error which is uniquely predicted by this model, is that she would either invert main verbs (if she generalized the behavior of auxiliary verbs) or not invert auxiliary verbs (if she generalized the behavior of main verbs).

Alternatively, she might notice that all auxiliaries exhibit subject-auxiliary inversion and from this conclude that all auxiliaries belong to a single subtype. If she did, she would not make the inversion errors above, However, we would still expect her to either fail to add the subject-verb agreement marker to nonmodals or add it to modals. She might notice from sentences like *she may have gone* that two auxiliaries can occur in a row and not realizing the different nature *may* and *have*, she might conclude that auxiliaries can be stacked. If she did, she might erroneously say **she may can go* to mean "she may be able to go".

Based on the fact that *be*, *have*, and *do* exhibit subject-verb agreement whereas the modals do not, the child might instead decide that *be*, *have*, and *do* all belong to a single subtype of auxiliaries distinct from

the modal subtype. If she did, she would not make the subject-verb agreement errors outlined above. However, she would still use illicit combinations and orderings of the *be*, *have*, and *do* auxiliaries. This would happen because sentences like *he has been waiting* she would lead her to conclude that a sentence can contain any two members of the *be-have-do* subtype. Hence she would produce ungrammatical sentences like **I has had going*, or **I was been going.*.

Based on the properties of auxiliary and main verbs, children could group the verbs in a number of different ways. Below, I outline some of the more plausible groupings (from most general) to most specific and the errors that would result from children adopting these groupings.[12]

1. All-Verb Category. All auxiliary and main verbs could be viewed as members of a single category, Verb. If children formed a single verb category, they would treat auxiliaries and main verbs alike. If they generalized the behavior of main verbs, they would apply incorrect inflections to auxiliaries (e.g., **she cans eat*), allow auxiliaries to appear by themselves in non-elliptical sentences(e.g., **she can*), overregularize the inflections of auxiliaries and main verbs (e.g., **she haves eaten* and **she haves dinner*) and fail to invert any auxiliaries (e.g., **what she can eat?*). If they generalized the behavior of auxiliaries, we would expect an inverse set of errors.

NonModal Verb and Modal Auxiliary Categories. Auxiliary *be*, *do*, and *have* and main verbs could constitute a single category distinct

---

[12] For the sake of simplicity, I am ignoring the passive *be* in the following categorization schemas. Obviously, even ignoring passive *be*, I cannot hope to discuss all of the logically possible groupings since there are 2^{20} possible groupings. See chapter 4 for a discussion of this matter.

from the Modal category. If children adopted this categorization, we would expect them to make all of the errors outlined for the All-Verb Category except they would make essentially no errors involving modal auxiliaries.

Main Verb and Aux Categories. All auxiliary verbs could be members of an Aux Category distinct from the Main Verb Category. If children adopted this categorization scheme, the only type of main verb error they would make would be overregularizing inflections. Children would still make many auxiliaries errors. They would allow illicit combinations of auxiliaries (e.g., **I might can go*). They would either apply subject-verb agreement to modals (e.g., **she mays eat*) or fail to apply subject-verb agreement to the nonmodal auxiliaries (e.g., **she be going*, **they be going*, **she do dishes*, **she have eaten*). They would never apply the progressive inflection to any auxiliary.

Main Verb, NonModal Aux, and Modal Aux Categories: Main verbs could be members of a a distinct main verb category, auxiliary *be*, *do*, and *have* could all be members of a a distinct nonmodal auxiliary category, and the modal auxiliaries could be members of a modal auxiliary category. With this categorization scheme, children would have the same pattern of main verb errors that they had for the Main Verb/Aux Verb categorization just described. Behavior on auxiliaries would also be similar to that described for the Main Verb/Aux Verb categorization However, with this Main Verb/NonModal Aux/Modal Aux categorization, children would not add the subject-verb agreement inflection to modals. Furthermore, while they might use several forms of *be*, *do*, or *have* in a sentence (e.g., **she has had going)*, they would never have more than one modal per sentence.

Main Verb Category, *Be&Have* Category,*Do* Category and Modal Category. This is the categorization scheme which is closest to that adopted by current GB theory. Notice that in the GB categorization scheme, auxiliary *be* and copula *be* are grouped together. Modals are distinct from *be*, *have*, or *do* because they are generated directly in I and never pass through AGR. This accounts for why modals do not exhibit subject-verb agreement or tense inflection. *Do* is distinct from *be* and *have* because it is only invoked when raising necessitates an element to support inflection. Main verbs differ from auxiliary verbs in having theta-roles to transmit. (Please see chapter 5 for a more complete discussion.)

Main Verb, *Be* , *Have* ,*Do* and Modal Categories. In this categorization scheme main verbs,*be*, *have*, *do*, and the modals each form a separate category. Under this categorization scheme, we would expect children to generalize from one member of the Main Verb, *Be*, *Do* , *Have* , Modal Categories to other members of the same category. So, for example, once a child began to invert *do*, he would immediately begin to invert *does* and *did*. He would not, however, start inverting all of the modals, perfect *have*, etc. There would still be errors such as negation errors which result from generalizing within a category (e.g., errors such as saying **amn't* because *aren't*, *isn't*, *wasn't*, and *weren't* are acceptable).

Main Verb, *Be* , *Have* ,*Do* and Individual Modal Categories. Because the modals are in some ways the least uniformly behaving of the types of auxiliaries, children might form general categories for all the forms of main verbs, *be*, *do*, and *have*, but form separate categories for each individual modal. Hence, depending on what we consider a modal auxiliary, this categorization scheme would involve a minimum of 9 categories to include those verbs which everyone agrees are

modal auxiliaries (i.e., *can*, *could*, *may*, *might*, *must*, *shall*, *should*, *will*, *would*) plus four categories to cover main verbs and auxiliary *be*, *do*, and *have*.

Main Verb Category and 20 Auxiliary Categories. Under this categorization scheme, main verbs would form a single category but each auxiliary would be the sole member of its own category. If children adopted this categorization scheme, children would not generalize anything they know about the behavior of one auxiliary to any other auxiliary. For each auxiliary, children would have to learn whether that auxiliary inverts and negates. For each auxiliary, children would have to learn which combinations and order restrictions apply to that auxiliary. This is basically the approach taken by GPSG.

The status of copula *be*. For the sake of simplicity, the above list of plausible categorizations of auxiliary and main verbs, ignores the question of what children might do with copula *be*. As is clear from the discussion thus far, copula *be* is an unusual case. It is the only main verb--if it is, in fact, a main verb--that can do every thing which auxiliaries can do, mainly invert and negate. On the other hand, copula *be* can do a number of things which auxiliaries can't do, mainly appear as the sole verb in a non-elliptical sentence and directly precede a noun phrase. Within current GB theory copula *be* is considered equivalent to auxiliary *be*. In a Lexical-Functionalist analysis of the auxiliary system, Falk (1984, p. 484) began by assuming that "aside from the '*be to*' construction, there is only one verb *be* in English--according to all available syntactic and semantic information. Syntactically, all uses of *be* undergo Subject-Auxiliary Inversion, are followed by negatives, etc. Semantically, Jackendoff 1976,

1978, 1983, has shown that a primitive BE exists which corresponds to all uses of the verb *be*."

Given the complex mixed behavior of copula *be*, I would assert that the matter is not nearly so settled and children might well choose to classify copula *be* as either a main verb or an auxiliary verb. If children treat copula *be* and auxiliary *be* the same(as evident from the errors they make and the pattern(s) of acquisition), this would suggest that a linguistic analysis that treats copula and auxiliary *be* the same might be preferred. If children treat copula *be* differently than auxiliary *be* (for example, if they invert or negate it less), this would support a linguistic analysis which distinguishes between the two.

The next chapter reviews the results of an extensive search for auxiliary errors in the spontaneous speech of children. Chapter 4 analyzes the patterns of acquisition of auxiliaries and the main verbs which are homophonous to the auxiliaries. Together, these analyses plus the in-depth analyses presented in chapters 5 and 6 suggest that children use the universal semantic and syntactic properties of auxiliaries to identify which elements are auxiliaries. Furthermore, these results are consistent with the view that children are essentially conservative in their acquisition of the auxiliaries, though they do make limited generalizations within the *be*, *do*, *have*, and modal subtypes of auxiliaries.

Chapter 3

Auxiliary Errors in Spontaneous Speech

As Maratsos and Kuczaj (Maratsos and Kuczaj, 1978; Maratsos,1984) and Pinker (1984) point out, most of the logically possible types of auxiliary errors discussed in chapter 2 have rarely or never been attested to in the literature. According to Maratsos (1984), the following sorts of errors have never been recorded: *he does be going*, *he did can go*, *he ising go*, *did can he go?*, *he had be singing*, *he does will like it*, *he ising going*, *he did could go*, *did would he like it*, and *he does is going*. According to Pinker (1984), essentially no examples of inflectional errors involving auxiliaries have been reported. For example, there have been no reports of incorrectly inflected modals (e.g., *canned*, *canning*, *cans*, *cannen*), perfect *have* (e.g., *have had gone*, *have haven/haved gone*, *am having gone*) or progressive *be* + *ing* (e.g., *am being going*). In addition, there have been no reports of the following illicit combinations of auxiliaries: Modal + Modal, *have* + Modal, *be* + Modal, *be* + perfect *have*, *be* + progressive *be*, and *do* + auxiliary (Pinker, 1984). Pinker (1984) concluded:

> Auxiliary inflection errors and auxiliary combination errors, then, are virtually absent from children's spontaneous speech, and (aside from questions with a surplus auxiliary in inverted position when the simple inverted form would be ungrammatical), auxiliary combination errors can be elicited in experimental situations only rarely and in special cases. (Pinker, 1984, p. 272)

With the results of Maratsos and Kuczaj and Pinker in mind, I began my largely futile search for auxiliary errors in the spontaneous speech transcripts of 14 children. In the last chapter, I outlined a number of different kinds of auxiliary errors that children might make. Depending on what the organization of children's auxiliary system is, different types of auxiliary errors should tend to occur. Basically, I looked for any errors involving auxiliaries, although I concentrated my search on those errors which are uniquely predicted to occur by one of the possible verb categorization schemes described in chapter 2, and those errors which Maratsos and Kuczaj and Pinker failed to find.

Error Analysis

1. Subjects

I examined all of the transcripts of the 14 children shown in Table 3.1. The children were between 0;11 and 2;10 when their speech was first recorded and between 0;11 and 7;10 when the last sample of their speech was recorded. In addition to the children being different ages when their speech was recorded and transcribed, the amount of speech recorded for each child varied from a low of 1,500 lines for Allison to a high of over 38,000 lines for Ross. There were a total of about 200,000 lines of transcribed speech said by the children. Many of these lines contained more than one utterance.

2. Auxiliary Corpus

I used the UNIX utility "fgrep" to search the computerized transcripts of 14 children which are available through the CHILDES system (MacWhinney and Snow, 1985) I used "fgrep" to cull all of the children's lines which contained the following letter strings: *am*, *are*, *be*, *better*, *can*, *could*, *do*, *did*, *gonna*, *gotta*, *had*, *has*, *have*, *is*, *may*, *might*, *must*, *need*, *never*, *no*, *not*, *n't*, *shall*, *should*, *was*, *were*, *will*, *won't*, and *would*. This yielded a corpus of children's speech which contained approximately 55,700 uses of auxiliaries[1] and 15,000 uses of negation markers.

The 14 children varied widely in the number of times they used auxiliaries. May used no auxiliaries and, therefore, the sample of errors collected is from the remaining 13 children. Allison and June used the next fewest auxiliaries (each used about 200 auxiliaries) and Adam and Ross used the most auxiliaries (each used over 11,500 auxiliaries). Table 3.1 lists the approximate number of auxiliaries used by each child.

Please Insert Table 3.1

---

[1]Note that occurrences of copula *be*, and main verb *have* and *do* are included in these tallies.

3. Procedure

I went through the corpus of utterances which contained auxiliaries or negation markers looking for mistakes involving auxiliaries or negation. In particular I looked for the following types of errors:

1. Illicit inflectional endings such as modals which had subject-verb agreement markers and auxiliaries with the progressive ending *-ing*.[2]

2. Examples of overregularization of inflectional forms for auxiliaries or their homophonous main verb counterparts (e.g., Eve17: what did you doed?)

3. Illicit combinations of auxiliaries (e.g., Adam28: Why was I did break it?)[3]

4. Illicit combinations of auxiliaries and inflections (e.g., Adam36: How does this broke?)[4]

5. Examples of unacceptable auxiliary + *-n't* combinations (i.e., **amn't*, **can-nt'*, **do-n't*, **mayn't*, **shalln't*, and **will'n't*).

6. Examples of scrambled auxiliary order (e.g., **I have could go*)

7. Examples of misplaced inflections (e.g., **she do goes to the store.*)

8. Examples of misplaced negational elements (e.g., Adam41 (3;11): did you take no cooking?)

---

[2] I did not record examples of subject-verb agreement errors (e.g, **me is happy?*) because it was not clear whether the error was due to difficulty with auxiliaries or difficulty with agreement.

[3] These errors are covered in more depth in Chapter 6.

[4] Examples of double-tensing errors are covered in more depth in Chapter 6.

In addition to recording examples of the above types of auxiliary errors, for all questions that contained both a subject and an auxiliary, I determined whether the subject appeared before or after the auxiliary.[5]

Utterances that had contracted, stuttered, or unclear auxiliaries and utterances which were obvious routines or imitations were not included in the error analyses or inversion analyses. In addition, I did not count examples which were ungrammatical because an inflection or an auxiliary was missing (e.g.,**she playing* or **she is play*) because such examples could well have been the result of limitations of the production system rather than limitations of the grammar.

4. Results

1. Inflectional Errors

Illicit Inflectional Endings

Children never applied the subject-verb agreement marker to modals (i.e., there were no examples like **he cans go*). Likewise, children never applied the progressive *-ing* inflection to a modal (i.e., there were no examples like **musting*). Furthermore, while the children said *doing*, *being*, and *having* 264 times, the following were the only anomalous examples[6]:

---

[5] Inversion errors are covered in more depth in Chapter 5.

[6] Again, aside from subject-verb agreement errors or errors of omitting an auxiliary

Eve13 (2;0) while me [/] while me being sitting on it.
Naomi21 (1;11) almost fall down [#] almost [#] going being careful [#] going being
Nathan01 (2;6) doing making dinner
Nathan01(2;6) having [#] making breakfast
Ross2125(2;7) Mommy doing sitting down and be dead.
Ross46 (3;11) Cause I think being mistakes is all right.
Shem19 (2;8) prob'ly I would doing dat game, anudder way.

Of the above examples, Naomi's example is one in which the progressive *-ing* was improperly attached to copula *be*. It does not, therefore, count as an example of attaching the progressive inflection to an auxiliary. The same is true for Ross's, *cause I think being mistakes is all right*. After he said this, Ross's father asked Ross whether he means to say "making mistakes" and Ross replied that he did. Even if that wasn't what Ross had intended to say, this would be an example of copula *being* and not auxiliary *being*. From context, it appears that Shem meant to say *I would like doing dat*, but he did not say *like*. Whether this is the case or not, it is clear from context that *doing* was a main verb and not an auxiliary.

Ross said *Mommy doing sitting down* in response to the question *what was Mommy doing?* Nathan said *doing making dinner* in response to the question *what are we doing?* Hence, these may be the equivalent of gerundive expressions in the adult grammar like *heavy lifting* in *did the heavy lifting*. In Nathan's other example, *having [#] making breakfast*, it appears from the context of the utterance that Nathan began to say *having breakfast* but switched and said *making breakfast* instead.

The one remaining example (Eve's *while me being sitting on it*) appears to be the only clear example of an auxiliary with a progressive

inflection, despite the fact that there were 264 examples of the homophonous main verbs appearing with the progressive inflection.

Overregularized auxiliaries

All of the auxiliaries and their homophonous main verb counterparts have irregular past tense and third person forms. The forms for both auxiliary and copula *be* are *am, is, are, was, and were* not *be, bees*, and *beed*. The forms of both auxiliary and main verb *do* are *do*, /duz/, and *did* not *do, doos*, and *doed*. The forms of both auxiliary and main *have* are *have, has* and *had* not *have, haves*, and *haved*. Thus, for example, one says *she was eating* and *she was happy* not **she be-d eating* and **she be-d happy*. *Can, may, shall,* and *will* may also have irregular past tense forms, if one accepts the somewhat controversial notion that *could* is the past tense of *can. might* is the past tense of *may, should* is the past tense of *shall*, and *would* is the past tense of *will* and (see Gazdar, Pullum and Sag, 1982; and Schacter, 1983, for arguments for treating *could, might, should*, and *would* as past tense modals).

It is well-known that children sometimes say "eated" for "ate" and "sayes" for /sez/. In fact, over-regularization of irregular past tense verbs is one of the most common errors made by children (see Pinker & Prince, 1988, for a review). Pinker, Prince, Hollander, Kim, Marcus, Prasada, and Ullman (1990) analyzed the transcripts for the same fourteen children I examined and found that they overregularized irregular past tense verbs about 8% of the time. Therefore, unless the children were completely conservative in their acquisition of auxiliaries, they should have occasionally produced overregularized past tense and subject-verb

agreement forms of the auxiliaries. I found no such examples. Children made no errors with modals (i.e., they never said "canned" for "could", etc.). I did find 32 overregularized forms of *do*, *have*, and *be*, but in all 32 cases the overregularized verb was a main verb and not an auxiliary verb. In Table 3.2, I list all of the examples of overregularization of main verbs that are homophonous with auxiliaries.

Please Insert Table 3.2

Notice that 11 of the 13 children made this kind of error at least once. Notice also that main verb *be*, *do*, and *have* are each overregularized a number of times. The breakdown of the overregularization errors is:

1 example of *ams* for *am*
1 example of *ares* for *are*
2 examples of *beed* for *was*
6 examples of *bees* for *is*
9 examples of *doed* for *did*
9 examples of *doos* for *does*
1 example of *hadded* for *had*
3 examples of *haves* for *has*

After I had finished the overregularization analysis for these 13 children, I wanted to know whether the general pattern of overregularizing *be*, *do*, and *have* only when they are main verbs was true of other children. I decided to check whether Stan Kuczaj's son, Abe, perhaps the most

famous and frequent over-regularizer[7], overregularized only main verb *do*, *be*, and *have*. Below are all of the examples from Abe's transcripts:

Abe50 (2;11) why you doed that? why you doed that? why you doed that>
Abe55 (2;11) yeah and the bear doed it read it you do it over again do that
Abe55 (2;11) yeah and [#] and they doed something [#] uh [#] they started.
Abe57 (3;0) I beed careful they didn't spilled.
Abe65 (3;1) yeah they doed something and they scared me.
Abe88 (3;4) hawks hadded bones and black fur?

Notice that Abe overregularized *be*, *do*, and *have* eight times, but none of the eight overregularizations were for auxiliary *be*, *do*, or *have*.

It's possible that the reason that overregularization errors occurred for main verbs but not for auxiliaries is that the children might not have acquired the correct auxiliaries forms of *be*, *do*, and *have*. If, at the age that the children were over-regularizing main verb *be*, *do*, and *have*, they hadn't yet begun to use auxiliary *be*, *do*, and *have*, this would explain why that there were no examples of overregularized auxiliary *be*, *do*, and *have*. However, this does not seem to be the explanation for the lack of auxiliary overregularizations.

As is shown in Table 3.3, with three exceptions, all of the children were correctly using the irregular forms of auxiliary *be*, *do*, or *have* at the age when they overregularized main verb *be*, *do*, or *have*. All three exceptions involved *have*. The first exception was Adam who had not yet used auxiliary *has* by his twenty-ninth transcript when he said *haves* for

[7] Pinker *et al.* (1990) reports that whereas the other 14 children analyzed in this study overregularized 8% or less of irregular past tense verbs, Abe overregularized 22% of irregular past tense verbs.

main verb *has*. The second exception was Peter who had not used auxiliary *has* by his tenth transcript when he said *haves* for main verb *has*. The third exception was Shem who had not yet used auxiliary *had* by his thirty-sixth transcript in which he said *hadded* for *had*.

Please Insert Table 3.3

Weak overregularization examples. In addition to the obvious cases of overregularization in which an illicit form is created, I examined cases in which children used *be* instead of correctly using *am*, *are* or *is*. Notice that this is form of overregularization in that *be* is not acceptable as a tensed auxiliary or copula. For example, *she could be sick* and *she could be dying* are fine, but **she be sick* and **she be dying* are not acceptable in standard dialects in English.[8], I found 134 cases in which children used *be* where they should have used *is*, *are*, or *am*. In 128 of 134 examples, *be* was a copula and not an auxiliary. The six exceptions were:

Adam42 (4;0) and den dey be looking for it.
Eve20 (2;3) we be coming down [#] Gloria
Eve20 (2;3) Gloria [#] you be sitting on the chair?
Sarah120 (4;8) Easter be coming too.
Shem15 (2;6) be going anudder one.
Shem27 (2;10) yeah, so- $- because he's be 'cratching for.

[8] It is acceptable in some vernacular Black English. However, none of the children in the transcript study were exposed to this dialect.

Summary of Inflectional Errors

In summary, the children whose transcripts I examined made essentially no inflectional errors with auxiliaries despite the fact that they inflected and overregularized main verbs which were homophonous to the auxiliaries. There was only one example in which the progressive inflection was attached to auxiliary *be*, *do*, or *have*, despite the fact that there were over 250 examples of main verb *being*, *doing*, and *having* in the transcripts. There were 32 examples of overregularized main verb *be*, *do*, or *have* (40 if we count Abe's examples), but there were no examples of overregularized auxiliary *be*, *do*, or *have*. Lastly, there were 134 examples in which children used *be* for *am, is, are*, *was*, or *were*. Only six of these examples involved auxiliary *be*.

2. Combination Errors

Illicit Combinations of Auxiliaries

As was discussed in chapter 2, most combinations of auxiliaries are not acceptable. Despite this fact, there were only a handful of utterances that contained illicit combinations of auxiliaries. There were only about 30 cases which could be examples of illicit combinations of auxiliaries in declaratives. All but two of these examples involved *do*-support. There were also only 32 cases of what appears to be illicit combinations of auxiliaries in questions. All of the illicit auxiliary combinations in questions involved having an auxiliary both before and after the subject.

Thirteen of the 32 examples actually involved the copula *be* rather than a "true" auxiliary. All but 12 of the 32 examples involved *do*-support.

Ungrammatical *do*-support

Because so many of the unacceptable combinations of auxiliaries seemed to involve *do*-support, I examined these errors more carefully. I divided the errors into examples of unnecessary *do*-support in non-negated declaratives, non-negated questions, negated declaratives, and negated questions.

Ungrammatical *do*-support in non-negated declaratives. Non-negated English declarative sentences which have *do*-support are ungrammatical unless the *do* is stressed. Hence, the sentence *I **do** like apples* is acceptable because *do* is stressed, whereas, **I do like apples* is not acceptable because the *do* is not stressed. Judging from context, there were fewer than 20 example in which the children incorrectly provided a *do* in an unemphatic declarative. In other words, there were very few examples like Sarah's (3;0): *I do taste dem* which contained an unemphatic auxiliary *do* and a main verb.

In addition to these examples there were two examples in which the children appeared to use an auxiliary *do* with another auxiliary in a non-negated declarative sentence. The two examples were *he was do Batman* (Ross28 (2;10)) and *I shall do pull it* (Nina30 (2;6)). From context, Nina's example appears to be a true example, but Ross's does not. It appears from context that Ross meant to say *he was doing Batman*, but failed to produce the progressive inflection *-ing* on the pro-verb *do*.

Ungrammatical *do*-support in non-negated questions.

English matrix questions that do not contain an auxiliary or copula require that auxiliary *do* be provided for the purposes of inversion. Questions that already contain an auxiliary or a copula do not require auxiliary *do*. If *do* is supplied in such questions, the result is ungrammatical. There were the following 11 examples of unnecessary *do*-support in non-negated questions:

Adam12 (2;9) yeah. <do? [/] <do> [/] do be stirs?
Adam28 (3;4) Why was I did break it?
Adam29 (3;5) does it be around it?
Adam34 (3;7) did was it be a comb?
Adam39 (3;10) I don't see nobody. did there be some. did it be
...
Adam39 (3;10) and does dese be in dere?
Mark68 (4;1) could you do untie the knot?
Nina50 (3;2) why did you did scare me?
Sarah121 (4;8) yeah. did I <be> [/] [#] be a good girl?
Shem02 (2;3) where does these pictures be taken?

If we examine the examples of unnecessary *do*-support shown above, we see that all but three examples involves copula *be*. The one exception involving auxiliary *be* (Adam's *why was I did break it?*), is a somewhat odd case of *do*-support (if that is, in fact, what it is) because *did* is in clause internal position, and *was* is in pre-subject position. Given the morphology of the verb, the redundant element seems to be the auxiliary *was* and not *do*. The second example which did not involve copula *be* was Mark's *could you do untie the knot.* I would argue that the *do* in this example is actually the pro-verb *do* and not auxiliary *do*. From context, it appears that Mark wanted someone to 'do' the knot, but he wanted to make it clear that what he wanted done was untying, not tying. The third

example of unnecessary *do*-support that involved an element other than a copula was Nina's *why did you did scare me?* From context it seems that Nina might have been trying to use *do*-support both for emphasis and for questioning.

Let's consider the remaining eight examples. In all eight examples in which the *do* appeared in pre-subject position, *do* carried the tense and the clause internal *be* was in infinitival form. Thus, these examples are consistent with the idea that the children treated copula *be* as if it were a standard non-inverting main verb which required *do*-support.

Ungrammatical *do*-support in negated declaratives. There was one example in which negated *do* appeared with the restricted modal *better* (Sarah101 (4;4): I think I don't better cut it). In addition to this example, I found six cases in which children included a negated *do* with an auxiliary or a copula. All six cases involved copula *be* rather than an auxiliary. In all six cases, negated *do* appeared first followed by infinitival copula *be*. Hence, the form of these examples is consistent with the possibility that the children did not realize that copula *be* is the only non-auxiliary which can be negated. If they did not realize that copula *be* negates, this would explain why they provided *do*-support. Thus, these six examples of unnecessary *do*-support in negated declarative sentences are very similar to the cases of unnecessary *do*-support in non-negated copula questions which were discussed in the previous section. The six examples of *do*-support with copula *be* were:

Adam10 (2;7) you don't be quiet
Adam33 (3;6) this doesn't be straight.
Peter18 (2;9) then don't be any room now.

Ross35 (3;2) that how we don't be go [= That isn't how it goes]
Ross52 (4;5) Because if you don't be a disney person then that means no fun.
Shem06 (2;4) didn't be nice t$ froggy.

There were also the following four examples in which Ross appeared to have used a negated form of *do* instead of simply using *not*:

Ross42 (3;7) Let's don't talk
Ross42 (3;7) Let's don't talk any more
Ross43 (3;7) Let's don't talk anymore about this
Ross48 (4;2) Let's don't teach anybody about this

These are exceptional cases because they are examples of negation in untensed clauses. In untensed clauses, negation is normally accomplished simply by using *not*. It's likely that Ross had not mastered this subtle point of grammar. It's also possible that he thought *don't* was a simple negation marker rather than a negated auxiliary as Maratsos and Kuczaj (1978) suggest.

Lastly, there were the following cases in which Nina and Sarah appeared to use a negated form of *do* instead of a negated form of the copula.

Nina12 (2;2) that doesn't in there (repeated three times in a row)
Nina14 (2;2) he don't heavy
Sarah109 (4;5) but I don't (sup)posed to be like that

It's conceivable that Nina and Sarah were trying to avoid using a negated copula because they were confused about why a non-auxiliary could negate. It's also possible that these are just random errors.

Ungrammatical *do*-support in negated questions. There were eight examples of unnecessary *do*-support in negated questions. One example was a question asked by Mark which was transcribed as unclear. The seven clear examples contained a non-negated *do* in sentence-initial position and a negated *do* of the same tense in clause-internal position. The examples are listed below:

Adam27 (3;4)	did I didn't mean to?
Adam31 (3;5)	did I didn't mean to?
Adam32 (3;6)	do she don't need that one?
Adam33 (3;6)	did I didn't mean to do that?
Mark71 (4;3)	oh [#] is it doesn't xxx everything xxx there?
Nina34 (2;10)	does it doesn't move?
Nina34 (2;10)	do they don't eat people up>
Ross36 (3;3)	why does Superman doesn't wear underroos on his bottom?

Maratsos and Kuczaj (1978) suggest that examples such as those shown above are the result of the children thinking negated *do* is a simple negation marker and not as auxiliary as . I have two problems with this explanation. First, why didn't the children make the same mistake for other auxiliaries? Why don't they think *can't* or *won't* is simply a negation marker? Second, if the children merely thought negated *do* was simply a negation marker and not an auxiliary, then why in all seven cases did they use the same tense for the clause-internal negated *do* as the clause-initial non-negated *do*?

Other illicit auxiliary combinations. There were on y 12 auxiliary combination errors in questions which did not involve *do*-support in one way or another. Four of the examples contained copula *be*., three

had progressive *be*, and one had passive *be*. The remaining four examples contained modals. The 12 examples were:

Adam28 (3;4)	is it was a snake?
Adam29 (3;5)	can dey should go down dat Massachusetts Avenue?
Mark78 (4;8)	would I may be excused?
Nathan26 (3;5)	why is he will stop?
Nathan26 (3;5)	why is [#] is the stove is on too high?
Nina33 (2;10)	is this is a dog?
Ross32 (3;0)	is my old baby blanket is clean?
Sarah59 (3;5)	it was get burned on my sore finger?
Sarah89 (4;1)	this [#] are you can make an elephant? huh?
Shem3 (2;3)	why is tony is closing $ door
Shem 10(2;5)	is the clock is working?
Shem40 (3;2)	wha' wh- why is wh- why is des clothes are blowing by?

There were only two examples of illicit combinations of auxiliaries in declarative sentences which did not involve *do*-support. At the age of 4;2, Mark said *so they could be hear it* and *they might can do it* (both examples from transcript 70).

***Can have* errors.** Given that the combination Modal+perfect *have* is acceptable for all modals except *can*, one might expect that errors like **I can have gone* would be particularly common. This was not the case. I found 40 utterances which contained a Modal and perfect *have* but none of them were *can have*. This is quite striking because the children used *can* more frequently than any other modal.

Illicit Combinations of Auxiliaries and Inflections

In addition to most combinations of auxiliaries being ungrammatical, most combinations of auxiliaries and main verb inflections are unacceptable. Therefore, I searched for ungrammatical combinations of auxiliaries and inflections. As mentioned above, I only counted examples which were ungrammatical because of the presence of an extraneous inflection or a non-contracted auxiliary.

Double-tensed utterances. I found that between 0.4% and 0.04% of children's questions were double-tensed. Over half of the double-tensed questions had irregular verbs, and many were negated. Because I discuss double-tensed questions at length in chapter 6, in this chapter, I will only discuss the double-tensed declaratives here.

There were 31 examples of double-tensed declarative sentences which contained a non-contracted auxiliary. These examples are listed in Table 3.4. Twenty-three of the 31 examples involved negated *do*. Eighteen of the 23 examples with negated *do* also had irregular past tense verbs (e.g., Sarah52 (3;3): I didn't did it). Six of the 31 examples were utterances by Peter and Ross which contained emphatic *do* (e.g., Peter14 (2;6): I locked it. I did locked it). All but one of these emphatic *do* cases had regular main verbs. The remaining two examples were also said by Peter and Ross. Peter (2;9) said *is this goes on the [///]* and Ross (6;0) said *is uses in school*. Peter's example is probably a question, and Ross's example is a mystery.

Please Insert Table 3.4

There were approximately 200,000 lines of the children's speech recorded in the transcripts. The children asked approximately 40,000 questions. Many if not most of the lines in the transcript contained more than one utterance. However, if we conservatively assume that each line of the transcripts contained a single utterance, there were at least 160,000 declarative utterances. In other words, there were 160,000 opportunities for double-tensed declarative sentences. This means that even if we count all of the negated examples and all of the examples with irregular past tense verb forms (see chapter 6 for reasons why we might want to discount such examples), double-tensed declarative sentences comprised less than 0.02% (31/160,000) of declarative utterances. Thus, the ratio of double-tensed questions to double-tensed declaratives is approximately 20:1. If we consider only those utterances which had auxiliaries, then double tensed utterances (both double-tensed questions and declaratives), constitituted about .2% of all utterances.

Auxiliary/inflection mismatches. There were 37 utterances in which the inflection of the main verb was inappropriate for the auxiliary present. Thus, utterances with auxiliary/inflection mismatches constituted less than .02% (37/200,000) of all utterances and .07% (37/55,700) of all utterances which contained auxiliaries . These utterances are listed in Table 3.5. In 13 utterances, children used auxiliary *do* with a progressive main verb. There were three examples in which a modal appeared with a progressive verb and four examples in which a modal appeared with a past

tense verb. There were two examples in which copula *be* was used instead of auxiliary *do* (one negative and one question). There was two examples of a modal, one example of auxiliary *be*, and one example with *do* occurring with the third person singular present tense inflection. There were three examples in which *be* occurred with a past tense verb. Shem interchanged *be* for *has* three times. Finally, one of Shem's utterances (*I gonna cook spaghetti and be to poi it*), I was unable to classify

---
Please Insert Table 3.5

Auxiliary + *n't*. The negation marker can attach onto most but not all auxiliaries; **amn't*, **can-n't*, **do-n't*, **mayn't*, **shalln't*, **willn't* are not acceptable. I found only one case of *amn't* and one case of *willn't* among the almost 9,400 uses of the contracted negation marker *-n't*. Mark was 4 years when old he said "I'm doing the puzzle well [#] amn't I" (from boys67 transcript) At age 2;9, Ross had the following argument with his father which culminated in him saying **willn't*:

Father Don't do that.now we are going to get dinner.
Ross No # we won't
Father No # we won't
Ross No # we don't. No # we don't. No # we won't.
Father Yes # we will
Ross No # we won't. No we willn't.
Father Marky # do you want to eat? (from Ross26 transcript)

In addition to these two examples, I found one example in which a child used a negated contracted form of the possessive *had* along with *do* (Sarah81 (3;10): Wince did hadn't any).

Summary of Combination Errors

Most of the logically possible combinations of auxiliaries and inflections are unacceptable, and each of the children's 200,000 lines of transcripts presented an opportunity to make such errors. Despite these two facts, I found very few examples of illicit combinations of auxiliaries and inflections. Most of the combination errors that did occur involved *do*-support. All but 14 of the examples of illicit combinations of auxiliaries involved at least one auxiliary *do*. In addition, there were only between 20 and 25 examples of illicit auxiliary/inflection combinations that did not involve *do*.

There were about 20 examples in which children used *do* in declarative sentences without emphasizing the *do*. There were about 30 examples of double-tensed declaratives and 75 examples of double-tensed questions (see chapter 6). The majority of the cases of double-tensing involved *do*-support, usually in conjunction with irregular past tense verbs. Thus, Maratsos and Kuczaj's observation that most double-tensed utterances have irregular past tense verbs is correct, although it is not true that there are no errors with regular past tense verbs as Maratsos (1984) claimed. The other errors involved providing unnecessary *do*-support in questions and declaratives that contained auxiliaries or copulas. I would like to discuss briefly the dozen and a half cases in which children used *do*-support with an auxiliary or a copula. There are several things that are

striking about these examples. Aside from one case with the restricted modal *better*, all of the examples involved copula *be*. There are no cases with auxiliary *be* or with any of the other auxiliaries. Second, all of the examples would have been perfectly grammatical if copula *be* were a standard main verb that didn't require *do*-support. Thus, I would argue that the children seemed to know the co-occurrence restrictions of auxiliaries, they just had trouble with *do*-support. Furthermore, if we examine the cases in which *do*-support was a problem, we see that many such cases involved negated *do*, verbs with irregular past tense forms, or copula *be*.

The last error I would like to discuss never occurred, namely the *can have* error. The children never made this error despite the fact that *can* was the most frequent modal used by the children and every other modal can appear with perfect *have*. One can say *she could have eaten fish yesterday*, *she will have eaten fish ten times by tomorrow*, even *she can't have even fish every day this week*, but one cannot say **she can have eaten fish*. Whether this restriction is semantic or syntactic, the fact that children obeyed it indicates that they did not blindly generalize the behavior of one modal auxiliary to another modal auxiliary.

3. Word Order Errors Other than Inversion Errors

Auxiliary Order. Of the approximately 6,700 utterances that had more than one auxiliary,[9] children never scrambled the relative order of the auxiliaries.

Misplaced Inflection. In English, the first auxiliary carries the tense inflection. I analyzed the children's utterances to determine whether they obeyed this constraint. In all but about 10 of the children's utterances which had non-contracted auxiliaries and tense or subject-verb agreement inflection, the inflectional marker appeared on the first auxiliary. The examples of misplaced inflection were:

Adam12 (2;9)	yeah <do> / <do> be stirs. (also a *do*-support example)
Mark83 (5;1)	I don't got [!] to do nutting@a
Ross28 (2;10)	I don't broke it in two bits.
Ross29 (2;11)	Do you baked turn it to jelly?
Sarah68 (3;7)	it doesn't broke. I won't ruin [#] her hair.
Sarah117 (4;7)	do you got a bow in your head?
Shem18 (2;8)	What do you did?
Shem22a (2;9)	What do you turned on?
Shem33 (3;0)	What do you wanted to say to dat man dere?

---

9 This number was calculated in the following way. First, I used the "fgrep" utility to count the number of lines that contained any one of the 20 auxiliaries. I then used fgrep to determine the number of lines which contained each of the auxiliaries (e.g., the number of lines that contained *am*, the number of lines that contained *are*, etc). I added up the total line count for each of the 20 auxiliaries and compared that with the overall line count for auxiliaries. For example, if the overall line count for a child was 150 and adding up the separate line counts for each of the auxiliaries gave a total of 175, this indicates that there were 25 lines which contained two different auxiliaries. This method of calculating the number of utterances with more than one auxiliary would inflate the number if a transcript had lines with two or more utterance that contained different auxiliaries. It would under-report the number of utterances with more than one auxiliary if children had utterances that spanned more than one transcript line or if children used the same auxiliary twice in one utterance (e.g., **can she can go?*).

Shem 37 (3;1) haft- hey- hey what do you said?

Of these 10 examples, Adam's has already been discussed. All of the remaining nine examples contained an untensed *do* and a tensed main verb. In six cases the tensed main verb had an irregular past tense form. It is possible that children did not intend to use a past tense verb, but they did not realize that these irregular verbs were in the past tense (e.g., Ross might have meant to say "I don't break it in two bits"). The remaining three examples are all are questions, and may represent some failure of affix-hopping.

Negation

Negation without an auxiliary Among the 15,000 negated utterances, there were many cases in which children used a *no!* or a *no* in a sentence which did not contain an auxiliary. In all but a handful of these examples, children incorrectly placed a negation marker before the main verb. In other words, they very rarely said things like **he taking not all of de walls down* (Adam, 3;5). This suggests that while the children frequently omitted an auxiliary, they seemed to know that only auxiliaries and copulas can precede a negation marker in simple single-clause sentences. They did this correctly despite the existence of grammatical sentences like *she tried not to scream.*

Negation in declaratives. The vast majority of the time in declarative sentences which contained an auxiliary or a copula, children

correctly placed the negation marker after the first auxiliary or copula. The following were the only counterexamples:

Nina56 (3;4)	Mommy [#] screw this off. so it could be not a branch
Ross41 (3;6)	No [#] you're supposed not to be the Wizard of Oz.
Ross43 (3;7)	You're supposed to not do [...] [= you're not supposed to do]
Ross49 (4;3)	That's the only one that might get not good.
Sarah99 (4;3)	but not you can. maybe I can [...]
Shem2 (2;3)	want see not should turn.
Shem2 2;3)	wanna see not should turning round

From context, Nina wanted to specifically negate the object *branch* (i.e., she wanted the branch to be not-a-branch). From context when Ross said "might get not good" it is apparent that what he wanted to say was "might get bad". Hence, he placed the *not* before *good* in order to specifically negate *good*. Sarah seemed to place *not* before the subject of the sentence in order to negate only the subject. Thus, in all three cases, the children attempted something quite reasonable that happens not to be acceptable in English.[10]

The remaining two examples by Ross involve a contracted auxiliary. It is possible that *you're* was a memorized routine and not *you* + *'re*.[11] If Ross did not realize *you're* contained an auxiliary, he would not have known to place *not* after *you're*. In quick succession, Shem uttered two sentences in which he placed *not* before an auxiliary. At the age of 2;3, he said *want see not should turn. wanna see not should turning round* (from

---

10 These examples could be explained via the children having some sort of Scope/Relevance Principle such as that described by Givon (1985) or Slobin's (1985) Relevance Operating Principle.

11 See Brown (1973) and Chapter 6 for a discussion of the status of contracted auxiliaries.

transcript 2). As one can see, both examples had many other grammatical failings, aside from the placement of *not*. My impression judging from the context of these utterances was that Shem was very much over his head linguistically and was trying valiantly (but unsuccessfully) to deal with many aspects of grammar that he had not yet mastered.

Negation in questions. Most of the time when children asked negated questions, they used a contracted negated auxiliary (e.g., *why can't I go?*). Occasionally, however, children used an uncontracted auxiliary *not* in a question. There was only one example in which a child incorrectly raised an auxiliary and *not* (Eve20 (2;3): why do not you want?). There were also five examples in which the children raised the auxiliary but left the *not* in clause-internal position. They were:

Naomi81 (3;2)	why are we not getting anymore?
Namoi92 (4;8)	will you not be in here?
Nina49 (3;1)	why am I not going go on the airplane with my daddy?
Ross36 (3;3)	why does he not have underroos on his bottom?

Notice that none of the examples are ungrammatical. The scope of negation in these examples is restricted to the verb phrase instead of encompassing the whole clause as is the case when the negation marker appears in clause-initial position. Thus, these examples are very similar to the examples of "misplaced" *not* in declaratives.

Summary of Placement Errors (other than Inversion Errors)

In the above section, I describe errors of misplaced inflection and errors of placement of negation markers. All of the examples of misplaced inflection involved *do*-support usually with irregular past tense verbs. Negation placement errors were quite rate and very rarely were utterances frankly ungrammatical because of negation placement. There were a number of examples in which children placed a *not* in a position where it had narrow scope rather than wide scope. Judging from context, in the majority of cases, children seemed to intend a narrow scope for negation.

4. Inversion Errors

Which Verbs Can Invert

Non-inverting verbs. There were no examples of the children inverting semi-auxiliaries and restricted modals. In other words, they never asked questions like **better I go?* I found no examples of these children who were learning American English inverting possessive *have* (i.e., there were no examples like *have you a match?*). In addition, I found only the following 6 questions in which children inverted a main verb.

Adam7 (2;6) Where go one?
Adam11 (2;8) Where go pencil
Adam45 (4;3) Where makes de noise?
Sarah111 (4;6) What means two?
Nathan28 (3;7) What's gonna be tomorrow?
Peter11 (2;3) Where da go frisbee go

I discuss these six questions at great length in chapter 6. For now, I would just point out that six examples out of 40,000 questions is very few (less than 0.01%)

***Aren't* and inversion.** *Aren't* has one extremely unusual, idiosyncratic property. *Aren't* can appear with *I* but only if it appears in inverted position. Compare, for example, *aren't I going? with *I aren't going*. It would seem that this seemingly arbitrary restriction would be a extremely difficult for the children to have mastered. Three children used *aren't I* a total of seven times, and no child ever said *I aren't*. Interestingly, there are no examples of any adults using *aren't I* while talking to Adam. Ross and Mark's parents used *aren't I* twice, once in transcript 57 and once in transcript 78.[12]

Adam49 (4;7) I'm getting good [#] aren't I?
Mark68 (4;1) I'm grown up # aren't I
Mark70 (4;2) I'm still working on a happy world, aren't I?
Ross2125 (2;7) I'm bad when I do it, [#] aren't I?
Ross78 (6;7) why aren't I saying dada googoo?
Ross81 (6;10) aren't I [!] supposed to be have the good guys?
Ross81 (6;10) aren't I supposed to have the good guys?

---

12 In transcript 78, there is the following fascinating interchange between Brian MacWhinney and Ross when MacWhinney incorrectly thought Ross said *I aren't* and tried to correct him:

ROS: then why aren't I saying dada googoo?
FAT: why am I not saying gaga googoo?
MOT: why he?
FAT: not why aren't I [#] why am I not saying gaga googoo. I aren't saying? can't say I aren't saying. you should say why aren't I saying.
ROS: ok ok ok.

Settings Which Allow Inversion

The children sometimes failed to invert in settings which require inversion. In about 7% of the matrix questions that had auxiliaries, children placed the subject before the auxiliary (e.g., Adam (3;3): What I will read?). Furthermore, all of the children who asked at least 25 questions with auxiliaries made this mistake at least once. The children also sometimes inverted where they shouldn't have. They inverted subject and auxiliary in 10% of the embedded questions that had auxiliaries (e.g., Adam (4;3): I wonder what are dese for?). In addition, they inverted subject and auxiliary in between 0% to 14% of how come questions that had auxiliaries (e.g., Adam (4;10): How come is dat?).

Summary of Inversion Errors

Inversion errors were by far the most common and pervasive type of auxiliary error. However, not all possible types of inversion errors occurred. For example, the children essentially never inverted verbs which are non-invertible. They did make two types of inversion errors. They neglected to invert in settings that required inversion and they permitted inversion in settings that do not allow inversion. Because inversion errors were the most common type of auxiliary error, yet even here the types of inversion errors were quite restricted, I decided to examine the problem of inversion more carefully. In chapter 5, I discuss how several linguistic theories have dealt with the problem of inversion. I also present much more extensive analyses of the inversion errors which did and did not occur.

Discussion

In some way my search for auxiliary errors was futile. Like Maratsos and Kuczaj and Pinker, I found very few of the possible auxiliary errors. My search was not, however, pointless. In the process of looking for auxiliary errors, I uncovered several interesting patterns to the auxiliary errors that did and did not occur. I would like to review some of the implications of these findings.

Main verbs and auxiliary verbs. In the above transcript analysis, I found that the children overregularized past tense and subject-verb agreement inflections for *be*, *do*, and *have* when *be*, *do*, and *have* were main verbs but not when they were auxiliary verbs. I would argue that this finding indicates that, from a very early age, the children distinguished main verbs from auxiliary verbs. It also indicates that they applied a productive learning strategy to main verbs and a conservative learning strategy to auxiliary verbs.

How this should best be fleshed out is still unclear. One possibility is that inflected forms of the auxiliaries are stored as is in the lexicon and inserted directly into the phrase structure trees, whereas for the homophonous main verbs, what is inserted into the phrase structure trees are items such as Main Verb + Past or Main Verb + Present. When it comes time for the phonological form of main verbs to be spelled out, the lexical entry which corresponds to, say, possessive *have* + Past is consulted and *had* is inserted. Perhaps, as Pinker *et al.*.(1990) have suggested, this

lexical look-up is not 100% perfect. If that were the case, sometimes main verb *had* would not be found and children would resort to the general rule of past tense or present tense verb formation. Because auxiliaries are stored and inserted fully inflected such errors are impossible. If the look-up is unsuccessful for auxiliaries, the result is the auxiliary simply isn't said.

What other possible accounts could be given to explain the overregularization of main verb *be*, *do*, and *have*, but not auxiliary *be*, *do*, and *have*. One possibility is that auxiliary *be*, *do*, and *have* are protected from over-regularization because of some structural property they share. One possibility is that they are protected by virtue of being in I. This seems an unlikely explanation for two reasons. First, it is ad hoc. There is no reason to predict or expect I to protect from generalization. Second, since the copula *be* also occupies I, for this explanation to be correct, there should have been no examples of overregularized copula *be*. However, there were 10 errors involving copula *be*. I would argue that no explanation that does not appeal to the distinction between auxiliary verbs and main verbs can account for these overregularization findings.

Let's assume that I am right and that the explanation hinges on the distinction between main verbs and auxiliary verbs. How are children able to distinguish main verbs from auxiliary verbs even when the two are homophonic? Perhaps, as part of Universal Grammar, children are born knowing that they may encounter two classes of verbs in their language. This will not solve their problems. It isn't enough that they come expecting two classes of verbs. It's not even enough to build into the children the knowledge that auxiliary verbs tend to convey information about tense, mood and aspect, that they tend to be few in number, small in

size, appear in first, second or last position and be involved in negation, questioning, etc.[13] These things certainly help, but the child still needs to do some heavy-duty bootstrapping. In the next chapters, I will outline how children approach the daunting task of learning the auxiliary system.

The conservative approach. In chapter 2, I argued that the linguistic behavior of auxiliaries is so complex that, if a child were to make most of the logically possible generalizations from one auxiliary to another, she would almost certainly make errors. In this chapter, I presented the results of an analysis of over 55,000 spontaneous uses of auxiliaries by children. The results of this analysis suggest that errors involving *do*-support and inversion were the only types of auxiliary errors that children made with any appreciable frequency. Because I extensively address the question of inversion in chapter 5 and *do*-support in chapter 6, I will only briefly comment on these errors.

***Do*-support.** Although the children made a number of errors involving *do*-support, they very rarely invoked *do*-support when it was not needed. All but one case of unnecessary *do*-support for negation or questioning involved utterances which contained the copula *be*. In each case, the children treated copula *be* as if it were a main verb which could neither negate or invert. Thus, these *do*-support errors seem to represent instances in which children treated copula *be* as a regular main verb, rather than treating it as the exceptional verb that it is. Notice that the majority

---

[13] In their cross-linguistic analysis of auxiliaries, Steele *et al* (1981) conclude that these and other traits of auxiliaries are universal to all auxiliaries. If these traits are universal, then it is possible that they are part of the innate linguistic endowment that all children receive.

of the examples of unnecessary *do*-support involve non-agentive main verb *be*.

Inversion. Likewise, while inversion errors were (relatively) common, only certain types of inversion errors ever occurred. Notice, for example, that the children essentially never inverted verbs that English does not allow to invert. Notice also that no child consistently failed to invert a verb which must invert in English. Thus, the only inversion errors that occurred were errors that involved deciding whether inversion was required in a particular setting.

The paucity of most types of auxiliary errors is consistent with children acquiring the auxiliary system in a conservative fashion. The lack of errors does not, however, prove that children are completely conservative learners. It could be that children do generalize, but that they are fairly conservative in the generalizations that they make. Notice that all the members of the *be* subtype, the *do* subtype, and the *have* subtype behave like the other members of that subtype. Children could fairly safely generalize within each of these subtypes.

Modals. The modal subtype, on the other hand, poses a number of problems for generalization. One problem with generalizing within the modal subtype is that the restricted modals (which by definition lack certain auxiliary-like traits) behave most like the modals. If children generalize the behavior of modals to restricted modals, they would do things like invert restricted modals, negate restricted modals, have restricted modals precede perfect *have*, etc. Alternatively, if they generalize the behavior of restricted modals to true modals, they would fail to invert true modals. A second problem with the modals is the fact that utterances with *can* + perfect *have* are unacceptable even though every

other modal can be combined with perfect *have*. If children were wary about generalizing within the modal category, they would be able to avoid most errors. Another way they could avoid modal errors would be if every modal is the sole member of its own subcategory. Hence, it would be impossible to generalize the behavior of one modal to any other modal.

Negation. Negation presents another serious problem for the non-conservative learner. First, some auxiliaries can only appear with negation markers. For example, one can say *you need not go*, but one cannot say **you need go*. Second, the negative marker cannot attach onto *am* or *may*, and *can*, *do*, *may*, *shall* and *will* have irregular negated forms. Third, *aren't* can appear before the pronoun *I* but not after it (e.g., *aren't I being good?* but not **I aren't being good*). Finally, certain negated forms of auxiliaries have different connotations depending on whether they appear before or after the subject. [14]

The children in the transcript analysis did extremely well with regard to negation and auxiliaries. They correctly placed the negation marker after the first auxiliary in the overwhelming majority of sentences that were negated. The rare utterances in which they did not place the negation marker after the first auxiliary were not ungrammatical, but resulted in narrow scope negation. In addition, to correctly placing the negation element *not*, children only made three errors in the almost 9,500 uses of the contracted negation marker *-n't*. Thus, it appears that children were quite conservative in the generalizations they made about negation.

In summary, the analysis of children's auxiliary errors provides little or no evidence of a productive learning strategy. However, it may be that

---

[14] Compare for example, the meanings of *won't* in *you won't go* and *won't you go?*

there are very few errors because children make the right generalizations. In addition to looking at auxiliary errors, one may be able to determine whether children are conservative or productive by examining the acquisitional time course for the various auxiliaries. For example, one could determine whether all members of an auxiliary subtype exhibit certain linguistic behaviors at the same time. One could also determine whether auxiliaries that are acquired late immediately exhibit all of the traits associated with that subtype, or whether these traits have to be learned anew for each auxiliary. The next chapter presents these sorts of analyses. Together, the error analyses and the acquisitional time-course analyses should reveal how conservative children are in their acquisition of the auxiliary system.

Chapter 4

First Uses of Auxiliaries and their Homophonous Main Verb Counterparts

In the last chapter, I looked for auxiliary errors in the spontaneous speech transcripts of 14 children. The motivation was that such errors provide insights into the verb generalizations made by children and that these generalizations are clues about the structure of the verb categories in their internal language or grammar. Because I found very few auxiliary errors with the exception of inversion errors and occasional double-tensing errors, there was little evidence of verb generalization in the way of auxiliary errors.

Studies of errors such as those reported in chapter 3 can simultaneously be too sensitive and too crude an indicator of categories of verbs. They may be too sensitive in that wrong generalizations are only one cause of auxiliary errors. Errors may also simply be slips of the tongue or production errors. The question is, what should we make of a mistake that occurs less than one time for every hundred or thousand opportunities? Does such a mistake provide a rare glimpse into the structure of a child's I-language or is it just the result of the child being distracted, tired, etc.? Looking at errors is also too crude an indicator of verb subcategories because it only lets us see the wrong categories children form.

In other words, the lack of auxiliary errors does not preclude the possibility that children make some generalizations. The lack of errors

only indicates that children don't make the sorts of overly broad generalizations which leave tell-tale errors in their wake. The lack of auxiliary errors says nothing about the correct, sufficiently narrow generalizations that children make.

Obviously, we cannot look to errors to find evidence of correct generalization since, by definition, a correct generalization is sufficiently narrow so as not to lead to errors. Instead we have to look for more subtle evidence of generalization in the pattern of correct acquisition of the various auxiliaries. If, for example, a child is willing to generalize what she knows about the auxiliary *does* to the auxiliaries *do* and *did*, then we might expect that she would acquire auxiliary *do*, *does*, and *did* at about the same age and that she would begin to use all three forms of auxiliary *do* in certain contexts (say, inverted or negated contexts) at approximately the same age. In effect, we might predict a clustering in the acquisition of auxiliaries which are related in the child's grammar.

Which particular clusterings we expect to find depends on which of the divisions of auxiliaries (and main verbs) outlined in chapter 2 we believe to be correct. If, for example, we do not believe there is any division between auxiliaries and main verbs (a position falsified by the results of chapter 3), we would predict that all the auxiliaries and main verbs would be acquired at the same time. A more refined prediction would be that a child will acquire auxiliary and main verb homophones (e.g., main verb and auxiliary *do*) at the same time. Imagine instead that we believe there is a division between main verbs and auxiliaries, but we don't think there are any further divisions within the auxiliary verbs (a position which also cannot be right given the lack of errors found in chapter 3). According to this theory, we would not predict clustering for

auxiliary and main verb homophones, but we would expect clustering of all of the auxiliary verbs.

Now let's imagine we think that there is a division between main verbs and auxiliary verbs and that there are further subdivisions among the auxiliary verbs. As discussed in chapter 2, linguists have suggested an array of possible divisions among the auxiliaries. Rather than discussing the clustering predictions of all of these theories, by way of example, I will just focus on the predictions made by two of them. The first is that there is a distinction between modal auxiliaries and nonmodal auxiliaries. According to this division, we would expect that there would be clustering for all of the modals and for all of the nonmodal auxiliaries. We would predict that children would acquire all of the nonmodal auxiliaries at the same time and that they would begin to invert or negate all nonmodal auxiliaries at the same time. Likewise, we would expect children to acquire all of the modals at the same age, and that they would begin inverting them and negating them all *en masse*.

Now let's leave the modal auxiliaries aside for the moment and concentrate on the behavior of auxiliary *be*, *do*, and *have*. Imagine that we think the various inflected forms of auxiliary *be*, *do*, and *have* form subcategories of auxiliaries. Accordingly, we would predict an acquisitional clustering of auxiliary *am*, *are*, *is*, *was*, and *were*, an acquisitional cluster of auxiliary *did*, *do*, and *does*, and an acquisitional cluster of *had*, *has*, and *have*. Each of the possible divisions of verbs discussed in chapter 2 makes different predictions about which verbs will form acquisitional clusters. By looking at the pattern of acquisition of the various auxiliaries--which auxiliaries come in at the same time, which

auxiliaries begin to invert and negate at the same time--we may be able to uncover signs of the divisions within which children generalize.

For the sake of clarity, I simplified matters slightly when I discussed acquisitional clustering. If we took what I said literally, we would predict that children would begin to use each member of a subcategory on the same day and that they would begin to negate or invert each member on the same day. Let's consider again a division system that has the modal auxiliaries as a distinct subcategory. A literal interpretation of clustering would require that a child begin to use the modal *can* and the modal *shall* on the same day.

There are two reasons why this strict requirement might not hold, yet the category could still be psychologically real. The first is that, in order to learn a word, a child must hear it. If a child has never heard the word *shall* at the point when she begins to use *can*, then obviously she cannot start using *shall* on the same day she begins to use *can*. The second reason there might be some acquisitional spread within a psychologically real subcategory is that the different members of a category can and do carry different information. This means that in every day conversation some words are more likely to be used than others because of the information they convey. Getting back to *shall* and *can*, a two-year old child probably has more opportunities to use the word *can* than the word *shall*. For the child, the word *will* probably suffices for most of the communicative needs of *shall*.

I do not mean to imply that one cannot, in practice, hope to use acquisitional clustering as a means of exploring the boundaries of verb subcategories. The hope is that, despite differences in input or communicative demands, we still may find evidence of clustering among members of a category. We should not, however, expect the clusterings to

be perfectly tight. I used analyses of variances to determine which cluster models accounted for the most variance in the data.

In the analyses which follow, I examined the pattern of acquisition of the auxiliaries and their homophonous main verb counterparts, looking for clusters. I decided that the most fruitful approach would be to determine when each auxiliary and main verb was acquired in each of its permissible contexts. So, for example, for each child and for each inflected form of each auxiliary and for each of the inflected forms of the copula *be*, I determined when that form of the auxiliary or copula began to appear in standard non-inverted position (e.g., *I should eat*), inverted position (e.g., *should I eat?*), preceding the negative element *not* (e.g., *I should not eat*), and with the contracted negation element *n't* (e.g., *I shouldn't eat*). For each inflected form of homophonous main verb *do* and *have*, I determined when each form first occurred in standard, non-inverted position (e.g., *I have two cookies*).

Measures of Acquisition. The question that arose was what criterion I should use to determine when a child acquired a particular auxiliary or main verb. One criterion for acquisition that is frequently used within the field of language acquisition is percent use in obligatory context. The particular percentage required varies, but the general idea is to establish a certain percentage as an arbitrary cutoff above which a child is said to have acquired whatever is under investigation.[1] The problem

[1] As I point out elsewhere (Stromswold, 1989), the particular percentage required is completely arbitrary. There is nothing magical about 90% (the percent used by Brown, 1973), 75% (a percentage used in a large number of studies) or 50%, for that matter. Obviously, language is not a fixed choice experiment where there are only two possible choices which can be made. However, this seems to be exactly the assumption that underlies many researchers choice of 75% use in obligatory contexts. What these

with this sort of measure is that it isn't clear how one would define the relevant obligatory context for each of the verbs.

Another measure is the age a child is when she first uses each of the verbs (henceforth, Age of First Use). It's possible, however, that the first use of a verb is actually a speech error, unanalyzed routine or imitation. These possibilities can be minimized--but not eliminated--by checking the context utterances. A third measure of acquisition is the age by which a child repeatedly uses a construction (Age of Repeated Use). A fourth possible measure is the age at which a child is regularly using a construction (Age of Regular Use). A problem with the Age of Repeated Use and Age of Regular Use as measures of acquisition is that for rare constructions (e.g., some of the modals), the age of acquisition may be inflated or, in the case of extremely rare constructions (e.g., *mightn't*), the child may never qualify as having acquired a construction she clearly uses.

The Age of First Use is inherently the most sensitive measure of acquisition since it is a measure of the earliest age at which a child could be said to have acquired a construction. The empirical question is, how accurate or reliable a measure is it? In previous work (Stromswold, 1989), I examined the relative merits of the Age of First Use, Repeated Use and Regular Use measures of acquisition. I determined when 12 children first used, repeatedly used, and regularly used each of five constructions. For the purposes of the three measures, I only counted utterances which were clear (i.e., not stuttered) and neither imitations nor routines.

---

researchers implicitly fail to realize is that there are almost an infinite number of choices a child could make about use of a construction, inflection, etc. All the child has to demonstrate is that she uses the construction more often than would be expected by chance. Imagine for example a child who only uses the *-ed* form in 1% of obligatory contexts. One could argue that, if she only uses *-ed* to mark the past tense on verbs and never applies it to an adjective, noun, adverb, etc., the child has acquired the inflection.

I defined Age of Repeated Use to be the age by which a construction had either appeared five times or had appeared twice in one month, whichever occurs earlier. Age of Repeated Use was the age at which a child began to use a construction frequently. The way I determined when this happened was to graph the number of occurrences of a construction and visually inspect the graph for the point at which the change of slope was greatest. I found that for the five constructions under investigation, the three measures were very highly correlated (all r's > .85, all p's < .01). Because of the high degree of correlation between the three measures, in the present study, I use Age of First Use as a measure of acquisition because it is the simplest and most sensitive measure of acquisition.

Age of First Use Transcript Analysis

1. Corpora

I searched the CHILDES transcripts (MacWhinney and Snow, 1985; 1990) of the 14 children shown in Table 4.1 below. These children were the same children used in the analyses reported in chapter 2, except that I replaced May's transcripts with the transcripts of Stan Kuczaj's son Abe. I used the UNIX utility "fgrep" to cull all of the children's lines which contained the letter strings: *am*, *are*, *be*, *can*, *could*, *did*, *do*, *does*, *had*, *has*, *have*, *is*, *may*, *might*, *must*, *n't*, *shall*, *should*, *was*, *were*, *will*, and *would*. This yielded a corpus which contained approximately 66,300 uses of auxiliaries.

Please Insert Table 4.1

---

2. Procedure

For each child, I went through the corpus of lines which contained an auxiliary or a main verb homophone. For each inflected form of each auxiliary and main verb, I recorded the first use of that word in standard position following the subject of the sentence. For auxiliaries and the copula *be*, I recorded the first appearance of the verb preceding the subject (inverted position), the first appearance preceding a *not*, and the first appearance with the contracted negation marker *n't*. Utterances that had contracted, stuttered, or unclear auxiliaries or their homophonous main verb counterparts were not counted. Utterances which were obviously routines or imitations were also not counted. If it was impossible to tell whether a particular word was an auxiliary or main verb, I did not count the sentence.

Lastly, in order for an utterance to count as the first use in inverted or noninverted position, the sentence in question had to contain a subject noun phrase. (Obviously, one can't tell if an auxiliary comes before or after a subject noun phrase if there is no subject noun phrase.) I did not require that the first usages in negated contexts (either *not* or *-n't*) contain a subject noun phrase. An undergraduate research assistant did the same analyses on half of the transcripts. Our concordance ratings were 95% or greater for all transcripts, all verbs, and all sentence types.

3. Results

3.1 Missing data. Not all of the children used all of the verbs in all possible contexts. Table 4.2 shows that the amount of missing data ranged from 14% to 91% for the 14 children with a mean of 43% missing data. Note that the children were different ages during the period when transcripts were recorded. Examination of Table 4.2 suggested that there might be a relationship between the amount of missing data and the ages of the children when the transcripts were stopped. I found that amount of missing data and age during last transcript session were highly correlated ($r = -.8081$; $p = .0005$). Children who were younger at the age of the last transcript session have greater amounts of missing data.[2] This is consistent with the view that missing data represent those forms that the children have yet to acquire at the date of the last transcript session.

Please Insert Table 4.2

---

[2] I did a multiple regression analysis using Age at last transcript, total line count in the transcripts, and number of auxiliaries in the transcripts as variables. This yielded the equation:

% missing data = -.5013Age + - 6.618 x e-5 Linecount + .002471Auxcount + 80.012

($F(3, 10) = 12.2603$, $p = .0011$). Age as predictor variable was marginally significant ($F(1,10) = 3.9025$, $p = .0764$) as was Auxcount ($F(1, 10) = 3.4158$, $p = .0943$). The total number of lines said by the child was not even marginally significant as a predictor variable ($F(1, 10) = .0122$, $p = .911$). The fact that the total number of lines was not a good predictor of the amount of missing data is important because it helps assure us that the Age of First Use data aren't simply a reflection of the size of the speech sample for each child (i.e., a reflection of frequency).

I decided to remove Allison and June's data because they had considerably more missing data (72% and 91%, respectively) than the other 12 children. The mean percent of missing data for the remaining 12 children was 36%. Henceforth, all of the analyses I describe are conducted on these 12 children. Table 4.3 shows the breakdown of the missing data by sentence modality and verb type. Overall, 42% of the auxiliary data are missing and 25% of the copula data are missing. There are no missing data for main verbs. There were more missing data for some sentence modalities than others. Overall, 13% of the declarative word order data, 32% of the inverted word order data, 39% of the contracted negation data, and 67% of the "verb + *not*" data were missing.

Please Insert Table 4.3

Table 4.4 shows that the percent of missing data for each auxiliary and copula form averaging across all permissible sentence modalities for that form.[3] The percent of missing data ranges from a low of 4% for copula *is* to a high of 85% for perfective *have*, with a mean of 39% missing data.

Please Insert Table 4.4

---

[3] In other words, the percentages shown take into account that *am* and *may* cannot appear with the contracted negation marker (**amn't* and **mayn't*) and that the contracted form of *shall not* is highly marked (*?shan't*).

3.2 Replaced data. Because many analyses were impossible to perform with so much missing data, I decided to replace each child's missing data with a dummy value. I assumed that missing data points represented constructions that a child had not acquired by the age of the last transcript session (see Footnote 2 of this chapter for support for this assumption). I made the further, very conservative, assumption that any such construction would have been acquired the very next day following the last transcript day.

Thus, for each child I replaced all missing data for that child with the age the child would have been the day after the last transcript session. For example, Shem had not yet used the modal *must* in noninverted position by the end of his last transcript when he was 3 years, 2 months and 2 days old. I assumed that had he been recorded the next day, he would have used *must*. Hence, I used 3 years, 2 months, and 3 days as the age at which Shem first used noninverted *must* (and the other 39 constructions he had not yet used by the last transcript session). In other words, I assumed that even though Shem had not used *must* for the first 1,159 days of his life, on the 1,160th day he would have begun to use it.

Replacing all of a child's missing data with a single very conservative dummy variable should have the effect of adding noise to the data. Replacing data may hide real effects (Type II error) but it should not yield spuriously significant effects (Type I error).

3.3 Acquisition of auxiliary verbs and main verbs. For the two analyses where it was possible, I did the same analyses of variance on the original, unreplaced data and on the data with replaced missing values. In the first analysis I did a one-way analysis of variance of the age of first

use of auxiliary and main verb forms of *be*, *do*, and *have* in declarative word order.[4] The mean age of first use was 2.32 years for the main verbs and 2.77 years for the auxiliary verbs. This main effect for Auxiliary/Main Verb was very significant and large ($F(1, 11) = 39.98$, $p < .0005$, eta = .89). I did the exact same analysis on the data with replaced missing values. For the replaced values data, the mean age of first use was 2.32 years for the main verbs (i.e., there was no missing datum) and 2.93 years for the auxiliary verbs. Analysis of variance of the data with replaced values revealed a main effect for Auxiliary/Main Verb which was approximately as significant and large as that which was found for the original, unreplaced data (for the replaced data, $F(1,11) = 65.69$, $p < .0005$, eta = .99).

3.4 Be/Do/Have. The only other analysis which could be done on the original, unreplaced data was an analysis of variance with the factors Auxiliary/Main Verb and Be/Do/Have.[5] In order to do this analysis, I had to remove data from Eve and Peter because they each had an empty cell (neither used auxiliary *have*). The analysis on the original, unreplaced data revealed a large, significant main effect for Auxiliary/Main Verb ($F(1, 9) = 72.61$, $p < .0005$, eta = .94) and a large, significant main effect for Be/Do/Have ($F(2, 18) = 12.18$, $p < .0005$, eta = .76). In addition there was a large, significant interaction between Auxiliary/Main Verb and

---

[4] I did not include auxiliaries or copulas in inverted or negated context, nor did I include the modals or auxiliary or copula *am* or *were*, because if I did the cell sizes would have been very unequal. Also, I wanted to be able separate out the effect of auxiliary vs. main verb with the effect of sentence modality. If I had included data for all sentence modalities, this would have greatly inflated the age for the auxiliaries but not the main verbs since only the copula can appear in contexts other than declaratives.

[5] Again, I did not include inverted uses or negated uses, nor did I include auxiliary or copula *am* or *were*.

Be/Do/Have ($F(2, 18) = 30.50$, $p < .0005$, eta = .77). On average, the children began using main verb forms of *be*, *do*, and *have* at the same age (2.41 years for copula *be*, 2.41 years for main verb *do*, and 2.40 years for main verb *do*), whereas this was not true for auxiliary *be*, *do*, and *have* (mean age of first use was 2.62 years for auxiliary *be*, 2.77 years for auxiliary *do*, and 3.48 years for auxiliary *have*). The interaction between Auxiliary/Main Verb and Be/Do/Have for the original, unreplaced data is shown in Figure 4.1

Please Insert Figure 4.1

A one-way analysis of variance of just the Main Verb data revealed no significant main effect of Be/Do/Have ($F(2, 18) < 1$) whereas the same analysis of the Auxiliary data revealed a large, significant main effect of Be/Do/Have ($F(2, 18) = 23.41$, $p < .0005$, eta = .85). This indicates that the significant main effect of Be/Do/Have in the original, unreplaced two factor analysis of variance is solely the result of an effect of Be/Do/Have for the auxiliaries.

I did the same two factor analysis of variance on the data with replaced missing values (including Eve and Peter's data). The results were essentially the same as they were for the original, unreplaced data. There were large, significant main effects of both Auxiliary/Main Verb ($F(1, 11) = 65.69$, $p < .0005$; eta = .93), and Be/Do/Have ($F(2, 22) = 15.64$, $p < .0005$, eta = .77). There was also a significant interaction between Auxiliary/Main Verb and Be/Do/Have ($F(2, 22) = 38.07$, $p < .0005$, eta =

.88). On average, the 12 children began using main verb forms of *be*, *do*, and *have* at the same age (2.33 years for copula *be*, 2.32 years for main verb *do*, and 2.31 years for main verb *do*), whereas this was not true for auxiliary *be*, *do*, and *have* (mean age of first use was 2.52 years for auxiliary *be*, 2.78 years for auxiliary *do*, and 3.47 years for auxiliary *have*). The interaction between Auxiliary/Main Verb and Be/Do/Have is shown in Figure 4.2.

Please Insert Figure 4.2

An analysis of just the Main Verb data again revealed no significant main effect of Be/Do/Have ($F(2, 22) < 1$), whereas the same analysis of the Auxiliary data revealed a large, significant main effect of Be/Do/Have ($F(2, 22) = 27.97$, $p < .0005$, eta = .85).

3.5 Comparison of Original and Replaced Data Sets. That the results of the analyses of variance of the original, unreplaced data and the data with replaced values are essentially identical suggest that the replaced values data are reasonable. Because of this and because many analyses were not possible if there are missing data, I used the data with replaced values in all further analyses. While the replaced data appeared adequate, I nonetheless avoided doing analyses which critically involved many replaced data points.

3.6 Tense. An analysis of variance with the factors Auxiliary/Main Verb, Be/Do/Have and Tense was performed on the

replaced values data. As in the previous analyses, this analysis of variance revealed a large significant main effect for Auxiliary/Main Verb ($F(1, 11) = 65.69$, $p < .0005$), and Be/Do/Have ($F(2, 22) = 15.64$, $p < .0005$). In addition, there was a significant main effect of Tense ($F(2, 22) = 13.29$, $p < .0005$). Children began using the *are*, *do*, and *have* forms earlier (mean age of first use = 2.47 years) than they began using either the *was*, *did*, or *had* forms (mean age = 2.69 years) or the *is*, *does*, or *has* forms (mean age = 2.71 years). As in previous analyses, there was a significant interactions between Auxiliary/Main Verb and Be/Do/Have ($F(2, 22) = 38.07$, $p < .0005$). In addition, there was a significant interaction between Be/Do/Have and Tense ($F(4, 44) = 11.45$, $p < .0005$). The interaction between Be/Do/Have and Tense is shown in Figure 4.3.

Please Insert Figure 4.3

Separate analyses of the Auxiliary and Main Verb data suggest that the main effect of Tense and the interaction between Be/Do/Have and Tense is mainly due to the Main Verb data, whereas the main effect of Be/Do/Have is due primarily to the Auxiliary data. For the Main Verb only data, there was a large, significant main effect for Tense ($F(2, 22) = 16.89$, $p < .0005$, eta = .78) and large, significant interaction between Tense and Be/Do/Have ($F(4, 44) = 18.31$, $p < .0005$, eta = .79) but no significant main effect for Be/Do/Have ($F(2, 22) < 1$). Figure 4.4 shows the interaction between Tense and Be/Do/Have for the Main Verb data.

Please Insert Figure 4.4

For the Auxiliary only data, there was a very large, significant main effect of Be/Do/Have ($F(2, 22) = 27.97$, $p < .0005$, eta = .85) but not for Tense($F(2, 22) = 3.191$, $p = .061$). There was a significant interaction between Be/Do/Have and Tense for the Auxiliary only data ($F(4, 44) = 3.150$, $p = .023$, eta = .47), but it is small in size and significance compared to that found for the Main Verb only data. Figure 4.3 shows the interactions between Be/Do/Have and Tense for the combined auxiliary and main verb data, while Figures 4.4 and 4.5 show the interactions for the main-verb-only and auxiliary-only data, respectively.

Please Insert Figure 4.5

To summarize, the results of the analyses described above compared age of first use of auxiliaries and their homophonous main verbs in declarative word order. The results of these analyses indicate that children began to use main verbs significantly before they used the homophonous auxiliaries. The results also suggest that for the auxiliaries, acquisition of *be* forms, *do* forms and *have* forms tended to cluster, whereas the same was not true for the homophonous main verbs. Acquisition of the main verbs tended to cluster along tense lines. Children acquired main verb *are*, *do*, and *have* before they acquired *was*, *did*, and *had*, or *is*, *does*, and *has*. This was not true for auxiliary verbs.

3.7 Types of Auxiliaries. The next set of analyses was designed to help determine what was the most natural groupings of the auxiliaries. I decided not to include auxiliary *am* or*were* in these analyses, so I would be better able to compare the results of these analyses with the previous analyses which did not include *am* or *were*. Hence, I included 18 auxiliaries in these analyses (*are, is, was, did, does, had, has, have, can, could, may, might, must, shall, should, will* and *would*). If I were to consider all possible combinations of auxiliaries as potential groups (i.e., if we allowed logically possible, though highly implausible groupings like *is-can-did-have*), I would have to consider 2^{18} or 262,144 possible groupings. I could no more do this than a child learning the language could. Even if I considered only the possible groupings of the *Be*, *Do*, *Have*, and Modal Groups, I would still have to consider 2^4 or 16 possible groupings.

Rather than do analyses on all 16 of the logically possible groupings of plausible groups, I restricted my analyses to examination of the following three possible groupings which are most plausible for either linguistic or learnability reasons:

1. Be/Do/Have/Modals: {Be forms} ≠ {Do forms} ≠ {Have forms} ≠ {Modals}
2. Be&Have/Do/Modals: {Be & Have forms} ≠ {Do forms} ≠ {Modals}
3. Nonmodals/Modals: {Be & Have & Do forms} ≠ {Modals}

Because I was only comparing one type of auxiliary to another, it was possible to include the data for inverted and negated uses. Because almost three-quarters of the AUX + *not* data were missing, however, I decided to analyze only the declarative, inverted and contracted negation

uses of each auxiliary.[6] I will discuss the results of each of the three groupings separately and then I will compare the results for the different models.

3.7.1 Be/Do/Have/Modal Analysis. When the data were grouped in this way, there was a large, significant main effect for Auxiliary Group ($F(3, 33) = 34.49$, $p < .0005$, eta = .87, $f = 1.77$). On average, children first used auxiliary *do* at 2.54 years of age, auxiliary *be* at 2.92 years, the modals at 3.22 years and auxiliary *have* at 3.69 years. There was also a large, significant main effect for Sentence Type ($F(2, 22) = 7.54$, $p = .003$, eta = .64, $f = .77$). Averaging across auxiliaries, auxiliaries began to appear in noninverted position at 2.94 years, in inverted position at 3.17 years, and with the contracted negation marker at 3.18 years. In addition, Auxiliary Group and Sentence Type interacted significantly ($F(6, 66) = 11.32$, $p < .0005$, eta = .71, $f = 1.01$). This interaction is shown in Figure 4.6.

Please Insert Figure 4.6

3.7.2 Be&Have/Do/Modal Analysis. When the data were grouped into these three groups, there was also a large, significant main effect for Auxiliary Group ($F(2, 22) = 41.29$, $p < .0005$, eta = .89, $f = 1.94$). On average, children first used auxiliary *do* forms at 2.54 years of

---

[6]If each of the auxiliaries could appear in each of the 3 settings, there would be 54 data points per child. Each child had only 52 data points because **mayn't* is unacceptable and **shan't* is not part of standard American English

age and modal auxiliaries at 3.22 years of age. On average, they began to use members of the Be&Have Group at 3.31 years of age. Sentence Type just sneaked in with a main effect which was significant at the $p < .05$ level ($F(2, 22) = 3.68$, $p = .042$, eta = .50, $f = .58$). There was also a significant interaction between Auxiliary Group and Sentence Type ($F(4, 44) = 15.33$, $p < .0005$, eta = .76, $f = 1.18$). This interaction in shown in Figure 4.7.

Please Insert Figure 4.7

3.7.3 Nonmodal/Modal Analysis. When the data were grouped as either being modal or nonmodal auxiliaries, there was a significant main effect for Auxiliary Group ($F(1, 11) = 20.435$, $p = .001$, eta = .81, $f = 1.36$). On average, children began to use the nonmodal auxiliaries about 2 months before they began to use modal auxiliaries (mean ages 3.05 years for nonmodals and 3.22 years for modals). There was also a significant main effect for Sentence Type ($F(2, 22) = .001$, $p = .001$, eta = .69, $f = .96$). There was not a significant interaction between Nonmodal/Modal and Sentence Type ($F(2, 22) = 1.95$, $p = .166$). This nonsignificant interaction (shown in Figure 4.8) can be compared with the significant interactions for Be/Do/Have/Modal and Sentence Type (shown in Figure 4.6) and for Be&Have/Do/Modal and Sentence Type (shown in Figure 4.7).

Please Insert Figure 4.8

All three groupings of the auxiliaries yielded significant main effects for Auxiliary Group and Sentence Type. In addition, the Be/Do/Have/Modal and Be&Have/Do/Modal groupings revealed a significant interaction between Auxiliary Group and Sentence Type. My first fear was that perhaps any of the 2^{18} logically possible grouping of individual auxiliaries would reveal significant effects.

3.7.4 Random Auxiliary Groups. I assured myself that this was not the case by grouping the 18 auxiliaries into three random equally sized groups and running the same analysis with these Random Auxiliary Groups. Random Group 1 contained the auxiliaries *are, may, have, will, should,* and *would*, Random Group 2 contained *could, did, does, had, is,* and *must* and Random Group 3 contained *can, do, has, might, shall,* and *was.*

The analysis of variance revealed neither a significant main effect of Random Auxiliary Group ($F(2, 22) = .682$, $p = .516$), nor a significant interaction between Random Auxiliary Group and Sentence Type ($F(4, 44) = .573$, $p = .684$). It did, however, reveal the expected significant main effect of Sentence Type ($F(2, 22) = 9.866$, $p = .001$, eta = .69, $f = .95$). The results of this Random Group analysis suggest that some groupings of auxiliaries are more natural than others.

3.7.5 Single-Member Auxiliary Groups. I decided to do another analysis on the data, this time treating each of the 18 auxiliaries as the sole member of its own group (i.e., the analysis of variance had one factor with 18 levels). Single-member groups represent the most fine-grained grouping possible of the data. Thus, this analysis represents one

extreme of the 2^{18} possible divisions. I could not include *may* or *shall* in this analysis because neither can occur with a contracted auxiliary (i.e., their inclusion creates empty cells that prevent the analysis).

The results of this analysis revealed significant main effects for Single-Member Auxiliary Group ($F(16,176) = 21.61$, $p < .0005$, eta = .81, $f = 1.40$) and Sentence Type ($F(2, 22) = 14.01$, p .0005, eta = .75, $f =$ 1.13). It also revealed a significant interaction between Single-Member Auxiliary Group and Sentence Type ($F(32, 352) = 11.69$, $p < .0005$, eta = .55, $f = 0.65$). The significant Main Effect of Single-Member Auxiliary Group suggests that each of the children learn the 16 auxiliaries in more or less the same order.

The existence of a significant interaction between Single-Member Auxiliary Group and Sentence Type concerned me. The significance of this interaction raised the possibility that not only did all 12 children began to use each of the 16 auxiliaries in approximately the same order, but that perhaps the order in which the children began to use each auxiliary in each sentence type was relatively fixed across children. To test this possibility, I did an analysis with a single factor that conflated Auxiliary and Sentence Type. The combined Auxiliary-Sentence Type factor had 52 levels. This analysis revealed a significant main effect for Auxiliary-Sentence Type ($F(51, 561) = 11.12$, $p < .0005$, eta = .71, $f = 1.00$).

3.7.6 Comparison of Auxiliary Groupings. Analyses of variance of the data grouped in various ways revealed that all auxiliary groupings other than the random grouping had significant main effects. In other words, the amount of within-group variance was significantly less than the between-group variance whether the groupings were as coarse-

grained as Modal/NonModal or as fined-grained as Individual-Auxiliaries-in-Specific-Sentence-Types (e.g., *could*-inverted). I will compare the goodness of the various groupings by comparing the effect-size, *f*, for each of the groupings.

The effect-sizes for the various groupings are (in order from largest to smallest effect-size):

1. Be&Have/Do/Modal Grouping effect size, $f = 1.94$
2. Be/Do/Have/Modal Grouping effect size:, $f = 1.77$
3. Single-MemberAuxiliary Grouping effect size, $f = 1.40$
4. Modal/Nonmodal Grouping effect size, $f = 1.36$
5. Each-Auxiliary-Each-Sentence-Type Grouping effect size, $f = 1.00$

Notice that all five of the groupings above have reasonably large effect sizes which suggests that all five groupings capture a good deal of the variance in the data. The grouping which captures the largest amount of the variance, however, is the Be&Have/Do/Modal Grouping with the Be/Do/Have/Modal Grouping a close second. Interestingly dividing the auxiliaries into modals and nonmodals is less good than having each auxiliary be the sole member of its own category. This suggests that there are some distinction(s) among the nonmodals which affect age of first use. Whatever distinctions are lost when all the nonmodal auxiliaries are lumped together, dividing auxiliaries into modals or nonmodals is still more explanatory than treating each auxiliary in each sentence type as a separate group.

3.8 Sentence Type. In all of the analyses above, including the Random Grouping analysis, there is a clear main effect for Sentence Type.

In general, the children begin to use auxiliaries in standard declarative word order before they use them in either inverted position or with a negation marker. The interaction that is seen between Auxiliary Group and Sentence Types for some of the groupings of auxiliaries suggests that the effect of Sentence Type is not uniform for all auxiliaries. That there is an effect even when the groupings of auxiliaries are random suggests that the main effect of Sentence Type isn't due solely to a "natural" subset of the auxiliaries.

3.9 Copula and Auxiliary *Be*. I next compared the data for auxiliary and copula *be*. I chose not to include the data for auxiliary or copula *were*, because over half of the data were missing. I could not include data for *am* because **amn't* is not acceptable (i.e., there is an inherently missing cell in the data). The analysis of variance on the remaining data revealed significant main effects for Copula/Auxiliary, Verb Form, and Tense factors. On average, children began using copula *be* forms about 3.3 months before they began to use auxiliary *be* forms ($F(1, 11) = 24.79$, $p < .0005$, eta = .83, $f = 1.50$;). On average, children began using copula and auxiliary *is* at 2.53 years of age, *are* at 2.79 years, and *was* at 3.02 years.($F(2, 22) = 68.96$, $p < .0005$, eta = .93, $f = 2.50$).On average, children began to use copula and auxiliary *be* in noninverted positions when they were 2.43 years old, in inverted positions when they were 2.73 years and in with contracted *-n't* when they were 3.19 years old ($F(2, 22) = 40.24$, $p < .0005$, eta = .89, $f = 1.91$).

In addition to the main effects, Tense interacted significantly with Auxiliary/Copula ($F(2, 22) = 5.46$, $p = .012$, eta = .58, $f = .70$). Tense also interacted significantly with Sentence Type ($F(4, 44) = 9.47$, $p <$

.0005, eta = .68, f =.93). The interaction between Tense and Auxiliary/Copula is shown in Figure 4.9 and the interaction between Tense and Sentence Type is shown in Figure 4.10.

Please Insert Figures 4.9 and 4.10

The three-way interaction among Auxiliary/Copula, Tense, and Sentence Type Factors just missed being significant ($F(4, 44) = 2.42$, p = .062, eta = .42, f = .47). In order to further investigate this marginal interaction, I did separate analyses of variance for the copula and auxiliary data. For the Copula-only data, the interaction between Sentence Type and Tense was significant ($F(4, 44) = 3.48$, $p = .016$). This interaction in shown in Figure 4.11. For the Auxiliary-only data, the interaction between Sentence Type and Tense was even more significant than that for the Copula data ($F(4, 44) = 7.16$, $p < .0005$). As is shown in Figure 4.12, children began using auxiliary *is*, *are*, and *was* in declarative contexts at about two-and-a-half years of age. About 8 months later they pretty much simultaneously began to use auxiliary *is*, *are*, and *was* attached to the negation marker *-n't* . Children began inverting auxiliary *is* and *are* when they were about two-and-a-half years old. They didn't begin inverting auxiliary *was* for another full year.

Please Insert Figures 4.11 and 4.12

---

4. Discussion

In this section, I will briefly discuss the implications of the acquisitional clustering results discussed above. I will restrict my comments to those that pertain to the questions raised in chapter 2 about the linguistics and acquisition of the auxiliaries and main verbs. I will argue that the Age of First Use data provide support for acquisitional divisions between auxiliary and main verbs, between auxiliary and copula forms of *be*, and between the various types of auxiliaries.

Main Verbs and Auxiliaries

A number of results presented in this chapter strongly support an acquisitional distinction between auxiliary and main verbs. The first is the finding that overall children use main verb *be*, *do*, and *have* significantly earlier than they use the corresponding auxiliary verbs. If children applied what they knew about main verbs to homophonous auxiliary verbs, we would expect no such difference.

A second finding is a sort of double-dissociation in the effects of Tense and Be/Do/Have on auxiliaries and their homophonous main verbs. The tense of a main verb significantly affected when that main verb was first used. Tense did not significantly affect when children acquired auxiliaries. Conversely, whether a particular auxiliary was *be*, *do*, or *have* affected when that particular auxiliary was acquired. The same was not true for the homophonous main verbs. The children acquired main verb *do*, *be*, and *have* verbs at approximately the same age.

I would like to argue that the reason that there is a main effect of Be/Do/Have for auxiliaries but not for main verbs is that the Be/Do/Have distinction is a (sub)categorical distinction for auxiliaries but not for main verbs. For the main verbs, the important distinction to be made among main verbs *am, are, did, do, does, had, has, have, is, was* and *were* is a distinction among the tenses of these verbs. For the auxiliary verbs, the *Be/Do/Have* subcategory takes precedence.

Copula and Auxiliary *Be*

As I discussed in chapter 2, copula *be* is unique among the non-auxiliary verbs in that it behaves like an auxiliary. At least in American English, copula *be* is the only non-auxiliary verb that can invert positions with a subject noun phrase or precede the negation markers *not* or *-nt*[7]. In fact, the only ways that the progressive auxiliary *be* and the copula *be* differ are that:

1) copula *be* can be the sole verb in a non-eliptical sentence, whereas auxiliary *be* cannot.
2) copula *be* can precede a noun phrase, adjective phrase, etc. but auxiliary *be* cannot
2) copula *be* can take the progressive inflection whereas the progressive auxiliary *be* cannot (e.g., compare *she is being mean.* with**she being hurting him*).

One might expect, therefore, that children would be tempted to generalize what they know about copulas to progressive *be*, and vice versa. It's even possible that children would initially think that there was only one

---

[7] Leaving aside, for the moment, cases such as *she tried not to laugh*, which involve untensed clauses.

type of *be*. If children do generalize, then they should acquire the copula and auxiliary forms of *be* at the same time. They should also begin to invert and negate copula *be* and auxiliary *be* at the same time.

The results of the Age of First Use analyses suggest that from the very earliest stages, children distinguish between copula *be* and auxiliary *be*. Recall that children acquired copula *be* over 3 months before auxiliary *be*. In addition, there was a significant interaction between Auxiliary/Copula and Tense and a near significant three-way interaction among Auxiliary/Copula, Tense and Sentence Type. This suggests that not only do children acquire auxiliary *be* later than copula *be*, but the pattern of acquisition differs for the two types of *be*. This finding is contrary to current thinking within the Government-Binding framework which considers auxiliary and copula *be* to be one and the same.

Distinctions among the Auxiliaries

The pattern of acquisition of main verbs and auxiliaries suggests that the two types of verbs are distinct in the children's grammar. I next looked for acquisitional clusters within the auxiliaries. I considered a number of different possible clusters. I found that even the smallest possible grouping (particular-auxiliary-in-particular-sentence-type) was significant. One could surmise that this shows that children don't make even the narrowest generalizations among the auxiliaries. To do so would be wrong. What this result suggests is merely that the pattern of acquisition for particular auxiliaries in particular sentence types is quite similar from child to child. This result does not mean that the data cannot

cluster more naturally (i.e., more significantly) into broader, higher-order clusters.

I tested a number of broader groupings which were plausible from either a linguistic or a learnability standpoint. All of these broader groupings captured a significant amount of the variance in the data. Of all of these groupings, the grouping that had the greatest effect size was one that grouped all of the modals in one group, all of the forms of auxiliary *do* in a second group, and all of the forms of auxiliary *do* and *have* in a third group. This is consistent with current thinking within the Government Binding Framework which suggests that auxiliary *be* and *have* share certain syntactic properties (Chomsky, 1989; Pollock, 1989).

5. A Comparison andReconciliation of the Results of the Error and First Use Analyses

In chapter 3, I hoped to learn about the nature of children's auxiliary and main verbs by studying the errors they made when they used these verbs. In this chapter, I adopted the opposite approach. Here I looked at what children did right. In other words, instead of looking for evidence of wrong generalizations, I looked for evidence for correct generalizations. Whenever one approaches a problem from two starting points, one runs the risk of ending up in different places depending on where one started. Fortunately, the analyses of what children did wrong and what they did right both led me to the same place.

Main verbs and auxiliary verbs. The results of both error and cluster analyses strongly indicate that children distinguish between

auxiliaries and their homophonous main verbs. As reported in chapter 3, children did not invert, negate, or 'stack' main verbs. The children also didn't inflect auxiliaries, have them appear as the sole verb in a sentence, etc. In chapter 4, children acquired auxiliaries significantly later than their main verb counterparts.

Types of auxiliaries. The results of the error analyses and the age of first use analyses suggest that children do not have an undifferentiated AUX category. They appear to distinguish between the various types of auxiliaries. For example, if the children had not distinguished between modal auxiliaries and nonmodal auxiliaries, they would have attached subject-verb agreement inflections to the modals, saying **cans*, **musts*, etc. They made no such errors. Because the children made no errors consistent with having applied what they knew about one auxiliary to another auxiliary, the results of the error analyses are consistent with children's auxiliary system having been organized in such a way that each auxiliary was the sole member of its own subcategory. The error analysis results were also consistent with their being larger categories (such as *do* auxiliaries, *be* auxiliaries, etc.) as long as these categories were such that is was not possible to make generalizations that led to errors.

The results of the patterns of acquisition analyses suggests that children do have larger-than-one-member subcategories of auxiliaries. I tested the goodness of a number of different plausible subcategories of auxiliaries ranging from single-member-categories to modal/nonmodal categories. The best subcategorizations was one that distinguished *be* and *have* from *do* from the modals. However, this was only slightly better than a subcategorization that distinguished *be* from *have* from *do* from the

modals. What is interesting is that single-member-categories were actually somewhat better than a categorization that only distinguished between modals and nonmodals. This suggests that there are important differences between *do*, *have*, and *be*.

Copula *be* and auxiliary *be*. The results of both studies suggest that children distinguish between the homophonous auxiliary *be* and copula *be* despite all of the similarities the exist between copula and auxiliary *be*. In chapter 3, I found that children applied the progressive inflection *-ing* to main verb *be* but not auxiliary *be*. I found children overregularized copula *be* but not auxiliary *be*. I found that the children usually inverted and negated copula *be* correctly, but over a dozen times, they incorrectly used *do*-support. I found no examples of *do*-support with auxiliary *be*.

In this chapter, I found children acquired copula *be* at a significantly younger age than they acquired auxiliary *be*. If we look at the use of copula and auxiliary *be* in declarative, inverted, and negated contexts, we see that it isn't just a matter of auxiliary *be* being on a delayed time schedule compared to copula *be*. The relative ages at which the children used *are* and *was* in declarative, inverted, and negated contexts varied according to whether *are* and *was* were copulas or auxiliaries.

One might predict that the children would have begun inverting copula *be* late compared to auxiliary *be* because copula *be* is the only non-auxiliary to invert. (For a full discussion of the inversion behavior of auxiliary and copula *be*, please see chapter 5.) This does not seem to be the case. In fact, children began inverting copula *be* before they began inverting auxiliary *be*. Thus, initially, children don't seem to have any difficulty inverting copula *be*. The *do*-support errors (e.g., Adam29 (3;5): does it be around it?) came later. One could argue that initially children

notice that copula *be* inverts, so they invert it. Later, as they realize the ways that copula *be* acts like a main verb, they occasionally get confused and use *do*-support instead of inverting.

Chapter 5

Subject-Auxiliary Inversion

I would now like to consider more carefully one of the types of auxiliary errors discussed in chapters 2 and 3, mainly errors of subject-auxiliary inversion. My reasons for doing so are three-fold. First, errors of inversion are (relatively) common. They were the most common type of auxiliary error in the transcripts I examined, and they are among the most frequent type of error mentioned in the acquisitional literature. Second, of all of the aspects of the auxiliaries, the question of when and what to invert is among the most challenging from a learnability standpoint. Lastly, the phenomena surrounding inversion are quite complex from a linguistic standpoint, and different linguistic theories have approached the problem in fundamentally different ways.

I will begin this chapter with a learning-theoretic discussion of the complexities surrounding inversion, drawing heavily on the work of Pinker (1984). I will then briefly discuss how three linguistic theories deal with the phenomena of inversion. Next, I will review the acquisitional literature which deals with inversion, focusing on two questions. First, I will determine how common inversion errors really are in the language of young children. Second, I will review the factors that have been suggested as reasons or causes for inversion errors. Having laid this backdrop, I will reexamine the corpus of inversion errors collected in chapter 3. I will try to determine whether the children's inversion errors were random, or

whether there was some hidden pattern in the errors which might provide a clue as to why inversion errors are (relatively) common in children's early language. I will then present the results of a judgment experiment, designed to determine if some or all children go through a stage during which inversion is optional or unacceptable in *wh*-questions.

1. Learnability and Inversion

There are no simple surface cues that the child can use to determine when to invert subject and auxiliary (see Pinker, 1984). Auxiliary placement, therefore, provides the child with a potentially serious learnability problem. In English, at least eight factors seem to affect auxiliary placement. They are: *wh*-word present or absent, *wh*-word location, *wh*-word choice, embedded or matrix clause, modality of sentence (e.g., declarative, interrogative, exclamative, negative polarity), prosody (e.g., rising, falling or steady pitch contour and stress), and auxiliary or verb choice. Before proceeding, I would like to give the reader a sense of the complexity of the interactions a child might have to consider by discussing just a few of the interactions that seem to be present in English.

1. Modality, Prosody, Presence of *Wh*-Word, and Inversion Interact. In English, there appears to be a 4-way interaction between modality, prosody, presence of *wh*-word, and inversion. Consider the declarative sentence *she can eat* which has no inversion and a flat prosody contour. The echo *yes/no* interrogative *she can eat?* also has no subject-auxiliary inversion though it does have a rising contour. The

interrogative *can she eat?* has subject-auxiliary inversion and a falling contour. Negative polarity sentences (e.g., *never have I seen such a mess*) and exclamatives (e.g., *can she eat?*) have inversion and rising-then-falling pitch contours. *Yes/no* exclamatives (e.g., *can she eat!*) and *wh* - exclamatives (e.g., *what foods she can eat!*) both have rising then falling contours, but only *yes/no* exclamatives allow inversion.

The picture is further complicated when one considers the interaction with the type of interrogative. The pitch contour is not the same for *wh*-questions and yes-no questions nor for 'standard' or echo interrogatives. For example, *wh*-word fronted questions like *what can she eat?* don't have the falling pitch contour of inverted yes-no questions (e.g., *can she eat?*). Both uninverted *yes/no* questions (e.g., *she can eat?*), and echo *wh*-questions (e.g., *she can eat what?*) lack subject-auxiliary inversion and have a rising contour. However, echo yes-no questions and *wh*-questions have different stress patterns. In the echo question *she can eat?* , the stress can be on any one of the three words. Which word is stressed depends on what element of previous statement is being questioned. If one wants to know **who**, the word *she* is stressed. If one wants to know what she can **do**, *eat* is stressed. If one wants to cast doubt on the fact that she can in fact eat, the stress pattern is *she **can** eat?*.

2. *Wh*-word Position, Prosody, and Inversion Interact.

There are two types of echo questions. The first type (discussed above) attempts to clarify an earlier statement. For example, if I say *she can eat apples* and you respond with *she can eat what?*, your response is an echo question which seeks to clarify (or confirm) **what** she can eat. The second type of echo question is a request to clarify or repeat a part of an earlier

yes-no question. For example, if I say *can she eat apples?* and you respond *can she eat **what?***, your response is an echo question of my question. You are asking me to repeat the last part of my question before you answer it.

Notice that subject-auxiliary inversion can occur when the *wh*-word appears in sentence initial position (e.g., *what can she eat?*) or when the *wh-* phrase appears *in situ* as in the second type of echo question (e.g., *can she eat what?*). However, when the *wh*-word appears *in situ* and there is subject-auxiliary inversion, the result is grammatical only with a certain sort of prosody. The *in situ wh*-phrase must be stressed and at a higher pitch. The importance of prosody is underscored by the fact that sometimes the echo question and the standard question have the same word order and differ only in prosody. Matrix and embedded *wh*-subject echo questions differ from standard questions only in that the *wh*-word is stressed and has a higher pitch (e.g., *who is running?* vs. ***who** is running?*). The same is true for embedded object questions (e.g., *you know what she can eat?* vs. *you know **what** she can eat?*).[1] The child must be cognizant of the interaction among *wh*-word placement, inversion and prosody shown in Table 5.1 below.

Insert Table 5.1 here

---

3. Matrix vs. Embedded Clause, Modality, and Inversion Interact. In English, the auxiliary appears before the subject in matrix

[1]You would ask *you know **what** she can eat?* if you are concerned that *what* wasn't really the *wh-* word I said.

questions like *what can she eat?* but not in embedded questions like *I wonder what she can eat.* In declaratives, the subject occurs before the verb in both matrix and embedded clauses. In negative polarity sentences, the verb occurs before the subject in both matrix and embedded clause. How does the child sort this out? The question of whether to invert in embedded questions is made more difficult by the fact that the child occasionally hears grammatical quasi-quotative utterances like "I wonder...who is that." If the child fails to consider prosody, he may use such utterances to conclude that inversion is required or at least optional in embedded questions. Because there is no fool-proof way to distinguish matrix from embedded clauses that does not involve considerations of prosody, the child might actually have to consider the 4-way interaction between prosody, clause-type, modality, and inversion rather than just the 3-way interaction between clause-type, modality, and inversion.

4. Choice of *Wh*-word and Inversion Interact. The child must learn that unlike all the other *wh*-words, there is never subject-auxiliary inversion with the *wh*-word *how come.*

5. Choice of Verb or Auxiliary and Inversion Interact. The child must learn that in American English, some but not all auxiliaries will undergo subject-auxiliary inversion and main verbs other than copula *be* never invert. In other languages, main verbs can invert. For example, *better I go?* and *go I?* are ungrammatical in English, but in German, French and Scandinavian languages similar verbs can invert. In American English, the perfect *have* can invert (e.g., *has she eaten?*) but the possessive *have* cannot (e.g., **has she food?*) while in British English both perfect and

possessive *have* can invert (though inversion is optional for possessive *have*).

Let's imagine for a moment a language in which +/-Inversion is determined by the 7 factors listed above, and that these factors are completely independent factors and all combinations are possible. Rather than try all possibilities, it would make sense to hunt for generalizations about which values of factors tend to allow inversion. A child who is unwilling to make generalizations about which combinations require inversion will have to consider 13,600 combinations.[2] Even if we assume that the child is born knowing that these are the 7 factors that determine whether inversion occurs and we assume the child is also born able to recognize these factors when she sees them (i.e., she can tell which words are question words, which clauses are matrix and which are embedded, which words are auxiliaries, etc.), considering 13,600 possibilities seems an impossibly daunting task to master in less than 5 years. There are four sorts of approaches one could take at this point. The first would be to argue that the child doesn't really have to learn all of these factors because certain of these factors are part of Universal Grammar. One might argue, for example, that children are born with the knowledge that, if their language allows inversion, it will never be used in standard declarative sentences or in embedded clauses.

---

2 The figure 13,600 comes from the following calculations:
For *-wh* -structures: 2 clause types (matrix vs embedded) x 4 sentence modalities (declarative, interrogative, negative polarity, exclamative) x 4 prosody contours x 25 auxiliaries and copulas = 800
For *+wh*-structures: 8 *wh*-words x 2 *wh*-positions (*+/- in situ*) x 2 clause types x 4 sentence modalities x 4 prosody contours x 25 auxiliaries & copulas = 12,800

A second approach one could take is to bite the bullet and say, 'yes, children consider all possible combinations of factors when deciding whether to invert or not.' The fact that inversion errors are relatively common could be taken as support for the consider-all-possibilities approach. Children frequently fail to invert because they take a very conservative approach and don't invert a structure until they have heard it inverted. The consider-all-possibilties-and-be-conservative approach predicts strongly that, while children may fail to invert where inversion is required, they should never make the complementary error of inverting where inversion is prohibited. Based on the results discussed in chapter 3, this does not seem to be the case.

A third approach one could take to the problem of when to invert would be to reduce the number of factors that need be considered. Grimshaw (1979) takes this approach when she suggests that modality of sentence is the only factor that determines whether there is subject-auxiliary inversion.[3] Grimshaw argues that subject and auxiliary can only invert in non-declarative sentences. Negative polarity sentences, interrogatives, exclamatives, etc., all invert. The reason that embedded questions don't have inversion is because they aren't really questions. She points out that the indirect question *the publisher told us when he had seen such a book* is not a request for information while the corresponding matrix question *when had he seen such a book?* is. She contrasts this with the parallel pair of matrix and embedded negative polarity sentences *(never in his life had he seen such a book*, and *the publisher told us that never in*

---

[3] I can't help but notice that Grimshaw's suggestion may just be a semantic/pragmatic reflection of the suggestion that a properly filled CP triggers inversion. (See the section and footnote connected to the Principles & Parameters approach section.)

his life h~d he seen such a book) which both have the force of negative polarity and both have subject-auxiliary inversion. If all the child has to do to know when to apply inversion is to decide whether a sentence is declarative or not (presumably by using prosodic and contextual information),[4] this would greatly simplify the learning task for the child.

The fourth approach one could adopt concerning inversion is to argue that it only looks like there are 7 factors affecting invertibility. One could argue that inversion is just the automatic reflex of a number of independently motivated principles of grammar acting in concert. In other words, one could argue that the child doesn't have to learn anything new in order to know when to invert.

2. Linguistic Analyses of Subject-Auxiliary Inversion

2.1 Principles & Parameters /Government-Binding Approach

Much of the recent technical and theoretical work within the principles-and-parameters framework adopts this fourth approach. Chomsky (1986, 1989) and Pollock (1989) argue that the pattern of inversion in English is the necessary outcome of the combined forces of the Empty Category Principle (ECP), theta-theory, the weakness of English agreement category AGR, and a principle of "Least Effort".

---

[4] Presumably the child knows the pragmatic force of the sentences he is about to utter. It is an open question as to whether he is consciously aware of the intonation contour of what he is about to say.

2.1.1 V-movement

Within the current principles-and-parameters framework, the only transformation is the move-alpha transformation. While this is the only transformation, it is an extremely powerful and vital part of the theory. The move-alpha transformation basically allows anything to move anywhere as long as no constraint is violated by the movement. It is used to generate long distance dependencies, passives and raising constructions. It is also what causes subject-auxiliary inversion.

In subject-auxiliary inversion, anything that can get to AGR-S head position of I' (head position T of TP in Pollock, 1989), can under the proper conditions move to the head position COMP (C) of the Complementizer Phrase because IP alone isn't a barrier. The question of what can invert is really a question of what can get to AGR-S position. The modal auxiliaries are usually thought to be generated in AGR-S position so their invertibility causes no inherent problems. For ECP reasons, the verb head (V) of the Verb Phrase (VP) can only reach C position by first passing through the head position AGR where it 'amalgamates' with AGR-O (AGR in Pollock's notation) to form V-AGR-O. V-AGR-O must then raise to head position AGR-S where it amalgamates again. This twice amalgamated V-AGR-O-I then is able to move to COMP forming a 4-member chain between the amalgamated inflected verb and its verb-traces left in AGR-S, AGR-O, and in the original position. If the AGR-S position is filled by a modal or by the trace of a previously moved V then no further element can move into that position on its way to C. This explains why one does not get more than one

verbal element inverting (i.e., it explains why questions like **what could have she eaten?* are bad).

The question that immediately comes to mind is: why are questions like **eats she cookies?* ungrammatical in English but acceptable in French, Spanish, etc.? The reason that is currently in vogue is that the nature of AGR differs parametrically across languages and this difference accounts for the acceptability of inversion of tensed main verbs in French but not in English. The claim is that inversion is unacceptable in languages with weak AGR because the weak AGR cannot transmit the theta-role from a verb to its trace. Thus, the reason that main verbs in English cannot invert is that they cannot get to AGR-S without violating the ECP. Because main verbs cannot move to I, a dummy element *do* is inserted in the modal position to bear the raised affix of the main verb. Once *do* amalgamates with the affix, *do*-I-AGR is free to raise to C. A related question is, why can one invert auxiliary and copula *be* and perfect (and in British English possessional) *have*. The reason that Pollock gives (and Chomsky accepts) is that copula and auxiliary *be* and perfect *have* can move to AGR-S (and hence to C) because they do not have any theta-role to transmit. Since they don't have any theta-roles to transmit, the weakness of AGR doesn't result in ECP violations (see Figure 5.1).

There are a number of questions about inversion that remain unanswered. Perhaps the most fundamental is what prompts the head of V to move to C in questions, *yes/no* exclamatives, and negative polarity sentences rather than merely staying at the head position of AGR-S. The reason Chomsky gives for the necessity of V-movement for *yes/no* questions is that if V stays in the head position of AGR-S in *yes/no* questions "the resulting form is indistinguishable from the declarative at

PF, and is, furthermore, illegitimate (at S-structure) if Q is a real element as postulated" (Chomsky, 1989, p. 9). While this may make sense for *yes/no* questions, it doesn't explain the need for V-movement to C in *wh*-questions. **What she can eat?* is distinguishable at PF from *she can eat cookies* (or the in situ *wh*-question, *she can eat what*), yet it is unacceptable. I would argue, therefore, that some further explanation is needed. For the purposes of this paper, I will assume that V-movement to C is triggered by having an appropriate element in CP. I will not attempt to settle on a more precise formulation of what counts as an "appropriate element in CP." There are two possible formulations: first, that any element in CP which has clausal scope causes V-movement to C; and second, that any element that is in CP automatically has sentential scope, and thereby, triggers V-movement.[5]

Insert Figure 5.1 here (V-raising tree-structure)

2.1.2 Acquisitional Issues

[5] With either formulation one would have to argue that in *wh*-exclamatives, the *wh*-phrase does not have clausal scope, but in *yes/no* exclamatives, the *yes/no* exclamative operator does have clausal scope. One would also have to argue that *in situ wh*-questions that have inversion (e.g., *can she eat what?*) are echo yes/no questions with an abstract Q morpheme in CP which force V-movement. *In situ wh*-questions that do not have inversion (e.g., *she can eat what?*) would have to be echo *wh*-questions in which the *wh*-element does not have clausal scope (at least at S-structure).

I would now like to address a few somewhat technical points about the account of V-raising given above. I do so because these points may have implications for acquisition.

Which verbs are theta-markers? While I agree that auxiliary *have* and *be* intuitively do not seem to assign theta-roles, I do not agree that the same is true for either copula *be* or possessive *have* in British English. Pollock argues that in sentences like *John is happy* and *Mary is in Florida*, the *John* is assigned a theta-role by the predicate adjective *happy* and *Mary* is assigned a theta role by the locative *in Florida*. But notice that the same sentences without the copulas are unacceptable. While one could argue that this is due to the lack of tense in the copula-less sentences, I would like to argue that, the child (and some linguists), might use the ungrammaticality of copula-less sentences as evidence that the copula assigns some sort of abstract theta role such as property.

The problem seems to be even more acute for possessive *have* in British English. In order for the "lack-of-theta-role-assignment" argument to go through, Pollock would be forced to argue that in Britain, *have* (at least occasionally) does not assign a theta-role while in America it always does. Obviously this is not something he wants to do. What Pollock is forced to assume is that the British English *have* is a linguistic "fossil" from Old English when main verbs could invert.

The exceptional nature of copula *be* and British possessive *have* might create difficulties for the language learner. If children think that copula *be* is a theta-marker or if they are confused about whether it is a theta-role assigner, they should be reluctant to raise it. This reluctance might be evident in more frequent failures to invert the copula *be* than

other auxiliaries. Perhaps most strikingly, we would predict that children would fail to invert the copula *be* more often than the auxiliary *be* (i.e., they would ask questions like **she is happy?* more frequently than **she is laughing?*). We might also predict that children would occasionally use *do*-support with copula *be*. More specifically, we would predict that use of *do*-support would be more frequent with copula *be* than with auxiliary *be*. We would make the same sorts of predictions concerning the use of possessive *have* among children learning British English.

Are modals, *be* , *have* , and *do*+I treated the same? Since there is one move-alpha rule that permits raising to COMP, we might expect that the child would learn raising for one auxiliary and then use this knowledge to raise all other auxiliaries. If this were the case, then we would expect that once the child recognizes that a particular lexical item is an auxiliary, he should invert as often on that auxiliary as his current raising performance on the auxiliaries he knows. If the child has this sort of general raising-to-COMP 'rule', we would expect that raising of auxiliaries that are acquired later would be learned faster and with fewer errors than the first-learned type of auxiliaries.

However, one might predict different patterns of acquisition for the different verbs that raise. If the problem is raising V to AGR-O, amalgamating and then raising V-AGR to AGR-S, we would predict children would correctly raise modals before other auxiliaries. If the problem is *do*-support, we would expect children to have particular trouble with sentences in which AGR is the head of VP. And as mentioned above, if the child is confused by whether copula *be* assigns a theta role, this could delay or impair raising in copula *be* questions.

Raising and Embedded Clauses. Within the principles-and-parameters approach described above, whether the move-alpha (or affect-alpha) transformation occurs or not depends on the interaction of the grammatical subsystems. However, it is not at all clear what interaction would require V-raising to COMP in matrix questions and prohibit it in embedded questions. The ECP prevents raising to AGR-S any element other than the initial verb of VP governed by I. However, the V of an embedded clause is no less governed by I than the V of a matrix clause.

According to Chomsky (1986a, p.6) "of the three elements that might appear in pre-IP position--namely inflected verb, *wh*-phrase (or similar element), and complementizer--only two may appear, one in specifier position and one in the head position." Only complementizers and V_I can occur in C because scope considerations demand that operators and *wh*-phrases appear in SPEC. Perhaps Chomsky would rule out V-raising in embedded clauses by saying that all embedded clauses have a complementizer that is either phonologically null or obligatorily deleted. (I will refer to this phonologically null/deleted complementizer as *COMP.) In English, only *for*, *that* and perhaps *COMP can occupy head position of CP.

What evidence is there for *COMP? The ungrammaticality of sentences (1), (2) and (3) and the acceptability of (4) would seem to be evidence for *COMP; (1) and (2) could be unacceptable because they have both *COMP and *that*. The acceptability of (8) argues against this explanation. If all embedded COMPS are filled with *COMP, then (8) should be ungrammatical. (Presumably, (1), (2) and (7) are

ungrammatical because they violate ECP, which requires that *wh*-traces be properly governed.)

(1) *I know who that is John.
(2) *I know who that John is.
(3) *I know who is John.
(4) I know who John is.
(5) Who$_i$ do you know t_i saw John?
(6) Who$_i$ do you know John saw t_i?
(7) *Who$_i$ do you know that t_i saw John?
(8) Who$_i$ do you know that John saw t_i

One prediction that the *COMP theory makes is that any language that accepts sentences like (1), (2), and (8) should allow subject-auxiliary inversion in embedded questions. If it is true that it was possible in Middle English to say such things as *I wonder who that you love*, and if it also was possible to omit the overt complementizer in embedded clauses, then the filled *COMP theory would predict that inversion was grammatical in Middle English embedded questions. The equivalent acquisitional prediction is that if a child inverts in embedded questions, we should expect her to also say sentences like (1), (2) and (7). Conversely, if a child says sentences like (1), (2) and (7), she should ask inverted embedded questions.

Instead of taking a *COMP approach, we could attempt to explain the failure of inversion in embedded questions by adopting some version of Emonds' Structure Preserving Constraint (Emonds, 1976). The Structure Preserving Constraint (SPC) is a universal constraint which requires that

so-called root transformations which are neither local[6] nor structure preserving,[7] only be performed on matrix clauses. Emonds claims that subject-auxiliary inversion never occurs in embedded questions because subject-auxiliary inversion is a root transformation. The SPC approach has the added advantage that because it is supposed to be a universal constraint, the child could innately know it. Reducing the possible types of transformations to 3 greatly simplifies the task of learning a language (Grimshaw, 1979). The child only formulates rules of these types, and, once the child identifies which of the 3 types a particular rule is, the child automatically knows a great deal about when to apply the rule. If the SPC is innate, one would expect that while the child might occasionally fail to invert subject and auxiliary, he would never make the mistake of inverting in embedded questions. The child would identify subject-auxiliary inversion as a root transformation and never attempt to perform it on embedded clauses.

A third approach is a sort of grammaticalization of the semantic/pragmatic approach to inversion suggested by Grimshaw (1979). If one assumes that it is the presence of an element with clausal scope in SPEC of CP which triggers inversion, one could argue that the reason embedded questions don't invert is because the question operator is not really in SPEC position. Of course, then the problem becomes one of explaining why this is so and how the child knows it is so.

---

[6]A local transformation affects only a single noun phrase node and an adjacent node (Emonds, 1976).

[7]A structure preserving transformation introduces or substitutes a constituent C into a position in a phrase-marker held by a node C (Emonds, 1976).

All of the accounts I have suggested have problems.[8] The idea of a structure preserving constraint that distinguishes matrix from embedded clauses doesn't fit well into the current principles-and-parameters framework. The grammaticality of question (8) argues against the proposal that embedded C position is filled with a non-overt complementizer. Furthermore, both proposals leave unexplained Grimshaw's observation that modern English allows inversion in embedded negative polarity sentences. It is unclear what the relevant difference between embedded questions and negative polarity sentences would be under either account. What is clear, however, is that both the SPC approach and the *COMP approach make specific predictions about the acquisition of embedded questions. The SPC approach predicts children will never invert in embedded questions. The *COMP approach predicts that any child who does invert in embedded questions should also say sentences like (1) and (2) above. The third account (the clausal-scope-element-in-SPEC-of-CP account) is not as precisely laid out as the *COMP or SPC accounts, so its problems and predictions are not as clear. Perhaps the biggest problem with it is that in its current form it is an unmotivated stipulation. For this reason, it should be difficult for the child to learn. This difficulty might be evident in errors of inverting in embedded questions.

Raising and How Come Questions. How would the principles-and-parameters approach bar subject-auxiliary inversion in *how come* questions? One cannot say that the lexical entry for *how come* specifically

[8]The fact that it is hard to come up with a good explanation for the lack of inversion in embedded questions suggests that this restriction might be hard for the child to master.

prohibits inversion of subject and auxiliary because this would require a kind of context sensitive syntactic rule that is not allowed in the Government-Binding framework. One must show, rather, that there is something about *how come* that interacts with the rules of the grammatical sub-theories and thereby prohibits V-Raising to COMP. What principled difference is there between other *wh*-questions and *how come* questions that would prohibit V-raising in *how come* questions?

One possibility is that V-raising is blocked because there is already a pseudo-auxiliary verb in COMP (the *come* in *how come*). Support for this hypothesis can be found in the vague acceptability of questions like *how came she upon such riches?* and *how goes it?* and the auxiliary-like qualities of *come* and *go* (e.g., *she comes/goes running into the room, she's gonna eat it all* , etc.). A second possible explanation is that *how come* is a linguistic anachronism from a time when English allowed inversion of main verbs. As the AGR system in English deteriorated, main verb inversion became impossible, but its traces are still seen in *how come* questions. A third possible explanation for why *how come* questions don't allow inversion is that *how come* is a phonologically reduced version of the clausal phrase *how did it come to be*. If this were the case, *how come she can eat* would actually be the embedded question *How did it come to be that she can eat* and subject-auxiliary inversion would be prohibited by the same mechanism that prohibits inversion in embedded questions. Some support for this third theory comes from the ungrammaticality of *how come* echo questions (e.g., **she can eat how come?*). The explanation would be that *how come* cannot remain *in situ* in echo questions because it was never there to begin with

Raising in Exclamatives. I can think of several reasons why there is (optional) raising in *yes/no* exclamatives but never in *wh*-exclamatives. To me at least, there seems to be a different force to *wh*-exclamatives than *yes/no* exclamatives. If this is true, then one possible explanation for both the difference in force and invertibility of *wh*-exclamatives and *yes/no* exclamatives would be that inversion is triggered by an element in CP which has clausal scope. One would argue that the reason *wh*-exclamatives have limited scope and do not invert is that the *wh*-phrase is in a TOPIC position rather than in CP. Another less tempting possibility is that perhaps V-raising in *yes/no* exclamatives is historically just a generalization of the fact that most *yes/no* exclamatives lack an overt subject (e.g., *do your best!* not **you do your best!*) and, hence, appear to have undergone V-raising. Alternatively, *wh*-exclamatives might be the exception. Perhaps *wh*-exclamatives of any sort are a fairly peripheral part of the grammar which is learned by rote. One thing that suggests that this might be the case is the fact that only *what NP* and *how AP* exclamatives (and not *which NPs*, etc.) are permitted.

Raising and *Better*. Consider the sentence *I better go*. If *better* is the head of the VP, why doesn't it raise to COMP to form the question **Better I go?* One possible explanation is that *better* isn't actually the head of VP because an optionally deleted perfect *have* heads the VP. At the relevant level of description *I better go* is actually *I had better go*. This seems like a plausible explanation since the grammatical question is *had I better go?*, not *did I better go?* Also note that *I'd better go* actually sounds better than *I better go*. Probably *I better go* is just has a phonologically reduced form of *I had better go* which contains the perfect *had*. The

reason *better* cannot invert, therefore, is that *better* is not the head of the verb phrase, and only the head of a verb phrase can raise.

2.2 Generalized Phrase Structure Grammar

GPSG is a homogeneous grammar. It posits only one level of syntactic representation, surface structure. GPSG does not have standard phrase structure rules. Instead it has rules that constrain what an admissible node is. One crucial aspect of GPSG is how information (encoded as features) is passed between nodes of a tree. These features can either be atom-valued or they can be categories like the category AGR which refers to the agreement between subject and verb and includes such features as gender, plurality and person. GPSG makes extensive use of *immediate dominance* (ID) and *linear precedence* (LP) rules. For example, in English, VP must immediately dominate V, NP and PP and a linear precedence rule requires that the NP follow the V. Information about LP and ID requirements are read directly off the nodes of the tree.

Metarules introduce a variable (denoted by *W*) which can represent any category in an ID rule. Metarules, therefore, can capture the patterns of similarities between different sentences (e.g. between active and passive sentences and declarative and interrogative sentences) that are usually cashed out in terms of transformational rules. The difference between meta-rules and traditional transformational rules is that meta-rules take rules and generate new rules, whereas transformations map one syntactic level of representation onto another.

2.2.1 Subject-Auxiliary Inversion. GPSG introduces a 'Subject-Auxiliary Inversion' Metarule to explain the existence of aux-initial structures. This metarule (shown below) captures the observation that what can follow the subject NP in an aux-initial sentence is that which can be the complement of the auxiliary in a VP.

'Subject-Auxiliary Inversion' (SAI) Metarule

V ^{2}[-SUBJ]----->*W*

V ^{2}[+INV, +SUBJ]----->*W*, NP

The SAI Metarule takes any VP ID-rule and turns it into an S (+SUBJ) ID-rule by adding the subject NP and the INV feature which indicates the VP occurs before the subject. Unless the +INV feature is explicitly represented, there is no inversion because the first Feature Specification Default (FSD 1) states that the unmarked or default value is for no inversion. The optional +INV feature is lacking in embedded and *how come* questions so there is no inversion. Ungrammatical questions like **ran she fast?* are blocked because of a Feature Coocurrence Restriction, which requires that can +INV only occur on +HEAD nodes which have the +AUX feature. Invoking +INV feature can also explain inversion in negative polarity sentences; in sentences like *never have I seen such a mess* and *I do not like lima beans, nor do I like liver* , the lexical entries for *nor* , *never* and *not only* have the +INV feature. It can also explain why some auxiliaries can only occur in inverted position. *Aren't I happy?* is grammatical but not **I aren't happy* because the lexical entry for *aren't* only allows *aren't* in singular sentences when the +INV feature is present.

2.2.2 Acquisitional Issues

Generalization. GPSG rules are essentially lexical, or word-specific. Therefore, there should be no generalization from one structure to another. A lexical theory like GPSG would predict that acquisition of correct auxiliary placement should be no faster or accurate for auxiliaries that are acquired later than for the first auxiliaries learned.

Subject-Auxiliary Placement. Because the +INV feature is added to lexical items one by one, what inverts is, in a sense, arbitrary. Embedded questions,[9] *how come* questions, *wh*-exclamatives, etc., aren't inverted because they lack the +INV feature. No explanation is given for why some lack +INV and others are +INV. In a sense, then, a child trying to learn a GPSG-like grammar is exactly in the position of trying to figure out which of the possible combinations of factors determine whether something inverts. In fact, in some ways, the child is in a worse situation than the hypothetical situation described above in which a child is trying to figure out which of the 7! possible structures allow inversion. There, at least the factors to be considered were spelled out. By its very nature GPSG has no such overriding factors. There is nothing in principle within the GPSG framework that would prevent a child from having to determine one by one whether each verb inverts. The GPSG account is frustrating because the very flexibility which allows it to stipulate what inverts and what doesn't invert, allows it to explain any pattern of errors that a child

[9]Actually it is the embedding verbs which lack the +INV feature.

might make. A child who inverts embedded questions has merely incorrectly added a +INV feature to an embedding verb. A child who inverts a main verb has merely incorrectly added +INV to that main verb.

In a sense what the GPSG account would predict would be that unless the child is painfully conservative, saying only what she has heard, she will make many errors such as inverting main verbs, etc. Since we showed in chapter 3 that children do not make such errors, then GPSG would have to argue for such extreme conservativism. What GPSG would predict, therefore, is that inversion would be acquired very slowly on a verb-by-verb basis. Furthermore, if a child does make errors such as inverting in embedded questions, her errors should be restricted to a particular embedding verb to which she has incorrectly appended the +INV feature.

2.3 Lexical Functional Grammar

LFG takes subject and object to be primitive grammatical functions. The structures that represent these functions are called functional structures or f-structures. F-structures state invariant grammatical constraints on anaphora and agreement and are thought to be fairly constant across languages. F-structures are only well-formed if (1) every f-structure is a true function (the uniqueness condition); (2) all subcategorization requirements are met (the completeness condition); and (3) all arguments are arguments of a predicate (the coherence condition). LFG also has constituent structures (C-structures) that are roughly analogous to Phonological Form in Government-Binding and surface structure in GPSG. C-structures have things like NPs and VPs and express word order and phrasal structure. C-structures exhibit considerable cross-linguistic

variation. The entries in the lexicon are the bridge between f-structures and s-structures, because they provide information about grammatical function, subcategorization, thematic structure, case, tense, person, and number.

2.3.1 Subject-Auxiliary Inversion. While Bresnan (1982) offers no account of auxiliary placement, Pinker (1984) provides an LFG account that draws heavily on the lexicalist proposals of Baker (1981) and Gazdar, Pullum, and Sag (1982). It's not surprising, therefore, that there are few differences between the Pinker's LFG account and the GPSG account. Like Baker and Gazdar *et al.*, Pinker treats auxiliaries as complement-taking main verbs that are subcategorized for complements with specific grammatical features. These subcategorization restrictions determine which auxiliaries can occur where. The lexical entries of auxiliaries determine what features their verb complements must have. (The Consistency condition assures that complements which do not match the requirements of an auxiliary are eliminated.)

Pinker (1984) assumes that all verbs (including auxiliaries) form two-dimensional "verb paradigms" of verb form (infinitival, participial, and finite) and modality (neutral, inverted, negated, and emphatic). Children learn the paradigms of auxiliaries conservatively; the empty cells of the auxiliary paradigm can only be filled when the auxiliary is heard in the particular verb form in a particular modality (Pinker, 1984).

Pinker deals with the problem of inversion by including in the lexical entries of *wh*-phrases and auxiliaries the constraining equation that sentence modality must be either +/-not inverted. *How come* questions, *wh*-exclamatives and uninvertible auxiliaries are not inverted because they can

only occur in +noninverted sentences. Inversion in embedded questions is ruled out by appending a +/- ROOT CLAUSE feature to every clause. The phrase structure rule which creates sentences with subject and auxiliary in 'inverted' positions can only occur in the presence of of the +ROOT CLAUSE value.

2.2.2 Acquisitional Issues

Generalization and Inversion. Like GPSG, LFG is a lexical theory and, therefore, LFG predicts that acquisition of correct auxiliary placement should be no faster or accurate for auxiliaries that are acquired later than for the first auxiliaries learned. LFG allows inversion by placing a constraining equation that requires that the complements of certain structures not have noninverted sentence modality. GPSG allows inversion by attaching a +INV feature to some structures and assuming a default of noninversion. One technical difference is that LFG prevents inversion in embedded questions by requiring that inversion only occur in matrix clauses, whereas GPSG accomplishes this by requiring that the complements of embedding verbs not have inversion. Thus, LFG should predict absolutely no errors of inverting in embedded questions, whereas GPSG might predict occasional, verb-specific inverted embedded questions. But the real difference between LFG and GPSG is that LFG tries to provide a principled reason for the lexical rules and restrictions it invokes. GPSG does not. Hence, the child learning a Pinker-LFG grammar is not in quite as desperate a position as the GPSG-learner. In LFG the dimensions which affect invertibility are spelled out (in the two-

dimensional verb paradigms), whereas in GPSG they are not.[10] Furthermore, Pinker argues that the factors which can determine invertibility are probably universal to all languages.

[10] Because it lacks any real basic categories, it's not clear to me that GPSG could provide such guidelines. The meta-rules which substitute for categories are just descriptive generalizations. There seems to be no limit on what can be a meta-rule (i.e., what verbal elements can be grouped together).

3. Acquisitional Predictions and Evidence

One of the important difference between the way subject-auxiliary inversion is dealt with in the principles-and-parameters approach of GB and the way it is dealt with in lexical theories such as GPSG and LFG is that in GB a transformation moves the auxiliary to clause-initial position whereas in the lexical theories, phrase structure rules directly generate the auxiliary in its "inverted" position. If auxiliaries are moved via a transformation, one might expect that occasionally when the child is acquiring English, he will fail to perform any transformation and produce sentences with the to-be-moved element in its original position. One might even predict that the child would go through a **stage** during which he consistently fails to perform the transformation for any auxiliary or for a certain type of auxiliary. In GPSG, one might expect that the child would occasionally fail to append a +INVERT feature to the lexical entry for a particular auxiliary and, therefore, he would produce questions with that auxiliary in noninverted position. In LFG, two possible types of failures could result in inversion errors. Like GPSG, a particular auxiliary might fail to get the +INVERT feature and, therefore, it would be ineligible to appear in inverted positions. The second possible cause would be if the child did not have the constraining equation which requires that all auxiliaries and *wh*-phrases be specified for +/- INVERT. In such a case, children would probably perform randomly on inversion, inverting all auxiliaries about half the time.

3.1 Previous Studies of Inversion Errors

How common are inversion errors or auxiliary *in situ* errors? Klima and Bellugi first reported auxiliary *in situ* errors in 1966. Inversion errors continue to be the most common type of English word order error reported or studied in the acquisitional literature. (Bellugi, 1971; Derwing and Smyth, 1988; Erreich, 1984; Ingram and Tyack, 1979; Klee, 1981, 1985; Klima and Bellugi (1966); Kuczaj and Brannick, 1979; Kuczaj and Maratsos, 1975, 1983; Labov and Labov, 1978; and Nakayama 1987). Despite the consensus that inversion errors do occur, there is considerable debate about the details of the phenomenon. In Table 5.2, I summarize and compare the results of 9 studies on subject-auxiliary inversion. To the best of my knowledge, these 9 studies represent the only published studies that address any of the following questions for "normal" children learning English: 1) do children go through a "stage" during which they fail to invert; 2) do children only fail to invert in *wh*-questions, or do they also fail to invert in *yes/no* questions; 3) do children invert for some *wh*-words and not others; and 3) do children fail to invert more often for some auxiliaries (or copulas) than others.[11]

Insert Table 5.2 here

---

[11] A disclaimer: The summaries I present in Table 5.2 do not necessarily represent the conclusions made by the authors. In a number of cases, the authors did not address the questions I put forth above. Whenever possible, I used what data they did present to answer the questions. Because I did not have the original data on which to perform the appropriate statistical tests, all of the observations and conclusions should be considered trends which are not necessarily significant.

Being able to compare the results of nine studies is a luxury not often available in syntax acquisition. The comparison of the results of these nine studies suggests that it is a luxury we may not want to do without in the study of syntax acquisition. It is not an exaggeration to say that most of the possible outcomes are represented in the data from these nine studies. If we just concern ourselves with the answer to the first question--is there are "failure-to-invert stage"--we find that seven of the nine studies suggest that no such stage exists. Six of the nine studies address the second question of whether failure-to-invert is specific to *wh*-questions. Two studies (Bellugi, 1971 and Kuczaj and Maratsos, 1975) found that failure to invert was restricted to *wh*-questions. Two studies (Derwing and Smyth, 1988; Klee, 1985) found failures-of-inversion to be just as common in *yes/no* questions as in *wh*-questions. Finally, two studies (Erreich, 1984; Ingram and Tyack, 1979) found failures-of-inversion to be more common in *yes/no* questions than *wh*-questions. Four of the nine studies considered the effect of *wh*-word choice on inversion. In all four studies, different *wh*-words had different rates of failure to invert. In general, children inverted the most for *what* and *where* and the least for *why* and *when*.

Only Kuczaj and Maratsos's study (1983) specifically addresses the fourth question of whether auxiliary choice affects frequency of inversion. In this study, Kuczaj and Maratsos found that children began inverting most auxiliaries at the same age. However, they found that *could* and *would* appeared in declarative contexts for some time before they began to appear in inverted position in questions. The results of two other studies also suggest that some auxiliaries may be more invertible than others. Labov and Labov (1978) noted that their daughter inverted *be*, *can*, and *should* less than other auxiliaries. In addition, examination of Klee's (1985)

results reveals that his subjects inverted copula *be* less frequently than any of the auxiliaries. The auxiliary choice effect that is most consistently found across the studies is that children are less likely to invert auxiliaries and copulas which are combined with the contracted negation element *-nt* than they are non-negated auxiliaries and copulas. This finding is explicitly mentioned in studies by Bellugi (1971), Derwing and Smyth (1988), Erreich (1984)and Labov and Labov (1978).

So here we have nine studies all looking at the same phenomenon, all getting slightly different results. This is undoubtably due in part to differences in the methodologies employed in the studies. Using embedded questions to elicit imitations or productions could likely lead to more failures to invert, since the eliciting sentence has a noninverted auxiliary. If an experimenter tries to elicit questions using sentences like "Ask the boy why he can eat" it would not be surprising if occasionally the child repeats back the second half of the sentence and produces incorrect noninverted questions like **what he can eat?*. Imagine what happens if the questions to be imitated are elicited by asking the child "can you say" or "can you repeat". The child could interpret such sentences as incorrect embedded questions (e.g., **can you say what can he eat?*) and attempt to repair them.

Analyses of children's spontaneous speech do not fall prey to the above problems. A longitudinal study of a single child is invaluable but it can be misleading because it only provides information about what a single child does. While it is true that within a nativist tradition, most acquisitional options are not available to the child, children can still vary widely in the extent to which they exploit the various options which are available. If the child being studied makes some type of error very frequently, it is more likely to be noticed and reported than the other

errors the child makes less frequently. Thus, the concern isn't so much that the child being studied will make an error that no other child would ever make. The problem is that the child may make many different types of errors but the only one that will be noticed and reported is the one that (for whatever reason) the child makes often. This is less likely to be a problem in cross-sectional studies in which one can study a large number of children. The main problem with cross-sectional studies, however, is that by only studying a child at one or two ages, it is possible to miss a transient stage in linguistic development. The best methodology would, of course, be a longitudinal study of a large number of children.

Let's put aside for a moment the differences and limitations of the different methodologies and assume that the rates of failure-to-invert reported in each study are accurate. How much variation is there from study to study? The answer is: a lot. Of the studies that report the failure-to-invert rate or provide the information necessary to calculate this rate, failure-to-invert rates ranged from a low of one to two percent in Klee's (1985) study to a high of 100% inversion failures for *wh*-questions in Bellugi's study of Adam (1971).[12] All of the studies report at least some inversion errors, however. If all of the reported inversion errors represent real failures of grammar and not just anomalous responses, we must conclude that something about some children's grammars allow them to fail to invert. Why the rate of failure varies so much from child to child will need to be addressed.

---

12 Please note that Bellugi's figure appears to be incorrect. She only looked at a subset of Adam's questions when she arrived at this figure. Judging from the examples she cites in her unpublished dissertation, Bellugi happened to sample a period when Adam was asking many *why* questions with the negated auxiliaries *can't* and *won't*. Of all possible combinations of *wh*-words and auxiliaries and copulas (including negated ones), Adam had the lowest rates of inversion for *why ...can't* and *why ....won't*.

I would now like to examine more carefully the inversion errors made by the 14 children whose transcripts I examined in chapter 3. Even with longitudinal data from 14 children, I cannot hope to consider all of the possible conditions that affect inversion. I will not, for example, consider production-related effects such as the relationship between NP length and failure to invert described by Crain and Nakayama (Crain and Nakayama, 1987; Nakayama, 1987). In addition, because I am examining written transcripts, I cannot examine the effects of prosody on inversion. What I would like to focus on is the four general questions raised by previous acquisitional research and by the predictions made by the three linguistic theories described above. The four general questions I will address are: 1.) is there a difference in the inversion rates for *yes/no*- versus *wh*-questions?; 2.) does *wh*-choice affect inversion rates?; 3.) does verb choice affect inversion rates?; 4.) are negated auxiliaries inverted less frequently than their affirmative counterparts? I would also like to address a fifth question: do children ever invert when they shouldn't?

1. Inversion rates in *yes/no* and *wh*-questions. This question is important to settle because the difference in inversion rates for yes/no and *wh*-questions is one of the most frequently cited findings in the acquisitional literature, despite the inconclusiveness of the results. This question has important linguistic implications because it suggests a distinction in type between *yes/no* and *wh*-questions. More specifically, within current principles-and-parameters work, a higher rate of inversion in *yes/no* questions could be taken as support for the suggestion made by

Chomsky (1989) that the reason there is inversion is to be able to distinguish questions from declaratives at S-structure (see above).

2. Inversion rates for different *wh*-questions. LFG and GPSG predict that inversion will be learned piece-meal *wh*-word-by-*wh*-word. The principles-and-parameters approach predicts that as long as the *wh*-word is definitely in the SPEC position of CP, the choice of *wh*-word should not affect inversion. Within the principles-and-parameters framework, therefore, any difference in inversion rates for a particular *wh*-word suggests something either about the location of the *wh*-word (in SPEC of CP vs. in a TOPIC position) or the nature of the *wh*-word (perhaps that *wh*-word does not have clausal scope).

3. Inversion rates for different auxiliaries. LFG and GPSG predict that inversion will be acquired auxiliary-by-auxiliary as the child appends the +INVERT feature to each auxiliary. Hence, differences in inversion rates for different auxiliaries are to be expected within LFG and GPSG frameworks and no particular pattern of differences is predicted to occur. Within the principles-and-parameters framework, any difference in inversion rates for the various auxiliaries must reflect important differences in the syntactic nature or behavior of those auxiliaries (see above). If the inversion rates vary for different auxiliaries, these differences should pattern along the natural divisions made within Government-Binding theory (e.g., the inversion rates for all the modals should same, etc).

4. Inversion rates for negated and non-negated auxiliaries. GPSG and LFG would not predict any systematic difference in the inversion rates for negated or non-negated auxiliaries. Whether a particular auxiliary is inverted depend on whether that auxiliary has the +INVERT feature yet. There is no reason to expect this to systematically occur later for negated auxiliaries than non-negated auxiliaries. Within the principles-and-parameters approach, one might predict that the inversion rate would be lower for negated auxiliaries than non-negated auxiliaries. This is because in order for the negation marker *-n't* to cliticize to the auxiliary or copula, the auxiliary or copula has to move to NegP. Modals differ in that either the modal has to move down to NegP or the negation marker has to move up to I. Thus, not only might the principles-and-parameters approach predict a systematic difference in inversion rates for negated auxiliaries as compared to non-negated auxiliaries, it might predict that the difference would be greater for modals than for the other auxiliary or copula elements.

Other questions. In addition to the 4 major questions outlined above, I will also examine inversion in embedded questions, negative polarity sentences and in *yes/no* and *wh*-exclamatives. I will examine more closely the children's errors of inverting in embedded questions (i.e., errors like **I wonder what can I eat*) for three reasons. First, one way GPSG and LFG differ is that LFG prohibits such errors from ever occurring whereas GPSG merely predicts that they will be specific to certain embedding verbs. Second, what predictions the principles-and-parameters approach makes depends on what is thought to prevent inversion in embedded questions (see above). Thus, while Government-

Binding theory doesn't make any firm predictions about inversion in embedded questions, the results may help decide between the various options I outline above.

The third reason for examining inversion in embedded questions is that the semantic/pragmatic explanation of inversion offered by Grimshaw (1979) makes specific predictions about inversion errors. Grimshaw (1979) suggests that modality of sentence is the only factor that determines whether there is subject-auxiliary inversion. She argues that subjects and auxiliaries can only invert in non-declarative sentences. A few examples, however, make it clear that Grimshaw's semantic/pragmatic guideline can't be the only thing that is operating to determine whether there is inversion. Consider for example a *wh*-question embedded within a matrix *yes/no* question. *Do you know who she is?* does not have inversion, yet it requires the same answer that the matrix question *who is she?* requires. In addition, *how come* questions and *wh*-questions echoing a statement never have subject-auxiliary inversion yet they certainly are requests for information.

But Grimshaw's distinction is a good first approximation of the inversion phenomena in English. Perhaps children initially distinguish declarative from nondeclarative sentences by prosodic and contextual information and use semantics/pragmatics to decide when to invert. If children do this, then they should place the auxiliary before the subject in embedded *wh*-questions that are introduced by interrogative verbs (e.g., *wonder, like to know,* etc.), or which are embedded within a *yes/no* question (e.g., *have you seen what he has written?*). Children would also invert in *how come* and echo questions.[13] I checked children's embedded

[13] Assuming that echo questions are basically the same as standard questions. See Janda (1985) for an alternative analysis of echo questions.

questions to determine whether they are more likely to invert embedded questions that are embedded in an interrogative matrix question.

I also looked at inversion in two other constructions which, though rare, are critical test cases for the various theories. The first type of construction was embedded negative polarity sentences. Only Grimshaw's pragmatic explanation predicts they will be inverted in the adult grammar, so only Grimshaw's theory predicts children will invert them. The second type of construction I looked at was exclamatives that contain auxiliaries. Do children distinguish *wh*-exclamatives from *yes/no* exclamatives and only invert in the latter type? Here, it seems like Grimshaw's pragmatic/semantic account would predict that children would invert all exclamatives. GPSG and LFG would not predict inversion of *wh*-exclamatives. As was the case with embedded questions, what a principles-and-parameters approach predicts depends on the principle or parameter used to bl ɔck inversion in *wh*-exclamatives. The acquisitional result could influence which choice Government-Binding adopts.

In the following section, I examine the inversion errors made by 14 children and try to answer a number of related questions about how children know when to invert subject and auxiliary. Are children conservative about when they invert or do they make generalizations beyond what they have heard? If children do generalize, what kinds of generalization do they make? Are they semantic/pragmatic or are they syntactic?

4. Transcript Analyses

4.1. Subjects

I used the same transcripts from the same 14 children that I used in chapter 3. For Ross, however, I only included those inversion errors made through Transcript 60. I did this because by Transcript 60, he was already 5;5. I felt that including the later transcripts which were recorded after he presumably had already acquired adult competence in inversion would only serve to dilute any evidence of an early "child" stage.

4.2 Corpora

***Wh*-Questions.** The corpus of *wh*-questions was collected in the following way. I used the Unix utility "fgrep" to cull all of the children's lines which the strings "how", "what", "when", "where", "which", "who", and "why" and put these lines into a new file. Because I was interested in auxiliary placement, I then used "fgrep" again to pull out those lines in the *wh*-word file which contained an auxiliary.[14]

***Yes/No* Questions.** The corpus of *yes/no* questions was collected by using "fgrep" to cull those lines which contained a question mark ("?")

---

[14] This second fgrep program searched for the following letter strings: *am*, *are*, *be*, *better*, *can*, *could*, *do*, *did*, *gonna*, *gotta*, *had*, *has*, *have*, *is*, *may*, *might*, *must*, *need*, *never*, *no*, *not*, *n't*, *shall*, *should*, *was*, *were*, *will*, *won't*, and *would*.

or the embedding *yes/no* complementizers *if* or *whether*.[15] *Wh*-questions were eliminated by eliminating all lines that contained a *wh*-word.[16] Because I was interested in auxiliary placement, I used "fgrep" to pull out only those lines in the *yes/no* file which contained an auxiliary.

Exclamative and Negative Polarity. In addition to searching for questions and declaratives with auxiliaries, I searched for *yes/no* and *wh*-exclamatives by searching for exclamation marks (!). I searched for negative polarity sentences by searching for *never*, *not one*, and *nor*.

4.3 Tallying Procedure

Questions. For every question that contained an auxiliary and a subject NP and was neither a routine nor a stuttered or unclear utterance (i.e., for every scoreable question), I determined whether the subject of the sentence occurred before or after the auxiliary. I kept separate tallies for:

-- each of the 14 children

--each transcript for each child (from 6 to 139 transcripts per child)

---

[15]Some yes/no questions were not found by this search. For example, some of the transcripts (e.g. Nathan's and Peter's transcripts) used a "?" inconsistently. When I was going through the +auxiliary declarative sentences for the analyses reported in Chapter 3, when I found utterances which were clearly questions (e.g., if the parent answered the child or if there was notations like "question intonation"), I included them in the tallies. I did not include utterances that were marked with a period (.), even if these utterances had subject-auxiliary inversion. The reason I chose not to include anything followed by a period is that if I used inversion as a criterion for question-hood, I would have inflated the inversion rates.

[16] Notice that yes/no questions that were on the same line as *wh*- questions were eliminated by this search procedure. If while going through the *wh*-corpus by hand, I noticed yes/no questions that contained auxiliaries, I added them to the yes/no corpus.

--each of the 9 question types (*yes/no*, *how*, *how come*, *what*, *when*, *where*, *which*, *who*, *why*)

--each of the two clause types (matrix or embedded)

--each of the 25 auxiliaries and copulas

In addition, for each of the 22 negated auxiliaries and copulas, I kept separate tallies for each transcript of each child. Because there were relatively few negated auxiliaries, I kept one tally for all 9 question types. Table 5.3 below shows how many questions containing an auxiliary and a subject NP each of the 14 children asked. Note that two children asked no such questions. Notice also that for the 12 children who asked scoreable questions, the number of questions asked ranged from 38 (Allison) to 1,749 (Adam).[17] The analyses that follow are based on 6,370 scoreable questions asked by 12 children.

Please Insert Table 5.3

Additional tallies. In addition to the question tallies, I recorded all examples of *wh-* and *yes/no* exclamatives and negative polarity sentences that had auxiliaries.

[17] As was the case in the Age of First Use Analyses presented in Chapter 4, in general, children who were older and had more speech recorded when they were older, had more scoreable data.

4.4 Results

Because of the large quantity of data and possible analyses, I will restrict my analyses to those that address the questions I raised earlier. For the sake of organization and clarity, I will address each question in turn and, wherever possible, supplement the text with graphs and tables that enable the reader to get an overall sense of the results. In the appendices, I provide more detailed results for the reader who is interested in specific numbers and breakdowns for certain constructions.

4.4.1 Overall inversion rates and individual differences.

Before addressing whether inversion rates are the same for *yes/no* and *wh*-questions, I would like to comment on the overall frequency of inversion errors.[18] If we consider questions with both negated and non-negated auxiliaries, overall children correctly inverted 91.5% of the time. The variation in inversion rates for the children was statistically significant significant ($\chi^2(11) = 607.9, p < .0005$), with rates ranging from a low of 50.6% for Nathan to a high of 97.4% for Allison, with a median inversion rate of about 94%. (Please See Table 5.4). If, however, we only consider questions with non-negated auxiliaries, then overall children correctly inverted 93.4% with individual inversion rates ranging from 54% (Nathan) to 98% (Ross) with a median inversion rate of about 95%. Thus, overall, failing to invert subject and auxiliary in matrix questions was neither

[18]Please note that throughout this chapter, when I give the rate of inversion, I am giving percentages for the number of correctly inverted questions divided by the number of scoreable questions. Unless I specify otherwise, all analyses and percentages are for all scoreable matrix questions--excluding *how come* questions--that had non-negated auxiliaries.

pervasive nor exceptionally rare. Some children made the error more than other children.

Figures 5.2 through 5.13 depict the developmental trend for inversion for each of the children. The Y-axis is the percent of scoreable questions that are inverted and the X-axis is age. Notice that Adam (Figure 5.2), Eve (Figure 5.5), Nathan (Figure 5.8), Nina (Figure 5.9), and Sarah (Figure 5.12) experienced a dip in inversion performance , while Allison (Figure 5.3), April (Figure 5.4), Mark (Figure 5.6), Naomi (Figure 5.7), Peter (Figure 5.10), Ross (Figure 5.11), and Shem (Figure 5.13) either inverted at a constant rate as they got older or improved slightly with age. In general, the children with the lower overall rates of inversion were the children who experienced a dip in performance.

Please Insert Table 5.4
and
Figures 5.2 through 5.13

4.4.2 Inversion rates in yes/no and *wh*-questions. If we compare the inversion rates for *yes/no* and *wh*-questions that have non-negated auxiliaries, [19] we find that overall the children inverted 93.0% of all*wh*-questions and 93.7% of all *yes/no* questions. This difference was not significant ($F(1, 22) < 1$). Table 5.5 shows that three of the children had a higher inversion rate for *yes/no* questions than *wh*-questions, six children

[19]Again, unless otherwise stated, all comparisons are for questions with non-negated auxiliaries. This is necessary because the number of questions with negated auxiliaries was small enough that I did not distinguish between usages in yes/no and *wh*-questions, let alone differences between the different *wh*-questions or between different children.

had virtually identical inversion rates for both types of questions, and four children had higher inversion rates for *wh*-questions than *yes/no* questions.

One could argue that failing to invert in *wh*-questions is a relatively early stage and, therefore, that only children who have just begun to invert will have a higher inversion rate for *yes/no* than *wh*-questions. If we assume overall inversion rate is an indicator of how completely children have mastered inversion, we can use the overall inversion rates shown in Table 5.4 to rank the children according to level of mastery. If we do this, there doesn't seem to be a discernable relation between level of mastery of inversion (as indicated by overall inversion rate) and whether a child inverted more in *yes/no* questions than *wh*-questions. Of the 6 children who had below median overall inversion rates, one child inverted more frequently in *yes/no* questions, two children inverted equally often in both types of questions, and three children inverted more often in *wh*-questions. For the six children with better than median inversion rates, two inverted more frequently in *yes/no* questions, three inverted equally often in both types, and one inverted more often in *wh*-questions. (Please see Table 5.5.)

Please insert Table 5.5

Perhaps there was a difference between *yes/no* question and *wh*-question inversion rates but performance on certain types of *wh*-questions obscured the overall *yes/no-wh* question difference. Figure 5.14 below suggests this was not the case. Half of the *wh*-words had inversion rates that are greater and half of the *wh*-words had inversion rates that are lower than the inversion rate for *yes/no* questions.

Please insert Figures 5.14

With the exception of Nathan, all of the children inverted 80% or more of their *yes/no* questions. Figure 5.16 depicts the overall developmental trend in inversion rate for *yes/no* questions. Note that the inversion rate is in the 90% to 100% range at all ages.

Please insert Figures 5.15 and 5.16

4.4.3 Inversion rates for different *wh*-questions.

Examination of Figure 5.14 leads to the question of whether the inversion rates for all *wh*-questions are the same. Overall, the children made more inversion errors in *when*, *which*, and *why* questions. Table 5.6 gives the percent inversion for each type of question for each child. It also gives the number of questions that were used to calculate these percentages. It's important to note that the children didn't ask the same number of scoreable questions for each type of question. Overall, the children asked considerably fewer scoreable *when* and *which* questions than any of the other types of *wh*-questions.

Nine children asked at least four of the seven types of *wh*-questions. Three of these nine children failed to ask one type of question. For each of these three children, I substituted that child's average *wh*-question inversion rate for the missing value. I did an analysis of variance on the

resulting data for these 9 children and found that there was a large and significant main effect for type of *wh*-question ($F(6, 48) = 3.788$, $p = .004$, eta = .57, $f = .69$). For these 9 children, the mean inversion rates ranged from a high of 100% for *who* questions to a low of 73.2% for *when* questions (mean inversion rates: *how* = 98.5%, *what* = 91.9%, *when* = 73.2%, *where* = 94.6%, *which* = 89.5%, *who* = 100%, *why* = 80.0%).

Please Insert Table 5.6 here

Figures 5.17 through 5.23 show the inversion rates for each type of *wh*-question for each child. Notice that all of the children inverted *who* (Figure 5.17) and *how* (Figure 5.18) approximately 100% of the time. There was increasingly more variation between children for *where* (Figure 5.19), *what* (Figure 5.20), and *which* (Figure 5.21) questions, respectively. The between-child variation was particularly striking for the *when* (Figure 5.22) and *why* (Figure 5.23) questions. If the *wh*-questions are ranked according to overall inversion rate and also according to between-subjects variation in inversion rate (as measured by the standard deviation), there was a near perfect match in the inversion rate ranks and the variation rate ranks. (*When* questions ranked sixth in inversion rate and fifth in variation, and *why* questions ranked fifth in inversion rate and sixth in variation rate.) What this suggests is that the amount of individual variation may reflect a ceiling effect.

Please Insert Figures 5.17 through 5.23

The inversion rates for the *wh*-questions that had the highest rates of inversion (the *who*, *how*, *where*, and *what* questions) are fairly constant over time. For *which*, *when*, and *why* the inversion rates are fairly labile, dipping down from perfect inversion to below 60% in each case.

Please Insert Figures 5.24 through 5.30

4.4.4 Inversion rates for different auxiliaries. Just as children do not invert equally often for all types of *wh*-questions, children are more likely to invert some auxiliaries than others. Collapsing across children, question-type, and across auxiliary and copula forms of *be*, the overall inversion rates ranged from a high of 100% for *am*, *does*, *has*, and *have* to a low of 0% inversion for *might* (0/3 questions inverted). Tables 5.7 and 5.8 list the inversion rates for each auxiliary in each type of question. These overall inversion rates, broken down by *yes/no* and *wh*-question, are shown graphically in Figures 5.31 through 5.34.

Please Insert Tables 5.7 and 5.8
and
Figures 5.31 through 5.34

Because of the number of empty cells, it wasn't possible to do an analysis of variance of the effect of auxiliary choice using each of the 20 auxiliaries as a level. Because of the results of the Age of First Use

analyses indicate that the auxiliaries form natural clusters, I decided to compare the inversion rates for the acquisitional auxiliary clusters. Overall, children inverted in 87.9% of *be* questions, 98.9% of *do* questions, 90.9% of *have* questions and 91.8% of modal questions. These inversion rates, broken down by *yes/no* and *wh*-question, are shown graphically in Figures 5.35

Please insert Figure 5.35

In order to determine whether the differences shown in Figure 5.35 were significant, I did an analysis of variance for the mean inversion rates for each of the following clusters of auxiliaries: *be* forms (auxiliary and copula forms combined), *do* forms, and modals.[20] This analysis revealed a large, significant main effect for Auxiliary Type ($F(2, 22) = 7.72$, $p = .003$, eta = .64, $f = .84$).

Recall that in chapter 4, I found that the age of acquisition of the various constructions differed for auxiliary and copula versions of *be*. For this reason, I decided to compare the inversion rates for copula and auxiliary *be*. Eleven of the 12 children asked at least one scoreable question that had an auxiliary *be* and one question that had a copula *be*. For the first analysis of variance, for each of the 11 children, I determined the overall rate of inversion for copula and auxiliary *be*. Analysis of variance of these overall rates of inversion revealed no significant main

[20] I did not include *have* forms because only 4 of the 12 children used *have* forms in scoreable questions.

effect for Auxiliary/Copula ($F(1,10) = 2.58$, $p = .142$). The average inversion rate was 88.4% for auxiliary *be* questions and 86.0% for copula *be* questions.

I was concerned that differences in inversion rates for *am*, *are*, *is*, *was*, and *were* might be obscuring a significant difference between copula and main verb be. For this reason, I did an analysis of variance with the factors Verb Form and Auxiliary/Copula. Because 6 of the 12 children lacked either copula or auxiliary *am* questions, 6 out of 12 lacked either copula or auxiliary *was* questions, and 10 out of 12 lacked either copula or auxiliary *were* questions, I decided to only include data for auxiliary and copula *is* and *are* questions. Of the 12 children, 10 asked both auxiliary and copula *is* and *are* questions. For these 10 children, I determined each child's inversion rates for copula and auxiliary forms of *is*, and *are*. The analysis of variance revealed a marginally significant main effect of Is/Are ($F(1, 9) = 3.83$, p =.082, eta = .54, $f = .64$). On average, children inverted 84.4% of *are* questions and 90.0% of *is* questions. More importantly, the Auxiliary/Copula factor just missed being significant at the $p < .05$ level ($F(1, 0) = 5.02$, $p = .052$, eta = .60, $f = .75$). On average, children inverted 90.2% of auxiliary *be* questions and 84.2% of copula *be* questions. The interaction between Is/Are and Auxiliary/Copula was not even marginally significant ($F(1, 9) < 1$).

4.4.5 Inversion rates for negated and non-negated auxiliaries. Having considered the inversion rates for the various groups of auxiliaries at length, I would like to turn to the question of whether the children I studied inverted negated and non-negated auxiliaries equally often. Averaging across all the data from all 12 children, the inversion rate

for non-negated auxiliaries was 93.4% while the inversion rate was 71.3% for negated auxiliaries. The difference between negated and non-negated inversion rates was greatest for the modals (inversion rate for non-negated modals 40 percentage point higher than for negated modals), intermediate for the *do* forms (inversion rate for non-negated *do* 30 percentage points higher than for negated *do*), and least for the *be* forms (inversion rate for non-negated *be* 5 percentage points lower than for negated *be*). The exact percentages and the numbers used to calculate them are shown in Table 5.9.

Please Insert Table 5.9

Is the difference in inversion rates between negated and non-negated auxiliaries significant? Eleven of the 12 children asked scoreable negated auxiliary question. (Allison asked none.) For each of these 11 children, I determined the inversion rates for negated and non-negated questions. An analysis of variance of these rates revealed a large significant effect of negation ($F(1, 10) = 17.28, p = .002$, eta = .80, $f = 1.31$). Overall, these 11 children inverted non-negated auxiliaries 90.7% of the time and negated auxiliaries 55.6% of the time.

I next wanted to determine whether the negation effect was significant for all types of auxiliaries. There weren't enough data to do the analysis with each individual auxiliary as a level. In addition, only 7 children had enough data to be analyzed by an analysis of variance with the factors Auxiliary Group and Negation. This analysis revealed a significant main effect of Negation ($F(1, 6) = 15.74, p = .007$, eta = .85, $f = 1.62$) and a marginally significant interaction between Negation and Auxiliary Group

($F(2, 12) = 3.09$, $p = .082$, eta = .22, $f = .22$). Figures 5.36 graphically depicts this interaction for all of the children (not just the 7 children whose data were included in the analysis of variance). In addition, Figures 5.37 through 5.48 shows each child's inversion rate for each type of negated and non-negated auxiliary.

Please Insert Figures 5.36 through 5.48

I decided to further investigate the marginal interaction between Negation and Auxiliary Group by analyzing each group of auxiliaries separately. Eight children asked scoreable questions with both negated and non-negated forms of *be*. [21] Analysis of these children's *be* data revealed not even a marginally significant main effect for Negation ($F(1,7) < 1$). The inversion rate was 92.6% for non-negated *be*, and 80.8% for negated *be*.[22] Ten children asked scoreable questions with both negated and non-negated forms of *do*. Analysis of these children's *do* data revealed a large significant main effect for Negation ($F(1,9) = 20.44$, $p < .0005$, eta = .83, $f = 1.51$). Children inverted non-negated *do* 98.7% of the time and negated

---

[21] I combined auxiliary and copula forms together because only 5 children asked scoreable questions with both negated and non-negated versions of both auxiliary and copula *be*.

[22] The overall inversion rates for all 12 children were 90% for non-negated *be* and 95% for negated *be* . However, the average inversion rates for the 8 children who asked both negated and non-negated *be* questions were 93% for non-negated *be* and 81% for negated *be*. The main cause of this "discrepancy" was the methods used to calculate overall inversion rate and average inversion rate, not differences in inversion behavior of the 8 children versus the 12 children. As you can see from the formulas given below, with the Overall Inversion Rate, the anomalous data of a very talkative child can swamp the data of a majority of taciturn children. In Average Inversion Rate, all children's data get equal weighting.

Overall inversion rate = total # inverted for 12 children/Total # for all 12 children

Avg inversion rate = (inversion rate for child 1 + inversion rate for child 2 + ... inversion rate for child 8)/8).

do 59.7% of the time. Ten children also asked scoreable questions with both negated and non-negated modals. Analysis of these children's modal data revealed a large significant main effect for Negation ($F(1,9) = 30.87$, $p < .0005$, eta = .88, $f = 1.85$). The difference in inversion rates for negated versus non-negated auxiliaries was most striking for the modals. Overall, non-negated modals were inverted 90.2% of the time, whereas negated models were inverted only 33.6% of the time.

4.4.6 *How come* questions. As was discussed above, *how come* questions are unique among *wh*-questions because subject-auxiliary inversion is not permitted in matrix *how come* questions. *How come* questions share three other features that are related to the lack subject-auxiliary inversion. The first corollary feature is that because *how come* questions do not have subject-auxiliary inversion, they do not require auxiliaries. Second, if *do*-support is provided, the question is starred unless special emphasis in given to the *do* just as is the case with declarative sentences (e.g., **how come she does go to the store?* versus *how come she does go to the store?*). Third, because there is no subject-auxiliary inversion, auxiliary and copula elements left in their initial position are able to reduce and cliticize to their subject NP. Notice, for example, that *how come she is going?* sounds somewhat stilted compared with *how come she's going?*. A final, and perhaps related feature is that, at least to my ears, *how come* questions with auxiliaries sound much more natural if the auxiliary is negated (compare, for example, *how come she can't go to the store?* with *how come she can go to the store?*).[23]

---

[23] There are two ways to make the second *how come* question sound more natural. The first is to add a negating tail to the sentence (e.g., *how come she can go to the store, and I*

The question is, did the children respect these differences between *how come* questions and the other types of questions? The rarity of *how come* questions with full auxiliaries suggests that the children acknowledged the differences and respected them. Overall, the children asked 165 *how come* questions. Of these questions, only 45 had full or contracted auxiliaries. Of these 45 questions, only 22 had non-contracted auxiliaries. Of these 22 questions, 17 of the questions had negated auxiliaries.

In practice, determining the true inversion rate was a bit trickier for *how come* questions than for the other types of questions. Overall, children incorrectly placed a full auxiliary before the subject NP 14% of the time. If we include contracted auxiliaries (both *in situ* and raised), children incorrectly inverted 11% of the time. Compared to some of the error rates I discussed in chapter 3, these error rates seem quite impressive.

I would like to argue, however, that if one looks at the actual inverted *how come* questions that were asked, one is led to the conclusion that children essentially never make inversion errors in *how come* questions. The following were the five examples of inverted *how come* questions:

Adam54 (4;10) why...how come is that?
Adam54 (4;10) how come is dat?
Adam54 (4;10) how come is dat?
Nathan29 (3;9) how comes he [...] cause he threw it and and this spilled and and and

---

can't). The second is to stress the subject NP (e.g., *how come **she** can go to the store*). That these two maneuvers work suggests that the stiltedness of *how come* questions with non-negated auxiliaries is due to pragmatic factors and not syntactic factors.

Nathan29 (3;9) eh dis [...] nose [#&] eh how comes he taking a big bite?

Notice that only two children asked inverted *how come* questions and each child asked all of their examples within a few lines of one another. Adam's examples are particularly enlightening. He began by asking a *why* question, hesitated and then switched to a *how come* question. He then repeated that same question twice more in the course of 5 lines. One can almost imagine that at age 4;10, Adam serendipitously noticed a syntactic inconsistency in the language, could think of no reason for its existence, and so chose to use a construction that he may have realized was a bit odd (in the same way you or I might use **amn't* for *am not*).

Nathan's examples are equally tantalizing. Given that *how comes* is not part of the adult language, we can assume that he never heard contracted *is* attached to *how come*. These may, therefore, be real bona fide examples of inversion with *how come*. They could also just be incorrectly transcribed *how come she*. It's even possible that they are examples of inverting inflected main verb *come* or restricted auxiliary *come*.

4.4.7 Inversion in embedded questions. I found no examples of scoreable embedded *yes/no* questions that were introduced by the complementizer *whether*. I found 46 scoreable embedded *yes/no* questions introduced by the complementizer *if*.[24] None of these 46 embedded *yes/no*

[24] I used the following test to distinguish embedded yes/no questions from embedded conditional clauses: I substituted *whether* for *if*. If the meaning conveyed by the *whether* sentence was the same as that conveyed by the *if* sentence, I counted the sentence as an embedded yes/no question. If the meaning was different, I did not count the sentence. For example, the sentence *what may happen if I put two wheels there?* (Adam39) was not counted because *what may happen whether I put two wheels there?* was different in

questions had an inverted auxiliary. In other words, I found no examples like **I wonder if can she go.* I did, however, find a lone example of *if* preceding an auxiliary in a matrix *yes/no* question. When Ross was 7;8 he said, "um@h or if do you want to read it with Mark?" (Ross93). I found approximately 364 scoreable embedded *wh*-questions.[25] Of these there were 36 examples in which a non-contracted auxiliary or copula appeared before the subject NP of the embedded *wh*-question.[26] The difference in inversion rate for embedded *yes/no* and *wh*-questions was statistically significant ($\chi^2(1) = 4.93, p < .05$). In Table 5.10, I list all of the 36 scoreable embedded *wh*-questions which I found in the 12 children's transcripts.

Please Insert Table 5.10

Perhaps the reason the children occasionally inverted subject and auxiliary in embedded questions is that, as Grimshaw (1979) suggests, they

---

meaning from the original sentence. The sentence *I wonder if I can find some place dey may have a nice race* (Adam 49) was counted as an embedded question because its meaning is the same as the substituted sentence *I wonder whether I can find some place* ....

[25] The number is approximate because it was not always possible to distinguish with certainty embedded *wh*-questions from relative clauses. I used the criteria outlined in Bresnan and Grimshaw (1978). If these criteria were not sufficient, I used my best judgment based on context, the speech patterns of the child, etc. If I still could not decide whether a particular example was a relative clause or an embedded question, I did not include it.

[26] I found another 50 possible cases of inversion in embedded *wh*-questions. Most of these cases were not considered scoreable because they contained a contracted auxiliary (e.g, Adam 41 (3;11): I don't know what's dat?). Some were stuttered or had hesitations marked which suggested that they were speech errors or quotations of matrix questions (e.g., Adam35 (3;8): I don't know where one place ... can I peak? Can I look?). I also found one embedded question that appeared to have two *wh*-words in Comp position (Mark74 (4;5): no [#1] I don't know why who're there). I did not include this example in the tallies. It's interesting to note, however, that Mark asked two of three of the inverted embedded *why* questions found in the transcripts.

used semantics/pragmatics to determine when to invert. If this is the reason, then we would expect the children would have inverted those embedded questions that were most question-like. If we adopt the intuitive criteria that a "question-y" embedded question is one that either (1) has question intonation as marked by a question mark, or (2) is a clear request for information, we find that 30% (11/36) of the children's inverted embedded questions are question-y. I did not determine what percentage of non-inverted embedded questions were question-y, so I cannot say whether 30% represents a larger percentage than one would expect based on chance alone. What I can say, however, is that at least some of the examples in which children inverted were distinctly un-question-like (e.g., Adam41 (3;11) I know what time is it.).

There are three other possible explanations for why the children inverted in embedded questions. The first explanation is to deny that such examples are real and argue that the children only looked like they inverted in embedded questions. According to this explanation all of the purported examples of inversion in embedded questions are actually improperly transcribed examples of quotative questions (e.g., "I asked, 'what is she doing here?' "). The second and third possibilities accept that the examples of embedded inverted questions are real. One possibility is that the children may not have always correctly distinguished matrix from embedded questions. The other possibility is that they might have been a little lax about obeying the Structure Preserving Constraint or whatever constraint normally prevented them from inverting subject and auxiliary in embedded questions. In other words, whatever constraint was operating may have been probabilistic rather than 100% effective.

All three explanations have in common that they rely on there being a confusion between matrix and embedded questions either on the part of the transcriber or the child. Therefore, if any of these last three explanations is right, then for each type of *wh*-question, the frequency of inversion in embedded questions should reflect the frequency of inversion in the corresponding type of matrix question. For example, since the children inverted matrix *who* questions 100% of the time and matrix *when* questions 77% of the time, we would predict that the children would have inverted embedded *who* questions more frequently than they would invert embedded *when* questions.

In order to test this, I determined how correlated the children's inversion rates were for matrix and embedded *wh*-questions. For each type of *wh*-question, I determined the frequency of inversion in matrix and embedded questions. I then ranked each type of matrix *wh*-question according to inversion rate. I did the same for each type of embedded *wh*-questions. The inversion ranks for matrix and embedded questions were highly and significantly correlated (Spearman rank correlation coefficient, $rho_s = .875$, $p < .01$). Please see Table 5.11 for the inversion rates and ranks for each type of matrix and embedded *wh*-question.

Please Insert Table 5.11

4.4.8 Inversion of Uninvertible Verbs. I found no examples of inversion of restricted or semi-auxiliaries in either *wh*-questions or *yes/no* questions. In other words, children never said things like *better I*

eat? , *hafta I eat?* or *needs she food?*. In previous analyses in which I examined all lines that contained a "?" or one of the *wh*-words, I only found the following examples in which a child placed a non-inverting verb in front of the subject NP:

Adam7 (2;6)	Where go one?
Adam11 (2;8)	Where go pencil
Adam45 (4;3)	Where makes de noise?
Sarah111 (4;6)	What means two?
Nathan28 (3;7)	What's gonna be tomorrow?
Peter11 (2;3)	Where da go frisbee go

Notice that 3 of the 6 examples involve the verb *go*. As mentioned above, *go* is an unusual verb because it has some auxiliary-like properties. It can precede a main verb (e.g., *let's go get ice cream, and then she goes running through the stop-light like she's the only car on the road, you're gonna pay for this,* etc.). It can be a pro-verb with much the same sort of limited semantic content as copula *be*. For example, compare *it goes like this* with *it's like this*, and *the dishes go here* with *the dishes are here.* More importantly for the issue of invertibility, *go* is perhaps the only non-auxiliary that can still invert (e.g., *how goes it?* or *how goes the thesis?*). Notice that even this very limited, somewhat stilted, inversion of *go* is limited to the deontic meaning of *go*. At least to my ears, inversion of the actional *go* is completely unacceptable, even as an example in the most literary style (compare, for example, **how goes she to the store?* with *how goes the thesis?*).[27]

---

[27] Notice that all of the examples with go are *where go* examples. It's at least possible that the only reason children say *where go* ___ is because of its phonological similarity to *there go* ____ and *here go* ____. It's also possible that these examples are just typographical errors in the transcripts and the children actually did say *there go* or *here go*.

Now consider Nathan's question, *what's gonna be tomorrow?* This could be an example of inverting *gonna be* along with the contracted auxiliary *'s*. However, the context in which the question was asked suggests that Nathan merely omitted the pronoun *it*. In the last clause-length sentence prior to Nathan's anomalous question, Nathan's mother asked him what the name of the next day of the week was by asking "what's it gonna be tomorrow?"

The context of Adam's question *where makes de noise?* is not as helpful in determining what is being said. The immediate context is:

Adam: Mommy # you know where de noise makes it?
Adam: Where makes de noise?
Adam: no # you turn dis light # and push dis over
%par <aft> siren noise from Adam.

Perhaps what Adam wanted to say was "you know where it makes de noise?" He might have been aware that his initial attempt had improper word order. He might have incorrectly attempted to fix the word order by moving *makes*, instead of correctly moving *it*.

Last we have Sarah's question, *what means two?* Sarah asked this question while coloring in a color-by-numbers coloring book. She wanted to know what color she should color something that was marked by a number two. Below is the context for this question:

Sarah: hey # what t ...
Sarah: what means two?
Sarah: green # does green t ...
Mother: two?
Sarah: yeah.

Mother: means brown.
Sarah: oh # I did it green.

It's possible that Sarah was aiming for the subject question *what means green*. From the context, however, it seems more likely that her target question was *what does green mean* and that rather than invoke *do*-support, she moved the main verb *means*. What is interesting is that of all the verbs she might have moved, she moved a main verb that has the kind of meaning that is among the meanings auxiliaries exhibit cross-linguistically (see Steele, *et al.* 1981).

4.4.9 Exclamatives. I found very few examples of *yes/no* exclamatives with subject-auxiliary inversion. Most sentences that were marked with an exclamation point were either non-sentential (e.g, *Cookie!*) or sentences with standard declarative word order (e.g., *I don't like fish!*). The only examples I found of *yes/no* exclamatives with auxiliary occurring before the subject were:[28]

Adam54 (4;10) Wasn't dat a fly!
Adam52 (5;2): Is that loud!
Eve19 (2;3) [C/Int:Was you man!], you boy. Papa's a boy
Sarah37 (3;0): Do me like dis!
Sarah81 (3;10) Now, don't you fall off!

Likewise, there were few utterances with *wh*-words that had sentence

[28] There have been a number of formatting changes to the transcripts since I did this search. Emphasis on particular words is noted by [!] in some of the recent versions. Some of the examples shown above no longer appear as exclamations and other examples now appear that didn't then. I have not attempted to re-do the analyses as the transcripts have changed. The fact that they have changed so just points out that it is very hard to represent intonation in print and that any analysis that relies heavily on intonation for it's validity is subject to criticism. I think analyses of exclamatives definitely fall into that class.

final exclamation points to indicate that they were *wh*-exclamatives. There were even fewer that contained an auxiliary and a subject NP. The only candidates for scoreable *wh*-exclamatives that I found were the following:

Peter 20 (3;2)	when was [/] this [!] day [!]
Ross35 (3;2)	where did that go!

Of the two potential examples of *wh*-exclamatives, the way Peter's example is transcribed suggests that most likely Peter just exuberantly asked a *wh*-question. From context, the same is likely to be true for Ross's exclamative, although in Ross's case, we don't have the transcriptional clues to support this decision.

4.4.10 Negative polarity sentences. I found no examples of negative polarity sentences, nor did this surprise me.

4.5 Discussion

In this section I will restrict my discussion to fairly specific details of the findings. I will summarize my results and compare them with the results reported in the literature, emphasizing the questions I posed earlier. I will save the more general discussion of the learnability and linguistic implications for the end of the chapter.

4.5.1 Overall inversion rate. Overall, 92% of the children's questions which contained an auxiliary had the auxiliary correctly placed before the subject. In addition, the median inversion rate for the children was about 94%. What these figures suggest is that failing-to-invert in matrix questions is neither pervasive nor exceptionally rare. All children except Nathan were right significantly more often than they are wrong. In addition, no child performed worse than chance. In other words, none of the children had a grammar which prohibited inversion. The overall rate of inversion for this study is comparable with the rates reported in other studies that examined inversion frequency in spontaneous speech (i.e., Ingram and Tyack, 1979; Klee,1985; Labov and Labov, 1978). The inversion rate found in this study is somewhat higher than that reported in elicitation studies (i.e., Erreich, 1984, Kuczaj and Brannick, 1979). This is consistent with Erreich's finding that the rate of inversion was significantly higher for children's spontaneous questions than for their elicited questions.

In addition to the overall inversion rates for the children being significantly different, two patterns of performance were observed. For about half of the children (the children with high inversion rates), inversion rates remained steady or improved with age. For the other half of the children (the children with lower overall inversion rates), there was a dip in inversion rate after an initial period of near-perfect performance. However, judging from the existence of significant variable effects and interactions in the analyses of variance, the relative performance on particular constructions (e.g., particular *wh*-questions, particular auxiliaries, etc.) was fairly uniform across children. What this suggests is that, although some of the children had a developmental dip in

performance, the relative rates of inversion for different constructions was the same for both the children whose performance was monotonic and for those that dipped.

4.5.2 Inversion rates in yes/no and *wh*-questions. Of the studies on subject-auxiliary inversion rates for *yes/no* and *wh*-questions that I reviewed above, basically a third found higher inversion rates for *yes/no* questions than *wh*-questions, a third found the inversion rates to be the same for *yes/no* and *wh*-questions, and a third found the inversion rate to be higher for *wh*-questions than *yes/no* questions.

My longitudinal, multi-child analysis replicated the overall findings of the previous studies. In my study, 93.0% of the children *wh*-questions were inverted and 93.7% of their *yes/no* questions were inverted. Three of the 6 children I studied inverted more frequently in *yes/no* questions than *wh*-questions, 4 inverted more often in *wh*-questions than *yes/no* questions, and 6 inverted with equal frequency for both types of sentences.

One possible explanation for the differences in inversion rates for the children is that perhaps these rates are really just a reflection of the fact that the children were various ages when their speech was transcribed. Perhaps Bellugi is right and there is a stage that all children go through during which they invert in *yes/no* questions but not in *wh*-questions, but we only managed to catch this stage in 3 of the children. This does not seem to be the case, for two reasons. First, the three children who inverted more in *yes/no* questions than *wh*-questions were not the youngest children. Second, even for the child with the most extreme difference in rates of inversion for *yes/no* questions versus *wh*-questions, the difference was only about 15 percentage points. There were no children who came even close

to being in Bellugi's stage during which children invert 100% of the time in *yes/no* questions and 0% of the time in *wh*-questions

In summary, I found no evidence to support Bellugi's claim of the existence of a stage during which children invert in *yes/no* questions but not *wh*-questions. My results reflected the pooled results of all the previous studies; some children invert more in *yes/no* questions than *wh*-questions, some invert more in *wh*-questions than *yes/no* questions and some invert equally often in both types of questions. I suspect that which pattern a particular child's behavior falls into depends more on extra-syntactic factors than the nature of that particular child's grammar.

4.5.3 Inversion rates for different *wh*-questions. Like Erreich (1984), Klee (1985), Kuczaj and Brannick (1979), and Labov and Labov (1978), I found that the rate of inversion was different for the different types of *wh*-questions. The relative inversion rates for the children in this study were similar to those described by the researchers above. The overall rates of inversion for this study were: 100% for *who* questions, 97% for *how* questions, 95% for *where* questions, 94% for *what* questions, 87% for *why* questions, 79% for *which* questions, and 77% for *when* questions. The only notable difference is that the children in this study inverted *how* more frequently than the children in Kuczaj and Brannick's (1979) study or in Erreich's (1984) study. This difference may be due to the fact that in Kuczaj and Brannick's study and in Erreich's study, children were asked to produce *how* questions along with *how come* questions. The children in these two studies might have occasionally confused *how* questions for *how come* questions.

4.5.4 Inversion rates for different auxiliaries. In this study, the overall inversion rate for all of the questions asked by all 12 children was 88% for *be* questions (copula and auxiliary combined), 99% for *do* questions, 91% for *have* questions and 92% of modal questions. The difference among inversion rates for *be*, *do*, and the modals was statistically significant.

If we consider only *is* and *are* questions (the only *be* questions asked by a majority of the children), the average inversion rate was 90% for auxiliary *be* questions and 84% for copula *be* questions. The finding that the children tend to invert auxiliary *be* more often than copula *be* is particularly striking given that they acquired auxiliary *be* later (see chapter 4), because questions that are acquired later tend to be inverted less (see Stromswold 1988a, 1988b).

Like Klee (1985), Kuczaj and Maratsos (1983), and Labov and Labov (1978), but unlike Erreich (1984), I found that the children inverted some types of auxiliaries more often than others. I found that inversion rate was highest for *do*, intermediate for auxiliary *be*, the modals and *have* and lowest for copula *be*. Of note, Klee (1978) also found a low rate of inversion for copula *be* and Labov and Labov reported less inversion for *be* (although they did not distinguish auxiliary *be* from copula *be*). I will return to the significance of the differences in inversion rates for the various types of auxiliaries later in this paper.

4.5.5 Inversion rates for negated and non-negated auxiliaries. Bellugi (1971), Derwing and Smyth (1988) Erreich (1984), and Labov and Labov (1978) all report lower rates of inversion for

negated auxiliaries. I found that the average inversion rate for negated auxiliaries (56%) was significantly less than that for non-negated auxiliaries (91%). The difference between average inversion rates for negated and non-negated *do* (60% and 99%, respectively) was large and significant, as was the difference between average inversion rates for negated and non-negated modals (34% and 90%, respectively). In addition, the average inversion rate for negated *be* was less than for non-negated *be* (81% and 93%, respectively), but this difference was not significant. In sum, for each type of auxiliary, the children inverted negated forms less than non-negated forms. However, the size and significance of the effect varied for the various types of auxiliaries. It was most striking for the modals and least striking for *be*.

4.5.6 Inversion in *how come* questions. Inverted *how come* questions were extremely rare in absolute number. They represent some percent of scoreable *how come* questions, ranging somewhere from 0% (if we reject all of the purported cases) to 14% (if we count each of Adam's repetitions of *how come is dat?*). I would argue that the actual frequency is closer to zero percent than 14 % for two reasons. First, only two of the 12 children produced even one example of an inverted *how come*. Second, the two children who did produce them produced all of their examples at a relatively late age within a few lines of each other. Adam produced the same example with a full auxiliary three times and Nathan produced two examples with contracted auxiliaries. I characterize the frequency of this type of question as being more frequent than hens' teeth, but not nearly as common as four-leaf clovers.

How do I reconcile my findings what those of Kuczaj and Brannick's (1979)? In a series of three studies, Kuczaj and Brannick had three-, four-, five-, and six-year old children imitate different types of grammatical and ungrammatical *wh*-questions. They found that the children incorrectly imitated sentences like *how come the girl will kiss the boy* as **how come will the girl kiss the boy* between 10% and 40% of the time. Differences in methodologies might account for the differences between my findings and their findings. I examined the frequency of inverted *how come* questions in the spontaneous speech of children, whereas Kuczaj and Brannick examined the frequency of inversion in elicited *how come* questions. It's not surprising, therefore, that I found fewer errors given that the frequency of inversion errors has been shown to be greater in elicited questions than in spontaneous questions (Erreich, 1984).

In addition, Kuczaj and Brannick had their subjects imitating about 3 questions per minute. Only about 10% of the questions imitated in the studies were *how come* questions with properly placed auxiliaries. It's no wonder that in the context of rapidly imitating questions, some of which were inverted *how come* questions, most of which were of the form *wh*-word-Aux-NP, that the children sometimes inverted in *how come* questions. What is important to notice is that even in a setting that greatly favors the production of inverted *how come* questions, the children in Kuczaj and Brannick's study never inverted in the majority of *how come* sentences.

4.5.7. Inversion in embedded versus matrix questions.

There were no examples of inverted embedded *yes/no* questions among some 46 scoreable embedded *yes/no* questions. There were 36 examples of

inverted embedded *wh*-questions among 364 scoreable embedded *wh*-questions. The difference in inversion rate for embedded *yes/no* questions versus embedded *wh*-questions was statistically significant ($\chi^2(1) = 4.93, p < .05$). Therefore, whatever explanation we use to account for the inversion in embedded *wh*-questions, it cannot entail that we also should have found inverted embedded *yes/no* questions.

Having said that, let me return to the inverted embedded *wh*-questions. In the children studied, the overall inversion rate of embedded *wh*-questions was about 10%. Thus, the results are somewhat different from those reported by Pinker (1984) who wrote:

> I have found only a single such error in Brown's data (*I don't know what are dey*, from Adam's nineteenth sample), as compared to five or so correct embedded questions with noninverted auxiliaries. [29] Thus children do seem to restrict inversion to root question, despite the fact ... that children ... are rather cavalier at inverting versus not inverting when the choice is dictated by other grammatical features (viz., presence of a *wh*-word). (Pinker, 1984, p. 285)

Why do children occasionally invert in embedded *wh*-questions (and not in embedded *yes/no* question)? On the one hand, it's clear that children treated embedded questions differently than matrix questions. If they hadn't, then instead of inverting 10% of embedded questions, they would have inverted 93% of the embedded *yes/no* and *wh*-questions (the approximate overall inversion rate for matrix *yes/no* and *wh*-questions). On the other hand, they made mistakes 10% of the time.

---

[29] As Table 5.10 shows I found 16 other examples of inversion in embedded *wh*-questions in Adam's transcripts. The reason for the discrepancy between Pinker's findings and my findings is probably due to the fact that I analyzed all of Adam's transcripts, whereas Pinker only examined 23 samples of Adam's speech catalogued by type (Pinker, 1984, p. 261).

What could account for generally good, but not perfect, performance? One possibility is that the children used semantics and pragmatics to determine when to invert (Grimshaw, 1979). According to this explanation, the 10% error rate represented those embedded questions where using pragmatics and semantics led the children astray. This doesn't seem likely because only about 30% of the examples of inversion in embedded questions seem "question-y", and some of the errors look distinctly un-question-y. Another problem with the semantic/pragmatic account is that it cannot explain the lack of inversion errors in embedded questions, since embedded *yes/no* questions seem potentially as question-y as embedded *wh* -questions. The questions *do you know whether I can go?* and *do you know where I can go?* seem equally question-like.

It is possible that all of the purported examples of inversion in embedded questions are actually improperly transcribed examples of quotative questions. If one scrutinizes the list of embedded inversion errors in Table 5.10, this seems like a possible explanation for some of the sentences but not others. It's hard to imagine how examples like the following are quotative:

Adam20 (3;0): Smith's who are you.
Adam36 (3;8): I better see what was this.
Adam47 (4;4): Look at what are dey building?

Another possibility is that the children may not have always correctly distinguished matrix from embedded questions. Under this account, their inversion errors represent the times they mistook an embedded question for a matrix question. It's also possible that they

correctly distinguished matrix from embedded questions 100% of the time, but they were sometimes a little lax about obeying whatever constraint normally prevented them from inverting subject and auxiliary in embedded questions. These last three explanations all assume a confusion between matrix and embedded questions either on the part of the transcriber or the child. All three predict, therefore, that the frequency of inversion in embedded questions should correlate with the frequency of inversion in the corresponding type of matrix question. This prediction is borne out; the inversion ranks for matrix and embedded questions were highly and significantly correlated.

But what about the lack of inversion errors in embedded *yes/no* questions? It's hard to imagine a probabilistic syntactic constraint that would operate 100% of the time in *yes/no* questions but only 90% of the time in *wh*-questions. The quotative explanation would also predict that there would be occasional examples of mis-transcribed *yes/no* questions (e.g., "I wonder, 'can I go?' " --> *"I wonder can I go?", "I asked, 'did she go?' " -->*"I asked did she go?"). Likewise, if the children were merely confused occasionally about whether they were asking a matrix or embedded *wh*-question, what kept them from being confused for *yes/no* questions? I will leave this as an open question, and will return to this question after I have described the results of an experiment which was designed to probe the nature of inversion in embedded and matrix questions.

4.5.8 Inversion and the uninvertible verbs. There were no examples of children inverting a restricted modal auxiliary (e.g., **better*

you go?) or a semi-auxiliary (e.g., **gonna I go?*). In over 40,000 questions, there were only 5 cases in which a child moved a main verb. In three of the 5 examples, it's the same verb, deontic *go*, which moves. It is interesting to note that *go* was the verb that moved in all three examples of main verb movement that Pinker (1984) found in his survey of the acquisitional literature. Hence, my results mirror exactly those reported in the literature. I would conclude, first, that errors of inverting main verbs are as rare as hens' teeth and, second, that only a very particular breed of hen can have teeth.

4.5.9 Exclamatives. There were between zero and two examples of *wh*-exclamatives with subject-auxiliary inversion. However, because there are so few examples of *yes/no* exclamatives with inverted auxiliaries, I am unwilling to place much importance on the paucity of *wh*-exclamatives with inverted auxiliaries.

5. Inversion Judgment Experiment

In this experiment, preschool children judged the grammaticality of matrix and embedded questions that did or did not exhibit subject-auxiliary inversion. Whether one is a linguist who relies on one's own judgment and that of other adults, or a researcher of child language who designs complicated experiments in which children judge sentences, the critical assumption is that when a person judges the grammaticality of a sentence, she must consult her internal language (I-language) or grammar and determine whether that sentence is acceptable in this language. In this case, the assumption is that if a child's grammar is different with regard to inversion, this will be apparent in her judgments.

The same basic problem arises whether one is asking a professional linguist or preschooler to judge a sentence. One has to be sure that the person judging the sentence is judging that aspect of the sentence which is under investigation. When linguists ask adults to judge a sentence, they give detailed instructions about which aspects of the sentence should be considered and which should be ignored. They tell the adult informants which semantic reading of the sentence they are interested in. They also may contrive situations in which a semantically marked sentence could occur in order to help the adult tease apart semantic anomalousness from true ungrammaticality.

Consider a linguist who is interested in the grammaticality of long-distance adjunct *how* questions like **?how$_i$ did John know to fix the car t_i?*. She would instruct her informants that she is interested in the reading for which a possible answer might be "by replacing the spark plug" and not in the reading for which "his mother told him it needed fixing" might be a

possible reason. She might also tell the informant about the many possible ways of fixing a car to ensure that the informant will find the question semantically plausible. Despite all her efforts to focus her informants' judgments to those under investigation, some people will make better informants because they are able to ignore extraneous syntactic and semantic details of the sentence.

If adults' grammaticality judgments can be skewed by semantics, we should expect children to have even greater problems with this. Children seem to have difficulty considering language as an abstract entity separate from the things which language refers to. Below a certain age, children tend to say that a word like *tree* which refers to a large object is a bigger word than a word like *dandelion* which refers to a small object. An analogous phenomenon applies to young children's grammaticality judgments. They are more likely to say that a sentence like *elephants are little* is ungrammatical than the sentence **elephants big are* (de Villiers and de Villiers, 1974). For this reason, when we ask children to judge the grammaticality of sentences we have to make sure theat they judge syntactic acceptability rather than semantic plausibility.

One way to do this is to give children pairs of sentences which differ only in the feature under investigation (minimal pairs). So, for example, if we are interested in inversion, we would present both **what Kermit is?* and *what is Kermit?* If a particular child judges the sentences along some dimension other than inversion (e.g., semantic plausibility, pronoun choice, number of arguments, etc.), we will not be led astray because she will judge both sentences as either good or bad.

There are two basic ways we can have children judge minimal pairs of sentences. I will call the first method the Absolute Judgment Paradigm.

In this paradigm, the child first judges whether the question **what Kermit is?* is good and, later in the experiment, judges *what is Kermit?*. In the second paradigm, the Relative Judgment Paradigm, the child hears pairs of sentences which differ only in a single feature and decides which of the two sentences is better. The advantage of the Absolute Judgment Paradigm is that it allows us to discover the full range of constructions which are acceptable to the child. We might be able to discover, for example, if a child's grammar admits both **what Kermit is?* and *what is Kermit?* One advantage of the Relative Judgment Paradigm is that, by presenting minimal pairs of sentences and forcing the children to choose between them, we can focus the child's attention to the feature under investigation. A second advantage is that, if the child's grammar admits both sentences but one is the preferred form, we would be able to discover this. There are two disadvantages to the Relative Judgment Paradigm. First, if the child is forced to choose between two types of sentences which are equally good to that child, the child will be forced to either choose one of the two sentences at random during each trial, or (worse) to choose at random a single sentence type which she then picks throughout the experiment. A second disadvantage is the Relevant Judgment Paradigm makes greater demands of the child than the Absolute Judgment Paradigm because it requires that the child remember and compare two sentences in order to decide which is better.

5.1 Overview

Using the Absolute Judgment Paradigm, preschool children judged the grammaticality of matrix and embedded questions that did or did not exhibit subject-auxiliary inversion. I chose to use the Absolute Judgment

Paradigm because I wanted to be able to discover whether the children's grammars admitted both inverted and noninverted versions of embedded and matrix questions.

5.2 Subjects

Twenty-two children between the ages 3 years, 0 months and 6 years, 1 month (mean age = 4;6) participated in the experiment. All children had normal vision and hearing and spoke only English. The children came from a small industrial city in New Hampshire. Because the children came from a wide range of socio-economic classes, I was concerned that the children's ages might not be a good indication of their relative linguistic development. Therefore, the mean length of utterance (MLU) in children's descriptions of seven pictures was calculated. Children were divided into three groups according to age and Mean Length of Utterance (MLU). The average elicited MLU for the three MLU groups were 2.94, 6.12 and 8.76. The mean ages for the three groups were 4;0, 4;6, and 5;0, respectively.

5.3 Stimuli.

3.1 Questions types. Half of the questions were matrix questions and half were embedded questions. Embedded questions all began with the matrix clause, *he wonders* . . . . I read all the questions with the appropriate intonation pattern. The auxiliary occurred before the subject in half of the questions and after the subject in the other half of the questions. The four types of questions were:

Matrix +Inverted: What is Kermit?
Matrix -Inverted: *What Kermit is?
Embedded +Inverted: *He wonders what is Kermit.
Embedded -Inverted: He wonders what Kermit is.

3.2 *Wh*-words used. The following 6 *wh*-words were used: *who*, *what*, *where*, *how*, *how come*, and *why*. *When* and *which* were not included because they are acquired quite late and rarely used by many preschool children (Stromswold, 1988a, 1988b). The *wh*-questions used in this study were:

How: How is Grover?
How come: How come Bert is mad?
What: What is Kermit?
Where: Where is Big Bird?
Who: Who is Ernie?
Why: Why is Oscar happy?

These questions were chosen because they were approximately equal in length and in semantic plausibility.

3.3 Order of Presentation. Each question type occurred with each *wh*-question twice during the experiment, once in the first half of the trials and once in the second half of the trials. The order of the trial questions was pseudo-random with the following conditions. The same *wh*-question or question type never occurred more than twice in row and no more than 3 grammatical or ungrammatical questions occurred in a row. Half of the subjects heard the questions in one order, and the other half heard the questions in the reverse order.

5.4. Procedure.

In a quiet room away from the other children, I had each subject judge the grammaticality of the *wh*-questions in the following manner. I manipulated a dog puppet which was "just learning how to talk" and, therefore, sometimes made mistakes. The child helped teach the dog to talk by deciding if what the dog puppet said sounded good or bad. If it sounded good to her, she gave the dog a bone. If it sounded bad to her, she gave the dog a rock. To familiarize the child with the experimental procedure, each child judged practice questions until she judged 3 questions in a row correctly or until she had judged all 10 practice questions. (All of the children correctly judged 3 sequential questions.) The practice questions are shown in Table 5.12. Notice that none of the ungrammatical practice questions were wrong because of auxiliary placement. After each practice trial, I told the child whether the question was good or bad. I gave no feedback on experimental trials. I gave each child a star sticker after every fourth judgment and a short break after every 12 judgments. At the end of the experiment, I gave each child a picture book. Most of the children took between 45 minutes and an hour to judge the 48 trial questions. All of the children completed the experiment.

Please Insert Table 5.12 here

Elicited Mean Length of Utterance (MLU). After each child had judged all of the questions, I asked the child to describe 7 simple color drawings taken from a picture book. I let her speak as long as she

wanted. I used a slightly modified version of the guidelines outlined in Brown (1973) to calculate an elicited MLU for each child's descriptions of the pictures.[30]

5.5 Results

Overall, the children judged 76.3 % of the grammatical questions as being good and 57.3% of the ungrammatical questions as being good. This difference is significantly greater than would be expected if the children had merely been responding randomly to the test questions ($\chi^2(1) = 40.23$, $p < .0005$). The children in the Low MLU group judged 66.5% of the grammatical questions as good and 62.5% of the ungrammatical questions as good. Children in the Middle MLU group judged 72.6% of the grammatical questions as grammatical and 55.4% of the ungrammatical questions as grammatical. Children in the High MLU group correctly labeled 81.2% of the grammatical questions as good, but still incorrectly labeled 50.0% of the ungrammatical questions as good. Thus, the procedure appears to have successfully gotten children to judge the

---

[30] These modifications were necessary because of the non-conversational nature of the speech sample. Because the children were not talking to anyone, there were no conversational restraints on how long they could or should talk. Some children seemed compelled to fill the conversational void, glancing at the experimenter for permission to stop talking. Rather than calculate the MLU per picture described, if the prosodic contour of a child's description of a picture suggested that the child was simply stringing together sentences with conjunctions, I calculated the mean length of sentential clauses. If a child's description was just a string of NPs or VPs, I used the prosodic contour to decide how to divide the description into "utterances".

While elicited MLUs seem to capture the linguistic sophistication of one child relative to another, elicited MLUs are not directly comparable to MLUs for spontaneous speech samples. The nature of the task seems to make children use longer, more complex sentences than they use in spontaneous speech. Furthermore, elicited MLUs seem to be highly dependent on the stimuli used to elicit the speech sample. Children tend to say more and have higher MLUs when ask to describe more interesting stimuli.

grammaticality of sentences, though as one might expect the procedure was more successful with the more linguistically advanced children.

Analyses of variance were performed on the data using first Age and then MLU as the between-subjects variable. Neither Age nor MLU had a significant main effect and only MLU entered into any significant interactions. Therefore, I will only discuss the results of the MLU analyses. There was a significant main effect for *wh*-word ($F(5, 95) = 10.46$, $p < .0005$, eta = .60, $f = .55$). Children judged as good 70.5% of *how* questions, 40.3% of *how come* questions , 75.6% of *what* questions, 72.2% of *where* questions, 73.3% of *who* questions, and 69.3% of *why* questions. These results are shown in Figure 5.49 below. Planned comparisons between *how come* questions and the other 5 types of *wh*-questions revealed that *how come* questions were judged bad significantly more often than were any of the other types of questions (all F's (1, 21) greater than 14.21, all p's < .001).

Please Insert Figure 5.49

There was also a significant main effect of the Matrix/Embedded factor ($F(1, 19) = 5.31$, $p = .033$, eta = .55, $f = .66$). Children judged embedded questions to be good 69.7% of the time and they judged matrix questions to be good 64.0 % of the time. There was a significant interaction between MLU and Matrix/Embedded factors ($F(2, 19) = 4.19$, $p = .031$). As Figure 5.50 shows, children in the High and Middle MLU Groups judged embedded questions to be good more often than they judged

matrix questions to be good, while children in the Low MLU group judged matrix questions to be good more often than they judged embedded questions to be good. Separate analyses of each MLU group revealed that the difference between embedded and matrix questions was only significant for the Middle MLU Group ($F(1, 7) = 10.19$, $p = .015$, eta = .77, $f = 1.21$).

Please Insert Figure 5.50

Inversion and Matrix/Embedded factors interacted significantly ($F(1, 19) = 35.55$, $p < .0005$, eta = .81, $f = 1.91$). Children correctly judged inverted matrix questions good 76.5% of the time and the noninverted embedded questions good 77.3% of the time. They incorrectly judged the ungrammatical (in the adult language) noninverted matrix questions good 51.5% of the time and the inverted embedded questions good 62.1% of the time. The following planned comparisons were made. The difference between inverted and noninverted embedded questions was significant ($F(1, 21) = 18.34$, $p < .0005$, eta = .68, $f = .68$), as was the difference between inverted and not inverted matrix questions ($F(1, 21) = 21.66$, $p < .0005$, eta = .71, $f = 1.00$). In addition, the difference between the ungrammatical questions was significant ($F(1, 21) = 7.31$, $p = .013$, eta = .51, $f = .59$), whereas the differences between the grammatical questions was not significant ($F(1, 21) < 1$). In other words, children had a significant preference for matrix questions that were inverted and embedded questions that were not inverted. Furthermore, they judged the ungrammatical inverted embedded questions to be good significantly more often than they

judged the also ungrammatical noninverted matrix questions to be good. The interaction between inversion and clause type is shown in Figure 5.51.

Please Insert Figure 5.51

MLU, Inversion and Matrix/Embedded factors also interacted significantly ($F(2, 19) = 4.60, p = .024$, eta = .57, $f = .70$). This interaction is shown in Figure 5.52. Separate analyses for the three MLU groups revealed that the interaction between Inversion and Matrix/Embedded Question was increasingly larger and more significant for each successive MLU group (Low MLU $F(1, 6) = 4.85, p = .070$, eta = .67, $f = .90$; Middle MLU $F(1,7) = 7.28, p = .031$, eta = .71, $f = 1.02$; High MLU $F(1, 6) = 39.98, p = .001$, eta = .93, $f = 2.58$).

Please Insert Figure 5.52

Finally, the four-way interaction among MLU, *Wh*-word, Inversion and Matrix/Embedded factors was significant ($F(10, 95) = 2.49, p = .011$, eta = .46, $f = .51$). Separate analyses of the three MLU groups revealed that the interaction among *Wh*-word, Inversion, and Matrix/Embedded was only significant for the High MLU group ($F(5, 30) = 3.09, p = .023$, eta = .58, $f = .72$). If one compares Figure 5.53 (which depicts the data for all questions except *how come*) with Figure 5.54 (which depicts just the *how come* data), it appears that this interaction may be due to the more

linguistically mature children having treated *how come* questions differently from the other *wh-* questions.

Please Insert Figure 5.53 and 5.54

Inversion and *how come*. Because *how come* questions are unusual with regard to inversion, analysis of variance was performed for all data except *how come* data. With the *how come* data removed, Inversion had a significant main effect ($F(1, 19) = 4.88, p = .040$). Children judged inverted questions to be good 75.9% of the time, whereas they judged noninverted questions good only 68.4% of the time. In addition, the interaction between *Wh*-Question and Matrix/Embedded Question became just barely significant ($F(4, 76) = 2.51, p = .049$). The interaction between Inversion and Matrix/Embedded was still significant without the *how come* questions, though surprisingly the effect size was somewhat diminished ($F(1, 19) = 31.728, p < .0005$, eta = .79, $f = 1.67$). This is surprising because *how come* questions are the only type of questions that don't invert in matrix questions. I expected that by removing *how come* questions from the analysis noise would be removed from the data and the effect size of the interaction would increase. The interaction among MLU, Inversion and Matrix/Embedded was significant without the *how come* data, and the effect size was larger than in the analysis that included *how come* data ($F(2, 19) = 7.52, p = .004$, eta = .66, $f = .88$). Perhaps most significantly, without the *how come* data, the interaction among MLU, *Wh*-word, Inversion, and Matrix/Embedded clause was not significant ($p > .25$). This supports the suggestion made

above that the *how come* questions may have accounted for the significant four-way interaction.

The analysis of variance of just the *how come* questions data failed to reveal a significant interactions among MLU, Inversion and Matrix/Embedded Question. However, there was a nonsignificant developmental trend in the predicted direction (shown in Figure 5.34 above). Children in MLU group 1 judged inverted matrix *how come* questions to be grammatical almost 2.5 times as often as noninverted matrix questions. MLU group 2 judged inverted matrix *how come* questions to be good almost 1.5 times as frequently as not inverted questions. The most linguistically mature children, however, judged inverted matrix *how come* questions grammatical 75% less often than inverted questions.

How come questions only exhibit anomalous inversion behavior for matrix questions. Therefore, one would expect no interaction between MLU, *Wh*-word, and Inversion if one looked only at embedded questions. An analysis of variance of just the embedded questions revealed no significant interaction between MLU, *Wh*-word and Inversion ($p < .801$). As predicted, MLU, *Wh* -word and Inversion did interact significantly when only data for matrix questions were analyzed ($F(10, 90) = 1.93$, $p = .050$, eta = .41, $f = .45$).

5.6 Discussion

In this section, I will briefly highlight and interpret the importance of the more crucial results, saving more general discussion for the end of the chapter. First, the experiment worked as a means of eliciting

grammaticality judgments from the children. If we use adult judgments as the standard of comparison, the children liked more of the grammatical sentences than the ungrammatical sentences. Even the children in the Low MLU group didn't respond randomly. They judged *how come* questions to be significantly worse than any of the other types of questions. One might argue that the Low MLU children were merely responding to the semantics of the different *wh-* questions. If this were true, then we would expect that the Low MLU children would have also judged the *why* questions bad because *why* and *how come* are near synonyms. It appears, therefore, that the children in the Low MLU group were responding to the syntactic markedness of a *how come* question that contained a non-negated auxiliary.

In addition to responding preferentially to certain *wh*-questions, even the children in the Low MLU group liked the grammatical questions more than they liked the ungrammatical questions. This is evident from the marginally significant interaction between Inversion and Matrix/Embedded Questions. As Figure 5.32 shows, in each successive MLU group, the number of grammatical questions that are judged good and the number of ungrammatical questions that are judged bad increases. More simply, the children in the High MLU group did better than the children in the Middle MLU group, who in turn did better than the children in the Low MLU group.

I would now like to turn to two surprising results of the experiment. The first is that, the children liked more inverted embedded questions (e.g., **he wonders who is Ernie.*) than noninverted matrix questions (e.g., **who Ernie is?*). In the next section, I will compare this experimental finding with the findings of the transcript analyses and offer a possible explanation for the results from both studies.

The second surprising result surrounded children's judgments of *how come* questions. Basically, the children didn't like *how come* questions (though only children in the Low and Middle MLU groups had a statistically significant distaste for them). However, if we look just at the children in the Low MLU group, we find that of the four types of *how come* questions they most preferred the inverted matrix questions. (They thought **how come is Bert mad?* was good 50% of the time.) Children in the Middle MLU group most preferred the noninverted embedded question (*he wonders how come Bert is mad* was rated good 62.5% of the time), but they still preferred matrix *how come* questions which were inverted. In fact, they disliked the grammatical *how come Bert is mad?* as often as they disliked the ungrammatical **he wonders how come is Bert mad?* (each was rated bad 68.7% of the time). It isn't until we get to the High MLU group that we find a preference for *how come* questions that are not inverted (50% good for noninverted versus 29% good for inverted versions). However, even the children in the High MLU group did not have adult-like grammars of *how come*. Like the children in the Middle MLU group, they judged *how come Bert is mad?* as no better than **he wonders how come is Bert mad?*. The children's overall poor judgments of grammatical and ungrammatical *how come* questions contrasts sharply with the essential lack of *how come* inversion errors in the spontaneous speech of the children in the transcripts.

6. Conclusions

1. Inversion rates and individual differences. The overall rate of subject-auxiliary inversion in questions was about 92%. There were, however, significant differences among children's overall rate of inversion. These differences among children may account for some of the disparity among the inversion results reported in the literature. If a researcher was only studying a single child, she would draw very different conclusions about the frequency of inversion errors depending on whether she studied a child like Nathan (who inverted about 50% of the time) or a child like Allison (who had near perfect performance).

What accounts for the differences in inversion rates between children? Perhaps children must learn to invert and the differences in their inversion rates just reflect their different level of knowledge. If this is the explanation for the differences among the children, then we would expect that the younger children would have had the lower inversion rates and the older children would have the higher inversion rates. On a gross level, this is not the case. Nathan was not one of the younger children yet he had the lowest inversion rate of any of the children. Allison was one of the younger children and she had the highest inversion rate. In general, the children who experienced a dip in performance had lower inversion rates than the children whose performance improved in a fairly monotonic fashion. Perhaps, therefore, the question of why some children had a higher rate of inversion than others should be rephrased as a question of why some children experienced a dip in inversion rates while other children did not.

Another possibiiity is that the differences in inversion rates (or differences in developmental trends) are due to something extrinsic to the children. For example, it might have been input-related. If some of the children heard more embedded questions than others, these children might have been more likely to fail to invert in matrix questions. Because the parents' speech was not examined, all that can be said is that it is possible that differences in the input to the children might have affected their rates of inversion.[31]

Other possible causes are intrinsic to the child but extrinsic to the linguistic system. For example, some children might, for some reason or other, be more willing to say things they are unsure of. Children might vary in their ability to detect abnormalities in what they say. Children,who didn't or couldn't monitor what they say, might make more errors (similar to the finding that adults make more speech errors when they can't hear their own speech). Children could have differences in their short-term memory. Such children might be more prone to get "stuck" in the middle of a sentence having forgotten how the sentence started and, hence, produce ungrammatical sentences. Other causes might be intrinsic to the language system but still extrinsic to the grammar or I-language per se. For example, some children might have better production systems than others.

Children could also have differing criteria for when they are willing to generalize within the linguistic domains. If this is true, we would expect

---

[31] In their study of the effects of maternal input on language development, Newport, Gleitman and Gleitman (1977) found that one of the only aspects of development that was related to input was the acquisition of auxiliaries. Newport *et al.* found that children who heard many *yes/no* questions acquired auxiliaries earlier. Since then a number of researchers have documented the effects of different types of input on the acquisition of auxiliaries. See Hoff-Ginsberg and Shatz (1982) and Shatz, Hoff-Ginsberg, and MacInver (1989) for a review of this literature.

to find a certain sort of linguistically error-prone child (perhaps a child who passes the *wug* test with flying colors but says things like *eated* all the time). Finally, as heretical as it might seem, children are bound to vary to some extent in their native abilities to deal with language. This is not contrary to the nativist position, nor is it a threat to Universal Grammar. What it does suggest is that whatever universals we propose to build into the child must be robust enough that a child who has a budget-class language ability will learn the language essentially as quickly and as completely as the child who is a budding linguist.

Deciding what differences among children cause some children to invert 100% of the time and other children to invert 50% of the time (or why some children get worse and others do not) is clearly beyond the scope of this paper. However, it is worth noting that the fact that there can be such variation between children suggests something about the nature of subject-auxiliary inversion. To the best of my knowledge, no one has reported gross variation in the frequency with which children do things such as put determiners before nouns in noun phrases. There is no variation because all children put the determiner before the noun phrase 100% of the time.

The only other type of error that comes to mind that displays as much individual variation as inversion errors is errors of overregularizing past tense verbs (i.e., saying *eated* for *ate*, etc.). What might inversion and irregular past tense forms have in common? One thing that they might share is that they may both be relatively superficial properties. One could argue that this superficiality is just a reflection of the fact that what verbs invert or what verbs have irregular past tense forms is lexically determined. One can notice overall patterns in which verbs invert (e.g.,

verbs which carry information about tense, mood or aspect and precede a main verb, etc.) and which verbs have irregular past tense forms (e.g., verbs that end in *d* or *t*), but whether a particular verb inverts or has an irregular past tense form is lexically determined.

2. Inversion in yes/no and *wh*-questions. The overall inversion rate was 93.0% for *wh*-questions and 93.7% for *yes/no* questions. However, some children inverted more in *yes/no* questions than *wh*-questions, some inverted more in *wh*-questions than *yes/no* questions, and some inverted equally often in both types of questions. It doesn't seem like which pattern a child adopts is age-dependent. Hence, there is no support for the hypothesis that there is an early stage during which children invert in *yes/no* questions but not *wh*-questions.

What might lead a child to exhibit each of these patterns? Perhaps, Chomsky is right and the children who inverted more often in *yes/no* questions did so (for whatever reason, linguistic or not) in order to distinguish their *yes/no* questions from declaratives. Because *wh*-questions are unmistakable whether they are inverted or not, these children might not have felt as compelled to invert in *wh*-questions. On the other hand, the children who inverted more in *wh*-questions than *yes/no* questions might have done so because one can ask a grammatical *yes/no* question merely by using intonation (e.g., *I can eat now?*), whereas the same is not true for *wh*-questions (e.g., **what I can eat?*).[32] The children might have adopted

[32] Parenthetically, Ultan (1978) did a survey of how diverse languages mark questions. He found that inversion was a common device for marking questions. He also found that more languages invert in *yes/no* questions than *wh*-questions. This may reflect the necessity to do something to make *yes/no* questions distinguishable from declarative sentences. It could, however, reflect some universal property of *yes/no* questions.

this strategy because they heard noninverted *yes/no* questions but not noninverted *wh*-questions, because they found inversion onerous and only did it in order to sound right, etc.

Given that there was no difference in overall inversion rate or in average inversion rate for *yes/no* and *wh*-questions and given that no pattern of responding predominated among the children, the variation between children is probably due to extra-syntactic differences between children (such as the ones suggested above) rather than to differences in their grammars as Bellugi claimed.

3. Inversion rates for the 7 types of inverting *wh*-questions. Do children treat all *wh*-words the same with regard to subject-auxiliary inversion? There are three logical possibilities. The first is that children have a single inversion rule that they apply regardless of the *wh*-word and some other constraint prevents inversion in *how come* questions. The second possibility is that they are conservative and decide whether or not to invert subject and auxiliary individually for each *wh*-word based on what they hear. The third possibility is that there are two (or more) groups of *wh*-words, one group triggering subject-auxiliary inversion while the other one doesn't. An example of this third possibility is that children might decide that all *wh*-words except *how come* trigger inversion. Alternatively, they might form semantic groups with *how come* and why in one group and all the other *wh*-words in the other group.

There might be inversion for groups of *wh*-words that cut across semantic boundaries and shared some syntactic or structural properties. The children wouldn't form these syntactic *wh*-groups, consciously inverting with one syntactic group but not the other. Rather, the

interaction of the principles of their grammar would mean that some *wh-questions* would have inversion and others would not. This would give the illusion that children had two classes of *wh*-words, some of which triggered inversion and some of which did not.

LFG and GPSG predict that inversion will be learned *wh*-word-by-*wh*-word (the first option). The principles-and-parameters/Government-Binding approach predicts that as long as a clausal-scope *wh*-word is definitely in the SPEC position of CP, the choice of *wh*-word should not affect inversion (the second option). Therefore, according to the principles-and-parameters approach, any differences in inversion rates for particular *wh*-words must reflect either differences in the location of *wh*-words (e.g., some *wh*-words might be in TOPIC position rather than in SPEC of CP) or in the nature of *wh*-words (perhaps some *wh*-words do not have clausal scope).

Like the results of all of the previous studies which have been designed to examine this question, the results of my transcript analyses suggest that there are significant differences in the rates of inversion for the different *wh*-questions. In my study as in previous studies, the children tended to invert most frequently for *who*, *what*, *where*, and *how*. They inverted less for *which*, *when*, and *why* (and of course *how come*).

What factors affect which *wh*-questions have high inversion rates and which *wh*-questions have low inversion rates?. One possibility is that the longer a child has been using a particular type of *wh*-question, the more likely that child is to invert that type of *wh*-question. To a first approximation, this seems to be true. In previous work (Stromswold, 1988a), I found a significant ($p < .05$) negative correlation between the

age at which children acquire a particular type of question and the rate of inversion for that particular type of question.[33]

Of course, the question is, why do children acquire some types of *wh*-questions before others? In Stromswold (1988a and 1988b), I argued that the differences in government of argument and adjunct *wh*-traces cause children to acquire argument questions before adjunct questions. What's of particular interest for the current discussion is that children inverted more frequently for argument questions than adjunct questions. The overall inversion rate for argument questions was almost 97% whereas for the adjunct question the overall inversion rate was 85%. Interestingly, children inverted more often for argument *where* questions than adjunct *where* questions (99% inversion for argument *where* questions versus 84% inversion for adjunct *where* questions). The same held for argument and adjunct *how* questions (100% inversion for argument *how* questions versus 96% for adjunct *how* questions).

The developmental inversion trend for argument and adjunct questions is shown in Figure 5.55 and the inversion trend for argument and adjunct *where* questions is shown in Figure 5.56. Notice that performance on argument questions remained fairly constant, whereas performance on adjunct questions dropped to about 60%. For argument*where* questions, children began inverting 100% of the time at about age 2;0 and continued at this level of performance through age 5;0. The children did not begin asking scoreable adjunct *where* questions until almost age 3;0 at which time

---

[33] This correlation is not simply a reflection of children using a wh-question for a while before they 'learn' to invert it. In Stromswold (1988a, 1988b), for each type of *wh*-question, I graphed the percent inversion at 3 month intervals. I found that initially children inverted in essentially 100% of the time and later, started to make inversion errors. In other words, I found a U-shaped curve in for inversion.

they inverted adjunct *where* questions about 50% of the time. The inversion rate for adjunct *where* questions fluctated quite a bit, but overall was considerably lower than the inversion rate for argument *where* questions.

Please Insert Figures 5.55 and 5.56 here

The observation that the children inverted argument *where* and *how* questions more than the corresponding adjunct questions is consistent with the notion that there is something fundamentally different about argument and adjunct questions. In other words, it is more consistent with a GB-like theory than a lexical theory like LFG or GPSG.

I would like to propose one possible structural explanation for the difference in inversion rates for argument and adjunct questions. Perhaps the difference might be that in argument questions, the *wh*-word must be in the SPEC position of C, whereas (at least in some) adjunct questions, the *wh*-word may be in a pre-SPEC position (perhaps TOPIC position). If having an element with clausal scope in the SPEC of CP triggers inversion, then the overall inversion rate would be higher for argument questions than adjunct questions.

Let me flesh this argument out. Within the Government-Binding framework, in order for a *wh*-trace to be properly governed, its *wh*-word must be in SPEC position. For argument questions, the *wh*-trace must be left in the original *wh*-word position, or else the theta-criterion is violated. Thus, in argument questions, there must be a trace in the original *wh*-word

position, and this *wh*-trace must be properly governed by its *wh*-word which must, therefore, be in SPEC position.

As is the case with argument questions, in adjunct questions, the Empty Category Principle (ECP) requires that if there is a *wh*-trace in a sentence, it must be properly governed. However, because an adjunct *wh*-word is not used to meet the argument requirements of a verb, if the *wh*-trace is missing, the theta criterion is not violated. Therefore, for adjunct questions, it is possible for the adjunct *wh*-word to be directly generated in clause-initial position, leaving no *wh*-trace to govern. By doing this, neither the theta criterion nor the ECP are violated. However, there seems to be a general tendency (perhaps a requirement; see Chomsky, 1989), that any operator in SPEC must bind a variable. Therefore, we would not want (could not have) the *wh*-adjunct to be in SPEC position. Instead, we would want it in a pre-SPEC position.

Now, if what triggers verb movement is having the proper element in SPEC position, then we would predict that all argument questions would trigger inversion, whereas only those adjunct questions in which the adjunct *wh*-word is in SPEC would trigger inversion. This may not be the right way to account for the difference in inversion rates for argument and adjunct questions. It demonstrates, however, that it is possible to account for the difference in inversion rates within the Government-Binding framework. It is not clear that it would be possible to formulate an account within most lexicalist theories.

4. Inversion rates for the auxiliary clusters. Like Klee (1985), Kuczaj and Maratsos (1983), and Labov and Labov (1978), I found that some auxiliaries were more likely to be inverted than others.

Furthermore, I found that the inversion rates of auxiliaries cluster according to those groups of auxiliaries described in chapter 4. The inversion rates were significantly different for the various auxiliary clusters. To summarize the results, non-negated forms of *do* were inverted virtually 100% of the time. The mean inversion rates for non-negated forms of *have*, the modals, and auxiliary *be* were all in the 90% range. The mean inversion rate for non-negated forms of copula *be* was about 84%.

Let me review the predictions made by three linguistic theories discussed above. LFG and GPSG predict that inversion will be acquired one auxiliary at a time as the child appends the +INVERT feature to each auxiliary. Hence, differences in inversion rates are to be expected within LFG and GPSG frameworks and no particular pattern of differences is privileged. Within the principles-and-parameters framework of GB , there is only one V-raising rule so any difference in inversion rates for the various auxiliaries must reflect important differences in the syntactic behavior of those auxiliaries. If the inversion rates vary for different auxiliaries, these differences should pattern along the natural divisions made within Government-Binding theory (e.g., the inversion rates for all the modals should be the same, etc.).

It would appear, therefore, that differences in inversion rates for the various auxiliaries are most consistent with a principles-and-parameters approach. However, the differences in inversion rates found are exactly what would be predicted for learnability reasons. Thus, the relative rates of inversion for the various auxiliaries do not prove that the latest Chomsky-Pollock version of Government-Binding is correct. They are merely consistent with this recent version of Government-Binding.

I would like to argue that the reason non-negated *do* is inverted 100% of the time is that the only reason *do* is in the question is in order that it can invert. If there is no inversion, there is no need for *do*. In a sense, the inversion rate for *do* is inflated because it is only invoked when inversion is imminent. The 90% inversion rate for auxiliary *be*, *have* and the modals represents the "real" inversion rate for auxiliaries. It is approximately the same for all of these auxiliaries because children have a single auxiliary-raising rule. Once children identify a particular word as an invertible auxiliary, they will invert it equally often regardless of which particular auxiliary it is. In other words, children may begin inverting the auxiliaries at different times (as the results in chapter 4 indicate they do) but once they begin to invert an auxiliary (i.e., once they realize they can invert that auxiliary), the rate of inversion should be the same for that auxiliary as for the first auxiliary they ever began inverting.

Now we come to inversion and copula *be*. Here I would argue that the children in some sense have a more sensible grammar than that found in adult English. Children noticed that copula *be* can do things that only main verbs can do. For example, they might have noticed that copula *be* can be the sole verb in a non-eliptical sentence (e.g., *I am happy*) and that copula *be* can precede an object NP (e.g., *I am a big boy*). Based on these observations, they might conclude that copula *be* is a main verb and, therefore, be reluctant to raise it.[34]

In addition to the differences in inversion rate for copula *be*, there are two other findings which suggest that children may have treated copula *be* as a main verb. The first finding,which was mentioned briefly in

[34] In Pollock's model, the children would conclude that copula *be* was a theta-marker, and therefore, could not be moved.

chapter 2, is the existence of 16 utterances which contained copula *be* and also concomitant *do*-support. There were 9 cases that involved *do*-support for inversion (e.g., Adam39 (3;10): and does dese be in dere?) and 7 cases that had *do*-support for negation (e.g, Adam10 (2;7): you don't be quiet). There were no such examples for auxiliary *be*. I suggest that the reason such examples exist for copula *be* but not auxiliary *be* is that the children did not doubt that auxiliary *be* is an auxiliary and, therefore, could invert and negate. They were not as confident about the invertibility or negatability of copula *be*, however, because in some ways it behaves like a main verb and not an auxiliary.

The second finding which suggests that children either considered copula *be* a main verb or were at least confused about its status are the examples of overregularized *be* discussed in chapter 3. There were 10 cases of overregularization of *be*. All 10 involved copula *be* (e.g., Ross37 (3;3): she bees bad [#] then she bees good). If we assume that children can only apply the general past tense formation to open class verbs, then the fact that the children overregularized copula *be* (and crucially never auxiliary *be*), is evidence that at least some of the time the children treated copula *be* as an open class main verb.

5. Inversion rates for negated and non-negated auxiliaries: why they aren't the same. Like Bellugi (1971), Derwing and Smyth (1988), Erreich (1984), and Labov and Labov (1978), I found that children inverted non-negated auxiliaries more often than they inverted negated auxiliaries. The difference in inversion rates for negated and non-negated auxiliaries was significant and large (56% inversion for negated and 91% inversion for non-negated auxiliaries). In addition to

there being a significant overall difference in inversion rates for negated and non-negated auxiliaries, there was a marginally significantly effect of auxiliary type. The difference was greatest and most significant for the modals (34% versus 90%), followed by *do* (60% versus 99%). The same trend held for *be* (81% versus 93%), though it was not significant.[35] (The children did not ask any questions with negated *have* .) Why was the inversion rate for negated auxiliaries significantly less than the inversion rate for non-negated auxiliaries?

GPSG and LFG would not predict any systematic difference in the inversion rates for negated or non-negated auxiliaries. Whether a particular auxiliary is inverted depend on whether that auxiliary has the +INVERT feature yet. There is no reason to expect this to systematically occur later for negated auxiliaries than non-negated auxiliaries.

Within the principles-and-parameters approach, one might explain why the inversion rate was lower for negated auxiliaries than non-negated auxiliaries. In order for the negation marker *-n't* to attach to the auxiliary or copula, the element either has to move to NegP or the *-n't* has to move to the auxiliary or copula.[36] One possibility is that when an auxiliary attaches to *-n't*, the trace of the auxiliary that is left in NegP has meaning at Logical Form level and, therefore, cannot be deleted. Under certain configurations, the trace might be ungoverned and, therefore, there would be an ECP violation (please see Chomsky, 1989 and Pollock, 1989).

---

[35]Perhaps the reason the trend was not significant for *be* is that, because of the small number of questions with negated *be*, I was forced to combine copula and auxiliary *be*.

[36]The generally accepted view is that *-n't* is a clitic. However, Zwicky and Pullum (1983) argue fairly convincingly that *-n't* behaves more like an inflectional affix than a clitic. For the purposes of this argument, the distinction is not crucial.

Another possibility is that children might perceive a filled NegP as a barrier. A third possibility is that children incorrectly consider negated elements in C position to have clausal scope, and they do not like sentences which have two items with clausal scope. Which, if any, of these possibilities is right is not important. What is interesting is that within the principles-and-parameters framework, it is possible to formulate structural differences between questions with negated and non-negated auxiliaries that would account for the finding that negated auxiliaries invert less.

Can the Chomsky/Pollock model account for the negation effect being larger for modals than for *do* or *be?* Within the Chomsky/Pollock model, *be*, *do*, and *have* all move to NegP where they cliticize with the negation element *-n't*. Negation of modals differs from negation of *be*, *do*, or *have*. For a modal to be negated, either the modal must move down to NegP and then back to I, or the *-n't* element must move to I. Thus, not only can the principles-and-parameters approach account for a systematic difference in inversion rates for negated auxiliaries, given certain assumptions about the relative ease of various forms of movement, the Chomsky/Pollock model can account for the difference being greater for modals than for *be*, *do*, or *have*.

6. How come children don't invert in *how come* questions but they think inverted *how come* questions are good? For unclear reasons, subject-auxiliary inversion is not acceptable in *how come* questions even though it is required in all other types of matrix *wh*-questions. The results of the transcript analyses suggest that the children realized and obeyed this distinction. At very most, children incorrectly

placed the auxiliary before the subject in only 14% of *how come* questions, rather than in the 90% of *how come* questions that they would have if they had treated *how come* questions like ordinary *wh*-questions. If, for the reasons outlined in Section 4.5.6 above), we reject all 5 examples of inverted *how come* questions, then we could argue that the children in the transcripts never inverted in matrix *how come* questions. Whether the children made inversion errors in 0% or 14% of matrix *how come* questions, the picture one gets from the transcript analyses is that the children had essentially mastered the intricacies of *how come* questions.

This is not the picture one gets if one looks at the performance on *how come* questions in the judgment experiment. The children in the Low and Middle MLU groups liked **how come is Bert mad?* more than they liked *how come Bert is mad?* Even the children in the highest MLU group (average age 5;0) did not have nearly the competency on *how come* questions that they had on the other *wh*-questions.

Why do the children in the transcript study appear so competent while the children in the judgment experiment appear so incompetent at *how come* questions? The difference is probably due in large part to differences in the methodologies employed. In the transcripts, all the children had to do was talk, while in the experiment children had to judge the grammaticality of many types of matrix and embedded *wh*-questions, a minority of which were *how come* questions. We know that children's ability to judge constructions lags considerably behind their ability to produce these same constructions (deVilliers and deVilliers 1978). Thus, it is possible that this is the chief cause of the difference in performance of the children in the transcript study and the children in the judgment experiment.

I think, however, that another factor is involved. In spontaneous speech, children can simply choose not to ask *how come* questions. There is some evidence that the children in the transcript study adopted this strategy. *How come* questions were the most rarely-asked type of *wh*-question. Furthermore, 75% of the time they asked *how come* questions, the children in the transcripts avoided the whole issue of inversion by not including an auxiliary. Seventy-five percent of the *how come* questions which contained auxiliaries had a negated auxiliary (either *do* or a modal).[37] Given that the frequency of inversion for negated *do* and modals is about 60%, even if the children had treated *how come* questions just like the other *wh*-questions, we would have only expected them to invert in only 14 *how come* questions. The children only inverted 3 times, so clearly they did distinguish *how come* questions from the other *wh*-questions. My point is, however, that they may have looked more competent than they really were.

If everything in the transcript study conspired to make the children appear more competent with *how come* questions than they were, virtually everything about the judgment experiment served to emphasize children's incompetence. Children had to judge *how come* questions alongside *how* questions. They had to judge matrix *how come* questions alongside embedded *how come* questions.[38] Finally, all of the *how come* questions they judged were a bit odd because they contained an auxiliary and the auxiliary was not negated. The results of the two studies might be reconciled as follows: the children in the experiment probably were not as

---

[37]Ninety percent of the non-inverted *how come* questions had negated auxiliaries. Put another way, the children asked a total of two non-inverted *how come* questions that had non-negated auxiliaries (both with copula *be*, the least frequently inverted element).

[38] The children in the transcripts asked no embedded *how come* questions.

incompetent as they seemed and the children in the transcript study probably were not as competent as they seemed.

7. Similarities and differences between inversion in matrix and embedded questions. Matrix questions (except *how come* questions) require inversion and embedded questions forbid inversion.When it comes to inversion, do children treat matrix and embedded questions the same? One possibility is that children have to learn to apply subject-auxiliary inversion to matrix questions like *who is Ernie?* but not to embedded questions like *I wonder who Ernie is.* If this is the case, one would expect a stage in which children invert subject and auxiliary equally often for matrix and embedded questions. If, on the other hand, children have some innate constraint which prevents inversion from occurring in embedded questions (or if the principles of grammar conspire to make inversion impossible in embedded questions), they should invert in matrix questions but never invert in embedded questions.

The results of the transcript analyses and the judgment experiment indicate that children distinguish matrix from embedded questions and know to invert in the former but not in the latter type of question. In the transcript studies, children inverted in 90% of matrix questions but only 10% of embedded questions. In the judgment experiment, children significantly preferred matrix questions that were inverted and embedded questions that were not inverted.

What initially seems surprising is that the children in the judgment experiment judged inverted embedded questions (e.g., **he wonders who is Ernie*) as being significantly better than noninverted matrix questions (e.g.,

**who Ernie is?*).This finding is only surprising if we assume that children have some constraint which prevents them from even considering inversion in embedded questions (e.g., the Structure Preserving Constraint) or some grammatical principle that makes inversion in embedded questions impossible (e.g., the filled COMP theory) but no similar constraint which requires inversion in matrix questions.

If we base our expectations solely on the results of the transcript studies, we might predict a relative preference for inverted embedded questions over noninverted matrix questions found in the judgment experiment. Notice that in the transcript study, children made mistakes in 10% of the embedded *wh*-questions but only 7% of the matrix *wh*-questions. It appears that, relative to the number of opportunities for errors, inversion errors are more common in embedded questions than matrix questions. If this is so, why is it the general perception in the acquisitional literature (e.g., Erreich, 1984, Pinker, 1984) that failing to invert in root questions is as an everyday occurrence, while inverting in embedded questions is a rare, reportable event?

I think the reason for the misconception is that of an availability bias. In terms of absolute numbers, noninverted matrix questions are much more common than inverted embedded questions. There are two reasons for this. First, children ask many more matrix questions than they do embedded questions. Second, embedded *yes/no* and *wh*-questions do not require auxiliaries while all matrix *yes/no* questions and most matrix *wh*-questions (all but subject questions and *how come* questions) require an auxiliary. If a question does not have an auxiliary, then there can't be an error in auxiliary placement. Consequently children have more opportunities to make inversion errors in matrix questions than embedded

questions. In the transcripts analyzed, children asked about 25 matrix questions with auxiliaries for every embedded question they asked which contained an auxiliary. Thus, in terms of shear numbers, non-inverted matrix questions are obviously much more common than embedded questions.

Turning now to the question of why children occasionally invert in embedded *wh*-questions and matrix *how come* questions but never in embedded *yes/no* questions, I would like to briefly review the possible explanations presented in Section 4.5.7. The first possibility considered was that children use semantics and pragmatics to determine when to invert. However, this possibility doesn't seem right, because if they did, they would have inverted in 100% of *how come* questions. They also would have inverted in any and only those embedded questions (both *wh*-questions and *yes/no* questions) that were question-y. They did neither of these things.

I suggested and rejected the possibility that all of the purported examples of inversion in embedded questions were actually improperly transcribed examples of quotative questions. I next suggested that children had some constraint that prevented them from inverting in embedded questions but that something went awry 10% of the time. Perhaps children have an innate principle that prohibits inversion in embedded questions but they do no always correctly distinguish embedded questions from matrix questions. It's also possible that they correctly distinguished matrix from embedded questions 100% of the time, but the constraint is probabilistic. All three explanations correctly predict that the inversion ranks for matrix and embedded questions were highly and significantly correlated.

None of the three explanations can account for the lack of inversion errors in embedded questions.

Now it's possible that the reason there were no inversion errors in embedded *yes/no* questions is just a reflection of the fact that there were only 46 embedded *yes/no* questions that contained auxiliaries. If the error rate was the same for embedded *yes/no* questions as *wh*-questions, we would only have expected to find 4 or 5 examples of inverted embedded *yes/no* questions. Perhaps this is just one of the 5% of the times that no examples would be found.

8. Inverting (?) uninvertible verbs. How does the child decide which verbs to invert? The approach suggested by LFG and GPSG is that children conservatively append a + INVERT marker to a lexical entries on an verb-by-verb basis when they hear a verb inverted. If this is what children do, they should never invert main verbs or auxiliaries that aren't inverted.[39] The second possibility is that because verbs and auxiliaries have similar properties and privileges of occurrence, children might form a general subject-verb inversion rule. If they did, they would invert main verbs and uninvertible auxiliaries.

There is a third option, somewhere between the first two options. Perhaps children are generally conservative but are willing to make certain generalizations (e.g., generalizing that *did* can invert from the inversion of *do*). For learnability reasons, let us assume that all syntactic rules (or the

[39] Unless, of course, children are unable to distinguish between auxiliaries and their homophonous main verbs. The verb-by-verb approach only works if children can do this. It is not at all obvious how children accomplish this task.

intersection of principles that imitate rules) are obligatory. In other words, let us assume that there is no rule R which can arbitrarily be applied or not. Let us further assume that children innately know this. Given this assumption, upon exposure to questions like *did she eat?*, children will automatically know that not all verbs can invert. If they find that some but not all elements in their language can invert, they will be forced to look for distinctions that characterize those verbs which can invert. If we assume that they also know innately that there can potentially be two classes of verbs, one being the main verbs and the other being the auxiliary verbs, they can then identify the potentially invertible elements by the special semantics and syntactic properties which auxiliary verbs exhibit (see Steele, *et al*, 1981).

Depending on what semantic and syntactic properties they use to distinguish main verbs from auxiliary verbs, they may misclassify certain verbs and, therefore, make mistakes. The four putative examples of inversion of deontic *go* may be the result of children using certain semantic properties (e.g., expression of tense, mood or aspect) and syntactic properties (can precede a verb) as criteria for invertibility. The 16 cases where children used *do*-support with copula *be* are evidence of the children using semantic and syntactic properties to determine auxiliary-hood. Occasionally *go* passes the auxiliary-hood test and occasionally copula *be* fails the auxiliary test. The result are the *go* and copula errors described. If this is true, then why didn't the children invert *better* or *gonna* or *hafta*. For *better*, perhaps the reason might simply be that they never used *better* in a question. The more general explanation might be that *hafta*, *gonna*, and *better* are not the head of V and, therefore, cannot raise. More

generally, they may have known that only the head of a VP (or perhaps only a +tense element) can move.

I would like to propose a possible account for why certain errors appear in the transcript while other errors do not. This account explains the lack of embedded *yes/no* questions, the lack of matrix *yes/no* questions introduced by *if* or *whether*, the essential absence of inverted *how come* questions or inverted main verbs. In addition, it accounts for the existence of non-inverted matrix questions, inverted embedded *wh*-questions and the high degree of correlation between inversion in matrix and embedded versions of particular *wh*-questions. What I propose is that children never utter clauses that are not subsets of possible clauses in their language. Thus, the reason there were no embedded *yes/no* questions is that the string *if/whether can she go* is never found in English. The fact that this string is never found in English also accounts for the nonexistence of errors like the matrix question **If can she go?*

Children inverted in embedded *wh*-questions, but not in embedded *yes/no* questions, because they heard clauses introduced by *wh*-words that had subject-auxiliary inversion, mainly matrix *wh*-questions. This could explain why for the different types of *wh*-words, the frequency of inversion in embedded and matrix questions is highly correlated.The reason they sometimes fail to invert in matrix questions is that they hear noninverted embedded questions.

The reason children never (or essentially never) invert in *how come* questions is that they never hear *how come* clauses which have inversion. The same explanation can be given for the lack of inversion of main verbs.

Children never (or essentially never) invert main verbs because they never hear clauses which have the main verbs inverted.

Now, in a real sense, this account is circular. It isn't enough to say that children don't say something because they don't hear it. The real question is why children don't hear certain clauses. A much more satisfying account would explain why certain types of clauses are impossible in any setting. Presumably, such an account would appeal to the principles of the grammar. However, it is worth noting that children do not use clauses that do not appear in the language. Perhaps the explanation is as simple as the purported inversion errors are just speech or transcription errors. Perhaps, the children pay extraordinary attention to what they hear and say only what they hear. Or, perhaps the reason the children only make errors that result in possible clauses because their grammar is unable to generate the impossible clauses.

In summary, auxiliary placement is the sort of linguistic phenomenon that phrase structure systems like GPSG can easily describe but that a principles-and-parameters system like GB has difficulty with. While a phrase structure system can describe auxiliary placement by appending features like +Invert to lexical items, it is hard to imagine how the child would learn such a system. In this paper, I proposed some ways that a principles-and-parameters system could explain auxiliary placement. I analyzed transcripts and collected judgment data from children to see whether acquisition is more consistent with a phrase structure account or a principles-and-parameters approach. The results are mixed but are generally in favor of a principles-and-parameters approach. Children seem to know certain things innately (e.g., only the head of the VP can move) and look for some interactions (e.g., the interaction between clause type

and inversion) while ignoring others (e.g., the interaction between *wh*-word and clause type).

Chapter 6

Tense and Auxiliaries: Difficulties with *Do*-Support

Questions which contained an auxiliary and a tensed main verb and questions which contained two or more tense-bearing auxiliaries were among the most common type of auxiliary errors found in the transcripts (see chapter 3). In this chapter, I examine these double-tensed questions more carefully. In addition, I will also examine a second general type of error which I did not mention in chapter 3. Sometimes in the transcripts the children asked questions that lacked an obligatory auxiliary. Therefore, in this chapter, I examine what children do when they ask questions that have extra auxiliaries or no auxiliaries.

I begin by examining more carefully the actual examples of double-tensed questions found in the transcripts. Next, I analyze those questions the children asked which lacked obligatory auxiliaries. I then describe the results of a judgment experiment which corroborate the results of the Double-Tensed Questions and Auxiliary-less Questions Transcript Analyses.

I will argue that auxiliary-less questions are much more common than double-tensed questions. I will further argue that the auxiliary-less questions were not simply production errors but, rather, were the result of the children failing to invert (probably connected to difficulties with *do*-support). I will argue that double-tensed questions were usually just production errors, though some non-negligible fraction of the double-

tensed questions probably resulted from the children's lack of mastery of *do*-support.

1. Double-tensed questions

As was the case with errors of inversion, a number of researchers have noticed the occurrence of double-tensed questions (Fay, 1978; Goodluck and Solan, 1979; Hurford, 1975; Kuczaj, 1976, Maratsos, 1984, Maratsos and Kuczaj, 1978; Mayer, Erreich and Valian, 1978; Prideaux, 1976). As was also the case with inversion errors, these researchers are at odds as to the frequency and significance of double-tensed questions.

Klima and Bellugi (1966) noted that Adam occasionally made errors such as *does it rolls?*. Ten years later, Hurford (1975) noted that between the ages of about 1;10 to 2;6, his daughter "frequently asked" questions such as:

> What's this is?
> Whose is this is?
> What did you bought?
> What did you did?
> Did you came home? (Hurford, 1975, p. 300)

Hurford (1975), Fay (1978) and Mayer, Erreich and Valian (1978) argue that double-tensed questions are evidence that the "movement" transformation that moves the auxiliary to sentence initial position is actually composed of two basic operations, a "copy" operation (which

copies the element into its S-structure position) followed by a "delete" operation (which deletes the element from its D-structure position). They argue that double-tensing errors result when children correctly perform the copy operation but fail to perform the delete operation.

Maratsos and Kuczaj (Kuczaj, 1976; Maratsos & Kuczaj, 1978, and Maratsos, 1984), Prideaux (1976), and Goodluck and Solan (1979) do not deny that double-tensed questions exist. They argue, however, that compared to how frequently double-tensing errors could occur, these errors are surprisingly rare. Furthermore, they point out that some copy-without-deletion errors are never reported. For example, the copy-without-deletion explanation predicts that children should produce *wh*-questions with the *wh*-word *in situ* (e.g., **I did what yesterday?*) and *wh*-questions with two *wh*-words, one *in situ*. and the other in clause initial position (e.g., **what can I eat what?*). Because the number and type of copy-and-delete errors are so limited, they conclude that there is no evidence to support a copy-delete view of movement transformations.

Maratsos and Kuczaj (Maratsos, 1984, Maratsos and Kuczaj, 1978) further note that double-tensing errors often involve irregular verbs (e.g., Hurford **did you came home?* shown above). They suggest that these errors are not due to copying-without-deletion. Rather, these errors are the result of children not knowing that *came* is a past tense verb (hence the reason for the existence of **camed*). If *came* is a present tense verb in children's lexicons, then *did you came?* is not really an example of a copying-without-deletion error. Maratsos and Kuczaj also point out that many of the purported examples of copying-without-deletions (perhaps all of the examples in declarative sentences) involve *didn't*. They argue that

the reason children say things like **didn't I kicked the ball?*or **I didn't kicked the ball*, is that children think *didn't* is simply a negation element.

As a way of summarizing the previous literature on double-tensing errors, I would like to point out two things. First, no one denies that children make double-tensing errors. Second, no one denies that children seem to make more of these errors in some contexts (e.g., questions with irregular past tense verbs) than others. The debate concerns the significance of these errors. And the significance of these errors depends critically on their frequency. It is not surprising that many of the double-tensed questions involve irregular past tense verbs. Many of the first verbs children use have irregular past tense forms. If it turns out that all or nearly all of the double-tensed verbs involve verbs with irregular past tense forms, this would support Maratsos and Kuczaj's explanation. It's also not surprising that many of the examples of double-tensing involved negated *do* because negated forms are *do* are among the more common auxiliaries in the speech of young children. So here again, the question is one of the frequency of *do+n't* errors, relative to their overall frequency in the transcripts.

Transcript Analysis I: Double-Tensed Questions

1. Overview

In this analysis, I examine the frequency of various sorts of double-tensed questions.

2. Subjects

I used the same transcripts from the same 14 children that were used in the Transcript Error Analyses in chapter 3. Table 6.1 lists each of the children and the overall number of questions asked by each child. Notice that the total number of questions asked by all the children was 40,600, with individual children asking from zero questions (May) to over 10,650 questions (Adam).

Please Insert Table 6.1 here

---

3. Corpora

I used the Unix utility "fgrep" to cull all of the lines which had a *wh*-word or a question mark ("?"). Because I was only interested in double-tensed questions, I used "fgrep" again to eliminate all of the lines which did not include an auxiliary.

4. Tallying Procedure

I recorded all examples of doubly tensed questions. For each example, I noted:

-- whether the main verb had a regular or irregular past tense verb
-- which auxiliary or copula was used
-- whether the auxiliary or copula was contracted or not
-- whether the auxiliary or copula was negated

-- whether the auxiliary and main form had the same tense or not
--whether the utterance was transcribed in a way that suggested it was stuttered or unclear, etc.

5. Results

5.1 Double-tensed questions with two or more auxiliaries

There were 72 examples of possibly double-tensed questions that had two or more auxiliaries. All 72 examples are listed in Table 6.2. Of these 72 examples, only six had the identical auxiliary appearing twice in the same sentence (e.g., Nina33 (2;10): is this is a dog?). (These six cases are shown in Table 6.5.) A descriptive breakdown of the other 66 examples follows:

Please Insert Table 6.2 here

Negated examples. Of these 72 examples, 23 involved at least one negated auxiliary. Of these 22 examples, seven had a negated form of *do* and its corresponding non-negated form of *do* (e.g., Adam27(3;4) did I didn't mean to?). In all of seven cases, the clause internal *do* was the negated *do*. There were four examples where the two auxiliaries were the same but the transcription indicated the sentence had been stuttered or crucial words were unclear (e.g., Nathan15 (3;1): <eh ba ba-i don't don't know the rest?>). These four examples did not share much else in common with one another.

There were 11 negated questions in which the children used two different auxiliaries. In six of these 10 examples, the non-negated auxiliary was contracted. In five of the 10 examples, the sentence was stuttered or unclear (e.g., April4 (2;10): isn't [#] is Snow White down here?). The only mixed negated example that was neither stuttered nor unclear was a rather complicated question asked by Ross (Ross41 (3;6): But why aren't the good people have guns are not afraid of them?).

Mixed double auxiliary questions. If we move onto the 2-auxiliary questions that were not negated, we find that 32 of the questions had two different auxiliaries. Of these 31 examples, 12 were clearly transcribed questions with non-contracted auxiliaries (e.g., Adam29 (3;5): can dey should go down dat Massachusetts Avenue). Eleven more examples were marked as stuttered or unclear (e.g., Eve6 (1;9): where di Eve (must) find it?). The remaining nine examples involved a contracted auxiliary plus a different non-contracted auxiliary (e.g., April4 (2;10) can she's nice?).

Contracted and stuttered auxiliaries. In the remaining 14 of the 72 examples, the same auxiliary appeared twice, but one of the uses was either contracted or marked as unclear (e.g., Nathan24 (3;4): when's is dis glove?).

5.2 Double-tensed questions with one auxiliary

There were 123 examples of possibly double-tensed questions that contained only one auxiliary. These examples are shown in Table 6.3

Please Insert Table 6.3 here

As was the case for the double-tensed questions with two auxiliaries, many of the examples were somewhat ambiguous. Of the 123 examples, 64 were clearly transcribed cases in which the main verb had the same tense as the non-contracted auxiliary. In 50 of these examples, the main verb had an irregular verb form (e.g., Adam7 (2;6): what did you did?). The main verb had a regular verb form in only 14 cases (e.g., Adam22 (3;1): does it opens?). The following is a descriptive breakdown of the 123 examples of possible double-tensed questions with one auxiliary:

Irregular main verbs. Sixty-five of the 123 examples had irregular main verb forms. All of the examples but one involved irregular past tense forms of the verb. (The one exception was *does* in Eve17 (2;2): what did you does?)

Fifty of the 62 irregular verb examples were clear, non-contracted examples of the *what did you did?* sort. In six of the 62 examples, the auxiliary was either contracted or unclear (e.g., Adam33 (3;6) why d'you left them in?). Nine of the 62 examples had differently tenses for the main verb and the auxiliary (e.g., Adam36 (3;8): how does this broke? How did it break?). For convenience, in Table 6.4, I list all of the double-tensed questions with irregular verbs which were neither stuttered nor had contracted auxiliaries.

Please Insert Table 6.4

Mixed tense questions. Ten of the 123 examples had regular main verbs that differed in tense from the auxiliary. Three of these 13 cases had clear, uncontracted auxiliaries (e.g., Ross51 (4;4): Mom [#] does this looked fixed?). The remaining seven had unclear or contracted auxiliaries (e.g., Adam19 (3;0): why it(s) came off? Why its came off?).

Contracted auxiliaries. There were 23 cases in which the contracted auxiliary or copula *'s* appeared with a verb which had the third person singular present tense agreement marker *'s* (e.g., Adam24 (3;2): it's breaks.). All of these examples are like Adam's example in that they are actually declarative sentences which snuck into the analysis because they just happened to appear on a line that had a question in it. All of the examples are also like the example given for Adam in that all involved the pronoun *it*.

The remainders. The rest of the examples were a hodge-podge of different things. There were three examples that are best interpreted as two-clause utterances (e.g., Nina36 (2;10): where is the other duck went daddy duck?). There were two cases in which a child tensed an already tensed main verb (e.g., Mark84 (5;2): how does rain [#2] get maded?). There was one case in which the child used the past tense form of a verb rather than the perfective participle (Ross54 (4;8): Mom [#] have you ever saw an outlong circle like this?). Twice, children incorrectly tensed an infinitival verb (Adam19 (3;0): I can let it [?] spilled; and Ross49 (4;3): and they used to fitted you?). There were two grammatical sentences with *used* and a past tense verb form (e.g., Ross71 (6;2): did we used to gargle). Finally, Adam asked two quite bizarre double-tensed questions. Judging from the contexts of these questions, the first (Adam34 (3;7): How

could she made in Italy?) was missing the passive *be*. The second (Adam39 (3;10): could you being some arrows?), I assume is a typographical or speech error of *being* for *bring*.

6. Discussion

Contracted auxiliaries. Brown (1973) argues persuasively that many of the examples of contracted auxiliaries in the speech of young children are actually unanalyzed units and do not represent true uses of the auxiliary. Brown noted that contracted auxiliaries appeared cliticized to only a very small fraction of the possible words they could have attached to. In other words, children said *it's* and *what's* but never *Bill's*, etc. He argued, therefore, that for young children *it's* is not a shortening of *it* + *is*, but rather *it's* is just a phonological variant of *it*.

Given Brown's argument, should we count those examples which appear to be double-tensed because of a contracted auxiliary? Let's review the examples. Among the 195 examples of double-tensed sentences (72 questions with more than one auxiliary and 123 questions with an auxiliary and a tensed main verb), there were 46 examples that appear to be double-tensed because of a contracted auxiliary. Thirty-five of these cases involve *it's* (33 by Adam and two by Shem). The only other examples of contracted auxiliaries are five uses of *he's* (one by Nathan and four by Shem), one example each of *she's* (April), *that's* (Sarah), *what's* (Sarah), *when's* (Nathan), and one example each of *we're* and *dey're*, both by Shem.

Hence, it is clear that the uses of contracted auxiliaries in the double-tensed questions were very restricted. I would like to argue, therefore,

that the questions that appear double-tensed because of a contracted auxiliary do not represent real examples of double tensing and should not be counted when we calculate the frequency of double-tensed questions.

Stuttered and unclear auxiliaries. Many of the possibly double-tensed questions are transcribed as being unclear or as having clause-internal pauses. I would argue that if the transcribers were unclear as to whether a child actually used an auxiliary or a tensed form of a main verb, we should be very hesitant to include these examples. It seems equally unwise to include examples in which the children hesitated. Many of these examples probably represent self-corrections or false starts on the part of the children. For example, when at the age of 3 years, Shem said in his thirty-fifth transcript, "I had a softball and I don't know where it is - was", it seems much more likely that he was correcting the tense of the copula than that he was using two forms of the copula.

For this reason I would argue that examples of stuttered or unclear utterances should not be included when determining the frequency of double-tensed questions. If we examine the lists of possibly double-tensed questions that appear in Tables 5.1 and 5.2, we notice that about half of the sentences are either stuttered or marked as unclear. Having spent the past five years analyzing these transcripts, my impression is that in the transcripts as a whole much fewer than 50% of the children's utterances are unclear or stuttered. This suggests that the double-tensed questions are not just a random sample of the utterances of children. Rather, in large part, the purportedly double-tensed questions are just speech errors.

Mixed tense examples. In a transformational theory, subject-auxiliary inversion results when the contents of the head of the verb phrase raises to pre-subject position. If one translates the notion of copying-and-deleting into present-day GB terms, one would say that the way the verb gets into pre-subject position is that the head of VP is copied into I which then moves to C (or gets copied in C). Once this is accomplished, the head of VP c .n delete leaving a trace. Now, the very notion of copying requires that an identical copy be made, i.e., that what appears in C be identical to what was in VP.

If the double-tensed questions are the result of the children having correctly copied the contents of the head of VP into C, but having failed to delete the material in VP, then in all the double-tensed questions the tense of the head of VP should be identical to the tense of the auxiliary that appears in C. Examples of questions where there tense of the head of VP differs from the tense of the auxiliary in C provide no support for the copy-without-deletion story. Since it is unclear how a transformational system could account for there being different tenses in C and V, one could even argue that mixed tense examples are evidence that children do not always use a movement transformation when they front an auxiliary. In other words, mixed tense examples might be evidence against a transformational account of inversion.

If we return to the possibly double-tensed questions, 52 examples had different tenses in the head of VP and in C. Twenty of the 52 examples were stuttered or unclear and 18 had contracted auxiliaries. An additional 11 involved irregular past tense verbs. The only three "pure" mixed-tense questions were:

Adam29 (3;5): can dey should go down dat Massachusetts Avenue.
Mark68 (4;1): could you do untie the knot
Ross51 (4;4): Mom [#] does this looked fixed?

Given that there are only three pure mixed-tense questions among the over 40,000 questions asked by the children, the existence of mixed-tense questions does not pose a threat to a transformational account of inversion. On the other hand, these mixed-tense questions in no way support any kind of transformational story, be it copy-and-delete or affect alpha (unless we have an overly broad view of what counts as 'affecting').

Negated auxiliaries. Now consider the double-tensed questions that had a negated auxiliary. All but seven of the cases involved unclear or stuttered auxiliaries, contracted auxiliaries or auxiliaries whose tense did not match that of the head of the verb phrase. For the reason mentioned above, let's consider these as being at best very weak evidence of a copy-without-deletion account of double-tensed questions. The remaining seven "good" examples of negated auxiliary double-tensed questions are repeated below for convenience:

Adam27 (3;4) did I didn't mean to?
Adam31 (3;5) did I didn't mean to?
Adam32 (3;6) do she don't need that one?
Adam33 (3;6) did I didn't mean to do that?
Ross36 (3;3) why does Superman doesn't wear underroos on his bottom?
Nina34 (2;10) does it doesn't move?
Nina34 (2;10) do they don't eat people up

Notice the following three things. First, all seven examples are of the form: *do NP do + n't*. Second, only three children made this sort of error. Third, the three children who do made this error, produced all of their examples within a fairly short time span (2 months for Adam and 1 day for Nina and Ross). Given these three facts, I agree with Maratsos and Kuczaj (Maratsos & Kuczaj, 1978, and Maratsos, 1984) that these examples probably are caused by Adam, Nina and Ross having mis-analyzed *didn't*, *don't*, and *doesn't* as untensed negation elements. They probably do not represent "real" cases of double-tensing in questions.

Irregular past tense verbs and double-tensed questions. If we discount all of the double-tensed examples that have contracted auxiliaries, are stuttered or unclear, have mixed tense forms, or involve negated *do*, we are left with 70 examples. Of these 70 examples, six have the same auxiliary replicated twice (e.g., *is this is a dog?*), 14 have main verbs with a regular tense forms (e.g., *does it rolls?*), and 50 have main verbs with irregular tense forms (e.g., *did I caught it?*).

As I mentioned above, Maratsos and Kuczaj (Maratsos & Kuczaj, 1978, and Maratsos, 1984) noted the large number of double-tensed questions that had irregular past tense verbs. They suggested that double-tensed questions with irregular verb forms should not be counted because children may not know that these verbs are in the past tense. As I mentioned above, what we make of double-tensed questions with irregular verbs depends crucially on how frequent this errors are. I found that 71% of 'good' double-tensed questions (50/70) were of the form *did + irregular past tense verb*. I would argue that instead of calculating the percentage of

all double-tensed questions that have irregular verbs (50/70 = 71%), the correct percentage to calculate is:

double-tensed irregular /# irregular + # regular = 50/64 = 78%

since the comparison that we want to make is the frequency of regular double-tensed questions and irregular double-tensed questions relative to the overall frequency of regular and irregular verbs in the speech of children.

How does this compare to the overall frequency of irregular verbs in the speech of young children? Pinker and Prince (1988) report that about 50% of the past tense verb types in samples of Adam, Eve, and Sarah's transcripts were irregular. This isn't really the comparison we want to make. Ideally, we want to compare the relative frequency of irregular past tense verb tokens and all regular verb forms (not just those in the past tense) in declarative contexts, in single-tensed questions and double-tensed questions. If double-tensing errors are the result of children not realizing that verbs like *said* are past tense forms, then we would expect that the frequency of double-tensed questions with irregular verbs would be greater than the frequency in the transcripts as a whole.

Unfortunately, I do not know what percentage of children's tensed verbs are irregular past tense verbs, but my suspicion is that the number is smaller than 78%. For the sake of argument, let's make the educated, but conservative, guess that one half of all of the children's inflected verbs tokens are irregular past tense verbs. If we assume this, then in the transcripts, the children asked about 18 more double-tensed questions with irregular verbs (about 50% more) than we would have expected.

The frequency of double-tensed questions. What is the frequency of double-tensed questions? If we eliminate the declarative sentences that snuck into the sample because they appeared on a transcript line which contained a question, there were 164 questions that might possibly be double-tensed. Thus, the highest possible frequency for double-tensed questions is about 0.4% (164/40,600 questions). If we discount all examples that were stuttered or unclear, approximately 0.2% (87/40,600) of questions were double-tensed. If we choose to eliminate the seven *did she didn't* cases and the 11 cases with mixed tense, we are left with only 70 examples or about 0.17% of all questions being double-tensed. If we choose to accept all of the double-tensed questions with irregular verb forms, 0.17% represents the frequency of double-tensed questions.

Let's say that we decide, based on the assumptions made above, that at least 18 of these irregular verb double-tensed questions are spurious examples. Then we say that the frequency of double-tensed questions in the transcripts was about 0.13% (52/40,600). We could take a more extreme position and argue that all of the irregular verb examples are discountable. If we did, we would argue that the frequency of double-tensed questions is as low as 0.05% (20/40,600). (Table 6.5 lists all 20 examples that would still be considered real under this most stringent set of requirements.)

Please Insert Table 6.5

Whether we accept the high figure of 0.4% of the low figure of 0.05%, the fact is that double-tensed questions are very rare, certainly not

frequent enough to be the sort of strong evidence for subject-auxiliary inversion being performed via a pair of copy-and-delete operations as Hurford (1975), Fay (1978) and Mayer, Erreich and Valian (1978) suggest. The fact that almost half of the double-tensed questions have mixed tenses and, hence, cannot have been generated by copying-without-deletion strengthens the case against Hurford, Fay and Mayer *et al.*

I would argue that the remarkable frequency of stuttered and unclear utterances among the purported double-tensed questions (about 50%) suggests that many of these examples are speech errors. Of the double-tensed questions that were neither stuttered nor unclear nor involved mixed tenses, a remarkably high percentage were questions where *do*-support had to be invoked.

***Do*-support and auxiliary errors.** Why did so many of the children's combination errors involve *do*? One possible explanation is that the reason *do*-support errors are most common because they tend to appear with irregular past tense verbs. But notice that this is not an explanation. There is no reason why children should use irregular past tense verbs more with *do* than with any other auxiliary. If the children do not realize that irregular past tense verbs are, in fact, past tense forms, they should use these irregular past tense verbs with modals (e.g., *can you ate that?*), progressive *be* (e.g., *are you ateing that?*), and perfect *have* (e.g., *have you aten that*). One possibility is that *do* errors predominate because they are the most common auxiliary. It is true that the children used auxiliary *do* very frequently. This cannot be the entire explanation, however, because *do* accounted for only about 40% of the auxiliary uses, yet *do* errors accounted for over 80% of combination errors involving auxiliaries.

Martasos and Kuczaj argue that children think *didn't*, *don't*, and *doesn't* are simply negation markers and not auxiliaries. Thus, children say *did she didn't cook?* and *she didn't cooked.* Again this is not an explanation for the preponderance of *do* combination errors because it doesn't explain why children don't mistake *can't, won't, isn't*, etc. for simple negation markers. Perhaps the children made more mistakes with *didn't, don't*, and *doesn't* because they used these much more than the other negated auxiliaries. This cannot be the whole explanation. Overall, the children used approximately 9,400 negated auxiliaries or copulas. Of these, approximately 900 were *didn't*, 4,230 were*don't*, 670 were *doesn't*, and 3,600 were some other negated auxiliary or copula. Thus, if the children made errors because they didn't realize that negated auxiliaries were auxiliaries, then about 40% of their errors should have involved negated auxiliaries and copulas other than *do*. Only two negation errors did not involve *do*.

According to Maratsos (1984) most combination errors involve *do* because "for the other auxiliary verbs, the relations among interrogatives, negatives, truncates, and declaratives are quite regular. *Do* is more of a special case, and this may introduce performance errors so that children mark with *do*, but they also lose track of this and also mark the main verb. This would result in a redundant main verb tensing" (p. 757). According to Pinker, the problem with *do* stems from the interaction between the lack of semantic content for *do* and the procedure by which co-occurrence restrictions are acquired. Because *do* is semantically empty, children don't realize that *do* is a complement-taking verb. Since co-occurence restrictions are realized as restrictions on what complements a particular verb can take, if children do not always realize *do* is a complement-taking

verb, they will not always apply to *do* the restriction that requires that complement-taking verbs take only infinitival complements.

While both Pinker and Maratsos's explanations can account for the preponderance of *do* with tensed main verbs, neither can account for the fact that the vast majority of the double-tensing errors have irregular past tense verbs with past tense *did*. Both Pinker and Maratsos attempt to deal with this by saying that the reason many examples involve irregular verbs is that children do not realize that these verbs are tensed. But if this is so, why do the vast majority of the examples with *do*-support have past tense *did* and *didn't* rather than *does*, *doesn't*, *do* or *don't*? Note, that again the explanation does not lie in the relative frequency of *did*, *do*, and *does*, because *did* accounted for only about 30% of uses of non-negated forms of *do*, while *do* and *does* accounted for 50% and 20% of non-negated uses, respectively. Likewise, *didn't* accounted for only 15% of negated uses of *do*, while *don't* and *doesn't* accounted for 75% and 10% of negated uses, respectively.

I would argue that the high incidence of *did* with irregular past tense verbs and the low incidence of *did* with any other inflection or *do* or *does* with irregular past tense verbs suggests that the children knew that irregular past tense verbs were past tense forms. Under this theory, double-tensed errors were the result of problems with the mechanics of *do*-support coupled with problems accessing the lexical entries for irregular verbs.

One possible model that would account for the data is the following. Assume that children know that the tense of the sentence they intend to utter. Assume also that because they know that they intend to utter a past tense sentence,they access a past tense verb form from the lexicon. If they

intend to utter a question, negative statement or sentence with emphasis, they next invoke *do*-support. Because they know that *do* is the head of the VP and, therefore, must be tensed, they obtain the past tense form *did* from the lexicon and inserted into I. At this point they have an utterance with a tensed auxiliary *do* and an irregular past tense verb. They realize that the past tense verb should be in infinitival form, but they fail in their attempts to access from the lexicon the infinitival form of the irregular verb.

Contrast this with what would happen with regular past tense verbs.

2. Questions that lack auxiliaries

Brown and his colleagues (Bellugi, 1965, 1971; Brown, 1968; Brown and Frazer, 1963; Klima and Bellugi, 1966) noted that children's earliest questions often lack auxiliaries. In his 1968 paper, Brown noted that some of these early auxiliary-less questions had tensed main verbs and others did not. I would like to address the question of why children produce questions that lack obligatory auxiliaries, and why they produce the type of auxiliary-less questions that they do.

There are at least three possible reasons why children might produce auxiliary-less questions. First, they might think that auxiliaries are optional elements in matrix questions. If they think auxiliaries are optional, sometimes they might simply choose not to say them. I will call this the Optional Auxiliary Account. Second, children might know that auxiliaries are required in matrix questions, yet occasionally fail to

produce them because of limitations in their production systems. I will call this explanation the Production Constraint Account. The third possibility is that children ask questions that lack auxiliaries because of some failure in verb-raising. I will call this the Verb-raising Account.

The Optional Auxiliary and Production Constraint Accounts predict that when children ask auxiliary-less questions, the main verb should not be tensed. According to the Optional Auxiliary Account, the child's grammar represents questions as:

Q-morpheme (AUX) NP VP?

Thus the grammar allows the child to ask questions such as *does she walk?* and questions such as *she walk?* Sometimes the child correctly chooses the former type, but sometimes she incorrectly chooses to use the later type of structure. Essentially, sometimes the child means to say *she walk?* and that's exactly what she says.

According to the Production Constraint Account, when the child says a question that lacks an auxiliary, she knows that an auxiliary is needed. In other words, the child means to and wants to say *does she walk?* but somehow the *does* gets lost in production, and she says *she walk?*.

The Verb-Raising Account argues that auxiliary-less questions are the result of problems involving verb raising. Hence, according to this account, auxiliary-less questions are just examples of a certain type of inversion error. There are several different kinds of problems with verb-raising that could results in questions that lack auxiliaries. They might think that verb-raising is optional. Perhaps, they do not always recognize the contexts that demand verb-raising. A third possibility is that children

might know that verb-raising is required in matrix questions, but they have not completely mastered the intricacies of *do*-support. Any of these versions of the Verb-Raising Account would predict that when children ask auxiliary-less questions, they will tense the main verb. Consider, for example, how the *do*-support version would account for auxiliary-less questions. According to this version, the child means to say *does she walk?* but has trouble with *do*-support. Since for phonological reasons, the tense inflection must be attached to something, the child leaves the tense marker on the main verb (or possible returns it there after *do*-support fails). The result is that the child produces the question *she walks?*.

Transcript Analysis II: Auxiliary-less Questions

1. Overview

In this transcript study, I analyzed the spontaneous questions asked by 14 children to determine whether they tensed the main verb in questions where they omitted an obligatory auxiliary.

2. Subjects

I used the spontaneous speech transcripts of the same 14 children that I used in the double-tensing analysis. (Please see Figure 6.1 above). Notice, however, that in the double-tensing study I only examined questions

that contained auxiliaries. In the current study, I only examined those questions that lacked an auxiliary.

3. Corpora

Auxiliary-less *wh*-questions. I used the Unix utility "fgrep" to cull all of the lines which had *wh*-words. Because I was only interested in questions that did not contain an auxiliary, I used "fgrep" again to eliminate all of the lines which included any of the following letter strings: *am*, *are*, *can*, *could*, *didn't*, *don't*, *doesn't*, *hadn't*, *hasn't*, *haven't*, *is*, *may*, *might*, *must*, *shall*, *should*, *was*, *were*, *will*, *won't*, and *would*. I did not eliminate lines that contained *did*, *do*, *does*, *had*, *has*, or *have* because these words can be main verbs.

Auxiliary-less yes/no questions. For the yes/no questions I again used the Unix utility "fgrep". This time I searched for all lines containing a question mark "?". I then used "fgrep" to eliminate those lines which contained *wh*-words. Finally, I used fgrep a third time to eliminate all of those lines which contained the letter strings: *am*, *are*, *can*, *could*, *didn't*, *don't*, *doesn't*, *hadn't*, *hasn't*, *haven't*, *is*, *may*, *might*, *must*, *shall*, *should*, *was*, *were*, *will*, *won't*, and *would*.

4. Tallying Procedure

For each transcript for each child,I recorded the number of number of "potentially informative" auxiliary-less questions that had tensed main

verbs and the number of potentially informative auxiliary-less questions that did not have tensed tensed main verbs. I kept separate tallies *how*, *how come*, *what*, *when*, *where*, *which*, *who*, *why*, and yes/no questions. The following rather stringent criteria were adopted to determine what counted as a potentially informative auxiliary-less question.

1. It had to contain a subject
2. It had to contain a verb
3. It could not contain an auxiliary (either contracted or not contracted)
4. It had to clearly be a question (as indicated by a question mark, notations in the transcript, parental response etc.). In other words, it could not be a declarative or a fragment of a relative clause[1]
5. It had to be transcribed in a way that suggested it was neither stuttered nor unclear.
6. It could not be a routine.
7. It had to be "informative" in that at least one of the three accounts described above had to make a different prediction about the form the auxiliary-less question would take.

Because the seventh criterion is crucial but somewhat difficult to put into words, I would like to flesh out what I mean by "informative" and "not informative" auxiliary-less questions, giving examples of each type. Basically, in order for a question to be considered informative, it had to be possible to distinguish between when the main verb was inflected and when it was not inflected. In other words, I did not include questions which had phonologically null inflections. Consider the auxiliary-less question *you bite?* Below I show what each of the three accounts would predict for the auxiliary-less version of this question:

---

[1] See Chapter 5, Footnote 27 for a description of the criteria used to distinguish relative clauses.

Optional Auxiliary Account:

Child's Rule	Child chooses to say:
(Do) you bite?	*you bite?*

Production Constraint Account:

Child means to say	Child fails to say *do* and, hence, says:
Do you bite?	*you bite?*

Verb-Raising Account:

Child means to say	V-raising fails and, hence, child says:
Do+ϕAGR *you bite?*	*you bite*+ϕAGR*?*

Notice that what is predicted by all three accounts is that the child will say *you bite?* Thus, this auxiliary-less question is not informative.

Now if we consider the question *does it bite?* (see below), we see the the Optional Auxiliary and the Production Constraint Accounts predict that the child will say *it bite?*, while the Verb-Raising Account predicts that the child will say *it bites?* Thus, *it bite(s)?* is an informative auxiliary-less question.

Optional Auxiliary Account:

Child's Rule	Child chooses to say:
(Does) it bite?	*it bite?*

Production Constraint Account:

Child means to say	Child fails to say *does* and, hence, says:
Does it bite?	*you bite?*

Verb-Raising Account:

mean to say:	V-raising fails and, hence, child says
Do+-s it bite?	*it bites?*

Now consider the progressive question *is it raining?* All three models predict that if the child omits the auxiliary, she will ask *it raining?*

Therefore, auxiliary-less questions that have a verb with the progressive -*ing* are uninformative.

Notice what happens if the child means to ask a question with a modal auxiliary like *can it bite?* but, for whatever reason, omits the *can*. Here, all three accounts predict that the child will say *it bite?* Thus, even the Verb-Raising Account does not predict that 100% of auxiliary-less question with a third person singular pronoun have overtly inflected verbs. The difference between the three accounts is that the Verb-Raising Account predicts that there will be a non-negligible number of auxiliary-less questions that have overtly inflected main verbs, while the Optional Auxiliary Account and the Production Constraint Account predict that there will be no examples of auxiliary-less questions with tensed main verbs.[2]

5 Results

Auxiliary-less yes/no questions. Three children (Allison, June and May) asked no informative auxiliary-less questions. For the 11 children who asked informative auxiliary-less yes/no questions, the overall rate of tensed main verbs was 55%, with rates for individual children ranging from a low of 26% (Naomi) to a high of 95% (Mark). The median rate of tensing main verbs was 57%. Table 6.6 shows the rates for each of the children.

---

[2] Actually, this is somewhat overstated,. The Optional Auxiliary and Production Constraint Accounts would allow examples like *it bites?* as long as such examples were echo questions. The reason is that in an echo question, the auxiliary has not been omitted. It just wasn't there in the first place.

Please Insert Table 6.6

In Figures 6.1 through 6.12, the overall percent of informative yes/no questions that had tensed main verbs is graphed on the Y-axis and age on the X-axis. Figure 6.1 depicts the combined performance for all 11 children. Figures 6.2 through 6.12 depict the performance of individual children. Note that in the graph depicting the performance of all the children(Figure 6.1), the curve is essentially monotonic (excluding a small blip at about 2;4 caused by Nina's performance), with approximately 70% of informative auxiliary-less questions having tensed main verbs by around age 3;0. The graphs of individual children's performance reveal that all of the children exhibit the same general pattern.

Please Insert Figures 6.1 -6.12

Auxiliary-less *wh*-questions. As was the case with the yes/no questions, Allison, June and May asked no informative auxiliary-less *wh*-questions. For the 11 children who asked informative auxiliary-less *wh*-questions, the overall rate of tensed main verbs was 69%, with rates for individual children ranging from a low of 31% (Shem) to a high of 94% (April). The median rate of tensing main verbs was 66%. Table 6.6 shows the rates for each of the children. The overall rate of tensing main verbs in informative questions was 74% for *how* questions, 93% for *how come*

questions, 63% for *what* questions, 67% for *when* questions, 35% for *where* questions, 79% for *who* questions, and 78% for *why* questions.[3]

Figure 6.13 shows the percent of tensed main verbs for all *wh*-questions except *who* and *how come*,[4] for all 11 children. The general appearance of the overall *wh*-question graph (Figure 6.13) is remarkably similar to that of the overall yes/no question graph (Figure 6.1). There is a steady increase in the percent of auxiliary-less questions with tensed main verbs (again, except for a blip around 2;4 which is due to Nina). By about age 3;0, children were tensing the main verb in approximately 55% to 60% of *wh*-questions that lacked auxiliaries. The percent of tensed main verbs at age 3;0 was slightly less for the *wh*-questions than for the yes/no questions, despite the fact that the overall percent of tensed main verbs was higher for auxiliary-less *wh*-questions than for yes/no questions.

Please Insert Figure 6.13

I suspected this difference was just due to the fact that I did not include the data for the *how come* and *who* questions in this graph. Nonetheless, I decided to graph each type of *wh*-question separately to see if all of the curves were similar for all of the *wh*-questions. (Please see Figures 6.14 through 6.19.) The curves for the individual *wh*-questions are probably about as similar as they can be given the small amount of data

---

[3] It was not possible to do an analysis to see if these differences were significant because of the large number of empty cells.

[4] I did not graph *how come* questions because it is acceptable (even preferred) to ask a matrix *how come* question that lacks an auxiliary. I also did not graph *who* questions because most of the auxiliary-less *who* questions were perfectly grammatical subject questions (e.g., *who wants gum?*)

represented in some of the graphs. The curves for the *wh*-questions for which there were more data (*what*, *where*, and *why* questions)are fairly similar to one another and to the overall *wh*-question curve.

Please Insert Figures 6.14-6.19

Older tensing rates. There is no reason to expect that the children would have tensed the main verbs in auxiliary-less questions any more often than they tensed main verbs in declarative sentences. The percentage of tensed main verbs in declaratives can be thought of as the ceiling for the percentage of auxiliary-less questions that can have tensed main verbs. Therefore, for each child, I determined what percent of the auxiliary-less questions had tensed main verbs starting from the age at which that child began using a fair number of inflected main verbs in declaratives. I did not have formal criteria for decided what age to start counting auxiliary-less questions, but rather just got a general impression from reading several pages of the transcripts for each child.[5] No child was older than 3;0 at the point when I began counting auxiliary-less questions.

Table 6.6 shows that when I just considered those yes/no auxiliary-less questions that were asked after the children had achieved some competency, the overall rate of tensing main verbs was 75%, with individual rates ranging from 50% (Eve) to 95% (Mark). For the

---

[5] For Adam I judged him to have achieved some competency at verb inflections at about 2;11, about 6 months before he passed Brown's (1973) 90% correct in obligatory context criterion for most verbal inflections. For Sarah, I decided she had achieved some competency on verbal inflections about 12 months before she was deemed competent by Brown (1973). I judged Eve to be competent by 2;3, the same age Brown passed her on the verbal inflections.

auxiliary-less *wh*-questions asked after the children had acquired some minimal competency in the verbal inflections, approximately 85% of the auxiliary-less *wh*-questions had tensed main verbs, with rates ranging from 50% (Eve) to 95% (Nina).

6. Discussion

What this transcript analysis demonstrates is that, when children ask a question that does not have an auxiliary, they tend to inflect the main verb. For the yes/no questions, children asked questions like *it bites?* 55% of the time and they ask questions like *it bite?*. about 45% of the time. If we just consider those questions asked after children have developed some familiarity with verbal inflections, they tense the main verb in about 75% of their auxiliary-less yes/no questions. About 70% of children's auxiliary-less *wh*-questions had tensed main verbs. If we only count those *wh*-questions asked after about the age of 3;0, the percent of auxiliary-less *wh*-questions with tensed main verbs climbs to about 85%. If we examine how each child did, we see that all of the children tensed the main verb in a non-negligible percent of the cases. The child who tensed the fewest number of verbs in auxiliary-less yes/no questions (Naomi) still tensed over a quarter of the informative questions. Shem tensed the fewest number of verbs in the *wh*-questions, and he still tensed almost a third of the verbs in informative *wh*-questions. If we examine the different *wh*-questions, even for *where* questions, the *wh*-question with the lowest rate of tensed main verbs, over a third of auxiliary-less questions had tensed main verbs.

Thus, the percent of informative questions with tensed main verbs is clearly greater than the near 0% that is predicted by the Optional Auxiliary

Account and the Production Constraint Account. I would argue, therefore, that at least half of the children's auxiliary-less questions (and perhaps as high as 80%) represent failures of verb-raising. In a sense, they are analogous to the subject-auxiliary inversion errors examined in chapter 5. The only difference is that they are examples that would have required *do*-support, if inversion had taken place.

If we go back to the results of the Double-Tensing Analysis, we see that of the 70 double-tensed questions that aren't stuttered or unclear, 65 involve *do*-support. In other words, 93% of the double-tensed questions involved *do*-support.[6] If we decide to discount all of the examples with irregular past tense verbs, 74% (14/19) of double-tensed questions still involved *do*-support. I would argue that children know much about verb-raising. What they have trouble with are the details of *do*-support. *Do*-support is a very odd feature of English. It is the jerry-rigged result of English allowing verb-raising while prohibiting main verbs from raising. Cross-linguistically, it is also a peculiar sort of rule. It is not surprising, therefore, that children find it difficult to master. It cannot be a part of Universal Grammar, *per se*.[7]

---

[6] If we consider those *do NP do+n't* cases, the figure is still 93% (72/77) of the double-tensed questions involved *do*-support.

[7] Though Chomsky points out that if *do*-support is in fact triggered by the existence of a weak, but present AGR in English, then it is possible that the ability to recognize the need for a some kind of a rule like *do* support could be innate.

3. Children's Judgments of Questions with Extra Auxiliaries or No Auxiliaries

1. Overview

Using the same procedure that was used in the Inversion Judgment Experiment described in chapter 5, children judged questions that had either, neither, or both an auxiliary and a tensed main verb.

2. Subjects

Twenty preschool-aged children between the ages 3 years, 1 month and 5 years, 6 months (mean age = 4;4) participated in the experiment. All children had normal vision and hearing and spoke only English. Children came from the same city and the same range of socio-economic classes as the children in the Inversion Judgment Experiment reported in chapter 5.

Children were divided into four groups according to age and MLU as calculated using the procedure described in the Inversion Judgment Experiment. The average elicited MLU for the four MLU groups were 2.81, 4.14, 5.55, and 8.91. The average age of the subjects in these four MLU groups were 3;10, 4;0, 4;11, and 4;8, respectively.

3. Stimuli

3.1 Question Types. There were four types of questions defined by the orthogonal presence or absence of an auxiliary and a tensed main verb. The four types of questions were:

+Auxiliary	-Tensed Main Verb	e.g. What does Kermit eat ?
+Auxiliary	+Tensed Main Verb	e.g. *What does Kermit eats?
-Auxiliary	-Tensed Main Verb	e.g. *What Kermit eat?
-Auxiliary	+Tensed Main Verb	e.g. *What Kermit eats?

Note that of the four question types, only one type (+Auxiliary -Tensed Main Verb) is grammatical in the adult grammar.[8] If I had presented the children with an equal number of each type of question, the children might have become biased to say that questions were ungrammatical. To prevent this from happening, children judged three times as many +Auxiliary -Tensed Main Verb questions as the other three question types. Thus, subjects judged an equal number of grammatical and ungrammatical questions.

I had to choose which auxiliary or auxiliaries to use in the test questions. I decided to use the present tense form of *do*. I chose *do* because the results of the transcripts analyses for double-tensed and auxiliary-less questions suggested that lack of mastery of *do*-support might have caused children to ask double-tensed questions and auxiliary-less questions. I chose to use the third person singular present tense form *does* rather than *did* because, as Maratsos and Kuczaj (Maratsos, 1984; Maratsos and Kuczaj, 1978) point out, many of the examples of double-tensed questions involved irregular past tense verbs. By using *does* rather than *did*, I could use present tense forms of verbs and thereby study the effects of *do*-support separate from the effect of irregular past tense verbs.

---

[8]-Auxiliary +Tense Main Verb *how come* questions (e.g., *how come she eats liver?*) are as grammatical or more so than +Auxiliary -Tensed Main Verb *how come* questions (e.g., *how come she does/can eat liver?*). Please note that for the *how come* questions, the +Auxiliary -Tensed Main Verb sentences did not have subject-auxiliary inversion because inversion is unacceptable in *how come* questions.

I decided that I wanted to be able to compare how children judged test questions that had *does* as an auxiliary with questions that had *can*. I decided to do this for two reasons. First, the children's performance on *can* questions could serve as a baseline to compare children's performance on *does* questions. Second, Pinker (1984) argued that the reason that many double-tensing errors involve *do* is that *do* is not a very prototypical auxiliary, and for this reason, children do not always recognize that *do* is an auxiliary. I chose *can* because my analyses of children's spontaneous speech transcripts (see chapter 4) and the work by Wells (1979) and Majors (1974) suggest that *can* is one of the earliest and most frequently used auxiliaries. Therefore, half of the +Auxiliary question had *can* and half had *do*. Each child heard questions with *can* and *do* equally often.

3.2 *Wh-* questions Used. The *wh*-question used were:

HOW	How (AUX) Big Bird fly(s)?
HOW COME	How come Ernie (AUX) talk(s)?
WHAT	What (AUX) Kermit eat(s)?
WHERE	Where (AUX) Oscar live(s)?
WHO	Who (AUX) Bert lick(s)?
WHY	Why AUX Grover sing(s)?

3.3 Order of Presentation. The design of the experiment is such that it requires 72 questions if both *can* and *do* are to appear with each question type and there are to be equal numbers of grammatical and ungrammatical questions. However, a pilot study revealed that children were unable (or unwilling) to judge 72 test questions. To cut the number of questions down to 36, each child did not judge each every *wh*-question in every question-type with both *can* and *do*. Each child did, however,

hear *can* and *do* equally often for each of the *wh*-questions. In addition, *can* and *do* appeared equally often for each *wh*-question and each question-type across subjects.

Order of presentation was pseudo-random with the following conditions: 1) the same *wh*-question never occurred more than twice in a row; 2) the same question type never occurred more than three times in a row; and 3) no more than three grammatical or ungrammatical questions occurred in a row.

4. Procedure

I used the same procedure and practice questions that I used in the Inversion Judgment Experiment which is reported in chapter 5.

5. Results

Overall, the children judged 71.4% of the grammatical questions as being good and 53.8% of the ungrammatical questions as being good. Children in the lowest MLU group (MLU Group 1) judged 56.8% of grammatical questions and 45.5% of the ungrammatical questions to be good. Children in MLU Group 2 liked 72.3% of the grammatical questions and 65.2% of the ungrammatical questions. Children in MLU Group 3 liked 75.5% of the grammatical questions and 62.2% of the ungrammatical questions. Children in the highest MLU group (MLU Group 4) liked 81 2% of the grammatical questions and 42.2% of the ungrammatical questions. Thus, the children in all the MLU groups liked the grammatical questions more than the ungrammatical questions, and this preference was greater among the children in the highest MLU group more

than in the lower three MLU groups. Thus, if we take adult judgments as the standard for comparison, it appears that the experiment was successful at getting children to judge questions.

I performed analyses of variance of the data using first Age and then MLU as the between-subjects variable. Neither Age nor MLU had a significant main effect and only MLU entered into any significant interactions. Therefore, I will only present the results for the MLU analyses. An analysis of variance with MLU as the between-subjects variable and *Wh-* word, +/-Auxiliary and +/-Tensed Main Verb as the within-subjects variables revealed that none of the factors had a significant main effect.

There were, however, a number of significant interactions among the factors. Figure 6.20 depicts the significant interaction between +/-Auxiliary and +/-Tensed Main Verb ($F(1, 16) = 12.42$, $p = .003$, eta = .66, $f = .88$). Children liked 48.3% of tenseless questions (e.g., **what Kermit eat?*), 50.0% of double-tensed questions (e.g., **what does Kermit eats?*), 65.8% of questions which contained only a tensed main verb (e.g., **what Kermit eats?*), and 68.7% of grammatical questions that contained only an auxiliary (e.g., *what does Kermit eat?*).

Please Insert Figure 6.20

I made the following planned comparisons. I found there to be no significant difference between questions which were ungrammatical because they had two tense markers and questions that were ungrammatical

because they had no tense markers ($F(1, 19) = .13, p = .723$). I also found that there was no significant difference between grammatical questions such as *what does Kermit eat?* which had only an auxiliary and ungrammatical questions such as **what Kermit eats?* which had only a tensed main verb ($F(1, 19) = .187, p = .670$). As predicted, the difference between the grammatical +Auxiliary -Tensed Main Verb questions and the double-tensed questions was significant ($F(1, 19) = 7.97, p = .011$, eta = .54, f = .65). The difference between the grammatical questions and the untensed questions was even greater and more significant ($F(1, 19) = 15.54, p$ = .001, eta = .67, f = .91).

In addition to these planned comparisons, I made post-hoc comparisons between the questions that had only a tensed main verb and the questions that had both or no tense markers. The difference between the -Auxiliary+Tensed Main Verb questions and the questions with no tense markers (-Auxiliary-Tensed Main Verb) was significant ($F(1, 19) = 5.31, p$ = .033, eta = .47, f = .53) and the difference between the -Auxiliary+Tensed Main Verb questions and the double-tensed questions (+Auxiliary+Tensed Main Verb) was marginally significant ($F(1, 19)$ = 4.04, p = .059, eta = .42, f = .46).

Descriptively, the above can be summarized by saying that children significantly preferred questions that had one tense marker (either an auxiliary or a main verb) over questions that had both tense markers or no tense markers. If we do an analysis of variance in which there is a single factor, Number of Tenses with three levels (zero, one, and two tense markers) instead of the two factors +/-Auxiliary and +/-Tensed Main Verbs, there is a significant main effect for Number of Tenses ($F(2, 32)$ =

7.38, p = .002, eta = .56, f= .68)[9]. Overall, the children liked 48.3% of questions with no tense markers, 67.25% of questions with one tense marker, and 50% of questions with two tense markers (see Figure 6.21).

Please Insert Figure 6.21

There was a significant interaction between MLU and +/-Auxiliary (3, 16) = 3.53, p = .039, eta = .63, f = .81). As is shown in Figure 6.22, all but the lowest MLU group had a slight preference for questions that had auxiliaries. Children in the lowest MLU group preferred auxiliary-less questions to questions that contained an auxiliary. Separate analyses of variance for each MLU group revealed that the difference in preference for questions that did and did not have an auxiliary was marginally significant for the lowest MLU group (F(1, 4) = 5.67, p = 076, eta = .77, f= 1.20) but not for any of the three higher MLU groups (MLU Group 2 F(1, 4) = 3.29, p = .144; MLU Group 3 F(1, 4) = .85, p = .409; MLU Group 4 F(1, 4) = 1.86, p = .245).

Please Insert Figure 6.22

There was also the significant interaction between *Wh-* word and +/-Auxiliary which is shown in Figure 6.23 (F(5, 80) = 2.42, p = .043, eta = .36, f = .39). Descriptively, children preferred *what* questions that had auxiliaries, and *why* and *how come* questions that lacked auxiliaries.

---

[9] Notice that the size of the main effect of Number of Tense Cures is smaller (eta = .56, f = .68) than the size of the equivalent interaction between +/-Auxiliary and +/-Tensed Main Verb (eta = .66, f = .88)

Please Insert Figure 6.23

Figure 6.24 depicts the significant interaction between *Wh-* word and +/-Tensed Main Verb ($F(5, 80) = 3.57$, p = .006, eta = .43, $f = .47$). This interaction seems to be due to children having preferred *how come* questions that had inflected main verbs and *why* questions that didn't have inflected main verbs.

Please Insert Figure 6.24

Because the *how come* questions differ from all other matrix questions in that *how come* questions do not require an auxiliary if the main verb is tensed, I suspected that the interactions between *Wh*-word and +/-Auxiliary and *Wh*-word and +/-Tensed Main Verb were caused by the *how come* questions. To test this, I performed separate analyses of variance on just the *how come* question data and on all the data except for the *how come* questions. The analysis of all the data but the *how come* data revealed no significant interaction between *Wh*-word and +/-Tensed Main Verb ($F(4, 64) = 1.78$, $p = .144$). The analysis of just the *how come* data revealed a significant main effect for +/-Tensed Main Verb ($F(1, 16) = 10.08$, $p = .006$, eta = .62, $f = .79$). Children liked *how come* questions with tensed main verbs 72.5% of the time and questions without tensed main verbs 49.2% of the time. This suggests that the significant interaction

between *Wh*-word and Tense is due in large part to the anomalous behavior of *how come* questions.

The analysis of all but the *how come* data revealed a significant interaction between *Wh*-word and +/-Auxiliary ($F(4, 64) = 3.13$, $p = .016$, eta = .41, $f = .46$). The analysis of just the *how come* data revealed no significant main effect for +/-Auxiliary ($F(1, 16) = .769$, $p = .39$). Thus, it is unlikely that the *how come* questions caused the interaction between *Wh*-word and +/-Auxiliary.

The interaction between +/-Auxiliary and +/-Tensed Main Verb was significant for both the *how come* data ($F(1, 16) = 8.05$, $p = .012$, eta = .58, $f = .71$) and for the all data except the *how come* data ($F(1, 16) = 9.67$, $p = .007$, eta = .61, $f = .77$). However, the interaction between MLU and +/-Auxiliary (which was significant to the $p < .05$ level when we analyzed data for all of the *wh*-words) was not significant for just the *how come* data ($F(3, 16) = 1.44$, $p = .27$) but was significant for the data without the *how come* questions ($F(3, 16) = 3.47$, $p = .041$, eta = .63, $f = .81$).

***Can* versus *Does*.** I did an analysis of variance of all of the +Auxiliary data to determine if there was a main effect or interactions for Auxiliary Choice. Auxiliary Choice did not have a significant main effect. Children liked 62.4% of *can* questions and 56.7% of *does* questions. In addition, auxiliary choice did not enter into any significant interactions. More specifically, I did not find a significant interaction between Auxiliary Choice and Tensed Main Verb ($F(1, 16) = 20$, $p = .658$). For the questions that lacked a tensed main verbs, children liked the *can* questions 70.6% of the time and the *does* questions 66.8% of the time. For the questions that

had tensed main verbs (i.e., ungrammatical questions), children liked the *can* questions 55.1% of the time and the *does* questions 46.7% of the time.

The only significant main effect or interaction in the *Can/Do* analysis was a significant main effect of +/-Tensed Main Verb ($F(1, 16) = 8.95$, $p = .009$, eta = .60, $f = .75$). Children liked the questions that lacked a tensed main verb (i.e., the grammatical questions) 68.7% of the time and they liked the questions that did have a tensed main verb (i.e., the ungrammatical questions) only 50.9% of the time.

6. Discussion

The most surprising result of the judgment experiment was that the children judged questions that had either one of the two tense markers to be equally good. In other words, they judged ungrammatical questions like **what Kermit eats?* to be just as good as grammatical questions like *what does Kermit eat?*. Both types of single-tensed questions were considered to be significantly better than questions that had either no tense markers (e.g., **what Kermit eat?*) or two tense markers (e.g., **what does Kermit eats?*). (The children liked questions that weren't tensed or were double-tensed equally little.)

The other rather surprising result was that there was no main effect of auxiliary choice. Children liked slightly more *can* questions than *does* questions but the difference was not significant. I also expected that the children would have preferred double-tensed questions which had *does* as the auxiliary to questions which had *can* as the auxiliary. Children actually liked questions such as *what can Kermit eats?* slightly more than questions

like *what does Kermit eats?* but the difference was not significant. This is surprising because in *can* questions, the tense of V and C do not match, whereas they do match in the *does* questions. Frankly, I am at somewhat of a loss to explain why the results came out this way. Perhaps it has to do with the differing semantics of the *can* and *does* versions of the sentences. Perhaps, it has to do with the fact that *how come* questions with *do* are syntactically ill-formed unless the *do* is stressed (Idid not stress *do*). These possibilities cannot be tested, however, because the design of the experiment does not allow us to do an analysis of variance with *wh*-word as a factor. I would simply like to point out that the difference was between the *can* and the *do* versions were not anywhere near significant.

4. Conclusion

In English, matrix clauses are marked exactly once for tense (henceforth, the English Tense Constraint or ETC). Thus, singly-tensed sentences like *she eats* are grammatical, whereas doubly-tensed sentences like **she does eats* and untensed sentences like **she eat* are ungrammatical. Two types of errors cited in the acquisitional literature suggest that children acquiring English disobey the ETC. Brown (1968) and others have noted that children often ask questions that lack obligatory auxiliaries. Hurford (1975) and others have noted that children asked questions that are tensed twice. In order to obey the ETC, children must know 1) what a matrix clause is, 2) what the tense markers of English are, and 3) what

lexical items can carry tense. Do children know these three things and obey the ETC, or do they frequently disobey the ETC as is suggested by the existence of the types of errors noted by Brown and Hurford?

The results of the first transcript analysis indicates that double-tensed questions are remarkably rare, representing somewhere between 0.04% to 0.4% of children's spontaneous questions. The results of the judgment experiment corroborate this. Children judged double-tensed questions to be significantly worse than questions that were tensed once. I would argue, therefore, that it is unwise to look to double-tensed questions as support for a copy-and-delete formulation of movement transformation. In the second transcript analysis, when children asked questions that lacked obligatory auxiliaries, usually they tensed the main verb. In the judgment experiment, children like questions that lacked an obligatory auxiliary but had a tensed main verb just as often as they liked grammatical questions. In addition, the results of the judgment experiment suggests that children consider double-tensed questions to be neither better nor worse than untensed questions.

Hence, the results of the two transcript analyses and the judgment experiment indicate that children do know and obey the ETC, but that initially they do not seem to know how to tense matrix questions in English. Specifically, they seem to have difficulty with the fact that English requires that only an auxiliary or copula carry tense in matrix questions. For the children, *do*-support seems to be a particular stumbling block. In the first transcript analysis, somewhere between 80-90% of the handful of double-tensed questions involved *do*-support. In the second analysis, I found most auxiliary-less questions looked like they were the result of failure to raise the verb. The results of both transcript analyses

are consistent with children merely having trouble with the mechanics of *do*-support. The results of the judgment experiment suggest that the problem may be more basic. That the children judged questions such as **what Kermit eats?* to be just as good as grammatical questions with *do*-support indicates that they did not realize that *do*-support is obligatory in matrix questions that lack an auxiliary.

We have an interesting combination of knowledge and ignorance. On the one hand children know that English requires that matrix clauses be tensed exactly once. They can identify a matrix clause and the inflectional markers. They know which words are verbs and they know that tense markers attach to verbs. They even know that while English has verb-raising, in English main verbs can't be raised. The one thing they haven't quite figured out is what to do about the fact that verb-raising is obligatory in questions but main verbs can't move. As the results of chapter 5 indicate, if a question had an auxiliary, children inverted it about 93% of the time. When they managed to invoke *do*-support, they placed the *do* in front of the subject NP about 95% of the time (99% of the time if *do* was not negated).

But the results of transcript analyses reported in this chapter suggest that children don't always manage to successfully invoke *do*-support. Furthermore, the results of the judgment experiment reported in this chapter suggest that they may not even realize that they should try. I would argue that the difference between the aspects of tensing children succeed at (obeying the ETC, not inverting main verbs, etc.) and *do*-support is that *do*-support is a rule which is specific to English, whereas the others aspects are reflections of the principles of Universal Grammar. Thus, *do*-support is one aspect of language that children must truly learn.

Chapter 7

Conclusions

The past six chapters have been a discussion of the acquisition of auxiliary verbs. Chapter 2 reviewed the linguistic behavior of auxiliary verbs and showed how the behavior of auxiliaries verbs is similar and different from the behavior of main verbs. The behavior of each type of auxiliary verb is restricted and these restrictions have been fleshed out in different ways by different linguistic theories. At one extreme there is the approach taken by GPSG which essentially lists those inflections, order, and combinations of auxiliaries are acceptable. Other theories have argued that the restrictions on auxiliaries are the result of selection restrictions based on the defective morphological nature of the auxiliaries. Others have argued that semantic restrictions are essential in determining which combinations and orders are acceptable.

Baker (1981) and Pinker (1984) had a slightly different agenda when they set out to describe the behavior of the auxiliaries. Whereas most of the linguistic theories of auxiliaries have merely been content to adequately describe the behavior of the auxiliaries within a one particular linguistic framework or another, Baker and Pinker wanted a description of the auxiliary system which would be learnable to a child. Baker and Pinker both assumed that the behavior of auxiliaries could be captured if one assumed that auxiliaries have defective paradigms which restrict the combinations, inflections, and orders auxiliaries can appear in. Both Baker and Pinker assumed that each auxiliary had separate different paradigms

for declarative and non-declarative contexts. They made the critical assumption that children would not use knowledge about the behavior of main verbs or other auxiliaries to fill out the paradigms for an auxiliaries. Most importantly, children would not not generalize what they know about main verb *be*, *do*, or *have* to auxiliary *be*, *do*, or *have*, despite the identical appearance of the two forms. In the last section of chapter 2, I outlined some of the groupings of verbs which are plausible for either linguistic or learning-theoretic reasons and described the types of errors and correct generalizations children would make if they had each type of grouping.

Chapter 3 examined the auxiliary errors in the spontaneous speech transcripts of 14 children. Among over 55,000 uses of auxiliaries and homophonous main verb and 15,000 uses of negation markers, there were very few errors aside from errors of inversion and errors that involved *do*-support. Most notably, over-regularization errors (e.g., *doed* for *did*, etc.) only occurred for main verb *be*, *do*, and *have*, and never for auxiliary *be*, *do*, and *have*. This suggests that, not only do children distinguish between auxiliary and main verb forms of *be*, *do*, and *have*, but that children treat main verbs and auxiliary verbs differently. They use a productive strategy of generalization for main verbs but not auxiliary verbs. In Baker and Pinker's terminology, children use general paradigms or rules to fill out the paradigms for main verbs but not auxiliary verbs. Another notable, systematic restriction concerned *do*-support. In 16 cases, *do*-support was unnecessarily invoked for negation or inversion in utterances which already contained an element which could invert or negate. In all of these examples, the element was copula *be*. It was never an auxiliary. This suggests that children have difficulty with the

exceptional behavior of copula *be*, and occasionally treat copula *be* as if it were a regular main verb.

In chapter 4 examined the patterns of acquisition of auxiliaries and homophonous main verbs in declaratives, interrogatives, and negatives. The acquisitional patterns were statistically different for auxiliaries than main verbs. In addition, within the auxiliaries, there were statistically significant acquisitional clusters. A number of possible groupings of auxiliaries were reasonably good, but the groupings of Be&Have/Do/Modals and Be/Have/Do/Modals were the best. Interesting, the grouping Modal/NonModal which did not distinguish among *be*, *do*, and *have* was not as good as a model which treated every form of every auxiliary as the sole member of its own category. Thus, it appears that there is(are) some acquisitionally important distinction(s) among *be*, *do*, and *have*. In addition to the differences between the various auxiliaries, copula and auxiliary *be* had statistically significant different patterns of acquisition.

Chapter 5 examined more carefully the subject-auxiliary inversion errors. While inversion errors were relatively common, not all inversion errors occurred with the same frequency. Some errors essentially never occurred. Children essentially never inverted verbs that cannot invert and no child ever consistently failed to invert an auxiliary or copula which does invert in English. In addition, there were statistically significant differences in the inversion rates for individual children, *wh*-words, and auxiliaries. Some, but not all children seemed to go through a period during which they were worse at inverting. These children who got worse at inverting were the same children who had lower overall rates of inversion. There was also a difference in inversion rates for the various

wh-words which seemed to reflect a division between argument questions and adjunct questions. The higher rates of inversion for argument questions than adjunct questions may reflect a structural difference between the two types of questions such as a difference in where argument and adjunct *wh*-words are placed.

There was also a statistically significant difference in inversion rates for the different types of auxiliaries. Copula *be* was inverted less often than the auxiliaries. I argued that the reason copula *be* is inverted less often is that the children were somewhat confused about why a non-auxiliary inverts and, therefore, they sometimes made mistakes. In addition to the differences in inversion rates for the various types of auxiliaries, there was a statistically significant difference in the inversion rates for negated and non-negated forms of auxiliaries. While this might just represent an increased processing load for inversion plus negation, the difference might reflect a structural difference between negated and non-negated auxiliaries such as that described in chapter 5.

In addition to comparing the inversion rates for various types of matrix questions, I compared the inversion rates for matrix questions and embedded questions. Contrary to reports in the literature, the error of inverting in embedded questions was more common than the error of failing-to-invert in matrix questions. In addition, for the 7 types of *wh*-questions, the rates of inversion in embedded and matrix *wh*-questions were highly correlated. Another surprising discovery was that there were no examples of inverted embedded *yes/no* questions.

Chapter 6 examined more carefully children's knowledge of auxiliaries and tense. Questions that had more than one tense inflection and questions that lacked obligatory auxiliaries were examined. Almost all of

the double-tensed questions contained *did* with irregular past tense verbs. This suggests that many double-tensed questions were the result of difficulty with *do*-support coupled with difficulty with the retrieval of infinitival forms of irregular past tense verbs. The majority of questions that lacked obligatory auxiliaries had tensed main verbs. This suggests that these auxiliary-less questions resulted from failure to invoke *do*-support for inversion. In a judgment experiment, children judged questions that had no tense markers or two tense markers to be equally bad and significantly worse than questions that had one tense marker. Surprisingly, children judged questions such as *what does Kermit eat?* and *what Kermit eats?* to be equally good. Thus, it would appear that children know that sentences must be marked exactly once for tense, but they do not realize that *do*-support must be invoked in questions that lack auxiliaries.

How can the pattern of findings described in the past 6 chapters be accounted for? Most of the logically possible auxiliary errors never occurred or occurred so rarely that they probably were just slips of the tongue. If we look at the errors that essentially never occurred or were judged as bad in the judgment experiments, we see that all of the errors resulted in clause structures that are not acceptable in any setting in English. Hence, this result could be the reflection of the children having adopted some form of Slobin's Fixed Word Order Operating Principle.[1] However, it may be that the children were obeying some constraint of Universal Grammar which prohibits certain types of clauses from occurring.

---

[1] Slobin's Fixed Word Order Operating Principle. "If you have determined that word order expresses basic semantic relations in your language, keep the order of morphemes in a clause constant" (Slobin, 1985, p. 1255).

But by far the most striking finding is that children distinguish main verbs from auxiliaries, even though there are main verbs which are homophonous to most of the auxiliaries. The results discussed in this thesis suggest that children distinguish main verbs from auxiliaries even before they have acquired much of the syntax which linguists use to distinguish auxiliaries from main verbs. This suggests the children must be innately endowed with knowledge that makes it possible for them to distinguish main verbs from auxiliary verbs.

To the extent that auxiliaries are universal to all languages and to the extent that the features auxiliaries exhibit across languages are similar, children could be born knowing much of what they need to know in order to distinguish auxiliary verbs from main verbs. Steele et al. (1981) identified a number of properties of auxiliaries which are universal. That children are able to distinguish main verbs from auxiliaries is evidence that the category Aux and its properties are universal That children use a productive strategy in acquiring the main verbs, and a conservative strategy in acquiring auxiliaries suggests that children have an innate knowledge of the existence of different kinds of linguistic categories, open-class categories which are productive and allow fairly free generalization and closed-class categories which are closed to generalizations and innovations.

The intricacies of the auxiliary system are such that most of the logically possible generalizations result in errors which are unrecoverable without negative evidence. The combinatorics of the auxiliary system are such that it is impossible to acquire the auxiliary system in a finite amount of time without making any generalizations or assumptions about what is being acquired. Thus, the child is forced

to make those generalizations which Universal Grammar allows. Some of these Universal Grammar-based generalizations are specific to the auxiliary system (e.g., Steele's universals) and some relate to more general aspects of language (e.g., universal aspects of tense). In those aspects of the auxiliary system which Universal Grammar can guide the child, her acquisition is early and essentially error-free. Where Universal Grammar cannot guide the child (i.e., in language-particular aspects such as *do*-support), acquisition is protracted and the grammar of the child may differ from that of the adult.

References

Akmakian, A., Steele, S., & Wasow, T. 1979. The category AUX in Universal Grammar. *Linguistic Inquiry, **10***, 1-64.

Anderson, S. 1985. Inflectional morphology. In T. Shopen (ed.), *Language typology and syntactic descriptions. Vol. 3: Grammatical categories and the lexicon*. Cambridge: Cambridge University Press.

Baker, C. 1981. Learnability and the English auxiliary system. In C. Baker and J. McCarthy (eds.) *The logical problem of language acquisition* . Cambridge: MIT Press.

Baker, C. 1984. Two observations on British English *Do* . *Linguistic Inquiry, **15,*** 155-157.

Bellugi, U. 1967. The acquisition of the system of negation in childrens speech. Unpublished doctoral thesis. Harvard School of Education.

Bellugi, U. 1971. Simplification in children's language. In R. Huxley and E. Ingram (Eds.), *Language acquisition: Models and methods.* New York: Academic Press.

Bellugi, U., & Brown, R. 1964. *The acquisition of language*. Report on the 4th Conference sponsored by the Committee on Intellective Processes Research of the Social Sciences Research Council. Chicago: University of Chicago Press.

Berman, R., & Dromi, E. 1984. On marking time without aspect in child language, *Papers and Reports in Child Language Development, **23***. Stanford.

Blackenmeyer, G., & Weston, A. 1987. Experimental analysis of response-class learning beween *is* copula and auxiliary and *are* copula and auxiliary. *Perceptual and Motor Skills, **64**(1)*, 27-43.

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

Bloom, L. 1973. *One word at a time: The use of single word utterances before syntax*. The Hague: Mouton.

Bloom, L., & Lakey, M. (Eds.) 1978. *Language development and language disorder*. New York: Wiley.

Bloom, L., Lifter, K., & Hafitz, J. 1980. Semantics of verbs and the development of verb inflections in child language. *Language, **56***, 386-412.

Bloom, L., Merkin, S., & Wootten, J. 1984. *Wh*-questions: Linguistic factors that contribute to the sequence of acquisition. *Child Development, **53***, 1084-1092.

Bloom, L., Rispoli, M., Gartner, B., & Hafitz, J. 1989. Acquisition of complementation. *Journal of Child Language, **16***, 101-129.

Bresnan, J. 1978. A realistic transformational grammar. In M. Halle, J. Bresnan and G. Miller (eds.), *Linguistic theory and psychological reality*. Cambridge: MIT Press.

Bresnan, J. 1982. *The mental representation of grammatical relations*. Cambridge, MA: MIT Press.

Bresnan, J., & Grimshaw, J. 1978. The syntax of free relatives in English. *Linguistic Inquiry, 9*, 331-391.

Brinton, L. 1988. *The development of English aspectual systems*. Cambridge: CUP

Brown, R. 1968. The development of *wh* -questions in child speech. *Journal of Verbal Learning and Verbal Behavior, 7*, 279-290.

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

Brown, R., Cazden, C., & Bellugi-Klima, U. 1969. The child's grammar from I to III. In J. Hull (Ed.), *Minnesota symposium on child psychology, Volume 2*. Minneapolis, MN: University of Minnesota Press.

Brown, R., & Fraser, C. 1964. The acquisition of syntax. In U. Bellugi & R. Brown (Eds.).*The acquisition of language*. Report on the 4th Conference sponsored by the Committee on Intellective Processes Research of the Social Sciences Research Council. Chicago: University of Chicago Press.

Brown, R., Frazar, C., & Bellugi, U. 1964. Explorations in grammar evaluation. In U. Bellugi & R. Brown (Eds.).*The acquisition of language*. Report on the 4th Conference sponsored by the Committee on Intellective Processes Research of the Social Sciences Research Council. Chicago: University of Chicago Press.

Brown, R., & Hanlon, C. 1970. Derivational complexity and order of acquisition. In J. Hayes (ed.), *Cognition and the development of language*. New York: Academic Press.

Cairns, H. & Hsu, J. 1978. Who, why, when and how: A developmental study. *Journal of Child Language, 5*, 477-488.

Chisholm Jr., W (ed.) 1984. *Interrogativity: Typological studies in language, volume 4*.. Amsterdam: John Benjamins Publishers.

Chomsky, N. 1957. *Syntactic structures*. The Hague: Mouton.

Chomsky, N. 1981. *Lectures on government and binding*. Dordrecht, Netherlands: Foris Publications.

Chomsky, N. 1982. *Some concepts and consequences of binding*. Cambridge, MA: MIT Press.

Chomsky, N. 1986a. *Barriers*. Cambridge, MA: MIT Press.

Chomsky, N. 1986b. *Knowledge of language: Its nature, origin and use*. New York: Praeger.

Chung, S. & Timberlane, A. 1985. Tense, aspect, and mood. In T. Shopen (ed.), *Language typology and syntactic description: Vol 3, Grammatical categories and the lexicon.* Cambridge: Cambridge University Press.

Clark, E. 1978. Strategies for communicating. *Child Development, **49**,* 953-959.

Comrie, B. 1981. *Language universals and linguistic typology.* Oxford, England: Basil Blackwell

Comrie, B. 1984. Interrogativity in Russian. In W. Chisholm Jr., L. Milic, & J. Greppin (eds.), *Interrogativity: A colloquium on the grammar, typology and pragmatics of questions in seven diverse languages.* Amsterdam: John Benjamins Publishing Co.

Comrie, B., 1985. *Tense.* Cambridge: CUP

Comrie, B., 198?. *Aspect.* Cambridge: CUP.

Crain, S., & Nakayama, M. 1987. Structure dependence in grammar formation. *Language, **63**,* 522-543.

Dahl, O. 1985. *Tense and aspect systems.* Oxford: Basil Blackwell,

Davis, L. 1987. Remarks on government and proper government. *Linguistic Inquiry, **18**,* 311-321.

de Villiers, J. 1981. The process of rule learning in child speech: A new look. In K. Nelson (ed.), *Children's language, Vol.2.* New York: Gardner Press.

de Villiers, J. 1984. Form and force interactions: The development of negatives and questions. In R. Schifelbusch & J. Pickar (Eds.), *Communication competence: Acquisition and intervention.* Baltimore, MD: University Park Press.

deVilliers, J. 1985. Learning how to use verbs: Lexical coding and the influence of the input. *Journal of Child Language,* 12, 587-595.

de Villiers, J., & de Villiers, P. 1973. Syntax and semantics in the first two years. In F. Minifie and L. Llyod (eds.), *Communicative and cognitive abilities: Early behavioral assessment.* Baltimore, MD: University Park Press.

de Villiers, J., & de Villiers, P. 1974. Competence and performance in child language: Are children able to judge? *Journal of Child Language, **1**,* 11-22.

de Villiers, J., & de Villiers, P. 1985. The accquisition of English. In D. Slobin (Ed.), *The crosslinguistic study of language acquisiton, volume 1: The data.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Derwing., & Smyth, R. 1988. Auxiliary placement revisited. *Journal of the Atlantic Provinces Association, **10**.* Offprint.

Dik, S. 1983. Auxiliary and copula *be* in Functional Grammar. In F. Heny & B. Richards (eds.), *Linguistic categories: Auxiliaries and related puzzles, Vol 2*. Dordrecht: Reidel Publishing Company.

Emonds, J. 1985. A unified theory of syntactic categories. Dordrect, Holland & Cinnaminson, NJ: Foris.

Emonds, J. 1976. *A transformational approach to English syntax*. New York: Academic Press.

Engdahl, E. 1986. *Constituent questions: Studies in linguistics and philosophy 27*. Dordecht, Holland: Reidel Publishers.

Erreich, A. 1984. Learning how to ask: Patterns of inversion in *yes-no* and *wh-* questions. *Journal of Child Language, **11***, 579-502.

Ervin-Tripp, S. 1970. Discourse agreement: How children answer questions. In J. Hayes (Ed.), *Cognition and the development of language*. New York: Wiley.

Falk, Y. 1984. The English auxiliary system: A Lexical Functional Analysis. *Language, **60***, 483-509.

Fasold, R. & Schiffrin, D. (eds.) 1989. *Language change and variation. Current Issues in Linguistic Theory, **52***. (Papers originally presented at NWAVE XI, Georgetown, 1982). Philadelphia: John Benjamins

Fay, D. 1978?. Transformations as mental operations: A reply to Kuczaj. *Journal of Child Language, 5*, 143-149.

Gazdar, G. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry, **12***, 155-184.

Gazdar, G., Klein, E., Pullum, G., & Sag, I. 1985. *Generalized phrase structure grammar*. Cambridge, MA: Harvard University Press.

Gazdar, G., Pullum, G., & Sag, I. 1982. Auxiliaries and related phenomena in a restrictive theory of grammar. *Language, **58***, 591-638.

Gerken, L. 1987. Telegraphic speaking does not imply telegraphic listening. *Papers and Reports in Child Language Development, **26***, 48-55.

Gibson, K. 1986. The ordering of auxiliary notions in Guyanese Creole. *Language, **62***, 571-586.

Givon, T. 1985. Function, structure and language acquisition. In D. Slovin (ed.), *The crosslinguistic study of language acquisistion, volume 2*. Hillsdale, NJ: Lawrence Erlbaum Associates

Givon, T. 1985 Iconicity, isomorphism, and non-arbitary coding in syntax. In J. Haiman (ed.), *Iconicity in syntax*. Amersterdam: John Benjamins.

Goodluck, H. & Solan, L. 1979. A reevaluation of the basic operations hypothesis. *Cognition, 7*, 85-91.

Green, G. 1980. Some wherefores of English inversion. *Language*, **56**, 582-600.

Grimshaw, J. 1979. The structure-preserving constraint: A review of *A transformational approach to English syntax* by J. E. Emonds. *Linguistic Analysis*, **5**, 313-343.

Grimshaw, J. 1981. Form, function and the language acquisition device. In C. Baker and J. McCarthy (eds.), *The logical problem of language acquisition*. Cambridge: MIT Press.

Grimshaw, J. 1985. *English wh-constructions and the theory of government*. NY: Garland Publishing Co. Published version of Grimshaw's 1977 UMass Amherst dissertation

Harris, M., & Ramat, P., ed., 1987. *Historical development of auxiliaries*. New York: Mouton de Gruyter.

Hawkins, J. 1983. *Word order universals*. New York: Academic Press.

Heny, F. & Richards, B. 1983, eds. *Linguistic categories: Auxiliaries and related puzzles. Vol 1-2*. Dordrecht: Reidel.

Higginson, R. 1985. Fixing: A demonstration of the child's active role in language acquisition. Unpublished doctoral dissertation. Washington State University.

Hiz, H (Ed.) 1978. *Questions: Syntheses language library, volume 1*. Dordecht, Holland: Reidel Publishers.

Hoeksema, J. 1985. Traces in COMP: Invisible or nonexistent? *Chicago Linguistic Societry Papers* **21**, 160-170.

Hoff-Ginsberg, E. & Shatz, M. 1982. Linguistic input and the child's acquisition of language. *Psychological Bulletin*, **92**, 3-26.

Hopper, P., ed., 1982. *Tense-Aspect: Between semantics and pragmatics*. Amsterdam: John Benjamins Publishing Co.

Hurford, J. 1975. A child and the English question formation rule. *Journal of Child Language*, **2**, 299-301.

Ingram, D. 1989. *First language acquisition: Method, description and explanation*. Cambridge: Cambridge University Press.

Ingram, D., & Tyack, D. 1979. Inversion of Subject NP and Aux in children's questions. *journal of Psycholinguistic Research*, **8**, 333-341.

Janda, R. 1985. Echo-questions are evidence for what? *Chicago Linguistics Society Papers* **21**, 171-188.

Kaisse, E. 1983. The syntax of auxiliary reduction in English. *Language*, **59**, 93-122.

Kaisse, E., 1984. Review of S. Steele *et al. An encyclopedia of AUX*. *Language*, **60**, 924-928.

Karttunen, L. 1977. Syntax and semantics of questions. *Linguistics and Philosophy*, **1**, 3-44.

Kayne, R. 1984. *Connectedness and binary branching*. Dordrecht-Holland: Foris Publications.
Keenan, E. 1985. Relative clauses.In T. Shopen (ed.), *Language typology and syntactic descriptions. Vol. 2: Complex constructions.* Cambridge: Cambridge University Press.
Kessel, F., ed., 1988. *The development of language and language researchers: Essays in honor of Roger Brown.* Hillsdale, NJ: LEA
Kilby, D. 1984. Descriptive syntax and the English verb. London: Croom Helm.
Klee, T. 1981. Semantic and syntactic aspects of children's questions at three linguistic stages. Unpublished Ph.D. dissertation. University of Wisconsin, Madison.
Klee, T. 1985. Role of inversion in children's question development. *Iournal of Speech and Hearing Research, **28**,* 225-232.
Kroch, A. 1989. Function and grammar in the history of English: Periphrastic *do*. In Fasold, R. & Schiffrin, D. (eds.) 1989. *Language change and variation. Current Issues in Linguistic Theory, **52**,* 133-172. Philadelphia: John Benjamins.
Kuczaj, S. 1976. Arguments agains Hurford's 'Aux Copying Rule'. *Journal of Child Language, **3**,* 423-427.
Kuczaj, S. 1985/86. General developmental patterns and individual differences in the acquisiton of copula and auxiliary *be* forms. *First Language, **6**,* 111-117.
Kuczaj, S. 198?. More on children's initial failure to relate specific acquisitions. *Journal of Child Language, **8**,* 485-487.
Kuczaj, S. 1977. The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior, **16**,* 589-600.
Kuczaj, S. 1978. Children's judgments of grammatical and ungrammaticalirregular past-tense verbs. *Child Development, **49**,* 319-326.
Kuczaj, S. 197?. Why do children fail to overgeneralize the progressive inflection? *Journal of Child Language, **5**,* 167-171.
Kuczaj, S. 1981. More on children's initial failure to relate specific acquisitions. *Journal of Child Language, **8**,* 485-487.
Kuczaj, S. & Brannick, N. 1979. Children's use of the *wh-* question modal auxiliary placement rule. *Journal of Experimental Child Psychology, **28**,* 43-67.
Kuczaj, S., & Maratsos, M. 1975. What children can say before they will. *Merrill-Palmer Quarterly, **21**,* 89-111.
Kuczaj, S., & Maratsos, M. 1983. The initial verbs of yes-no questions: A different kind of general grammatical category. *Developmental Psychology, **19**,* 440-444.

Lapointe, S. 1980. A lexical analysis of the English auxiliary verb system. In T. Hoekstra *et al.* (eds.), *Lexical grammar*. Dordrecht: Foris.

Lasnik, H. 1981. Learnability , restrictiveness and the evaluation metric. In C. Baker and J. McCarthy (eds.), *The logical problem of language acquisition*. Cambridge, MA: MIT Press.

Lasnik, H. & Saito, M. 1984. On the nature of proper government. *Linguistic Inquiry, **15**,* 235-289.

Labov, W. & Labov, T. 1978. Learning the syntax of questions. In R. Campbell & P. Smith (Eds.), *Recent advance in the psychology of language: Formal and experimental approaches*. New York: Plenum Press.

Labov, W. & Harris, W. 1986. De Facto segregaton of black and white vernaculars. In D. Sankoff (ed.) *Diversity and diachrony. Current Issues in Linguistic Theory, **53***. Philadelphia: John Benjamins.

Lightfoot, D. 1979. *Principles of diachronic syntax*. Cambridge, England: Cambridge University Press.

Lightfoot, D. 1982. *The language lotterty*. Cambridge, MA: MIT Press.

Lüdtke, H. 1987. Auxiliary verbs in the universal theory of language change. In A. Ramat, O. Carruba, & G. Bernini (Eds.), *Papers from the 7th International Conference on Historical Linguististics* . p. 349-354. Amsterdam: John Benamin Publishing Co..

McShane, J, & Whittaker, S. 1988. The encoding of tense and aspect by 3-to-5-year-old children. *Journal of Experimental Child Psychology, **45***, 52-70.

MacWhinney, B., & Snow, C. 1985. The Child Language Data Exchange System. *Journal of Child Language, **12***, 271-296.

Majors, D. 1974. *The acquisition of modal auxiliaries in the language of children*. The Hague: Mouton

Mapstone, E. & Harris, P. 1985. Is the English present progressive unique. *Journal of Child Language, **12***, 433-441.

Maratsos, M. 1984. Some current issues in the study of the acquisition of grammar. In J. Flavell & E. Markman (Eds.), *Handbook of child psychology, fourth edition. Vol. 3: Cognitive development.*

Maratsos, M. 1988. The acquisition of formal word classes. In Y. Levy, I. Schlesinger & M. Braine (eds.), *Categories and processes in language acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Maratsos, M. 1988. Crosslinguistic analysis, universals, and language acquisition. In F. Kessel, (ed.), *The development of language and language researchers: Essays in honor of Roger Brown*. Hillsdale, NJ: LEA

Maratsos, M. 1989. Innateness and plasticity in language acquisition. In M. Rice & R. Schiefelbusch, eds., *The teachability of language*.

Maratsos, M. & Kuczaj, S. 1978. Against the transformationalist account: A simpler analysis of auxiliary overmarkings. *Journal of Child Language, 5,* 337-345.

Mayer, J. W., Erreich, A., & Valian, V. 1978. Transformations, basic operations and languge acquisition. *Cognition, 6,* 1-13.

Menyuk, P. 1969. *Sentences children use.* Cambridge, MA: MIT Press.

Myhill, J. & Harris, W. 1986. The use of the verbal -s inflection in BEV. In D. Sankoff (ed.) 1986. *Diversity & diachrony. Current Issues in Linguistic Theory, vol 53.* Philadelphia: John Benjamins.

Nakayama, M., & Crain, S. 1985. Performance factors in children's sentence production. Paper presented at the 10th Annual BU Language Conference.

Newport, E., Gleitman, H., & Gleitman, L. 1977. Mother, I'd rather do it myself: Some effects and non-effects of maternal speech style. In C. Snow and C, Ferguson (eds.), *Talking to children.* Cambridge: Cambridge University Press.

Parnell, M., Patterson, S., & Harding, M. 1984. Answers to *wh-* questions: A developmental study. *Journal of Speach and Hearing Research, 27,* 297-305.

Payne, J. 1985b. Negation. In T. Shopen (ed.), *Language typology and syntactic descriptions. Vol. 1: Clause structure.* Cambridge: Cambridge University Press.

Pinker, S. 1981. On the acquisition of grammatical morphemes. *Journal of Child Language, 8,* 477-484.

Pinker, S. 1984. *Language learnability and language development.* Cambridge, MA: Harvard University Press.

Pinker, S. & Prince, A. 1988. On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition, 28,* 73-193.

Pinker, S., Prince, A., Hollander, M., Kim, J., Marcus, G., Prasada, S., and Ullman, M. 1990. Why do children overregularize. Paper presented at the 20th annual meeting of the Cognitive Sciences Society, July, 1990/

Prideaux, G. 1976. A functional analysis of English questions acquisition: A response to Hurford. *Journal of Child Language, 3,* 417-422.

Perkins, M. 1983. Modal expressions in English. London: Frances Pinter Publishers.

Pullum, G., & WIlson, D. 1977. Autonomous syntax and the analysis of auxiliaries. *Language, 53,* 741-788.

Radford, A. 1981. *Transformational syntax.* Cambridge, England: Cambridge University Press.

Rissanen, M. 1986. Variation and the study of English historical syntax. In D. Sankoff(ed.), *Diversity & diachrony. Current Issues in Linguistic Theory, vol 53*. Philadelphia: John Benjamins .

Rice, M., & Schiefilbusch, eds. 1989 *The teachability of language*. Baltimore: Paul H. Brookes, Publishing Co..

Ross, J. 1969. Auxiliaries as main verbs. *Studies in Philosophical Linguistics, 1*, 77-102.

Sachs, J. 1983. Talking about here and then: The emergence of displaced reference in parent-child discourse. In K. Nelson (Ed.), *Children's language, Volume 4*. Hillsdale, NJ: Lawwrence Erlbaum Press.

Sankoff, D. (ed.) 1986. *Diversity & diachrony. Current Issues in Linguistic Theory, vol 53*. Philadelphia: John Benjamins.

Schachter, P. 1983. Explaining auxiliary order. In F. Heny & B. Richards (eds.), *Linguistic categories: Auxiliaries and related puzzles, Vol 2*. Dordrecht: Reidel Publishing Company.

Shatz, M., Hoff-Ginsverg, E., Maciver, D. 1989. Induction and the acquisition of English auxiliaries: the effects of differentially enriched input. *Journal of Child Language*, 16, 121-140.

Sells, P. 1985. *Lectures on contemporary syntactic theories: An introduction to Government-Binding Theory, Generalized Phrase Structure Grammar, and Lexical Functional Grammar. CSLI Lecture Notes 3*. Stanford, CA: Center for the Study of Language and Information.

Slobin, D. 1985. Crosslinguistic evidence for the language-making capacity. In D. Slobin (Ed.), *The crosslinguistic study of language acquisiton, volume 2: Theoretical issues*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Spears, A. 1982. The Black English semi-auxiliary "come". *Language, 58*, 850-872.

Steele, S. et. al. 1981. *An encyclopedia of AUX: A study in cross-linguistic equivalence*. Cambridge, MA: MIT Press.

Stemberger, J., & MacWhinney, B. 1988. Are inflected forms stored in the lexicon? In *Theoretical Morphology*, NY: Academic Press.

Stewart, J.,& Sinclair, H. 1975. Comprehension of questions by children between 5 and 9. *International Journal of Psycholinguistics, 3*, 17-26.

Stromswold, K. 1989. Using naturalistic data: Methodological and theoretical issues (or How to lie with naturalistic data) Invited Paper at the 14th Annual Boston University Child Language Conference, October, 1989.

Stromswold, K. 1989. Productivity, conservativism and the acquisition of auxiliaries. Paper to be presented at the 21st

Annual Stanford Child Language Research Forum, April 7-9, 1989.

Stromswold, K. 1989. Children's knowledge of tense. Paper to be presented at the 60th Annual Meeting of the Eastern Psychologyical Association, March 30- April 2, 1989.

Stromswold, K. 1988. The structure of children's *wh*-questions. Invited Paper at the 13th Annual Boston University Child Language Conference, October 22, 1988. .

Stromswold, K. 1988. Linguistic representations of children's *wh*-questions. Presented at the 20th Annual Stanford Child Language Research Forum, April 9, 1988

Torrey, J. 1983. Black children's knowledge of Standard English. *American Educational Research Journal, **20***, 626-643.

Tyack D., & Ingrams, D. 1977. Children's production and comprehension of questions. *Journal of Child Language, **4,*** 211-228.

Ultan, R. 1978. Some general characteristics of interrogative systems. In J. Greenberg (Ed.), *Universals of human language, Volume 4: Syntax*. Stanford, CA: Stanford University Press.

van Riemsdijk, H., & Williams, E. 1986. *Introduction to the theory of grammar*. Cambridge, MA: MIT Press.

Wachowicz, K. 1974. Against the unversality of a single *wh*- question movement. *Foundations of Linguistics, **11***, 155-166.

Wells, G. 1979. Learning and using the auxiliary verb in English. In V. Lee (Ed.), *Language development*. New York: Wiley.

Weist, R. 1984. Tense and aspect: temporal systems in child language. In P. Fletcher and M. Garman, eds., Language acquisition: Studies in first language development, second edition. CUP.

Wolff, J. 1988. Learning syntax and meanings through optimization and distributional analyses. In Y. Levy, I. Schlesinger & M. Braine (eds.), *Categories and processes in language acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates

Wolfrman, W. 1990. Reexamining Vernacular Black English. Review article of E. Schneider, 1989, *American earlier Black English: Morphologicla and syntactic variables* (Tuscaloosa, AL: U of Alabama Press) and R. Butter, 1989, *The death of Black English: Divergence and convergence in black and white vernaculars* (Frankfurt am Main: Peter Lang). *Language, **66***, 121-133.

Zwicky, A., & Pullum, G. 1983. Cliticization vs. inflection: English *n't*. *Language, **59***, 502-513.

Table 3.1
Transcripts Used in Error Analyses

Child	Transcripts collected by	Age Range	Approx. # Aux's*
Adam	Brown (1973)	2;3 - 5;2	11,600
Allison	Bloom (1973)	1;4 - 2;10	200
April	Higginson (1985)	1;10 - 2;11	1,200
Eve	Brown (1973)	1;6 - 2;3	1,800
June	Higginson (1985)	1;3 - 1;9	200
Mark	MacWhinney & Snow (1985)	1;5 - 6;0	2,100
Naomi	Sachs (1983)	1;2 - 4;9	3,100
Nathan	MacWhinney & Snow (1985)	2;6-3;9	1,700
Nina	Suppes (1973)	2;0 - 3;3	7,500
Peter	Bloom (1973)	1;10 - 3;2	5,200
Ross	MacWhinney & Snow (1985)	2;10 - 7;10	11,700
Sarah	Brown (1973)	2;3 - 5;1	6,900
Shem	Clark (1978)	2;3 - 3;2	2,500
Total			55,700

* The numbers given are the number of lines which contained strings of letters that could be auxiliaries. In other words, the figures given include lines that contain main verbs which are homophonous with auxiliaries.

Table 3.2

Overregularized Verbal Inflections

Adam29 (3;5)	he just haves [?] a cold.
Adam30 (3;5)	you do dis kind. you dos [?] dis kind of bowling ball
Adam31 (3;5)	from de store. de milkman dos [#] from de store.
Adam32 (3;6)	And Mommy doos two.
Adam35 (3;8)	somebody teach me to sing. I almost <do> [/] [#] doed a
Adam35 (3;8)	a drill. what a drill dos?
Adam37 (3;10)	Mommy [#] why it dos like that?
Eve17 (;22)	what did you doed?
Eve17 (2;2)	and wha- [/] what did you doed?
Eve19 (2;3)	xxx on floor. what what does [C/int: ie duwz]
April04 (2;9)	what does he dos?
Mark59.(3;5)	He said "I ams"
Naomi27 (2;0)	I doed it
Naomi75.(3;0)	[#] cuz it doed it under my arm [#]
Naomi89.(3;5)	and he did [#] he beed careful [#] he looked both ways
Naomi89.(3;5)	that was [#] he beed careful he looked both ways.
Nathan09 (2;7)	enne Nathaniel doos it
Nathan15 (3;1)	de de de de en eh ee ee ee eh ee dos like like ih ee ee ee ih ee
Nina33 (2;10)	see what he doed?
Peter10 (2;3)	Humpty Dumpty have [#] have [#] have [#] um Humpty Dumpty haves a great
Ross37.(3;4)	No [#] she bees bad [#] then she bees good [#] okay?
Ross43 (3;8)	Turn them out haves all three.
Ross75 (6;6)	Daddy [#] first Daddy bes first.
Ross75 (6;6)	word? bes was a good word? who said that? who said that bes
Sarah121 (4;8)	she dos what her Mother tells her.
Sarah128 (4;11)	mmhm. he bes all baloneys.
Shem02.(2;3)	you doed it $ turning, there.
Shem03.(2;3)	i doed a star!
Shem27 (2,10)	my daddy doed it.
Shem36.(3;1)	um- but first i hadded to get like dis..
Shem40.(3;2)	yeah, jus' d$ clowns are- some are scary an' some ares funny an' just

Table 3.3

Overregularization BASELINES

ABE: 1 MV *beed*	# Correct Aux *was*	# Correct MV *was*
By transcript 57	28	40
ABE: 6 MV *doed*	# Correct Aux *did*	# Correct MV *did*
By transcript 65	40	49
ABE: 1 MV *hadded*	# Correct Aux *had*	# Correct MV *had*
By transcript 88	1	63
ADAM: 5 MV *do/doos*	# Correct Aux *does*	# Correct MV *does*
By transcript 37	88	7
ADAM: 1 MV *doed*	# Correct Aux *did*	# Correct MV *did*
By transcript 35	77	61
ADAM: 1 MV *haves*	# Correct Aux *has*	# Correct MV *has*
By transcript 29	0	12
APRIL: 1 MV *dos*	# Correct Aux *does*	# Correct MV *does*
By transcript 4	13	1
EVE: 2 MV *doed*	# Correct Aux *did*	# Correct MV *did*
By transcript 17	9	4
EVE: 1 MV /duwz/	# Correct Aux *does*	# Correct MV *does*
By transcript 19	11	1
MARK: 1 MV *ams*	# Correct Aux *am*	# Correct MV *am*
By transcript 59	22	15
NAOMI: 2 MV *beed*	# Correct Aux *was*	# Correct MV *was*
By transcript 89	7	26
NAOMI: 2 MV *doed*	# Correct Aux *did*	# Correct MV *did*
By transcript 75	14	17
NATH: 2 MV *doos/dos*	# Correct Aux *does*	# Correct MV *does*
By transcript 15	2	0
NINA: 1 MV *doed*	# Correct Aux *did*	# Correct MV *did*
By transcript 33	13	18
PETER: 1 MV *haves*	# Correct Aux *has*	# Correct MV *has*
By transcript 10	0	4
ROSS: 4 MV *bes/bees*	# Correct Aux *is*	# Correct MV *is*
through age 3;6	57	296
ROSS: 1 MV *haves*	# Correct Aux *has*	# Correct MV *has*
through age 3;6	2	8
SARAH: 1 MV *bes*	# Correct Aux *is*	# Correct MV *is*

Table 3.3
Auxiliary and Main Verb Overregularization Baselines

By transcript 128	24	659
SARAH: 1 MV *dos*	# Correct Aux *does*	# Correct MV *does*
By transcript 121	63	3
SHEM: 1 MV *ares*	# Correct Aux *are*	# Correct MV *are*
By transcript 40	48	58
SHEM: 3 MV *doed*	# Correct Aux *did*	# Correct MV *did*
By transcript 27	15	36
SHEM: 1 MV *hadded*	# Correct Aux *had*	# Correct MV *had*
By transcript 36	0	16

Table 3.4
Double-tensed Declarative Utterances

Mark56	Wait [#] you didn't got all your space ships.
Nathan8	elephant go [#] moves so slowly [#&] doesn't moves so fast
Nina13	he don't [#] don't broke her bed.
Peter14	I locked it. I did locked it. I did locked. I did locked it. <I like>
Peter14	I'm gon (t)a lock the door. I did locked a door.
Peter17	that didn't went down. you do it right there.
Peter17	is this goes on the [///]
Peter19	I didn't did it. I didn't did it. I
Peter19	one fell down and two fell down and [#] one doesn't fell down.
Peter19	and this one doesn't fell down either. ACOMB
Ross32	He's getting unhappy [#] and he doesn't likes to be unhappy.
Ross34	I did fell
Ross35	But my boots does tickles.
Ross35	And he did jumped on there.
Ross35	I didn't disappeared.
Ross39	Yeah [#] but for a long time it did worked.
Ross55	You ask me [#] it doesn't exist.
Ross68	is uses in schools.
Sarah52	I didn't did it.
Sarah63	you didn't caught me.
Sarah70	I didn't did nothin(g)
Sarah81	Wince did hadn't any.
Sarah82	I didn't broke this.
Sarah86	I didn't spilled it.
Sarah95	you didn't gave me too much.
Sarah96	we didn't saw cartoons.
Sarah107	because [#] nobody didn't broke it.
Sarah131	somone didn't did it.
Shem16	dat (didn't) fell in.
Shem30a	no, you didn't left to go to work
Shem34	(h)e has $ hat, but my daddy has- doesn't has.

Table 3.5
Auxiliary/Inflection Mismatches

Adam9 (2;7)	Don't fighting me
Adam9 (2;7)	cowboy did fighting me
Adam14 (2;10)	don't crying. don't crying
Adam49 (4;7)	She won't makes me.
Eve6 (1;9)	do running fast.
Mark57 (3;1)	my tummy will fell off
Mark80 (4;10)	no I didn't gets to do very much stuff
Naomi37 (2;0)	do Georgie going xxx?
Nathan18 (3;3)	goodnight [#&] en does eh he saying goodnight?
Nina7 (2;1)	this is my doggy crys (repeated three times in a row)
Nina27 (2;6)	Maggie and Leila did going go to Maggie's house
Nina28 2;6)	he was cried
Peter14 (2;6)	that could goin(g) a be a bedroom
Peter18 ()	I can building a tower with you. ACOMB
Peter20 (3;2)	do you lookin(g) at the wheel
Ross26 (2;9)	This goona will hurt me.
Ross35 (3;2)	Because if you catch him too big he will stung young.
Ross35 (3;2)	If you do this it won't came off but it you do this it will come off.
Ross54 (4;8)	And the first thing he did # is looked in my mouth
Ross54 (4;8)	Then he looked is [#] in my ears
Ross54 (4;8)	And Marky could have went to school with Mom
Ross64 (5;9)	it wasn't have friends with it
Sarah42 (3;1)	was she had no pants?
Sarah64 (3;6)	that will means
Sarah109 (4;5)	does a mouth (sup) posed to be right dere?
Sarah115 (4;7)	does she (sup)posed to go home at?
Sarah121 (4;8)	oh [#] you should saw it!
Sarah126 (4;10)	do you supposed to put this around shoulders?
Sarah128 (4;11)	he will wound it up.
Shem6 (2;4)	this one do riding
Shem6 (2;4)	and do sliding on d$ box
Shem9 (2;5)	the -(s) letter man is came and get letters
Shem16 (2;7)	dat be I bought
Shem18 (2;8)	this is got holes in it
Shem19 (2;8)	so you can not squished it up.
Shem20 (2;8)	I gonna cook spaghetti and be to pot it.
Shem34 (3;0)	he has $ sick.

Table 4.1
Transcripts Used in Age of First Use Analyses

Child	Corpus collected by	Age Range	Approx. # Aux*s*
Abe	Kuczaj (1976)	2;5 - 5;0	10,600
Adam	Brown (1973)	2;3 - 5;2	11,600
Allison	Bloom (1973)	1;4 - 2;10	200
April	Higginson (1985)	1;10 - 2;11	1,200
Eve	Brown (1973)	1;6 - 2;3	1,800
June	Higginson (1985)	1;3 - 1;9	200
Mark	MacWhinney & Snow (1985)	1;5 - 6;0	2,100
Naomi	Sachs (1983)	1;2 - 4;9	3,100
Nathan	MacWhinney & Snow (1985)	2;6-3;9	1,700
Nina	Suppes (1973)	2;0 - 3;3	7,500
Peter	Bloom (1973)	1;10 - 3;2	5,200
Ross	MacWhinney & Snow (1985)	2;10 - 7;10	11,700
Sarah	Brown (1973)	2;3 - 5;1	6,900
Shem	Clark (1978)	2;3 - 3;2	2,500
Total			66,300

* The numbers given are the number of lines which contained strings of letters that could be auxiliaries. In other words, the figures given include lines that contain main verbs which are homophonous with auxiliaries.

Table 4.2
Number of Missing Data per Child

Child	# Missing	% Missing
Abe	15	15%
Adam	22	21%
Allison	74	72%
April	61	59%
Eve	57	55%
June	94	91%
Mark	28	27%
Naomi	43	42%
Nathan	51	50%
Nina	34	33.%
Peter	46	45%
Ross	14	14%
Sarah	39	38%
Shem	40	39%
Total	618	43%
Total for best 12 children	440	36%

Table 4.3
Number of Missing Data per Sentence Modality

Modality	Aux's	Copula *Be*	Main Verbs	Total
Declarative	46 (19.%)	3 (5%)	0 (0%)	49 (13%)
Inverted	81 (34 %)	15 (25%)	--	96 (32%)
Verb+ *n't*	92 (43%)	11 (23%)	--	103 (39%)
Verb+ *not*	174 (73%)	28 (47%)	--	202 (67%)
TOTAL	393 (42%)	57 (25%)	0 (0%)	450 (36%)

Table 4.4
Number of Missing Data per Auxiliary and Copula Form

	# Missing	% Missing
Aux-Am	12	(33%)
Aux-Are	14	(29%)
Aux-Is	7	(15%)
Aux-Was	16	(33%)
AuxWere	25	(52%)
Did	9	(19%)
Do	7	(15%)
Does	12	(25%)
Had	41	(85%)
Has	34	(71%)
Have	22	(46%)
Can	3	(6%)
Could	16	(33%)
May	20	(56%)
Might	36	(75%)
Must	37	(77%)
Shall	26	(72%)
Should	21	(44%)
Will	7	(15%)
Would	16	(33%)
Copula-Am	14	(39%)
Copula-Are	6	(13%)
Copula-Is	2	(4%)
Copula-Was	10	(21%)
Copula-Were	25	(52%)

Table 5.1
Interaction Among *Wh*-placement, Inversion and Prosody

Matrix	+Inv	*In situ*	Prosody	Example:
+	+	+	High, stressed WH	Can he eat WHAT?
+	+	-	Rising & falling	What can he eat?
+	-	+	falling, stressed WH	He can eat WHAT?
+	-	-	------	*What he can eat?
-	+	+	------	*You know can he eat what?
-	+	-	------	*You know what can he eat?
-	-	+	Rising, stressed WH	You know he can eat WHAT?
-	-	-	Rise & fall, high WH	You know WHAT he can eat?

Table 5.2

Comparison of Published Studies on Subject/Auxiliary Inversion

Study	Method	Ages	-INV Stage	% Inverted	*Wh*-word effect	Aux effect
Bellugi (1971)	Longitudinal spontan.& elicit. production, N = 1	2;3 - 5;1	Yes for *Wh*	Y/N: 100% *Wh* : 0%	None noted	*-n't* worse
Derwing & Smyth (1988)	Elicited imitation & production, N = 24	3;0 - 4;6*	No	Y/N = *Wh*	None noted	*-n't* worse
Erreich (1984)	Elicited production N=18	2;5 - 3;10 (x = 2;8)	No	Y/N: 50% *Wh* : 64%	*Why, how,* & *when* worse	*-n't* worse
Ingram & Tyack (1979)	Cross-sectional, diary study, N = 21	2;0 - 3;11 (x = 2;11)	No	Young: *Wh* > Y/N Older: *Wh* = Y/N	None noted	None noted
Klee (1985)	Cross-sectional, diary study, N = 18	2;1 - 3;11 (x = 2;11)	No	Y/N = *Wh* Very few errors	*Why* worse for low MLU	Copula *be* worse
Kuczaj & Brannick (1979)	Elicit. imitation, N = 20 Judgment, N = 60	x = 4;0 x = 5;0	No	Only *wh*-questions studied	*What* & *where* best	None noted
Kuczaj & Maratsos (1975)	Elicited imitation from 1 non-inverting child	I: 2;5 - 2;7 II: 2;9	Yes for *Wh*	Y/N > *Wh*	None noted	None noted
Kuczaj & Maratsos (1983)	Longitudinal, N = 2 Cross-sectional, N = 14	2;5 - 4;0 1;11 - 3;0 2;6 - 5;6 *	No	Only Y/N questions studied	Only Y/N questions studied	*Could* & *would* worse
Labov & Labov (1978)	Diary of all spontaneous *wh*-questions, N = 1	2;2 - 4;8	No. Multi-factorial	Only *wh*-questions recorded	*Why* & *when* inverted later	*-n't, can,be* ** *should* bad

* Mean ages of subjects not provided.

** Labov & Labov (1978) do not distinguish between copula and auxiliary *be*.

Table 5.3
Number of Scoreable Questions per Child

Child	*Wh*-Questions	Yes/No Questions	Auxn't Questions	Total
Adam	642	879	228	1,749
Allison	7	31	0	38
April	24	17	13	54
Eve	67	39	2	108
June	0	0	0	0
Mark	282	377	56	715
May	0	0	0	0
Naomi	157	154	14	325
Nathan	118	106	17	241
Nina	337	410	21	768
Peter	126	269	8	403
Ross	439	530	104	1,073
Sarah	154	371	90	615
Shem	180	95	6	281
Total	2,533	3,278	559	6,370

Table 5.4
Overall Frequency of Inversion Errors

(Negated & Non-negated Aux Combined)

Child	All Inverted	All Scorable Sentences	%Correctly Inverted
Adam	1,574	1,749	90.0%
Allison	37	38	97.4%
April	51	54	94.4%
Eve	98	108	90.7%
Mark	691	715	96.6%
Naomi	303	325	93.2%
Nathan	122	241	50.6%
Nina	730	768	95.1%
Peter	384	403	95.3%
Ross	1,034	1,073	96.4%
Sarah	553	615	89.9%
Shem	249	281	88.6%
Total	5,826	6,370	91.5%

Table 5.5
Children's Inversion Rates for *Yes/No* and *Wh*-Questions

Child	WH	Y/N	Y/N > WH	Y/N = WH	Y/N < WH
Adam	88.3 %	96.6%	+	-	-
Allison	85.7%	100%	+	-	-
April	91.7%	94.1%	-	+	-
Eve	95.5%	87.2%	-	-	+
Mark	97.9%	97.6%	-	+	-
Naomi	96.2%	94.2%	-	+	-
Nathan	60.1%	46.2%	-	-	+
Nina	98.5%	93.9%	-	-	+
Peter	92.1%	98.5%	+	-	-
Ross	99.3%	97.0%	-	+	-
Sarah	92.9%	91.9%	-	+	-
Shem	95.6%	79.0%	-	-	+
MEAN	93.0%	93.7%	-	+	-

Table 5.6
Inversion Rates for Each Child and Each Question Type

	Adam	Allison	April	Eve	Mark	Naomi	Nathan	Nina	Peter	Ross	Sarah	Shem	TOTAL
HowInv	154	0	0	0	46	8	3	24	2	38	30	2	307
How+aux	159	0	0	0	48	8	3	24	2	38	32	2	316
%HowInv	96.86	#N/A	#N/A	#N/A	95.83	100.00	100.00	100.00	100.00	100.00	93.75	100.00	97.15
WhatInv	297	3	16	38	134	93	30	127	42	127	71	101	1079
What+aux	325	3	16	41	135	94	54	130	48	128	73	101	1148
%WhatInv	91.38	100.00	100.00	92.68	99.26	98.94	55.56	97.69	87.50	99.22	97.26	100.00	93.99
WhenInv	1	0	0	0	11	1	0	7	5	5	2	4	36
When+aux	2	0	2	0	14	2	0	7	5	5	5	5	47
%WhenInv	50.00	#N/A	0.00	#N/A	78.57	50.00	#N/A	100.00	100.00	100.00	40.00	80.00	76.60
WhereInv	64	1	4	26	26	31	24	100	51	48	27	43	445
Where+aux	74	1	4	26	26	35	24	100	54	48	30	47	469
%WhereInv	86.49	100.00	100.00	100.00	100.00	88.57	100.00	100.00	94.44	100.00	90.00	91.49	94.88
WhichInv	3	0	0	0	7	1	3	2	0	6	4	1	27
Which+aux	9	0	0	0	7	1	3	2	0	6	5	1	34
%WhichInv	33.33	#N/A	#N/A	#N/A	100.00	100.00	100.00	100.00	100.00	#N/A	100.00	80.00	100.00	79.41
WhoInv	22	0	2	0	5	3	4	14	13	6	9	7	85
Who+aux	22	0	2	0	5	3	4	14	13	6	9	7	85
%WhoInv	100.00	#N/A	100.00	#N/A	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
WhyInv	26	2	0	0	47	14	7	58	3	206	0	14	377
Why+aux	51	3	0	0	47	14	30	60	4	208	0	17	434
%WhyInv	50.98	66.67	#N/A	#N/A	100.00	100.00	23.33	96.67	75.00	99.04	#N/A	82.35	86.87
AllWhInv	567	6	22	64	276	151	71	332	116	436	143	172	2356
AllWh+aux	642	7	24	67	282	157	118	337	126	439	154	180	2533
%AllWhInv	88.32	85.71	91.67	95.52	97.87	96.18	60.17	98.52	92.06	99.32	92.86	95.56	93.01
Y/NINV	849	31	16	34	368	145	49	385	265	514	341	75	3072
Y/N+aux	879	31	17	39	377	154	106	410	269	530	371	95	3278
%Y/NINV	96.59	100.00	94.12	87.18	97.61	94.16	46.23	93.90	98.51	96.98	91.91	78.95	93.72
AllQuesInv	1416	37	38	98	644	296	120	717	381	950	484	247	5428
AllQues+au	1521	38	41	106	659	311	224	747	395	969	525	275	5811
%AllQuesIr	93.10	97.37	92.68	92.45	97.72	95.18	53.57	95.98	96.46	98.04	92.19	89.82	93.41

Table 5.7
Overall Inversion Rates for Each Auxiliary Type

AUX-TYPE	All Wh-questions	All Yes/No Questions
am	0.00	100.00
are	90.33	90.43
can	85.00	95.59
could	92.50	95.83
did	99.40	98.92
do	97.84	99.23
does	100.00	100.00
had	50.00	#N/A
has	#N/A	100.00
have	100.00	100.00
is	92.23	90.16
may	25.00	92.00
might	0.00	0.00
must	#N/A	#N/A
shall	94.44	100.00
should	82.48	87.50
was	74.95	66.67
were	43.67	72.73
will	66.51	93.15
would	85.00	96.15
All Be	84.81	88.73
All Do	98.30	99.28
All Have	50.00	100.00
All Modals	80.74	94.98
All	89.84	93.72

Table 5.8
Inversion Rates for Each Auxiliary in Each Question Type

Aux-Type	How	What	When	Where	Which	Who	Why	Overall Wh	YN %	Overall %
am									100.00	100.00
are	92.86	96.46	71.43	93.59	100.00	100.00	77.97	90.33	90.43	91.73
can	90.91	74.07	100.00	100.00			60.00	85.00	95.59	94.14
could	100.00	62.50	100.00	100.00		100.00		92.50	95.83	94.85
did	100.00	98.68	100.00	100.00	100.00	100.00	97.12	99.40	98.92	98.71
do	100.00	99.15	85.71	100.00	100.00	100.00	100.00	97.84	99.23	99.32
does	100.00	100.00		100.00	100.00	100.00	100.00	100.00	100.00	100.00
had				100.00			0.00	50.00		50.00
has									100.00	100.00
have				100.00				100.00	100.00	100.00
is	84.21	93.29	100.00	94.80	100.00	100.00	73.33	92.23	90.16	91.08
may		50.00		0.00				25.00	92.00	83.33
might							0.00	0.00	0.00	0.00
must										0.00
shall	100.00	83.33		100.00				94.44	100.00	97.22
should	100.00	84.62		77.78	50.00		100.00	82.48	87.50	87.23
was	100.00	94.12	22.22	100.00	50.00	100.00	58.33	74.95	66.67	76.33
were		73.33	50.00	75.00	0.00		20.00	43.67	72.73	61.54
will	88.89	72.22	100.00	0.00			71.43	66.51	93.15	90.59
would	100.00	100.00		100.00	25.00		100.00	85.00	96.15	94.59
All Be	88.57	93.78	67.74	94.38	76.92	100.00	72.29	84.81	88.73	89.88
All Do	100.00	99.36	90.00	100.00	100.00	100.00	98.73	98.30	99.28	99.30
All Have				100.00			0.00	50.00	100.00	90.91
All Modals	94.12	75.00	100.00	85.29	33.33	100.00	77.42	80.74	94.98	92.98
All	97.15	93.99	76.60	94.88	79.41	100.00	86.87	89.84	93.72	93.41

Table 5.9
Overall Inversion Rates for Negated and Non-Negated Auxiliaries

	AUX (Invert/Total)	AUX+*N'T* (Invert/Total)
BE	89.9% (2390/2659)	95.2% (119/125)
DO	99.3% (1691/1703)	70.1% (202/288)
HAVE	90.9% (10/11)	N/A (0/0)
MODAL	93.0% (1337/1438)	53.1% (75/145)
TOTAL	93.4% (5428/5811)	71.3% (398/558)

Table 5.10
Inverted Embedded Questions

Yes/no Questions

Ross93	um@h or if do you want to read it with Mark?

Wh- Questions

Who

Adam20	Smith's who are you.
Adam29	I don't know. I don't know [#] I don't know who is dat.
Adam36	I don't know who is she. I getting tired.
Mark74	no [#1] I don't know why who're they

What

Adam19	I don't know what are dey.
Adam24	I don't xxx what is dis?
Adam35	I don't know what is a finger toe.
Adam36	I better see what was this.
Adam36	What's that do you see on there?
Adam41	I don't know what is dat.
Adam41	I know what time is it.
Adam44	I wonder what are dese for?
Adam47	Look at what are dey builing?
Mark63	and he said what are you named?
Mark77	I don't know what is GIJoe [possible subject ?]
Naomi77	know what xxx was it upside down?
Nina37	no let me see who is that.
Nina55	"tell me what do you wanna@a do this morning?"
Peter15	know what is it
Peter17	know what is it Mom
Peter19	now know what is it now (five times)
Ross54	What I want to know [#] is how does the coolen pee get down there.
Ross48	Is that what's called it? [i.e. is that what it's called]
Ross41	Do you know what am I?
Ross57	I don't know what is his name?
Sarah98	Tell me what is it.
Shem17	Wha(t)'s - you tell him what is she doing with d$ s- t-cricter
Shem30a	you - you - you tell me what is it?

Table 5.10
Inverted Embedded Questions

Where	
Adam36	I know where de is de 'A'.
Peter18	yeah [#& know where is it
Peter18	yeah [#&] know where is it was right here
Peter18	know where'd is it?
Ross44	They don't say where is my Great Pumpkin book any more.
Shem32	yeah. no dat's where is it?
How	
Adam29	Look how good boy am I.
Adam42	I wonder how is he gonna come back in.
Adam52	This how shall we begin.
Adam52	I said how is Rinny doing?
Adam52	This how shall we begin. Shall we begin . . .
Adam54	Now I wonder how can you make # dose steps.
Sarah120	Dis is how (will) it go.
Why	
Mark78	I don't understand why is the grass poisoned up?
Mark87	I wonder why did you <give em/them@a> [//] get [...] [\]
Shem24	cause how she go faster fast. dat's why do cwock goes faster, faster.

Table 5.11
Inversion Rates for Matrix and Embedded *Wh*-Question

	Matrix Questions		Embedded Questions	
	%Inversion	Inversion Rank	%Inversion	Inversion Rank
How	97.2%	2	14.6%	2
What	94.0%	4	10.7%	3
When	76.6%	7	0.0%	6
Where	94.9%	3	5.2%	5
Which	79.4%	6	0.0%	6
Who	100%	1	66.7%	1
Why	86.9%	5	7.7%	4

Spearman's rank correlation coefficient, rho = .875, $p < .01$

Table 5.12
Practice Questions Used in Judgment Experiments

a. Will the cat sing?	(Good)
b. Who onion tree?	(Bad)
c. House the cat touches?	(Bad)
d. Are the cows brown	(Good)
e. Will the cow sleep?	(Good)
f. Where ball happy?	(Bad)
g. Are you good?	(Good)
h. Runs he fast?	(Bad)
i. Which apples is red?	(Good)
j. When water the blue?	(Bad)

Table 6.1
Transcripts Used in Tensing Analyses

Child	Transcripts collected by	Age Range	# Questions
Adam	Brown (1973)	2;3 - 5;2	~10,650
Allison	Bloom (1973)	1;4 - 2;10	~50
April	Higginson (1985)	1;10 - 2;11	~400
Eve	Brown (1973)	1;6 - 2;3	~1,700
June	Higginson (1985)	1;3 - 1;9	~150
Mark	MacWhinney & Snow (1985)	1;5 - 6;0	~3,600
May	Higginson (1985)	0;11-0;11	~0
Naomi	Sachs (1983)	1;2 - 4;9	~2,150
Nathan	MacWhinney & Snow (1985)	2;6-3;9	~2,700
Nina	Suppes (1973)	2;0 - 3;3	~3,050
Peter	Bloom (1973)	1;10 - 3;2	~3,250
Ross	MacWhinney & Snow (1985)	2;10 - 7;10	~5,150
Sarah	Brown (1973)	2;3 - 5;1	~5,350
Shem	Clark (1978)	2;3 - 3;2	~2,400
Total			~40,600

Table 6.2
Double-Tensed Questions with More than One Auxiliary

Adam12	yeah. <do> [/] <do> do be stirs?
Adam17	d(o) you want me [#] should be careful?
Adam24	Where dere's a heel? It's cant fit with dis.
Adam26	it's [?] doesn't write.
Adam27	did I didn't mean to?
Adam27	no [#] it's doesn't go.
Adam28	is it was a snake [#] huh?
Adam28	Why was I did break it?
Adam29	can dey should go down dat Massachusetts avenue?
Adam29	Paul go to be a boy. can [#] is Paul a boy?
Adam29	does it be around it?
Adam31	did I didn't mean to?
Adam32	do she don't need that one?
Adam33	did I didn't mean to do that?
Adam34	did was it be a comb?
Adam34	it's will knock Paul?
Adam35	I almost caught myself. see? it's will hold it.
Adam37	what (was) dis one is about?
Adam39	I don't see nobody. did there be some. did it be
Adam39	and does dese be in dere?
Adam39	like dis. it's could go in de tunnel like dat. could it [#] huh?
Adam41	it's doesn't talk
Adam42	it's won't hurt.
Adam45	it's is real [?] [#]
Adam50	whats is dis? (repeated 4 times)
April4	can she's nice?
April4	isn't [#] is Snow White down here?
Eve6	Where di Eve (must) ... find it
Mark68	could you do untie the knot
Mark71	oh [#] is it doesn't xxx everything xxx there?
Mark78	oh wait. would I may be excused?
Naomi76	those are [#] how could writed? I writed the same ....
Nathan15	why you don't don't want it in the butter?
Nathan16	<eh ba ba-i don't don't know the rest?> [overlap]
Nathan22	why don't he can't he get his supper on the rock?
Nathan24	when's is dis glove?
Nathan26	why is not wathcing <where> he's goin'?

Table 6.2
Double-Tensed Questions with More than One Auxiliary

Nathan26	why is he will stop?
Nathan26	why is [#] is the stove is on too high?
Nina33	is this is a dog?
Nina34	does it doesn't move?
Nina34	do they don't eat people up?
Nina34	are she is she gonna@a put on her hair?
Nina36	what can [...] this baby one doesn't stand up.
Nina36	what were those are?
Nina43	is that house that's on fire?
Nina45	why do [#] did you get it in there?
Nina50	why did you did scare me?
Peter19	Mommy [#&] is a can we have ahook on the tractor
Ross32	is my old baby blanket is clean?
Ross36	why does Superman doesn't wear underroos on his bottom?
Ross41	But why aren't the good people have guns are not afraid of them
Ross45	She wouldn't# didn't go trick-or-treating with us?
Sarah42	was she had no pants?
Sarah59	it was get burned on my sore finger?
Sarah70	I <didn't did> [//] [#] caught it [#]
Sarah89	this [#] are you can make an elephant? huh?
Sarah111	Mommy, what ('s) ... is three mean? Three, three .
Sarah120	what does thats begin with?
Sarah121	did I <be> [/] [#] be a good girl?
Shem2	where does these pictures be taken?
Shem3	why is tony is clossing $ door?
Shem3	wha(t) do we're going $ record?
Shem8	you can- can do it?
Shem10	is the clock is working?
Shem10	ann' why on $- an' it's- is it floating in $ water?
Shem17	wha(t)'s is dis fing?
Shem17	an' what is he's doing>
Shem21a	what is he's tired for?
Shem21a	what i' he's gonna take away?
Shem21b	where is dey're inside d$ house?
Shem28b	and what is he's doing??
Shem28b	what is he - dat s an - $fox
Shem28b	and what di' he- why- what did he's showing?
Shem35	i had a softball an' i didn't know where it is - was.
Shem36	$small- why did- is dat m- mine?
Shem40	wha' wh- why is- wh- why is des clothes are blowing by?

Table 6.3
Double-Tensed Questions with One Auxiliary

Adam7	what did you did?
Adam9	What kitty do went (ing) place? Read dat
Adam16	What me write? Ct: Writing with pen] # It work # Iss works #
Adam19	Why it (s) came off? Why its came off?
Adam19	Dose, (are) ball. Why iss came
Adam19	I can let it [?] spilled.
Adam22	does it opens?
Adam22	dat one opens. does it opens?
Adam24	it's breaks.
Adam26	two will [?] works [?].
Adam27	xxx. see [#] it's opens.
Adam28	did he took his pencil sharpener?
Adam28	does it rolls? xxx.
Adam28	Adam said: # When it go outside its moves?
Adam29	do you got some more?
Adam29	oh [#] did I caught it?
Adam30	does his nose goes dere?
Adam30	does monkey climbs on it? SVA
Adam30	did you broke that part?
Adam30	does President Kennedy shot a new
Adam31	dey have accident. did the eggs broke?
Adam31	did you made a mistake?
Adam33	it's fell so hard.
Adam33	did someone [?] almost took it off?
Adam33	did we went to somebody's house?
Adam33	Why d'you left them in?
Adam33	Why its flies all by itself?
Adam34	How could she made in Italy?
Adam34	Its popped. Its looks like a popper who pop
Adam34	What d'you use fishing rods for do?
Adam35	Its pulls it. What goes first?
Adam35	Dat's does goes like dis, see?
Adam35	Why dey not big enough? xxx watch. Itc comes apart.
Adam35	it's keeps. did [?] the street light says go or stop?
Adam36	did somebody broke it?
Adam36	How does this broke? How did it break?
Adam36	Ok # I like # Its makes like a (sword)
Adam36	What movie did I saw?

Table 6.3
Double-Tensed Questions with One Auxiliary

Adam37	I can. it's comes off.
Adam37	Its hurts.
Adam37	it's runs away. it's a bug. does animals like bugs? SVA
Adam37.	no. drop. it's breaks
Adam39	does dis house <has> [/] [#] have a full house?
Adam39	could you being some arrows? [being=bring?]
Adam39	Why you told me to back up? Its keeps falling off.
Adam39	Paul [#] did I shot you down?
Adam40	did you moved [?] (same as one below?)
Adam40	did you moved (d) dat?
Adam40	can my truck crash? it's turns.
Adam42	you can make a fish caught?
Adam42	See its makes some more colors.
Adam42	Did I missed it?
Adam43	Its stopped OK
Adam45	what did I almost ran in?
Adam47	it's writes [...] dis doesn't come out?
Adam49	Did you heard me?
April4	what does he dos?
Eve17	what did you does?
Eve17	and wha what did you doed?
Naomi93	Mommy did I used to have a [#]bottle [#] didn't I?
Nina35	did you got a big piece?
Nina36	did we saw a lion at the zoo?
Nina36	where is the other duck went daddy duck?
Nina39	did we [#] did we [#] did we left them at Maggie's house?
Nina40	is the doctor gave that to me?
Nina46	what did jimmy gave them? (twice)
Mark84	how does rain [#2] get maded?
Peter18	did the bell ringed?
Ross2125	Did you did that thing? Mark's toy bag.
Ross26	Did you bought me my motorcycle?
Ross27	Did Grandpa bought this for me?
Ross32	Do you got the right money?
Ross37	And why did you thought I wanted it?
Ross40	And why did they sucked him up?
Ross40	So Hans got Princess Lela and ran away. Why did they ran
Ross40	But why did Obi Wan Kenobi came?

Table 6.3
Double-Tensed Questions with One Auxiliary

Ross41	Why did all those Germans ran
Ross46	Would it helped if I got out and pushed? (-have)
Ross49	And they used to fitted you? . tensedinfitival
Ross49	Did you said a bad man? You did.
Ross51	Mom [#] does this looked fixed?
Ross51	Did you won today?
Ross54	Mom [#] have you ever saw an outlong circle like this?
Ross71	did we used to gargle?
Sarah59	it was get burned on my sore finger?
Sarah65	did I lost one?
Sarah69	did you bought this?
Sarah80	did the chief said to go on the fire engine?
Sarah101	did you [#] did you heard of a fallin(g) star?
Sarah104	did you saw the orange?
Sarah104	did you got it?
Sarah111	I had green [#] too? did I had green?
Sarah112	did I did five in there?
Sarah113	did I lost it?
?Sarah114	did <my breath> [/] my breath went in there?
Sarah119	did I said that the same time?
Sarah122	did he said that?
Sarah122	what did he said?
Sarah124	What color camed out? Red.
Sarah127	How did you caught him?
Sarah137	What d'you did that for?
Sarah123	did I did six?
Sarah130	did you saw him before?
Sarah132	did you forgot?
Sarah134	did you took this off?
Sarah136	did you went to [#] nursery school?
Sarah137	does this one works right?
Sarah138	don't ya heard of that?
Shem19	can i said swam?
Shem20	When did you bought
Shem20	no, i(t) would go - where did it goes?
Shem24	what did da' did?
Shem28b	why did it- why did it bwoke dere?

Table 6.4

Double-Tensed Questions with Irregular Verbs ("good" cases)

Adam7	what did you did?
Adam28	did he took his pencil sharpener?
Adam29	oh [#] did I caught it?
Adam30	did you broke that part?
Adam31	dey have accident. did the eggs broke?
Adam31	did you made a mistake?
Adam33	did someone [?] almost took it off?
Adam33	did we went to somebody's house?
Adam36	did somebody broke it?
Adam36	What movie did I saw?
Adam39	Paul [#] did I shot you down?
Adam45	what did I almost ran in?
Adam49	Did you heard me?
Nina35	did you got a big piece?
Nina36	did we saw a lion at the zoo?
*Nina39	did we [#] did we [#] did we left them at Maggie's house?
Nina46	what did jimmy gave them? (twice)
Ross2125	Did you did that thing? Mark's toy bag.
Ross26	Did you bought me my motorcycle?
Ross27	Did Grandpa bought this for me?
Ross32	Do you got the right money?
Ross37	And why did you thought I wanted it?
Ross40	So Hans got Princess Lela and ran away. Why did they ran
Ross40	But why did Obi Wan Kenobi came?
Ross41	Why did all those Germans ran
Ross49	Did you said a bad man? You did.
Ross51	Did you won today?
Sarah65	did I lost one?
Sarah69	did you bought this?
Sarah80	did the chief said to go on the fire engine?
*Sarah101	did you [#] did you heard of a fallin(g) star?
Sarah104	did you saw the orange?
Sarah104	did you got it?
Sarah111	I had green [#] too? did I had green?
Sarah112	did I did five in there?
Sarah113	did I lost it?
*Sarah114	did <my breath> [/] my breath went in there?
Sarah119	did I said that the same time?
Sarah122	did he said that?

Table 6.4
Double-Tensed Questions with Irregular Past Tense Verbs

Sarah122	what did he said?
Sarah127	How did you caught him?
Sarah123	did I did six?
Sarah130	did you saw him before?
Sarah132	did you forgot?
Sarah134	did you took this off?
Sarah136	did you went to [#] nursery school?
Shem20	When did you bought
Shem24	what did da' did?
Shem28b	why did it- why did it bwoke dere?

Table 6.5
"Good" Examples of Double-Tensed Questions

Examples involving *do*-support

Adam22 (3;1)	does it opens?
Adam22 (3;1)	dat one opens. does it opens?
Adam28 (3;4)	does it rolls? xxx.
Adam30 (3;5)	does his nose goes dere?
Adam30 (3;5)	does monkey climbs on it?
Adam35 (3;8)	Dat's does goes like dis, see?
Adam40 (3;11)	did you moved [?] (same as one below?)
Adam40 (3;11)	did you moved (d) dat? 2xtense
Adam42 (4;0)	Did I missed it?
April4 (2;10)	what does he dos?
Eve17 (2;2)	and wha what did you doed?
Nina50 (3;2)	why did you did scare me?
Peter18 (2;9)	did the bell ringed?
Ross40 (3;5)	And why did they sucked him up?
Sarah137 ((5;1)	does this one works right?

Examples involving other auxiliaries:

Nathan26 (3;5)	why is [#] is the stove is on too high?
Nina33 (2;10)	is this is a dog?
Ross32 (3;0)	is my old baby blanket is clean?
Shem3 (2;3)	why is tony is clossing $ door?
Shem10 (2;5)	is the clock is working?

Table 6.6
Percent of Auxiliary-less Questions with Tensed Main Verbs

Child	Y/N	Y/N (old*)	Wh	Wh (old*)
Adam	44%	~85%	60%	~80%
Allison	n/a	n/a	n/a	n/a
April	50%	~70%	94%	~95%
Eve	36%	~50%	35%	~50%
June	n/a	n/a	n/a	n/a
Mark	95%	~95%	89%	~90%
Naomi	26%	~75%	67%	~85%
Nathan	65%	~75%	64%	~75%
Nina	43%	~90%	94%	~95%
Peter	57%	~60%	65%	~90%
Ross	67%	~90%	90%	~90%
Sarah	65%	~80%	66%	~85%
Shem	81%	~85%	31%	~85%
Mean	55%	~75%	69%	~85%

* Approximate percentages from age at which child demonstrates knowledge of verbal inflections.

Figure 4.1
Original, Unreplaced Data: Interaction between Auxiliary/Main Verb and Be/Do/Have

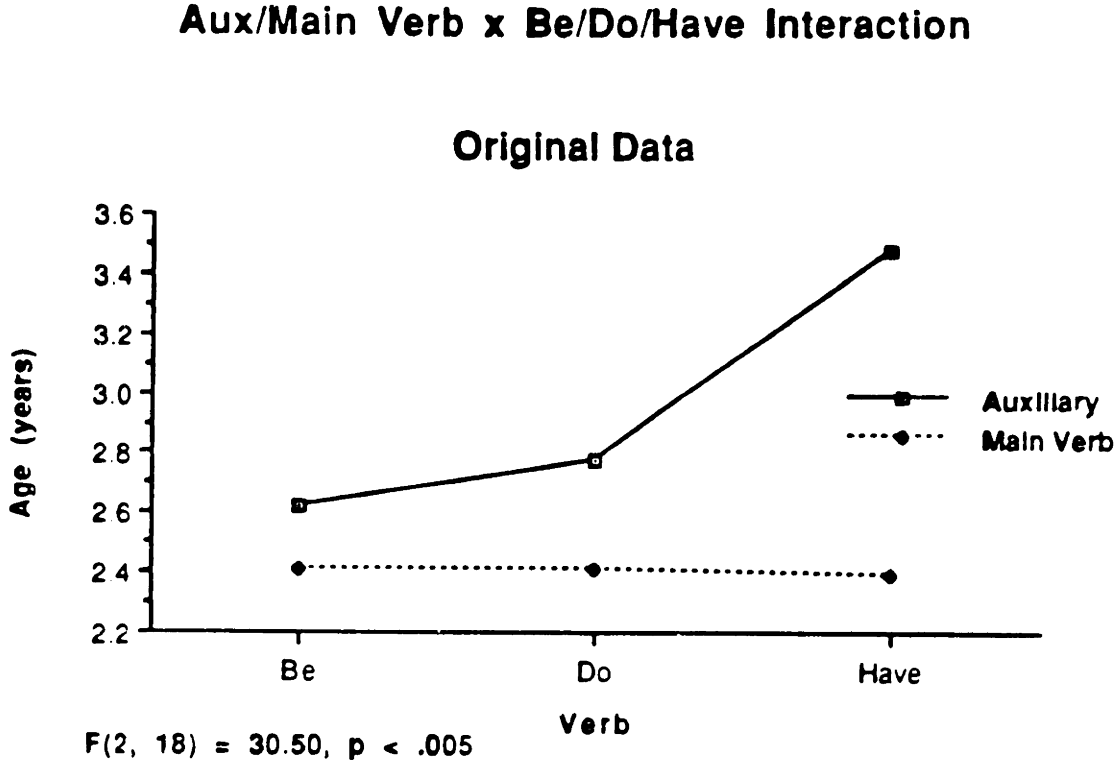


Figure 4.2
Replaced Data: Interaction between Auxiliary/Main Verb and Be/Do/Have

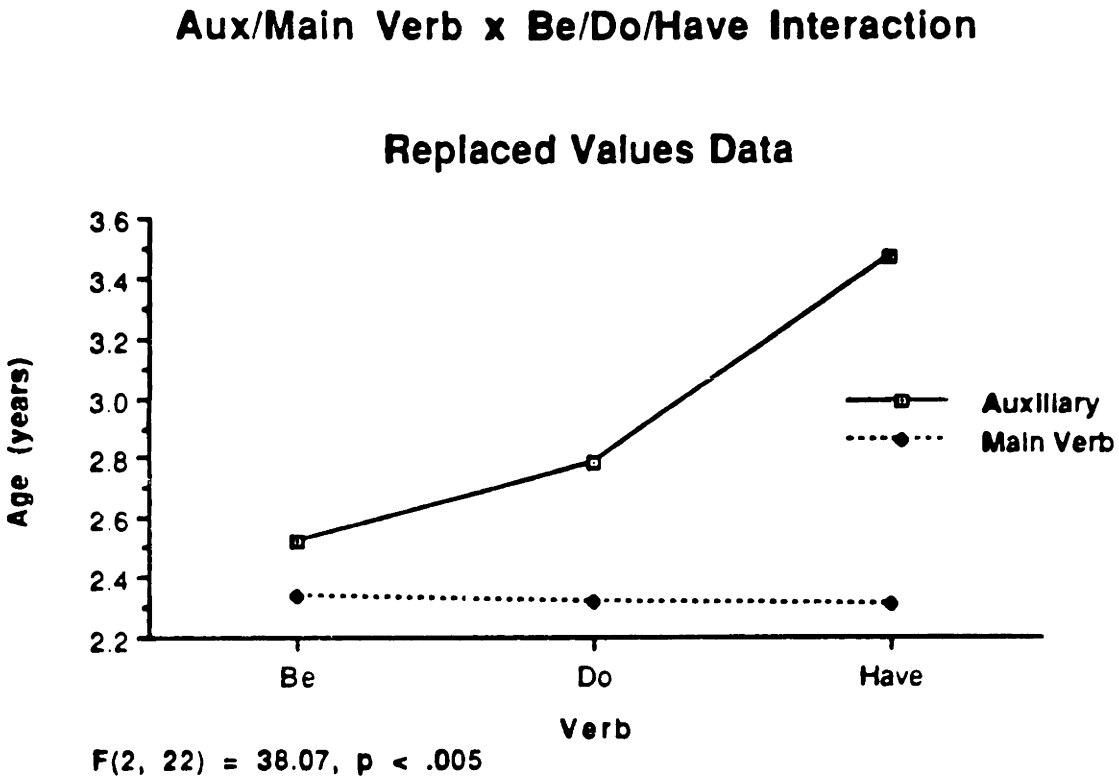
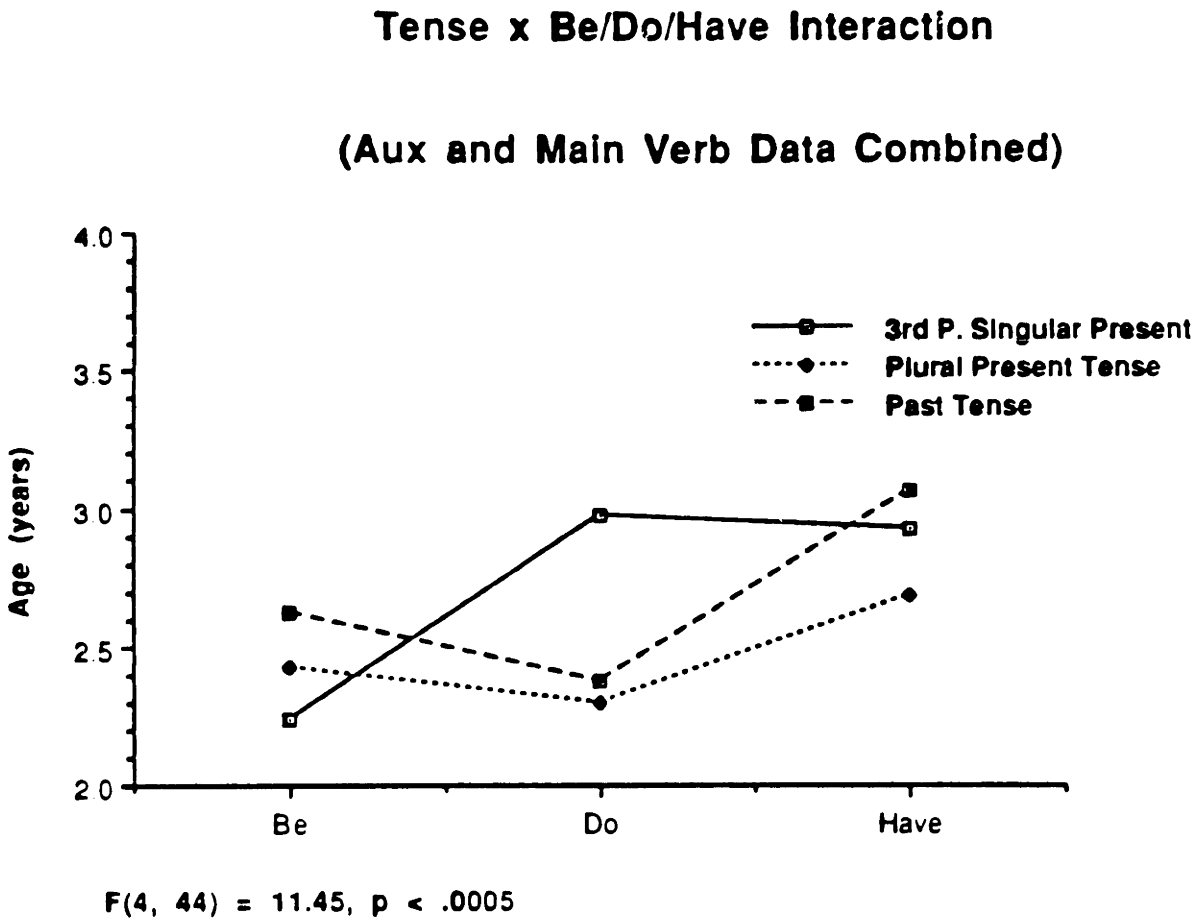
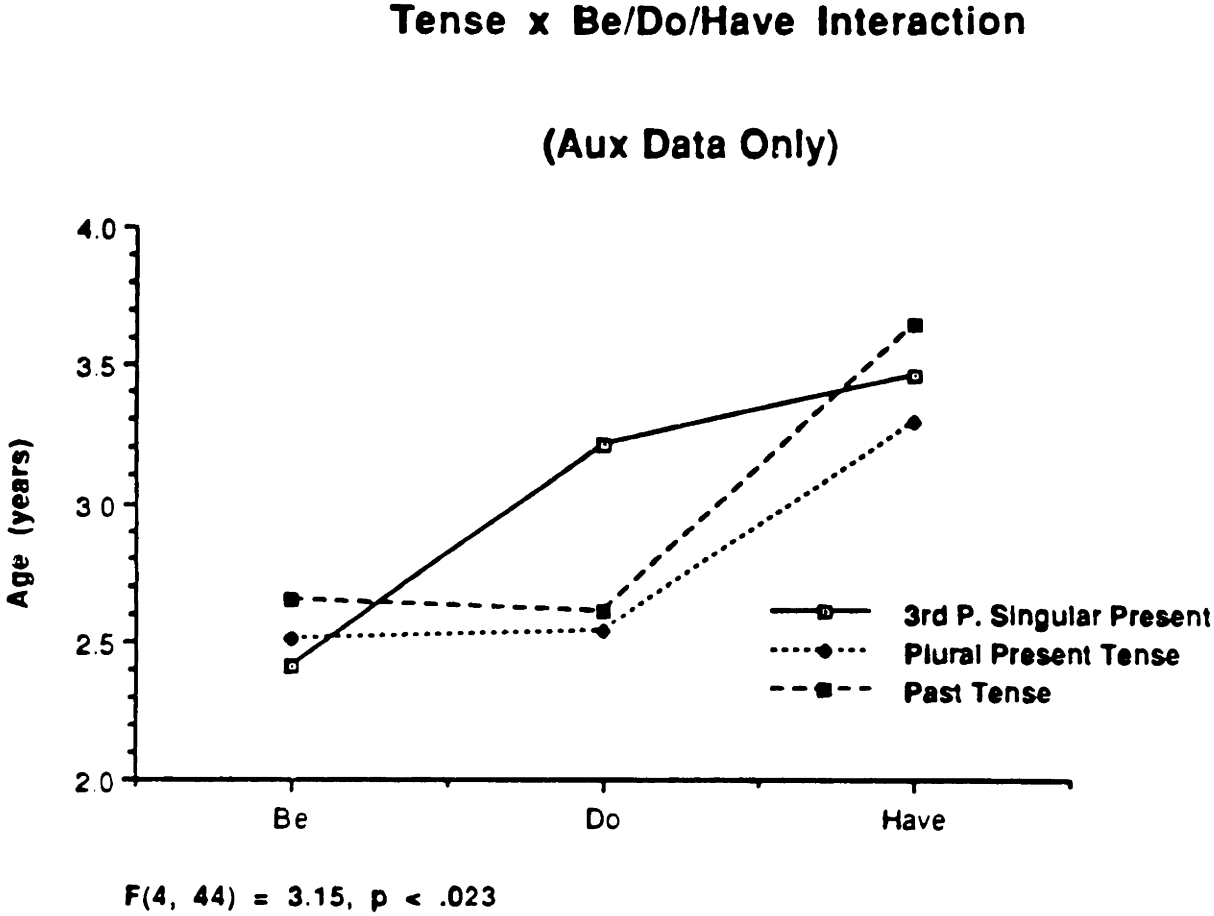


Figure 4.3
Interaction between Tense and Be/Do/Have (All Data)



F(4, 44) = 11.45, p < .0005

Figure 4.4
Interaction between Tense and Be/Do/Have (Auxiliary-Only Data)



$F(4, 44) = 3.15, p < .023$

Figure 4.5
Interaction between Tense and Be/Do/Have (Main Verb-Only Data)

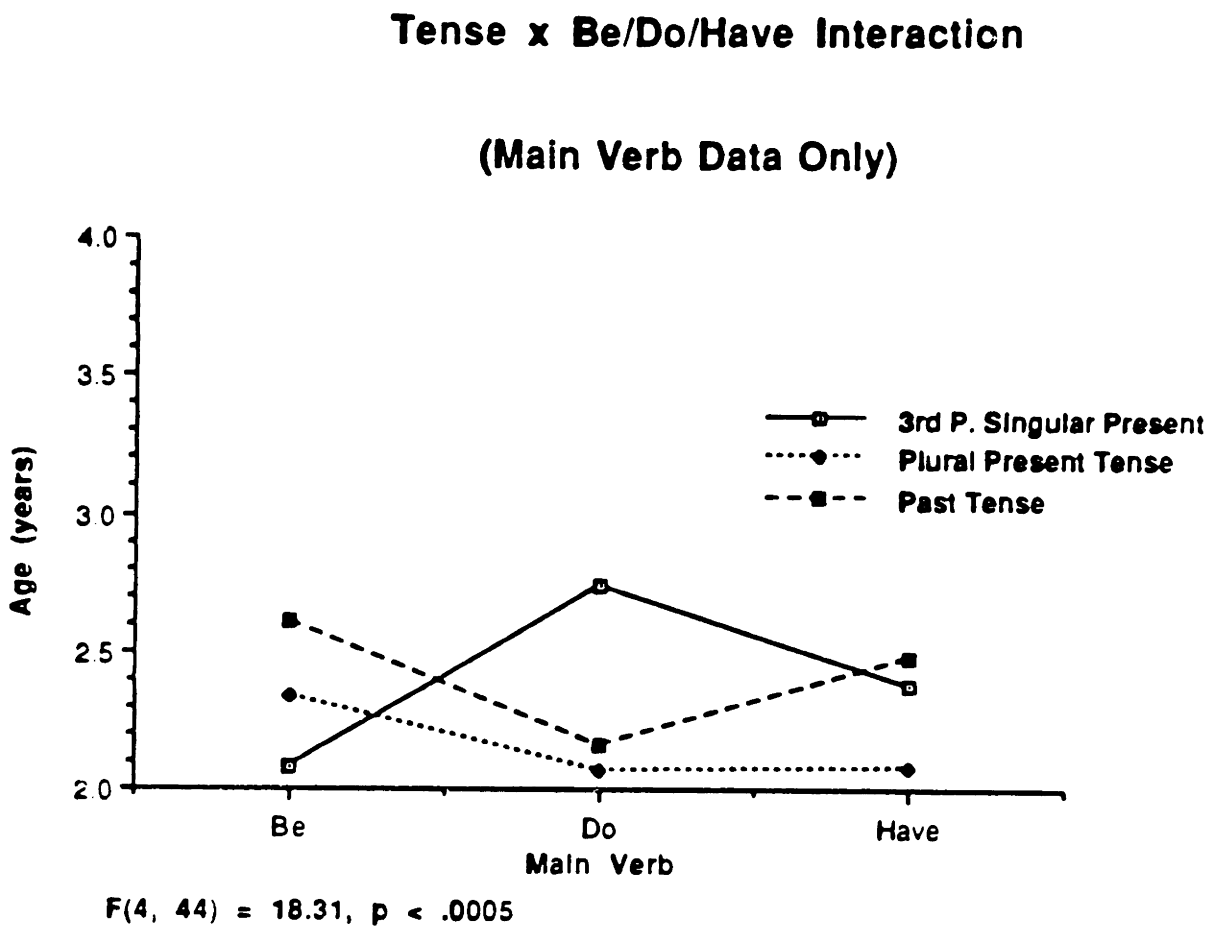


Figure 4.6
Interaction between Be/Do/Have/Modal and Sentence Type

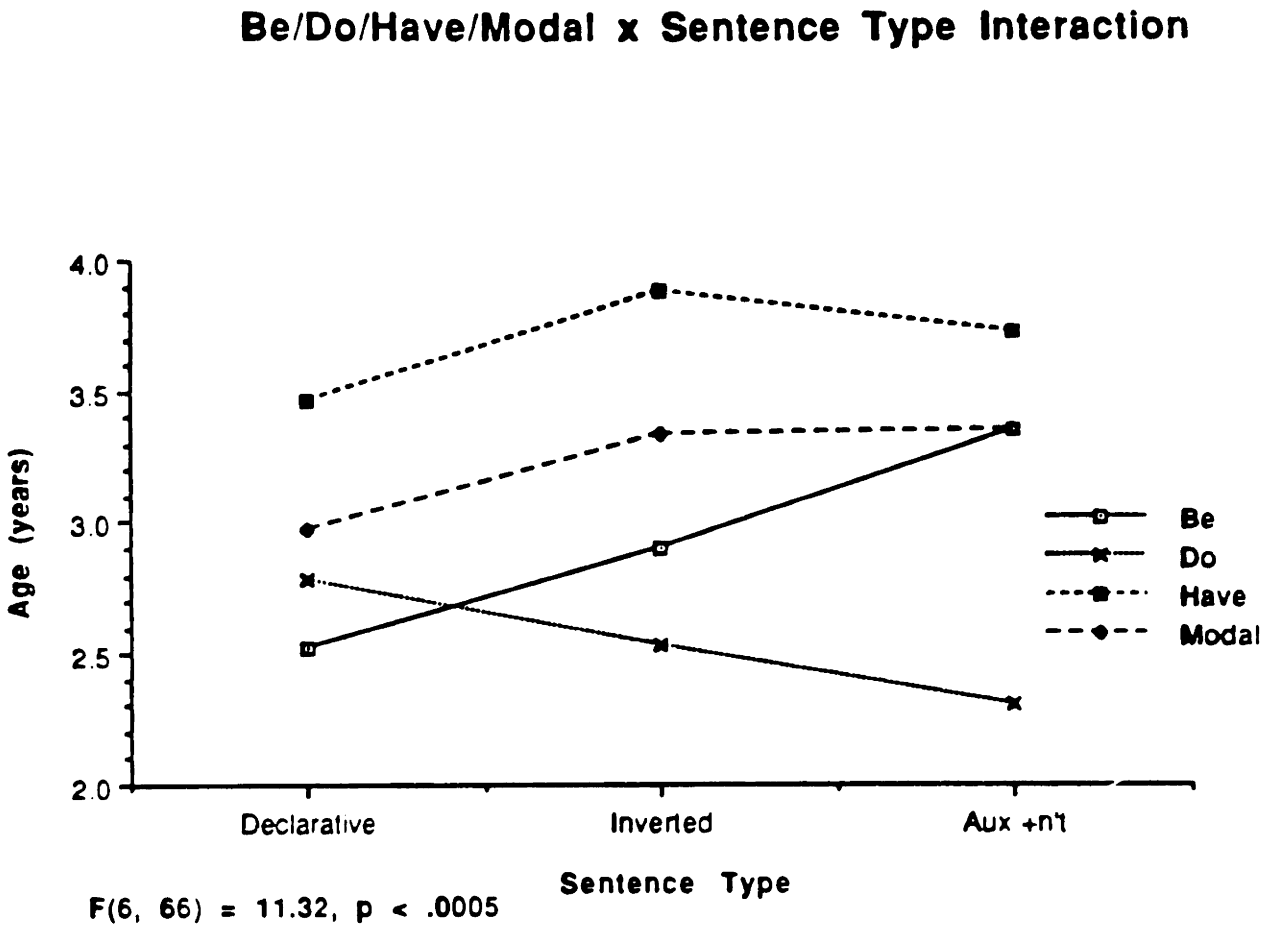
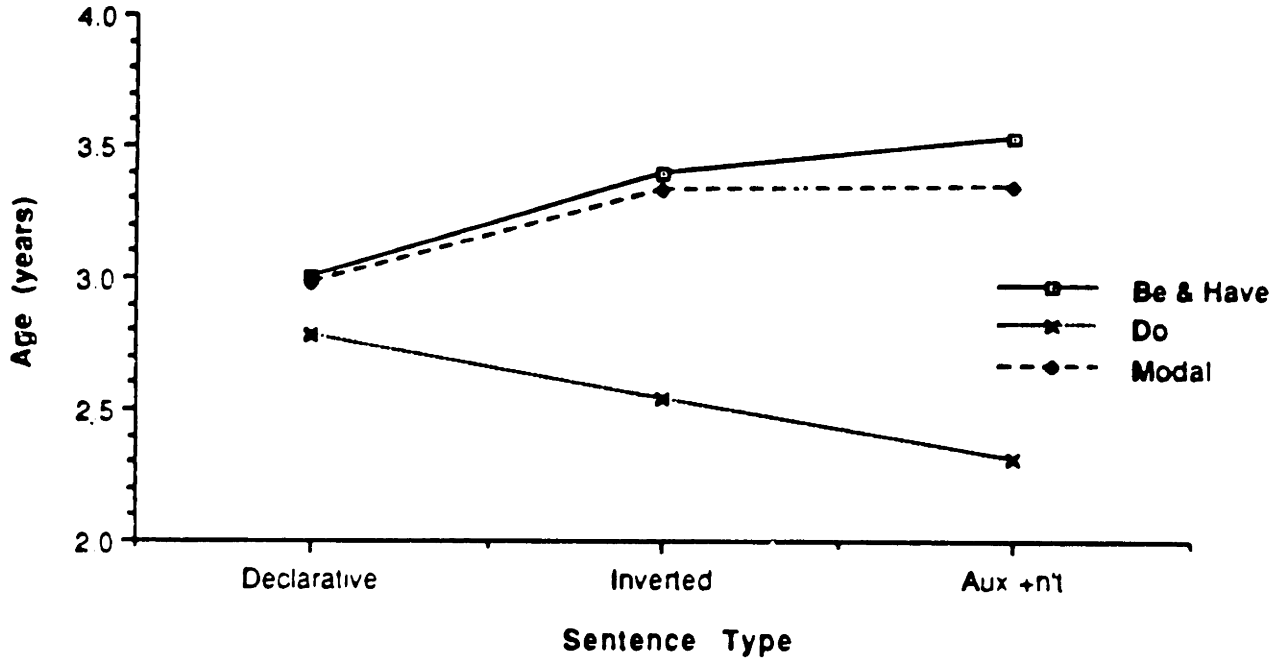


Figure 4.7
Interaction between Be&Have/Do/Modal and Sentence Type



F(4, 44) = 15.33, p < .0005

Figure 4.8
Interaction between Modal/NonModal and Sentence Type

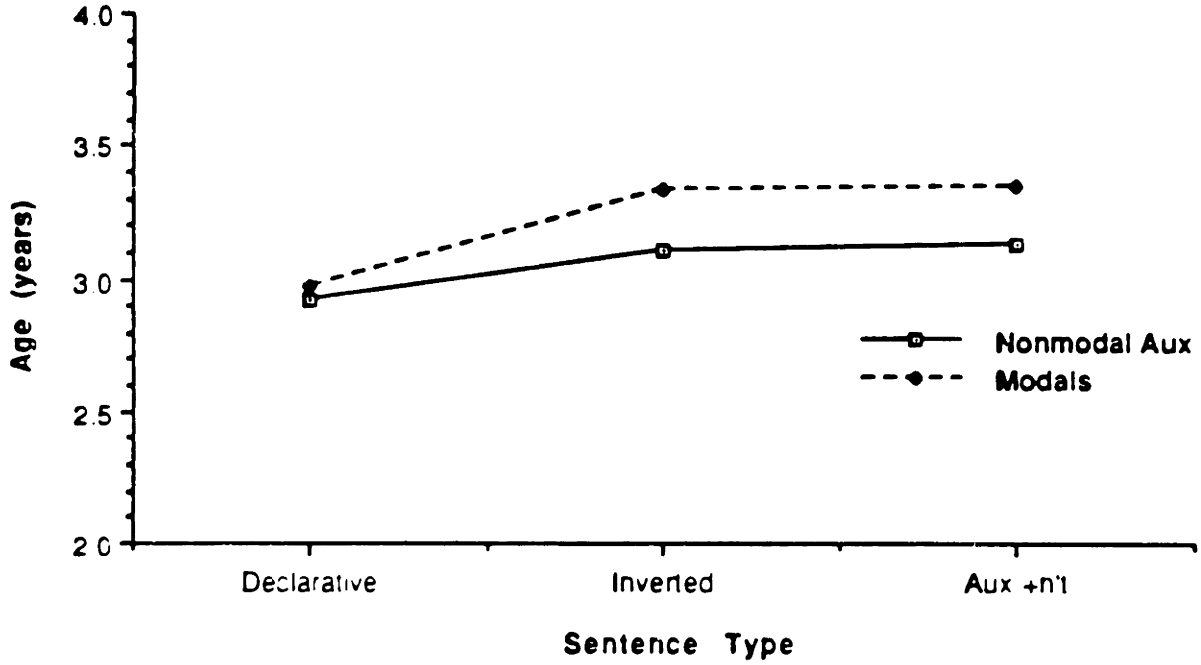


Figure 4.9
Interaction between Auxiliary/Copula *Be* and Tense

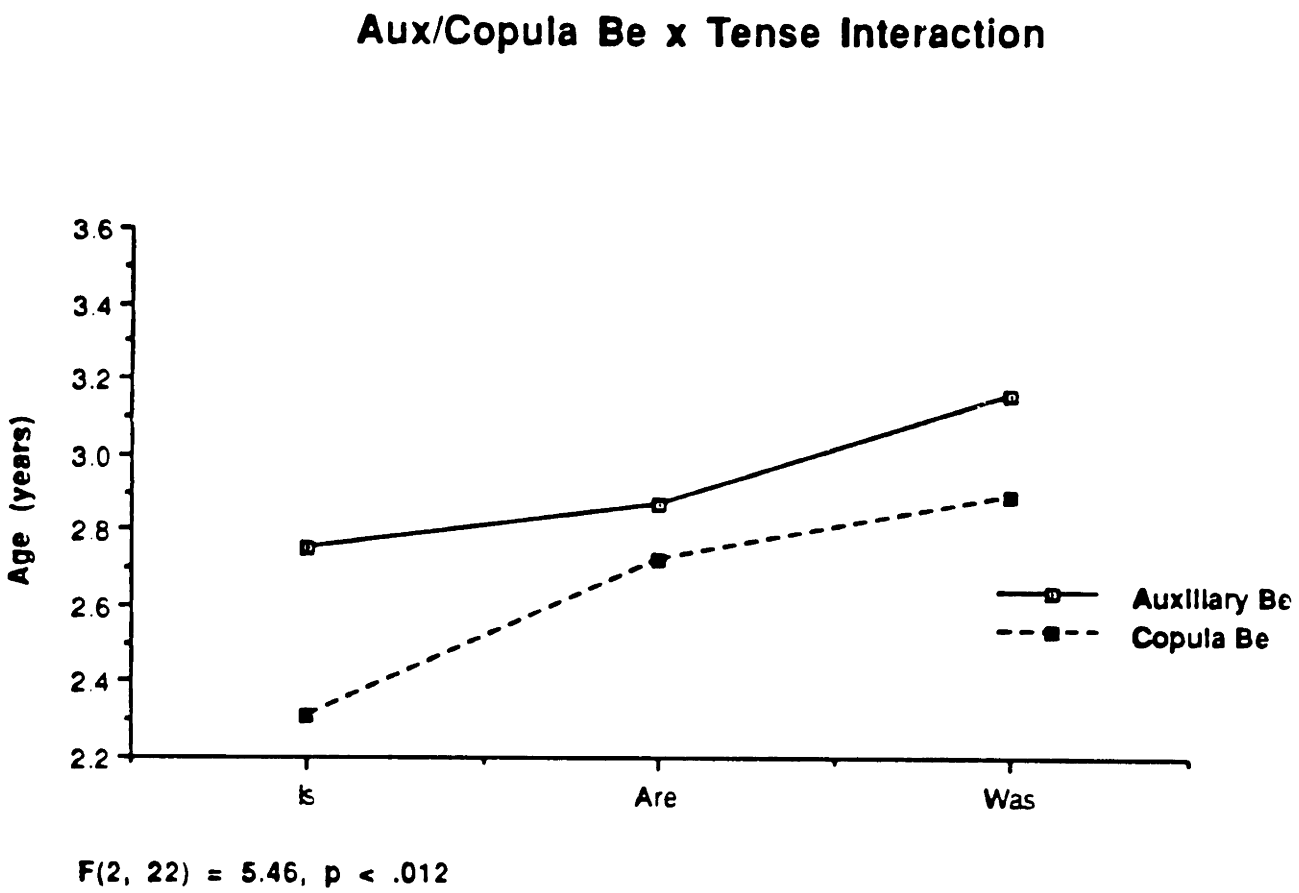
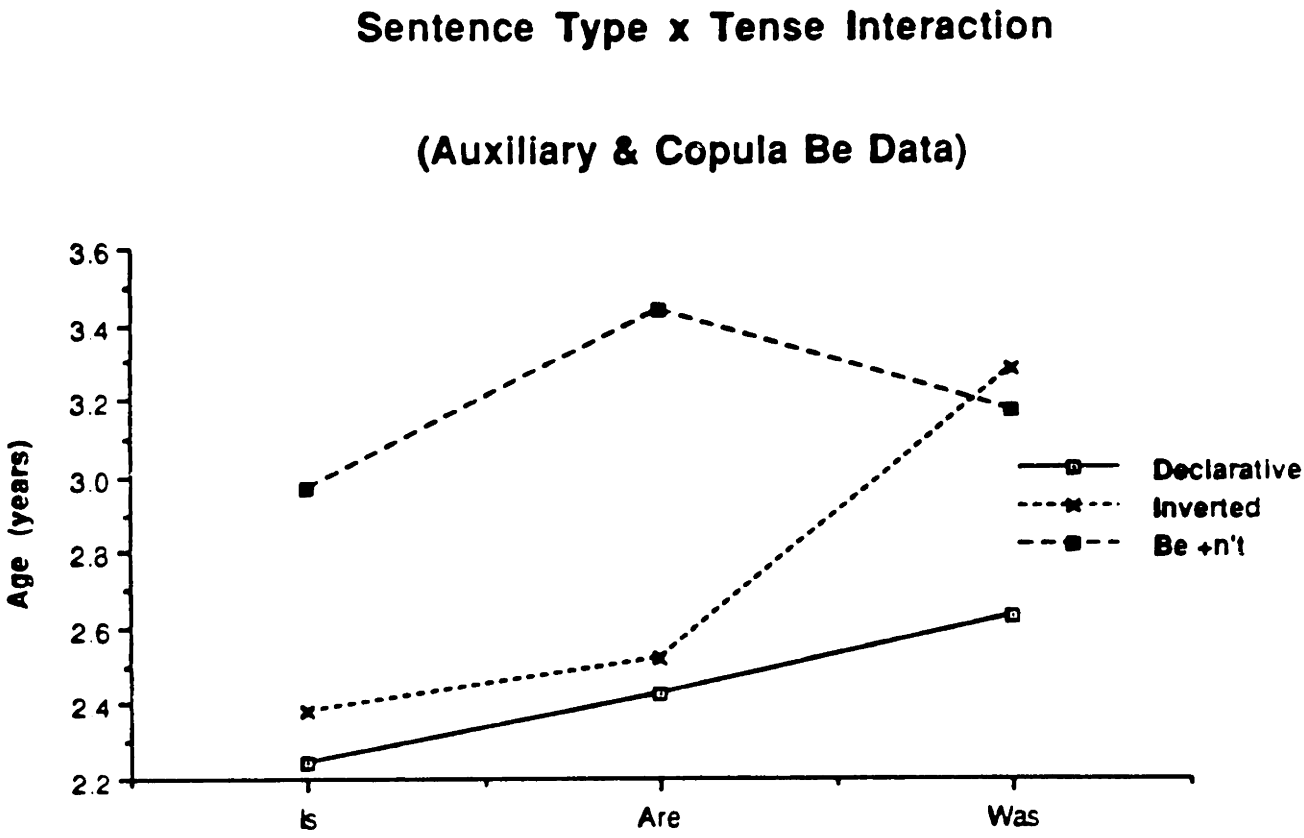


Figure 4.10
Interaction between Sentence Type and Tense (Aux & Copula Data)



$F(4, 44) = 9.47, p < .0005$

Figure 4.11
Interaction between Sentence Type and Tense (Copula-Only Data)

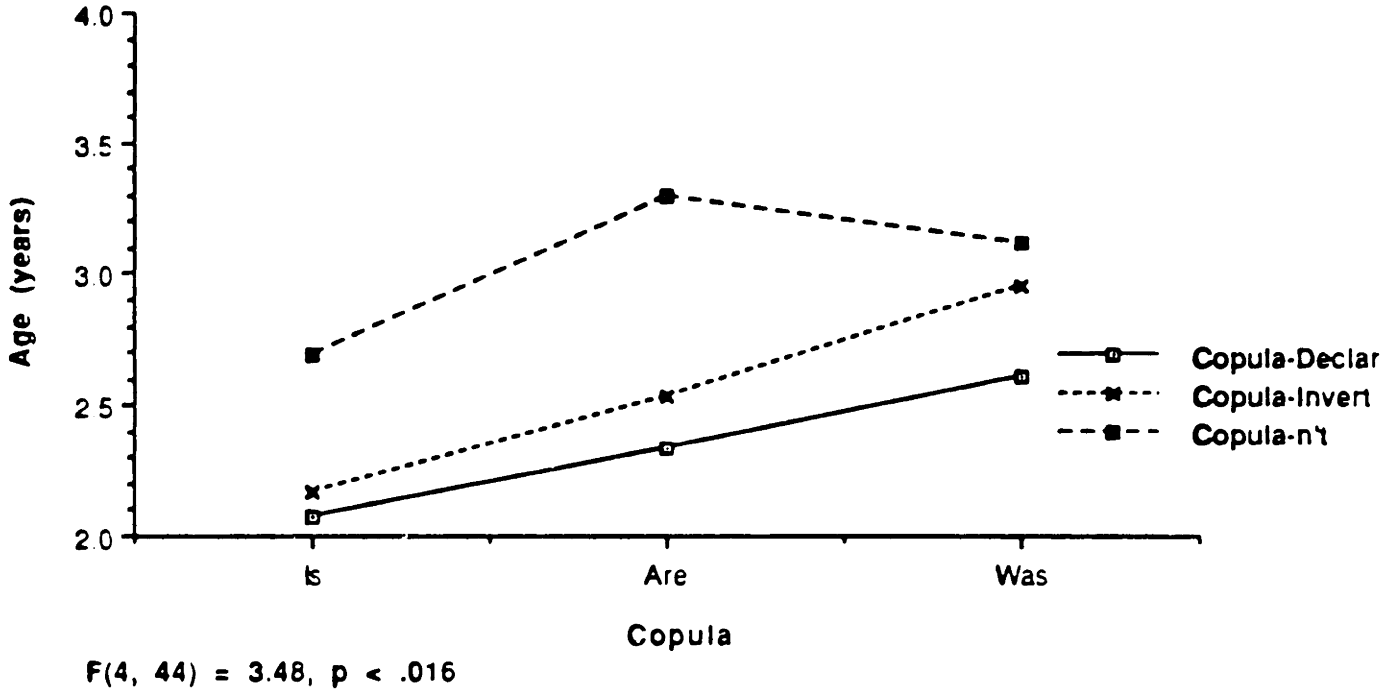


Figure 4.12
Interaction between Sentence Type and Tense (Auxiliary-Only Data)

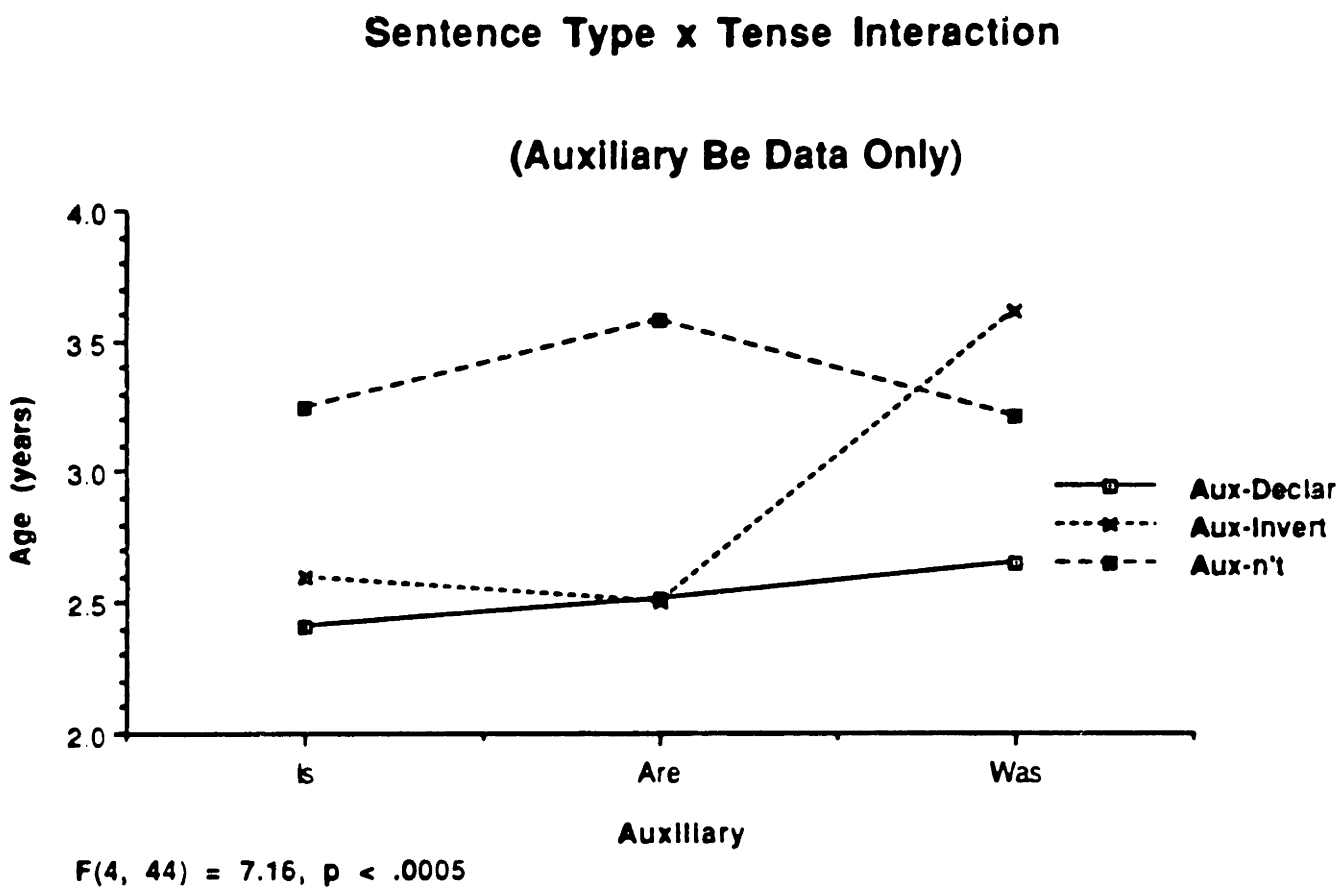


Figure 5.1
Pollock/Chomsky V-Raising Account of Inversion

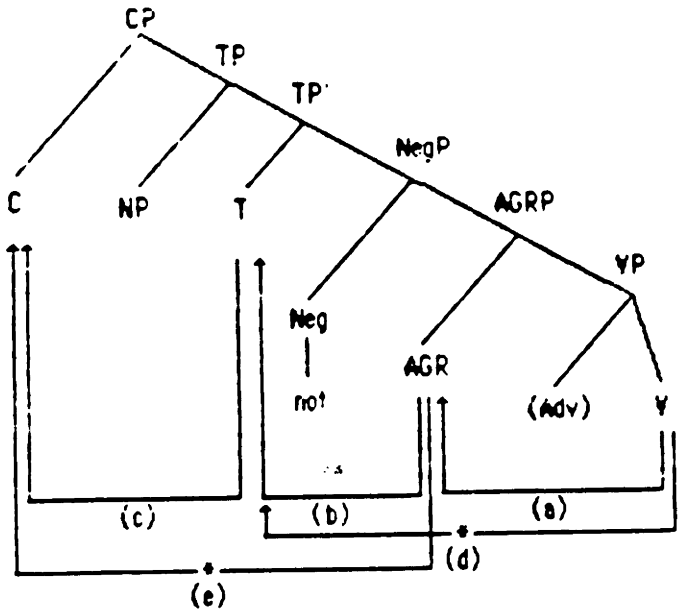


Figure 5.2
Developmental Trend: Adam

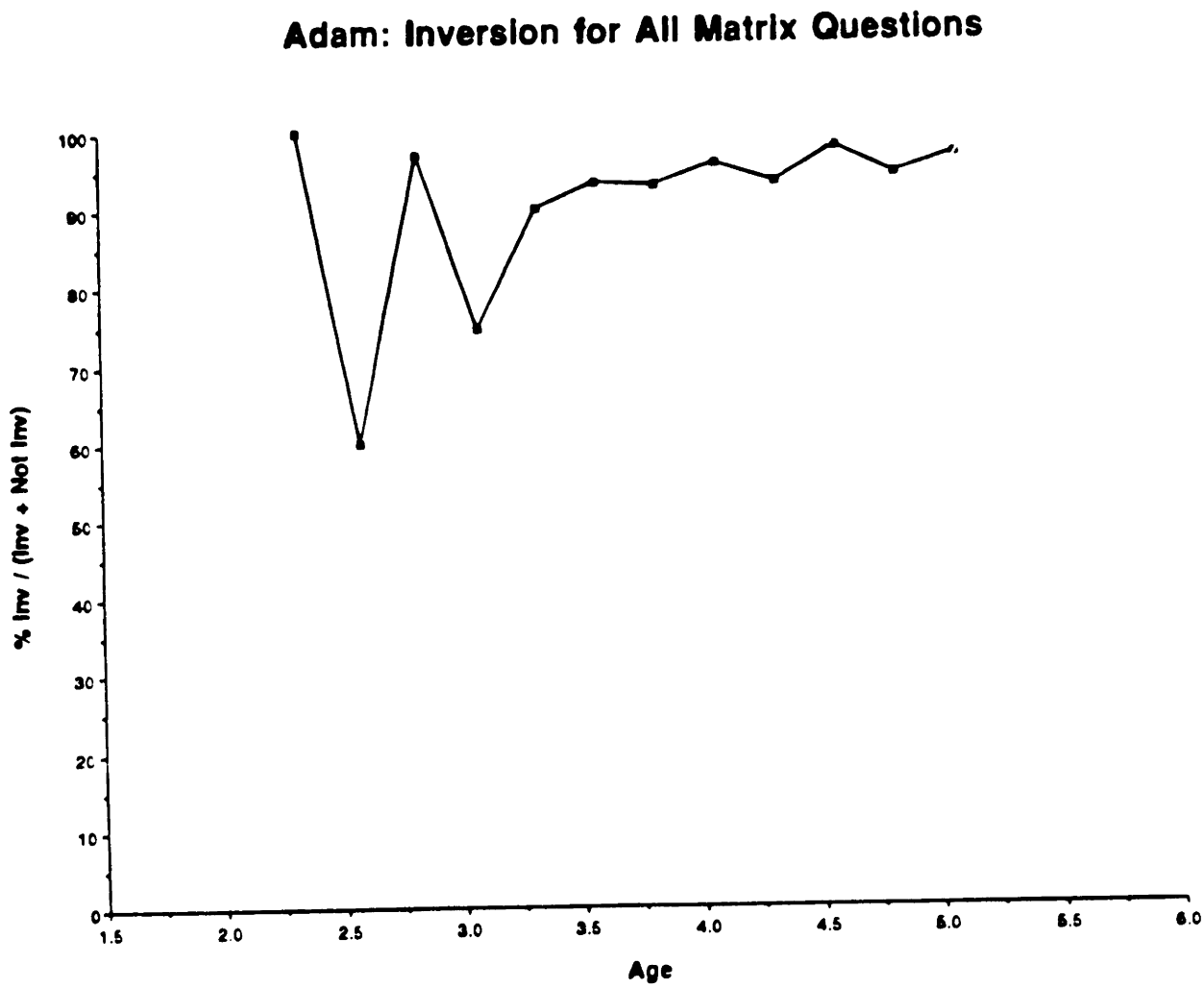


Figure 5.3
Developmental Trend: Allison

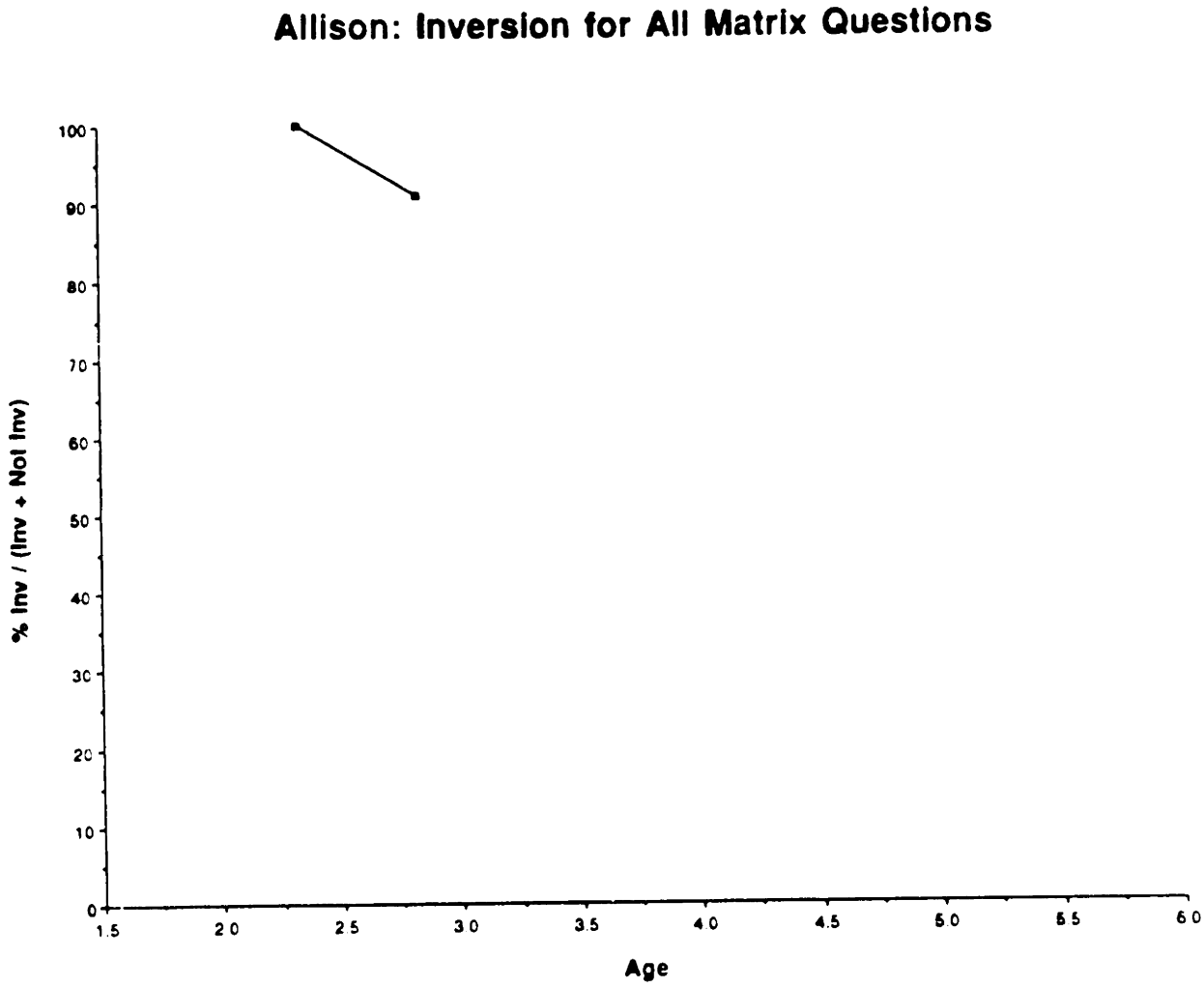


Figure 5.4
Developmental Trend: April

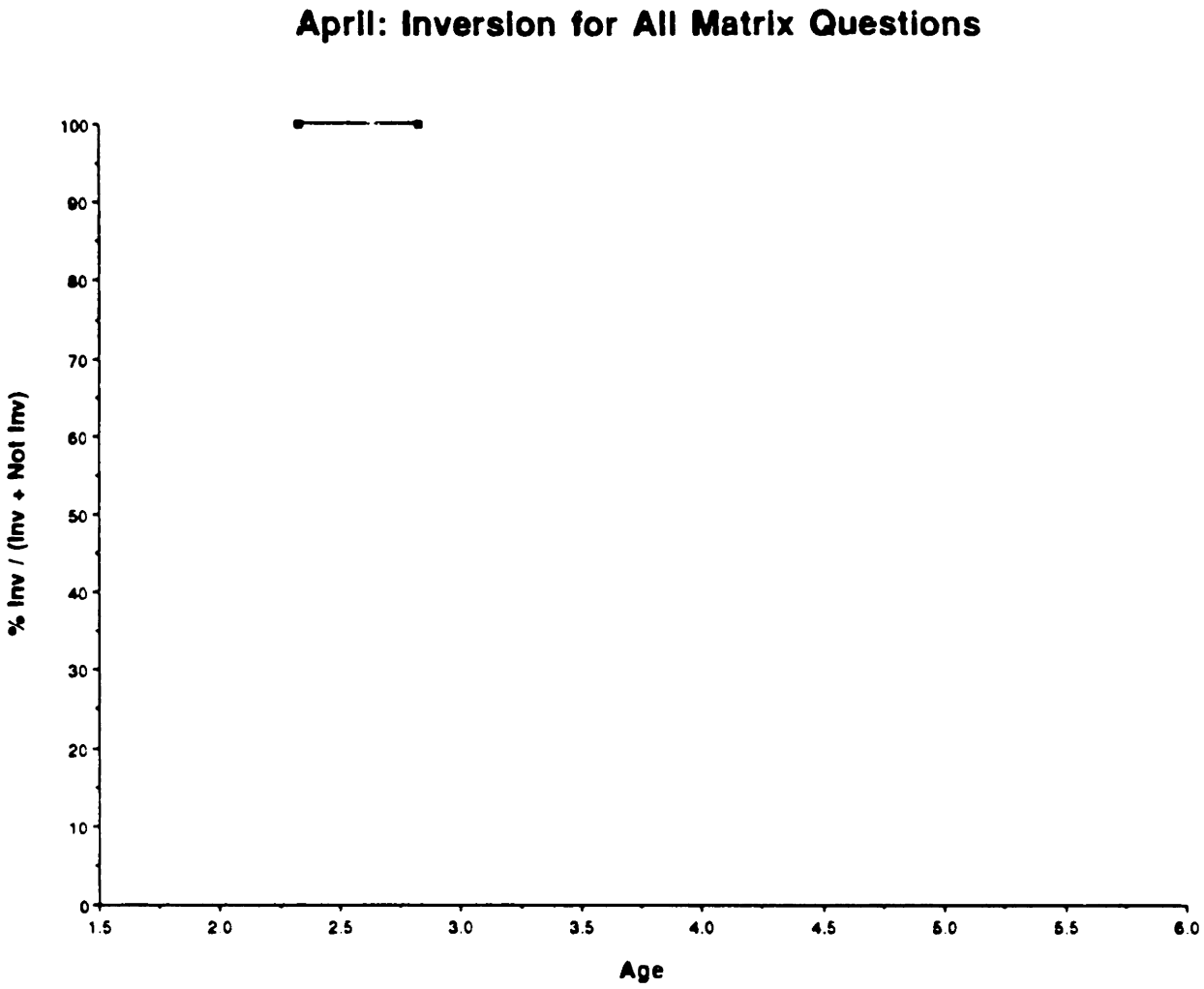


Figure 5.5
Developmental Trend: Eve

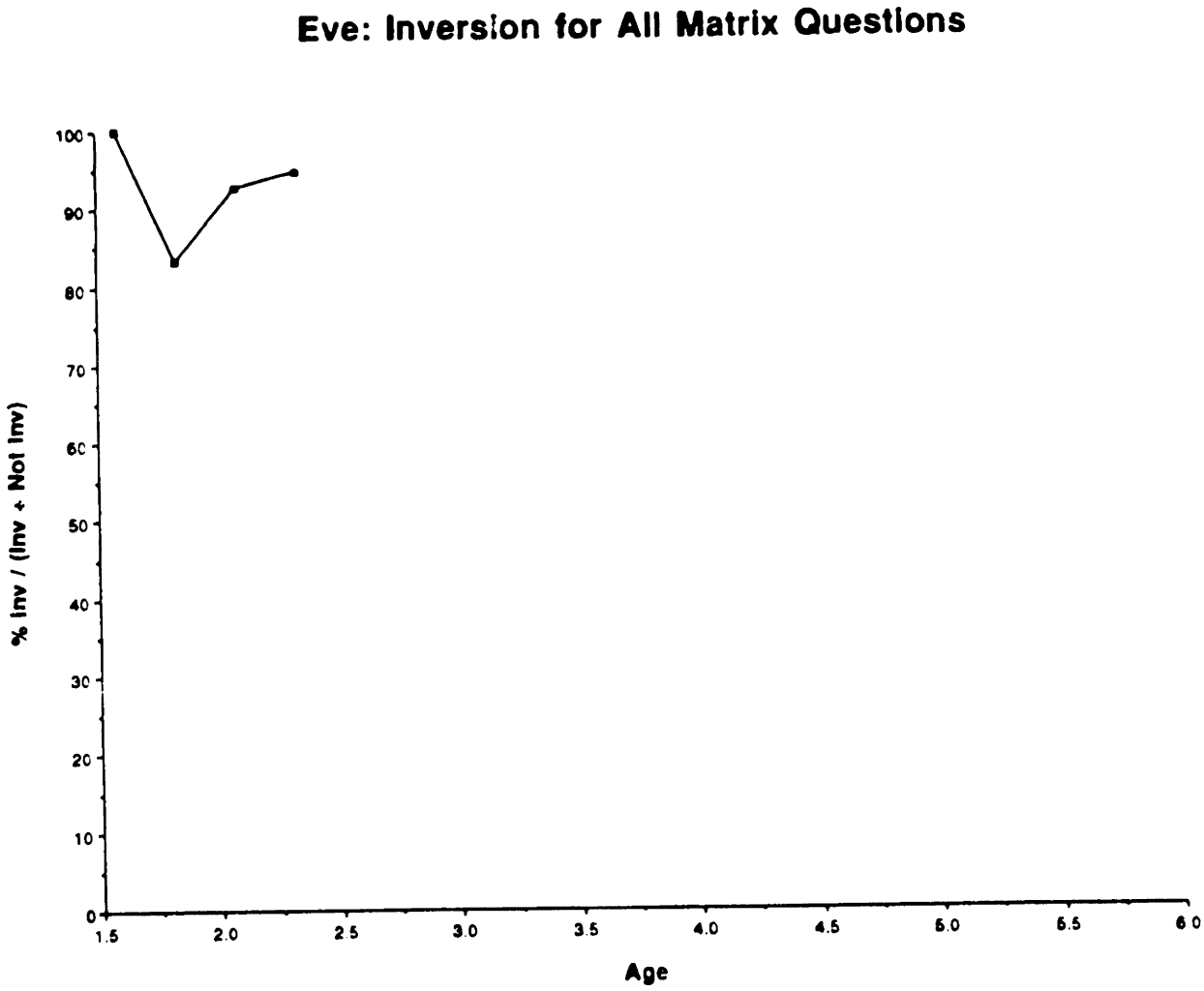


Figure 5.6
Developmental Trend: Mark

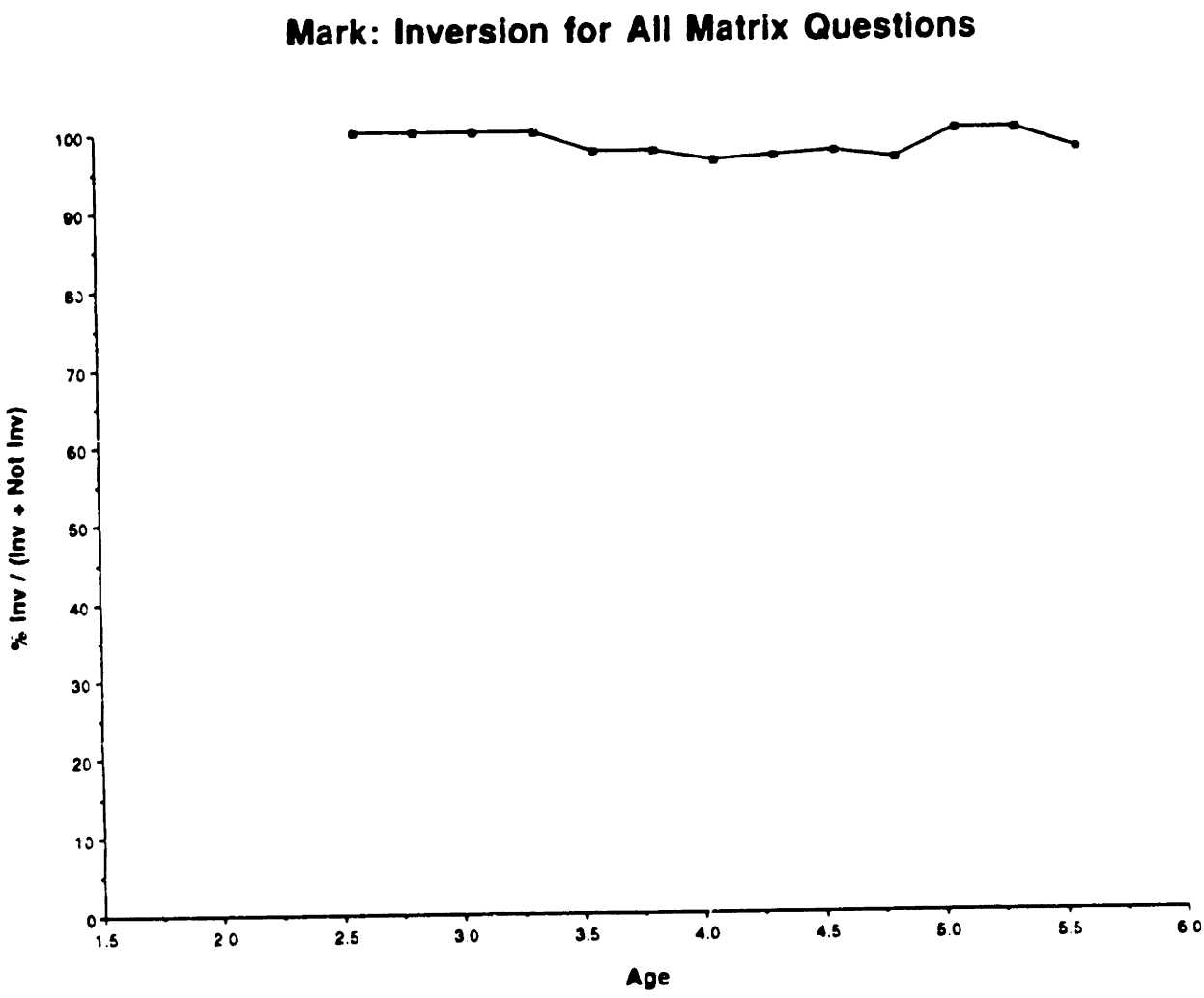


Figure 5.7
Developmental Trend: Naomi

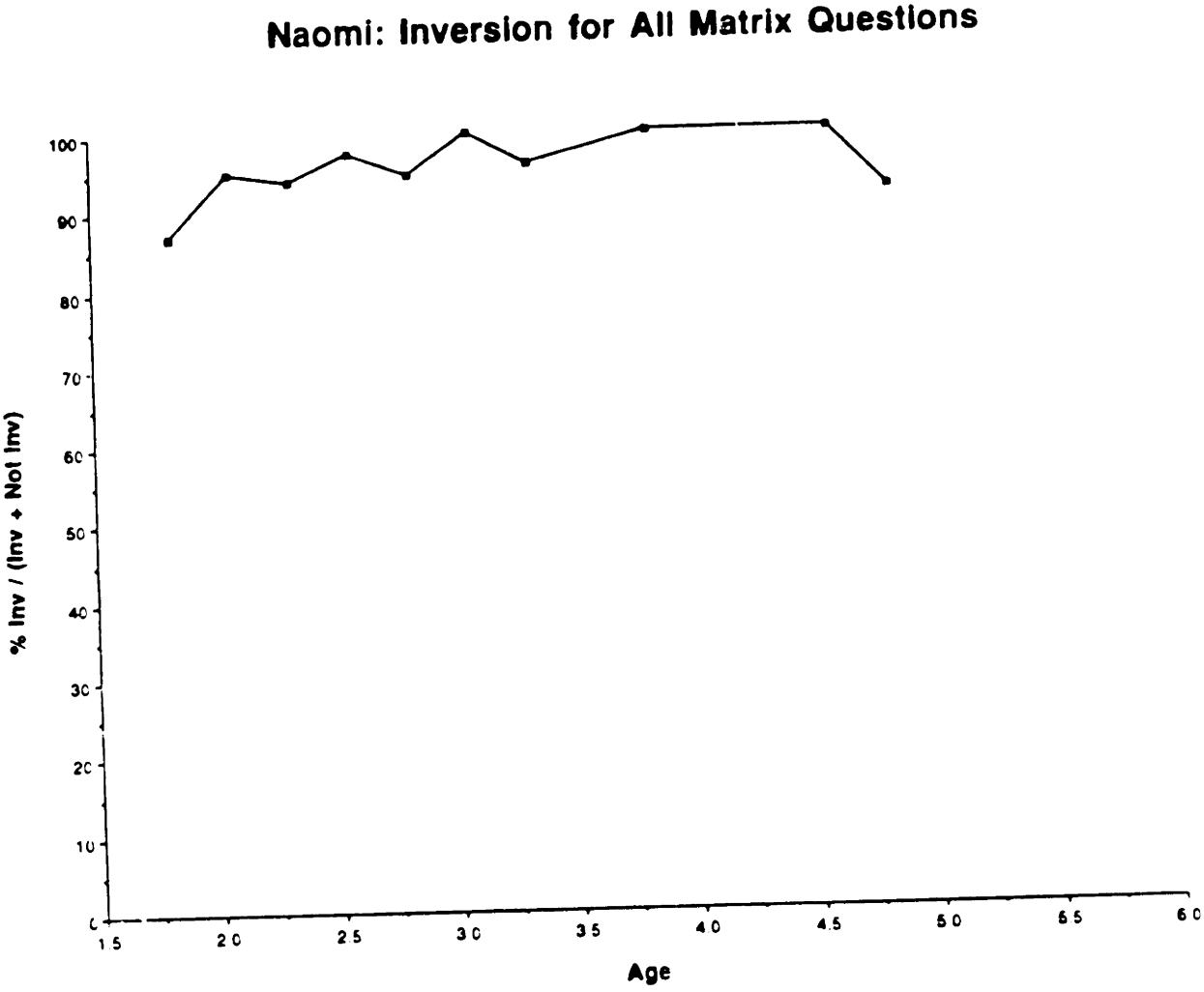


Figure 5.8
Developmental Trend: Nathan

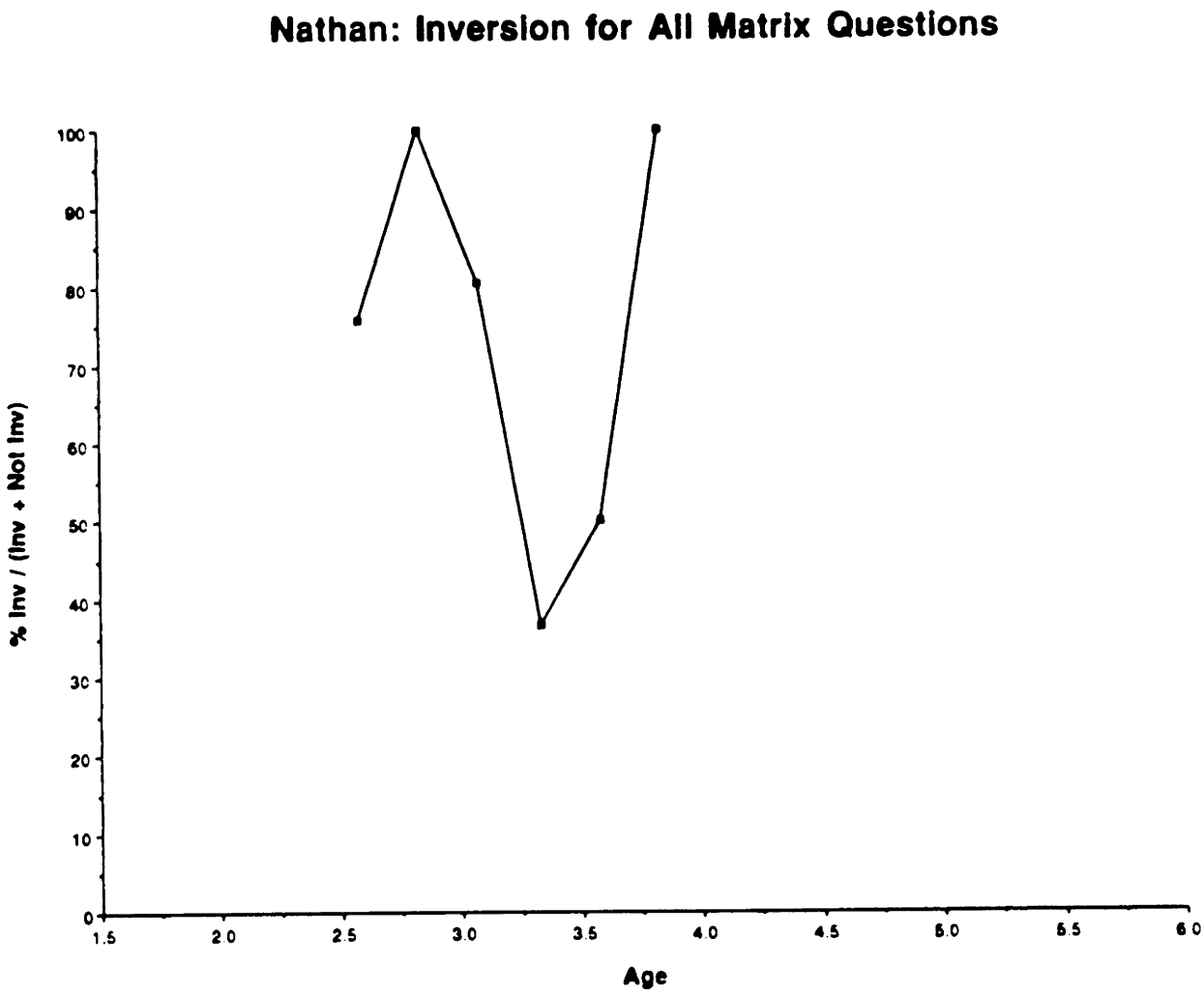


Figure 5.9
Developmental Trend: Nina

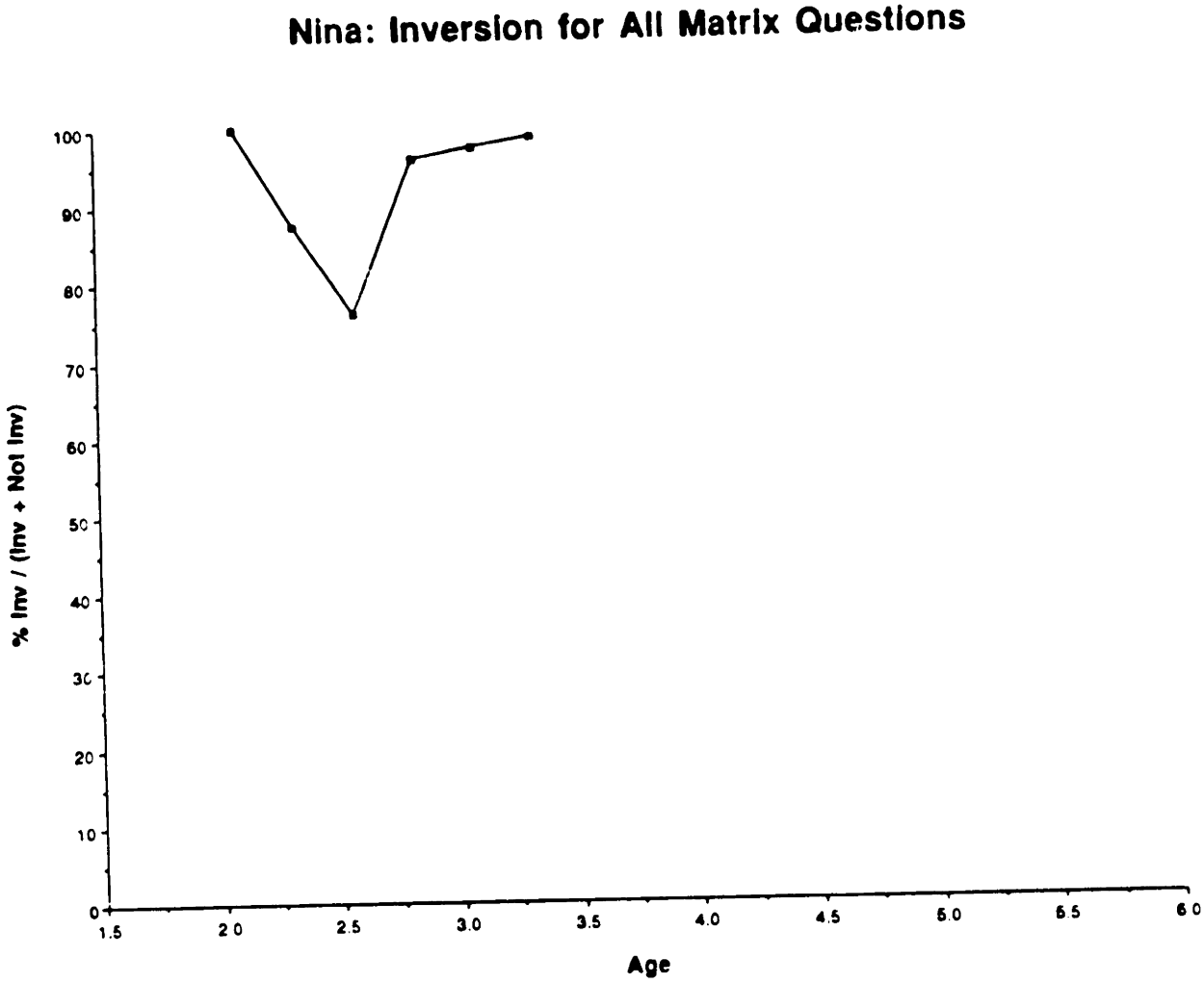


Figure 5.10
Developmental Trend: Peter

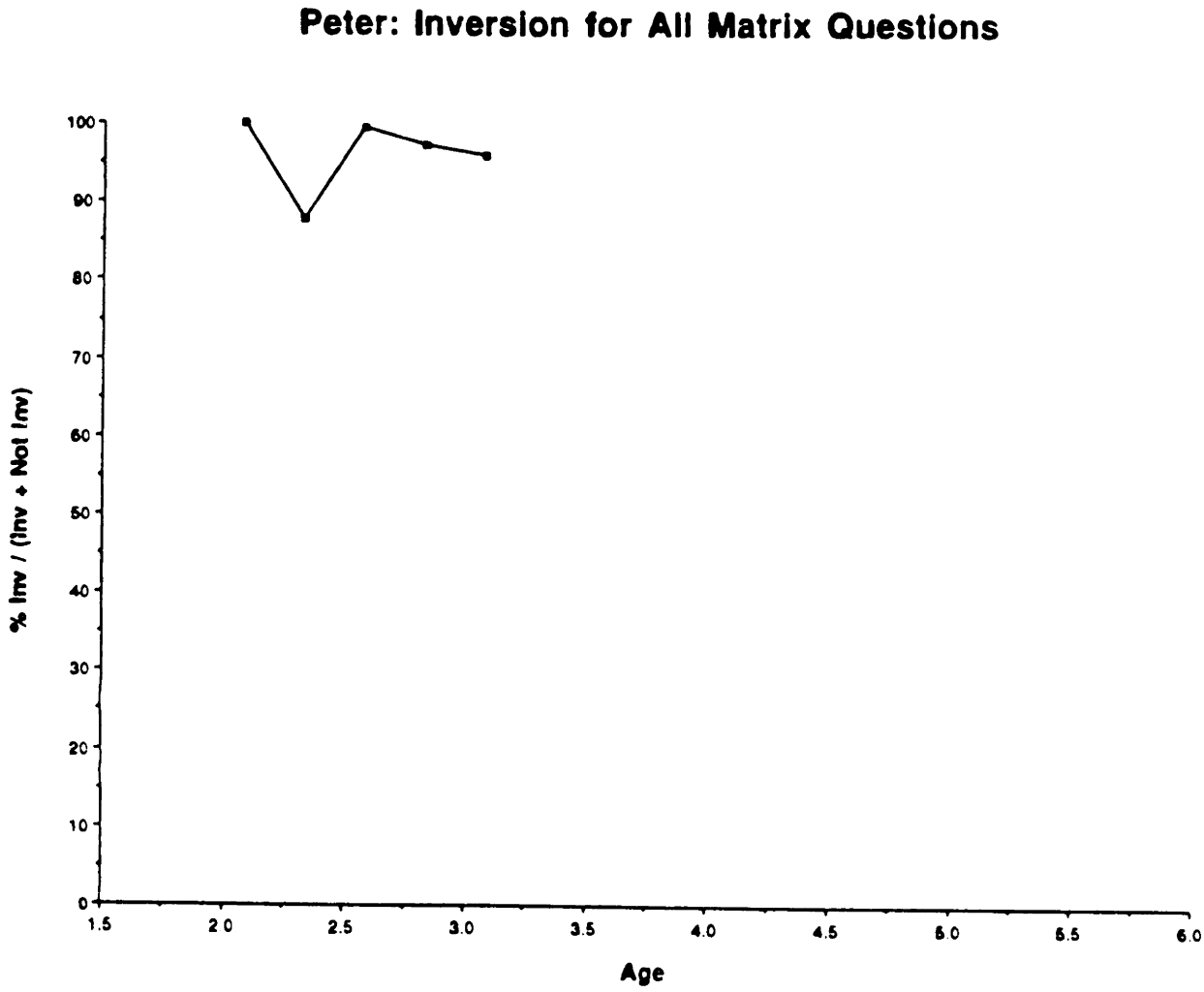


Figure 5.11
Developmental Trend: Ross

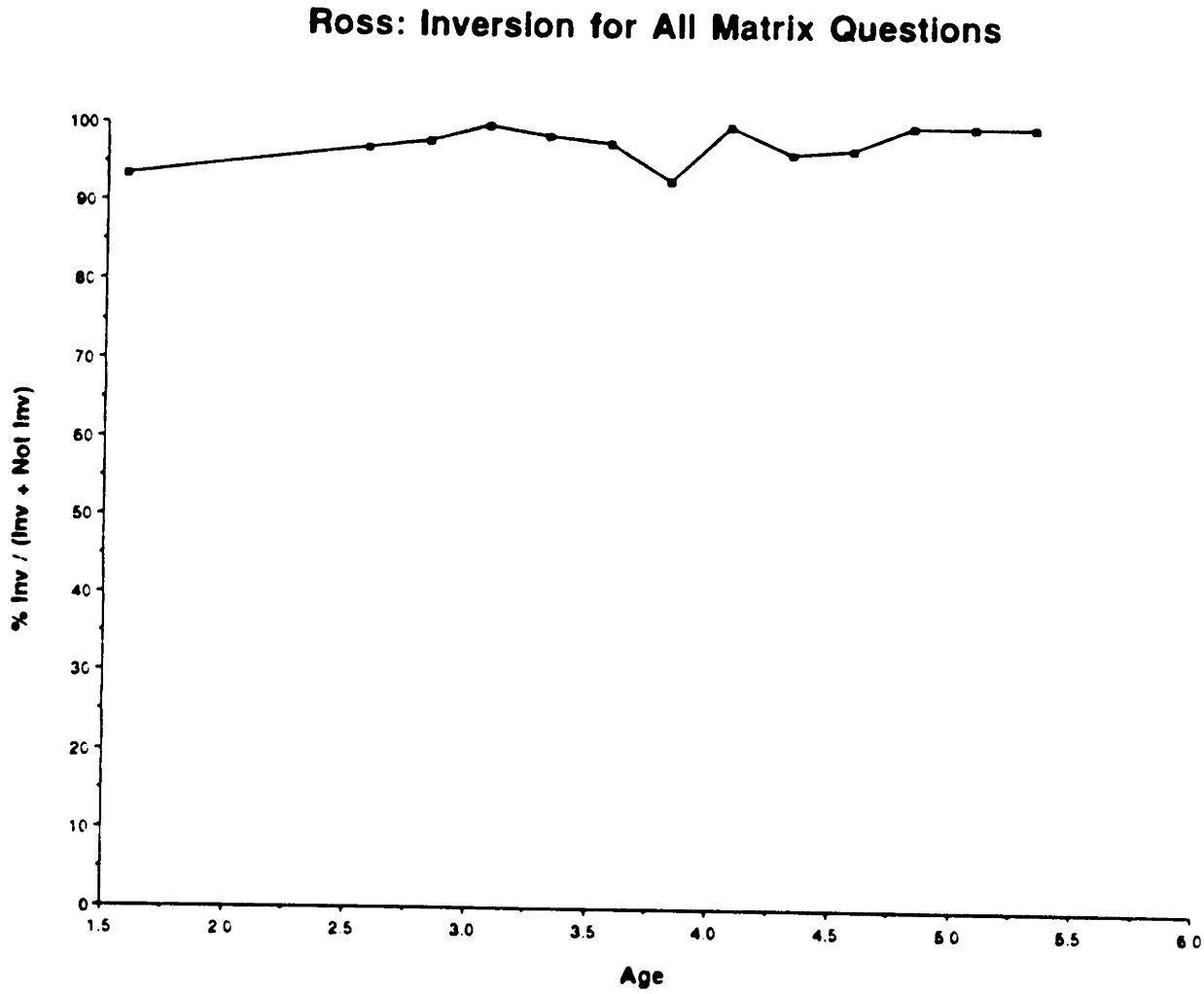


Figure 5.12
Developmental Trend: Sarah

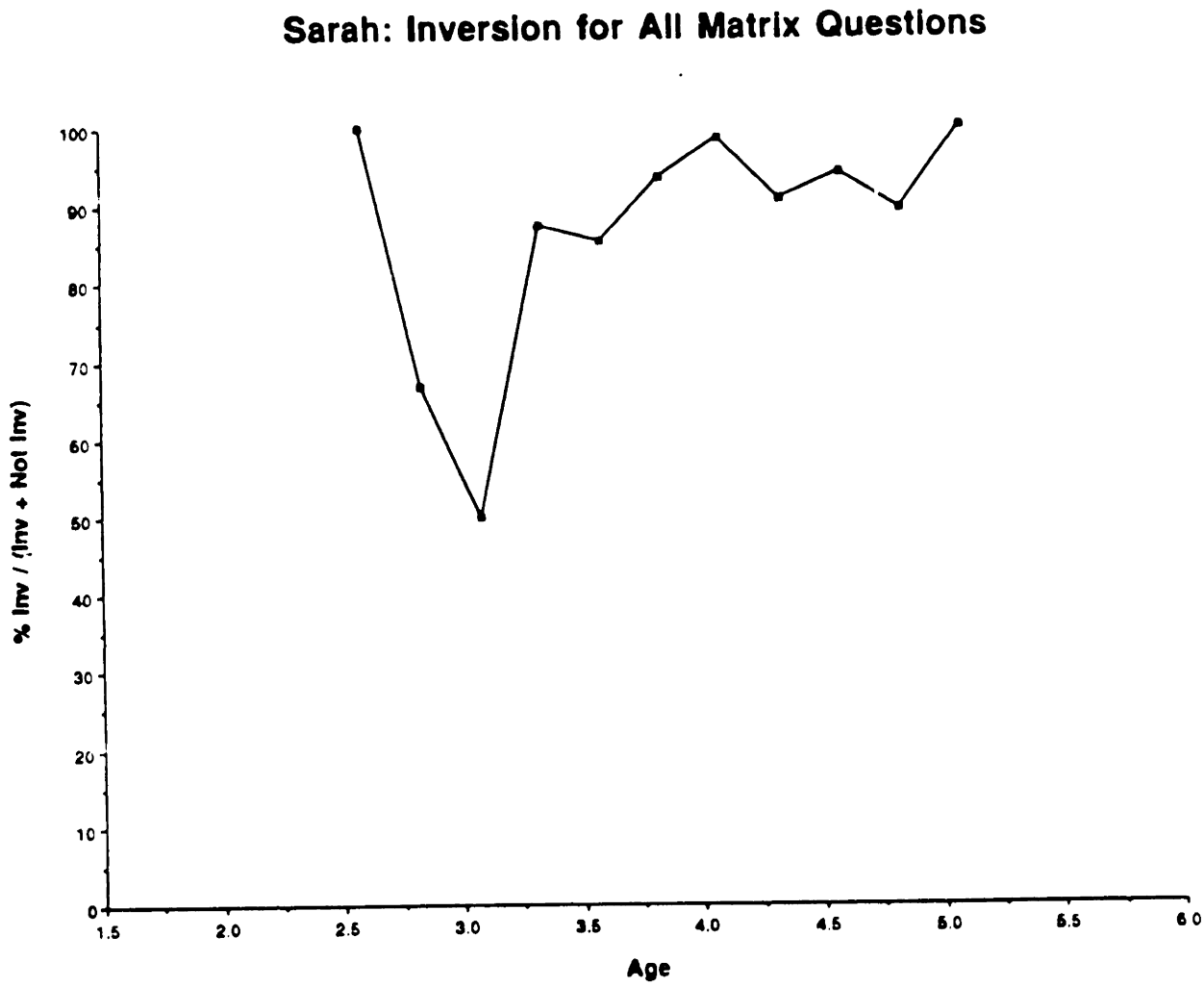


Figure 5.13
Developmental Trend: Shem

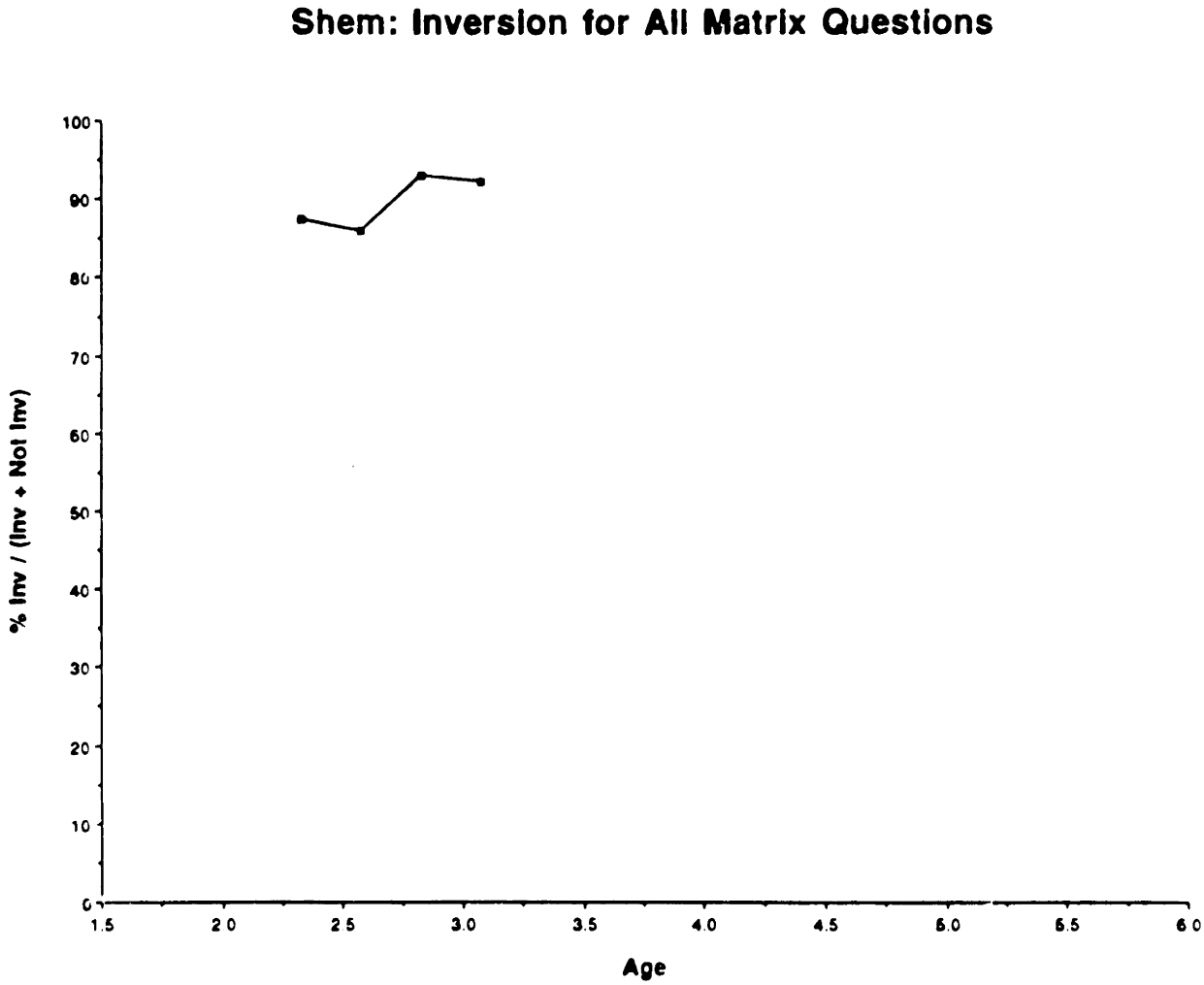


Figure 5.14
Overall Inversion Rate for Each Type of Question

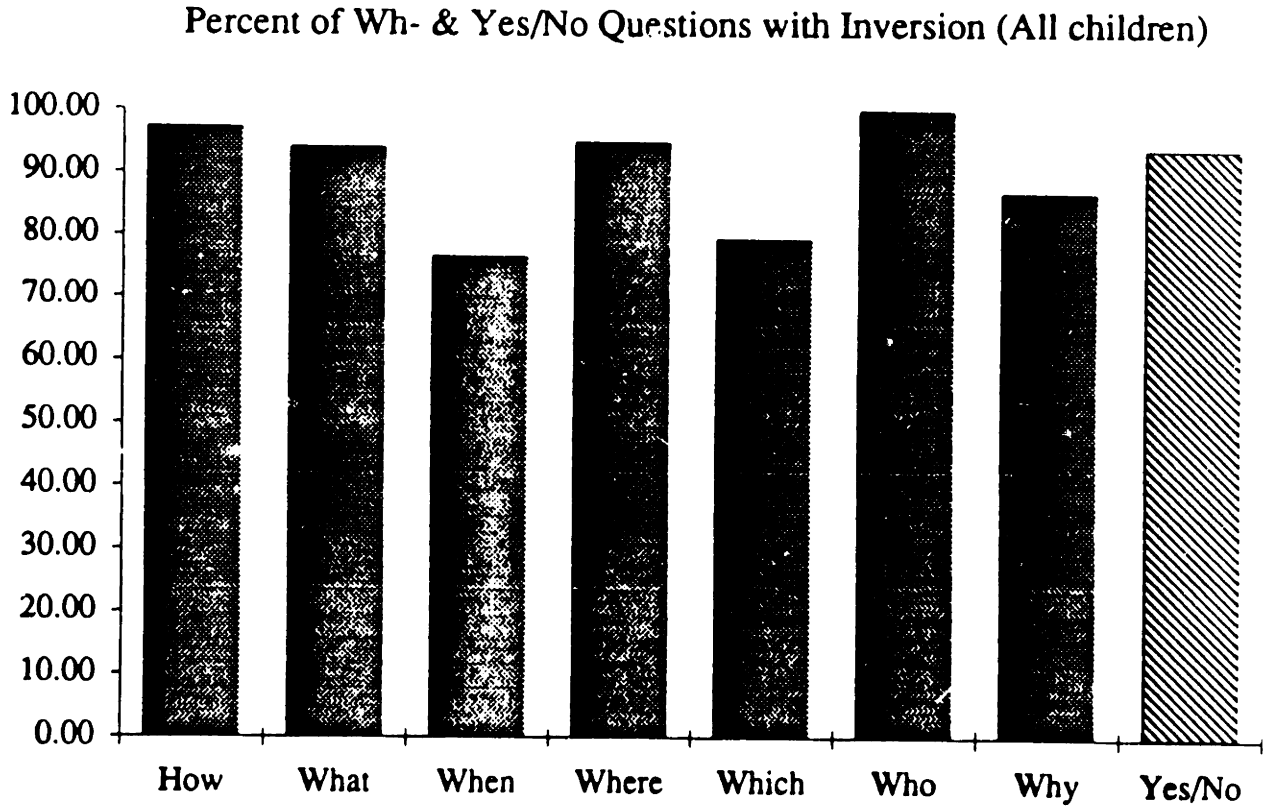


Figure 5.15
Individual Children's Inversion Rates for *Yes/No* Questions

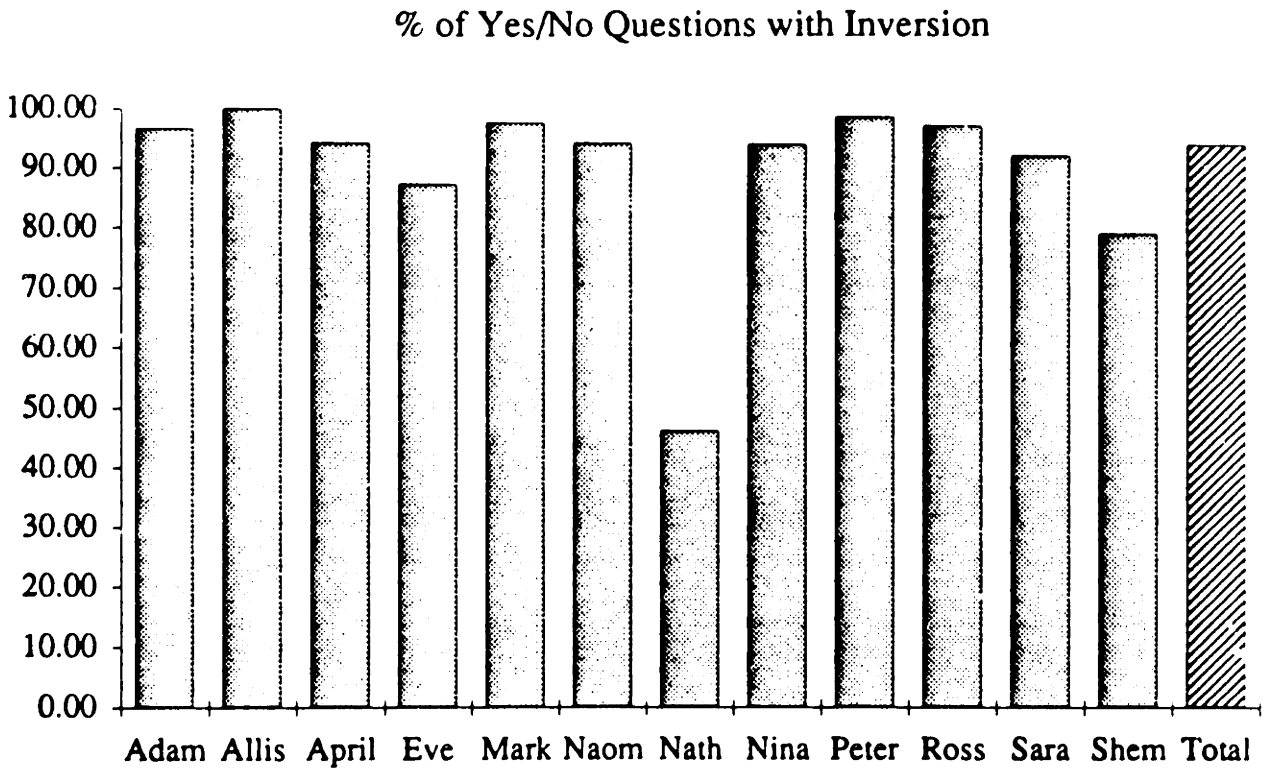


Figure 5.16
Overall Developmental Trend: *Yes/No* Questions

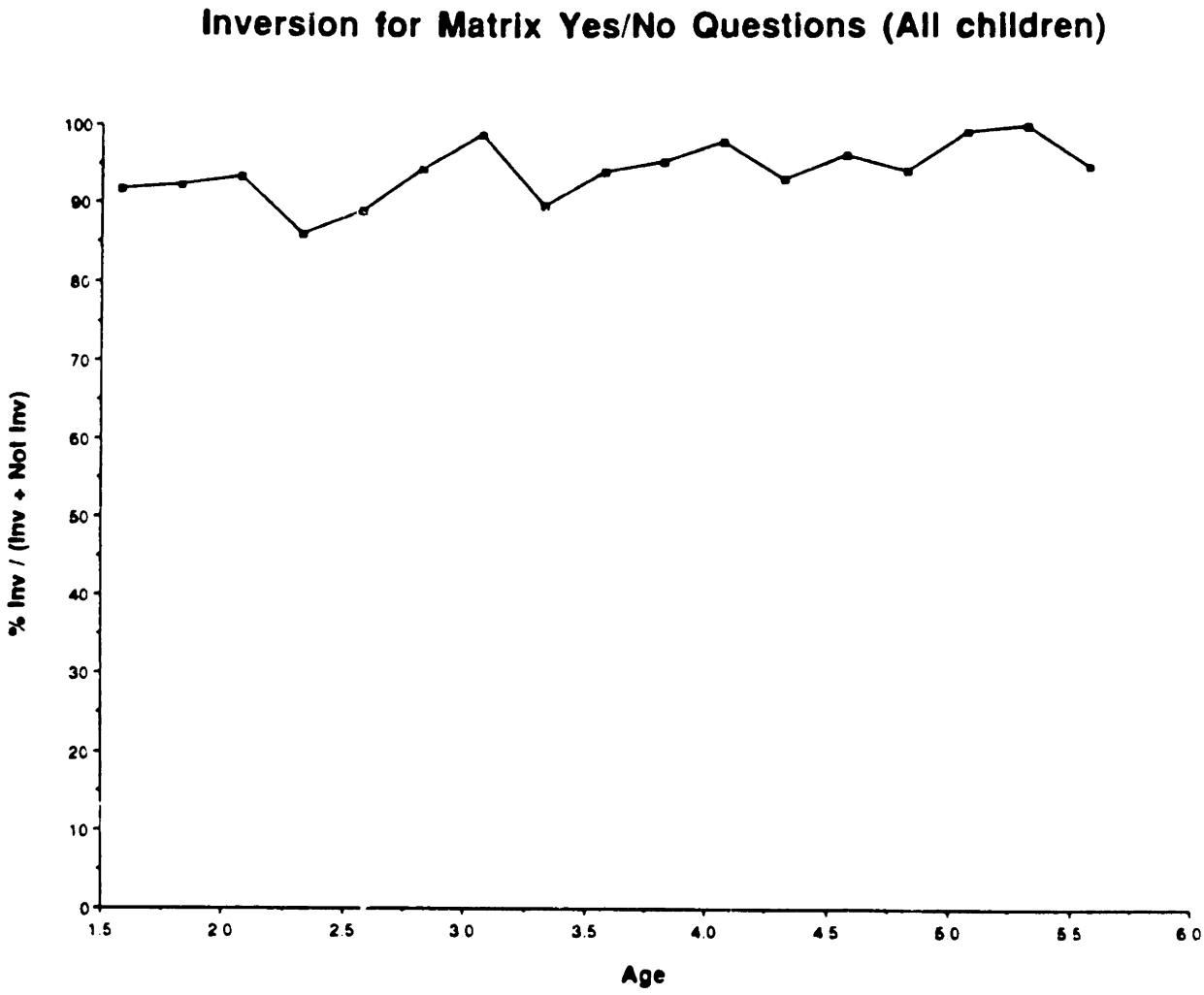


Figure 5.17
Individual Children's Inversion Rates for *Who* Questions

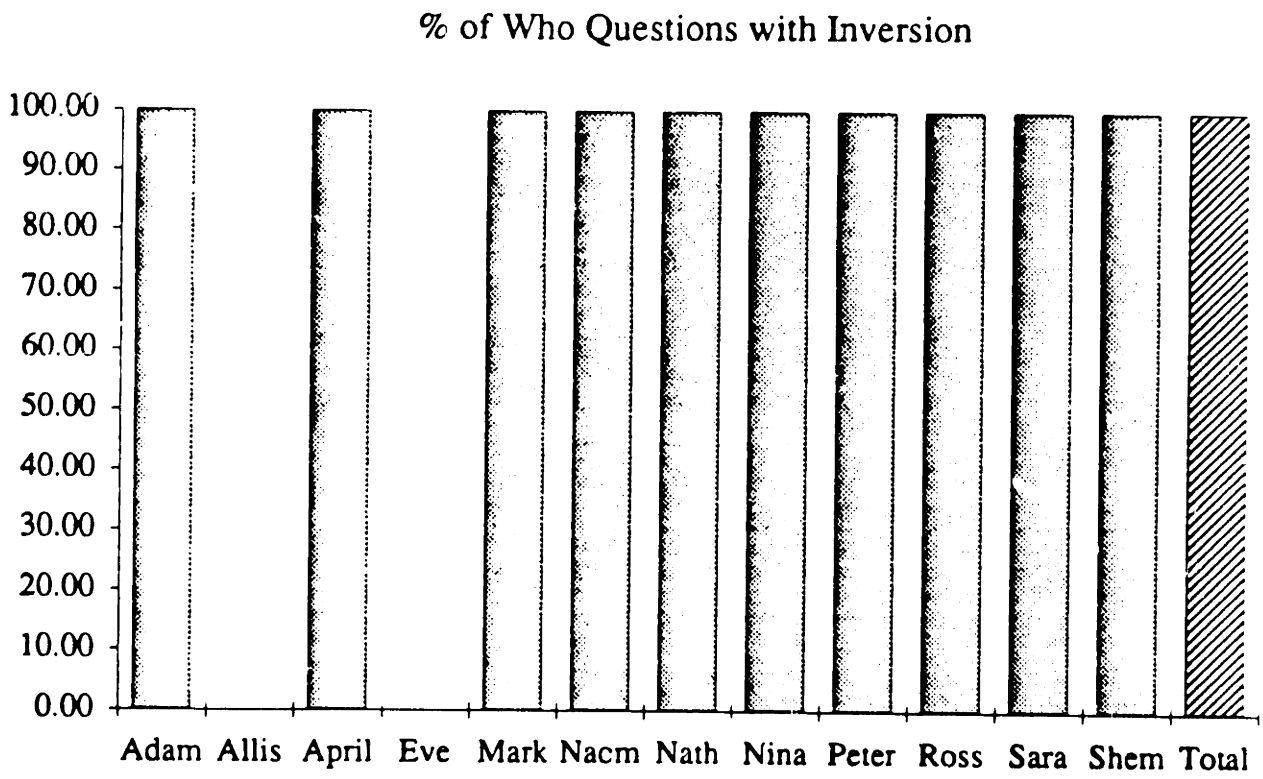


Figure 5.18
Individual Children's Inversion Rates for *How* Questions

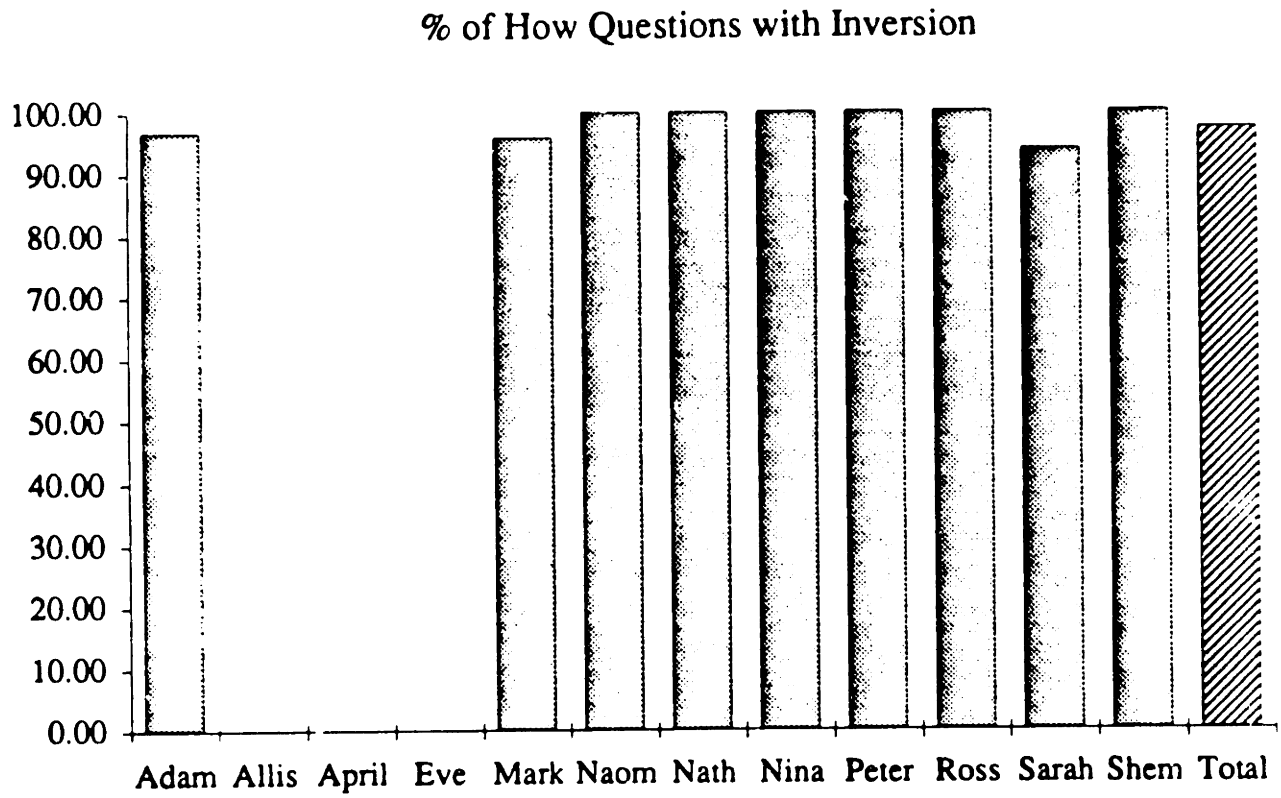


Figure 5.19
Individual Children's Inversion Rates for *Where* Questions

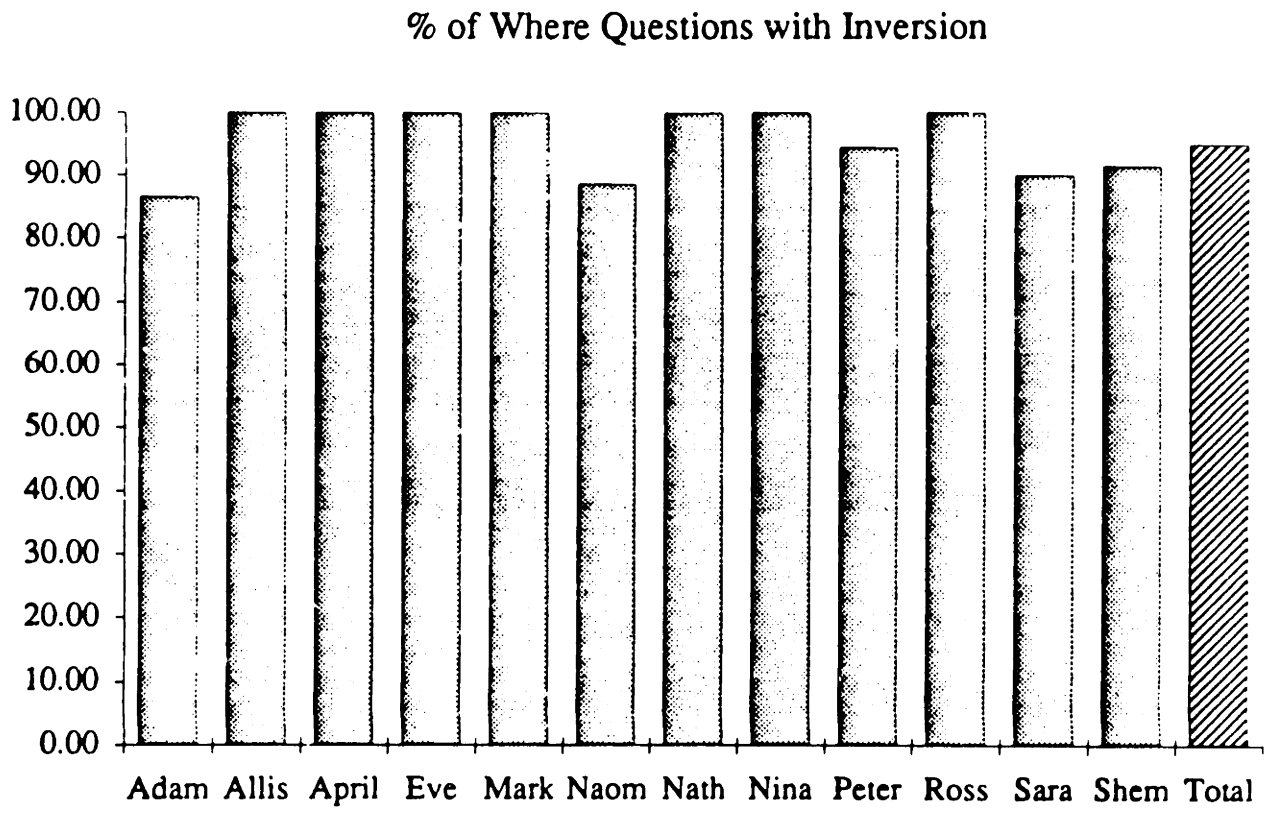


Figure 5.20
Individual Children's Inversion Rates for *What* Questions

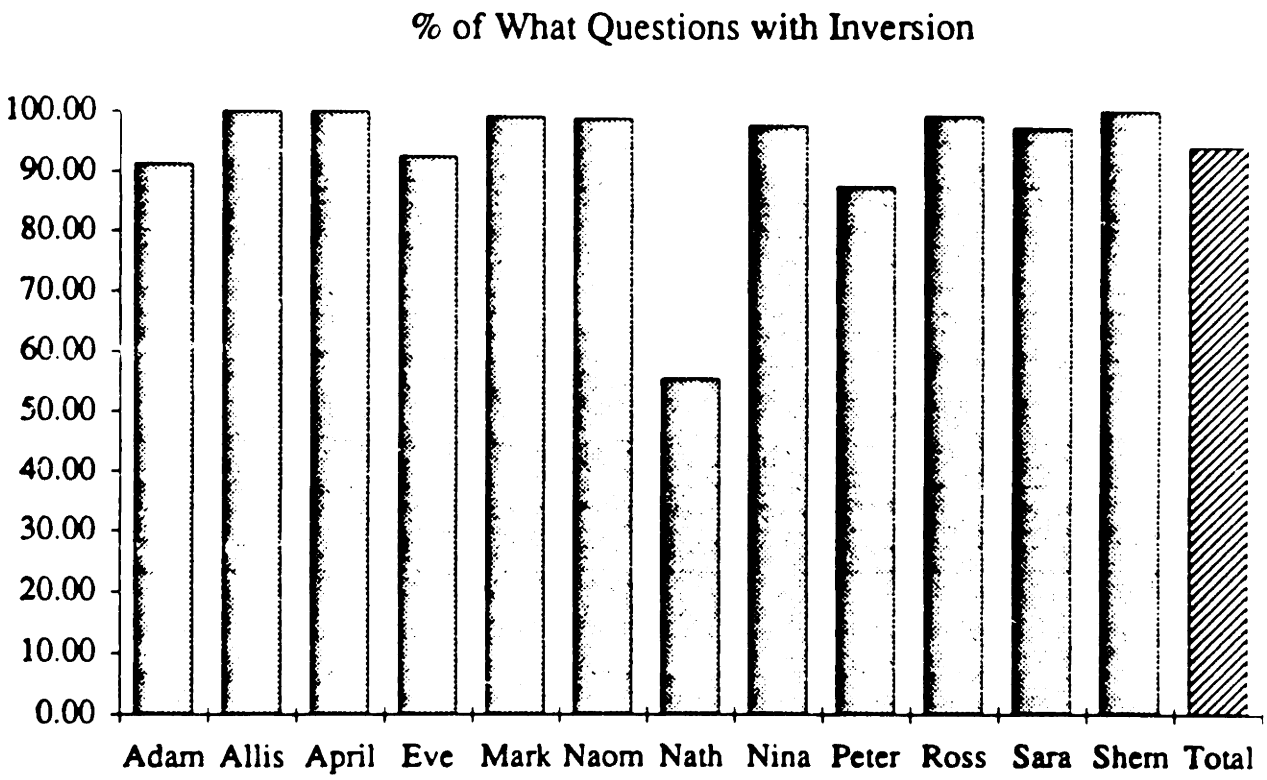


Figure 5.21
Individual Children's Inversion Rates for *Which* Questions

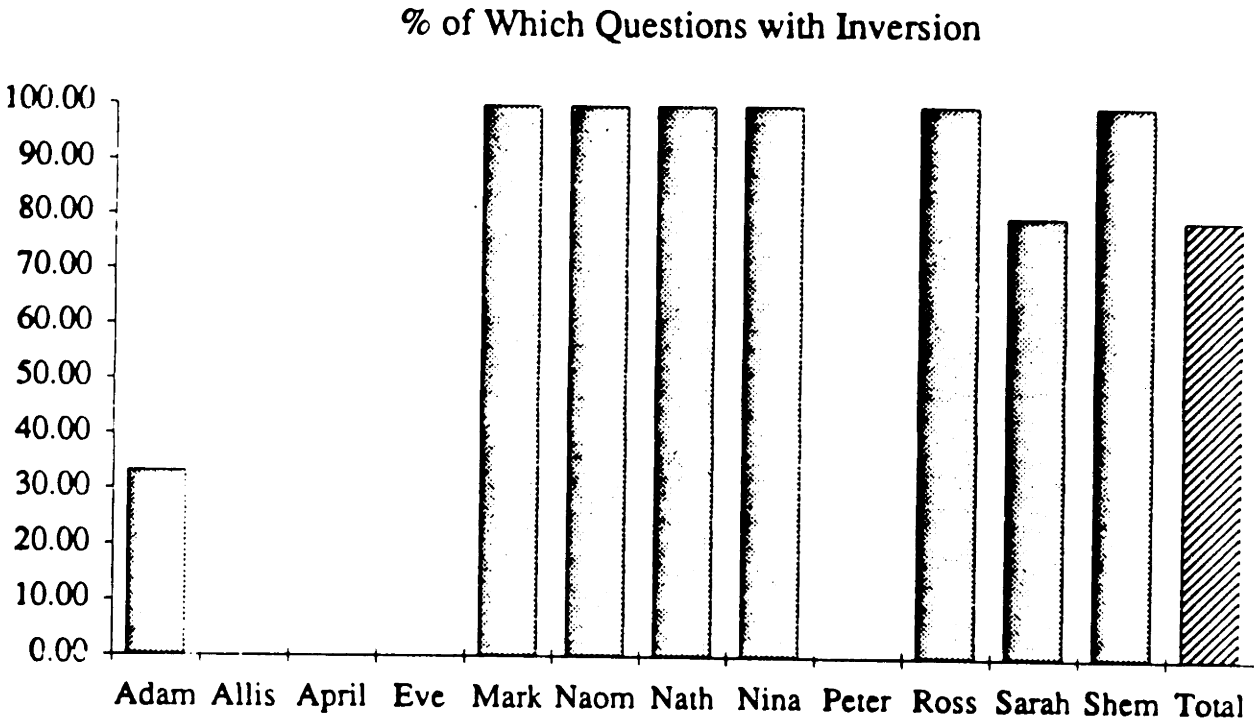


Figure 5.22
Individual Children's Inversion Rates for *When* Questions

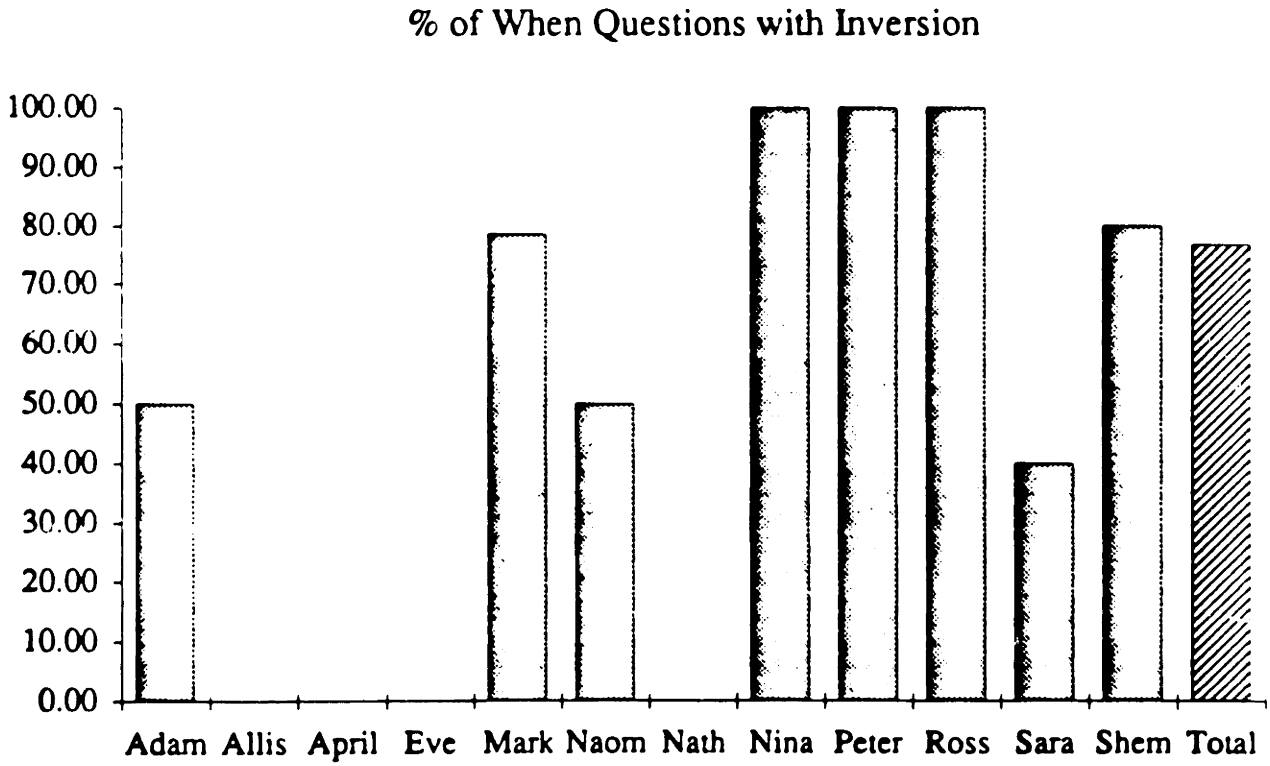


Figure 5.23
Individual Children's Inversion Rates for *Why* Questions

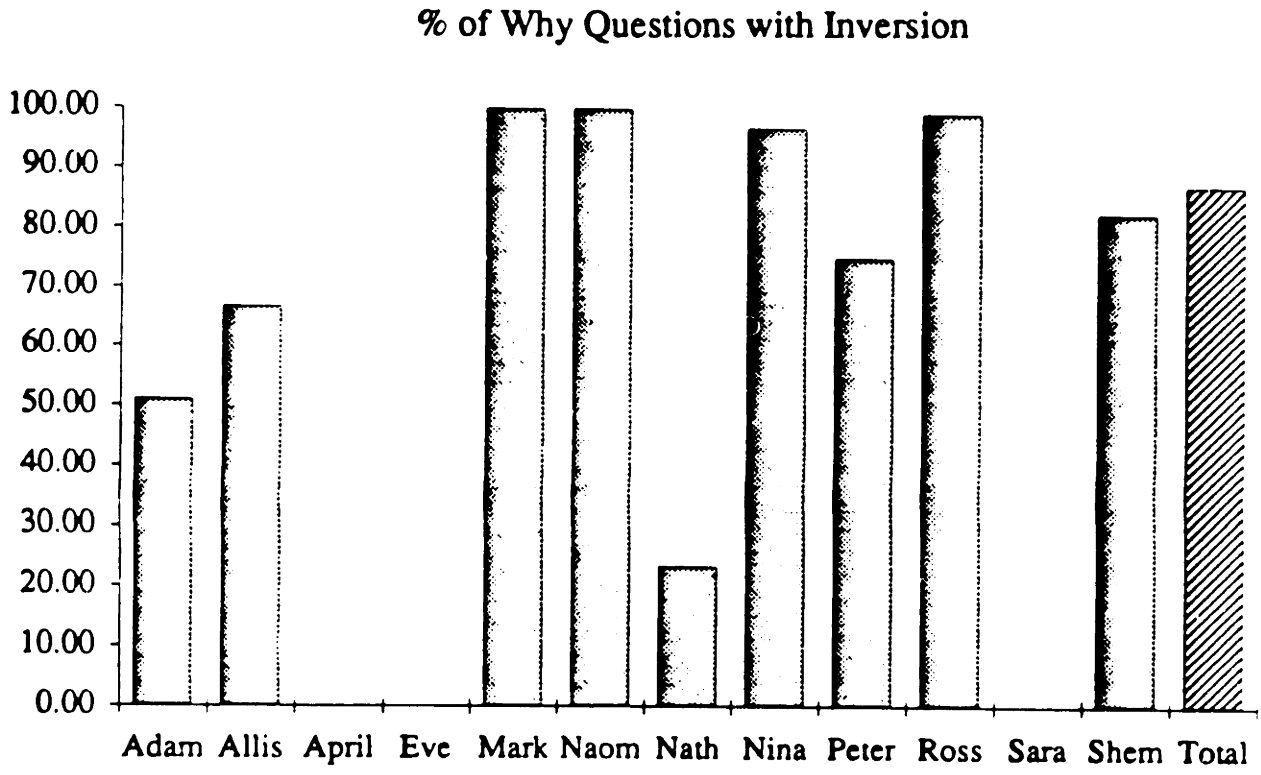


Figure 5.24
Overall Developmental Trend: *Who* Questions

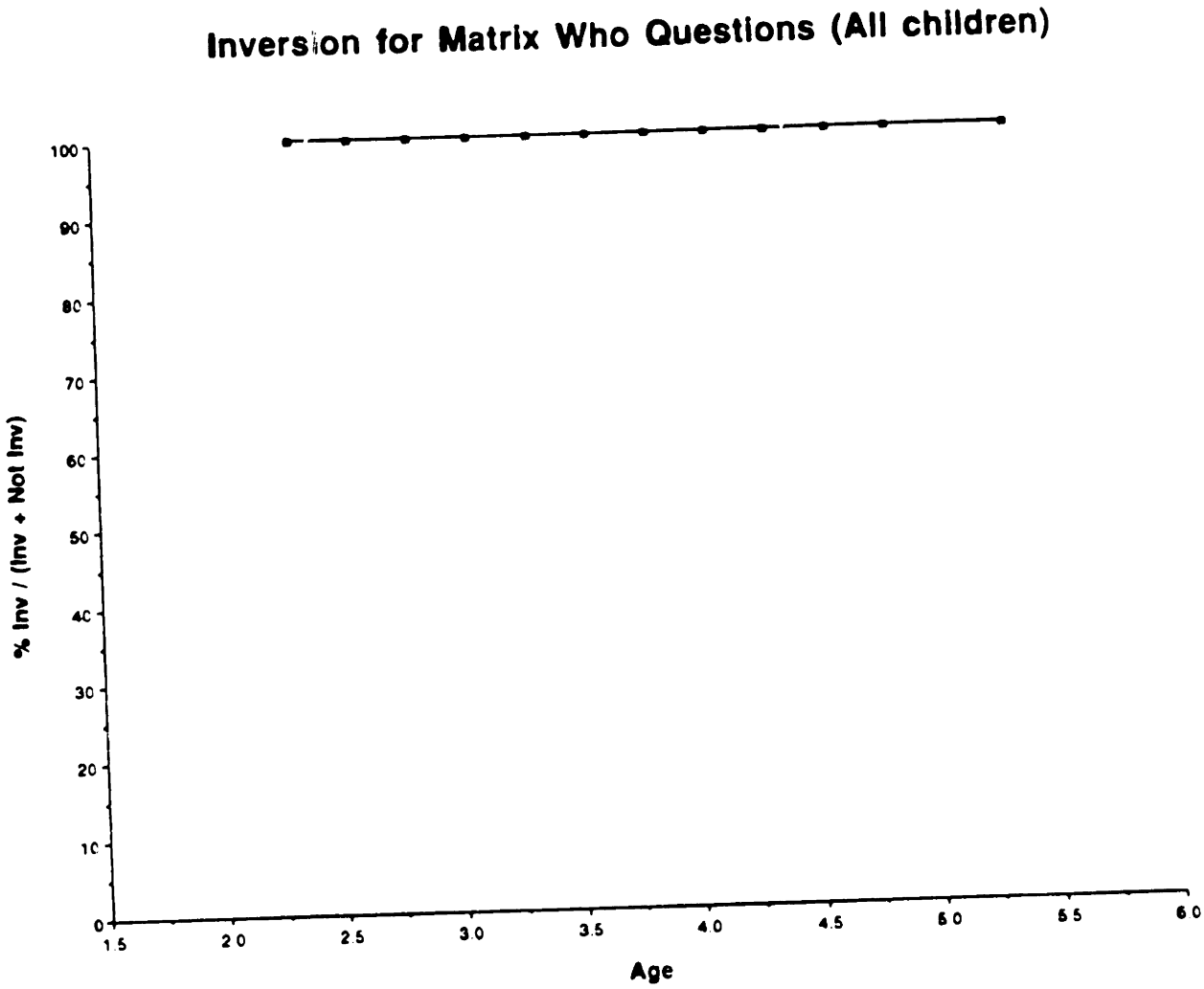


Figure 5.25
Overall Developmental Trend: *How* Questions

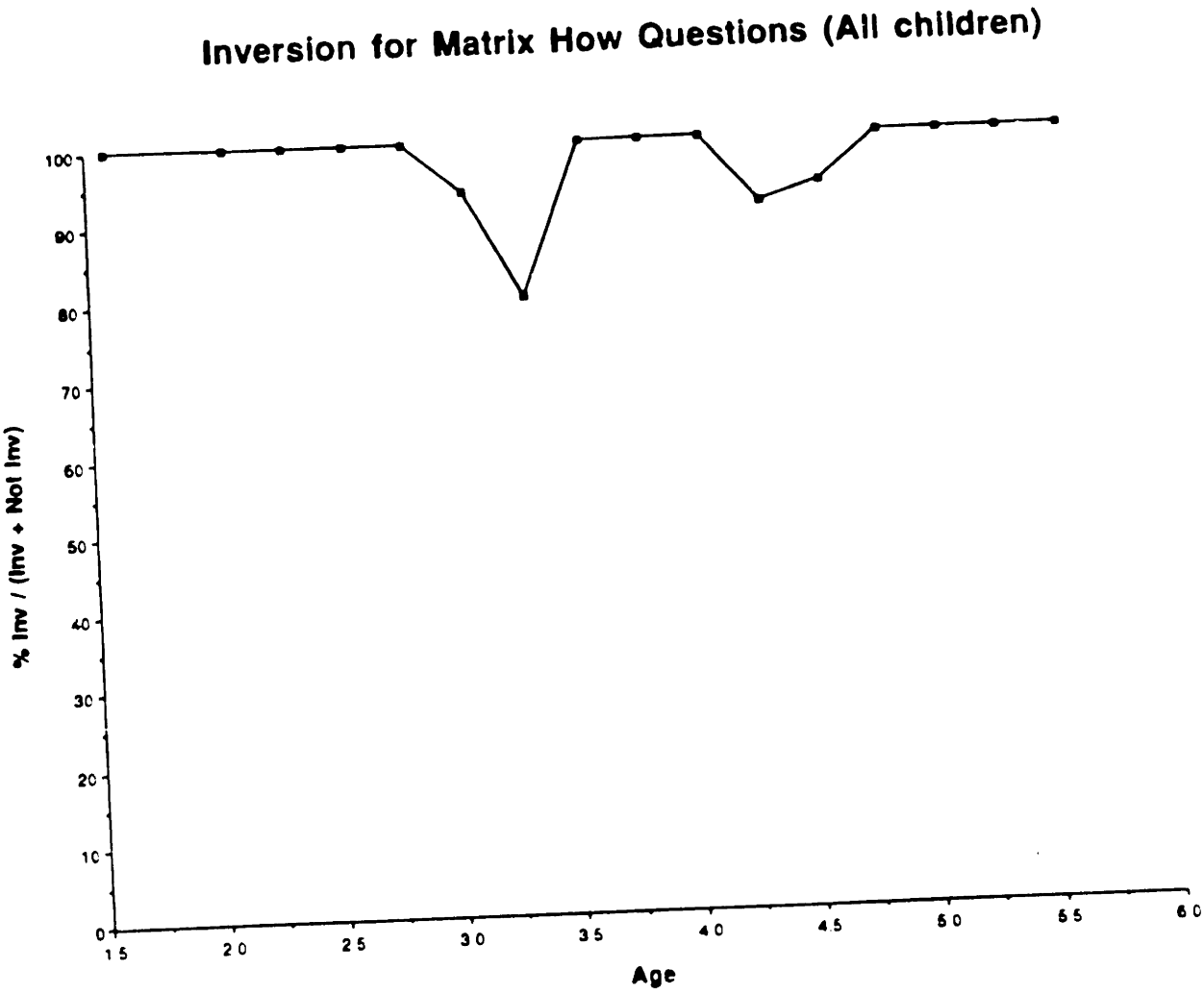


Figure 5.26
Overall Developmental Trend: *Where* Questions

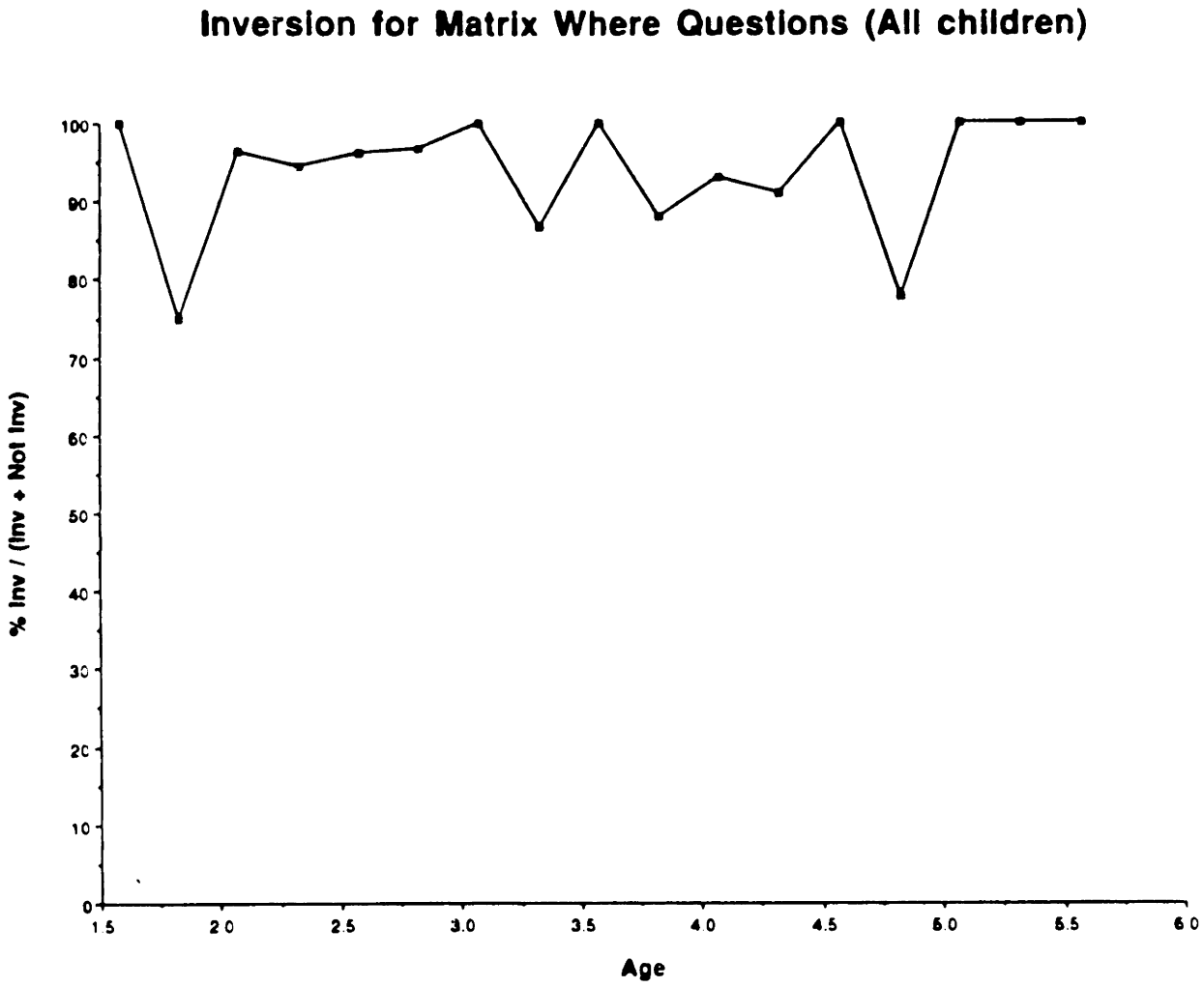


Figure 5.27
Overall Developmental Trend: *What* Questions

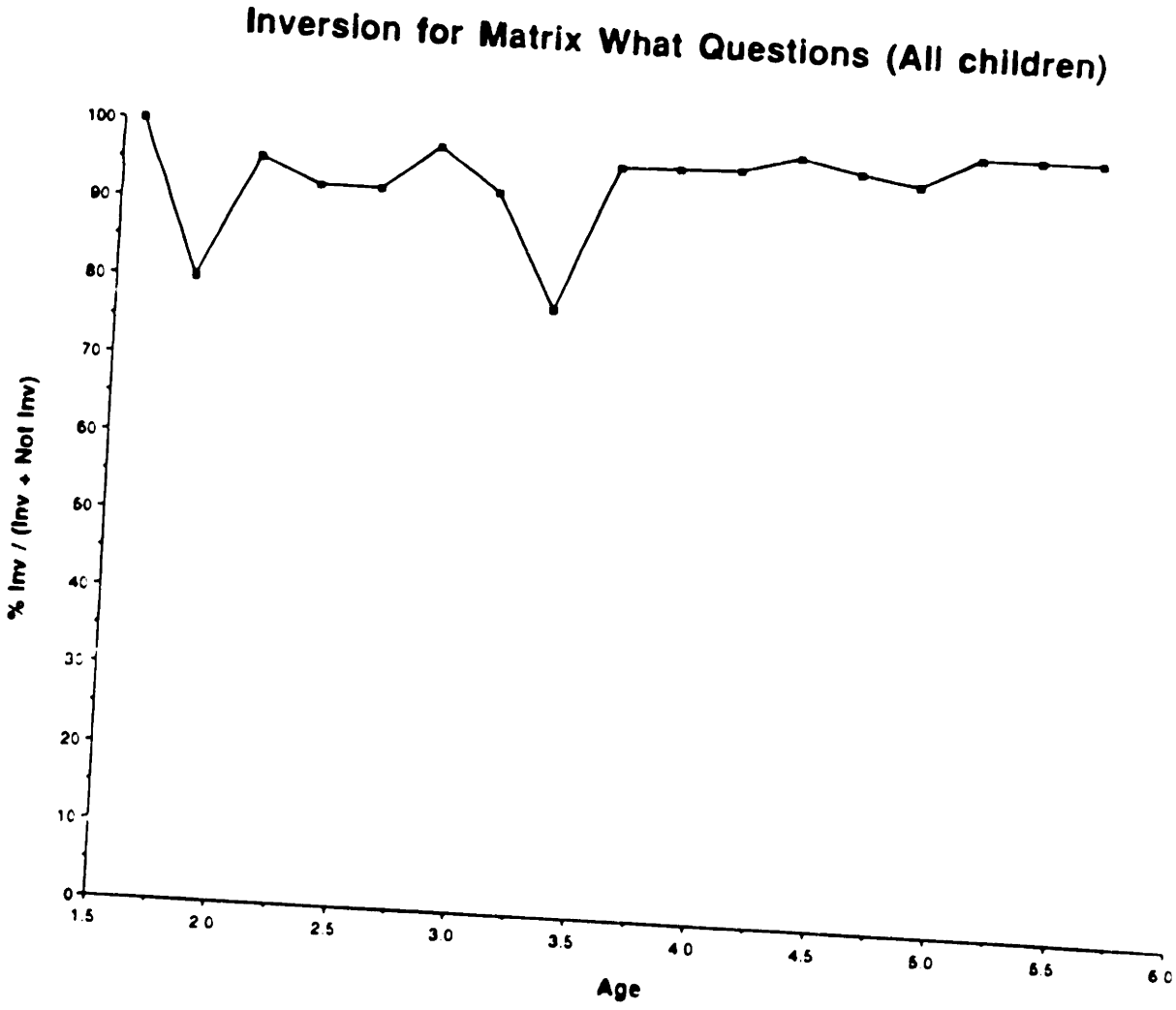


Figure 5.28
Overall Developmental Trend: *Which* Questions

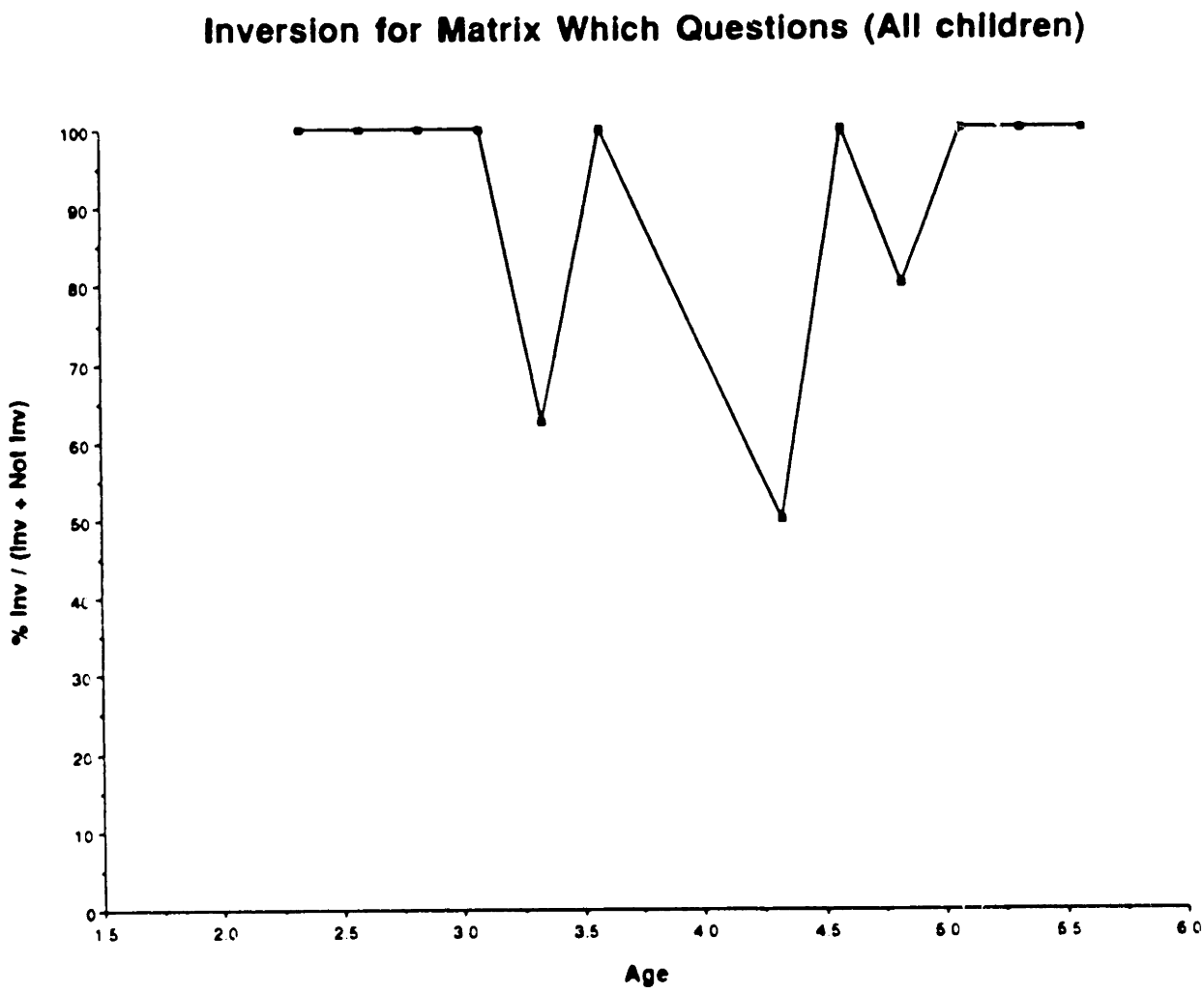


Figure 5.29
Overall Developmental Trend: *When* Questions

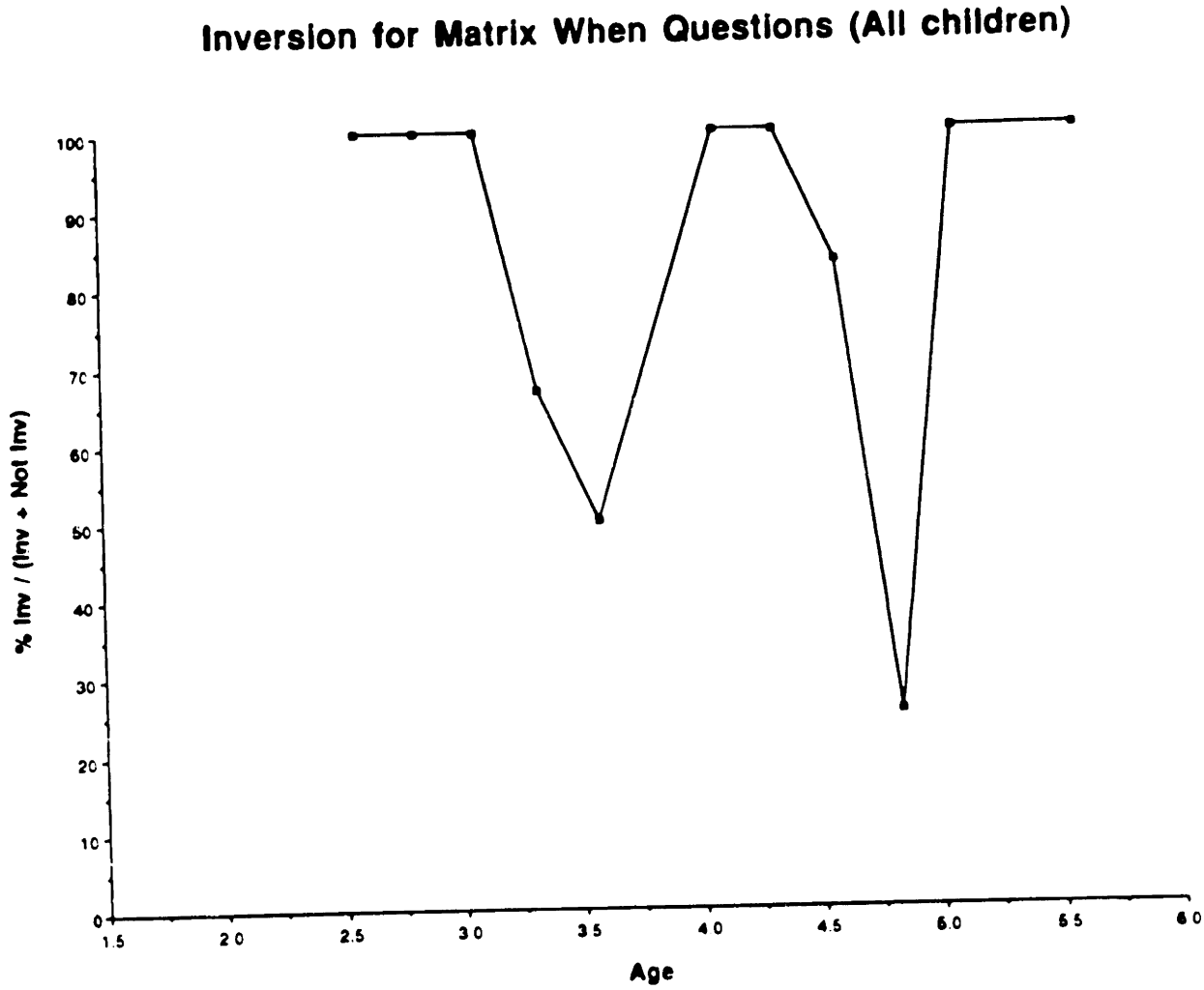


Figure 5.30
Overall Developmental Trend: *Why* Questions

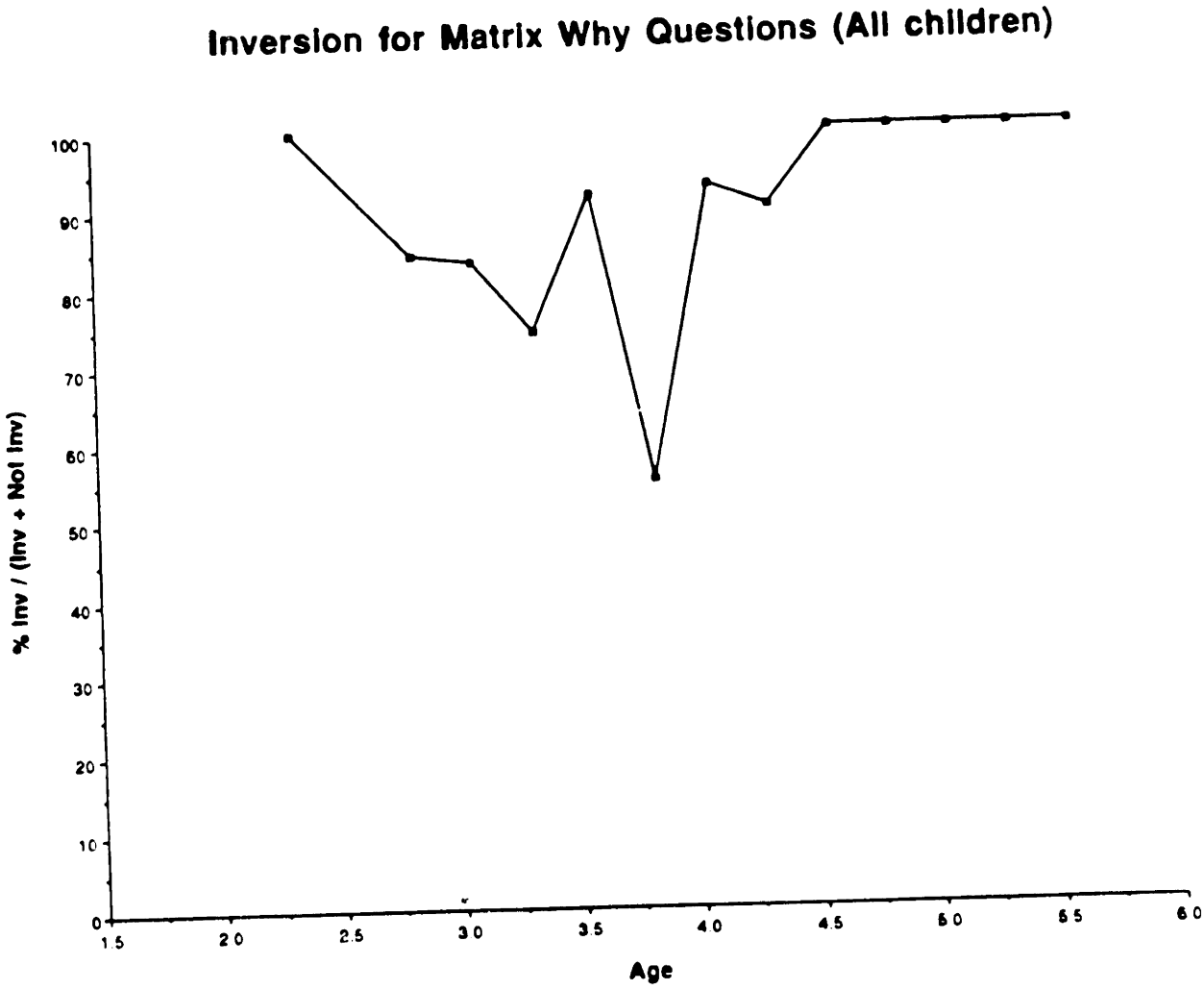


Figure 5.31
Overall Inversion Rates for Forms of *Be*

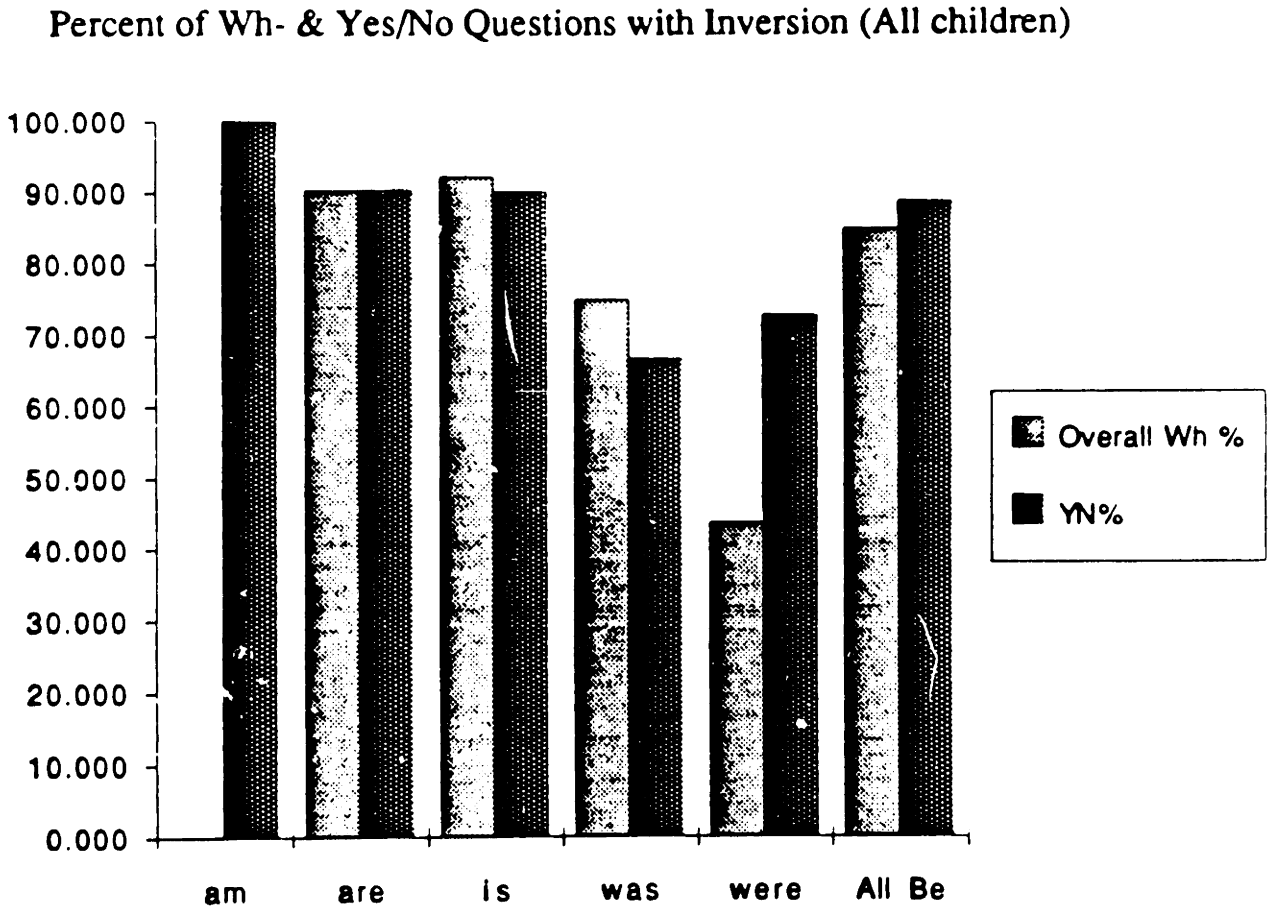


Figure 5.32
Overall Inversion Rates for Forms of Auxiliary *Do*

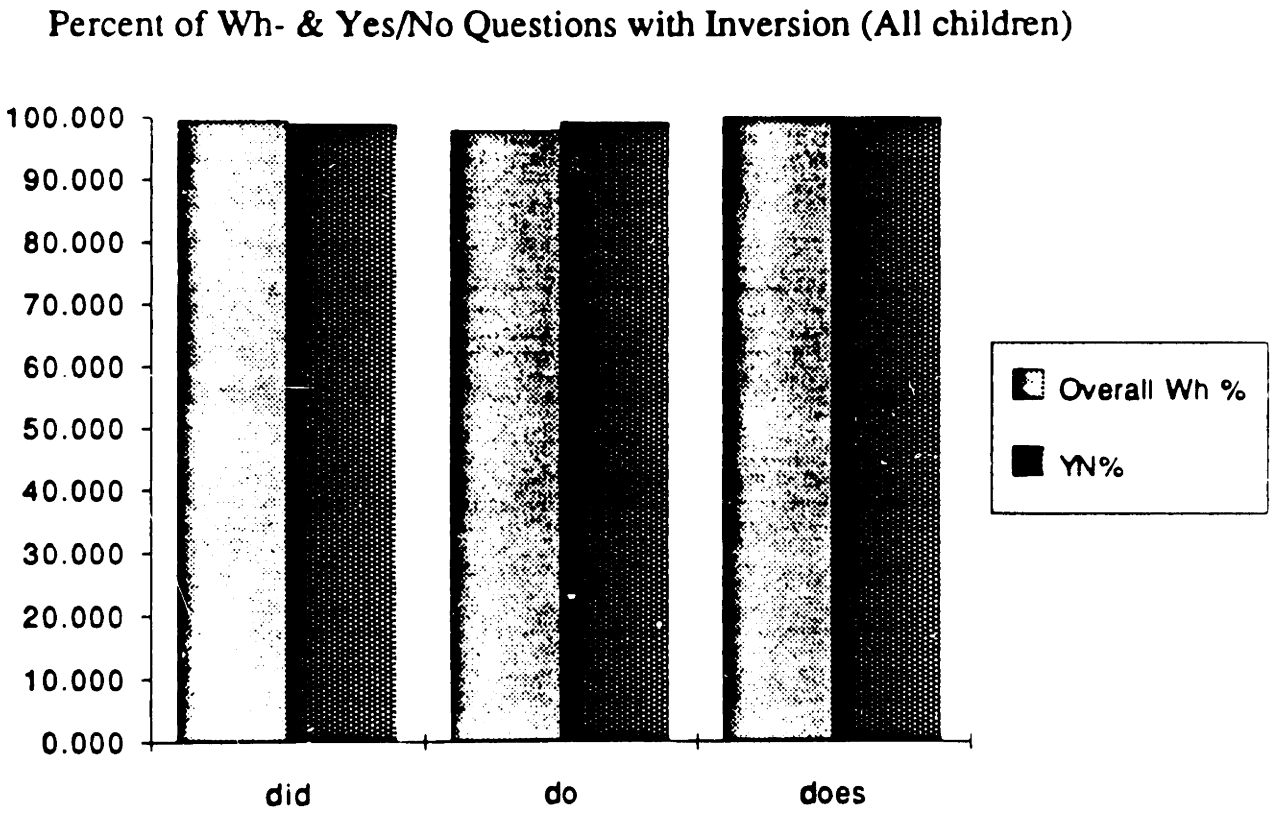
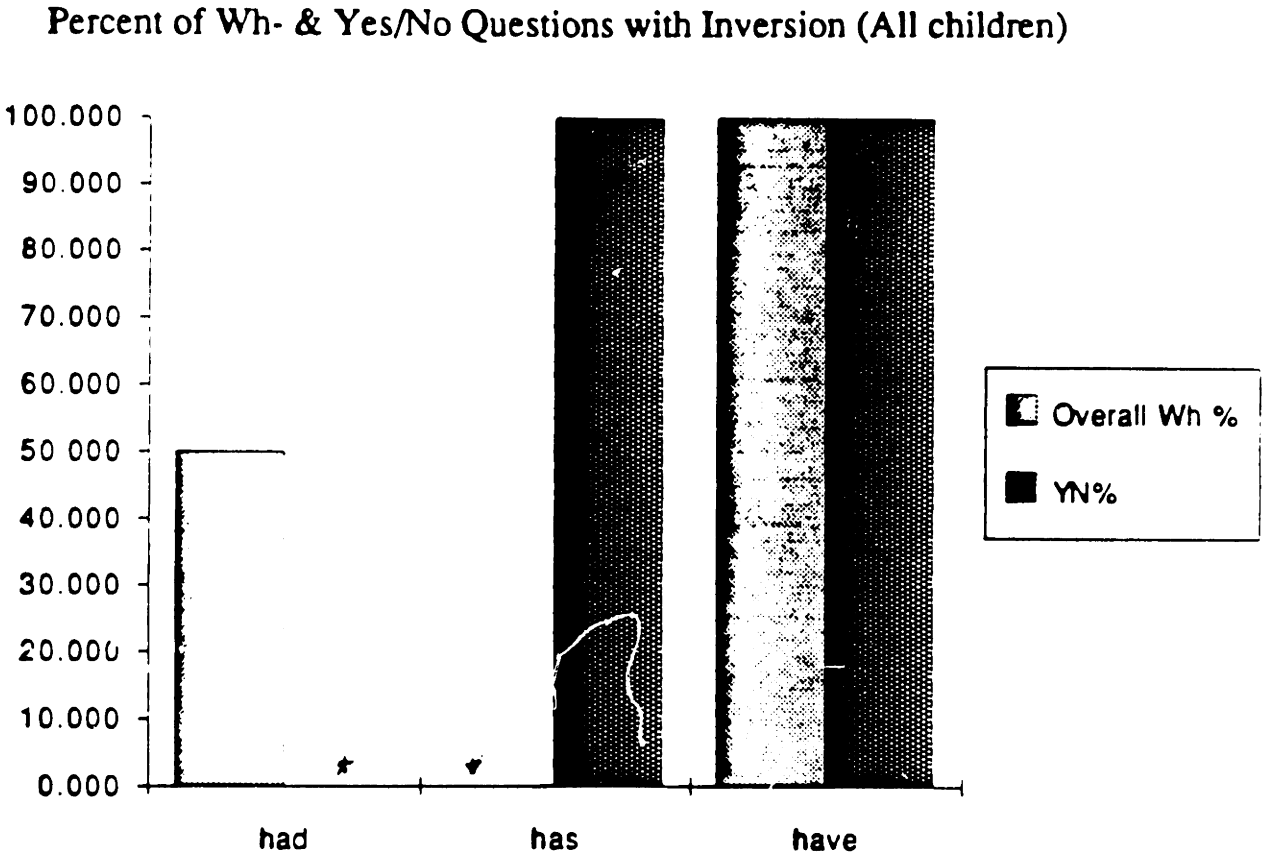
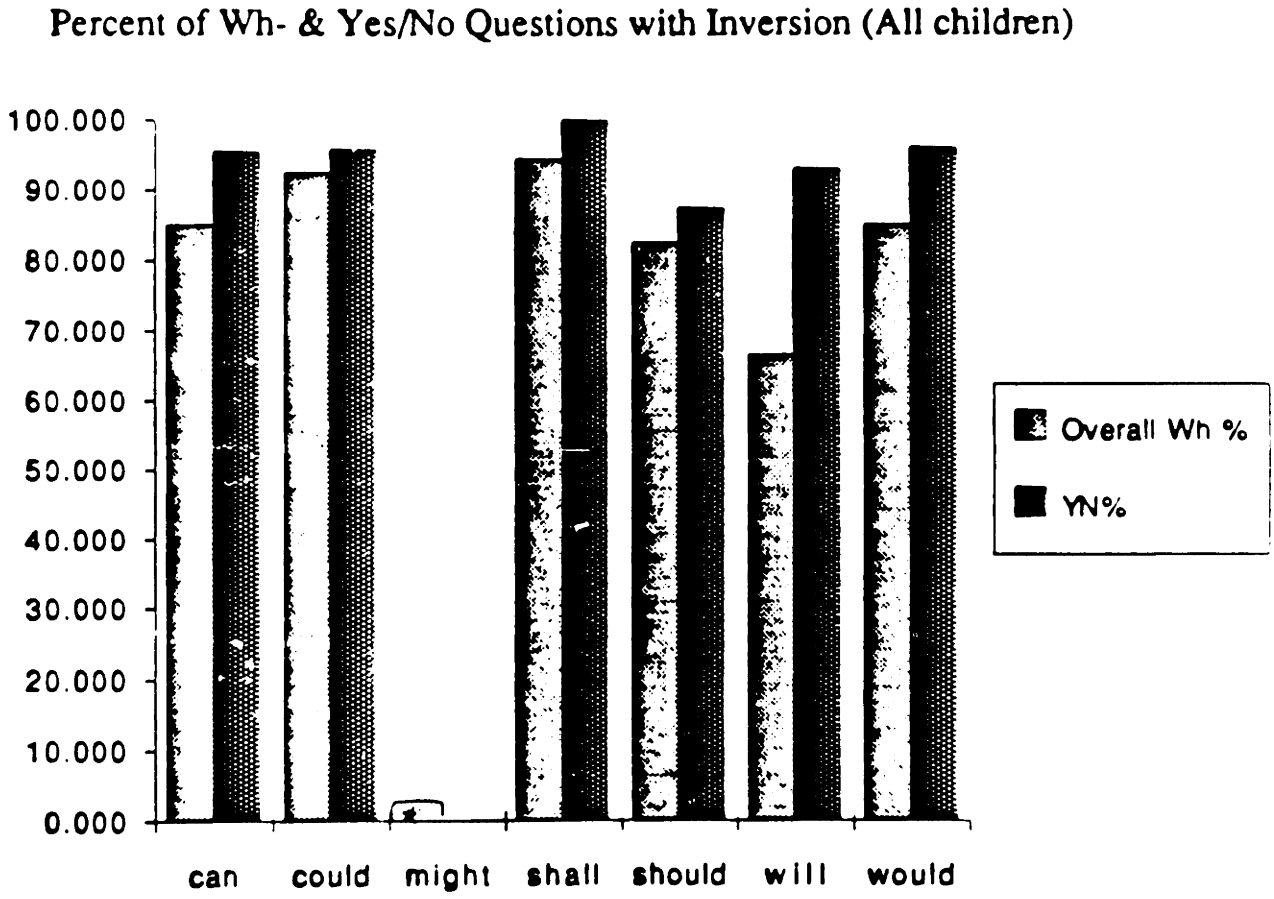


Figure 5.33
Overall Inversion Rates for Forms of Auxiliary *Have*



* = no examples (i.e., 0/0)

Figure 5.34
Overall Inversion Rates for Modal Auxiliaries



* = no examples (i.e., 0/0)

Figure 5.35
Overall Inversion Rates for *Be*, *Do*, *Have*, and Modal Auxiliaries

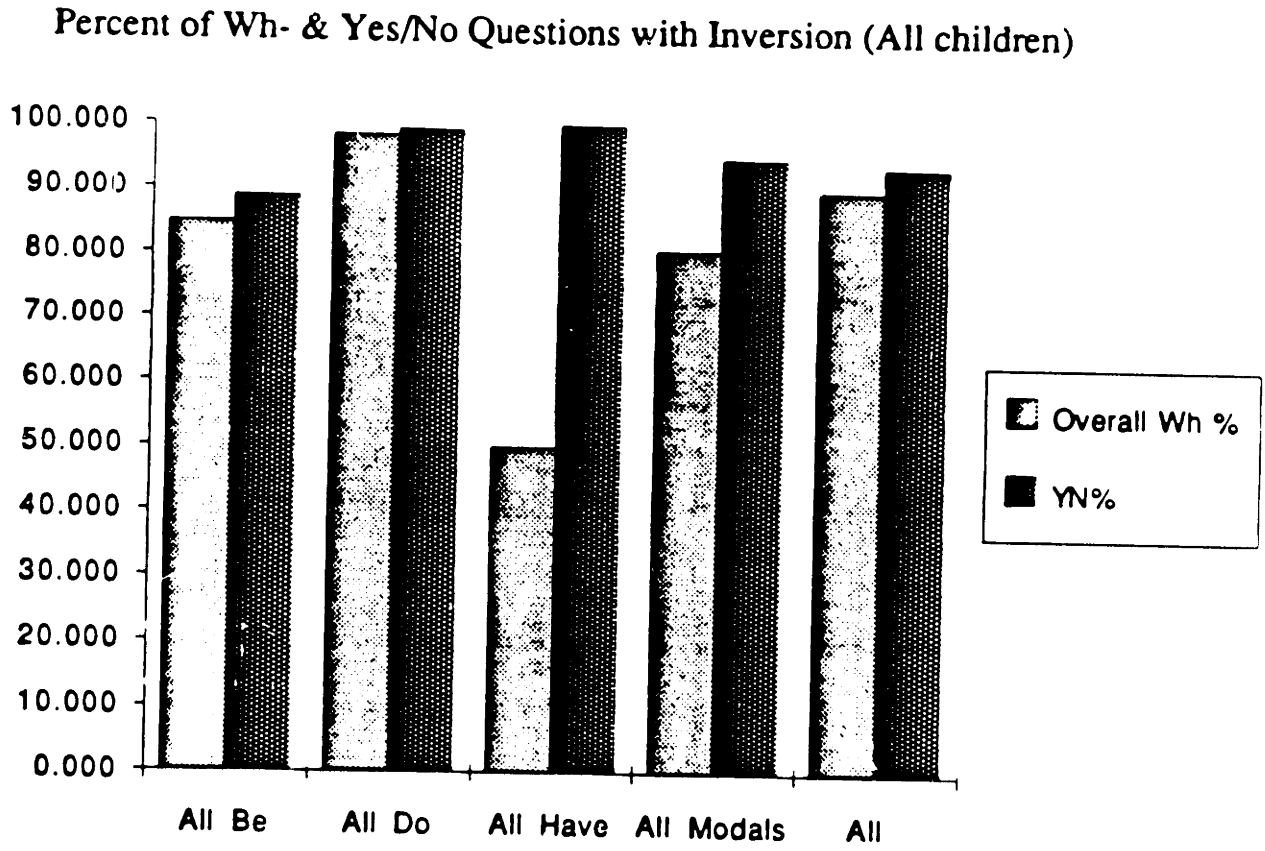
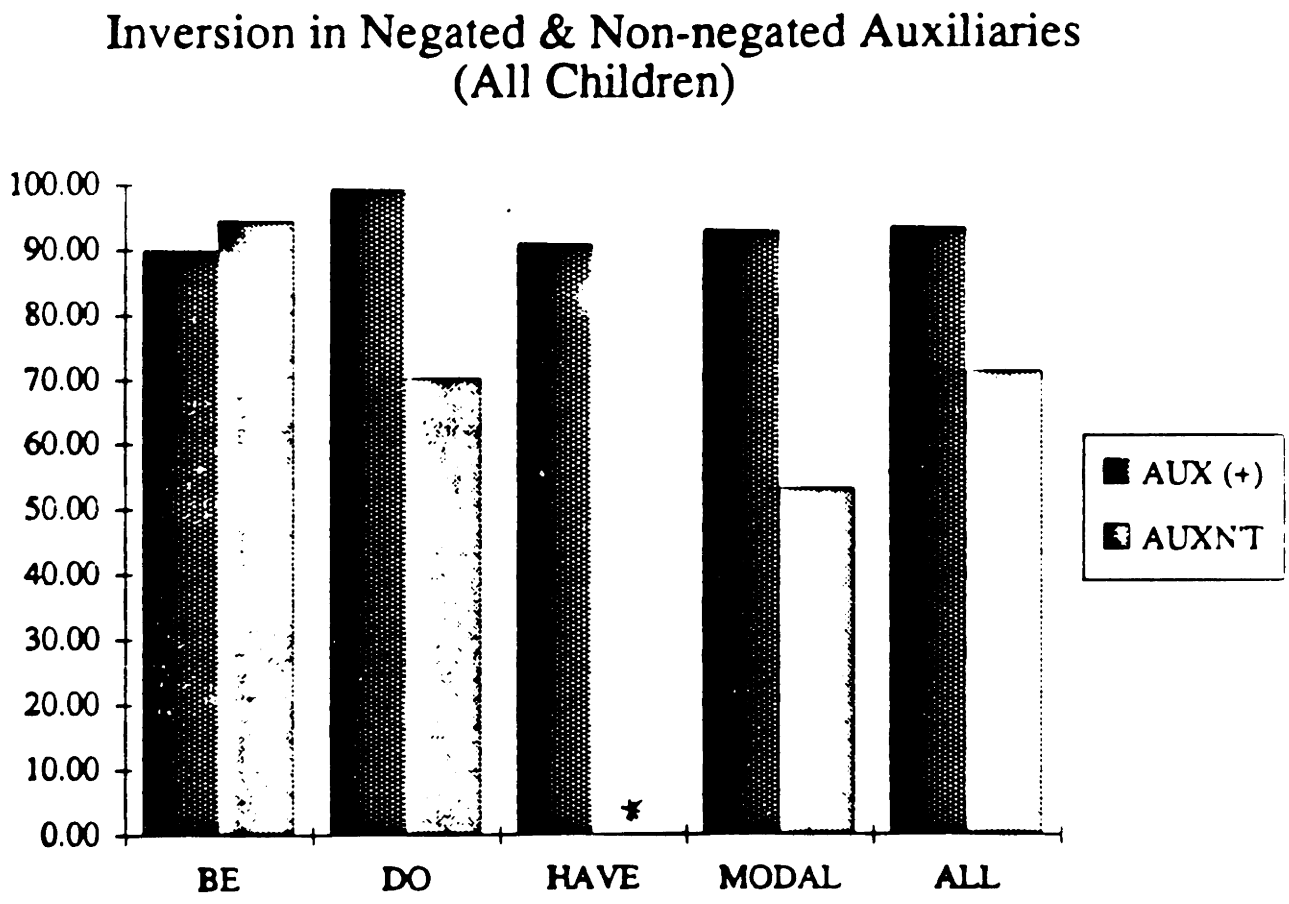


Figure 5.36
Overall Inversion Rates for Negated and Non-Negated Auxiliaries



* = no examples (i.e., 0/0)

Figure 5.37
Inversion Rates for Negated and Non-Negated Auxiliaries: Adam

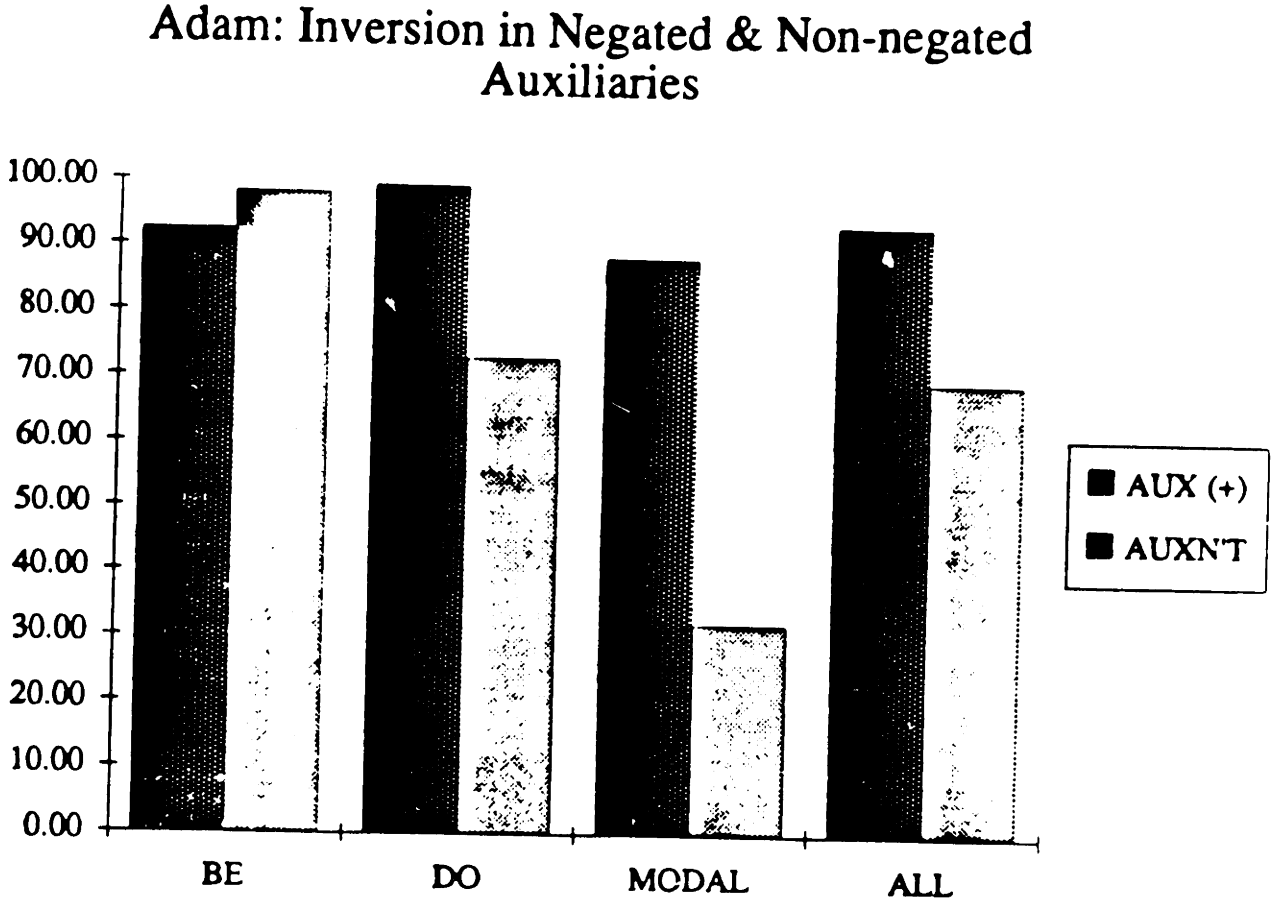
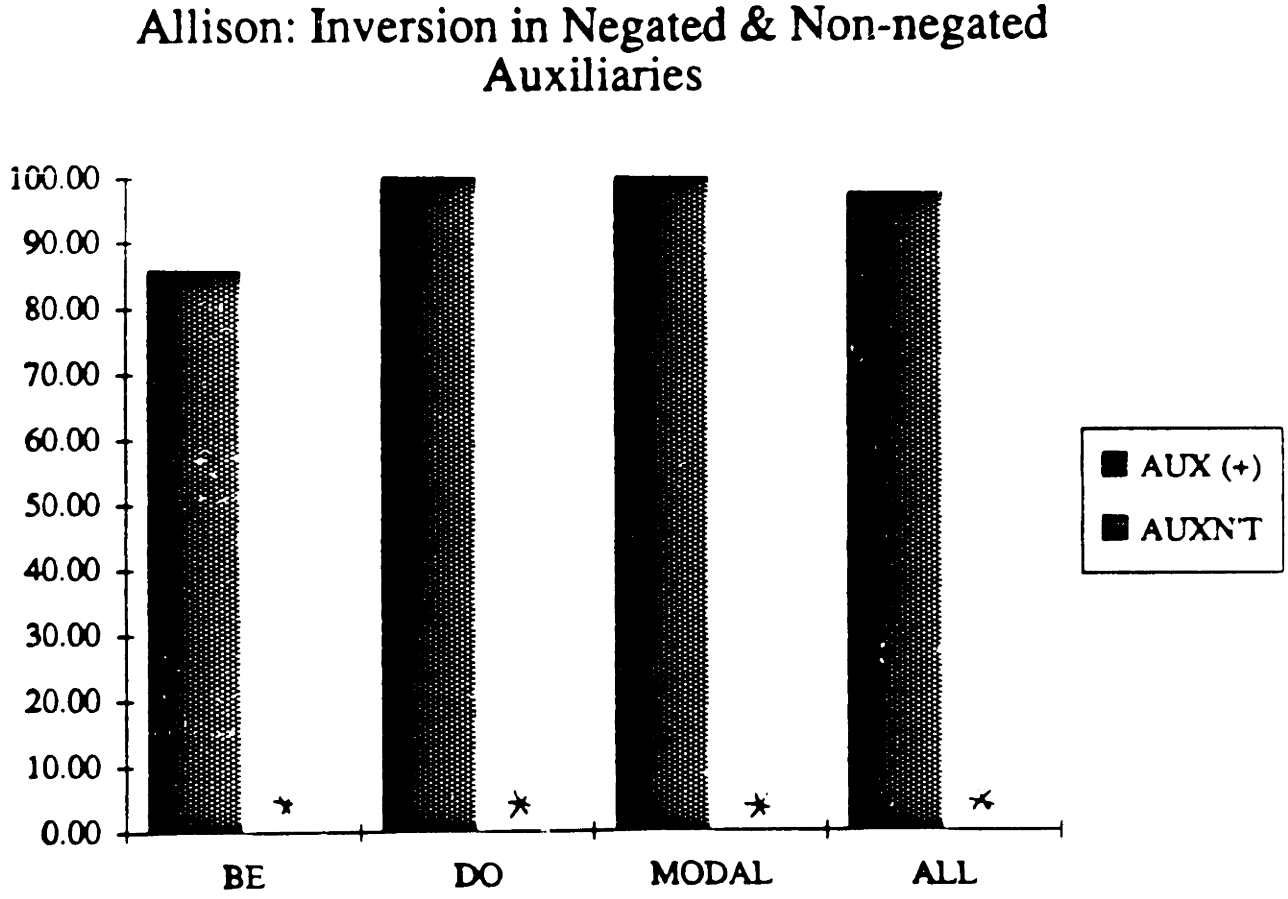
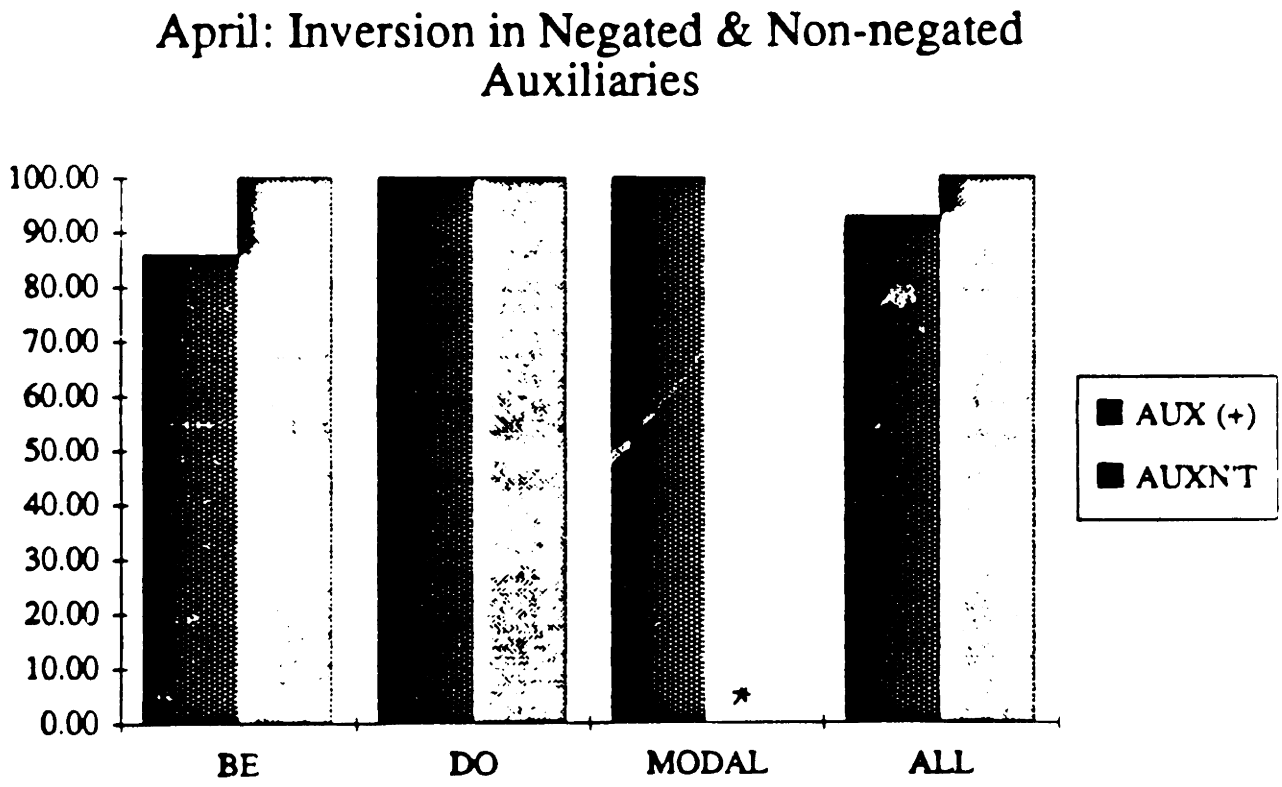


Figure 5.38
Inversion Rates for Negated and Non-Negated Auxiliaries: Allison



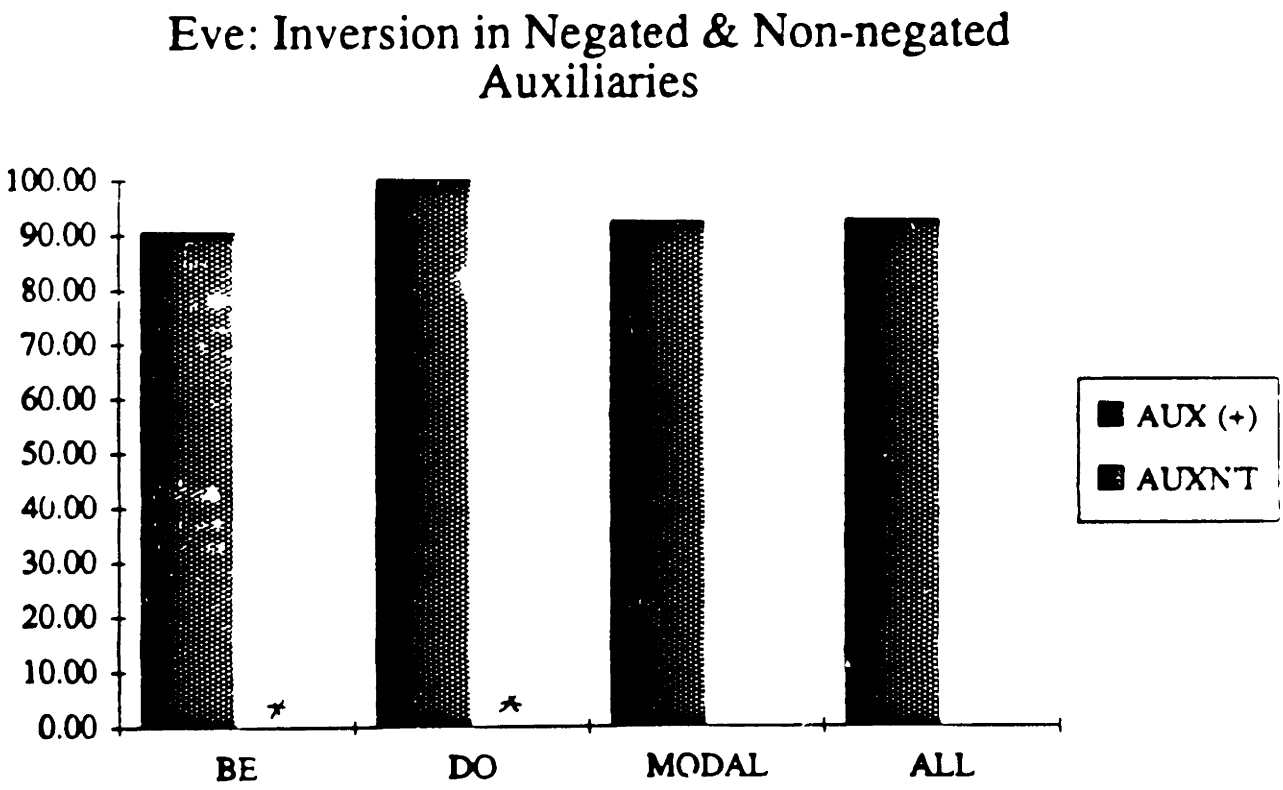
* = no examples (i.e., 0/0)

Figure 5.39
Inversion Rates for Negated and Non-Negated Auxiliaries: April



* = no examples (i.e., 0/0)

Figure 5.40
Inversion Rates for Negated and Non-Negated Auxiliaries: Eve



* = no examples (i.e., 0/0)

Figure 5.41
Inversion Rates for Negated and Non-Negated Auxiliaries: Mark

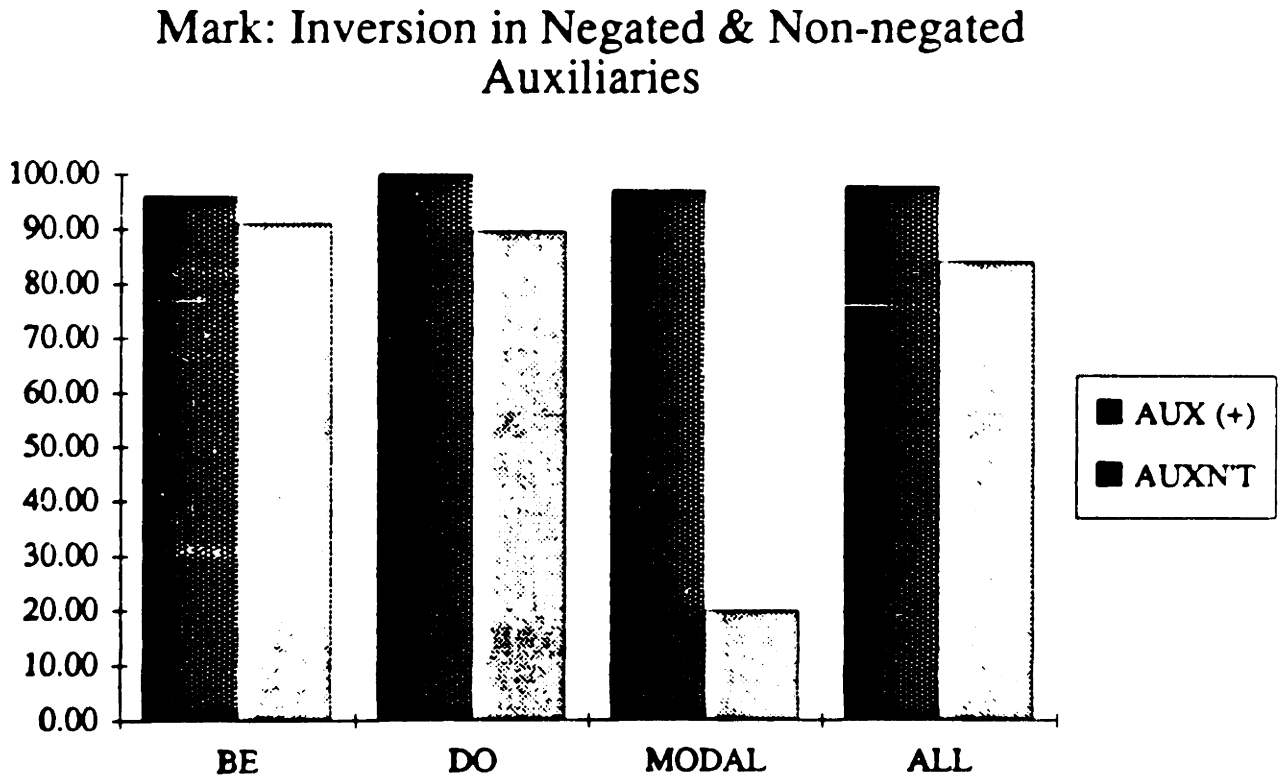


Figure 5.42
Inversion Rates for Negated and Non-Negated Auxiliaries: Naomi

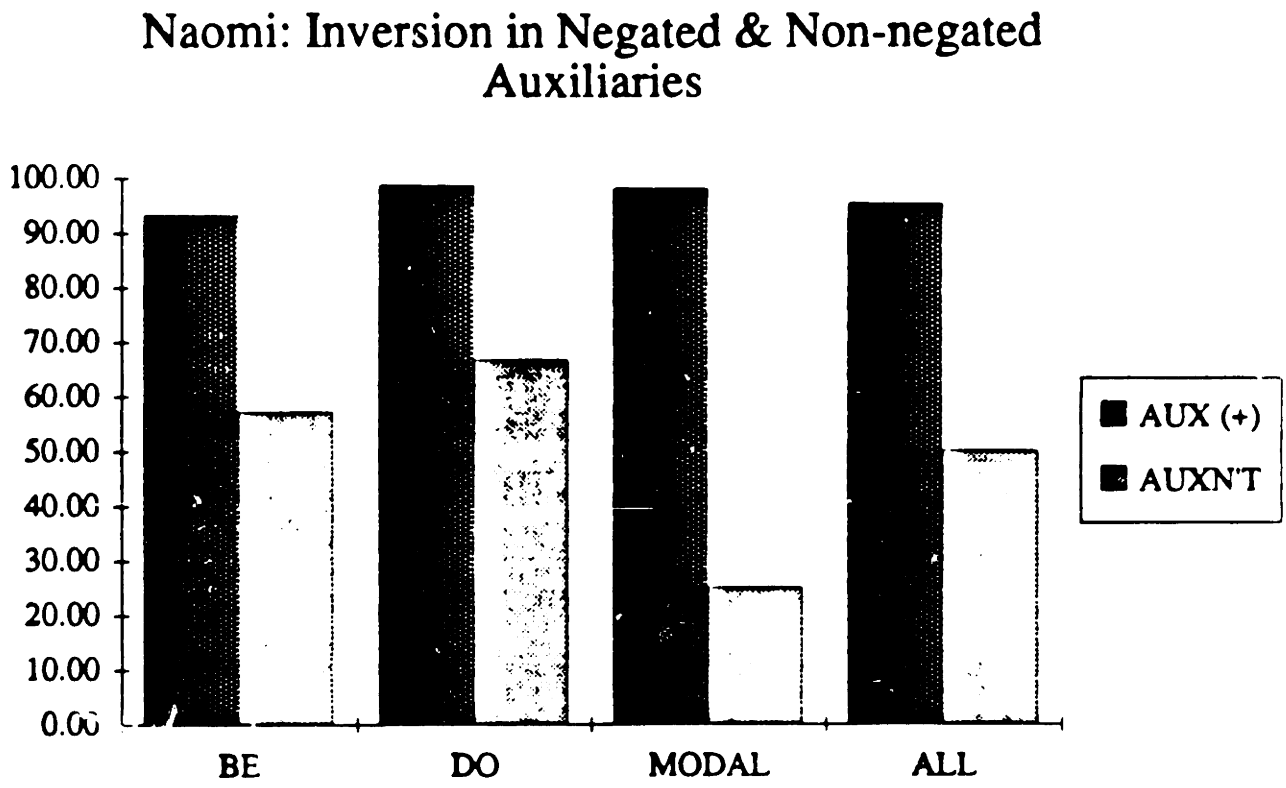
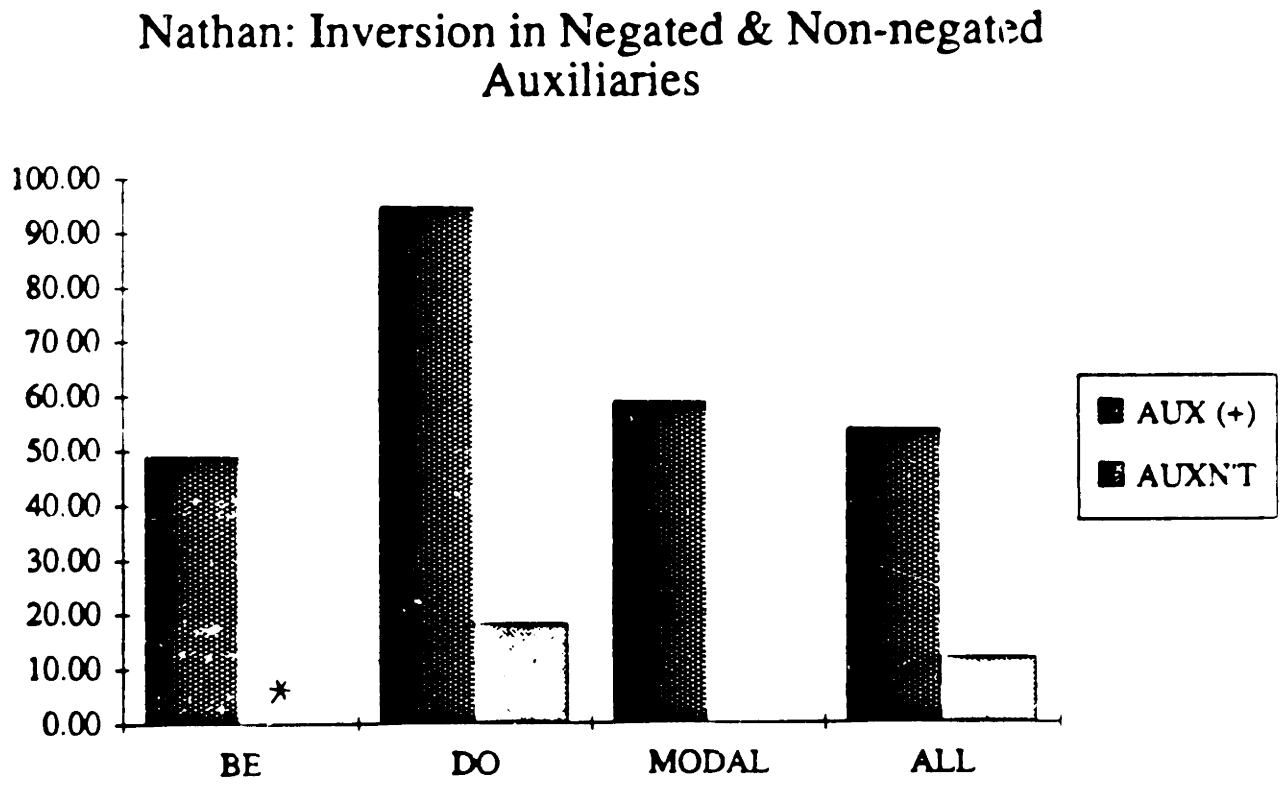


Figure 5.43
Inversion Rates for Negated and Non-Negated Auxiliaries: Nathan



* = no examples (i.e., 0/0)

Figure 5.44
Inversion Rates for Negated and Non-Negated Auxiliaries: Nina

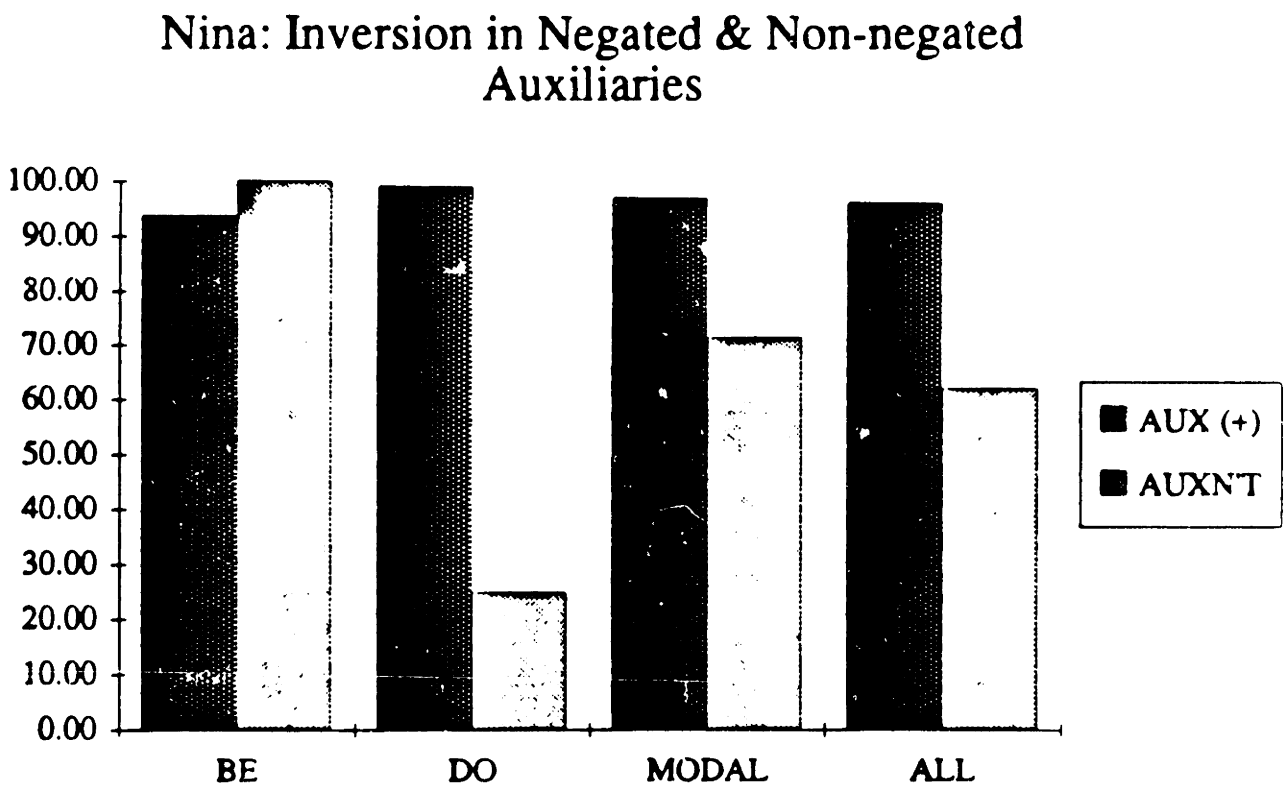


Figure 5.45
Inversion Rates for Negated and Non-Negated Auxiliaries: Peter

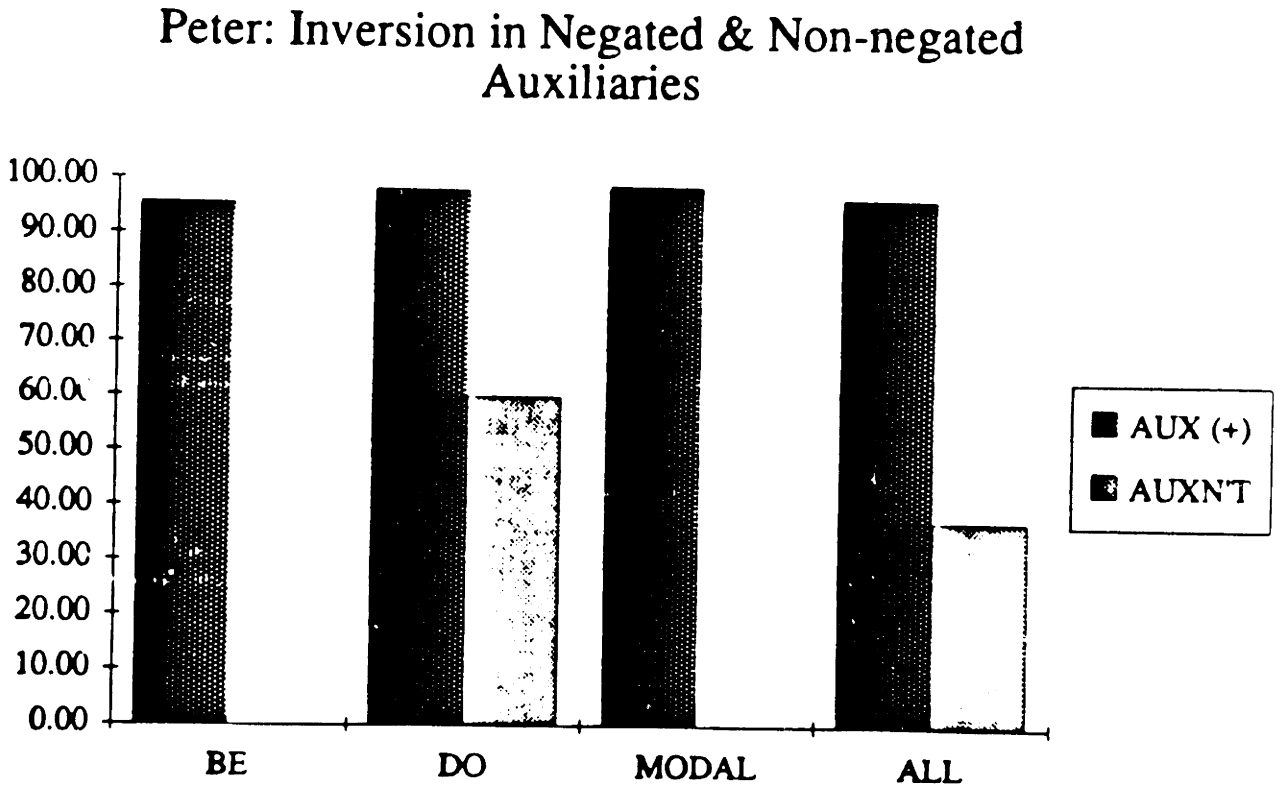


Figure 5.46
Inversion Rates for Negated and Non-Negated Auxiliaries: Ross

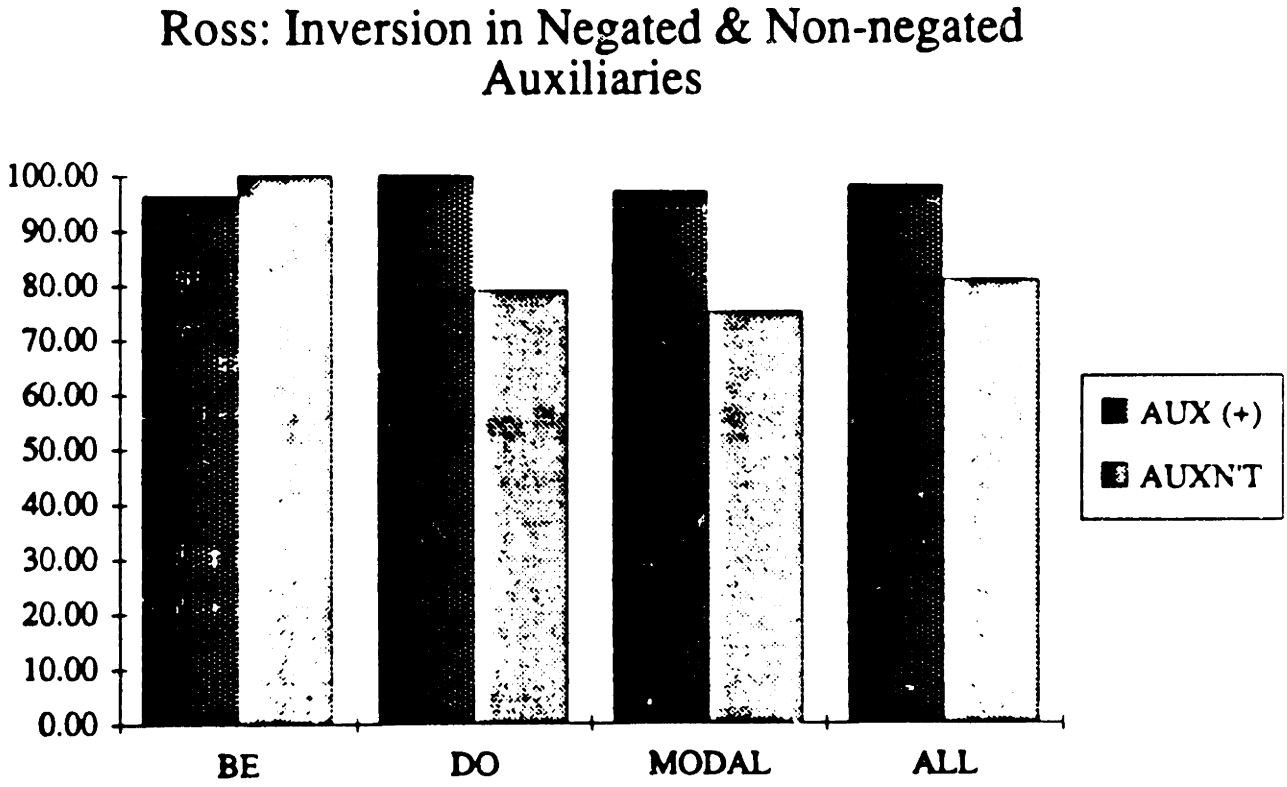


Figure 5.47
Inversion Rates for Negated and Non-Negated Auxiliaries: Sarah

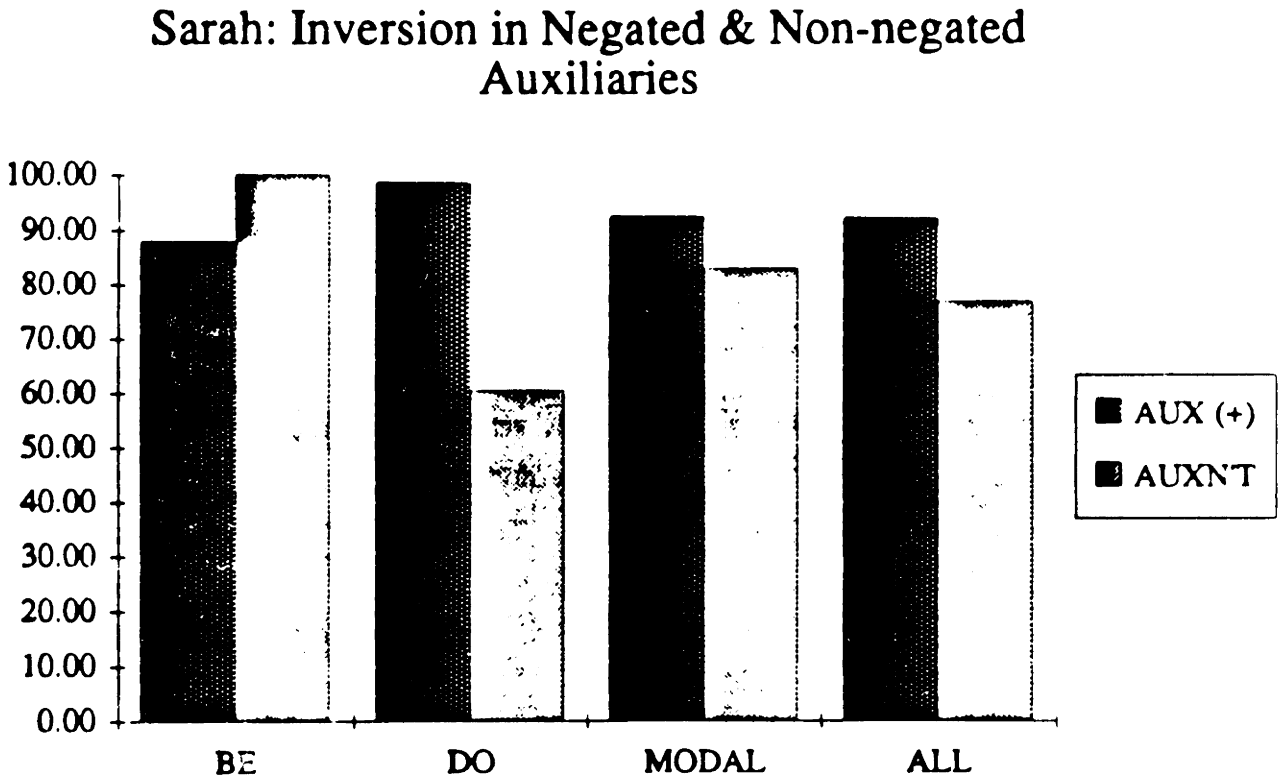
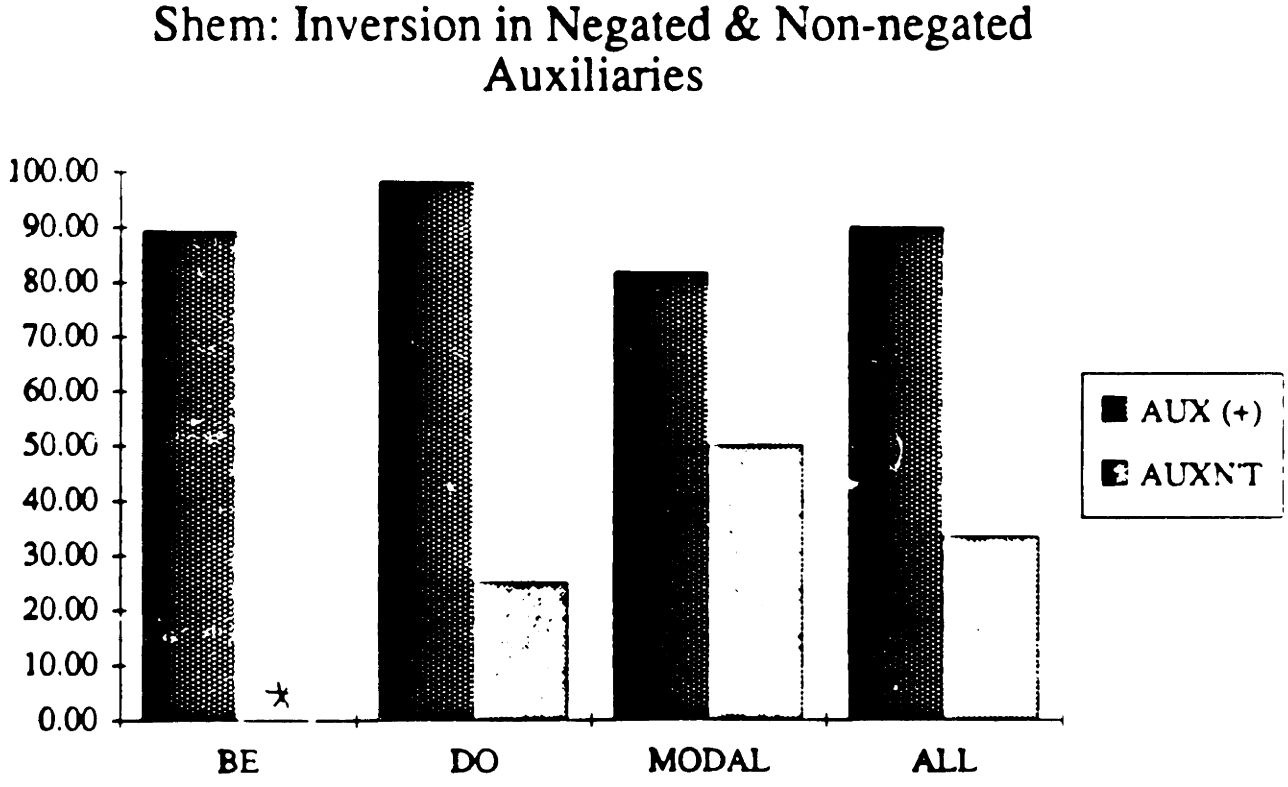


Figure 5.48
Inversion Rates for Negated and Non-Negated Auxiliaries: Shem



* = no examples (i.e., 0/0)

Figure 5.49
Inversion Study: Percent of *Wh*-Questions Judged Grammatical

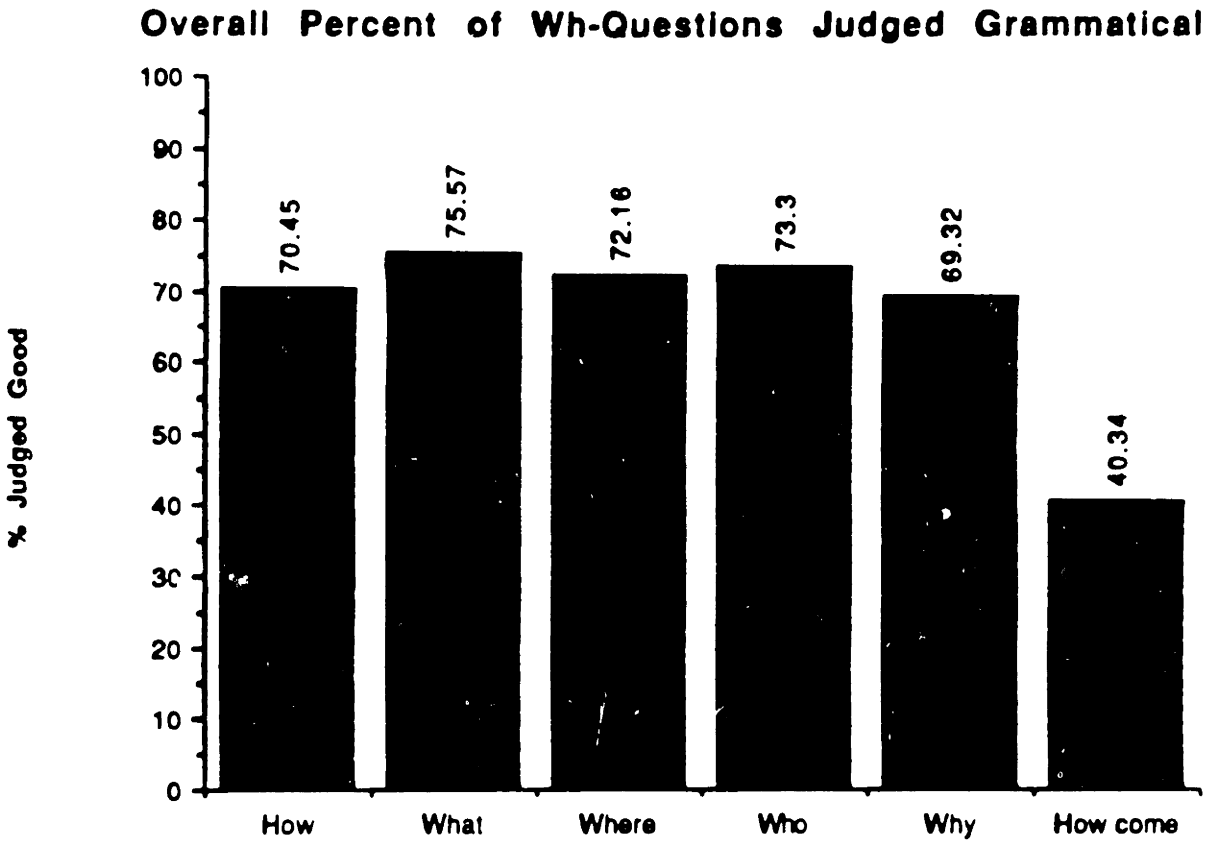


Figure 5.50
Interaction Between MLU and Matrix/Embedded Question

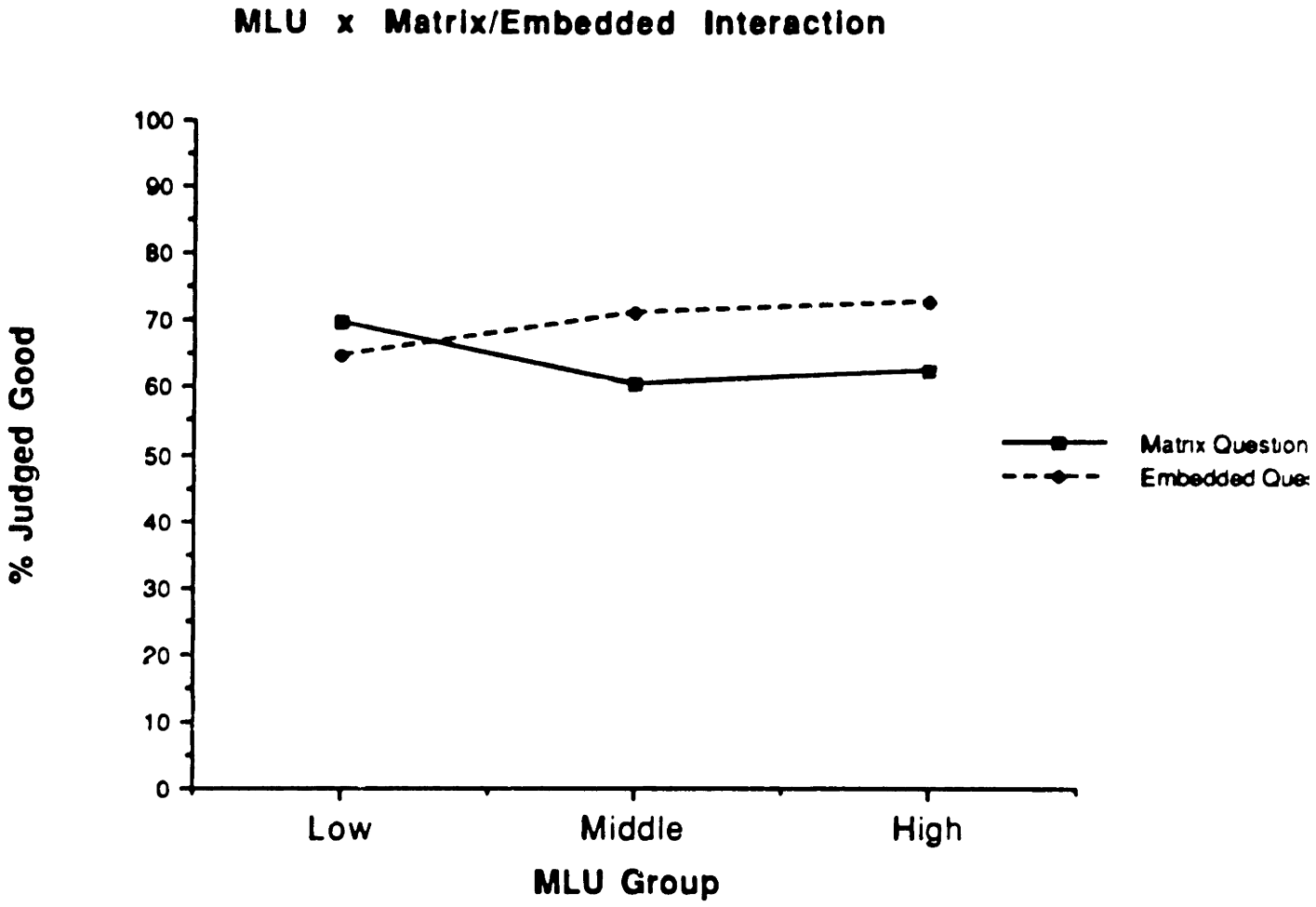
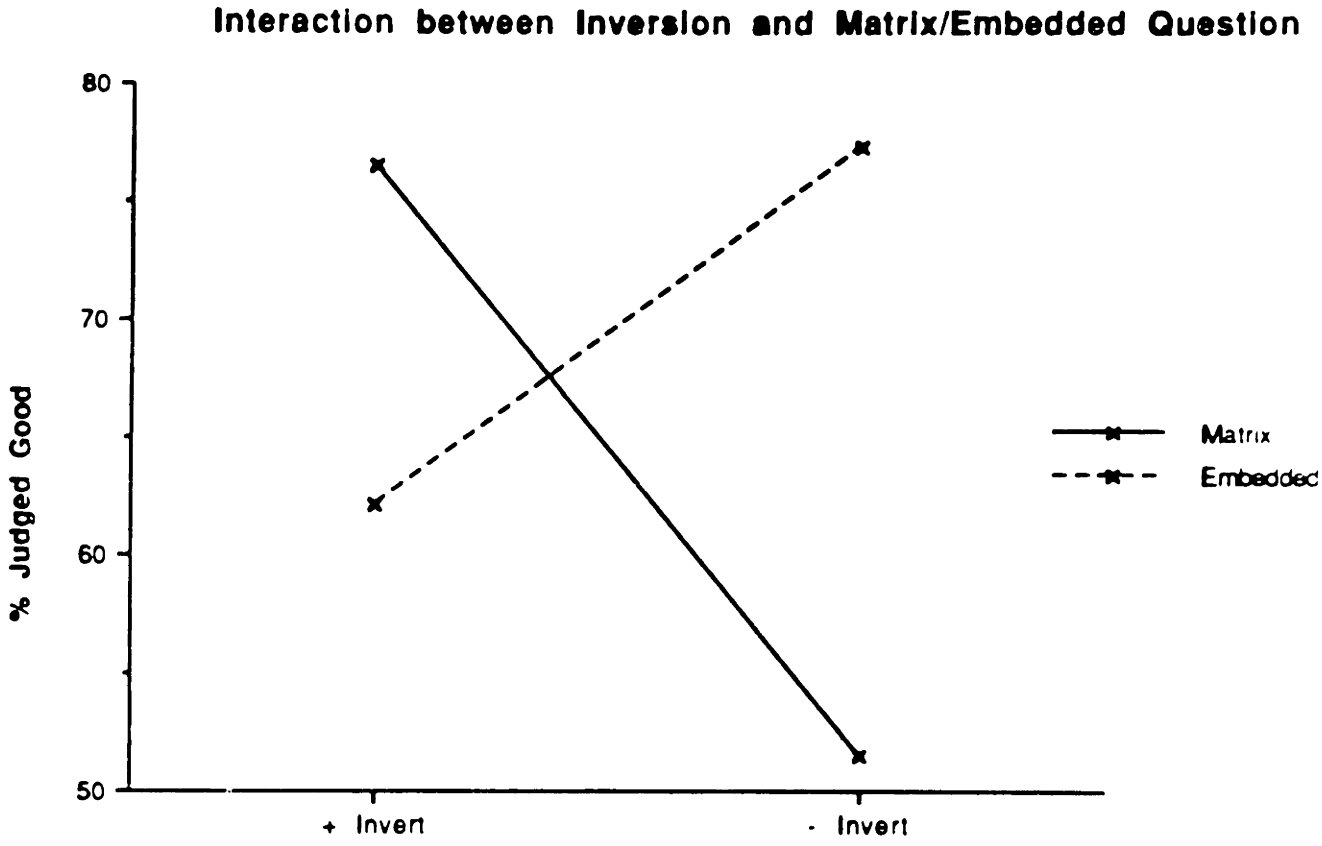


Figure 5.51
Interaction Between Inversion and Matrix/Embedded Question



F(1, 19) = 35.52, p < .0005

Figure 5.52
Interaction Among MLU, Inversion, and Matrix/Embedded Question (All *Wh*-Questions)

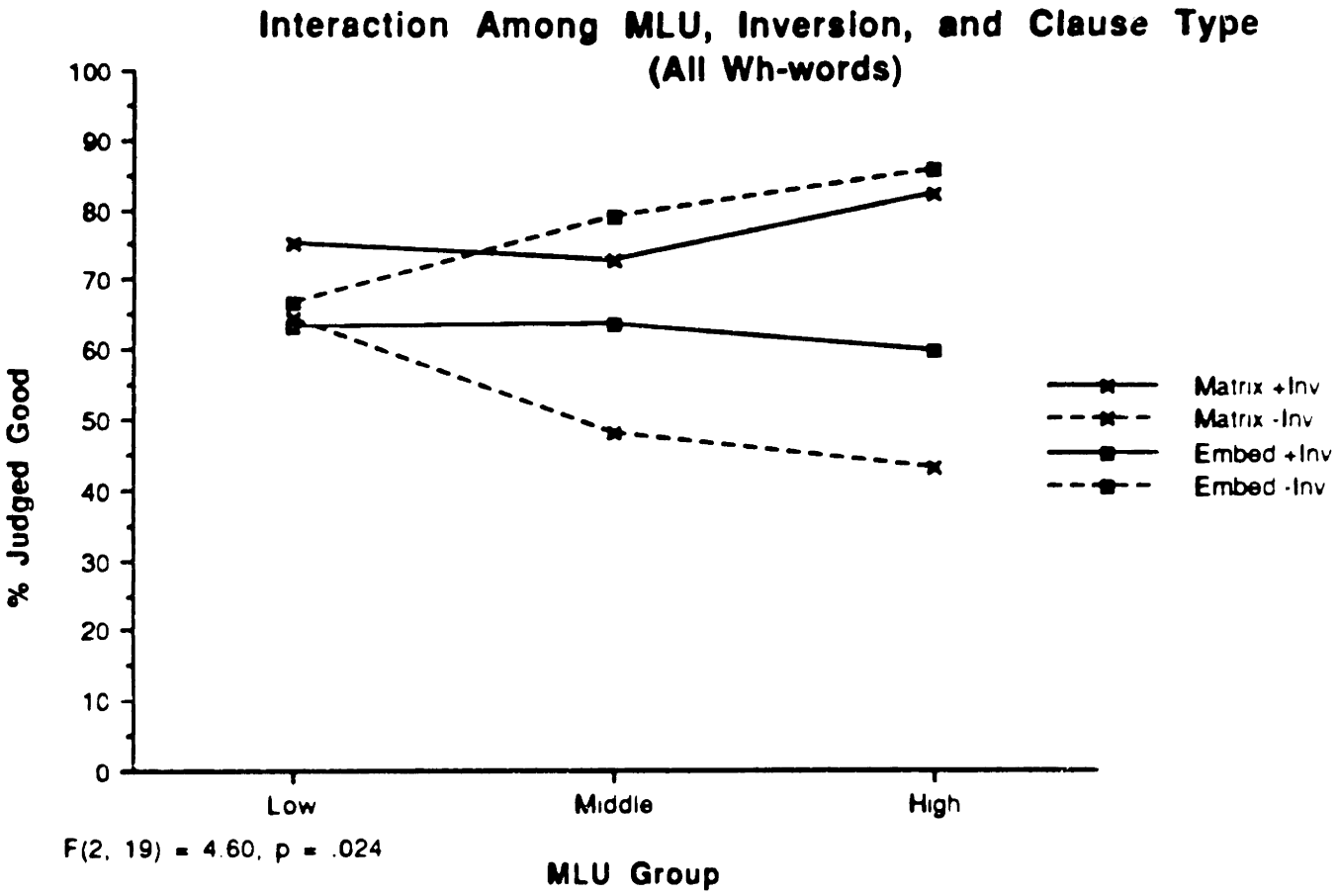


Figure 5.53
Interaction Among MLU, Inversion, and Matrix/Embedded Question (All *Wh*-Questions except *How come* Questions)

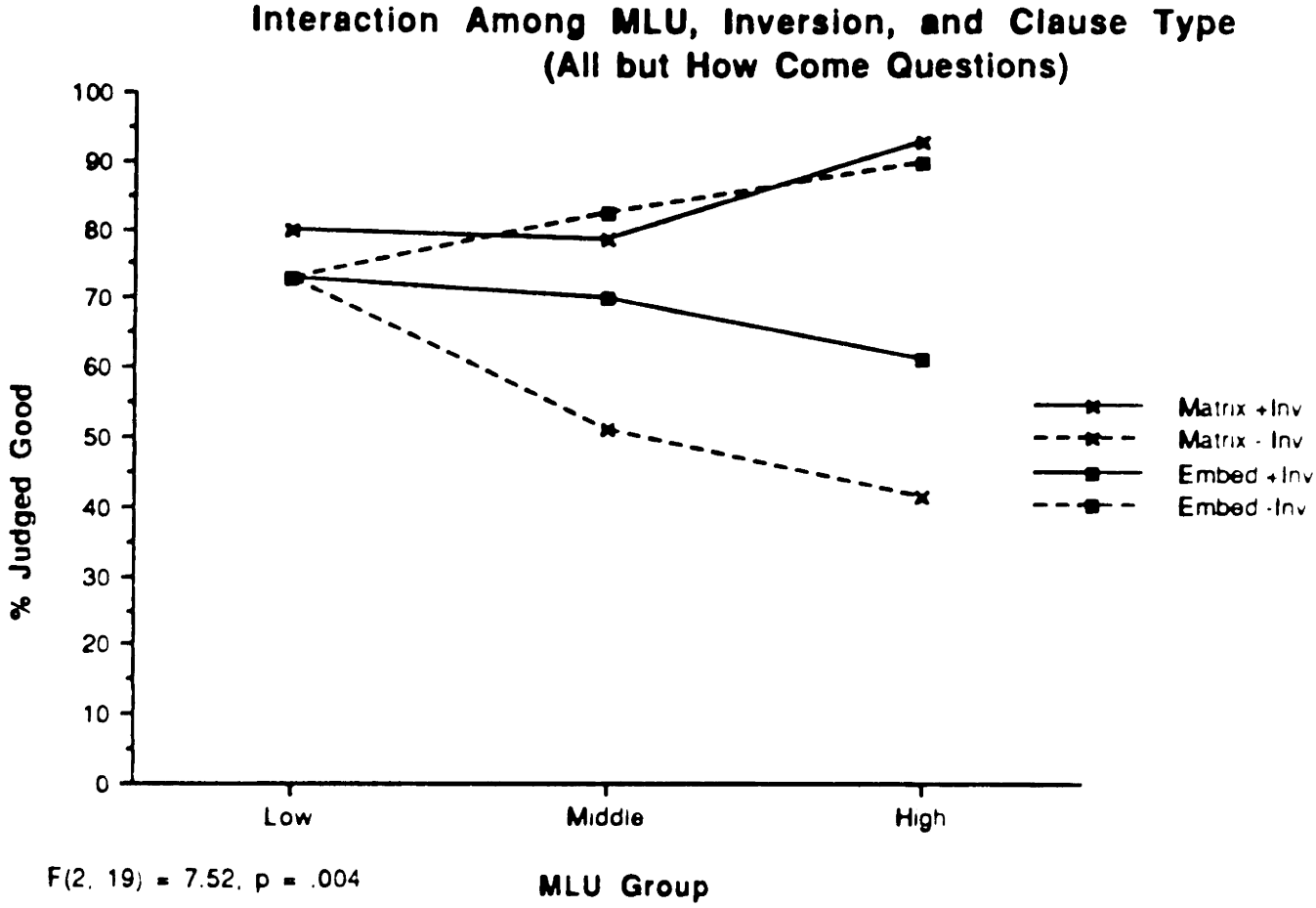
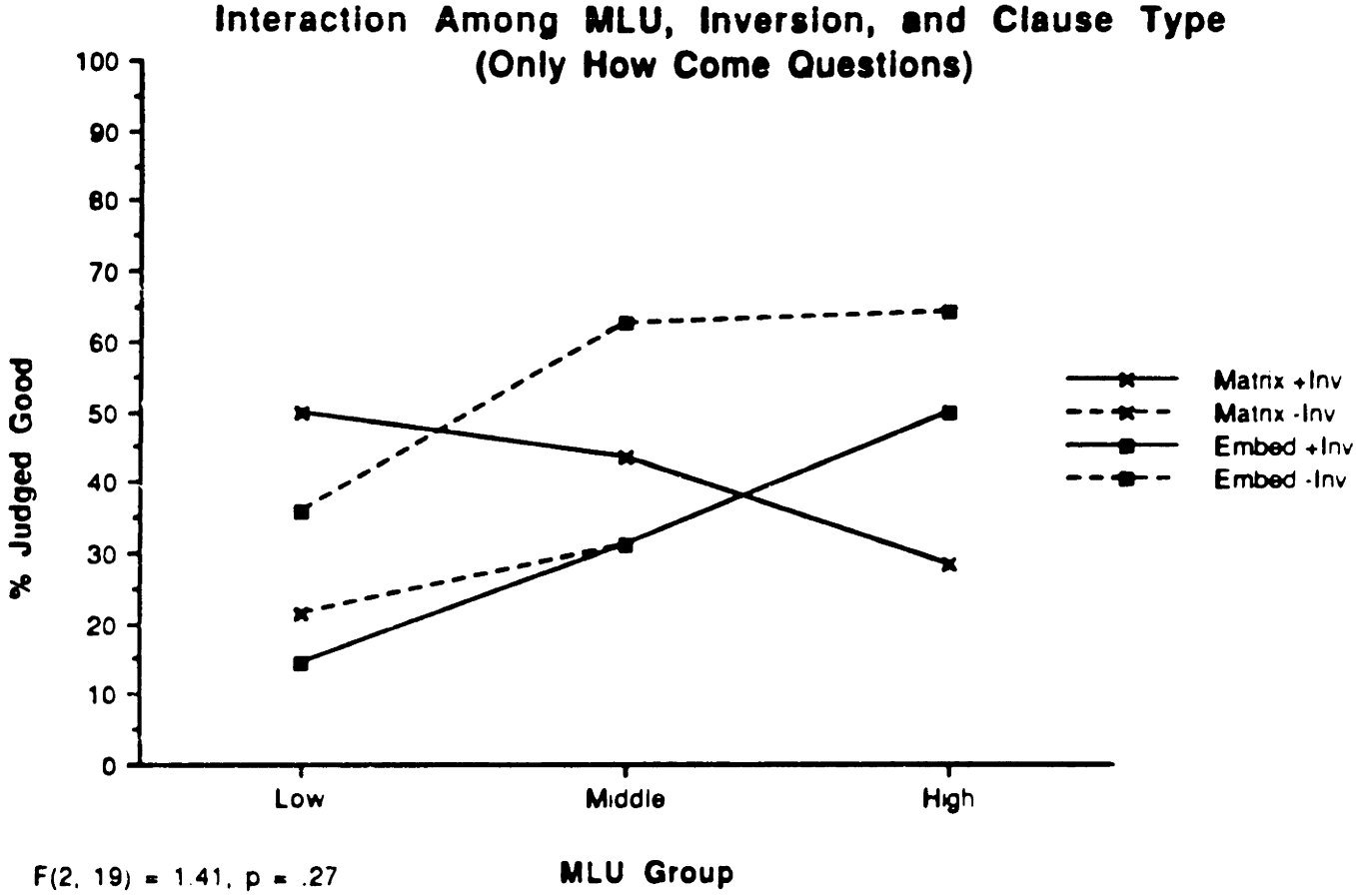


Figure 5.54
Interaction Among MLU, Inversion, and Matrix/Embedded Question (Only *How come* Questions)



F(2, 19) = 1.41, p = .27

Figure 5.55
Developmental Trend: Argument vs. Adjunct *Wh*- Questions

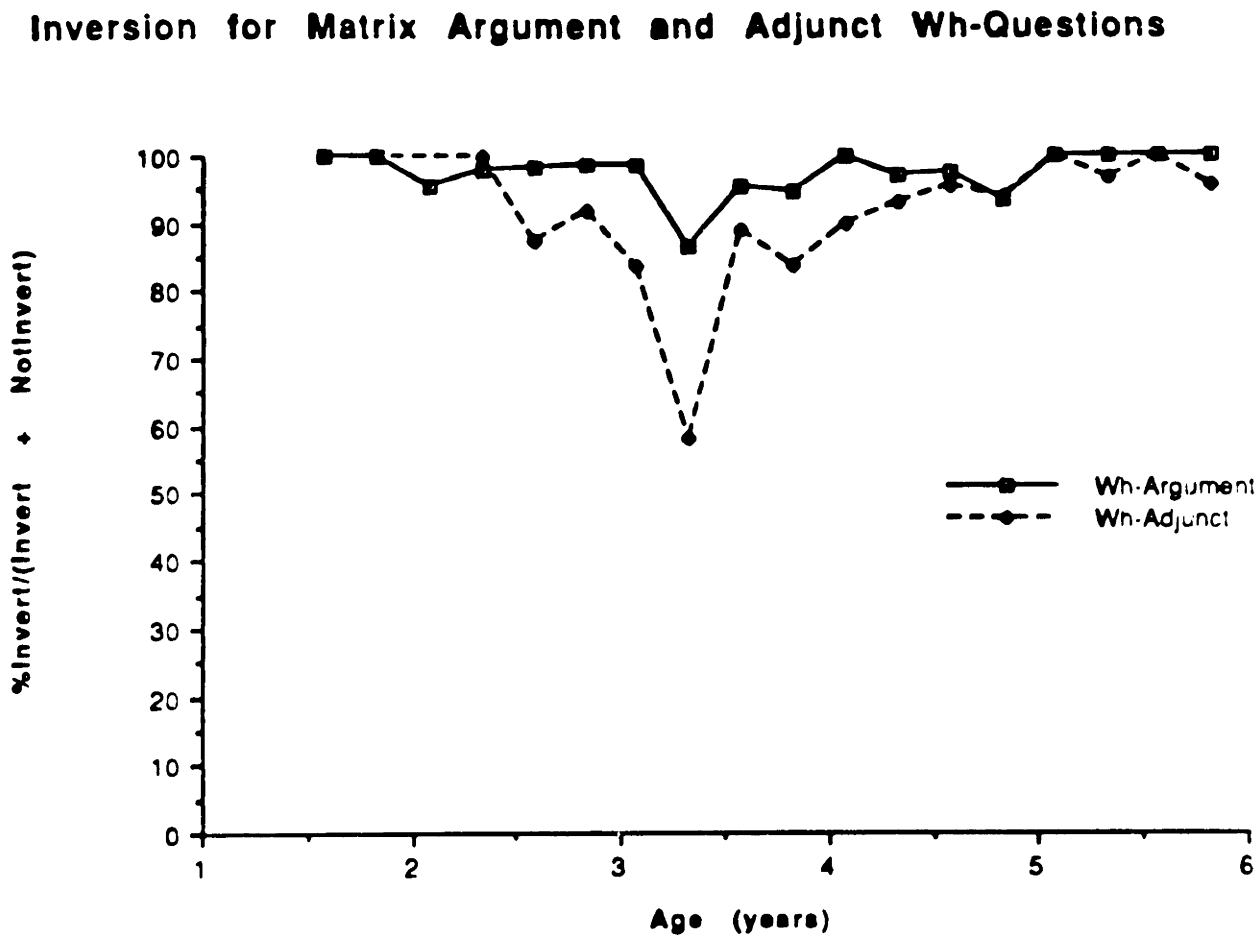


Figure 5.56
Developmental Trend: Argument vs. Adjunct *Where* Questions

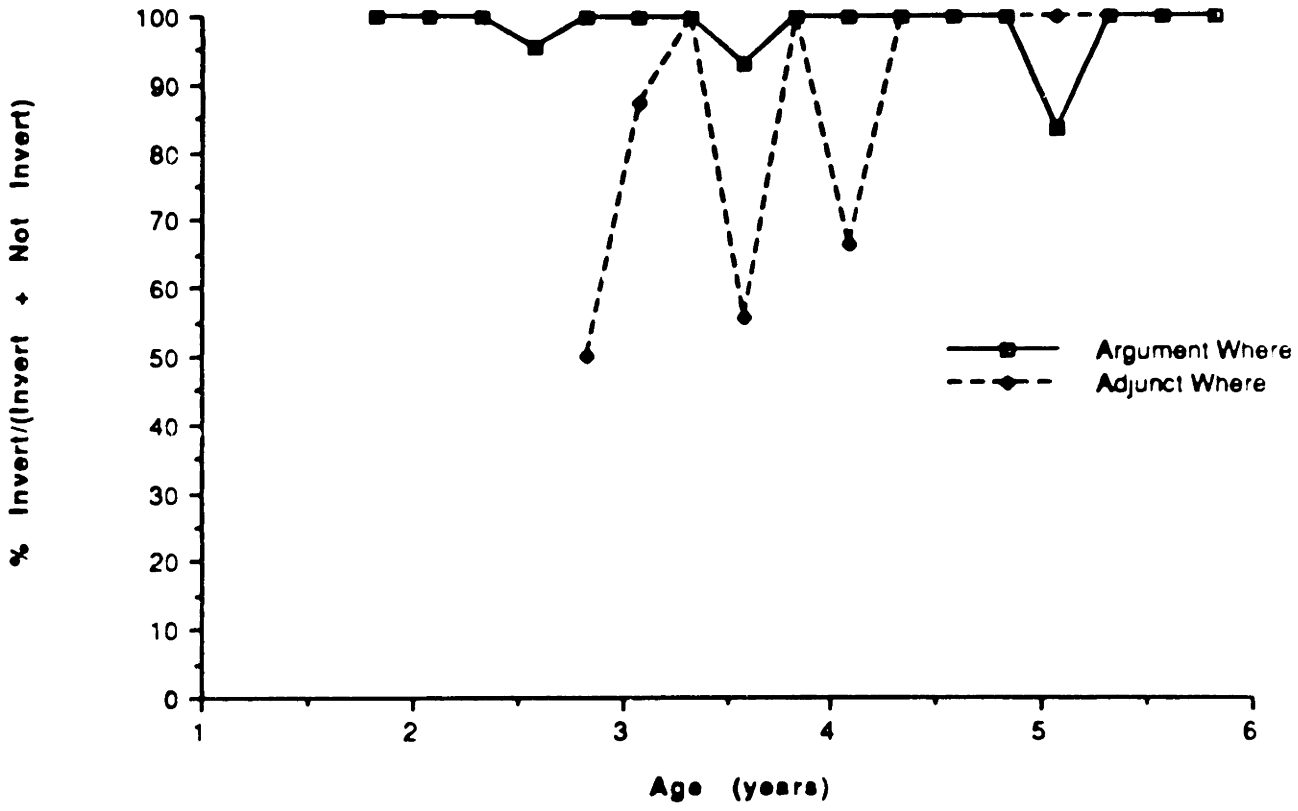


Figure 6.1
Aux-less *Yes/No* Questions with Tensed Main Verbs: All Children

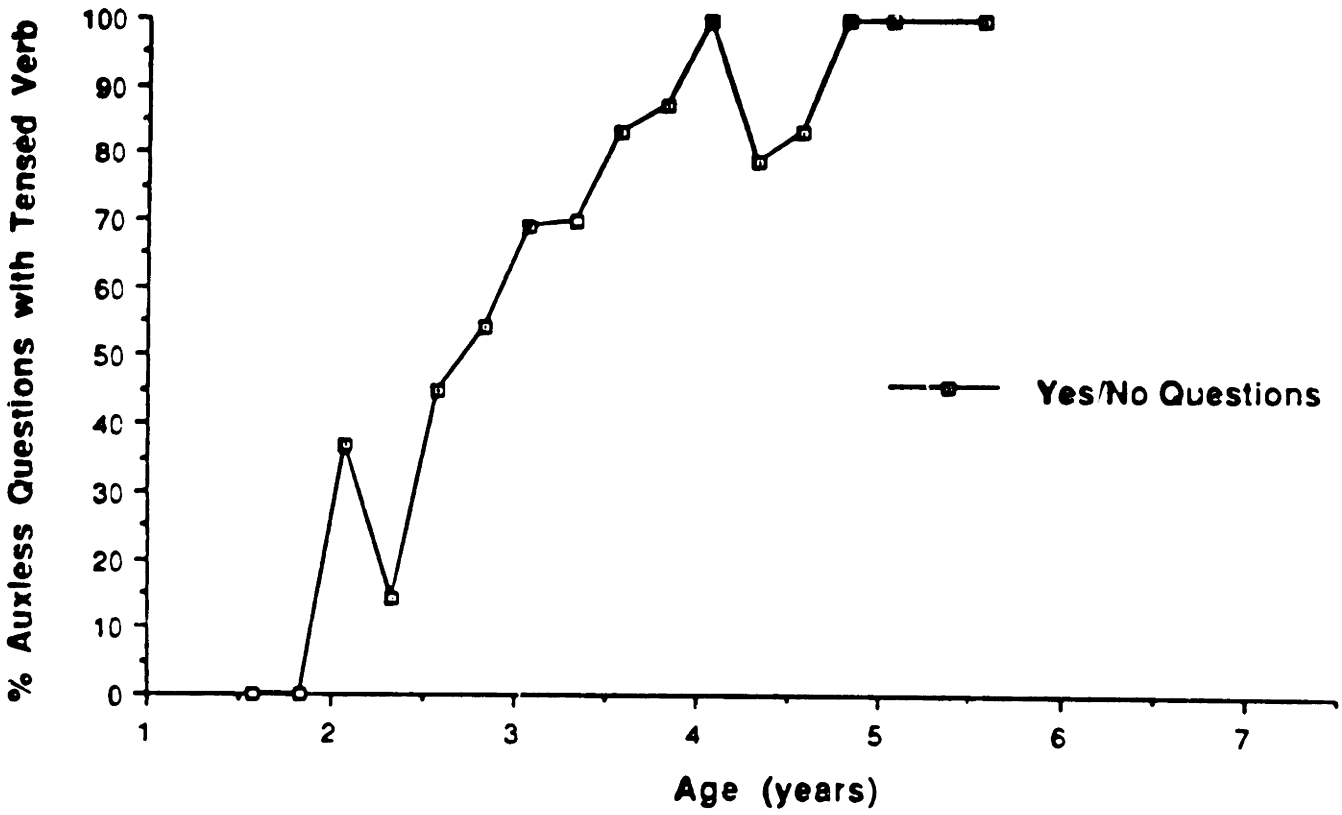


Figure 6.2
Aux-less *Yes/No* Questions with Tensed Main Verbs: Adam

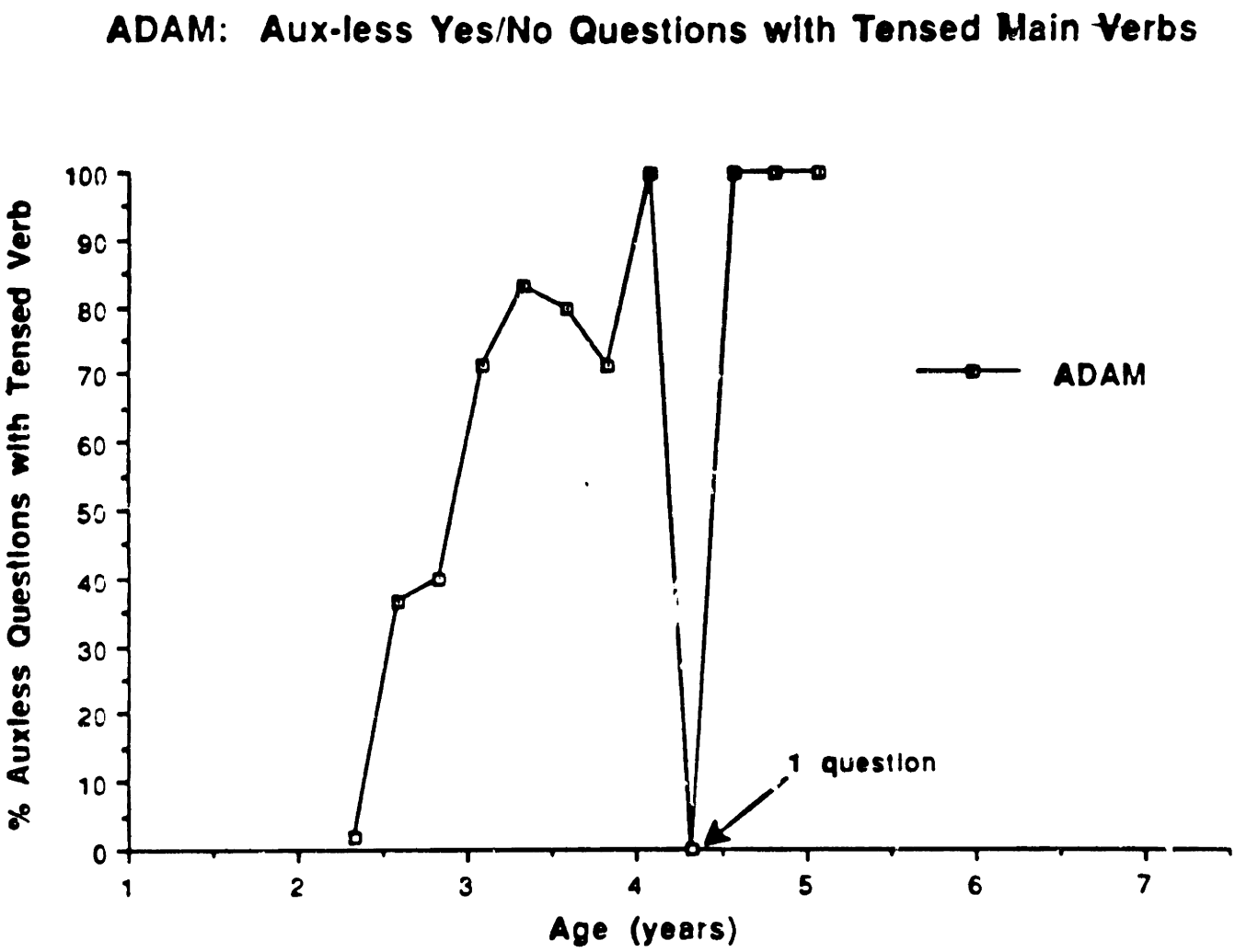


Figure 6.3
Aux-less *Yes/No* Questions with Tensed Main Verbs: April

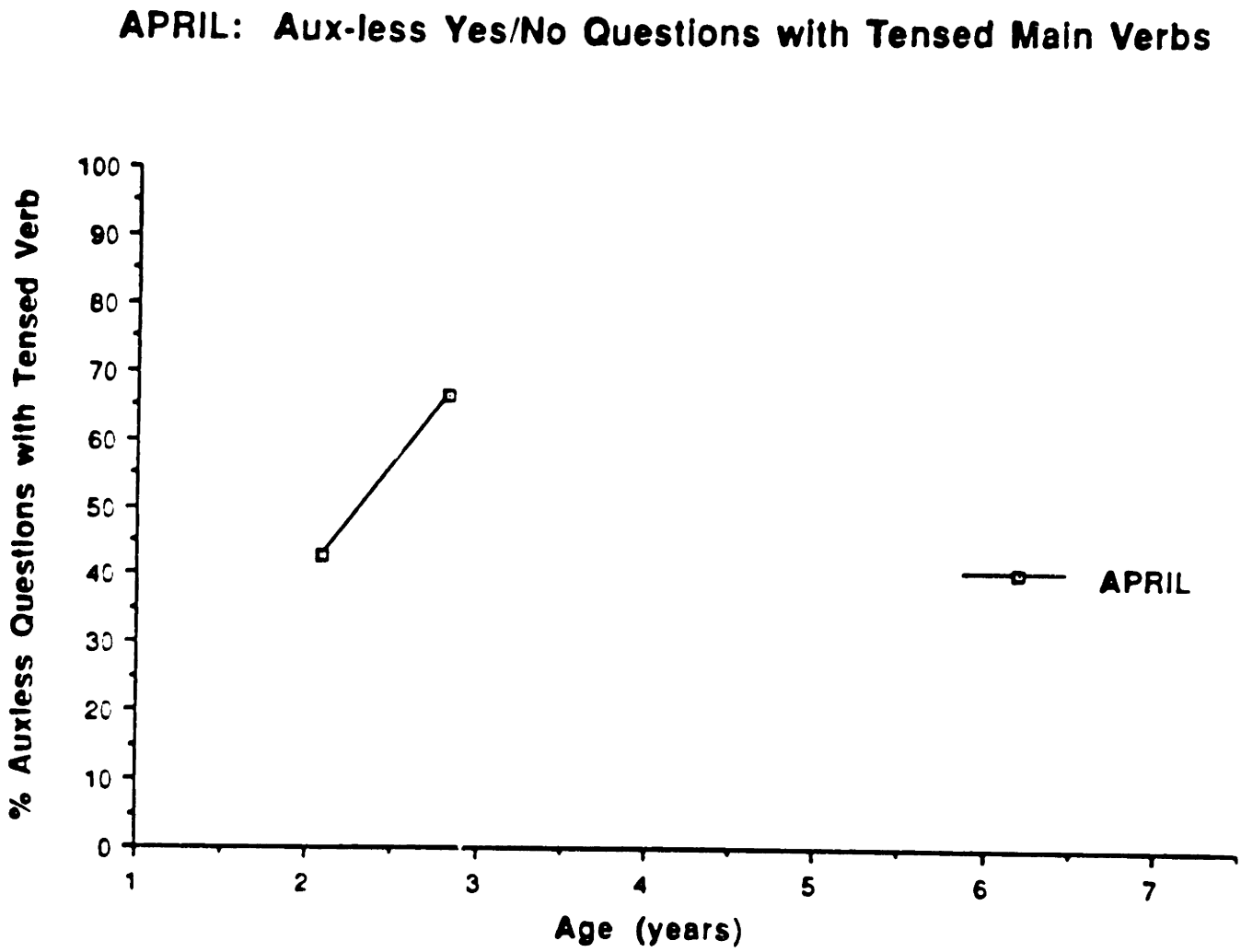


Figure 6.4
Aux-less *Yes/No* Questions with Tensed Main Verbs: Eve

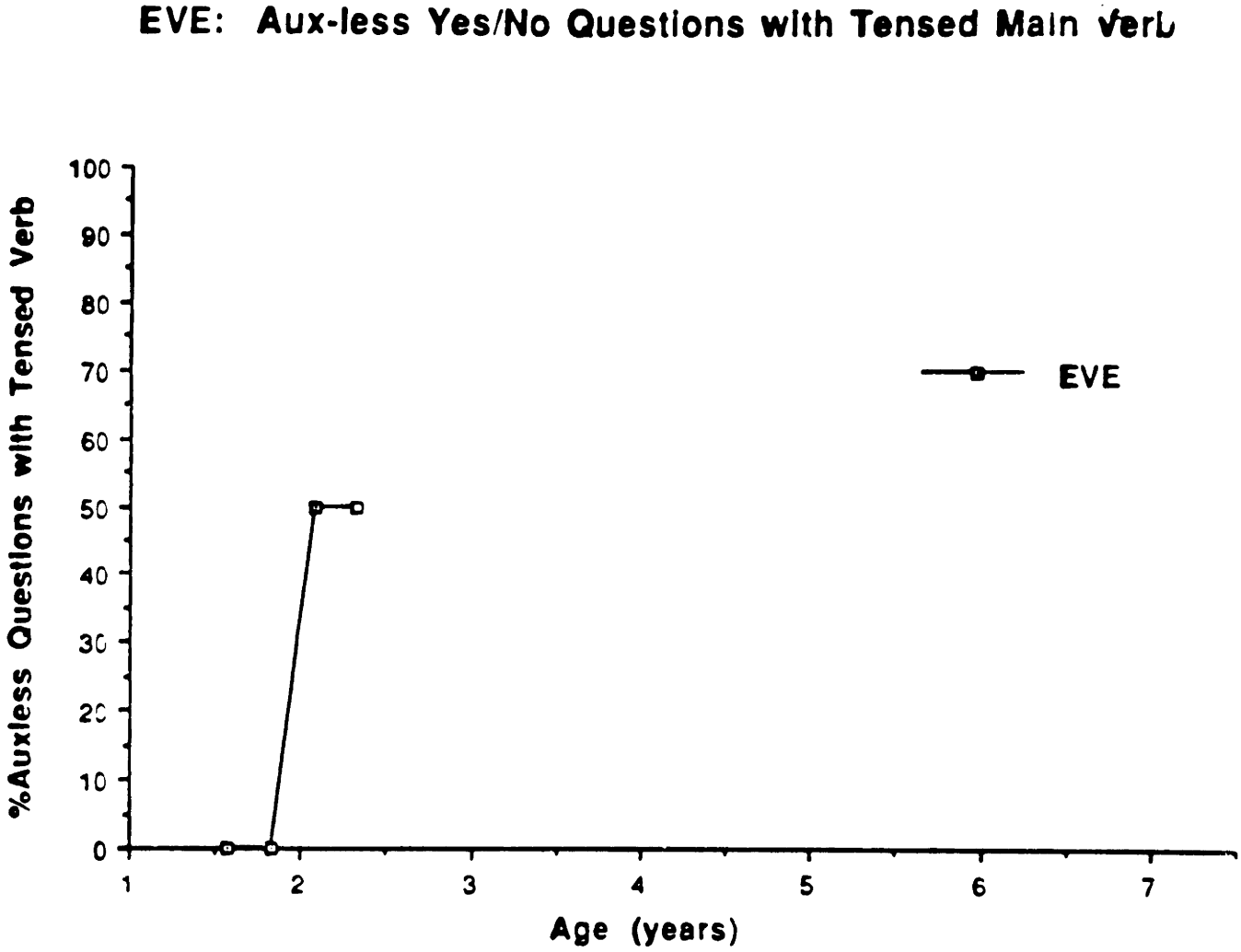


Figure 6.5
Aux-less *Yes/No* Questions with Tensed Main Verbs: Mark

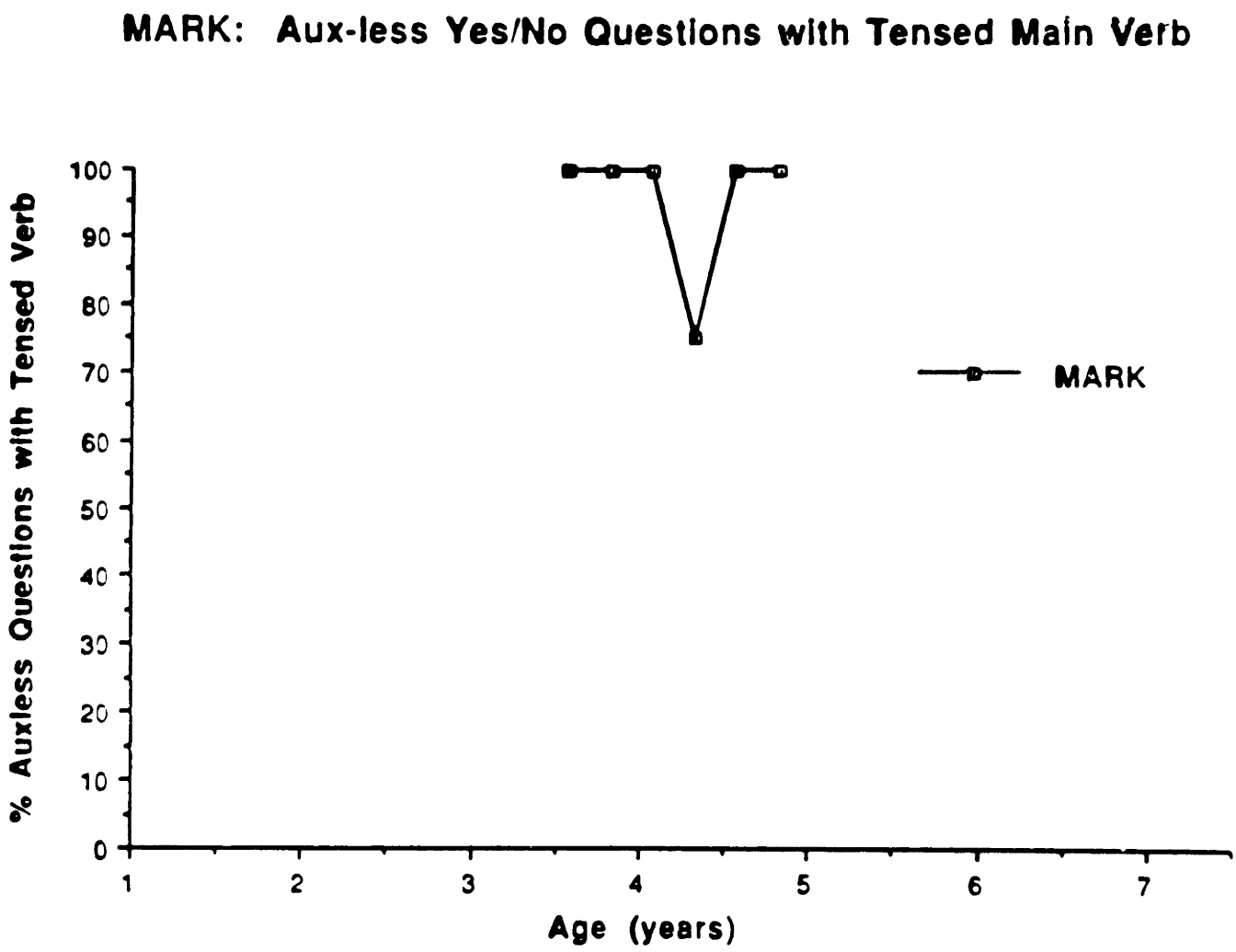


Figure 6.6
Aux-less *Yes/No* Questions with Tensed Main Verbs: Naomi

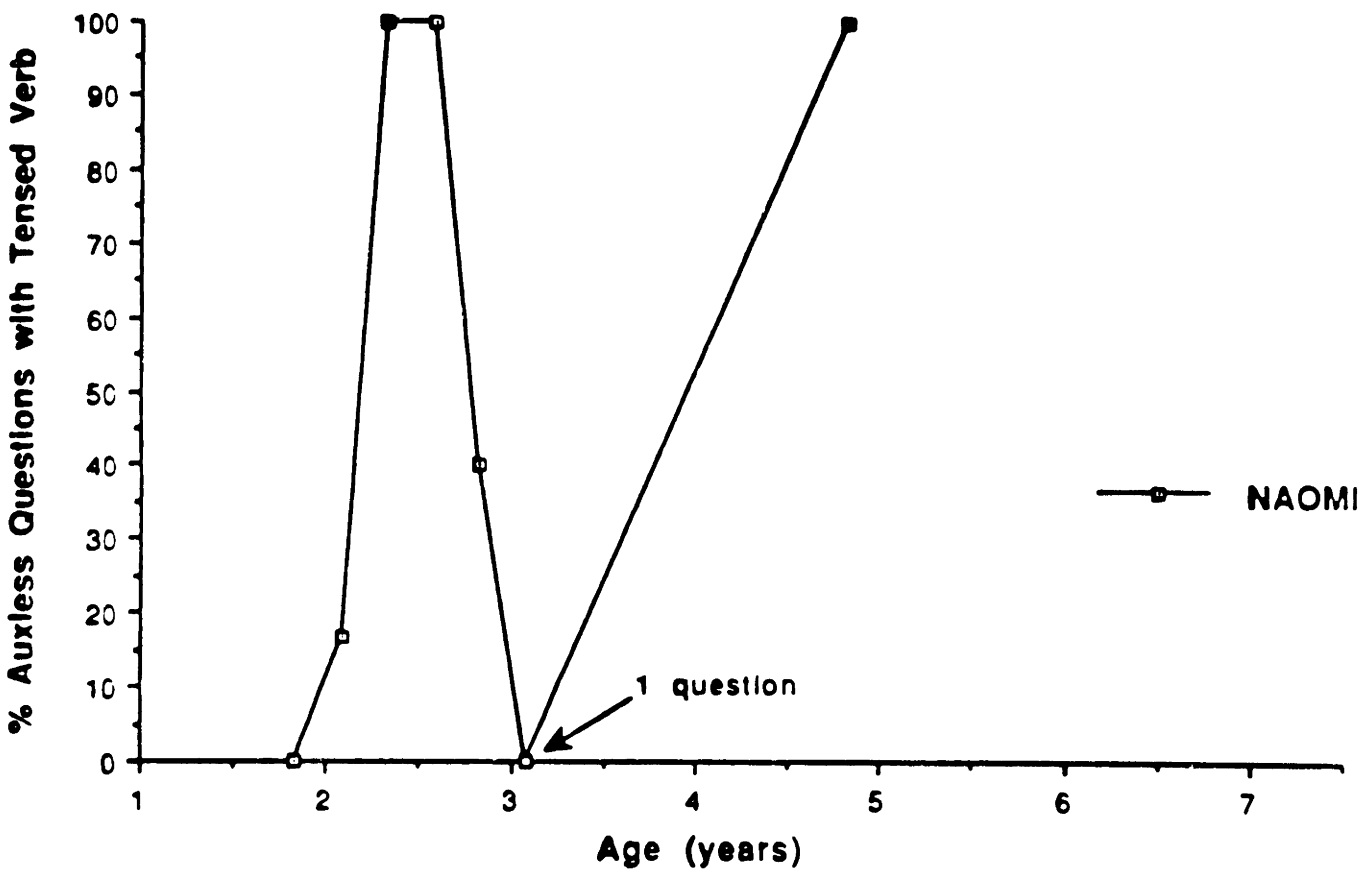


Figure 6.7
Aux-less *Yes/No* Questions with Tensed Main Verbs: Nathan

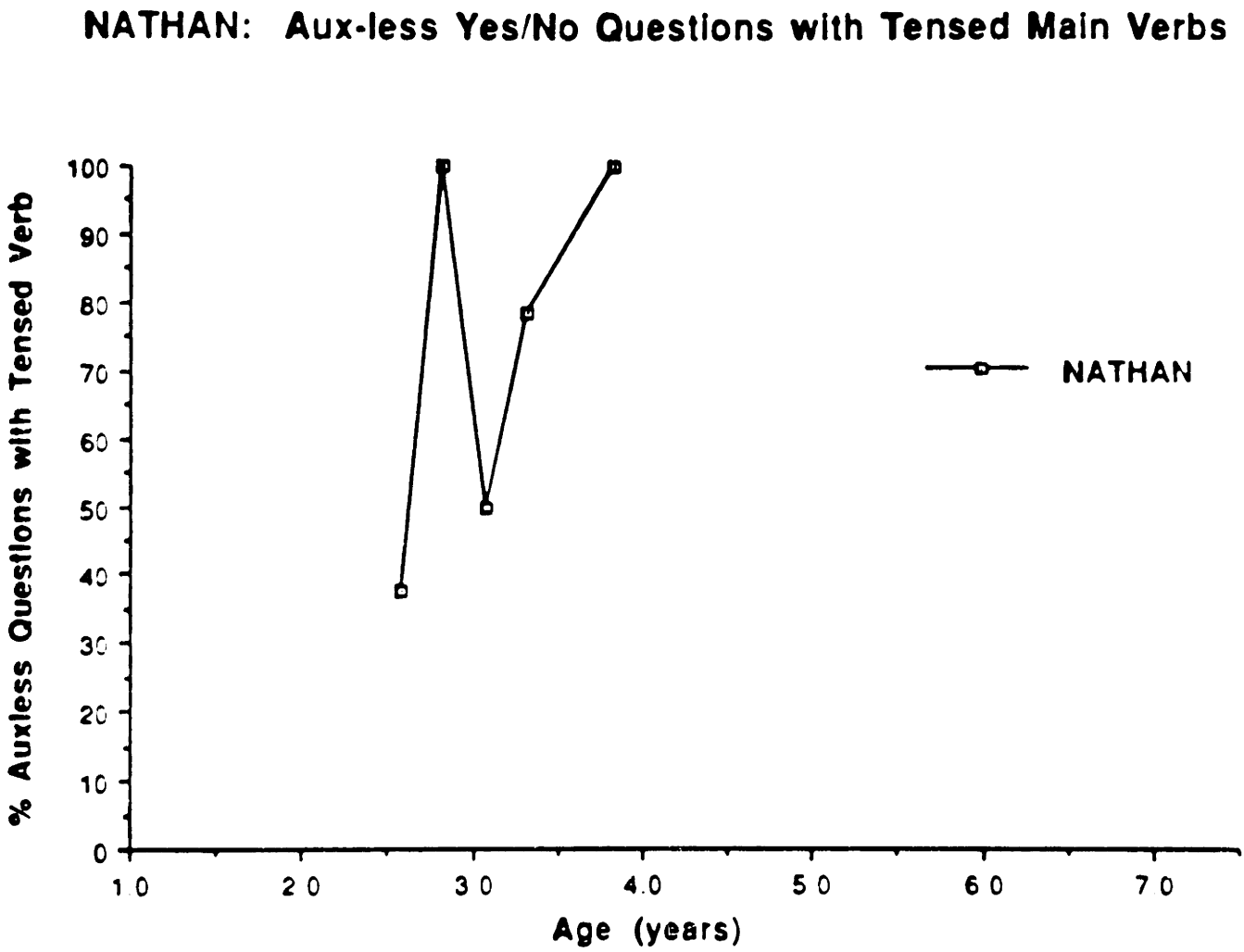


Figure 6.8
Aux-less *Yes/No* Questions with Tensed Main Verbs: Nina

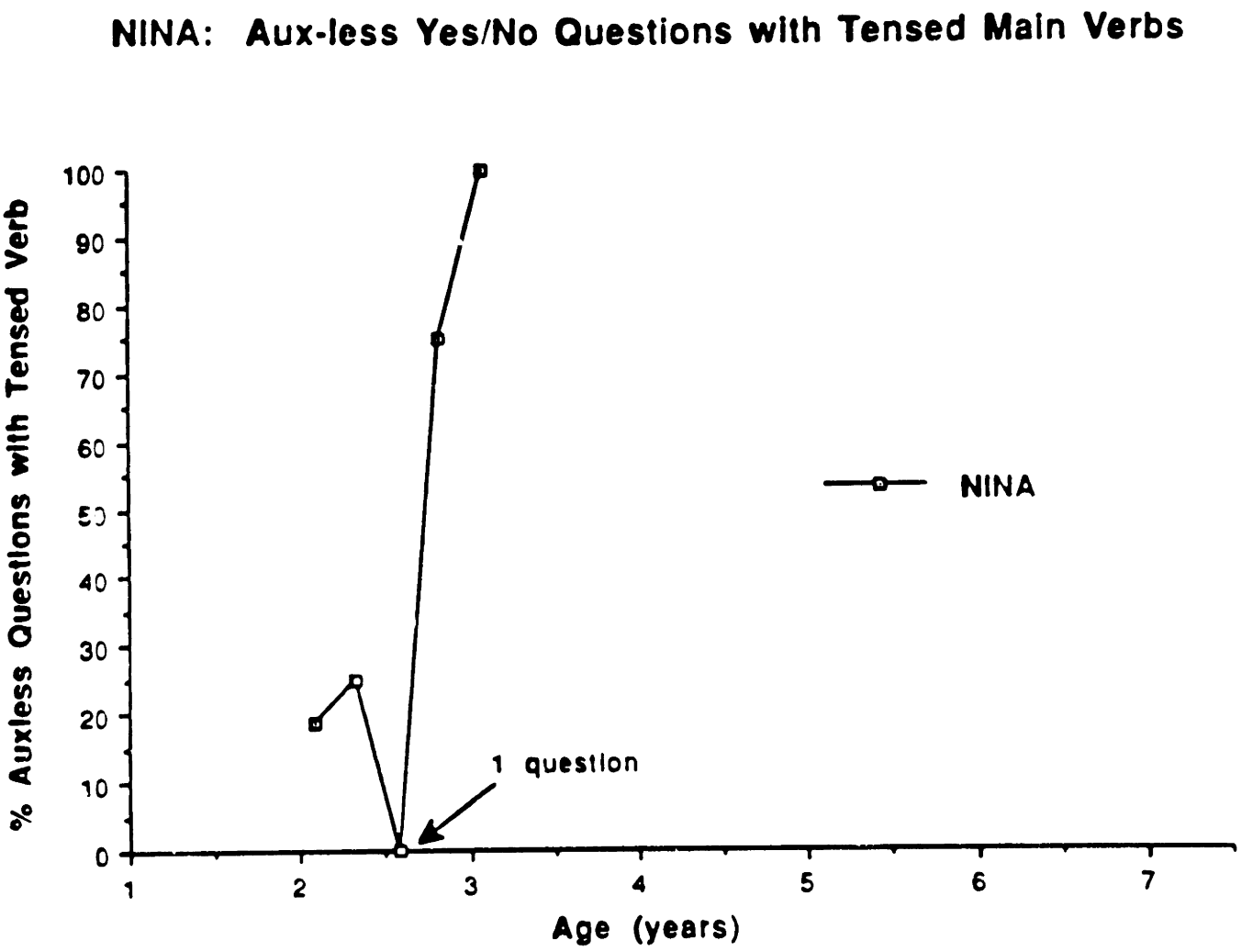


Figure 6.9
Aux-less *Yes/No* Questions with Tensed Main Verbs: Peter

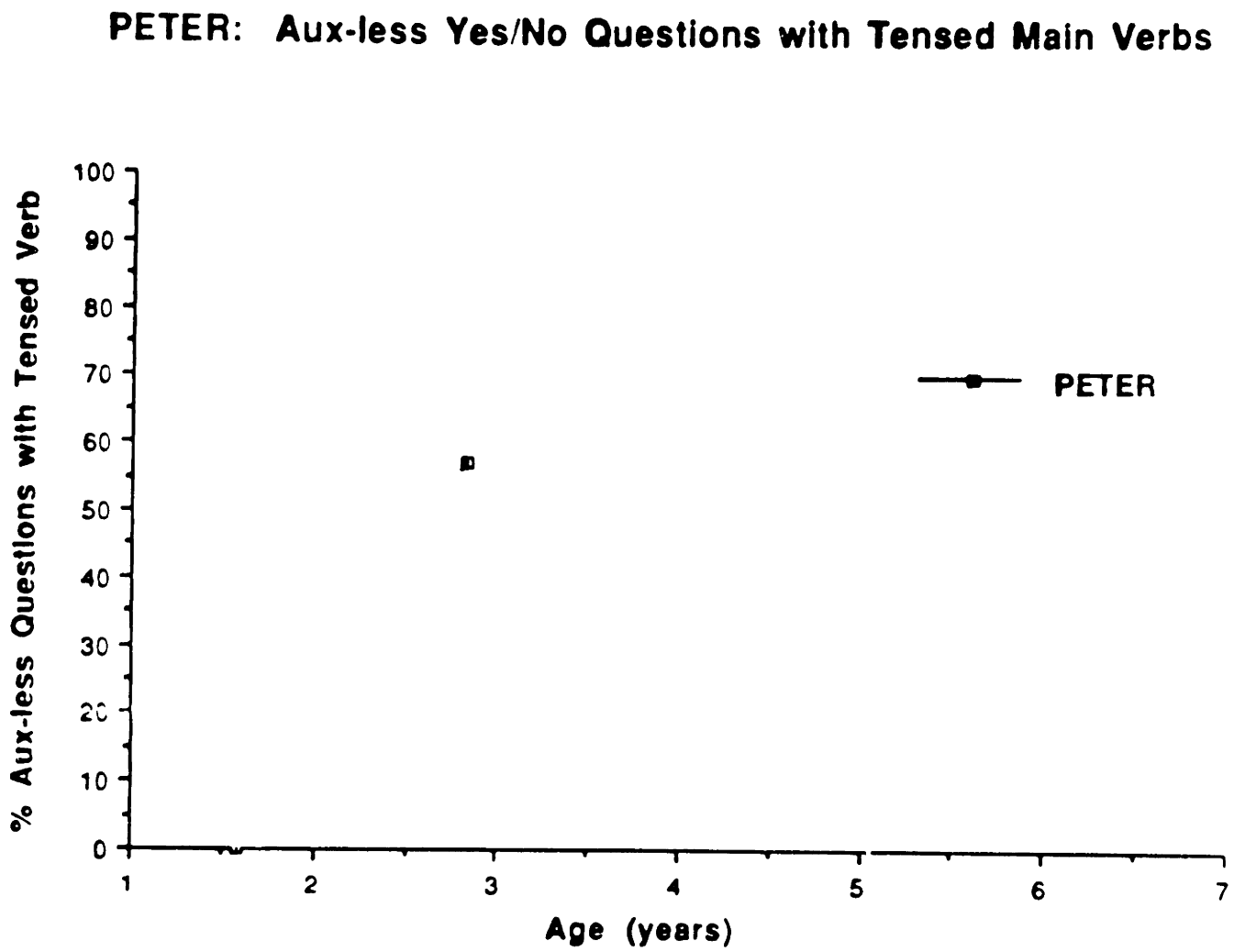


Figure 6.10
Aux-less *Yes/No* Questions with Tensed Main Verbs: Ross

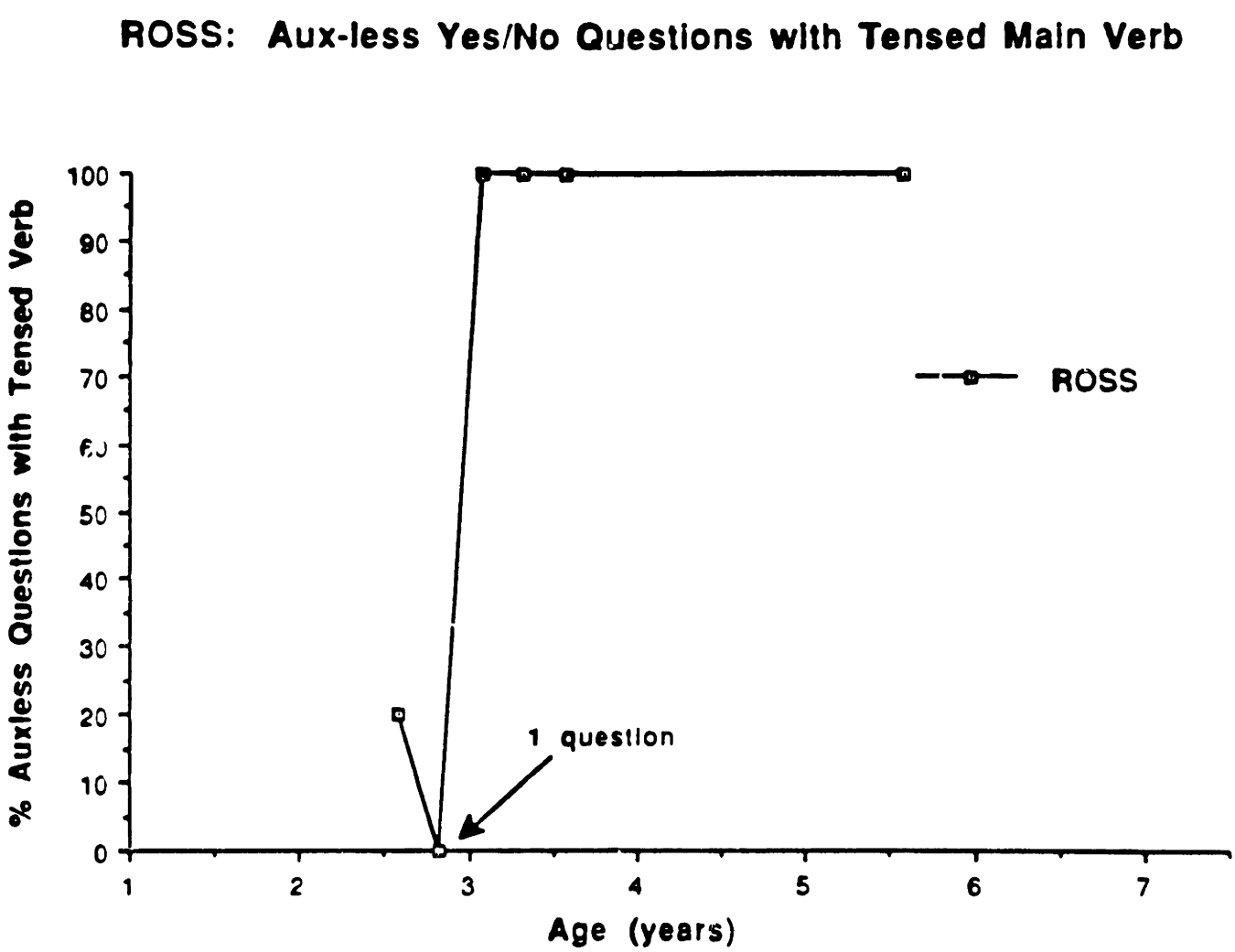


Figure 6.11
Aux-less *Yes/No* Questions with Tensed Main Verbs: Sarah

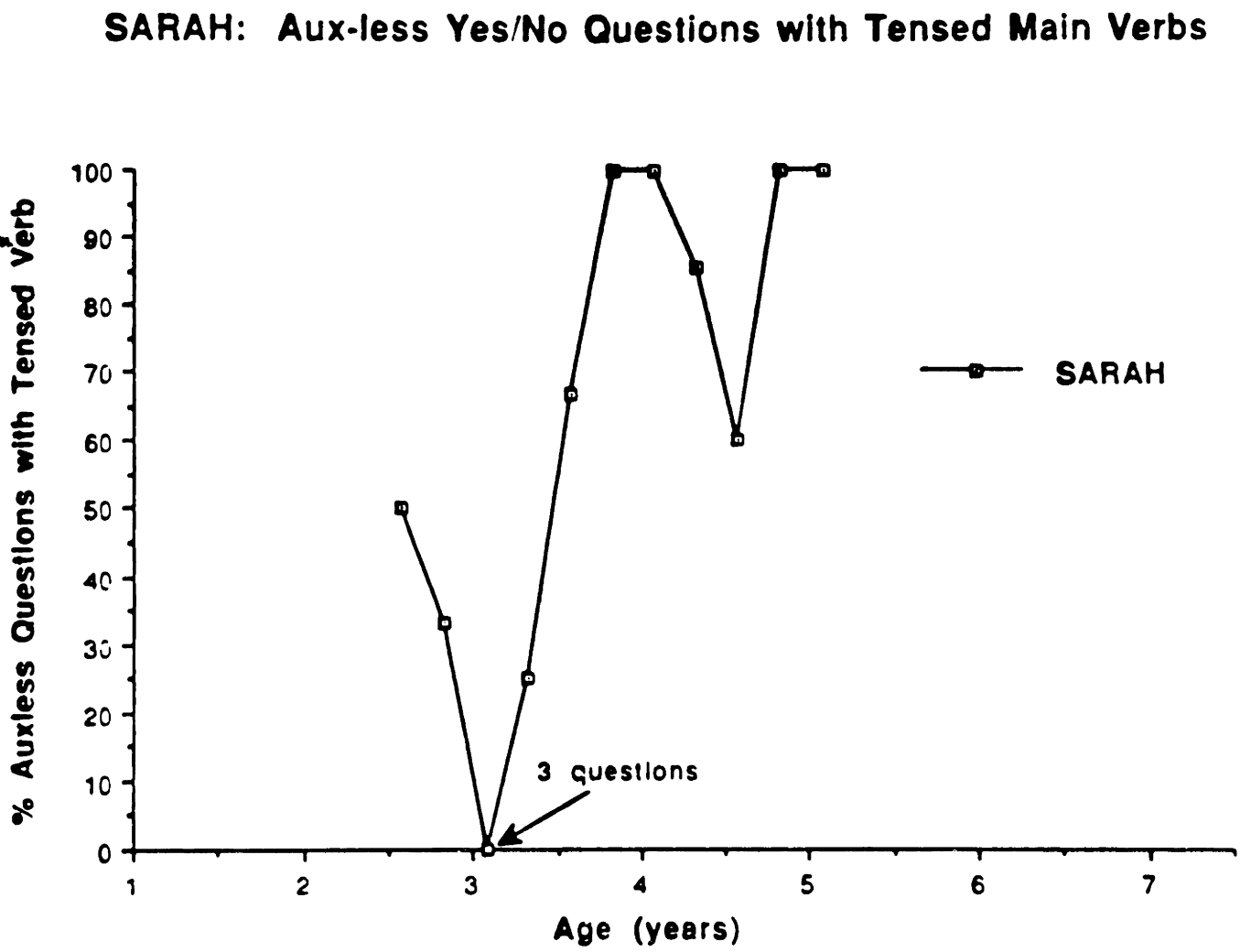


Figure 6.12
Aux-less *Yes/No* Questions with Tensed Main Verbs: Shem

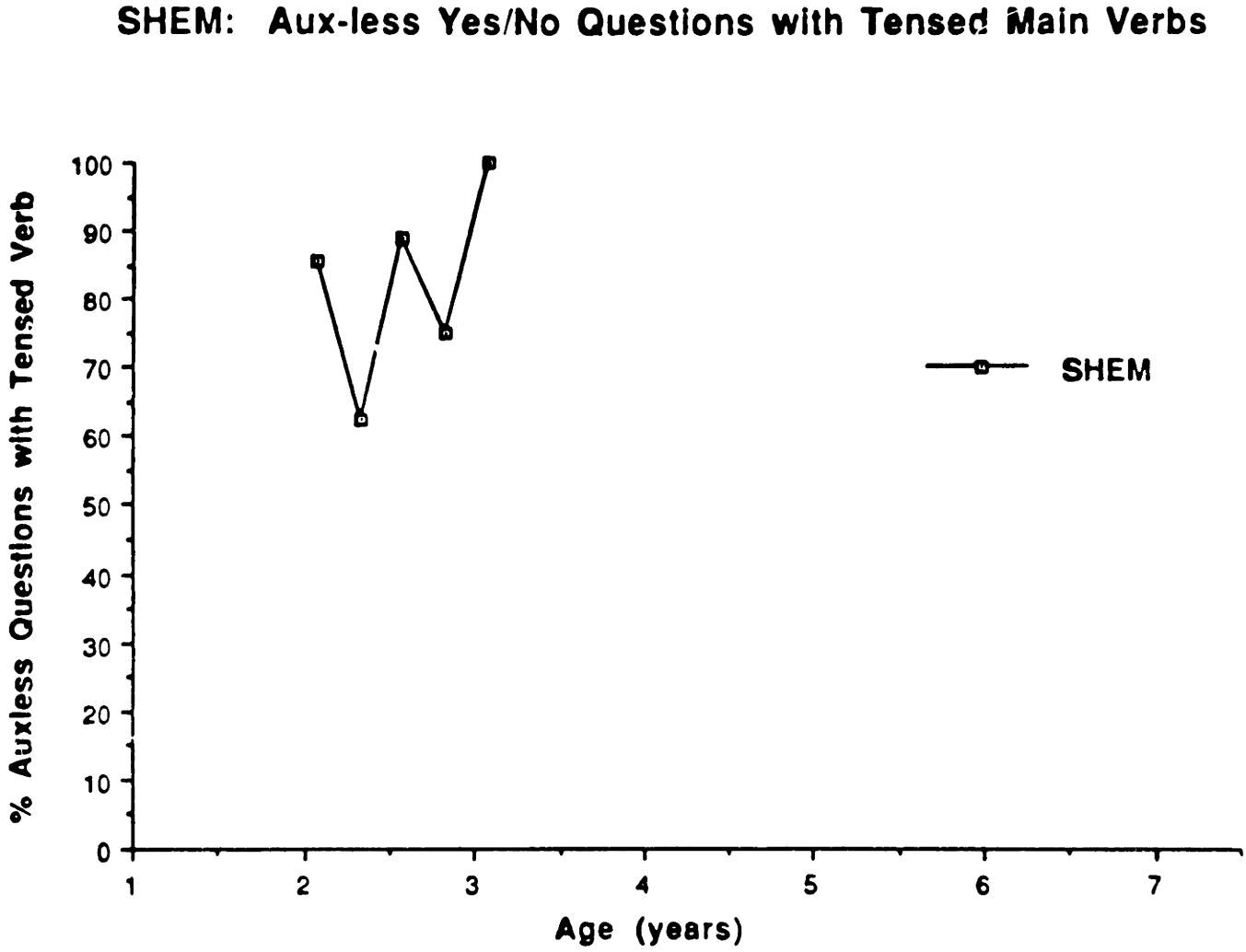


Figure 6.13
Aux-less *Wh*-Questions with Tensed Main Verbs: All Children

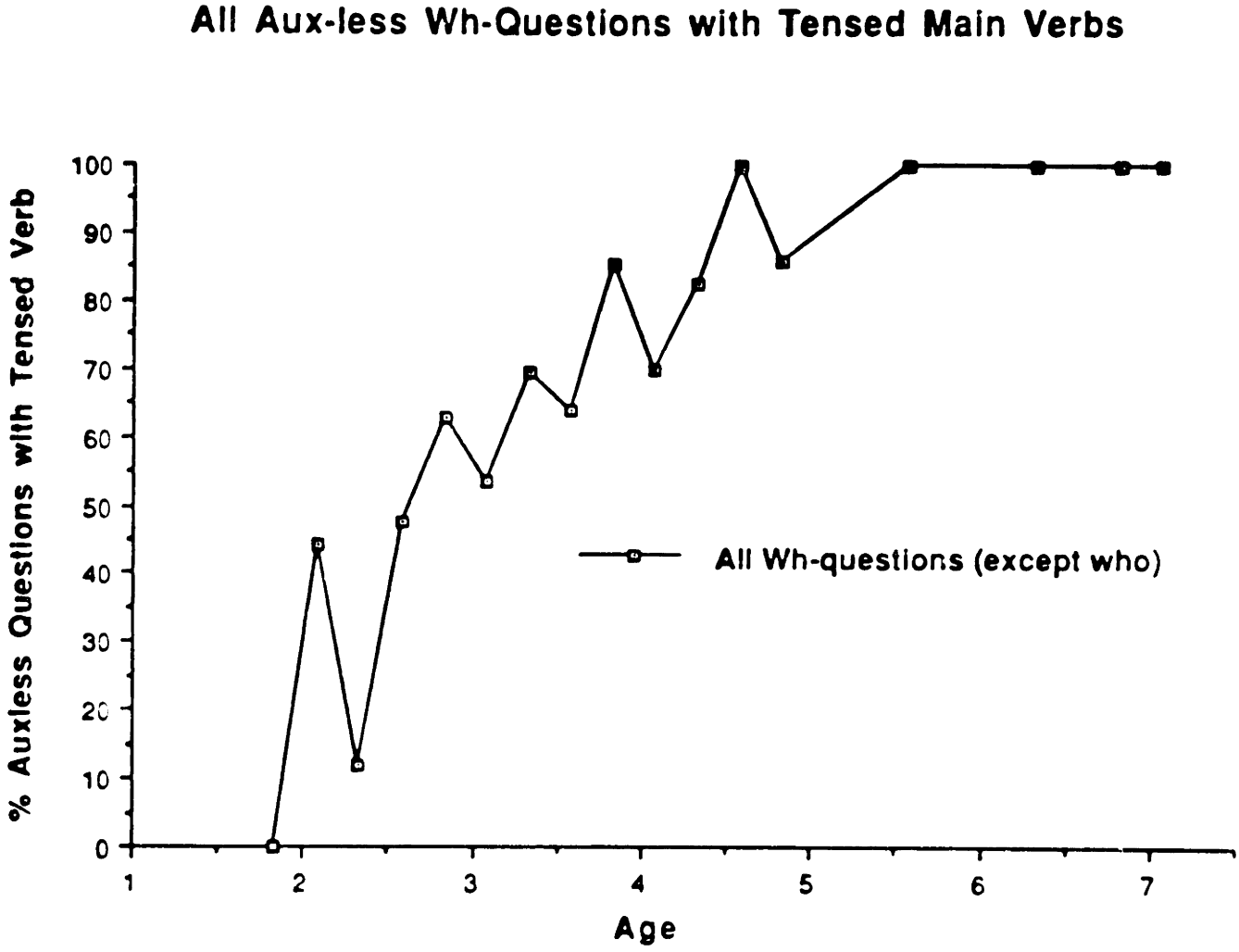


Figure 6.14
Aux-less *How* Questions with Tensed Main Verbs: All Children

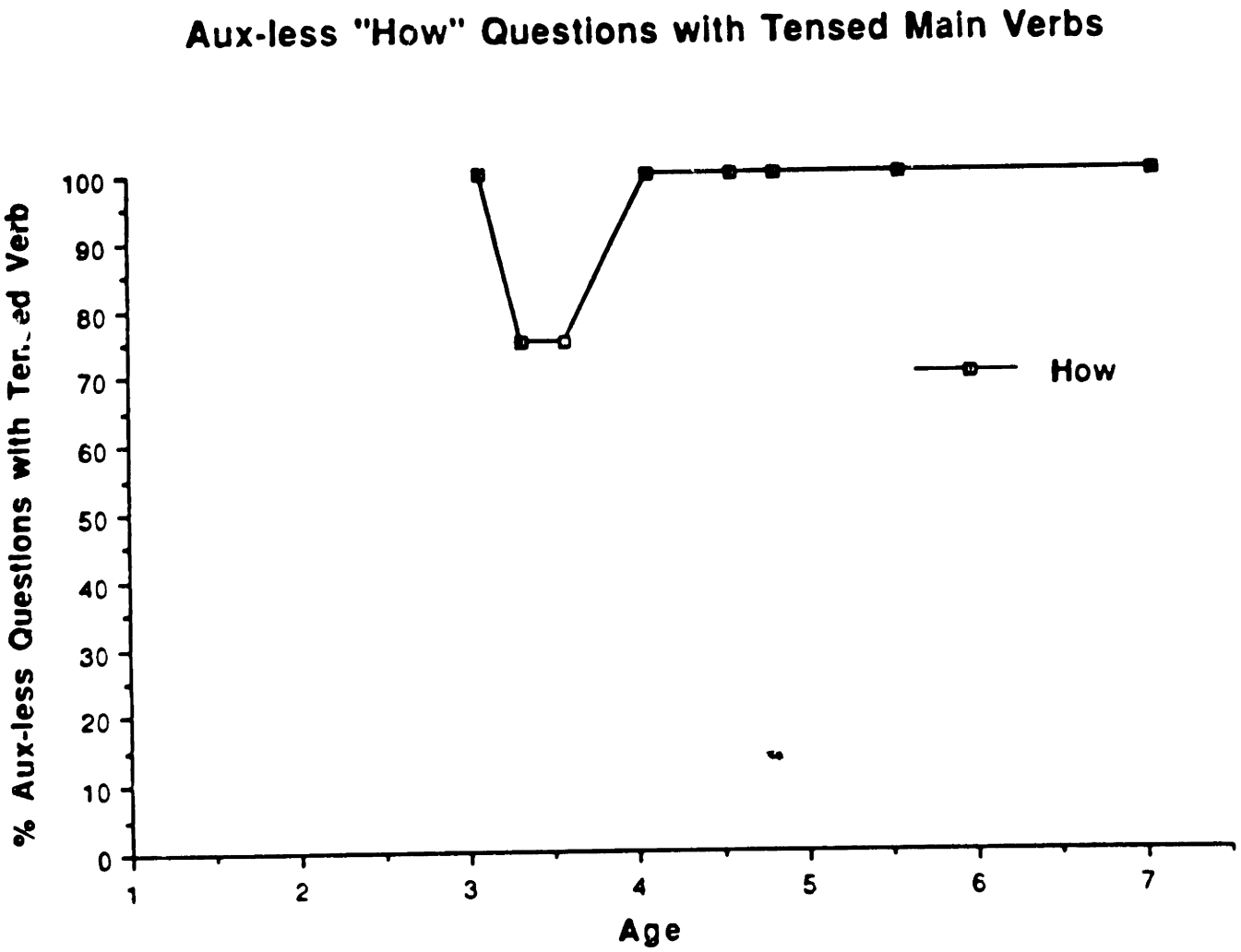


Figure 6.15
Aux-less *What* Questions with Tensed Main Verbs: All Children

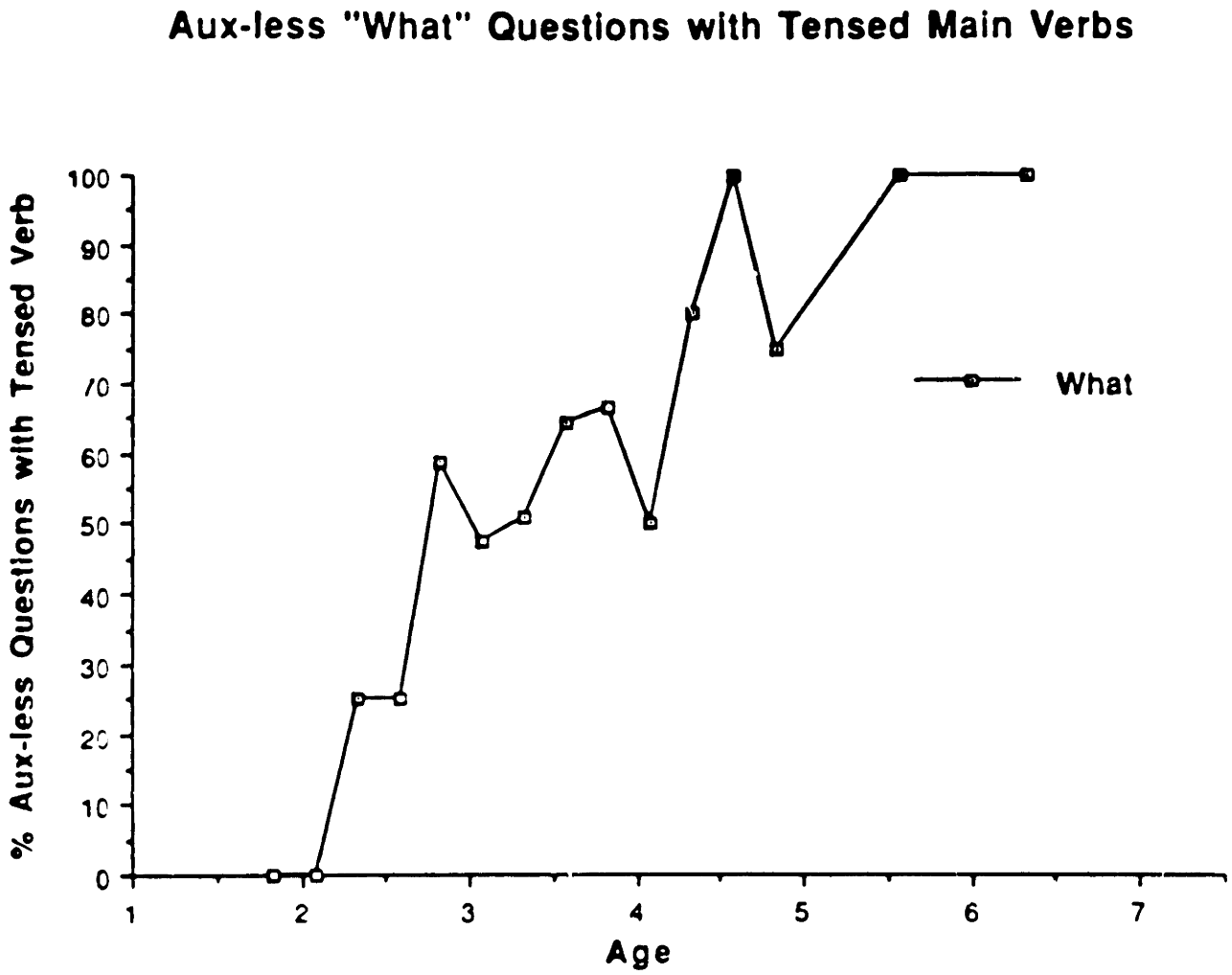


Figure 6.16
Aux-less *When* Questions with Tensed Main Verbs: All Children

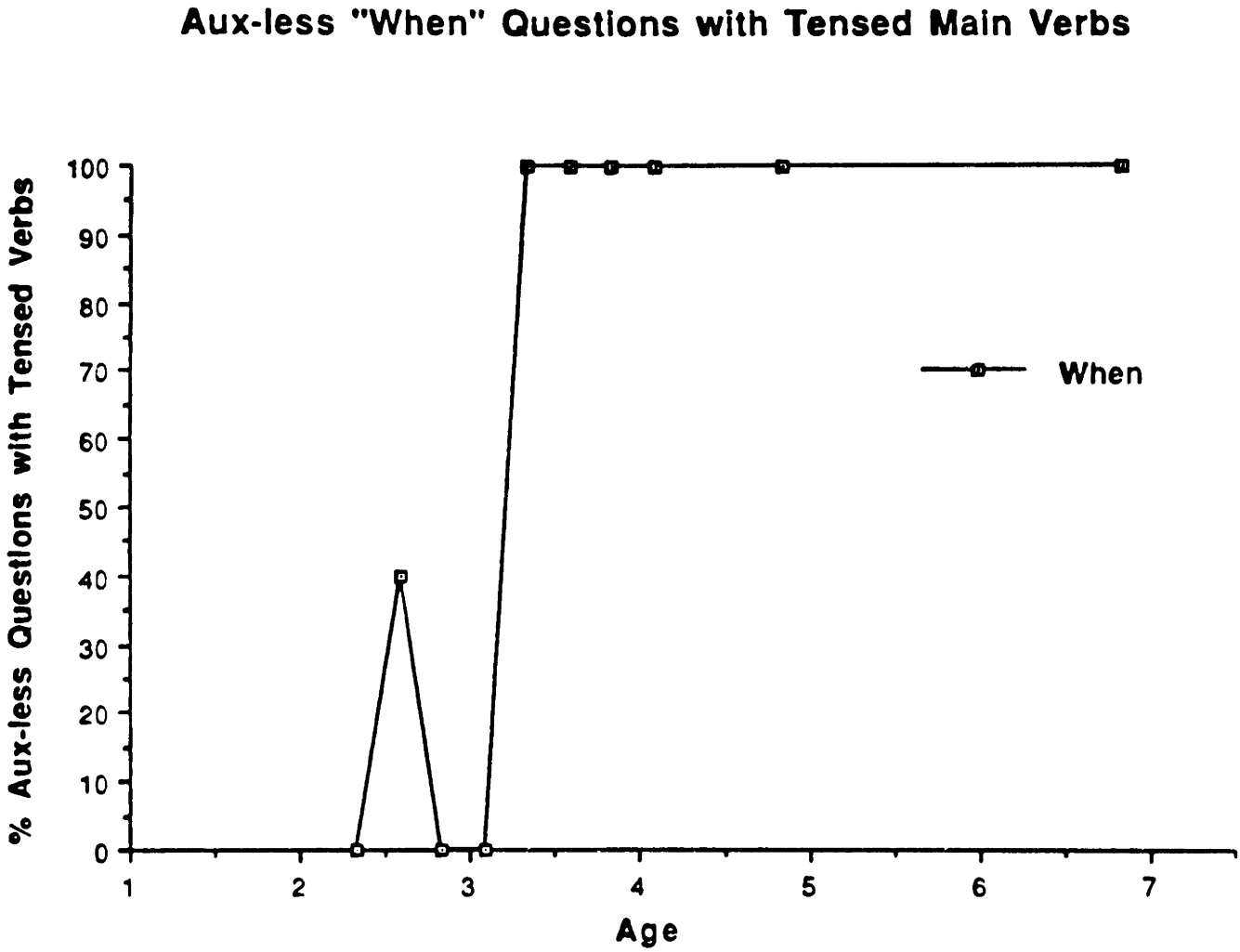


Figure 6.17
Aux-less *Where* Questions with Tensed Main Verbs: All Children

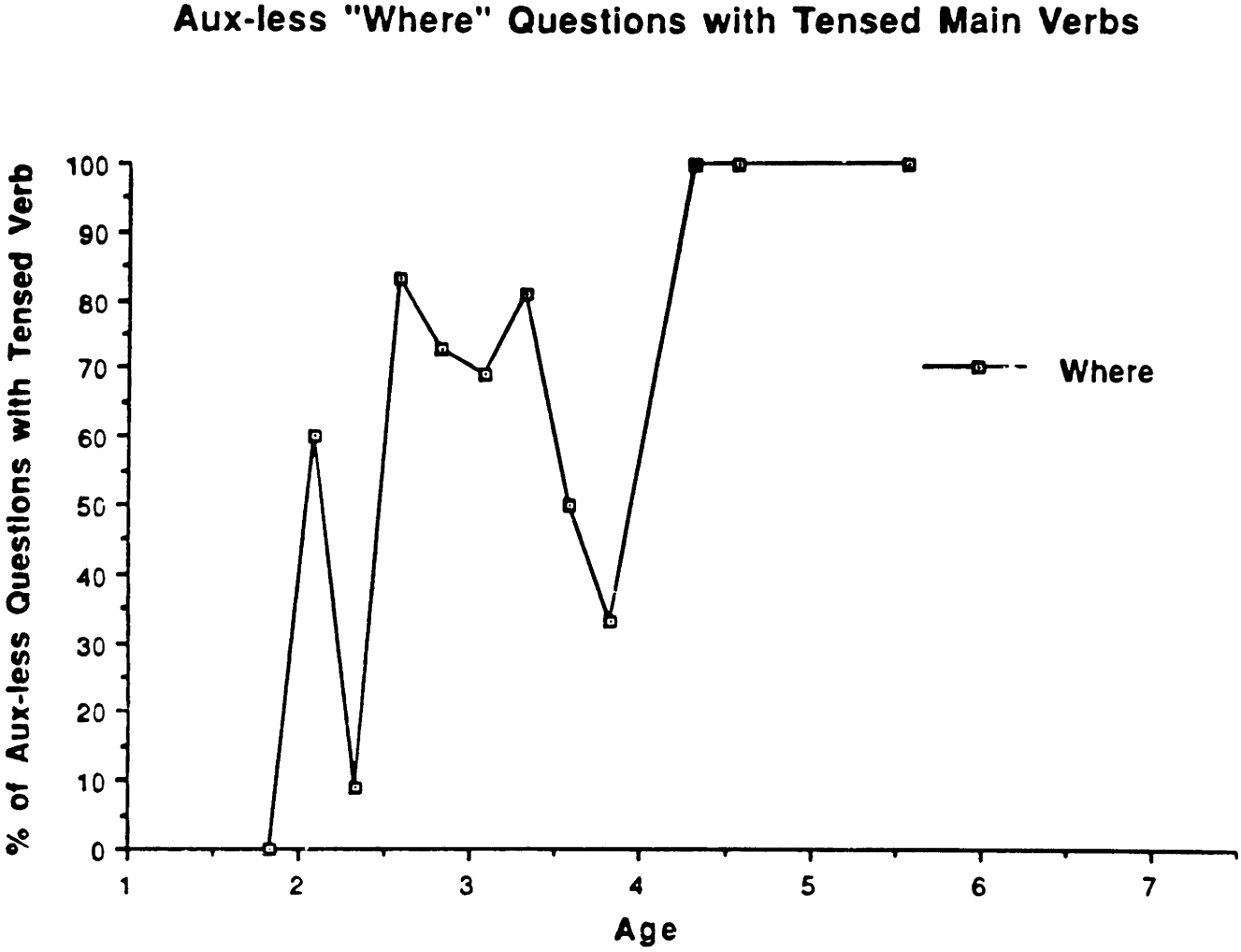


Figure 6.18
Aux-less *Which* Questions with Tensed Main Verbs: All Children

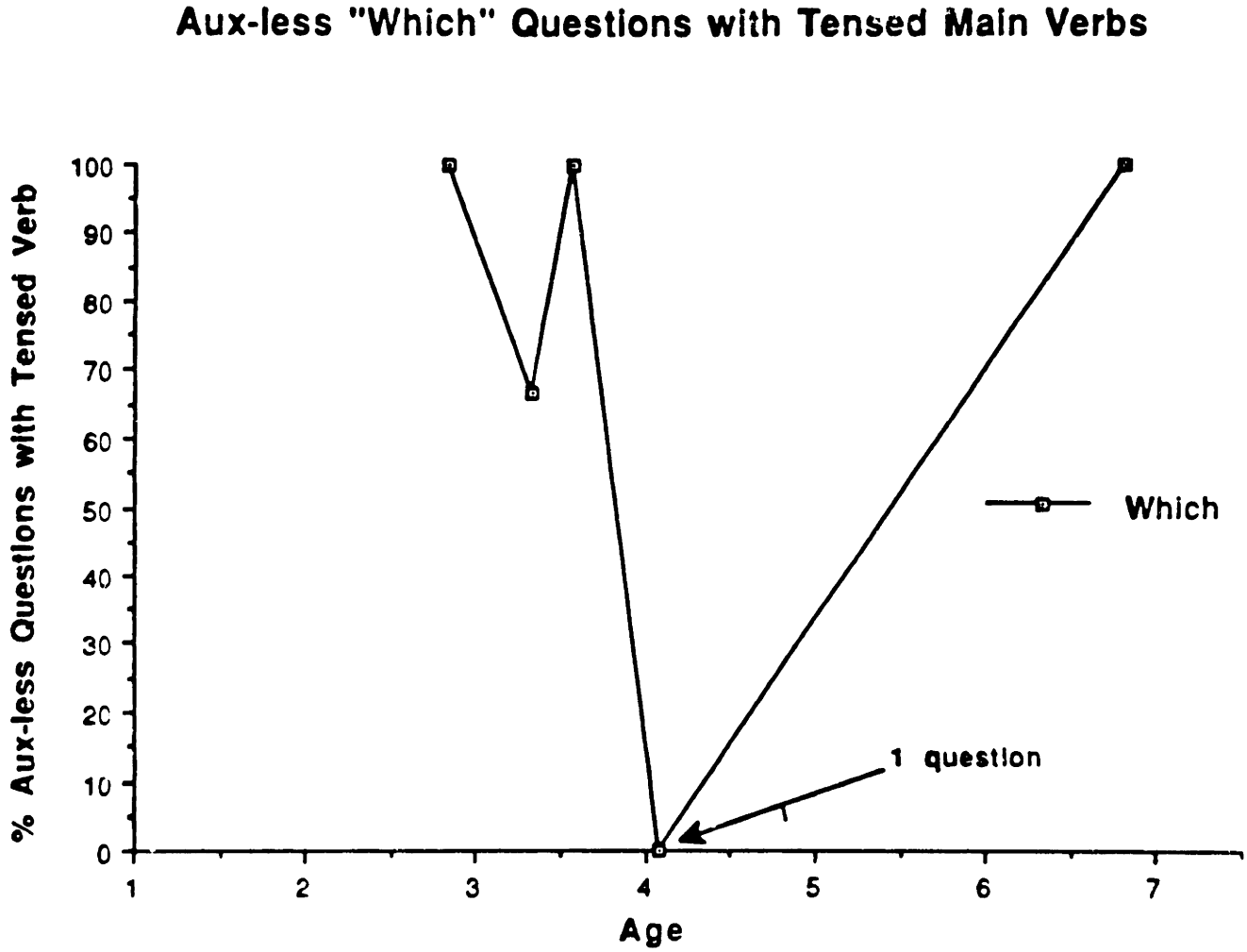


Figure 6.19
Aux-less *Why* Questions with Tensed Main Verbs: All Children

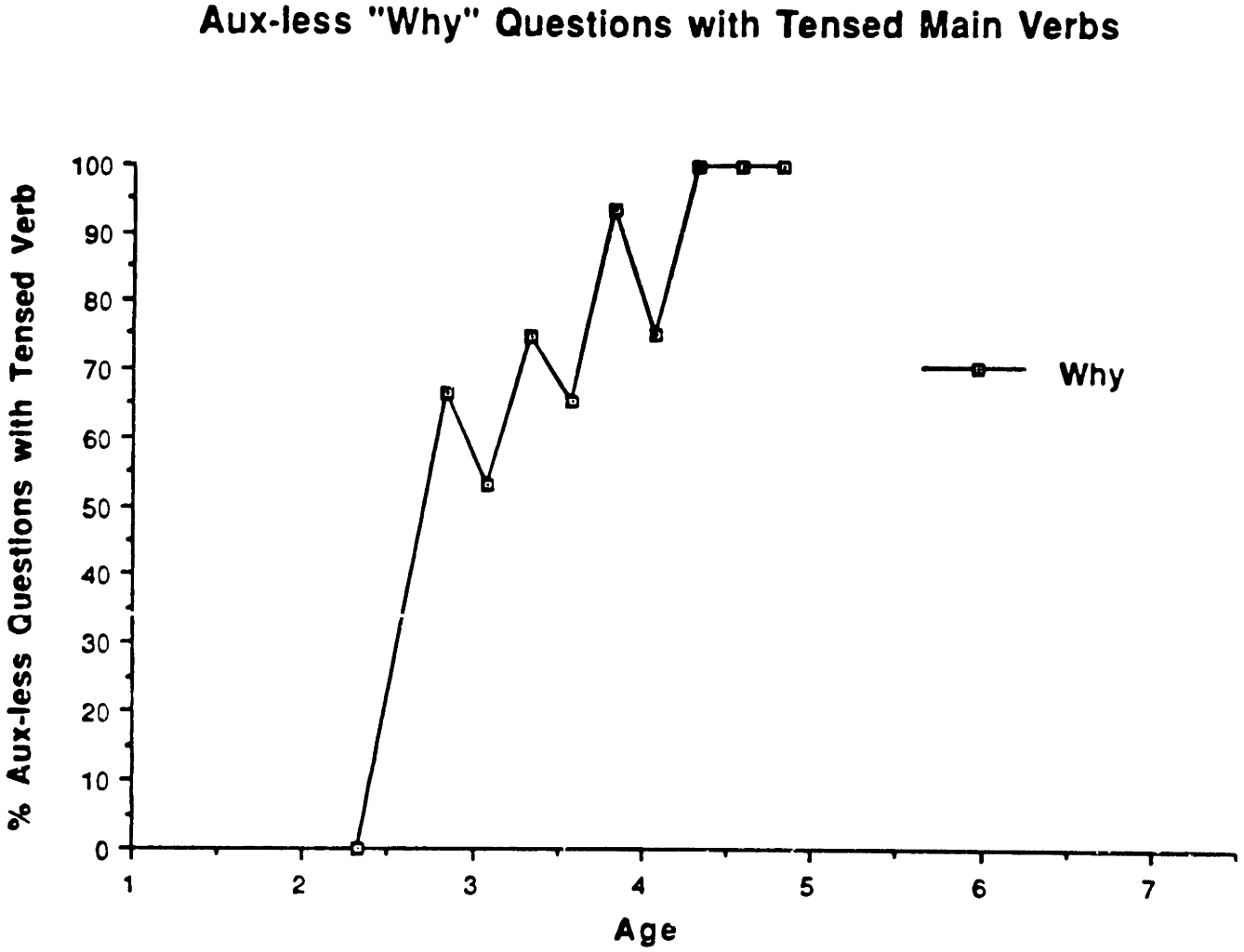
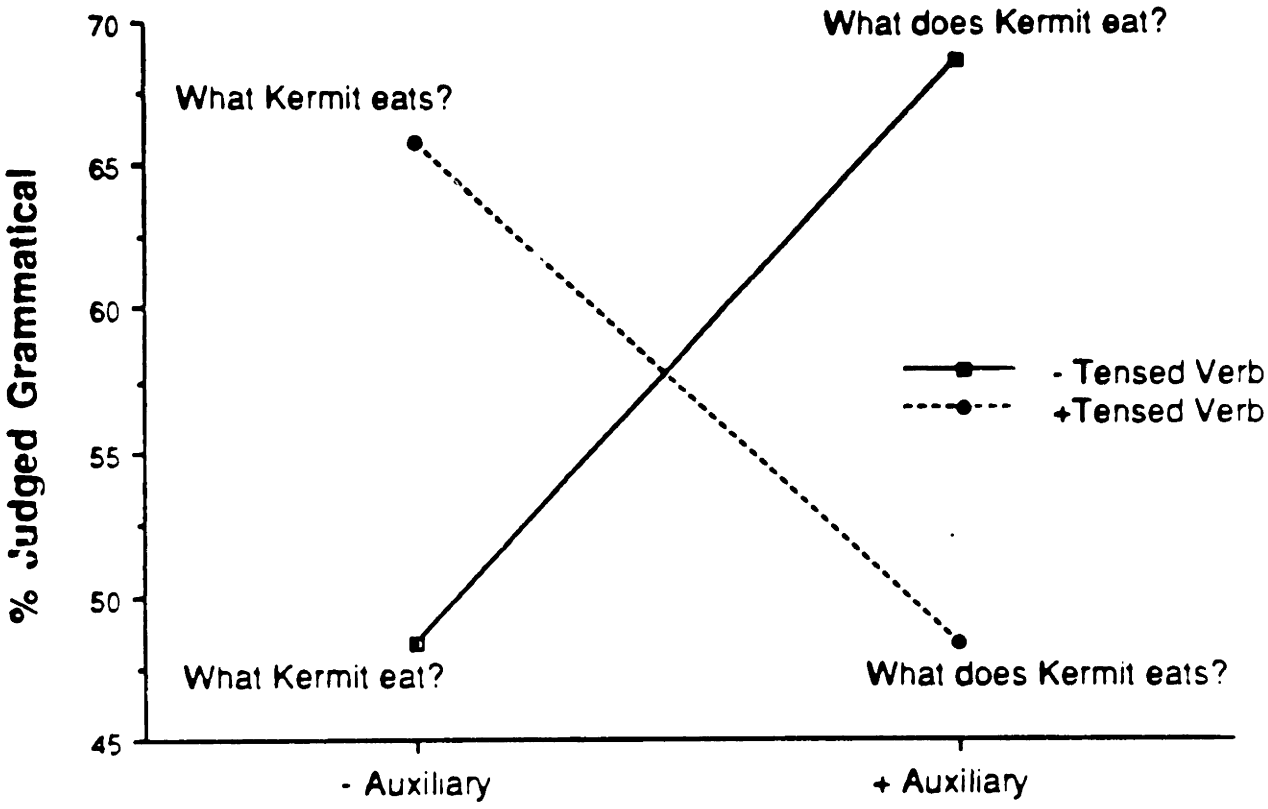
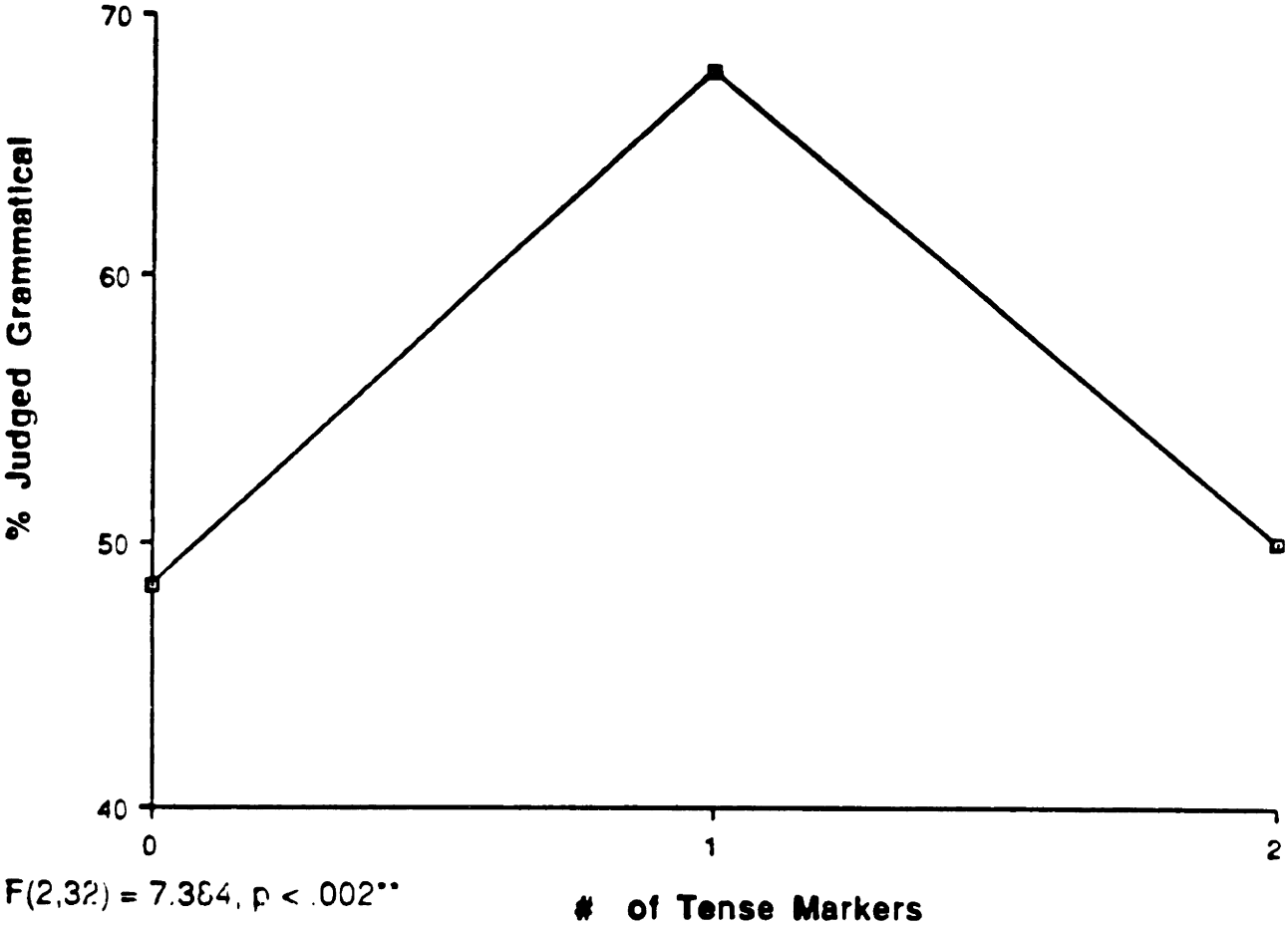


Figure 6.20
Interaction between +/-Auxiliary and +/- Tensed Main Verb



$F(1,16) = 3.924, p < .003^{**}$

Figure 6.21
Grammaticality Judgments and the Number of Tense Markers



$F(2,32) = 7.384, p < .002^{**}$

Figure 6.22
Interaction between MLU and +/-Auxiliary

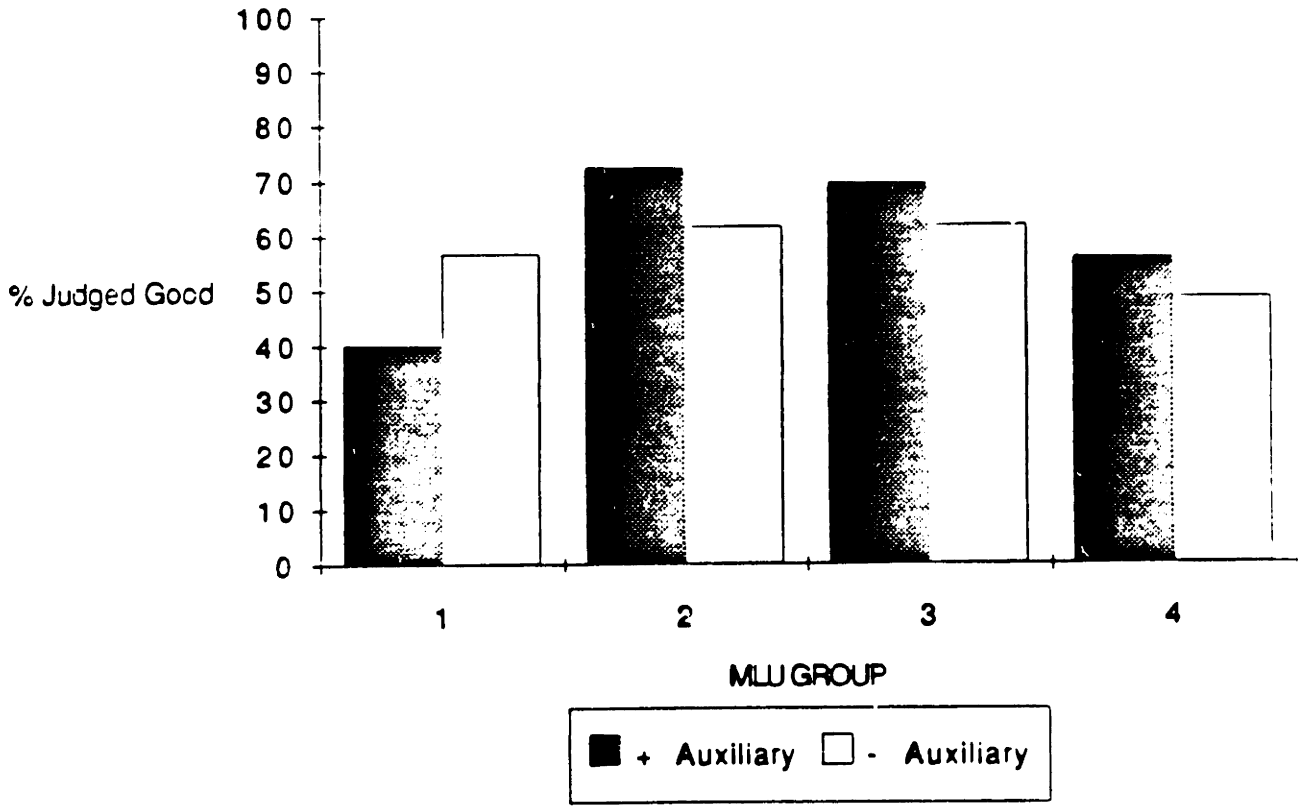


Figure 6.23
Interaction between *Wh*-word and +/- Auxiliary

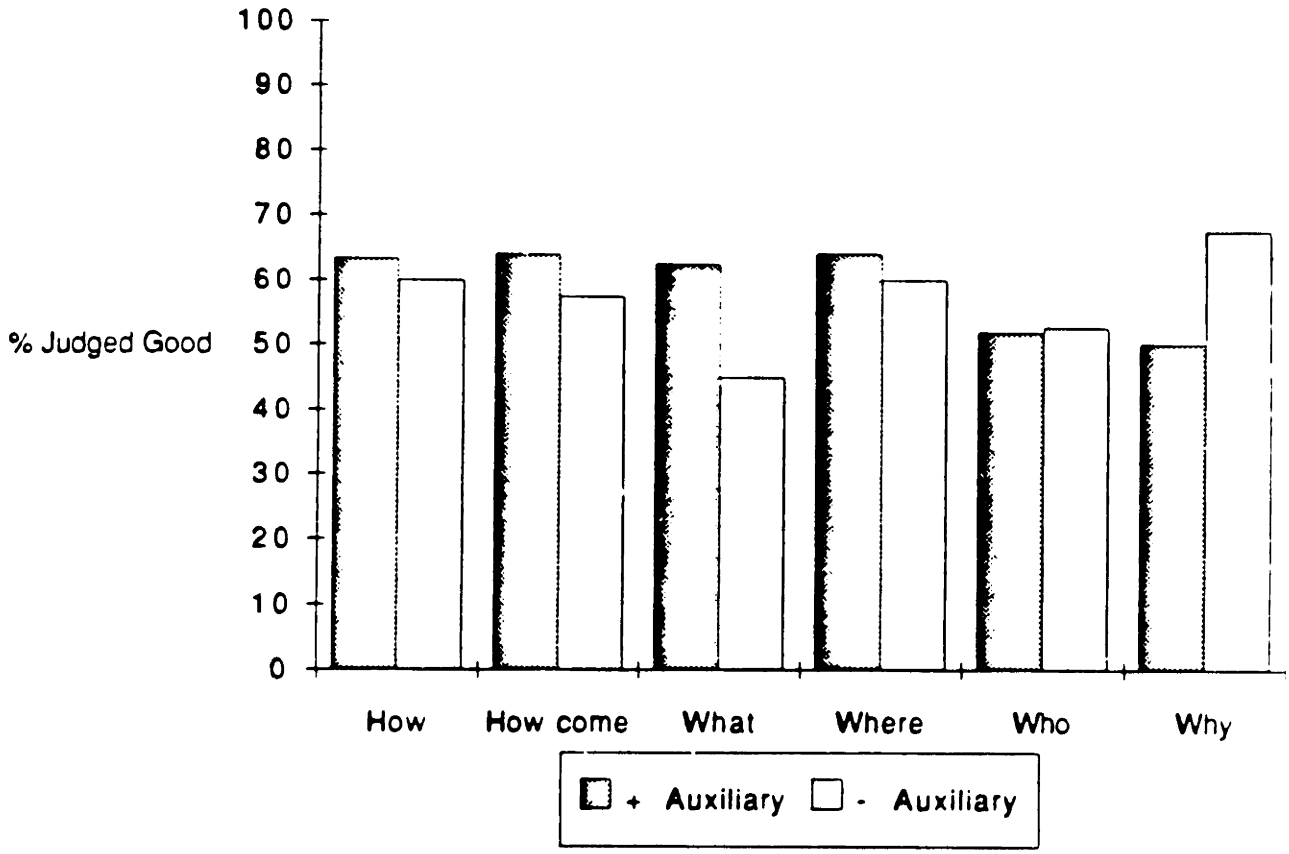


Figure 6.24
Interaction between *Wh*-word and +/- Tensed Main Verb