## INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand comer and continuing from left to right in equal sections with small overiaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality $6^{\prime \prime} \times 9^{\prime \prime}$ black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell \& Howell Information and Leaming 300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA 800-521-0600
UMI

# Accessing Linguistic Competence: <br> EVIDENCE FROM CHILDREN'S AND ADULTS' ACCEPTABILITY JUDGMENTS 

Kazuko Hiramatsu, Ph.D.<br>University of Connecticut, 2000

In order to Iearn about grammar, linguists primarily rely on acceptability judgments from native speakers of the language under investigation. Our hope is that these judgments allow us to tap into people's competence, or their knowledge of the language, and allow us to investigate the grammar. However, there has been some criticism raised regarding the use of judgments and what they tell us about competence and performance.

First, many researchers have argued that grammaticality judgments are not appropriate for studying children's competence since they are not able to perform metalinguistic tasks. However, McDaniel and her colleagues have argued that children, as young as $2 ; 11$, are capable of providing consistent and reliable judgments if they are trained.

In this study, I provide additional evidence that children are able to give reliable grammaticality judgments, and show that a combination of production and judgment data may reveal more about the child's grammar than production data alone.

## Kazuko Hiramatsu-University of Connecticut, 2000

In particular, I investigate children's non-adult negative questions with doubled auxiliary verbs, as in (1).

## (1) What did the smurf didn't buy?

My studies show that children produced 2Aux questions sentences, yet judged them to be ungrammatical. I argue that these children do in fact have the adult grammar, contrary to recent proposals, and that their production of 2Aux questions is a performance error related to knowledge about constituent negation.

A second concern has been raised with respect to the study of the adult grammar. Linguists have noticed anecdotally that certain types of island violations become increasingly acceptable after repeated exposure. In order to determine whether this socalled "syntactic satiation" is a general performance phenomenon or constrained by syntax, Stromswold (1986) and Snyder $(1994,2000)$ have investigated it experimentally. In this study, I replicate Snyder (1994) and test additional types of island violations. I also examine whether subject-related, such as handedness or linguistic training, and taskrelated factors, such as general reading ability, response time and presentation method, are associated with satiation. The evidence from the studies suggests that syntactic satiation is constrained by syntax and that it is a reflection of competence.

# Accessing Linguistic Competence: Evidence from children's and adults' acceptability judgments 

Kazuko Hiramatsu

B.A., Northwestern University, 1993

M.A., University of Connecticut, 1998

## A Dissertation <br> Submitted in Partial Fulfillment of the <br> Requirements for the Degree of <br> Doctor of Philosophy <br> at the <br> University of Connecticut

2000

All rights reserved.

## UMI

UMI Microform9988043
Copyright 2001 by Bell \& Howell Information and Leaming Company.
All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

Bell \& Howell Information and Leaming Company 300 North Zeeb Road
P.O. Box 1346

Ann Arbor, Mi 48106-1346

## Copyright by

## Kazuko Hiramatsu

## APPROVAL PAGE

## Doctor of Philosophy Dissertation

## Accessing Linguistic Competence: Evidence from children's and adults' acceptability judgments

## Presented by

Kazuko Hiramatsu. B.A., M.A.


Co-Major Advisor


University of Connecticut
2000
ii

## Acknowledgements

I would like to thank my advisors, Diane Lillo-Martin and William Snyder for their guidance throughout my graduate studies. The task of helping me learn how to carry out research fell to Diane, and I am grateful for her patience and encouragement. She spent countless hours with me at the Child Development Laboratories, and I learned a lot about conducting research and formulating hypotheses, as well as how to interact with children. William Snyder has also been very generous with his time. I would often stop by his office to ask a "quick question," only to emerge three hours later with a fresh set of questions to think about. He introduced me to both of the research topics in my dissertation, children's doubled auxiliary questions and syntactic satiation, and has constantly reminded me to look at the big picture. Being the statistical expert in the department, he has also been extremely patient with my endless questions about ANOVAs and paired t-tests.
[ am thankful to my other committee member, Howard Lasnik, for teaching me about clear argumentation and for his general enthusiasm about linguistics. I would also like to thank the other faculty members who have made the department a friendly and intellectually stimulating environment: Mamoru Saito, David Michaels, Ignatius Mattingly, Zeljko Boškovic, Sigrid Beck, Eva Bar-Shalom, Andrea Calabrese, Yael Sharvit, Arthur Abramson, Mona Anderson, and Michael Hegarty, who was a visiting professor at UConn, as well as Letty Naigles and Whit Tabor in the psychology department. I would also like to thank Beatrice Santorini, who was my undergraduate advisor at Northwestern University, for encouraging to pursue a graduate degree in linguistics, and Rudy Troike at the University of Arizona, who suggested attending UConn. Thanks are also due to Judy Marcus, our departmental secretary, for heip with various administrative tasks, as well as conversations about how to keep plants alive.

My acquisition research would not have been possible without the wonderful people at the Child Development Laboratories: Charlotte Madison, Sue Spencer, Mary Cox, Donna Thibault, and of course, the children. I am indebted to many people for assistance with my experiments: Dave Braze, who generously allowed me to use the webbased experiment he designed, and who provided me with general computer and statistical help; Debbie Chen, Anne Halbert, Laurel LaPorte-Grimes, and Stephanie Storrs, for being puppets; Toshiko Hiramatsu, for coordinating the experiments at Northwestern University; Diane Lillo-Martin, Yael Sharvit, Eva Bar-Shalom, Roger Chaffin, Bernard Grela, and Arthur Stepanov, for allowing me to recruit subjects from their classes.

My graduate studies were made possible by funding from the Department of Linguistics and an NIH-NIDCD Grant \#DC00183 to (Principal Investigator) Diane LilloMartin and (Investigator) William Snyder. Funding for the research on syntactic satiation
was supported by a Faculty Large Research Grant to William Snyder from the University of Connecticut Research Foundation, and an Extraordinary Expense Award from the University of Connecticut Research Foundation.

I have enjoyed the company of many students who have been a part of the department. I had the good fortune of having wonderful classmates, both academically and socially: Dave Braze, Tina Hsin, Satoshi Oku, Masao Ochi, and Eric Shortt, and especially Sandra Stjepanovic. I benefited greatly from the many people working on psycholinguistics research: Debbie Chen. Maki Yamane. Koji Sugisaki. Nobuhiro Miyoshi, Kelly Inman, and Yutaka Kudo. I spent much of my time in the lab, and was very lucky to have lab mates (and visitors) who enjoyed good food and conversation: Douglas Wharram, Cedric Boeckx, Marcela Depiante, Doreen Simons-Marques, Laura Levesque, Stephanie Storrs, Kelly Inman, Sarah Felber, Emma Ticio, Bosook Kang, Mary Bowen, Barbara Way, and especially Debbie Chen. And I am grateful to the people who showed me the ropes: Laurel LaPorte-Grimes, Laura Conway Palumbo, Elizabeth Laurençot, Ayumi Matsuo, Kazumi Matsuoka, Keun-Won Sohn, Miyuki Yamashina, and Kazuko Yatsushiro.

Finally, I would like to thank my parents, Miyoko and Yutaka Hiramatsu, who have always encouraged me to pursue my dreams; my sister, Toshiko Hiramatsu, who always knew the right thing to say to keep me going; and Adolfo Ausin, for his endless support and encouragement.

## Table of Contents

I INTRODUCTION ..... I
1.1 ACCEPTABILTY JUDGMENTS ..... I
1.2 Chlidren's metalingustic abilities ..... 4
1.3 Changing Judgments in adults ..... 8
2 CHILDREN'S QUESTIONS WITH DOUBLED AUXILIARY VERBS ..... 13
2.I INTRODUCTION ..... 13
2.2 Previous studies of 2AuX Questions ..... 17
2.2.I Stromswold (1990) ..... 17
2.2.2 Thornton (1993) ..... 19
2.2.3 Guasti, Thornton and Wexler (I995) ..... 23
2.2.3.1 Negative questions with not ..... 28
2.2.3.2 Negative questions with no inversion ..... 29
2.2.3.3 Negative 2Aux questions. ..... 30
2.23.4 2Aux/Neg questions ..... 32
2.2.f Summary ..... 33
23 EXPERIMENT I: ChILDREN'S PRODUCTION OF 2AUX QUESTIONS ..... 35
2.3.1 Methodology ..... 35
2.3.2 Subjects. ..... 38
2.3.3 Results from the adults. ..... 39
2.3.4 Results from the children ..... 39
2.4 EXPERIMENT 2: THE ROLE OF EMPHATIC STRESS IN CHILDREN'S 2AUX QUESTIONS ..... 46
2.4.I Methodology ..... 47
2.4.2 Subjects. ..... 49
2.4.3 Results from the actults. ..... 49
2.4.4 Results from the children. ..... 50
2.5 EXPERMMENT 3: CHILDREN'S JUDGMENT OF 2AUX QUESTIONS ..... 54
2.5.I Methodology ..... 54
2.5.1.1 Training and pretest sessions. ..... 55
2.5.1.2 Test sessions and follow-up sessions. ..... 57
2.5.2 Stimuli. ..... 58
2.5.2.I Stimuli for training and pretest sessions. ..... 58
2.5.2.2 Stimuli for test sessions ..... 59
2.5.2.3 Stimuli for follow-up sessions. ..... 61
2.5.3 Subjects ..... 62
2.5.f Results from the adults. ..... 63
2.5.5 Results from the children ..... 63
2.6 DISCUSSION ..... 65
2.6.1 2Aux questions and children's competence ..... 67
2.6.2 The atult grammar ..... 69
2.6.2. I Sentential and constituent negation in English ..... 69
2.6.2.2 Verbal morphology and I to $C$ movement ..... 74
2.6.2.3 Derivations of adult questions ..... 76
2.6.3 Children's production of 2 Aux questions ..... 80
2.7 SUMMARY ..... 87
3 SYNTACTIC SATLATION EFFECTS ..... 89
3.1 Previous treatment of island violations ..... 89
3.2 A REVIEW OF SNYDER $(1994,2000)$ ..... 95
3.3 EXPERIMENT 1 ..... 98
3.3.I Goals. ..... 98
3.3.2 Subjects ..... 99
3.3.3 Methodology ..... 100
3.4 Results from Experiment 1 ..... 103
3.4.I Comparison of satiation effects with Snyder (1994) ..... 103
3.4.1.1 Subject island violations. ..... 104
3.4.1.2 CNPC violations. ..... 105
3.f. I. 3 That-trace and want-for constructions ..... 107
3.4. 2 Post-test results. ..... 108
3.4. 3 Porential contingencies with other variables. ..... 109
3.5 DISCUSSION OF EXPERIMENT I ..... 111
3.5.I Huang (I982). ..... 112
3.5.2 Chomsky (1986) ..... 112
3.5.3 Takahashi (1994) ..... 115
3.5. $\ddagger$ Summary ..... $1 / 7$
3.6 EXPERIMENT 2 ..... 119
3.6.I Goals. ..... 119
3.6.2 Subjects ..... 120
3.6.3 Methodology. ..... 120
3.7 RESULTS FROM EXPERIMENT 2 ..... 125
3.7.1 Satiation effects. ..... 125
3.7.2 Filler and control items ..... 130
3.7.3 Response time differences. ..... 133
3.8 FOLLOW-UPTO EXPERIMENT 2 ..... 136
3.8.1 Methodology ..... 137
3.8.2 Subjects ..... 137
3.8.3 Results and discussion ..... 138
3.9 SUMMARY ..... 142
4 CONCLUDING REMARKS ..... 145
APPENDIX 1 ..... 150
APPENDIX 2 ..... 152
REFERENCES ..... 154

## List of Tables

Table I: Types of positive questions produced by children ..... 42
Table 2: Types of negative questions produced by children ..... 43
Table 3: Summary of children's elicited production results ..... 44
Table 4: No. of questions with emphasis (in response to emphatic prompts) ..... 51
Table 5: Summary of satiation results ..... 104
Table 6: Fisher's exact test results for Wh-islands ..... 110
Table 7: Fisher's exact test results for Subject islands ..... 110
Table 8: Fisher's exact test results for that-trace ..... III
Table 9: Mean number of yes responses (beginning vs. end) for subjects with $90 \%$ correct filler/control items. ..... 127
Table 10: Mean number of yes responses (beginning vs. end) for subjects with at least median score on filler/control items ..... 129
Table II: Range of scores on Fast Reading task, Author Recognition task, and filler/control items. ..... 131
Table 12: Mean score (\% of correct responses) by filler/control type. ..... 131
Table 13: Mean response times by test sentence type ..... 134
Table 14: Mean response times (beginning vs. end) by test sentence type ..... 135
Table 15: Mean response times for test vs. control type. ..... 136
Table 16: Comparison of mean score (\% of correct responses) by filler/control typel ..... 140

Table 17: Children's percentage adult-like responses (Experiment 3)
Table 18: Children's percentage adult-like responses (follow-up to Experiment 3)... 151

## List of Figures

Figure 1: Children's question production ..... 40
Figure 2: Children's question production by sentence type. ..... 41
Figure 3: Satiation effects for subjects with at least $90 \%$ correct filler/control items 126
Figure 4: Satiation effects for subjects with at least median score on filler/control items128

## Chapter 1

## Introduction

## 1.I Acceptability judgments

Chomsky (1965: p. 4) makes a fundamental distinction between "competence (the speaker-hearer's knowledge of his language) and performance (the actual use of language in concrete situations)." One of the tasks of linguists is to describe competence using evidence from performance. The primary linguistic data we use to investigate competence are acceptability judgments from native speakers of the language under investigation.

Acceptability judgments are often incorrectly considered to be a direct reflection of competence. Schütze (1996) suggests that this view may be the result of confusion regarding "intuition", which is part of competence, and "judgment", which is a product of performance. He notes that we may have an "intuitive sense of grammaticalness" (Chomsky 1985 [1955-56]), but we cannot judge grammaticalness. Schütze (p. 24, fin. 10) also reminds us of Chomsky's original intent in using the term "grammaticalness":

About the term "grammaticalness,"...I purposely chose a neologism in the hope that it would be understood that the term was to be regarded as a technical term, with exactly the meaning that was given to it, and not assimilated to some term of ordinary discourse with a sense of connotations not to point in this context. (Noam Chomsky, quoted in Paikeday (1985), p.14)

To add to the confusion of terminology, the notions "acceptable" and "grammatical" are frequently used interchangeably. However, Chomsky (1965) makes the following distinction between these two notions, which I assume throughout the thesis.

Acceptability is a concept that belongs to the study of performance, whereas grammaticalness belongs to the study of competence. (p. 11)

Linguists elicit judgments of acceptability from informants, and then try to construct arguments about the (un)grammaticality of those sentences.

A direct mapping between grammaticality and acceptability does not always exist; the grammaticality of a sentence is not the only factor that determines the acceptability of the sentence. For example, the sentence in (1) is nonsensical and might be judged as unacceptable, even though it is grammatical.
(1) Colorless green ideas sleep furiously.

Similarly, sentences with multiple center-embedding, such as (2), are grammatical, but are judged as unacceptable. It has been argued that the unacceptability of these sentences is due to short-term memory limitations. (See for example Chomsky (1965).)
(2) The patient the nurse the clinic had hired admitted met Jack. (Frazier (1985))

Schütze (1996) and Cowart (1997) review several criticisms regarding the use of acceptability judgments. Householder (1965) regards judgments to be too variable
among speakers and also between occasions, while Labov (1975) found disagreement among informants, as well as between linguists and naive informants. Some researchers, such as Bever (1970), Birdsong (1989), and Gleitman and Gleitman (1979) have suggested that the grammar linguists are constructing based on judgments may not be the same as the grammar that underlies production and comprehension. Other researchers find the informal collection of judgments to be troublesome (Derwing (1979), Sampson (1975), Newmeyer (1983), Bradac et al. (1980)), although Chomsky (1969) acknowledges this.

Schütze cites four key reasons for the use of acceptability judgments. First, acceptability judgments allow us to examine the status of sentences that do not occur naturally in spontaneous speech or recorded corpora. Second, they allow us to examine negative information: sentences that are not part of the language. Third, if we were to just observe people's speech, it would be difficult to reliably distinguish performance errors, such as false starts and unfinished sentences, and grammatical production. Finally, under the assumption that communicative functions might cloud our insight into competence, the use of acceptability judgments minimizes the role of communicative functions since we ask the informant to judge a sentence with no real function.

In this thesis, I investigate some issues that have been raised about the use of acceptability judgments to understand competence. In the next section, I discuss some arguments that have been made with respect to children's metalinguistic abilities to
provide acceptability judgments. In section I.3, I turn to issues regarding the stability of acceptability judgments from adults, and discuss a phenomenon in which certain types of ungrammatical sentences become more acceptable after repeated exposure.

### 1.2 Children's metalinguistic abilities

In studying language acquisition, there are many approaches to discovering children's competence (McDaniel, McKee and Cairns (1990), Crain and Thornton (1998)). Although many linguists do not typically use production data from adults to investigate the grammar, this is a common strategy in the study of the child's grammar. From what the child does and does not produce, we hope to discover something about the development of language. The child does not utter fully grammatical sentences from the beginning; there are ways in which the child's utterances are systematically different from those allowed by the adult language. We study how the child's utterances are similar to or different from the adult language and attempt to answer why this might be so.

The study of children's production is important, especially in young children who would not be able to participate in an experimental task. However, we may be over-estimating or under-estimating the child's grammar based on her spontaneous production. The lack of some sentence structure may simply be the result of the child
choosing not to produce it. For example, long-distance questions, such as (3), are generally not found in spontaneous speech data.
(3) What do you think the smurf bought?

It would be incorrect to assume that children do not know how to form long-distance questions based on this finding. If children are put in an experimental situation where long-distance questions are appropriate and plausible, children do produce them (Thornton 1990).

Similarly, the lack of a particular sentence structure does not necessarily mean that they do not have knowledge about that structure. Thomton (1995) used a comprehension task in addition to a production task, and discovered that although children seldom produce grammatical negative questions with $n^{\wedge} t$, such as (10), they are able to understand them.
(4) What didn't the smurf buy?

Therefore, other ways of assessing the child's competence, such as elicited production, comprehension tasks and grammaticality judgment tasks, are also important. When we have data from different tasks, we hope to have a better picture of the child's grammar.

Until recently, researchers studying language acquisition did not elicit acceptability judgments from children since it was believed that they would not be able to provide them. The task is commonly referred to as the "grammaticality
judgment' task, but as I discussed earlier, this does not make sense given Chomsky's definitions of "acceptable" and "grammatical". The elicitation of judgments is part of performance, not competence. Therefore, it only makes sense to talk about "acceptability judgments". For ease of exposition, I will continue to refer to the task and aspects of the methodology as involving "grammaticality" judgments. However, in discussing the results from the grammaticality judgment task, I rely on the distinction between the acceptability and grammaticality of a sentence.

Many researchers have studied the development of metalinguistic abilities in children and have proposed reasons for the delay in the ability to provide acceptability judgments. (See Chaudron (1983), Ryan and Ledger (1984), Birdsong (1989) and Gombert (1992) for literature reviews.) For example, Hakes (1980) argues that the metalinguistic ability to give acceptability judgments, the ability to explain judgments of space and number, and the ability to develop intentional memorization strategies all emerge developmentally around the same age. (See also Van Kleeck (1982).) Hakes suggests that all of these abilities involve controlled processes, and hence emerge later than production and comprehension abilities, which he argues are automatic processes. Piaget attributes the ability to use controlled processes to evaluate a situation to middle childhood.

However, many researchers have used acceptability judgments successfully with young children. For example, de Villiers and de Villiers (1974) were able to elicit judgments from four year olds; Schlisselberg (1988) investigated children 3;6 to

6;6 and found that even some of her youngest subjects were able to give reliable acceptability judgments. McDaniel and her colleagues (McDaniel and Cairns (1990), McDaniel et al. (1995), McDaniel and Cairns (1996)) have been successfully eliciting acceptability judgments from young children for the past decade. (See McDaniel and Cairns (1990, 1996) among others.) Other researchers, such as Stromswold (1990) and Smith-Lock and Rubin (1993) have also used grammaticality judgments task.

In Chapter 2, [ investigate children's non-adult negative questions with doubled auxiliary verbs (2Aux questions), where tense and/or agreement is marked on both auxiliary verbs.
(5) What did the smurf didn't buy?

Using an elicited production task, I determined what types of grammatical and ungrammatical sentence structures children use in producing positive and negative questions. I also investigated whether children would produce positive 2Aux questions if emphatic stress was involved. The same children also participated in a grammaticality judgment task, successfully, in order to determine whether they would accept the 2Aux questions they had produced in the elicited production sessions. From the series of three experiments, we have evidence that although the children produce 2Aux questions, they do not accept them. Based on this evidence, I argue that with respect to 2 Aux questions, the grammaticality judgment data are a better indication of the children's competence than the elicited production data.

### 1.3 Changing judgments in adults

Acceptability judgments can be unstable and unreliable, but the source of this instability can be minimized. Many researchers have studied how differences between subjects, as well as differences in experimental manipulations, can cause variability in acceptability judgments. (See Schütze (1996) for a detailed discussion of these factors.) Subject-related factors that have been argued to be relevant for the judgment task include handedness (Bever, Carrithers and Townsend (1987), Cowart (1989)), linguistic training (Spencer (1973), Ross (1979), Greenbaum (1988), Rose (1973), Snow and Meijer (1977), Valian (1982)), literacy training and general schooling (Birdsong (1989), Bialystok and Ryan (1985), Masny and d'Angeljan (1985), Scribner and Cole (1981), Scholes and Willis (1987), Heeschen (1978)). Some factors, such as fatigue and attentiveness, have nothing to do with the subject's linguistic representation, yet may affect judgments nonetheless (Bradac et al. (1980)).

Researchers have also found that experimental manipulations related to the procedure or the stimuli can also cause variability in judgments. Procedural factors that have been argued to influence judgments include instructions (Hill (1961), Carden (1970), Bley-Vroman, Felix and Ioup (1988), Cowart (1997), Van Kleeck (1982)), order of presentation (Greenbaum (1973, 1976), Elliot, Legum and Thompson (1969), Greenbaum and Quirk (1970)), repetition of the same sentences (Nagata (1988, 1987a, 1987b, 1989d), Carroll (1979)), and speed of judgment (Bialystok (1979), Warner and Glass (1987), Mistler-Lachman (1972)). Stimulus factors that have been argued to
influence judgments include context (Snow (1975), Crain and Steedman (1985), Pesetsky (1987), Levelt (1974, vol. 3), Bever (1970, 1974), Greenbaum (1976), parsability (Clark and Haviland (1974), Warner and Glass (1987), Van Kleeck (1982)), and lexical choice (Leveit et al. (1977), Greenbaum (1977).

When we encounter unstable judgments, one (or more) of these factors may be responsible. One phenomenon involving judgments that change deserves further investigation since it does not seem to be due to performance factors. Linguists have noticed anecdotally that certain island violations become increasingly acceptable after repeated exposure to the construction. This so-called "syntactic satiation" effect has been studied experimentally by Stromswold (1986) and Snyder (1994, 2000). Syntactic satiation cannot just be the result of performance factors since the types of sentences that are affected are constrained by the syntax. Not all types of ungrammatical sentences show satiation effects.

The study of satiation is slightly different from the repetition effects studied by Nagata. In his studies, Nagata investigates two possible reactions to repetition of sentences: judgments may become more lenient due to habituation, or they may become more stringent if subjects notice more things wrong with it. In Nagata (1988), the subject first judged a series of 48 sentences. Then during a repetition phase, each sentence was presented nine times in a row. During this time the subject was told to think about the grammaticality of the sentence. After a tenth presentation of the sentence, the subject was asked to judged the sentence (for the second time). Nagata
found that the grammaticality rating became better after repetition. However, in Nagata (1989d), he compared sentences marked with a question mark in the literature, and those marked with an asterisk. He found that both groups of sentences were judged worse after repetition. Nagata is concerned with the general effect of repetition on judgments, and does not distinguish between sentence types. The subject judges the same sentence twice, rather than similar sentences of a given sentence type.

In contrast, the satiation studies are concerned with repetition effects for specific syntactic constructions and the theoretical implications of the phenomenon. For example, in the study by $\operatorname{Snyder}(1994,2000)$, as well as in the experiments presented in Chapter 3, the subjects judged several different examples of Wh-islands, as in (6), rather than judging the same Wh-island example many times, as in (7).
(6) a. Who did Sandra wonder whether Howard had scolded?
b. Who did Patrick wonder whether Sally had seen?
c. What did Kathy wonder whether Larry had eaten?
d. What did Sigrid wonder whether Debbie had seen?
e. What did Morris wonder whether Jane had bought?
f. What did Kelly wonder whether Tina had read?
g. Who did Douglas wonder whether Susan had met?
(7) a. Who did Sandra wonder whether Howard had scolded?
b. Who did Sandra wonder whether Howard had scolded?
c. Who did Sandra wonder whether Howard had scolded?
d. Who did Sandra wonder whether Howard had scolded?
e. Who did Sandra wonder whether Howard had scolded?
f. Who did Sandra wonder whether Howard had scolded?
g. Who did Sandra wonder whether Howard had scolded?

In Chapter 3, I continue to investigate syntactic satiation experimentally to gain more insight into how the phenomenon is tied to syntactic constraints. In the first
experiment, I replicated Snyder's $(1994,2000)$ study and found satiation effects for Wh-islands, and also found that with additional blocks of items, a sentence type (Subject islands) that only showed marginal satiation effects in Snyder's study showed significant satiation effects. Furthermore, I investigated whether satiation on one island violation would trigger satiation on another island violation. One goal in searching for carry-over effects is to determine whether we can identify natural classes of island violations.

In a second experiment, I tested two additional island violations for satiation effects, and measured the subject's response times to look at whether changes in judgments are associated with changes in response time. The results suggest that there is no association between satiation and changes in response time. I also explored whether general reading ability and familiarity with reading is associated with the subject's ability to provide acceptability judgments (measured by how well the subject scores on the filler items). Again, [ found no such association.

The satiation effects for Wh-islands in Experiment 2 were not as strong as in Experiment 1, and the subjects in Experiment 2 did not perform as well on the filler items. Some possible reasons for the difference in results include the addition of control items in Experiment 2, and the presentation method of the stimuli (printed questionnaire for Experiment I, computer for Experiment 2). In order to determine whether the presentation method of the stimuli might be causing the difference in results, I conducted a follow-up experiment in which the same stimuli from

Experiment 2 were presented in a printed questionnaire. Although the subjects performed better on some of the filler and control items in the follow-up experiment, there was still no significant satiation effect for $\mathbf{W h}$-islands. We cannot rule out presentation method as a contributing factor, but it seems that the addition of the control sentences, which were minimally different from the test sentences, might be a more relevant factor.

## Chapter 2

## Children's questions with doubled auxiliary verbs

### 2.1 Introduction

Over the past three decades, children's use of auxiliary verbs in questions has been investigated by many researchers (including Hurford (1975), Kuczaj (1976), Maratsos and Kuczaj (1978), Stromswold (1990), Thomton (1993, 1995), Guasti et al. (1994, 1995), Allen (1995)). Researchers have observed that children acquiring English often produce non-adult sentences with doubled tense and/or agreement (henceforth T/Agr).
(8) Doubled T/Agr with one auxiliary verb (IAux)
a. Did you came home?
(Hurford (1975))
b. That didn't went down.
(Peter 2;8) (Stromswold (1990))
(9) Doubled T/Agr with two auxiliary verbs (2Aux)
a. Does it doesn't move?
(Nina 2;9) (Stromswold (1990))
b. What did he didn't wanna bring to school? (Darrell $4 ; 1$ ) (Guasti et al. (1995))

Some utterances have T/Agr marked on one auxiliary verb as well as the main verb (henceforth IAux), and both have the correct tense and/or agreement inflection, as in (8). These cases seem to occur primarily with irregular verbs (Kuczaj (1976), Maratsos and Kuczaj (1978), and Maratsos (1984)). In other cases, particularly in negative questions, the utterance has two auxiliary verbs and T/Agr is marked on both auxiliary verbs (henceforth 2Aux), as in (9).

In the past, there have been two main proposals for accounting for 2Aux constructions. Hurford (1975), Fay (1978) and Mayer, Erreich and Valian (1978) have argued that children are failing to perform the "delete" operation of "copy and delete". Others, such as Maratsos and Kuczaj (1978) and Goodluck and Solan (1979) argue against such a "copy without deletion" analysis. They observed that many IAux questions occurred with irregular verbs, and that many 2Aux questions occurred with 'didn't'. Therefore, they argue that children do not know that irregular past tense verbs (in IAux questions) or do $+n$ 't (in 2Aux questions) are tensed elements, and hence that for the child these doubled T/Agr constructions in fact are only marked once for tense and agreement.

More recently, Stromswold (1990), Thomton (1993, 1995), Guasti et al. (1994, 1995) have studied 2Aux questions. Stromswold (1990) investigated spontaneous speech data from the CHILDES database and found that children produce very few 2Aux negative questions. She argues that children know that tense can only be marked once in a matrix clause, and she adopts the previous analyses discussed above proposed by researchers such as Maratsos and Kuczaj (1978) to explain these apparent counterexamples.

Thornton (1993, 1995) and Guasti et al. (1995) used elicited production tasks and discovered that children ( $3 ; 0-5 ; 0$ ) produced negative 2 Aux questions. Based on these results, Thomton and Guasti et al. account for these non-adult productions by proposing that the child has a grammar that is different from the aduit grammar.

Thornton (1993) proposes that children produce 2Aux questions because of a constraint against an auxiliary supporting more than one affix, such as the Q or Neg morpheme. Guasti et al. (1995) propose a "copy without deletion" analysis.

One potential problem with some of these studies is that they relied solely on production data. Although it is very common in the study of language acquisition to rely on production data as a way to discover something about the child's grammar, this may be over-estimating or under-estimating the child's grammar. In order to overcome the limitations of production data, we can supplement it by using other methodologies. For example, Thornton (1995) used a comprehension task in addition to the production task, and discovered that although children seldom produce grammatical negative questions with do+n't, such as (10), they are able to understand them.
(10) What didn't the smurf buy?

Although Stromswold (1990) used a grammaticality judgment task, in addition to spontaneous speech data, to investigate children's IAux constructions, none of these previous studies have tested children's acceptability judgments of 2 Aux questions.

In this study, the same group of children participate in both an elicited production task and grammaticality judgment task to investigate 2 Aux questions. ${ }^{\text {I }} \mathrm{By}$ combining production data with acceptability judgment data we can obtain a better

[^0]picture of the child's grammar with respect to 2Aux questions. Although acceptability judgments are the primary source of data for studying the adult grammar, they are seldom used with children. However, McDaniel and her colleagues have had success using this task with children as young as 2;11 (McDaniel, Chiu, and Maxfield (1995)).

There are at least three possible outcomes from using both production and judgment data. First, children may judge what they produce to be acceptable. This would be an expected result since we typically produce grammatical sentences, with the exception of speech errors. Another possible result is that children accept sentences that they do not produce. The judgment data in this scenario are important since they provide additional data that were previously unavailable. A third possible outcome is that children judge what they produce to be unacceptable.

This third case is exactly what we find from the three experiments in this study. Some of the children produced 2Aux questions yet nonetheless judged them to be unacceptable. This is an interesting case to investigate since we assume that we normally produce only grammatical sentences. In the current case, we are faced with the task of determining whether it is the production data or the judgment data that reflects the child's grammar. In this study, the judgment data are consistent with the adult grammar. I argue that the children have the adult grammar, based on their judgments, and that their production of 2Aux questions is the result of incorrect lexical information regarding constituent negation.

I first review some recent analyses that have been proposed to account for 2Aux questions in section 2.2. In section 2.3, I explain the methodology and present results from the first experiment, which elicits positive and negative questions from children learning English. In section 2.4, I present methodology and results from the second experiment, which elicits positive and negative questions with emphatic stress. In section 2.5 , I present the methodology and results from the third experiment, which asks children to judge positive and negative questions, including 2Aux questions. In section 2.6, I discuss the results from the three experiments, and propose an analysis to account for the production/judgment asymmetry.

### 2.2 Previous studies of 2Aux questions

### 2.2.I Stromswold (1990)

Stromswold (1990) examined spontaneous speech transcripts from the CHILDES database (McWhinney and Snow 1985). She investigated transcripts of fourteen English-speaking children, between the ages of 0;11 and 2;10 when they were first recorded, and between $0 ; 11$ and $7 ; 10$ at the end of recording. She excluded any utterances that included an unclear or stuttered auxiliary verb, or a contracted auxiliary verb.

Stromswold found some 2Aux questions, they were produced infrequently. Of the 40,600 questions she examined, she found only seven negative 2 Aux questions, as in (11), and six positive 2Aux questions, as in (12).
(II) a. Did I didn't meant to?
(Adam 3;4)
b. Did I didn't mean to?
(Adam 3;5)
c. Do she don't need that one?
(Adam 3;6)
d. Did I didn't mean to do that?
(Adam 3;6)
e. Why does Superman doesn't wear underroos on his bottom?
(Ross 3;3)
f. Does it doesn't move?
(Nina 2;10)
g. Do they don't eat people up?
(Nina 2;10)
(12) a. Why did you did scare me?
(Nina 3;2)
b. Is my old baby blanket is clean?

Only one of these six positive 2Aux questions involved do as in (12)a. The other five questions involved $b e$, as in (12)b. From the context, it seems that Nina might have been trying to use emphasis in the positive 2Aux question involving do.

Give the fact that only three children produced negative 2Aux questions, and did so in a relatively short time period (two months for Adam and one day for Nina and Ross), Stromswold does not treat these productions as "real" cases of doubled tense. Instead, she concludes that these children have probably misanalyzed do $+n$ ' $t$ as untensed negation elements, as proposed by Maratsos and Kuczaj (1978) and Maratsos (1984).

One problem with proposing that the lower auxiliary verb in negative 2Aux questions is a frozen negative form is that then we would expect didn't, doesn't, and don't to be in free variation as the lower negative element. However, this is not the
case. The data show that the lower auxiliary matches the higher auxiliary for tense and/or agreement.

Although 2Aux questions are limited in spontaneous speech data, they are commonly found in experimental situations. Let us now turn to two elicited production studies.

### 2.2.2 Thomton (1993)

Thornton (1993) investigated children's questions by studying nine children ( $4 ; 3$ to $5 ; 0$ ) using an elicitation task. Of the nine children, three had adult forms of both positive and negative questions. The remaining six children produced adult-like positive questions and subject extraction questions, but non-adult negative questions. $87 \%$ of these six children's positive questions were adult-like, whereas only $21 \%$ of their negative questions were adult-like.

For the remainder of the study, she focused on one child, Curly $(4 ; 8)$, who produced very consistent data over six sessions, during which she elicited 173 questions, both positive and negative. She produced mainly 2Aux questions for object extraction ( $88 \%$ of object extraction questions) and non-inversion questions for adjunct extraction ( $75 \%$ of adjunct extraction questions).

Thornton found three types of non-adult negative questions: negative 2Aux questions, questions without inversion, and questions with not. ${ }^{2}$
(13) a. What do you don't eat?
b. Why you don't like cheese?
c. What do you not like?

She argues that all three types of non-adult questions are the result of a constraint on the child's grammar. She proposes that children prohibit $T$ from supporting more than one affix in addition to the tense affix itself, such as a question morpheme $Q$, or a negative affix, $n^{\prime} t$. Given this constraint, when both a question morpheme and a negative affix must be supported, one of these three non-adult constructions is produced.

One strategy that a child can use is to provide two different hosts, one to support each affix, thereby producing a 2Aux question. For example, consider the question in (14).
a. What does he doesn't like?
b. [CP what ${ }_{1}\left[\mathrm{C}^{\prime}\right.$ does $_{3}-\mathrm{Q}$ [AGRsP he [AGRs $^{\prime} \mathrm{t}_{3}\left[\mathrm{TPP}_{\mathrm{t}} \mathrm{t}_{3}\left[\right.\right.$ AspP doesn't $_{2}\left[\mathrm{NegP}^{\mathrm{t}_{2}}\right.$ [vp like $t_{\mathrm{L}}$ 171]

Under Thomton's analysis both the question morpheme, Q , and the negative affix, $n^{\prime} t$, must be supported by its own lexical item. She assumes that Infl is divided into

[^1]several fumctional projections: AgrP, TP, AspP, ${ }^{3}$ and NegP. The higher do is inserted in Tense and raises to $C$ via AGRs to support the question morpheme. The negative affix is supported by a different lexical item. Asp is morphologically realized as do to support the negative affix, which raises to Asp from its base-generated position in Neg.

A child may also decide to host just one of the two affixes, either the question morpheme or the negative affix, rather than trying to host both. If a child uses the "full form" of negation, she will produce a question with not, as shown below.
(15) a. What does he not like?
b. [CP what ${ }_{l}$ [ $C^{\prime}$ does $_{2}-\mathrm{Q}$ [AGRsP he [AGRs' $\mathrm{t}_{2}$ [TP $\mathrm{t}_{2}$ [NegP not [VP like $\mathrm{t}_{\mathrm{l}}$ [I]J]I]

In (15), do is inserted in Tense and raises to $C$ via AGRs to support the question morpheme, similar to the 2Aux example given in (14). However, under this strategy, T only has to host one affix, Q , since not does not need a host. ${ }^{+}$

A child may also decide to just host the negative. If a child chooses this strategy, she will produce a non-inverted question, as shown below.
a. What he doesn't like?
b. [CP what ${ }_{1}\left[C^{\prime}\right.$ Q [AGRsP he [AGRs ${ }^{\prime}\left[\text { does }+\mathrm{n}^{\prime} \mathrm{t}_{2}\right]_{3}\left[\mathrm{TP} \mathrm{t}_{3}\left[\mathrm{NegP} \mathrm{t}_{2}\right.\right.$ [vp like $t_{1}$ I]]I]I]]

[^2]In (10), the child has chosen to satisfy the negative affix but not the Q morpheme. Hence, T does not raise to C to support Q and instead remains low to support $n$ 't.

A major problem with Thornton's analysis is her classification of questions with not as non-adult. Descriptively, questions with not are not necessarily negative questions using the "full form" of negation; there are restrictions on when not and the negative affix $n^{\prime} t$ can be used. In analyzing negative questions, and any structure involving negation, a distinction needs to be made between two types of negation: sentential negation and constituent negation. Consider the examples of each type, given below.
(17) a. The smurf could not buy the strawberries. [constituent negation] 'The smurf has the option of buying or not buying the strawberries.'
b. The smurf couldn't buy the strawberries. [sentential negation] 'It was not possible for the smurf to buy the strawberries.'

In (17)a, not (with emphasis and a pause before it) is negating the act of buying, while in (17)b, the affix $n$ 't is negating the possibility of buying. The use of not is appropriate and grammatical in adult English given a situation involving constituent negation. (See section 2.6 .2 for a more detailed discussion of sentential and constituent negation.)

One of the weaknesses of Thomton's analysis is that it does not explain why both auxiliary verbs in 2Aux questions have the correct tense and agreement inflection. Like Stromswold, Thomton's analysis predicts free variation of the lower negative element. In addition, there is a problem with the child's constraint which
prohibits T from hosting more than one affix. If this is true for children, they should not be allowing $T$ to host AGRs in addition to the question morpheme.

There is also another problem with respect to the analysis of non-inverted negative questions. The Q morpheme requires a host to support it. Under normal circumstances, the $Q$ morpheme in $C$ is supported by $T$, which raises to $C$. When children produce a non-inverted negative question, $T$ does not raise to $C$, and the $Q$ morpheme has no host. However, these same children are correctly inverting positive questions, which suggests that they do satisfy the Q morpheme when producing positive questions. It is not obvious why children who know that the Q morpheme requires a host for positive questions violate this requirement in negative questions. Thomton is unable to account for this asymmetry found in positive vs. negative questions.

### 2.2.3 Guasti, Thomton and Wexier (1995)

Guasti et al. (1995) also used elicitation tasks to study negative declaratives and questions (positive and negative) of ten children ( $3 ; 8$ to $4 ; 7$ ). The target questions included Yes/No questions and Wh-questions with subject, object and adjunct extraction. They used negative declaratives and positive questions as controls, and
these were all adult-like. The negative questions, however, were not adult-like. ${ }^{5}$ The findings are similar to Thornton's (1993) study. The children produced negative 2Aux questions, questions without inversion, and questions with not.
(18) a. What did he didn't wanna bring to school? (Darrell 4;1)
b. How can Emie can't sit?
(Emily 4;2)
Guasti et al. (1995) also found two variants of the constructions in (18). First, instead of the two auxiliary verbs with just one negative element, the Neg \& Aux doubling variant ( $2 \mathrm{Aux} / \mathrm{Neg}$ ) has two auxiliary verbs and two negative elements, as shown in (19). ${ }^{6}$
a. Why can't she can't go underneath? (Kathy 4;0)
b. What didn't Miss Pigo don't like to do? (Matt 4;3)
b. What didn't Miss Piggy don't like to do? (Matt 4;3)

The fact that the higher auxiliary in 2Aux questions is always inflected correctly with respect to the subject and tense of the situation suggests that children are correctly inverting the auxiliary verb (performing subject auxiliary inversion or SAI).

The second variant of the 2Aux question is one where the tense/agreement of the two auxiliary verbs are not identical to each other ( 2 Aux -m ) ${ }^{\top}$ Recall that Thomton (1993) found 2Aux questions with both auxiliary verbs with the same the

[^3]tense/agreement inflection $\left(2 A u x_{+m}\right)$. The pattems Guasti et al. observed for agreement (3rd person singular) and tense (past) are shown below.

| (20) | Agreement ${ }^{8,9}$ | does <br> does | doesn't <br> don't | $\begin{aligned} & \text { \# } \\ & \# \end{aligned}$ | $\begin{aligned} & \text { do } \\ & \text { do } \end{aligned}$ | doesn't <br> don't |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| (21) | Tense | did | didn't |  | could | couldn't |
|  |  | did | don't | \# | could | can't |
|  |  | \# do | didn't | \# | can | couldn't |
|  |  | \# do | don't |  |  |  |

With respect to agreement, they found that it was always correctly inflected on the higher auxiliary but not necessarily on the lower one, as shown in (22).
(22) Why does the snail don't have eyelashes? (Matt 4;3) (Guasti et al. (1994)) With respect to tense, there seems to be a contrast between main verbs and modals. In negative questions with main verbs, tense was always correctly inflected on the higher auxiliary, but not necessarily on the lower auxiliary verb, as shown in (23).
(23) What did he don't like? ${ }^{10}$
(Guasti et al. (1994))
This pattern is the same as the one observed for agreement. However, in negative questions with modals, the tense on both the higher and lower auxiliary was always correct, as in (24). ${ }^{\text {II }}$

[^4](24) a. Why could Snoopy couldn't fit in the boat? (Kathy 4;0)
b. \#Why could Snoopy can't iit in the boat?

Guasti et al., like Thornton (1993), try to account for the same three types of non-adult negative questions, 2Aux, non-inversion, and not, as well as another nonadult negative question type, 2Aux/Neg. Guasti et al. suggest that all four non-adult question types are the result of the interaction of two well-formedness conditions on questions and negation: the Wh-Criterion and the Neg-Criterion, given below.
(25) Wh-Criterion (May 1985, Rizzi 1991)
a. A wh-operator must be in a spec-head relation with a [+wh] head.
b. A $[+w h]$ head must be in a spec-head relation with a wh-operator.
(26) Neg-Criterion (Rizzi 1991; Haegeman and Zanuttini 1991)
a. A neg-operator must be in a spec-head relation with a [+neg] head.
b. A [+neg] head must be in a spec-head relation with a neg-operator.

Let us first take a look at how the Wh-Criterion applies to questions. Following Rizzi (1991), Guasti et al. assume that the [ +wh$]$ feature is generated in I. In order for the [ + wh] feature to be in a spec-head relation with a Wh-operator, I must raise to C , and the Wh-phrase must raise to [spec, CP]. In this way, the Wh-Criterion forces I to C movement (or SAl). Consider the following example.
(27) a. What did the smurf buy?
b. [CP what ${ }_{1}\left[C^{\prime} \quad I_{2} \quad\left[\mathbb{P}\right.\right.$ the smurf $f_{3}\left[r^{\prime} \mathrm{t}_{2}\left[\mathrm{vp}_{\mathrm{t}} \mathrm{t}_{3}\right.\right.$ buy $\left.\left.\left.\left.\left.\mathrm{t}_{1}\right]\right]\right]\right]\right]$ [ +wh ]

[^5]In (27), the Wh-operator, 'what' in [spec, CP], and the [ +wh ] head, Infl raised to C , are in a spec-head relation, thus satisfying the Wh-Criterion.

Let us now take a look at a negative sentence. Consider the following sentences.
(28) a. The smurf didn't buy strawberries.
b. [IP the smurfill $\left[r^{\prime}\right.$ did $+n^{\prime} t_{2}\left[{ }_{\text {NegP }} \mathrm{Op}\left[\mathrm{Neg}^{*} \mathrm{t}_{2}\left[\mathrm{VP}_{\mathrm{I}} \mathrm{t}_{\mathrm{l}}\right.\right.\right.$ buy strawberries $\left.\left.\left.]\right] 7\right]\right]$

(29) a. The smurf did not buy strawberries.
b. [ip the smurfil [r' did [NegP Op [Neg' not [vp $\mathrm{t}_{\mathrm{I}}$ buy strawberries]]]]] In the negative declarative in (28), the negative affix $n$ 't has raised to Infl. ${ }^{12}$ The NegCriterion is satisfied within the IP by the null negative operator in [spec, NegP], and the chain formed by $n ' t$ and its trace in Neg, which are in a spec-head relation. In the negative declarative in (29), not remains in Neg. The Neg-Criterion is satisfied within NegP by the null negative operator in [spec, NegP] and not in Neg.

In an adult-like negative question, such as in (30), we find that the NegCriterion can be satisfied within the $C P$.
(30) a. What didn't the smurf buy?
 [+wh]

[^6]The Wh-operator, what, in [spec, CP] and the [ + wh] head in C are in a spec-head relation, and the Wh -Criterion is satisfied. The negative affix $n$ 't raises to Infl, and Infl $+n$ 't in turn raises to $C$. The Neg-Criterion is satisfied between the null operator in [spec, NegP] and the chain headed by $n^{\prime} t$ in C.

Guasti et al. propose that all four types of non-adult negative questions (2Aux, non-inversion, not, and $2 \mathrm{Aux} / \mathrm{Neg}$ ) reveal a stage of grammatical development in which the Neg Parameter is set at the default (positive) setting, which is incorrect for English.
(31) Neg Parameter (Guasti 1994)

The Neg-Criterion must be satisfied in the V-related (IP) projection. $\{1,0\}$

They argue that children produce grammatical negative declaratives and non-adult negative questions (2Aux, non-inversion, not, and 2Aux/Neg) since in all of these cases the Neg-Criterion is satisfied within IP. Once children reset the parameter to zero so that the Neg-Criterion is satisfied within CP, they will produce adult negative questions as in (30). In the next sections, let us consider their derivations for each of the four non-adult negative question types.

### 2.2.3.1 Negative questions with not

Consider the following negative question with not.
(32) a. What did the smurf not buy?
 [ +wh ]

In (32), the Wh-Criterion is satisfied by the Wh-operator, what, in [spec, CP] and the [ +wh ] head in $C$. The Neg-Criterion is satisfied within the $\mathbb{P}$ by the null negative operator in [spec, NegP] and not in Neg.

Like Thomton (1993), Guasti et al. incorrectly treat questions with not as nonadult. However, as discussed in section 2.2.2, questions with not are a grammatical structure for constituent negation. (See section 2.6.2 for further discussion.)

### 2.2.3.2 Negative questions with no inversion

Consider now a negative question with no inversion.
(33) a. What the smurf didn't buy?
 [ +wh ]

The Neg-Criterion is satisfied within IP by the null negative operator in [spec, NegP] and the negative affix $n$ 't, which remains in Neg. In this sentence, Infl has not raised to C . Although the [twh] head, Infl, is not in a spec-head relation with a Whoperator, Guasti et al. argue that this does not violate the Wh-Criterion. They claim that the Wh-Criterion is satisfied by 'dynamic agreement', which is grammatical for some languages, such as French. Under dynamic agreement, "the head in Comp receives wh-features from the wh-operator instead of from the inflection which is raised to C" (Guasti et al. 1995: 235-236).

If dynamic agreement is available to the child, it isn't clear why she wouldn't prefer it over raising Infl to $C$, since dynamic agreement seems to be a more economical way of satisfying the Wh-Criterion. Furthermore, it isn't clear how the child would leam that dynamic agreement is not an option in English. There is also the problem of the positive/negative asymmetry in the production of non-inverted questions. Most children of this age produce adult (i.e. inverted) positive questions and non-inverted negative questions, among other types of non-adult negative questions. It is curious that children use dynamic agreement in negative questions, but not in positive questions.

### 2.2.3.3 Negative 2Aux questions

Let us now turn to negative 2Aux questions. Consider the following example.
(34) a. What did the smurf didn't/don't buy?
 [ +wh ]

The Wh-criterion is satisfied by the Wh-operator, 'what', in [spec, CP] and the [+wh] head in C. In order to satisfy the Neg-Criterion within IP, the child has kept the negative affix in Neg, similar to the non-inversion case. The Neg-Criterion is satisfied by the null negative operator in [spec, NegP] and $n$ 't in Neg. In order for the negative affix to have a host, the trace of Infl is spelled out instead of being deleted. The trace
of Infl may be spelled out with "the same or fewer features," thus accounting for both the matching and non-matching auxiliary verbs.

Guasti et al. do not motivate why children would spell out a copy with fewer features. ${ }^{13}$ If the child does not know the "delete" portion of the "copy and delete" process, the null hypothesis would be that the copy of the moved element would be spelled out with identical features. A better approach might be to say that the copy of the moved auxiliary can be spelled out as an auxiliary with the same features, or as an unspecified 'do'.

It is crucial to the analysis that only the auxiliary trace be spelled out, since no other traces are spelled out. However, if we assume that a child does not know the general process of deleting traces of movement, we would expect the child to spell out the trace of any moved element to produce utterances such as (35), among other unattested forms.
(35) What does he doesn't like what?

In (35), the copy of 'what' is spelled out. Since children don't produce utterances like this, a strict "copy without deletion" analysis must rely on the stipulation that only auxiliary traces are spelled out.

However, even this stipulation doesn't correctly generate the child's utterances since it incorrectly predicts that children produce sentences like.
a. What is he isn't is eating?
b. [CP what [ c ' is [ $[\mathrm{P}$ he [ r isn't [ vp is eating]]]]]


In this example, the copy of 'is' in $V$, after 'is' is raised from $V$ to $I$, is spelled out. Even though only auxiliary traces are spelled out, this kind of utterance is not produced by children. The stipulation that only auxiliary traces are spelled out still incorrectly allows sentences like (36).

### 2.2.3.4 2Aux/Neg questions

Guasti et al. argue that $2 \mathrm{Aux} /$ Neg constructions are a transitional form resulting from an attempt to satisfy both the default and adult English setting (i.e. zero setting) for the Neg Parameter. Consider the following example.
(37) a. What didn't the smurf didn't buy?



The Wh-Criterion is satisfied by the Wh-operator in [spec, CP] and the [+wh] head in C. To satisfy the Neg-Criterion outside P , the child raises the negative affix $n^{\prime} t$ to $I$, which raises to C. The child also wants to satisfy the Neg-Criterion within $\operatorname{IP}$, so she spells out the trace of $\left[1+n^{\prime} t\right]$.

[^7]Let us assume that the child does want to satisfy both the default and zero setting. The Neg Parameter is repeated here as (38).
(38) Neg Parameter (Guasti 1994)

The Neg-Criterion must be satisfied in the V -related (IP) projection. $\{1,0\}$ The way in which the Neg Parameter is stated suggests that it is a "light switch" type parameter: the parameter is either on or off. If the Neg Parameter is "on", the positive or default setting is chosen, and the Neg-Criterion must be satisfied within IP. However, if the Neg Parameter is "off", the zero setting is chosen. Presumably, in this "off" setting, the Neg Parameter should now be irrelevant, and hence where the NegCriterion is satisfied should be irrelevant as weil.

If we assume this "light switch" interpretation of the Neg Parameter, it is difficult to imagine a transitional stage which would account for 2Aux $N$ Neg questions. The parameter should be either "on" or "off." Even if the child were able to switch back and forth between the two settings, there is nothing that would force her to satisfy the Neg-Criterion outside IP and simultaneously inside IP. Satisfying the NegCriterion within IP alone is always an option for the zero setting of the Neg Parameter.

### 2.2.4 Summary

Although Stromswold (1990) found only a limited number of 2Aux questions in spontaneous speech, Thormton (1993) and Guasti et al. (1994, 1995) elicited many 2Aux questions in their experimental studies. There were some differences in the
details of the data. Stromswold (1990) found both positive and negative 2Aux questions, although limited in number, and Guasti et al. (1994) found negative 2 Aux $\mathrm{m}_{\mathrm{m}}$ questions.

In order to obtain a clearer picture of what the T/Agr matching pattems are for 2Aux questions, and whether positive 2Aux questions are produced, we investigate 2Aux questions using both elicited production and grammaticality judgment tasks. The disagreeing data regarding 2Aux questions makes it necessary to do more than just a replication of the previous elicitation studies.

With elicited production tasks or spontaneous speech studies, we can examine only what the child produces. We are unable to determine anything about utterances that the child does not produce. Although the elicited production task is designed to target certain sentence structures that may occur only rarely in children's spontaneous speech, they may still not produce that structure. Recall the although Guasti et al. (1995) found 2Aux $x_{m}$ questions, Thornton (1995) did not. By using a judgment task in conjunction with an elicited production task, we may be able investigate in more detail the patterns of tense and agreement in doubled T/Agr constructions. Even if a child does not produce any $2 A u x_{-m}$ questions, we can still have her judge $2 A u x_{-m}$ questions.

We also examine whether emphatic stress induces 2Aux questions in positive enviromments. Previous experimental studies have found that the production of 2Aux questions occurs primarily with negation, and it has been assumed that negation is the cause of 2Aux questions. However, Stromswold (1990) found a few cases of positive

2Aux questions (one with contextual indication for emphasis) and there is some anecdotal evidence for the production of positive 2Aux questions with emphatic stress. If emphatic stress is relevant for positive 2Aux questions, we may be able to elicit such questions using an elicited production task.

In the next three sections, 2.3-2.5, I discuss the two elicited production experiments I conducted, followed by the acceptability judgment experiment. The first elicited production experiment replicates the main findings by Thornton (1993, 1995) and Guasti et al. (1994, 1995). However, we did not find any 2 Aux $_{-\mathrm{m}}$ questions. The results from the second elicited production experiment show that children do not produce positive 2Aux questions, even when emphatic stress is involved. The grammaticality judgment task results were different from what we had expected: children did not judge the non-adult 2Aux questions they produced to be acceptable. We discuss the implications of this contrast in section 2.6.

### 2.3 Experiment 1: Children's production of 2Auxx questions

### 2.3.1 Methodology

In order to determine what types of positive and negative questions children produce, I used an elicited production task. Before starting any of the experiments, we played with the children at school several times to get to know them. Once the children were comfortable with us, we invited them to "play the games" with us.

Two experimenters were involved in the elicited production task: one experimenter played the role of a puppet, a colorful bird named CanTou, and the other manipulated toys and told stories which were designed to be felicitous for the production of questions. During the story, the puppet, who loves guessing games, was told to hide so that she would not see the entire story. The child then watched the rest of the story. At the end of the story, the storyteller prompted the child to ask the puppet a question so that the puppet could guess what had happened in the story. In response to the child's question, the puppet made a guess about what happened. The answer that the puppet gave was randomly varied for correctness. The child's role in the game was to feed the puppet either a donut, if the puppet guessed correctly, or an apple to "make him smarter," if he guessed incorrectly. The methodology hides the fact that the child is being tested; the child's role is to help the puppet play the guessing game and to evaluate his guesses. A sample story is shown below.
(39) Exp: This is a story about going outside in the cold.

Bart: I want to go outside, but it's really cold outside. I have to dress warm. I don't want to freeze outside. Let's see what I can put on. I have a brown scarf with flowers, and a purple hat. I want my neck to be warm, so I'm going to put on my scarf. (puts on scarf) Maybe I'll wear the hat too. It'll keep my head warm.
Exp: Okay CanTou, close your eyes. You can guess something about the story in a little bit.
Bart: (tries on hat) Oh no! It doesn't fit. I guess I won't wear the hat. It's too small for my head.
Exp: (to child) We know Bart doesn't wear the hat for a reason. Ask CanTou why.
Target: Why doesn't Bart wear the hat? / Why does Bart not wear the hat?
CanTou: Because it's too small. (correct answer)

We also used conversational situations where the experimenter prompted the child to ask the puppet a question about various characteristics pertaining to the puppet. An example of this type of situation is given below.
(40) Exp: I heard CanTou eats bugs, but I'm not sure. Ask him if he does? Target:Do you eat bugs?
(41) Exp: CanTou looks hungry. I bet he wants to eat something. Ask him what? Target: What do you want to eat?

The conversation situations were used mainly to elicit yes/no questions, since it was difficult to create plausible stories to ask negative yes/no questions. This situation was also used to elicit Wh-object questions from children who had difficulty producing such questions in a story situation. Some children produced Wh-adjunct questions with why instead of the targeted Wh-object questions. ${ }^{\text {I4 }}$

I elicited positive and negative questions with Wh -object and Wh-adjunct extraction, as well as positive and negative yes/no questions. At least two items for each question type were elicited from each subject. ${ }^{15}$ Sample target questions are shown below.

[^8](42) Sample positive question targets
a. Who did the tiger chase?
b. Where did Jasmine hide the jewels?
c. Do you eat bugs?
(43) Sample negative question targets
a. What didn't the smurf buy? / What did the smurf not buy?
b. Why doesn't Bart wear the hat? ! Why does Bart not wear the hat?
c. Didn't you sleep last night? / Did you not sleep last night?

### 2.3.2 Subjects

Fifteen native English speaking children attending the Child Development Laboratories at the University of Connecticut participated in the study. ${ }^{16}$ At the beginning of the elicited production sessions, the children ranged in age from $3 ; 10$ to $4 ; 11$. The children were tested individually in a room separate from their classroom. Each session lasted no more than thirty minutes, and each session was audio taped for later transcription. We had some difficulty eliciting the complete set of questions from five of the younger children ( $4 ; 5$ and under) in the first session, so they participated in a second session. Each child was given a pseudonym which will be used to refer to him throughout the discussion of the results.

We also tested four college-aged native English speakers as adult controls. One experimenter performed the role of both the storyteller and the puppet, but the

[^9]rest of the methodology remained the same. We told the same stories using toys, and prompted the adults using the same lead-ins.

### 2.3.3 Results from the adults

The adult subjects produced a total of 84 questions, and all of the questions produced were grammatical questions, as expected. ${ }^{17}$ Of the 41 negative questions, sixteen (39\%) were with the full form of negation, not, and 25 (61\%) were with the contracted form of negation. These results support the treatment of questions with the contracted form of negation as the preferred structure for adults, as was assumed by Thornton (1993, 1995) and Guasti et al. (1995). ${ }^{18}$ However, I will also treat questions with not as being grammatical.

### 2.3.4 Results from the children

The children produced a total of 207 questions: 127 positive questions and 80 negative questions. ${ }^{19}$ Of the 127 positive questions, 116 (91\%) were grammatical and

[^10]II (9\%) were ungrammatical. Of the 80 negative questions, 38 (48\%) were grammatical and 42 (53\%) were ungrammatical. (See graph in (44).) Children produced both the contracted form and the full form of negation in the grammatical negative questions.
(44) Figure 1: Children's question production


The percentage of ungrammatical questions in positive vs. negative contexts is significant by paired $t-\operatorname{test}(t(14)=4.031, p=.001)$. The mean percentage of ungrammatical positive and negative questions are $8 \%$ and $50 \%$, respectively. The children produced significantly more ungrammatical questions with negative questions than with positive questions. This result is consistent with previous findings from Guasti et al. (1995) and Thomton (1995).

This pattern for the number of grammatical vs. ungrammatical questions for positive and negative questions also holds for each sentence type, as shown in the graph below.
(45) Figure 2: Children's question production by sentence type


There were more grammatical than ungrammatical positive yes/no questions, $92 \%$ ( $55 / 60$ ) vs. $8 \%(5 / 60)$, and there were an equal number of grammatical and ungrammatical negative yes/no questions, $50 \%(8 / 16)$ vs. $50 \%(8 / 16)$. There were more grammatical than ungrammatical positive Wh-object questions, $97 \%$ ( $33 / 34$ ) vs. $3 \%$ ( $1 / 34$ ), and there were slightly fewer grammatical than ungrammatical negative Wh-object questions, $43 \%$ (13/30) vs. $57 \%$ (17/30). There were more grammatical than ungrammatical positive Wh-adjunct questions, $85 \%$ (28/33) vs. $15 \%$ (5/33), and
an equal number of grammatical and ungrammatical negative Wh-adjunct question, $50 \%(17 / 34)$ vs. $50 \%(17 / 34)$.

The types of positive and negative questions produced by each child are shown in the tables in (46) and (47).
(46) Table 1: Types of positive questions produced by children

| Child | Adult-like | IAux | non-inv. | 2Aux $_{+\mathrm{m}}$ | 2Aux $_{\mathrm{m}}$ | *Agr | Total |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Andrew | $9 \quad(90 \%)$ |  |  |  |  | $1(10 \%)$ | 10 |
| Anne | $10(83 \%)$ |  |  |  |  | $2(17 \%)$ | 12 |
| David | $12(100 \%)$ |  |  |  |  |  | 12 |
| Jennifer | $8(80 \%)$ |  | $1(10 \%)$ |  |  | $1(10 \%)$ | 10 |
| Judy | $5(83 \%)$ | $1(17 \%)$ |  |  |  |  | 6 |
| Kendra | $9(90 \%)$ | $1(10 \%)$ |  |  |  |  | 10 |
| Kim | $6(86 \%)$ | $1(14 \%)$ |  |  |  |  | 7 |
| Kristen | $5(100 \%)$ |  |  |  |  |  | 5 |
| Linda | $11(100 \%)$ |  |  |  |  |  | 11 |
| Marilyn | $5(83 \%)$ |  |  | $1(17 \%)$ |  |  | 6 |
| Maureen | $8(89 \%)$ |  |  |  |  | $1(11 \%)$ | 9 |
| Norbert | $6(100 \%)$ |  |  |  |  |  | 6 |
| Phillip | $8(89 \%)$ | $1(11 \%)$ |  |  |  |  | 9 |
| Rita | $8(100 \%)$ |  |  |  |  |  | 8 |
| Scott | $6(100 \%)$ |  |  |  |  |  | 6 |
| Total | 116 <br> $(91 \%)$ | 4 <br> $(3 \%)$ | 1 <br> (1\%) | 1 <br> $(1 \%)$ | 0 | 5 | 127 |

IAux = T/Agr marked on auxiliary verb and main verb
non-inv. = no "subject-auxiliary inversion"
$2 A u x_{+m}=$ two auxiliary verbs marked with the same T/Agr
$2 A u x_{\mathrm{m}}=$ two auxiliary verbs marked with different T/Agr
*Agr $=$ incorrect subject-verb agreement
(47) Table 2: Types of negative questions produced by children

| Child | not | n't | Adultlike | non-inv. | 2 Aux $_{+m}$ | 2Aux-m | 2Aux diff. aux | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Andrew | 3 |  | 3 (75\%) |  |  | 1 (25\%) |  | 4 |
| Anne |  |  | 0 (0\%) | 4 (57\%) | 3 (43\%) |  |  | 7 |
| David | 4 | 1 | 5 (71\%) |  | 2 (29\%) |  |  | 7 |
| Jennifer |  |  | 0 (0\%) |  | 6 (86\%) |  | 1 (14\%) | 7 |
| Judy |  |  | 0 (0\%) |  | 4 (80\%) |  | 1 (20\%) | 5 |
| Kendra |  |  | 0 (0\%) | 3 (75\%) | 1 (25\%) |  |  | 4 |
| Kim | 2 | 3 | 5 (83\%) |  | 1 (17\%) |  |  | 6 |
| Kristen | 4 |  | 4 (80\%) |  | 1 (20\%) |  |  | 5 |
| Linda | 1 |  | 1 (17\%) |  | 4 (67\%) |  | 1 (17\%) | 6 |
| Marilyn | 5 |  | 5 (83\%) |  | 1 (17\%) |  |  | 6 |
| Maureen | 1 | 1 | 2 (67\%) |  |  | 1 (33\%) |  | 3 |
| Norbert | 2 | 4 | 6(100\%) |  |  |  |  | 6 |
| Phillip |  | 4 | 4(100\%) |  |  |  |  | 4 |
| Rita |  | 3 | 3 (75\%) |  | 1 (25\%) |  |  | 4 |
| Scott |  |  | 0 (0\%) |  | 6(100\%) |  |  | 6 |
| Total | 22 | 16 | $\begin{array}{\|l} \hline 38 \\ (48 \%) \end{array}$ | $\begin{aligned} & \hline 7 \\ & (9 \%) \end{aligned}$ | $\begin{aligned} & 30 \\ & (38 \%) \end{aligned}$ | $\begin{aligned} & 2 \\ & (3 \%) \end{aligned}$ | $\begin{aligned} & 3 \\ & (4 \%) \end{aligned}$ | 80 neg. Qs |

non-inv. $=$ no "subject-auxiliary inversion"
2Aux ${ }_{\text {tm }}=$ two auxiliary verbs, both marked with the same T/Agr
2Aux. ${ }_{\text {m }}=$ two auxiliary verbs, both marked with different T/Agr
2 Aux diff aux. $=$ two different auxiliary verbs, both marked with T/Agr

Of the fifteen children, six children produced only grammatical positive questions.
The remaining nine produced mostly grammatical questions. With respect to the negative questions, only two children produced just grammatical negative questions.

Most of the children produced mainly ungrammatical questions. In fact, five children
produced no grammatical negative questions, and one child produced only one grammatical negative question (17\%).

Let us now tum to the types of ungrammatical questions that the children produced. The constructions produced for each type of question are summarized in the following table.
(48) Table 3: Summary of children's elicited production results

|  | adult <br> -like <br> (pos) | not | nt | gr. | LAux | non- <br> inv. | 2aux <br> +m | 2aux <br> -m | 2aux <br> diff <br> aux | ${ }^{\text {aggr }}$ | ungr | Total |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| pos yn | 55 |  |  | 55 |  |  | 1 |  |  | 4 | 5 | 60 |
| neg yn |  | 8 | 0 | 8 |  |  | 8 |  |  |  | 8 | 16 |
| pos obj | 33 |  |  | 33 | 1 |  |  |  |  |  | 1 | 34 |
| neg obj |  | 7 | 6 | 13 |  | 3 | 13 | 1 |  |  | 17 | 30 |
| pos adj | 28 |  |  | 28 | 3 | 1 |  |  |  | 1 | 5 | 33 |
| neg adj |  | 7 | 10 | 17 |  | 4 | 9 | 1 | 3 |  | 17 | 34 |
| Total | 116 | 22 | 16 | 15 | 4 | 8 | 31 | 2 | 3 | 5 | 53 | 207 |

IAux $\quad=\mathrm{T} /$ Agr marked on auxiliary verb and main verb
non-inv. $=$ no "subject-auxiliary inversion"
$2 \mathrm{Aux}_{+\mathrm{m}}=$ two auxiliary verbs, both marked with the same T/Agr
2 Aux $_{\mathrm{m}}=$ two auxiliary verbs, both marked with different T/Agr
2Aux diff aux. $=$ two different auxiliary verbs, both marked with T/Agr
*Agr $\quad=$ incorrect subject-verb agreement

Of the Il ungrammatical positive questions, there was one $2 \mathrm{Aux}_{+\mathrm{m}}$ question, four IAux questions, one question with non-inversion, and five with incorrect agreement. Examples of each type are shown below. ${ }^{20}$
(49) a. Do you do eat bugs?
b. Where did Jasmine \# hid the jewels?
c. Why he washes someone everyday? inversion
d. Does \# does you eat some bugs?
*Agreement
(Marilyn 3;11) 2Aux ${ }_{+m}$
(Phillip 4:6) IAux
(Jennifer 3;10) non-
(Andrew 4;3)

Of the ungrammatical negative questions, the majority (35/42) were 2Aux questions. Of the 13 children who produced ungrammatical negative questions, all of them produced at least one 2Aux question. There were $302 \mathrm{Aux} \mathrm{x}_{\mathrm{m}}$ questions (including one with the full form of negation), two $2 A u x_{-m}$ questions, and three $2 A u x$ questions with different auxiliary verbs (is/can't, is/doesn't, and did/couldn't). ${ }^{21}$ The other seven ungrammatical negative questions all involved non-inversion.
a. Why did he didn't brush the dog? (Scott 4;5) 2Aux ${ }_{+\mathrm{m}}$
b. What do eagles do not eat?
(David 4;7) 2Aux ${ }_{+m}$ with not
c. What does Cookie Monster don't like?
(Andrew 4;3) 2Aux $\mathrm{m}_{\mathrm{m}}$
d. Why is this hat doesn't fit him?
(Linda 4;1) 2Aux, different aux.
e. Why he didn't wear the hat?
(Kendra 4;5) non-inversion
The results from this experiment differ from previous studies in two respects. First, the children in this study did not produce any Neg/Aux doubling questions, such as in (51).

[^11](51) What didn't Miss Piggy don't like to do? (Matt 4;3) (Guasti et al. (1995)) In Guasti et al. (1995), there were 31 Neg/Aux doubling questions out of a total of 414 negative questions ( $7.5 \%$ ). Second, only two $2 \mathrm{Aux}_{\mathrm{m}}$ questions were produced in this study.

I discuss the results from this experiment together with the results from the other two experiments in section 2.6. Let us turn to the second elicited production experiment.

### 2.4 Experiment 2: The role of emphatic stress in children's 2Aux questions

Previous experimental studies have found that the production of 2 Aux questions occurs primarily with negation, and it has been assumed that negation is somehow related to the production of 2Aux questions. However, a few cases of positive 2Aux questions have been found. For example, Stromswold (1990) found six examples in spontaneous speech data. One invoived do, as in (52), and the other five involved be.
(52) Why did you did scare me? (Nina 3;2)

There has also been anecdotal evidence that children sometimes produce positive 2aux questions with emphatic stress on the second auxiliary verb.

[^12]In this experiment, I investigate whether emphatic stress will induce the production of positive 2Aux questions. By looking at emphatic stress, we can tease apart whether it is the child's attempt to use negation or the inherent focus/emphasis associated with negation which causes the production of 2Aux questions. If emphasis is a relevant factor in the production of 2 Aux questions, we expect children to produce positive and negative 2Aux questions with emphatic stress. If negation is the primary factor in the production of 2Aux questions, we expect 2Aux questions to be limited to negative contexts.

### 2.4.1 Methodology

I conducted another elicited production experiment to determine what types of questions children would produce with emphatic stress. In this experiment, the experimenter manipulating the puppet was also the storyteller. The other experimenter prompted the child to ask the puppet some questions. The puppet (a slug named Benny) likes talking to kids but not to adults, so the child's role was to help the experimenter find things out from the puppet. The puppet talked to the child about the things he did, but did not mention the exact details. The experimenter, curious about the details, prompts the child to ask the puppet a question to clarify what he did. Unfortunately, the puppet is frequently confused, and tells the child the opposite
answer of the one desired. The experimenter prompts the child again, this time using emphatic stress.

Sample stories eliciting a positive and negative question are shown below.
(53) Benny: You know what? I baked some yummy things yesterday.

Exp: Benny baked something. Ask him what.
Target: What did you bake?
Benny: I didn't bake a chocolate cake or cookies or an apple pie.
Exp: Benny keeps telling us about the things he didn't bake, but I want to know about the things he did bake. Ask him what.
Target: What did you bake?
Benny: I baked a cupcake and a donut.
(54) Benny: Today is my gardening day. I was going to plant flowers all over the yard, but I didn't finish because it got too late.
Exp: Benny didn't plant the flowers somewhere. Ask him where.
Target: Where didn't you plant flowers? / Where did you not plant flowers?
Benny: I planted flowers by the fence and by the shed and by the house.
Exp: Benny keeps telling us about the places he did plant flowers, but I want to know about the places he didn't plant flowers. Ask him where.
Target: Where didn't you plant flowers? / Where did you not plant flowers?
Benny: I didn't plant flowers by the trees or the bushes. I'll have to do that another day.

There were eight stories that elicited a total of sixteen questions (positive and negative), with Wh-object and Wh-adjunct extraction. Each story had a pair of target questions: the first without emphatic stress, and the second with emphatic stress. The complete set of target items is given below.
(55) a. What did you bake? / What did you bake?
b. What did you find? / What did you find?
c. Where did you hide the jewels? / Where did you hide the jewels?
d. Where did you leave your scarf? / Where did you leave your scarf?
e. What didn't you buy? / What didn't you buy?
f. What didn't you want? / What didn't you want?
g. Where didn't you chase Buzz? / Where didn't you chase Buzz?
h. Where didn't you plant flowers? / Where didn't vou plant flowers?

The positive questions were presented together, followed by the negative questions.
The items were randomly ordered and each child saw the same random order.

### 2.4.2 Subjects

Ten children $(3 ; 5-4 ; 7)$ attending the Child Development Laboratories at the University of Connecticut participated in the experiment. The youngest child $(3 ; 5)$ could not attend to the task, and was excluded from the study. The results are from the remaining nine children, who were each given pseudonyms. We also tested three adult native English speakers as control subjects.

### 2.4.3 Results from the adults

All three adult subjects performed as expected. They produced grammatical positive questions and grammatical negative questions, all with $n^{\prime} t$. They produced
questions with emphatic stress only in response to prompts with emphatic stress. ${ }^{22}$ In all of these questions, the emphatic stress was always on the auxiliary verb (do or $d o+n ' t)$.

### 2.4.4 Results from the children

The emphatic stress manipulation succeeded in getting children to produce positive and negative questions with emphatic stress. ${ }^{23}$ The production of questions with emphatic stress in response to prompts with and without emphatic stress is significantly different by paired $t$-test $(t(8)=5.09, p<001)$. The mean number of questions with emphatic stress in response to prompts with and without emphatic stress is 4.4 and 0.3 , respectively. Children produced questions using emphatic stress more often in response to the prompts with emphatic stress ( 40 out of 72 , or $56 \%$ ), than those without ( 3 out of 72 , or $4 \%$ ).

However, the children did not always produce a question with emphatic stress in response to a prompt with emphatic stress. Overall, the children produced questions with emphasis $56 \%$ of the time in response to prompts with emphatic stress. If we

[^13]look at the children individually, we find that five children produced questions with emphasis more than $75 \%$ of the time, and four children produced questions with emphasis less than $40 \%$ of the time. See (56) for a summary of each child's production of questions with emphasis in response to prompts with emphasis.
(56) Table 4: No. of questions with emphasis (in response to emphatic prompts)

|  | \# of questions with emphasis |  |
| :--- | :---: | :---: |
| Isacc $(3 ; 8)$ | $7 / 8$ | $(88 \%)$ |
| Mary $(4 ; 2)$ | $7 / 8$ | $(88 \%)$ |
| Michael $(3 ; 9)$ | $2 / 7$ | $(29 \%)$ |
| Barbara $(4 ; 4)$ | $6 / 8$ | $(75 \%)$ |
| Troy $(4 ; 5)$ | $3 / 8$ | $(38 \%)$ |
| Emma $(4 ; 5)$ | $6 / 8$ | $(75 \%)$ |
| Trevor $(4 ; 3)$ | $7 / 8$ | $(88 \%)$ |
| Eva $(4 ; 4)$ | $1 / 8$ | $(13 \%)$ |
| Dustin $(4 ; 7)$ | $1 / 8$ | $(13 \%)$ |

Let us turn to the types of grammatical and ungrammatical questions that the children produced. Overall, the children's positive questions were mainly adult-like, supporting previous findings. ${ }^{24}$ Of the 70 positive questions that were produced, 57 (79\%) were grammatical. Ten of the positive questions (14\%) were I Aux questions, all without emphasis, and all with irregular verbs (find, hide, leave). The children produced very few positive 2Aux questions, as was found in previous production

## (ii) What did you bake for real? (Trevor 4;3)

${ }^{24}$ Perhaps due to the way the prompts with emphatic stress were phrased, three children produced ungrammatical questions with part of the prompt (the things or the places) repeated.
(i) What did the things you did find?
(Isaac 3;8)
(ii) Where did the places you didn't chase him? (Emma $4 ; 5$ )

Eleven questions of this type were excluded ( 2 positive and 9 negative).
studies. There were only three positive 2Aux questions, and all three of these 2Aux questions involved emphatic stress on the lower auxiliary verb, as shown in (57).
What did you do find?
(Barbara 4;4)

The children's production of negative questions also corresponds to previous findings. ${ }^{25}$ Of the 46 negative questions that the children produced, 23 (50\%) were grammatical questions, as in (58), and 23 (50\%) were 2Aux questions. ${ }^{26}$ Of the negative 2Aux questions, 10 were produced with emphatic stress, and 16 were produced without emphatic stress. A pair of non-emphatic and emphatic questions is given in (59).
(58) a. What didn't you want?
(Dustin 4;7)
b. What didn't you want?
(Dustin 4;7)
c. What did you not want?
(Eva 4;4)
d. Where did you not plant flowers?
(Barbara 4;4)
a. Where did you didn't chase Buzz? (Mary $4 ; 2$ )
b. Where did you didn't chase Buzz? (Mary $4 ; 2)^{27}$

The production of 2Aux questions in positive vs. negative contexts is significantly different by paired $t$-test $(t(8)=2.93, p=.019)$. The mean number of 2 Aux questions produced in positive and negative contexts are 0.3 and 2.6 , respectively. There were significantly more 2Aux questions in negative contexts (23 2Aux

[^14]questions out of 46 negative questions, or 50\%) than in positive contexts (three 2Aux questions out of 70 positive questions, or $4 \%$ ).

We can use these results to evaluate the role of emphatic stress in 2Aux questions. Although the use of emphatic stress did result in the production of some positive 2 Aux questions, there were only three such questions out of the 70 positive questions produced. Furthermore, children produced more 2Aux questions without emphatic stress ( 16 out of 26 2Aux questions) than with emphatic stress (10 out of 26 2Aux questions). Of the 6 children who produced 2Aux questions, this difference is not significant by paired $t$-test, $t(5)=1.483, p=198$. Therefore, emphatic stress alone cannot explain the production of 2Aux questions. Instead, we must appeal to negation as the primary factor.

We now turn to the third experiment, which elicits acceptability judgments from children regarding grammatical and ungrammatical questions, in particular 2Aux questions. First, I review the methodology in section 2.5.1, followed by a summary of the stimuli in section 2.5.2. The results from the adults and children are presented in sections 2.5.4 and 2.5.5.

[^15]
### 2.5 Experiment 3: Children's judgment of 2Aux questions

### 2.5.1 Methodology

To determine how children judge ungrammatical 2Aux questions, I used a grammaticality judgment task similar to MicDaniel and Cairns (1996). McDaniel and Cairns suggest that the researcher first establish language as a topic, emphasizing the difference between form and content. After providing some background about language, the experimenter acts out scenes with toys and describes the scene by uttering a sentence. The experimenter then asks the child if the sentence sounds right or wrong. The child is instructed to pretend that the experimenter does not know English and that she needs help learning it.

In order to further hide the fact that the child is being tested and to make the task more similar to the elicited production task, Diane Lillo-Martin and I modified the grammaticality judgment task in the following ways. We eliminated the discussion about language by incorporating it into the characteristics of a puppet, and made the task more interactive in general. We introduced the child to a fish puppet from the moon (named Lulu), and explained that she talks "moon talk," which is different from "earth talk" or English. Since the puppet is still learning how to talk English, she sometimes says things in a silly way. By having the puppet speak a different language, the child no longer needs to pretend that the experimenter does not
know English. Also, the fact that the child is being tested is hidden; the point of the game is for the child to help the puppet learn how to talk the right way.

The modified grammaticality judgment task is now very similar to the production task: one experimenter tells stories using toys, and the other plays the role of a puppet, who gets treats depending on whether the puppet is right or wrong. ${ }^{28}$

Each session was audio taped, and the children's responses were recorded on a score sheet by the experimenter playing the role of the puppet. Any comments that may be relevant, such as distraction or hesitance on a particular response, were also noted. Each session lasted no longer than twenty-five minutes, with a maximum of sixteen test sentences.

### 2.5.1.1 Training and pretest sessions

Before beginning the test sessions, each child completed at least one training session. If the child made a mistake on the training session, we gave them an additional training session. At the beginning of the training session, the child practiced judging words rather than sentences. The puppet would call an object by the wrong name, or jumble the sounds in a word, as in pig-Latin. For example, the puppet might call the child "aureen may" instead of "Maureen", or substitute "chair" for

[^16]"table". Once the child understood these types of mistakes, we presented her with sentences.

During the training sessions, we provided as much feedback as necessary to help the children understand the task. For example, if a child responded that the sentence in (60) is correct, we asked him to describe what the dog was doing. Most of the time, the child answered "drinking water".
(60) The dog water drinking is.

We pointed out to the child that what the puppet said, "water drinking", was different from "drinking water". If the child was reluctant to describe what the dog was doing, the storyteller explained how she would describe the situation, and how her way of saying it was different from the puppet's statement.

We discovered that children who had a lot of difficulty with the task became less enthusiastic about getting feedback when they were wrong, and they seemed to be aware of the fact that they were being tested somehow. Therefore, feedback was kept to a minimum for these children. The children who understood the task had no problems with receiving feedback.

Although we did not require children to give the correct form of a statement if the puppet gave the incorrect form, a few children volunteered corrections to the puppet's silly sentences. However, the majority of children could not (or were hesitant to) provide the correct way to say a sentence, even if they knew that the puppet's sentence was wrong.

In order to participate in the test sessions, a child had to pass a pretest. Each child had two opportunities to pass the pretest. If she did not pass the first pretest, she was given a detailed reminder about the game and the puppet's moon-talk before the second pretest. In addition, we had the child practice giving judgments on a few sentences produced spontaneously by the puppet in a warm-up period prior to the second pretest.

### 2.5.1.2 Test sessions and follow-up sessions

At the beginning of each test session, we reminded the child about the game and the difference between moon-talk and earth-talk. We also had the child practice giving judgments for the puppet's spontaneous utterances.

As with the elicited production task, the storyteller manipulated toys and told a story. Part of the story took place behind a cardboard wall, and thus was hidden from both the child and the puppet. Since the puppet could not see what was happening, she asked a question about the story. The rest of the story was told after the child judged the puppet's utterance. By hiding the action, we tried to prevent the children from answering the puppet's questions instead of judging them. Two sample stories are shown below.
(61) Exp: This is a story about Cabbage Patch Girl.

CPG: Today's my shopping day. I'm going to go to the mall. (Go behind wall.)
Look at all those things. I like this one, but I don't like this other one.
Lulu: I have a question. What did Cabbage Patch Girl don't like?
Exp: Did Lulu say it right?
(62) Exp: This is a story about the Tin Man.

Tin Man: I go fishing every day in my brown boat. I take my bucket and my fishing pole, and go fishing in the river. I wonder what I'll catch today.
(Go behind wall.)
This looks like a good spot. I'm gonna be really quiet and maybe I'II catch something.
Lulu: I have a question. What does the Tin Man does catch today?
Exp: Did Lulu say it right?

### 2.5.2 Stimuli

### 2.5.2.1 Stimuli for training and pretest sessions

In the training and pretest sessions, children judged positive declarative
sentences and positive questions. An example of a grammatical positive question is shown below.
(63) What did Big Bird see? Positive question

The ungrammatical positive questions involved one of two types of errors. We chose phenomena that we knew children had mastered, rather than those that children had possibly not mastered, such as inversion. One type of error was incorrect word order;
the other type was the inclusion of additional lexical items. Some examples of each error type are shown in (64) and (65).
(64) Incorrect word order
a. Who the monster tickling is?
b. Why stay you home yesterday?
(65) Additional lexical item
a. What does Belle get food?
b. Who did the fairy give the balloons to the person?

The ungrammatical positive questions presented in these sessions were crucially not the type of questions that were the target of the test sessions, such as 2 Aux questions.

The pretest consisted of six positive questions (three grammatical and three ungrammatical). We followed McDaniel and Cairns' (1996) criterion for scoring the pretest. In order to pass, the child had to reject all three ungrammatical sentences, and had to accept at least two of the three grammatical sentences. Although this criterion seemed a bit rigorous at the beginning, it became clear after testing several children that such a strict criterion was needed to ensure a child's consistency and reliability.

### 2.5.2.2 Stimuli for test sessions

Each of the four test sessions consisted of two practice items, four test items, and four control items. There were also filler items (all positive) that were inserted after a certain number of yes or no responses. An ungrammatical filler was inserted after three yes responses, and a grammatical filler was inserted after two no responses.

Four stories were prepared with both grammatical and ungrammatical options ready for each story. The control and filler items were used to determine whether the children were paying attention to the task.

The practice items consisted of two each of the following two sentence types.
(66) a. Positive ungrammatical deciarative
b. Positive grammatical declarative

The stimuli for the test items consisted of two each of eight sentence types with doubled auxiliary verbs. The eight sentence types are shown below.
(67) a. Positive $2 A u x_{+m}$ Wh-object
b. Positive $2 A u x+m$ Wh-adjunct
c. Negative $2 A u x_{+m}$ Wh-object
d. Negative $2 A u x_{+m}$ Wh-adjunct
e. Positive $2 \mathrm{Aux}_{-\mathrm{m}}$ Wh-object
f. Positive 2 Aux -mh -adjunct
g. Negative 2Aux-m Wh-object
h. Negative 2Aux-m Wh-adjunct

The twelve control items consisted of three each of the following four sentence types.
(68) a. Positive grammatical question
b. Negative grammatical question
c. Positive ungrammatical question
d. Negative ungrammatical question

All of the negative control items were questions with the contracted form of negation, as shown below.
(69) What didn't the Martian taste? Negative question with raised $n$ 't

Negative questions with the full form of negation not were tested in the follow-up sessions.

All of the ungrammatical practice, control and filler items involved the same types of errors that were used for the training and pretest items (incorrect word order or additional lexical items).

The practice items were always presented first for each session. The rest of the items (test items and control items) were randomly ordered and divided into four sessions. ${ }^{29}$ The same random order was used for every child.

### 2.5.2.3 Stimuli for follow-up sessions

In the follow-up sessions, the children judged negative questions with the full form of negation, not. There were six negative questions with not: three each of the following two sentence types.
(70) a. Negative Wh-object question with not
b. Negative Wh-adjunct question with not

Each follow-up session included two practice positive declarative items (ungrammatical and grammatical) and four control items, as in the test sessions. The twelve control items consisted of three each of the following four sentence types.
(71) a. Positive grammatical question
b. Negative grammatical question
c. Positive ungrammatical question
d. Negative ungrammatical question

[^17]All twelve ungrammatical control items involved incorrect word order.
Each session began with the practice items. The control items and test items were randomly ordered and divided into three sessions. The same random order was used for each child.

### 2.5.3 Subjects

The same children who participated in the elicited production sessions of experiment l participated in the grammaticality judgment sessions. At the beginning of the sessions, the children ranged in age from $4 ; 0$ to $5 ; 0$. Twelve of the fifteen children also participated in three follow-up sessions. The same pseudonyms that were used in experiment 1 are used to report the acceptability judgment results.

The same four adult subjects who participated in the elicited production sessions of experiment I also participated in the grammaticality judgment sessions. The pretest sessions, test sessions and follow-up sessions were all tested in one session. The same methodology was used, but the adult subjects were tested in a group and recorded their own responses. ${ }^{30}$

[^18]
### 2.5.4 Results from the adults

All four adult subjects performed as expected on the pretest, test sessions, and follow-up sessions. They gave the expected adult judgments on the three ungrammatical questions and the three grammatical questions on the pretest. They also correctly rejected all of the ungrammatical test sentences, and accepted the grammatical questions with not. The declaratives and control questions were also judged correctly.

### 2.5.5 Results from the children

Overall, the children responded adult-like on $94 \%$ of the declarative, control, and filler items. This shows that children were consistently cooperating with the rules of the game. The children responded adult-like on $73 \%$ of the test items (positive $2 A u x_{+m}$ positive $2 A u x_{-m y}$ negative $2 A u x_{+m}$ negative $2 A u x_{-m}$ ). The percentage of adult-like responses for control items (declarative, control, and filler) vs. test items is significantly different by paired $t-\operatorname{test}(t(14)=3.967, p=.001)$. In general, children performed more adult-like on the declarative, control and filler items than on the test items.

To determine whether the children performed more adult-like on positive vs. negative 2Aux questions, or on 2Aux questions with matching vs. non-matching auxiliary verbs, I performed a two-way repeated measures ANOVA. There was a
significant main effect of matching vs. non-matching auxiliary verbs $(F(1,14)=5.064$, $p=.041$ ). Children responded more adult-like on $2 \mathrm{Aux}_{-\mathrm{m}}$ questions than on $2 \mathrm{Aux}_{\mathrm{rm}}$ questions (mean percentage of adult-like responses: $79 \%$ vs. $67 \%$, respectively). Although the children responded more adult-like on positive 2Aux questions (77\%) than negative 2Aux questions ( $69 \%$ ), the difference of the mean values among the different levels of positive/negative was not statistically significant $(F(1,14)=2.272$, $p=.154)$. There was also no statistically significant interaction effect $(F(1,14)=0.132$, $p=.722$ ). The mean percentage of adult-like responses was $62 \%$ for negative $2 A u x_{+m}$ questions, $77 \%$ for negative 2 Aux $_{m}$ questions, $73 \%$ for positive $2 \mathrm{Aux}_{+\mathrm{m}}$ questions, and $82 \%$ for positive $2 \mathrm{Aux}_{\mathrm{m}}$ questions.

With respect to the grammatical negative questions with not, which were tested in the follow-up sessions, the twelve children who received this question type accepted all of them. The children were cooperating with the rales of the game for the followup sessions also, as shown by their adult-like responses on $96 \%$ of the declarative and filler items. (See Appendix I for individual responses on the items.)

Let us summarize the grammaticality judgment results. Although the children's responses on the test items are not as adult-like as the control items, they are still correctly rejecting a majority ( $73 \%$ ) of the test items. Notice that the children as a group are rejecting the 2 Aux constructions that they had produced during the elicited production sessions. Of the different types of 2Aux questions, the children
reject the 2 Aux $_{-m}$ questions more often than the $2 A_{x_{+m}}$ questions. Recall that in Experiment 1 , the children produced essentially no 2 Aux $_{-m}$ questions.

If the children have a non-adult grammar, we would expect them to produce and also accept 2 Aux questions. We would not expect them to reject any of the 2 Aux questions. However, these children are rejecting most of the 2Aux questions. Therefore, positing a non-adult grammar for the children does not account for the data.

We now turn to a discussion of the results from all three experiments. In the next section, let us consider some possible models of the grammar that might account for the results, and argue that the children have the adult grammar. I then offer a proposal to explain why the children produce 2Aux questions despite having the aduit grammar.

### 2.6 Discussion

The results from the first elicited production experiment mainly support previous findings (Guasti et al. (1994, 1995), Thornton (1993)). Children's positive questions are $91 \%$ grammatical, while their negative questions include many (53\%) ungrammatical constructions. Of the grammatical negative questions produced by ten children, two children produced questions only with the contracted form of negation, four children produced questions only with the full form of negation, and four children produced questions with both the contracted form and full form of negation. Of the 13
children who produced ungrammatical negative questions, every child produced at least one 2 Aux question. The findings from this experiment differ from previous studies in that we did not find any Neg/Aux doubling questions, and we found only two cases of the $2 \mathrm{Aux}_{-\mathrm{m}}$ construction.

From the second elicited production experiment, we have evidence that suggests that negation, rather than emphatic stress, is the primary trigger for the production of 2Aux questions. Only three positive 2Aux questions were produced, compared with 23 negative 2Aux questions.

When we add the results from the grammaticality judgment task, we find an interesting pattern. Although not all of the children produced grammatical questions with the full form of negation, all of the children judged them to be acceptable. Even though the children may not have produced this type of grammatical negative question, they have the knowledge that it is grammatical. On the other hand, although almost all of the children produced at least one 2Aux question, they did not judge them to be acceptable. We had expected children to judge what they produce to be acceptable, but there seems to be a discrepancy between production and judgment. Given this production/judgment asymmetry, we must try to determine whether we should take the production or judgment data to be reflecting the competence.

### 2.6.1 2Aux questions and children's competence

Let us first assume that the children's production of 2Aux questions is reflecting their competence. However, under this approach, we will see that we cannot account for the data.

If we assume a Principles and Parameters model (Chomsky and Lasnik (1993)), we might try to account for the production/judgment asymmetry by allowing an optional setting for parameters. Suppose there is a parameter $Z$ that is relevant to the production of negative questions. If parameter Z is set incorrectly for English, negative 2Aux questions are produced, while if it is set correctly, adult negative questions are produced.

Given this, we might propose that the production/judgment asymmetry is the result of an optional setting of parameter $Z$ : the child has not chosen between the two settings. However, if the child had an optional setting of parameter $Z$, we would expect her to produce more adult negative questions with do+negation in $C$, and to judge some, if not all, of the negative 2Aux questions as being grammatical. Our data show that all of the children produced non-adult negative questions and only some of them produced adult negative questions. Furthermore, the children had adult judgments with respect to the negative 2Aux questions. Therefore, an optional parameter setting cannot account for the data.

Even without relying on a parameter, if the child has a grammar that is nonadult and allows both grammatical negative questions and 2 Aux questions, the data
cannot be accounted for. We would expect children to produce all three types of negative questions and judge all three types to be grammatical. However, the children produced many more 2Aux questions than adult negative questions, and more importantly, they judged 2Aux questions to be ungrammatical. Furthermore, it is not obvious what would force the child to change to the adult grammar which allows negative questions with $n$ 't and not, but not 2Aux questions.

We have seen that under the assumption that the production data are reflecting the children's competence, the judgment data cannot be explained. If the child's grammar allows the production of 2Aux questions, there is no way to account for why the children reject 2Aux questions. The analyses proposed by Thornton (1993) and Guasti et al. (1995), which are based on the assumption that the child's grammar is different from the adult grammar, do not explain the production/judgment asymmetry.

In order to account for the production/judgment asymmetry, we must take the judgment data to be reflecting the children's competence, and assume that the children do not allow 2 Aux questions. However, we must now explain why the children produce 2Aux questions. I propose that the children's production of 2Aux questions is a performance error resulting from incorrect lexical information.

In the next section, we first outline the adult grammar and how negation questions are formed. We then discuss how the children rely on the same adult grammar to form grammatical negative questions, and how the children's incorrect
lexical information regarding negation may trigger a "last resort" PF operation that results in the production of 2Aux questions.

### 2.6.2 The adult grammar

### 2.6.2. 1 Sentential and constituent negation in English

The previous analyses of children's negative questions do not distinguish between constituent and sentential negation, and hence incorrectly conclude that questions with not are non-adult or less preferred structures. However, questions with not are grammatical when the situation is appropriate for constituent negation.

Let us take a closer look at the characteristics of these two types of negation. In general, as the names suggest, the negative element negates the entire sentence in sentential negation, while it negates only a constituent in constituent negation.

Compare the declarative sentences below with constituent negation and sentential negation.
(72) a. Kerry can [not talk to Pat]
'Kerry has the option of talking or not talking to Pat.'
b. Kerry cannot talk to Pat
c. Kerry can't talk to Pat
'It is not possible for Kerry to talk to Pat.'
In the example with constituent negation in (72)a, not is negating the act of talking, while in the examples with sentential negation in (72)b and (72)c, not and $n^{\prime} t$ are negating the possibility (the modal verb can) of talking. Constituent negation is
marked by a slight pause before the negative element, not, which is stressed. (I will be referring to this emphasis or focus on not as "stress".) The negative affix $n$ 't cannot be used for constituent negation, as shown by the impossibility of interpreting (72)c as (72)a. (72)c does not mean that 'Kerry has the option of talking or not talking to Pat.'

The different interpretations of sentential and constituent negation can be attributed to different syntactic positions of negation. With respect to the structural position of sentential negation, I will adopt Laka's (1990) proposal of a functional projection, $\Sigma \mathrm{P}$, located between $\operatorname{Mod}(\mathrm{al}) \mathrm{P}$ and VP. She proposes that there is a negation morpheme and an affirmation morpheme ${ }^{31}$, which belong to the same syntactic category $\Sigma$. The negative/affirmation morpheme heads its own functional projection, NegP and AffP, respectively. In other words, NegP and AffP are different instantiations of $\Sigma P$. $^{32}$

[^19](73)


I will assume that constituent negation is adjoined to the XP it modifies. ${ }^{33}$
Notice that constituent negation is not restricted to modifying VPs, as shown below.
(74) a. Kerry can [Vp not [Vp talk to Pat]]
b. What did Terry [vp not [vp buy]]

In both (74) a and (74)b, not modifies a VP. Notice also, that it is possible to have both sentential and constituent negation in the same sentence.
(75) Kerry can't [not talk to Pat]
'It is not possible for Kerry to [not talk to Pat].'

[^20]In (75), the affix $n$ 't negates the possibility of talking, and not negates the act of talking. The sentence would be felicitous given a situation in which even though Kerry would rather not talk to Pat, Kerry sees Pat every day at work, and hence does not have the option of not talking to Pat.

Let us now characterize the constraints on the use of constituent and sentential negation. We have already seen how the two types of negation behave in declarative sentences. Consider the following examples.
(76) a. The smurf can't [vp buy the strawberries]
b. The smurf cannot [vp buy the strawberries]
c. The smurf can [Vp not [Vp buy the strawberries]] (= The smurf has the option of not buying the strawberries.)

The sentences in (76)a and (76)b contain sentential negation marked by the negative affix and the full negative form not, respectively. The sentence in (76)c contains constituent negation marked by not, which has stress and a slight pause before it.

The two types of negation behave the same way in questions. Consider the following three questions in (77), and the possible interpretations they may have.
(77) a. What couldn't the smurf buy?
b. What could the smurf not buy?
c. What could the smurf [not buy]?
(78) a. The smurf couldn't buy the strawberries. [sentential negation] (Because the smurf didn't have enough money.)
b. The smurf could [not buy the strawberries]. [constituent negation] (Because Mom said the smurf could decide whether or not to buy the strawberries.)

The questions in (77) a and (77) b are grammatical with a sentential negation interpretation, and can take the possible answer given in(78)a, but not (78)b. The question in (77) c is grammatical with a constituent negation interpretation, and can take the answer in (78)b, but not (78)a. Again, the question in (77)c involves stress on not, and pause before it.

All of the examples we have examined have involved modals. When modals are not used, we find that the constituent negation interpretation cannot be obtained. Consider the following declarative sentences without modals.
(79) a. The smurf didn't [vp buy the strawberries]
b. The smurf did.not [Vp buy the strawberries]
c. *The smurf did [vp not [vp buy the strawberries]] (= What the smurf did, was not buy the strawberries.)
d. The smurf did [vp not [Vp buy the strawberries]] (= What the smurf did do, was not buy the strawberries.)

The sentences in (79)a and (79)b both involve sentential negation. The sentence in (79)c has the structure of constituent negation, but the constituent negation interpretation cannot be obtained. There seems to be no difference in interpretation between (79)b and (79)c. The constituent negation interpretation is possible only if emphatic do is involved, as in (79)d.

To summarize, sentential negation can be expressed by either negative marker, n't or not. However, constituent negation can only be expressed by not, with stress and a pause before it. It cannot be expressed by the affix $n, r^{34}$

### 2.6.2.2 Verbal morphology and I to C movement

In addition to the characteristics of constituent and sentential negation, our analysis of the adult grammar will assume the minimalist program (Chomsky (1993), (1995)). The computational system takes elements from the lexicon to form the numeration. The computational system then uses these elements to generate derivations and structural descriptions. There are two interface levels, LF and PF, which are constrained by principles and well-formedness conditions. Economy principles compare derivations with the same numeration to eliminate all derivations that are not the most economical. A derivation converges if it generates a legitimate structural description, and crashes if it does not. A "copy theory" of movement is assumed in which the trace of a moved element is a copy of that element.

There are two other issues we need to consider in order to discuss negative questions: a theory of verbal morphology, and an analysis of Wh -movement and I-toC movement. Lasnik (1995) proposes a theory of verbal morphology in which French

[^21]verbs and English have and be are fully inflected in the lexicon, while English main verbs are bare. The following assumptions are made.
(80) a. Infl is freely an affix or a set of abstract features.
b. Finite featural Infl is strong in both French and English.
c. Affixal Infl must merge with a V, a PF process (distinct from head movement) demanding adjacency.

Under Lasnik's theory, finite featural Infl in English must have its strong features checked before spell-out in order to have a converging derivation. Strong features are visible at PF , while weak features are not. Therefore, in order to have a legitimate PF object and hence a converging derivation, strong features must be checked before spell-out, while weak features are forced by Procrastinate to survive until PF.

Main verbs are bare and not fully inflected. Therefore, for main verbs it does not matter whether featural Infl is strong or weak; only the affixal Infl is relevant. In order to have a converging derivation, a bare main verb must be adjacent to an affixal Infl to merge at PF. If the main verb is not adjacent to the affixal Infl, Infl is spelled out as do in order not to violate the Stranded Affix Filter. Do-support is a last resort operation at PF to prevent the derivation from crashing. English have and be are both fully inflected so a strong finite featural Infl will correctly necessitate overt raising, like all French verbs. Consider the following example.
(81) a. [PP Mandy [r I not [vp is crazy]]] $[+F] \quad[+F]$
b. Mandy is not crazy
c. *Mandy not is crazy

In (81) a, Infl is featural, and the auxiliary verb is is taken from the lexicon fully inflected. The auxiliary verb is raised overtly, to generate (81)b. If the auxiliary verb does not raise overtly, as in (81)c, the strong feature of Infl will survive into PF, causing a PF crash.

In Wh-questions, we will assume that Wh -movement is forced overtly in English due to a strong [ +wh ] feature. In English, the strong [ + wh] feature drives overt Wh-movement: the Wh-phrase must move overtly to [spec, CP] in order to eliminate the strong [ + wh] feature of $C$ through checking via spec-head agreement.

In addition, we will assume that I to C movement is triggered by some other strong $Q$ feature in C. Boškovic (1995) adopts Chomsky's (1957) proposal that the [ +wh ] C in English is a phonological affix that needs a tensed verbal eiement as a host. (See also Halle and Marantz (1993), Lasnik (1995) and Bobaljik (1995)). It is not important for our analysis of Wh-questions whether it is a feature that triggers I to C movement, or the affixal property of $C$.

### 2.6.2.3 Derivations of adult questions

Let us now take a look at the derivations for adult questions with sentential and constituent negation. First, consider the following question with sentential negation involving the affix $n^{\prime} t$.
(82) a. What didn't the smurf buy? (Sentential negation with $n$ 't)
b.


The Wh-phrase raises overtly to [spec, CP] to check off the strong [+wh] feature in C. The affix $n$ 't raises to Infl so that it has a host, and Infl ([+past, 3psg]) along with the negative affix raises overtly to $C$ to check off the strong $Q$ feature in $C$. Assuming the VP-internal subject hypothesis, the subject NP raises from [spec, VP] to [spec, IP] in order to check off the strong EPP feature of Infl. At PF, the affixal Infl in C and the bare main verb are not adjacent due to the intervening NP the smurf, and hence cannot merge. In order not to violate the Stranded Affix Filter (SAF), do-support occurs, and Infl is spelled out as did $+n^{\prime} t$.

The derivation for a question with sentential negation involving not is very similar to the derivation for a question with sentential negation involving $n$ 't.
(83) a. What did the smurf not buy? (Sentential negation with not)
b.


The Wh-phrase raises overtly to [spec, CP] to check off the strong [ + wh] feature in C , and Infl ([ + past, 3psg]) raises overtly to $C$ to check off the strong $Q$ feature in $C$. The subject NP raises from [spec, VP] to [spec, IP] in order to check off the strong EPP feature of Infl. At PF, the affixal Infl in $C$ and the bare main verb are not adjacent due to the intervening NP the smurf, and hence cannot merge. In order not to violate the Stranded Affix Filter (SAF), do-support occurs, and Infl is spelled out as did. The main difference between this derivation and the previous derivation for sentential
negation is that the negative marker not in this derivation does not raise to Infl. There is no need for not to raise to Infl since it is not an affix that requires a host; it can remain in $\Sigma$.

Let us now consider a question with constituent negation, as in (84).
(84) a. What did the smurf not buy? (Constituent negation with not)
b.


Like the two previous derivations, the Wh-phrase raises to [spec, CP] to check off the strong [+wh] feature in C, and affixal Infl ([+past, 3psg]) raises to C to check off the strong $Q$ feature in C. The subject NP raises to [spec, IP] to check off the strong EPP feature of Infl. At PF, the affixal Infl and the bare main verb are again not adjacent because of the intervening subject in [spec, $\mathbb{P}]$, and hence cannot merge. Do-support is invoked, and Infl is spelled out as did.

We have seen the adult derivations for negative questions involving sentential and constituent negation. Let us now turn to the child's grammar and see how it is similar to and different from the adult grammar.

### 2.6.3 Children's production of 2Aux questions

The results from the experiments show that children produce 2Aux questions but nevertheless judge them to be unacceptable. With respect to grammatical negative questions, only some children produced them, but all of the children accepted them. I propose that the children's derivations for grammatical sentential negation questions with with $n$ 't and not and grammatical constituent negation questions with not are the same as the derivations for the adults discussed in the previous section. This explains why they produce these grammatical negative questions and also judge them to be acceptable.

In order to explain why children judge 2Aux questions to be unacceptable even though they produce them, I propose that the children's production of 2Aux questions is a performance error due to incorrect lexical information regarding constituent negation. Suppose that children think that both forms of negation, $n$ 't and not, can be used for sentential and constituent negation. This assumption holds for sentential negation, but not for constituent negation. In English, only not can be used for constituent negation. When the child is required to generate a constituent negation
question, she may incorrectly choose the negative affix $n^{\prime} t$. This will force her into a situation that must be resolved by using a "last resort" operation and will result in the production of a 2Aux question. However, when the child is only required to judge a sentence, she is able to correctly reject 2Aux questions.

Let us go through the derivation for 2Aux questions in more detail. Consider the derivation for the following child utterance.
(85) *What did the smurf didn't buy?

The derivation for this sentence is the same as for the adult using a constituent negation structure until PF. (See (84) for the adult derivation.) The child correctly raises the Wh-word to [spec, CP] to check off the strong [ +wh ] feature in C, and raises Infl to $C$ to check off the strong $Q$ feature in C. Furthermore, the subject NP is raised to [spec, IP] to check off the strong EPP feature of Infl. The negative affix in this constituent negation structure does not raise to Infl as it does in a sentential negation structure since the negative affix is in a VP adjunct position here. Given Emond's (1976) Structure-Preserving Hypothesis and the theory of movement proposed in Chomsky (1986), only X ${ }^{0}$ can move to a head position. ${ }^{35}$

At PF, the child is left with the following structure.

+past +past

[^22]Under the "copy theory" of movement, the trace of a moved element is a copy of that element. The child knows the PF principle that deletes any copies of overtly moved elements, hence she correctly deletes the copies of the moved subject NP and the Whword.


| +past | +past |
| ---: | ---: |
| 3psg | 3psg |

Under normal circumstances, the child would also delete the PF copy of the moved Infl. However, the child has incorrectly chosen $n$ 't from the lexicon. The child has two competing derivations to choose from. One derivation, DI , correctly deletes the PF copy of Infl, which would result in the following structure.

+past +past

3psg 3psg

The affixal Infl in C is spelled out as did to satisfy the SAF. However, the affix $n^{\prime} t$ does not have a lexical host. Therefore, derivation D1 results in a PF crash.

In another derivation, D2, the child uses a "last resort" strategy and spells out both copies of Infl in order to satisfy the SAF and to provide a host for the affix. ${ }^{36}$
(89) [CP what [C' did [iP the smurf[r did [VP n't [vp buy]]l]] [D2]

[^23]If the child chooses the derivation, DI, she would either have to pronounce the affix without a host, which is phonologically difficult, or not pronounce it at all, which would result in important information being lost. If the child chooses the derivation, D2, the SAF is satisfied, but the PF principle that deletes any copies of overtly moved elements is violated. Given these two derivations, the child chooses the second derivation, D 2 .

This "last resort" strategy is similar to recent proposals arguing that there is a choice about which copy of the moved element to pronounce. (See Nunes (1995), Groat and O'Neil (1996), Bobaljik (1995), Pesetsky (1997), Richards (1997), Roberts (1997), Boškovic (1999).) For example, Boškovic (1999) presents data from multipleWh questions in Serbo-Croatian. In general, all Wh-phrases must be fronted in the syntax in Serbo-Croatian, as shown by the ungrammaticality of (90)b-(90)d.

| a. | ko sta <br> who <br> what | gdje <br> where | kupuje? <br> buys |
| :--- | :--- | :--- | :--- | :--- |
| 'Who is buying what where? |  |  |  |

However, there are exceptions to the obligatory fronting of Wh-phrases. If a Whphrase is phonologically identical to another fronted Wh-phrase, both Wh-phrases must be fronted, as shown in (91). ${ }^{37}$
(91) šta uslovljava šta?
what conditions what

Boskovic proposes that both Wh-phrases are fronted in the syntax, and that the tail of the second Wh-phrase chain is pronounced so that there is not a violation of the PF constraint against sequences of homophonous Wh -words.
(92) [šta sta [uslovljava šta ${ }_{1}$ ]]
what what conditions what ${ }_{1}$

If there is an adverb, the two Wh -phrases no longer form a sequence of homophonous Wh-words. Since there is no violation of the PF constraint against sequences of homophonous Wh-words, the second Wh-phrase must be fronted, as in (93)a, and may not remain in-situ, as in (93)b.
a. šta neprestano šta uslovljava?
what constantly what conditions 'What constantly conditions what?'

A Wh-phrase can be left in-situ only as a last resort to prevent a violation of the PF constraint.

[^24]To summarize, children know that there are two negative syntactic structures. The derivation of questions with sentential negation is the same as for the adult, and hence they produce grammatical sentential negation questions and judge them to be acceptable. However, the children have incorrect lexical information with respect to constituent negation. They incorrectly think that both negative elements, $n^{\prime} t$ and not, can be used for constituent negation. The derivation of grammatical constituent negation questions with not is the same as for adults. The children's production and judgment of constituent negation questions with not are adult-like. The children's non-adult production of negative 2Aux questions, as in (94), is a performance error driven by choosing the affix $n^{\prime} t$ in a constituent negation structure.
(94) *What did the smurf didn't buy?

The use of $n^{\prime} t$ results in an affix without a host, and in an attempt to save the derivation from crashing at PF , the child spells out both copies of Infl at PF . Since the moved Infl and the copy of the moved Infl are both [+past, 3psg], this analysis correctly predicts that children only produce 2Aux questions with matching auxiliary verbs. When the child is only required to judge a 2Aux question, she correctly rejects the sentence.

[^25]Now that we have a proposal to account for the production and judgment data, we must determine whether it is learnable by the child. How would the child leam that only not can be used for constituent negation?

In order to differentiate between the sentential and constituent interpretation of a sentence with not, the stress on not and the pause before it is crucial. Without this stress and intonation pattern, it is difficult, if not impossible, to differentiate the two interpretations. If the child is unsure of the stress and intonation requirement for constituent negation, she may use $n^{\prime} t$ for constituent negation. This would hold even if the child knew the general restriction that affixes such as $n$ 't cannot be stressed. When the child learns that the negative element in constituent negation must be stressed and have a pause before it, she will presumably stop using $n$ 't for constituent negation. Once this happens, the child will stop producing 2Aux questions, since under our proposal, as long as $n ' t$ is not chosen from the lexicon for constituent negation, a 2Aux question will not be generated. One question that arises is why children don't choose to violate the obligatory stress requirement once they know it, and continue to produce 2Aux questions. One possibility is that certain requirements are more easily violable than others during production. For example, the requirement to pronounce only one copy seems to be more easily violated than the requirement for obligatory stress.

The idea that young children have not mastered stress and intonation has been observed by many researchers. (See for example, Halbert (1997) and Vogel and

Raimy (2000).) We have some support for this from Experiment 2. Even though children did use emphatic stress correctly, they did not always use it when it was required. Although the adult subjects always used emphatic stress in a question in response to a prompt with emphatic stress, the children did not. Recall that four of the nine children used emphatic stress less than $40 \%$ of the time in response to prompts with emphasis. (See (56).) Also, when the children used emphatic stress, it was not always on the appropriate element. (See f.n. 23.)

### 2.7 Summary

In this chapter, I explored one of the criticisms raised with respect to eliciting acceptability judgments from children. Many researchers have questioned children's metalinguistic abilities, while others, such as McDaniel and her colleagues have had much success obtaining acceptability judgments from children. From the series of experiments regarding children's 2Aux questions, we have additional evidence supporting McDaniel that children can be trained to give reliable and consistent judgments.

We also found that different methodologies can uncover different aspects of the child's competence. Of course, we cannot categorically claim that one methodology is better than another at revealing the child's grammar. We must evaluate each study separately and determine which methodology is giving us a better
picture of the child's competence. With respect to 2Aux questions, we found that the previous studies relying on just the elicited production task was not capturing the child's knowledge about these questions. Previous researchers had proposed that children produce such questions because they have a non-adult grammar. However, the results from the grammaticality judgment task (Experiment 3) indicate that children do not accept such questions. By combining methodologies, we were able to discover that the data from the grammaticality judgment experiment revealed a different picture about the children's competence regarding 2Aux questions.

Based on this difference between the production and judgment data, I proposed that the children think both $n^{\prime} t$ and not can be used for constituent negation. When they generate a constituent negation question, they may incorrectly choose $n$ 't. In order to avoid a PF crash, the child must rely on a "last resort" operation and spell out more than one copy of Infl. If the child is only required to judge a 2 Aux question, she correctly rejects the sentence.

The children's derivations of sentential negation questions and constituent negation questions with not are the same as the adult derivations. Since the children's derivations are adult-like, their production and judgment of these questions are adultlike.

## Chapter 3

## Syntactic satiation effects

### 3.1 Previous treatment of island violations

The sentence types that have been noticed to show satiation effects are mainly island violations. The four types of island violations that Snyder $(1994,2000)$ tested were Wh-island, Complex NP Constraint (CNPC), Subject island and Adjunct island violations.
a. ?? Who does John wonder whether Mary likes ( t )?

Wh-island
b. ?? Who does Mary believe the claim that John likes ( $t$ )?

CNPC
c. ?* What does John think that a bottle of ( $t$ ) fell on the floor? Subject island
d. ?* Who did John talk with Mary after seeing ( t )?

Adjunct island
Wh-island and CNPC violations are generally taken to be milder violations than Subject and Adjunct island violations, and these are precisely the sentence types that showed satiation effects in Snyder's study. Snyder also found that Subject island violations showed a marginally significant satiation effect.

Given these experimental results, we might expect the two milder subjacency violations to form a natural class, while the status of Subject island violations remains unclear. Let us consider some possible interpretations of three recent syntactic theories, and what predictions they might make based on these interpretations.

Ross (1967) first discussed a list of separate island constraints, such as the CNPC, which Chomsky (1973) tried to mify with the Wh-island and Subject island Constraint. All of these subjacency violations were accounted for by a proposal that a single instance of movement could cross no more than one bounding node, where bounding nodes in English are NP and S.
(96) No rule may move an element from the position $Y$ to either position of $X$ or conversely in the following configuration:
...X...[ $\alpha \ldots[\beta \ldots Y$...]...]...X...
where $\alpha$ and $\beta$ are bounding nodes. (Chomsky (1977))

Huang (1982) noted another island effect, the Adjunct island, and proposed two different accounts for extraction out of various islands. He treats Wh-island and CNP effects as subjacency violations, as before, and proposes the Condition on Extraction Domain (CED) to account for the ban on extraction out of subjects and adjuncts.
(97) Condition on Extraction Domain (CED)

A phrase A may be extracted out of a domain B only if B is properly governed.

Given Huang's treatment of these different island violations, pure subjacency violations, such as Wh-island and CNPC violations, might then pattern together with respect to satiation effects, whereas CED violations, such as Subject and Adjunct island violations, constitute a different natural class and behave differently from simple subjacency violations.

Chomsky (1986) unifies traditional cases of subjacency violations, such as Wh-island and CNPC violations, with those subsumed under Huang's (1982) Subject
and Adjunct Condition of the CED, using the notion of barriers. The key terms are defined below.
(98) $\gamma$ is a blocking category ( BC ) for $\beta$ iff $\gamma$ is not $L$-marked and $\gamma$ dominates $\beta$.
(99) Where $\alpha$ is a lexical category, $\alpha$ L-marks $\beta$ iff $\beta$ agrees with the head of $\gamma$ that is $\theta$-governed by $\alpha$.
A L-marks $B$ if $B$ is a complement of $A$, A lexical
(100) $\gamma$ is a barrier for $\beta$ iff (a) or (b):
a. $\gamma$ immediately dominates $\delta, \delta$ a BC for $\beta$;
b. $\gamma$ is a $B C$ for $\beta, \gamma \neq \mathbb{P}$.
(101) $\beta$ is $n$-subjacent to $\alpha$ iff there are fewer than $n+1$ barriers for $\beta$ that exclude $\alpha$.

For full grammaticality, 0-Subjacency (or fewer than one barrier crossed) is required. Wh-island and CNPC violations are weak subjacency violations since movement crosses only one barrier, as shown in (102) and (103), while movement out of Subject and Adjunct islands crosses two barriers and hence results in a stronger subjacency violation, as shown in (104) and (105). ${ }^{38}$
(102) [CP who does [ $\mathbb{P}$ John [vp t [vp wonder [CP whether [ $\mathbb{P}$ Mary [ vp t [vp likes t[]I]I]I]
(103) [CP who does [rp Mary [vp t [vp believe [ NP the claim [ CP that John likes t]]]]]

[^26] [ vp fell]] 1 ] $] 1]$
(105) [CP who did [ ${ }_{\mathbb{P}}$ John talk with Mary [xp after seeing t t]]

If we take the number of barriers crossed as the criterion for distinguishing natural classes of subjacency violations, Chomsky might make the same predictions as Huang (1982). Weak subjacency violations, such as Wh-island and CNPC violations, should pattern together with respect to satiation effects and form a natural class. Strong subjacency violations, such as Subject and Adjunct islands, should form a different natural class and behave differently from the weak subjacency violations.

Takahashi (1994), like Chomsky (1986), uniformly accounts for subjacency violations and CED effects. He eliminates barriers, and proposes a more elegant account within the Minimalist framework using the Shortest Move Condition (SMC) proposed by Chomsky and Lasnik (1993), and the Uniformity Corollary on Adjunction (UCA).
(106) Shortest Move Condition (SMC) Make the shortest move.

## (107) Uniformity Corollary on Adjunction (UCA)

Adjunction is impossible to a proper subpart of a uniform group, where a uniform group is a nontrivial chain or a coordination.

If the SMC is violated, the offending trace is marked with a star, which results in mild ungrammaticality. A starred trace remaining at LF will result in complete ungrammaticality. Following Chomsky (1991), Takahashi assumes that legitimate LF
objects include umiform chains resulting from head movement, A-movement and $A^{\prime}$ movement of adjuncts, as well as operator-variable pairs.

In $A^{\prime}$-movement of arguments, any starred trace ( ${ }^{*} \mathrm{t}$ ) resulting from a SMC violation is deleted before LF in order to convert it into a legitimate LF object. This results in mild ungrammaticality. Movement out of a Wh-island violates the SMC, and movement out of CNP's, Subject islands and Adjunct islands all violate the SMC due to the UCA.
 [ vp likes ti]]][]]I]]
 [ip John [vp t [vp likes t]7]]I]]]]]]]
 [ NP bottle of t$]\left[\mathrm{vp} \mathrm{t}_{\mathrm{DP}}\right.$ fell 17$\left.\left.][7] 717\right]\right]$
(111) [ CP who did [ $\mathbb{P}$ John talk with Mary [ PP _ [pp after [ $\mathbb{P}$ * $\mathbf{t}[\mathbb{P}$ seeing $t]]]]]$ In the Wh-island violation in (108), movement of the Wh-element from the embedded $\mathbb{P}$-adjoined position passes over the embedded [spec, CP] position, which is already filled by whether. This forces a violation of the SMC; the starred trace is deleted before LF, resulting in mild ungrammaticality.

The other three island effects are accounted for by the UCA and the SMC. In the case of CNPC violations, Takahashi assumes, following Stowell (1981) and Grimshaw (1992), that apparent noun complement clanses are appositive and hence
adjuncts, rather than complements. Takahashi also adopts Higginbotham's (1985) idea that adjuncts involve coordination. Given these two assumptions, the UCA prohibits the Wh-phrase from adjoining to $C^{\prime}$ in (109) because the noun complement is a type of coordination, and the movement forces a violation of the SMC. ${ }^{39}$

In Subject islands, as in (110), the UCA prohibits the Wh-phrase from adjoining to $\mathrm{D}{ }^{\text {, }}$, since it is a nontrivial chain, and forces a violation of the SMC. Takahashi assumes that all subjects are base-generated within VP. Therefore, all subjects are nontrivial chains.

In Adjunct islands, as in (111), the UCA prohibits the Wh-phrase from adjoining to PP because it is a type of coordination, and forces a violation of the SMC.

Takahashi might predict two different classes of island violations based on whether the UCA is used. Wh-island violations should behave differently from CNPC, Subject island and Adjunct island violations with respect to satiation effects, since the second group of sentence types all rely on the UCA, while Wh-island violations do not. ${ }^{40}$

To summarize the recent theories highlighted here, Huang (1982) and Chomsky (1986) group Wh-island and CNPC violations as a natural class, and Subject

[^27]and Adjunct islands as a different natural class. Takahashi (1994) treats Wh-island violations differently from CNPC, Subject island and Adjunct island violations.

In the next section, I review the experimental study by $\operatorname{Snyder}(1994,2000)$. In section 3.3, I present the goals and methodology for Experiment 1 , followed by the results in section 3.4. I discuss the results from Experiment 1 in section 3.5. In sections 3.6 and 3.7, I present the goals and methodology of Experiment 2, followed by the results. In section 3.8 , I discuss the follow-up experiment. I summarize the chapter in section 3.9.

### 3.2 A review of Snyder $(1994,2000)$

Snyder (1994) investigated the possibility of experimentally inducing syntactic satiation effects that had been noticed anecdotally by various linguists. He designed an experiment to determine whether subjects would change their acceptability judgments on various violations, and replicated what had been noticed anecdotally. Mild island violations, such as Wh-island and CNPC violations, showed satiation effects, while stronger island violations, such as Subject and Adjunct island violations, did not.

Twenty-two paid undergraduate native English speakers participated in a yes/no acceptability judgment task consisting of (58) sentences. Each sentence was presented sequentially on separate pages of a printed questionnaire. The items were
set up so that a "context" sentence describing a situation was followed by a "test" sentence for that particular context.
(112) Context: Maria believes the claim that Beth found a $\$ 50$ bill.

Test sentence: "What does Maria believe the claim that Beth found?"

The questionnaire included (4) practice items, (50) experimental items (all in the form of a Wh-question), and a (4) item post-test. The experimental items were randomly divided into five blocks, each containing (3) acceptable items and (7) unacceptable sentence types varying in degree from mild to severe. The seven unacceptable sentence types are given below. ${ }^{41}$
(113) a. Who does John wonder whether Mary likes ( $t$ )?

Wh-island
b. Who does Mary believe the claim that John likes ( $t$ )?
c. What does John think that a bottle of ( $t$ ) fell on the floor?
d. Who did John talk with Mary after seeing ( $t$ )?
e. Who does John want for Mary to meet ( $t$ )?
f. Who does Mary think that ( t ) likes John? Complex NP
Subject island Adjunct island

Want-for ${ }^{42}$
That-trace
g. How many did John buy ( $t$ ) books?

Left Branch Condition

In order to measure syntactic satiation for each sentence type, each subject's number of yes responses in the first two blocks was compared to the number of yes responses in the last two blocks. If the number of yes responses in the last two blocks exceeded the number of yes responses in the first two blocks, the subject was considered to have shown a syntactic satiation effect. For each sentence type, the number of subjects who showed satiation were compared to the number of subjects who did not show such an

[^28]effect (i.e. subjects who had at least as many no responses in the last two blocks as the first two blocks). Satiation effects were measured in this way for each particular sentence type. A sentence type was judged to show a statistically significant satiation effect if the preponderance of satiating (no to yes) subjects over non-satiating subjects was significant ( $p<.05$ ) by sign test.

The Wh-island and CNPC violations were expected to be subject to syntactic satiation, given the anecdotal evidence of satiation. On the other hand, that-trace effects and Left Branch Condition (LBC) violations were predicted not to be subject to syntactic satiation. There was no prediction made for the other sentence types, which were included primarily as fillers and also to provide a range of degrees of unacceptability.

As predicted, there was a statistically significant satiation effect for both Whisland ( $p=.003$ ) and CNPC violations ( $p=.031$ ). There was no significant satiation effect for violations of that-trace or LBC violations. Of the remaining three sentence types, none showed a statistically significant satiation effect. There was, however, a marginally significant no to yes response shift for Subject islands.

The post-test was used to investigate whether satiation on Wh-island and CNPC violations using one lexical item would extend to other lexical items. All of the experimental Wh-island items used the matrix verb wonder and all of the experimental CNPC items used the phrase "believe the claim." The post-test consisted of four items, including one Wh-island item with the matrix verb ask, and one CNPC item using the
phrase "accept the idea." Subjects who had initially rejected Wh-island items with wonder and had satiated on these items were analyzed to determine whether this satiation carried over to the post-test Wh-island item with ask (binomial $p=.019$ ). Subjects who had satiated on CNPC items were analyzed to determine if this satiation carried over to the post-test CNPC item with "accept the idea" (binomial $p=.013$ ).

The results from the post-test analysis show that syntactic satiation on one lexical item does carry over to other lexical items. Syntactic satiation effects were not tied to specific lexical items, such as wonder whether or believe the claim. It does not seem that satiation is just learning believe the claim as a lexical variant of believe.

### 3.3 Experiment 1

### 3.3.1 Goals

In this experiment, we focus on two main questions that arise from Snyder's (1994) study. First, the results for the satiation of Subject islands were not conclusive. Subjects satiated somewhat, but the effect was only marginally significant. One possible explanation for this is that satiation for Subject islands is slower than for Whisland and CNPC violations. One way to test this is to have more Subject island items by having more blocks of experimental items. By looking at the changes across more blocks, we may be able to obtain clearer results.

Second, Snyder found that subjects satiated on Wh-island and CNPC violations, which are both weak subjacency violations. However, we cannot yet conclude that subjacency is necessarily the relevant factor. Subjects may have satiated on these two subjacency violations by coincidence. In order to determine whether subjacency is indeed the relevant factor, we investigated whether syntactic satiation for one type of subjacency violation carries over to another type of subjacency violation. This may also help us to determine whether certain subtypes of subjacency violations constitute a natural class, while other subtypes traditionally handled as subjacency violations, perhaps reflect a distinct phenomenon.

### 3.3.2 Subjects

In order to keep the effect of age to a minimum, the subject pool was restricted to undergraduate students. Forty paid undergraduate students from the University of Connecticut and Northwestern University participated in the study. Subjects were excluded if they did not respond correctly to at least $90 \%$ of the filler items (all grammatical sentences). In other words, any subject who responded incorrectly on more than two of the filler items was excluded. Seven subjects were excluded using this criterion. The results reported in section 3.4 are from the remaining thirty-three subjects. Of the thirty remaining subjects, eighteen participated in a follow-up study one to three weeks after the experiment.

### 3.3.3 Methodology

Each subject was first given a background form to fill out, which asked questions about their gender (male/female), handedness (leftright), knowledge of a second language (yes/no), exposure to linguistics classes (yes/no), and previous participation in a judgment task (yes/no).

Before beginning the experiment, subjects were given written instructions on how to provide judgments. ${ }^{43}$ I reviewed the instructions with each subject to confirm her understanding of the task. Subjects were asked to judge whether the test sentence was "grammatically possible" for them as a native speaker of English, with the following explanation of "grammatically possible." The subjects were told that we were not concerned about whether the sentence would be acceptable to a writing teacher. I explained that points of style and clarity, who versus whom, or ending a sentence in a preposition were not the issue. I told them that we wanted to know whether the sentence could have the intended meaning and still be accepted as English. Subjects were also informed that this was not a memory test, and that they weren't required to remember the judgments they gave for previous items. Furthermore, subjects were told that many of the items would be similar to one another, and that they should ignore this fact and provide an independent judgment on
each individual item. Finally, I asked subjects to work through the questionnaire as quickly as possible, going with their initial reaction for each sentence. (See Appendix 2 for the set of instructions the subjects received.)

I gave the subjects a clear example of a grammatically possible sentence, as in (114), and a grammatically impossible sentence, as in (115). I explained to the subjects that not all of the test sentences would be this clear cut, but that they should do their best to provide a Yes or No judgment.
(114) Context: Chris peeled an apple.

Test sentence: "What did Chris peel?"
(115) Context: Susi wants very much for Jonathan to meet Diana. Test sentence: "Who does Susi want very much whether for who to meet Diana?"

The design of the experiment is similar to Snyder (1994). The questionnaire consisted of a series of (70) sentences which were printed sequentially, one item per page. Each item had a context sentence describing the situation, paired with a test sentence that was appropriate for that particular context, as shown in (116).
(116) Context: Maria believes the claim that Beth found a $\$ 50$ bill. Test sentence: "What does Maria believe the claim that Beth found?" Judgment: $\qquad$ (Y/N)

The context was provided to insure that each subject was judging the test sentence based on the same context.

[^29]There were (4) practice items, (63) experimental items, and a (3) item post-test. The sixty-three experimental items were randomly ordered and arranged into seven blocks, which is two blocks more than in Snyder (1994).

Each block contained three acceptable items, as in (II7), one weak subjacency violation, as in (118), and one each of five unacceptable sentence types, as in (119). Ergative predicates were chosen for Subject island violations because these verbs lent themselves to inanimate subjects. ${ }^{44}$ Not all of the unacceptable sentence types were island violations.
(117) Examples of acceptable items
a. What does Sharyl claim that Peter studies?
b. What does Alex claim to believe that Sandy concocted?
c. How many cats does Paula think that Wanda brushed?
(118) Examples of weak subjacency violations
a. Who does John wonder whether Mary likes (t)?

Wh-island
b. Who does Mary believe the claim that John likes ( t )?

CNPC
(I19) Examples of other unacceptable sentence types
a. What does John know that a bottle of ( $t$ ) fell on the floor?

Subject island
b. Who did John talk with Mary after seeing ( t )?

Adjunct island
c. Who does John want for Mary to meet ( $t$ )?

Want-for
d. Who does Mary think that ( $t$ ) likes John? That-trace
e. How many did John buy ( t ) books? Left Branch Condition

With respect to the weak subjacency violation items, fifteen subjects judged only Wh-island violations, and eighteen subjects judged only Complex NP Constraint violations. The Wh-island violations all had finite embedded interrogative clauses,

[^30]and none of the Complex NP Constraint violations involved relative clauses, as in (120). All of the experimental items were in the form of a Wh -question.
(120) ?* What did John meet [NP a child [CP who ate t ]] CNPC (relative clause)

The post-test contained (3) items, including one weak subjacency item: either a CNPC or Wh-island violation. The subjects who were tested on Wh-island violations judged a CNPC violation on the post-test, while the subjects who were tested on CNPC violations judged a Wh-island violation on the post-test.

As a follow-up study, the subjects were asked to judge a set of ten sentences one to three weeks later to determine whether the satiation effects were still present.

In the next section, I first present the satiation results for each sentence type, followed by a discussion of possible associations between individual characteristics and satiation effects.

### 3.4 Results from Experiment I

### 3.4.1 Comparison of satiation effects with Snyder (1994)

A paired t-test was conducted on the values for the first three blocks and the last three blocks for each subject who received the given sentence type in her questionnaire. A significant satiation effect was obtained for: Wh-islands $(t(13)=2.223, p=.045)$, Subject islands $(t(32)=2.031, p=.051)$, That-trace $(t(31)=2.436$, $p=.021)$, and $W$ ant-for $(t(32)=3.672, p<.001)$. No satiation effect was obtained for

CNPC $(t(17)=369, p=.717)$, Adjunct islands $(t(31)=197, p=.845)$, or LBC ( $t(32)=.702, p=.488$ ). A summary of the results from Experiment 1 and Snyder (1994) is given below.
(121) Table 5: Summary of satiation results

| Sentence type | Snyder (1994) | Experiment 1 |  |  |
| :--- | :---: | :---: | :---: | :---: |
| Wh-island | $\checkmark$ | $\checkmark$ |  |  |
| CNPC | $\checkmark$ |  |  |  |
| Subject island | $(\checkmark)$ | $\checkmark$ |  |  |
| Adjunct island |  |  |  |  |
| Want-for |  | $\checkmark$ |  |  |
| That-trace |  | $\checkmark$ |  |  |
| LBC |  |  |  |  |
| =subjects showed satiation |  |  |  | $(\checkmark)$ =marginally significant |

### 3.4.1.I Subject island violations

One goal of this experiment was to obtain clearer data for Subject islands. Snyder (1994) found that subjects satiated somewhat, but the effect was not statistically significant. One possible explanation for this is that satiation for Subject island violations is slower than for Wh-island and CNPC violations. We can test this hypothesis by looking at the first five blocks of Subject islands (as opposed to the full seven blocks) to see if there was a significant satiation effect. A paired t-test was conducted on the values for the first two blocks and the third and fourth blocks for each subject, and there was no significant satiation effect $(\mathrm{t}(32)=1.030, \mathrm{p}=.311)$. Since there was a significant satiation effect for seven blocks but not for five blocks, it
seems that having two additional blocks of experimental items did enable subjects to satiate on Subject island violations.

Given the predictions that the various syntactic theories make, it is unexpected to find Subject islands, but not Adjunct islands, showing significant satiation. Huang (1982), Chomsky (1986) predict Subject island and Adjunct island violations to behave similarly. If two sentence types both show satiation effects, we cannot conclude that they form a natural class; it could be coincidence that they both behaved similarly. However, if two sentence types that are assumed to be treated the same by syntax show different satiation effects, then we can conclude that they do not form a natural class. We will discuss this in more detail in section 3.5.

### 3.4.1.2 CNPC violations

One major difference between the two studies is the satiation effect for CNPC violations. Snyder (1994) found it to be statistically significant, while this study did not. There are two potential interfering factors for the difference in results. First, the difference in methodology may be contributing to the difference in results. In this study, subjects were exposed to only Wh-island or only CNPC violations, while subjects in Snyder's study were exposed to both Wh-island and CNPC violations. The exposure that Snyder's subjects had to both types of weak subjacency violations may have allowed satiation more easily. However, if both types are equally hard to satiate,
then subjects in this study should not have shown satiation effects for Wh-island violations either.

Another factor to consider is the fact that of the subjects who were exposed to Wh-island violations, $60 \%$ accepted the first test item that they saw. Of the subjects who were exposed to CNPC violations, only $17 \%$ accepted the first test item that they saw. This suggests that the subjects in this study found CNPC violations less acceptable overall than Wh-island violations. Subjects may be satiating more easily on Wh-island violations than CNPC violations because Wh-island violations are generally milder than CNPC violations. Another possibility is that because CNPC violations are initially stronger, it takes a greater degree of satiation before a shift from no to yes is observed.

Continued research might clarify these different results, and may also shed some light on the status of CNP's: whether noun complement clauses are not complements at all but rather appositives/adjuncts. If further research indicates that CNPC violations do not show satiation effects, it might be taken as evidence for treating the noun complement clause as an adjunct, since Adjunct island violations don't show satiation effects either. However, if we find that CNPC violations do show satiation effects, it might be taken as evidence for treating the noun complement clause as a complement.

### 3.4.1.3 That-trace and want-for constructions

Another difference between the two studies is the results for that-trace and -want-for sentences. Snyder (1994) did not find any satiation effects, while this study did. The results from my study are unexpected since linguists do not satiate on these items. One possible interfering factor is that the judgments for these sentence types differ dialectally. Sixty percent of the subjects (21 out of 33 ) were from the University of Connecticut, which seems to allow that-trace constructions, while Snyder's subjects were geographically more diverse. ${ }^{45}$

Also, a few subjects had crossed out the word that on the questionnaire, suggesting that they were not judging these sentences, and perhaps want-for sentences, in the same way that they were judging other sentences. The fact that some of the subjects made explicit corrections on the questionnaire might also be an indication that there may have been subjects who made mental corrections to the sentences.

It is not clear what to conclude about the difference in satiation results. Since there are some factors that might have been affecting these two sentence types, but not other sentence types, I will leave that-trace and want-for sentences as an area for further research.

[^31]
### 3.4.2 Post-test results

The goal of the post-test in this study was to determine whether satiation on one type of weak subjacency violation carries over to another type of weak subjacency violation. By looking at carry-over effects we can test whether two sentence types are showing satiation effects by coincidence, or are showing satiation effects because the two sentence types reflect a natural class. Fifteen subjects were exposed only to Whisland violations and one CNPC post-test item; eighteen subjects were exposed only to CNPC violations and one Wh -island post-test item.

To determine whether satiation on CNPC violations carried over to the Whisland post-test item, I conduced a repeated measures one-within (beginning vs. end), one-between (yes vs. no on the post-test item) ANOVA. There was a marginally significant interaction between beginning vs. end for CNPC violations and Wh-island post-test response $(F(1,16)=3.900, p=.0658)$. Subjects who accepted the Wh-island post-test item show a higher degree of satiation on CNPC violations than subjects who didn't accept the Wh-island post-test item. To determine whether satiation on Whislands carried over to the CNPC post-test item, I conducted another repeated measures one-within, one-between ANOVA. However, there was no significant interaction effect ( $F=(1,12)=.002, p=9677$ ).

### 3.4.3 Potential contingencies with other variables

Since subject-related characteristics may affect the judgment process, it is important to establish whether there is an association between satiation and these variables. I conducted Fisher's exact tests to determine whether any of the measured variables were associated with an increase in the number of yes responses in the last three blocks (syntactic satiation). We only considered subjects with 0,1 , or 2 yes responses in the first three blocks. Subjects with all yes responses in the first three blocks were excluded since they do not have the potential of increasing the number of yes responses in the last three blocks. Subjects were considered to have shown a satiation effect if there was an increase in the number of yes responses in the last three blocks.

The variables that were analyzed are given in (122): gender, handedness, knowledge of a second language, exposure to linguistics classes, and previous participation in a judgment task.
(122) a. gender (male/female)
b. handedness (leftright)
c. knowledge of a second language (yes/no)
d. exposure to linguistics classes (yes/no)
e. previous participation in a judgment task (yes/no)

Eighteen subjects also participated in a follow-up study one to three weeks later, in which they were asked to judge a set of ten sentences of the types given in ((117)-(119)). The purpose of the follow-up test was to determine whether satiation effects would continue over time.

I considered the four test types that showed a significant satiation effect: Whislands, Subject islands, want-for, and that-trace sentences. There were six subjects who showed an increase in the number of yes responses for Wh-islands, and three subjects who did not show an increase. None of the measured variables were associated with satiation on Wh-islands by Fisher's Exact test.
(123) Table 6: Fisher's exact test results for Wh-islands

| Gender | Handedness | Second <br> language | Linguistics <br> classes | Previous <br> Judg. Task | Follow-up <br> experiment |
| :--- | :--- | :--- | :--- | :--- | :--- |
| $p=1.0$ | $p=1.0$ | $p=1.0$ | $p=.226$ | $p=1.0$ | $p=1.0$ |

There were twelve subjects who showed an increase in the number of yes responses for Subject islands, and seventeen subjects who did not show an increase. Again, none of the measured variables were associated with satiation on Subject islands by Fisher's Exact test.
(124) Table 7: Fisher's exact test results for Subject islands

| Gender | Handedness | Second <br> language | Linguistics <br> classes | Previous <br> Judg. Task | Follow-up <br> experiment |
| :--- | :--- | :--- | :--- | :--- | :--- |
| $p=.449$ | $p=.624$ | $p=.717$ | $p=1.0$ | $p=1.0$ | $p=1.0$ |

With respect to want-for sentences, there were eleven subjects who showed an increase in the number of yes responses, but only one subject who did not. Therefore, I did not test for potential associations.

Finally, there were eleven subjects who showed an increase in the number of yes responses for that-trace sentences, and thirteen subjects who did not show an increase. There was no significant association between any of the measured variables and satiation for this test type either.

## (125) Table 8: Fisher's exact test results for that-trace

| Gender | Handedness | Second <br> language | Linguistics <br> classes | Previous <br> Judg. Task | Follow-up <br> experiment |
| :--- | :--- | :--- | :--- | :--- | :--- |
| $p=1.0$ | $p=.481$ | $p=.679$ | $p=386$ | $p=1.0$ | $p=.266$ |

### 3.5 Discussion of Experiment I

As we saw in the previous section, we do not have a clear picture of the results for CNPC violations and that-trace and want-for sentences. However, we did find an unexpected difference between satiation effects for Subject and Adjunct islands, and in this section, we will focus on this difference. We found that Wh-islands and Subject islands show satiation effects, while Adjunct islands do not. Although none of the syntactic theories we reviewed in section 3.1 were constructed to handle satiation data, we can still look at how they might be modified to incorporate such data.

### 3.5.1 Huang (1982)

If we take the approach that the kind of violation (subjacency vs. the CED) distinguishes natural classes of island effects, Huang (1982) predicts Wh-islands to behave differently from both Subject and Adjunct islands. He accounts for both Subject and Adjunct islands under the CED, and hence predicts that these violations should form a natural class.

If we were to consider individual parts of the CED separately and distinguish Subject islands from Adjunct islands in this way, Huang's theory would capture the satiation data but would then lose the argument adjunct asymmetry. Subjects who satiate on the CED and thereby allow argument extraction out of Subject islands, should also allow adjunct extraction out of Subject islands. Although this needs further testing, this possibility is unlikely since linguists do not show any tendency to satiate on adjunct extraction in general. (In Experiment 2, we test argument vs. adjunct extraction out of Wh-islands. See section 3.6.) It seems difficult to modify Huang's theory to incorporate the fact that Subject islands but not Adjunct islands show satiation effects.

### 3.5.2 Chomsky (1986)

If we take the number of barriers a movement crosses to be the criterion determining natural classes of subjacency violations, Chomsky (1986) makes
predictions similar to Huang (1982) with respect to Subject and Adjunct islands. Modifications to the theory are possible to account for the difference in satiation for these two islands, but not without problems.

Chomsky distinguishes between weak subjacency violations (Wh-islands) and strong subjacency violations (Subject and Adjunct islands) based on the number of barriers the movement crosses. There seems to be a correlation between the number of barriers crossed and whether there are satiation effects. When one barrier is crossed, as in argument extraction out of a Wh-island, there is a satiation effect. When two barriers are crossed, as in argument extraction out of an Adjunct island, there are no satiation effects.

As with Huang's theory, the problem with Chomsky's theory is that Subject islands are predicted not to show satiation effects. Since argument extraction out of a Subject island crosses two barriers, the theory predicts there to be no satiation effect.
 floor[]I]I]

In (126), the movement of the Wh-element out of the subject NP to the lower [spec, CP] crosses two barriers: the NP is not L-marked and is both a blocking category and a barrier, the IP "inherits" barrierhood from the NP (by (100)a).

One possibility to account for the Subject island satiation effect is to find a way to reduce the number of barriers crossed to one. Let us incorporate the VP internal subject hypothesis into Chomsky's theory, and assume that the subject NP is
base-generated in [spec, VP], as opposed to [spec, IP]. ${ }^{46}$ Suppose that the Wh-element moves out of the subject first, before the subject raises to [spec, $\mathbb{P}]$, and adjoins to the VP.


The NP in (127) is a blocking category and a barrier since it is not L-marked by the lower verb. The subject NP and the Wh-element would move alternately following the cycle. Given this derivation, movement of the Wh-element out of the subject crosses only one barrier. ${ }^{47}$

The modifications necessary to reduce the number of barriers to allow extraction out of Subject islands may be contributing to why subjects take longer to satiate on this sentence type. Subjects required two additional test items (a total of seven test items) to show significant satiation effects for Subject islands. In comparison, Wh-islands, which cross one barrier without any modification to the structure, showed significant satiation effects after exposure to five test items (Snyder (1994)).

[^32]Although this modification seems to be a possible explanation for why Subject island violations show satiation effects, there is a problem with this solution. If this alternative derivation is available, it should be the preferred derivation overall since it crosses only one bamier. However, only $24 \%$ of the subjects accepted the first Subject island violation that they saw. In order for the derivation that crosses two barriers to be the initially preferred derivation, we would have to assume that there is some constraint X that favors movement of the entire subject NP to [spec, IP] first. We can then define satiation as a weakening of constraint X , which makes the alternative derivation that crosses only one barrier the preferred derivation. ${ }^{48}$

### 3.5.3 Takahashi (1994)

Although he makes predictions similar to Chomsky (1986), Takahashi's theory is more appealing since there is no need to rely on the notion of 'government', which isn't a part of the current Minimalist framework, and we can use a derivation for Subject island violations similar to the one discussed in the previous section. If we assume that relying on the UCA is relevant in distinguishing natural classes of subjacency violations, Takahashi predicts Wh-islands to behave differently from Subject and Adjunct islands. Argument extraction out of a Wh-island violates the

[^33]SMC independently of the UCA, while argument extraction out of a Subject or Adjunct island violates the SMC due to the UCA.

Contrary to Takahashi's predictions, this study found Subject islands but not Adjunct islands to show satiation effects. One way to account for this unexpected split is to take the UCA as a disjunction. However, like the disjunction we considered for Huang's CED, a disjunctive interpretation of the UCA will predict adjunct extraction out of Subject islands to be acceptable. If subjects satiate on the first disjunct of the UCA (prohibition of adjunction to a proper subpart of a nontrivial chain), then neither argument nor adjunct extraction out of Subject islands violates anything. Given the anecdotal evidence that linguists do not satiate on adjunct extraction in general, this is an undesired consequence.

Another way to account for this unexpected subject/adjunct split is by modifying the derivation of extraction out of Subject islands so that we do not rely on the UCA. Consider a derivation similar to the one we used to modify Chomsky's account. Let us assume that the subject NP is base-generated as the complement of the lower verb since the verbs that were used were all unaccusative or passive.
 The Wh-element can first move out of the subject and adjoin to the NP. The subject NP can then raise to [spec, VP] by adjoining to $V^{\prime}$, followed by the raising of the Whelement, which can adjoin to VP. The derivation would proceed in this fashion, with the Wh-element and the subject NP alternating steps.

Since nothing is violated under this derivation, including the SMC, this alternative derivation should be the preferred one. However, Chomsky (1995:365) strengthens the condition on argument chains, given in (129), with the provision given in (130), so that this type of derivation is raled out.
(129) Only the head of CH can be attracted by K. [199] (130) $\alpha$ can be attracted by $K$ only if it contains no trace. [200]

If the Wh-element raises out of the subject NP first, the subject NP can no longer be attracted by AGRs to check the EPP feature since it contains a trace. We could define satiation as a weakening of this condition: satiation effects for Subject island violations are reflecting a weakening of the condition on argument chains for some speakers. ${ }^{49}$

### 3.5.4 Summary

In Experiment 1, we continued to investigate satiation effects for various islands that Snyder $(1994,2000)$ had first explored. We replicated Snyder's findings for Wh-islands, Adjunct islands, and LBC violations, and confirmed that with additional test items, Subject island violations do show satiation effects. Furthermore, we found an unexpected contrast between Subject islands and Adjunct islands, which

[^34]none of the syntactic theories seem to predict. The results for CNPC violations, Thattrace and want-for sentences in were inconclusive, and we leave this for further research.

We cannot conclude that all of the sentence types that show satiation effects constitute a natural class. The fact that they show satiation effects could simply be coincidence. For example, argument extraction out of Wh-islands and Subject-island violations both showed satiation effects, yet they do not necessarily constitute a natural class. There would have to be further investigation of carry-over effects to provide stronger evidence that Wh-island and Subject islands are part of the same natural class.

On the other hand, a difference in satiation effects between two sentence types can be taken as strong evidence that those sentence types do not constitute a natural class if we assume that members of the same natural class behave similarly with respect to satiation effects. This is the case with Subject and Adjunct islands. Contrary to the prediction made by most syntactic theories, Subject islands show satiation effects while Adjunct islands do not. Therefore, these two islands should not be treated the same syntactically. The current theories would have to be modified, perhaps as we suggested in this section, in order to account for the distinction between Subject islands and Adjunct islands.

[^35]
### 3.6 Experiment 2

### 3.6.1 Goals

In Experiment 2, I tested two island violations (adjunct extraction out of Whislands, Subject islands with transitive verbs) which were not tested in previous experiments. Snyder (1994, to appear) and Experiment I tested argument extraction out of Wh-islands, and in this experiment, I included adjunct extraction out of Whislands. If satiation effects are somehow related to subjacency violations but not the ECP, then adjunct extraction out of Wh-islands, which violate both the ECP and subjacency, should not show any satiation effectsSubject islands with transitive verbs were included in this experiment to compare the results with Subject islands with ergative verbs from Experiment 1 .

Some aspects of the methodology for Experiment 2 is also different from previous experiments. First, I added control items for the test questions, and included more fillers to balance ungrammatical and grammatical items. Second, the questionnaire was presented on the computer rather than on a printed questionnaire. Finally, there were two auxiliary tests that measured the subjects' reading ability and knowledge of authors.

### 3.6.2 Subjects

Forty-three native English-speaking adults participated in the experiment. I first calculated the mean percentage of correct responses on the filler/control items and excluded any subject whose mean score was more than two standard deviations away from the mean ( $52.77 \%$ correct filler/control items). Two of the 43 subjects were excluded based on this criterion.

Using the same criterion as in Experiment 1, I included subjects who had at least $90 \%$ correct responses on the filler/control items. This excluded $73.17 \%$ of the subjects ( 30 subjects), and left only 11 subjects for analysis. In section 3.8, I discuss possible reasons for the increase in the number of subjects who had more than $10 \%$ incorrect responses.

Since there were so few subjects who correctly responded to at least $90 \%$ of the filler/control items, I also considered another criterion. I also considered a second group of subjects, the 22 subjects who had at least the median score (percentage of correct responses) on the filler/control items.

### 3.6.3 Methodology

We used a yes/no acceptability judgment task again in Experiment 2, but instead of a printed questionnaire, a web-based computer questionnaire (developed by

David Braze) was used. This method allowed administering the questionnaire to groups of subjects, rather than individuals.

The subjects first filled in a background questionnaire, similar to the one used in Experiment 1, and then read a set of instructions and examples on how to judge the acceptability of a sentence. The instructions made it clear to the subject that s/he is not to change any aspect of the context or the test question. This was emphasized since in Experiment 1, there was evidence that some of the subjects had altered the test question. (See section 3.4.1.3.)

For each item, there was a context sentence describing a situation, followed by a test question for that particular context, as shown below.
(131) Context: Yesterday, Kelly wondered whether Tina had read 'Amistad'. Question: What did Kelly wonder whether Tina had read? This question is: Good Bad

The experimental items were randomly ordered and arranged into seven blocks. The items were balanced across subjects for forward/backward order of presentation. Each block contained one each of five unacceptable sentence types (not all island violations), three acceptable control items and six filler items. All of the experimental items were in the form of a Wh-question.

Each of the three unacceptable sentence types in (132), (134), (136) had a control item, as in (133), (135), and (137). The control items are minimally different from the unacceptable sentences. $\operatorname{In}$ (133) and (137), the context remains the same for
the control and test items but the question is slightly different. The question in both cases is an acceptable question with when. In (135), the context is modified slightly, while the question remains the same. For the control item, the temporal adverb this year is associated with the matrix clause, rather than the embedded clause, and the question is acceptable. An example of the two other test questions (Subject island and LBC) is shown in (138) and (139).
(132) Extraction of an argument out of Wh -island

Context: Yesterday, Kelly wondered whether Tina had read 'Amistad'.
Question: What did Kelly wonder whether Tina had read?
(133) Control item for extraction of an argument out of Wh-island

Context: Yesterday, Kelly wondered whether Tina had read 'Amistad'.
Question: When did Kelly wonder whether Tina had read 'Amistad'?
(134) Extraction of an adjunct out of Wh-island

Context: Gary wondered whether, this year, Jordan would learn French.
Question: When did Gary wonder whether Jordan would learn French?
(135) Control item for extraction of an adjunct out of Wh-island

Context: Gary wondered, this year, whether Jordan would learn French.
Question: When did Gary wonder whether Jordan would learn French?
(136) Extraction of an argument out of Adjunct island

Context: This moming, Frasier cleaned the bathroom, while Judy vigorously mopped the kitchen.
Question: What did Frasier clean the bathroom while Judy vigorously mopped this moming?
(137) Control item for extraction of an argument out of Adjunct island

Context: This moming, Frasier cleaned the bathroom, while Judy vigorously mopped the kitchen.
Question: When did Frasier clean the bathroom while Judy vigorously mopped the kitchen?
(138) Extraction of an argument out of Subject island with a transitive verb Context: Clinton worries that an ally of the labor union will boycott the airline.
Question: What does Clinton worry that an ally of will boycott the airline?
(139) LBC violation

Context: Karen believes that the cook peeled twenty potatoes.
Question: How many does Karen believe that the cook peeled potatoes?
The six filler items were included to balance the number of grammatical and ungrammatical questions. The context sentence for the filler item is similar to the context sentences for the relevant test item. For example, the context sentences for the test question with argument extraction out of Wh-island, as in (132), and the filler item, as in (140), both involve "wonder whether" in the matrix clause.
(140) Grammatical filler for Wh-island (argument extraction)

Context: Candace occasionally wonders whether Donald Trump still makes a lot of money.
Question: What does Candace wonder occasionaily?
(141) Ungrammatical filler for Wh-island (adjunct extraction)

Context: Every moming, Justin wonders whether the publishers will accept the book.
Question: When does wonder Justin whether the publishers will accept the book?
(142) Grammatical filler for Adjunct island

Context: After the trip, Steve unpacked the suitcases, while Diane noisily vacuumed the van.
Question: What did Diane noisily vacuum while Steve unpacked the suitcases?
(143) Ungrammatical filler for Adjunct island

Context: Morgan looked for seashells, while Christina energetically dug for clams.
Question: When did Christina energetically dug for clams?
(144) Grammatical filler for Subject island

Context: Frank hoped that a team of reporters would investigate the issue.
Question: What did Frank hope that a team of reporters would investigate?
(145) Grammatical filler for LBC

Context: Hillary guessed that Carol had wrapped forty presents last Christmas.
Question: How many presents did Hillary guess that Carol had wrapped?

In addition to the judgment task, there were two auxiliary tests: a fast reading task, and an author recognition task. Schütze (1996) discusses possible subject-related factors that may influence the judgment of sentences. Some are organismic factors, such as handedness, while others are experiential factors, such as linguistic training, literacy, or education. Since reading carefully is essential in judging the sentence types used in this study, I decided to focus on how well and how often the subject reads.

The fast reading task (Karlsen and Gardner (1984)) measures how quickly and accurately a subject reads. The subject is required to choose the most appropriate word for a sentence from a group of three words. The subject is measured on how many items she can complete correctly in three minutes. The author recognition task (Stanovich and West (1989), modified by Dave Braze) measures how many authors
the subject can identify out of a list of eighty names in two minutes. The list includes popular writers (of books, magazine articles, and/or newspaper columns), as well as non-authors.

We can compare how well subjects perform on these two tasks, and how well they perform on the filler/control items. This will allow us to examine whether individual differences, such as accuracy in reading or familiarity with reading (based on their awareness of well-known authors), correlates with how they perform on the judgment task in general.

In the next section, I present the results from Experiment 2. I first report the results with respect to satiation effects, followed by the results for the filler/control items and response time differences.

### 3.7 Results from Experiment 2

### 3.7.1 Satiation effects

Let us first consider the II subjects who had at least $90 \%$ correct responses on the filler/control items. For each sentence type, I compared the number of subjects who showed a shift in response from no to yes (in the first three blocks vs. the last three blocks), to the number of subjects who showed a shift in response from yes to no.
(146) Figure 3: Satiation effects for subjects with at least $\mathbf{9 0 \%}$ correct filler/control items


The satiation effect is statistically significant for argument extraction out of Whislands $(p=.035)$ by sign test, but not for any of the other sentence types. I also conducted a two-way repeated measures ANOVA of block (the number of yes responses in the beginning vs. the end) and test sentence type. There was no significant interaction effect between block and question type $(F(4,8)=.422, p=.791)$. However, the mean number of yes responses (out of three total) for beginning vs. end are showing the right trend for Wh-islands, as shown below.
(147) Table 9: Mean number of yes responses (beginning vs. end) for subjects with $\mathbf{9 0 \%}$ correct filler/control items

|  | Beginning | End |
| :--- | :---: | :---: |
| Argument extraction out of Wh-island | 2.2 | 2.7 |
| Adjunct extraction out of Wh-island | 2.1 | 2.2 |
| Subject island (transitive verb) | 0.1 | 0.4 |
| Adjunct island | 0.1 | 0.3 |
| LBC | 0.1 | $3.701 \mathrm{E}-017$ |

The means for adjunct extraction out of Wh-islands suggest that the subjects may not have been interpreting the adjunct correctly. The subjects were accepting these questions, which suggests that they were associating the adjunct with the matrix verb, rather than the embedded verb. However, despite the high acceptance rate ( 2.111 out of 3 at the beginning), there is still no significant satiation effect by sign test or by two-way repeated-measures ANOVA. The means for the Subject islands with transitive verbs show that the subjects did not accept them overall.

We now turn to the 22 subjects whose score on the filler/control items was equivalent or better than the median score ( $85.71 \%$ correct responses). The satiation effect for argument extraction out of Wh-islands was marginally significant by sign test $(p=.059)$, but not for any of the other sentence types.
(148) Figure 4: Satiation effects for subjects with at least median score on filler/control items


By a two-way repeated measures ANOVA of block and sentence type, I found no significant interaction effect $(F(4,16)=.455, p=.769)$. Again, the mean number of yes responses (out of three total) for beginning vs. end are showing the right trend for Whislands, as shown below.
(149) Table 10: Mean number of yes responses (beginning vs. end) for subjects with at least median score on filler/control items

|  | Beginning | End |
| :--- | :---: | :---: |
| Argument extraction out of Wh-island | 2.2 | 2.4 |
| Adjunct extraction out of Wh -island | 2.0 | 2.2 |
| Subject island (transitive verb) | 0.4 | 0.6 |
| Adjunct island | 0.1 | 0.3 |
| LBC | 0.2 | 0.2 |

To summarize the results from Experiment 2, the satiation effect for extraction out of Wh-islands was significant by sign test although not by ANOVA. There was no significant satiation effect for Adjunct islands or Left Branch Condition. This replicates the results for these sentence types from $\operatorname{Snyder}(1994,2000)$ and Experiment 1. There was no significant satiation effect for either of the two new test types: adjunct extraction out of Wh -islands and Subject islands with transitive verbs.

Although the results for satiation effects were mainly as we had expected them to be, the satiation effect for argument extraction out of Wh-islands was not as strong as in Experiment 1. However, the acceptance rates for the first Wh-island (argument extraction) presented were similar for Experiments I and $2: 60 \%$ vs. $61 \%$ respectively.

One possible explanation for the difference between the two experiments may be related to the presentation of materials. In Experiment 1, the subjects judged the stimuli on a written questionnaire, while in Experiment 2, they judged the stimuli on the computer. We examined this possible explanation by having subjects judge the
same stimuli from Experiment 2 on a written questionnaire. The methodology and results from this follow-up experiment are discussed in section 3.8.

Another possible explanation is related to another methodological difference between the two experiments. In experiment 2, we added control items which were minimally different from the test questions in order to balance the number of grammatical and ungrammatical items. In the next section, I analyze the subjects' performance on the filler/control items and discuss how the addition of control items may have affected the subjects' performance on the test items.

### 3.7.2 Filler and control items

First, let us compare how well subjects performed on the two auxiliary tasks and how well they performed on the filler/control items to examine whether individual differences, such as accuracy in reading or familiarity with reading, correlate with how they perform on the judgment task in general. Using a Pearson Product Moment correlation, there was no significant correlation between performance on the Fast Reading task and the filler/control items ( $\mathrm{N}=41, r=.0314, p=.845$ ), nor between performance on the Author Recognition task and the filler/control items ( $\mathrm{N}=41$, $r=.00689, p=.966$ ). Individual differences in reading ability do not seem to influence the subjects' overall performance on the filler/control items. The range of scores for the two tasks and for the filler/control items is shown below.
(150) Table 11: Range of scores on Fast Reading task, Author Recognition task, and filler/control items

|  | Lowest score | Highest score | Range |
| :--- | :---: | :---: | :---: |
| Fast Reading task | $47 \%$ | $100 \%$ | 53 |
| Author Recognition task | $8 \%$ | $95 \%$ | 88 |
| Filler/control items | $411 \%$ | $100 \%$ | 59 |

We will now examine the results for the filler/control items by sentence type. In order to determine whether the subjects performed differently on the nine filler/control types, I coded the percentage of correct responses for each of the nine sentence types for each subject. There was a statistically significant main effect of sentence types by one-way repeated-measures ANOVA $(F(8,40)=11.062, p<.001)$. The means for each of the nine sentence types are shown below in (151). In general, the scores were worse on the control items than the filler items.
(151) Table 12: Mean score (\% of correct responses) by filler/control type

| Filler/Control type | Mean score |
| :--- | :---: |
| Grammatical Filler, Wh-island, argument extraction | $98 \%$ |
| Grammatical filler, LBC | $95 \%$ |
| Grammatical filler, Adjunct island | $93 \%$ |
| Ungrammatical filler, Adjunct island | $90 \%$ |
| Grammatical filler, Subject island | $90 \%$ |
| Grammatical Control, Wh-island (argument extraction) | $86 \%$ |
| Ungrammatical filler, Wh-island, adjunct extraction | $80 \%$ |
| Grammatical Control, Wh-island (adjunct extraction) | $79 \%$ |
| Grammatical Control, Adjunct island | $70 \%$ |

In a post-hoc analysis by Tukey's test, a significant difference was found for the following pairs. The scores for the three control items, (152)-(154), as well as the ungrammatical filler for Wh-islands, (155), were significantly worse than some of the other filler/control types.
(152) Grammatical Control (Wh-island, argument extraction) vs.

Grammatical Filler (Wh-island, argument extraction), $p<.049$
(153) Grammatical Control (Wh-island, adjunct extraction) vs.
a. Grammatical Filler (Wh-island, argument extraction), $p<.001$
b. Grammatical Filler (Adjunct island), $p=.009$
c. Grammatical Filler (LBC), $p=.001$
(154) Grammatical Control (Adjunct island) vs.
a. Grammatical Control (Wh-island, argument extraction), $p<.001$
b. Grammatical Filler (Wh-island, argument extraction), p.<001
c. Grammatical Filler (Adjunct island), $p<.001$
d. Ungrammatical Filler (Adjunct island), $p<.001$
e. Grammatical Filler (Subject island), $p<.001$
f. Grammatical Filler (LBC), $p<.001$
(155) Ungrammatical Filler (Wh-island, adjunct extraction) vs.
a. Grammatical Filler (Wh-island, argument extraction), $p<.001$
b. Grammatical Filler (Adjunct island), $p=.022$
c. Grammatical Filler (LBC), $p=.003$

In order to determine whether subjects performed differently on the beginning
vs. the end of the questionnaire, I also conducted a two-way repeated measures
ANOVA for block (beginning vs. end) and sentence type. There was a significant interaction between block and sentence type $(F(8,40)=3.214, p=.002)$. A post-hoc analysis using Tukey's Test revealed a significant difference by block for the grammatical control items for argument extraction out of Wh-islands ( $p=.016$ ),
grammatical control items for adjunct extraction out of Wh-islands ( $p=.007$ ), and grammatical filler items for Adjunct islands ( $p=.037$ ). There was a marginally significant difference by block for ungrammatical filler items for Wh-islands ( $p=00$ ).

### 3.7.3 Response time differences

The response time data, reported in seconds, may reveal whether the subjects responded differently by test sentence type. For example, the subjects may have responded more quickly or slowly on one particular test sentence type. The data may also reveal whether response times were different for the beginning vs. the end of the questionnaire on a particular test sentence type. In particular, we are interested in whether the response times for argument extraction out of Wh-islands changes over the course of the questionnaire. Snyder (1994) suggests that Wh-islands become better after repeated exposure because subjects find a better way to parse the sentence. If this is the case, we might expect response times on these items to be different between the beginning and the end of the questionnaire.

To answer these questions, [ conducted a two-way repeated measures ANOVA using block (first block vs. last block) and test sentence type (argument extraction out of Wh-islands, adjunct extraction out of Wh-islands, Subject islands, Adjunct islands, LBC) as within-subject variables. I considered only the eleven subjects who scored at least $90 \%$ correct on the filler and control items, since this group showed a significant
satiation effect by sign test. There was a main effect of test sentence type, $F(4,10)=4.645, p=004$. A post-hoc analysis by Tukey's test shows that the subjects responded significantly more quickly on the LBC items ( 6.1 seconds) than on Subject islands ( 8.9 seconds.) or argument extraction out of Wh-islands ( 8.8 seconds). The mean response time for each test sentence type is shown below.
(156) Table 13: Mean response times by test sentence type

| Test sentence type | Mean response time (seconds) |
| :--- | :---: |
| argument extraction out of Wh -islands | 8.8 |
| adjunct extraction out of Wh -islands | 8.2 |
| Subject islands | 8.9 |
| Adjunct islands | 7.0 |
| LBC | 6.1 |

There was no significant interaction effect between block and test sentence type, $F(4,10)=1.930, p-124$. The mean response time for argument extraction out of Wh-islands in the beginning vs. the end is 7.3 seconds vs. 10.2 seconds. The subjects are responding more slowly on the later items for argument extraction out of Whislands. The mean response time for each test sentence type in the beginning and the end is shown below.
(157) Table 14: Mean response times (beginning vs. end) by test sentence type

| Test sentence type | Beg. (seconds) | End (seconds) |
| :--- | :---: | :---: |
| argument extraction out of Wh-islands | 4.3 | 10.2 |
| adjunct extraction out of Wh -islands | 6.2 | 10.2 |
| Subject islands | 8.5 | 9.2 |
| Adjunct isiands | 5.4 | 8.4 |
| LBC | 5.8 | 6.3 |

Another question to ask about response times is whether the subjects responded differently on the test sentence vs. the control sentence. Recall that the test and control sentence were minimal pairs and were similar, if not the same, in length. (See (132)-(137) for some examples.) If the subjects find the ungrammatical test sentences more difficult to judge than the grammatical control sentences, the subjects may take longer to respond to the test sentences.

I calculated the mean response times for each subject on each sentence type. I then conducted a one-way repeated measures ANOVA for each of the three pairs of test and control sentence types. There was no significant difference in response time between argument extraction out of Wh-islands and the control type, $F(1$, $10)=.0000133, p=.997$, between adjunct extraction out of Wh-islands and the control type, $F(\mathrm{I}, 10)=1.129, p=313$, or between Adjunct islands and the control type, $F(1$, $10)=1.487, p=251$. The mean response times for each pair of test and control types are given below.
(158) Table 15: Mean response times for test vs. control type

| Sentence type | test type (seconds) | control type (seconds) |
| :--- | :---: | :---: |
| argument extraction out of Wh-islands <br> vs. control | 7.6 | 7.6 |
| adjunct extraction out of Wh-islands <br> vs. control | 7.8 | 7.4 |
| Adjunct islands vs. control | 7.7 | 8.5 |

To summarize the finding from the response time data, the subjects responded significantly more quickly on the LBC items, which are unarguably ungrammatical, than on Subject islands or argument extraction out of Wh-islands, which are perhaps ungrammatical to a lesser degree. There was no significant difference in response times from the beginning to the end of the questionnaire for argument extraction out of Wh-islands. If satiation on this sentence type occurs as a result of finding a better way to parse the sentence, we might expect subjects to respond differently between the beginning and the end of the questionnaire. However, this was not the case. Finally, there was no significant difference between the response times for a test sentence and its control sentence. This suggests that the subjects did not find it more difficult to respond to the test sentences than the control sentences.

### 3.8 Follow-up to Experiment 2

There were some important differences between the satiation results from Experiment I and 2. Although Wh-islands showed a significant satiation effect in

Experiment 1, it was not as strong in Experiment 2. Furthermore, the subjects overall performance on the filler items was considerably worse in Experiment 2. Of the 43 subjects who participated in the experiment, only 11 subjects judged at least $90 \%$ of the filler items correctly.

One possible reason for the difference in results is the way the materials were presented. In Experiment 1, the stimuli were presented in a printed questionnaire, while in Experiment 2, they were presented on a computer. In order to determine whether the computer presentation of the materials might be relevant, I conducted a follow-up experiment in which the materials were presented in a printed format.

### 3.8.1 Methodology

The stimuli used for the follow-up experiment were the same as for Experiment 2. The only difference in the methodology was that instead of presenting the sentences on the computer, they were presented on paper, as in Experiment I.

### 3.8.2 Subjects

Fifteen native English speaking adults participated in the follow-up experiment. I again calculated the mean percentage of correct responses on the filler/control items and excluded any subject whose mean score was more than two
standard deviations away from the mean (68\%). One subject was excluded based on this criterion.

I then excluded any subject who did not have at least $90 \%$ correct on the filler/control items. This excluded 79\% of the subjects (13 subjects) and left only three subjects for analysis. I also calculated the median score (84\%) and considered the seven subject who scored better than the median.

### 3.8.3 Results and discussion

Only three subjects met the $90 \%$ criterion for the filler/control items, so it is difficult to consider these results seriously. Of these three subjects, none of them showed a change either way for the argument extraction out of Wh -islands, Adjunct islands, or LBC violations. With respect to adjunct extraction out of Wh-islands, one subject showed no change, and two showed an increase in the number of no responses. With respect to Subject islands, two subjects showed no change, and one showed an increase in the number of no responses.

Of the seven subjects who scored better than the median, four subjects showed no change for the argument extraction out of Wh-islands, one subject showed an increase in the number of yes responses, and two subjects showed an increase in the number of no responses. For adjunct extraction out of Wh-islands, two subjects showed no change, one subject showed an increase in the number of yes responses,
and four subjects showed an increase in the number of no responses. For Subject islands, four subjects showed no change, one subject showed an increase in the number of yes responses, and two subjects showed an increase in the number of no responses. For Adjunct islands and LBC violations, all seven subjects rejected all the items and showed no change.

The results with respect to satiation are similar to those of Experiment 2. There was still no significant satiation effect for argument extraction out of Whislands, although fewer subjects were considered in the follow-up experiment. There was also no significant satiation effect for Adjunct islands or LBC violations.

The percentage of subjects who did not meet the $90 \%$ criterion for the fillet/control items for Experiment 2 and for the follow-up experiment are similar ( $73 \%$ vs. $79 \%$, respectively). The subjects' performance on the filler and control items were also similar to that of the subjects in Experiment 2. To determine whether the subjects performed differently on the nine filler/control types, I conducted a one-way repeated measures ANOVA on the percentage of correct responses for each type. There was a significant interaction effect between the sentence types $(F(8,14)=6.465$, $p<.001$ ). The mean percentages for each of the nine sentence types from this experiment, as well as from Experiment 2, is shown below in (159).
(159) Table 16: Comparison of mean score (\% of correct responses) by filler/control type

| Filler/Control type | Follow-up | Exp. 2 |
| :--- | :---: | :---: |
| Grammatical Filler, Wh-island, argument extraction | $98 . \%$ | $98 \%$ |
| Grammatical filler, Adjunct island | $95 \%$ | $93 \%$ |
| Grammatical filler, LBC | $94 \%$ | $95 \%$ |
| Grammatical Control, Wh-island (argument extraction) | $91 \%$ | $86 \%$ |
| Grammatical Control, Wh-island (adjunct extraction) | $91 \%$ | $79 \%$ |
| Ungrammatical filler, Adjunct island | $87 \%$ | $90 \%$ |
| Grammatical filler, Subject island | $87 \%$ | $90 \%$ |
| Grammatical Control, Adjunct island | $73 \%$ | $70 \%$ |
| Ungrammatical filler, Wh-island, adjunct extraction | $78 \%$ | $80 \%$ |

A post-hoc analysis by Tukey's test revealed the following significant interactions.
(160) Grammatical Filler (Wh-island, argument extraction) vs.
a. Grammatical Control (Adjunct island), $p<.001$
b. Ungrammatical Filler (Wh-island, adjunct extraction), $p<.001$
(I6I) Grammatical Filler (Adjunct island, argument extraction) vs.
a. Grammatical Control (Adjunct island), $p<.001$
b. Ungrammatical Filler (Wh-island, adjunct extraction), $p=.008$
(162) Grammatical Filler (LBC) vs.
a. Grammatical Control (Adjunct island), $p<.001$
b. Ungrammatical Filler (Wh-island, adjunct extraction), $p=.015$
(163) Grammatical Control (Wh-island, argument extraction) vs.

Grammatical Control (Adjunct island), $p=.004$
(164) Grammatical Control (Wh-island, adjunct extraction) vs.

Grammatical Control (Adjunct island), $p=.004$

A major difference between the follow-up experiment and Experiment 2 is the subjects' performance on the grammatical control items for Wh-islands. The subjects from the follow-up experiment performed better on the grammatical control items for both types of Wh-islands. The subjects in the follow-up experiment performed better on the grammatical control items for both types of Wh-islands. The subjects in both experiments performed poorly on the grammatical control items for Adjunct islands and the ungrammatical fillers for Wh-islands (adjunct extraction). Given these results, the presentation method (computer vs. printed questionnaire) may be relevant to how well the subjects perform on the control items. The subjects who received the printed questionnaire seem to perform better on the control items for both Wh-islands than the subjects who received the computer presentation.

Method of presentation cannot be the primary reason subjects are not showing satiation effects for argument extraction out of Wh-islands since the subjects in both experiments did not show satiation effects. Also, the number of subjects who did not meet the $90 \%$ criterion for the filler/control items were very similar for both experiments. The addition of control items and additional filler items may be a more relevant factor.

### 3.9 Summary

In this chapter, I investigated the phenomenon of syntactic satiation, which involves certain island violations which become increasingly acceptable after repeated exposure. Satiation effects cannot simply be due to performance factors since they are constrained by the syntax. We have evidence from Snyder $(1994,2000)$ and Experiment I that Wh-islands, CNPC, and Subject islands show satiation effects, but that Adjunct islands do not. Furthermore, not all types of ungrammatical sentences show satiation effects. For example, LBC violations do not.

The study of satiation effects is promising since it may help us identify natural classes of island violations. We found a contrast between Subject and Adjunct islands, which is unexpected given that many syntactic theories treat these two islands uniformly. The satiation results suggest that we may want to modify the syntactic theories to treat the two islands differently. There is also evidence that satiation is a sensitive test, since we found satiation effects for Subject islands with ergative predicates, but not with transitive verbs.

I also examined whether subject-related variables might be associated with satiation, and found that they were not. In Experiment 1, there was no evidence of gender, handedness, knowledge of a second language, exposure to linguistics classes, or previous participation in a judgment task being associated with satiation. In Experiment 2, I tested whether general reading ability was associated with performance on the filler/control items. For example, a more careful reader might
perform better on the judgment task. However, there was no association between the subject-related variables and performance on the filler/control items.

For the computer-based experiment (Experiment 2), I explored whether there were any differences in response time among the various sentence types. I found that the subjects responded significantly more quickly on the LBC items, which are unmistakably ungrammatical, than on Subject islands (transitive verb) or Wh-islands (argument extraction), which have a more intermediate status of unacceptability. I also looked at whether subjects responded differently between the beginning and the end of the questionnaire for Wh-islands (argument extraction). Snyder $(1994,2000)$ suggests that satiation may be the result of finding a better way to parse the sentence. If this is the case, we might expect subjects respond differently over the course of the experiment. However, there was no response time difference between beginning vs. end of the questionnaire.

When we used the computer-based presentation method in Experiment 2, the satiation effects for Wh-islands (argument extraction) were not as strong as in Experiment 1. I examined the role of presentation method by administering the same set of stimuli on a printed questionnaire in a follow-up experiment. The results from the follow-up experiment were similar to Experiment 2, suggesting that the presentation method is not a major contributing factor for the difference between Experiment I and 2. The addition of the control sentences and extra filler items in Experiment 2 might be a more relevant factor since more than $70 \%$ of the subjects in

Experiment 2 and the follow-up did not meet the $90 \%$ criterion for the filler/control items.

## Chapter 4

## Concluding remarks

In this thesis. I explored some issues that have been raised concerning the use of acceptability judgments to understand competence. In Chapter 2, I examined one of the criticisms raised about the ability of young children to provide reliable acceptability judgments. Many researchers have avoided eliciting acceptability judgments from young children because of the belief that children have not developed the metalinguistic ability necessary to give acceptability judgments. However, McDaniel and her colleagues have successfully elicited acceptability judgments from children.

From the series of experiments I conducted regarding children's 2Aux questions, we have additional evidence supporting McDaniel that children can be trained to give reliable and consistent acceptability judgments. In addition, we discovered that the elicited product task was not reflecting the child's competence concerning negative questions. Previous studies by Thomton (1993, 1995) and Guasti et al. (1994, 1995) had assumed that children have a non-adult grammar since they produce non-adult negative 2Aux questions. However, the children did not accept 2Aux questions in the grammaticality judgment task (Experiment 3). I argued that the
data from the grammaticality judgment task, rather than the elicited production task, are reflecting the child's competence.

Based on the difference between the children's production and judgment data for 2Aux questions, I proposed that the children think both $n$ 't and not can be used for sentential and constituent negation. The children's derivations of sentential negation questions are the same as the adult derivations. Since the children's derivations are adult-like, their production and judgment of these questions are adult-like. With respect to constituent negation, if the child chooses not, the derivation proceeds in the same way as the adult derivation, and their production and judgment of these questions are adult-like. However, when the children is required to generate a constituent negation question and they incorrectly choose $n$ 't, they are faced with a derivation resulting in a PF crash. In order to avoid this, the child must rely on a "last resort' operation and spell out more than one copy of Infl. If the child is only required to judge a 2 Aux question, she correctly rejects the sentence.

In chapter 3, I turned to a concern that has been raised regarding the stability of acceptability judgments from adults. There has been extensive research on the subject-related and experimental factors that may cause variability in acceptability judgments. When we are faced with unstable acceptability judgments, one of these factors may be responsible. I investigated the phenomenon of syntactic satiation experimentally, following Snyder $(1994,2000)$, to try to determine whether the
increased acceptability of certain island violations is due to a performance factor or is reflecting a change in the competence.

The main finding from the series of experiments is that satiation effects seem to be part of the competence, rather than performance. Although we cannot completely rule out the relevance of performance factors without further research, the initial findings suggest that subject-related are not associated with satiation. There was no evidence of gender, handedness, knowledge of a second language, exposure to linguistics classes, or previous participation in a judgment task being associated with satiation. Nor was general reading ability associated with performance on the filler/control items. However, further investigation of the effects of experimental variables is necessary to determine whether presentation method or the use of control sentences affects satiation.

Another reason that satiation effects cannot simply be due to performance factors is that they are constrained by the syntax. We have evidence from Snyder (1994, 2000) and Experiment I that Wh-islands, CNPC, and Subject islands show satiation effects, but that Adjunct islands do not. Furthermore, not all types of ungrammatical sentences show satiation effects. For example, LBC violations do not. One possible reason for the change in judgment may be that our grammar remains flexible in some respects. For example, syntactic constraints for some islands may weaken, while not for other islands. We might also take satiation effects as evidence that judgments are continuous, rather than dichotomous.

However, there is still the possibility that satiation is due, at least partly, to processing (and performance). Snyder $(1994,2000)$ suggests that subjects may be finding a better way to parse the sentences that show satiation effect. There have been several processing-based accounts of subjacency violations, rather than syntactic accounts. For example, Berwick and Weinberg (1984) propose that the umacceptability of subjacency violations is due to limitations of sentence processing. Data from evoked response potentials (ERP) have also shown that subjacency violations behave differently from other types of unacceptability. Based on this evidence, Kluender and Kutas (1993) argue that subjacency is a processing, rather than a syntactic phenomenon.

Even if we do not yet understand what syntactic satiation is, it can still be used as a tool to identify natural classes of grammatical phenomena. In chapter 3, we discovered that Subject and Adjunct islands behave differently with respect to satiation, and hence may not form a natural class as many syntactic theories have been assuming. We also found that satiation is a sensitive tool since Subject islands showed satiation with ergative predicates but not with transitive verbs.

In this thesis, I have explored the use of acceptability judgments to access competence and presented additional ways to use them. First, I have provided additional support for using acceptability judgments with children, a population that many researchers have avoided. With proper training, many children can be trained to give reliable and consistent acceptability judgments. By expanding the methodologies
we can use with children, we increase the possibility of accessing the children's competence. Second, I have shown that the instability of judgments for certain island violations does not indicate that we must abandon the use of acceptability judgments. To the contrary, syntactic satiation is a promising tool that can be used to investigate competence.

## Appendix 1

## Results from children's acceptability judgments

Table 17: Children's percentage adult-like responses (Experiment 3)

| Child | Neg <br> 2Aux+m | Neg <br> 2Aux-m | Pos <br> 2Aux+m | Pos <br> 2Aux-m | Decl. | Control | Filler |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Judy | 50 | 75 | 75 | 100 | 100 | 100 | 100 |
| Marilyn | 50 | 100 | 100 | 50 | 100 | 100 | 100 |
| Rita | 50 | 75 | 100 | 100 | 100 | 100 | 100 |
| Kim | 0 | 50 | 50 | 75 | 88 | 100 | 100 |
| Phillip | 50 | 75 | 75 | 100 | 100 | 100 | 100 |
| Norbert | 25 | 75 | 25 | 75 | 100 | 100 | 67 |
| Anne | 75 | 100 | 100 | 75 | 100 | 92 | 100 |
| David | 75 | 100 | 100 | 100 | 100 | 92 | 100 |
| Kendra | 75 | 100 | 100 | 100 | 100 | 83 | 91 |
| Linda | 75 | 50 | 75 | 25 | 100 | 83 | 88 |
| Scott | 100 | 0 | 0 | 100 | 100 | 92 | 67 |
| Jennifer | 75 | 100 | 100 | 100 | 100 | 75 | 71 |
| Kristen | 75 | 75 | 25 | 50 | 100 | 100 | 100 |
| Maureen | 100 | 75 | 67 | 75 | 100 | 78 | 38 |
| Andrew | 50 | 100 | 100 | 100 | 100 | 100 | 100 |

Table 18: Children's percentage adult-like responses (follow-up to Experiment 3)

| Child | not | decl. | control |
| :--- | :--- | :--- | :--- |
| Judy | 100 | 100 | 100 |
| Marilyn | 100 | 100 | 100 |
| Rita | 100 | 83 | 100 |
| Kim | 100 | 67 | 100 |
| Phillip | 100 | 100 | 92 |
| Norbert | 100 | 83 | 83 |
| Anne | 100 | 100 | 92 |
| David | 100 | 100 | 100 |
| Kendra | 100 | 100 | 100 |
| Linda | 100 | 100 | 92 |
| Scott | 100 | 100 | 100 |
| Jennifer | 100 | 100 | 100 |
| Kristen | - | - | - |
| Maureen | - | - | - |
| Andrew | - | - | - |

## Appendix 2

## Instructions for Syntactic Satiation Experiment 1

In this study, you are asked to provide a judgment of grammatical acceptability for each of (70) sentences. Every test sentence is in the form of a question. On each page of the experiment, you will see a declarative sentence (the "context" sentence), followed by a question based on the declarative sentence (the "test" sentence). We would like you to judge whether each test sentence is a "grammatically possible" sentence of English, given the meaning that fits the preceding context sentence.

By "grammatically possible" we mean possible for you as a native speaker of English. We are not concerned about whether the test sentence would be acceptable to a writing teacher: points of style and clarity are not the issue, and we are not concerned about who versus whom, or ending a sentence in a preposition. Instead, we are interested in whether the test sentence could have the intended meaning and still be accepted as "English", in your opinion.

Some examples follow. We would expect most English-speakers to answer "yes" (grammatically possible) on the first example, and "no" (grammatically not possible) on the second example. Many of the test sentences in this study are likely to fall between these two extremes. Even so, we would like you to give a yes/no judgment for each sentence.

## Example I

[Context: Chris peeled an apple.]
Test sentence: "What did Chris peel?"
Judgment:
(Y/N)

## Example II

[Context: Susi wants very much for Jonathan to meet Diana.]
Test sentence: "Who does Susi want very much whether for who to meet Diana?"
Judgment: (Y/N)

Please do not tum back to previous items. This is not a memory test. You are not required to remember the judgments that you gave for previous items. Many of the items will be quite similar to one another, to the best of your ability, you should simply ignore this fact and provide an "independent" judgment on each individual item.

Please work through the questionnaire as quickly as possible. Try not to dwell on a sentence; instead try to go with your initial reaction for each sentence. If you have any questions, especially questions about the intended meaning of a given sentence, please ask.

## References

Allen, Joe. 1995. 'Tense' doubling in early grammars. In Proceedings of the 19th Annual Boston University Conference on Language Development, eds. Dawn MacLaughlin and Susan McEwen, 48-59. Somerville, MA: Cascadilla Press.
Bach, Emmon, and George M. Horn. 1976. Remarks on "Conditions on Transformations." Linguistic Inquiry 7.2, 265-361.
Berwick, Robert C., and Amy S. Weinberg. 1984. The grammatical basis of linguistic performance: Language use and acquisition Cambridge, Mass.: MIT Press.
Bever, Thomas G. 1970. The cognitive basis for linguistic structures. In Cognition and the development of language, ed. John R. Hayes, 279-362. New York: John Wiley and Sons.
Bever, Thomas G. 1974. The ascent of the specious; or, There's a lot we don't know about mirrors. In Explaining linguistic phenomena, ed. David Cohen. Washington, DC: Hemisphere.
Bever, Thomas G., Caroline Carrithers, and David J. Townsend. 1987. A tale of two brains; or, The sinistral quasimodularity of language. In Cognitive Science Society, Program of the ninth anmual conference of the Cognitive Science Society, 764-73. Hillsdale, New Jersey: Erlbaum.
Bever, Thomas G., John Carroll, and Richard Hurtig. 1976. Analogy or ungrammatical sequences that are utterable and comprehensible are the origins of new grammars in language acquisition and linguistic evolution. In An integrated theory of linguistic ability, eds. Thomas G. Bever, Jerrold Katz, and D. Terence Langendoen, 149-182. New York: T. Y. Crowell Press.
Bialystok, Ellen. 1979. Explicit and implicit judgments of L2 grammaticality. Language Learning 29.1, 81-103.
Bialystok, Ellen, and Ellen Bouchard Ryan. 1985. A metacognitive framework for the development of first and second language skills. In Metacognition, cognition, and human performance, Volume 1: Theoretical perspectives, eds. Donna Lynn Forrest-Pressley, G. E. MacKinnon, and T. Gary Waller, 207-52. Orlando: Academic Press.
Birdsong, David. 1989. Metalinguistic performance and interlinguistic competence. New York: Springer-Verlag.
Bley-Vroman, Robert W., Sascha W. Felix, and Georgette L. Ioup. 1988. The accessibility of universal grammar in adult language learning. Second Language Research 4.1, 1-32.
Bobaljik, Jonathan. 1995. Morphosyntax: The syntax of verbal inflection. Doctoral dissertation, MIT, Cambridge, Mass.

Bobaljik, Jonathan. 1998. The autonomy of syntax and the typology of movement. Paper read at GLOW.
Boeckx, Cedric. 1999. Syntactic satiation really is syntactic. Talk presented at Psychology-Linguistics Lunch Seminar, University of Connecticut, November 22, 1999.

Boškovic, Zeljko. 1995. Superiority effects with multiple wh-fronting in SerboCroatian. Ms., University of Connecticut, Storrs.
Boškovic, Zeljko. 1999. The syntax-phonology interface: A view from the syntay. Book ms. University of Connecticut, Storrs.
Bradac, James J., Larry W. Martin, Norman D. Elliott, and Charles H. Tardy. 1980. On the neglected side of linguistic science: Multivariate studies of sentence judgment. Linguistics 18, 967-95.
Carden, Guy. 1970. Discussion of Heringer 1970. In Chicago Linguistic Society, Papers from the sixth regional meeting, 296. Chicago: Chicago Linguistic Society.
Carroll, John M. 1979. Complex compounds: Phrasal embedding in lexical structures. Linguistics 17, 863-77.
Carroll, John M., Thomas G. Bever, and Chava R. Pollack. 1981. The non-uniqueness of linguistic intuitions. Language 57, 368-383.
Chaudron, Craig. 1983. Research on metalinguistic judgments: A review of theory, methods and results. Language Learning 33.3: 343-77.
Chomsky, Noam. 1957. Syntactic structures. The Hague: Mouton.
Chomsky, Noam. 1964. Current Issues in Linguistic Theory. The Hague: Mouton.
Chomsky, Noam. 1965. Aspects of the theory of syntax. Cambridge, Mass.: MIT Press.
Chomsky, Noam. 1969. Language and philosophy. In Language and Philosophy: A symposium, ed. Sidney Hook, 51-94. New York: New York University Press.
Chomsky, Noam. 1973. Conditions on Transformations. In A Festschrift for Morris Halle, eds. Stephen R. Anderson and Paul Kiparsky. New York: Holt, Rinehart and Winston.
Chomsky, Noam. 1977. On Wh-movement. In Formal Syntax, eds. Peter Culicover, Thomas Wasow, and Adrian Akmajian. New York: Academic Press.
Chomsky, Noam. 1981. Lectures on Government and Binding. Dordrecht: Foris.
Chomsky, Noam. 1985 [1955-56]. The logical structure of linguistic theory. New York: Plenum, 1975. Reprint, Chicago: University of Chicago Press.
Chomsky, Noam. 1986. Barriers. Cambridge, Mass.: MIT Press.
Chomsky, Noam. 1993. A minimalist program for linguistic theory. In The view from Building 20, eds. Ken Hale and Samuel J. Keyser, 1-52. Cambridge, MA: MIT Press.
Chomsky, Noam. 1994. Bare Phrase Structure. MIT Occasional Papers in Linguistics 5. Cambridge, Mass.: MITWPL. [Published in Government and Binding Theory and the Minimalist Program, ed. Gert Webelhuth, 383-439. Oxford: Blackwell. (1995).I

Chomsky, Noam. 1995. The Minimalist Program. Cambridge, Mass.: MIT Press.
Chomsky, Noam. 1998. Some observations on economy in generative grammar. In Is the best good enough? Optimality and competition in syntax, eds. Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky, II5-127. Cambridge, Mass.: MIT Press.
Chomsky, Noam, and Howard Lasnik. 1977. Filters and control. Linguistic Inquiry II, 1-46.
Chomsky. Noam, and Howard Lasnik. 1993. The theory of principles and parameters. In Syntax: An international handbook of contemporary research, eds. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann. Berlin: de Gruyter.
Cinque, Guglielmo. 1999. Adverbs and functional heads: A cross-linguistic perspective. New York: Oxford University Press.
Clark, Herbert H., and Susan E. Haviland. 1974. Psychological processes as linguistic explanation. In Explaining linguistic phenomena, ed. David Cohen, 91-124. Washington, D. C.: Hemisphere.
Corver, Norbert. 1992. On deriving certain left branch extraction asymmetries: A case study in parametric syntax. In Proceeding of Northeast Linguistic Society (NELS) 22, 67-84. Amherst, Mass.: GLSA.
Cowart, Wayne. 1989. Notes on the biology of syntactic processing. Journal of Psycholinguistic Research 18.1, 89-103.
Cowart, Wayne. 1994. Anchoring and grammar effects in judgments of sentence acceptability. Perceptual and Motor Skills 79, 1171-1182.
Cowart, Wayne. 1997. Experimental syntax: Applying objective metionds to sentence judgments. Thousand Oaks, Califomia: Sage Publications.
Crain, Stephen, and Mark Steedman. 1985. On not being led up the garden path: The use of context by the psychological syntax processor. In Natural language processing: Psychological, computational, and theoretical perspectives, eds. David D. Dowty, Lauri Kartunen, and Arnold M. Zwicky, 320-58. Cambridge: Cambridge University Press.
Crain, Stephen, and Rosalind Thornton. 1998. Investigations in universal grammar: $A$ guide to experiments on the acquisition of syntax and semantics. Cambridge, Mass.: MIT Press.
Derwing, Bruce L. 1979. Against autonomous linguistics. In Evidence and argumentation in linguistics, ed. Thomas A. Perry, 163-89. Berlin: de Gruyter.
de Villiers, Jill G. and Peter A. de Villiers. 1974. Competence and performance in child language: Are children really competent to judge? Journal of Child Language 1, 11-22.

Elliot, Dale, Stanley Legum, and Sandra Annear Thompson. 1969. Syntactic variation as linguistic data. In Papers from the fifth regional meeting of the Chicago Linguistic Society, eds. Robert I. Binnick, Alice Davison, Georgia M. Green, and Jerry L. Morgan, 52-59. Chicago: Chicago Linguistic Society.
Emonds, Joseph. 1976. A transformational approach to English syntax. New York: Academic Press.
Ernst, Thomas. 1992. The phrase structure of English negation. The Linguistic Review 9. 109-144.

Fay, David. 1978. Transformations as mental operations: A reply to Kuczaj. Journal of Child Language 5.1, 143-149.
Frazier, Lyn. 1985. Syntactic complexity. In Natural language processing: Psychological, computational, and theoretical perspectives, eds. David D. Dowty, Lauri Karttunen and Amold M. Zwicky, 129-89. Cambridge: Cambridge University Press.
Gerken, LouAnn, and Thomas G. Bever. 1986. Linguistic intuitions are the result of interactions between perceptual processes and linguistic universals. Cognitive Science 10, 457-476.
Gleitman, Henry, and Lila R. Gleitman. 1979. Language use and language judgment. In Individual differences in language ability and language behavior, eds. Charles J. Fillmore, Daniel Kempler and William S-Y. Wang, 103-26. New York: Academic Press.
Gombert, Jean Émile. 1992. Metalinguistic development. Chicago: University of Chicago Press.
Goodluck, Helen and Lawrence Solan. 1979. A reevaluation of the basic operations hypothesis. Cognition 7, 85-91.
Greenbaum, Sidney. 1973. Informant elicitation of data on syntactic variation. Lingua 31,201-12.
Greenbaum, Sidney. 1976. Contextual influence on acceptability judgments. Linguistics 187, 5-11.
Greenbaum, Sidney. 1977. Judgments of syntactic acceptability and frequency. Studia Linguistica 31.2, 83-105.
Greenbaum, Sidney. 1988. Good English and the grammarian. London: Longman.
Greenbaum, Sidney, and Randolph Quirk. 1970. Elicitation experiments in English: Linguistic studies in use and attitude. Coral Gables, Florida: University of Miami Press.
Grimshaw, Jane. 1992. Argument structure. Cambridge, Mass.: MIT Press.
Groat, Erich, and John O'Neil. 1996. Spell-out at the LF interface. In Wemer Abraham, Samuel David Epstein, Hoskuldur Thrainsson, and C. Jan-Wouter Zwart (eds.), Minimal ideas, 113-39. Amsterdam: John Benjamins.
Guasti, Maria Teresa. 1994. Acquisition of Italian interrogatives. Ms., University of Geneva.

Guasti, Maria Teresa, Rosalind Thornton, and Kenneth Wexler. 1994. Negation in children's questions: The case of English. Ms., University of Geneva.
Guasti, Maria Teresa, Rosalind Thornton, and Kenneth Wexler. 1995. Negation in children's questions: The case of English. In Proceedings of the 19th Annual Boston University Conference on Language Development, eds. Dawn MacLaughlin and Susan McEwen, 228-239. Somerville, Mass.: Cascadilla Press.
Haegeman, Lilliane and Raffaella Zanuttini. 1991. Negative heads and the NegCriterion. The Linguistic Review 8, 233-51.
Hakes, David T. 1980. The development of metalinguistic abilities in children. New York: Springer-Verlag.
Halbert, Anne. 1997. Children's use of prosodic cuing in ambiguity resolution. Masters thesis, University of Connecticut, Storrs.
Halle, Morris and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Ken Hale and Samuel J. Keyser (eds.), The view from Building 20. Cambridge, MA: MIT Press.
Heeschen, Volker. 1978. The metalinguistic vocabulary of a speech community in the highlands of Irian Jaya (West New Guinea). In The child's conception of language, eds. A. Sinclair, R. J. Jarvella and W. J. M. Levelt, 155-87. Berlin: Springer-Verlag.
Higginbotham, James. 1985. On semantics. Linguistic Inquiry 16, 547-594.
Hill, Archibald. 1961. Grammaticality. Word 17.1, 1-10.
Hiramatsu, Kazuko. 1998. Natural classes of subjacency violations: Evidence from syntactic satiation. General examination paper, University of Connecticut, Storrs.
Hiramatsu, Kazuko. 1999a. Subject and adjunct island asymmetries: Evidence from syntactic satiation. In Proceedings of the 18th West Coast Conference on Formal Linguistics, eds. Sonya Bird, Andrew Camie, Jason D. Haugen, and Peter Norquest,183-192. Somerville, Mass.: Cascadilla Press.
Hiramatsu, Kazuko. 1999b. What does Move doesn't delete? In Cranberry Linguistics, UConn Working Papers in Linguistics IO, eds. David Braze, Kazuko Hiramatsu, Yutaka Kudo. University of Connecticut, Storrs.
Hiramatsu, Kazuko, and Diane Lillo-Martin. 1998. Children who judge what they produce ungrammatical. In Proceedings of the 22nd Annual Boston University Conference on Language Development, eds. Annabel Greenhill, Mary Hughes, Heather Littlefield, and Hugh Walsh, 337-347. Somerville, Mass.: Cascadilla Press.
Householder, Fred W., Jr. 1965. On some recent claims in phonological theory. Journal of Linguistics 1.1, 13-34.
Huang, C.-T. James. 1982. Logical relations in Chinese and the theory of grammar. Doctoral dissertation, MIT, Cambridge, Mass.
Hurford, James. 1975. A child and the English question formation rule. Journal of Child Language 2, 299-301.

Karlsen, B. and E. F. Gardner. 1984. Stanford Diagnostic Reading Test: Stanford Measurement Series. San Anotonio, Texas: Harcourt Brace Educational Measurement.
Kayne, Richard. 1989. Notes on English agreement. CIEFL Bulletin. Hyderabad. 1, 41-67.
Kayne, Richard. 1993. The antisymmetry of syntax. Ms., City University of New York.
Kuczaj, Stan. 1976. Arguments against Hurford's "Aux Copying Rule.' Journal of Child Language 3, 423-427.
Kluender, Robert and Marta Kutas. 1993. Subjacency as a processing phenomenon. Language and Cognitive Processes 8.4, 573-633.
Labov, William. 1975. What is a linguistic fact? Lisse: Peter de Ridder. [Also published as Empirical foundations of linguistic theory. In The scope of American linguistics: Papers of the first Golden Anniversary Symposium of the Linguistic Society of America, held at the University of Massachusetts, Amherst on July 24 and 25, 1974, ed. Robert P. Austerlitz, 77-133. Lisse: Peter de Ridder.]
Laka, Itziar. Negation in syntax: On the nature of functional categories and projections. Doctoral dissertation, MIT.
Langendoen, D. Terrence, and T. Bever. 1973. Can a not unhappy person be called a not sad one? In A festschrift for Morris Halle, eds. Stephen R. Anderson and Paul Kiparsky, 392-409. New York: Holt, Rinehart \& Winston. [Reprinted in Bever, Thomas G., Jerrold Katz and D. Terence Langendoen, eds. 1976. An integrated theory of linguistic ability. 239-60. New York: Crowell.]
Lasnik, H. 1995. Verbal morphology: Syntactic Structures meets the Minimalist Program. In Paula Kempchinsky and Hector Campos (eds.), Evolution and revolution in linguistic theory: Essays in honor of Carlos Otero. Washington, D.C.: Georgetown University Press.

Lasnik, Howard. 1999. On the locality of movement. In Functionalism and Formalism in Linguistics (Studies in Language Companion Series, volume 41), eds. Mike Darnell, Edith A. Moravscik, Frederick Newmeyer, Michael Noonan, and Kathleen Wheatley, 33-54. Amsterdam: John Benjamins.
Lasnik, Howard, and Mamoru Saito. 1992. Move $\alpha$ Cambridge, Mass.: MIT Press.
Lasnik, Howard, and Juan Uriagereka. 1988. A Course in GB Syntax. Cambridge, Mass.: MIT Press.
Levelt, W. J. M. 1974. Formal grammar in linguistics and psycholinguistics. The Hague: Mouton.
Levelt, W. J. M., J. A. W. M. van Gent, A. F. J. Haans, and A. J. A. Meijers. 1977. Grammaticality, paraphrase, and imagery. In Acceptability in language, ed. Sidney Greenbaum, 87-10I. The Hague: Mouton.

Maratsos, Michael. 1984. Some current issues in the study of acquisition of grammar. In Handbook of child psychology, VoL. 3: Cognitive development, eds. John H. Flavell and Ellen M. Markman0. 4th edition. New York: Wiley.
Maratsos, Michael, and Stan Kuczaj. 1978. Against the transformationalist account: A simpler analysis of auxiliary overmarkings. Journal of Child Language 5, 337345.

Masny, Diana, and Alison d'Angeljan. 1985. Language, cognition, and second language grammaticality judgments. Journal of Psycholinguistic Research 14.2: 175-97.
May, Robert. 1985. Logical Form: Its structure and derivation. Cambridge, Mass.: MIT Press.
Mayer, Judith Winzemer, Anne Erreich, and Virginia Valian. 1978. Transformations, basic operations and language acquisition. Cognition 6, 1-13.
McDaniel, Dana and Helen Smith Cairns. 1990. The child as informant: Eliciting intuitions from young children. Journal of Psycholinguistic Research 19, 331-344.
McDaniel, Dana, and Helen Smith Caims. 1996. Eliciting judgments of grammaticality and reference. In Methods for assessing children's syntox, eds. Dana McDaniel et al., 233-254. Cambridge, Mass.: MIT Press.
McDaniel, Dana, Bonnie Chiu, and Tom Maxfield. 1995. Parameters for whmovement types. Natural Language \& Linguistic Theory, 13, 709-753.
McDaniel, Dana, Cecile McKee, and Helen Smith Cairns. 1996. Methods for assessing children's syntax. Cambridge, Mass.: MIT Press.
McWhinney, Brian, and Catherine Snow. 1985. The child language data exchange system. Journal of Child Language 12, 271-296.
Mistler-Lachman, Janet L. 1972. Levels of comprehension in processing of normal and ambiguous sentences. Journal of Verbal Learning and Verbal Behavior 11, 614-23.
Nagata, Hiroshi. 1987a. Change in the modulus of judgmental scale: An inadequate explanation for the repetition effect in judgments of grammaticality. Perceptual and Motor Skills 65.3, 907-10.
Nagata, Hiroshi. 1987b. Long-term effect of repetition on judgments of grammaticality. Perceptual and Motor Skills 65.1, 295-99.
Nagata, Hiroshi. 1988. The relativity of linguistic intuition: The effect of repetition on grammaticality judgments. Journal of Psycholinguistic Research 17.1, 1-17.
Nagata, Hiroshi. 1989a. Effect of repetition on grammaticality judgments under objective and subjective self-awareness conditions. Journal of Psycholinguistic Research 18.3, 255-69.
Nagata, Hiroshi. 1989b. Judgments of sentence grammaticality and field dependence of subjects. Perceptual and Motor Skills 69.3, 739-47.
Nagata, Hiroshi. 1989c. Judgments of sentence grammaticality with differentiation and enrichment strategies. Perceptual and Motor Skills 68.2, 463-69.

Nagata, Hiroshi. 1989d. Repetition effect in judgments of grammaticality of sentences: Examination with ungrammatical sentences. Perceptual and Motor Skills 68.1, 275-82.
Nagata, Hiroshi. 1990a. On-line judgments of grammaticality of sentences. Perceptual and Motor Skills 70.3, 987-94.
Nagata, Hiroshi. 1990b. Speaker's sensitivity to rule violations in sentences. Psychologia 33.3, 179-184.
Nagata. Hiroshi. 1992. Anchoring effects in judging grammaticality of sentences. Perceptual and Motor Skills 75, 159-64.
Newmeyer, Frederick J. 1983. Grammatical theory. Chicago: University of Chicago Press.
Nunes, Jairo. 1995. The copy theory of movement and linearization of chains in the Minimalist Program. Doctoral dissertation, University of Maryland, College Park.
Ouhalla, Jamal. [988. Syntax of head movement: A study of Berber. Doctoral dissertation, University College, London.
Ouhalla, Jamal. 1990. Sentential negation, Relativized Minimality and the aspectual status of auxiliaries. Linguistic Review 7, 183-231.
Ouhalla, Jamal. 1991. Functional categories and parametric variation. London: Routeledge.
Ouhalla, Jamal. 1993. Negation, Focus, and Tense. Rivista di Linguistica 5: 275-300.
Paikeday, Thomas M. 1985. The native speaker is dead! Toronto: Paikeday Publishing.
Pesetsky, David. 1987. Wh-in-situ: Movement and unselective binding. In The representation of (in)definiteness, eds. Eric J. Reuland and Alice G. B. ter Meulen, 98-129. Cambridge, Mass.: MIT Press.
Pesetsky, David. 1997. Some optimality principles of sentence pronunciation. In Is the best good enough, eds. Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky. Cambridge, MA: MIT Press and MITWPL.
Pritchett, Bradley L. 1992. Parsing with grammar: islands, heads, and garden paths. In Island constraints: Theory, acquisition and processing, eds. Helen Goodluck and Michael Rochemont. Dordrecht: Kluwer.
Richards, Norvin. 1997. What moves where when in which language. Doctoral dissertation, MIT.
Rizzi, Luigi. 1990. Relativized Minimality. Cambridge, Mass.: MIT Press.
Rizzi, Luigi. 1991. Residual verb second and the Wh-Criterion. Technical reports in formal and computational linguistics. University of Geneva.
Roberts, Ian G. 1993. Verbs and Diachronic Syntax. Dordrecht: Kluwer Academic Publishers.
Roberts, Ian G. 1994. Two types of head movement in Romance. In Verb movement, eds. David Lightfoot and Norbert Hornstein. Cambridge: Cambridge University Press.

Roberts, Ian G. 1997. Restructuring, head movement, and locality. Linguistic Inquiry 28, 423-460.
Rose, Robert G. 1973. Linguist and nonlinguist agreement concerning surface structures. The Journal of General Psychology 89.2, 325-26.
Ross, John Robert. 1967. Constraints on variables in syntax. Unpublished doctoral dissertation, MIT, Cambridge, Mass.
Ross, John Robert. 1979. Where's English? In Individual differences in language ability and language behavior. eds. Charles J. Fillmore, Daniel Kemper, and William S.-Y. Wang, 127-63. New York: Academic Press.
Rudin, Catherine. 1988. On multiple questions and multiple Wh fronting. Natural Language \& Linguistic Theory 6, 455-501.
Ryan, Ellen Bouchard, and George W. Ledger. 1984. Learning to attend to sentence structure: Links between metalinguistic development and reading. In Language awareness and learning to read, eds. John Downing and Renate Valtin, 149-71. New York: Springer-Veriag.
Sampson, Geoffrey. 1975. The form of language. London: Weidenfeld and Nicolson.
Sarma, Jaya. 1991. The acquisition of Wh-questions in English. Doctoral dissertation, University of Connecticut, Storrs.
Schlisselberg, G. 1988. Development of selected conservation skills and the ability to judge sentence well-formedness in young children. Doctoral dissertation, City University of New York.
Scholes, Robert J., and Brenda J. Willis. 1987. Literacy and language. Journal of Literary Semantics 16.1, 3-11.
Schütze, Carson T. 1996. The empirical basis of linguistics: Grammaticality judgments and linguistic methodology. Chicago: University of Chicago Press.
Scribner, Sylvia, and Michael Cole. 1981. The psychology of literacy. Cambridge, Mass.: Harvard University Press.
Smith-Lock, Karen M., and Hyla Rubin. 1993. Phonological and morphological analysis skills in young children. Journal of Child Language 20, 437-454.
Snow, Catherine. 1975. Linguists as behavioral scientists: Towards a methodology for testing linguistic intuitions. In Linguistics in the Netherlands 1972-1973, ed. A. Kraak, 27I-75. Assen: Van Gorcum.
Snow, Catherine, and Guus Meijer. 1977. On the secondary nature of syntactic intuitions. In Acceptability in Language, ed. Sidney Greenbaum, 163-77. The Hague: Mouton.
Snyder, William. 1994. A psycholinguistic investigation of weak crossover, scope, and syntactic satiation effects: Implications for distinguishing competence from performance. Presentation at CUNY Human Sentence Processing Conference poster Session, March 17, 1994.
Snyder, William. 2000. An experimental investigation of syntactic satiation effects. Linguistic Inquiry 31.3, 575-82.

Spencer, N. J. 1973. Differences between linguists and nonlinguists in intuitions of grammaticality-acceptability. Journal of Psycholinguistic Research 2.2, 83-98.
Stanovich, Keith E. and Richard F. West. 1989. Exposure to print and orthographic processing. Reading Research Quarterly 24.4, 402-433.
Stowell, Timothy. 1981. Origins of Phrase Structure. Doctoral dissertation, MIT, Cambridge, Mass.
Stromswold, Karin. 1986. Syntactic satiation. Ms., MIT, Cambridge, Mass.
Stromswold, Karin. 1988. Linguistic representations of children's Wh-questions. Papers and reports on child language development 27. Stanford University.
Stromswold, Karin. 1990. Learnability and the acquisition of auxiliaries. Doctoral dissertation, MIT, Cambridge, Mass.
Takahashi, Daiko. 1994. Minimality of movement. Doctoral dissertation, University of Connecticut, Storrs.
Thornton, Rosalind. 1990. Adventures in long-distance moving: The acquisition of complex wh-questions. Doctoral dissertation, University of Connecticut, Storrs.
Thornton, Rosalind. 1993. Children who don't raise the negative. Paper presented at the 67th Annual Meeting of the Linguistic Society of America, Los Angeles, California.
Thornton, Rosalind. 1995. Children's negative questions: A production/comprehension asymmetry. In Proceedings of the Eleventh Eastern States Conference on Linguistics (1994), eds. Janet M. Fuller, Ho Han and David Parkinson, 306-317. Ithaca, NY: DMLL Publications.
Valian, Virginia. 1982. Psycholinguistic experiment and linguistic intuition. In Language, mind, and brain, eds. Thomas W. Simon and Robert J. Scholes, 179-88. Hillsdale, New Jersey: Erlbaum.
Van Kleeck, Anne. 1982. The emergence of linguistic awareness: A cognitive framework. Merrill-Palmer Quarterly 28.2: 237-65.
Vogel, Irene and Eric Raimy. 2000. Compound \& phrasal stress: A case of late acquisition. Paper presented at 74th Annual Meeting of the Linguistic Society of America, Chicago, IL, January 6-9, 2000.
Warner, John, and Arnold L. Glass. 1987. Context and distance-to-disambiguation effects in ambiguity resolution: Evidence from grammaticality judgments of garden path sentences. Journal of Memory and Language 26, 714-38.
Zanuttini, Raffaella. 1990. Syntactic properties of sentential negation: A comparative study of Romance languages. Doctoral dissertation, University of Pennsylvania.
Zanuttini, Raffaella. 1997. Negation and clausal structure. New York: Oxford University Press.


[^0]:    'We do not investigate IAux questions in this study.

[^1]:    2 In both Thornton (1993) and Guasti et al. (1995), there were a very small number of questions with mixed auxiliary verbs, such as do and a modal.
    (i) What do you can't eat?
    (ii) What did Snoopy couldn't do?
    (Guasti et al. 1994)
    (iii) What can Cookie Monster doesn't want to share?
    (Guasti et al. 1994)
    I will ignore these cases in this study.

[^2]:    ${ }^{5}$ See Ouhalla (1990).

    * Although Thormton treats questions with nor as non-adult forms, I will be treating them as adult. See sections 2.3 .3 and 2.5.4 for results from adult control subjects for Experiment I (elicited production) and Experiment 2 (grammaticality judgment).

[^3]:    s There were two types of negative questions that were adult in form: how-come questions and Whquestions with subject extraction. Presumably, how-come questions and Wh-subject questions were adult because neither involves subject-auxiliary inversion.
    ${ }^{6}$ The Neg \& Aux doubling constructions were not very common, and Guasti et al. treat it as a transitional construction.
    ${ }^{T}$ The ' $+\mathrm{m}^{\prime}$ in ' 2 Aux +m ' refers to 'matching' auxiliary verbs. The ' $-\mathrm{m}^{\prime}$ in ' 2 Aux. $m$ ' refers to 'nonmatching' auxciliary verbs.

[^4]:    ${ }^{8}$ Agreement inflections are not obvious for modals or for past tense in English, and therefore are excluded from the paradigm.
    ${ }^{9}$ The patterms that were unattested are marked with a "\#" sign.
    ${ }^{10}$ It is unclear whether this utterance is a general example of the question type umder consideration or one produced by a child during an experiment.

[^5]:    ${ }^{\text {II }}$ Guasti et al. do not provide the data for the Neg-Aux doubling constructions (2Aux/Neg), but they claim that the same tense patterns hold for these cases.

[^6]:    12 Guasti et al. assume that the negative affix raises to Infl, but do not motivate this movement Another possibility is that the negative affix remains in Neg, and undergoes PF merger with the stranded Infl, spelled out as 'did' at PF. However, this possibility cannot account for negative questions since there is I to $C$ movement, and the subject NP intervenes between Infl and the negative affix.
    

[^7]:    ${ }^{13}$ The idea of spelling out fewer features is reminiscent of resumptive pronoums. I thank Howard Lasnik for pointing this out.

[^8]:    ${ }^{14}$ Some children also produced Wh-subject questions instead of Wh-object questions. For example, given the lead-in, "We know the lady rubbed someone. Ask the puppet who," one child asked "Who rubbed the dog?" Sarma (1991) observed the same phenomenon in two of her subjects. This is an interesting fact in light of Stromswold's (1988) observation that children produce more object questions than subject questions in spontaneous speech. It is curious that they have trouble producing such questions in the experimental sessions.
    ${ }^{15}$ The only exception is with the positive yes/no questions with five children. Since these children had already produced several positive yes/no questions in response to the negative yesino question prompts,

[^9]:    they received fewer than two positive yes/no question prompts. Also one child ( $\mathbf{3} ; \mathrm{II}$ ) did not receive the complete set of items for Wh-questions.
    ${ }_{16}$ One child who did not attend the Child Development Laboratories was tested at home.

[^10]:    ${ }^{17}$ Two of the adult subjects were presented with the full set of stimuli with the extra conversational items that some of the children received. The other two subjects were presented with the stories and the conversational items for just the yes/no questions.
    ${ }^{18}$ It is unclear on what basis Guasti et al (1995) and Thomton (1995) decided that the questions with not were dispreferred by adults. Neither smdy reports resuits from adult subjects.
    19 There were 16 questions excluded from this total. There were 6 yes/no questions that were both positive and negative, and 10 grammatical Wh-subject questions.
    (i) Did you or didn't you sleep last night? (Ame 4;1)
    (ii) Who brushed the sheep?
    (Maureen 3;11)

[^11]:    ${ }^{20}$ Of the four positive questions with non-inversion, three were with the irregular verb hide, and one was with the regular verb rub.

[^12]:    ${ }^{21}$ The two children who produced 2Aux. . $_{\text {a }}$ questions also produced questions with incorrect agreement.

[^13]:    2 Two adult subjects produced a question with emphatic stress in response to a prompt without emphatic stress, but it was obvious that the subject was anticipating that the puppet's incorrect response. Some of the children did this as well.
    ${ }^{2}$ Some of the children placed emphatic stress on the wrong lexical tem, such as the main verb rather than that auxiliary verb, as in (i), or added extra lexical items to mark emphasis, as in (ii). I considered questions like these to involve emphatic stress.
    (i) Where did you HIDE them? (Isaac 3;8)

[^14]:    ${ }^{5}$ There was some difficulty eliciting negative questions, as in the first elicited production study. Fifteen positive questions were produced in response to a prompt for a negative question. Each child avoided at least one negative question target in his way.
    ${ }^{26}$ Of the 23 negative 2 Aux questions, two used the full form of negation.
    (i) Where did you do not plant the flowers? (Barbara 4;4)
    (ii) Where did you did not chase him?
    (Troy 4;5)

[^15]:    ${ }^{27}$ The questions produced are from the same child, Mary, in response to the pair of unstressed and stressed prompts.

[^16]:    38 We had difficulty using the task with the younger children and found that some children could not do the task. This result is different from McDaniel's, since she has been successful with children as young as 2;11.

[^17]:    ${ }^{25}$ The items were divided into three sessions for the first five children who were tested (Amne, Kendra, David, Linda and Scott). However, these sessions seemed too long, so we divided the items into four sessions for the remaining ten children.

[^18]:    ${ }^{30}$ For the adult subjects, one experimenter played the role of both the storyteller and the puppet, as with the elicited production task.

[^19]:    ${ }^{31}$ See also Chomsky (1957).
    ${ }^{32}$ One problem with assuming that negation is a head in English is that $V$ to I movement over negation violates the Head Movement Constraint (HMC).
    (i) Mandy is ${ }_{I}$ not $t_{I}$ crazy

    In ( $i$ ), the auxiliary verb 'is' has raised from $V$ to I , crossing negation. I will assume a possibility that Roberts (1993, 1994) proposes, which is discussed in Lasnik (1995). Roberts (1993, 1994) proposes that negation is an $A^{\prime}$ head while $V$ and $[$ are $A$ heads, and that Rizi's (1990) relativized minimality applies to head movement as well as XP movement He proposes that a head only blocks the movement of the same type of head. Hence negation, an $A^{\prime}$ head, will not block the movement of $V$ or $I$, both of which are A heads.

[^20]:    ${ }^{35}$ See Ernst (1992) for arguments based on stranding and distributional facts for treating constituent negation differently from sentential negation.

[^21]:    34 Since constituent negation involves stress on the negative marker, it follows that $n$ 't camnot be used for constituent negation since affixes usually cannot be stressed.

[^22]:    ${ }^{35}$ However, see Chomsky (1994, 1995) for arguments that an item, such as clitics in Romance languages, can be both an $\mathrm{X}^{0}$ and an XP.

[^23]:    ${ }^{36}$ The idea of pronouncing multiple copies of a moved element has been proposed by Nunes (1995). He argues that both the head and the tail of a non-trivial chain can be pronomiced, if the chain nivolves $\mathrm{X}^{0}$-elements.

[^24]:    ${ }^{37}$ The sentence is marginally acceptable with the second Wh-phrase fronted if it is very heavily

[^25]:    stressed.

[^26]:    ${ }^{38}$ None of Chomsky's accoumts, exciuding Subject island violations, are straightforward. First, he remains vague about the position of the complementizer whether in the Wh-island violation in (8). Second, according to his analysis, movement out of a CNP does not cross any barriers, yet he proposes that there must be one barrier, either the NP or CP, due to its "intermediate" status. Finally, Chomsky claims that movement out of an adjunct crosses two barriers, but does not make explicit what the adjumet must be adjoined to in order to ensure the crossing of two barriers.

[^27]:    ${ }^{39}$ Takashi assumes that a specifier position is created by adjunction to X .
    ${ }^{50}$ This second group of sentence types could be broken down into two further groups (Subject island violations vs. CNPC and Adjunct island violations) based on which part of the UCA is used. This same approach could be used for Huang's CED to distinguish Subject and Adjunct island violations. However, we will see in section 6 that this disjunction approach is not able to accoumt for the satiation data.

[^28]:    41 All of the examples of type a-f involve $A^{\prime}$-movement of an argument.
    42 This construction is acceptable in some dialects of English.

[^29]:    *The instructions were the same as those used by Snyder (1994).

[^30]:    44 This was also the case in Snyder (1994). Subject island violations with animate sabjects and transitive verbs are tested in Experiment 2.

[^31]:    *5 It has been reported anecdotally that some people from Comecticut accept that-trace violations.

[^32]:    ${ }^{46}$ Since all of the Subject island examples involved either an unaccusative or passive verb, one might argue that the subject NP originated in the complement position of the lower verb. Let us assume that if the subject NP is base-generated in the complement position, it can then raise to [spec, VP].
    ${ }^{47}$ There is another derivation we might consider. Suppose that the subject NP is base-generated in the complement position of the lower verb (see fin. 46) and the Wh-element moves out of this position. There will be no barriers crossed. The NP is no longer a blocking category (or a barrier) since it is Lmarked by the lower verb. However, this type of derivation is not allowed. See Collins (1994) and references therein for arguments against this type of derivation.

[^33]:    ${ }^{43}$ This possibility of weakening constraints does not explain the satiation effects for Wh-islands.

[^34]:    *9 One major problem remains with respect to modifying Takahashi's theory to incorporate the satiation data. It is unclear how to accoumt for the Wh-island satiation effects. We do not want to say

[^35]:    that subjects are satiating on the SMC because this would predict that all of the things that Takahashi tries to unify under the SMC should show satiation effects.

